DISSERTATION

BUEHLER CONFIDENCE REGIONS AND THE COMPARISON OF DIAGNOSTIC TESTS

TODOR DINEV

zur Erlangung des akademischen Grades »Doktor der Naturwissenschaften«
dem Fachbereich IV der Universität Trier im Dezember 2015 vorgelegt

Betreuer:	Prof. Dr. Lutz Mattner
Berichterstatter:	Prof. Dr. Ludwig Baringhaus

Todor Dinev: *Buehler confidence regions and the comparison of diagnostic tests,* PhD Thesis. December 2015

ABSTRACT

In the first part of this work we generalize a method of building optimal confidence bounds provided in Buehler (1957) by specializing an exhaustive class of confidence regions inspired by Sterne (1954). The resulting confidence regions, also called *Buehlerizations,* are valid in general models and depend on a "designated statistic" that can be chosen according to some desired monotonicity behaviour of the confidence region. For a fixed designated statistic, the thus obtained family of confidence regions indexed by their confidence level is nested. Buehlerizations have furthermore the optimality property of being the smallest (w.r.t. set inclusion) confidence regions that are increasing in their designated statistic. The theory is eventually applied to normal, binomial, and exponential samples.

The second part deals with the statistical comparison of pairs of diagnostic tests and establishes relations 1. between the sets of lower confidence bounds, 2. between the sets of pairs of comparable lower confidence bounds, and 3. between the sets of admissible lower confidence bounds in various models for diverse parameters of interest.

ZUSAMMENFASSUNG

Der erste Teil dieser Arbeit widmet sich der Verallgemeinerung eines Verfahrens von Buehler (1957) zur Konstruktion optimaler Konfidenzschranken, ausgehend von einer von Sterne (1954) inspirierten, in naheliegendem Sinne universellen Klasse von Konfidenzbereichen. Die dabei gebildeten Konfidenzbereiche, auch *Buehlerisierungen* genannt, sind in allgemeinen Modellen gültig und hängen von einer sog. »designierten Statistik« ab, welche gemäß eines gewünschten Monotonieverhaltens des Konfidenzbereiches gewählt werden kann. Für eine feste designierte Statistik besitzt die durch Indizierung durch das Konfidenzniveau entstandene Familie die Schachtelungseigenschaft. Buehlerisierungen besitzen ferner folgende Optimalitätseigenschaft: Sie sind die (bzgl. mengentheoretischer Inklusion) kleinsten Konfidenzbereiche, welche bzgl. der designierten Statistik wachsen. Die Theorie wird schließlich auf Normal-, Binomial- und Exponentialverteilungsmodelle angewandt.

Der zweite Teil befasst sich mit dem statistischen Vergleich von Paaren diagnostischer Tests und stellt Beziehungen her 1. zwischen den Mengen unterer Konfidenzschranken, 2. zwischen den Mengen von Paaren vergleichbarer unterer Konfidenzschranken und 3. zwischen den Mengen zulässiger unterer Konfidenzschranken in mehreren Modellen für diverse interessierende Parameter.

The greatest good you can do for another
is not just to share your riches,
but to reveal to him his own.

— Benjamin Disraeli

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my doctoral supervisor, Prof. Dr. Lutz Mattner, for his excellent guidance, in mathematics as well as numerous other branches, through the past four years. As a member of his scientific staff at the University of Trier I had the opportunity to work in a highly rewarding environment and gain invaluable experience.

I also wish to very kindly thank Prof. Dr. Ludwig Baringhaus for his willingness to referee this thesis on rather short notice.

It goes without saying how deeply grateful I am to my family for supporting me in every imaginable way, and to my friends for being wonderful companions. Oliver Hauke and Jona Schulz deserve special mentioning for carefully proofreading the manuscript.

Trier, December 2015

SYMBOLS AND ABBREVIATIONS

LHS, RHS	left-hand side, right-hand side
$\lvert A\rvert$	cardinality of a set A
$\bigvee$, $\bigwedge$	maximum, minimum
x^-, x^+	negative, positive part of a real number x
$\lvert m\rvert$	the sum of the components of a multi-index m
$f[A]$	image of a set A under a function f, i.e., $\{f(x)\colon x \in A\}$
$f^{-1}[B]$	preimage of a set B under a function f, i.e., $\{x\colon f(x) \in B\}$
pr_i	projection onto the ith coordinate
$\mathbf{1}(\mathsf{p})$	indicator of a proposition p; equals 1 if p is true, otherwise 0
δ_x	unit mass at a point x
B_p	Bernoulli law with success probability p
$\mathrm{B}_{n,p}$	binomial law with sample size n and success probability p
E_λ	exponential law with rate λ
$\mathrm{N}_{\mu,\sigma^2}$	normal law with mean μ and variance σ^2
Φ	distribution function of the standard normal law $\mathrm{N}_{0,1}$
φ	usual density of the standard normal law $\mathrm{N}_{0,1}$
$\operatorname{supp} P$	support of a law P
$\Longrightarrow$, $\rightsquigarrow$, $\twoheadrightarrow$	relations between models, introduced in Definition 5.3.2

In most of the remarks and many examples, the assumptions about the objects occurring therein have been omitted for the sakes of brevity and a smoother readability. In such cases, the assumptions of the immediately preceding definition, theorem, or lemma are tacitly presupposed.

CONTENTS

INTRODUCTION

This work consists of two independent parts. The first part (Chapters 1–4) investigates a class of confidence regions introduced in Buehler (1957), studied and generalized in a multitude of papers such as Sudakov (1974), Winterbottom (1984), Harris and Soms (1991), Reiser and Jaeger (1991), Revyakov (1992), Kabaila and Lloyd (1997), Kabaila and Lloyd (2000), Kabaila (2001), Kabaila and Lloyd (2002), Kabaila and Lloyd (2003), Lloyd and Kabaila (2003), Kabaila and Lloyd (2004), Kabaila and Lloyd (2006), Kabaila (2013), and recently reinvented (see Lloyd and Kabaila, 2010) in Wang (2010). In the literature these confidence regions have mainly been studied as confidence *bounds,* and are therefore known under the names "Buehler bounds," "tight confidence limits," or "smallest upper/greatest lower confidence bound."

Before developing the theory of Buehler bounds, we introduce a class of confidence regions in a general model $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ on an arbitrary measurable space $(\mathcal{X}, \mathfrak{A})$ for the identity id_Θ as parameter of interest. These confidence regions, very similar in nature to the ones introduced in Sterne (1954), are given for $\beta \in [0, 1]$ by

$$R_{\mathcal{T},\beta}(x) := \{\vartheta \in \Theta \colon P_\vartheta(T_\vartheta > T_\vartheta(x)) < \beta\} \qquad \text{for } x \in \mathcal{X}$$

and depend on a family $\mathcal{T} = (T_\vartheta \colon \vartheta \in \Theta)$ of statistics T_ϑ taking values in a separable totally preordered set (which is practically always the real line). The family consisting of the $R_{\mathcal{T},\beta}$ turns out to be 1. nested if β varies, and 2. exhaustive if $\mathcal{T}$ varies, meaning that every confidence region for the identity in $\mathcal{P}$ can be written as $R_{\mathcal{T},\beta}$ for some $\mathcal{T}$ as above. The latter universality property might appear interesting but it also makes this class too wide to exhibit any optimality properties. If we restrict our focus, however, to families $\mathcal{T}$ consisting of a single statistic T, henceforth called "designated statistic," the then resulting confidence region $R_{T,\beta}$ can be shown to be the smallest (with respect to set inclusion) confidence region (with level β and for the identity in $\mathcal{P}$) that is increasing in T. Under certain natural conditions, $R_{T,\beta}$ becomes a confidence ray and can indeed be considered a generalization of Buehler bounds, which explains the designation "Buehler confidence region." If, instead of its general aspect as a set-valued function admitting a certain confidence property, its designated statistic is emphasized, $R_{T,\beta}$ is called "Buehlerization of T."

The theory of Buehler confidence regions is presented here in a generality that may be uncommon for some parts of statistics. We believe, however, that this approach simplifies the comprehension of some proofs and does not noticeably hinder the reader's grasping of the essential notions. As for designated statistics, they can be assumed to take real values for two reasons: firstly, the examples in the subsequent chapters employ solely real-valued designated statistics; secondly, Theorem A.1.74 states that every separable totally preordered set is in essence a subset of the real line.

Buehler's theory is exemplarily applied to normal, binomial, and exponential samples. In the normal setting, we start with the rather general model

$$\left(\bigotimes_{i=1}^{n} \mathrm{N}_{\mu_i,\sigma_i^2}^{\otimes m_i} \colon (\mu, \sigma) \in \mathbb{R}^n \times \left]0, \infty\right[^n \right)$$

of n independent samples of known sizes m_i and both unknown means μ_i and variances σ_i^2, and buehlerize several designated statistics. These examples, though most of them not yielding useful confidence regions, prove fruitful in the more specialized models that follow, where either variances σ_i^2 or means μ_i become known. In the binomial model

$$\left(\bigotimes_{i=1}^{n} \mathrm{B}_{m_i,p_i} : p \in [0,1]^n\right)$$

of n independent binomial samples of known sizes m_i and unknown success probabilities p_i, we consider variations of one designated statistic, namely the usual estimator for the success probabilities vector. Several more specialized models appear for mainly illustrative purposes. Buehlerization in the exponential model

$$\left(\bigotimes_{i=1}^{n} \mathrm{E}_{\lambda_i}^{\otimes m_i} : \lambda \in \,]0,\infty[^n\right)$$

of n independent exponential samples of known sizes m_i and unknown rates λ_i yields useful confidence regions with minimal effort and very straight-forward calculations. If the reader wishes to obtain a glimpse into the practical application of the theory from Chapter 1, this might be the right place to start.

The second part of this thesis (Chapter 5) investigates statistical relations between several models for pairs of diagnostic tests. After a short informal introduction explaining the terms “diagnostic test,” “sensitivity/specificity,” and “predictive values,” we state a multinomial model by Gart and Buck (1966) that allows the study of pairs of diagnostic tests when true states of the members of the population are unobservable and the prevalence of the condition being examined is unknown. The main result establishes relations

- between the sets of lower confidence bounds,
- between the sets of pairs of comparable lower confidence bounds,
- between the sets of admissible lower confidence bounds

in various models for diverse parameters of interest. The proof of the result rests on a number of auxiliary results of essentially two different kinds: propositions allowing a (sometimes partial) reduction of a statement in a certain model to one in a similar, already covered model, and lemmas where images under certain linear maps of semialgebraic sets are computed in an elementary manner (that is to say, without tools from real algebraic geometry).

In short, the outline of this thesis is thus as follows. Chapter 1 presents the theory behind Buehler's ideas, beginning with the general class of confidence regions $R_{\mathcal{T},\beta}$ mentioned above, specializing to Buehlerizations $R_{T,\beta}$, and closing with some remarks on important work published in that area. Chapters 2–4 apply this theory to normal, binomial, and exponential models. The examples provided in these chapters assume some knowledge of the concepts and terminology from Chapter 1, but can be read independently of each other. Chapter 5 deals with the statistical comparison of pairs of diagnostic tests. It does not rely on the previous chapters and can be read independently. Appendix A recapitulates commonly used notions (such as functions and families, preorders, and topologies) and some basic results, provides a counterexample to a conjectured inequality from Chapter 1 employing ordinal numbers, and recalls several basic statistical concepts.

1

BUEHLER CONFIDENCE REGIONS

- In 1957 Robert J. Buehler presented an increasing upper confidence bound in a product binomial model for the product of the success probabilities. This confidence bound has the optimality property of being smaller than any other such bound. Buehler's method of construction was readily generalizable and has since found wide application in reliability theory. In statistics, however, this method remained until recently largely unknown despite its interesting features and potential widespread use. This chapter generalizes and develops some of the ideas published in the field of Buehler bounds.
- Outline of this chapter:
  - Section 1.1 introduces a class of confidence regions closely associated to both Sterne's (1954) confidence intervals and Buehler's (1957) method. It turns out that *every* confidence region is a member of this class by suitably selecting some parameter. The necessary order theoretic and statistical background is recapitulated in Sections A.1 and A.3, respectively, of Appendix A.
  - Section 1.2 specializes the confidence regions introduced in the previous section, introducing thus Buehler's concept in a general setting. Examples and applications to the theory developed in this section are presented in Chapters 2–4.
  - The notes in Section 1.3 briefly sketch some of the most important work in the field of Buehler bounds.

1.1 A CLASS OF CONFIDENCE REGIONS

1.1.1 Remark 1. Before considering the general problem of constructing confidence regions for a parameter of interest κ, we shall focus on the special case $\kappa = \mathrm{id}_\Theta$. Confidence regions for κ based on ones for id_Θ can be obtained (up to measurability issues) using Theorem A.3.21.

2. We generalize Sterne's (1954) construction of a confidence region for the binomial model $(\mathrm{B}_{n,p}\colon p \in [0,1])$ to arbitrary models. This generalized confidence region depends on an additional parameter, namely a family

$$\mathcal{T} = (T_\vartheta\colon \vartheta \in \Theta)$$

of statistics, which makes it encompass, by suitably varying this parameter, classes of well-known confidence regions.

3. As pointed out in Remark A.1.69, part 2, totally preordered sets are in the following always endowed with their order topologies.
4. The two next lemmas lay the foundations for the confidence property of the function considered in Definition 1.1.5.

1.1.2 Lemma *Let*

- P *be a law on a measurable space* $(\mathcal{X}, \mathfrak{A})$,
- $(\mathcal{Y}, \leq)$ *a separable totally preordered set,*
- $T\colon \mathcal{X} \to \mathcal{Y}$ *a statistic,*
- $F\colon \mathcal{X} \to [0,1]$, $x \mapsto P(T \leq T(x))$.

Then

- F *is measurable,*
- $P(F \leq F(x)) = F(x)$ *for* $x \in \mathcal{X}$.

If, moreover,

- Q *is a further law on* $(\mathcal{X}, \mathfrak{A})$,
- $G\colon \mathcal{X} \to [0,1]$, $x \mapsto Q(T \leq T(x))$,

then

- G *is measurable,*
- $P(F \leq F(x)) \leq P(G \leq G(x))$ *for* $x \in \mathcal{X}$.

Proof. 1. F is well-defined since $\{T \leq y\} = \mathcal{X} \setminus \{T > y\} \in \mathfrak{A}$ for $y \in \mathcal{Y}$ due to the measurability of T.

2. Let us consider $\tilde{F}\colon \mathcal{Y} \to [0,1]$, $y \mapsto P(T \leq y)$. Since $\tilde{F}$ is increasing, it is measurable by Remark A.1.69, part 6. The measurability of F thus follows from $F = \tilde{F} \circ T$.
3. Let now $x \in \mathcal{X}$. For $\xi \in \mathcal{X}$, the inequality $F(\xi) \leq F(x)$ is equivalent to either $T(\xi) \leq T(x)$ or both $T(\xi) > T(x)$ and $P(T(x) < T \leq T(\xi)) = 0$. Thus, by setting

$$\begin{aligned} A_F &\coloneqq \{\xi \in \mathcal{X}\colon T(\xi) > T(x),\ P(T(x) < T \leq T(\xi)) = 0\} \\ &= \{T > T(x)\} \setminus \{F > F(x)\}, \end{aligned}$$

we obtain $A_F \in \mathfrak{A}$ and

$$P(F \leq F(x)) = F(x) + P(A_F).$$

4. We now show $P(A_F) = 0$. Let us first assume $T[A_F]$ has a greatest element, say $T(\xi)$ with $\xi \in A_F$. Then $A_F = \{T(x) < T \leq T(\xi)\}$:
   - If $\tilde{x} \in A_F$, then $T(\tilde{x}) > T(x)$ by definition of A_F, and $T(\tilde{x}) \leq \max T[A_F] = T(\xi)$ by definition of ξ, hence $\tilde{x} \in \{T(x) < T \leq T(\xi)\}$.
   - Let $\tilde{x} \in \{T(x) < T \leq T(\xi)\}$. Then $T(\tilde{x}) > T(x)$ trivially, and $F(\tilde{x}) = P(T \leq T(\tilde{x})) \leq P(T \leq T(\xi)) = F(\xi) \leq F(x)$, where the first inequality is due to $T(\tilde{x}) \leq T(\xi)$, which holds by assumption, combined with the monotonicity of measures, and the second inequality follows from $\xi \in A_F$ and the representation of A_F in part 3. This yields $\tilde{x} \in A_F$.

   From this follows $P(A_F) = 0$ by the definition of A_F.

   Let us now suppose that $T[A_F]$ has no greatest element, i.e., let us assume the existence of a function $g\colon A_F \to A_F$ such that $T(g(\xi)) > T(\xi)$ for $\xi \in A_F$. The separability of $\mathcal{Y}$ implies the existence of a function $y\colon A_F \to \mathcal{Y}$ with countable image such that $T(\xi) < y(\xi) < T(g(\xi))$ for $\xi \in A_F$. This yields $A_F = \bigcup_{\xi \in A_F} \{T(x) < T \leq y(\xi)\}$:
   - Let $\tilde{x} \in A_F$. Then $T(\tilde{x}) > T(x)$ trivially, and $T(\tilde{x}) < y(\tilde{x})$ by definition of y. This implies $\tilde{x} \in \{T(x) < T \leq y(\tilde{x})\} \subseteq \bigcup_{\xi \in A_F} \{T(x) < T \leq y(\xi)\}$.
   - Let now $\tilde{x} \in \{T(x) < T \leq y(\xi)\}$ for some $\xi \in A_F$. This means $T(\tilde{x}) > T(x)$ and $T(\tilde{x}) \leq y(\xi)$. Since $y(\xi) < T(g(\xi))$ by definition of y, the latter inequality

combined with the monotonicity of measures yields $P(T(x) < T \leq T(\tilde{x})) \leq P(T(x) < T \leq T(g(\xi))) = 0$, the latter equality following from $g(\xi) \in A_F$ and the definition of A_F. This implies $\tilde{x} \in A_F$.

Since y has countable image, there is a countable subset $A'_F \subseteq A_F$ such that $\{\{T(x) < T \leq y(\xi)\} \colon \xi \in A_F\} = \{\{T(x) < T \leq y(\xi)\} \colon \xi \in A'_F\}$. This yields

$$\begin{aligned} P(A_F) &= P\left(\bigcup_{\xi \in A_F} \{T(x) < T \leq y(\xi)\}\right) \\ &= P\left(\bigcup_{\xi \in A'_F} \{T(x) < T \leq y(\xi)\}\right) \\ &\leq P\left(\bigcup_{\xi \in A'_F} \{T(x) < T \leq T(g(\xi))\}\right) \\ &\leq 0, \end{aligned}$$

where the first inequality follows from the definition of y and the monotonicity of measures, and the second inequality from the countability of A'_F, the σ-subadditivity of measures, and $g(\xi) \in A_F$ for $\xi \in A'_F$.

5. By applying the above to Q and G instead of P and F, we obtain the measurability of G and

$$P(G \leq G(x)) = F(x) + P(A_G) \geq F(x) = P(F \leq F(x)). \qquad \square$$

1.1.3 Remark 1. The equality $P(F \leq F(x)) = F(x)$ in the conclusion of the last result yields an explicit formula for the effective levels of the confidence regions defined later.

2. Not even the weaker result $P(F \leq F(x)) \leq F(x)$ holds without presupposing
   - separability on $\mathcal{Y}$: Lemma A.2.6 yields a counterexample relying on ordinal numbers;
   - totality of the preorder $\leq$ on $\mathcal{Y}$: if $\mathcal{X} := \mathcal{Y} := \{0,1\}^2$ are endowed with the product order, $P := \mathrm{U}_{\{0,1\}^2}$ is the uniform distribution on $\{0,1\}^2$, and $T := \mathrm{id}_{\{0,1\}^2}$ the identity on $\{0,1\}^2$, then $P(F \leq F(1,0)) = 3/4 > 1/2 = F(1,0)$.
3. It may seem straightforward to characterize unbiasedness of the confidence regions studied in Section 1.2. The inequality $P(F \leq F(x)) \leq P(G \leq G(x))$ does not, however, yield unbiasedness; an inequality of the type $P(F \leq t) \leq P(G \leq t)$ for $t \in [0,1]$ would be required to this end. The resulting confidence regions from Section 1.2 turn out to be biased, as Remark 1.2.3, part 12, shows.
4. The set $\mathcal{Y}$ is practically always a subset of $\mathbb{R}$ or of the extended real line $\overline{\mathbb{R}}$. Theorems A.1.74 and A.1.71 allow a reduction of the general setting to the real one in any case.

1.1.4 Lemma *Let*

- *P be a law on a measurable space $(\mathcal{X}, \mathfrak{A})$,*
- *$(\mathcal{Y}, \leq)$ a separable totally preordered set,*
- *$T \colon \mathcal{X} \to \mathcal{Y}$ a statistic,*
- *$\beta \in [0,1]$.*

Then $\{x \in \mathcal{X} \colon P(T > T(x)) < \beta\} \in \mathfrak{A}$ and

$$P(\{x \in \mathcal{X} \colon P(T > T(x)) < \beta\}) = \inf\{P(T > T(x)) \colon x \in \mathcal{X},\ P(T > T(x)) \geq \beta\}.$$

Proof. For brevity, let us set again

$$F\colon \mathcal{X} \to [0,1], \quad x \mapsto P(T \leq T(x)).$$

Lemma 1.1.2 implies the measurability of F, and thus

$$\{x \in \mathcal{X}\colon P(T > T(x)) < \beta\} = (1-F)^{-1}\big[[0,\beta[\big] \in \mathfrak{A}. \tag{$*$}$$

We now show

$$P(F \leq 1-\beta) = \sup F[F \leq 1-\beta]. \tag{$**$}$$

Let us first assume the existence of some $x \in \mathcal{X}$ with $F(x) = \sup F[F \leq 1-\beta]$. Then $\{F \leq 1-\beta\} = \{F \leq F(x)\}$:

- If $\xi \in \{F \leq 1-\beta\}$, then $F(\xi) \leq \sup F[F \leq 1-\beta] = F(x)$ by the definition of a supremum.
- If $\xi \in \{F \leq F(x)\}$, then $F(\xi) \leq F(x) = \sup F[F \leq 1-\beta] \leq 1-\beta$.

Lemma 1.1.2 thus yields $P(F \leq 1-\beta) = P(F \leq F(x)) = F(x) = \sup F[F \leq 1-\beta]$.

Let us now assume $F(x) \neq \sup F[F \leq 1-\beta]$ for $x \in \mathcal{X}$. Then $\{F \leq 1-\beta\} = \{F < \sup F[F \leq 1-\beta]\}$:

- Let $\xi \in \{F \leq 1-\beta\}$. Then $F(\xi) \leq \sup F[F \leq 1-\beta]$ by the definition of a supremum. Since $F(\xi) \neq \sup F[F \leq 1-\beta]$ by assumption, we have $\xi \in \{F < \sup F[F \leq 1-\beta]\}$.
- Let $\xi \in \{F < \sup F[F \leq 1-\beta]\}$. Then $F(\xi) < \sup F[F \leq 1-\beta] \leq 1-\beta$.

If $\{F \leq 1-\beta\} = \emptyset$, then $P(F \leq 1-\beta) = 0 = \sup F[F \leq 1-\beta]$, as the latter supremum is taken in the unit interval, where $\sup \emptyset = 0$ holds. Let now $\{F \leq 1-\beta\} \neq \emptyset$ and let us pick a sequence $(x_n\colon n \in \mathbb{N}) \in \{F \leq 1-\beta\}^{\mathbb{N}}$ such that $(F(x_n)\colon n \in \mathbb{N})$ is increasing with $F(x_n) \to \sup F[F \leq 1-\beta]$ for $n \to \infty$. Then

$$\begin{aligned} P(F \leq 1-\beta) &= P(F < \sup F[F \leq 1-\beta]) \\ &= P(F < \sup\nolimits_{n\in\mathbb{N}} F(x_n)) \\ &= \sup\nolimits_{n\in\mathbb{N}} P(F \leq F(x_n)) \\ &= \sup\nolimits_{n\in\mathbb{N}} F(x_n) \\ &= \sup F[F \leq 1-\beta], \end{aligned}$$

where the first equality follows from what has just been shown, the second by construction of the sequence (x_n) and by assumption, the third from the continuity from below of measures, the fourth from Lemma 1.1.2, and the last one by construction of (x_n) again.

The equation ($**$) is thus shown. Using ($*$) in the first step and ($**$) in the second, we obtain

$$\begin{aligned} P(\{x \in \mathcal{X}\colon P(T > T(x)) < \beta\}) &= 1 - P(F \leq 1-\beta) \\ &= \inf(1-F)[1-F \geq \beta]. \end{aligned}$$ □

1.1.5 Definition (A general confidence procedure) Let

- $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$,
- $(\mathcal{Y}, \leq)$ a separable totally preordered set,
- $\mathcal{T} = (T_\vartheta \colon \vartheta \in \Theta)$ a family of statistics $T_\vartheta \colon \mathcal{X} \to \mathcal{Y}$,
- $\beta \in [0, 1]$.

Let us define

$$R_{\mathcal{T},\beta} \colon \mathcal{X} \to 2^\Theta, \quad x \mapsto \{\vartheta \in \Theta \colon P_\vartheta(T_\vartheta > T_\vartheta(x)) < \beta\}.$$

If $\mathcal{T}$ is constant, say $\mathcal{T} = (T \colon \vartheta \in \Theta)$, we write $R_{T,\beta}$ for $R_{\mathcal{T},\beta}$. If the spaces $\mathcal{X}$ and $\mathcal{Y}$ and their inherent preorders and topologies coincide, we write $R_{\mathrm{id},\beta}$ instead of $R_{\mathrm{id}_\mathcal{X},\beta}$. If different models are considered in the same context, we occasionally append the model to the subscript and write $R_{\mathcal{T},\beta,\mathcal{P}}$ and $R_{T,\beta,\mathcal{P}}$, respectively.

1.1.6 Theorem (Nested confidence regions) *Let*

- $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ *be a model on a measurable space* $(\mathcal{X}, \mathfrak{A})$,
- $(\mathcal{Y}, \leq)$ *a separable totally preordered set,*
- $\mathcal{T} = (T_\vartheta \colon \vartheta \in \Theta)$ *a family of statistics* $T_\vartheta \colon \mathcal{X} \to \mathcal{Y}$.

Then,

(i) *for* $\beta \in [0, 1]$, $R_{\mathcal{T},\beta}$ *is a confidence region for* id_Θ *with level* β *and effective level*

$$\beta_{\mathrm{eff}}(R_{\mathcal{T},\beta}) = \inf\{P_\vartheta(T_\vartheta > T_\vartheta(x)) \colon x \in \mathcal{X},\ \vartheta \in \Theta,\ P_\vartheta(T_\vartheta > T_\vartheta(x)) \geq \beta\}.$$

(ii) $(R_{\mathcal{T},\beta} \colon \beta \in [0, 1])$ *is a nested family, i.e.,*

$$[0, 1] \to 2^\Theta, \quad \beta \mapsto R_{\mathcal{T},\beta}(x),$$

is increasing for fixed $x \in \mathcal{X}$.

Proof. Lemma 1.1.4 implies

$$\{R_{\mathcal{T},\beta} \ni \vartheta\} = \{x \in \mathcal{X} \colon P_\vartheta(T_\vartheta > T_\vartheta(x)) < \beta\} \in \mathfrak{A} \qquad \text{for } \vartheta \in \Theta$$

and

$$\begin{aligned}\inf_{\vartheta \in \Theta} P_\vartheta(R_{\mathcal{T},\beta} \ni \vartheta) &= \inf_{\vartheta \in \Theta} P_\vartheta(\{x \in \mathcal{X} \colon P_\vartheta(T_\vartheta > T_\vartheta(x)) < \beta\}) \\ &= \inf\{P_\vartheta(T_\vartheta > T_\vartheta(x)) \colon x \in \mathcal{X},\ \vartheta \in \Theta,\ P_\vartheta(T_\vartheta > T_\vartheta(x)) \geq \beta\} \\ &\geq \beta.\end{aligned}$$

Part (ii) follows immediately from the definition of $R_{\mathcal{T},\beta}$. □

1.1.7 Remark 1. The confidence region $R_{\mathcal{T},\beta}$ is a direct generalization of the confidence region by Sterne (1954), where $\mathcal{T}$ consists of densities of the model $\mathcal{P}$.

2. $R_{\mathcal{T},\beta}$ has effective level β if, and only if, there are sequences $(\vartheta_n \colon n \in \mathbb{N})$ and $(\xi_n \colon n \in \mathbb{N})$ taking values in Θ and $\mathcal{X}$, respectively, with

$$\inf_{n \in \mathbb{N}} P_{\vartheta_n}(T_{\vartheta_n} > T_{\vartheta_n}(\xi_n)) = \beta.$$

Let us first show the "if" part. Let $(\vartheta_n \colon n \in \mathbb{N})$ and $(\xi_n \colon n \in \mathbb{N})$ be two sequences with values in Θ and $\mathcal{X}$, respectively, such that $\inf_{n \in \mathbb{N}} P_{\vartheta_n}(T_{\vartheta_n} > T_{\vartheta_n}(\xi_n)) = \beta$. Then

$$\{P_\vartheta(T_\vartheta > T_\vartheta(x))\colon x \in \mathcal{X},\ \vartheta \in \Theta,\ P_\vartheta(T_\vartheta > T_\vartheta(x)) \geq \beta\} \\ \supseteq \{P_{\vartheta_n}(T_{\vartheta_n} > T_{\vartheta_n}(\xi_n))\colon n \in \mathbb{N}\},$$

hence

$$\begin{aligned}\beta_{\mathrm{eff}}(R_{\mathcal{T},\beta}) &= \inf\{P_\vartheta(T_\vartheta > T_\vartheta(x))\colon x \in \mathcal{X},\ \vartheta \in \Theta,\ P_\vartheta(T_\vartheta > T_\vartheta(x)) \geq \beta\} \\ &\leq \inf_{n\in\mathbb{N}} P_{\vartheta_n}(T_{\vartheta_n} > T_{\vartheta_n}(\xi_n)) \\ &= \beta.\end{aligned}$$

Since also $\beta_{\mathrm{eff}}(R_{\mathcal{T},\beta}) \geq \beta$, we obtain $\beta_{\mathrm{eff}}(R_{\mathcal{T},\beta}) = \beta$.
The "only if" part follows from the separability of the unit interval or, more precisely, from the existence, given a non-empty set $A \subseteq [0,1]$, of a sequence $(a_n\colon n \in \mathbb{N})$ with values in A such that $\inf A = \inf_{n\in\mathbb{N}} a_n$.

3. For $\beta = 0$ we obtain $R_{\mathcal{T},0} \equiv \emptyset$. This confidence region obviously also has effective level β. We therefore often presuppose $\beta > 0$ in the following calculations.
4. If $\mathcal{X}$ is preordered and each statistic T_ϑ of the family $\mathcal{T}$ increasing, then so is $R_{\mathcal{T},\beta}$. In other words, if $(\mathcal{X}, \leq)$ is a preordered set and
$$x_1 \leq x_2 \quad\Longrightarrow\quad T_\vartheta(x_1) \leq T_\vartheta(x_2) \qquad \text{for } x_1, x_2 \in \mathcal{X} \text{ and } \vartheta \in \Theta,$$
then
$$x_1 \leq x_2 \quad\Longrightarrow\quad R_{\mathcal{T},\beta}(x_1) \subseteq R_{\mathcal{T},\beta}(x_2) \qquad \text{for } x_1, x_2 \in \mathcal{X}.$$
5. The property of nestedness states the implication
$$\beta_1 \leq \beta_2 \quad\Longrightarrow\quad R_{\mathcal{T},\beta_1}(x) \subseteq R_{\mathcal{T},\beta_2}(x) \qquad \text{for } \beta_1, \beta_2 \in [0,1] \text{ and } x \in \mathcal{X}.$$
6. Remark A.3.25 implies that
$$f\colon [0,1] \to [0,1], \quad \beta \mapsto \beta_{\mathrm{eff}}(R_{\mathcal{T},\beta}),$$
is increasing. Furthermore, $f(0) = 0$, $f(1) = 1$, and $f(\beta) \geq \beta$ for $\beta \in [0,1]$. The Bernoulli example in Section 3.3, combined with Theorem 1.1.8, shows that f need not be continuous.
7. Even if $\mathcal{T} = (T_\vartheta\colon \vartheta \in \Theta)$ and $\mathcal{S} = (S_\vartheta\colon \vartheta \in \Theta)$ consist of densities T_ϑ and S_ϑ of P_ϑ with respect to measures μ and ν, respectively, we need not have $R_{\mathcal{T},\beta} = R_{\mathcal{S},\beta}$. In fact, if $\{0,1\}$ is the sample space, $P := \mu := \delta_0$ the unit mass at 0, and $\nu := \delta_0 + \delta_1$ the counting measure on $\{0,1\}$, then $T := \mathbf{1}_{\{0,1\}}$ and $S := \mathbf{1}_{\{0\}}$ yield $P(T > T(1)) = 0$ and $P(S > S(1)) = 1$, which implies
$$R_{T,\beta}(1) = \{P\} \quad\text{and}\quad R_{S,\beta}(1) = \emptyset$$
for $\beta \in \,]0,1[$ in the model $\mathrm{id}_{\{P\}}$ consisting of P alone.
8. Theorem 1.1.6 also applies to
$$\tilde{R}_{\mathcal{T},\beta}\colon \mathcal{X} \to 2^\Theta, \quad x \mapsto \{\vartheta \in \Theta\colon P_\vartheta(T_\vartheta < T_\vartheta(x)) < \beta\},$$
upon reversing, of course, the strict inequality sign "<" in part (i); this follows from considering the dual order on $\mathcal{Y}$. Theorem 1.1.8, however, states that we can restrict our (theoretical) focus to the investigation of $R_{\mathcal{T},\beta}$.

9. The intersection $R_{\mathcal{T},\beta_1} \cap \tilde{R}_{\mathcal{T},\beta_2}$ is thus also a confidence region with level β whenever $\beta_1, \beta_2 \in [0,1]$ are such that $\beta_1 + \beta_2 = 1 + \beta$ (due to Lemma A.3.23). These confidence regions are obviously also nested, that is,

$$\beta_1 \le \beta_1' \text{ and } \beta_2 \le \beta_2' \quad \Longrightarrow \quad R_{\mathcal{T},\beta_1}(x) \cap \tilde{R}_{\mathcal{T},\beta_2}(x) \subseteq R_{\mathcal{T},\beta_1'}(x) \cap \tilde{R}_{\mathcal{T},\beta_2'}(x)$$

for $\beta_1, \beta_2, \beta_1', \beta_2' \in [0,1]$ and $x \in \mathcal{X}$. We cannot, however, express the effective level of a confidence region built by intersection in terms of the effective levels of the individual confidence regions.

10. If $T_\vartheta \,\square\, P_\vartheta$ is continuous for every $\vartheta \in \Theta$, then

$$\tilde{R}_{\mathcal{T},\beta}(x) = \{\vartheta \in \Theta \colon P_\vartheta(T_\vartheta > T_\vartheta(x)) > 1 - \beta\} \qquad \text{for } x \in \mathcal{X},$$

and the effective level of $\tilde{R}_{\mathcal{T},\beta}$ is given by

$$\beta_{\text{eff}}(\tilde{R}_{\mathcal{T},\beta}) = 1 - \sup\{P_\vartheta(T_\vartheta > T_\vartheta(x)) \colon x \in \mathcal{X},\ \vartheta \in \Theta, \\ P_\vartheta(T_\vartheta > T_\vartheta(x)) \le 1 - \beta\}.$$

11. As the Bernoulli example in Section 3.3 shows, the effective levels of $R_{\mathcal{T},\beta}$ and $\tilde{R}_{\mathcal{T},\beta}$ need not be equal.
12. The next result relates $\tilde{R}_{\cdot,\beta}$ to $R_{\cdot,\beta}$.

1.1.8 Theorem *Let*

- $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ *be a model on a measurable space* $(\mathcal{X}, \mathfrak{A})$*,*
- $(\mathcal{Y}, \le)$ *a separable totally preordered set,*
- $\mathcal{T} = (T_\vartheta \colon \vartheta \in \Theta)$ *a family of statistics* $T_\vartheta \colon \mathcal{X} \to \mathcal{Y}$*,*
- $\beta \in [0,1]$.

Then there exist

- *a separable totally preordered set* $(\mathcal{Z}, \le)$*,*
- *a family* $\mathcal{S} = (S_\vartheta \colon \vartheta \in \Theta)$ *of statistics* $S_\vartheta \colon \mathcal{X} \to \mathcal{Z}$

such that $|\{S_\vartheta \colon \vartheta \in \Theta\}| = |\{T_\vartheta \colon \vartheta \in \Theta\}|$ *and*

$$\tilde{R}_{\mathcal{T},\beta} = R_{\mathcal{S},\beta}.$$

Proof. This follows after endowing $\mathcal{Z} := \mathcal{Y}$ with the dual $\ge$ of the order $\le$ on $\mathcal{Y}$. □

1.1.9 Remark 1. The part concerning the cardinality of the ranges of $\mathcal{T}$ and $\mathcal{S}$ makes this theorem also applicable in Section 1.2, where $\mathcal{T}$ consists of a single statistic.

2. In case $(\mathcal{Y}, \le)$ possesses a decreasing involution f (Remark A.1.9, part 3, defines the term "involution"), the cumbersome construction of $(\mathcal{Z}, \le)$ can be avoided by defining $\mathcal{S} := (f \circ T_\vartheta \colon \vartheta \in \Theta)$. Such is the case with $\mathcal{Y} := \mathbb{R}$ or $\mathcal{Y} :=]0, \infty[$ (take, e.g., $f(x) := -x$ or $f(x) := 1/x$, respectively).
3. The next result strengthens the statement from Theorem 1.1.8: Every confidence region with level $\beta > 0$ for id_Θ is attained by some $R_{\mathcal{T},\beta}$ by suitable (and simple) choice of $\mathcal{Y}$ and $\mathcal{T}$.

1.1.10 Theorem (Universality) *Let* $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ *be a model on a measurable space* $(\mathcal{X}, \mathfrak{A})$ *and* $\beta \in]0,1]$*. If* R *is a confidence region for* id_Θ*, then there is a family* $\mathcal{T} = (T_\vartheta \colon \vartheta \in \Theta)$ *of measurable indicators* $T_\vartheta \colon \mathcal{X} \to \{0,1\}$ *such that*

$$R = R_{\mathcal{T},\beta'} \qquad \textit{for every } \beta' \in]0, \beta_{\text{eff}}(R)].$$

Proof. Let R be a confidence region for id_Θ and

$$T_\vartheta\colon \mathcal{X} \to \{0,1\}, \quad x \mapsto \mathbf{1}(\vartheta \in R(x)),$$

for $\vartheta \in \Theta$. Since R is a confidence region, we have $\{R \ni \vartheta\} \in \mathfrak{A}$, which implies the measurability of T_ϑ, for $\vartheta \in \Theta$.

Let $\beta' \in \,]0, \beta_{\mathrm{eff}}(R)]$. For $\vartheta \in \Theta$ and $x \in \mathcal{X}$, we have equivalence between $\vartheta \in R(x)$ and $P_\vartheta(T_\vartheta > T_\vartheta(x)) < \beta'$. In fact, $\vartheta \in R(x)$ implies $T_\vartheta(x) = 1$, hence, T_ϑ being $\{0,1\}$-valued, $\{T_\vartheta > T_\vartheta(x)\} = \emptyset$, and thus $P_\vartheta(T_\vartheta > T_\vartheta(x)) = 0 < \beta'$; conversely, $\vartheta \notin R(x)$ implies $T_\vartheta(x) = 0$, so $\{T_\vartheta > T_\vartheta(x)\} = \{R \ni \vartheta\}$, whence $P_\vartheta(T_\vartheta > T_\vartheta(x)) = P_\vartheta(R \ni \vartheta) \geq \beta'$.

This equivalence yields $R = R_{\mathcal{T},\beta'}$. □

1.1.11 Remark 1. The above result cannot hold for $\beta = 0$ in view of $R_{\mathcal{T},0}(x) = \emptyset$ for $x \in \mathcal{X}$, which holds independently of $\mathcal{T}$.

2. In light of Theorem 1.1.6, part (ii), the latter result might insinuate that, since every confidence region for id_Θ is in fact some $R_{\mathcal{T},\beta}$, every family of confidence regions is nested. This is, of course, not true since the construction of the family $\mathcal{T}$ in Lemma 1.1.10 inevitably depends on R and thus on β. Theorem 1.1.6, part (ii), merely claims the monotonicity of $\beta \mapsto R_{\mathcal{T},\beta}$ with *fixed* $\mathcal{T}$.
3. Theorem 1.1.10 is mostly of theoretical interest. Indeed, to many well-known confidence regions correspond canonical families $\mathcal{T}$ which do not necessarily consist solely of indicators. For instance, the famous Clopper-Pearson confidence regions are obtained by considering $R_{T,\beta}$ and $\tilde{R}_{T,\beta}$ for

$$T := \mathrm{id}_{\{0,\ldots,n\}};$$

Sterne's confidence regions are obtained as $R_{\mathcal{D},\beta}$ by considering the family $\mathcal{D}$ consisting of the densities with respect to counting measure, i.e.,

$$\mathcal{D} = (\mathrm{b}_{n,p}\colon p \in [0,1]).$$

It would be misleading to restrict our attention to families of $\{0,1\}$-valued statistics. This justifies the general setting of Theorem 1.1.6.
4. Theorem 1.1.6 and the proof of Theorem 1.1.10 show that the effective level of a confidence region R for id_Θ with level $\beta \in \,]0,1]$ is also given by

$$\beta_{\mathrm{eff}}(R) = \inf\{P_\vartheta(R \setminus R(x) \ni \vartheta)\colon x \in \mathcal{X},\ \vartheta \in \Theta,\ P_\vartheta(R \setminus R(x) \ni \vartheta) \geq \beta\}.$$

5. The following result shows that our considering the special parameter of interest id_Θ does not entail any loss in generality.

1.1.12 Theorem (General universality) *Let $\mathcal{P} = (P_\vartheta\colon \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$, $\kappa\colon \Theta \to \Gamma$ a parameter of interest, and $\beta \in \,]0,1]$. If* K *is a confidence region for κ, then there is a family $\mathcal{T} = (T_\vartheta\colon \vartheta \in \Theta)$ of measurable indicators $T_\vartheta\colon \mathcal{X} \to \{0,1\}$ such that*

$$\mathrm{K} = \kappa[R_{\mathcal{T},\beta'}] \qquad \text{for every } \beta' \in \,]0, \beta_{\mathrm{eff}}(\mathrm{K})].$$

Proof. Theorem A.3.21, part (ii), yields the confidence region $R := \kappa^{-1}[\mathrm{K}]$ for id_Θ with $\beta_{\mathrm{eff}}(R) = \beta_{\mathrm{eff}}(\mathrm{K})$. Theorem 1.1.10 yields a family $\mathcal{T}$ of measurable indicators with $R = R_{\mathcal{T},\beta'}$ for $\beta' \in \,]0, \beta_{\mathrm{eff}}(R)]$. This yields $\kappa[R_{\mathcal{T},\beta'}] = \mathrm{K}$ for $\beta' \in \,]0, \beta_{\mathrm{eff}}(\mathrm{K})]$. □

1.2 BUEHLERIZATION

1.2.1 Remark We now focus on constant families $\mathcal{T} = (T : \vartheta \in \Theta)$.

1.2.2 Definition (Buehlerization) Let

- $\mathcal{P} = (P_\vartheta : \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$,
- $(\mathcal{Y}, \leq)$ a separable totally preordered set,
- $T : \mathcal{X} \to \mathcal{Y}$ a statistic,
- $\beta \in [0, 1]$.

The confidence region $R_{T,\beta} = R_{T,\beta,\mathcal{P}}$ from Definition 1.1.5, i.e.,

$$R_{T,\beta} : \mathcal{X} \to 2^\Theta, \quad x \mapsto \{\vartheta \in \Theta : P_\vartheta(T > T(x)) < \beta\},$$

shall be called **Buehlerization** of the **designated statistic** T (in the model $\mathcal{P}$).

1.2.3 Remark (Properties of Buehlerizations) 1. Let us endow $\mathcal{X}$ with the total preorder $\leq_T$ induced by T (defined in Lemma A.1.44, part 1) and its order topology (from Definition A.1.68). If $\mathcal{X}$ is separable, then Remark A.1.46, part 4, yields $R_{T,\beta}(x) \supseteq R_{\mathrm{id},\beta}(x)$ and $\tilde{R}_{T,\beta}(x) \supseteq \tilde{R}_{\mathrm{id},\beta}(x)$ for $x \in \mathcal{X}$, with equality everywhere if T is injective. If, furthermore, $\mathfrak{A}$ contains all Borel sets in $\mathcal{X}$ (defined in Remark A.1.69, part 3), then $R_{\mathrm{id},\beta}$ and $\tilde{R}_{\mathrm{id},\beta}$ are also confidence regions with level β for id_Θ.

2. We have $R_{T,\beta,\mathcal{P}} = R_{\mathrm{id},\beta,T\square\mathcal{P}} \circ T$ (the "pushforward model" $T \square \mathcal{P}$ is defined in Remark A.3.9).
3. The confidence region $\tilde{R}_{T,\beta}$ shall in view of Theorem 1.1.8 also be called *Buehlerization of T*. It possesses dual properties to $R_{T,\beta}$.
4. If $T \square P_\vartheta$ is continuous for every $\vartheta \in \Theta$, then

$$\tilde{R}_{T,\beta}(x) = \{\vartheta \in \Theta : P_\vartheta(T > T(x)) > 1 - \beta\}.$$

5. Let us interpret events $A \in \mathfrak{A}$ with $P_\vartheta(A) \geq \beta$ as *probable* and ones with $P_\vartheta(A) \leq 1 - \beta$ as *improbable* under P_ϑ. Let us also call an observation $x_1 \in \mathcal{X}$ *more extreme* (with respect to T) than an observation $x_2 \in \mathcal{X}$ if $T(x_1) > T(x_2)$.
   - $R_{T,\beta}(x)$ consists of those $\vartheta \in \Theta$ that make the occurrence of an observation that is more extreme than x not probable under P_ϑ.
   - If $T \square P_\vartheta$ is continuous for $\vartheta \in \Theta$, then $\tilde{R}_{T,\beta}(x)$ consists of those parameters $\vartheta \in \Theta$ that make the occurrence of an observation that is more extreme than x not improbable under P_ϑ.
6. If $\mathcal{X}$ is a topological space and T continuous and unbounded below on $\operatorname{supp} P_\vartheta$ for $\vartheta \in \Theta$ (unboundedness and the support of a law are defined in Definitions A.1.43 and A.3.2, respectively), then $R_{T,1}(x) = \Theta$ for $x \in \mathcal{X}$. (The same result holds with "below" and "$R_{T,1}$" replaced by "above" and "$\tilde{R}_{T,1}$," respectively.) In fact, the continuity of T yields the openness of $\{T < T(x)\}$, the unboundedness assumption yields $\{T < T(x)\} \cap \operatorname{supp} P_\vartheta \neq \emptyset$, and the claim now follows from Remark A.3.3, part 1.
7. If $\kappa : \Theta \to \Gamma$ is a parameter of interest for $\mathcal{P}$, then $\kappa[R_{T,\beta}]$ is increasing in T and $\kappa[\tilde{R}_{T,\beta}]$ is decreasing in T (monotonicity in T is defined in Definition A.1.40). In fact, if $x_1, x_2 \in \mathcal{X}$ with $T(x_1) \leq T(x_2)$, then for $\vartheta \in \Theta$

$$P_\vartheta(T > T(x_1)) \geq P_\vartheta(T > T(x_2)) \quad \text{and} \quad P_\vartheta(T < T(x_1)) \leq P_\vartheta(T < T(x_2)),$$

hence $\kappa[R_{T,\beta}(x_1)] \subseteq \kappa[R_{T,\beta}(x_2)]$ and $\kappa[\tilde{R}_{T,\beta}(x_1)] \supseteq \kappa[\tilde{R}_{T,\beta}(x_2)]$.

8. It follows from the respective definitions that $\kappa[R_{T,\beta}]$ has the same properties in the set of all confidence regions for κ with level β that are increasing in T as $\kappa[\tilde{R}_{T,\beta}]$ in the set of all confidence regions for κ with level β that are decreasing in T.
9. Theorem 1.1.8 states that we can again focus on $R_{T,\beta}$.
10. $R_{T,\beta}$ is strictly increasing in T if, and only if, $x_1, x_2 \in \mathcal{X}$ with $T(x_1) \leq T(x_2)$ implies the existence of some $\vartheta \in \Theta$ with $P_\vartheta(T \leq T(x_1)) \leq 1 - \beta < P_\vartheta(T \leq T(x_2))$.
11. Let $\mathcal{P}$ be injective and invariant (invariance of models is defined in Definition A.3.30) over a transformation group $\mathcal{G}$ (transformation groups are considered in Definition A.3.27) on the sample space $\mathcal{X}$. The Buehlerization $R_{T,\beta}$ of T is then equivariant over $\mathcal{G}$ (equivariance of parameters of interest is defined in Definition A.3.33) if, and only if, the following equivalence holds:
$$P_\vartheta(T > T(x)) < \beta \quad \Longleftrightarrow \quad P_{\overline{g}(\vartheta)}(T > T(g(x))) < \beta \quad \text{for } x \in \mathcal{X} \text{ and } g \in \mathcal{G}.$$
12. Buehlerizations $R_{T,\beta}$ can be biased. In fact, the Buehlerization of the identity $\mathrm{id}_{\{0,1\}}$ in the Bernoulli model from Section 3.3 is given by
$$R_{\mathrm{id},\beta}(x) = \begin{cases} [0, \beta[& \text{if } x = 0 \\ [0, 1] & \text{if } x = 1, \end{cases}$$
which yields for $p, p' \in [0, 1]$ the coverage probability
$$\mathrm{B}_p(R_{\mathrm{id},\beta} \ni p') = \begin{cases} p & \text{if } p' \in [\beta, 1] \\ 1 & \text{if } p' \in [0, \beta[. \end{cases}$$
We thus obtain $\inf_{p \in [0,1]} \mathrm{B}_p(R_{\mathrm{id},\beta} \ni p') = 1$ if $p' < \beta$.
13. The following observation can be used to verify the measurability requirement for confidence regions in parts (ii) and (iii) of the next theorem. Let us endow $\mathcal{X}$ with the total preorder $\leq_T$ induced by T (defined in Lemma A.1.44, part 1). If $\mathfrak{A}$ contains all downrays in $\mathcal{X}$, then $\{\kappa[R_{T,\beta}] \ni \gamma\}$, $\{\lceil \kappa[R_{T,\beta}] \rceil \ni \gamma\} \in \mathfrak{A}$ for $\gamma \in \Gamma$. In fact, let $\gamma \in \Gamma$ and $x \in \mathcal{X}^2$ with $x_1 \leq_T x_2$. Due to Remark 1.2.3, part 7, $\kappa[R_{T,\beta}(x_1)] \ni \gamma$ implies $\kappa[R_{T,\beta}(x_2)] \ni \gamma$, and hence $\lceil \kappa[R_{T,\beta}(x_1)] \rceil \ni \gamma$ implies $\lceil \kappa[R_{T,\beta}(x_2)] \rceil \ni \gamma$. Thus, $\{\kappa[R_{T,\beta}] \ni \gamma\}$ and $\{\lceil \kappa[R_{T,\beta}] \rceil \ni \gamma\}$ are uprays in $\mathcal{X}$, and, due to Remark A.1.35, part 1, members of $\mathfrak{A}$.
14. Strictly monotonic transformations of $\overline{\mathbb{R}}$-valued designated statistics are easily expressed in terms of the original Buehlerization: Let
    - $\mathcal{P} = (P_\vartheta : \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$,
    - $T: \mathcal{X} \to \overline{\mathbb{R}}$ a statistic,
    - $f: \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ a strictly monotonic function,
    - $\beta \in [0, 1]$.

    Then
$$R_{f \circ T,\beta} = \begin{cases} R_{T,\beta} & \text{if } f \text{ is strictly increasing} \\ \tilde{R}_{T,\beta} & \text{if } f \text{ is strictly decreasing.} \end{cases}$$
15. Buehlerizations of designated statistics that are monotonic in each other are ordered: Let

- $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$,
- $T, S \colon \mathcal{X} \to \overline{\mathbb{R}}$ two statistics,
- $\beta \in [0, 1]$.

Remark A.1.46, parts 4 and 8, then implies the following:

(i) If T is increasing in S, then $R_{S,\beta}(x) \subseteq R_{T,\beta}(x)$ and $\tilde{R}_{S,\beta}(x) \subseteq \tilde{R}_{T,\beta}(x)$ for $x \in \mathcal{X}$.

(ii) If T is decreasing in S, then $R_{S,\beta}(x) \subseteq \tilde{R}_{T,\beta}(x)$ and $\tilde{R}_{S,\beta}(x) \subseteq R_{T,\beta}(x)$ for $x \in \mathcal{X}$.

16. In Chapters 2–4 we buehlerize merely point estimators since this simplifies computations and sometimes allows a representation of the resulting confidence regions in a closed form. Buehler (1957) suggested the Buehlerization of confidence bounds. The trend in most of the applications nowadays, however, is towards the Buehlerization of approximate confidence bounds as these seemingly promise less conservatism. A truly systematic study as to the choice of the designated statistic in specific situations is still missing.

1.2.4 Theorem (Optimality of Buehlerizations) *Let*

- $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ *be a model on a measurable space* $(\mathcal{X}, \mathfrak{A})$,
- $(\mathcal{Y}, \leq)$ *a separable totally preordered set,*
- $T \colon \mathcal{X} \to \mathcal{Y}$ *a statistic,*
- $\beta \in [0, 1]$.

Then the following holds:

(i) $R_{T,\beta}$ *is the least confidence region for* id_Θ *with level* β *that is increasing in* T*, i.e., if* R *is any confidence region in the model* $\mathcal{P}$ *for* id_Θ *with level* β *that is increasing in* T*, then* $R_{T,\beta}(x) \subseteq R(x)$ *for* $x \in \mathcal{X}$.

(ii) *If* $\kappa \colon \Theta \to \Gamma$ *is a parameter of interest for* $\mathcal{P}$ *and* $\{\kappa[R_{T,\beta}] \ni \gamma\} \in \mathfrak{A}$ *for* $\gamma \in \Gamma$*, then, analogously,* $\kappa[R_{T,\beta}]$ *is the least confidence region for* κ *with level* β *that is increasing in* T.

(iii) *If* $\kappa \colon \Theta \to \Gamma$ *is a parameter of interest for* $\mathcal{P}$*,* Γ *preordered, and* $\{\lceil \kappa[R_{T,\beta}] \rceil \ni \gamma\} \in \mathfrak{A}$ *for* $\gamma \in \Gamma$*, then, analogously,* $\lceil \kappa[R_{T,\beta}] \rceil$ *is the least confidence downray for* κ *with level* β *that is increasing in* T*. (*$\lceil \cdot \rceil$ *is defined in Remark A.1.35, part 4.)*

Proof. (i) follows from part (ii) since $\{R_{T,\beta} \ni \vartheta\} \in \mathfrak{A}$ for $\vartheta \in \Theta$ due to Theorem 1.1.6.

(ii) Remark 1.2.3, part 7, and Theorem A.3.21, part (i), yield that $\kappa[R_{T,\beta}]$ is a confidence region for κ with level β that is increasing in T. Let now K be a confidence region for κ with level β that is increasing in T, and let us assume the existence of some $x \in \mathcal{X}$ with $\kappa[R_{T,\beta}(x)] \nsubseteq \mathrm{K}(x)$. Let us pick $\vartheta \in R_{T,\beta}(x)$ with $\kappa(\vartheta) \notin \mathrm{K}(x)$. As K is increasing in T, we obtain $\{T \leq T(x)\} \subseteq \{\mathrm{K} \not\ni \kappa(\vartheta)\}$. Since $\vartheta \in R_{T,\beta}(x)$, we obtain

$$P_\vartheta(\mathrm{K} \not\ni \kappa(\vartheta)) \geq P_\vartheta(T \leq T(x)) > 1 - \beta,$$

which yields $\beta_{\mathrm{eff}}(\mathrm{K}) < \beta$, contradicting the confidence property of K.

(iii) Part (ii), Remark A.1.35, part 4, and $\lceil G \rceil \supseteq G$ for $G \in 2^\Gamma$ yield that $\lceil \kappa[R_{T,\beta}] \rceil$ is a confidence downray for κ with level β that is increasing in T. Let now K be a confidence downray for κ with level β that is increasing in T, and let us assume the existence of some $x \in \mathcal{X}$ with $\lceil \kappa[R_{T,\beta}(x)] \rceil \nsubseteq \mathrm{K}(x)$. Remark A.1.35, part 5, yields $\kappa[R_{T,\beta}(x)] \nsubseteq \mathrm{K}(x)$. The proof now proceeds exactly as in part (ii). □

1.2.5 Remark (Optimality and admissibility) 1. Theorem 1.2.4 not only states minimality, but optimality of $\kappa[R_{T,\beta}]$.

2. Remark 1.2.3, part 3, yields the following result: $\tilde{R}_{T,\beta}$ is the least confidence region for id_Θ with level β that is decreasing in T, and, if $\{\kappa[R_{T,\beta}] \ni \gamma\} \in \mathfrak{A}$ for $\gamma \in \Gamma$, then $\kappa[\tilde{R}_{T,\beta}]$ is the least confidence region for κ with level β that is decreasing in T.
3. Let $\mathcal{K}$ be a set of confidence regions for κ admitting a least element $\min \mathcal{K}$. Then

$$(\min \mathcal{K})(x) = \bigcap_{\mathrm{K} \in \mathcal{K}} \mathrm{K}(x) \qquad \text{for } x \in \mathcal{X}.$$

Most statisticians usually call "$\mathcal{K}$-admissible" and "$\mathcal{K}$-optimal" (and sometimes append Buehler's name) what we call "minimal in $\mathcal{K}$" and "least in $\mathcal{K}$," respectively. We believe this terminology to be clearer to readers from other fields since the order $\subseteq$ naturally occurs when investigating confidence regions in all generality.
4. $R_{T,\beta}$ need not be minimal in the set of all confidence regions in the model $\mathcal{P}$ for id_Θ with level β; in fact, Buehlerizations of $\mathcal{P}$-a.s. constant statistics are trivial.
5. The assumption of totality of the preorder on $\mathcal{Y}$ cannot be weakened considerably: If $\mathcal{X} := \mathcal{Y} := \mathbb{R}^2$ are equipped with the product order and $\mathcal{P} := (\mathrm{N}_{\mu,1}^{\otimes 2} \colon \mu \in \mathbb{R})$, then $R_{\mathrm{id},\beta}(x) \ni \mu$ is equivalent to $(1 - \Phi(x_1 - \mu))(1 - \Phi(x_2 - \mu)) < \beta$ for $x \in \mathbb{R}^2$ and $\mu \in \mathbb{R}$, hence the set-valued sequence $(\{R_{\mathrm{id},\beta} \ni \mu\} \colon \mu \in \mathbb{N})$ is decreasing with limit $\bigcap_{\mu \in \mathbb{N}} \{R_{\mathrm{id},\beta} \ni \mu\} = \emptyset$, yielding $\lim_{\mu \to \infty} \mathrm{N}_{\mu,1}^{\otimes 2}(R_{\mathrm{id},\beta} \ni \mu) = 0$.
6. The next theorem gives sufficient conditions for a Buehlerization to be a confidence down- or upray.

1.2.6 Theorem (Buehlerizations and down-/uprays) *Let*

- $(\Theta, \leq)$, $(\mathcal{X}, \leq)$, *and* $(\Gamma, \leq)$ *be preordered sets,*
- $(\mathcal{Y}, \leq)$ *a separable totally preordered set,*
- $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ *a stochastically monotonic model on* $\mathcal{X}$ *(see Remark A.1.69),*
- $\kappa \colon \Theta \to \Gamma$ *a monotonic parameter of interest for* $\mathcal{P}$,
- $T \colon \mathcal{X} \to \mathcal{Y}$ *a monotonic statistic,*
- $\beta \in [0, 1]$.

Then the following holds:

(i) *Let* T *and* $\mathcal{P}$ *be both increasing or both decreasing. Then* $R_{T,\beta}$ *is a confidence downray and* $\tilde{R}_{T,\beta}$ *a confidence upray for* id_Θ *with level* β. *Furthermore,*
   - *if* $\kappa[\mathrm{id}_\Theta \leq \vartheta] = \{\mathrm{id}_\Gamma \leq \kappa(\vartheta)\}$ *for* $\vartheta \in \Theta$, *then* $\kappa[R_{T,\beta}]$ *is a confidence downray and* $\kappa[\tilde{R}_{T,\beta}]$ *a confidence upray for* κ *with level* β*;*
   - *if* $\kappa[\mathrm{id}_\Theta \leq \vartheta] = \{\mathrm{id}_\Gamma \geq \kappa(\vartheta)\}$ *for* $\vartheta \in \Theta$, *then* $\kappa[R_{T,\beta}]$ *is a confidence upray and* $\kappa[\tilde{R}_{T,\beta}]$ *a confidence downray for* κ *with level* β.

(ii) *Let* T *be increasing and* $\mathcal{P}$ *decreasing or vice versa. Then the conclusions in* (i) *hold with "downray" and "upray" interchanged.*

Proof. Let us assume T increasing, $\mathcal{P}$ stochastically increasing, and $x \in \mathcal{X}$. The downray property of $R_{T,\beta}(x)$ in Θ follows from $P_{\vartheta_1}(T > T(x)) \leq P_{\vartheta_2}(T > T(x))$ for $\vartheta_1, \vartheta_2 \in \Theta$ with $\vartheta_1 \leq \vartheta_2$, by Theorem A.3.48. The rest now follows from Theorem 1.1.6 and by considering all combinations of preorders and their duals on Θ and $\mathcal{Y}$, while Lemma A.1.53, parts (iii) and (iv), yields the "furthermore" claims. $\square$

1.2.7 Remark (Buehlerized confidence bounds) 1. The condition "$\kappa[\mathrm{id}_\Theta \le \vartheta] = \{\mathrm{id}_\Gamma \le \kappa(\vartheta)\}$ for $\vartheta \in \Theta$" in the "furthermore" claim is stronger than κ simply being increasing; analogously for the second condition and "decreasing." In general, surjectivity of κ does not suffice for its validity either.

2. In case Γ is furthermore complete (completeness of preordered sets is defined in Definition A.1.39), the following implications hold for $x \in \mathcal{X}$:

$$\mathrm{K}(x) \text{ downray in } \Gamma \quad\Longrightarrow\quad \mathrm{K}(x) = \begin{cases} \{\mathrm{id}_\Gamma \le \max \mathrm{K}(x)\} & \text{if } \max \mathrm{K}(x) \text{ exists} \\ \{\mathrm{id}_\Gamma < \sup \mathrm{K}(x)\} & \text{otherwise} \end{cases}$$

$$\mathrm{K}(x) \text{ upray in } \Gamma \quad\Longrightarrow\quad \mathrm{K}(x) = \begin{cases} \{\mathrm{id}_\Gamma \ge \min \mathrm{K}(x)\} & \text{if } \min \mathrm{K}(x) \text{ exists} \\ \{\mathrm{id}_\Gamma > \inf \mathrm{K}(x)\} & \text{otherwise.} \end{cases}$$

Many statisticians are not interested in whether the boundaries $\inf \mathrm{K}(x)$ and $\sup \mathrm{K}(x)$ are contained in the confidence region or not, and thus simply consider the confidence bounds

$$\underline{\kappa}\colon \mathcal{X} \to \Gamma, \quad x \mapsto \inf \mathrm{K}(x), \qquad \text{and} \qquad \overline{\kappa}\colon \mathcal{X} \to \Gamma, \quad x \mapsto \sup \mathrm{K}(x).$$

3. If, in the just considered situation, we set $\overline{\kappa}_{T,\beta} \coloneqq \sup \kappa[R_{T,\beta}]$, and if $(\Gamma, \le) = (\mathcal{Y}, \le)$, an immediate question is whether we gain something by buehlerizing the Buehlerization of T or not. In other words, does $\overline{\kappa}_{\overline{\kappa}_{T,\beta},\beta} \le \overline{\kappa}_{T,\beta}$ hold? This is not the case, in general. Remark 1.2.5, part 1, however, states the validity of the reverse inequality: Since $\overline{\kappa}_{T,\beta}$ is increasing in T, we have $R_{T,\beta}(x) \subseteq R_{\overline{\kappa}_{T,\beta},\beta}(x)$, and thus $\overline{\kappa}_{T,\beta}(x) \le \overline{\kappa}_{\overline{\kappa}_{T,\beta},\beta}(x)$ for $x \in \mathcal{X}$.

4. The next example yields Buehlerizations of maxima and minima of several designated statistics in product experiments.

1.2.8 Example (Product experiments) Let

- $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$,
- $n \in \mathbb{N}$, $m \in \mathbb{N}^n$, $|m| \coloneqq \sum_{i=1}^n m_i$, and $\mathcal{P}_m \coloneqq (\bigotimes_{i=1}^n P_{\vartheta_i}^{\otimes m_i} \colon \vartheta \in \Theta^n)$,
- $(\mathcal{Y}, \le)$ a separable totally preordered lattice,
- $T_i \colon \mathcal{X}^{m_i} \to \mathcal{Y}$ a statistic for $i \in \{1, \dots, n\}$,
- $\beta \in [0, 1]$.

We are interested in the experiment $\mathcal{P}_m$. Its sample space is $\prod_{i=1}^n \mathcal{X}^{m_i}$, its parameter space Θ^n. Let us interpret an observation x as an n-tuple $(x_1, \dots, x_n)$ of vectors $x_i \in \mathcal{X}^{m_i}$ of possibly different lengths with components $x_{i,1}, \dots, x_{i,m_i} \in \mathcal{X}$. Let us furthermore define the projections

$$\mathrm{pr}_k \colon \prod_{i=1}^n \mathcal{X}^{m_i} \to \mathcal{X}^{m_k}, \quad x \mapsto x_k, \qquad\qquad \mathrm{pr}_{k,l} \colon \prod_{i=1}^n \mathcal{X}^{m_i} \to \mathcal{X}, \quad x \mapsto x_{k,l},$$

for $k \in \{1, \dots, n\}$ and $l \in \{1, \dots, m_k\}$. Since

$$\bigotimes_{i=1}^n P_{\vartheta_i}^{\otimes m_i}\Big(\textstyle\bigvee_{k=1}^n (T_k \circ \mathrm{pr}_k) \begin{Bmatrix} < \\ \le \end{Bmatrix} \bigvee_{k=1}^n T_k(x_k)\Big) = \displaystyle\prod_{i=1}^n P_{\vartheta_i}^{\otimes m_i}\Big(T_i \begin{Bmatrix} < \\ \le \end{Bmatrix} \textstyle\bigvee_{k=1}^n T_k(x_k)\Big)$$

$$\bigotimes_{i=1}^n P_{\vartheta_i}^{\otimes m_i}\Big(\textstyle\bigwedge_{k=1}^n (T_k \circ \mathrm{pr}_k) \begin{Bmatrix} > \\ \ge \end{Bmatrix} \bigwedge_{k=1}^n T_k(x_k)\Big) = \displaystyle\prod_{i=1}^n P_{\vartheta_i}^{\otimes m_i}\Big(T_i \begin{Bmatrix} > \\ \ge \end{Bmatrix} \textstyle\bigwedge_{k=1}^n T_k(x_k)\Big)$$

for $\vartheta \in \Theta^n$ and $x \in \prod_{i=1}^n \mathcal{X}^{m_i}$, buehlerizing

- $\bigvee_i (T_i \circ \mathrm{pr}_i)\colon \prod_{i=1}^n \mathcal{X}^{m_i} \to \mathcal{Y}$, $x \mapsto \bigvee_{i=1}^n T_i(x_i)$, yields

$$R_{\bigvee_i (T_i\circ \mathrm{pr}_i),\beta}(x) = \left\{\vartheta \in \Theta^n\colon \prod_{i=1}^n P_{\vartheta_i}^{\otimes m_i}\left(T_i \le \textstyle\bigvee_{k=1}^n T_k(x_k)\right) > 1-\beta\right\}$$

$$\tilde{R}_{\bigvee_i (T_i\circ \mathrm{pr}_i),\beta}(x) = \left\{\vartheta \in \Theta^n\colon \prod_{i=1}^n P_{\vartheta_i}^{\otimes m_i}\left(T_i < \textstyle\bigvee_{k=1}^n T_k(x_k)\right) < \beta\right\}$$

- $\bigwedge_i (T_i \circ \mathrm{pr}_i)\colon \prod_{i=1}^n \mathcal{X}^{m_i} \to \mathcal{Y}$, $x \mapsto \bigwedge_{i=1}^n T_i(x_i)$, yields

$$R_{\bigwedge_i (T_i\circ \mathrm{pr}_i),\beta}(x) = \left\{\vartheta \in \Theta^n\colon \prod_{i=1}^n P_{\vartheta_i}^{\otimes m_i}\left(T_i > \textstyle\bigwedge_{k=1}^n T_k(x_k)\right) < \beta\right\}$$

$$\tilde{R}_{\bigwedge_i (T_i\circ \mathrm{pr}_i),\beta} = \left\{\vartheta \in \Theta^n\colon \prod_{i=1}^n P_{\vartheta_i}^{\otimes m_i}\left(T_i \ge \textstyle\bigwedge_{k=1}^n T_k(x_k)\right) > 1-\beta\right\}$$

for $x \in \prod_{i=1}^n \mathcal{X}^{m_i}$. In particular, if $T\colon \mathcal{X} \to \mathcal{Y}$ is a statistic and $T_i(x_i) = \bigvee_{j=1}^{m_i} T(x_{i,j})$ or $T_i(x_i) = \bigwedge_{j=1}^{m_i} T(x_{i,j})$ for $x_i \in \mathcal{X}^{m_i}$ and $i \in \{1,\dots,n\}$, then the above also yields the Buehlerizations of $\bigvee_{i=1}^n \bigvee_{j=1}^{m_i} (T\circ \mathrm{pr}_{i,j})$, $\bigvee_{i=1}^n \bigwedge_{j=1}^{m_i} (T\circ \mathrm{pr}_{i,j})$, $\bigwedge_{i=1}^n \bigvee_{j=1}^{m_i} (T\circ \mathrm{pr}_{i,j})$, and $\bigwedge_{i=1}^n \bigwedge_{j=1}^{m_i} (T\circ \mathrm{pr}_{i,j})$.

