Westfälische
Wilhelms-Universität
Münster

Nichtparametrische Statistik - Abschnitt 29: Die Pitmanschen Permutationstests

Seminarvortrag
Blockseminar zur Wahrscheinlichkeitstheorie
Master of Science

Westfälische Wilhelms-Universität Münster
Fachbereich Mathematik und Informatik
Institut für Mathematische Statistik

Betreuung:
Prof. Dr. Matthias Löwe
Andrea Winkler

Referent:
Jörn Borrink

Münster, den 15. Februar 2013

Inhaltsverzeichnis

29. Die Pitmanschen Permutationstests

Literatur Die einzige Quelle dieses Seminars ist das Kapitel 5 des Skriptes zur Mathematischen Statistik von Gerold Alsmeyer (3. Auflage 2009).

Einleitendes Beispiel: Für die Behandlung einer Krankheit liegen zwei neue und völlig verschiedene Methoden (Medikamente) I und II vor. Zur Überprüfung, ob Methode I zur Behandlung besser geeignet ist als Methode II, könnten Patienten mit den Methoden behandelt werden und der Erfolg der Behandlung durch eine Kennzahl dargestellt werden, wobei höhere Zahlen für bessere Ergebnisse stehen. Denkbar wären zum Beispiel folgende Szenarien:

1. Die Patienten bilden eine inhomogene Gruppe. Daher werden mittels dem Zufallsprinzip von n Patienten n_1 mit Methode I und $n_2 = n - n_1$ mit Methode II behandelt. Jeder Patient liefert am Ende der Behandlung einen Messwert, so dass zum Beispiel folgende Tabelle zustande kommen könnte:

I	6,8	10,8	5,9	7,2	5,3
II	12,6	6,9	10,3	13,8	

2. Die Patienten besitzen bestimmte identische Ausprägungen (z.B Zwillinge) und können daher zu Paaren zusammengefügt werden. Nun wird der eine Patient eines Paares mit Methode I und der andere Patient mit Methode II behandelt. Jedes Paar liefert dann am Ende der Behandlung einen Messwert, wobei Werte größer 0 für Methode I und Werte kleiner 0 für Methode II sprechen. Eine Beobachtung könnte zum Beispiel sein:

I-II	-5,8	3,9	-4,4	-6,6

Da die beiden Produkte neuartig sind, ist keinerlei Aussage über die Verteilung des Genesungsprozesses möglich, d.h. wir können keine sinnvolle (gerechtfertigte) Annahme über den vorliegenden Verteilungstyp machen. Damit ist die Anwendung von bekannten guten Tests, wie zum Beispiel dem *t-Test* oder $\mathcal{X}^2$*-Test*, nicht möglich.

Interessante Fragen: Ist es möglich (zu einen Niveau $\alpha \in (0,1)$) die Behauptung 'Methode I ist zur Behandlung besser geeignet als Methode II' zu überprüfen? Und wie würden in diesem Fall die optimalen Tests aussehen und was heißt in diesem Kontext eigentlich optimal? Lassen sich für das Testproblem zum Beispiel gleichmäßig beste unverfälschte oder gleichmäßig beste invariante Tests zum Niveau $\alpha \in (0,1)$ herleiten?

Ziel dieses Vortrags ist die Beantwortung dieser Fragen, also die Entwicklung 'optimaler' Tests unter Verwendung einiger Restriktionen in den beiden obigen Szenarien.

29.1. Der stetige Fall

Wir gehen in diesem Abschnitt davon aus, dass die Beobachtungen einer stetigen Verteilung unterliegen. Diese Annahme ist durchaus kritisch zu hinterfragen, welches wir kurz im Ausblick auf den diskreten Fall erläutern werden (siehe 29.2).

29.1.1. Zweistichprobenvergleich

Problemstellung 29.1 (Zweistichprobenvergleich)
Zum Vergleich der Güte zweier Verfahren (Behandlungsmethoden) I und II werden n_1 Messungen unter Verwendung von Verfahren I und n_2 Messungen unter Verwendung von Verfahren II durchgeführt. Auf der Grundlage der Ergebnisse, repräsentiert durch $X_{11}, \ldots, X_{1n_1}$ und $X_{21}, \ldots, X_{2n_2}$ für Verfahren I bzw. II, soll entschieden werden, ob Verfahren I besser ist als II, wobei sich größere Güte durch höhere Meßwerte ausdrückt.