1.3 NOTES

Buehler confidence bounds have been established well before Buehler's seminal work from 1957. An example are the renowned confidence bounds of Clopper and Pearson (1934) with confidence level $\beta \in [0,1]$ in the binomial model

$$(\mathrm{B}_{n,p}\colon p \in [0,1])$$

(with fixed $n \in \mathbb{N}$) for the parameter of interest $\mathrm{id}_{[0,1]}$. The upper bound is given by

$$u_{\mathrm{CP},\beta}\colon \{0,\dots,n\} \to [0,1], \quad x \mapsto \sup\{p \in [0,1]\colon \mathrm{B}_{n,p}(\{0,\dots,x\}) > 1-\beta\}.$$

Since, for fixed $x \in \{0,\dots,n-1\}$, the function

$$f\colon [0,1] \to [0,1], \quad p \mapsto \mathrm{B}_{n,p}(\{0,\dots,x\}),$$

is continuous (being a polynomial function), strictly decreasing (due to $f'(p) = -n\mathrm{b}_{n-1,p}(x)$), and surjective (due to $f(0) = 1$, $f(1) = 0$, and the intermediate value theorem), $u_{\mathrm{CP},\beta}(x)$ can also be thought of as the unique $p \in [0,1]$ such that $\mathrm{B}_{n,p}(\{0,\dots,x\}) = 1-\beta$ in case $x < n$, while $u_{\mathrm{CP},\beta}(n) = 1$.

The lower confidence bound $l_{\mathrm{CP},\beta}$ with level β can be obtained by replacing each occurrence of $\mathrm{B}_{n,p}(\{0,\dots,x\})$ and "sup" with $\mathrm{B}_{n,p}(\{x,\dots,n\})$ and "inf," respectively:

$$l_{\mathrm{CP},\beta}\colon \{0,\dots,n\} \to [0,1], \quad x \mapsto \inf\{p \in [0,1]\colon \mathrm{B}_{n,p}(\{x,\dots,n\}) > 1-\beta\}.$$

Since, for fixed $x \in \{1,\dots,n\}$, the function

$$g\colon [0,1] \to [0,1], \quad p \mapsto \mathrm{B}_{n,p}(\{x,\dots,n\}),$$

is continuous (again being a polynomial function), strictly increasing (due to $g'(p) = n\mathrm{b}_{n-1,p}(x-1)$), and surjective (due to $g(0) = 0$, $g(1) = 1$, and the intermediate value theorem), $l_{\mathrm{CP},\beta}(x)$ can also be thought of as the unique $p \in [0,1]$ such that $\mathrm{B}_{n,p}(\{x, \ldots, n\}) = 1 - \beta$ in case $x > 0$, while $l_{\mathrm{CP},\beta}(0) = 0$.

Together with the identity $\mathrm{b}_{n,p}(x) = \mathrm{b}_{n,1-p}(n-x)$, this yields the relation

$$l_{\mathrm{CP},\beta}(x) = 1 - u_{\mathrm{CP},\beta}(n-x), \tag{1}$$

a property closely connected to *equivariance* (see Example A.3.34, part 2, for equivariance in this binomial model). It is occasionally also used for defining the lower bound in terms of the upper (or vice versa).

In their original paper Clopper and Pearson did not concentrate as much on the one-sided setting (i.e., on confidence bounds) as on the two-sided situation (i.e., on confidence intervals of the form $]l_{\mathrm{CP},(1+\beta)/2}, u_{\mathrm{CP},(1+\beta)/2}[$). This may be the reason why their work does not mention the monotonicity of either confidence bound. The nesting property (i.e., $[0,1] \to [0,1]$, $\beta \mapsto l_{\mathrm{CP},\beta}(x)$, is decreasing and $[0,1] \to [0,1]$, $\beta \mapsto u_{\mathrm{CP},\beta}(x)$, increasing for fixed $x \in \{0, \ldots, n\}$) is not brought up either, which, however, is more likely to be due to the simplicity in this particular case.

If we append the sample size n to the Clopper and Pearson confidence bounds in order to emphasize their dependence on the model $(\mathrm{B}_{n,p}\colon p \in [0,1])$, then the monotonicity of the above functions f and g and the monotonicity of $f(p)$ and $g(p)$ with respect to n for $p \in [0,1]$ yield that $u_{\mathrm{CP},\beta,n}$ and $l_{\mathrm{CP},\beta,n}$ are decreasing in $n \in \mathbb{N}$.

Sterne (1954) proposed the confidence region

$$R^{\mathrm{S}}\colon \{0, \ldots, n\} \to 2^{[0,1]}, \quad x \mapsto \{p \in [0,1]\colon \mathrm{B}_{n,p}(\mathrm{b}_{n,p} \le \mathrm{b}_{n,p}(x)) \ge 1 - \beta\}.$$

Dümbgen (2004) notes that $R^{\mathrm{S}}(x)$ is not always an interval. Indeed, for $\beta \coloneqq 0.928$, $n \coloneqq 10$, $x \coloneqq 0$, and $(p, r, q) \coloneqq (0.25, 0.27, 0.29)$ we obtain using R, version 3.0.2,

$$\begin{aligned}
&\mathrm{B}_{n,p}(\mathrm{b}_{n,p} \le \mathrm{b}_{n,p}(x)) \approx 0.0760 > 1 - \beta, \\
&\mathrm{B}_{n,r}(\mathrm{b}_{n,r} \le \mathrm{b}_{n,r}(x)) \approx 0.0717 < 1 - \beta, \\
&\mathrm{B}_{n,q}(\mathrm{b}_{n,q} \le \mathrm{b}_{n,q}(x)) \approx 0.0729 > 1 - \beta,
\end{aligned}$$

hence $p, q \in R^{\mathrm{S}}(x)$, but $r \notin R^{\mathrm{S}}(x)$. (Similar examples can be constructed for almost every $n \in \mathbb{N}$ and for β's in certain sets having 1 as an accumulation point.) This led Sterne to the consideration of the confidence interval $[\min R^{\mathrm{S}}, \max R^{\mathrm{S}}]$ (the occurring minimum and maximum exist since $R^{\mathrm{S}}(x)$ is a closed set; in fact, $[0,1] \to [0,1]$, $p \mapsto \mathrm{B}_{n,p}(\mathrm{b}_{n,p} \le \mathrm{b}_{n,p}(x))$, is upper semicontinuous, for which Sterne fails to provide an argument). Sterne (1954) argues in favor of his confidence intervals over the ones given by Clopper and Pearson (1934) regarding their size at the extremal observations 0 and n and their coverage probabilities.

Crow (1956) proposes a modification of Sterne's confidence interval, and shows that both Sterne's and his confidence intervals have minimal total length by proving that inverting a family of tests (using Theorem A.3.39) with acceptance intervals of minimal length yields a confidence interval with minimal total length.

Blyth and Still (1983) provide equivariant confidence intervals for the identity $\mathrm{id}_{[0,1]}$ in the classical binomial model $(\mathrm{B}_{n,p}\colon p \in [0,1])$ whose endpoints are increasing in the observation for fixed n, and decreasing in n for a fixed observation. They list the three possibilities that can occur regarding uniqueness and total interval length, and classify the confidence interval from Clopper and Pearson (1934) in this list. Newcombe (1998) and Brown *et al.* (2001) compare several confidence intervals for the success probability in binomial samples. Agresti and Coull (1998) argue for the usage of approximate confidence intervals in terms of the behaviour of the coverage probabilities.

Buehler (1957) established the formula for $\sup R_{T,\beta}$ and gave tables for upper confidence bounds in the model

$$(\mathrm{B}_{n_1,p_1} \otimes \mathrm{B}_{n_2,p_2} \colon (p_1, p_1) \in [0,1]^2)$$

for the parameter of interest $(p_1, p_2) \mapsto p_1 p_2$. His tables are based on a Poisson approximation, which makes the values usable for a whole range of sample sizes n_1 and n_2. Buehler used the designated statistic

$$T \colon \{0,\ldots,n_1\} \times \{0,\ldots,n_2\} \to [0,1], \quad (x_1,x_2) \mapsto u_{\mathrm{CP},\sqrt{\beta}}(x_1)u_{\mathrm{CP},\sqrt{\beta}}(x_2),$$

where $u_{\mathrm{CP},\beta}$ denotes the upper confidence bound with level β from Clopper and Pearson. At the end of his paper, Buehler (1957) generalizes his method to arbitrary discrete models (the discreteness condition can, however, also be dropped) with existing confidence bounds as designated statistics.

Sudakov (1974) provides upper and lower confidence bounds for the parameter of interest $[0,1]^n \to [0,1]$, $p \mapsto \prod_{i=1}^n p_i$, in the binomial model $(\prod_{i=1}^n \mathrm{B}_{m_i,p_i} \colon p \in [0,1]^n)$ considered in Section 3.1 of Chapter 3. This is motivated by interpreting the model as a representation of a sequential system with n independent components each of which possesses a reliability p_i (i.e., a probability of failure $1 - p_i$) and is tested in m_i trials; the above parameter of interest then expresses the reliability index, i.e., the probability of non-failure of the entire system. His confidence bounds arise by buehlerizing the maximum likelihood estimator $\prod_{i=1}^n \{0,\ldots,m_i\} \to [0,1]$, $x \mapsto \prod_{i=1}^n x_i/m_i$. According to Lloyd and Kabaila (2003), these results are obtained independently of the work by Buehler (1957).

Winterbottom (1984) summarizes some of the methods, among them Buehler's (1957) and some Bayesian ones, that have been applied on the general problem of finding lower confidence bounds for the reliability index of a system consisting of multiple components. The test data considered do not necessarily follow a binomial distribution, nor are other properties of the underlying system assumed.

Harris and Soms (1991) prove some of the results from Buehler (1957) and Sudakov (1974) in a more general setting by the inversion of families of tests, and disproves an inequality from Sudakov (1974) involving the incomplete beta function.

Reiser and Jaeger (1991) consider a two-component series system with binomially distributed test data and buehlerize the maximum likelihood estimator. They illustrate the anomaly (due, according to them, to the discreteness of the model) that additional success results can result in a decrease of the lower confidence bound. A similar peculiarity of Buehlerizations, the existence of so-called "ties," is remarked by Harris and Soms (1983) and investigated by Kabaila and Lloyd (2003) and, in more detail, Kabaila and Lloyd (2006).

Revyakov (1992) reformulates and generalizes some of the results developed in Buehler (1957), and applies them to a number of reliability problems. Jobe and David (1992) prove, among other things, those fundamental results from Buehler's theory in greater generality for the first time.

Pfanzagl (1994, Theorems 5.3.3, p. 167) presents, under certain continuity assumptions, upper confidence bounds which possess a certain similarity to Buehlerizations and are randomized subsequently.

Bagdonavičius *et al.* (1997) presents a similar result to our Lemma 1.1.2, which is then used to establish a theorem much alike Buehler's (1957) main result. Both results are due to Bolshev (1965), who seems unaware of Buehler's (1957). The second part illustrates the results with examples using Poisson, exponential, Bernoulli, geometric, normal, and some other samples.

Kabaila and Lloyd (1997) provide a theory of Buehler confidence bounds (similar in structure to what is presented here) for discrete models with approximate confidence bounds as designated statistics. Their approach is slightly different from Buehler's (1957), and their results rely on the validity of a few supplementary regularity conditions, which, however, also allow statements on coverage probabilities for certain parameter values. The parameters of interest considered are real-valued, the parameter space is a finite-dimensional vector space. Since the resulting Buehlerization is least subject to the conditions of 1. sustaining a prescripted confidence level, and 2. being ordered the same way as the approximate confidence bound they start with, Buehler bounds are called "tight." The authors argue heuristically in favor of employing approximate confidence limits instead of estimators as designated statistics, an issue taken up more formally in Kabaila (2001), in Kabaila and Lloyd (2002), where approximate confidence bounds based on the likelihood ratio statistic are recommended as designated statistics, in Kabaila and Lloyd (2003), and in Kabaila and Lloyd (2004), where consequences on the nestedness of Buehlerizations by a possible dependence of the designated statistic on the confidence level are investigated.

Lloyd and Moldovan (2000) employ Buehler confidence bounds in a medical context to investigate the difference between two correlated proportions. According to them, considering confidence bounds rather than two-sided confidence intervals yields less conservative confidence statements.

Kabaila and Lloyd (2000) show that smallest upper and greatest lower confidence bounds with prescribed confidence level exist only in some trivial or unusual models (like the Bernoulli or the translated symmetric Bernoulli models from Sections 3.3 and 3.4 of Chapter 3, respectively). In their main result, they establish an assumption that implies nonexistence of such a confidence bound, and verify that assumption in the binomial models of sample size not less than two, and in models consisting of two independent binomial samples of equal size with the difference of their success probabilities as parameter of interest. Looking for "best" confidence bounds makes thus sense after a restricting somehow the class of considered confidence bounds—Buehler (1957) does so by adding the requirement of monotonicity.

Lloyd and Kabaila (2003) prove the optimality of Buehler bounds in more generality and claim that a modification is in order when a certain set inside a supremum is empty—a problem that only arises if sup is, contrary to common practice, not defined as the least element in the underlying completely ordered set. They furthermore show that Buehlerization in general linear models yields trivial confidence bounds.

Wang (2006) rediscovers Buehler's (1975) method, restricted to the classical binomial model and citing only works by Bolshev and Nikulin, and establishes a condition allowing to conclude whether or not confidence intervals are least among the ones having increasing end-points that satisfy the property (1). This is used to investigate for which confidence levels the equal-tailed two-sided confidence interval by Clopper and Pearson (1934) is least in that class. This line of work is continued in Wang (2010), where a more generalized setting is considered, but still restating for the greater part results from already published works, as pointed out by Lloyd and Kabaila (2010).

Applications of Buehler's theory to a medical setting can be found in, e.g., Lloyd and Moldovan (2000), Lloyd and Moldovan (2007), and Lloyd (2015).

2 APPLICATION: NORMAL SAMPLES

- In this chapter and the next ones, the set $\mathcal{Y}$ from Definition 1.1.5 is mostly $\mathbb{R}$ or a subset thereof, the order being the usual. Its interval topology is thus the usual Euclidean topology and $\mathfrak{B}$ is the Borel σ-algebra, as agreed in Remark A.1.69, parts 2 and 4.
- Let us remember that $\beta \in]0,1[$ (in view of Remarks 1.1.7, part 3, and 1.2.3, part 6), unless stated otherwise.
- Outline of this chapter:
  - Section 2.1 considers several normal samples with unknown means and variances. This is the most general model considered in this chapter. We buehlerize minimum and maximum of the different samples' means (beginning with Example 2.1.3), minimum and maximum of the sample mean divided by the sample standard deviation (beginning with Example 2.1.8), overall minimum and maximum (beginning with Example 2.1.13), and the sample mean of the different samples' means (beginning with Example 2.1.18).
  - Section 2.2 deals with several normal samples with unknown means but known variances. This is a submodel of the model from the previous section, which means that some results can be taken over with just a few adjustments. We determine the Buehlerization of minimum and maximum of the different samples' means (beginning with Example 2.2.3), overall minimum and maximum (beginning with Example 2.2.8), and the sample mean of the different samples' means (beginning with Example 2.2.13).
  - Section 2.3 considers several normal samples with unknown but equal means and known variances. This is a submodel of the model from the previous section. We calculate the Buehlerization of the sample mean of the different samples' means (Example 2.3.3).
  - Section 2.4 treats several normal samples with known means but unknown variances. This is a submodel of the general model from Section 2.1. We determine the Buehlerization of minimum and maximum of the different samples' variances and sample variances.
  - Section 2.5 treats several normal samples with known means and unknown but equal variances. This is a submodel of the model from the previous section. The Buehlerization of the designated statistics from the previous section are derived from the preceding results.

2.1 SEVERAL SAMPLES

2.1.1 Definition Let $n \in \mathbb{N}$ and $m \in \mathbb{N}^n$. Let us consider the n-sample normal model

$$\mathcal{P}_1 := \left(\bigotimes_{i=1}^{n} \mathrm{N}_{\mu_i,\sigma_i^2}^{\otimes m_i} : (\mu,\sigma) \in \mathbb{R}^n \times]0,\infty[^n \right)$$

with known individual sample sizes $m_1, \ldots, m_n$.

2.1.2 Remark The sample space is $\mathcal{X} = \prod_{i=1}^n \mathbb{R}^{m_i}$, the parameter space $\Theta = \mathbb{R}^n \times]0,\infty[^n$. Let us interpret an observation x as an n-tuple $(x_1,\ldots,x_n)$ of vectors $x_i \in \mathbb{R}^{m_i}$ of possibly different lengths with components $x_{i,1},\ldots,x_{i,m_i}$.

2.1.3 Example Let us consider $\bigwedge_i \overline{X_i} \colon \prod_{i=1}^n \mathbb{R}^{m_i} \to \mathbb{R}$, $x \mapsto \bigwedge_{i=1}^n \overline{x_i}$, as designated statistic, where, for $i \in \{1,\ldots,n\}$,

$$\overline{X_i} \colon \prod_{j=1}^{n} \mathbb{R}^{m_i} \to \mathbb{R}, \quad x \mapsto \overline{x_i} := \frac{1}{m_i} \sum_{j=1}^{m_i} x_{i,j},$$

denotes the arithmetic mean of the ith sample. Since $(\mathbb{R}^r \to \mathbb{R},\ x \mapsto \sum_{k=1}^r x_k/r) \,\square\, \mathrm{N}_{\nu,\tau^2}^{\otimes r} = \mathrm{N}_{\nu,\tau^2/r}$ for $r \in \mathbb{N}$, $\nu \in \mathbb{R}$, and $\tau \in]0,\infty[$, we obtain

$$\bigotimes_{i=1}^{n} \mathrm{N}_{\mu_i,\sigma_i^2}^{\otimes m_i} \left(\bigwedge_{i=1}^{n} \overline{X_i} > t \right) = \prod_{i=1}^{n} \Phi\left(\sqrt{m_i} \frac{\mu_i - t}{\sigma_i} \right) \tag{2}$$

for $t \in \mathbb{R}$. Since $(\bigwedge_i \overline{X_i}) \,\square\, \bigotimes_{i=1}^n \mathrm{N}_{\mu_i,\sigma_i^2}^{\otimes m_i}$ is continuous, Remark 1.1.7, part 10, yields the confidence regions given by

$$R_{\bigwedge_i \overline{X_i},\beta}(x) = \left\{ (\mu,\sigma) \in \mathbb{R}^n \times]0,\infty[^n \colon \prod_{i=1}^{n} \Phi\left(\sqrt{m_i} \frac{\mu_i - \bigwedge_{k=1}^n \overline{x_k}}{\sigma_i} \right) < \beta \right\}$$

$$\tilde{R}_{\bigwedge_i \overline{X_i},\beta}(x) = \left\{ (\mu,\sigma) \in \mathbb{R}^n \times]0,\infty[^n \colon \prod_{i=1}^{n} \Phi\left(\sqrt{m_i} \frac{\mu_i - \bigwedge_{k=1}^n \overline{x_k}}{\sigma_i} \right) > 1-\beta \right\}$$

for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$. Since

$$\prod_{i=1}^{n} \Phi\left(\sqrt{m_i} \frac{\mu_i - \bigwedge_{k=1}^n \overline{X_k}}{\sigma_i} \right) \colon \prod_{i=1}^{n} \mathbb{R}^{m_i} \to]0,1[$$

is surjective, the effective levels of $R_{\bigwedge_i \overline{X_i},\beta}$ and $\tilde{R}_{\bigwedge_i \overline{X_i},\beta}$ are given due to Theorem 1.1.6, part (i), by

$$\beta_{\text{eff}}(R_{\bigwedge_i \overline{X_i},\beta}) = \beta_{\text{eff}}(\tilde{R}_{\bigwedge_i \overline{X_i},\beta}) = \beta \qquad \text{for } \beta \in [0,1].$$

2.1.4 Remark 1. We shall use in the proofs of the next few lemmas the fact that

$$\mathbb{R}^n \to]0,1[, \quad \mu \mapsto \prod_{i=1}^{n} \Phi\left(\sqrt{m_i} \frac{\mu_i - t}{\sigma_i} \right),$$

is strictly increasing for $t \in \mathbb{R}$ and $\sigma \in]0,\infty[^n$, and

$$]0,\infty[\to]0,1[, \quad \sigma_k \mapsto \prod_{i=1}^{n} \Phi\left(\sqrt{m_i} \frac{\mu_i - t}{\sigma_i} \right),$$

is, for $k \in \{1,\ldots,n\}$, $t \in \mathbb{R}$ and $\mu \in \mathbb{R}^n$, strictly increasing if $\mu_k < t$, strictly decreasing if $\mu_k > t$, and constant if $\mu_k = t$.

2. The next result yields confidence regions for the following projections as parameters of interest:

$$\mathrm{pr}_1 \colon \mathbb{R}^n \times]0,\infty[^n \to \mathbb{R}^n, \qquad (\mu,\sigma) \mapsto \mu,$$
$$\mathrm{pr}_2 \colon \mathbb{R}^n \times]0,\infty[^n \to]0,\infty[^n, \quad (\mu,\sigma) \mapsto \sigma.$$

2.1.5 Lemma *Let us consider the projections* pr_1 *and* pr_2 *from Remark 2.1.4, part 2. For* $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$ *then*

$$\mathrm{pr}_1[R_{\bigwedge_i \overline{X_i},\beta}(x)] = \begin{cases} \mathbb{R}^n & \text{if } \beta \in]\frac{1}{2^n}, 1] \\ \bigcup_{k=1}^n \{\mu \in \mathbb{R}^n \colon \mu_k < \bigwedge_{i=1}^n \overline{x_i}\} & \text{if } \beta \in]0, \frac{1}{2^n}] \end{cases} \tag{3}$$

$$\mathrm{pr}_1[\tilde{R}_{\bigwedge_i \overline{X_i},\beta}(x)] = \begin{cases} \mathbb{R}^n & \text{if } \beta \in]1-\frac{1}{2^n}, 1] \\ \{\mu \in \mathbb{R}^n \colon |\{\mu > \bigwedge_{i=1}^n \overline{x_i}\}| > n-k\} & \text{if } \beta \in]1-\frac{1}{2^{k-1}}, 1-\frac{1}{2^k}] \\ & \text{for some } k \in \{1,\dots,n\} \end{cases} \tag{4}$$

$$\mathrm{pr}_2[R_{\bigwedge_i \overline{X_i},\beta}(x)] = \mathrm{pr}_2[\tilde{R}_{\bigwedge_i \overline{X_i},\beta}(x)] =]0,\infty[^n. \tag{5}$$

Proof. Let $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$, and let us assume w.l.o.g. $\beta \in]0,1[$ (due to Remarks 1.1.7, part 3, and 1.2.3, part 6).

(3) Let $\beta \in]\frac{1}{2^n}, 1[$. The inclusion LHS $\subseteq$ RHS is clear. It therefore remains to show LHS $\supseteq$ RHS. To this end, let $\mu \in \mathbb{R}^n$, and let us define $t \coloneqq (\bigvee_{k=1}^n \mu_k) \vee (\bigwedge_{k=1}^n \overline{x_k}) + 1$, $M \coloneqq \bigvee_{k=1}^n m_k$, and

$$\sigma_i \coloneqq M \frac{t - \bigwedge_{k=1}^n \overline{x_k}}{\Phi^{-1}(\beta^{1/n})} \qquad \text{for } i \in \{1,\dots,n\}.$$

Then $\sigma \in]0,\infty[^n$ and , implies

$$\prod_{i=1}^n \Phi\left(\sqrt{m_i}\frac{\mu_i - \bigwedge_{k=1}^n \overline{x_k}}{\sigma_i}\right) < \Phi^n\left(M\frac{t - \bigwedge_{k=1}^n \overline{x_k}}{\sigma_1}\right) = \beta$$

due to Remark 2.1.4, part 1. Example 2.1.3 yields $(\mu,\sigma) \in R_{\bigwedge_i \overline{X_i},\beta}(x)$. Since $\mathrm{pr}_1(\mu,\sigma) = \mu$, we obtain $\mu \in$ LHS.

Let now $\beta \in]0, \frac{1}{2^n}]$. We first show LHS $\subseteq$ RHS. To this end, let $(\mu,\sigma) \in R_{\bigwedge_i \overline{X_i},\beta}(x)$, and let us assume $\mu \notin$ RHS. This means $\bigwedge_{i=1}^n \mu_i \geq \bigwedge_{i=1}^n \overline{x_i}$, which implies

$$\prod_{i=1}^n \Phi\left(\sqrt{m_i}\frac{\mu_i - \bigwedge_{k=1}^n \overline{x_k}}{\sigma_i}\right) \geq \Phi^n(0) = \frac{1}{2^n} \geq \beta,$$

due to Remark 2.1.4, part 1, a contradiction to $(\mu,\sigma) \in R_{\bigwedge_i \overline{X_i},\beta}(x)$ in view of Example 2.1.3.

We now show LHS $\supseteq$ RHS. Let $\mu \in$ RHS, and let us pick $i_0 \in \{1,\dots,n\}$ such that $\mu_{i_0} < \bigwedge_{k=1}^n \overline{x_k}$. Let us define $\sigma_i \coloneqq 1$ for $i \in \{1,\dots,n\} \setminus \{i_0\}$,

$$\varepsilon \coloneqq \prod_{\substack{i=1 \\ i \neq i_0}}^n \Phi\left(\sqrt{m_i}(\mu_i - \textstyle\bigwedge_{k=1}^n \overline{x_k})\right) \quad \text{and} \quad \sigma_{i_0} \coloneqq \begin{cases} 1 & \text{if } \varepsilon \leq 2\beta \\ \sqrt{m_{i_0}}\frac{\mu_{i_0} - \bigwedge_{k=1}^n \overline{x_k}}{2\Phi^{-1}(\beta/\varepsilon)} & \text{if } \varepsilon > 2\beta. \end{cases}$$

Then $\sigma \in]0,\infty[^n$ and

$$\prod_{i=1}^{n} \Phi\left(\sqrt{m_i}\frac{\mu_i - \bigwedge_{k=1}^{n} \overline{x_k}}{\sigma_i}\right) \le \varepsilon \Phi\left(\sqrt{m_{i_0}}\frac{\mu_{i_0} - \bigwedge_{k=1}^{n} \overline{x_k}}{\sigma_{i_0}}\right) < \beta$$

due to Remark 2.1.4, part 1. Example 2.1.3 yields $(\mu,\sigma) \in R_{\bigwedge_i \overline{X_i},\beta}(x)$. Since $\mathrm{pr}_1(\mu,\sigma) = \mu$, we obtain $\mu \in$ LHS.

(4) Let $\beta \in]1-\frac{1}{2^n}, 1]$. The inclusion LHS $\subseteq$ RHS is clear. It thus remains to show LHS $\supseteq$ RHS. To this end, let $\mu \in \mathbb{R}^n$, and let us define $t := (\bigwedge_{k=1}^{n} \mu_k) \wedge (\bigwedge_{k=1}^{n} \overline{x_k}) - 1$, $M := \bigvee_{k=1}^{n} m_k$, and

$$\sigma_i := M\frac{t - \bigwedge_{k=1}^{n} \overline{x_k}}{\Phi^{-1}((1-\beta)^{1/n})} \qquad \text{for } i \in \{1,\dots,n\}.$$

Then $\sigma \in]0,\infty[^n$ and

$$\prod_{i=1}^{n} \Phi\left(\sqrt{m_i}\frac{\mu_i - \bigwedge_{k=1}^{n} \overline{x_k}}{\sigma_i}\right) > \Phi^n\left(M\frac{t - \bigwedge_{k=1}^{n} \overline{x_k}}{\sigma_1}\right) = 1-\beta$$

due to Remark 2.1.4, part 1. Example 2.1.3 yields $(\mu,\sigma) \in R_{\bigwedge_i \overline{X_i},\beta}(x)$. Since $\mathrm{pr}_1(\mu,\sigma) = \mu$, we obtain $\mu \in$ LHS.

Let now $\beta \in]1-\frac{1}{2^{k-1}}, 1-\frac{1}{2^k}]$ for some $k \in \{1,\dots,n\}$. We first show LHS $\subseteq$ RHS. To this end, let $(\mu,\sigma) \in \tilde{R}_{\bigwedge_i \overline{X_i},\beta}(x)$, and let us assume $\mu \notin$ RHS. This means $|\{\mu \le \bigwedge_{i=1}^{n} \overline{x_i}\}| \ge k$, i.e., $\mu_j \le \bigwedge_{i=1}^{n} \overline{x_i}$ for at least k of the indices $j \in \{1,\dots,n\}$. Remark 2.1.4, part 1, implies

$$\prod_{i=1}^{n} \Phi\left(\sqrt{m_i}\frac{\mu_i - \bigwedge_{j=1}^{n} \overline{x_j}}{\sigma_i}\right) \le \Phi^k(0) = \frac{1}{2^k} \le 1-\beta,$$

a contradiction to $(\mu,\sigma) \in \tilde{R}_{\bigwedge_i \overline{X_i},\beta}(x)$ in view of Example 2.1.3.

We now show LHS $\supseteq$ RHS. Let $\mu \in$ RHS, and let us define $I := \{\mu > \bigwedge_{j=1}^{n} \overline{x_j}\} = \{i \in \{1,\dots,n\} \colon \mu_i > \bigwedge_{j=1}^{n} \overline{x_j}\}$, $J := \{1,\dots,n\} \setminus I$, and $\varepsilon := \beta - 1 + \frac{1}{2^{k-1}} \in]0, \frac{1}{2^k}[$. If $k \ge 2$, let furthermore $M := \bigvee_{j=1}^{n} m_j$, $t := \bigwedge_{j=1}^{n} \mu_j$, and, noting that $(1-\beta)^{\frac{1}{k-1}} < \frac{1}{2}$,

$$\sigma_i := 1 + M\frac{t - \bigwedge_{j=1}^{n} \overline{x_j}}{\Phi^{-1}((1-\beta)^{\frac{1}{k-1}})} \qquad \text{for } i \in J.$$

If $J \ne \emptyset$, then $k \ge 2$ and, with arbitrary $j \in J$,

$$\eta := \prod_{i\in J} \Phi\left(\sqrt{m_i}\frac{\mu_i - \bigwedge_{l=1}^{n} \overline{x_l}}{\sigma_i}\right) \ge \Phi^{|J|}\left(M\frac{t - \bigwedge_{l=1}^{n} \overline{x_l}}{\sigma_j}\right) > 1-\beta$$

due to Remark 2.1.4, part 1. Let us pick $N \in \mathbb{N}$ such that $(\frac{1-\beta}{\eta})^{1/N} > \frac{1}{2}$, and let us define $s := \bigwedge_{i\in I} \mu_i$ and

$$\sigma_i := \frac{t - \bigwedge_{l=1}^{n} \overline{x_l}}{2\Phi^{-1}((\frac{1-\beta}{\eta})^{1/N})} \qquad \text{for } i \in I.$$

Then $\sigma \in \,]0,\infty[^n$ and, with arbitrary $j \in I$,

$$\prod_{i=1}^{n} \Phi\left(\sqrt{m_i}\frac{\mu_i - \bigwedge_{l=1}^{n} \overline{x_l}}{\sigma_i}\right) \geq \eta \Phi^{|I|}\left(\frac{t - \bigwedge_{l=1}^{n} \overline{x_l}}{\sigma_j}\right) > \eta \left(\frac{1-\beta}{\eta}\right)^{|I|/N} \geq 1-\beta$$

due to Remark 2.1.4, part 1. Example 2.1.3 implies $(\mu,\sigma) \in R_{\bigwedge_i \overline{X_i},\beta}(x)$. Since $\mathrm{pr}_1(\mu,\sigma) = \mu$, we obtain $\mu \in$ LHS.

(5) The inclusions $\mathrm{pr}_2[R_{\bigwedge_i \overline{X_i},\beta}(x)] \subseteq \,]0,\infty[^n \supseteq \mathrm{pr}_2[\tilde{R}_{\bigwedge_i \overline{X_i},\beta}(x)]$ are clear. We first show the inclusion $\mathrm{pr}_2[R_{\bigwedge_i \overline{X_i},\beta}(x)] \supseteq \,]0,\infty[^n$. To this end, let $\sigma \in \,]0,\infty[^n$, and let us define

$$\mu_i := \bigwedge_{k=1}^{n} \overline{x_k} - \frac{\bigvee_{k=1}^{n} \sigma_k}{\bigvee_{k=1}^{n} m_k}\left(\Phi^{-1}(\beta^{1/n})\right)_{-} - 1 \qquad \text{for } i \in \{1,\dots,n\}.$$

Then $\mu \in \mathbb{R}^n$ and

$$\prod_{i=1}^{n} \Phi\left(\sqrt{m_i}\frac{\mu_i - \bigwedge_{k=1}^{n} \overline{x_k}}{\sigma_i}\right) \leq \Phi^n\left(M\frac{\mu_1 - \bigwedge_{k=1}^{n} \overline{x_k}}{\bigvee \sigma}\right) < \beta$$

due to Remark 2.1.4, part 1. Example 2.1.3 yields $(\mu,\sigma) \in R_{\bigwedge_i \overline{X_i},\beta}(x)$. Since $\mathrm{pr}_2(\mu,\sigma) = \sigma$, we obtain $\sigma \in \mathrm{pr}_2[R_{\bigwedge_i \overline{X_i},\beta}(x)]$.

We now show $\mathrm{pr}_2[\tilde{R}_{\bigwedge_i \overline{X_i},\beta}(x)] \supseteq \,]0,\infty[^n$. To this end, let $\sigma \in \,]0,\infty[^n$, and let us define

$$\mu_i := \bigwedge_{k=1}^{n} \overline{x_k} + \bigvee_{k=1}^{n} \sigma_k \Phi^{-1}((1-\beta)^{1/n}) + 1 \qquad \text{for } i \in \{1,\dots,n\}.$$

Then $\mu \in \mathbb{R}^n$ and

$$\prod_{i=1}^{n} \Phi\left(\sqrt{m_i}\frac{\mu_i - \bigwedge_{k=1}^{n} \overline{x_k}}{\sigma_i}\right) \geq \Phi^n\left(\frac{\mu_1 - \bigwedge_{k=1}^{n} \overline{x_k}}{\bigvee_{k=1}^{n} \sigma_k}\right) > 1-\beta$$

due to Remark 2.1.4, part 1. Example 2.1.3 implies $(\mu,\sigma) \in \tilde{R}_{\bigwedge_i \overline{X_i},\beta}(x)$. Together with $\mathrm{pr}_2(\mu,\sigma) = \sigma$, this yields $\sigma \in \mathrm{pr}_2[R_{\bigwedge_i \overline{X_i},\beta}(x)]$. □

2.1.6 Remark 1. If we generalize the model $\mathcal{P}_1$ to $\mathcal{P}_1|_{\mathrm{M}\times\Sigma}$, with $\mathrm{M} \subseteq \mathbb{R}^n$ and $\Sigma \subseteq \,]0,\infty[^n$, Example 2.1.3 yields $R_{\bigwedge_i \overline{X_i},\beta,\mathcal{P}_1|_{\mathrm{M}\times\Sigma}}(x) = R_{\bigwedge_i \overline{X_i},\beta,\mathcal{P}_1}(x) \cap (\mathrm{M}\times\Sigma)$ for $x \in \prod_{i=1}^{n} \mathbb{R}^{m_i}$, but an analogous version of Lemma 2.1.5 cannot be easily established in such generality.

2. Lemma 2.1.5 yields for the parameter of interest $\mathbb{R}^n \times \,]0,\infty[^n \to \mathbb{R}$, $(\mu,\sigma) \mapsto \bigwedge_{i=1}^{n} \mu_i$, the confidence regions given by

$$\bigwedge \circ \mathrm{pr}_1[R_{\bigwedge_i \overline{X_i},\beta}(x)] = \begin{cases} \mathbb{R} & \text{if } \beta \in \,]\frac{1}{2^n}, 1] \\ ]-\infty, \bigwedge_{i=1}^{n} \overline{x_i}[& \text{if } \beta \in \,]0, \frac{1}{2^n}] \end{cases}$$

$$\bigwedge \circ \mathrm{pr}_1[\tilde{R}_{\bigwedge_i \overline{X_i},\beta}(x)] = \begin{cases} \mathbb{R} & \text{if } \beta \in \,]\frac{1}{2}, 1] \\ ]\bigwedge_{i=1}^{n} \overline{x_i}, \infty[& \text{if } \beta \in \,]0, \frac{1}{2}] \end{cases}$$

for $x \in \prod_{i=1}^{n} \mathbb{R}^{m_i}$. This suggests that Buehlerization of $\bigwedge_i \overline{X_i}$ in $\mathcal{P}_1$ is rather useless (compared to the results obtained in the model $\mathcal{P}_2$ from the next section together with the remarks on location-scale models in Lloyd and Kabaila (2003), Section 4).

2.1.7 Example If we consider $\bigvee_i \overline{X_i}\colon \prod_{i=1}^n \mathbb{R}^{m_i} \to \mathbb{R}$, $x \mapsto \bigvee_{i=1}^n \overline{x_i}$, as designated statistic and define

$$f\colon \prod_{i=1}^n \mathbb{R}^{m_i} \to \prod_{i=1}^n \mathbb{R}^{m_i}, \quad x \mapsto -x,$$
$$g\colon \mathbb{R}^n \times]0,\infty[^n \to \mathbb{R}^n \times]0,\infty[^n, \quad (\mu,\sigma) \mapsto (-\mu,\sigma),$$

then $(\bigvee_i \overline{X_i}) \,\square\, \bigotimes_{i=1}^n \mathrm{N}_{\mu_i,\sigma_i^2}^{\otimes m_i} = (\bigwedge_i \overline{X_i} \circ f) \,\square\, \bigotimes_{i=1}^n \mathrm{N}_{-\mu_i,\sigma_i^2}^{\otimes m_i}$, which yields the confidence regions given by

$$\begin{aligned} R_{\bigvee_i \overline{X_i},\beta}(x) &= g^{-1}[\tilde{R}_{\bigwedge_i \overline{X_i},\beta}(f(x))] \\ &= \left\{(\mu,\sigma) \in \mathbb{R}^n \times]0,\infty[^n\colon \prod_{i=1}^n \Phi\left(\sqrt{m_i}\frac{\bigvee_{k=1}^n \overline{x_k} - \mu_i}{\sigma_i}\right) > 1-\beta\right\} \\ \tilde{R}_{\bigvee_i \overline{X_i},\beta}(x) &= g^{-1}[R_{\bigwedge_i \overline{X_i},\beta}(f(x))] \\ &= \left\{(\mu,\sigma) \in \mathbb{R}^n \times]0,\infty[^n\colon \prod_{i=1}^n \Phi\left(\sqrt{m_i}\frac{\bigvee_{k=1}^n \overline{x_k} - \mu_i}{\sigma_i}\right) < \beta\right\} \end{aligned}$$

for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$, with effective levels

$$\beta_{\text{eff}}(R_{\bigvee_i \overline{X_i},\beta}) = \beta_{\text{eff}}(\tilde{R}_{\bigwedge_i \overline{X_i},\beta}) = \beta = \beta_{\text{eff}}(R_{\bigwedge_i \overline{X_i},\beta}) = \beta_{\text{eff}}(\tilde{R}_{\bigvee_i \overline{X_i},\beta}) \qquad \text{for } \beta \in [0,1].$$

Lemma 2.1.5 furthermore yields for the parameters of interest pr_1 and pr_2 from Remark 2.1.4, part 2, the confidence regions given for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$ by

$$\begin{aligned} \mathrm{pr}_1[R_{\bigvee_i \overline{X_i},\beta}(x)] &= \begin{cases} \mathbb{R}^n & \text{if } \beta \in]1-\frac{1}{2^n},1] \\ \{\mu \in \mathbb{R}^n\colon |\{\mu < \bigvee_{i=1}^n \overline{x_i}\}| > n-k\} & \text{if } \beta \in]1-\frac{1}{2^{k-1}}, 1-\frac{1}{2^k}] \\ & \text{for some } k \in \{1,\ldots,n\} \end{cases} \\ \mathrm{pr}_1[\tilde{R}_{\bigvee_i \overline{X_i},\beta}(x)] &= \begin{cases} \mathbb{R}^n & \text{if } \beta \in]\frac{1}{2^n},1] \\ \bigcup_{k=1}^n \{\mu \in \mathbb{R}^n\colon \mu_k > \bigvee_{i=1}^n \overline{x_i}\} & \text{if } \beta \in]0,\frac{1}{2^n}] \end{cases} \\ \mathrm{pr}_2[R_{\bigvee_i \overline{X_i},\beta}(x)] &= \mathrm{pr}_2[\tilde{R}_{\bigvee_i \overline{X_i},\beta}(x)] =]0,\infty[^n. \end{aligned}$$

2.1.8 Example If we consider $\overline{X_i}$ from Example 2.1.3 and define

$$S_i\colon \prod_{k=1}^n \mathbb{R}^{m_k} \to \mathbb{R}, \quad x \mapsto \sqrt{\sum_{k=1}^{m_i} (x_{i,k} - \overline{x_i})^2/(m_i-1)}, \qquad \text{for } i \in \{1,\ldots,n\},$$

then, since $(\sqrt{m_i}\overline{X_i}/S_i) \,\square\, \mathrm{N}_{\mu_i,\sigma_i^2}^{\otimes m_i} = \mathrm{t}_{m_i-1,\mu_i/\sigma_i}$ is continuous (the noncentral t-distribution is introduced in Definition A.3.53), Remark 1.1.7, part 10, yields as Buehlerization of $\bigwedge_i \sqrt{m_i}\overline{X_i}/S_i$ the confidence regions given by

$$\begin{aligned} R_{\bigwedge_i \sqrt{m_i}\overline{X_i}/S_i,\beta}(x) &= \left\{(\mu,\sigma) \in \mathbb{R}^n \times]0,\infty[^n\colon \prod_{i=1}^n \mathrm{t}_{m_i-1,\mu_i/\sigma_i}\left(\left]\sqrt{m_i}\frac{\overline{x_i}}{S_i(x)},\infty\right[\right) < \beta\right\} \\ \tilde{R}_{\bigwedge_i \sqrt{m_i}\overline{X_i}/S_i,\beta}(x) &= \left\{(\mu,\sigma) \in \mathbb{R}^n \times]0,\infty[^n\colon \prod_{i=1}^n \mathrm{t}_{m_i-1,\mu_i/\sigma_i}\left(\left]\sqrt{m_i}\frac{\overline{x_i}}{S_i(x)},\infty\right[\right) > 1-\beta\right\} \end{aligned}$$

for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$. Since

$$\prod_{i=1}^n \mathrm{t}_{m_i-1,\mu_i/\sigma_i}\left(\left]\sqrt{m_i}\frac{\overline{X_i}}{S_i},\infty\right[\right) : \prod_{i=1}^n \mathbb{R}^{m_i} \to \,]0,1[$$

is surjective in view of Remark A.3.54, part 4, the effective levels of the above confidence regions are given by

$$\beta_{\text{eff}}(R_{\bigwedge_i \sqrt{m_i}\overline{X_i}/S_i,\beta}) = \beta_{\text{eff}}(\tilde{R}_{\bigwedge_i \sqrt{m_i}\overline{X_i}/S_i,\beta}) = \beta \qquad \text{for } \beta \in [0,1].$$

2.1.9 Remark For $\alpha \in \,]0,1[$ and $t \in \mathbb{R}$ there is exactly one

$$\overline{\bigwedge \mu/\sigma}_\alpha(t) \in \mathbb{R} \quad \text{such that} \quad \mathrm{t}_{n-1,\overline{\bigwedge \mu/\sigma}_\alpha(t)} = \alpha.$$

This follows from Remark A.3.54, part 6.

2.1.10 Lemma *Let us consider the parameter of interest*

$$\bigwedge \frac{\mathrm{pr}_1}{\mathrm{pr}_2} : \mathbb{R}^n \times \,]0,\infty[^n \to \mathbb{R}, \quad (\mu,\sigma) \mapsto \bigwedge_{i=1}^n \frac{\mu_i}{\sigma_i}.$$

For $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$ then

$$\bigwedge \frac{\mathrm{pr}_1}{\mathrm{pr}_2}[R_{\bigwedge_i \sqrt{m_i}\overline{X_i}/S_i,\beta}(x)] = \left]-\infty, \overline{\bigwedge \mu/\sigma}_{\beta^{1/n}}\left(\sqrt{m_i}\frac{\overline{x_i}}{S_i(x)}\right)\right[\tag{6}$$

$$\bigwedge \frac{\mathrm{pr}_1}{\mathrm{pr}_2}[\tilde{R}_{\bigwedge_i \sqrt{m_i}\overline{X_i}/S_i,\beta}(x)] = \left]\overline{\bigwedge \mu/\sigma}_{1-\beta}\left(\sqrt{m_i}\frac{\overline{x_i}}{S_i(x)}\right), \infty\right[. \tag{7}$$

Proof. Let $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$, and let us assume w.l.o.g. $\beta \in \,]0,1[$ (due to Remarks 1.1.7, part 3, and 1.2.3, part 6).

(6) We first show LHS $\subseteq$ RHS. To this end, let $(\mu,\sigma) \in R_{\bigwedge_i \sqrt{m_i}\overline{X_i}/S_i,\beta}(x)$, and let us assume $\bigwedge_{i=1}^n \mu_i/\sigma_i \geq \sup \text{RHS}$. Remark A.3.54, part 6, then yields

$$\prod_{i=1}^n \mathrm{t}_{m_i-1,\mu_i/\sigma_i}\left(\left]\sqrt{m_i}\frac{\overline{x_i}}{S_i(x)},\infty\right[\right) \geq \prod_{i=1}^n \mathrm{t}_{m_i-1,\sup \text{RHS}}\left(\left]\sqrt{m_i}\frac{\overline{x_i}}{S_i(x)},\infty\right[\right) \geq \beta,$$

a contradiction to $(\mu,\sigma) \in R_{\bigwedge_i \sqrt{m_i}\overline{X_i}/S_i,\beta}(x)$ in view of Example 2.1.8.
We now show LHS $\supseteq$ RHS. Let $t \in$ RHS, and let us define $\mu_i := t$ and $\sigma_i := 1$ for $i \in \{1,\ldots,n\}$. Then $\bigwedge_{i=1}^n \mu_i/\sigma_i = t$. Remark A.3.54, part 6, yields

$$\prod_{i=1}^n \mathrm{t}_{m_i-1,\mu_i/\sigma_i}\left(\left]\sqrt{m_i}\frac{\overline{x_i}}{S_i(x)},\infty\right[\right) < \prod_{i=1}^n \mathrm{t}_{m_i-1,\sup \text{RHS}}\left(\left]\sqrt{m_i}\frac{\overline{x_i}}{S_i(x)},\infty\right[\right) = \beta,$$

implying $(\mu,\sigma) \in R_{\bigwedge_i \sqrt{m_i}\overline{X_i}/S_i,\beta}(x)$ due to Example 2.1.8.

(7) We first show LHS $\subseteq$ RHS. To this end, let $(\mu,\sigma) \in \tilde{R}_{\bigwedge_i \sqrt{m_i}\overline{X_i}/S_i,\beta}(x)$, and let us assume $\bigwedge_{i=1}^n \mu_i/\sigma_i \leq \inf \text{RHS}$. Remark A.3.54, part 6, then yields

$$\prod_{i=1}^n \mathrm{t}_{m_i-1,\mu_i/\sigma_i}\left(\left]\sqrt{m_i}\frac{\overline{x_i}}{S_i(x)},\infty\right[\right) \leq \prod_{i=1}^n \mathrm{t}_{m_i-1,\inf \text{RHS}}\left(\left]\sqrt{m_i}\frac{\overline{x_i}}{S_i(x)},\infty\right[\right) \leq 1-\beta,$$

a contradiction to $(\mu,\sigma) \in \tilde{R}_{\bigwedge_i \sqrt{m_i}\overline{X_i}/S_i,\beta}(x)$ in view of Example 2.1.8.

We now show LHS $\supseteq$ RHS. Let $t \in$ RHS, and let us define $\mu_1 := t$ and $\sigma_i := 1$ for $i \in \{1,\dots,n\}$. Since

$$\lim_{\mu_2,\dots,\mu_n\to\infty} \prod_{i=2}^{n} \mathrm{t}_{m_i-1,\mu_i/\sigma_i}\left(\left]\sqrt{m_i}\frac{\overline{x_i}}{S_i(x)},\infty\right[\right) = 1,$$

we can pick $\mu_2,\dots,\mu_n \in [t,\infty[$ such that

$$\prod_{i=2}^{n} \mathrm{t}_{m_i-1,\mu_i/\sigma_i}\left(\left]\sqrt{m_i}\frac{\overline{x_i}}{S_i(x)},\infty\right[\right) > \frac{1-\beta}{\varepsilon} \wedge (1-\beta),$$

where $\varepsilon := \mathrm{t}_{m_1-1,\mu_1/\sigma_1}(]\sqrt{m_1}\cdot\overline{x_1}/S_1(x),\infty[) \in]0,1[$. Then $\bigwedge_{i=1}^n \mu_i/\sigma_i = t$. Remark A.3.54, part 6, yields

$$\prod_{i=1}^{n} \mathrm{t}_{m_i-1,\mu_i/\sigma_i}\left(\left]\sqrt{m_i}\frac{\overline{x_i}}{S_i(x)},\infty\right[\right) \geq \varepsilon \prod_{i=2}^{n} \mathrm{t}_{m_i-1,\mu_i/\sigma_i}\left(\left]\sqrt{m_i}\frac{\overline{x_i}}{S_i(x)},\infty\right[\right) > 1-\beta,$$

implying $(\mu,\sigma) \in \tilde{R}_{\bigwedge_i \sqrt{m_i}\overline{X_i}/S_i,\beta}(x)$ in view of Example 2.1.8. □

2.1.11 Definition For $\alpha \in]0,1[$ and $t \in \mathbb{R}$ let

$$\underline{\bigvee \mu/\sigma}_{\alpha}(t) := -\overline{\bigwedge \mu/\sigma}_{\alpha}(-t),$$

where $\overline{\bigwedge \mu/\sigma}_{\alpha}$ is the confidence bound given in Remark 2.1.9.

2.1.12 Example If we consider $\bigvee_i \sqrt{m_i}\overline{X_i}/S_i$ as designated statistic and set f and g as in Example 2.1.7, then $(\bigvee_i \sqrt{m_i}\overline{X_i}/S_i) \,\square\, \bigotimes_{i=1}^n \mathrm{N}_{\mu_i,\sigma_i^2}^{\otimes m_i} = (\bigwedge_i \sqrt{m_i}\overline{X_i}/S_i \circ f) \,\square\, \bigotimes_{i=1}^n \mathrm{N}_{-\mu_i,\sigma_i^2}^{\otimes m_i}$, which yields the confidence regions given by

$$\begin{aligned} R_{\bigvee_i \sqrt{m_i}\overline{X_i}/S_i,\beta}(x) &= g^{-1}[\tilde{R}_{\bigwedge_i \sqrt{m_i}\overline{X_i}/S_i,\beta}(f(x))] \\ &= \left\{(\mu,\sigma) \in \mathbb{R}^n \times]0,\infty[^n \colon \prod_{i=1}^{n} \mathrm{t}_{m_i-1,-\mu_i/\sigma_i}\left(\left]-\sqrt{m_i}\frac{\overline{x_i}}{S_i(x)},\infty\right[\right) > 1-\beta\right\} \end{aligned}$$

and

$$\begin{aligned} \tilde{R}_{\bigvee_i \sqrt{m_i}\overline{X_i}/S_i,\beta}(x) &= g^{-1}[R_{\bigwedge_i \sqrt{m_i}\overline{X_i}/S_i,\beta}(f(x))] \\ &= \left\{(\mu,\sigma) \in \mathbb{R}^n \times]0,\infty[^n \colon \prod_{i=1}^{n} \mathrm{t}_{m_i-1,-\mu_i/\sigma_i}\left(\left]-\sqrt{m_i}\frac{\overline{x_i}}{S_i(x)},\infty\right[\right) < \beta\right\} \end{aligned}$$

for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$, with effective levels

$$\begin{aligned} \beta_{\mathrm{eff}}(R_{\bigvee_i \sqrt{m_i}\overline{X_i}/S_i,\beta}) &= \beta_{\mathrm{eff}}(\tilde{R}_{\bigwedge_i \sqrt{m_i}\overline{X_i}/S_i,\beta}) = \beta \\ &= \beta_{\mathrm{eff}}(R_{\bigwedge_i \sqrt{m_i}\overline{X_i}/S_i,\beta}) = \beta_{\mathrm{eff}}(\tilde{R}_{\bigvee_i \sqrt{m_i}\overline{X_i}/S_i,\beta}) \qquad \text{for } \beta \in [0,1]. \end{aligned}$$

Lemma 2.1.10 furthermore yields for the parameter of interest

$$\bigvee \frac{\mathrm{pr}_1}{\mathrm{pr}_2} \colon \mathbb{R}^n \times]0,\infty[^n \to \mathbb{R}, \quad (\mu,\sigma) \mapsto \bigvee_{i=1}^{n} \frac{\mu_i}{\sigma_i},$$

the confidence regions given for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$ by

$$\bigvee \frac{\mathrm{pr}_1}{\mathrm{pr}_2}[R_{\bigvee_i \sqrt{m_i}\overline{X_i}/S_i,\beta}(x)] = \left]-\infty, \underline{\bigvee \mu/\sigma}_{1-\beta}\left(\sqrt{m_i}\frac{\overline{x_i}}{S_i(x)}\right)\right[$$
$$\bigvee \frac{\mathrm{pr}_1}{\mathrm{pr}_2}[\tilde{R}_{\bigvee_i \sqrt{m_i}\overline{X_i}/S_i,\beta}(x)] = \left]\underline{\bigvee \mu/\sigma}_{\beta^{1/n}}\left(\sqrt{m_i}\frac{\overline{x_i}}{S_i(x)}\right), \infty\right[,$$

where $\overline{\bigwedge \mu/\sigma}_\beta$ is the confidence bound from Definition 2.1.11.

2.1.13 Example If we consider the designated statistic

$$\textstyle\bigwedge_{i,k} X_{i,k} \colon \prod_{i=1}^n \mathbb{R}^{m_i} \to \mathbb{R}, \quad x \mapsto \bigwedge_{i=1}^n \bigwedge_{k=1}^{m_i} x_{i,k},$$

then, as $(\bigwedge_{i,k} X_{i,k}) \square \bigotimes_{i=1}^n \mathrm{N}_{\mu_i,\sigma_i^2}^{\otimes m_i}$ is continuous, Remark 1.1.7, part 10, yields the confidence regions given by

$$R_{\bigwedge_{i,k} X_{i,k},\beta}(x) = \left\{(\mu,\sigma) \in \mathbb{R}^n \times]0,\infty[^n \colon \prod_{i=1}^n \Phi^{m_i}\left(\frac{\mu_i - \bigwedge_{j=1}^n \bigwedge_{k=1}^{m_j} x_{j,k}}{\sigma_i}\right) < \beta\right\}$$
$$\tilde{R}_{\bigwedge_{i,k} X_{i,k},\beta}(x) = \left\{(\mu,\sigma) \in \mathbb{R}^n \times]0,\infty[^n \colon \prod_{i=1}^n \Phi^{m_i}\left(\frac{\mu_i - \bigwedge_{j=1}^n \bigwedge_{k=1}^{m_j} x_{j,k}}{\sigma_i}\right) > 1-\beta\right\}$$

for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$. Since

$$\prod_{i=1}^n \Phi^{m_i}\left(\frac{\mu_i - \bigwedge_{j,k} X_{j,k}}{\sigma_i}\right) \colon \prod_{i=1}^n \mathbb{R}^{m_i} \to]0,1[$$

is surjective, their effective levels are given by

$$\beta_{\mathrm{eff}}(R_{\bigwedge_{i,k} X_{i,k},\beta}) = \beta_{\mathrm{eff}}(\tilde{R}_{\bigwedge_{i,k} X_{i,k},\beta}) = \beta \qquad \text{for } \beta \in [0,1].$$

2.1.14 Remark

1. We shall use in the proof of the next lemma the fact that

   $$\mathbb{R}^n \to]0,1[, \quad \mu \mapsto \prod_{i=1}^n \Phi^{m_i}\left(\frac{\mu_i - t}{\sigma_i}\right),$$

   is strictly increasing for $t \in \mathbb{R}$ and $\sigma \in]0,\infty[^n$, and

   $$]0,\infty[\to]0,1[, \quad \sigma_k \mapsto \prod_{i=1}^n \Phi^{m_i}\left(\frac{\mu_i - t}{\sigma_i}\right),$$

   is, for $k \in \{1,\ldots,n\}$, $t \in \mathbb{R}$ and $\mu \in \mathbb{R}^n$, strictly increasing if $\mu_k < t$, strictly decreasing if $\mu_k > t$, and constant if $\mu_k = t$.
2. The next result yields confidence regions for the parameters of interest pr_1 and pr_2 from Remark 2.1.4, part 2.

2.1.15 Lemma *Let* pr_1 *and* pr_2 *denote the projections from Remark 2.1.4, part 2, and* $|m| := \sum_{i=1}^n m_i$. *For* $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$ *then*

$$\mathrm{pr}_1[R_{\bigwedge_{i,k} X_{i,k},\beta}(x)] = \begin{cases} \mathbb{R}^n & \text{if } \beta \in]\frac{1}{2^{|m|}}, 1] \\ \bigcup_{k=1}^n \{\mu \in \mathbb{R}^n \colon \mu_k < \bigwedge_{i=1}^n \bigwedge_{j=1}^{m_i} x_{i,j}\} & \text{if } \beta \in]0, \frac{1}{2^{|m|}}] \end{cases} \tag{8}$$

$$\mathrm{pr}_1[\tilde{R}_{\bigwedge_{i,k} X_{i,k},\beta}(x)] = \mathbb{R}^n \qquad \text{if } \beta \in]1 - \tfrac{1}{2^{|m|}}, 1] \tag{9}$$

$$\mathrm{pr}_2[R_{\bigwedge_{i,k} X_{i,k},\beta}(x)] = \mathrm{pr}_2[\tilde{R}_{\bigwedge_{i,k} X_{i,k},\beta}(x)] =]0,\infty[^n. \tag{10}$$

Proof. Let $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$, and let us assume w.l.o.g. $\beta \in]0,1[$ (due to Remarks 1.1.7, part 3, and 1.2.3, part 6).

(8) Let $\beta \in]\frac{1}{2^{|m|}}, 1]$. The inclusion LHS $\subseteq$ RHS is clear. It thus remains to show LHS $\supseteq$ RHS. To this end, let $\mu \in \mathbb{R}^n$. Since

$$\lim_{\sigma_1,\ldots,\sigma_n \to \infty} \prod_{i=1}^n \Phi^{m_i}\left(\frac{\mu_i - \bigwedge_{j=1}^n \bigwedge_{k=1}^{m_j} x_{j,k}}{\sigma_i}\right) = \frac{1}{2^{|m|}},$$

we can pick $\sigma \in]0,\infty[^n$ such that

$$\prod_{i=1}^n \Phi^{m_i}\left(\frac{\mu_i - \bigwedge_{j=1}^n \bigwedge_{k=1}^{m_j} x_{j,k}}{\sigma_i}\right) = \frac{1}{2^{|m|}} < \beta.$$

Example 2.1.13 implies $(\mu,\sigma) \in R_{\bigwedge_{i,k} X_{i,k},\beta}(x)$.

Let now $\beta \in]0, \frac{1}{2^{|m|}}]$. We first show LHS $\subseteq$ RHS. To this end, let $(\mu,\sigma) \in R_{\bigwedge_{i,k} X_{i,k},\beta}(x)$, and let us assume $\bigwedge_{k=1}^n \mu_k \geq \bigwedge_{i=1}^n \bigwedge_{j=1}^{m_i} x_{i,j}$. Remark 2.1.14, part 1, then implies

$$\prod_{i=1}^n \Phi^{m_i}\left(\frac{\mu_i - \bigwedge_{j=1}^n \bigwedge_{k=1}^{m_j} x_{j,k}}{\sigma_i}\right) \geq \frac{1}{2^{|m|}} \geq \beta,$$

a contradiction to $(\mu,\sigma) \in R_{\bigwedge_{i,k} X_{i,k},\beta}(x)$ in view of Example 2.1.13.

We now show LHS $\supseteq$ RHS. Let $\mu \in$ RHS, and let us pick $r \in \{1,\ldots,n\}$ such that $\mu_r < \bigwedge_{i=1}^n \bigwedge_{j=1}^{m_i} x_{i,j}$. Since

$$\lim_{\sigma_r \to 0+} \prod_{i=1}^n \Phi^{m_i}\left(\frac{\mu_i - \bigwedge_{j=1}^n \bigwedge_{k=1}^{m_j} x_{j,k}}{\sigma_i}\right) = 0,$$

we can pick $\sigma \in]0,\infty[^n$ such that

$$\prod_{i=1}^n \Phi^{m_i}\left(\frac{\mu_i - \bigwedge_{j=1}^n \bigwedge_{k=1}^{m_j} x_{j,k}}{\sigma_i}\right) < \beta.$$

Example 2.1.13 yields $(\mu,\sigma) \in R_{\bigwedge_{i,k} X_{i,k},\beta}(x)$.

(9) Let $\beta \in]1 - \frac{1}{2^{|m|}}, 1]$. The inclusion LHS $\subseteq$ RHS is clear. It thus remains to show LHS $\supseteq$ RHS. To this end, let $\mu \in \mathbb{R}^n$. Since

$$\lim_{\sigma_1,\ldots,\sigma_n \to \infty} \prod_{i=1}^n \Phi^{m_i}\left(\frac{\mu_i - \bigwedge_{j=1}^n \bigwedge_{k=1}^{m_j} x_{j,k}}{\sigma_i}\right) = \frac{1}{2^{|m|}},$$

we can pick $\sigma \in]0,\infty[^n$ such that

$$\prod_{i=1}^n \Phi^{m_i}\left(\frac{\mu_i - \bigwedge_{j=1}^n \bigwedge_{k=1}^{m_j} x_{j,k}}{\sigma_i}\right) = \frac{1}{2^{|m|}} > 1 - \beta.$$

Example 2.1.13 yields $(\mu,\sigma) \in \tilde{R}_{\bigwedge_{i,k} X_{i,k},\beta}(x)$.

(10) The inclusions $\mathrm{pr}_2[R_{\bigwedge_{i,k} X_{i,k},\beta}(x)] \subseteq\ ]0,\infty[^n \supseteq \mathrm{pr}_2[\tilde{R}_{\bigwedge_{i,k} X_{i,k},\beta}(x)]$ are clear. It thus remains to show $]0,\infty[^n \subseteq \mathrm{pr}_2[R_{\bigwedge_{i,j} X_{i,j},\beta}(x)] \cap \mathrm{pr}_2[\tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(x)]$. To this end, let $\sigma \in\ ]0,\infty[^n$. Since

$$\lim_{\mu_1,\dots,\mu_n\to\mathsf{Inf}} \prod_{i=1}^{n} \Phi^{m_i}\left(\frac{\mu_i - \bigwedge_{j=1}^{n}\bigwedge_{k=1}^{m_j} x_{j,k}}{\sigma_i}\right) = \begin{cases} 0 & \text{if } \mathsf{Inf} = -\infty, \\ 1 & \text{if } \mathsf{Inf} = \infty, \end{cases}$$

we can pick $\mu \in \mathbb{R}^n$ such that

$$\prod_{i=1}^{n} \Phi^{m_i}\left(\frac{\mu_i - \bigwedge_{j=1}^{n}\bigwedge_{k=1}^{m_j} x_{j,k}}{\sigma_i}\right) = \frac{1}{2^{|m|}} < \beta \quad \text{resp.} \quad > 1-\beta$$

Example 2.1.13 yields $(\mu,\sigma) \in R_{\bigwedge_{i,k} X_{i,k},\beta}(x)$ resp. $(\mu,\sigma) \in \tilde{R}_{\bigwedge_{i,k} X_{i,k},\beta}(x)$. □

2.1.16 Remark The preceding result yields for the parameter of interest $\bigwedge \circ \mathrm{pr}_1 \colon \mathbb{R}^n \times\ ]0,\infty[^n \to \mathbb{R}$, $(\mu,\sigma) \mapsto \bigwedge_{i=1}^n \mu_i$, the confidence regions given by

$$\bigwedge \circ \mathrm{pr}_1[R_{\bigwedge_{i,j} X_{i,j},\beta}(x)] = \begin{cases} \mathbb{R} & \text{if } \beta \in\ ]\frac{1}{2^{|m|}},1] \\ ]-\infty, \bigwedge_{i=1}^{n}\bigwedge_{j=1}^{m_i} x_{i,j}[& \text{if } \beta \in\ ]0,\frac{1}{2^{|m|}}] \end{cases}$$
$$\bigwedge \circ \mathrm{pr}_1[\tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(x)] = \mathbb{R} \qquad \text{if } \beta \in\ ]1-\tfrac{1}{2^{|m|}},1]$$

for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$. This suggests that Buehlerization of $\bigwedge_{i,j} X_{i,j}$ in $\mathcal{P}_1$ is rather useless.

2.1.17 Example If we consider

$$\textstyle\bigvee_{i,j} X_{i,j} \colon \prod_{i=1}^{n} \mathbb{R}^{m_i} \to \mathbb{R}, \quad x \mapsto \bigvee_{i=1}^{n}\bigvee_{j=1}^{m_i} x_{i,j},$$

as designated statistic and set f and g as in Example 2.1.7, then $(\bigvee_{i,j} X_{i,j}) \,\square \bigotimes_{i=1}^n \mathrm{N}_{\mu_i,\sigma_i^2}^{\otimes m_i} = (\bigwedge_{i,j} X_{i,j} \circ f) \,\square \bigotimes_{i=1}^n \mathrm{N}_{-\mu_i,\sigma_i^2}^{\otimes m_i}$, which yields the confidence regions given by

$$\begin{aligned} R_{\bigvee_{i,j} X_{i,j},\beta}(x) &= g^{-1}[\tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(f(x))] \\ &= \left\{(\mu,\sigma) \in \mathbb{R}^n \times\ ]0,\infty[^n \colon \prod_{i=1}^{n} \Phi^{m_i}\left(\frac{\bigwedge_{j=1}^{n}\bigwedge_{k=1}^{m_j} x_{j,k} - \mu_i}{\sigma_i}\right) > 1-\beta\right\} \\ \tilde{R}_{\bigvee_i \overline{X_i},\beta}(x) &= g^{-1}[R_{\bigwedge_i \overline{X_i},\beta}(f(x))] \\ &= \left\{(\mu,\sigma) \in \mathbb{R}^n \times\ ]0,\infty[^n \colon \prod_{i=1}^{n} \Phi^{m_i}\left(\frac{\bigwedge_{j=1}^{n}\bigwedge_{k=1}^{m_j} x_{j,k} - \mu_i}{\sigma_i}\right) < \beta\right\} \end{aligned}$$

for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$, with effective levels

$$\beta_{\mathrm{eff}}(R_{\bigvee_{i,j} X_{i,j},\beta}) = \beta_{\mathrm{eff}}(\tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}) = \beta = \beta_{\mathrm{eff}}(R_{\bigwedge_{i,j} X_{i,j},\beta}) = \beta_{\mathrm{eff}}(\tilde{R}_{\bigvee_{i,j} X_{i,j},\beta})$$

for $\beta \in [0,1]$. Lemma 2.1.15 furthermore yields for the parameters of interest pr_1 and pr_2 from Remark 2.1.4 the confidence regions given for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$ by

$$\mathrm{pr}_1[R_{\bigvee_{i,j} X_{i,j},\beta}(x)] = \mathbb{R}^n \qquad \text{if } \beta \in\ ]1-\tfrac{1}{2^{|m|}},1]$$

$$\mathrm{pr}_1[\tilde{R}_{\bigvee_{i,j} X_{i,j},\beta}(x)] = \begin{cases} \mathbb{R}^n & \text{if } \beta \in \left]\frac{1}{2^{\lceil m \rceil}}, 1\right] \\ \bigcup_{k=1}^n \{\mu \in \mathbb{R}^n : \mu_k > \bigvee_{i=1}^n \bigvee_{j=1}^{m_i} x_{i,j}\} & \text{if } \beta \in \left]0, \frac{1}{2^{\lceil m \rceil}}\right] \end{cases}$$
$$\mathrm{pr}_2[R_{\bigvee_{i,j} X_{i,j},\beta}(x)] = \mathrm{pr}_2[\tilde{R}_{\bigvee_{i,j} X_{i,j},\beta}(x)] =]0,\infty[^n.$$

2.1.18 Example If we consider the mean

$$\overline{X}\colon \prod_{i=1}^n \mathbb{R}^{m_i} \to \mathbb{R}, \quad x \mapsto \overline{x} := \frac{1}{n}\sum_{i=1}^n \overline{x_i} = \sum_{i=1}^n \sum_{j=1}^{m_i} \frac{x_{i,j}}{nm_i},$$

of the different samples' means as designated statistic, then $\overline{X} \,\square\, \bigotimes_{i=1}^n \mathrm{N}_{\mu_i,\sigma_i^2}^{\otimes m_i} = \mathrm{N}_{\overline{\mu},\overline{\sigma^2/m}/n}$, which yields

$$\bigotimes_{i=1}^n \mathrm{N}_{\mu_i,\sigma_i^2}^{\otimes m_i}(\overline{X} > \overline{x}) = \Phi\left(\sqrt{n}\frac{\overline{\mu} - \overline{x}}{\sqrt{\overline{\sigma^2/m}}}\right)$$

for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$. Here, $\overline{\mu} = \sum_{i=1}^n \mu_i/n$ and $\overline{\sigma^2/m} = \sum_{i=1}^n \sigma_i^2/(nm_i)$. Since $\overline{X} \,\square\, \bigotimes_{i=1}^n \mathrm{N}_{\mu_i,\sigma_i^2}^{\otimes m_i}$ is continuous, Remark 1.1.7, part 10, yields the confidence regions given by

$$R_{\overline{X},\beta}(x) = \left\{(\mu,\sigma) \in \mathbb{R}^n \times]0,\infty[^n : \sqrt{n}\frac{\overline{\mu} - \overline{x}}{\sqrt{\overline{\sigma^2/m}}} < \Phi^{-1}(\beta)\right\}$$
$$\tilde{R}_{\overline{X},\beta}(x) = \left\{(\mu,\sigma) \in \mathbb{R}^n \times]0,\infty[^n : \sqrt{n}\frac{\overline{x} - \overline{\mu}}{\sqrt{\overline{\sigma^2/m}}} < \Phi^{-1}(\beta)\right\}$$

for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$. Since

$$\Phi\left(\sqrt{n}\frac{\overline{\mu} - \overline{X}}{\sqrt{\overline{\sigma^2/m}}}\right) : \prod_{i=1}^n \mathbb{R}^{m_i} \to]0,1[$$

is surjective, the effective levels of $R_{\overline{X},\beta}$ and $\tilde{R}_{\overline{X},\beta}$ are given by

$$\beta_{\text{eff}}(R_{\overline{X},\beta}) = \beta_{\text{eff}}(\tilde{R}_{\overline{X},\beta}) = \beta \qquad \text{for } \beta \in [0,1].$$

2.1.19 Lemma *Let us consider the parameter of interest*

$$\kappa\colon \mathbb{R}^n \to \mathbb{R}, \quad (\mu,\sigma) \mapsto \overline{\mu} = \frac{1}{n}\sum_{i=1}^n \mu_i.$$

For $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$ *then*

$$\kappa[R_{\overline{X},\beta}(x)] = \begin{cases} \mathbb{R} & \textit{if } \beta \in]\frac{1}{2}, 1] \\ ]-\infty, \overline{x}[& \textit{if } \beta \in]0, \frac{1}{2}] \end{cases} \tag{11}$$

$$\kappa[R_{\overline{X},\beta}(x)] = \begin{cases} \mathbb{R} & \textit{if } \beta \in]\frac{1}{2}, 1] \\ ]\overline{x}, \infty[& \textit{if } \beta \in]0, \frac{1}{2}]. \end{cases} \tag{12}$$

Proof. Let $x \in \prod_{i=1}^{n} \mathbb{R}^{m_i}$, and let us assume w.l.o.g. $\beta \in]0,1[$ (due to Remarks 1.1.7, part 3, and 1.2.3, part 6).

(11) Let $\beta \in]\frac{1}{2}, 1]$. The inclusion LHS $\subseteq$ RHS is clear. It thus remains to show LHS $\supseteq$ RHS. To this end, let $t \in \mathbb{R}$, and let us define $\mu_i := t$ for $i \in \{1, \ldots, n\}$. Since

$$\lim_{\sigma_1 \to \infty} \sqrt{n} \frac{\overline{x} - \overline{\mu}}{\sqrt{\overline{\sigma^2/m}}} = 0 < \Phi^{-1}(\beta),$$

we can pick $\sigma \in]0, \infty[^n$ such that $(\mu, \sigma) \in R_{\overline{X},\beta}(x)$ in view of Example 2.1.18.
Let now $\beta \in]0, \frac{1}{2}]$. We first show LHS $\subseteq$ RHS. Let $(\mu, \sigma) \in R_{\overline{X},\beta}(x)$, and let us assume $\overline{\mu} \geq \overline{x}$. Then

$$\sqrt{n} \frac{\overline{\mu} - \overline{x}}{\sqrt{\overline{\sigma^2/m}}} \geq 0 \geq \Phi^{-1}(\beta),$$

a contradiction to $(\mu, \sigma) \in R_{\overline{X},\beta}(x)$ in view of Example 2.1.18.
We now show LHS $\supseteq$ RHS. Let $t \in]-\infty, \overline{x}[$, and let us define $\mu_i := t$ for $i \in \{1, \ldots, n\}$. Since

$$\lim_{\sigma_1, \ldots, \sigma_n \to 0+} \sqrt{n} \frac{\overline{\mu} - \overline{x}}{\sqrt{\overline{\sigma^2/m}}} = -\infty < \Phi^{-1}(\beta),$$

we can pick $\sigma \in]0, \infty[^n$ such that $(\mu, \sigma) \in R_{\overline{X},\beta}(x)$ in view of Example 2.1.18.

(12) Let $\beta \in]\frac{1}{2}, 1]$. The inclusion LHS $\subseteq$ RHS is clear. It thus remains to show LHS $\supseteq$ RHS. To this end, let $t \in \mathbb{R}$, and let us define $\mu_i := t$ for $i \in \{1, \ldots, n\}$. Since

$$\lim_{\sigma_1 \to \infty} \sqrt{n} \frac{\overline{x} - \overline{\mu}}{\sqrt{\overline{\sigma^2/m}}} = 0 < \Phi^{-1}(\beta),$$

we can pick $\sigma \in]0, \infty[^n$ such that $(\mu, \sigma) \in \tilde{R}_{\overline{X},\beta}(x)$ in view of Example 2.1.18.
Let now $\beta \in]0, \frac{1}{2}]$. We first show LHS $\subseteq$ RHS. To this end, let $(\mu, \sigma) \in \tilde{R}_{\overline{X},\beta}(x)$, and let us assume $\overline{\mu} \leq \overline{x}$. Then

$$\sqrt{n} \frac{\overline{x} - \overline{\mu}}{\sqrt{\overline{\sigma^2/m}}} \geq 0 \geq \Phi^{-1}(\beta),$$

a contradiction to $(\mu, \sigma) \in \tilde{R}_{\overline{X},\beta}(x)$ in view of Example 2.1.18.
We now show LHS $\supseteq$ RHS. Let $t \in]\overline{x}, \infty[$, and let us define $\mu_i := t$ for $i \in \{1, \ldots, n\}$. Since

$$\lim_{\sigma_1, \ldots, \sigma_n \to 0+} \sqrt{n} \frac{\overline{x} - \overline{\mu}}{\sqrt{\overline{\sigma^2/m}}} = -\infty < \Phi^{-1}(\beta),$$

we can pick $\sigma \in]0, \infty[^n$ such that $(\mu, \sigma) \in \tilde{R}_{\overline{X},\beta}(x)$ in view of Example 2.1.18. $\square$

2.2 SEVERAL SAMPLES WITH KNOWN VARIANCES

2.2.1 Definition Let $n \in \mathbb{N}$, $m \in \mathbb{N}^n$, and $\sigma \in \,]0,\infty[^n$, and let

$$\mathcal{P}_2 := \left(\bigotimes_{i=1}^{n} \mathrm{N}_{\mu_i,\sigma_i^2}^{\otimes m_i} : \mu \in \mathbb{R}^n \right)$$

be the n-sample normal model with known variances $\sigma_1^2, \ldots, \sigma_n^2$ and known sample sizes $m_1, \ldots, m_n$.

2.2.2 Remark 1. The sample space is $\mathcal{X} = \prod_{i=1}^n \mathbb{R}^{m_i}$, the parameter space $\Theta = \mathbb{R}^n$.

2. $\mathcal{P}_2$ is stochastically increasing.
3. The following is a special case of $\mathcal{P}_2$:

$$\mathcal{P}_2' := \left(\bigotimes_{i=1}^{n} \mathrm{N}_{\mu_i,\sigma^2}^{\otimes m_i} : \mu \in \mathbb{R}^n \right) \qquad \text{for } \sigma \in \,]0,\infty[,$$

the n-sample homoscedastic normal model with known variance σ^2 and known sample sizes $m_1, \ldots, m_n$.

2.2.3 Example If we consider the designated statistic $\bigwedge_i \overline{X_i}$ from Example 2.1.3, then the calculations there yield

$$R_{\bigwedge_i \overline{X_i},\beta}(x) = \left\{ \mu \in \mathbb{R}^n \colon \prod_{i=1}^n \Phi\left(\sqrt{m_i} \frac{\mu_i - \bigwedge_{k=1}^n \overline{x_k}}{\sigma_i} \right) < \beta \right\}$$

$$\tilde{R}_{\bigwedge_i \overline{X_i},\beta}(x) = \left\{ \mu \in \mathbb{R}^n \colon \prod_{i=1}^n \Phi\left(\sqrt{m_i} \frac{\mu_i - \bigwedge_{k=1}^n \overline{x_k}}{\sigma_i} \right) > 1-\beta \right\},$$

for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$, with effective levels $\beta_{\text{eff}}(R_{\bigwedge_i \overline{X_i},\beta}) = \beta_{\text{eff}}(\tilde{R}_{\bigwedge_i \overline{X_i},\beta}) = \beta$ for $\beta \in [0,1]$.

2.2.4 Remark For $\alpha \in \,]0,1[$ and $t \in \mathbb{R}$ there is exactly one

$$\overline{\bigwedge \mu}_\alpha(t) \in \mathbb{R} \quad \text{such that} \quad \prod_{i=1}^n \Phi\left(\sqrt{m_i} \frac{\overline{\bigwedge \mu}_\alpha(t) - t}{\sigma_i} \right) = \alpha.$$

In fact, $f\colon \mathbb{R} \to \,]0,1[$, $s \mapsto \prod_{i=1}^n \Phi(\sqrt{m_i}\frac{s-t}{\sigma_i})$, is bijective since it is strictly increasing and continuous with $\lim_{s\to-\infty} f(s) = 0$ and $\lim_{s\to\infty} f(s) = 1$.

2.2.5 Lemma *Let us consider the parameter of interest*

$$\bigwedge \colon \mathbb{R}^n \to \mathbb{R}, \quad \mu \mapsto \bigwedge_{i=1}^n \mu_i.$$

For $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$ *then*

$$\bigwedge[R_{\bigwedge_i \overline{X_i},\beta}(x)] = \left]-\infty, \overline{\bigwedge \mu}_\beta(\textstyle\bigwedge_{i=1}^n \overline{x_i})\right[\tag{13}$$

$$\bigwedge[\tilde{R}_{\bigwedge_i \overline{X_i},\beta}(x)] = \begin{cases} \left]\bigwedge_{i=1}^n \overline{x_i} - \Phi^{-1}(\beta) \bigvee_{i=1}^n \frac{\sigma_i}{\sqrt{m_i}}, \infty\right[& \textit{if } \beta \in [\frac{1}{2}, 1] \\ \left]\bigwedge_{i=1}^n \overline{x_i} - \Phi^{-1}(\beta) \bigwedge_{i=1}^n \frac{\sigma_i}{\sqrt{m_i}}, \infty\right[& \textit{if } \beta \in [0, \frac{1}{2}]. \end{cases} \tag{14}$$

Proof. Let $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$, and let us assume w.l.o.g. $\beta \in]0,1[$ (due to Remarks 1.1.7, part 3, and 1.2.3, part 6).

(13) We first show LHS $\subseteq$ RHS. Let $\mu \in R_{\bigwedge_i \overline{X_i},\beta}(x)$ and assume $\bigwedge_{i=1}^n \mu_i \geq \sup \text{RHS}$. Remarks 2.1.4, part 1, and 2.2.4 imply

$$\prod_{i=1}^n \Phi\left(\sqrt{m_i}\frac{\mu_i - \bigwedge_{k=1}^n \overline{x_k}}{\sigma_i}\right) \geq \prod_{i=1}^n \Phi\left(\sqrt{m_i}\frac{\sup \text{RHS} - \bigwedge_{k=1}^n \overline{x_k}}{\sigma_i}\right) = \beta,$$

a contradiction to $\mu \in R_{\bigwedge_i \overline{X_i},\beta}(x)$ in view of Example 2.2.3.
We now show LHS $\supseteq$ RHS. Let $t \in$ RHS, and let us define $\mu_i := t$ for $i \in \{1,\dots,n\}$. Remark 2.1.4, part 1, yields

$$\prod_{i=1}^n \Phi\left(\sqrt{m_i}\frac{\mu_i - \bigwedge_{k=1}^n \overline{x_k}}{\sigma_i}\right) < \prod_{i=1}^n \Phi\left(\sqrt{m_i}\frac{\sup \text{RHS} - \bigwedge_{k=1}^n \overline{x_k}}{\sigma_i}\right) = \beta,$$

hence $\mu \in R_{\bigwedge_i \overline{X_i},\beta}(x)$ due to Example 2.2.3. Together with $\bigwedge_{k=1}^n \mu_k = t$, this yields $t \in$ LHS.

(14) We first show LHS $\subseteq$ RHS. To this end, let $\mu \in \tilde{R}_{\bigwedge_i \overline{X_i},\beta}(x)$, and let us assume $\bigwedge_{i=1}^n \mu_i \leq \inf \text{RHS}$. Pick $i_0 \in \{1,\dots,n\}$ such that $\mu_{i_0} = \bigwedge_{i=1}^n \mu_i$. Remark 2.1.4, part 1, implies

$$\begin{aligned}\prod_{i=1}^n \Phi\left(\sqrt{m_i}\frac{\mu_i - \bigwedge_{k=1}^n \overline{x_k}}{\sigma_i}\right) &\leq \Phi\left(\sqrt{m_{i_0}}\frac{\mu_{i_0} - \bigwedge_{k=1}^n \overline{x_k}}{\sigma_{i_0}}\right) \\ &\leq \begin{cases}\Phi\left(\Phi^{-1}(1-\beta)\frac{\sqrt{m_{i_0}}}{\sigma_{i_0}} \bigvee_{i=1}^n \frac{\sigma_i}{\sqrt{m_i}}\right) & \text{if } \beta \in [\frac{1}{2},1[\\ \Phi\left(\Phi^{-1}(1-\beta)\frac{\sqrt{m_{i_0}}}{\sigma_{i_0}} \bigwedge_{i=1}^n \frac{\sigma_i}{\sqrt{m_i}}\right) & \text{if } \beta \in]0,\frac{1}{2}]\end{cases} \\ &\leq 1-\beta,\end{aligned}$$

a contradiction to $\mu \in \tilde{R}_{\bigwedge_i \overline{X_i},\beta}(x)$ in view of Example 2.2.3.
We now show LHS $\supseteq$ RHS. Let $t \in$ RHS, let us pick $i_0 \in \{1,\dots,n\}$ such that

$$\frac{\sigma_{i_0}}{\sqrt{m_{i_0}}} = \begin{cases}\bigvee_{i=1}^n \frac{\sigma_i}{\sqrt{m_i}} & \text{if } \beta \in [\frac{1}{2},1[\\ \bigwedge_{i=1}^n \frac{\sigma_i}{\sqrt{m_i}} & \text{if } \beta \in]0,\frac{1}{2}[,\end{cases}$$

and let us define $\mu_{i_0} := t$. Then

$$\begin{aligned}\varepsilon := \Phi\left(\sqrt{m_{i_0}}\frac{t - \bigwedge_{k=1}^n \overline{x_k}}{\sigma_{i_0}}\right) &> \begin{cases}\Phi\left(\Phi^{-1}(1-\beta)\frac{\sqrt{m_{i_0}}}{\sigma_{i_0}} \bigvee_{k=1}^n \frac{\sigma_k}{\sqrt{m_k}}\right) & \text{if } \beta \in [\frac{1}{2},1[\\ \Phi\left(\Phi^{-1}(1-\beta)\frac{\sqrt{m_{i_0}}}{\sigma_{i_0}} \bigwedge_{k=1}^n \frac{\sigma_k}{\sqrt{m_k}}\right) & \text{if } \beta \in]0,\frac{1}{2}]\end{cases} \\ &= 1-\beta.\end{aligned}$$

If we further define

$$\mu_i := \left(\bigwedge_{k=1}^n \overline{x_k} + \Phi^{-1}\left(\left(\frac{1-\beta}{\varepsilon}\right)^{1/n}\right) \bigvee_{k=1}^n \sigma_k\right) \vee t \qquad \text{for } i \in \{1,\dots,n\} \setminus \{i_0\},$$

then, with arbitrary $j \in \{1, \dots, n\} \setminus \{i_0\}$,

$$\prod_{i=1}^{n} \Phi\left(\sqrt{m_i}\frac{\mu_i - \bigwedge_{k=1}^{n} \overline{x_k}}{\sigma_i}\right) > \varepsilon \Phi^{n-1}\left(\frac{\mu_j - \bigwedge_{k=1}^{n} \overline{x_k}}{\bigvee_{k=1}^{n} \sigma_k}\right) \geq \varepsilon \left(\frac{1-\beta}{\varepsilon}\right)^{\frac{n-1}{n}} > 1 - \beta.$$

Example 2.2.3 implies $\mu \in \tilde{R}_{\bigwedge_i \overline{X_i},\beta}$. Together with $\bigwedge_{k=1}^{n} \mu_k = t$, this yields $t \in$ LHS. □

2.2.6 Definition For $\alpha \in]0, 1[$ and $t \in \mathbb{R}$ let

$$\underline{\bigvee \mu}_{\alpha}(t) := -\overline{\bigwedge \mu}_{\alpha}(-t),$$

where $\overline{\bigwedge \mu}_{\alpha}$ is the confidence bound from Remark 2.2.4.

2.2.7 Example Let us consider the designated statistic $\bigvee_i \overline{X_i}$ from Example 2.1.17, and let us set f as in Example 2.1.7 and $g\colon \mathbb{R}^n \to \mathbb{R}^n$, $\mu \mapsto -\mu$. Then $(\bigvee_i \overline{X_i}) \mathbin{\square} \bigotimes_{i=1}^{n} \mathrm{N}_{\mu_i,\sigma_i^2}^{\otimes m_i} = (\bigwedge_i \overline{X_i} \circ f) \mathbin{\square} \bigotimes_{i=1}^{n} \mathrm{N}_{-\mu_i,\sigma_i^2}^{\otimes m_i}$, which yields the confidence regions given by

$$\begin{aligned} R_{\bigvee_i \overline{X_i},\beta}(x) &= g^{-1}[\tilde{R}_{\bigwedge_i \overline{X_i},\beta}(f(x))] \\ &= \left\{\mu \in \mathbb{R}^n \colon \prod_{i=1}^{n} \Phi\left(\sqrt{m_i}\frac{\bigvee_{k=1}^{n} \overline{x_k} - \mu_i}{\sigma_i}\right) > 1 - \beta\right\} \\ \tilde{R}_{\bigvee_i \overline{X_i},\beta}(x) &= g^{-1}[R_{\bigwedge_i \overline{X_i},\beta}(f(x))] \\ &= \left\{\mu \in \mathbb{R}^n \colon \prod_{i=1}^{n} \Phi\left(\sqrt{m_i}\frac{\bigvee_{k=1}^{n} \overline{x_k} - \mu_i}{\sigma_i}\right) < \beta\right\} \end{aligned}$$

for $x \in \prod_{i=1}^{n} \mathbb{R}^{m_i}$, with effective levels

$$\beta_{\mathrm{eff}}(R_{\bigvee_i \overline{X_i},\beta}) = \beta_{\mathrm{eff}}(\tilde{R}_{\bigwedge_i \overline{X_i},\beta}) = \beta = \beta_{\mathrm{eff}}(R_{\bigwedge_i \overline{X_i},\beta}) = \beta_{\mathrm{eff}}(\tilde{R}_{\bigvee_i \overline{X_i},\beta}) \qquad \text{for } \beta \in [0, 1].$$

Lemma 2.2.10 furthermore yields for the parameter of interest

$$\bigvee\colon \mathbb{R}^n \to \mathbb{R}, \quad \mu \mapsto \bigvee_{i=1}^{n} \mu_i,$$

the confidence regions given for $x \in \prod_{i=1}^{n} \mathbb{R}^{m_i}$ by

$$\begin{aligned} \bigvee[R_{\bigvee_i \overline{X_i},\beta}(x)] &= \bigvee[g^{-1}[\tilde{R}_{\bigwedge_i \overline{X_i},\beta}(f(x))]] \\ &= -\bigwedge[\tilde{R}_{\bigwedge_i \overline{X_i},\beta}(f(x))] \\ &= \begin{cases} \left]-\infty, \bigvee_{i=1}^{n} \overline{x_i} + \Phi^{-1}(\beta) \bigvee_{i=1}^{n} \frac{\sigma_i}{\sqrt{m_i}}\right[& \text{if } \beta \in [\frac{1}{2}, 1] \\ \left]-\infty, \bigvee_{i=1}^{n} \overline{x_i} + \Phi^{-1}(\beta) \bigwedge_{i=1}^{n} \frac{\sigma_i}{\sqrt{m_i}}\right[& \text{if } \beta \in [0, \frac{1}{2}] \end{cases} \\ \bigvee[\tilde{R}_{\bigvee_i \overline{X_i},\beta}(x)] &= \bigvee[g^{-1}[R_{\bigwedge_i \overline{X_i},\beta}(f(x))]] \\ &= -\bigwedge[R_{\bigwedge_i \overline{X_i},\beta}(f(x))] \\ &= \left]\underline{\bigvee \mu}_{\beta}(\bigvee_{i=1}^{n} \overline{x_i}), \infty\right[. \end{aligned}$$

2.2.8 Example If we consider the designated statistic $\bigwedge_{i,j} X_{i,j}$ from Example 2.1.13, then the calculations there yield

$$R_{\bigwedge_{i,k} X_{i,k},\beta}(x) = \left\{\mu \in \mathbb{R}^n\colon \prod_{i=1}^{n} \Phi^{m_i}\left(\frac{\mu_i - \bigwedge_{j=1}^{n}\bigwedge_{k=1}^{m_j} x_{j,k}}{\sigma_i}\right) < \beta\right\}$$
$$\tilde{R}_{\bigwedge_{i,k} X_{i,k},\beta}(x) = \left\{\mu \in \mathbb{R}^n\colon \prod_{i=1}^{n} \Phi^{m_i}\left(\frac{\mu_i - \bigwedge_{j=1}^{n}\bigwedge_{k=1}^{m_j} x_{j,k}}{\sigma_i}\right) > 1-\beta\right\}$$

for $x \in \prod_{i=1}^{n} \mathbb{R}^{m_i}$, with effective levels $\beta_{\text{eff}}(R_{\bigwedge_{i,k} X_{i,k},\beta}) = \beta_{\text{eff}}(\tilde{R}_{\bigwedge_{i,k} X_{i,k},\beta}) = \beta$ for $\beta \in [0,1]$.

2.2.9 Remark For $\alpha \in \,]0,1[$ and $t \in \mathbb{R}$ there is exactly one

$$\overline{\bigwedge\bigwedge\mu}_\alpha(t) \in \mathbb{R} \quad \text{such that} \quad \prod_{i=1}^{n} \Phi^{m_i}\left(\frac{\overline{\bigwedge\bigwedge\mu}_\alpha(t) - t}{\sigma_i}\right) = \alpha.$$

In fact, $f\colon \mathbb{R} \to \,]0,1[$, $s \mapsto \prod_{i=1}^{n} \Phi^{m_i}(\frac{s-t}{\sigma_i})$, is bijective since it is strictly increasing and continuous with $\lim_{s\to-\infty} f(s) = 0$ and $\lim_{s\to\infty} f(s) = 1$.

2.2.10 Lemma *Let us consider the parameter of interest*

$$\bigwedge\colon \mathbb{R}^n \to \mathbb{R}, \quad \mu \mapsto \bigwedge_{i=1}^{n} \mu_i.$$

For $x \in \prod_{i=1}^{n} \mathbb{R}^{m_i}$ then

$$\bigwedge[R_{\bigwedge_{i,j} X_{i,j},\beta}(x)] = \left]-\infty, \overline{\bigwedge\bigwedge\mu}_\beta(\textstyle\bigwedge_{i=1}^{n}\bigwedge_{j=1}^{m_i} x_{i,j})\right[\tag{15}$$
$$\bigwedge[\tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(x)] = \left]\bigwedge_{i=1}^{n}\bigwedge_{j=1}^{m_i} x_{i,j} + \bigwedge_{i=1}^{n} \sigma_i \Phi^{-1}\left((1-\beta)^{1/m_i}\right), \infty\right[. \tag{16}$$

Proof. Let $x \in \prod_{i=1}^{n} \mathbb{R}^{m_i}$, and let us assume w.l.o.g. $\beta \in \,]0,1[$ (due to Remarks 1.1.7, part 3, and 1.2.3, part 6).

(15) We first show LHS $\subseteq$ RHS. To this end, let $\mu \in R_{\bigwedge_{i,j} X_{i,j},\beta}(x)$, and let us assume $\bigwedge_{i=1}^{n} \mu_i \geq \sup \text{RHS}$. Remarks 2.1.14, part 1, and 2.2.9 imply

$$\prod_{i=1}^{n} \Phi^{m_i}\left(\frac{\mu_i - \bigwedge_{k=1}^{n}\bigwedge_{r=1}^{m_k} x_{k,r}}{\sigma_i}\right) \geq \prod_{i=1}^{n} \Phi^{m_i}\left(\frac{\sup \text{RHS} - \bigwedge_{k=1}^{n}\bigwedge_{r=1}^{m_k} x_{k,r}}{\sigma_i}\right) = \beta,$$

a contradiction to $\mu \in R_{\bigwedge_{i,j} X_{i,j},\beta}(x)$ in view of Example 2.2.8.

We now show LHS $\supseteq$ RHS. Let $t \in$ RHS, and let us define $\mu_i := t$ for $i \in \{1,\ldots,n\}$. Remarks 2.1.14, part 1, and 2.2.9 imply

$$\prod_{i=1}^{n} \Phi^{m_i}\left(\frac{\mu_i - \bigwedge_{k=1}^{n}\bigwedge_{r=1}^{m_k} x_{k,r}}{\sigma_i}\right) < \prod_{i=1}^{n} \Phi^{m_i}\left(\frac{\sup \text{RHS} - \bigwedge_{k=1}^{n}\bigwedge_{r=1}^{m_k} x_{k,r}}{\sigma_i}\right) = \beta,$$

hence $\mu \in R_{\bigwedge_{i,j} X_{i,j},\beta}(x)$ due to Example 2.2.8. Together with $\bigwedge_{k=1}^{n} \mu_k = t$, this yields $t \in$ LHS.

(16) We first show LHS $\subseteq$ RHS. To this end, let $\mu \in \tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(x)$, and let us assume $\bigwedge_{i=1}^n \mu_i \le \inf \text{RHS}$. Let us pick $i_0 \in \{1,\dots,n\}$ such that $\mu_{i_0} = \bigwedge_{i=1}^n \mu_i$. Remark 2.1.14, part 1, implies

$$\begin{aligned}\prod_{i=1}^n \Phi^{m_i}\left(\frac{\mu_i - \bigwedge_{k=1}^n \bigwedge_{r=1}^{m_k} x_{k,r}}{\sigma_i}\right) &\le \Phi^{m_{i_0}}\left(\frac{\mu_{i_0} - \bigwedge_{k=1}^n \bigwedge_{r=1}^{m_k} x_{k,r}}{\sigma_{i_0}}\right)\\ &\le \Phi^{m_{i_0}}\left(\frac{\bigwedge_{i=1}^n \sigma_i \Phi^{-1}\left((1-\beta)^{1/m_i}\right)}{\sigma_{i_0}}\right)\\ &\le \Phi^{m_{i_0}}\left(\Phi^{-1}\left((1-\beta)^{1/m_{i_0}}\right)\right)\\ &= 1-\beta,\end{aligned}$$

a contradiction to $\mu \in \tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(x)$ in view of Example 2.2.8.
We now show LHS $\supseteq$ RHS. Let $t \in$ RHS, let us pick $i_0 \in \{1,\dots,n\}$ such that

$$\sigma_{i_0}\Phi^{-1}\left((1-\beta)^{1/m_{i_0}}\right) = \bigwedge_{i=1}^n \sigma_i \Phi^{-1}\left((1-\beta)^{1/m_i}\right),$$

and let us define $\mu_{i_0} := t$. Then

$$\begin{aligned}\varepsilon &:= \Phi^{m_{i_0}}\left(\frac{t - \bigwedge_{k=1}^n \bigwedge_{r=1}^{m_k} x_{k,r}}{\sigma_{i_0}}\right)\\ &> \Phi^{m_{i_0}}\left(\frac{\bigwedge_{i=1}^n \sigma_i \Phi^{-1}\left((1-\beta)^{1/m_i}\right)}{\sigma_{i_0}}\right)\\ &= \Phi^{m_{i_0}}\left(\Phi^{-1}\left((1-\beta)^{1/m_{i_0}}\right)\right)\\ &= 1-\beta.\end{aligned}$$

Since

$$\lim_{\substack{\mu_i \to \infty \\ \text{for } i \ne i_0}} \prod_{\substack{i=1 \\ i \ne i_0}}^n \Phi^{m_i}\left(\frac{\mu_i - \bigwedge_{k=1}^n \bigwedge_{r=1}^{m_k} x_{k,r}}{\sigma_i}\right) = 1,$$

we can pick $\mu_i \in]t,\infty[$ for $i \in \{1,\dots,n\} \setminus \{i_0\}$ such that

$$\prod_{\substack{i=1 \\ i \ne i_0}}^n \Phi^{m_i}\left(\frac{\mu_i - \bigwedge_{k=1}^n \bigwedge_{r=1}^{m_k} x_{k,r}}{\sigma_i}\right) > \frac{1-\beta}{\varepsilon} \wedge (1-\beta).$$

Example 2.2.8 implies $\mu \in \tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(x)$. Together with $\bigwedge_{k=1}^n \mu_k = t$, this yields $t \in$ LHS. □

2.2.11 Remark For $\alpha \in]0,1[$ and $t \in \mathbb{R}$ let

$$\underline{\bigvee\bigvee\mu}_{\alpha}(t) := -\overline{\bigwedge\bigwedge\mu}_{\alpha}(-t),$$

where $\overline{\bigwedge\bigwedge\mu}_{\alpha}$ is the confidence bound from Remark 2.2.9.

2.2.12 Example Let us consider the designated statistic $\bigvee_{i,j} X_{i,j}$ from Example 2.1.17, and let us set f and g as in Example 2.2.7. Then $(\bigvee_i \overline{X_i})\square\bigotimes_{i=1}^n \mathrm{N}_{\mu_i,\sigma_i^2}^{\otimes m_i} = (\bigwedge_i \overline{X_i}\circ f)\square\bigotimes_{i=1}^n \mathrm{N}_{-\mu_i,\sigma_i^2}^{\otimes m_i}$, which yields the confidence regions given by

$$\begin{aligned}
R_{\bigvee_{i,j} X_{i,j},\beta}(x) &= g^{-1}[\tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(f(x))] \\
&= \left\{\mu \in \mathbb{R}^n \colon \prod_{i=1}^n \Phi^{m_i}\left(\frac{\bigvee_{j=1}^n \bigvee_{k=1}^{m_j} x_{j,k} - \mu_i}{\sigma_i}\right) > 1-\beta\right\} \\
\tilde{R}_{\bigvee_{i,j} X_{i,j},\beta}(x) &= g^{-1}[R_{\bigwedge_{i,j} X_{i,j},\beta}(f(x))] \\
&= \left\{\mu \in \mathbb{R}^n \colon \prod_{i=1}^n \Phi^{m_i}\left(\frac{\bigvee_{j=1}^n \bigvee_{k=1}^{m_j} x_{j,k} - \mu_i}{\sigma_i}\right) < \beta\right\}
\end{aligned}$$

for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$, with effective levels

$$\beta_{\text{eff}}(R_{\bigvee_{i,j} X_{i,j},\beta}) = \beta_{\text{eff}}(\tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}) = \beta = \beta_{\text{eff}}(R_{\bigwedge_{i,j} X_{i,j},\beta}) = \beta_{\text{eff}}(\tilde{R}_{\bigvee_{i,j} X_{i,j},\beta})$$

for $\beta \in [0,1]$. Lemma 2.2.10 furthermore yields for the parameter of interest

$$\bigvee\colon \mathbb{R}^n \to \mathbb{R}, \quad \mu \mapsto \bigvee_{i=1}^n \mu_i,$$

the confidence regions given for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$ by

$$\begin{aligned}
\bigvee[R_{\bigvee_{i,j} X_{i,j},\beta}(x)] &= \bigvee[g^{-1}[\tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(f(x))]] \\
&= -\bigwedge[\tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(f(x))] \\
&= \left]-\infty, \bigvee_{i=1}^n \bigvee_{j=1}^{m_i} x_{i,j} - \bigwedge_{i=1}^n \sigma_i \Phi^{-1}\left((1-\beta)^{1/m_i}\right)\right[\\
\bigvee[\tilde{R}_{\bigvee_{i,j} X_{i,j},\beta}(x)] &= \bigvee[g^{-1}[R_{\bigwedge_{i,j} X_{i,j},\beta}(f(x))]] \\
&= -\bigwedge[R_{\bigwedge_{i,j} X_{i,j},\beta}(f(x))] \\
&= \left]\underline{\bigvee\bigvee\mu}_{\beta}(\bigvee_{i=1}^n \bigvee_{j=1}^{m_i} x_{i,j}), \infty\right[.
\end{aligned}$$

2.2.13 Example If we consider the designated statistic $\overline{X}$ from Example 2.1.18, then the calculations there yield

$$\begin{aligned}
R_{\overline{X},\beta}(x) &= \left\{\mu \in \mathbb{R}^n \colon \overline{\mu} < \overline{x} + \sqrt{\frac{\overline{\sigma^2/m}}{n}}\Phi^{-1}(\beta)\right\} \\
\tilde{R}_{\overline{X},\beta}(x) &= \left\{\mu \in \mathbb{R}^n \colon \overline{\mu} > \overline{x} - \sqrt{\frac{\overline{\sigma^2/m}}{n}}\Phi^{-1}(\beta)\right\}
\end{aligned}$$

for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$, with effective levels $\beta_{\text{eff}}(R_{\overline{X},\beta}) = \beta_{\text{eff}}(\tilde{R}_{\overline{X},\beta}) = \beta$ for $\beta \in [0,1]$. This immediately yields for the parameter of interest $\kappa\colon \mathbb{R}^n \to \mathbb{R}$, $\mu \mapsto \overline{\mu} = \sum_{i=1}^n \mu_i/n$, the confidence regions given for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$ by

$$\kappa[R_{\overline{X},\beta}(x)] = \left]-\infty, \overline{x} + \sqrt{\frac{\overline{\sigma^2/m}}{n}}\Phi^{-1}(\beta)\right[$$

$$\kappa[\tilde{R}_{\overline{X},\beta}(x)] = \left]\overline{x} - \sqrt{\frac{\sigma^2/m}{n}}\Phi^{-1}(\beta), \infty\right[.$$

2.3 SEVERAL HOMOGENEOUS SAMPLES WITH KNOWN VARIANCES

2.3.1 Definition Let $n \in \mathbb{N}$, $m \in \mathbb{N}^n$, and $\sigma \in]0,\infty[^n$, and let

$$\mathcal{P}_3 := \left(\bigotimes_{i=1}^{n} \mathrm{N}_{\mu,\sigma_i^2}^{\otimes m_i} : \mu \in \mathbb{R}\right)$$

be the n-sample homogeneous normal model with known variances $\sigma_1^2, \ldots, \sigma_n^2$ and known sample sizes $m_1, \ldots, m_n$.

2.3.2 Remark 1. The sample space is $\mathcal{X} = \prod_{i=1}^n \mathbb{R}^{m_i}$, the parameter space $\Theta = \mathbb{R}$.
2. $\mathcal{P}_3$ is stochastically increasing.
3. The following model is a special case of $\mathcal{P}_3$:

$$\mathcal{P}_3' := (\mathrm{N}_{\mu,\sigma^2}^{\otimes n} : \mu \in \mathbb{R}) \qquad \text{for } \sigma \in]0,\infty[,$$

the one-sample normal model with known variance σ^2 and known sample size n. The sample space in this case is $\mathcal{X} = \mathbb{R}^n$.

2.3.3 Example If we consider the designated statistic $\overline{X}$ from Example 2.1.18, then the calculations in Example 2.2.13 yield for the identity $\mathrm{id}_{\mathbb{R}}$ as parameter of interest the confidence regions given by

$$R_{\overline{X},\beta}(x) = \left]-\infty, \overline{x} + \sqrt{\sigma^2/m}\Phi^{-1}(\beta)\right[$$
$$\tilde{R}_{\overline{X},\beta}(x) = \left]\overline{x} - \sqrt{\sigma^2/m}\Phi^{-1}(\beta), \infty\right[$$

for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$, with effective levels $\beta_{\mathrm{eff}}(R_{\overline{X},\beta}) = \beta_{\mathrm{eff}}(\tilde{R}_{\overline{X},\beta}) = \beta$ for $\beta \in [0,1]$.

In the model $\mathcal{P}_3'$ from the previous remark, the Buehlerizations of $\overline{X} \colon \mathbb{R}^n \to \mathbb{R}$, $x \mapsto \overline{x} := \sum_{i=1}^n x_i/n$, yield for the same parameter of interest the well-known confidence regions given for $x \in \mathbb{R}^n$ by

$$R_{\overline{X},\beta,\mathcal{P}_3'}(x) = \left]-\infty, \overline{x} + \frac{\sigma}{\sqrt{n}}\Phi^{-1}(\beta)\right[\qquad \tilde{R}_{\overline{X},\beta,\mathcal{P}_3'}(x) = \left]\overline{x} - \frac{\sigma}{\sqrt{n}}\Phi^{-1}(\beta), \infty\right[.$$

2.4 SEVERAL SAMPLES WITH KNOWN MEANS

2.4.1 Definition Let $n \in \mathbb{N}$, $m \in (\mathbb{N} \setminus \{1\})^n$, and $\mu \in \mathbb{R}^n$, and let

$$\mathcal{P}_4 := \left(\bigotimes_{i=1}^{n} \mathrm{N}_{\mu_i,\sigma_i^2}^{\otimes m_i} \colon \sigma \in \left]0,\infty\right[^n \right)$$

be the n-sample normal model with known means $\mu_1, \ldots, \mu_n$ and known sample sizes $m_1, \ldots, m_n$.

2.4.2 Remark 1. The sample space is $\mathcal{X} = \prod_{i=1}^n \mathbb{R}^{m_i}$, the parameter space $\Theta = \left]0,\infty\right[^n$.
2. $\mathcal{P}_2$ is not stochastically monotonic.
3. The following model is a special case of $\mathcal{P}_4$:

$$\mathcal{P}_4' := \left(\bigotimes_{i=1}^{n} \mathrm{N}_{\mu,\sigma_i^2}^{\otimes m_i} \colon \sigma \in \left]0,\infty\right[^n \right) \qquad \text{for } \mu \in \mathbb{R},$$

the n-sample homogeneous normal model with known mean μ and known sample sizes $m_1, \ldots, m_n$.

2.4.3 Example If we set S_i as in Example 2.1.8, then Remark A.3.54, part 1, yields $((m_i - 1)S_i^2/\sigma_i^2) \,\square\, \mathrm{N}_{\mu_i,\sigma_i^2}^{\otimes m_i} = \chi^2_{m_i-1}$ for $i \in \{1, \ldots, n\}$. Since these laws are continuous, the Buehlerizations of $\bigwedge_i S_i^2 \colon \prod_{i=1}^n \mathbb{R}^{m_i} \to [0,\infty[$, $x \mapsto \bigwedge_{i=1}^n S_i^2(x)$, are given by

$$R_{\bigwedge_i S_i^2,\beta}(x) = \left\{ \sigma \in \left]0,\infty\right[^n \colon \prod_{i=1}^{n} \chi^2_{m_i-1} \left(\left] \frac{m_i - 1}{\sigma_i^2} \bigwedge_{k=1}^{n} S_k^2(x), \infty \right[\right) < \beta \right\}$$
$$\tilde{R}_{\bigwedge_i S_i^2,\beta}(x) = \left\{ \sigma \in \left]0,\infty\right[^n \colon \prod_{i=1}^{n} \chi^2_{m_i-1} \left(\left] \frac{m_i - 1}{\sigma_i^2} \bigwedge_{k=1}^{n} S_k^2(x), \infty \right[\right) > 1 - \beta \right\}$$

for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$ in view of Remark 1.1.7, part 10. Let us note that $R_{\bigwedge_i S_i^2,\beta}(x) = \emptyset$ and $\tilde{R}_{\bigwedge_i S_i^2,\beta}(x) = \left]0,\infty\right[^n$ if x contains at least one constant vector $x_i = (x_{i,1}, \ldots, x_{i,m_i})$. Since

$$\prod_{i=1}^{n} \chi^2_{m_i-1} \left(\left] \frac{m_i - 1}{\sigma_i^2} \bigwedge_{k=1}^{n} S_k^2, \infty \right[\right) \colon \prod_{i=1}^{n} \mathbb{R}^{m_i} \to \left]0,1\right[$$

is surjective, the effective levels of these confidence regions are given by

$$\beta_{\mathrm{eff}}(R_{\bigwedge_i S_i^2,\beta}) = \beta_{\mathrm{eff}}(\tilde{R}_{\bigwedge_i S_i^2,\beta}) = \beta \qquad \text{for } \beta \in [0,1].$$

2.4.4 Remark 1. We shall employ in the proof of Lemma 2.4.5 the fact that

$$\left]0,\infty\right[^n \to \left]0,1\right[, \quad \sigma \mapsto \prod_{i=1}^{n} \chi^2_{m_i-1} \left(\left] \frac{m_i - 1}{\sigma_i^2} \bigwedge_{k=1}^{n} S_k^2(x), \infty \right[\right),$$

is, for fixed $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$,

- strictly increasing if each vector $x_i = (x_{i,1}, \ldots, x_{i,m_i})$ is not constant,
- constantly 1 if there is a constant vector $x_i = (x_{i,1}, \ldots, x_{i,m_i})$.

2. For $\alpha \in]0,1[$ and $t \in]0,\infty[$ there is exactly one

$$\overline{\bigwedge\sigma}_{\alpha}(t) \in]0,\infty[\quad \text{such that} \quad \prod_{i=1}^{n} \chi^2_{m_i-1}\left(\left]\frac{m_i-1}{(\overline{\bigwedge\sigma}_{\alpha}(t))^2}t,\infty\right[\right) = \alpha.$$

In fact, $f\colon]0,\infty[\to]0,1[$, $s \mapsto \prod_{i=1}^{n} \chi^2_{m_i-1}(]\frac{m_i-1}{s^2}t,\infty[)$, is bijective since it is strictly increasing and continuous with $\lim_{s\to 0} f(s) = 0$ and $\lim_{s\to\infty} f(s) = 1$. Let us furthermore define $\overline{\bigwedge\sigma}_{\alpha}(0) := 0$.

3. For $i \in \{1,\dots,n\}$, $\alpha \in]0,1[$, and $t \in]0,\infty[$ there is exactly one

$$\underline{\sigma_i}_{\alpha}(t) \in]0,\infty[\quad \text{such that} \quad \chi^2_{m_i-1}\left(\left]\frac{m_i-1}{(\underline{\sigma_i}_{\alpha}(t))^2}t,\infty\right[\right) = \alpha,$$

namely $\underline{\sigma_i}_{\alpha}(t) := \sqrt{(m_i-1)t/F_i^{-1}(1-\alpha)}$, where F_i denotes the distribution function of the law $\chi^2_{m_i-1}$. Let us furthermore define $\underline{\sigma_i}_{\alpha}(0) = 0$.

2.4.5 Lemma *Let us consider the parameters of interest*

$$\bigwedge\colon]0,\infty[^n \to]0,\infty[, \quad \sigma \mapsto \bigwedge_{i=1}^{n} \sigma_i, \qquad \bigvee\colon]0,\infty[^n \to]0,\infty[, \quad \sigma \mapsto \bigvee_{i=1}^{n} \sigma_i.$$

For $x \in \prod_{i=1}^{n} \mathbb{R}^{m_i}$ *then*

$$\bigwedge[R_{\bigwedge_i S_i^2,\beta}(x)] = \left]0, \overline{\bigwedge\sigma}_{\beta}(\textstyle\bigwedge_{i=1}^{n} S_i^2(x))\right[\tag{17}$$

$$\bigwedge[\tilde{R}_{\bigwedge_i S_i^2,\beta}(x)] = \left]\bigwedge_{i=1}^{n} \underline{\sigma_i}_{1-\beta}(\textstyle\bigwedge_{k=1}^{n} S_k^2(x)),\infty\right[\tag{18}$$

$$\bigvee[R_{\bigwedge_i S_i^2,\beta}(x)] = \begin{cases}]0,\infty[& \textit{if } n \geq 2 \textit{ and no } x_i \textit{ is constant} \\ \emptyset & \textit{if } n \geq 2 \textit{ and some } x_i \textit{ is constant} \\ \left]0, \underline{\sigma_1}_{\beta}(S_1^2(x))\right[& \textit{if } n = 1 \end{cases} \tag{19}$$

$$\bigvee[\tilde{R}_{\bigwedge_i S_i^2,\beta}(x)] = \left]\overline{\bigwedge\sigma}_{1-\beta}(\textstyle\bigwedge_{k=1}^{n} S_k^2(x)),\infty\right[. \tag{20}$$

Proof. Let $x \in \prod_{i=1}^{n} \mathbb{R}^{m_i}$, and let us assume w.l.o.g. $\beta \in]0,1[$ (due to Remarks 1.1.7, part 3, and 1.2.3, part 6). If $x_i = (x_{i,1},\dots,x_{i,m_i})$ is constant for some $i \in \{1,\dots,n\}$, then $\bigwedge_{k=1}^{n} S_k^2(x) = 0$, hence the claims are clear by Example 2.4.3 and Remark 2.4.4, parts 2 and 3. Let us therefore suppose that each vector $x_i = (x_{i,1},\dots,x_{i,m_i})$ is not constant, that is, $\bigwedge_{k=1}^{n} S_k^2(x) > 0$.

(17) We first show LHS $\subseteq$ RHS. Let $\sigma \in R_{\bigwedge_i S_i^2,\beta}(x)$, and let us assume $\bigwedge_{i=1}^{n} \sigma_i \geq \sup \text{RHS}$. Remark 2.4.4, parts 1 and 2, implies

$$\prod_{i=1}^{n} \chi^2_{m_i-1}\left(\left]\frac{m_i-1}{\sigma_i^2}\bigwedge_{k=1}^{n} S_k^2(x),\infty\right[\right) \geq \prod_{i=1}^{n} \chi^2_{m_i-1}\left(\left]\frac{m_i-1}{(\sup \text{RHS})^2}\bigwedge_{k=1}^{n} S_k^2(x),\infty\right[\right) = \beta,$$

a contradiction to $\sigma \in R_{\bigwedge_i S_i^2,\beta}(x)$ in view of Example 2.4.3.

We now show LHS $\supseteq$ RHS. Let $t \in$ RHS, and let us define $\sigma_i := t$ for $i \in \{1, \ldots, n\}$. Remark 2.4.4, part 1, then yields

$$\prod_{i=1}^{n} \chi^2_{m_i-1}\left(\left]\frac{m_i-1}{\sigma_i^2}\bigwedge_{k=1}^{n} S_k^2(x), \infty\right[\right) < \prod_{i=1}^{n} \chi^2_{m_i-1}\left(\left]\frac{m_i-1}{(\sup \mathrm{RHS})^2}\bigwedge_{k=1}^{n} S_k^2(x), \infty\right[\right) = \beta,$$

hence $\sigma \in R_{\bigwedge_i S_i^2,\beta}(x)$ due to Example 2.4.3. Together with $\bigwedge_{k=1}^n \sigma_k = t$, this yields $t \in$ LHS.

(18) We first show LHS $\subseteq$ RHS. Let $\sigma \in \tilde{R}_{\bigwedge_i S_i^2,\beta}(x)$, and let us assume $\bigwedge_{i=1}^n \sigma_i \leq \inf$ RHS. Let us pick $i_0 \in \{1, \ldots, n\}$ such that $\sigma_{i_0} = \bigwedge_{i=1}^n \sigma_i$. Remark 2.4.4, parts 1 and 3, yields

$$\begin{aligned}\prod_{i=1}^{n} \chi^2_{m_i-1}\left(\left]\frac{m_i-1}{\sigma_i^2}\bigwedge_{k=1}^{n} S_k^2(x), \infty\right[\right) &\leq \chi^2_{m_{i_0}-1}\left(\left]\frac{m_{i_0}-1}{\sigma_{i_0}^2}\bigwedge_{k=1}^{n} S_k^2(x), \infty\right[\right)\\ &\leq \chi^2_{m_{i_0}-1}\left(\left]\frac{m_{i_0}-1}{(\inf \mathrm{RHS})^2}\bigwedge_{k=1}^{n} S_k^2(x), \infty\right[\right)\\ &\leq \chi^2_{m_{i_0}-1}\left(\left]\frac{(m_{i_0}-1)\bigwedge_{k=1}^{n} S_k^2(x)}{(\underline{\sigma_{i_0}}_{1-\beta}(\bigwedge_{k=1}^n S_k^2(x)))^2}, \infty\right[\right)\\ &= 1-\beta,\end{aligned}$$

a contradiction to $\sigma \in \tilde{R}_{\bigwedge_i S_i^2,\beta}(x)$ due to Example 2.4.3.
We now show LHS $\supseteq$ RHS. Let $t \in$ RHS, and let us pick $i_0 \in \{1, \ldots, n\}$ such that

$$\underline{\sigma_{i_0}}_{1-\beta}(\textstyle\bigwedge_{k=1}^n S_k^2(x)) = \displaystyle\bigwedge_{i=1}^{n} \underline{\sigma_i}_{1-\beta}(\textstyle\bigwedge_{k=1}^n S_k^2(x)).$$

Since

$$\begin{aligned}\lim_{\substack{\sigma_i\to\infty\\ \text{for } i\neq i_0}} \prod_{i=1}^{n} \chi^2_{m_i-1}\left(\left]\frac{m_i-1}{\sigma_i^2}\bigwedge_{k=1}^{n} S_k^2, \infty\right[\right) &= \chi^2_{m_{i_0}-1}\left(\left]\frac{m_{i_0}-1}{\sigma_{i_0}^2}\bigwedge_{k=1}^{n} S_k^2, \infty\right[\right)\\ &> \chi^2_{m_{i_0}-1}\left(\left]\frac{m_{i_0}-1}{(\inf \mathrm{RHS})^2}\bigwedge_{k=1}^{n} S_k^2, \infty\right[\right)\\ &= 1-\beta\end{aligned}$$

in view of Remark 2.4.4, part 1, we can choose $\sigma_i \in]t, \infty[$ for $i \in \{1, \ldots, n\} \setminus \{i_0\}$ such that $\sigma \in \tilde{R}_{\bigwedge_i S_i^2,\beta}(x)$ due to Example 2.4.3. Together with $\bigwedge_{i=1}^n \sigma_i = t$, this yields $t \in$ LHS.

(19) If $n \geq 2$ and $x = (x_1, \ldots, x_n)$ contains at least one constant vector $x_i = (x_{i,1}, \ldots, x_{i,m_i})$, then Example 2.4.3 yields $R_{\bigwedge_i S_i^2,\beta}(x) = \emptyset$, which implies the claim. Let us now consider the case $n \geq 2$, and let us assume furthermore that x contains no constant vector x_i. The inclusion LHS $\subseteq$ RHS is clear. It thus remains to show LHS $\supseteq$ RHS. To this end, let $t \in]0, \infty[$, and let us define $\sigma_1 := t$. Since

$$\lim_{\sigma_2\to 0+} \prod_{i=1}^{n} \chi^2_{m_i-1}\left(\left]\frac{m_i-1}{\sigma_i^2}\bigwedge_{k=1}^{n} S_k^2(x), \infty\right[\right) = 0 < \beta,$$

for $\sigma_3, \dots, \sigma_n \in]0,\infty[$, we can choose $\sigma_2, \dots, \sigma_n \in]0,t[$ such that $\sigma \in R_{\bigwedge_i S_i^2,\beta}(x)$ in view of Example 2.4.3. Together with $\bigvee_{i=1}^n \sigma_i = t$, this yields $t \in$ LHS.
Let us now consider the case $n = 1$. Example 2.4.3 and Remark 2.4.4, part 1, yield the equivalence

$$\sigma \in R_{S_1^2,\beta}(x) \quad \iff \quad \sigma < \underline{\sigma_1}_\beta(S_1^2(x)),$$

which implies the claim.

(20) We first show LHS $\subseteq$ RHS. To this end, let $\sigma \in \tilde{R}_{\bigwedge_i S_i^2,\beta}(x)$, and let us assume $\bigvee_{i=1}^n \sigma_i \le \inf \text{RHS}$. Remark 2.4.4, parts 1 and 3, then yields

$$\begin{aligned}\prod_{i=1}^n \chi^2_{m_i-1}\left(\left]\frac{m_i-1}{\sigma_i^2}\bigwedge_{k=1}^n S_k^2(x),\infty\right[\right) &\le \prod_{i=1}^n \chi^2_{m_i-1}\left(\left]\frac{m_i-1}{(\inf \text{RHS})^2}\bigwedge_{k=1}^n S_k^2(x),\infty\right[\right)\\ &= 1-\beta,\end{aligned}$$

a contradiction to $\sigma \in \tilde{R}_{\bigwedge_i S_i^2,\beta}(x)$ due to Example 2.4.3.
We now show LHS $\supseteq$ RHS. Let $t \in$ RHS, and let us define $\sigma_i \coloneqq t$ for $i \in \{1,\dots,n\}$. Remark 2.4.4, parts 1 and 3, then yields

$$\begin{aligned}\prod_{i=1}^n \chi^2_{m_i-1}\left(\left]\frac{m_i-1}{\sigma_i^2}\bigwedge_{k=1}^n S_k^2(x),\infty\right[\right) &> \prod_{i=1}^n \chi^2_{m_i-1}\left(\left]\frac{m_i-1}{(\inf \text{RHS})^2}\bigwedge_{k=1}^n S_k^2(x),\infty\right[\right)\\ &= 1-\beta,\end{aligned}$$

which implies $\sigma \in \tilde{R}_{\bigwedge_i S_i^2,\beta}(x)$ due to Example 2.4.3. Together with $\bigvee_{i=1}^n \sigma_i = t$, this yields $t \in$ LHS. □

2.4.6 Example If we set S_i as in Example 2.1.8, then Remark A.3.54, part 1, yields $((m_i-1)S_i^2/\sigma_i^2) \,\square\, \mathrm{N}_{\mu_i,\sigma_i^2}^{\otimes m_i} = \chi^2_{m_i-1}$ for $i \in \{1,\dots,n\}$. Since these laws are continuous, the Buehlerizations of $\bigvee_i S_i^2 \colon \prod_{i=1}^n \mathbb{R}^{m_i} \to [0,\infty[,\ x \mapsto \bigvee_{i=1}^n S_i^2(x)$, are given by

$$\begin{aligned}R_{\bigvee_i S_i^2,\beta}(x) &= \left\{\sigma \in]0,\infty[^n \colon \prod_{i=1}^n \chi^2_{m_i-1}\left(\left]0,\frac{m_i-1}{\sigma_i^2}\bigvee_{k=1}^n S_k^2(x)\right[\right) > 1-\beta\right\}\\ \tilde{R}_{\bigvee_i S_i^2,\beta}(x) &= \left\{\sigma \in]0,\infty[^n \colon \prod_{i=1}^n \chi^2_{m_i-1}\left(\left]0,\frac{m_i-1}{\sigma_i^2}\bigvee_{k=1}^n S_k^2(x)\right[\right) < \beta\right\}\end{aligned}$$

for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$ in view of Remark 1.1.7, part 10. Let us note that $R_{\bigvee_i S_i^2,\beta}(x) = \emptyset$ and $\tilde{R}_{\bigvee_i S_i^2,\beta}(x) =]0,\infty[^n$ if x consists of solely constant vectors $x_i = (x_{i,1},\dots,x_{i,m_i})$. Since

$$\prod_{i=1}^n \chi^2_{m_i-1}\left(\left]0,\frac{m_i-1}{\sigma_i^2}\bigvee_{k=1}^n S_k^2\right[\right) \colon \prod_{i=1}^n \mathbb{R}^{m_i} \to]0,1[$$

is surjective, the effective levels of these confidence regions are given by

$$\beta_{\text{eff}}(R_{\bigvee_i S_i^2,\beta}) = \beta_{\text{eff}}(\tilde{R}_{\bigvee_i S_i^2,\beta}) = \beta \qquad \text{for } \beta \in [0,1].$$

2.4.7 Remark 1. We shall employ in the proof of Lemma 2.4.8 the fact that

$$]0,\infty[^n \to]0,1[, \quad \sigma \mapsto \prod_{i=1}^n \chi^2_{m_i-1}\left(\left]0,\frac{m_i-1}{\sigma_i^2}\bigvee_{k=1}^n S_k^2(x)\right[\right),$$

is, for fixed $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$,

- strictly decreasing if at least one vector $x_i = (x_{i,1}, \dots, x_{i,m_i})$ is not constant,
- constantly 0 if every vector $x_i = (x_{i,1}, \dots, x_{i,m_i})$ is constant.

2. For $\alpha \in]0,1[$ and $t \in]0,\infty[$ there is exactly one

$$\underline{\bigvee \sigma}_\alpha(t) \in]0,\infty[\quad \text{such that} \quad \prod_{i=1}^n \chi^2_{m_i-1}\left(\left]0, \frac{m_i-1}{(\underline{\bigvee\sigma}_\alpha(t))^2} t\right[\right) = \alpha.$$

In fact, $f\colon]0,\infty[\to]0,1[$, $s \mapsto \prod_{i=1}^n \chi^2_{m_i-1}(]0, \frac{m_i-1}{s^2}t[)$, is bijective since it is strictly decreasing and continuous with $\lim_{s\to 0} f(s) = 1$ and $\lim_{s\to\infty} f(s) = 0$. Let us furthermore define $\underline{\bigvee\sigma}_\alpha(0) \coloneqq 0$.

2.4.8 Lemma *Let us consider the parameters of interest $\bigwedge$ and $\bigvee$ from Lemma 2.4.5. With the confidence bounds $\underline{\sigma_i}_\beta$ from Remark 2.4.4, part 3, we have for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$*

$$\bigwedge[R_{\bigvee_i S_i^2,\beta}(x)] = \left]0, \underline{\bigvee\sigma}_{1-\beta}(\textstyle\bigvee_{i=1}^n S_i^2(x))\right[\tag{21}$$

$$\bigwedge[\tilde{R}_{\bigvee_i S_i^2,\beta}(x)] = \begin{cases}]0,\infty[& \text{if } n \geq 2 \\ \left]\underline{\sigma_1}_{1-\beta}(S_1^2(x)), \infty\right[& \text{if } n = 1 \end{cases} \tag{22}$$

$$\bigvee[R_{\bigvee_i S_i^2,\beta}(x)] = \left]0, \bigvee_{i=1}^n \underline{\sigma_i}_\beta(\textstyle\bigvee_{i=1}^n S_i^2(x))\right[\tag{23}$$

$$\bigvee[\tilde{R}_{\bigvee_i S_i^2,\beta}(x)] = \left]\underline{\bigvee\sigma}_\beta(\textstyle\bigvee_{k=1}^n S_k^2(x)), \infty\right[. \tag{24}$$

Proof. Let $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$, and let us assume w.l.o.g. $\beta \in]0,1[$ (due to Remarks 1.1.7, part 3, and 1.2.3, part 6). If $x_i = (x_{i,1}, \dots, x_{i,m_i})$ is constant for every $i \in \{1,\dots,n\}$, then $\bigvee_{k=1}^n S_k^2(x) = 0$, hence the claims are clear by Example 2.4.6 and Remarks 2.4.4, parts 2 and 3, and 2.4.7. Let us therefore suppose that at least one vector $x_i = (x_{i,1}, \dots, x_{i,m_i})$ is not constant, that is, $\bigvee_{k=1}^n S_k^2(x) > 0$.

(21) We first show LHS $\subseteq$ RHS. To this end, let $\sigma \in R_{\bigvee_i S_i^2,\beta}(x)$, and let us assume $\bigwedge_{i=1}^n \sigma_i \geq \sup \text{RHS}$. Remark 2.4.7 then implies

$$\begin{aligned}\prod_{i=1}^n \chi^2_{m_i-1}\left(\left]0, \frac{m_i-1}{\sigma_i^2} \bigvee_{k=1}^n S_k^2(x)\right[\right) &\leq \prod_{i=1}^n \chi^2_{m_i-1}\left(\left]0, \frac{m_i-1}{(\sup \text{RHS})^2} \bigvee_{k=1}^n S_k^2(x)\right[\right) \\ &= 1-\beta,\end{aligned}$$

a contradiction to $\sigma \in R_{\bigvee_i S_i^2,\beta}(x)$ in view of Example 2.4.6.

We now show LHS $\supseteq$ RHS. Let $t \in$ RHS, and let us define $\sigma_i \coloneqq t$ for $i \in \{1,\dots,n\}$. Remark 2.4.7 then implies

$$\begin{aligned}\prod_{i=1}^n \chi^2_{m_i-1}\left(\left]0, \frac{m_i-1}{\sigma_i^2} \bigvee_{k=1}^n S_k^2(x)\right[\right) &> \prod_{i=1}^n \chi^2_{m_i-1}\left(\left]0, \frac{m_i-1}{(\sup \text{RHS})^2} \bigvee_{k=1}^n S_k^2(x)\right[\right) \\ &= 1-\beta,\end{aligned}$$

yielding $\sigma \in R_{\bigvee_i S_i^2,\beta}(x)$ due to Example 2.4.6. Together with $\bigwedge_{i=1}^n \sigma_i = t$, this yields $t \in$ LHS.

(22) Let us first consider the case $n \geq 2$. The inclusion LHS $\subseteq$ RHS is clear. It therefore remains to show LHS $\supseteq$ RHS. To this end, let $t \in]0,\infty[$ and let us define $\sigma_1 = t$. Since

$$\lim_{\sigma_2\to\infty} \prod_{i=1}^{n} \chi^2_{m_i-1}\left(\left]0, \frac{m_i-1}{\sigma_i^2}\bigvee_{k=1}^{n} S_k^2(x)\right]\right) = 0 < \beta,$$

for $\sigma_3,\dots,\sigma_n \in]0,\infty[$, we can choose $\sigma_2,\dots,\sigma_n \in]t,\infty[$ such that $\sigma \in \tilde{R}_{\bigvee_i S_i^2,\beta}(x)$ in view of Example 2.4.6. Together with $\bigwedge_{i=1}^{n} \sigma_i = t$, this yields $t \in$ LHS.
Let us now consider the case $n = 1$. Example 2.4.6 and Remark 2.4.7, part 1, yield the equivalence

$$\sigma \in R_{S_1^2,\beta}(x) \quad \Longleftrightarrow \quad \sigma > \underline{\sigma_1}_{1-\beta}(S_1^2(x)),$$

which implies the claim.

(23) We first show LHS $\subseteq$ RHS. To this end, let $\sigma \in R_{\bigvee_i S_i^2,\beta}(x)$, and let us assume $\bigvee_{i=1}^{n} \sigma_i \geq \sup \text{RHS}$. Let us pick $i_0 \in \{1,\dots,n\}$ such that $\sigma_{i_0} = \bigvee_{i=1}^{n} \sigma_i$. Remarks 2.4.4, part 3, and 2.4.7, part 1, then imply

$$\begin{aligned}
&\prod_{i=1}^{n} \chi^2_{m_i-1}\left(\left]0, \frac{m_i-1}{\sigma_i^2}\bigvee_{k=1}^{n} S_k^2(x)\right]\right)\\
&\quad\leq \chi^2_{m_{i_0}-1}\left(\left]0, \frac{m_{i_0}-1}{\sigma_{i_0}^2}\bigvee_{k=1}^{n} S_k^2(x)\right]\right)\\
&\quad\leq \chi^2_{m_{i_0}-1}\left(\left]0, \frac{m_{i_0}-1}{(\sup \text{RHS})^2}\bigvee_{k=1}^{n} S_k^2(x)\right]\right)\\
&\quad\leq \chi^2_{m_{i_0}-1}\left(\left]0, \frac{m_{i_0}-1}{(\underline{\sigma_{i_0}}_\beta(\bigvee_{k=1}^{n} S_k^2(x)))^2}\bigvee_{k=1}^{n} S_k^2(x)\right]\right)\\
&\quad= 1 - \chi^2_{m_{i_0}-1}\left(\left]\frac{m_{i_0}-1}{(\underline{\sigma_{i_0}}_\beta(\bigvee_{k=1}^{n} S_k^2(x)))^2}\bigvee_{k=1}^{n} S_k^2(x), \infty\right]\right)\\
&\quad= 1-\beta,
\end{aligned}$$

a contradiction to $\sigma \in R_{\bigvee_i S_i^2,\beta}(x)$ in view of Example 2.4.6.
We now show LHS $\supseteq$ RHS. Let $t \in$ RHS, and let us pick $i_0 \in \{1,\dots,n\}$ such that

$$\underline{\sigma_{i_0}}_\beta(\textstyle\bigvee_{k=1}^{n} S_k^2(x)) = \displaystyle\bigvee_{i=1}^{n} \underline{\sigma_i}_\beta(\textstyle\bigvee_{k=1}^{n} S_k^2(x)).$$

Let us define $\sigma_{i_0} := t$. Since

$$\begin{aligned}
&\lim_{\substack{\sigma_i\to 0+\\ \text{for } i\neq i_0}} \prod_{i=1}^{n} \chi^2_{m_i-1}\left(\left]0, \frac{m_i-1}{\sigma_i^2}\bigvee_{k=1}^{n} S_k^2(x)\right]\right)\\
&\qquad= \chi^2_{m_{i_0}-1}\left(\left]0, \frac{m_{i_0}-1}{\sigma_{i_0}^2}\bigvee_{k=1}^{n} S_k^2(x)\right]\right)\\
&\qquad> \chi^2_{m_{i_0}-1}\left(\left]0, \frac{m_{i_0}-1}{(\sup \text{RHS})^2}\bigvee_{k=1}^{n} S_k^2(x)\right]\right)
\end{aligned}$$

$$
\begin{aligned}
&= \chi^2_{m_{i_0}-1}\left(\left]0, \frac{m_{i_0}-1}{(\underline{\sigma_{i_0}}_\beta(\bigvee_{k=1}^n S_k^2(x)))^2}\bigvee_{k=1}^n S_k^2(x)\right[\right) \\
&= 1-\chi^2_{m_{i_0}-1}\left(\left]\frac{m_{i_0}-1}{(\underline{\sigma_{i_0}}_\beta(\bigvee_{k=1}^n S_k^2(x)))^2}\bigvee_{k=1}^n S_k^2(x), \infty\right[\right) \\
&= 1-\beta,
\end{aligned}
$$

we can pick $\sigma_i \in\,]0,t[$ for $i\{1,\dots,n\}\setminus\{i_0\}$ such that $\sigma \in R_{\bigvee_i S_i^2,\beta}(x)$ due to Example 2.4.6. Together with $\bigvee_{i=1}^n \sigma_i = t$, this yields $t \in \mathrm{LHS}$.

(24) We first show LHS $\subseteq$ RHS. To this end, let $\sigma \in \tilde{R}_{\bigvee_i S_i^2,\beta}(x)$, and let us assume $\bigvee_{i=1}^n \sigma_i \leq \inf \mathrm{RHS}$. Remark 2.4.7 then yields

$$
\begin{aligned}
\prod_{i=1}^n \chi^2_{m_i-1}\left(\left]0, \frac{m_i-1}{\sigma_i^2}\bigvee_{k=1}^n S_k^2(x)\right[\right) &\geq \prod_{i=1}^n \chi^2_{m_i-1}\left(\left]0, \frac{m_i-1}{(\inf \mathrm{RHS})^2}\bigvee_{k=1}^n S_k^2(x)\right[\right) \\
&= \beta,
\end{aligned}
$$

a contradiction to $\sigma \in \tilde{R}_{\bigvee_i S_i^2,\beta}(x)$ due to Example 2.4.6.
We now show LHS $\supseteq$ RHS. Let $t \in \mathrm{RHS}$, and let us define $\sigma_i \coloneqq t$ for $i \in \{1,\dots,n\}$. Remark 2.4.7 then yields

$$
\begin{aligned}
\prod_{i=1}^n \chi^2_{m_i-1}\left(\left]0, \frac{m_i-1}{\sigma_i^2}\bigvee_{k=1}^n S_k^2(x)\right[\right) &< \prod_{i=1}^n \chi^2_{m_i-1}\left(\left]0, \frac{m_i-1}{(\inf \mathrm{RHS})^2}\bigvee_{k=1}^n S_k^2(x)\right[\right) \\
&= \beta,
\end{aligned}
$$

which implies $\sigma \in \tilde{R}_{\bigvee_i S_i^2,\beta}(x)$ due to Example 2.4.6. Together with $\bigvee_{i=1}^n \sigma_i = t$, this yields $t \in \mathrm{LHS}$. □

2.4.9 Example If we set

$$
\tilde{S}_i\colon \prod_{k=1}^n \mathbb{R}^{m_k} \to \mathbb{R}, \quad x \mapsto \sqrt{\frac{1}{m_i}\sum_{k=1}^{m_i}(x_{i,k}-\mu_i)^2}, \qquad \text{for } i \in \{1,\dots,n\},
$$

then Remark A.3.54, part 2, yields $(m_i\tilde{S}_i^2/\sigma_i^2) \,\square\, \mathrm{N}_{\mu_i,\sigma_i^2}^{\otimes m_i} = \chi^2_{m_i}$ for $i \in \{1,\dots,n\}$. The calculations in Examples 2.4.3 and 2.4.6 and Lemmas 2.4.5 and 2.4.8 with m_i and $\tilde{S}_i$ instead of $m_i - 1$ and S_i, respectively, yield the confidence regions for the parameters of interest $\bigwedge$ and $\bigvee$ from Lemma 2.4.5 based on the Buehlerizations of $\bigwedge_i \tilde{S}_i^2$ and $\bigvee_i \tilde{S}_i^2$ given by

$$
\begin{aligned}
\bigwedge[R_{\bigwedge_i \tilde{S}_i^2,\beta}(x)] &= \left]0, \overline{\bigwedge\tilde{\sigma}}_\beta(\textstyle\bigwedge_{i=1}^n \tilde{S}_i^2(x))\right[\\
\bigwedge[\tilde{R}_{\bigwedge_i \tilde{S}_i^2,\beta}(x)] &= \left]\bigwedge_{i=1}^n \underline{\tilde{\sigma}_i}_{1-\beta}(\textstyle\bigwedge_{k=1}^n \tilde{S}_k^2(x)), \infty\right[\\
\bigvee[R_{\bigwedge_i \tilde{S}_i^2,\beta}(x)] &= \begin{cases}]0,\infty[& \text{if } n \geq 2 \text{ and no } x_i \text{ is constantly } \mu_i \\ \emptyset & \text{if } n \geq 2 \text{ and some } x_i \text{ is constantly } \mu_i \\ \left]0, \underline{\tilde{\sigma}_1}_\beta(\tilde{S}_1^2(x))\right[& \text{if } n = 1 \end{cases}
\end{aligned}
$$

$$\bigvee[\tilde{R}_{\bigwedge_i \tilde{S}_i^2,\beta}(x)] = \Big]\overline{\bigwedge\tilde{\sigma}}_{1-\beta}(\textstyle\bigwedge_{k=1}^n \tilde{S}_k^2(x)), \infty\Big[$$
$$\bigwedge[R_{\bigvee_i \tilde{S}_i^2,\beta}(x)] = \Big]0, \underline{\bigvee\tilde{\sigma}}_{1-\beta}(\textstyle\bigvee_{i=1}^n \tilde{S}_i^2(x))\Big[$$
$$\bigwedge[\tilde{R}_{\bigvee_i \tilde{S}_i^2,\beta}(x)] = \begin{cases}]0,\infty[& \text{if } n \geq 2 \\ \Big]\underline{\tilde{\sigma}_1}_{1-\beta}(\tilde{S}_1^2(x)), \infty\Big[& \text{if } n = 1 \end{cases}$$
$$\bigvee[R_{\bigvee_i \tilde{S}_i^2,\beta}(x)] = \Big]0, \bigvee_{i=1}^n \underline{\tilde{\sigma}_i}_\beta(\textstyle\bigvee_{i=1}^n \tilde{S}_i^2(x))\Big[$$
$$\bigvee[\tilde{R}_{\bigvee_i \tilde{S}_i^2,\beta}(x)] = \Big]\underline{\bigvee\tilde{\sigma}}_\beta(\textstyle\bigvee_{k=1}^n \tilde{S}_k^2(x)), \infty\Big[.$$

for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$, where, for $\alpha \in]0,1[$, $t \in]0,\infty[$, and $i \in \{1,\dots,n\}$,

$$\overline{\bigwedge\tilde{\sigma}}_\alpha(t) \text{ is the unique } s \in]0,\infty[\text{ with } \prod_{i=1}^n \chi^2_{m_i}\left(\Big]\frac{m_i}{s^2}t, \infty\Big[\right) = \alpha, \qquad \overline{\bigwedge\tilde{\sigma}}_\alpha(0) := 0,$$
$$\underline{\bigvee\tilde{\sigma}}_\alpha(t) \text{ is the unique } s \in]0,\infty[\text{ with } \prod_{i=1}^n \chi^2_{m_i}\left(\Big]0, \frac{m_i}{s^2}t\Big[\right) = \alpha, \qquad \underline{\bigvee\tilde{\sigma}}_\alpha(0) := 0,$$
$$\underline{\tilde{\sigma}_i}_\alpha(t) \text{ is the unique } s \in]0,\infty[\text{ with } \chi^2_{m_i}\left(\Big]\frac{m_i}{s^2}t, \infty\Big[\right) = \alpha, \qquad \underline{\tilde{\sigma}_i}_\alpha(0) := 0.$$

2.5 SEVERAL HOMOSCEDASTIC SAMPLES WITH KNOWN MEANS

2.5.1 Definition Let $n \in \mathbb{N}$, $m \in \mathbb{N}^n$, and $\mu \in \mathbb{R}^n$, and let

$$\mathcal{P}_5 := \left(\bigotimes_{i=1}^n \mathrm{N}_{\mu_i,\sigma^2}^{\otimes m_i} : \sigma \in]0,\infty[\right)$$

be the n-sample homoscedastic normal model with known means $\mu_1, \dots, \mu_n$ and known sample sizes $m_1, \dots, m_n$.