Modell 29.2 (Zweistichprobenvergleich)
Wir machen folgende Annahmen:

- $(X_{ij})_{i=1,2,j=1,\ldots,n_i}$ *sind stochastisch unabhängig*
- $X_{ij} \sim \mathbb{Q}_i$ *für* $i = 1,2$ *und* $j = 1, \ldots, n_i$
- $(\mathbb{Q}_1, \mathbb{Q}_2) \in \Theta$ *mit dem Parameterraum*

$$\Theta := \{(\mathbb{Q}_1, \mathbb{Q}_2) : \mathbb{Q}_i \textit{ stetiges W'Maß auf } (\mathbb{R}, \mathcal{B}) \textit{ und } \mathbb{Q}_1 \succeq \mathbb{Q}_2 \textit{ oder } \mathbb{Q}_1 \preceq \mathbb{Q}_2\}$$

Erläuterung der Annahmen: Zum einen sichern wir mit der Wahl von Θ, dass die beiden Methoden vergleichbar sind, wodurch eine Untersuchung überhaupt sinnvoll ist. Zum anderen forderen wir gleichzeitig, dass die einzelnen Versuchsergebnisse der Patientenbehandlungen unabhängig und identisch verteilt sind ($X_{ij} \sim \mathbb{Q}_i$). Dies impliziert, dass der Behandlungserfolg unabhängig vom heutigen Zustand des Patienten ist. Lezteres ist eine sehr wünschenswerte Voraussetzung, da Abhängigkeiten das Ergebnis verfälschen könnten:
Stellen wir uns zum Beispiel die 20-jährige Sportstudentin Sarah vor, die gerade ihren dritten Marathon für dieses Jahr absolviert hat und der die Ernährung unglaublich wichtig ist (ausgewogen, vitaminreich) und den 70-jährigen Bankchef Harald, der vor lauter Schreibtischarbeit sowie Stress dem Rauchen und den Süßigkeiten im überdimensionalem Maße zuspricht, die 100kg Marke mit Leichtigkeit übertrumpft sowie chronisch hohen Blutdruck besitzt. Nun werden Sarah und Harald am selben Tag mit einer Krebsdiagnose eingeliefert. Der Stationschef entscheidet sich dazu, Sarah mit Methode I und Harald mit Methode II zu behandeln, und wird in der Folge die Methode verwenden, welche schneller zur vollständigen Heilung geführt hat. Durch die Missachtung der unterschiedlichen Voraussetzungen von Sarah (jung, sportlich, gut ernährt) und Harald (alt, übergewichtig, seit Jahren Raucher) kann dieses Vorgehen des Stationschefs nun aber zu einer Fehlentscheidung führen, wenn Sarah zuerst geheilt ist. In diesem Fall kann trotzdem Methode II die bessere sein, da sie den 'schwierigen' Patienten zu heilen hatte.
Insbesondere ist ein Testergebnis im Bezug auf Abhängigkeiten zwischen den Patienten bzw. zwischen Patient und Krankheit/Behandlung kritisch zu hinterfragen.

Gehen wir nun zurück zum Modell und setzen $X = (X_{11}, \dots, X_{1n_1}, X_{21}, \dots, X_{2n_2})$ und $n = n_1 + n_2$, so erhalten wir das statistische Experiment

$$\mathcal{E} = (\mathbb{R}^n, \mathcal{B}^n, (\mathbb{Q}_1^{n_1} \otimes \mathbb{Q}_2^{n_2})_{(\mathbb{Q}_1, \mathbb{Q}_2) \in \Theta}).$$

Das Testproblem lässt sich nun präzisieren in der Form

$$\begin{array}{lcl} H = \{(\mathbb{Q}_1, \mathbb{Q}_2) \in \Theta : \mathbb{Q}_1 \preceq \mathbb{Q}_2\} & \text{gegen} & K = \{(\mathbb{Q}_1, \mathbb{Q}_2) \in \Theta : \mathbb{Q}_1 \succ \mathbb{Q}_2\} \\ \hat{=} \text{ I ist nicht besser als II} & & \hat{=} \text{ I ist besser als II} \end{array}$$

Lemma 29.3
In der Problemstellung 29.1 gilt

- $\mathbb{E}_{(\mathbb{Q}_1,\mathbb{Q}_2)}(\varphi) = \alpha$ *für alle* $(\mathbb{Q}_1,\mathbb{Q}_2) \in J = \overline{H} \cap \overline{K}$ *und* $\varphi \in \Phi_\alpha^u$
- $J = \Theta_= := \{(\mathbb{Q}_1,\mathbb{Q}_2) \in \Theta : \mathbb{Q}_1 = \mathbb{Q}_2\}$, *d.h.*

$$\{\mathbb{P}_\theta^X : \theta \in J\} = \{\mathbb{Q}_1^{n_1} \otimes \mathbb{Q}_2^{n_2} : (\mathbb{Q}_1,\mathbb{Q}_2) \in \Theta_=\} = \{\mathbb{Q}^n : \mathbb{Q} \text{ stetiges W'Maß auf } (\mathbb{R},\mathcal{B})\}$$