2.5.2 Remark 1. The sample space is $\mathcal{X} = \prod_{i=1}^n \mathbb{R}^{m_i}$, the parameter space $\Theta =]0,\infty[$.

2. $\mathcal{P}_2$ is not stochastically monotonic.
3. The following is a special case of $\mathcal{P}_5$:

$$\mathcal{P}_5' := (\mathrm{N}_{\mu,\sigma^2}^{\otimes n} : \sigma \in]0,\infty[) \qquad \text{for } \mu \in \mathbb{R},$$

the one-sample normal model with known mean μ and known sample size n. The sample space in this case is $\mathcal{X} = \mathbb{R}^n$.

2.5.3 Example If we set S_i as in Example 2.1.8, then the Buehlerizations of $\bigwedge_i S_i^2$ and $\bigvee_i S_i^2$ are given, using the confidence bounds $\overline{\bigwedge\sigma}_\beta$ and $\underline{\bigvee\sigma}_\beta$ from Remarks 2.4.4, part 3, and 2.4.7, part 2, respectively, by

$$R_{\bigwedge_i S_i^2,\beta}(x) = \Big]0, \overline{\bigwedge\sigma}_\beta(\textstyle\bigwedge_{i=1}^n S_i^2(x))\Big[\qquad \tilde{R}_{\bigwedge_i S_i^2,\beta}(x) = \Big]\overline{\bigwedge\sigma}_{1-\beta}(\textstyle\bigwedge_{i=1}^n S_i^2(x)), \infty\Big[$$
$$R_{\bigvee_i S_i^2,\beta}(x) = \Big]0, \underline{\bigvee\sigma}_\beta(\textstyle\bigvee_{i=1}^n S_i^2(x))\Big[\qquad \tilde{R}_{\bigvee_i S_i^2,\beta}(x) = \Big]\underline{\bigvee\sigma}_{1-\beta}(\textstyle\bigvee_{i=1}^n S_i^2(x)), \infty\Big[$$

for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$ due to Examples 2.4.3 and 2.4.6. The effective levels of these confidence regions are given by

$$\beta_{\text{eff}}(R_{\bigwedge_i S_i^2,\beta}) = \beta_{\text{eff}}(\tilde{R}_{\bigwedge_i S_i^2,\beta}) = \beta_{\text{eff}}(R_{\bigvee_i S_i^2,\beta}) = \beta_{\text{eff}}(\tilde{R}_{\bigvee_i S_i^2,\beta}) = \beta \qquad \text{for } \beta \in [0,1].$$

In the model $\mathcal{P}_5'$ from the previous remark, the Buehlerization of

$$S\colon \mathbb{R}^n \to [0,\infty[, \quad x \mapsto \sqrt{\sum_{i=1}^n (x_i - \overline{x})^2/(n-1)},$$

yield for the identity $\mathrm{id}_{]0,\infty[}$ as parameter of interest the confidence regions given for $x \in \mathbb{R}^n$ by

$$R_{S^2,\beta,\mathcal{P}_5'}(x) = \left]0, \sqrt{\frac{n-1}{F^{-1}(1-\beta)}}S(x)\right[\qquad \tilde{R}_{S^2,\beta,\mathcal{P}_5'}(x) = \left]\sqrt{\frac{n-1}{F^{-1}(\beta)}}S(x), \infty\right[,$$

where F denotes the distribution function of the law χ^2_{n-1}.

2.5.4 Example If we set $\tilde{S}_i$ as in Example 2.4.9, then the Buehlerizations of $\bigwedge_i \tilde{S}_i^2$ and $\bigvee_i \tilde{S}_i^2$ are given, using the confidence bounds $\overline{\bigwedge \tilde{\sigma}}_\beta$ and $\underline{\bigvee \tilde{\sigma}}_\beta$ from Example 2.4.9 by

$$R_{\bigwedge_i \tilde{S}_i^2,\beta}(x) = \left]0, \overline{\bigwedge \tilde{\sigma}}_\beta(\wedge_{i=1}^n \tilde{S}_i^2(x))\right[\qquad \tilde{R}_{\bigwedge_i \tilde{S}_i^2,\beta}(x) = \left]\overline{\bigwedge \tilde{\sigma}}_{1-\beta}(\wedge_{i=1}^n \tilde{S}_i^2(x)), \infty\right[$$

$$R_{\bigvee_i \tilde{S}_i^2,\beta}(x) = \left]0, \underline{\bigvee \tilde{\sigma}}_\beta(\vee_{i=1}^n \tilde{S}_i^2(x))\right[\qquad \tilde{R}_{\bigvee_i \tilde{S}_i^2,\beta}(x) = \left]\underline{\bigvee \tilde{\sigma}}_{1-\beta}(\vee_{i=1}^n \tilde{S}_i^2(x)), \infty\right[$$

for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$ due to Example 2.4.9. The effective levels of these confidence regions are given by

$$\beta_{\text{eff}}(R_{\bigwedge_i \tilde{S}_i^2,\beta}) = \beta_{\text{eff}}(\tilde{R}_{\bigwedge_i \tilde{S}_i^2,\beta}) = \beta_{\text{eff}}(R_{\bigvee_i \tilde{S}_i^2,\beta}) = \beta_{\text{eff}}(\tilde{R}_{\bigvee_i \tilde{S}_i^2,\beta}) = \beta \qquad \text{for } \beta \in [0,1].$$

In the model $\mathcal{P}_5'$ from Remark 2.5.2, part 3, the Buehlerization of

$$\tilde{S}\colon \mathbb{R}^n \to [0,\infty[, \quad x \mapsto \sqrt{\sum_{i=1}^n (x_i - \mu)^2/n},$$

yield for the identity $\mathrm{id}_{]0,\infty[}$ as parameter of interest the confidence regions given for $x \in \mathbb{R}^n$ by

$$R_{\tilde{S}^2,\beta,\mathcal{P}_5'}(x) = \left]0, \sqrt{\frac{n}{F^{-1}(1-\beta)}}\tilde{S}(x)\right[\qquad \tilde{R}_{\tilde{S}^2,\beta,\mathcal{P}_5'}(x) = \left]\sqrt{\frac{n}{F^{-1}(\beta)}}\tilde{S}(x), \infty\right[,$$

where F denotes the distribution function of the law χ^2_n.

3 APPLICATION: BINOMIAL SAMPLES

- In this chapter, the set $\mathcal{Y}$ is a subset of the integers, the order is the usual. Its interval topology and the induced Borel σ-algebra are both the power set of $\mathcal{Y}$.
- Let us remember that $\beta \in]0,1[$ due to Remark 1.1.7, part 3, unless stated otherwise.
- Outline of this chapter:
  - Section 3.1 deals with several binomial samples with known sample sizes and unknown success probabilities. This is the most general model considered in this chapter. We buehlerize minimum and maximum of the scaled samples. The Buehlerization of the sample mean of the scaled samples lies unfortunately out of reach.
  - Section 3.2 investigates single binomial samples on the basis of the results from the previous section. The identity is buehlerized and Sterne's (1954) confidence region is derived.
  - Section 3.3 deals with the simple but nevertheless instructive Bernoulli model. We buehlerize again the identity, but this time on the basis of a general result that yields every confidence region with level β for the identity on the parameter space.
  - Section 3.4 investigates a translated symmetric version of the model from the previous section. The latter two models are particular in the sense that they allow for the existence of minimal resp. least confidence down- and uprays.

3.1 SEVERAL SAMPLES

3.1.1 Definition Let $n \in \mathbb{N}$, $m \in \mathbb{N}^n$, and let

$$\mathcal{P} := \left(\bigotimes_{i=1}^{n} \mathrm{B}_{m_i,p_i} : p \in [0,1]^n \right)$$

be the n-sample binomial model with known sample sizes $m_1, \ldots, m_n$.

3.1.2 Remark The sample space is $\mathcal{X} = \prod_{i=1}^{n}\{0,\ldots,m_i\}$, the parameter space $\Theta = [0,1]^n$.

3.1.3 Example Let us consider the designated statistic

$$\bigwedge X/m\colon \prod_{i=1}^{n}\{0,\ldots,m_i\} \to [0,1], \quad x \mapsto \bigwedge_{i=1}^{n} \frac{x_i}{m_i}.$$

Its Buehlerization is then given for $x \in \prod_{i=1}^{n}\{0,\ldots,m_i\}$ by

$$R_{\bigwedge X/m,\beta}(x) = \left\{ p \in [0,1]^n\colon \prod_{i=1}^{n} \mathrm{B}_{m_i,p_i}\left(\left\{ \left\lfloor m_i \bigwedge_{k=1}^{n} \frac{x_k}{m_k} \right\rfloor + 1, \ldots, m_i \right\} \right) < \beta \right\}$$

$$\tilde{R}_{\bigwedge X/m,\beta}(x) = \left\{ p \in [0,1]^n\colon \prod_{i=1}^{n} \mathrm{B}_{m_i,p_i}\left(\left\{ \left\lceil m_i \bigwedge_{k=1}^{n} \frac{x_k}{m_k} \right\rceil, \ldots, m_i \right\} \right) > 1-\beta \right\}.$$

Since, for $i \in \{1, \ldots, n\}$,

$$m_i \bigwedge_{k=1}^{n} \frac{x_k}{m_k} = \begin{cases} 0 & \text{if } x_k = 0 \text{ for some } k \in \{1, \ldots, n\} \\ m_i & \text{if } x = m, \end{cases}$$

we have

$$[0,1]^n = \begin{cases} R_{\bigwedge X/m,\beta}(x) & \text{if } x = m \\ \tilde{R}_{\bigwedge X/m,\beta}(x) & \text{if } x_k = 0 \text{ for some } k \in \{1, \ldots, n\}. \end{cases}$$

Due to the above and the surjectivity of

$$[0,1] \to [0,1], \quad p \mapsto \prod_{i=1}^{n} \mathrm{B}_{m_i,p}(\{1, \ldots, m_i\}) = \prod_{i=1}^{n} (1 - (1-p)^{m_i}),$$

$$[0,1] \to [0,1], \quad p \mapsto \prod_{i=1}^{n} \mathrm{B}_{m_i,p}(\{m_i\}) = p^{\sum_{i=1}^{n} m_i},$$

the effective levels of $R_{\bigwedge X/m,\beta}$ and $\tilde{R}_{\bigwedge X/m,\beta}$ are

$$\beta_{\text{eff}}(R_{\bigwedge X/m,\beta}) = \beta_{\text{eff}}(\tilde{R}_{\bigwedge X/m,\beta}) = \beta \qquad \text{for } \beta \in [0,1].$$

3.1.4 Remark 1. We shall use the fact that, for $r \in \mathbb{N}$ and $k \in \mathbb{N}^r$,

$$[0,1]^n \to [0,1], \quad p \mapsto \prod_{i=1}^{r} \mathrm{B}_{k_i,p_i}(\{\lfloor k_i t \rfloor + 1, \ldots, k_i\}),$$

is strictly increasing for $t \in [0,1[$, and constantly 0 for $t = 1$, and

$$[0,1]^n \to [0,1], \quad p \mapsto \prod_{i=1}^{r} \mathrm{B}_{k_i,p_i}(\{\lceil k_i t \rceil, \ldots, k_i\}),$$

is strictly increasing for $t \in]0,1]$, and constantly 1 for $t = 0$.

2. For $r \in \mathbb{N}$, $k \in \mathbb{N}^r$, $\alpha \in [0,1]$, and $t \in [0,1[$ there is exactly one

$$\overline{\bigwedge p(k)}_\alpha(t) \in [0,1] \quad \text{such that} \quad \prod_{i=1}^{r} \mathrm{B}_{k_i, \overline{\bigwedge p(k)}_\alpha(t)}(\{\lfloor k_i t \rfloor + 1, \ldots, k_i\}) = \alpha.$$

In fact, $f\colon [0,1] \to [0,1]$, $p \mapsto \prod_{i=1}^{r} \mathrm{B}_{k_i,p}(\{\lfloor k_i t \rfloor + 1, \ldots, k_i\})$, is bijective since it is strictly increasing and continuous with $f(0) = 0$ and $f(1) = 1$.

3. For $r \in \mathbb{N}$, $k \in \mathbb{N}^r$, $\alpha \in [0,1]$, and $t \in]0,1]$ there is exactly one

$$\underline{\bigvee p(k)}_\alpha(t) \in [0,1] \quad \text{such that} \quad \prod_{i=1}^{r} \mathrm{B}_{k_i, \underline{\bigvee p(k)}_\alpha(t)}(\{0, \ldots, \lceil k_i t \rceil - 1\}) = \alpha,$$

namely $\underline{\bigvee p(k)}_\alpha(t) := 1 - \overline{\bigwedge p(k)}_\alpha(1-t)$.

4. For $r \in \mathbb{N}$, $k \in \mathbb{N}^r$, $\alpha \in [0,1]$, and $t \in]0,1]$ there is exactly one

$$\underline{\bigvee \tilde{p}(k)}_\alpha(t) \in [0,1] \quad \text{such that} \quad \prod_{i=1}^{r} \mathrm{B}_{k_i, \underline{\bigvee \tilde{p}(k)}_\alpha(t)}(\{\lceil k_i t \rceil, \ldots, k_i\}) = 1 - \alpha.$$

In fact, $f\colon [0,1] \to [0,1]$, $p \mapsto \prod_{i=1}^{r} \mathrm{B}_{k_i,p}(\{\lceil k_i t \rceil, \ldots, k_i\})$, is bijective since it is strictly increasing and continuous with $f(0) = 0$ and $f(1) = 1$.

5. For $r \in \mathbb{N}$, $k \in \mathbb{N}^r$, $\alpha \in [0,1]$, and $t \in [0,1[$ there is exactly one
$$\overline{\bigwedge \tilde{p}(k)}_\alpha(t) \in [0,1] \quad \text{such that} \quad \prod_{i=1}^{r} \mathrm{B}_{k_i, \overline{\bigwedge \tilde{p}(k)}_\alpha(t)}(\{0,\dots,\lfloor k_i t \rfloor\}) = 1-\alpha,$$
namely $\overline{\bigwedge \tilde{p}(k)}_\alpha(t) := 1 - \underline{\bigvee \tilde{p}(k)}_\alpha(1-t)$.
6. For $k \in \mathbb{N}$, $\alpha \in [0,1]$, and $t \in]0,1]$ there is exactly one
$$\underline{p(k)}_\alpha(t) \in [0,1] \quad \text{such that} \quad \mathrm{B}_{k,\underline{p(k)}_\alpha(t)}(\{\lceil kt \rceil,\dots,k\}) = 1-\alpha.$$
In fact, $f\colon [0,1] \to [0,1]$, $p \mapsto \mathrm{B}_{k,p}(\{\lceil kt \rceil,\dots,k\})$, is bijective since it is strictly increasing and continuous with $f(0)=0$ and $f(1)=1$.
7. For $k \in \mathbb{N}$, $\alpha \in [0,1]$, and $t \in [0,1[$ there is exactly one
$$\overline{p(k)}_\alpha(t) \in [0,1] \quad \text{such that} \quad \mathrm{B}_{k,\overline{p(k)}_\alpha(t)}(\{0,\dots,\lfloor kt \rfloor\}) = 1-\alpha,$$
namely $\overline{p(k)}_\alpha(t) := 1 - \underline{p(k)}_\alpha(1-t)$.
8. The functions $\overline{\bigwedge p(k)}_\alpha$, $\underline{\bigvee p(k)}_\alpha$, $\overline{\bigwedge \tilde{p}(k)}_\alpha$, $\underline{\bigvee \tilde{p}(k)}_\alpha$, $\underline{p(k)}_\alpha$, and $\overline{p(k)}_\alpha$ are increasing with
$$\begin{aligned}
\overline{\bigwedge p(k)}_\alpha(0) &= p^*(\alpha) & \lim_{t\to 1} \overline{\bigwedge p(k)}_\alpha(t) &= 1\\
\lim_{t\to 0} \underline{\bigvee p(k)}_\alpha(t) &= 0 & \underline{\bigvee p(k)}_\alpha(1) &= p_*(\alpha)\\
\lim_{t\to 0} \underline{\bigvee \tilde{p}(k)}_\alpha(t) &= p_*(1-\alpha) & \underline{\bigvee \tilde{p}(k)}_\alpha(1) &= (1-\alpha)^{1/|k|}\\
\overline{\bigwedge \tilde{p}(k)}_\alpha(0) &= 1-(1-\alpha)^{1/|k|} & \lim_{t\to 1} \overline{\bigwedge \tilde{p}(k)}_\alpha(t) &= p^*(1-\alpha)\\
\lim_{t\to 0} \underline{p(k)}_\alpha(t) &= 0 & \underline{p(k)}_\alpha(1) &= (1-\alpha)^{1/k}\\
\overline{p(k)}_\alpha(0) &= 1-(1-\alpha)^{1/k} & \lim_{t\to 1} \overline{p(k)}_\alpha(t) &= 1,
\end{aligned}$$
where, for $\gamma \in [0,1]$,
$$\left\{\begin{matrix} p^*(\gamma) \\ p_*(\gamma) \end{matrix}\right\} \text{ denotes the unique } p \in [0,1] \text{ satisfying } \begin{cases} \prod_{i=1}^{r}(1-(1-p)^{k_i}) = \gamma \\ \prod_{i=1}^{r}(1-p^{k_i}) = \gamma. \end{cases}$$
9. If $k \in \mathbb{N}$, and $l_{\mathrm{CP},\beta}$ and $u_{\mathrm{CP},\beta}$ denote the lower and upper confidence bound, respectively, with level β of Clopper and Pearson (1934) for $(\mathrm{B}_{k,p}\colon p \in [0,1])$, and the above functions $\underline{p(k)}_\beta$ and $\overline{p(k)}_\beta$ are extended to $[0,1]$ by continuity, then
$$\underline{p(k)}_\beta\left(\frac{x}{k}\right) = l_{\mathrm{CP},\beta}(x) \quad \text{and} \quad \overline{p(k)}_\beta\left(\frac{x}{k}\right) = u_{\mathrm{CP},\beta}(x) \qquad \text{for } x \in \{0,\dots,k\}.$$

3.1.5 Lemma *Let us consider the parameters of interest*
$$\bigwedge\colon [0,1]^n \to [0,1], \quad p \mapsto \bigwedge_{i=1}^{n} p_i, \qquad \bigvee\colon [0,1]^n \to [0,1], \quad p \mapsto \bigvee_{i=1}^{n} p_i.$$
For $x \in \prod_{i=1}^{n}\{0,\dots,m_i\}$ *then*
$$\bigwedge[R_{\bigwedge X/m,\beta}(x)] = \begin{cases} \left[0, \overline{\bigwedge p(m)}_\beta(\bigwedge_{i=1}^{n} \frac{x_i}{m_i})\right[& \textit{if } x \neq m \\ [0,1] & \textit{if } x = m \end{cases} \tag{25}$$

$$\bigwedge[\tilde{R}_{\bigwedge X/m,\beta}(x)] = \begin{cases} \left]\bigwedge_{i=1}^n \underline{p(m_i)}_\beta(\bigwedge_{k=1}^n \frac{x_k}{m_k}), 1\right] & \textit{if } x_i > 0 \textit{ for } i \in \{1,\dots,n\} \\ [0,1] & \textit{if } x_i = 0 \textit{ for some } i \in \{1,\dots,n\} \end{cases} \tag{26}$$

$$\bigvee[R_{\bigwedge X/m,\beta}(x)] = \begin{cases} \left[0, \overline{p(m)}_\beta(\frac{x}{m})\right[& \textit{if } n = 1 \textit{ and } x < m \\ [0,1] & \textit{if } n = 1 \textit{ and } x = m, \textit{ or } n \geq 2 \end{cases} \tag{27}$$

$$\bigvee[\tilde{R}_{\bigwedge X/m,\beta}(x)] = \begin{cases} \left]\underline{\bigvee \tilde{p}(m)}_\beta(\bigwedge_{i=1}^n \frac{x_i}{m_i}), 1\right] & \textit{if } x_i > 0 \textit{ for } i \in \{1,\dots,n\} \\ [0,1] & \textit{if } x_i = 0 \textit{ for some } i \in \{1,\dots,n\}. \end{cases} \tag{28}$$

Proof. Let $x \in \prod_{i=1}^n \{0,\dots,m_i\}$. Example 3.1.3 then yields $R_{\bigwedge X/m,\beta}(m) = [0,1]^n$ and $\tilde{R}_{\bigwedge X/m,\beta}(x) = [0,1]^n$ if some $x_i = 0$.

(25) It remains to consider the case $x \neq m$. Let us first show LHS $\subseteq$ RHS. To this end, let $p \in R_{\bigwedge X/m,\beta}(x)$, and let us assume $\bigwedge_{i=1}^n p_i \geq \sup \mathrm{RHS}$. Remark 3.1.4, part 1, then implies

$$\prod_{i=1}^n \mathrm{B}_{m_i,p_i}\left(\left\{\left\lfloor m_i \bigwedge_{k=1}^n \frac{x_k}{m_k}\right\rfloor + 1,\dots,m_i\right\}\right) \geq \prod_{i=1}^n \mathrm{B}_{m_i,\sup \mathrm{RHS}}\left(\left\{\left\lfloor m_i \bigwedge_{k=1}^n \frac{x_k}{m_k}\right\rfloor + 1,\dots,m_i\right\}\right) = \beta,$$

a contradiction to $p \in R_{\bigwedge X/m,\beta}(x)$ in view of Example 3.1.3.

Let us now show LHS $\supseteq$ RHS. Let $t \in \mathrm{RHS}(x)$, and let us define $p_i := t$ for $i \in \{1,\dots,n\}$. Remark 3.1.4, parts 1 and 2, then implies

$$\prod_{i=1}^n \mathrm{B}_{m_i,p_i}\left(\left\{\left\lfloor m_i \bigwedge_{k=1}^n \frac{x_k}{m_k}\right\rfloor + 1,\dots,m_i\right\}\right) < \prod_{i=1}^n \mathrm{B}_{m_i,\sup \mathrm{RHS}}\left(\left\{\left\lfloor m_i \bigwedge_{k=1}^n \frac{x_k}{m_k}\right\rfloor + 1,\dots,m_i\right\}\right) = \beta,$$

hence $p \in R_{\bigwedge X/m,\beta}(x)$ due to Example 3.1.3. Together with $\bigwedge_{i=1}^n p_i = t$, this yields $t \in \mathrm{LHS}(x)$.

(26) It remains to consider the case where $x_i > 0$ for some $i \in \{1,\dots,n\}$. Let us pick $i_0 \in \{1,\dots,n\}$ such that $\underline{p(m_{i_0})}_\beta(\bigwedge_{k=1}^n \frac{x_k}{m_k}) = \bigwedge_{i=1}^n \underline{p(m_i)}_\beta(\bigwedge_{k=1}^n \frac{x_k}{m_k})$.

We first show the inclusion LHS $\subseteq$ RHS. To this end, let $p \in \tilde{R}_{\bigwedge X/m,\beta}(x)$, and let us assume $\bigwedge_{k=1}^n p_k \leq \inf \mathrm{RHS}$. Remark 3.1.4, parts 1 and 6, then implies

$$\prod_{i=1}^n \mathrm{B}_{m_i,p_i}\left(\left\{\left\lceil m_i \bigwedge_{k=1}^n \frac{x_k}{m_k}\right\rceil,\dots,m_i\right\}\right) \leq \mathrm{B}_{m_{i_0},\inf \mathrm{RHS}}\left(\left\{\left\lceil m_{i_0} \bigwedge_{k=1}^n \frac{x_k}{m_k}\right\rceil,\dots,m_{i_0}\right\}\right) = 1-\beta,$$

contradicting $p \in \tilde{R}_{\bigwedge X/m,\beta}(x)$ in view of Example 3.1.3.

We now show LHS $\supseteq$ RHS. To this end, let $t \in$ RHS, and let us define

$$\begin{aligned} p_{i_0} &:= t \\ \varepsilon &:= \mathrm{B}_{m_{i_0},p_{i_0}}\left(\left\{\left\lceil m_{i_0}\bigwedge_{k=1}^{n}\frac{x_k}{m_k}\right\rceil,\dots,m_{i_0}\right\}\right) \qquad (\in\,]1-\beta,1]) \\ p_i &:= t \vee \left(\frac{1}{2}\left(1+\overline{\bigwedge p(m|_{\{1,\dots,n\}\setminus\{i_0\}})}_{\frac{1-\beta}{\varepsilon}}\left(\bigwedge_{k=1}^{n}\frac{x_k}{m_k}\right)\right)\right) \quad \text{for } i\in\{1,\dots,n\}\setminus\{i_0\}. \end{aligned}$$

Remark 3.1.4, parts 1 and 2, implies

$$\begin{aligned} &\prod_{i=1}^{n}\mathrm{B}_{m_i,p_i}\left(\left\{\left\lceil m_i\bigwedge_{k=1}^{n}\frac{x_k}{m_k}\right\rceil,\dots,m_i\right\}\right) \\ &> \varepsilon\prod_{\substack{i=1\\ i\neq i_0}}^{n}\mathrm{B}_{m_i,\overline{\bigwedge p(m|_{\{1,\dots,n\}\setminus\{i_0\}})}_{\frac{1-\beta}{\varepsilon}}(\bigwedge_{k=1}^{n}x_k/m_k)}\left(\left\{\left\lceil m_i\bigwedge_{k=1}^{n}\frac{x_k}{m_k}\right\rceil,\dots,m_i\right\}\right) = 1-\beta, \end{aligned}$$

hence $p \in \tilde{R}_{\bigwedge X/m,\beta}(x)$ in view of Example 3.1.3. Together with $\bigwedge_{k=1}^{n} p_k = t$, this yields $t \in$ LHS.

(27) The case $n = 1$ is a special case of (25) since $\overline{\bigwedge p(m)}_\beta = \overline{p(m)}_\beta$ in that case. Let us therefore assume $n \geq 2$. The inclusion LHS $\subseteq$ RHS is clear. It thus remains to show LHS $\supseteq$ RHS. To this end, let $t \in [0,1]$, and let us define $p_1 := t$ and $p_i := 0$ for $i \in \{2,\dots,n\}$. Remark 3.1.4, part 1, then implies

$$\begin{aligned} &\prod_{i=1}^{n}\mathrm{B}_{m_i,p_i}\left(\left\{\left\lfloor m_i\bigwedge_{k=1}^{n}\frac{x_k}{m_k}\right\rfloor+1,\dots,m_i\right\}\right) \\ &\qquad\qquad \leq \mathrm{B}_{m_2,0}\left(\left\{\left\lfloor m_2\bigwedge_{k=1}^{n}\frac{x_k}{m_k}\right\rfloor+1,\dots,m_2\right\}\right) = 0, \end{aligned}$$

hence $p \in R_{\bigwedge X/m,\beta}$ due to Example 3.1.3. Together with $\bigvee_{k=1}^{n} p_k = t$, this yields $t \in$ LHS.

(28) It remains to consider the case where $x_i > 0$ for every $i \in \{1,\dots,n\}$. We first show LHS $\subseteq$ RHS. To this end, let $p \in \tilde{R}_{\bigwedge X/m,\beta}(x)$, and let us assume $\bigvee_{k=1}^{n} p_k \leq \inf \mathrm{RHS}$. Remark 3.1.4, parts 1 and 4, then implies

$$\begin{aligned} &\prod_{i=1}^{n}\mathrm{B}_{m_i,p_i}\left(\left\{\left\lceil m_i\bigwedge_{k=1}^{n}\frac{x_k}{m_k}\right\rceil,\dots,m_i\right\}\right) \\ &\qquad\qquad \leq \prod_{i=1}^{n}\mathrm{B}_{m_i,\inf \mathrm{RHS}}\left(\left\{\left\lceil m_i\bigwedge_{k=1}^{n}\frac{x_k}{m_k}\right\rceil,\dots,m_i\right\}\right) = 1-\beta, \end{aligned}$$

contradicting $p \in \tilde{R}_{\bigwedge X/m,\beta}(x)$ in view of Example 3.1.3.

We now show LHS $\supseteq$ RHS. Let $t \in$ RHS, and let us define $p_i := t$ for $i \in \{1,\dots,n\}$. Remark 3.1.4, parts 1 and 4, then implies

$$\prod_{i=1}^{n}\mathrm{B}_{m_i,p_i}\left(\left\{\left\lceil m_i\bigwedge_{k=1}^{n}\frac{x_k}{m_k}\right\rceil,\dots,m_i\right\}\right)$$

$$> \prod_{i=1}^{n} \mathrm{B}_{m_i,\inf \mathrm{RHS}}\left(\left\{\left\lceil m_i \bigwedge \frac{x}{m}\right\rceil, \ldots, m_i\right\}\right) = 1-\beta,$$

hence $p \in \tilde{R}_{\bigwedge X/m,\beta}(x)$ due to Example 3.1.3. Together with $\bigvee_{k=1}^{n} p_k = t$, this yields $t \in \mathrm{LHS}$. □

3.1.6 Example Let us consider the designated statistic

$$\bigvee X/m\colon \prod_{i=1}^{n}\{0,\ldots,m_i\} \to [0,1], \quad x \mapsto \bigvee_{i=1}^{n} \frac{x_i}{m_i}.$$

Since $(m - \mathrm{id}_{\mathcal{X}}) \mathbin{\square} \bigotimes_{i=1}^{n} \mathrm{B}_{m_i,p_i} = \bigotimes_{i=1}^{n} \mathrm{B}_{m_i,1-p_i}$ and $\bigvee X/m = 1 - (\bigwedge X/m) \circ (m - \mathrm{id}_{\mathcal{X}})$, Example 3.1.3 yields the confidence regions given for $x \in \prod_{i=1}^{n}\{0,\ldots,m_i\}$ by

$$\begin{aligned} R_{\bigvee X/m,\beta}(x) &= 1 - \tilde{R}_{\bigwedge X/m,\beta}(m-x) \\ &= \left\{p \in [0,1]^n\colon \prod_{i=1}^{n} \mathrm{B}_{m_i,p_i}\left(\left\{0,\ldots,\left\lfloor m_i \bigvee_{k=1}^{n} \frac{x_k}{m_k}\right\rfloor\right\}\right) > 1-\beta\right\} \\ \tilde{R}_{\bigvee X/m,\beta}(x) &= 1 - R_{\bigwedge X/m,\beta}(m-x) \\ &= \left\{p \in [0,1]^n\colon \prod_{i=1}^{n} \mathrm{B}_{m_i,p_i}\left(\left\{0,\ldots,\left\lceil m_i \bigvee_{k=1}^{n} \frac{x_k}{m_k}\right\rceil - 1\right\}\right) < \beta\right\} \end{aligned}$$

and, using Lemma 3.1.5,

$$\bigvee[R_{\bigvee X/m,\beta}(x)] = \begin{cases} \left[0, \bigvee_{i=1}^{n} \overline{p(m_i)}_\beta(\bigvee_{k=1}^{n} \frac{x_k}{m_k})\right[& \text{if } x_i < m_i \text{ for every } i \in \{1,\ldots,n\} \\ [0,1] & \text{if } x_i = m_i \text{ for some } i \in \{1,\ldots,n\} \end{cases}$$

$$\bigvee[\tilde{R}_{\bigvee X/m,\beta}(x)] = \begin{cases} \left]\underline{\bigvee p(m)}_\beta(\bigvee_{k=1}^{n} \frac{x_k}{m_k}), 1\right] & \text{if } x_i > 0 \text{ for some } i \in \{1,\ldots,n\} \\ [0,1] & \text{if } x_i = 0 \text{ for every } i \in \{1,\ldots,n\} \end{cases}$$

$$\bigwedge[R_{\bigvee X/m,\beta}(x)] = \begin{cases} \left[0, \overline{\bigwedge \tilde{p}(m)}_\beta(\bigvee_{k=1}^{n} \frac{x_k}{m_k})\right[& \text{if } x_i < m_i \text{ for every } i \in \{1,\ldots,n\} \\ [0,1] & \text{if } x_i = m_i \text{ for some } i \in \{1,\ldots,n\} \end{cases}$$

$$\bigwedge[\tilde{R}_{\bigvee X/m,\beta}(x)] = \begin{cases} \left]\underline{p(m)}_\beta(\frac{x}{m}), 1\right] & \text{if } n = 1 \text{ and } x > 0 \\ [0,1] & \text{if } n = 1 \text{ and } x = 0, \text{ or } n \geq 2. \end{cases}$$

3.1.7 Remark The Buehlerization of the designated statistic

$$\overline{X/m}\colon \prod_{i=1}^{n}\{0,\ldots,m_i\} \to [0,1], \quad x \mapsto \frac{1}{n}\sum_{i=1}^{n} \frac{x_i}{m_i},$$

involves rather complex calculations and does not seem to be establishable without investing considerably more effort. Mattner and Tasto (2014), however, have shown that for $\beta \geq 3/4$ the confidence bounds $u_{\mathrm{CP},\beta} = \overline{p(n)}_\beta(\cdot/n)$ and $l_{\mathrm{CP},\beta} = \underline{p(n)}_\beta(\cdot/n)$ are, if modified at only $n-1$ resp. 1, valid for the parameter of interest

$$[0,1]^n \to [0,1], \quad p \mapsto \overline{p} := \frac{1}{n}\sum_{k=1}^{n} p_k,$$

in the model $(\ast_{i=1}^{n} \mathrm{B}_{p_i}\colon p \in [0,1]^n)$ of Bernoulli convolutions $\ast_{i=1}^{n} \mathrm{B}_{p_i} := \mathrm{B}_{p_1} \ast \ldots \ast \mathrm{B}_{p_n}$.

3.2 ONE SAMPLE

3.2.1 Definition Let $n \in \mathbb{N}$. Let us consider the one-sample binomial model

$$\mathcal{P} := (\mathrm{B}_{m,p} \colon p \in [0,1])$$

with known sample size m.

3.2.2 Remark The sample space is $\mathcal{X} = \{0, \dots, m\}$, the parameter space $\Theta = [0,1]$.

3.2.3 Example Let us consider the family

$$\mathcal{D} := (\mathrm{b}_{m,p} \colon p \in [0,1])$$

of densities $\mathrm{b}_{m,p} = \sum_{k=1}^{m} \binom{m}{k} p^k (1-p)^{m-k} \mathbf{1}_{\{k\}}$ of $\mathrm{B}_{m,p}$ with respect to counting measure on $\{0, \dots, m\}$. The terms $\mathrm{B}_{m,p}(\mathrm{b}_{m,p} > \mathrm{b}_{m,p}(x))$ and $\mathrm{B}_{m,p}(\mathrm{b}_{m,p} < \mathrm{b}_{m,p}(x))$ cannot be expressed in a more explicit form in general. The confidence region $R_{\mathcal{D},\beta}$ differs from Sterne's (1954) proposed one R^{S}_{β} only by the fact that it does not contain its boundary points, that is, $R_{\mathcal{D},\beta} = R^{\mathrm{S}}_{\beta} \setminus \partial R^{\mathrm{S}}_{\beta}$.

Since

$$[0,1] \times \{0, \dots, n\} \to [0,1], \quad (p, x) \mapsto \mathrm{b}_{m,p}(x),$$

is surjective (which follows from the surjectivity of its restriction to $[0,1] \times \{n\}$, i.e., $[0,1] \to [0,1]$, $p \mapsto \mathrm{b}_{n,p}(n) = p^n$) the same applies to

$$[0,1] \times \{0, \dots, m\} \to [0,1], \quad (p, x) \mapsto \mathrm{B}_{m,p}(\mathrm{b}_{m,p} \sim \mathrm{b}_{m,p}(x)),$$

where $\sim \in \{<, >\}$, due to the unimodality of every $\mathrm{b}_{m,p}$. The effective levels of $R_{\mathcal{D},\beta}$ and $\tilde{R}_{\mathcal{D},\beta}$ are thus

$$\beta_{\mathrm{eff}}(R_{\mathcal{D},\beta}) = \beta_{\mathrm{eff}}(\tilde{R}_{\mathcal{D},\beta}) = \beta \qquad \text{for } \beta \in [0,1].$$

3.2.4 Example If we consider the identity $\mathrm{id}_{\{0,\dots,m\}}$ as designated statistic, then Example 3.1.6 for instance yields the confidence regions given for $x \in \{0, \dots, m\}$ by

$$R_{\mathrm{id},\beta}(x) = \left\{\begin{matrix} [0,1] & \text{if } x = n \\ [0, \overline{p(m)}_\beta(\tfrac{x}{m})[& \text{otherwise} \end{matrix}\right\} \quad \tilde{R}_{\mathrm{id},\beta}(x) = \left\{\begin{matrix} [0,1] & \text{if } x = 0 \\ ]\underline{p(m)}_\beta(\tfrac{x}{m}), 1] & \text{otherwise} \end{matrix}\right\}.$$

Their effective levels are

$$\beta_{\mathrm{eff}}(R_{\mathrm{id},\beta}) = \beta_{\mathrm{eff}}(\tilde{R}_{\mathrm{id},\beta}) = \beta \qquad \text{for } \beta \in [0,1].$$

3.2.5 Remark One can show that $R_{\mathrm{id},\beta}$ and $\tilde{R}_{\mathrm{id},\beta}$ are minimal in the set of all confidence down- resp. uprays with level β for $\mathrm{id}_{[0,1]}$.

3.3 BERNOULLI MODEL

3.3.1 Definition Let $\mathcal{P} := (\mathrm{B}_p \colon p \in [0,1])$ be the Bernoulli model, i.e., the one-sample binomial model with sample size 1.

3.3.2 Remark The sample space is $\mathcal{X} = \{0,1\}$, the parameter space $\Theta = [0,1]$.

3.3.3 Example We would like to find a general expression of an arbitrary confidence bound for the identity $\mathrm{id}_{[0,1]}$ in this model. Let us therefore consider a general family

$$\mathcal{T} = (T_p \colon p \in [0,1])$$

of functions $T_p \colon \{0,1\} \to \mathcal{Y}$ (the set $\mathcal{Y}$ being as required in the statement of Theorem 1.1.6), and let $\beta \in \,]0,1]$. The confidence regions $R_{\mathcal{T},\beta}$ and $\tilde{R}_{\mathcal{T},\beta}$ from Definition 1.1.5 then depend on $\mathcal{T}$ solely via the two sets

$$A := \{p \in [0,1] \colon T_p(0) < T_p(1)\} \quad \text{and} \quad B := \{p \in [0,1] \colon T_p(0) > T_p(1)\}.$$

In fact,

$$\mathrm{B}_p(T_p > T_p(x)) = \begin{cases} p & \text{if } p \in A \text{ and } x = 0, \\ 1-p & \text{if } p \in B \text{ and } x = 1, \\ 0 & \text{otherwise,} \end{cases}$$

and the same result holds for $\mathrm{B}_p(T_p < T_p(x))$ with A and B interchanged. What follows is therefore also valid for $\tilde{R}_{\mathcal{T},\beta}$ after switching A and B. We obtain

$$R_{\mathcal{T},\beta}(0) = (A \cap [0,\beta[) \cup ([0,1] \setminus A) \tag{29}$$
$$R_{\mathcal{T},\beta}(1) = (B \cap \,]1-\beta,1]) \cup ([0,1] \setminus B). \tag{30}$$

The effective level of $R_{\mathcal{T},\beta}$ is given by

$$\beta_{\mathrm{eff}}(R_{\mathcal{T},\beta}) = \inf\big((A \cup (1-B)) \cap [\beta,1]\big) \qquad \text{for } \beta \in [0,1].$$

3.3.4 Remark By varying $A, B \in 2^{[0,1]}$ with $A \cap B = \emptyset$ in (29) and (30), we exhaust the set of all confidence regions with level β for $\mathrm{id}_{[0,1]}$ in this model in view of Theorem 1.1.10.

3.3.5 Example Let us consider the family

$$\mathcal{D} := (\mathrm{b}_p \colon p \in [0,1])$$

of densities $\mathrm{b}_p = p\mathbf{1}_{\{1\}}$ of B_p with respect to counting measure on $\{0,1\}$. We then have $A = \,]\frac{1}{2},1]$ and $B = [0,\frac{1}{2}[$, hence

$$\begin{array}{lll} \text{for } \beta \in \,]\frac{1}{2},1]: & R_{\mathcal{D},\beta}(0) = [0,\beta[& R_{\mathcal{D},\beta}(1) = \,]1-\beta,1] \\ \text{for } \beta \in \,]0,\frac{1}{2}]: & R_{\mathcal{D},\beta}(0) = [0,\frac{1}{2}] & R_{\mathcal{D},\beta}(1) = [\frac{1}{2},1] \end{array}$$

and

$$\begin{array}{lll} \text{for } \beta \in [\frac{1}{2},1]: & \tilde{R}_{\mathcal{D},\beta}(0) = [0,1] & \tilde{R}_{\mathcal{D},\beta}(1) = [0,1] \end{array}$$

for $\beta \in]0, \frac{1}{2}]$: $\tilde{R}_{\mathcal{D},\beta}(0) = [0, \beta[\cup [\frac{1}{2}, 1]$ $\tilde{R}_{\mathcal{D},\beta}(1) = [0, \frac{1}{2}] \cup]1-\beta, 1]$.

The effective levels of $R_{\mathcal{D},\beta}$ and $\tilde{R}_{\mathcal{D},\beta}$ are given for $\beta \in [0,1]$ by

$$\beta_{\text{eff}}(R_{\mathcal{D},\beta}) = \begin{Bmatrix} \beta & \text{if } \beta \in \{0\} \cup [\frac{1}{2}, 1] \\ \frac{1}{2} & \text{if } \beta \in]0, \frac{1}{2}] \end{Bmatrix} \qquad \beta_{\text{eff}}(\tilde{R}_{\mathcal{D},\beta}) = \begin{Bmatrix} 1 & \text{if } \beta \in [\frac{1}{2}, 1] \\ \beta & \text{if } \beta \in [0, \frac{1}{2}[\end{Bmatrix}.$$

3.3.6 Example If we consider the identity $\text{id}_{\{0,1\}}$ as designated statistic, we obtain $A = [0,1]$ and $B = \emptyset$, which yields the confidence regions given by

$$R_{\text{id},\beta}(0) = [0, \beta[\qquad \tilde{R}_{\text{id},\beta}(0) = [0, 1]$$
$$R_{\text{id},\beta}(1) = [0, 1] \qquad \tilde{R}_{\text{id},\beta}(1) =]1-\beta, 1],$$

with effective levels

$$\beta_{\text{eff}}(R_{\text{id},\beta}) = \beta_{\text{eff}}(\tilde{R}_{\text{id},\beta}) = \beta \qquad \text{for } \beta \in [0,1].$$

3.3.7 Remark One can show that $R_{\text{id},\beta}$ and $\tilde{R}_{\text{id},\beta}$ are least in the set of all confidence down- resp. uprays with level β for $\text{id}_{[0,1]}$.

3.4 TRANSLATED SYMMETRIC BERNOULLI MODEL

3.4.1 Definition Let

$$P_\vartheta := \frac{1}{2}(\delta_{\lfloor\vartheta\rfloor} + \delta_{\lfloor\vartheta\rfloor+1}) = \delta_{\lfloor\vartheta\rfloor} * \mathrm{B}_{1/2} \qquad \text{for } \vartheta \in \mathbb{R}$$

and $\mathcal{P} := (P_\vartheta : \vartheta \in \mathbb{R})$ be a translated version of the symmetric Bernoulli model.

3.4.2 Remark The sample space is $\mathcal{X} = \mathbb{Z}$, the parameter space $\Theta = \mathbb{R}$.

3.4.3 Example If we consider the identity $\text{id}_{\mathbb{Z}}$ on the integers $\mathbb{Z}$ as designated statistic, then

$$R_{\text{id},\beta}(x) = \{\vartheta \in \mathbb{R} : P_\vartheta(]x, \infty[\cap \mathbb{Z}) < \beta\}$$
$$\tilde{R}_{\text{id},\beta}(x) = \{\vartheta \in \mathbb{R} : P_\vartheta(]x-1, \infty[\cap \mathbb{Z}) > 1-\beta\}$$

for $x \in \mathbb{Z}$. Since

$$P_\vartheta(]x, \infty[\cap \mathbb{Z}) = \begin{cases} 0 & \text{if } \lfloor\vartheta\rfloor < x \\ \frac{1}{2} & \text{if } \lfloor\vartheta\rfloor = x \\ 1 & \text{if } \lfloor\vartheta\rfloor > x \end{cases}$$

for $x \in \mathbb{Z}$, we have

$$R_{\text{id},\beta}(x) = \begin{Bmatrix}]-\infty, x[& \text{if } \beta \in]0, \frac{1}{2}] \\ ]-\infty, x+1[& \text{if } \beta \in]\frac{1}{2}, 1] \end{Bmatrix}, \quad \tilde{R}_{\text{id},\beta}(x) = \begin{Bmatrix} [x, \infty[& \text{if } \beta \in]0, \frac{1}{2}] \\ [x-1, \infty[& \text{if } \beta \in]\frac{1}{2}, 1] \end{Bmatrix}$$

for $x \in \mathbb{Z}$. The effective levels of $R_{\text{id},\beta}$ and $\tilde{R}_{\text{id},\beta}$ are

$$\beta_{\text{eff}}(R_{\text{id},\beta}) = \beta_{\text{eff}}(\tilde{R}_{\text{id},\beta}) = \begin{cases} \frac{1}{2} & \text{if } \beta \in]0, \frac{1}{2}] \\ 1 & \text{if } \beta \in]\frac{1}{2}, 1]. \end{cases}$$

3.4.4 Remark One can show that $R_{\mathrm{id},\beta}$ and $\tilde{R}_{\mathrm{id},\beta}$ are least in the set of all confidence down- resp. uprays with level β for $\mathrm{id}_{\mathbb{R}}$.

4 APPLICATION: EXPONENTIAL SAMPLES

- In this chapter and the next ones, the set $\mathcal{Y}$ is $]0,\infty[$, the order being the usual. Its interval topology is thus the usual Euclidean topology and $\mathfrak{B}$ is the Borel σ-algebra on $]0,\infty[$, as agreed in Remark A.1.69, part 4.
- Let us remember that $\beta \in\ ]0,1[$ due to Remarks 1.1.7, part 3, and 1.2.3, part 6, unless stated otherwise.
- Outline of this chapter:
  - Section 4.1 considers several exponential samples with known sample sizes and unknown rates. This is the most general model considered in this chapter. We buehlerize overall minimum and maximum of the samples.
  - The short Section 4.2 specializes to one exponential sample with known size m and unknown rate. The results from the previous section yield confidence bounds for the identity on the parameter space.

4.1 SEVERAL SAMPLES

4.1.1 Definition Let $n \in \mathbb{N}$ and $m \in \mathbb{N}^n$. Let us consider the n-sample exponential model

$$\mathcal{P}_1 := \left(\bigotimes_{i=1}^{n} \mathrm{E}_{\lambda_i}^{\otimes m_i} \colon \lambda \in\ ]0,\infty[^n \right)$$

with known individual sample sizes $m_1, \ldots, m_n$.

4.1.2 Remark The sample space is $\mathcal{X} = \prod_{i=1}^{n}]0,\infty[^{m_i}$, the parameter space $\Theta =]0,\infty[^n$. Let us interpret an observation x as an n-tuple $(x_1,\ldots,x_n)$ of vectors $x_i \in \mathbb{R}^{m_i}$ of possibly different lengths with components $x_{i,1},\ldots,x_{i,m_i}$. In the case $n = 1$, let us interpret $x = (x_1,\ldots,x_m)$ as a vector of length m of strictly positive numbers.

4.1.3 Example Let us consider

$$\textstyle\bigwedge_{i,j} X_{i,j} \colon \prod_{i=1}^{n}]0,\infty[^{m_i} \to\ ]0,\infty[, \quad x \mapsto \bigwedge_{i=1}^{n} \bigwedge_{j=1}^{m_i} x_{i,j},$$

as designated statistic. Since $(\bigwedge_{i,j} X_{i,j}) \,\Box\, \bigotimes_{i=1}^{n} \mathrm{E}_{\lambda_i}^{\otimes m_i} = \mathrm{E}_{\sum_{i=1}^{n} m_i\lambda_i}$, we obtain

$$\bigotimes_{i=1}^{n} \mathrm{E}_{\lambda_i}^{\otimes m_i} \left(\textstyle\bigwedge_{i,j} X_{i,j} > \displaystyle\bigwedge_{i=1}^{n} \bigwedge_{j=1}^{m_i} x_{i,j} \right) = \mathrm{E}_{\sum_{i=1}^{n} m_i\lambda_i} \left(\left] \bigwedge_{i=1}^{n} \bigwedge_{j=1}^{m_i} x_{i,j}, \infty \right[\right) \tag{31}$$

$$= \exp\left(-\bigwedge_{i=1}^{n} \bigwedge_{j=1}^{m_i} x_{i,j} \sum_{k=1}^{n} m_k\lambda_k \right) \tag{32}$$

for $x \in \prod_{i=1}^{n}]0,\infty[^{m_i}$. Since exponential distributions are continuous, Remark 1.1.7, part 10, yields the confidence regions given by

$$R_{\bigwedge_{i,j} X_{i,j},\beta}(x) = \left\{ \lambda \in\ ]0,\infty[^n \colon \sum_{i=1}^{n} m_i\lambda_i > \frac{-\log(\beta)}{\bigwedge_{i=1}^{n} \bigwedge_{j=1}^{m_i} x_{i,j}} \right\}$$

$$\tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(x) = \left\{ \lambda \in \,]0,\infty[^n \colon \sum_{i=1}^n m_i \lambda_i < \frac{-\log(1-\beta)}{\bigwedge_{i=1}^n \bigwedge_{j=1}^{m_i} x_{i,j}} \right\}$$

for $x \in \prod_{i=1}^n]0,\infty[^{m_i}$. Since

$$\exp\left(-\bigwedge_{i,j} X_{i,j} \sum_{k=1}^n m_k \lambda_k \right) \colon \prod_{i=1}^n]0,\infty[^{m_i} \to \,]0,1[$$

is surjective, the effective levels of $R_{\bigwedge_{i,j} X_{i,j},\beta}$ and $\tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}$ are given by

$$\beta_{\text{eff}}(R_{\bigwedge_{i,j} X_{i,j},\beta}) = \beta_{\text{eff}}(\tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}) = \beta \qquad \text{for } \beta \in [0,1].$$

4.1.4 Lemma *Let us consider the parameters of interest*

$$\begin{aligned} \mathrm{pr}_k \colon & \prod_{i=1}^n]0,\infty[^{m_i} \to \,]0,\infty[, \quad \lambda \mapsto \lambda_k, \qquad \textit{for } k \in \{1,\dots,n\} \\ \bigwedge \colon & \prod_{i=1}^n]0,\infty[^{m_i} \to \,]0,\infty[, \quad \lambda \mapsto \bigwedge_{i=1}^n \lambda_i, \\ \bigvee \colon & \prod_{i=1}^n]0,\infty[^{m_i} \to \,]0,\infty[, \quad \lambda \mapsto \bigvee_{i=1}^n \lambda_i. \end{aligned}$$

For $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$ and $k \in \{1,\dots,n\}$ then

$$\mathrm{pr}_k[R_{\bigwedge_{i,j} X_{i,j},\beta}(x)] = \begin{cases}]0,\infty[& \textit{if } n \geq 2 \\ \left] \frac{-\log(\beta)}{m \bigwedge_{j=1}^m x_j}, \infty \right[& \textit{if } n = 1 \end{cases} \tag{33}$$

$$\mathrm{pr}_k[\tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(x)] = \left] 0, \frac{-\log(1-\beta)}{m_k \bigwedge_{i=1}^n \bigwedge_{j=1}^{m_i} x_{i,j}} \right[\tag{34}$$

$$\bigwedge[R_{\bigwedge_{i,j} X_{i,j},\beta}(x)] = \begin{cases}]0,\infty[& \textit{if } n \geq 2 \\ \left] \frac{-\log(\beta)}{m \bigwedge_{j=1}^m x_j}, \infty \right[& \textit{if } n = 1 \end{cases} \tag{35}$$

$$\bigwedge[\tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(x)] = \left] 0, \frac{-\log(1-\beta)}{\bigwedge_{i=1}^n \bigwedge_{j=1}^{m_i} x_{i,j} \sum_{r=1}^n m_r} \right[\tag{36}$$

$$\bigvee[R_{\bigwedge_{i,j} X_{i,j},\beta}(x)] = \left] \frac{-\log(\beta)}{\bigwedge_{i=1}^n \bigwedge_{j=1}^{m_i} x_{i,j} \sum_{r=1}^n m_r}, \infty \right[\tag{37}$$

$$\bigvee[\tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(x)] = \left] 0, \frac{-\log(1-\beta)}{\bigwedge_{r=1}^n m_r \bigwedge_{i=1}^n \bigwedge_{j=1}^{m_i} x_{i,j}} \right[. \tag{38}$$

Proof. Let $x \in \prod_{i=1}^n]0,\infty[^{m_i}$ and $k \in \{1,\dots,n\}$, and let us assume w.l.o.g. $\beta \in \,]0,1[$ (due to Remarks 1.1.7, part 3, and 1.2.3, part 6).

(33) Let us first consider the case $n \geq 2$. The inclusion LHS $\subseteq$ RHS is clear. It thus remains to show LHS $\supseteq$ RHS. To this end, let $t \in \,]0,\infty[$, and let us define $\lambda_k := t$ and $\lambda_i := -\log(\beta)/\bigwedge_{r=1}^n \bigwedge_{j=1}^{m_r} x_{r,j}$ for $i \in \{1,\dots,n\} \setminus \{k\}$. Then

$$\sum_{i=1}^n m_i \lambda_i \geq m_k t + (n-1) \frac{-\log(\beta)}{\bigwedge_{i=1}^n \bigwedge_{j=1}^{m_i} x_{i,j}} > \frac{-\log(\beta)}{\bigwedge_{i=1}^n \bigwedge_{j=1}^{m_i} x_{i,j}},$$

which yields $\lambda \in R_{\bigwedge_{i,j} X_{i,j},\beta}(x)$ due to Example 4.1.3. Since $\mathrm{pr}_k(\lambda) = t$, we have $t \in \mathrm{LHS}$.

Let us now consider the case $n = 1$. This also implies $k = 1$. Example 4.1.3 yields the equivalence

$$\lambda \in R_{\bigwedge_{i,j} X_{i,j},\beta}(x) \qquad \Longleftrightarrow \qquad \lambda > \frac{-\log(\beta)}{m \bigwedge_{j=1}^{m} x_j},$$

which implies the claim.

(34) We first show the inclusion LHS $\subseteq$ RHS. To this end, let $\lambda \in \tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(x)$. Example 4.1.3 yields $\sum_{r=1}^{n} m_r \lambda_r < -\log(\beta) / \bigwedge_{i=1}^{n} \bigwedge_{j=1}^{m_i} x_{i,j}$, which implies

$$\lambda_k < \frac{-\log(1-\beta)}{m_k \bigwedge_{i=1}^{n} \bigwedge_{j=1}^{m_i} x_{i,j}} - \sum_{\substack{r=1 \\ r \neq k}}^{n} \frac{m_r}{m_k} \lambda_r < \frac{-\log(1-\beta)}{m_k \bigwedge_{i=1}^{n} \bigwedge_{j=1}^{m_i} x_{i,j}},$$

i.e., $\mathrm{pr}_k(\lambda) \in \mathrm{RHS}$.

We now show LHS $\supseteq$ RHS. Let $t \in \mathrm{RHS}$, and let us set $\lambda_k := t$. Since

$$\lim_{\substack{\lambda_r \to 0+ \\ \text{for } r \neq k}} \sum_{r=1}^{n} m_r \lambda_r = m_k \lambda_k < \frac{-\log(1-\beta)}{\bigwedge_{i=1}^{n} \bigwedge_{j=1}^{m_i} x_{i,j}},$$

we can pick $\lambda_r \in\,]0, \infty[$ with $\lambda \in \tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(x)$ due to Example 4.1.3. Since $\mathrm{pr}_k(\lambda) = t$, we obtain $t \in \mathrm{LHS}$.

(35) Let us first consider the case $n \geq 2$. The inclusion LHS $\subseteq$ RHS is clear. It thus remains to show LHS $\supseteq$ RHS. To this end, let $t \in\,]0, \infty[$, and let us define $\lambda_1 := t$ and $\lambda_i := t \vee (-\log(\beta) / \bigwedge_{r=1}^{n} \bigwedge_{j=1}^{m_r} x_{r,j})$ for $i \in \{2, \ldots, n\}$. Then $\sum_{i=1}^{n} m_i \lambda_i > -\log(\beta) / \bigwedge_{i=1}^{n} \bigwedge_{j=1}^{m_i} x_{i,j}$, hence $\lambda \in R_{\bigwedge_{i,j} X_{i,j},\beta}(x)$ due to Example 4.1.3. Since $\bigwedge_{i=1}^{n} \lambda_i = \lambda_1 = t$, we have $t \in \mathrm{LHS}$.

The case $n = 1$ behaves identically to the one corresponding to (33): Example 4.1.3 yields the equivalence

$$\lambda \in R_{\bigwedge_{i,j} X_{i,j},\beta}(x) \qquad \Longleftrightarrow \qquad \lambda > \frac{-\log(\beta)}{m \bigwedge_{j=1}^{m} x_j},$$

which implies the claim.

(36) We first show the inclusion LHS $\subseteq$ RHS. To this end, let $\lambda \in \tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(x)$, and let us assume $\bigwedge_{i=1}^{n} \lambda_i \geq \sup \mathrm{RHS}$. Then

$$\sum_{i=1}^{n} m_i \lambda_i \geq \sup \mathrm{RHS} \sum_{r=1}^{n} m_r = \frac{-\log(1-\beta)}{\bigwedge_{i=1}^{n} \bigwedge_{j=1}^{m_i} x_{i,j}},$$

contradicting Example 4.1.3.

We now show LHS $\supseteq$ RHS. Let $t \in \mathrm{RHS}$, and let us define $\lambda_i := t$ for $i \in \{1, \ldots, n\}$. Then

$$\sum_{i=1}^{n} m_i \lambda_i = t \sum_{i=1}^{n} m_i < \frac{-\log(1-\beta)}{\bigwedge_{i=1}^{n} \bigwedge_{j=1}^{m_i} x_{i,j}},$$

i.e., $\lambda \in \tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(x)$ due to Example 4.1.3. Since $\bigwedge_{i=1}^{n} \lambda_i = t$, we have $t \in \mathrm{LHS}$.

(37) We first show the inclusion LHS $\subseteq$ RHS. To this end, let $\lambda \in R_{\bigwedge_{i,j} X_{i,j},\beta}(x)$, and let us assume $\bigvee_{i=1}^n \lambda_i \leq \inf \mathrm{RHS}$. Then

$$\sum_{i=1}^n m_i\lambda_i \leq \inf \mathrm{RHS} \sum_{r=1}^n m_r = \frac{-\log(\beta)}{\bigwedge_{i=1}^n \bigwedge_{j=1}^{m_i} x_{i,j}},$$

contradicting Example 4.1.3.

We now show LHS $\supseteq$ RHS. Let $t \in \mathrm{RHS}$, and let us define $\lambda_i \coloneqq t$ for $i \in \{1,\ldots,n\}$. Then

$$\sum_{i=1}^n m_i\lambda_i = t\sum_{i=1}^n m_i > \frac{-\log(\beta)}{\bigwedge_{i=1}^n \bigwedge_{j=1}^{m_i} x_{i,j}},$$

i.e., $\lambda \in R_{\bigwedge_{i,j} X_{i,j},\beta}(x)$ due to Example 4.1.3. Since $\bigvee_{i=1}^n \lambda_i = t$, we have $t \in \mathrm{LHS}$.

(38) We first show the inclusion LHS $\subseteq$ RHS. To this end, let $\lambda \in \tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(x)$, and let us assume $\bigvee_{i=1}^n \lambda_i \geq \sup \mathrm{RHS}$. Let us pick $i_0 \in \{1,\ldots,n\}$ with $\lambda_{i_0} = \bigvee_{i=1}^n \lambda_i$. Then

$$\sum_{i=1}^n m_i\lambda_i > m_{i_0}\lambda_{i_0} \geq m_{i_0} \sup \mathrm{RHS} = \frac{m_{i_0}}{\bigwedge_{r=1}^n m_r} \cdot \frac{-\log(1-\beta)}{\bigwedge_{i=1}^n \bigwedge_{j=1}^{m_i} x_{i,j}} \geq \frac{-\log(1-\beta)}{\bigwedge_{i=1}^n \bigwedge_{j=1}^{m_i} x_{i,j}},$$

contradicting Example 4.1.3.

We now show LHS $\supseteq$ RHS. Let $t \in \mathrm{RHS}$. Let us pick $i_0 \in \{1,\ldots,n\}$ such that $m_{i_0} = \bigwedge_{i=1}^n m_i$, and let us define $\lambda_{i_0} \coloneqq t$. Since

$$\lim_{\substack{\lambda_i \to 0+ \\ \text{for } i \neq i_0}} \sum_{i=1}^n m_i\lambda_i = m_{i_0} t < \frac{-\log(1-\beta)}{\bigwedge_{i=1}^n \bigwedge_{j=1}^{m_i} x_{i,j}},$$

we can pick $\lambda_i \in \,]0,\infty[$ such that $\lambda \in \tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(x)$ in view of Example 4.1.3. Since $\bigvee_{i=1}^n \lambda_i = t$, we have $t \in \mathrm{LHS}$. □

4.1.5 Remark In case $m_1 = \ldots = m_n =: M$, Example 4.1.3 immediately yields for the parameter of interest $\kappa\colon]0,\infty[^n \to]0,\infty[$, $\lambda \mapsto \sum_{k=1}^n \lambda_k$ the confidence regions given for $x \in \prod_{i=1}^n]0,\infty[^M$ by

$$\kappa[R_{\bigwedge_{i,j} X_{i,j},\beta}(x)] = \left]\frac{-\log(\beta)}{M\bigwedge_{i=1}^n \bigwedge_{j=1}^{M} x_{i,j}}, \infty\right[$$
$$\kappa[\tilde{R}_{\bigwedge_{i,j} X_{i,j},\beta}(x)] = \left]0, \frac{-\log(1-\beta)}{M\bigwedge_{i=1}^n \bigwedge_{j=1}^{M} x_{i,j}}\right[.$$

4.1.6 Example Let us consider

$$\textstyle\bigvee_{i,j} X_{i,j}\colon \displaystyle\prod_{i=1}^n]0,\infty[^{m_i} \to]0,\infty[, \quad x \mapsto \bigvee_{i=1}^n \bigvee_{j=1}^{m_i} x_{i,j},$$

as designated statistic. The situation is now similar to, but cannot be readily reduced to the one from Example 4.1.3. We have

$$\bigotimes_{i=1}^n \mathrm{E}_{\lambda_i}^{\otimes m_i}\left(\textstyle\bigvee_{i,j} X_{i,j} < \displaystyle\bigvee_{i=1}^n \bigvee_{j=1}^{m_i} x_{i,j}\right) = \prod_{k=1}^n \left(1 - \mathrm{e}^{-\lambda_k \bigwedge_{i=1}^n \bigwedge_{j=1}^{m_i} x_{i,j}}\right)^{m_k} \tag{39}$$

for $x \in \prod_{i=1}^n]0,\infty[^{m_i}$. Since exponential distributions are continuous, Remark 1.1.7, part 10, yields the confidence regions given by

$$R_{\bigvee_{i,j} X_{i,j},\beta}(x) = \left\{\lambda \in]0,\infty[^n \colon \prod_{k=1}^n \left(1 - \mathrm{e}^{-\lambda_k \bigwedge_{i=1}^n \bigwedge_{j=1}^{m_i} x_{i,j}}\right)^{m_k} > 1-\beta\right\}$$
$$\tilde{R}_{\bigvee_{i,j} X_{i,j},\beta}(x) = \left\{\lambda \in]0,\infty[^n \colon \prod_{k=1}^n \left(1 - \mathrm{e}^{-\lambda_k \bigwedge_{i=1}^n \bigwedge_{j=1}^{m_i} x_{i,j}}\right)^{m_k} < \beta\right\}$$

for $x \in \prod_{i=1}^n]0,\infty[^{m_i}$. Since

$$\prod_{k=1}^n \left(1 - \mathrm{e}^{-\lambda_k \bigwedge_{i,j} X_{i,j}}\right)^{m_k} \colon \prod_{i=1}^n]0,\infty[^{m_i} \to]0,1[$$

is surjective, the effective levels of $R_{\bigvee_{i,j} X_{i,j},\beta}$ and $\tilde{R}_{\bigvee_{i,j} X_{i,j},\beta}$ are given by

$$\beta_{\mathrm{eff}}(R_{\bigvee_{i,j} X_{i,j},\beta}) = \beta_{\mathrm{eff}}(\tilde{R}_{\bigvee_{i,j} X_{i,j},\beta}) = \beta \qquad \text{for } \beta \in [0,1].$$

4.1.7 Remark Let us notice that

$$]0,\infty[^n \to]0,1[, \quad \lambda \mapsto \prod_{k=1}^n \left(1 - \mathrm{e}^{-\lambda_k t}\right)^{m_k},$$

is strictly increasing for $t \in]0,\infty[$.

4.1.8 Lemma *Let us consider the parameters of interest*

$$\begin{aligned}
\mathrm{pr}_k \colon & \prod_{i=1}^n]0,\infty[^{m_i} \to]0,\infty[, \quad \lambda \mapsto \lambda_k, \qquad \textit{for } k \in \{1,\ldots,n\}\\
\bigwedge \colon & \prod_{i=1}^n]0,\infty[^{m_i} \to]0,\infty[, \quad \lambda \mapsto \bigwedge_{i=1}^n \lambda_i,\\
\bigvee \colon & \prod_{i=1}^n]0,\infty[^{m_i} \to]0,\infty[, \quad \lambda \mapsto \bigvee_{i=1}^n \lambda_i.
\end{aligned}$$

For $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$ and $k \in \{1,\ldots,n\}$ then

$$\mathrm{pr}_k[R_{\bigvee_{i,j} X_{i,j},\beta}(x)] = \left]\frac{-\log(1-(1-\beta)^{1/m_k})}{\bigvee_{i=1}^n \bigvee_{j=1}^{m_i} x_{i,j}}, \infty\right[\tag{40}$$

$$\mathrm{pr}_k[\tilde{R}_{\bigvee_{i,j} X_{i,j},\beta}(x)] = \begin{cases}]0,\infty[& \textit{if } n \geq 2\\ \left]0, \frac{-\log(1-\beta^{1/m})}{\bigvee_{j=1}^m x_j}\right[& \textit{if } n = 1 \end{cases} \tag{41}$$

$$\bigwedge[R_{\bigvee_{i,j} X_{i,j},\beta}(x)] = \left]\frac{-\log(1-(1-\beta)^{1/\bigwedge_{r=1}^n m_r})}{\bigvee_{i=1}^n \bigvee_{j=1}^{m_i} x_{i,j}}, \infty\right[\tag{42}$$

$$\bigwedge[\tilde{R}_{\bigvee_{i,j} X_{i,j},\beta}(x)] = \left]0, \frac{-\log(1-\beta^{1/\sum_{r=1}^n m_r})}{\bigvee_{i=1}^n \bigvee_{j=1}^{m_i} x_{i,j}}\right[\tag{43}$$

$$\bigvee[R_{\bigvee_{i,j} X_{i,j},\beta}(x)] = \left]\frac{-\log(1-(1-\beta)^{1/\sum_{r=1}^n m_r})}{\bigvee_{i=1}^n \bigvee_{j=1}^{m_i} x_{i,j}}, \infty\right[\tag{44}$$

$$\bigvee[\tilde{R}_{\bigvee_{i,j} X_{i,j},\beta}(x)] = \begin{cases}]0,\infty[& \text{if } n \geq 2 \\ \left]0, \frac{-\log(1-\beta^{1/m})}{\bigvee_{j=1}^{m} x_j}\right[& \text{if } n = 1. \end{cases} \tag{45}$$

Proof. Let $x \in \prod_{i=1}^{n}]0,\infty[^{m_i}$ and $k \in \{1,\ldots,n\}$, and let us assume w.l.o.g. $\beta \in]0,1[$ (due to Remarks 1.1.7, part 3, and 1.2.3, part 6).

(40) We first show LHS $\subseteq$ RHS. To this end, let $\lambda \in R_{\bigvee_{i,j} X_{i,j},\beta}(x)$. Example 4.1.6 then yields

$$\left(1 - \mathrm{e}^{-\lambda_k \bigvee_{i=1}^{n} \bigvee_{j=1}^{m_i} x_{i,j}}\right)^{m_k} > (1-\beta) \prod_{\substack{r=1 \\ r \neq k}}^{n} \left(1 - \mathrm{e}^{-\lambda_r \bigvee_{i=1}^{n} \bigvee_{j=1}^{m_i} x_{i,j}}\right)^{-m_r} > 1 - \beta,$$

i.e., $\lambda_k \in$ RHS.

We now show LHS $\supseteq$ RHS. Let $t \in$ RHS, and let us define $\lambda_k := t$. Since

$$\lim_{\substack{\lambda_r \to \infty \\ \text{for } r \neq k}} \prod_{r=1}^{n} \left(1 - \mathrm{e}^{-\lambda_r \bigvee_{i=1}^{n} \bigvee_{j=1}^{m_i} x_{i,j}}\right)^{m_r} = \left(1 - \mathrm{e}^{-\lambda_k \bigvee_{i=1}^{n} \bigvee_{j=1}^{m_i} x_{i,j}}\right)^{m_k} > 1 - \beta,$$

we can pick $\lambda_r \in]0,\infty[$ for $r \in \{1,\ldots,n\} \setminus \{k\}$ such that $\lambda \in R_{\bigvee_{i,j} X_{i,j},\beta}(x)$ due to Example 4.1.6. Since $\mathrm{pr}_k(\lambda) = t$, this implies $t \in$ LHS.

(41) We first consider the case $n \geq 2$. The inclusion LHS $\subseteq$ RHS is clear. It thus remains to show LHS $\supseteq$ RHS. To this end, let $t \in]0,\infty[$, and let us define $\lambda_k := t$. Since, for any $l \in \{1,\ldots,n\} \setminus \{k\}$,

$$\lim_{\lambda_l \to 0+} \prod_{r=1}^{n} \left(1 - \mathrm{e}^{-\lambda_r \bigvee_{i=1}^{n} \bigvee_{j=1}^{m_i} x_{i,j}}\right)^{m_r} = 0 < \beta,$$

we can pick $\lambda_r \in]0,\infty[$ for $r \in \{1,\ldots,n\} \setminus \{k\}$ such that $\lambda \in \tilde{R}_{\bigvee_{i,j} X_{i,j},\beta}(x)$ due to Example 4.1.6. Since $\mathrm{pr}_k(\lambda) = t$, we obtain $t \in$ LHS.

Let us now consider the case $n = 1$. This also implies $k = 1$. Example 4.1.6 yields the equivalence

$$\lambda \in \tilde{R}_{\bigvee_{i,j} X_{i,j},\beta}(x) \quad \iff \quad \lambda < \frac{-\log(1-\beta^{1/m})}{\bigvee_{j=1}^{m} x_j},$$

which implies the claim.

(42) We shall first show the inclusion LHS $\subseteq$ RHS. To this end, let $\lambda \in R_{\bigvee_{i,j} X_{i,j},\beta}(x)$. For $l \in \{1,\ldots,n\}$ then

$$\left(1 - \mathrm{e}^{-\lambda_l \bigvee_{i=1}^{n} \bigvee_{j=1}^{m_i} x_{i,j}}\right)^{m_l} > \prod_{r=1}^{n} \left(1 - \mathrm{e}^{-\lambda_r \bigvee_{i=1}^{n} \bigvee_{j=1}^{m_i} x_{i,j}}\right)^{m_r} > 1 - \beta,$$

hence $\lambda_l > \bigwedge_{r=1}^{n}(-\log(1-(1-\beta)^{1/m_r})/\bigvee_{i=1}^{n} \bigvee_{j=1}^{m_i} x_{i,j}) = \inf \text{RHS}$, i.e., $\bigwedge_{l=1}^{n} \lambda_l \in$ RHS.

We now show LHS $\supseteq$ RHS. Let $t \in$ RHS. Let us pick $r_0 \in \{1, \dots, n\}$ such that $m_{r_0} = \bigwedge_{r=1}^n m_r$, and let us define $\lambda_{r_0} := t$. Since

$$\lim_{\substack{\lambda_r \to \infty \\ \text{for } r \neq r_0}} \prod_{r=1}^n \left(1 - \mathrm{e}^{-\lambda_r \bigvee_{i=1}^n \bigvee_{j=1}^{m_i} x_{i,j}}\right)^{m_r} = \left(1 - \mathrm{e}^{-\lambda_{r_0} \bigvee_{i=1}^n \bigvee_{j=1}^{m_i} x_{i,j}}\right)^{m_{r_0}} > 1 - \beta,$$

we can choose $\lambda_r \in \,]t, \infty[$ for $r \in \{1, \dots, n\} \setminus \{r_0\}$ such that $\lambda \in R_{\bigvee_{i,j} X_{i,j}, \beta}(x)$ due to Example 4.1.6. Since $\bigwedge_{r=1}^n \lambda_r = t$, we obtain $t \in$ LHS.

(43) We first show the inclusion LHS $\subseteq$ RHS. To this end, let $\lambda \in \tilde{R}_{\bigvee_{i,j} X_{i,j}, \beta}(x)$, and let us assume $\bigwedge_{i=1}^n \lambda_i \geq \sup$ RHS. Remark 4.1.7 then implies

$$\prod_{r=1}^n \left(1 - \mathrm{e}^{-\lambda_r \bigvee_{i=1}^n \bigvee_{j=1}^{m_i} x_{i,j}}\right)^{m_r} \geq \left(1 - \mathrm{e}^{-\sup \mathrm{RHS} \bigvee_{i=1}^n \bigvee_{j=1}^{m_i} x_{i,j}}\right)^{\sum_{r=1}^n m_r} = \beta,$$

contradicting Example 4.1.6.

We now show LHS $\supseteq$ RHS. Let $t \in$ RHS, and let us define $\lambda_i := t$ for $i \in \{1, \dots, n\}$. Remark 4.1.7 then implies

$$\prod_{r=1}^n \left(1 - \mathrm{e}^{-\lambda_r \bigvee_{i=1}^n \bigvee_{j=1}^{m_i} x_{i,j}}\right)^{m_r} < \left(1 - \mathrm{e}^{-\sup \mathrm{RHS} \bigvee_{i=1}^n \bigvee_{j=1}^{m_i} x_{i,j}}\right)^{\sum_{r=1}^n m_r} = \beta$$

i.e., $\lambda \in \tilde{R}_{\bigvee_{i,j} X_{i,j}, \beta}(x)$ due to Example 4.1.6. Since $\bigwedge_{i=1}^n \lambda_i = t$, we have $t \in$ LHS.

(44) We first show the inclusion LHS $\subseteq$ RHS. To this end, let $\lambda \in R_{\bigvee_{i,j} X_{i,j}, \beta}(x)$, and let us assume $\bigvee_{i=1}^n \lambda_i \leq \inf$ RHS. Remark 4.1.7 then implies

$$\prod_{r=1}^n \left(1 - \mathrm{e}^{-\lambda_r \bigvee_{i=1}^n \bigvee_{j=1}^{m_i} x_{i,j}}\right)^{m_r} \leq \left(1 - \mathrm{e}^{-\inf \mathrm{RHS} \bigvee_{i=1}^n \bigvee_{j=1}^{m_i} x_{i,j}}\right)^{\sum_{r=1}^n m_r} = 1 - \beta,$$

contradicting Example 4.1.6.

We now show LHS $\supseteq$ RHS. Let $t \in$ RHS, and let us define $\lambda_i := t$ for $i \in \{1, \dots, n\}$. Remark 4.1.7 then implies

$$\prod_{r=1}^n \left(1 - \mathrm{e}^{-\lambda_r \bigvee_{i=1}^n \bigvee_{j=1}^{m_i} x_{i,j}}\right)^{m_r} > \left(1 - \mathrm{e}^{-\sup \mathrm{RHS} \bigvee_{i=1}^n \bigvee_{j=1}^{m_i} x_{i,j}}\right)^{\sum_{r=1}^n m_r} = 1 - \beta$$

i.e., $\lambda \in R_{\bigvee_{i,j} X_{i,j}, \beta}(x)$ due to Example 4.1.6. Since $\bigvee_{i=1}^n \lambda_i = t$, we have $t \in$ LHS.

(45) We first the case $n \geq 2$. The inclusion LHS $\subseteq$ RHS is clear. It thus remains to show LHS $\supseteq$ RHS. To this end, let $t \in \,]0, \infty[$, and let us define $\lambda_1 := t$. Since, for any $l \in \{2, \dots, n\}$,

$$\lim_{\lambda_l \to 0+} \prod_{r=1}^n \left(1 - \mathrm{e}^{-\lambda_r \bigvee_{i=1}^n \bigvee_{j=1}^{m_i} x_{i,j}}\right)^{m_r} = 0 < \beta,$$

we can pick $\lambda_2, \dots, \lambda_n \in \,]0, t[$ such that $\lambda \in \tilde{R}_{\bigvee_{i,j} X_{i,j}, \beta}(x)$ in view of Example 4.1.6. Since $\bigvee_{r=1}^n \lambda_r = t$, we obtain $t \in$ LHS.

The case $n = 1$ behaves identically to the one corresponding to (41): Example 4.1.6 yields the equivalence

$$\lambda \in \tilde{R}_{\bigvee_{i,j} X_{i,j},\beta}(x) \quad \Longleftrightarrow \quad \lambda < \frac{-\log(1-\beta^{1/m})}{\bigvee_{j=1}^{m} x_j},$$

which implies the claim. □

4.1.9 Remark 1. We have not been able to establish an analogue to Remark 4.1.5 for the designated statistic $\bigvee_{i,j} X_{i,j}$ and the sum as parameter of interest without a considerable amount of calculations.

2. Determining the Buehlerization of the sum

$$S\colon \prod_{i=1}^{n}]0,\infty[^{m_i} \to]0,\infty[, \quad x \mapsto \sum_{i=1}^{n}\sum_{j=1}^{m_i} x_{i,j},$$

is in simple cases possible using the density of the convolution of exponential distributions provided by Akkouchi (2008).

4.2 ONE SAMPLE

4.2.1 Definition Let $m \in \mathbb{N}$. Let us consider the one-sample exponential model

$$\mathcal{P}_2 := (\mathrm{E}_\lambda^{\otimes m} \colon \lambda \in]0,\infty[)$$

with known sample size m.

4.2.2 Remark The sample space is $\mathcal{X} =]0,\infty[^m$, the parameter space $\Theta =]0,\infty[$.

4.2.3 Example If we buehlerize $\bigwedge\colon]0,\infty[^m \to]0,\infty[$, $x \mapsto \bigwedge_{i=1}^{m} x_i$, and $\bigvee\colon]0,\infty[^m \to]0,\infty[$, $x \mapsto \bigvee_{i=1}^{m} x_i$, then Examples 4.1.3 and 4.1.6 yield for the identity on the parameter space the confidence regions given for $x \in]0,\infty[^m$ by

$$R_{\bigwedge,\beta}(x) = \left]\frac{-\log(\beta)}{m\bigwedge_{i=1}^{m} x_i},\infty\right[\qquad \tilde{R}_{\bigwedge,\beta}(x) = \left]0,\frac{-\log(1-\beta)}{m\bigwedge_{i=1}^{m} x_i}\right[$$

$$R_{\bigvee,\beta}(x) = \left]\frac{-\log(1-(1-\beta)^{1/m})}{\bigvee_{i=1}^{m} x_i},\infty\right[\qquad \tilde{R}_{\bigvee,\beta}(x) = \left]0,\frac{-\log(1-\beta^{1/m})}{\bigvee_{i=1}^{m} x_i}\right[.$$

5 COMPARING PAIRS OF DIAGNOSTIC TESTS

- This chapter's (unattained) aim is the construction, in analogy to Mattner and Mattner (2013), of confidence bounds in models describing pairs of diagnostic tests. Let us assume that we would like to statistically compare two diagnostic tests without being able to observe the true states of the members of the underlying population. Let us also assume no knowledge about the prevalence of the condition that is being examined or any kind of independence between the two diagnostic tests. In some cases it might be reasonable to assume that one diagnostic test is, e.g., more sensitive than the other. A confidence bound for the difference of the two diagnostic tests' specificities can in such a case yield superiority of one test over the other.
- The main result of this chapter is the diagram in Theorem 5.3.6. It relates several models for pairs of diagnostic tests. The notation is similar to the one employed in Mattner and Mattner (2013). Most results in this chapter rely on computations made "by hand" and are incomplete in the sense that central questions remain unfortunately unanswered.
- Outline of this chapter:
  - Section 5.1 gives an brief informal introduction to the notion of diagnostic test.
  - Section 5.2 presents a statistical model for handling pairs of diagnostic tests due to Gart and Buck (1966).
  - Section 5.3 contains the above mentioned main result, which establishes relations
    - between the sets of lower confidence bounds,
    - between the sets of pairs of comparable lower confidence bounds,
    - between the sets of admissible lower confidence bounds

    in various models for diverse parameters of interest.
  - The proof of the main result rests on a number of auxiliary results, which are provided in Section 5.4. These results are of essentially two kinds: propositions allowing a (sometimes partial) reduction of a statement in a certain model to one in a similar, already covered model, and lemmas where images under certain linear maps of semialgebraic sets are computed in an elementary manner (that is to say, without tools from real algebraic geometry).

5.1 INFORMAL INTRODUCTION

5.1.1 Definition By a **(dichotomous) diagnostic test** we mean any procedure for classifying objects of a fixed set, called **population,** into two states.

5.1.2 Remark 1. The state space is most often taken to be $\{0, 1\}$, with 1 being interpreted as "condition present" or "positive", and 0 as "condition absent" or "negative."

2. Although the term "diagnostic test" is customarily used in a mainly medical context, it naturally occurs in numerous other areas, among them psychology, quality assurance, financial engineering, and everyday life.

5.1.3 Example 1. From medicine: measurement of body temperature to diagnose influenza; checking for tooth ache to diagnose dental caries; performance of an ELISA test to determine the presence of HIV; using a breathalyzer to determine alcohol consumption.

2. From psychology: performance of an IQ test to determine above average intelligence; taking a personality test to diagnose schizophrenia.
3. From quality assurance: "trying out" specimina to determine whether ot not a product is durable; immersion into fluids to diagnose watertightness.
4. From financial engineering: looking at the distribution of up- and downcrossings of a financial product's value to detect a high volatility.
5. From everyday life: measuring air humidity to predict whether or not it will rain in the next few hours.

5.1.4 Remark Is it preferable in order to diagnose influenza to simply measure one's body temperature or to pay a visit to the physician? The latter diagnostic test may be more reliable, but is also more expensive and more time-consuming than the former. In order to decide whether a more sophisticated test is worth its money and effort, we need some way of quantifying its features of interest to us.

5.1.5 Definition (Sensitivity and specificity) Given a diagnostic test for some condition, its **sensitivity** is the probability of a positive test result given the presence of the condition, and its **specificity** is the probability of a negative test result given the absence of the condition. In symbols and with $\mathbb{P}$ denoting "the underlying" probability measure,

$$\begin{aligned} \text{sensitivity} &:= \mathbb{P}(\text{test} = 1 \mid \text{state} = 1) \\ \text{specificity} &:= \mathbb{P}(\text{test} = 0 \mid \text{state} = 0). \end{aligned}$$

5.1.6 Remark 1. The pair consisting of the sensitivity and specificity of a diagnostic test represent the *accuracy* of that diagnostic test.

2. Of greater interest to practitioners, who tend to be more oriented towards a prognosis, are two different figures: the **positive predictive value,** which is the probability of the condition being present given a positive test result, and the **negative predictive value,** which is the probability of the condition being absent given a negative test result:

$$\begin{aligned} \text{positive predictive value} &:= \mathbb{P}(\text{state} = 1 \mid \text{test} = 1) \\ \text{negative predictive value} &:= \mathbb{P}(\text{state} = 0 \mid \text{test} = 0). \end{aligned}$$

3. The pair consisting of the predictive values of a diagnostic test represent the *usefulness* of that diagnostic test.
4. In defining the predictive values we treated these numbers as attributes of the diagnostic test solely. This is, however, not the case—they also depend on another figure, intrinsic to the population: the **prevalence** of the condition, which is the probability that the condition is present in the population at hand. Using the predictive values in order to compare two diagnostic tests is thus improper since one of the tests may dominate the other in terms of their predictive values inside one population but may be inferior inside a different population. This is the reason we chose to compare diagnostic tests according to their accuracy and not their usefulness.

5.2 A MODEL FOR TWO DIAGNOSTIC TESTS

5.2.1 Remark In the following we shall often need to sum over some of the indices of a multi-indexed family. For instance, if $x \in [0,\infty[^{\{0,1\}^2}$, then we would like to have $x_{+0} = x_{0,0} + x_{1,0}$, $x_{1+} = x_{1,0} + x_{1,1}$, and $x_{++} = x_{+0} + x_{+1} = x_{0,0} + x_{0,1} + x_{1,0} + x_{1,1}$. This motivates the following, somewhat unusual definition.

5.2.2 Definition (Summation notation for arrays) Let $(I_j : j \in J)$ be a finite family of finite sets I_j, $I \coloneqq \prod_{j\in J} I_j$ its cartesian product, and $\mathcal{X}$ a subset of an abelian group with operation $+$ (that is not contained in either set I_j). Given $i' \in \prod_{j\in J}(I_j \cup \{+\})$ and $x \in \mathcal{X}$, we set $x_{i'} \coloneqq \sum x_i$, where the sum is taken over all $i \in I$ with $i|_{\{i'\neq +\}} = i'|_{\{i'\neq +\}}$.

5.2.3 Definition (Notation for counting densities) Given two sets $\mathcal{X}$ and $\mathcal{Y}$, we denote by

$$\operatorname{prob}(\mathcal{X}) \coloneqq \left\{ f \in [0,1]^{\mathcal{X}} \colon \sum_{x\in\mathcal{X}} f_x = 1 \right\}$$

the set of probability counting densities on $\mathcal{X}$ and by

$$\operatorname{mark}(\mathcal{X},\mathcal{Y}) \coloneqq \{ f \in [0,1]^{\mathcal{X}\times\mathcal{Y}} \colon f(x,\,\cdot\,) \in \operatorname{prob}(\mathcal{Y}) \text{ for } x \in \mathcal{X} \}$$

the set of Markov (transition) counting densities from $\mathcal{X}$ to $\mathcal{Y}$. In this context we write $f_{y|x}$ for $f(x,y)$ if $f \in \operatorname{mark}(\mathcal{X},\mathcal{Y})$.

5.2.4 Definition (Multinomial distribution) Let $\mathcal{X}$ be a finite set. We denote by $\mathrm{M}_{n,p}$ the **multinomial distribution** with *sample size* $n \in \mathbb{N}$ and *outcome probabilities* $p \in \operatorname{prob}(\mathcal{X})$, given by the probability counting density

$$\{0,\dots,n\}^{\mathcal{X}} \to [0,1], \quad k \mapsto \binom{n}{k} p^k,$$

where

$$\binom{n}{k} \coloneqq \frac{n!}{\prod_{x\in\mathcal{X}} k_x!} \cdot \mathbf{1}\left(\sum_{x\in\mathcal{X}} k_x = n \right)$$

denotes the **multinomial coefficient** for $n \in \mathbb{N}_0$ and $k \in \mathbb{N}_0^{\mathcal{X}}$.

5.2.5 Remark The mapping

$$\operatorname{prob}(\mathcal{X}) \to \operatorname{Prob}(\{0,\dots,n\}^{\mathcal{X}}), \quad p \mapsto \mathrm{M}_{n,p},$$

is injective: if $p, q \in \operatorname{prob}(\mathcal{X})$ with $p \neq q$, there is $x \in \mathcal{X}$ with $p_x \neq q_x$, hence $k \coloneqq n\mathbf{1}_{\{x\}} \in \{0,\dots,n\}^{\mathcal{X}}$ satisfies $\mathrm{M}_{n,p}(\{k\}) = p_x^n \neq q_x^n = \mathrm{M}_{n,q}(\{k\})$.

5.2.6 Remark Let us consider a pair of diagnostic tests applied each to a population of size $n \in \mathbb{N}$. The prevalence of the diagnosis within the underlying population is a number $\pi_1 \in [0,1]$ and can thus be extended to a probability counting density $\pi \in \operatorname{prob}(\{0,1\})$. The two diagnostic tests can be described by a Markov counting density $\chi \in \operatorname{mark}(\{0,1\},\{0,1\}^2)$, where, e.g., $\chi_{0,1|1}$ stands for the probability that an individual, chosen randomly from the part of the population consisting of positive individuals, tests negatively by means of the first test and positively by means of the second.

The two diagnostic tests may also be considered separately by defining two Markov counting densities χ', $\chi'' \in \mathrm{mark}(\{0,1\},\{0,1\})$ by $\chi'_{j|i} := \chi_{j+|i} = \chi_{j,0|i} + \chi_{j,1|i}$ and $\chi''_{j|i} := \chi_{+j|i} = \chi_{0,j|i} + \chi_{1,j|i}$ for $i,j \in \{0,1\}$ (see Definition 5.2.3 for the notation involving a subscript "+"). The sensitivity and specificity of the first test are then given by $\chi'_{1|1}$ and $\chi'_{0|0}$, respectively, and analogously for the second test.

Let us stress that knowledge of the two individual diagnostic tests χ' and χ'' does not suffice in order to recover the original pair of diagnostic tests χ, except in the case of conditional independence of the two tests.

5.2.7 Definition (Pairs of diagnostic tests) Let us consider the parameter space

$$\Theta := \mathrm{prob}(\{0,1\}) \times \mathrm{mark}(\{0,1\},\{0,1\}^2).$$

Given a parameter $(\pi,\chi) \in \Theta$, its *joint density* is given by $\pi \otimes \chi\colon \{0,1\} \times \{0,1\}^2 \to [0,1]$, $(i,j) \mapsto \pi_i \chi_{j|i}$. We will denote the second marginal density of $\pi \otimes \chi$ by $\mu(\pi,\chi)$, i.e., μ maps Θ onto $\mathrm{prob}(\{0,1\}^2)$ and is given by

$$\mu(\pi,\chi)_j = \pi_0 \chi_{j|0} + \pi_1 \chi_{j|1} \qquad \text{for } j \in \{0,1\}^2.$$

The model is then taken to be

$$\mathcal{P} := (\mathrm{M}_{n,\mu(\pi,\chi)}\colon (\pi,\chi) \in \Theta),$$

consisting thus of all multinomial distributions $\mathrm{M}_{n,\mu(\pi,\chi)}$ with sample size n and outcome probabilities given by the second marginal density $\mu(\pi,\chi)$.

5.2.8 Remark 1. This model is in essence due to Gart and Buck (1966).

2. Many parameters of interest are non-identifiable in $\mathcal{P}$ (identifiability is defined in Definition A.3.12). Mattner and Mattner (2013, Lemma 2.8) yields for instance the non-identifiability of $\Theta \to [0,1]$, $(\pi,\chi) \mapsto \pi_1$.
3. The consideration of $\mu(\pi,\chi)$ rather than (π,χ) makes sense and is even required by the fact that true states are *de facto* unobservable.
4. A higher specificity of the second test over the first is expressible as $\chi''_{0|0} \geq \chi'_{0|0}$ or, equivalently, $\chi_{1,0|0} \geq \chi_{0,1|0}$; a higher sensitivity of the second test over the first as $\chi''_{1|1} \geq \chi'_{1|1}$ or, equivalently, $\chi_{0,1|1} \geq \chi_{1,0|1}$.
5. Certain situations allow postulating some relation between the individual diagnostic tests. Suppose, for instance, that the first test constitutes a part of the second and that the second then yields a positive result if already the first one does. It is then plausible to assume the first test at most as sensitive as the second one, i.e., $\chi'_{1|1} \leq \chi''_{1|1}$. This suggests considering submodels $\mathcal{P}_R := \mathcal{P}|_{\Theta_R} = (\mathrm{M}_{n,\mu(\pi,\chi)}\colon (\pi,\chi) \in \Theta_R)$ obtained by restricting the parameter space to

$$\Theta_R := \{(\pi,\chi) \in \Theta\colon (\chi'_{0|0}, \chi'_{1|1})\, R\, (\chi''_{0|0}, \chi''_{1|1})\},$$

where R is a binary relation on $\mathbb{R}^2$. The relation in the above example, for instance, would be $R = \mathbb{R}^2 \otimes \leq$ (we refer to Definition A.1.22, part 3, for product relations).
6. The mapping $R \mapsto \Theta_R$ is increasing, i.e., $R \subseteq S \subseteq \mathbb{R}^4$ implies $\Theta_R \subseteq \Theta_S$.
7. The set $\Theta_{2,\leq}$ from Mattner and Mattner (2013) is represented here by $\Theta_{\leq \otimes \mathbb{R}^2}$.

5.2.9 Example This numerical example justifies the consideration of submodels $\mathcal{P}_R$. Aydiner *et al.* (2012) consider three tests for screening for methicillin-resistant *Staphylococcus* *aureus* (MRSA): LightCycler® Advanced MRSA, Detect-Ready® MRSA, and CHROMagar MRSA. The latter plays the role of gold standard (reference test) for comparing the first two tests. It is thus reasonable to work in the model $\mathcal{P}_{\leq\otimes\leq}$ when the first test is either LightCycler® Advanced MRSA or Detect-Ready® MRSA, and the second test is CHROMagar MRSA. The summary of their data can be found in Tables 2 and 3, where, e.g., $k_{0,1} = 5$ from Table 2 means that five patients were diagnosed negatively by LightCycler® Advanced MRSA and positively by CHROMagar MRSA.

$k_{0,0} = 1000$	$k_{0,1} = 5$	$k_{0+} = 1005$
$k_{1,0} = 15$	$k_{1,1} = 27$	$k_{1+} = 42$
$k_{+0} = 1015$	$k_{+1} = 32$	$k_{++} = 1047$

Table 2: LightCycler® Advanced MRSA vs CHROMagar MRSA

$k_{0,0} = 978$	$k_{0,1} = 11$	$k_{0+} = 989$
$k_{1,0} = 4$	$k_{1,1} = 15$	$k_{1+} = 19$
$k_{+0} = 982$	$k_{+1} = 26$	$k_{++} = 1008$

Table 3: Detect-Ready® MRSA vs CHROMagar MRSA

5.3 RELATING MODELS FOR TWO DIAGNOSTIC TESTS

5.3.1 Definition For a real number $x \in \mathbb{R}$, let

$$x^+ := x \vee 0 \qquad \text{and} \qquad x^- := -x \wedge 0$$

denote its **positive** and **negative parts,** respectively, so that $x = x^+ - x^-$ and $|x| = x^+ + x^-$.

5.3.2 Definition (Relations between experiments) Let $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ and $\mathcal{Q} = (Q_\eta \colon \eta \in \mathrm{H})$ be two models on a common measurable space $(\mathcal{X}, \mathfrak{A})$, let $\kappa \colon \Theta \to \overline{\mathbb{R}}$ and $\lambda \colon \mathrm{H} \to \overline{\mathbb{R}}$ be parameters of interest for $\mathcal{P}$ and $\mathcal{Q}$, respectively, and let $\beta \in [0, 1]$. We write

$(\mathcal{P}, \kappa) \Longrightarrow (\mathcal{Q}, \lambda)$ if every lower confidence bound with level β for the parameter of interest κ in the model $\mathcal{P}$ is also one for the parameter of interest λ in the model $\mathcal{Q}$;

$(\mathcal{P}, \kappa) \rightsquigarrow (\mathcal{Q}, \lambda)$ if, given two lower confidence bounds L and L' such that

- both are valid for the parameter of interest κ in the model $\mathcal{P}$,
- both are valid for the parameter of interest λ in the model $\mathcal{Q}$,
- both have level β in both of the above settings,

the superiority (see Definition A.3.19) of L over L' with respect to κ in $\mathcal{P}$ implies the superiority of L over L' with respect to λ in $\mathcal{Q}$;

$(\mathcal{P}, \kappa) \twoheadrightarrow (\mathcal{Q}, \lambda)$ if, given a lower confidence bound L with level β both for the parameter of interest κ in the model $\mathcal{P}$ and for the parameter of interest λ in the model $\mathcal{Q}$, admissibility of L with respect to κ in $\mathcal{P}$ implies admissibility of L with respect to λ in $\mathcal{Q}$;

$(\mathcal{P}, \kappa) \Leftrightarrow (\mathcal{Q}, \lambda)$ if $(\mathcal{P}, \kappa) \Rightarrow (\mathcal{Q}, \lambda)$ and $(\mathcal{Q}, \lambda) \Rightarrow (\mathcal{P}, \kappa)$;

$(\mathcal{P}, \kappa) \leftrightsquigarrow (\mathcal{Q}, \lambda)$ if $(\mathcal{P}, \kappa) \rightsquigarrow (\mathcal{Q}, \lambda)$ and $(\mathcal{Q}, \lambda) \rightsquigarrow (\mathcal{P}, \kappa)$;

$(\mathcal{P}, \kappa) \twoheadleftarrow\!\!\twoheadrightarrow (\mathcal{Q}, \lambda)$ if $(\mathcal{P}, \kappa) \twoheadrightarrow (\mathcal{Q}, \lambda)$ and $(\mathcal{Q}, \lambda) \twoheadrightarrow (\mathcal{P}, \kappa)$.

If more than one relation holds between $(\mathcal{P}, \kappa)$ and $(\mathcal{Q}, \lambda)$, we shall write the relations more succinctly on top of each other, e.g., $(\mathcal{P}, \kappa) \overset{\Leftarrow}{\twoheadrightarrow} (\mathcal{Q}, \lambda)$.

5.3.3 Remark The relations $\Rightarrow$, $\rightsquigarrow$, and $\twoheadrightarrow$ are reflexive, $\Rightarrow$, $\overset{\Rightarrow}{\rightsquigarrow}$, $\overset{\Rightarrow}{\leftsquigarrow}$, $\overset{\Rightarrow}{\twoheadrightarrow}$, and $\overset{\Rightarrow}{\twoheadleftarrow}$ are transitive, and in general no combination of them is symmetric (except the obvious $\Leftrightarrow$, $\leftrightsquigarrow$, $\twoheadleftarrow\!\!\twoheadrightarrow$, and combinations thereof) or antisymmetric.

5.3.4 Lemma *Let $\mathcal{X}$ be a set, and $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ and $\mathcal{Q} = (Q_\eta \colon \eta \in \mathrm{H})$ two models on the measurable space $(\mathcal{X}, 2^{\mathcal{X}})$. Let furthermore $\kappa \colon \Theta \to \overline{\mathbb{R}}$ and $\lambda \colon \mathrm{H} \to \overline{\mathbb{R}}$ be parameters of interest for $\mathcal{P}$ and $\mathcal{Q}$, respectively, and let $\beta \in [0, 1]$.*

(i) *If to every $\eta \in \mathrm{H}$ corresponds some $\vartheta \in \Theta$ with $Q_\eta = P_\vartheta$ and $\lambda(\eta) \geq \kappa(\vartheta)$, then $(\mathcal{P}, \kappa) \Rightarrow (\mathcal{Q}, \lambda)$.*

(ii) *If to every $\eta \in \mathrm{H}$ corresponds some $\vartheta \in \Theta$ with $Q_\eta = P_\vartheta$ and $\lambda(\eta) \leq \kappa(\vartheta)$, then $(\mathcal{P}, \kappa) \rightsquigarrow (\mathcal{Q}, \lambda)$.*

(iii) *If $(\mathcal{P}, \kappa) \leftrightsquigarrow (\mathcal{Q}, \lambda)$, then also $(\mathcal{P}, \kappa) \twoheadleftarrow\!\!\twoheadrightarrow (\mathcal{Q}, \lambda)$.*

Proof. (i) Let L be a lower confidence bound for κ in $\mathcal{P}$ with level β. Let $\eta \in \mathrm{H}$ and pick $\vartheta \in \Theta$ with $Q_\eta = P_\vartheta$ and $\lambda(\eta) \geq \kappa(\vartheta)$. Then $Q_\eta(L \leq \lambda(\eta)) = P_\vartheta(L \leq \lambda(\eta)) \geq P_\vartheta(L \leq \kappa(\vartheta)) \geq \beta$. L is thus a lower confidence bound for λ in $\mathcal{Q}$ with level β.

(ii) Let L and L' be two lower confidence bounds for both κ in $\mathcal{P}$ and λ in $\mathcal{Q}$ with levels β, and assume L' better than L with respect to κ in $\mathcal{P}$. Let $\eta \in \mathrm{H}$, $t \in \,]-\infty, \lambda(\eta)[$ and pick $\vartheta \in \Theta$ with $Q_\eta = P_\vartheta$ and $\lambda(\eta) \leq \kappa(\vartheta)$. Then $t \in \,]-\infty, \kappa(\vartheta)[$ also, hence $Q_\eta(L' \geq t) = P_\vartheta(L' \geq t) \geq P_\vartheta(L \geq t) = Q_\eta(L \geq t)$. L' is thus also better than L with respect to λ in $\mathcal{Q}$.

(iii) For ease of expression, let us denote by C the set of lower confidence bounds with level β for both κ in $\mathcal{P}$ and λ in $\mathcal{Q}$. Let $L \in \mathsf{C}$ be admissible with respect to κ in $\mathcal{P}$, and let $L' \in \mathsf{C}$ be better than L with respect to λ in $\mathcal{Q}$. Due to $(\mathcal{P}, \kappa) \leftsquigarrow (\mathcal{Q}, \lambda)$, L' is better than L with respect to κ in $\mathcal{P}$. Since L is admissible in C with respect to κ in $\mathcal{P}$, L is better than L' with respect to κ in $\mathcal{P}$. Due to $(\mathcal{P}, \kappa) \rightsquigarrow (\mathcal{Q}, \lambda)$, L is better than L' with respect to $(\mathcal{Q}, \lambda)$. This yields the admissibility of L in C with respect to λ in $\mathcal{Q}$. The remaining part of the claim follows by symmetry. □

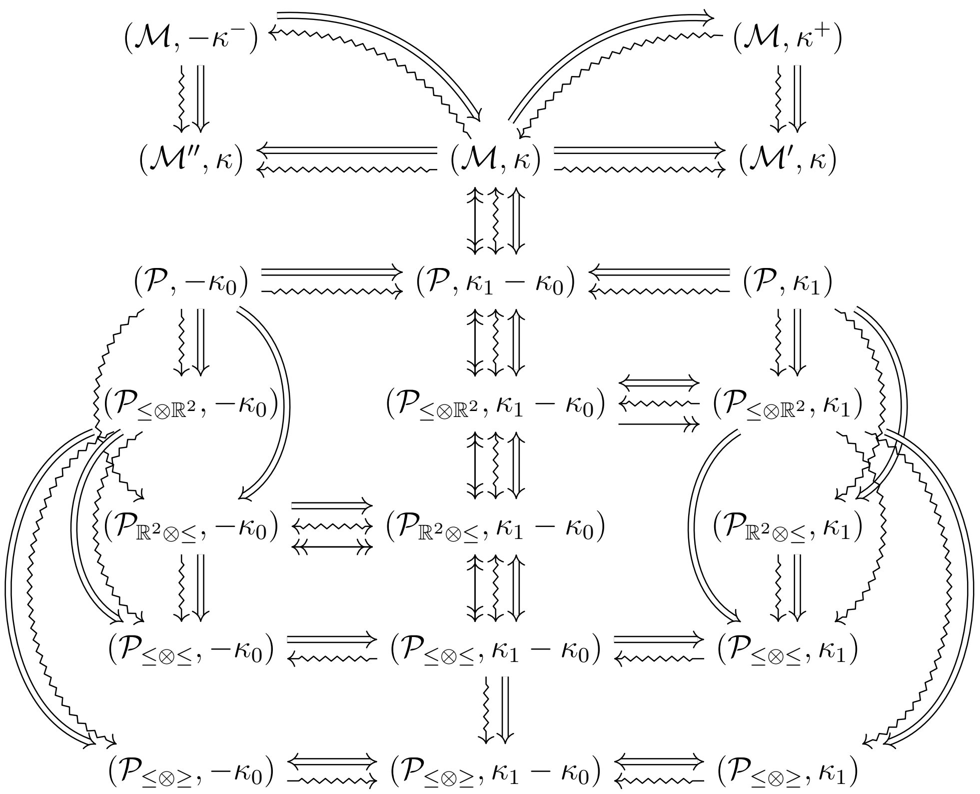
5.3.5 Remark 1. Lemma 5.3.4 is a restatement of the fundamental Lemma 4.1 from Mattner and Mattner (2013).

2. Lemma 5.3.4, parts (i) and (ii), implies in particular: If to every $(\vartheta_1, \eta_2) \in \Theta \times \mathrm{H}$ corresponds some $(\eta_1, \vartheta_2) \in \mathrm{H} \times \Theta$ with $P_{\vartheta_i} = Q_{\eta_i}$ and $\kappa(\vartheta_i) \leq \lambda(\eta_i)$ for $i \in \{1, 2\}$, then $(P, \kappa) \rightleftarrows (\mathcal{Q}, \lambda)$. These assumptions are trivially fulfilled if $\mathcal{P} = \mathcal{Q}$ and $\kappa \leq \lambda$.
3. The following theorem, relating different submodels and parameters of interest using the notions introduced in Definition 5.3.2, constitutes the main result of this chapter.

5.3.6 Theorem (Relations between models of diagnostic tests) *Let*

- $n \in \mathbb{N}$,
- $\kappa \colon \mathrm{prob}(\{0,1\}^2) \to [-1, 1]$, $q \mapsto q_{0,1} - q_{1,0}$,
- $\kappa_i \colon \Theta \to [-1, 1]$, $(\pi, \chi) \mapsto \pi_i(\chi''_{i|i} - \chi'_{i|i})$ *for* $i \in \{0, 1\}$,
- $\mathcal{M} := (\mathrm{M}_{n,q} \colon q \in \mathrm{prob}(\{0,1\}^2))$,
- $\mathcal{M}' := (\mathrm{M}_{n,q} \colon q \in \{\kappa \geq 0\})$,
- $\mathcal{M}'' := (\mathrm{M}_{n,q} \colon q \in \{\kappa \leq 0\})$.

Then the following diagram holds:



Proof. The proof uses Lemma 5.3.4, Remark 5.3.5, part 2, and the following results from the next section: Lemmas 5.4.6, 5.4.7, 5.4.12, 5.4.13, and Remark 5.4.5, part 1.

Let us first make the following observation: Given a parameter of interest $\lambda \colon \Theta \to \overline{\mathbb{R}}$ for $\mathcal{P}$ and binary relations R, S on $\mathbb{R}^2$ with $S \supseteq R$, Lemma 5.3.4, parts (i) and (ii), implies $(\mathcal{P}_S, \lambda) \rightrightarrows (\mathcal{P}_R, \lambda)$.

- $(\mathcal{M}, \kappa) \rightrightarrows (\mathcal{P}, \kappa_1 - \kappa_0)$: Let $(\pi, \chi) \in \Theta$, and let us define $q := \mu(\pi, \chi) \in \mathrm{prob}(\{0,1\}^2)$. Lemma 5.3.4, parts (i) and (ii), and Remark 5.4.5, part 1, then yield the assertion.

- $(\mathcal{M}, \kappa) \leftleftarrows (\mathcal{P}, \kappa_1 - \kappa_0)$: Let $q \in \mathrm{prob}(\{0,1\}^2)$, and let us pick, using Lemma 5.4.7, $(\pi, \chi) \in \mu^{-1}[\{q\}]$. Remark 5.4.5, part 1, and Lemma 5.3.4, parts (i) and (ii), then part (iii), yield the assertion.

- $(\mathcal{P}, \kappa_1 - \kappa_0) \overset{\Leftrightarrow}{\underset{\twoheadleftarrow\!\twoheadrightarrow}{\leftrightsquigarrow}} (\mathcal{P}_{\leq\otimes\mathbb{R}^2}, \kappa_1 - \kappa_0)$ follows from $\Theta_{\leq\otimes\mathbb{R}^2} \supseteq \Theta_{\leq\otimes\leq}$, Lemma 5.4.6 and a double application of Lemma 5.3.4, parts (i) and (ii), then part (iii).

- $(\mathcal{P}_{\leq\otimes\mathbb{R}^2}, \kappa_1 - \kappa_0) \overset{\Leftrightarrow}{\underset{\twoheadleftarrow\!\twoheadrightarrow}{\leftrightsquigarrow}} (\mathcal{P}_{\mathbb{R}^2\otimes\leq}, \kappa_1 - \kappa_0)$ follows from $\Theta_{\leq\otimes\mathbb{R}^2} \cap \Theta_{\mathbb{R}^2\otimes\leq} \supseteq \Theta_{\leq\otimes\leq}$ and a double application of Lemmas 5.4.6 and 5.3.4, parts (i) and (ii), then part (iii).

- $(\mathcal{P}_{\mathbb{R}^2\otimes\leq}, \kappa_1 - \kappa_0) \overset{\Leftrightarrow}{\underset{\twoheadleftarrow\!\twoheadrightarrow}{\leftrightsquigarrow}} (\mathcal{P}_{\leq\otimes\leq}, \kappa_1 - \kappa_0)$ follows from $\Theta_{\mathbb{R}^2\otimes\leq} \supseteq \Theta_{\leq\otimes\leq}$ and a double application of Lemmas 5.4.6 and 5.3.4, parts (i) and (ii), then part (iii).

- $(\mathcal{P}_{\leq\otimes\leq}, \kappa_1 - \kappa_0) \overset{\Rightarrow}{\rightsquigarrow} (\mathcal{P}_{\leq\otimes\geq}, \kappa_1 - \kappa_0)$ follows from Lemmas 5.3.4, parts (i) and (ii), and 5.4.6.

- $(\mathcal{P}_{\leq\otimes\mathbb{R}^2}, \kappa_1 - \kappa_0) \overset{\Leftrightarrow}{\underset{\twoheadrightarrow}{\leftsquigarrow}} (\mathcal{P}_{\leq\otimes\mathbb{R}^2}, \kappa_1)$ is shown in Mattner and Mattner (2013, Theorem 1.1, parts A–C).

- $(\mathcal{P}, \lambda) \overset{\Rightarrow}{\rightsquigarrow} (\mathcal{P}_{\leq\otimes\mathbb{R}^2}, \lambda)$ for $\lambda \in \{\kappa_1, -\kappa_0\}$ follows from $\Theta \supseteq \Theta_{\leq\otimes\mathbb{R}^2}$ and the observation at the beginning of the proof.

- $(\mathcal{P}, \lambda) \overset{\Rightarrow}{\rightsquigarrow} (\mathcal{P}_{\mathbb{R}^2\otimes\leq}, \lambda)$ for $\lambda \in \{\kappa_1, -\kappa_0\}$ follows from $\Theta \supseteq \Theta_{\mathbb{R}^2\otimes\leq}$ and the observation at the beginning of the proof.

- $(\mathcal{P}_{\leq\otimes\mathbb{R}^2}, \lambda) \overset{\Rightarrow}{\rightsquigarrow} (\mathcal{P}_{\leq\otimes\leq}, \lambda)$ for $\lambda \in \{\kappa_1, -\kappa_0\}$ follows from $\Theta_{\leq\otimes\mathbb{R}^2} \supseteq \Theta_{\leq\otimes\leq}$ and the observation at the beginning of the proof.

- $(\mathcal{P}_{\leq\otimes\mathbb{R}^2}, \lambda) \overset{\Rightarrow}{\rightsquigarrow} (\mathcal{P}_{\leq\otimes\geq}, \lambda)$ for $\lambda \in \{\kappa_1, -\kappa_0\}$ follows from $\Theta_{\leq\otimes\mathbb{R}^2} \supseteq \Theta_{\leq\otimes\geq}$ and the observation at the beginning of the proof.

- $(\mathcal{P}_{\mathbb{R}^2\otimes\leq}, \lambda) \overset{\Rightarrow}{\rightsquigarrow} (\mathcal{P}_{\leq\otimes\leq}, \lambda)$ for $\lambda \in \{\kappa_1, -\kappa_0\}$ follows from $\Theta_{\mathbb{R}^2\otimes\leq} \supseteq \Theta_{\leq\otimes\leq}$ and the observation at the beginning of the proof.

- $(\mathcal{P}_{\mathbb{R}^2\otimes\leq}, -\kappa_0) \overset{\Rightarrow}{\leftsquigarrow} (\mathcal{P}_{\mathbb{R}^2\otimes\leq}, \kappa_1 - \kappa_0)$ follows from $-\kappa_0 \leq \kappa_1 - \kappa_0$ on $\Theta_{\mathbb{R}^2\otimes\leq}$ and Remark 5.3.5, part 2.

- $(\mathcal{P}_{\mathbb{R}^2\otimes\leq}, -\kappa_0) \overset{\rightsquigarrow}{\twoheadleftarrow\!\twoheadrightarrow} (\mathcal{P}_{\mathbb{R}^2\otimes\leq}, \kappa_1 - \kappa_0)$: Let $(\pi, \chi) \in \Theta_{\mathbb{R}^2\otimes\leq}$. If we set $q := \mu(\pi, \chi)$, $\mathrm{Pr} := 0$, $\mathrm{Se}_1, \mathrm{Se}_2 \in [0,1]$ arbitrary with $\mathrm{Se}_1 \leq \mathrm{Se}_2$, and $\mathrm{Sp}_1 := q_{0+}$ and $\mathrm{Sp}_2 := q_{+0}$, then equations (27)–(30) of Lemma 2.3 from Mattner and Mattner (2013) are satisfied, yielding $(\tilde{\pi}, \tilde{\chi}) \in \Theta_{\mathbb{R}^2\otimes\leq}$ with $\mu(\tilde{\pi}, \tilde{\chi}) = q = \mu(\pi, \chi)$ and $\tilde{\pi} = (1, 0)$. This implies $\kappa_1(\tilde{\pi}, \tilde{\chi}) = 0$, yielding $(\kappa_1 - \kappa_0)(\pi, \chi) = \kappa \circ \mu(\pi, \chi) = \kappa \circ \mu(\tilde{\pi}, \tilde{\chi}) = (\kappa_1 - \kappa_0)(\tilde{\pi}, \tilde{\chi}) = -\kappa_0(\tilde{\pi}, \tilde{\chi})$. Lemma 5.3.4, part (ii), then part (iii) and the previous point, yields the claim.

- $(\mathcal{P}_{\leq\otimes\leq}, \kappa_1 - \kappa_0) \overset{\Rightarrow}{\leftsquigarrow} (\mathcal{P}_{\leq\otimes\leq}, \kappa_1)$ follows from $\kappa_1 - \kappa_0 \leq \kappa_1$ on $\Theta_{\leq\otimes\leq}$ and Remark 5.3.5, part 2.

- $(\mathcal{P}_{\leq\otimes\leq}, -\kappa_0) \overset{\Rightarrow}{\leftsquigarrow} (\mathcal{P}_{\leq\otimes\leq}, \kappa_1 - \kappa_0)$ follows from $-\kappa_0 \leq \kappa_1 - \kappa_0$ on $\Theta_{\leq\otimes\leq}$ and Remark 5.3.5, part 2.

- $(\mathcal{P}_{\leq\otimes\geq}, \kappa_1 - \kappa_0) \overset{\Rightarrow}{\leftsquigarrow} (\mathcal{P}_{\leq\otimes\geq}, \kappa_1)$ follows from $\kappa_1 - \kappa_0 \leq \kappa_1$ on $\Theta_{\leq\otimes\geq}$ and Remark 5.3.5, part 2.

- $(\mathcal{P}_{\leq\otimes\geq}, \kappa_1 - \kappa_0) \Leftarrow (\mathcal{P}_{\leq\otimes\geq}, \kappa_1)$: Let $(\pi, \chi) \in \Theta_{\leq\otimes\geq}$. Lemma 5.4.12, part (b), with $q := \mu(\pi, \chi)$ and $\Pr := 1$ then yields $(\tilde{\pi}, \tilde{\chi}) \in \Theta_{\leq\otimes\geq}$ with $\mu(\tilde{\pi}, \tilde{\chi}) = q = \mu(\pi, \chi)$ and $\tilde{\pi} = (0, 1)$. This implies $\kappa_0(\tilde{\pi}, \tilde{\chi}) = 0$, yielding $(\kappa_1 - \kappa_0)(\pi, \chi) = \kappa \circ \mu(\pi, \chi) = \kappa \circ \mu(\tilde{\pi}, \tilde{\chi}) = (\kappa_1 - \kappa_0)(\tilde{\pi}, \tilde{\chi}) = \kappa_1(\tilde{\pi}, \tilde{\chi})$. Lemma 5.3.4, part (i), yields the claim.

- $(\mathcal{P}_{\leq\otimes\geq}, -\kappa_0) \substack{\Leftarrow \\ \rightsquigarrow} (\mathcal{P}_{\leq\otimes\geq}, \kappa_1 - \kappa_0)$ follows from $\kappa_1 - \kappa_0 \leq -\kappa_0$ on $\Theta_{\leq\otimes\geq}$ and Remark 5.3.5, part 2.

- $(\mathcal{P}_{\leq\otimes\geq}, -\kappa_0) \Longrightarrow (\mathcal{P}_{\leq\otimes\geq}, \kappa_1 - \kappa_0)$: Let $(\pi, \chi) \in \Theta_{\leq\otimes\geq}$. Lemma 5.4.13, part (b), with $q := \mu(\pi, \chi)$ and $\Pr := 0$ then yields $(\tilde{\pi}, \tilde{\chi}) \in \Theta_{\leq\otimes\geq}$ with $\mu(\tilde{\pi}, \tilde{\chi}) = q = \mu(\pi, \chi)$ and $\tilde{\pi} = (1, 0)$. This implies $\kappa_1(\tilde{\pi}, \tilde{\chi}) = 0$, yielding $(\kappa_1 - \kappa_0)(\pi, \chi) = \kappa \circ \mu(\pi, \chi) = \kappa \circ \mu(\tilde{\pi}, \tilde{\chi}) = (\kappa_1 - \kappa_0)(\tilde{\pi}, \tilde{\chi}) = -\kappa_0(\tilde{\pi}, \tilde{\chi})$. Lemma 5.3.4, part (i), yields the claim.

- $(\mathcal{P}, -\kappa_0) \substack{\Leftarrow \\ \leftsquigarrow} (\mathcal{P}, \kappa_1 - \kappa_0) \substack{\Rightarrow \\ \rightsquigarrow} (\mathcal{P}, \kappa_1)$ follows from the transitivity of $\substack{\Rightarrow \\ \rightsquigarrow}$.

- $(\mathcal{M}, \kappa) \substack{\Rightarrow \\ \rightsquigarrow} (\mathcal{M}', \kappa)$ follows from the observation at the beginning of the proof.

- $(\mathcal{M}, \kappa) \substack{\Rightarrow \\ \rightsquigarrow} (\mathcal{M}'', \kappa)$ follows from the observation at the beginning of the proof.

- $(\mathcal{M}, \kappa) \substack{\Rightarrow \\ \leftsquigarrow} (\mathcal{M}, \kappa^+)$ follows from $\kappa \leq \kappa^+$ and Remark 5.3.5, part 2.

- $(\mathcal{M}, \kappa) \substack{\Leftarrow \\ \rightsquigarrow} (\mathcal{M}, -\kappa^-)$ follows from $-\kappa^- \leq \kappa$ and Remark 5.3.5, part 2.

- $(\mathcal{M}'', \kappa) \substack{\Leftarrow \\ \leftsquigarrow} (\mathcal{M}, -\kappa^-)$ follows from $-\kappa^- = \kappa$ on $\{\kappa \leq 0\}$ and the observation at the beginning of the proof.

- $(\mathcal{M}', \kappa) \substack{\Leftarrow \\ \leftsquigarrow} (\mathcal{M}, \kappa^+)$ follows from $\kappa^+ = \kappa$ on $\{\kappa \geq 0\}$ and the observation at the beginning of the proof. □

5.3.7 Remark 1. The relation $(\mathcal{P}_{\leq\otimes\mathbb{R}^2}, \kappa_1) \Leftarrow (\mathcal{P}_{\leq\otimes\leq}, \kappa_1)$ does not hold. In fact, if $L \equiv 0$, then L trivially is a lower confidence bound for κ_1 in $\mathcal{P}_{\leq\otimes\leq}$, but not a lower confidence bound for κ_1 in $\mathcal{P}_{\leq\otimes\mathbb{R}^2}$ since $\mathrm{M}_{n,\mu(\pi,\chi)}(L \leq \kappa_1(\pi, \chi)) = \mathrm{M}_{n,\mu(\pi,\chi)}(\emptyset) = 0$ for $(\pi, \chi) \in \Theta_{\leq\otimes>}$ obtained with Lemma 5.4.12, part (b), with $\Pr := 1$ and arbitrary $\Delta\mathrm{Se} < 0$.

2. Theorem 5.3.6 allows reducing certain inference problems for non-identifiable parameters of interest (e.g., κ_1 in $\mathcal{P}_{\leq\otimes\leq}$) to corresponding ones for identifiable parameters of interest (e.g., κ in $\mathcal{M}$ in view of Remark 5.2.5).

5.4 AUXILIARY RESULTS

5.4.1 Remark We establish in the following first some results on isomorphisms that prove useful for extending assertions regarding one submodel $\mathcal{P}_R$ to another one $\mathcal{P}_{R'}$. The lemmas that follow are often incomplete and proved "by hand" (instead of a more systematic approach using real algebraic geometry or the theory on functional equations as presented in, e.g., Aczél (1966), Aczél (1984), or Aczél and Dhombres (1989)). They are needed for the proof of Theorem 5.3.6.

5.4.2 Lemma *Let $f_1 \colon \mathcal{X} \to \mathcal{X}$, $f_2, g_1 \colon \mathcal{X} \to \mathcal{Y}$, and $g_2 \colon \mathcal{Y} \to \mathcal{Y}$ be four functions. Then the following implication holds:*

$$f_1[\cdot] \circ f_2^{-1}[\cdot] = g_1^{-1}[\cdot] \circ g_2[\cdot] \quad \Longrightarrow \quad g_1 \circ f_1 = g_2 \circ f_2.$$

The converse implication holds if f_i is surjective and g_{3-i} is injective for some $i \in \{1,2\}$.

Proof. Let us assume $f_1[\cdot] \circ f_2^{-1}[\cdot] = g_1^{-1}[\cdot] \circ g_2[\cdot]$ and let $x \in \mathcal{X}$. Then $x \in f_2^{-1}[\{f_2(x)\}]$ and hence $f_1(x) \in f_1[f_2^{-1}[\{f_2(x)\}]] = g_1^{-1}[g_2[\{f_2(x)\}]] = g_1^{-1}[\{g_2 \circ f_2(x)\}]$, which implies $g_1 \circ f_1(x) = g_2 \circ f_2(x)$ since $h[h^{-1}[\{y\}]] \in \{\emptyset, \{y\}\}$ for any function h.

Let us now assume that $g_1 \circ f_1 = g_2 \circ f_2$ and let $A \in 2^{\mathcal{Y}}$. The inclusion $f_1[f_2^{-1}[A]] \subseteq g_1^{-1}[g_2[A]]$ holds without additional assumptions since $x \in f_2^{-1}[A]$ implies $g_1 \circ f_1(x) = g_2 \circ f_2(x) \in g_2[A]$. Let now $i \in \{1,2\}$ and suppose f_i is surjective and g_{3-i} is injective. Applying $f_i^{-1}[\cdot]$ from the right and $g_{3-i}^{-1}[\cdot]$ from the left to $g_1 \circ f_1 = g_2 \circ f_2$ yields

$$g_{3-i}^{-1}[\cdot] \circ g_i[\cdot] \circ f_i[\cdot] \circ f_i^{-1}[\cdot] = g_{3-i}^{-1}[\cdot] \circ g_{3-i}[\cdot] \circ f_{3-i}[\cdot] \circ f_i^{-1}[\cdot],$$

i.e., $f_{3-i}[\cdot] \circ f_i^{-1}[\cdot] = g_{3-i}^{-1}[\cdot] \circ g_i[\cdot]$ or, equivalently, $f_i[\cdot] \circ f_{3-i}^{-1}[\cdot] = g_i^{-1}[\cdot] \circ g_{3-i}[\cdot]$ due to Remark A.1.15. □

5.4.3 Definition A function $f\colon \mathcal{X} \to \mathcal{Y}$ **factorizes over** a function $g\colon \mathcal{X} \to \mathcal{Z}$ if there is a function $h\colon \mathcal{Z} \to \mathcal{Y}$ such that $f = h \circ g$.

5.4.4 Remark

1. f factorizes over g if, and only if, to every $z \in \mathcal{Z}$ corresponds some $y \in \mathcal{Y}$ with $g^{-1}[\{z\}] \subseteq f^{-1}[\{y\}]$. (This condition implies that f is constant wherever g is constant.) In fact, if $f = h \circ g$, $z \in Z$, and $y := h(z)$, then, for $\xi \in g^{-1}[\{z\}]$, we have $f(\xi) = h \circ g(\xi) = h(z) = y$. Conversely, for every $z \in \mathcal{Z}$, let $h(z)$ stand for some $y \in \mathcal{Y}$ with $g^{-1}[\{z\}] \subseteq f^{-1}[\{y\}]$. Then, for $x \in \mathcal{X}$, we have $x \in g^{-1}[\{g(x)\}] \subseteq f^{-1}[\{h(g(x))\}]$, which yields $f(x) = h(g(x))$.
2. One could analogously define: A function $f\colon \mathcal{X} \to \mathcal{Y}$ *factorizes under* a function $g\colon \mathcal{Z} \to \mathcal{Y}$ if there is a function $h\colon \mathcal{X} \to \mathcal{Z}$ such that $f = g \circ h$.
3. Let $\Phi\colon \Theta \to \Theta$ and $\psi\colon \mathrm{prob}(\{0,1\}^2) \to \mathrm{prob}(\{0,1\}^2)$ be bijective. By Lemma 5.4.2 (applied to $f_1 := \Phi$, $f_2 := g_1 := \mu$, and $g_2 := \psi$), we have the equivalence

$$\Phi \circ \mu^{-1} = \mu^{-1} \circ \psi \quad\iff\quad \mu \circ \Phi \text{ factorizes over } \mu,$$

in which case $\psi[\cdot] = \mu[\cdot] \circ \Phi[\cdot] \circ \mu^{-1}[\cdot]$ by Remark A.1.15 since μ is surjective.
4. Further below, we are interested in sets of the form $T[\mu^{-1}[\{q\}] \cap \Theta_R]$ for different maps T defined on Θ.
5. Let $s\colon \{1,2\} \to \{1,2\}$, $k \mapsto 3-k$, and let us consider the involutions φ_j on $\mathrm{mark}(\{0,1\},\{0,1\}^2)$, ψ_j on $\mathrm{prob}(\{0,1\}^2)$, and Φ_j on Θ given by

$$\begin{aligned} \varphi_1(\chi)_{\iota|i} &:= \chi_{\mathbf{1}-\iota|1-i}, & \varphi_2(\chi)_{\iota|i} &:= \chi_{\iota\circ s|i}, \\ \psi_1(q)_\iota &:= q_{\mathbf{1}-\iota}, & \psi_2(q)_\iota &:= q_{\iota\circ s}, \\ \Phi_1(\pi,\chi) &:= (\mathbf{1}-\pi, \varphi_1(\chi)), & \Phi_2(\pi,\chi) &:= (\pi, \varphi_2(\chi)) \end{aligned}$$

for $i \in \{0,1\}$, $\iota \in \{0,1\}^2$, $(\pi,\chi) \in \Theta$, $q \in \mathrm{prob}(\{0,1\}^2)$, and $\mathbf{1} := (1,1)$. Let furthermore

$$\sigma\colon \bigcup_{n\in\mathbb{N}} \mathbb{R}^n \to \bigcup_{n\in\mathbb{N}} \mathbb{R}^n, \quad (x_1,\ldots,x_n) \mapsto (1-x_1, x_2, \ldots, x_n) \quad \text{for } n \in \mathbb{N}.$$

Due to

$$(\varphi_1(\chi)'_{\iota|i}, \varphi_1(\chi)''_{\iota|i}) = (\chi'_{1-\iota|1-i}, \chi''_{1-\iota|1-i}) \quad \text{and} \quad (\varphi_2(\chi)', \varphi_2(\chi)'') = (\chi'', \chi')$$

for $\iota, i \in \{0,1\}$, we obtain

$$\Phi_1[\Theta_{S\otimes R}] = \Theta_{R\otimes S} = \Phi_2[\Theta_{R^{\mathrm{op}}\otimes S^{\mathrm{op}}}]$$

for binary relations R, S on $\mathbb{R}$. From $\mu^{-1}[\{q\}] = \Phi_j[\mu^{-1}[\{\psi_j(q)\}]]$ and the bijectivity of Φ_j for $j \in \{1,2\}$ it follows that

$$\begin{aligned}
T[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\leq}] &\subseteq T[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\mathbb{R}^2}] \cap T[\Phi_1[\mu^{-1}[\{\psi_1(q)\}] \cap \Theta_{\leq\otimes\mathbb{R}^2}]], \\
T[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\geq}] &\subseteq T[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\mathbb{R}^2}] \\
&\qquad \cap T[\Phi_1[\Phi_2[\mu^{-1}[\{\psi_1 \circ \psi_2(q)\}] \cap \Theta_{\leq\otimes\mathbb{R}^2}]]], \\
T[\mu^{-1}[\{q\}] \cap \Theta_{\mathbb{R}^2\otimes\leq}] &= T[\Phi_1[\mu^{-1}[\{\psi_1(q)\}] \cap \Theta_{\leq\otimes\mathbb{R}^2}]]
\end{aligned}$$

for maps T defined on Θ.

5.4.5 Remark 1. Let us note that $\kappa_1 - \kappa_0 = \kappa \circ \mu$. In fact, for $(\pi, \chi) \in \Theta$, we have

$$\begin{aligned}
(\kappa_1 - \kappa_0)(\pi, \chi) &= \pi_1(\chi''_{1|1} - \chi'_{1|1}) - \pi_0(\chi''_{0|0} - \chi'_{0|0}) \\
&= \pi_1(\chi_{0,1|1} - \chi_{1,0|1}) - \pi_0(\chi_{1,0|0} - \chi_{0,1|0}) \\
&= \mu(\pi, \chi)_{0,1} - \mu(\pi, \chi)_{1,0}.
\end{aligned}$$

2. The next result unifies certain steps in the proof of Theorem 5.3.6.

5.4.6 Lemma *For every $(\pi, \chi) \in \Theta$ there exists $(\tilde{\pi}, \tilde{\chi}) \in \Theta_{\leq\otimes\leq}$ with $\mu(\tilde{\pi}, \tilde{\chi}) = \mu(\pi, \chi)$ and $(\kappa_1 - \kappa_0)(\tilde{\pi}, \tilde{\chi}) = (\kappa_1 - \kappa_0)(\pi, \chi)$.*

Proof. Let $\tilde{\pi}_0 := \mathbf{1}(\kappa(\mu(\pi, \chi)) \leq 0)$ and $\tilde{\pi}_1 := 1 - \tilde{\pi}_0$, as well as

$$(\tilde{\chi}_{j|\tilde{\pi}(0)}, \tilde{\chi}_{j|\tilde{\pi}(1)}) := (\mathbf{1}(j = \mathbf{1} - \tilde{\pi}), \mu(\pi, \chi)_j) \qquad \text{for } j \in \{0,1\}^2.$$

Then $(\tilde{\pi}, \tilde{\chi}) \in \Theta_{\leq\otimes\leq}$ and $\mu(\tilde{\pi}, \tilde{\chi})_j = \tilde{\chi}_{j|\tilde{\pi}(1)} = \mu(\pi, \chi)_j$ for $j \in \{0,1\}^2$. The rest now follows from Remark 5.4.5, part 1. □

5.4.7 Lemma *Let $q \in \mathrm{prob}(\{0,1\}^2)$. Then*

$$\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\mathbb{R}^2} \neq \emptyset, \tag{46}$$
$$\mu^{-1}[\{q\}] \cap \Theta_{\mathbb{R}^2\otimes\leq} \neq \emptyset, \tag{47}$$
$$\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\leq} \neq \emptyset, \tag{48}$$
$$\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\geq} \neq \emptyset \quad \iff \quad q_{0,1} \leq q_{1,0}. \tag{49}$$

Proof. (48) follows from Lemma 5.4.6, while (46) and (47) follow from (48) and $\Theta_{\leq\otimes\leq} \subseteq \Theta_{\mathbb{R}^2\otimes\leq} \cap \Theta_{\leq\otimes\mathbb{R}^2}$. As to (49), if $(\pi, \chi) \in \mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\geq}$, then

$$q_{0,1} = \pi_0\chi_{0,1|0} + \pi_1\chi_{0,1|1} \leq \pi_0\chi_{1,0|0} + \pi_1\chi_{1,0|1} = q_{1,0}$$

since $\chi_{+0|0} \geq \chi_{0+|0}$ and $\chi_{+1|1} \leq \chi_{1+|1}$. If $q_{0,1} \leq q_{1,0}$, then $\pi_i := 1/2$ and $\chi_{\iota|i} := q_\iota$ for $i \in \{0,1\}$ and $\iota \in \{0,1\}^2$ satisfy $(\pi, \chi) \in \Theta_{\leq\otimes\geq}$ and $\mu(\pi, \chi) = q$. □

5.4.8 Remark The following results are analogues to Lemmas 2.3–2.16 from Mattner and Mattner (2013). Not every one of them is needed for the proof of Theorem 5.3.6, but each one can be of interest or might prove useful for possible future extensions of that theorem.

5.4.9 Lemma (Analogue to Lemma 2.3 from Mattner and Mattner, 2013) *Let* $q \in \mathrm{prob}(\{0,1\}^2)$,

$$T_A\colon \Theta \to \mathbb{R}^5, \quad (\pi,\chi) \mapsto (\pi_1, \chi'_{0|0}, \chi'_{1|1}, \chi''_{0|0}, \chi''_{1|1}),$$

and $A_R \coloneqq T_A[\mu^{-1}[\{q\}] \cap \Theta_R]$ *for binary relations* R *on* $\mathbb{R}^2$. *Then*

(a) $A_{\leq\otimes\leq} = \{(\mathrm{Pr}, \mathrm{Sp}_1, \mathrm{Se}_1, \mathrm{Sp}_2, \mathrm{Se}_2) \in [0,1]^5 \colon \mathrm{Sp}_1 \leq \mathrm{Sp}_2,\ \mathrm{Se}_1 \leq \mathrm{Se}_2,\ (50)\text{–}(53)\} \eqqcolon A'$, *where*

$$(1-\mathrm{Pr})\mathrm{Sp}_1 + \mathrm{Pr}(1-\mathrm{Se}_1) = q_{0+} \tag{50}$$
$$(1-\mathrm{Pr})\mathrm{Sp}_2 + \mathrm{Pr}(1-\mathrm{Se}_2) = q_{+0} \tag{51}$$
$$(1-\mathrm{Pr})(\mathrm{Sp}_1 + \mathrm{Sp}_2 - 1)^+ + \mathrm{Pr}(1-\mathrm{Se}_1-\mathrm{Se}_2)^+ \leq q_{0,0} \tag{52}$$
$$(1-\mathrm{Pr})\mathrm{Sp}_1 + \mathrm{Pr}(1-\mathrm{Se}_2) \geq q_{0,0} \tag{53}$$

(b) $A_{\leq\otimes\geq} = \{(\mathrm{Pr}, \mathrm{Sp}_1, \mathrm{Se}_1, \mathrm{Sp}_2, \mathrm{Se}_2) \in [0,1]^5 \colon \mathrm{Sp}_1 \leq \mathrm{Sp}_2,\ \mathrm{Se}_1 \geq \mathrm{Se}_2,\ (50)\text{–}(52)\} \eqqcolon A''$.

(c) $A_{\mathbb{R}^2\otimes\leq} = \{(\mathrm{Pr}, \mathrm{Sp}_1, \mathrm{Se}_1, \mathrm{Sp}_2, \mathrm{Se}_2) \in [0,1]^5 \colon \mathrm{Se}_1 \leq \mathrm{Se}_2,\ (50),\ (51),\ (54),\ (55)\} \eqqcolon A'''$, *where*

$$(1-\mathrm{Pr})(1-\mathrm{Sp}_1-\mathrm{Sp}_2)^+ + \mathrm{Pr}(\mathrm{Se}_1+\mathrm{Se}_2-1)^+ \leq q_{1,1} \tag{54}$$
$$(1-\mathrm{Pr})(1-\mathrm{Sp}_1 \vee \mathrm{Sp}_2) + \mathrm{Pr}\,\mathrm{Se}_1 \geq q_{1,1} \tag{55}$$

Proof. By Remark 5.4.4 and in view of Lemma 2.3 from Mattner and Mattner (2013), it remains to show $A' \subseteq A_{\leq\otimes\leq}$ and $A'' \subseteq A_{\leq\otimes\geq}$. Let $a \coloneqq (\mathrm{Pr}, \mathrm{Sp}_1, \mathrm{Se}_1, \mathrm{Sp}_2, \mathrm{Se}_2) \in A'$ resp. A'' and set $\pi \coloneqq (1-\mathrm{Pr}, \mathrm{Pr})$. Since

$$f\colon [(\mathrm{Sp}_1+\mathrm{Sp}_2-1)^+, \mathrm{Sp}_1] \times [(1-\mathrm{Se}_1-\mathrm{Se}_2)^+, (1-\mathrm{Se}_1)\wedge(1-\mathrm{Se}_2)] \eqqcolon M \to \mathbb{R},$$
$$(x,y) \mapsto \pi_0 x + \pi_1 y,$$

is continuous, its domain connected, and $f(\min M) \leq q_{0,0} \leq f(\max M)$ by (52) and (53) resp. (52) and (50), depending on whether $a \in A'$ or A'', there exists $(\chi_{0,0|0}, \chi_{0,0|1}) \in M$ with $\pi_0\chi_{0,0|0} + \pi_1\chi_{0,0|1} = q_{0,0}$. With

$$\chi_{0,1|0} \coloneqq \mathrm{Sp}_1 - \chi_{0,0|0} \qquad \chi_{0,1|1} \coloneqq 1 - \mathrm{Se}_1 - \chi_{0,0|1}$$
$$\chi_{1,0|0} \coloneqq \mathrm{Sp}_2 - \chi_{0,0|0} \qquad \chi_{1,0|1} \coloneqq 1 - \mathrm{Se}_2 - \chi_{0,0|1}$$
$$\chi_{1,1|0} \coloneqq 1 - \mathrm{Sp}_1 - \mathrm{Sp}_2 + \chi_{0,0|0} \qquad \chi_{1,1|1} \coloneqq \mathrm{Se}_1 + \mathrm{Se}_2 - 1 + \chi_{0,0|1}$$

we obtain $(\pi,\chi) \in \Theta_{\leq\otimes\leq}$ resp. $\Theta_{\leq\otimes\geq}$, again depending on whether $a \in A'$ or A''. We furthermore have $\mu(\pi,\chi) = q$ and $F(\pi,\chi) = a$, i.e., $a \in A_{\leq\otimes\leq}$ resp. $A_{\leq\otimes\geq}$.

Part (c) follows from Lemma 2.3 from Mattner and Mattner (2013) and

$$A_{\mathbb{R}^2\otimes\leq} = \tau[T_A[\Phi_1[\mu^{-1}[\{q\}] \cap \Theta_{\mathbb{R}^2\otimes\leq}]]] = \tau[T_A[\mu^{-1}[\{\psi_1(q)\}] \cap \Theta_{\leq\otimes\mathbb{R}^2}]] = A'''$$

since $T_A = \tau \circ T_A \circ \Phi_1$ with $\tau\colon \mathbb{R}^5 \to \mathbb{R}^5$, $x \mapsto (1-x_1, x_3, x_2, x_5, x_4)$. □

5.4.10 Lemma (Analogue to Lemma 2.9 from Mattner and Mattner, 2013) *Let* $q \in \mathrm{prob}(\{0,1\}^2)$,

$$T_B\colon \Theta \to \mathbb{R}^3, \quad (\pi,\chi) \mapsto (\pi_1, \chi'_{1|1}, \chi''_{1|1}),$$

and $B_R \coloneqq T_B[\mu^{-1}[\{q\}] \cap \Theta_R]$ *for binary relations* R *on* $\mathbb{R}^2$. *Then*

(a) $B_{\leq\otimes\leq} = \{(\mathrm{Pr}, \mathrm{Se}_1, \mathrm{Se}_2) \in [0,1]^3 \colon \mathrm{Se}_1 \leq \mathrm{Se}_2, (56)\text{–}(60)\} =: B'$, *where*

$$\mathrm{Pr} - q_{1+} \leq \mathrm{Pr}(1 - \mathrm{Se}_1) \leq q_{0+} \tag{56}$$
$$\mathrm{Pr} - q_{+1} \leq \mathrm{Pr}(1 - \mathrm{Se}_2) \leq q_{+0} \tag{57}$$
$$-q_{0,0} \leq \mathrm{Pr}(\mathrm{Se}_1 + \mathrm{Se}_2 - 1) \leq q_{1,1} \tag{58}$$
$$\mathrm{Pr}(\mathrm{Se}_2 - \mathrm{Se}_1) \geq q_{0,1} - q_{1,0} \tag{59}$$
$$\mathrm{Pr}(\mathrm{Se}_2 - \mathrm{Se}_1) \leq q_{0,1} \tag{60}$$

(b) $B_{\leq\otimes\geq} = \{(\mathrm{Pr}, \mathrm{Se}_1, \mathrm{Se}_2) \in [0,1]^3 \colon \mathrm{Se}_1 \geq \mathrm{Se}_2, (56)\text{–}(59)\} =: B''$.

Proof. It remains to show $B' \subseteq B_{\leq\otimes\leq}$ and $B'' \subseteq B_{\leq\otimes\geq}$. Let $(\mathrm{Pr}, \mathrm{Se}_1, \mathrm{Se}_2) \in B'$ resp. B''. In view of (56) and (57) we can choose $(\mathrm{Sp}_1, \mathrm{Sp}_2) \in [0,1]^2$ satisfying (50) and (51), and furthermore $\mathrm{Sp}_1 \leq \mathrm{Sp}_2$ if $\mathrm{Pr} = 1$. By subtracting (51) from (50) we obtain using (59) $(1-\mathrm{Pr})(\mathrm{Sp}_1 - \mathrm{Sp}_2) + \mathrm{Pr}(\mathrm{Se}_2 - \mathrm{Se}_1) = q_{0,1} - q_{1,0} \leq \mathrm{Pr}(\mathrm{Se}_2 - \mathrm{Se}_1)$, yielding $\mathrm{Sp}_1 \leq \mathrm{Sp}_2$ if $\mathrm{Pr} < 1$. (52) follows using (58) by adding (50) and (51), while (53) is implied by subtracting (60) from (50). Lemma 5.4.9 yields $(\mathrm{Pr}, \mathrm{Sp}_1, \mathrm{Se}_1, \mathrm{Sp}_2, \mathrm{Se}_2) \in A_{\leq\otimes\leq}$ resp. $A_{\leq\otimes\geq}$ and thus $(\mathrm{Pr}, \mathrm{Se}_1, \mathrm{Se}_2) \in B_{\leq\otimes\leq}$ resp. $B_{\leq\otimes\geq}$. □

5.4.11 Lemma (Analogue to Lemma 2.9 from Mattner and Mattner, 2013) *Let* $q \in \mathrm{prob}(\{0,1\}^2)$,

$$T_{\tilde{B}} \colon \Theta \to \mathbb{R}^3, \quad (\pi, \chi) \mapsto (\pi_1, \chi'_{0|0}, \chi''_{0|0}),$$

and $\tilde{B}_R := T_{\tilde{B}}[\mu^{-1}[\{q\}] \cap \Theta_R]$ *for binary relations* R *on* $\mathbb{R}^2$. *Then*

(a) $\tilde{B}_{\leq\otimes\leq} = \{(\mathrm{Pr}, \mathrm{Sp}_1, \mathrm{Sp}_2) \in [0,1]^3 \colon \mathrm{Sp}_1 \leq \mathrm{Sp}_2, (61)\text{–}(65)\} =: \tilde{B}'$, *where*

$$q_{0+} - \mathrm{Pr} \leq (1-\mathrm{Pr})\mathrm{Sp}_1 \leq q_{0+} \tag{61}$$
$$q_{+0} - \mathrm{Pr} \leq (1-\mathrm{Pr})\mathrm{Sp}_2 \leq q_{+0} \tag{62}$$
$$-q_{1,1} \leq (1-\mathrm{Pr})(\mathrm{Sp}_1 + \mathrm{Sp}_2 - 1) \leq q_{0,0} \tag{63}$$
$$(1-\mathrm{Pr})(\mathrm{Sp}_2 - \mathrm{Sp}_1) \geq q_{1,0} - q_{0,1} \tag{64}$$
$$(1-\mathrm{Pr})(\mathrm{Sp}_2 - \mathrm{Sp}_1) \leq q_{1,0} \tag{65}$$

(b) $\tilde{B}_{\leq\otimes\geq} = \{(\mathrm{Pr}, \mathrm{Sp}_1, \mathrm{Sp}_2) \in [0,1]^3 \colon \mathrm{Sp}_1 \leq \mathrm{Sp}_2, (61)\text{–}(63), (64')\} =: \tilde{B}''$, *where*

$$(1-\mathrm{Pr})(\mathrm{Sp}_2 - \mathrm{Sp}_1) \leq q_{1,0} - q_{0,1} \tag{64'}$$

Proof. (a) Since $T_{\tilde{B}} = \sigma \circ T_B \circ \Phi_1$, we have $\tilde{B}_{\leq\otimes\leq} = \sigma[T_B[\Phi_1[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\leq}]]] = \sigma[T_B[\mu^{-1}[\{\psi_1(q)\}] \cap \Theta_{\leq\otimes\leq}]] = \tilde{B}'$ due to Lemma 5.4.10 (a).

(b) Let $\tau \colon \mathbb{R}^3 \to \mathbb{R}^3$, $(x_1, x_2, x_3) \mapsto (x_1, x_3, x_2)$. Since $T_{\tilde{B}} = \sigma \circ \tau \circ T_B \circ \Phi_1 \circ \Phi_2$, we have

$$\begin{aligned}\tilde{B}_{\leq\otimes\geq} &= \sigma[\tau[T_B[\Phi_1[\Phi_2[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\geq}]]]]] \\ &= \sigma[\tau[T_B[\mu^{-1}[\{\psi_1 \circ \psi_2(q)\}] \cap \Theta_{\leq\otimes\geq}]]] \\ &= \tilde{B}''\end{aligned}$$

due to Lemma 5.4.10 (b). □

5.4.12 Lemma (Analogue to Lemma 2.10 from Mattner and Mattner, 2013) *Let* $q \in \operatorname{prob}(\{0,1\}^2)$,

$$T_C\colon \Theta \to \mathbb{R}^2, \quad (\pi,\chi) \mapsto (\pi_1, \chi''_{1|1} - \chi'_{1|1}),$$

and $C_R := T_C[\mu^{-1}[\{q\}] \cap \Theta_R]$ *for binary relations* R *on* $\mathbb{R}^2$. *Then*

(a) $C_{\leq\otimes\leq} = \{(\mathrm{Pr}, \Delta\mathrm{Se}) \in [0,1]^2 \colon (66)\} =: C'$, *where*

$$q_{0,1} - q_{1,0} \leq \mathrm{Pr}\,\Delta\mathrm{Se} \leq q_{0,1} \wedge (q_{0,1} - q_{1,0} + 1 - \mathrm{Pr}) \tag{66}$$

(b) $C_{\leq\otimes\geq} = \{(\mathrm{Pr}, \Delta\mathrm{Se}) \in [0,1] \times [-1,0] \colon (66')\} =: C''$, *where*

$$q_{0,1} - q_{1,0} \leq \mathrm{Pr}\,\Delta\mathrm{Se} \leq q_{0,1} - q_{1,0} + 1 - \mathrm{Pr} \tag{66'}$$

Proof. It remains to show $C' \subseteq C_{\leq\otimes\leq}$ and $C'' \subseteq C_{\leq\otimes\geq}$. Let $(\mathrm{Pr}, \Delta\mathrm{Se}) \in C'$ resp. C''. If $\mathrm{Pr} = 0$, then $(\mathrm{Se}_1, \mathrm{Se}_2) := (0, \Delta\mathrm{Se})$ resp. $(\mathrm{Se}_1, \mathrm{Se}_2) := (-\Delta\mathrm{Se}, 0)$ fulfills (56)–(60). Lemma 5.4.10 yields $(\mathrm{Pr}, \mathrm{Se}_1, \mathrm{Se}_2) \in B_{\leq\otimes\leq}$ resp. $B_{\leq\otimes\geq}$ and thus $(\mathrm{Pr}, \Delta\mathrm{Se}) \in C_{\leq\otimes\leq}$ resp. $C_{\leq\otimes\geq}$.

If $\mathrm{Pr} > 0$, we can pick

$$\begin{aligned} \mathrm{Se}'_1 &\in [(1 - q_{0+}/\mathrm{Pr})^+, 1 \wedge (q_{1+}/\mathrm{Pr})] =: M_1 \\ \mathrm{Se}'_2 &\in [(1 - q_{+0}/\mathrm{Pr})^+, 1 \wedge (q_{+1}/\mathrm{Pr})] =: M_2 \end{aligned}$$

such that $\Delta\mathrm{Se} = \mathrm{Se}'_2 - \mathrm{Se}'_1$. Let us note that $M_1, M_2 \neq \emptyset$ since

$$\begin{aligned} 1 \wedge (q_{1+}/\mathrm{Pr}) - (1 - q_{0+}/\mathrm{Pr})^+ &= \min\{1/\mathrm{Pr} - 1, q_{0+}/\mathrm{Pr}, q_{1+}/\mathrm{Pr}, 1\} \geq 0 \\ 1 \wedge (q_{+1}/\mathrm{Pr}) - (1 - q_{+0}/\mathrm{Pr})^+ &= \min\{1/\mathrm{Pr} - 1, q_{+0}/\mathrm{Pr}, q_{+1}/\mathrm{Pr}, 1\} \geq 0. \end{aligned}$$

This is possible since

$$f\colon M_1 \times M_2 \to \mathbb{R}, \quad (x,y) \mapsto y - x,$$

is continuous, its domain connected, and

$$\begin{aligned} \mathrm{Pr}\, f(\min M_1, \max M_2) &= q_{+1} \wedge \mathrm{Pr} - (\mathrm{Pr} - q_{0+})^+ \\ &= \min\{q_{+1} + q_{0+} - \mathrm{Pr}, q_{+1}, q_{0+}, \mathrm{Pr}\} \\ &\geq \min\{q_{0,1} - q_{1,0} + 1 - \mathrm{Pr}, q_{0,1}, \mathrm{Pr}\,\Delta\mathrm{Se}\} \\ &= \mathrm{Pr}\,\Delta\mathrm{Se} \\ &= \max\{q_{0,1} - q_{1,0}, \mathrm{Pr}\,\Delta\mathrm{Se}\} \\ &\geq \max\{q_{0,1} - q_{1,0} + \mathrm{Pr} - 1, -q_{+0}, -q_{1+}, -\mathrm{Pr}\} \\ &= \max\{\mathrm{Pr} - q_{1+} - q_{+0}, -q_{+0}, -q_{1+}, -\mathrm{Pr}\} \\ &= (\mathrm{Pr} - q_{+0})^+ - q_{1+} \wedge \mathrm{Pr} \\ &= \mathrm{Pr}\, f(\max M_1, \min M_2) \end{aligned}$$

by (66) resp. (66′).

By considering translations of $(\mathrm{Se}'_1, \mathrm{Se}'_2)$ along the diagonal we can choose $(\mathrm{Se}_1, \mathrm{Se}_2) \in M_1 \times M_2$ such that (58) holds in addition to $\Delta\mathrm{Se} = \mathrm{Se}'_2 - \mathrm{Se}'_1 = \mathrm{Se}_2 - \mathrm{Se}_1$. This is possible since

$$M_1 \times M_2 \to \mathbb{R}, \quad (x,y) \mapsto x + y - 1,$$

is continuous, its domain connected, and

$$\begin{aligned}\Pr g(\min M_1, y) &\leq \Pr g(\min M_1, \max M_2)\\ &= (\Pr - q_{0+})^+ + q_{+1} \wedge \Pr - \Pr\\ &= \max\{q_{+1} \wedge \Pr - q_{0+}, q_{+1} \wedge \Pr - \Pr\}\\ &\leq (q_{+1} - q_{0+})^+\\ &= (q_{1,1} - q_{0,0})^+\\ &\leq q_{1,1}\end{aligned}$$

and

$$\begin{aligned}\Pr g(x, \max M_2) &\geq \Pr g(\min M_1, \max M_2)\\ &= (\Pr - q_{0+})^+ + q_{+1} \wedge \Pr - \Pr\\ &= \min\{(\Pr - q_{0+})^+ + q_{+1} - \Pr, (\Pr - q_{0+})^+\}\\ &\geq \min\{q_{+1} - q_{0+}, 0\}\\ &= \min\{q_{1,1} - q_{0,0}, 0\}\\ &\geq -q_{0,0}\end{aligned}$$

for $(x, y) \in M_1 \times M_2$.

Since $(\mathrm{Se}_1, \mathrm{Se}_2) \in M_1 \times M_2$, we have

$$\begin{aligned}\Pr - q_{1+} \leq (\Pr - q_{1+})^+ = \Pr - \Pr \wedge q_{1+} = \Pr(1 - \max M_1) \leq \Pr(1 - \mathrm{Se}_1) \leq\\ \leq \Pr(1 - \min M_1) = \Pr - (\Pr - q_{0+})^+ = q_{0+} \wedge \Pr \leq q_{0+}\end{aligned}$$

and

$$\begin{aligned}\Pr - q_{+1} \leq (\Pr - q_{+1})^+ = \Pr - \Pr \wedge q_{+1} = \Pr(1 - \max M_2) \leq \Pr(1 - \mathrm{Se}_2) \leq\\ \leq \Pr(1 - \min M_2) = \Pr - (\Pr - q_{+0})^+ = q_{+0} \wedge \Pr \leq q_{+0},\end{aligned}$$

i.e., (56) and (57). (66) resp. (66′) implies (59) and (60), hence Lemma 5.4.10 yields $(\Pr, \mathrm{Se}_1, \mathrm{Se}_2) \in B_{\leq\otimes\leq}$ resp. $B_{\leq\otimes\geq}$, and thus $(\Pr, \Delta\mathrm{Se}) \in C_{\leq\otimes\leq}$ resp. $C_{\leq\otimes\geq}$. □

5.4.13 Lemma (Analogue to Lemma 2.10 from Mattner and Mattner, 2013) *Let* $q \in \mathrm{prob}(\{0,1\}^2)$,

$$T_{\tilde{C}} \colon \Theta \to \mathbb{R}^2, \quad (\pi, \chi) \mapsto (\pi_1, \chi''_{0|0} - \chi'_{0|0}),$$

and $\tilde{C}_R := T_{\tilde{C}}[\mu^{-1}[\{q\}] \cap \Theta_R]$ *for binary relations* R *on* $\mathbb{R}^2$. *Then*

(a) $\tilde{C}_{\leq\otimes\leq} = \{(\Pr, \Delta\mathrm{Sp}) \in [0,1]^2 \colon (67)\} =: \tilde{C}'$, *where*

$$q_{1,0} - q_{0,1} \leq (1 - \Pr)\Delta\mathrm{Sp} \leq q_{1,0} \wedge (q_{1,0} - q_{0,1} + \Pr) \tag{67}$$

(b) $\tilde{C}_{\leq\otimes\geq} = \{(\Pr, \Delta\mathrm{Sp}) \in [0,1]^2 \colon (67')\} =: \tilde{C}''$, *where*

$$q_{1,0} - q_{0,1} - \Pr \leq (1 - \Pr)\Delta\mathrm{Sp} \leq q_{1,0} - q_{0,1} \tag{67′}$$

Proof. (a) Since $T_{\tilde{C}} = \sigma \circ T_C \circ \Phi_1$, we have

$$\tilde{C}_{\leq\otimes\leq} = \sigma[T_C[\Phi_1[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\leq}]]] = \sigma[T_C[\mu^{-1}[\{\psi_1(q)\}] \cap \Theta_{\leq\otimes\leq}]] = \tilde{C}'$$

due to Lemma 5.4.12 (a).

(b) Let $\tau\colon \mathbb{R}^2 \to \mathbb{R}^2$, $(x_1, x_2) \mapsto (x_1, -x_2)$. Since $T_{\tilde{C}} = \sigma \circ \tau \circ T_C \circ \Phi_1 \circ \Phi_2$, we have

$$\begin{aligned}\tilde{C}_{\leq\otimes\geq} &= \sigma[\tau[T_C[\Phi_1[\Phi_2[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\geq}]]]]] \\ &= \sigma[\tau[T_C[\mu^{-1}[\{\psi_1 \circ \psi_2(q)\}] \cap \Theta_{\leq\otimes\geq}]]] \\ &= \tilde{C}''\end{aligned}$$

due to Lemma 5.4.12 (b). □

5.4.14 Lemma (Analogue to Lemma 2.11 from Mattner and Mattner, 2013) *Let* $q \in \mathrm{prob}(\{0,1\}^2)$,

$$T_D\colon \Theta \to \mathbb{R}^2, \quad (\pi, \chi) \mapsto (\pi_1, \chi'_{1|1}),$$

and $D_R \coloneqq T_D[\mu^{-1}[\{q\}] \cap \Theta_R]$ *for binary relations* R *on* $\mathbb{R}^2$. *Then*

(a) $D_{\leq\otimes\leq} = \{(\mathrm{Pr}, \mathrm{Se}_1) \in [0,1]^2 \colon (68)\} \eqqcolon D'$, *where*

$$(\mathrm{Pr} - q_{1+}) \vee (\mathrm{Pr} - q_{+1}) \vee \frac{\mathrm{Pr} + q_{0,1} - q_{1+}}{2} \vee \frac{\mathrm{Pr} - q_{1,1}}{2} \vee \\ \vee (q_{0,1} - q_{1,0}) \leq \mathrm{Pr}(1 - \mathrm{Se}_1) \leq q_{0+} \quad (68)$$

(b) $D_{\leq\otimes\geq} = \{(\mathrm{Pr}, \mathrm{Se}_1) \in [0,1]^2 \colon (68'), q_{0,1} \leq q_{1,0}\} \eqqcolon D''$, *where*

$$(\mathrm{Pr} - q_{1+}) \vee \frac{\mathrm{Pr} + q_{0,1} - q_{1+}}{2} \leq \mathrm{Pr}(1 - \mathrm{Se}_1) \leq q_{0+} \wedge q_{+0} \wedge \frac{\mathrm{Pr} + q_{0,0}}{2} \quad (68')$$

Proof. We first show $D_{\leq\otimes\leq} \subseteq D'$ and $D_{\leq\otimes\geq} \subseteq D''$. Let $(\mathrm{Pr}, \mathrm{Se}_1) \in D_{\leq\otimes\leq}$ resp. $D_{\leq\otimes\geq}$. By Lemma 5.4.10 there exists $\mathrm{Se}_2 \in [0,1]$ with $\mathrm{Se}_2 \geq \mathrm{Se}_1$ resp. $\mathrm{Se}_2 \leq \mathrm{Se}_1$, satisfying (57) and (58). In the case $(\mathrm{Pr}, \mathrm{Se}_1) \in D_{\leq\otimes\leq}$, we have

$$\mathrm{Pr}(1 - \mathrm{Se}_1) \geq \mathrm{Pr}(1 - \mathrm{Se}_2) \geq \mathrm{Pr} - q_{+1}$$

due to $\mathrm{Se}_1 \leq \mathrm{Se}_2$ and (57), and

$$\mathrm{Pr}(1 - \mathrm{Se}_1) = \frac{\mathrm{Pr}(\mathrm{Se}_2 - \mathrm{Se}_1) - \mathrm{Pr}(\mathrm{Se}_1 + \mathrm{Se}_2 - 1) + \mathrm{Pr}}{2} \geq \frac{\mathrm{Pr} - q_{1,1}}{2}$$

due to $\mathrm{Pr} \leq 1$, $\mathrm{Se}_1 \leq \mathrm{Se}_2$, and (58). In the case $(\mathrm{Pr}, \mathrm{Se}_1) \in D_{\leq\otimes\geq}$, we have

$$\mathrm{Pr}(1 - \mathrm{Se}_1) \leq \mathrm{Pr}(1 - \mathrm{Se}_2) \leq q_{+0}$$

due to $\mathrm{Se}_1 \geq \mathrm{Se}_2$ and (57), and

$$\mathrm{Pr}(1 - \mathrm{Se}_1) = \frac{\mathrm{Pr}(\mathrm{Se}_2 - \mathrm{Se}_1) - \mathrm{Pr}(\mathrm{Se}_1 + \mathrm{Se}_2 - 1) + \mathrm{Pr}}{2} \leq \frac{\mathrm{Pr} + q_{0,0}}{2}$$

due to $\mathrm{Se}_1 \geq \mathrm{Se}_2$ and (58). The remaining inequalities in (68) resp. (68′) follow from Lemma 2.11 of Mattner and Mattner (2013).

We now show $D' \subseteq D_{\leq\otimes\leq}$ and $D'' \subseteq D_{\leq\otimes\geq}$. Let $(\mathrm{Pr},\mathrm{Se}_1) \in D'$ resp. D''. Either of (68) and (68′) implies (56). If $\mathrm{Pr} = 0$, then Se_1 and $\mathrm{Se}_2 \coloneqq 1$ resp. $\mathrm{Se}_2 \coloneqq 0$ fulfill (57)–(60). Lemma 5.4.10 yields $(\mathrm{Pr},\mathrm{Se}_1,\mathrm{Se}_2) \in B_{\leq\otimes\leq}$ resp. $B_{\leq\otimes\geq}$ and thus $(\mathrm{Pr},\mathrm{Se}_1) \in D_{\leq\otimes\leq}$ resp. $D_{\leq\otimes\geq}$.

If $\mathrm{Pr} > 0$, let us note that (57)–(60) and $0 \leq \mathrm{Se}_2 \leq 1$ are equivalent to $\mathrm{Se}_2 \in \bigcap_{i=1}^4 R_i$ with

$$R_1 \coloneqq \left[1 - \frac{q_{+0}}{\mathrm{Pr}}, \frac{q_{+1}}{\mathrm{Pr}}\right], \qquad R_2 \coloneqq \left[1 - \mathrm{Se}_1 - \frac{q_{0,0}}{\mathrm{Pr}}, 1 - \mathrm{Se}_1 + \frac{q_{1,1}}{\mathrm{Pr}}\right],$$
$$R_3 \coloneqq \left[\mathrm{Se}_1 + \frac{q_{0,1} - q_{1,0}}{\mathrm{Pr}}, \mathrm{Se}_1 + \frac{q_{0,1}}{\mathrm{Pr}}\right], \qquad R_4 \coloneqq [0,1].$$

Furthermore, (68) is equivalent to

$$\mathrm{Pr} - q_{0+} \leq \mathrm{Pr}\,\mathrm{Se}_1 \leq q_{1+} \wedge q_{+1} \wedge \frac{\mathrm{Pr} + q_{1+} - q_{0,1}}{2} \wedge \frac{\mathrm{Pr} + q_{1,1}}{2} \wedge (\mathrm{Pr} + q_{1,0} - q_{0,1}) \quad (69)$$

while (68′) is equivalent to

$$(\mathrm{Pr} - q_{0+}) \vee (\mathrm{Pr} - q_{+0}) \vee \frac{\mathrm{Pr} - q_{0,0}}{2} \leq \mathrm{Pr}\,\mathrm{Se}_1 \leq q_{1+} \wedge \frac{\mathrm{Pr} + q_{1+} - q_{0,1}}{2} \quad (69')$$

We shall now show that $\bigcap_{i=1}^4 R_i \neq \emptyset$ or, equivalently, $\max R_i \geq \min R_j$ for $i,j \in \{1,2,3,4\}$. Once this has been established, we may set $\mathrm{Se}_2 \coloneqq \min\{\max R_i : i \in \{1,2,3,4\}\}$ resp. $\mathrm{Se}_2 \coloneqq \max\{\min R_i : i \in \{1,2,3,4\}\}$ to also obtain $\mathrm{Se}_1 \leq \mathrm{Se}_2$ resp. $\mathrm{Se}_1 \geq \mathrm{Se}_2$, as will be shown afterwards.

By (68) or (69) resp. (68′) or (69′), we have

- $\mathrm{Pr}(\max R_1 - \min R_1) = 1 - \mathrm{Pr} \geq 0,$
  $\mathrm{Pr}(\max R_1 - \min R_2) = q_{1,1} + q_{0+} - \mathrm{Pr}(1 - \mathrm{Se}_1) \geq q_{1,1} \geq 0,$
  $\mathrm{Pr}(\max R_1 - \min R_3) = q_{1+} - \mathrm{Pr}\,\mathrm{Se}_1 \geq 0,$
  $\max R_1 \geq 0 = \min R_4,$
- $\mathrm{Pr}(\max R_2 - \min R_1) = \mathrm{Pr}(1 - \mathrm{Se}_1) - \mathrm{Pr} + q_{1+} + q_{0,0} \geq q_{0,0} \geq 0,$
  $\mathrm{Pr}(\max R_2 - \min R_2) = q_{1,1} + q_{0,0} \geq 0,$
  $\mathrm{Pr}(\max R_2 - \min R_3) = \mathrm{Pr} + q_{1+} - q_{0,1} - 2\mathrm{Pr}\,\mathrm{Se}_1 \geq 0,$
  $\max R_2 \geq 0 = \min R_4,$
- $\mathrm{Pr}(\max R_3 - \min R_1) = \mathrm{Pr}\,\mathrm{Se}_1 - \mathrm{Pr} + q_{0+} + q_{1,0} \geq q_{1,0} \geq 0,$
  $\mathrm{Pr}(\max R_3 - \min R_2) = 2\mathrm{Pr}\,\mathrm{Se}_1 - \mathrm{Pr} + q_{0+} \geq \mathrm{Pr}\,\mathrm{Se}_1 \geq 0,$
  $\mathrm{Pr}(\max R_3 - \min R_3) = q_{1,0} \geq 0,$
  $\max R_3 \geq 0 = \min R_4,$
- $\mathrm{Pr}(\max R_4 - \min R_3) = \mathrm{Pr} + q_{1,0} - q_{0,1} - \mathrm{Pr}\,\mathrm{Se}_1 \geq 0$
  $\max R_4 = 1 \geq \min R_1 \vee \min R_2 \vee \min R_4.$

In the case $(\mathrm{Pr},\mathrm{Se}_1) \in D'$, (69) implies

$$\mathrm{Pr}(\max R_1 - \mathrm{Se}_1) = q_{+1} - \mathrm{Pr}\,\mathrm{Se}_1 \geq 0,$$

$$\begin{aligned}
\Pr(\max R_2 - \mathrm{Se}_1) &= \Pr + q_{1,1} - 2\Pr\mathrm{Se}_1 \geq 0,\\
\Pr(\max R_3 - \mathrm{Se}_1) &= q_{0,1},\\
\max R_4 &= 1 \geq \mathrm{Se}_1,
\end{aligned}$$

hence $\mathrm{Se}_1 \leq \mathrm{Se}_2$. In the case $(\Pr, \mathrm{Se}_1) \in D''$, (69′) implies

$$\begin{aligned}
\Pr(\mathrm{Se}_1 - \min R_1) &= \Pr\mathrm{Se}_1 - \Pr + q_{+0} \geq 0,\\
\Pr(\mathrm{Se}_1 - \min R_2) &= 2\Pr\mathrm{Se}_1 - \Pr + q_{0,0} \geq 0,\\
\Pr(\mathrm{Se}_1 - \min R_3) &= q_{1,0} - q_{0,1} \geq 0,\\
\mathrm{Se}_1 &\geq \min R_4 = 0,
\end{aligned}$$

hence $\mathrm{Se}_1 \geq \mathrm{Se}_2$. Lemma 5.4.10 yields $(\Pr, \mathrm{Se}_1, \mathrm{Se}_2) \in B_{\leq\otimes\leq}$ resp. $B_{\leq\otimes\geq}$, i.e., $(\Pr, \mathrm{Se}_1) \in D_{\leq\otimes\leq}$ resp. $D_{\leq\otimes\geq}$. □

5.4.15 Lemma (Analogue to Lemma 2.11 from Mattner and Mattner, 2013) *Let* $q \in \mathrm{prob}(\{0,1\}^2)$,

$$T_{\tilde{D}} \colon \Theta \to \mathbb{R}^2, \quad (\pi, \chi) \mapsto (\pi_1, \chi'_{0|0}),$$

and $\tilde{D}_R := T_{\tilde{D}}[\mu^{-1}[\{q\}] \cap \Theta_R]$ *for binary relations* R *on* $\mathbb{R}^2$. *Then*

(a) $\tilde{D}_{\leq\otimes\leq} = \{(\Pr, \mathrm{Sp}_1) \in [0,1]^2 \colon (70)\} =: \tilde{D}'$, *where*

$$q_{0+} - \Pr \leq (1-\Pr)\mathrm{Sp}_1 \leq q_{0+} \wedge q_{+0} \wedge \frac{1 - \Pr - q_{1,0} + q_{0+}}{2} \wedge \\ \wedge \frac{1 - \Pr + q_{0,0}}{2} \wedge (1 - \Pr - q_{1,0} + q_{0,1}) \quad (70)$$

(b) $\tilde{D}_{\leq\otimes\geq} = \{(\Pr, \mathrm{Sp}_1) \in [0,1]^2 \colon (70'), q_{0,1} \leq q_{1,0}\} =: \tilde{D}''$, *where*

$$(q_{0+} - \Pr) \vee \frac{q_{0,1} + q_{0+} - \Pr}{2} \leq (1-\Pr)\mathrm{Sp}_1 \leq q_{0+} \wedge q_{+0} \wedge \frac{q_{0,0} + 1 - \Pr}{2} \quad (70')$$

Proof. Since $T_{\tilde{D}} = \sigma \circ T_D \circ \Phi_1$, we have $\tilde{D}_{\leq\otimes\leq} = \sigma[T_D[\Phi_1[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\leq}]]] = \sigma[T_D[\mu^{-1}[\{\psi_1(q)\}] \cap \Theta_{\leq\otimes\leq}]] = \tilde{D}'$ due to Lemma 5.4.14 (a), which proves (a).

To prove (b), we first show $\tilde{D}_{\leq\otimes\geq} \subseteq \tilde{D}''$. Let $(\Pr, \mathrm{Sp}_1) \in \tilde{D}_{\leq\otimes\geq}$. By Lemma 5.4.11 there exists $\mathrm{Sp}_2 \in [\mathrm{Sp}_1, 1]$ with (62) and (63). We thus have

$$(1-\Pr)\mathrm{Sp}_1 \leq (1-\Pr)\mathrm{Sp}_2 \leq q_{+0}$$

due to $\mathrm{Sp}_1 \leq \mathrm{Sp}_2$ and (62), and

$$\begin{aligned}
(1-\Pr)\mathrm{Sp}_1 &= \frac{(1-\Pr)(\mathrm{Sp}_1 + \mathrm{Sp}_2 - 1) - (1-\Pr)(\mathrm{Sp}_2 - \mathrm{Sp}_1) + 1 - \Pr}{2}\\
&\leq \frac{q_{0,0} + 1 - \Pr}{2}
\end{aligned}$$

due to $\mathrm{Sp}_1 \leq \mathrm{Sp}_2$ and (63). Together with

$$\tilde{D}_{\leq\otimes\geq} = \sigma[T_E[\Phi_1[\Phi_2[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\geq}]]]]$$

$$= \sigma[T_E[\mu^{-1}[\{\psi_1 \circ \psi_2(q)\}] \cap \Theta_{\leq\otimes\geq}]]$$
$$\subseteq \sigma[T_E[\mu^{-1}[\{\psi_1 \circ \psi_2(q)\}] \cap \Theta_{\leq\otimes\mathbb{R}^2}]]$$

and Lemma 2.12 from Mattner and Mattner (2013), we obtain (70′), i.e., $(\mathrm{Pr}, \mathrm{Sp}_1) \in \tilde{D}''$.

We now show $\tilde{D}'' \subseteq \tilde{D}_{\leq\otimes\geq}$. Let $(\mathrm{Pr}, \mathrm{Sp}_1) \in \tilde{D}''$. (70′) implies (61). If $\mathrm{Pr} = 1$, then Sp_1 and $\mathrm{Sp}_2 \coloneqq 1$ satisfy (62), (63), and (64′). Lemma 5.4.11 yields $(\mathrm{Pr}, \mathrm{Sp}_1, \mathrm{Sp}_2) \in \tilde{B}_{\leq\otimes\geq}$ and thus $(\mathrm{Pr}, \mathrm{Sp}_1) \in \tilde{D}_{\leq\otimes\geq}$.

If $\mathrm{Pr} < 1$, let us note that (62), (63), (64′), and $\mathrm{Sp}_1 \leq \mathrm{Sp}_2 \leq 1$ are equivalent to $\mathrm{Sp}_2 \in \bigcap_{i=1}^4 \tilde{R}_i$ with

$$\tilde{R}_1 \coloneqq \left[\frac{q_{+0} - \mathrm{Pr}}{1 - \mathrm{Pr}}, \frac{q_{+0}}{1 - \mathrm{Pr}}\right], \qquad \tilde{R}_2 \coloneqq \left[1 - \mathrm{Sp}_1 - \frac{q_{1,1}}{1 - \mathrm{Pr}}, 1 - \mathrm{Sp}_1 + \frac{q_{0,0}}{1 - \mathrm{Pr}}\right],$$
$$\tilde{R}_3 \coloneqq \left[\mathrm{Sp}_1, \mathrm{Sp}_1 + \frac{q_{1,0} - q_{0,1}}{1 - \mathrm{Pr}}\right], \qquad \tilde{R}_4 \coloneqq [0, 1].$$

Furthermore, (70′) is equivalent to

$$(q_{1+} - \mathrm{Pr}) \vee (q_{+1} - \mathrm{Pr}) \vee \frac{1 - \mathrm{Pr} - q_{0,0}}{2} \leq (1 - \mathrm{Pr})(1 - \mathrm{Sp}_1) \leq$$
$$\leq q_{1+} \wedge \frac{1 - \mathrm{Pr} + q_{1+} - q_{0,1}}{2} \quad (71)$$

We shall now show that $\bigcap_{i=1}^4 \tilde{R}_i \neq \emptyset$ or, equivalently, $\max \tilde{R}_i \geq \min \tilde{R}_j$ for $i, j \in \{1, 2, 3, 4\}$. Once this has been established, we may, for instance, set $\mathrm{Sp}_2 \coloneqq \min\{\max \tilde{R}_i : i \in \{1, 2, 3, 4\}\}$.

By (70′) or (71), we have

- $(1 - \mathrm{Pr})(\max \tilde{R}_1 - \min \tilde{R}_1) = \mathrm{Pr} \geq 0,$
  $(1 - \mathrm{Pr})(\max \tilde{R}_1 - \min \tilde{R}_2) = q_{1+} + q_{0,0} - (1 - \mathrm{Pr})(1 - \mathrm{Sp}_1) \geq q_{0,0} \geq 0,$
  $(1 - \mathrm{Pr})(\max \tilde{R}_1 - \min \tilde{R}_3) = q_{+0} - (1 - \mathrm{Pr})(1 - \mathrm{Sp}_1) \geq 0,$
  $\max \tilde{R}_1 \geq 0 = \min \tilde{R}_4,$
- $(1 - \mathrm{Pr})(\max \tilde{R}_2 - \min \tilde{R}_1) = \mathrm{Pr} - q_{1,0} + (1 - \mathrm{Pr})(1 - \mathrm{Sp}_1) \geq (1 - \mathrm{Pr})(1 - \mathrm{Sp}_1) - q_{1+} + \mathrm{Pr} \geq 0,$
  $(1 - \mathrm{Pr})(\max \tilde{R}_2 - \min \tilde{R}_2) = q_{0,0} + q_{1,1} \geq 0,$
  $(1 - \mathrm{Pr})(\max \tilde{R}_2 - \min \tilde{R}_3) = q_{0,0} + 1 - \mathrm{Pr} - 2(1 - \mathrm{Pr})(1 - \mathrm{Sp}_1) \geq 0,$
  $\max \tilde{R}_2 \geq 0 = \min \tilde{R}_4,$
- $(1 - \mathrm{Pr})(\max \tilde{R}_3 - \min \tilde{R}_1) = \mathrm{Pr} - q_{0+} + (1 - \mathrm{Pr})\mathrm{Sp}_1 \geq 0,$
  $(1 - \mathrm{Pr})(\max \tilde{R}_3 - \min \tilde{R}_2) = \mathrm{Pr} - q_{0+} - q_{0,1} + 2(1 - \mathrm{Pr})\mathrm{Sp}_1 \geq 0,$
  $(1 - \mathrm{Pr})(\max \tilde{R}_3 - \min \tilde{R}_3) = q_{1,0} - q_{0,1} \geq 0,$
  $\max \tilde{R}_3 \geq 0 = \min \tilde{R}_4,$
- $(1 - \mathrm{Pr})(\max \tilde{R}_4 - \min \tilde{R}_2) = (1 - \mathrm{Pr})\mathrm{Sp}_1 + q_{1,1} \geq 0,$
  $\max \tilde{R}_4 = 1 \geq \min \tilde{R}_1 \vee \min \tilde{R}_3 \vee \min \tilde{R}_4.$

Lemma 5.4.11 yields $(\mathrm{Pr}, \mathrm{Sp}_1, \mathrm{Sp}_2) \in \tilde{B}_{\leq\otimes\geq}$, i.e., $(\mathrm{Pr}, \mathrm{Sp}_1) \in \tilde{D}_{\leq\otimes\geq}$. □

5.4.16 Lemma (Analogue to Lemma 2.12 from Mattner and Mattner, 2013) *Let* $q \in \operatorname{prob}(\{0,1\}^2)$,

$$T_E\colon \Theta \to \mathbb{R}^2, \quad (\pi,\chi) \mapsto (\pi_1, \chi''_{1|1}),$$

and $E_R \coloneqq T_E[\mu^{-1}[\{q\}) \cap \Theta_R]$ *for binary relations* R *on* $\mathbb{R}^2$. *Then*

(a) $E_{\leq\otimes\leq} = \{(\mathrm{Pr}, \mathrm{Se}_2) \in [0,1]^2 \colon (72)\} =: E'$, *where*

$$(\mathrm{Pr} - q_{+1}) \vee \frac{\mathrm{Pr} - q_{+1}}{2} \leq \mathrm{Pr}(1 - \mathrm{Se}_2) \leq q_{+0} \wedge q_{0+} \wedge \\ \wedge (\mathrm{Pr} + q_{1,0} - q_{0,1}) \wedge \frac{\mathrm{Pr} + q_{+0} - q_{0,1}}{2} \wedge \frac{\mathrm{Pr} + q_{0,0}}{2} \quad (72)$$

(b) $E_{\leq\otimes\geq} = \{(\mathrm{Pr}, \mathrm{Se}_2) \in [0,1]^2 \colon (72'), q_{0,1} \leq q_{1,0}\} =: E''$, *where*

$$q_{+0} \vee \frac{\mathrm{Pr} + q_{+0} - q_{0,1}}{2} \leq \mathrm{Pr}(1 - \mathrm{Se}_2) \leq (\mathrm{Pr} - q_{+1}) \wedge (\mathrm{Pr} - q_{1+}) \wedge \frac{\mathrm{Pr} - q_{1,1}}{2} \quad (72')$$

Proof. Since $T_E = \sigma \circ T_{\tilde{D}} \circ \Phi_1 \circ \Phi_2$, we have $E_{\leq\otimes\geq} = \sigma[T_{\tilde{D}}[\Phi_1[\Phi_2[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\geq}]]]] = \sigma[T_{\tilde{D}}[\mu^{-1}[\{\psi_1 \circ \psi_2(q)\}] \cap \Theta_{\leq\otimes\geq}]] = E''$ due to Lemma 5.4.15 (b), which proves (b).

To prove (a), we first show $E_{\leq\otimes\leq} \subseteq E'$. Let $(\mathrm{Pr}, \mathrm{Se}_2) \in E_{\leq\otimes\leq}$. By Lemma 5.4.10 there exists $\mathrm{Se}_1 \in [0, \mathrm{Se}_2]$ with (56) and (58). We thus have

$$(1 - \mathrm{Pr})\mathrm{Se}_2 \leq (1 - \mathrm{Pr})\mathrm{Se}_1 \leq q_{0+}$$

due to $\mathrm{Se}_1 \leq \mathrm{Se}_2$ and (56), as well as

$$(1 - \mathrm{Pr})\mathrm{Se}_2 = \frac{\mathrm{Pr} - \mathrm{Pr}(\mathrm{Se}_2 - \mathrm{Se}_1) - \mathrm{Pr}(\mathrm{Se}_1 + \mathrm{Se}_2 - 1)}{2} \geq \frac{\mathrm{Pr} - q_{+1}}{2}$$

due to (58) and (60), and

$$(1 - \mathrm{Pr})\mathrm{Se}_2 = \frac{\mathrm{Pr} - \mathrm{Pr}(\mathrm{Se}_2 - \mathrm{Se}_1) - \mathrm{Pr}(\mathrm{Se}_1 + \mathrm{Se}_2 - 1)}{2} \leq \frac{\mathrm{Pr} + q_{0,0}}{2}$$

due to $\mathrm{Se}_1 \leq \mathrm{Se}_2$ and (58). Together with $E_{\leq\otimes\leq} \subseteq T_E[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\mathbb{R}^2}]$ and Lemma 2.12 from Mattner and Mattner (2013), we obtain (72), i.e., $(\mathrm{Pr}, \mathrm{Se}_2) \in E'$.

We now show $E' \subseteq E_{\leq\otimes\leq}$. Let $(\mathrm{Pr}, \mathrm{Se}_2) \in E'$. (72) implies (57). If $\mathrm{Pr} = 0$, then Se_2 and $\mathrm{Se}_1 \coloneqq 0$ satisfy (56) and (58)–(60). Lemma 5.4.10 yields $(\mathrm{Pr}, \mathrm{Se}_1, \mathrm{Se}_2) \in B_{\leq\otimes\leq}$ and thus $(\mathrm{Pr}, \mathrm{Se}_2) \in E_{\leq\otimes\leq}$.

If $\mathrm{Pr} > 0$, let us note that (56), (58)–(60), and $0 \leq \mathrm{Se}_1 \leq \mathrm{Se}_2$ are equivalent to $\mathrm{Se}_1 \in \bigcap_{i=1}^4 S_i$ with

$$S_1 \coloneqq \left[1 - \frac{q_{0+}}{\mathrm{Pr}}, \frac{q_{1+}}{\mathrm{Pr}}\right], \qquad S_2 \coloneqq \left[1 - \mathrm{Se}_2 - \frac{q_{0,0}}{\mathrm{Pr}}, 1 - \mathrm{Se}_2 + \frac{q_{1,1}}{\mathrm{Pr}}\right],$$
$$S_3 \coloneqq \left[\mathrm{Se}_2 - \frac{q_{0,1}}{\mathrm{Pr}}, \mathrm{Se}_2 + \frac{q_{1,0} - q_{0,1}}{\mathrm{Pr}}\right], \qquad S_4 \coloneqq [0, \mathrm{Se}_2].$$

Furthermore, (72) is equivalent to

$$(\mathrm{Pr}-q_{+0})\vee(\mathrm{Pr}-q_{0+})\vee(q_{0,1}-q_{1,0})\vee\frac{\mathrm{Pr}+q_{0,1}-q_{+0}}{2}\vee \\ \vee\frac{\mathrm{Pr}-q_{0,0}}{2}\leq \mathrm{PrSe}_2\leq q_{+1}\wedge\frac{\mathrm{Pr}+q_{+1}}{2} \quad (73)$$

We shall now show that $\bigcap_{i=1}^4 S_i \neq \emptyset$ or, equivalently, $\max S_i \geq \min S_j$ for $i,j \in \{1,2,3,4\}$. Once this has been established, we may, for instance, set $\mathrm{Sp}_1 \coloneqq \max\{\min S_i\colon i \in \{1,2,3,4\}\}$.

By (72) or (73), we have

- $\mathrm{Pr}(\max S_1-\min S_1)=1-\mathrm{Pr}\geq 0,$
  $\mathrm{Pr}(\max S_1-\min S_2)=q_{1,1}+q_{+0}-\mathrm{Pr}(1-\mathrm{Se}_2)\geq q_{1,1}\geq 0,$
  $\mathrm{Pr}(\max S_1-\min S_3)=q_{1,0}+q_{+1}-\mathrm{Pr}\,\mathrm{Se}_2\geq q_{1,0}\geq 0,$
  $\max S_1\geq 0=\min S_4,$
- $\mathrm{Pr}(\max S_2-\min S_1)=\mathrm{Pr}(1-\mathrm{Se}_2)-\mathrm{Pr}+q_{+1}+q_{0,0}\geq q_{0,0}\geq 0,$
  $\mathrm{Pr}(\max S_2-\min S_2)=q_{0,0}+q_{1,1}\geq 0,$
  $\mathrm{Pr}(\max S_2-\min S_3)=\mathrm{Pr}+q_{+1}-2\mathrm{Pr}\,\mathrm{Se}_2\geq 0,$
  $\max S_2\geq 0=\min S_4,$
- $\mathrm{Pr}(\max S_3-\min S_1)=\mathrm{Pr}\,\mathrm{Se}_2+q_{+0}-\mathrm{Pr}\geq 0,$
  $\mathrm{Pr}(\max S_3-\min S_2)=2\mathrm{Pr}\,\mathrm{Se}_2+q_{+0}-q_{0,1}-\mathrm{Pr}\geq 0,$
  $\mathrm{Pr}(\max S_3-\min S_3)=q_{1,0}\geq 0,$
  $\max S_3\geq 0=\min S_4,$
- $\mathrm{Pr}(\max S_4-\min S_1)=\mathrm{Pr}\,\mathrm{Se}_2-\mathrm{Pr}+q_{0+}\geq 0,$
  $\mathrm{Pr}(\max S_4-\min S_2)=2\mathrm{Pr}\,\mathrm{Se}_2+q_{0,0}-\mathrm{Pr}\geq 0,$
  $\mathrm{Pr}(\max S_4-\min S_3)=q_{0,1}\geq 0,$
  $\max S_4\geq 0=\min S_4.$

Lemma 5.4.10 yields $(\mathrm{Pr},\mathrm{Se}_1,\mathrm{Se}_2)\in B_{\leq\otimes\leq}$, i.e., $(\mathrm{Pr},\mathrm{Se}_2)\in E_{\leq\otimes\leq}$. □

5.4.17 Lemma (Analogue to Lemma 2.12 from Mattner and Mattner, 2013) *Let* $q \in \mathrm{prob}(\{0,1\}^2)$,

$$T_{\tilde{E}}\colon \Theta\to\mathbb{R}^2,\quad (\pi,\chi)\mapsto(\pi_1,\chi''_{0|0}),$$

and $\tilde{E}_R \coloneqq T_{\tilde{E}}[\mu^{-1}[\{q\}]\cap\Theta_R]$ *for binary relations* R *on* $\mathbb{R}^2$. *Then*

(a) $\tilde{E}_{\leq\otimes\leq}=\{(\mathrm{Pr},\mathrm{Sp}_2)\in[0,1]^2\colon (74)\}\eqqcolon\tilde{E}'$, *where*

$$(q_{+0}-\mathrm{Pr})\vee\frac{q_{+0}+q_{1,0}-\mathrm{Pr}}{2}\vee(q_{1,0}-q_{0,1})\leq(1-\mathrm{Pr})\mathrm{Sp}_2\leq q_{+0} \quad (74)$$

(b) $\tilde{E}_{\leq\otimes\geq}=\{(\mathrm{Pr},\mathrm{Sp}_2)\in[0,1]^2\colon (74')\}\eqqcolon\tilde{E}''$, *where*

$$(q_{+1}-\mathrm{Pr})\vee\frac{q_{+1}+q_{0,1}-\mathrm{Pr}}{2}\leq(1-\mathrm{Pr})(1-\mathrm{Sp}_2)\leq q_{+1}\wedge q_{1+}\wedge \\ \wedge\frac{1-\mathrm{Pr}+q_{1,1}}{2} \quad (74')$$

Proof. (a) Since $T_{\tilde{E}} = \sigma \circ T_E \circ \Phi_1$, we have

$$\tilde{E}_{\leq\otimes\leq} = \sigma[T_E[\Phi_1[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\leq}]]] = \sigma[T_E[\mu^{-1}[\{\psi_1(q)\}] \cap \Theta_{\leq\otimes\leq}]] = \tilde{E}'$$

due to Lemma 5.4.16 (a).

(b) Since $T_{\tilde{E}} = \sigma \circ T_D \circ \Phi_1 \circ \Phi_2$, we have

$$\begin{aligned}\tilde{E}_{\leq\otimes\geq} &= \sigma[T_D[\Phi_1[\Phi_2[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\geq}]]]]\\ &= \sigma[T_D[\mu^{-1}[\{\psi_1 \circ \psi_2(q)\}] \cap \Theta_{\leq\otimes\geq}]]\\ &= \tilde{E}''\end{aligned}$$

due to Lemma 5.4.14 (b). □

5.4.18 Lemma (Analogue to Lemma 2.13 from Mattner and Mattner, 2013) *Let* $q \in \operatorname{prob}(\{0,1\}^2)$,

$$T_F \colon \Theta \to \mathbb{R}, \quad (\pi,\chi) \mapsto \chi''_{1|1} - \chi'_{1|1},$$

and $F_R \coloneqq T_F[\mu^{-1}[\{q\}] \cap \Theta_R]$ *for binary relations* R *on* $\mathbb{R}^2$. *Then*

(a) $F_{\leq\otimes\leq} = [(q_{0,1} - q_{1,0})^+, 1] \eqqcolon F'$.

(b) $F_{\leq\otimes\geq} = \left\{\begin{array}{cc}[-1,0] & \text{if } q_{0,1} \leq q_{1,0}\\ \emptyset & \text{otherwise}\end{array}\right\} \eqqcolon F''$.

Proof. (a) It remains to show $F' \subseteq F_{\leq\otimes\leq}$. Let $\Delta\mathrm{Se} \in F'$ and consider $\mathrm{Pr} \coloneqq \mathbf{1}(q_{1,0} < q_{0,1})(q_{0,1} - q_{1,0})/\Delta\mathrm{Se}$. Lemma 5.4.12 yields $(\mathrm{Pr}, \Delta\mathrm{Se}) \in C_{\leq\otimes\leq}$ and thus $\Delta\mathrm{Se} \in F_{\leq\otimes\leq}$.

(b) If $q_{0,1} \leq q_{1,0}$, then $F'' = [-1,0] \supseteq F_{\leq\otimes\geq}$. For $\Delta\mathrm{Se} \in F''$ and $\mathrm{Pr} \coloneqq 0$ we obtain $(\mathrm{Pr}, \Delta\mathrm{Se}) \in C_{\leq\otimes\geq}$ by Lemma 5.4.12 and thus $\Delta\mathrm{Se} \in F_{\leq\otimes\geq}$.

If $q_{0,1} > q_{1,0}$, then Lemma 5.4.7 implies $F_{\leq\otimes\geq} = \emptyset = F''$. □

5.4.19 Lemma (Analogue to Lemma 2.13 from Mattner and Mattner, 2013) *Let* $q \in \operatorname{prob}(\{0,1\}^2)$,

$$T_{\tilde{F}} \colon \Theta \to \mathbb{R}, \quad (\pi,\chi) \mapsto \chi''_{0|0} - \chi'_{0|0},$$

and $\tilde{F}_R \coloneqq T_{\tilde{F}}[\mu^{-1}[\{q\}] \cap \Theta_R]$ *for binary relations* R *on* $\mathbb{R}^2$. *Then*

(a) $\tilde{F}_{\leq\otimes\leq} = [(q_{1,0} - q_{0,1})^+, 1] \eqqcolon \tilde{F}'$.

(b) $\tilde{F}_{\leq\otimes\geq} = \left\{\begin{array}{cc}[0,1] & \text{if } q_{0,1} \leq q_{1,0}\\ \emptyset & \text{otherwise}\end{array}\right\} \eqqcolon \tilde{F}''$.

Proof. (a) Since $T_{\tilde{F}} = T_F \circ \Phi_1$, we have

$$\tilde{F}_{\leq\otimes\leq} = T_F[\Phi_1[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\leq}]] = T_F[\mu^{-1}[\{\psi_1(q)\}] \cap \Theta_{\leq\otimes\leq}] = \tilde{F}'$$

due to Lemma 5.4.18 (a).

(b) Since $T_{\tilde{F}} = -T_F \circ \Phi_1 \circ \Phi_2$, we have

$$\begin{aligned}\tilde{F}_{\leq\otimes\geq} &= -T_F[\Phi_1[\Phi_2[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\geq}]]] \\ &= -T_F[\mu^{-1}[\{\psi_1 \circ \psi_2(q)\}] \cap \Theta_{\leq\otimes\geq}] \\ &= \tilde{F}''\end{aligned}$$

due to Lemma 5.4.18 (b). □

5.4.20 Lemma (Analogue to Lemma 2.14 from Mattner and Mattner, 2013) *Let* $q \in \operatorname{prob}(\{0,1\}^2)$,

$$T_G \colon \Theta \to \mathbb{R}, \quad (\pi, \chi) \mapsto \chi'_{1|1},$$

and $G_R := T_G[\mu^{-1}[\{q\}] \cap \Theta_R]$ *for binary relations* R *on* $\mathbb{R}^2$*. Then*

(a) $G_{\leq\otimes\leq} = G'$, *where*

$$G' := \left[0, \left(\frac{q_{1+}}{q_{1+} + q_{0,1}} \vee \frac{q_{1,1}}{(q_{+1} - q_{1,0})^+}\right) \wedge 1\right] \quad \text{with } \frac{0}{0} := 1$$

(b) $G_{\leq\otimes\geq} = \left\{\begin{array}{ll} [0,1] & \text{if } q_{0,1} \leq q_{1,0} \\ \emptyset & \text{otherwise} \end{array}\right\} =: G''$

Proof. (a) It remains to show $G' \subseteq G_{\leq\otimes\leq}$. Let $\mathrm{Se}_1 \in G'$. If $q_{1,0} \geq q_{0,1}$, then $q_{+1} - q_{1,0} \leq q_{1,1}$ and thus $G' = [0,1]$. With $\Pr := 0$ we obtain $(\Pr, \mathrm{Se}_1) \in D_{\leq\otimes\leq}$ by Lemma 5.4.14 and thus $\mathrm{Se}_1 \in G_{\leq\otimes\leq}$.

If $q_{1,0} < q_{0,1}$, let

$$f \colon]0,1] \to \mathbb{R}, \quad x \mapsto \min\left\{\frac{q_{1+}}{x}, \frac{1}{2} + \frac{q_{1+} - q_{0,1}}{2x}, 1 + \frac{q_{1,0} - q_{0,1}}{x}\right\},$$

as well as $x_1 := q_{1+} + q_{0,1}$ and $x_2 := q_{+1} - q_{1,0}$, both belonging to $]0,1]$ in view of $q_{1,0} < q_{0,1}$ and $q \in \operatorname{prob}(\{0,1\}^2)$. Since

$$\begin{aligned} f(x_1) &= \min\left\{\frac{q_{1+}}{q_{1+} + q_{0,1}}, \frac{1}{2} + \frac{q_{1+} - q_{0,1}}{2q_{1+} + 2q_{0,1}}, 1 + \frac{q_{1,0} - q_{0,1}}{q_{1+} + q_{0,1}}\right\} \\ &= \min\left\{\frac{q_{1+}}{q_{1+} + q_{0,1}}, \frac{q_{1+} + q_{1,0}}{q_{1+} + q_{0,1}}\right\} \\ &= \frac{q_{1+}}{q_{1+} + q_{0,1}} \end{aligned}$$

and

$$\begin{aligned} f(x_2) &= \min\left\{\frac{q_{1+}}{q_{+1} - q_{1,0}}, \frac{1}{2} + \frac{q_{1+} - q_{0,1}}{2q_{+1} - 2q_{1,0}}, 1 + \frac{q_{1,0} - q_{0,1}}{q_{+1} - q_{1,0}}\right\} \\ &= \min\left\{\frac{q_{1+}}{q_{+1} - q_{1,0}}, \frac{q_{1,1}}{q_{+1} - q_{1,0}}\right\} \\ &= \frac{q_{1,1}}{q_{+1} - q_{1,0}} \end{aligned}$$

$$= \frac{q_{1,1}}{(q_{+1} - q_{1,0})^+},$$

we have $\lim_{x\to 0} f(x) = -\infty < \mathrm{Se}_1 \le \max G' \le f(x_1) \vee f(x_2)$. The function f being continuous, we can pick $\mathrm{Pr} \in]0,1]$ such that $\mathrm{Se}_1 = f(\mathrm{Pr})$. Lemma 5.4.14 yields $(\mathrm{Pr}, \mathrm{Se}_1) \in D_{\le\otimes\le}$ and thus $\mathrm{Se}_1 \in G_{\le\otimes\le}$.

(b) If $q_{0,1} \le q_{1,0}$, then $G'' = [0,1] \supseteq G_{\le\otimes\ge}$. For $\mathrm{Se}_1 \in G''$ and $\mathrm{Pr} := 0$ we obtain $(\mathrm{Pr}, \mathrm{Se}_1) \in D_{\le\otimes\ge}$ by Lemma 5.4.14 and thus $\mathrm{Se}_1 \in G_{\le\otimes\ge}$.

If $q_{0,1} > q_{1,0}$, then Lemma 5.4.7 implies $G_{\le\otimes\ge} = \emptyset = G''$. □

5.4.21 Lemma (Analogue to Lemma 2.14 from Mattner and Mattner, 2013) *Let* $q \in \mathrm{prob}(\{0,1\}^2)$,

$$T_{\tilde{G}} \colon \Theta \to \mathbb{R}, \quad (\pi, \chi) \mapsto \chi'_{0|0},$$

and $\tilde{G}_R := T_{\tilde{G}}[\mu^{-1}[\{q\}] \cap \Theta_R]$ *for binary relations* R *on* $\mathbb{R}^2$*. Then*

(a) $\tilde{G}_{\le\otimes\le} = \tilde{G}'$, *where*

$$\tilde{G}' := \left[0, \left(\frac{q_{0+}}{q_{0+} + q_{1,0}} \vee \frac{q_{0,0}}{(q_{+0} - q_{0,1})^+}\right) \wedge 1\right] \quad \textit{with } \frac{0}{0} := 1$$

(b) $$\tilde{G}_{\le\otimes\ge} = \left\{ \begin{array}{cc} [0,1] & \textit{if } q_{0,1} \le q_{1,0} \\ \emptyset & \textit{otherwise} \end{array} \right\} =: \tilde{G}''$$

Proof. (a) Since $T_{\tilde{G}} = T_G \circ \Phi_1$, we have

$$\tilde{G}_{\le\otimes\le} = T_G[\Phi_1[\mu^{-1}[\{q\}] \cap \Theta_{\le\otimes\le}]] = T_G[\mu^{-1}[\{\psi_1(q)\}] \cap \Theta_{\le\otimes\le}] = \tilde{G}'$$

due to Lemma 5.4.20 (a).

(b) If $q_{0,1} \le q_{1,0}$, then $\tilde{G}'' = [0,1] \supseteq \tilde{G}_{\le\otimes\ge}$. For $\mathrm{Sp}_1 \in \tilde{G}''$ and $\mathrm{Pr} := 1$ we obtain $(\mathrm{Pr}, \mathrm{Se}_1) \in \tilde{D}_{\le\otimes\ge}$ by Lemma 5.4.15 and thus $\mathrm{Sp}_1 \in \tilde{G}_{\le\otimes\ge}$.

If $q_{0,1} > q_{1,0}$, then Lemma 5.4.7 implies $\tilde{G}_{\le\otimes\ge} = \emptyset = \tilde{G}''$. □

5.4.22 Remark Let us note that in the next results $0/0 := 0$, as opposed to the preceding Lemmas. We would like to point out that this definition (instead of their $0/0 := 1$) ought to be used in Lemma 2.15 of Mattner and Mattner (2013).

5.4.23 Lemma (Analogue to Lemma 2.15 from Mattner and Mattner, 2013) *Let* $q \in \mathrm{prob}(\{0,1\}^2)$,

$$T_H \colon \Theta \to \mathbb{R}, \quad (\pi, \chi) \mapsto \chi''_{1|1},$$

and $H_R := T_H(\mu^{-1}(\{q\}) \cap \Theta_R)$ *for binary relations* R *on* $\mathbb{R}^2$*. Then*

(a) $H_{\le\otimes\le} = H'$, *where*

$$H' := \left[\frac{q_{0,1}}{q_{0+} + q_{1,0}} \wedge \frac{(q_{0,1} - q_{1,0})^+}{q_{0+} - q_{1,0}}, 1\right] \quad \textit{with } \frac{0}{0} := 0$$

(b) $H_{\leq\otimes\geq} = \left\{ \begin{array}{ll} [0,1] & \textit{if } q_{0,1} \leq q_{1,0} \\ \emptyset & \textit{otherwise} \end{array} \right\} =: H''$

Proof. (a) It remains to show $H' \subseteq H_{\leq\otimes\leq}$. Let $\mathrm{Se}_2 \in H'$. If $q_{1,0} \geq q_{0,1}$, then $H' = [0,1]$. With $\mathrm{Pr} := 0$ we obtain $(\mathrm{Pr}, \mathrm{Se}_2) \in E_{\leq\otimes\leq}$ by Lemma 5.4.16, and thus $\mathrm{Se}_2 \in H_{\leq\otimes\leq}$.

If $q_{1,0} < q_{0,1}$, let

$$f\colon]0,1] \to \mathbb{R}, \quad x \mapsto \max\left\{1 - \frac{q_{+0}}{x}, \frac{1}{2} + \frac{q_{0,1} - q_{+0}}{2x}, \frac{q_{0,1} - q_{1,0}}{x}\right\},$$

as well as $x_1 := q_{+0} + q_{0,1}$ and $x_2 := q_{0+} - q_{1,0}$, both belonging to $]0,1]$ in view of $q_{1,0} < q_{0,1}$ and $q \in \mathrm{prob}(\{0,1\}^2)$. Since

$$\begin{aligned} f(x_1) &= \max\left\{1 - \frac{q_{+0}}{q_{+0} + q_{0,1}}, \frac{q_{0,1} - q_{1,0}}{q_{+0} + q_{0,1}}, \frac{1}{2} + \frac{q_{0,1} - q_{+0}}{2q_{+0} + 2q_{0,1}}\right\} \\ &= \max\left\{\frac{q_{0,1}}{q_{+0} + q_{0,1}}, \frac{q_{0,1} - q_{1,0}}{q_{+0} + q_{0,1}}\right\} \\ &= \frac{q_{0,1}}{q_{+0} + q_{0,1}} \end{aligned}$$

and

$$\begin{aligned} f(x_2) &= \max\left\{1 - \frac{q_{+0}}{q_{0+} - q_{1,0}}, \frac{q_{0,1} - q_{1,0}}{q_{0+} - q_{1,0}}, \frac{1}{2} + \frac{q_{0,1} - q_{+0}}{2q_{0+} - 2q_{1,0}}\right\} \\ &= \max\left\{\frac{q_{0,1} - 2q_{1,0}}{q_{0+} - q_{1,0}}, \frac{q_{0,1} - q_{1,0}}{q_{0+} - q_{1,0}}\right\} \\ &= \frac{q_{0,1} - q_{1,0}}{q_{0+} - q_{1,0}}, \end{aligned}$$

we have $\lim_{x\to 0} f(x) = \infty > \mathrm{Se}_2 \geq \min H' \geq f(x_1) \wedge f(x_2)$. The function f being continuous, we can pick $\mathrm{Pr} \in]0,1]$ such that $\mathrm{Se}_2 = f(\mathrm{Pr})$. Lemma 5.4.16 yields $(\mathrm{Pr}, \mathrm{Se}_2) \in E_{\leq\otimes\leq}$, and thus $\mathrm{Se}_2 \in H_{\leq\otimes\leq}$.

(b) Since $T_H = T_{\tilde{G}} \circ \Phi_1 \circ \Phi_2$, we have

$$H_{\leq\otimes\geq} = T_{\tilde{G}}[\Phi_1[\Phi_2[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\geq}]]] = T_{\tilde{G}}[\mu^{-1}[\{\psi_1 \circ \psi_2(q)\}] \cap \Theta_{\leq\otimes\geq}] = H''$$

due to Lemma 5.4.21 (b). □

5.4.24 Lemma (Analogue to Lemma 2.15 from Mattner and Mattner, 2013) *Let* $q \in \mathrm{prob}(\{0,1\}^2)$,

$$T_{\tilde{H}}\colon \Theta \to \mathbb{R}, \quad (\pi, \chi) \mapsto \chi''_{0|0},$$

and $\tilde{H}_R := T_{\tilde{H}}[\mu^{-1}[\{q\}] \cap \Theta_R]$ *for binary relations* R *on* $\mathbb{R}^2$*. Then*

(a) $\tilde{H}_{\leq\otimes\leq} = \tilde{H}'$*, where*

$$\tilde{H}' := \left[\frac{q_{1,0}}{q_{1+} + q_{0,1}} \wedge \frac{(q_{1,0} - q_{0,1})^+}{q_{1+} - q_{0,1}}, 1\right] \quad \textit{with } \frac{0}{0} := 0$$

(b) $\tilde{H}_{\leq\otimes\geq} = \left\{ \begin{array}{cc} [0,1] & \textit{if } q_{0,1} \leq q_{1,0} \\ \emptyset & \textit{otherwise} \end{array} \right\} =: \tilde{H}''$

Proof. (a) Since $T_{\tilde{H}} = T_H \circ \Phi_1$, we have

$$\tilde{H}_{\leq\otimes\leq} = T_H[\Phi_1[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\leq}]] = T_H[\mu^{-1}[\{\psi_1(q)\}] \cap \Theta_{\leq\otimes\leq}] = \tilde{H}'$$

due to Lemma 5.4.23 (a).

(b) Since $T_{\tilde{H}} = T_G \circ \Phi_1 \circ \Phi_2$, we have

$$\tilde{H}_{\leq\otimes\geq} = T_G[\Phi_1[\Phi_2[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\geq}]]] = T_G[\mu^{-1}[\{\psi_1 \circ \psi_2(q)\}] \cap \Theta_{\leq\otimes\geq}] = \tilde{H}''$$

due to Lemma 5.4.20 (b). □

5.4.25 Lemma (Analogue to Lemma 2.16 from Mattner and Mattner, 2013) *Let* $q \in \mathrm{prob}(\{0,1\}^2)$,

$$T_I \colon \Theta \to \mathbb{R}, \quad (\pi, \chi) \mapsto \pi_1,$$

and $I_R := T_I[\mu^{-1}[\{q\}] \cap \Theta_R]$ *for binary relations* R *on* $\mathbb{R}^2$. *Then*

(a) $I_{\leq\otimes\leq} = [(q_{0,1} - q_{1,0})^+, 1 - (q_{1,0} - q_{0,1})^+] =: I'$

(b) $I_{\leq\otimes\geq} = \left\{ \begin{array}{cc} \left[\frac{1}{2} - \left|q_{1,0} - q_{0,1} - \frac{1}{2}\right|, \frac{1}{2} + \left|q_{1,0} - q_{0,1} - \frac{1}{2}\right|\right] & \textit{if } q_{0,1} \leq q_{1,0} \\ \emptyset & \textit{otherwise} \end{array} \right\} =: I''$

Proof. (a) Lemma 2.16 from Mattner and Mattner (2013) implies

$$\begin{aligned} I_{\leq\otimes\leq} &= T_I[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\mathbb{R}^2} \cap \Phi_1[\Theta_{\leq\otimes\mathbb{R}^2}]] \\ &\subseteq T_I[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\mathbb{R}^2}] \cap T_I[\Phi_1[\mu^{-1}[\{\psi_1(q)\}] \cap \Theta_{\leq\otimes\mathbb{R}^2}]] \\ &= [(q_{0,1} - q_{1,0})^+, \max\{1 - (q_{1,0} - q_{0,1})^+, q_{1,0} - q_{0,1}\}] \\ &\quad \cap [\min\{(q_{0,1} - q_{1,0})^+, 1 - q_{0,1} + q_{1,0}\}, 1 - (q_{1,0} - q_{0,1})^+] \\ &= [(q_{0,1} - q_{1,0})^+, 1 - (q_{1,0} - q_{0,1})^+] \\ &= I'. \end{aligned}$$

Let now $\mathrm{Pr} \in I'$. If $q_{0,1} \leq q_{1,0}$, then $\mathrm{Pr} \in I'$ implies $q_{0,1} - q_{1,0} + 1 - \mathrm{Pr} \geq 0$, so that $\Delta\mathrm{Se} := 0$ fulfills (66). If $q_{0,1} > q_{1,0}$, then $\mathrm{Pr} \in I'$ implies $\Delta\mathrm{Se} := (q_{0,1} - q_{1,0})/\mathrm{Pr} \in \,]0,1]$ and (66). In both cases Lemma 5.4.12 (a) yields $(\mathrm{Pr}, \Delta\mathrm{Se}) \in C_{\leq\otimes\leq}$ and thus $\mathrm{Pr} \in I_{\leq\otimes\leq}$.

(b) If $q_{0,1} > q_{1,0}$, then Lemma 5.4.7 implies $I_{\leq\otimes\geq} = \emptyset = I''$. Suppose therefore $q_{0,1} \leq q_{1,0}$ from now on.

Lemma 2.16 from Mattner and Mattner (2013) implies

$$\begin{aligned} I_{\leq\otimes\geq} &= T_I[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\mathbb{R}^2} \cap \Phi_1[\Phi_2[\Theta_{\leq\otimes\mathbb{R}^2}]]] \\ &\subseteq T_I[\mu^{-1}[\{q\}] \cap \Theta_{\leq\otimes\mathbb{R}^2}] \cap T_I[\Phi_1[\Phi_2[\mu^{-1}[\{\psi_1 \circ \psi_2(q)\}] \cap \Theta_{\leq\otimes\mathbb{R}^2}]]] \\ &= [0, \max\{1 - q_{1,0} + q_{0,1}, q_{1,0} - q_{0,1}\}] \cap [\min\{q_{1,0} - q_{0,1}, 1 - q_{1,0} + q_{0,1}\}, 1] \\ &= [\min\{q_{1,0} - q_{0,1}, 1 - q_{1,0} + q_{0,1}\}, \max\{q_{1,0} - q_{0,1}, 1 - q_{1,0} + q_{0,1}\}] \end{aligned}$$

$$= I''.$$

Let now $\Pr \in I''$. If $q_{0,1} + 1/2 \geq q_{1,0}$, then $\Pr \in I'' = [q_{1,0} - q_{0,1}, 1 - q_{1,0} + q_{0,1}]$ implies $\Delta\mathrm{Se} \coloneqq (q_{0,1} - q_{1,0})/\Pr \in [-1, 0[$ and (66′). If $q_{0,1} + 1/2 < q_{1,0}$, then $\Pr \in I'' = [1 - q_{1,0} + q_{0,1}, q_{1,0} - q_{0,1}]$ implies $\Delta\mathrm{Se} \coloneqq (q_{0,1} - q_{1,0} + 1)/\Pr - 1 \in [-1, 0]$ and (66′). In both cases Lemma 5.4.12 (b) yields $(\Pr, \Delta\mathrm{Se}) \in C_{\leq\otimes\geq}$ and thus $\Pr \in I_{\leq\otimes\geq}$. □

A

BASIC NOTIONS

- This appendix defines some frequently used notions and terminology, and recapitulates some fundamental results. It does not introduce the concepts gently or with much explanation, nor does it always provide proofs for the results (sources containing proofs are, however, mentioned in the remarks following them).
- Outline of this appendix:
  - Section A.1 deals with cartesian products, functions and families, relations (mostly preorders), and their connections.
  - Section A.2 presents a counterexample taken from Rudin (1986) employing ordinal numbers.
  - Section A.3 introduces fundamental statistical concepts and results, such as models, confidence regions, tests, and the duality between them, P-variables, stochastic monotonicity, and monotone likelihood ratios.

A.1 FUNCTIONS AND RELATIONS

A.1.1 Definition The **power set** of a set A is denoted by 2^A.

A.1.2 Definition Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be three sets. Given elements $x \in \mathcal{X}$, $y \in \mathcal{Y}$, and $z \in \mathcal{Z}$ we define the **pair** with components x and y and the **triplet** with components x, y, and z as

$$(x, y) \coloneqq \{x, \{x, y\}\} \qquad \text{and} \qquad (x, y, z) \coloneqq (x, (y, z)),$$

respectively. The set of all pairs (x, y) with $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ constitutes the **cartesian product** of $\mathcal{X}$ and $\mathcal{Y}$, denoted by

$$\mathcal{X} \times \mathcal{Y} \coloneqq \{(x, y) \in 2^{\mathcal{X} \cup 2^{\mathcal{X} \cup \mathcal{Y}}} : x \in \mathcal{X},\ y \in \mathcal{Y}\}.$$

A.1.3 Remark If $\mathcal{X}$ or $\mathcal{Y}$ is empty, then so is $\mathcal{X} \times \mathcal{Y}$ (and vice versa).

A.1.4 Definition Let $\mathcal{X}$ and I be two sets. A triplet $(I, f, \mathcal{X})$ where

$$f \subseteq I \times \mathcal{X}$$

is such that for every $i \in I$ there exists one, and only one, $x \in \mathcal{X}$ with $(i, x) \in f$ is called **function** or **mapping** (from I to $\mathcal{X}$). The sets I and $\mathcal{X}$ are usually referred to as **domain** and **codomain,** respectively, of the function.

A.1.5 Remark 1. We denote a function $(I, f, \mathcal{X})$ more commonly by

$$f \colon I \to \mathcal{X}$$

or, more succinctly, by f whenever I and $\mathcal{X}$ are fixed, clear from the setting, or simply irrelevant.

2. It is customary to write $f(i) = x$ instead of $(i, x) \in f$.
3. Writing $i \mapsto f(i)$ means that f is defined by mapping each i from its domain to $f(i)$.
4. It is common to omit the name of a function when defining it (especially if it is not relevant or if the function does not occur later on), as in

$$I \to \mathcal{X}, \quad x \mapsto f(x).$$

5. One way of specifying a function without giving a specific definition or name is to write, e.g., $\mathbb{R} \to \mathbb{C}$. This stands for any element of the set $\mathbb{C}^{\mathbb{R}}$ (see Definition A.1.17).
6. A usual way of defining a function is thus, e.g.,

$$f\colon \mathbb{R} \to \mathbb{C}, \quad x \mapsto \exp(\mathrm{i}x).$$

7. In the above definition, $I = \emptyset$ is allowed, in which case $f = \emptyset$ is named the **empty function.** The case $\mathcal{X} = \emptyset$ can only occur if $I = \emptyset$.

A.1.6 Remark 1. Functions are often regarded as "black boxes," producing an output based on some kind of input. Whenever a function $f\colon I \to \mathcal{X}$ is interpreted as a means to index objects, it is rather called **family.** In such situations, we often write f_i instead of $f(i)$ and denote the family f then by

$$(f_i\colon i \in I).$$

We occasionally call a finite family a **tuple,** and a finite family whose members are numbers a **vector.**

2. When the codomain of a family is a set of functions we often use a stylized letter like $\mathcal{F}$ to denote the family and f_i (instead of $\mathcal{F}_i$) for its values. The reason behind this is the conventional usage of the minuscule f for a family's functions and stylized versions of this letter for families, sets, or classes of functions, like $\mathcal{F} = (f_i\colon i \in I)$.
3. Families and sets are sometimes used interchangeably, often without doing any real harm:
   - given a family $f\colon I \to \mathcal{X}$, its **range** $\{f_i\colon i \in I\}$ is a set encompassing all members of f;
   - given a set $\mathcal{X}$, the **identity function** $\mathrm{id}_{\mathcal{X}}\colon \mathcal{X} \to \mathcal{X}$, $x \mapsto x$, is a family whose members coincide with those of $\mathcal{X}$.

A.1.7 Definition Let $f\colon \mathcal{X} \to \mathcal{Y}$ be a function and $A \subseteq \mathcal{X}$. The function

$$f|_A\colon A \to \mathcal{Y}, \quad x \mapsto f(x),$$

is called **restriction** of f to A.

A.1.8 Definition A function $f\colon \mathcal{X} \to \mathcal{Y}$ is called

- **injective** (or **one-to-one**), if $f(x_1) = f(x_2)$ implies $x_1 = x_2$ for $x_1, x_2 \in \mathcal{X}$;
- **surjective** (or **onto**), if for every $y \in \mathcal{Y}$ there is an $x \in \mathcal{X}$ with $f(x) = y$;
- **bijective,** if it is both injective and surjective.

A.1.9 Remark Let $f\colon \mathcal{X} \to \mathcal{Y}$ be a function.

1. f is bijective if, and only if, to each $y \in \mathcal{Y}$ corresponds exactly one $x \in \mathcal{X}$ with $f(x) = y$. This unique x is denoted by $f^{-1}(y)$, and the thus defined function $f^{-1}\colon \mathcal{Y} \to \mathcal{X}$ is called **inverse** of f.

2. The mere mentioning of the inverse f^{-1} entails the claim that f is bijective.
3. If $f = f^{-1}$, then f is called an **involution.**
4. Injective functions can be made bijective by narrowing down their codomain to their range: if f is injective, then $\mathcal{X} \to f[\mathcal{X}]$, $x \mapsto f(x)$, is bijective.

A.1.10 Definition If p is a logical proposition, then

$$\mathbf{1}(\mathsf{p}) := \begin{cases} 1 & \text{if } \mathsf{p} \text{ is true} \\ 0 & \text{otherwise} \end{cases}$$

is called **indicator** of the proposition p. A close companion is the **indicator function** of a subset A of a given set $\mathcal{X}$:

$$\mathbf{1}_A \colon \mathcal{X} \to \{0, 1\}, \quad x \mapsto \mathbf{1}(x \in A).$$

A.1.11 Remark An indicator function of a set $A \subseteq \mathcal{X}$ is

- injective if, and only if, it is the empty function or $|A| + 1 = |\mathcal{X}| \leq 2$,
- surjective if, and only if, $A \notin \{\emptyset, \mathcal{X}\}$.

A.1.12 Definition Let $f \colon \mathcal{X} \to \mathcal{Y}$ and $g \colon \mathcal{Y} \to \mathcal{Z}$ be two functions. Their **composition** $g \circ f$ is then given by

$$g \circ f \colon \mathcal{X} \to \mathcal{Z}, \quad x \mapsto g(f(x)).$$

A.1.13 Remark In some situations it is customary to write the composition $g \circ f$ of two functions as $g(f)$. If, e.g., $g \colon \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$, and $f := \cos$, then $g \circ f$ is usually expressed as $\cos^2$. This covers also the case of, e.g., $\kappa[R]$, occurring in Theorem A.3.21, or $\bigwedge[R_{\bigwedge X/m,\beta}]$, occurring in Lemma 3.1.5. The value of the latter function at x does, by the way, not stand for the least element of $R_{\bigwedge X/m,\beta}(x)$, but for the set $\{\bigwedge_{i=1}^n p_i \in [0,1] \colon p \in R_{\bigwedge X/m,\beta}(x)\}$.

A.1.14 Definition Let $f \colon \mathcal{X} \to \mathcal{Y}$ be a function. The functions

$$\begin{aligned} f[\cdot] \colon 2^{\mathcal{X}} \to 2^{\mathcal{Y}}, &\quad A \mapsto \{f(x) \in \mathcal{Y} \colon x \in A\}, \\ f^{-1}[\cdot] \colon 2^{\mathcal{Y}} \to 2^{\mathcal{X}}, &\quad B \mapsto \{x \in \mathcal{X} \colon f(x) \in B\}, \end{aligned}$$

are called **image** and **preimage functions,** respectively, of f. Instead of $f[\cdot](A)$ and $f^{-1}[\cdot](B)$ we write $f[A]$ and $f^{-1}[B]$, respectively. In order to contrast f from its image function $f[\cdot]$, f is sometimes written as $f(\cdot)$.

A.1.15 Remark

1. We have

$$f[f^{-1}[B]] \subseteq B \quad \text{and} \quad f^{-1}[f[A]] \supseteq A \qquad \text{for } A \in 2^{\mathcal{X}} \text{ and } B \in 2^{\mathcal{Y}},$$

with equality if f is surjective or injective, respectively. Conversely, if there is equality for all $B \in \{\{y\} \colon y \in \mathcal{Y}\}$ or for all $A \in \{\{x\} \colon x \in \mathcal{X}\}$, then f is surjective or injective, respectively.

2. The image function of the preimage function of f

$$f^{-1}[[\cdot]] := (f^{-1}[\cdot])[\cdot] \colon 2^{2^{\mathcal{Y}}} \to 2^{2^{\mathcal{X}}}, \quad \mathfrak{B} \mapsto \{f^{-1}[B] \in 2^{\mathcal{X}} \colon B \in \mathfrak{B}\},$$

and the preimage function of the image function of f

$$[f^{-1}[\cdot]] := (f[\cdot])^{-1}[\cdot] \colon 2^{2^{\mathcal{Y}}} \to 2^{2^{\mathcal{X}}}, \quad \mathfrak{B} \mapsto \{A \in 2^{\mathcal{X}} \colon f[A] \in \mathfrak{B}\},$$

are in general different functions. In fact, $f\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$, and $\mathfrak{B} := \{\{1\}\}$ yield $f^{-1}[[\mathfrak{B}]] = \{\{-1,1\}\}$ and $[f^{-1}[\mathfrak{B}]] = \{\{-1\},\{1\},\{-1,1\}\}$. More precisely, this remark's first part implies

$$f^{-1}[[\cdot]] = [f^{-1}[\cdot]] \qquad \iff \qquad f \text{ is bijective.}$$

3. The image function of the image function of f

$$f[[\cdot]] := (f[\cdot])[\cdot]\colon 2^{2^{\mathcal{X}}} \to 2^{2^{\mathcal{Y}}}, \quad \mathfrak{A} \mapsto \{f[A] \in 2^{\mathcal{Y}}\colon A \in \mathfrak{A}\},$$

and the preimage function of the preimage function of f

$$[f[\cdot]] := (f^{-1}[\cdot])^{-1}[\cdot]\colon 2^{2^{\mathcal{X}}} \to 2^{2^{\mathcal{Y}}}, \quad \mathfrak{A} \mapsto \{B \in 2^{\mathcal{Y}}\colon f^{-1}[B] \in \mathfrak{A}\},$$

are in general different functions, too. In fact, the above f and $\mathfrak{A} := \{\{-1\}\}$ yield $f[[\mathfrak{A}]] = \{\{1\}\}$ and $[f[\mathfrak{A}]] = \emptyset$. More precisely, this remark's first part implies

$$f[[\cdot]] = [f[\cdot]] \qquad \iff \qquad f \text{ is bijective.}$$

4. For two functions $f, g\colon \mathcal{X} \to \mathcal{Y}$ the following equivalences hold:

$$\begin{array}{rcccc} f = g & \iff & f[\cdot] = g[\cdot] & \iff & f^{-1}[\cdot] = g^{-1}[\cdot] \\ f \text{ injective} & \iff & f[\cdot] \text{ injective} & \iff & f^{-1}[\cdot] \circ f[\cdot] = \mathrm{id}_{2^{\mathcal{X}}} \\ f \text{ surjective} & \iff & f[\cdot] \text{ surjective} & \iff & f[\cdot] \circ f^{-1}[\cdot] = \mathrm{id}_{2^{\mathcal{Y}}}. \end{array}$$

5. For two functions $f\colon \mathcal{X} \to \mathcal{Y}$ and $g\colon \mathcal{Y} \to \mathcal{Z}$ we have

$$(g \circ f)[\cdot] = g[\cdot] \circ f[\cdot] \qquad \text{and} \qquad (g \circ f)^{-1}[\cdot] = f^{-1}[\cdot] \circ g^{-1}[\cdot].$$

A.1.16 Definition If $f\colon \mathcal{X} \to \mathcal{Y}$ is a function and $\mathsf{p}(y)$ is a logical assertion whose truth depends on $y \in \mathcal{Y}$, we often write

$$\{\mathsf{p}(f)\} := \{x \in \mathcal{X}\colon \mathsf{p}(f(x))\}$$

and omit the curly braces if $\{\mathsf{p}(f)\}$ is the argument of a function. For instance, if $B \in 2^{\mathcal{Y}}$, then $\{f \in B\} = f^{-1}[B]$.

A.1.17 Definition The **cartesian product** of a family $(\mathcal{X}_i\colon i \in I)$ of sets is

$$\prod_{i\in I} \mathcal{X}_i := \left\{ x\colon I \to \bigcup_{i\in I} \mathcal{X}_i\colon x_i \in \mathcal{X}_i \text{ for } i \in I \right\}.$$

If $\mathcal{X}_i = \mathcal{X}$ for all $i \in I$, we write $\mathcal{X}^I$ instead of $\prod_{i\in I} \mathcal{X}$. If the index set I is finite, say $I = \{1, \ldots, n\}$, then we write $\prod_{i=1}^n \mathcal{X}_i$, or $\mathcal{X}^n$ if all $\mathcal{X}_i$ are equal to, say, $\mathcal{X}$, for its cartesian product.

A.1.18 Remark 1. Identifying pairs with their canonical functional representatives via

$$\mathcal{X} \times \mathcal{Y} \to (\mathcal{X} \cup \mathcal{Y})^{\{1,2\}}, \qquad (x,y) \mapsto \big(\{1,2\} \to \mathcal{X} \cup \mathcal{Y}, \quad 1 \mapsto x,\ 2 \mapsto y\big)$$

makes this definition of cartesian products compatible with the one given for a family of length two in Definition A.1.2. Let us note that the above expression to the right is not to be read as a triplet, but as the definition of a mapping (which, formally, is a triplet anyway).

2. We therefore also define $\mathcal{X}^n := \mathcal{X}^{\{1,\dots,n\}}$ for $n \in \mathbb{N}$, and $\mathcal{X}^0 := \mathcal{X}^\emptyset = \{\emptyset\}$.
3. The **axiom of choice** states that arbitrary cartesian products of non-empty sets are non-empty.

A.1.19 Definition Let $\mathcal{X}$ be a set and $n \in \mathbb{N}$. Any subset of $\mathcal{X}^n$ is called **n-ary relation** on $\mathcal{X}$ and n its **order.** Relations of order 2 and 3 are commonly called **binary** and **ternary,** respectively.

A.1.20 Remark 1. The term "relation," without further specification of an order, denotes in the following always one of order 2, i.e., a binary relation.
2. Given a relation R, it is customary to write $x\,R\,y$ rather than $(x, y) \in R$ in order to stress the existence of a relationship between x and y.

A.1.21 Example Examples for relations on a set $\mathcal{X}$ are 1. $\mathcal{X}^2$, the all-relation, 2. $\emptyset$, the empty relation, and 3. $\Delta_\mathcal{X} := \{(x, x) \in \mathcal{X}^2 : x \in \mathcal{X}\}$, the diagonal of $\mathcal{X}$.

A.1.22 Definition 1. If R is a relation on a set $\mathcal{X}$ and $\mathcal{X}_0 \subseteq \mathcal{X}$ is a subset, then the relation

$$R|_{\mathcal{X}_0} := R \cap \mathcal{X}_0^2$$

on $\mathcal{X}_0$ is called **induced** by R on $\mathcal{X}_0$.
2. If R is a relation on a set $\mathcal{X}$, then the relation

$$R^{\mathrm{op}} := \reflectbox{R} := \{(x, y) \in \mathcal{X}^2 : y\,R\,x\}$$

on $\mathcal{X}$ is called **dual relation** of R.
3. If $(R_i : i \in I)$ is a family of relations R_i on sets $\mathcal{X}_i$, then the relation

$$\bigotimes_{i \in I} R_i := \left\{ (x, y) \in \left(\prod_{i \in I} \mathcal{X}_i \right)^2 : x_i\,R_i\,y_i \text{ for } i \in I \right\}$$

on $\prod_{i \in I} \mathcal{X}_i$ is called **product relation** of $(R_i : i \in I)$. For finite families $(R_1, \dots, R_n)$ we also employ the notation $R_1 \otimes \ldots \otimes R_n$. If all relations R_i are equal, say R, we write $R^{\otimes I}$ and, in the case of a finite family of length n, $R^{\otimes n}$ for the product relation.
4. If $(R_1, \dots, R_n)$ is a finite family of relations on a set $\mathcal{X}$, then the relation

$$R_1 \cdot \ldots \cdot R_n := \{(x_0, x_n) \in \mathcal{X}^2 : \text{there are } x_1, \dots, x_{n-1} \in \mathcal{X} \text{ with } x_{i-1}\,R_i\,x_i \text{ for } i \in \{1, \dots, n\}\}$$

is called their **relational product.** If all relations R_i are equal, say R, we write $R^{\cdot n}$ for their relational product.

A.1.23 Remark 1. Subsets and cartesian products of sets endowed with a relation shall in the following always be endowed with the induced or product relation, respectively.
2. The mapping $R \mapsto R^{\mathrm{op}}$ is an involution (on the set of all relations on a fixed set), i.e., the dual of the dual relation yields the original relation (hence the practice of denoting duals by reflecting the symbol that designates the original relation).
3. If R is a relation on a set $\mathcal{X}$ and $\mathcal{X}_0 \subseteq \mathcal{X}$ a subset, then $R^{\mathrm{op}}|_{\mathcal{X}_0} = (R|_{\mathcal{X}_0})^{\mathrm{op}} =: R|_{\mathcal{X}_0}^{\mathrm{op}}$.

A.1.24 Definition Let $\mathcal{X}$ be a set and $\Delta_\mathcal{X} := \{(x, x) \in \mathcal{X}^2 : x \in \mathcal{X}\}$ its diagonal. A relation R on $\mathcal{X}$ is called **reflexive** if $\Delta_\mathcal{X} \subseteq R$, **irreflexive** if $\Delta_\mathcal{X} \cap R = \emptyset$, **symmetric** if $R \subseteq R^{\mathrm{op}}$, **antisymmetric** if $R \cap R^{\mathrm{op}} \subseteq \Delta_\mathcal{X}$, **transitive** if $R^{\cdot 2} \subseteq R$, and **total** if $R \cup R^{\mathrm{op}} \cup \Delta_\mathcal{X} = \mathcal{X}^2$.

A.1.25 Remark 1. Each of the above properties is passed on to induced and dual relations, and each one but totality is passed on to products. The latter means: if $(R_i\colon i \in I)$ is a family of relations R_i on sets $\mathcal{X}_i$, each having one and the same of the above properties except totality, then the product $\bigotimes_{i\in I} R_i$ also possesses this property.

2. Let $(R_i\colon i \in I)$ be a family of relations on a set $\mathcal{X}$. Then $\bigcap_{i\in I} R_i$ is the greatest and $\bigcup_{i\in I} R_i$ the least (see Definition A.1.37) relation on $\mathcal{X}$ contained in or containing, respectively, every R_i. Furthermore, $(\bigcup_{i\in I} R_i)^{\mathrm{op}} = \bigcup_{i\in I} R_i^{\mathrm{op}}$ and $(\bigcap_{i\in I} R_i)^{\mathrm{op}} = \bigcap_{i\in I} R_i^{\mathrm{op}}$, as well as $(\bigcup_{i\in I} R_i)^{\cdot 2} = \bigcup_{i,j\in I} R_i \cdot R_j$ and $(\bigcap_{i\in I} R_i)^{\cdot 2} \subseteq \bigcap_{i\in I} R_i^{\cdot 2}$. This yields the implications

$$R_i \text{ reflexive for } i \in I \implies \begin{cases} \bigcap_{i\in I} R_i \text{ reflexive} \\ \bigcup_{i\in I} R_i \text{ reflexive if } I \neq \emptyset \end{cases}$$

$$R_i \text{ irreflexive for } i \in I \implies \begin{cases} \bigcap_{i\in I} R_i \text{ irreflexive if } I \neq \emptyset \\ \bigcup_{i\in I} R_i \text{ irreflexive} \end{cases}$$

$$R_i \text{ symmetric for } i \in I \implies \bigcap_{i\in I} R_i \text{ and } \bigcup_{i\in I} R_i \text{ symmetric}$$

$$R_i \text{ antisymmetric for } i \in I \implies \bigcap_{i\in I} R_i \text{ and } \bigcup_{i\in I} R_i \text{ antisymmetric}$$

$$R_i \text{ transitive for } i \in I \implies \bigcap_{i\in I} R_i \text{ transitive}$$

$$R_i \text{ total for } i \in I \implies \begin{cases} \bigcap_{i\in I} R_i \text{ total} \\ \bigcup_{i\in I} R_i \text{ total if } I \neq \emptyset. \end{cases}$$

3. Given a relation R on a set $\mathcal{X}$, we can thus define 1. the **reflexive hull** $R^{\mathrm{refl}} := \bigcap\{S \subseteq \mathcal{X}^2\colon S \supseteq R \text{ reflexive}\} = R \cup \Delta_{\mathcal{X}}$ of R, the smallest reflexive relation containing R, 2. the **irreflexive core** $R^{\mathrm{irrefl}} := \bigcup\{S \subseteq R\colon S \text{ irreflexive}\} = R \setminus \Delta_{\mathcal{X}}$, the greatest irreflexive relation contained in R, 3. the **symmetric hull** $R^{\mathrm{symm}} := \bigcap\{S \subseteq \mathcal{X}^2\colon S \supseteq R \text{ symmetric}\} = R \cup R^{\mathrm{op}}$, the smallest symmetric relation containing R, 4. the **transitive hull** $R^{\mathrm{trans}} := \bigcap\{S \subseteq \mathcal{X}^2\colon S \supseteq R \text{ transitive}\} = \bigcup_{n\in\mathbb{N}} R^{\cdot n}$, the smallest transitive relation containing R, and the **total hull** $R^{\mathrm{tot}} := \bigcap\{S \subseteq \mathcal{X}^2\colon S \supseteq R \text{ total}\}$, the smallest total relation containing R.

A.1.26 Definition 1. A reflexive, symmetric, and transitive relation is called an **equivalence relation;** such relations are often denoted by $\sim$ or $\equiv$.

2. A reflexive and transitive relation is called a **preorder;** preorders are often denoted by symbols like $\leq$ or $\preceq$.
3. An antisymmetric preorder is called an **order.**
4. An irreflexive and transitive relation is called a **strict order;** strict orders are most often denoted by $<$ or $\prec$.
5. The pair $(\mathcal{X}, \leq)$ is called a **preordered** or **ordered set** if $\leq$ is a preorder or order, respectively, on $\mathcal{X}$.
6. The pair $(\mathcal{X}, <)$ is called **strictly ordered set** if $<$ is a strict order on $\mathcal{X}$.

A.1.27 Remark 1. Subsets of $\mathbb{R}$ are in the following always endowed with the usual order, unless stated otherwise.

2. The product of a single preorder or of a single strict order shall be denoted by the same symbol as the preorder or strict order, respectively, if no confusion arises (see Remark A.1.30 for an example of a confusing situation).

A.1.28 Example Let $\mathcal{X}$ be a set. The diagonal $\Delta_{\mathcal{X}}$ is the smallest reflexive relation on $\mathcal{X}$. It is moreover a symmetric order. The smallest irreflexive relation on $\mathcal{X}$ is $\emptyset$. It is also a symmetric and antisymmetric strict order.

A.1.29 Theorem *Let $\mathcal{X}$ be a set and $\Delta_{\mathcal{X}} \coloneqq \{(x,x) \in \mathcal{X}^2 : x \in \mathcal{X}\}$ its diagonal.*

(i) *Let $\leq$ be an order on $\mathcal{X}$. Then*

$$< \coloneqq \leq^{\mathrm{str}} \coloneqq \leq \setminus \Delta_{\mathcal{X}}$$

defines a strict order on $\mathcal{X}$.

(ii) *Let $<$ be a strict order on $\mathcal{X}$. Then*

$$\leq \coloneqq <^{\mathrm{unstr}} \coloneqq < \cup \Delta_{\mathcal{X}}$$

defines an order on $\mathcal{X}$.

(iii) *Let $\leq$ be a preorder on $\mathcal{X}$. Then*

$$< \coloneqq \leq^{\mathrm{str}} \coloneqq \leq \setminus \leq^{\mathrm{op}}$$

defines a strict order on $\mathcal{X}$.

Proof.

(i) 1. Irreflexivity: This follows at once from the definition of $<$.
2. Transitivity: Let $x, y, z \in \mathcal{X}$ with $x < y$ and $y < z$. The $x \leq y$ and $y \leq z$, which yields $x \leq z$ due to the transitivity of $\leq$. Furthermore, $x \neq z$ since, otherwise, $x \leq y$ and $y \leq z = x$ together with the antisymmetry of $\leq$ imply $x = y$, a contradiction to $x < y$. This yields $x < z$.

(ii) 1. Reflexivity: This follows at once from the definition of $\leq$.
2. Antisymmetry: Let $x, y \in \mathcal{X}$ with $x \leq y$ and $y \leq x$. If $x \neq y$, then $x < y$ and $y < x$, and the transitivity of $<$ implies $x < x$, a contradiction to the irreflexivity of y.
3. Transitivity: Let $x, y, z \in \mathcal{X}$ with $x \leq y$ and $y \leq z$. In case $x = y = z$ we obtain $x \leq z$ trivially, while in all other cases we obtain $x < z$, thus also $x \leq z$.

(iii) 1. Irreflexivity: Let $x \in \mathcal{X}$. We have $x < x$ if, and only if, both $x \leq x$ and its negation hold. Thus, $x \not< x$.
2. Transitivity: Let $x < y$ and $y < z$. This means $x \leq y$, $y \not\leq x$, $y \leq z$, and $z \not\leq y$. We have to show $x \leq z$ and $z \not\leq x$. From the transitivity of $\leq$ follows already $x \leq z$. If $z \leq x$ were valid, then the transitivity of $\leq$ would imply $z \leq y$, a contradiction to $z \not\leq y$. □

A.1.30 Remark Let $\leq$ be an order and $<$ a strict order.

1. $\leq \setminus \leq^{\mathrm{op}} = \leq \setminus \Delta_{\mathcal{X}}$, implying that part (iii) generalizes part (i). In fact, $\leq \cap \leq^{\mathrm{op}} = \Delta_{\mathcal{X}}$ due to the antisymmetry of $\leq$.
2. $\leq^{\mathrm{str}} \subset \leq$ and $<^{\mathrm{unstr}} \supset <$, and $(\leq^{\mathrm{str}})^{\mathrm{unstr}} = \leq$ and $(<^{\mathrm{unstr}})^{\mathrm{str}} = <$.
3. $(\leq^{\mathrm{op}})^{\mathrm{str}} = (\leq^{\mathrm{str}})^{\mathrm{op}}$ and $(<^{\mathrm{op}})^{\mathrm{unstr}} = (<^{\mathrm{unstr}})^{\mathrm{op}}$.
4. If I is a set, then $(\leq^{\otimes I})^{\mathrm{str}} \supseteq (\leq^{\mathrm{str}})^{\otimes I}$ and $(<^{\otimes I})^{\mathrm{unstr}} \subseteq (<^{\mathrm{unstr}})^{\otimes I}$. Both inclusions are in general strict.

A.1.31 Remark Given an order $\leq$, then $< := \leq^{\text{str}}$ if the symbol $<$ is unspecified. Similarly, given a strict order $<$, then $\leq := <^{\text{unstr}}$ if $\leq$ is unspecified.

A.1.32 Example 1. $\leq^{\text{op}} = \geq$ and $<^{\text{op}} = >$, considered as relations on $\mathbb{R}$.
2. $\subseteq^{\text{op}} = \supseteq$, considered as relations on some set of sets.

A.1.33 Definition Let $(\mathcal{X}, \leq)$ be a preordered set. A set $D \subseteq \mathcal{X}$ is called **downray** if

$$x \in D, \quad y \in \mathcal{X}, \quad y \leq x \qquad \Longrightarrow \qquad y \in D, \tag{75}$$

and **upray** if (75) holds with $\leq$ replaced by its dual $\geq$. Arbitrary intersections of up- with downrays are called **intervals.**

A.1.34 Remark 1. $D \subseteq \mathcal{X}$ is a downray if, and only if, $D \supseteq \bigcup_{x\in D}\{\mathrm{id}_{\mathcal{X}} \leq x\}$.
2. Due to the order's reflexivity we have $D \subseteq \bigcup_{x\in D}\{\mathrm{id}_{\mathcal{X}} \leq x\}$ for every $D \subseteq \mathcal{X}$.
3. $D \subseteq \mathcal{X}$ is a downray or upray if, and only if, its indicator function $\mathbf{1}_D$ is decreasing or increasing (see Definition A.1.40), respectively.
4. Other frequently encountered names for the terms "down-" and "upray" are *lower set, decreasing set, initial segment, downward closed set* and *upper set, increasing set, upward closed set,* respectively.
5. In certain of the following results we restrict our attention to downrays since a consideration of the dual order (which amounts to replacing each occurrence of "$\leq$" with "$\geq$," and vice versa) yields analogous results for uprays.

A.1.35 Remark 1. Complements of downrays are uprays. This follows from $\mathbf{1}_{A^c} = 1 - \mathbf{1}_A$ for $A \in 2^{\mathcal{X}}$.
2. Any union or intersection of downrays remains a downray. This follows from $\mathbf{1}_{\bigcup_{i\in I} D_i} = \sup_{i\in I} \mathbf{1}_{D_i}$ and $\mathbf{1}_{\bigcap_{i\in I} D_i} = \inf_{i\in I} \mathbf{1}_{D_i}$ and part 2 of Remark A.1.34. (These sup and inf are to be understood pointwise and to be formed in $\{0,1\}$.)
3. Given a set $A \subseteq \mathcal{X}$, we therefore call

$$\lceil A \rceil := \lceil A \rceil_{\mathcal{X}} := \bigcap \{D \in 2^{\mathcal{X}} : D \text{ downray}, \; D \supseteq A\}$$

the **downray generated** by A in $\mathcal{X}$. Another common notation for $\lceil A \rceil$ is $\downarrow A$.
4. Generated downrays admit the representation

$$\lceil A \rceil = \bigcup_{x\in A} \{\mathrm{id}_{\mathcal{X}} \leq x\}$$

since $A \subseteq \bigcup_{x\in A}\{\mathrm{id}_{\mathcal{X}} \leq x\}$ is a downray and since any downray $D \subseteq \mathcal{X}$ with $D \supseteq A$ fulfills $D = \bigcup_{x\in D}\{\mathrm{id}_{\mathcal{X}} \leq x\} \supseteq \bigcup_{x\in A}\{\mathrm{id}_{\mathcal{X}} \leq x\}$.
5. Let us note that

$$2^{\mathcal{X}} \to 2^{\mathcal{X}}, \quad A \mapsto \lceil A \rceil,$$

is increasing (see Definition A.1.40) with respect to $\subseteq$ and preserves arbitrary unions due to the preceding remark (the latter meaning $\lceil \bigcup_{i\in I} A_i \rceil = \bigcup_{i\in I} \lceil A_i \rceil$ for families $(A_i : i \in I)$ of subsets of $\mathcal{X}$). It does not preserve even finite intersections (for instance, $\mathcal{X} := [0,1]$, $A := \{1\}$, and $B := \{1/2\}$ yield $\lceil A \cap B \rceil = \emptyset \neq \lceil A \rceil \cap \lceil B \rceil = [0, 1/2]$), but we have $\lceil \bigcap_{i\in I} A_i \rceil \subseteq \bigcap_{i\in I} \lceil A_i \rceil$ for families $(A_i : i \in I)$ of subsets of $\mathcal{X}$.
6. Generated downrays make apparent that rays are not necessarily totally ordered sets. If we endow, e.g., $\mathcal{X} := \{0,1\}^2$ with the product order, then $\mathcal{X} \setminus \{(1,1)\} = \lceil \{(1,0),(0,1)\} \rceil$ is not totally ordered.

7. Not every downray admits a countable generator, as shows the example of $\mathcal{X} \coloneqq D \coloneqq \mathbb{R}$ endowed with the order $=$. Assuming even a total order does not remedy this deficit: the least uncountable ordinal ω_1 (see Section A.2 for details on ordinal numbers), which is a well-ordered (hence, totally ordered) set, possesses no countable generator since all downrays in ω_1 are ordinals strictly less than ω_1 and, as such, countable.

A.1.36 Lemma *Let $(\mathcal{X}, \leq)$ be a preordered set and $\mathcal{X}_0 \subseteq \mathcal{X}$. Then the downrays in $\mathcal{X}_0$ are precisely the intersections of $\mathcal{X}_0$ with downrays in $\mathcal{X}$.*

Proof. If D_0 is a downray in $\mathcal{X}_0$, then considering the downray $D \coloneqq \lceil D_0 \rceil_{\mathcal{X}}$ in $\mathcal{X}$ yields $D_0 = D \cap \mathcal{X}_0$. If, conversely, D is a downray in $\mathcal{X}$, then $D_0 \coloneqq D \cap \mathcal{X}_0$ is a downray in $\mathcal{X}_0$ since

$$\bigcup_{x \in D_0} \{\mathrm{id}_{\mathcal{X}_0} \leq x\} = \mathcal{X}_0 \cap \bigcup_{x \in D_0} \{\mathrm{id}_{\mathcal{X}} \leq x\} \subseteq \mathcal{X}_0 \cap \bigcup_{x \in D} \{\mathrm{id}_{\mathcal{X}} \leq x\} \subseteq D_0$$ □

A.1.37 Definition Let $(\mathcal{X}, \leq)$ be a preordered set and $\mathcal{X}_0 \subseteq \mathcal{X}$. An element $\xi \in \mathcal{X}$ is called 1. a **lower bound** for $\mathcal{X}_0$ if $\{\mathrm{id}_{\mathcal{X}_0} \geq \xi\} = \mathcal{X}_0$, 2. **minimal** if $\{\mathrm{id}_{\mathcal{X}} < \xi\} = \emptyset$, 3. **least** or **smallest,** and denoted by $\min \mathcal{X}$ in case of uniqueness (see the next remark), if $\{\mathrm{id}_{\mathcal{X}} \geq \xi\} = \mathcal{X}$; it is called an **upper bound** for $\mathcal{X}_0$, **maximal,** or **greatest** if it is a lower bound for $\mathcal{X}_0$, minimal, or least, respectively, with respect to the dual preorder. A **supremum** of $\mathcal{X}_0$ is a least element in the set of upper bounds for $\mathcal{X}_0$; an **infimum** is a supremum with respect to the dual preorder.

A.1.38 Remark 1. In what follows, we consider merely lower bounds, minimal, and least elements. Analogous results about the notions “upper bound,” “maximal,” and “greatest” follow by considering the dual preorder.

2. Minimal elements need not be unique, not even in ordered sets: in the ordered set $(2^{\{0,1\}} \setminus \{\emptyset\}, \subseteq)$ both $\{0\}$ and $\{1\}$ are minimal.
3. In preordered sets, least elements need not be unique either: if $\mathcal{X} \coloneqq \{0,1\}^2 \setminus \{(0,0)\}$ and $x \preceq y$ is defined to hold whenever $x_1 \leq y_1$ or $x_2 \leq y_2$ for $x, y \in \mathcal{X}$, then $(\mathcal{X}, \leq)$ is a totally preordered set in which both $(0,1)$ and $(1,0)$ are least (note that $\leq\, = \mathcal{X}^2$ since $1 \in \{x_1, x_2\} \cap \{y_1, y_2\}$ for $x, y \in \mathcal{X}$). In ordered sets, however, least elements are unique due to the order’s antisymmetry.
4. Suprema need not exist, not even in totally ordered sets: the set $\mathcal{X}_0 \coloneqq \{x \in \mathbb{Q} \colon x^2 < 2\}$ admits no supremum in $\mathcal{X} \coloneqq \mathbb{Q}$.

A.1.39 Definition A preordered set $(\mathcal{X}, \leq)$ is called 1. **complete** if every subset of $\mathcal{X}$ has a supremum and an infimum, 2. **conditionally complete** if every non-empty subset of $\mathcal{X}$ having an upper or lower bound also has a supremum or infimum, respectively, 3. **well-ordered** if every non-empty subset of $\mathcal{X}$ has a least element.

A.1.40 Definition Let $\mathcal{X}$ be a set, $(\mathcal{Y}, \leq)$ and $(\mathcal{Z}, \leq)$ be two preordered sets, and $T \colon \mathcal{X} \to \mathcal{Y}$ a function. A function $f \colon \mathcal{X} \to \mathcal{Z}$ is called **increasing** in T if

$$T(x_1) \leq T(x_2) \quad \Longrightarrow \quad f(x_1) \leq f(x_2)$$

for $x_1, x_2 \in \mathcal{X}$. It is called

- **decreasing** in T if it is increasing in T after the preorder $\leq$ on (either $\mathcal{Y}$ or) $\mathcal{Z}$ is switched with its dual order $\geq$,

- **strictly increasing** in T if it is increasing after the orders $\leq$ on $\mathcal{Y}$ and $\mathcal{Z}$ are switched with their corresponding strict orders $<$,
- **strictly decreasing** in T if it is increasing after the orders $\leq$ on $\mathcal{Y}$ and $\mathcal{Z}$ are switched with their strict $<$ and strict dual orders $>$, respectively.

Functions that are (strictly) increasing or decreasing in T are called **(strictly) monotonic** in T. We employ these terms without appending "in T" if $\mathcal{X} = \mathcal{Y}$ and $T = \mathrm{id}_{\mathcal{X}}$.

A.1.41 Remark 1. If $(\mathcal{X}, \leq)$, $(\mathcal{Y}, \leq)$, and $(\mathcal{Z}, \leq)$ are preordered sets, and $f\colon \mathcal{X} \to \mathcal{Y}$ and $g\colon \mathcal{Y} \to \mathcal{Z}$ are both increasing or both decreasing, then their composition $g \circ f$ is increasing. If one of them is increasing and the other decreasing, then their composition $g \circ f$ is decreasing.
2. f is increasing in T if, and only if, there is a function $g\colon \mathcal{Y} \to \mathcal{Z}$ such that $g|_{T[\mathcal{X}]}$ is increasing and $f = g \circ T$. In fact, the previous remark yields the "if" part, while the "only if" part follows by picking $g \in \prod_{y \in T[\mathcal{X}]} f[T = y]$ and extending it arbitrarily to $\mathcal{Y}$.
3. Strictly monotonic functions from totally ordered sets to preordered sets are injective.

A.1.42 Lemma *A preordered set $(\mathcal{X}, \leq)$ is well-ordered if, and only if, its preorder is total and there is no strictly decreasing sequence in $\mathcal{X}$.*

Proof. Let $(\mathcal{X}, \leq)$ be well-ordered. Since $\{x, y\}$ has a least element for $x, y \in \mathcal{X}$, $\leq$ is total. Let now $x \in \mathcal{X}^{\mathbb{N}}$ be decreasing. Since $x[\mathbb{N}]$ has a least element, there is some $N \in \mathbb{N}$ with $x_N \leq x_n$ for $n \in \mathbb{N}$. The monotonicity of x implies $x_n = x_N$ for $n \geq N$.

Let us now assume that $\leq$ is total and that there is some non-empty $\mathcal{X}_0 \subseteq \mathcal{X}$ with no least element. We then construct a strictly decreasing sequence in $\mathcal{X}$ recursively as follows. Let $x_1 \in \mathcal{X}_0$. Given $x_n \in \mathcal{X}_0$ for some $n \in \mathbb{N}$, we pick $x_{n+1} \in \mathcal{X}_0$ with $x_{n+1} < x_n$ (this is possible since, otherwise, $\{\mathrm{id}_{\mathcal{X}_0} \geq x_n\} = \mathcal{X}_0$, making x_n least in $\mathcal{X}_0$). □

A.1.43 Definition Let $(\mathcal{X}, \leq)$ and $(\mathcal{Y}, \leq)$ be two preordered sets and $\mathcal{X}_0 \subseteq \mathcal{X}$. A function $T\colon \mathcal{X} \to \mathcal{Y}$ is called **unbounded above** or **below** on $\mathcal{X}_0$ if for each $y \in \mathcal{Y}$ there is some $x \in \mathcal{X}_0$ with $f(x) > y$ or $f(x) < y$, respectively.

A.1.44 Lemma *Let $\mathcal{X}$ and $\mathcal{Y}$ be two sets and $T\colon \mathcal{X} \to \mathcal{Y}$ a function.*

1. *If $\leq$ is a (total) preorder on $\mathcal{Y}$, then*

$$\leq_T \coloneqq \{x \in \mathcal{X}^2 \colon T(x_1) \leq T(x_2)\}$$

is a (total) preorder on $\mathcal{X}$.

2. *If $<$ is a strict order on $\mathcal{Y}$, then*

$$<_T \coloneqq \{x \in \mathcal{X}^2 \colon T(x_1) < T(x_2)\}$$

is a strict order on $\mathcal{X}$.

Proof.
- Reflexivity of $\leq_T$: For $x \in \mathcal{X}$ we have $T(x) \leq T(x)$ due to the reflexivity of $\leq$, hence $x \leq_T x$.
- Irreflexivity of $<_T$: For $x \in \mathcal{X}$ we have $T(x) \not< T(x)$ due to the irreflexivity of $<$, hence $x \not<_T x$.
- Transitivity of $\leq_T$ and $<_T$: Let $(\prec, \prec_T) \in \{(\leq, \leq_T), (<, <_T)\}$ and let $x \in \mathcal{X}^3$ with $x_1 \prec_T x_2$ and $x_2 \prec_T x_3$. Then $T(x_1) \prec T(x_2)$ and $T(x_2) \prec T(x_3)$, so $T(x_1) \prec T(x_3)$ due to the transitivity of $\prec$, which means $x_1 \prec_T x_3$.

- Totality of $\leq_T$: Let $x \in \mathcal{X}^2$. The totality of $\leq$ implies $T(x_1) \leq T(x_2)$ or $T(x_2) \leq T(x_1)$, yielding $x_1 \leq_T x_2$ or $x_2 \leq_T x_1$. □

A.1.45 Definition The preorder $\leq_T$ and the strict order $<_T$ from the preceding lemma shall be referred to as the preorder and the strict order, respectively, **induced** by T.

A.1.46 Remark 1. The strict order $<_T$ induced by T is total if, and only if, $< \cap\, T[\mathcal{X}]^2$ is total (on $T[\mathcal{X}]$) and T is injective.

Proof. Let $<_T$ be total and let $x \in \mathcal{X}^2$ with $x_1 \neq x_2$. The totality of $<_T$ implies $x_1 <_T x_2$ or $x_2 <_T x_1$, which means $T(x_1) < T(x_2)$ or $T(x_2) < T(x_1)$. This yields the totality of $< \cap\, T[\mathcal{X}]^2$. The irreflexivity of $<$ implies $T(x_1) \neq T(x_2)$, yielding the injectivity of T.
Let now $< \cap\, T[\mathcal{X}]^2$ be total, T injective, and let $x \in \mathcal{X}^2$ with $x_1 \neq x_2$. The injectivity of T implies $T(x_1) \neq T(x_2)$, and the totality of $<$ then implies $T(x_1) < T(x_2)$ or $T(x_2) < T(x_1)$, which means $x_1 \leq_T x_2$ or $x_2 \leq_T x_1$. □

2. The preorder $\leq_T$ induced by T is an order if, and only if, $\leq \cap\, T[\mathcal{X}]^2$ is an order and T is injective.

Proof. Let $\leq_T$ be an order and let $x \in \mathcal{X}^2$ with $T(x_1) \leq T(x_2)$ and $T(x_2) \leq T(x_1)$. Then $x_1 \leq_T x_2$ and $x_2 \leq_T x_1$, and the antisymmetry of $\leq_T$ implies $x_1 = x_2$, hence $T(x_1) = T(x_2)$. The injectivity of T follows from the above and the reflexivity of $\leq$.
Let now $\leq \cap\, T[\mathcal{X}]^2$ be an order, T injective and $x \in \mathcal{X}^2$ with $x_1 \leq_T x_2$ and $x_2 \leq_T x_1$. Then $T(x_1) \leq T(x_2)$ and $T(x_2) \leq T(x_1)$, and the antisymmetry of $\leq \cap\, T[\mathcal{X}]^2$ implies $T(x_1) = T(x_2)$. The injectivity of T implies $x_1 = x_2$. □

3. $(\mathcal{X}, \leq_T)$ is well-ordered if, and only if, $(T[\mathcal{X}], \leq \cap\, T[\mathcal{X}]^2)$ is well-ordered.

Proof. Let $(\mathcal{X}, \leq_T)$ be well-ordered and $\emptyset \neq B \subseteq T[\mathcal{X}]$. Set $A := T^{-1}[B]$. Then $\emptyset \neq A \subseteq \mathcal{X}$, hence there is some $\xi \in A$ with $\xi \leq_T x$ for $x \in A$. Thus, $T(\xi) \leq T(x)$ for $x \in A$, which means that $T(\xi) \leq z$ for $z \in B$.
Let now $(T[\mathcal{X}], \leq \cap\, T[\mathcal{X}]^2)$ be well-ordered and $\emptyset \neq A \subseteq \mathcal{X}$. Then $\emptyset \neq T[A] \subseteq T[\mathcal{X}]$, hence there is some $\xi \in A$ with $T(\xi) \leq T(x)$ for $x \in A$. Thus, $\xi \leq_T x$ for $x \in A$. □

4. We have $(\leq_T)^{\mathrm{str}} \supseteq (\leq^{\mathrm{str}})_T$ and $(<_T)^{\mathrm{unstr}} \subseteq (<^{\mathrm{unstr}})_T$, with equality in both cases if T is injective. In fact,

$$(\leq_T)^{\mathrm{str}} = \{x \in \mathcal{X}^2 \colon T(x_1) \leq T(x_2),\ x_1 \neq x_2\}$$
$$(\leq^{\mathrm{str}})_T = \{x \in \mathcal{X}^2 \colon T(x_1) \leq T(x_2),\ T(x_1) \neq T(x_2)\}$$

and

$$(<_T)^{\mathrm{unstr}} = \{x \in \mathcal{X}^2 \colon T(x_1) < T(x_2) \text{ or } x_1 = x_2\}$$
$$(<^{\mathrm{unstr}})_T = \{x \in \mathcal{X}^2 \colon T(x_1) < T(x_2) \text{ or } T(x_1) = T(x_2)\}.$$

5. We have $(\leq_T)^{\mathrm{op}} = (\leq^{\mathrm{op}})_T =: \leq_T^{\mathrm{op}}$ and $(<_T)^{\mathrm{op}} = (<^{\mathrm{op}})_T =: <_T^{\mathrm{op}}$.
6. Every preorder can be regarded as being induced by some function (the identity for instance).
7. Let $(\mathcal{X}, \leq_{\mathcal{X}})$ and $(\mathcal{Y}, \leq_{\mathcal{Y}})$ be two preordered sets. A function $T \colon \mathcal{X} \to \mathcal{Y}$ is increasing if, and only if, $\leq_{\mathcal{X}} \subseteq \leq_T$. Analogous results hold for "decreasing" and the strict variants.

8. Let $\mathcal{X}$ be a set, $(\mathcal{Y}, \leq_{\mathcal{Y}})$ and $(\mathcal{Z}, \leq_{\mathcal{Z}})$ two preordered sets, and $T\colon \mathcal{X} \to \mathcal{Y}$ and $S\colon \mathcal{X} \to \mathcal{Z}$ two functions. T is increasing or decreasing in S if, and only if, $\leq_S \subseteq \leq_T$ or $\leq_S \subseteq \leq_T^{\mathrm{op}}$, respectively.

A.1.47 Example Let $\mathcal{X} \coloneqq \{1/n\colon n \in \mathbb{N}\}$. Then $\mathrm{id}_{\mathcal{X}}$ is strictly increasing, but there is no strictly decreasing function $f\colon \mathcal{X} \to \mathcal{X}$ since, otherwise, $1/k = f(1) < f(1/n)$ for some $k \in \mathbb{N}$ and every $n \geq 2$, which implies $\{f(1/n)\colon n \geq 2\} \subseteq \{1/n\colon n \in \{1, \dots, k-1\}\}$, a contradiction since the left set is infinite due to Remark A.1.41.

A.1.48 Definition Let $(\mathcal{X}, \leq)$ and $(\mathcal{Y}, \leq)$ be two preordered sets. A function $f\colon \mathcal{X} \to \mathcal{Y}$ is called **(strictly) unimodal** if there is some $\xi \in \mathcal{X}$ such that

$$\left\{\begin{matrix} f|_{\{\mathrm{id}_{\mathcal{X}} \leq \xi\}} \\ f|_{\{\mathrm{id}_{\mathcal{X}} \geq \xi\}} \end{matrix}\right\} \text{ is (strictly) } \left\{\begin{matrix} \text{increasing} \\ \text{decreasing} \end{matrix}\right\}.$$

A.1.49 Remark Let $f\colon \mathbb{Z} \to]0, \infty[$ be a function with

$$\frac{f(n+1)}{f(n)} \left\{\begin{matrix} > \\ < \end{matrix}\right\} 1 \quad \iff \quad n \left\{\begin{matrix} < \\ > \end{matrix}\right\} \xi$$

for some $\xi \in \mathbb{R}$. Then

$$\operatorname{Arg\,max} f = \{\lceil \xi \rceil, \lfloor \xi \rfloor + 1\}$$

and f is unimodal (even strongly if $\xi \notin \mathbb{Z}$).

A.1.50 Example

1. For $f = \mathrm{b}_{n,p}$ with $n \in \mathbb{N}$ and $p \in]0, 1[$, we obtain

$$\operatorname{Arg\,max} \mathrm{b}_{n,p} = \{\lceil (n+1)p \rceil - 1, \lfloor (n+1)p \rfloor\}.$$

If $p \in \{0, 1\}$, then $\operatorname{Arg\,max} \mathrm{b}_{n,p} = \{np\}$.

2. For $f = \mathrm{p}_\lambda$ with $\lambda \in]0, \infty[$, we obtain

$$\operatorname{Arg\,max} \mathrm{p}_\lambda = \{\lceil \lambda \rceil - 1, \lfloor \lambda \rfloor\}.$$

A.1.51 Definition For $n \in \mathbb{N}$ and $x \in \mathbb{R}^n$ we denote by

$$x_{(i)} \coloneqq \min\left\{\xi \in \mathbb{R}\colon \sum_{k=1}^{n} \mathbf{1}(\xi \geq x_k) \geq i\right\} \qquad \text{for } i \in \{1, \dots, n\}$$

the ith **order statistic** of x.

A.1.52 Remark We have $\bigwedge_{k=1}^n x_k = x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)} = \bigvee_{k=1}^n x_k$.

A.1.53 Lemma *Let $(\mathcal{X}, \leq)$ and $(\mathcal{Y}, \leq)$ be two preordered sets and $f\colon \mathcal{X} \to \mathcal{Y}$ a function.*

(i) *If f is increasing, then preimages of down- or uprays in $\mathcal{Y}$ are down- or uprays in $\mathcal{X}$, respectively.*

(ii) *If f is decreasing, then preimages of down- or uprays in $\mathcal{Y}$ are up- or downrays in $\mathcal{X}$, respectively.*

(iii) *If $f[\mathrm{id}_{\mathcal{X}} \leq x] = \{\mathrm{id}_{\mathcal{Y}} \leq f(x)\}$ for $x \in \mathcal{X}$, then images of down- or uprays in $\mathcal{X}$ are down- or uprays in $\mathcal{Y}$, respectively.*

(iv) *If* $f[\mathrm{id}_{\mathcal{X}} \leq x] = \{\mathrm{id}_{\mathcal{Y}} \geq f(x)\}$ *for* $x \in \mathcal{X}$*, then images of down- or uprays in* $\mathcal{X}$ *are up- or downrays in* $\mathcal{Y}$*, respectively.*

A.1.54 Remark We have the implications

$$\begin{aligned} f \text{ increasing} \quad &\Longrightarrow \quad f[\mathrm{id}_{\mathcal{X}} \leq x] \subseteq \{\mathrm{id}_{\mathcal{Y}} \leq f(x)\} \quad \text{for } x \in \mathcal{X} \\ f \text{ decreasing} \quad &\Longrightarrow \quad f[\mathrm{id}_{\mathcal{X}} \leq x] \subseteq \{\mathrm{id}_{\mathcal{Y}} \geq f(x)\} \quad \text{for } x \in \mathcal{X}, \end{aligned}$$

but surjectivity of f alone does not suffice for the converse inclusions to hold.

Proof of Lemma A.1.53. The claim concerning preimages follows from $\mathbf{1}_{f^{-1}[B]} = \mathbf{1}_B \circ f$ and Remark A.1.41.

The claim concerning images follows from

$$f[D] = f\left[\bigcup_{x \in D} \{\mathrm{id}_{\mathcal{X}} \leq x\}\right] = \bigcup_{x \in D} f[\mathrm{id}_{\mathcal{X}} \leq x] = \bigcup_{x \in D} \{\mathrm{id}_{\mathcal{Y}} \leq f(x)\} = \lceil f[D] \rceil_{\mathcal{Y}}$$

for downrays D in $\mathcal{X}$, and the consideration of all remaining three combinations of preorders and their duals on $\mathcal{X}$ and $\mathcal{Y}$. □

A.1.55 Remark The following result is needed in the chapters with applications to well-known distribution classes.

A.1.56 Lemma *Let* $(\mathcal{X}, \leq)$ *and* $(\mathcal{Y}, \leq)$ *be two preordered sets and* $f, g \colon \mathcal{X} \to \mathcal{Y}$ *two functions with* $f \leq g$*. Then*

$$f^{-1}[D] \supseteq g^{-1}[D] \quad \text{and} \quad f^{-1}[U] \subseteq g^{-1}[U] \qquad \text{for downrays } D \text{ and uprays } U \text{ in } \mathcal{Y}.$$

Moreover,

$$\begin{aligned} g[\mathrm{id}_{\mathcal{X}} \leq x] = \{\mathrm{id}_{\mathcal{Y}} \leq g(x)\} \text{ for } x \in \mathcal{X} \quad &\Longrightarrow \quad f[D] \subseteq g[D] \text{ for downrays } D \text{ in } \mathcal{X} \\ f[\mathrm{id}_{\mathcal{X}} \leq x] = \{\mathrm{id}_{\mathcal{Y}} \leq f(x)\} \text{ for } x \in \mathcal{X} \quad &\Longrightarrow \quad f[U] \supseteq g[U] \text{ for uprays } U \text{ in } \mathcal{X} \\ g[\mathrm{id}_{\mathcal{X}} \leq x] = \{\mathrm{id}_{\mathcal{Y}} \geq g(x)\} \text{ for } x \in \mathcal{X} \quad &\Longrightarrow \quad f[D] \supseteq g[D] \text{ for downrays } D \text{ in } \mathcal{X} \\ f[\mathrm{id}_{\mathcal{X}} \leq x] = \{\mathrm{id}_{\mathcal{Y}} \geq f(x)\} \text{ for } x \in \mathcal{X} \quad &\Longrightarrow \quad f[U] \subseteq g[U] \text{ for uprays } U \text{ in } \mathcal{X}. \end{aligned}$$

Proof. If D is a downray in $\mathcal{Y}$ and $x \in g^{-1}[D]$, then $f(x) \leq g(x) \in D$, which implies $f(x) \in D$, i.e., $x \in f^{-1}[D]$. If U is an upray in $\mathcal{Y}$ and $x \in f^{-1}[U]$, then $U \ni f(x) \leq g(x)$, which implies $g(x) \in U$, i.e., $x \in g^{-1}[U]$.

From the implications concerning preimages we prove merely the first one since the others follow from a consideration of all three remaining combinations of dual orders on $\mathcal{X}$ and $\mathcal{Y}$. For a downray D in $\mathcal{X}$ and $x \in D$ we have $f(x) \leq g(x)$, and, since Lemma A.1.53 yields that $g[D]$ is a downray in $\mathcal{Y}$, we obtain $f(x) \in g[D]$. □

A.1.57 Definition Let $(\mathcal{X}, \leq)$ be a totally preordered set. The sets $\{\mathrm{id}_{\mathcal{X}} < x\}$ and $\{\mathrm{id}_{\mathcal{X}} > x\}$ are called **open downray** and **open upray,** respectively, with *endpoint* $x \in \mathcal{X}$. **Open intervals** are finite (i.e., possibly empty) intersections of open downrays with open uprays.

A.1.58 Remark If $(\mathcal{X}, \leq)$ is a totally preordered set, then open down-, uprays, and intervals are down-, uprays, or intervals, respectively.

A.1.59 Definition Let $\mathcal{X}$ be a set. A system $\mathfrak{T} \subseteq 2^{\mathcal{X}}$ is called a **topology** on $\mathcal{X}$ if

1. $\bigcap \mathfrak{T}_0 \in \mathfrak{T}$ for $\mathfrak{T}_0 \subseteq \mathfrak{T}$ with $|\mathfrak{T}_0| < \infty$,
2. $\bigcup \mathfrak{T}_0 \in \mathfrak{T}$ for $\mathfrak{T}_0 \subseteq \mathfrak{T}$,

in which case $(\mathcal{X}, \mathfrak{T})$ is called a **topological space.**

A.1.60 Remark

1. If $(\mathcal{X}, d)$ is a metric space, then the system of open sets is a topology on $\mathcal{X}$, called **induced** by d.
2. Metric spaces shall in the following always be equipped with their induced topologies.

A.1.61 Remark Let $\mathcal{X}$ be a set.

1. If $\mathcal{T} \subseteq 2^{2^{\mathcal{X}}}$ is a set of topologies $\mathfrak{T}$ on $\mathcal{X}$, then $\bigcap \mathcal{T}$ is a topology on $\mathcal{X}$.
2. Therefore, given a system $\mathfrak{T}_0 \subseteq 2^{\mathcal{X}}$,

   $$\tau(\mathfrak{T}_0) := \bigcap \{\mathfrak{T} \in 2^{2^{\mathcal{X}}} : \mathfrak{T} \supseteq \mathfrak{T}_0 \text{ is a topology on } \mathcal{X}\}$$

   is the least topology on $\mathcal{X}$ containing $\mathfrak{T}_0$. $\mathfrak{T}_0$ is called a **subbase** of $\tau(\mathfrak{T}_0)$.
3. $\mathfrak{T}_0 \subseteq 2^{\mathcal{X}}$ is called a **base** of $\tau(\mathfrak{T}_0)$ if $\bigcap \mathfrak{T}_1 \in \mathfrak{T}_0$ for $\mathfrak{T}_1 \subseteq \mathfrak{T}_0$ with $|\mathfrak{T}_1| < \infty$.
4. The topology $\tau(\mathfrak{T}_0)$ can also be expressed as the set of arbitrary unions of finite intersections of members of $\mathfrak{T}_0$.

A.1.62 Definition Let $(\mathcal{X}, \mathfrak{T})$ be a topological space and $\mathcal{X}_0 \subseteq \mathcal{X}$. Then $\mathfrak{T} \cap \mathcal{X}_0 := \{U \cap \mathcal{X}_0 : U \in \mathfrak{T}\}$ is a topology on $\mathcal{X}_0$, called **subspace topology** or **induced** by $\mathfrak{T}$ on $\mathcal{X}_0$.

A.1.63 Definition Let $(\mathcal{X}, \mathfrak{T})$ be a topological space. A subset $\mathcal{X}_0 \subseteq \mathcal{X}$ is called **dense** (in $\mathcal{X}$) if $U \cap \mathcal{X}_0 \neq \emptyset$ for $U \in \mathfrak{T} \setminus \{\emptyset\}$.

A.1.64 Definition Let $(\mathcal{X}, \mathfrak{T})$ be a topological space. The **density** $\operatorname{dens}(\mathcal{X}, \mathfrak{T})$ of $(\mathcal{X}, \mathfrak{T})$ is the least cardinality of a dense subset of $\mathcal{X}$, i.e.,

$$\operatorname{dens}(\mathcal{X}, \mathfrak{T}) := \min\{|\mathcal{X}_0| : \mathcal{X}_0 \subseteq \mathcal{X} \text{ is dense}\}.$$

The **hereditary density** of $(\mathcal{X}, \mathfrak{T})$ is

$$\operatorname{heredens}(\mathcal{X}, \mathfrak{T}) := \sup\{\operatorname{dens}(\mathcal{X}_0, \mathfrak{T} \cap \mathcal{X}_0) : \mathcal{X}_0 \subseteq \mathcal{X}\}.$$

A.1.65 Remark

1. The minimum in the definition of $\operatorname{dens}(\mathcal{X}, \mathfrak{T})$ exists since every cardinal number is an ordinal number and every set of ordinal numbers is well-ordered.
2. Obviously, $\operatorname{dens}(\mathcal{X}, \mathfrak{T}) \leq \operatorname{heredens}(\mathcal{X}, \mathfrak{T})$.
3. If the topology $\mathfrak{T}$ is clear from the context (as is the case with, e.g., totally preordered sets according to Remark A.1.69), we merely write $\operatorname{dens}(\mathcal{X})$ and $\operatorname{heredens}(\mathcal{X})$.

A.1.66 Definition A topological space $(\mathcal{X}, \mathfrak{T})$ is called **separable** if $\operatorname{dens}(\mathcal{X}, \mathfrak{T}) \leq \aleph_0$, i.e., if there exists a countable dense subset of $\mathcal{X}$.

A.1.67 Example We have $\operatorname{dens}(\mathbb{R}) = \aleph_0$ since the rationals are countable and dense in the reals. Even the stronger result $\operatorname{heredens}(\mathbb{R}) = \aleph_0$ holds by Theorem A.1.71.

A.1.68 Definition Let $(\mathcal{X}, \leq)$ be a totally preordered set. The topology $\mathfrak{I}(\mathcal{X})$ on $\mathcal{X}$ having as subbase the system of all open intervals in $\mathcal{X}$ is called **order topology** on $\mathcal{X}$.

A.1.69 Remark Let $(\mathcal{X}, \leq)$ be a totally preordered set.

1. The system of all open intervals in $\mathcal{X}$ is a base of the order topology $\mathfrak{I}(\mathcal{X})$ on $\mathcal{X}$.
2. Unless stated otherwise, $\mathcal{X}$ shall in the following always be endowed with its order topology—and not its *Alexandrov topology* (which is the system of all uprays in $\mathcal{X}$), as is usually the case with preordered sets.

3. We can thus speak of the Borel σ-algebra $\mathfrak{B}(\mathcal{X}) := \sigma(\mathfrak{I}(\mathcal{X}))$ on $\mathcal{X}$, i.e., the σ-algebra generated by all open intervals.
4. When considered as a measurable space, $\mathcal{X}$ shall in the following always be endowed with its Borel σ-algebra.
5. Intervals are Borel sets. This follows from Remark A.1.35 and the observation that every downray D that is not open can be expressed as $D = \{\mathrm{id}_{\mathcal{X}} \leq x\} = \mathcal{X} \setminus \{\mathrm{id}_{\mathcal{X}} > x\} \in \mathfrak{B}(\mathcal{X})$ for some $x \in \mathcal{X}$.
6. If $(\mathcal{Y}, \leq)$ is a further totally preordered set and $f\colon \mathcal{X} \to \mathcal{Y}$ is monotonic, then Lemma A.1.53, Remark A.1.58, and the above imply the measurability of f.
7. $\mathcal{X}_0 \subseteq \mathcal{X}$ is dense if, and only if, the following implication holds for every open interval I in $\mathcal{X}$:

$$I \neq \emptyset \quad \Longrightarrow \quad I \cap \mathcal{X}_0 \neq \emptyset.$$

8. If $\mathcal{X}_0 \subseteq \mathcal{X}$ is dense, then so is $\mathcal{X}_1 := \mathcal{X}_0 \cup \{\xi \in \mathcal{X} \colon \xi = \min \mathcal{X} \text{ or } \xi = \max \mathcal{X}\}$, and $|\mathcal{X}_1| = |\mathcal{X}_0|$. In fact, if $\mathcal{X}$ is finite, then the preceding part yields $\min \mathcal{X}, \max \mathcal{X} \in \mathcal{X}_0$, and if $\mathcal{X}$ is infinite, then so is $\mathcal{X}_0$, hence $|\mathcal{X}_1| = |\mathcal{X}_0|$.

A.1.70 Example The order topology $\mathfrak{I}(\mathcal{X})$ on an interval $\mathcal{X} \subseteq \mathbb{R}$ (endowed with the usual order) coincides with the usual (induced) Euclidean topology on $\mathcal{X}$. This does not, however, hold for arbitrary subsets $\mathcal{X} \subseteq \mathbb{R}$, as shows $\mathcal{X} := \{-1\} \cup \{1/n \colon n \in \mathbb{N}\}$.

A.1.71 Theorem *If $(\mathcal{X}, \leq)$ is a totally preordered set, then* $\mathrm{heredens}(\mathcal{X}) = \mathrm{dens}(\mathcal{X})$.

Proof. A proof can be found in Bridges and Mehta (1995) or Scott (2012). □

A.1.72 Definition Let $(\mathcal{X}, \leq)$ and $(\mathcal{Y}, \leq)$ be two preordered sets. A function $f\colon \mathcal{X} \to \mathcal{Y}$ is called **order-preserving** if the following equivalence holds:

$$x_1 \leq x_2 \quad \Longleftrightarrow \quad f(x_1) \leq f(x_2) \qquad \text{for } x_1, x_2 \in \mathcal{X}.$$

A.1.73 Remark 1. The term "order-preserving" usually denotes a function that preserves the order in merely one direction, i.e., an increasing function. Our definition follows the one from Debreu (1954, bottom of p. 160).
2. If a function $f\colon \mathcal{X} \to \mathcal{Y}$ is order-preserving, then the following equivalence holds:

$$x_1 < x_2 \quad \Longleftrightarrow \quad f(x_1) < f(x_2) \qquad \text{for } x_1, x_2 \in \mathcal{X}.$$

This follows from the following chain of equivalences:

$$\begin{aligned} x_1 < x_2 \quad &\Longleftrightarrow \quad x_1 \leq x_2 \text{ and } x_2 \not\leq x_1 \\ &\Longleftrightarrow \quad f(x_1) \leq f(x_2) \text{ and } f(x_2) \not\leq f(x_1) \\ &\Longleftrightarrow \quad f(x_1) < f(x_2). \end{aligned}$$

3. Order-preserving functions between ordered sets are thus injective.
4. If $\mathcal{X}$ is totally preordered, then f is order-preserving if, and only if, merely the following implication holds:

$$x_1 \leq x_2 \quad \Longrightarrow \quad f(x_1) \leq f(x_2) \qquad \text{for } x_1, x_2 \in \mathcal{X}.$$

In fact, let us assume the above implication. Then the following chain of implications holds:

$$\begin{aligned} x_1 \not\leq x_2 \quad &\Longrightarrow \quad x_2 \leq x_1 \\ &\Longrightarrow \quad x_2 < x_1 \\ &\Longrightarrow \quad f(x_2) < f(x_1) \\ &\Longrightarrow \quad f(x_1) \not\leq f(x_2), \end{aligned}$$

where the first implication follows from the totality of the preorder $\leq$ on $\mathcal{X}$, the second one by definition of the strict order $< \, = \, \leq^{\mathrm{str}}$ on $\mathcal{X}$, the third one by part 1, and the last one by definition of the strict order $< \, = \, \leq^{\mathrm{str}}$ on $\mathcal{Y}$.

A.1.74 Theorem *Let $(\mathcal{X}, \leq)$ be a totally preordered set. Then* $\operatorname{dens}(\mathcal{X}) \leq \aleph_0$ *if, and only if, there is an order-preserving function $f\colon \mathcal{X} \to \mathbb{R}$.*

Proof. The following proof is in essence due to Greinecker (2012).

- Let us assume $\operatorname{dens}(\mathcal{X}) \leq \aleph_0$, and let us choose a countable dense set $\mathcal{X}_0 = \{x_n \colon n \in \mathbb{N}\}$ in $\mathcal{X}$. If we define

$$f\colon \mathcal{X} \to \mathbb{R}, \qquad x \mapsto \sum_{\substack{n \in \mathbb{N} \\ x_n \leq x}} \frac{1}{2^n} - \sum_{\substack{n \in \mathbb{N} \\ x_n \geq x}} \frac{1}{2^n},$$

then f is order-preserving due to part 3 of the last remark and the following implications:

$$\begin{aligned} x \leq y \quad &\Longrightarrow \quad \begin{cases} \{n \in \mathbb{N} \colon x_n \leq x\} \subseteq \{n \in \mathbb{N} \colon x_n \leq y\} \\ \{n \in \mathbb{N} \colon x_n \geq x\} \supseteq \{n \in \mathbb{N} \colon x_n \geq y\} \end{cases} \\ &\Longrightarrow \quad f(x) \leq f(y). \end{aligned}$$

- Let now $f\colon \mathcal{X} \to \mathbb{R}$ be an order-preserving function. Let us set

$$G := \Big\{ (\alpha, \beta) \in \mathbb{Q}^2 \colon \alpha < \beta,\ f^{-1}\big[ ]\alpha, \beta[\big] \neq \emptyset \Big\}.$$

Then G, being a subset of $\mathbb{Q}^2$, is countable. Let us consider a choice function $g\colon G \to \mathcal{X}$ with $g(\alpha, \beta) \in f^{-1}\big[ ]\alpha, \beta[\big]$. Then $\mathcal{X}_0 := g[G]$ is dense in $\mathcal{X}$. To see this, we shall apply Remark A.1.69, part 7. Let I be a non-empty open interval in $\mathcal{X}$. Let us first suppose $I = \{x < \mathrm{id}_{\mathcal{X}} < y\}$ for some $x, y \in \mathcal{X}$. Part 1 of the previous remark then yields for $z \in \mathcal{X}$ the equivalence

$$x < z < y \qquad \Longleftrightarrow \qquad f(x) < f(z) < f(y).$$

Let now $z \in I$. The denseness of $\mathbb{Q}$ in $\mathbb{R}$ yields $\alpha, \beta \in \mathbb{Q}$ with $f(x) < \alpha < f(z) < \beta < f(y)$. This implies $(\alpha, \beta) \in G$, and thus $g(\alpha, \beta) \in f^{-1}\big[ ]\alpha, \beta[\big] \subseteq I$, the inclusion following from the above equivalence. This yields $I \cap \mathcal{X}_0 \neq \emptyset$. The cases where the interval I is of the form $\{\mathrm{id}_{\mathcal{X}} > x\}$ or $\{\mathrm{id}_{\mathcal{X}} < x\}$ are handled the same way. □

A.1.75 Remark Debreu (1954) gives necessary conditions for the existence of a continuous order-preserving function $f\colon \mathcal{X} \to \mathbb{R}$. Cantor (1895, § 11) does similarly for order-isomorphisms $f\colon \mathcal{X} \to [0, 1]$.

A.2 ORDINAL NUMBERS

A.2.1 Remark While the inequality

$$P(F \leq t) \leq t \qquad \text{for } t \in [0,1]$$

holds for every law P on $\mathbb{R}$ with corresponding distribution function F, this may not be so upon replacing $\mathbb{R}$ with an arbitrary totally ordered set. Lemma A.2.6, taken from Rudin (1986, Chapter 2, exercise 18), illustrates this.

A.2.2 Definition A set α is called an **ordinal (number)** if $(\alpha, \in^{\mathrm{unstr}})$ is well-ordered and every element of α is also a subset of α.

A.2.3 Remark In order to understand the rest of this section it suffices to know of the existence of a least uncountable ordinal number, denoted by ω_1. The interested reader may consult, e.g., Dugundji (1966, Chapter 2, Section 6) or Jech (2003, Chapter 1, Section 2) for a rigorous introduction to the theory of ordinal numbers. Special attention to the first uncountable ordinal number is given in Dugundji (1966, Chapter 2, Section 9), where it is denoted by Ω.

A.2.4 Definition Let

$$\mathcal{X} := \omega_1 + 1 = [0, \omega_1]$$

be the second uncountable ordinal (see, e.g., Jech, 2003) and

$$\begin{aligned} \mathfrak{A}_1 &:= \bigcup_{\alpha \in [0,\omega_1[} \{A \subseteq \mathcal{X} \colon [\alpha, \omega_1[\subseteq A\} \\ \mathfrak{A}_0 &:= \{A^{\mathrm{c}} \colon A \in \mathfrak{A}_1\} \\ \mathfrak{A} &:= \mathfrak{A}_1 \cup \mathfrak{A}_0. \end{aligned}$$

A.2.5 Remark
1. $\mathfrak{A}_1$ is an upray and $\mathfrak{A}_0$ a downray in $(2^{\mathcal{X}}, \subseteq)$.
2. $\mathfrak{A}_1 \cap \mathfrak{A}_0 = \emptyset$. In fact, if $A \in \mathfrak{A}_1 \cap \mathfrak{A}_0$, then there exist two ordinals $\alpha_1, \alpha_0 \in [0, \omega_1[$ with $[\alpha_1, \omega_1[\subseteq A$ and $[\alpha_0, \omega_1[\subseteq A^{\mathrm{c}}$, hence $[\alpha_1 \vee \alpha_0, \omega_1[\subseteq A \cap A^{\mathrm{c}} = \emptyset$, yielding $\alpha_1 \vee \alpha_0 = \omega_1$, a contradiction.

A.2.6 Lemma
(i) *$\mathfrak{A}$ is a σ-algebra on $\mathcal{X}$.*
(ii) *$P := \mathbf{1}_{\mathfrak{A}_1}|_{\mathfrak{A}}$ is a law on $\mathcal{X}$.*
(iii) *If F denotes the distribution function of P, then $\{F \leq 0\} \in \mathfrak{A}$ and $P(F \leq 0) = 1$.*

Proof. (i), (ii) We have $\mathcal{X} \in \mathfrak{A}_1$, which yields $P(\emptyset) = 0$. By definition of $\mathfrak{A}_0$, the equivalence $A \in \mathfrak{A}_1 \iff A^{\mathrm{c}} \in \mathfrak{A}_0$ holds. Given a sequence $(A_n \colon n \in \mathbb{N}) \in \mathfrak{A}^{\mathbb{N}}$, we distinguish the following two cases:

- There is some $N \in \mathbb{N}$ with $A_N \in \mathfrak{A}_1$. The previous remark then yields $\bigcup_{n\in\mathbb{N}} A_n \in \mathfrak{A}_1$. If the sequence is furthermore pairwise disjoint, then $A_n \in \mathfrak{A}_0$ for $n \in \mathbb{N} \setminus \{N\}$, which implies $P(\bigcup_{n\in\mathbb{N}} A_n) = 1 = P(A_N) = \sum_{n=1}^{\infty} P(A_n)$.
- $A_n \in \mathfrak{A}_0$ for every $n \in \mathbb{N}$. Pick, for each $n \in \mathbb{N}$, an ordinal $\alpha_n \in [0, \omega_1[$ with $[\alpha_n, \omega_1[\subseteq A_n^{\mathrm{c}}$. Then $\alpha := \bigcup_{n\in\mathbb{N}} \alpha_n$ is, as a countable union of countable sets, a countable ordinal, i.e., $\alpha \in [0, \omega_1[$. Furthermore, $[\alpha, \omega_1[\subseteq \bigcap_{n\in\mathbb{N}} A_n^{\mathrm{c}}$, i.e., $\bigcup_{n\in\mathbb{N}} A_n \in \mathfrak{A}_0$. This also implies $P(\bigcup_{n\in\mathbb{N}} A_n) = 0 = \sum_{n=1}^{\infty} P(A_n)$.

(iii) The following equivalences hold for $x \in \mathcal{X}$:

$$\begin{aligned} F(x) = 0 \quad &\iff \quad]x, \omega_1] \in \mathfrak{A}_1 \\ &\iff \quad x < \omega_1. \end{aligned}$$

Thus, $\{F \leq 0\} = [0, \omega_1[\in \mathfrak{A}_1$ and $P(F \leq 0) = 1$. □

A.3 STATISTICAL NOTIONS

A.3.1 Definition Let $(\mathcal{X}, \mathfrak{A})$ and $(\mathcal{Y}, \mathfrak{B})$ be two measurable spaces. A function $T \colon \mathcal{X} \to \mathcal{Y}$ is called **measurable** (with respect to $\mathfrak{A}$ and $\mathfrak{B}$) if

$$T^{-1}[B] \in \mathfrak{A} \qquad \text{for } B \in \mathfrak{B},$$

that is, more succinctly and using the notation "$f^{-1}[[\cdot]]$" established in Remark A.1.15, if $T^{-1}[[\mathfrak{B}]] \subseteq \mathfrak{A}$. In a statistical context, measurable functions defined on the sample space are called **statistics.**

A.3.2 Definition If $(\mathcal{X}, \mathfrak{T})$ is a topological space and μ a measure on $\mathcal{X}$, then

$$\operatorname{supp} \mu := \mathcal{X} \setminus \bigcup \{U \in \mathfrak{T} \colon \mu(U) = 0\}$$

denotes the **support** of μ.

A.3.3 Remark 1. We have $\operatorname{supp} \mu = \{x \in \mathcal{X} \colon \mu(U) > 0 \text{ for } U \in \mathfrak{T} \text{ with } U \ni x\}$.

2. $\operatorname{supp} \mu$ is the largest closed set $A \subseteq \mathcal{X}$ such that $\mu(U \cap A) > 0$ for $U \in \mathfrak{T}$ with $U \cap A \neq \emptyset$.
3. If the topology is metrizable, then $\operatorname{supp} \mu$ is the smallest closed subset $A \subseteq \mathcal{X}$ with $\mu(\mathcal{X} \setminus A) = 0$ (see, e.g., Parthasarathy, 2005, Theorem 2.21, p. 12).

A.3.4 Definition A measure μ on a measurable space $(\mathcal{X}, \mathfrak{A})$ is called **continuous** if $\mu(A) = 0$ for every countable $A \in \mathfrak{A}$.

A.3.5 Definition Let $(\mathcal{X}, \mathfrak{A}, \mu)$ be a measure space, $(\mathcal{Y}, \mathfrak{B})$ a measurable space, and $T \colon \mathcal{X} \to \mathcal{Y}$ a measurable function (with respect to $\mathfrak{A}$ and $\mathfrak{B}$). We write

$$T \mathbin{\square} \mu := \mu \circ (T^{-1}[\cdot]|_{\mathcal{B}})$$

for the **image measure** of μ under T. If μ is a probability measure, we say **distribution** of T under μ rather than image measure of μ under T.

A.3.6 Definition Let $(\mathcal{X}, \mathfrak{A}, \mu)$ be a measure space. A measurable function $f \colon \mathcal{X} \to \mathbb{R}$ is called **semi-integrable** if $(\int_{\mathcal{X}} f_+ \,\mathrm{d}\mu) \wedge (\int_{\mathcal{X}} f_- \,\mathrm{d}\mu) < \infty$. Its **integral** is then as usual the well-defined quantity $\int_{\mathcal{X}} f \,\mathrm{d}\mu := \int_{\mathcal{X}} f_+ \,\mathrm{d}\mu - \int_{\mathcal{X}} f_- \,\mathrm{d}\mu \in \mathbb{R} \cup \{-\infty, \infty\}$. A function $f \colon \mathcal{X} \to \mathbb{R}^n$ is called **semi-integrable** if each of its components $f_k \colon \mathcal{X} \to \mathbb{R}$ is, and its integral is then $\int_{\mathcal{X}} f \,\mathrm{d}\mu := (\int_{\mathcal{X}} f_k \,\mathrm{d}\mu \colon k \in \{1, \ldots, n\}) \in (\mathbb{R} \cup \{-\infty, \infty\})^n$.

A.3.7 Definition Let $(\mathcal{X}, \mathfrak{A}, \mu)$ be a measurable space and f semi-integrable. We write

$$\mu(f) := \int_{\mathcal{X}} f \,\mathrm{d}\mu$$

for the integral of f.

A.3.8 Definition Let $(\mathcal{X}, \mathfrak{A})$ be a measurable space. A family

$$\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$$

of probability measures P_ϑ on $(\mathcal{X}, \mathfrak{A})$ is called a **model** or **experiment** on $(\mathcal{X}, \mathfrak{A})$.

A.3.9 Remark Let $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$, and $(\mathcal{Y}, \mathfrak{B})$ another measurable space. A statistic $T \colon \mathcal{X} \to \mathcal{Y}$ then induces the model

$$T \mathbin{\square} \mathcal{P} := (T \mathbin{\square} P_\vartheta \colon \vartheta \in \Theta)$$

on $(\mathcal{Y}, \mathfrak{B})$.

A.3.10 Definition Let $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$ and Γ a set. A mapping

$$\kappa \colon \Theta \to \Gamma$$

is called a **parameter of interest** in $\mathcal{P}$.

A.3.11 Remark Parameters of interest are often given by maps κ with $\operatorname{dom}(\kappa) \supseteq \Theta$. We then, too, designate by κ instead of $\kappa|_\Theta$ the parameter of interest.

A.3.12 Definition Let $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$. A parameter of interest $\kappa \colon \Theta \to \Gamma$ is called **identifiable** if the following implication holds:

$$P_\vartheta = P_\eta \quad \Longrightarrow \quad \kappa(\vartheta) = \kappa(\eta) \qquad \text{for } \vartheta, \eta \in \Theta.$$

A.3.13 Remark Identifiability of id_Θ is the same as injectivity of the model $\mathcal{P}$.

A.3.14 Example In the model $\mathcal{P} := (\mathrm{N}^{\otimes n}_{\mu+\nu,\sigma^2} \colon (\mu, \nu) \in \mathbb{R}^2)$, with known sample size $n \in \mathbb{N}$ and variance $\sigma^2 \in \,]0, \infty[$, the parameter of interest $\mathbb{R}^2 \to \mathbb{R}$, $(\mu, \nu) \mapsto \mu + \nu$, is identifiable, whereas $\mathbb{R}^2 \to \mathbb{R}$, $(\mu, \nu) \mapsto \mu - \nu$, is not. The first claim follows from $\mathrm{N}_{\mu+\nu,\sigma^2}(\mathrm{id}_\mathbb{R}) = \mu + \nu$, the second one from $\mathrm{N}_{1-1,\sigma^2} = \mathrm{N}_{0,\sigma^2}$. Examples of non-identifiable parameters of interest in multinomial models can be found in Section 5.3 of Chapter 5.

A.3.15 Definition Let $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$ and $\kappa \colon \Theta \to \Gamma$ a parameter of interest. A function

$$\mathrm{K} \colon \mathcal{X} \to 2^\Gamma$$

such that $\{\mathrm{K} \ni \gamma\} \in \mathfrak{A}$ for $\gamma \in \Gamma$ is called a **confidence region** for κ (in $\mathcal{P}$). Its **effective level** is the number

$$\beta_{\mathrm{eff}}(\mathrm{K}) := \inf_{\vartheta \in \Theta} P_\vartheta(\mathrm{K} \ni \kappa(\vartheta)).$$

Given $\beta \in [0, 1]$, the confidence region K is said to have **level** β if $\beta_{\mathrm{eff}}(\mathrm{K}) \geq \beta$.

A.3.16 Remark In contrast to many statisticians, we explicitly do not exclude the possibility of $\mathrm{K}(x) = \emptyset$. In fact, such an occurrence is rather informative since it tells us that the data $x \in \mathcal{X}$ correspond to the $(1 - \beta)100\%$ (or less) of cases where K does not cover the true parameter $\kappa(\vartheta)$.

A.3.17 Definition Let $\mathcal{P} = (P_\vartheta\colon \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$, $(\Gamma, \leq)$ a preordered set, and $\kappa\colon \Theta \to \Gamma$ a parameter of interest. A function $L\colon \mathcal{X} \to \Gamma$ such that $\{L \leq \mathrm{id}_\Gamma\} = \{\gamma \in \Gamma\colon L \leq \gamma\}$ is a confidence region is called **lower confidence bound.** If $\leq$ is replaced by its dual preorder $\geq$, then it is called **upper confidence bound.** Confidence regions whose values are downrays, uprays, or intervals in Γ (see Definition A.1.33) are called **confidence downrays, uprays,** or **intervals,** respectively.

A.3.18 Remark

1. The reason for introducing confidence rays instead of merely confidence bounds lies, apart from the obvious and profitable generalization appreciable in higher dimensional sets Γ, in the ability to distinguish between confidence regions that include their boundary (as in our definition) and ones that do not (i.e., $\{L < \mathrm{id}_\Gamma\}$ is required to be a confidence region).
2. Classical cases of confidence downrays are thus $]\underline{\kappa}, \infty[$ and $[\underline{\kappa}, \infty[$ in the case $\Gamma = \mathbb{R}$, $\underline{\kappa}$ being a lower confidence bound.

A.3.19 Definition (Comparison of confidence bounds) Let

- $\mathcal{P} = (P_\vartheta\colon \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$,
- $\kappa\colon \mathcal{X} \to \overline{\mathbb{R}}$ a parameter of interest,
- L and L' two lower confidence bounds for κ with level $\beta \in [0, 1]$.

L' is called **better** than (or **superior** to) L if

$$P_\vartheta(L' \geq t) \geq P_\vartheta(L \geq t) \qquad \text{for } \vartheta \in \Theta \text{ and } t \in \,]-\infty, \kappa(\vartheta)[,$$

strictly better (or **strictly superior**) if additionally strict inequality holds for at least one such pair (ϑ, t), and **equivalent** if each is better than the other. L is called **admissible** if there is no strictly better confidence bound for κ to the level β.

A.3.20 Remark Admissibility of L is the same as superiority of L over L' for all confidence bounds L' (for κ and to level β) that are better than L.

A.3.21 Theorem *Let $\mathcal{P} = (P_\vartheta\colon \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$, $\kappa\colon \Theta \to \Gamma$ a parameter of interest for $\mathcal{P}$, and $\beta \in [0, 1]$.*

(i) *Let R be a confidence region for id_Θ, $\mathrm{K} \coloneqq \kappa[R]$, and $\{\mathrm{K} \ni \kappa(\vartheta)\} \in \mathfrak{A}$ for $\vartheta \in \Theta$. Then K is a confidence region for κ with effective level $\beta_{\mathrm{eff}}(\mathrm{K}) \geq \beta_{\mathrm{eff}}(R)$, and the following implication holds:*

$$\kappa^{-1}[\mathrm{K}(x)] \subseteq R(x) \text{ for } x \in \mathcal{X} \quad \Longrightarrow \quad \beta_{\mathrm{eff}}(\mathrm{K}) = \beta_{\mathrm{eff}}(R).$$

(ii) *Let K be a confidence region for κ and $R \coloneqq \kappa^{-1}[\mathrm{K}]$. Then R is a confidence region for id_Θ with effective level $\beta_{\mathrm{eff}}(R) = \beta_{\mathrm{eff}}(\mathrm{K})$.*

Proof. (i) Since $\vartheta \in R(x)$ implies $\kappa(\vartheta) \in \kappa[R(x)] = \mathrm{K}(x)$ for $\vartheta \in \Theta$ and $x \in \mathcal{X}$, we have $\{\mathrm{K} \ni \kappa(\vartheta)\} \supseteq \{R \ni \vartheta\}$ and thus $P_\vartheta(\mathrm{K} \ni \kappa(\vartheta)) \geq P_\vartheta(R \ni \vartheta)$ for $\vartheta \in \Theta$. This yields the first claim. The second claim follows from $\{\mathrm{K} \ni \kappa(\vartheta)\} \subseteq \{R \ni \vartheta\}$ (in addition to the reverse inclusion just shown) since $\kappa(\vartheta) \in \mathrm{K}(x)$ implies $\vartheta \in \kappa^{-1}[\mathrm{K}(x)] \subseteq R(x)$ for $x \in \mathcal{X}$ and $\vartheta \in \Theta$.

(ii) Since $\vartheta \in R(x) = \kappa^{-1}[\mathrm{K}(x)]$ is equivalent to $\kappa(\vartheta) \in \mathrm{K}(x)$ for $\vartheta \in \Theta$ and $x \in \mathcal{X}$, we have $\{R \ni \vartheta\} = \{\mathrm{K} \ni \kappa(\vartheta)\} \in \mathfrak{A}$ and $P_\vartheta(R \ni \vartheta) = P_\vartheta(\mathrm{K} \ni \kappa(\vartheta))$ for $\vartheta \in \Theta$. □

A.3.22 Remark

1. Remark A.1.15 yields the converse inclusion $\kappa^{-1}[\mathrm{K}(x)] \supseteq R(x)$ for $x \in \mathcal{X}$ in Theorem A.3.21(i).

2. If $(\mathcal{X}, \leq)$ is preordered, the monotonicity behaviours of R and K with respect to set inclusion on 2^Θ and 2^Γ, respectively, coincide.

A.3.23 Theorem *Let $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$, $\kappa \colon \Theta \to \Gamma$ a parameter of interest, and $\mathcal{K}$ a countable set of confidence regions K for κ with respective levels $\beta(\mathrm{K}) \in [0,1]$ such that*

$$\alpha := \sum_{\mathrm{K} \in \mathcal{K}} (1 - \beta(\mathrm{K}))$$

converges in $[0,1]$. Then

$$\bigcap \mathcal{K} \colon \mathcal{X} \to 2^\Gamma, \quad x \mapsto \bigcap_{\mathrm{K} \in \mathcal{K}} \mathrm{K}(x),$$

is a confidence region for κ with level $\beta := 1 - \alpha$.

Proof. For $\vartheta \in \Theta$ we have

$$\left\{ \bigcap \mathcal{K} \ni \kappa(\vartheta) \right\} = \bigcap_{\mathrm{K} \in \mathcal{K}} \{\mathrm{K} \ni \kappa(\vartheta)\} \in \mathfrak{A}$$

and

$$P_\vartheta \left(\bigcap \mathcal{K} \not\ni \kappa(\vartheta) \right) = P_\vartheta \left(\bigcup_{\mathrm{K} \in \mathcal{K}} \{\mathrm{K} \not\ni \kappa(\vartheta)\} \right) \leq \sum_{\mathrm{K} \in \mathcal{K}} (1 - \beta(\mathrm{K})). \qquad \square$$

A.3.24 Definition Let $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$, $\kappa \colon \Theta \to \Gamma$ a parameter of interest for $\mathcal{P}$, and $\mathrm{B} \subseteq [0,1]$. A family $(\mathrm{K}_\beta \colon \beta \in \mathrm{B})$ of confidence regions K_β for κ is called **nested** if

$$\mathrm{B} \to 2^\Gamma, \quad \beta \mapsto \mathrm{K}_\beta(x),$$

is increasing for $x \in \mathcal{X}$, i.e., if

$$\beta_1, \beta_2 \in \mathrm{B}, \quad \beta_1 \leq \beta_2 \qquad \Longrightarrow \qquad \mathrm{K}_{\beta_1}(x) \subseteq \mathrm{K}_{\beta_2}(x) \quad \text{for } x \in \mathcal{X}.$$

A.3.25 Remark If a family $(\mathrm{K}_\beta \colon \beta \in \mathrm{B})$ of confidence regions K_β is nested, then

$$\mathrm{B} \to [0,1], \quad \beta \mapsto \beta_{\mathrm{eff}}(\mathrm{K}_\beta),$$

is increasing.

A.3.26 Definition Let $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$ and $\kappa \colon \Theta \to \Gamma$ a parameter of interest. A confidence region $\mathrm{K} \colon \mathcal{X} \to 2^\Gamma$ for κ is called **unbiased** if

$$P_\vartheta(\mathrm{K} \ni \kappa(\vartheta)) \geq P_\vartheta(\mathrm{K} \ni \kappa(\vartheta')) \qquad \text{for } \vartheta, \vartheta' \in \Theta.$$

A.3.27 Definition Let $\mathcal{X}$ be a set. A set $\mathcal{G} \subseteq \mathcal{X}^{\mathcal{X}}$ of bijective functions $\mathcal{X} \to \mathcal{X}$ that is a group with respect to composition $\circ$ of functions is called a **transformation group** on $\mathcal{X}$. In such a case, members of the group $\mathcal{G}$ are called **transformations** of $\mathcal{X}$.

A.3.28 Example 1. A frequently occurring transformation group on $\mathbb{R}$ is given by

$$\mathcal{N} := \{\mathbb{R} \to \mathbb{R},\ x \mapsto \sigma x + \mu \colon (\mu, \sigma) \in \mathbb{R} \times]0, \infty[\};$$

its generalization to $\mathbb{R}^n$ is

$$\{\mathbb{R}^n \to \mathbb{R}^n,\ x \mapsto \Sigma x + \mu \colon \\ \mu \in \mathbb{R}^n,\ \Sigma \in \mathbb{R}^{n \times n} \text{ symmetric and positive definite}\}.$$

A transformation group on $\mathbb{R}^n$ that is more easily to handle is

$$\mathcal{N}' := \{\mathbb{R}^n \to \mathbb{R}^n,\ x \mapsto x + \mu \colon \mu \in \mathbb{R}^n\}.$$

2. A transformation group on $\{0, \ldots, n\}$ is given by

$$\mathcal{B} := \{\mathrm{id}_{\{0,\ldots,n\}}, \quad (\{0, \ldots, n\} \to \{0, \ldots, n\},\ x \mapsto n - x)\}.$$

A.3.29 Remark Let $\mathcal{G}$ be a transformation group on a set $\mathcal{X}$.

1. A function $f \colon \mathcal{X} \to \mathcal{Y}$ is called **invariant** over $\mathcal{G}$ if $f \circ g = f$ for every $g \in \mathcal{G}$; it is called **equivariant** over $\mathcal{G}$, if $\mathcal{G}$ “induces” in some way a second transformation group $\mathcal{G}'$ on $\mathcal{Y}$ such that to every $g \in \mathcal{G}$ corresponds some $g' \in \mathcal{G}'$ with $f \circ g = g' \circ f$.
2. If we define

$$x \equiv_{\mathcal{G}} y \quad :\Longleftrightarrow \quad g(x) = y \quad \text{for some } g \in \mathcal{G},$$

then $\equiv_{\mathcal{G}}$ is an equivalence relation on $\mathcal{X}$. In fact, the existence of a neutral element in $\mathcal{G}$ yields the reflexivity, and the closedness with respect to building inverses and compositions the symmetry and the transitivity, respectively.
3. The *equivalence class* $\{\xi \in \mathcal{X} \colon \xi \equiv_{\mathcal{G}} x\}$ of $x \in \mathcal{X}$ is called the **orbit** of x.
4. Functions that are invariant over $\mathcal{G}$ are obviously constant on every orbit.
5. A function $F \colon \mathcal{X} \to \mathcal{Y}$ is called **maximal invariant** over $\mathcal{G}$ if the following equivalence holds:

$$F(x_1) = F(x_2) \quad \Longleftrightarrow \quad x_1 \equiv_{\mathcal{G}} x_2 \qquad \text{for } x_1, x_2 \in \mathcal{X}.$$

6. Let $F \colon \mathcal{X} \to \mathcal{Y}$ be maximal invariant over $\mathcal{G}$. A function $f \colon \mathcal{X} \to \mathcal{Z}$ is then invariant over $\mathcal{G}$ if, and only if, there is a function $f' \colon \mathcal{Y} \to \mathcal{Z}$ with $f = f' \circ F$.

Proof. Let us first assume the invariance of f, and let us define

$$f' \colon \mathcal{Y} \to \mathcal{Z}, \quad y \mapsto \begin{cases} f(x) & \text{if } y = F(x) \text{ for some } x \in \mathcal{X} \\ z_0 & \text{otherwise,} \end{cases}$$

with an arbitrary $z_0 \in \mathcal{Z}$. Then f' is well-defined since $F(x_1) = F(x_2)$ implies $x_1 \equiv_{\mathcal{G}} x_2$ (due to the maximal invariance of F), which in turn implies $f(x_1) = f(x_2)$ (due to the invariance of f). Trivially, we have $f' \circ F = f$.
Let us now assume the existence of a function f' as above, and let $g \in \mathcal{G}$ and $x \in \mathcal{X}$. Since $g(x) \equiv_{\mathcal{G}} x$ and F is maximal invariant, we have $F(g(x)) = F(x)$, hence $f(g(x)) = f'(F(g(x))) = f'(F(x)) = f(x)$. Thus, f is invariant over $\mathcal{G}$. □

A.3.30 Definition Let $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$ and $\mathcal{G}$ a transformation group of measurable functions on $\mathcal{X}$. The model $\mathcal{P}$ is called **invariant** over $\mathcal{G}$ if for $g \in \mathcal{G}$ and $\vartheta \in \Theta$ the distribution $g \,\square\, P_\vartheta$ of g under P_ϑ is again a member of the model $\mathcal{P}$, i.e., if to every $g \in \mathcal{G}$ and every $\vartheta \in \Theta$ corresponds a $\overline{g}(\vartheta) \in \Theta$ with $g \,\square\, P_\vartheta = P_{\overline{g}(\vartheta)}$.

If $\mathcal{P}$ is injective and invariant over $\mathcal{G}$, then, given a transformation $g \in \mathcal{G}$, the thus well-defined mapping $\overline{g} \colon \Theta \to \Theta$ is called **induced** by g; if, furthermore, $\overline{\mathcal{G}} := \{\overline{g} \in \Theta^\Theta \colon g \in \mathcal{G}\}$ is a transformation group on Θ, then $\overline{\mathcal{G}}$ is called **induced** by $\mathcal{G}$.

A.3.31 Remark If an injective model $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ is invariant over a transformation group $\mathcal{G}$ on the sample space which induces a transformation group $\overline{\mathcal{G}}$ on the parameter space Θ, then

$$\overline{\mathrm{id}_{\mathcal{X}}} = \mathrm{id}_\Theta, \quad \overline{g \circ h} = \overline{g} \circ \overline{h}, \quad \text{and} \quad \overline{g}^{-1} = \overline{g^{-1}} \qquad \text{for } g, h \in \mathcal{G}.$$

The first equation is obvious. The second one follows from $P_{\overline{g \circ h}(\vartheta)} = (g \circ h) \,\square\, P_\vartheta = g \,\square\, (h \,\square\, P_\vartheta) = g \,\square\, P_{\overline{h}(\vartheta)} = P_{\overline{g}(\overline{h}(\vartheta))} = P_{\overline{g} \circ \overline{h}(\vartheta)}$ for $\vartheta \in \Theta$ and the injectivity of the model $\mathcal{P}$. The following equivalences for $\vartheta, \eta \in \Theta$ yield the third one:

$$\begin{aligned} \overline{g}^{-1}(\eta) = \vartheta \quad &\iff \quad \overline{g}(\vartheta) = \eta \\ &\iff \quad g \,\square\, P_\vartheta = P_\eta \\ &\iff \quad g^{-1} \,\square\, (g \,\square\, P_\vartheta) = g^{-1} \,\square\, P_\eta \\ &\iff \quad P_\vartheta = g^{-1} \,\square\, P_\eta \\ &\iff \quad \overline{g^{-1}}(\eta) = \vartheta. \end{aligned}$$

A.3.32 Example 1. The normal model $(\mathrm{N}_{\mu,\sigma} \colon (\mu, \sigma) \in \mathbb{R} \times]0, \infty[)$ is injective and invariant over the transformation group $\mathcal{N}$ from Example A.3.28. The induced transformation group is

$$\overline{\mathcal{N}} = \{\mathbb{R} \times]0, \infty[\to \mathbb{R} \times]0, \infty[,\ (\nu, \tau) \mapsto (\mu + \nu, \sigma^2 \tau^2) \colon (\mu, \sigma) \in \mathbb{R} \times]0, \infty[\}.$$

2. The normal model $(\bigotimes_{i=1}^n \mathrm{N}_{\mu_i, \sigma_i^2}^{\otimes m_i} \colon \mu \in \mathbb{R}^n)$ from Section 2.2 of Chapter 2 is also injective and invariant over the transformation group $\mathcal{N}'$ from Example A.3.28. The induced transformation group is

$$\overline{\mathcal{N}'} = \mathcal{N}'.$$

3. The binomial model $(\mathrm{B}_{n,p} \colon p \in [0, 1])$ is injective and invariant over the transformation group $\mathcal{B}$ from Example A.3.28. The induced transformation group is

$$\overline{\mathcal{B}} = \{\mathrm{id}_{[0,1]}, \quad ([0, 1] \to [0, 1],\ p \mapsto 1 - p)\}.$$

A.3.33 Definition Let $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ be an injective model on a measurable space $(\mathcal{X}, \mathfrak{A})$, $\mathcal{G}$ a transformation group on $\mathcal{X}$ inducing a transformation group $\overline{\mathcal{G}}$ on Θ, and $\kappa \colon \Theta \to \Gamma$ a parameter of interest. A confidence region $\mathrm{K} \colon \mathcal{X} \to 2^\Gamma$ is called **equivariant** over $\mathcal{G}$ if

$$\mathrm{K}(g(x)) = \overline{g}[\mathrm{K}(x)] \qquad \text{for } x \in \mathcal{X} \text{ and } g \in \mathcal{G}.$$

A.3.34 Example 1. The confidence regions from Example 2.2.13 given by

$$R_{\overline{X},\beta}(x) = \left\{\mu \in \mathbb{R}^n : \overline{\mu} < \overline{x} + \sqrt{\frac{\overline{\sigma^2/m}}{n}}\Phi^{-1}(\beta)\right\}$$

$$\tilde{R}_{\overline{X},\beta}(x) = \left\{\mu \in \mathbb{R}^n : \overline{\mu} > \overline{x} - \sqrt{\frac{\overline{\sigma^2/m}}{n}}\Phi^{-1}(\beta)\right\}$$

for $x \in \prod_{i=1}^n \mathbb{R}^{m_i}$ in the model $(\bigotimes_{i=1}^n \mathrm{N}_{\mu_i,\sigma_i^2}^{\otimes m_i} : \mu \in \mathbb{R}^n)$ from Section 2.2 of Chapter 2 are clearly equivariant over the transformation group $\mathcal{N}'$ from Example A.3.28.

2. A confidence region $\mathrm{K}\colon \{0,\dots,n\} \to 2^{[0,1]}$ for the identity $\mathrm{id}_{[0,1]}$ in the binomial model $(\mathrm{B}_{n,p}\colon p \in [0,1])$ is equivariant if, and only if, it satisfies

$$\mathrm{K}(n-x) = \{1-p \in [0,1] \colon p \in \mathrm{K}(x)\} \qquad \text{for } x \in \{0,\dots,n\}.$$

A.3.35 Definition Let $\mathcal{P} = (P_\vartheta\colon \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X},\mathfrak{A})$ and $\Theta_0 \subseteq \Theta$. A measurable function (with respect to $\mathfrak{A}$ and $2^{\{0,1\}}$)

$$\psi\colon \mathcal{X} \to \{0,1\}$$

is called a **test** for the **hypothesis** Θ_0 (in $\mathcal{P}$). Its **effective level** is the number

$$\alpha_{\mathrm{eff}}(\psi) := \sup_{\vartheta\in\Theta_0} P_\vartheta(\psi).$$

The complement $\Theta \setminus \Theta_0$ of the hypothesis is called **alternative.** Given $\alpha \in [0,1]$, the test ψ is said to **preserve** the (or simply **have**) **level** α if $\alpha_{\mathrm{eff}}(\psi) \le \alpha$.

A.3.36 Remark 1. Tests are used to reject hypotheses as follows: after observing $x \in \mathcal{X}$, the hypothesis is rejected if $\psi(x) = 1$.

2. A **randomized test** is a measurable function $\psi\colon \mathcal{X} \to [0,1]$. After observing $x \in \mathcal{X}$, the hypothesis is rejected with probability $\psi(x)$. The decision is thus dependent on the outcome of yet another experiment.
3. The next result is an analogue of Theorem A.3.23.

A.3.37 Theorem *Let $\mathcal{P} = (P_\vartheta\colon \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X},\mathfrak{A})$ and Ψ a countable set of tests ψ for the respective hypotheses $\Theta_0(\psi)$ with respective levels $\alpha(\psi) \in [0,1]$ such that*

$$\alpha := \sum_{\psi\in\Psi} \alpha(\psi)$$

converges in $[0,1]$. *Then*

$$\sup_{\psi\in\Psi} \psi = \mathbf{1}_{\bigcup_{\psi\in\Psi}\{\psi=1\}}$$

is a test for the hypothesis $\bigcap_{\psi\in\Psi}\Theta_0(\psi)$ with level α.

Proof. For $\vartheta \in \bigcap_{\psi\in\Psi}\Theta_0(\psi)$ we have

$$P_\vartheta\left(\sup_{\psi\in\Psi}\psi\right) = P_\vartheta\left(\bigcup_{\psi\in\Psi}\{\psi=1\}\right) \le \sum_{\psi\in\Psi}\alpha(\psi). \qquad \square$$

A.3.38 Definition Let $\mathcal{P} = (P_\vartheta : \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$ and $\Theta_0 \subseteq \Theta$. A test ψ for the hypothesis Θ_0 is called **unbiased** if

$$P_\vartheta(\psi) \leq P_{\vartheta'}(\psi) \qquad \text{for } \vartheta \in \Theta_0 \text{ and } \vartheta' \in \Theta \setminus \Theta_0.$$

A.3.39 Theorem *Let $\mathcal{P} = (P_\vartheta : \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$, $\kappa : \Theta \to \Gamma$ a parameter of interest for $\mathcal{P}$, and $\alpha \in [0, 1]$.*

(i) *To every confidence region* K *for κ with level $1 - \alpha$ corresponds a family* $\mathsf{test}(\mathrm{K}) = (\mathsf{test}(\mathrm{K})_\gamma : \gamma \in \Gamma)$ *of tests*

$$\mathsf{test}(\mathrm{K})_\gamma : \mathcal{X} \to \{0, 1\}, \quad x \mapsto \mathbf{1}(\gamma \notin \mathrm{K}(x)),$$

for the respective hypotheses $\kappa^{-1}[\{\gamma\}]$ with level α.

(ii) *To every family $\psi = (\psi_\gamma : \gamma \in \Gamma)$ of tests ψ_γ for the respective hypotheses $\kappa^{-1}[\{\gamma\}]$ with respective levels α corresponds a confidence region*

$$\mathsf{conf}(\psi) : \mathcal{X} \to 2^\Gamma, \quad x \mapsto \{\gamma \in \Gamma : \psi_\gamma(x) = 0\},$$

for κ with level $1 - \alpha$.

(iii) *The thus well-defined maps* test *and* conf *are bijective and inverse to each other.*

Proof. (i) Let K be a confidence region for κ with level $1 - \alpha$, $\gamma \in \Gamma$, and

$$\psi : \mathcal{X} \to \{0, 1\}, \quad x \mapsto \mathbf{1}(\gamma \notin \mathrm{K}(x)).$$

Since $\{\mathrm{K} \not\ni \gamma\} \in \mathfrak{A}$, ψ is measurable. If $\kappa^{-1}[\{\gamma\}] = \emptyset$, then $\alpha_{\mathrm{eff}}(\psi) = 0 \leq \alpha$. In the other case we obtain

$$P_\vartheta(\psi) = P_\vartheta(\psi = 1) = P_\vartheta(\mathrm{K} \not\ni \kappa(\vartheta)) \leq \alpha \qquad \text{for } \vartheta \in \kappa^{-1}[\{\gamma\}],$$

yielding $\alpha_{\mathrm{eff}}(\psi) \leq \alpha$.

(ii) Let, for $\gamma \in \Gamma$, ψ_γ be a test for the hypothesis $\kappa^{-1}[\{\gamma\}]$ and let us define

$$\mathrm{K} : \mathcal{X} \to 2^\Gamma, \quad x \mapsto \{\gamma \in \Gamma : \psi_\gamma(x) = 0\}.$$

We have $\{\mathrm{K} \ni \gamma\} = \{\psi_\gamma = 0\} \in \mathfrak{A}$ for $\gamma \in \Gamma$. Moreover,

$$P_\vartheta(\mathrm{K} \ni \kappa(\vartheta)) = P_\vartheta(\psi_{\kappa(\vartheta)} = 0) = 1 - P_\vartheta(\psi_{\kappa(\vartheta)}) \geq 1 - \alpha \qquad \text{for } \vartheta \in \Theta,$$

yielding $\beta_{\mathrm{eff}}(\mathrm{K}) \geq 1 - \alpha$.

(iii) $\mathsf{test}^{-1} = \mathsf{conf}$ is obvious. □

A.3.40 Remark 1. Let K be a confidence region for κ. After observing $x \in \mathcal{X}$, a hypothesis $\Theta_0 \subseteq \Theta$ is rejected by $\mathsf{test}(\mathrm{K})$ if $\mathrm{K}(x) \cap \kappa[\Theta_0] = \emptyset$. In fact, $\Theta_0 \subseteq \bigcup_{\vartheta \in \Theta_0} \kappa^{-1}[\{\kappa(\vartheta)\}]$ is rejected if $\kappa^{-1}[\{\kappa(\vartheta)\}]$ is for $\vartheta \in \Theta_0$, i.e., if $\kappa(\vartheta) \notin \mathrm{K}(x)$ for $\vartheta \in \Theta_0$.

2. $\mathrm{K}(x) \cap \kappa[\Theta_0] = \emptyset$ is equivalent to $\kappa^{-1}[\mathrm{K}(x)] \cap \Theta_0 = \emptyset$:
   - if $\vartheta \in \kappa^{-1}[\mathrm{K}(x)] \cap \Theta_0$, then $\vartheta \in \Theta_0$ and $\kappa(\vartheta) \in \mathrm{K}(x)$, i.e., $\kappa(\vartheta) \in \mathrm{K}(x) \cap \kappa[\Theta_0]$;
   - if $\gamma \in \mathrm{K}(x) \cap \kappa[\Theta_0]$, then there is some $\vartheta \in \Theta_0$ with $\gamma = \kappa(\vartheta)$, i.e., $\vartheta \in \kappa^{-1}[\mathrm{K}(x)] \cap \Theta_0$.

A.3.41 Definition Let $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$ and $\Theta_0 \subseteq \Theta$. A statistic $\widehat{\alpha} \colon \mathcal{X} \to [0,1]$ such that

$$\sup_{\vartheta \in \Theta_0} P_\vartheta(\widehat{\alpha} \le u) \le u \qquad \text{for } u \in [0,1]$$

is called **P-variable** for the hypothesis Θ_0.

A.3.42 Remark 1. Using the language of stochastic preorders (see Definition A.3.46), P-variables are statistics $\widehat{\alpha}$ that are, under the hypothesis, stochastically greater than uniformly distributed statistics on $[0,1]$, i.e., $\mathrm{U}_{[0,1]} \le_{\mathrm{st}} \widehat{\alpha} \,\square\, P_\vartheta$ for $\vartheta \in \Theta_0$.

2. After having fixed a level $\alpha \in [0,1]$, the hypothesis Θ_0 is rejected on the basis of an observation $x \in \mathcal{X}$ if $\widehat{\alpha}(x) \le \alpha$.
3. As Theorem A.3.45 shows, P-variables correspond to certain families of tests. Testing a hypothesis can therefore usually be done with a multitude of different P-variables.
4. Many statisticians call "P-value" what we have defined as P-variable. We reserve the term **P-value** for a realization $\widehat{\alpha}(x)$ of a P-variable $\widehat{\alpha}$.

A.3.43 Definition Let $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$, $\Theta_0 \subseteq \Theta$, and $\mathrm{A} \subseteq [0,1]$. A family $(\psi_\alpha \colon \alpha \in \mathrm{A})$ of tests ψ_α for the hypothesis Θ_0 is called **nested** if

$$\mathrm{A} \to \{0,1\}, \quad \alpha \mapsto \psi_\alpha(x),$$

is increasing for $x \in \mathcal{X}$, i.e., if

$$\alpha_1, \alpha_2 \in \mathrm{A}, \quad \alpha_1 \le \alpha_2 \qquad \Longrightarrow \qquad \psi_{\alpha_1}(x) \le \psi_{\alpha_2}(x) \quad \text{for } x \in \mathcal{X}.$$

A.3.44 Remark If a family $(\psi_\alpha \colon \alpha \in \mathrm{A})$ of tests ψ_α is nested, then

$$\mathrm{A} \to [0,1], \quad \beta \mapsto \alpha_{\mathrm{eff}}(\psi_\alpha),$$

is increasing.

A.3.45 Theorem *Let $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ be a model on a measurable space $(\mathcal{X}, \mathfrak{A})$, $\Theta_0 \subseteq \Theta$, and $\mathrm{A} \subseteq [0,1]$.*

(i) *To every P-variable $\widehat{\alpha}$ for the hypothesis Θ_0 corresponds a nested family* $\mathsf{test}(\widehat{\alpha}) = (\mathsf{test}(\widehat{\alpha})_\alpha \colon \alpha \in \mathrm{A})$ *of tests*

$$\mathsf{test}(\widehat{\alpha})_\alpha \colon \mathcal{X} \to [0,1], \quad x \mapsto \mathbf{1}(\widehat{\alpha}(x) \le \alpha),$$

for the hypothesis Θ_0 with respective levels α.

(ii) *To every nested family $\psi = (\psi_\alpha \colon \alpha \in \mathrm{A})$ of tests ψ_α for the hypothesis Θ_0 with respective levels α corresponds a P-variable*

$$\mathsf{p\text{-}var}(\psi) \colon \mathcal{X} \to [0,1], \quad x \mapsto \inf\{\alpha \in \mathrm{A} \colon \psi_\alpha(x) = 1\},$$

for the hypothesis Θ_0.

Proof. See Mattner (2014). □

A.3.46 Definition Let $(\mathcal{X}, \leq)$ be a preordered set, endowed with the σ-algebra generated by the order topology on $\mathcal{X}$. For two laws P and Q on $\mathcal{X}$ let

$$\begin{aligned} P \leq_{\mathrm{st}} Q \quad &:\iff \quad P(U) \leq Q(U) \quad \text{for uprays } U \subseteq \mathcal{X} \\ P <_{\mathrm{st}} Q \quad &:\iff \quad P(U) < Q(U) \quad \text{for uprays } U \subseteq \mathcal{X} \text{ with } P(\mathcal{X} \setminus U) \wedge Q(U) > 0. \end{aligned}$$

$\leq_{\mathrm{st}}$ and $<_{\mathrm{st}}$ are called **stochastic preorder** and **strict stochastic preorder** on $\mathcal{X}$, respectively.

A.3.47 Remark

1. $\leq_{\mathrm{st}}$ and $<_{\mathrm{st}}$ of course depend on the underlying set $\mathcal{X}$, as well as its preorder $\leq$. The former is, however, not reflected in our notation since it is in all considered cases clear from the context.
2. $\leq_{\mathrm{st}}$ and $<_{\mathrm{st}}$ are a preorder and a strict preorder, respectively, on the set of laws on $\mathcal{X}$.
3. We have $(\leq_{\mathrm{st}})^{\mathrm{str}} \supseteq <_{\mathrm{st}}$ and $(<_{\mathrm{st}})^{\mathrm{unstr}} \subseteq \leq_{\mathrm{st}}$.
4. Models $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ on preordered sets $(\mathcal{X}, \leq)$ having preordered parameter sets $(\Theta, \leq)$ can thus possess monotonicity properties. For strict monotonicity, however, the strict stochastic preorder $<_{\mathrm{st}}$ instead of $(\leq_{\mathrm{st}})^{\mathrm{str}}$ is considered. For instance, $\mathcal{P}$ is said to be **stochastically increasing** if

$$\vartheta_1 \leq \vartheta_2 \quad \Longrightarrow \quad P_{\vartheta_1} \leq_{\mathrm{st}} P_{\vartheta_2} \qquad \text{for } \vartheta_1, \vartheta_2 \in \Theta$$

and **stochastically strictly increasing** if

$$\vartheta_1 < \vartheta_2 \quad \Longrightarrow \quad P_{\vartheta_1} <_{\mathrm{st}} P_{\vartheta_2} \qquad \text{for } \vartheta_1, \vartheta_2 \in \Theta.$$

A.3.48 Theorem *Let $(\mathcal{X}, \leq)$ and $(\mathcal{Y}, \leq)$ be two preordered sets, $\prec \in \{\leq_{\mathrm{st}}, <_{\mathrm{st}}\}$, and let P and Q be two laws on $\mathcal{X}$ with $P \prec Q$ and $T \colon \mathcal{X} \to \mathcal{Y}$ a monotonic function. Then*

$$\begin{aligned} T \text{ increasing} \quad &\Longrightarrow \quad T \,\square\, P \prec T \,\square\, Q \\ T \text{ decreasing} \quad &\Longrightarrow \quad T \,\square\, Q \prec T \,\square\, P. \end{aligned}$$

A.3.49 Theorem *Let P_i and Q_i be laws on preordered sets $(\mathcal{X}_i, \leq_i)$ with $P_i \prec Q_i$ for $i \in \{1, \ldots, n\}$ and $\prec \in \{\leq_{\mathrm{st}}, <_{\mathrm{st}}\}$. Then $\bigotimes_{i=1}^n P_i \prec \bigotimes_{i=1}^n Q_i$.*

A.3.50 Example Let $n \in \mathbb{N}$ and $m \in \mathbb{N}^n$.

1. The family $(\bigotimes_{i=1}^n \mathrm{B}_{m_i, p_i} \colon p \in [0,1]^n)$ is stochastically strictly increasing. This follows from Theorems A.3.48 and A.3.49, the monotonicity of $\{0,1\}^n \to \{0, \ldots, n\}$, $x \mapsto \sum_{k=1}^n x_k$, and $\mathrm{B}_p <_{\mathrm{st}} \mathrm{B}_q$ for $p, q \in [0,1]$ with $p < q$.
2. For $\sigma \in \,]0, \infty[$, the family $(\bigotimes_{i=1}^n \mathrm{N}_{\mu_i, \sigma^2}^{\otimes m_i} \colon \mu \in \mathbb{R}^n)$ is stochastically strictly increasing. This follows from Theorem A.3.49 and $\mathrm{N}_{\mu, \sigma^2} <_{\mathrm{st}} \mathrm{N}_{\nu, \sigma^2}$ for $\mu, \nu \in \mathbb{R}$ with $\mu < \nu$.
3. The family $(\bigotimes_{i=1}^n \mathrm{E}_{\lambda_i}^{\otimes m_i} \colon \lambda \in \,]0, \infty[^n)$ is stochastically strictly decreasing. This follows from Theorem A.3.49, and $\mathrm{E}_\lambda <_{\mathrm{st}} \mathrm{E}_\mu$ for $\lambda, \mu \in \,]0, \infty[$ with $\lambda > \mu$.
4. The family $(\bigotimes_{i=1}^n \mathrm{P}_{\lambda_i}^{\otimes m_i} \colon \lambda \in \,]0, \infty[^n)$ is stochastically strictly increasing. This follows from Theorem A.3.49, and $\mathrm{P}_\lambda <_{\mathrm{st}} \mathrm{P}_\mu$ for $\lambda, \mu \in \,]0, \infty[$ with $\lambda < \mu$.

A.3.51 Definition Let $(\Theta, \leq)$ and $(\mathcal{Y}, \leq)$ be preordered sets and $T \colon \mathcal{X} \to \mathcal{Y}$ a statistic. A model $\mathcal{P} = (P_\vartheta \colon \vartheta \in \Theta)$ on a measurable space $(\mathcal{X}, \mathfrak{A})$ is said to have **(strictly) increasing/decreasing likelihood ratios** in T if for $\vartheta_0, \vartheta_1 \in \Theta$ with $\vartheta_0 < \vartheta_1$ there are densities f_{ϑ_i} of P_{ϑ_i} for $i \in \{0,1\}$ with respect to some measure μ on $(\mathcal{X}, \mathfrak{A})$ such that

$$\frac{f_{\vartheta_1}}{f_{\vartheta_0}} \quad \text{is } (P_{\vartheta_0} + P_{\vartheta_1})\text{-a.s. (strictly) increasing/decreasing in } T.$$

A.3.52 Remark If $\mathcal{P}$ has increasing/decreasing likelihood ratios in T, then $T \Box \mathcal{P}$ is stochastically increasing/decreasing.

A.3.53 Definition Let $\mu \in \mathbb{R}$, $n \in \mathbb{N} \setminus \{1\}$, $X \coloneqq \mathrm{id}_{\mathbb{R}^n}$, and

$$\overline{X} \coloneqq \frac{1}{n} \sum_{i=1}^{n} X_i \quad \text{and} \quad S \coloneqq \left(\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2 \right)^{1/2}.$$

The laws

$$\chi^2_{n-1} \coloneqq ((n-1)S^2) \Box \mathrm{N}_{0,1}^{\otimes n} \qquad \text{and} \qquad \mathrm{t}_{n-1,\mu} \coloneqq \sqrt{n}\frac{\overline{X}}{S} \Box \mathrm{N}_{\mu,1}^{\otimes n}$$

are called χ^2**-distribution** (with $n-1$ **degrees of freedom**) and **noncentral t-distribution** (with $n-1$ **degrees of freedom** and **noncentrality parameter** μ), respectively.

A.3.54 Remark

1. We have $((n-1)S^2/\sigma^2) \Box \mathrm{N}_{\mu,\sigma^2}^{\otimes n} = \chi^2_{n-1}$ and $(\sqrt{n} \cdot \overline{X}/S) \Box \mathrm{N}_{\mu,\sigma^2}^{\otimes n} = \mathrm{t}_{n-1,\mu/\sigma}$ for $\mu \in \mathbb{R}$ and $\sigma \in \,]0, \infty[$.
2. If $\tilde{S} \colon \mathbb{R}^n \to \mathbb{R}$, $x \mapsto \sqrt{\sum_{i=1}^n (x_i - \mu)^2}$, then $(\tilde{S}^2/\sigma^2) \Box \mathrm{N}_{\mu,\sigma^2}^{\otimes n} = \chi^2_n$. The law χ^2_n is usually defined this way.
3. If $T \colon \mathbb{R} \times \,]0, \infty[\,\to \mathbb{R}$, $(x, y) \mapsto x/\sqrt{y/n}$, then $T \Box (\mathrm{N}_{\mu,1} \otimes \chi^2_n) = \mathrm{t}_{n,\mu}$. The law $\mathrm{t}_{n,\mu}$ is usually defined this way.
4. The laws χ^2_{n-1} and $\mathrm{t}_{n-1,\mu}$ are continuous and with support $\operatorname{supp} \chi^2_{n-1} = [0, \infty[$ and $\operatorname{supp} \mathrm{t}_{n-1,\mu} = \mathbb{R}$, respectively. For $\mu \in \mathbb{R} \setminus \{0\}$, $\mathrm{t}_{n-1,\mu}$ is asymmetric.
5. We have $\chi^2_{n-1} \rightsquigarrow \delta_\infty$ for $n \to \infty$, and $t_{n-1,\mu} \rightsquigarrow \delta_{\pm\infty}$ for $\mu \to \pm\infty$. This follows straightly from the definition and $\mathrm{t}_{n-1,\mu} = (\sqrt{n}\frac{\overline{X}+\mu}{S}) \Box \mathrm{N}_{0,1}^{\otimes n}$, respectively.
6. It follows that $\mathbb{R} \to \,]0, 1[$, $\mu \mapsto \mathrm{t}_{n-1,\mu}(]-\infty, x])$, is strictly increasing and surjective for $x \in \mathbb{R}$.

BIBLIOGRAPHY

Aczél J. (1966). *Lectures on Functional Equations and Their Applications.* Academic Press, New York.

Aczél J. (1984), editor. *Functional Equations: History, Applications and Theory.* D. Reidel Publishing Company, Dordrecht.

Aczél J. and Dhombres J. (1989). *Functional Equations in Several Variables.* Cambridge University Press, Cambridge.

Agresti A. and Coull B.A. (1998). Approximate is better than “exact” for interval estimation of binomial proportions. *Amer. Statist.,* **52**(2), 119–126.

Agresti A. and Gottard A. (2005). Comment: Randomized confidence intervals and the mid-P approach. *Statist. Sci.,* **20**(4), 367–371.

Akkouchi M. (2008). On the convolution of exponential distributions. *J. Chungcheong Math. Soc.,* **21**(4), 501–510.

Altman D.G. and Bland J.M. (1994). Diagnostic tests 1: sensitivity and specificity. *BMJ: British Medical Journal,* **308**(6943), 1552.

Altman D.G. and Bland J.M. (1994). Diagnostic tests 2: predictive values. *BMJ: British Medical Journal,* **309**(6947), 102.

Altman D.G. and Bland J.M. (1994). Diagnostic tests 3: receiver operating characteristic plots. *BMJ: British Medical Journal,* **309**(6948), 188.

Anderson T.W. and Samuels S.M. (1967). Some inequalities among binomial and Poisson probabilities. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability,* Volume 1: Statistics, 1–12.

Arnold B.C. and Shavelle R.M. (1998). Joint confidence sets for the mean and variance of a normal distribution. *Amer. Statist.,* **52**(2), 133–140.

Aydiner A., Lüsebrink J., Schildgen V., Winterfeld I., Knüver O., Schwarz K., Messler S., Schildgen O., and Mattner F. (2012). Comparison of two commercial PCR methods for methicillin-resistant Staphylococcus aureus (MRSA) screening in a tertiary care hospital. *PLoS One,* **7,** e43935.

Bagdonavičius V., Nikoulina V., and Nikulin M. (1997). Bolshev’s method of confidence limit construction. *Qüestiió,* **21**(3), 549–562.

Bapat R.B. and Beg M.I. (1989). Order statistics for nonidentically distributed variables and permanents. *Sankhyā,* **51**(1), 79–93.

Berger R.L. and Boos D.D. (1994). P values maximized over a confidence set for the nuisance parameter. *J. Amer. Statist. Assoc.,* **89**(427), 1012–1016.

Blaker H. (2000). Confidence curves and improved exact confidence intervals for discrete distributions. *Canad. J. Statist.,* **28**(4), 783–798.

Blyth C.R. and Still H.A. (1974). Binomial confidence intervals. *J. Amer. Statist. Assoc.,* **78,** 108–116.

Bolshev L.N. (1965). On the construction of confidence limits. *Theory Probab. Appl.,* **10,** 173–177.

Bridges D.S. and Mehta G.B. (1995). *Representations of Preferences Orderings.* Springer, Lecture Notes in Economics and Mathematical Systems.

Brown L., Cai T.T., and DasGupta A. (2001). Interval estimation for a binomial proportion. *Statist. Sci.,* **16**(2), 101–117.

Buehler R.J. (1957). Confidence intervals for the product of two binomial parameters. *J. Amer. Statist. Assoc.,* **52,** 482–493.

Cantor G. (1895). Beiträge zur Begründung der transfiniten Mengenlehre. *Mathematische Annalen,* **46**(4), 481–512.

Casella G. (1986). Refining binomial confidence intervals. *Canad. J. Statist.,* **14**(2), 113–129.

Casella G. and Berger R.L. (2001). *Statistical Inference.* Duxbury Press, Pacific Grove.

Clopper C.J. and Pearson E.S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika,* **26,** 404–413.

Crow E.L. (1956). Confidence intervals for a proportion. *Biometrika,* **43**(3), 423–435.

Debreu G. (1954). Rerpresentation of a preference ordering by a numerical function. *Decision processes,* **3,** 159–165.

Dugundji, J. (1966). *Topology.* Allyn and Bacon, Boston, MA.

Dümbgen L. (2004). Exact confidence bounds in discrete models – algorithmic aspects of Sterne's method. *Preprint,* retrieved from `http://www.imsv.unibe.ch`.

Enøe C., Georgiadis M.P., and Johnson W.O. (2000). Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when true disease state is unknown. *Prev. Vet. Med.,* **45,** 61–81.

Finner H. (1994). Two-sided tests and one-sided confidence bounds. *Ann. Statist.,* **22**(3), 1502–1516.

Gart J.J. and Buck A.A. (1966). Comparison of a screening test and a reference test in epidemiologic studies II: a probabilistic model for the comparison of diagnostic tests. *Am. J. Epidemiol.,* **83**(3), 593–602.

Geyer C.J. and Meeden G.D. (2005). Fuzzy and randomized confidence intervals and P-values. *Statist. Sci.,* **20**(4), 358–366.

Geyer C.J. and Meeden G.D. (2005). Rejoinder: Fuzzy and randomized confidence intervals and P-values. *Statist. Sci.,* **20**(4), 384–387.

Gibbons, J.D. (1982). “*P values*” in Johnson & Kotz (1982): *Encyclopedia of Statistical Sciences,* Vol. 7, 366–368.

Gibbons, J.D. and Pratt, J.W. (1975). P-values: Interpretation and methodology. *Amer. Statist.,* **29**(1), 20–25.

Greenhalgh T. (1997). How to read a paper. Papers that report diagnostic or screening tests. *BMJ: British Medical Journal,* **315**(7107), 540–543.

Greinecker, M. (2012). Answer to: “Preorders induced by continuous functions to the reals.” *StackExchange (Mathematics),* `http://math.stackexchange.com/questions/179461/preorders-induced-by-continuous-functions-to-the-reals`, accessed on May 8, 2015.

Hande S. (1994). A note on order statistics for nonidentically distributed variables. *Sankhyā,* **56**(2), 365–368.

Harris B. and Soms A.P. (1991). Recent advances in statistical methods for system reliability using Bernoulli sampling of components. in: DePriest D.J. and Launer R.L. (Eds.): *Reliability in the Acquisitions Process.* Marcel Dekker, New York, 55–68.

Harris B. and Soms A.P. (1991). Theory and counterexamples for confidence limits on system reliability. *Statist. Probab. Lett.,* **11,** 411–417.

Hoeffding W. (1956). On the distribution of the number os successes in independent trials. *Ann. Math. Statist.,* **27**(3), 713–721.

Hui S.L. and Walter S.D. (1980). Estimating the error rates of diagnostic tests. *Biometrics,* **36**(1), 167–171.

Hui S.L. and Zhou X.H. (1998). Evaluation of diagnostic tests without gold standards. *Stat. Methods Med. Res.,* **7,** 354–370.

Jech, T. (2003). *Set Theory.* Springer, Berlin.

Jobe J.M. and David H.T. (1992). Buehler confidence bounds for a reliability-maintainability measure. *Technometrics,* **34**(2), 214–222.

Kabaila P. (2001). Better Buehler confidence limits. *Statist. Probab. Lett.,* **52,** 145–154.

Kabaila P. (2013). On randomized confidence intervals for the binomial probability. *Preprint,* arXiv:1302.6659.

Kabaila P. and Lloyd C.J. (1997). Tight upper confidence limits from discrete data. *Austral. J. Statist.,* **39**(2), 193–204.

Kabaila P. and Lloyd C.J. (2000). When do best confidence limits exist? *Statist. Probab. Lett.,* **50,** 115–120.

Kabaila P. and Lloyd C.J. (2002). The importance of the designated statistic on Buehler upper limits on a system failure probability. *Technometrics,* **44**(4), 390–395.

Kabaila P. and Lloyd C.J. (2003). The efficiency of Buehler confidence limits. *Statist. Probab. Lett.,* **65,** 21–28.

Kabaila P. and Lloyd C.J. (2004). Buehler confidence limits and nesting. *Aust. N. Z. J. Stat.,* **46**(3), 463–469.

Kabaila P. and Lloyd C.J. (2006). Improved Buehler limits based on refined designated statistics. *J. Statist. Plann. Inference,* **136,** 3145–3155.

Kolmogoroff A.N. (1941). Confidence limits for an unknown distribution function. *Ann. Math. Statist.,* **12**(4), 461–463.

Lalkhen A.G. and McCluskey A. (2008). Clinical tests: sensitivity and specificity. *Continuing Education in Anaesthesia, Clinical Care & Pain,* **8**(6), 221–223.

Lecoutre B. and Poitevineau J. (2012). New results for computing exact confidence intervals for one parameter discrete distributions. *Preprint,* hal-00911488.

Lehmann E.L. and Casella G. (1998). *Theory of Point Estimation.* Springer, New York.

Lehmann E.L. and Romano J.P. (2006). *Testing Statistical Hypotheses.* Springer, New York.

Lloyd C.J. (2015). Computing exact one-sided confidence limits for treatment effect in clinical trials. *Comm. Statist. Simulation Comput.,* just accepted.

Lloyd C.J. and Kabaila P. (2003). On the optimality and limitations of Buehler bounds. *Aust. N. Z. J. Stat.,* **45**(2), 167–174.

Lloyd C.J. and Kabaila P. (2010). Letter to the editor: Some comments on "On construction of the smallest one-sided confidence interval for the difference of two proportions". *Ann. Statist.,* **38**(6) 3840–3841.

Lloyd C.J. and Moldovan M.V. (2000). Exact one-sided confidence limits for the difference between two correlated proportions. *Statist. Med.,* **26,** 3369–3384.

Lloyd C.J. and Moldovan M.V. (2007). Exact confidence bounds for the risk ratio in 2×2 tables with structural zero. *Biom. J.,* **49**(6), 952–963.

Loong T.-W. (2003). Understanding sensitivity and specificity with the right side of the brain. *BMJ: British Medical Journal,* **327**(7417), 716–719.

Marshall A.W., Olkin I., and Arnold B.C. (2011). *Inequalities: Theory of Majorization and Its Applications.* Springer, New York.

Mattner L. and Mattner F. (2013). Confidence bounds for the sensitivity lack of a less specific diagnostic test, without gold standard. *Metrika,* **76,** 239–263.

Mattner F., Winterfeld I., and Mattner L. (2012). Diagnosing toxigenic Clostridium difficile: New confidence bounds show culturing increases sensitivity of the toxin A/B enzyme immuniassay and refute gold standards. *Scand. J. Infect. Diseases,* **44**(8), 578–585.

Mattner L. (2014). *Lecture Notes on Mathematical Statistics,* lecture held at Universität Trier.

Mattner L. and Tasto Ch. (2014). Confidence bounds for average success probabilities. *Probab. Math. Statist.,* **52**(2), 301–312.

Monk J.D. (2009). *Cardinal Invariants on Boolean Algebras.* Birkhäuser, Modern Birkhäuser Classics.

Newcombe R.G. (1998). Two-sided confidence intervals for the single proportion: comparison of seven methods. *Statist. Med.,* **17,** 857–872.

Newcombe R.G. (1998). Improved confidence intervals for the difference between binomial proportions based on paired data. *Statist. Med.,* **17,** 2635–3650.

Parthasarathy, K.L. (2005). *Probability measures on metric spaces.* AMS Chelsea Publishing, Providence, RI.

Pepe M.S. (2003). *The statistical evaluation of medical tests for classification and prediction.* Oxford University Press, Oxford.

Pfanzagl, J. (1994). *Parametric Statistical Theory.* De Gruyter, Berlin.

Reiser B. and Jaeger M. (1991). A comment on Buehler optimal confidence bounds for series systems reliability. *Statist. Probab. Lett.,* **11,** 65–67.

Revyakov M.I. (1992). Development of the Buehler approach to the construction of a confidence interval for a function of the parameters of a discrete distribution. *Cybernet. Systems Anal.,* **28**(6), 868–874.

Rudin W. (1986). *Real and Complex Analysis.* McGraw-Hill, New York.

Rüschendorf, L. (2014). *Mathematische Statistik.* Springer, Berlin.

Schervish, M.J. (1996). P values: What they are and what they are not. *Amer. Statist.,* **50**(3), 203–206.

Schulzer, M. (1994). Diagnostic tests: a statistical review. *Muscle & Nerve,* **17,** 815–819.

Scott, B.M. (2012). Answer to: "What are the requirements for separability inheritance?" *StackExchange (Mathematics),* `http://math.stackexchange.com/questions/180740/what-are-the-requirements-for-separability-inheritance`, accessed on May 8, 2015.

Sterne T.E. (1954). Some remarks on confidence of fiducial limits. *Biometrika,* **41**(1), 275–278.

Sudakov R.S. (1974). On the question of interval estimation of the index of reliability of a sequential system. *Engrg. Cybernetics,* **12,** 55–63.

Tango T. (1998). Equivalence test and confidence interval for the difference in proportions for the paired-sample design. *Stat. Med.,* **17,** 891–908.

Thompson E.A. (2005). Comment: Fuzzy and randomized confidence intervals and P-values. *Statist. Sci.,* **20**(4), 382–383.

Thulin M. (2014). The cost of using exact confidence intervals for a binomial proportion. *Electron. J. Statist.,* **8**(1), 817–840.

Vernersson A. (2015). *Continuous Order-Isomorphisms. A Debreu open gap approach.* PhD thesis, available under `http://umu.diva-portal.org/smash/get/diva2:818584/FULLTEXT01.pdf`.

Vind K. (2003). *Independence, Uncertainty, Additivity.* Springer, Studies in Economic Theory, vol. 14.

Wang W. (2006). Smallest confidence intervals for one binomial proportion. *J. Statist. Plann. Inference,* **136,** 4293–4306.

Wang W. (2010). On construction of the smallest one-sided confidence interval for the difference of two proportions. *Ann. Statist.,* **38**(2) 1227–1243.

Wang W. (2012). An inductive order construction for the difference of two dependent proportions. *Statist. Probab. Lett.,* **82,** 1623–1628.

Winterbottom A. (1984). The interval estimation of system reliability from component test data. *Oper. Res.,* **32**(3), 628–640.

Zhou X.-H., Obuchowski N.A., and McClish D.K. (2002). *Statistical Methods in Diagnostic Medicine.* Wiley, New York.

INDEX