Bemerkung 29.4
Bezeichnen $T_{1,(\cdot)}, T_{2,(\cdot)}$ *die OS der ersten und zweiten Teilstichprobe von* x*, genauer*

$$T_{i,(\cdot)}(x) = (x_{i,(1)}, \dots, x_{i,(n_i)}), \quad i = 1, 2,$$

und definieren wir $T = (T_{1,(\cdot)}, T_{2,(\cdot)})$*, so ist die Statistik* T *suffizient für das Experiment* $\mathcal{E} = (\mathbb{R}^n, \mathcal{B}^n, (\mathbb{Q}_1^{n_1} \otimes \mathbb{Q}_2^{n_2})_{(\mathbb{Q}_1,\mathbb{Q}_2)\in\Theta})$.

Bemerkung 29.5
Die Statistik V

$$V = T_{(\cdot)} : (x_{11}, \dots, x_{2n_2}) \mapsto (x_{(11)}, \dots, x_{(2n_2)}) := x_{(\cdot)}$$

ist vollständig und suffizient für $(\mathbb{P}_\theta^X)_{\theta\in J}$.

Optimalität im Zweistichprobenvergleich: Bisher sind für das vorgestellte Testproblem weder gleichmäßig beste unverfälschte noch gleichmäßig beste invariante Tests zum Niveau $\alpha \in (0,1)$ bekannt. Aus diesem Grund werden wir uns mit der Angabe eines unverfälschten Tests zum Niveau α begnügen, der lediglich auf einer *interessanten* Teilmenge K_1 der Alternative K den Fehler 2. Art gleichmäßig unter allen J-ähnlichen Tests zum Niveau α minimiert.
Wir suchen also eine Lösung φ^* von

$$\begin{aligned} &\varphi^* \in \Phi_\alpha^u \\ &\mathbb{E}_\theta \varphi^*(X) = \max_{\varphi\in\Phi_\alpha^u} \mathbb{E}_\theta \varphi(X) \quad \text{für alle } \theta \in K_1 \subset K \end{aligned}$$

Dies bedeutet, die Hypothese H wird gegen eine eingeschränkte Alternative $K_1 \subset K$ getestet, wobei die Menge K_1 im Folgenden genauer spezifiziert wird.

Satz 29.6 (Einseitiger Pitman-Zweistichprobentest)
Sei

$$K_1 := \left\{(\mathbb{Q}_1, \mathbb{Q}_2) \in K : \exists \beta_1 > \beta_2, g \geq 0 : \frac{d\mathbb{Q}_i}{d\lambda\!\!\lambda}(x) = C_{\beta_i} e^{\beta_i x} g(x) \text{ für } i = 1, 2\right\}.$$

Zum Niveau $\alpha \in (0,1)$ wird der Pitman-Test durch

$$\varphi^*(x) = \begin{cases} 1 & > \\ \gamma_\nu^* \quad \text{falls } \sum_{j=1}^{n_1} x_{1j} & = c_\nu^* \\ 0 & < \end{cases} \tag{29.1}$$

definiert, $\nu = V(x) = x_{(\cdot)}$ für $x \in \mathbb{R}^n_{\neq}$, wobei sich $\gamma_\nu^* \in [0,1)$ und c_ν^* aus der Gleichung

$$\frac{1}{n!}\left|\left\{x \in \mathbb{R}^n_{\neq} : x_{(\cdot)} = \nu, \sum_{j=1}^{n_1} x_{1j} > c_\nu^*\right\}\right| + \frac{\gamma_\nu^*}{n!}\left|\left\{x \in \mathbb{R}^n_{\neq} : x_{(\cdot)} = \nu, \sum_{j=1}^{n_1} x_{1j} = c_\nu^*\right\}\right| = \alpha \tag{29.2}$$

bestimmen. Dann gilt:

(a) φ^* ist gleichmäßig bester Test für H gegen K_1 unter allen J-ähnlichen Tests zum Niveau α, d.h.

$$\varphi^* \in \Phi_{J,\alpha} \quad \text{und} \quad \mathbb{E}_\theta(\varphi^*) = \max_{\varphi \in \Phi_{J,\alpha}} \mathbb{E}_\theta(\varphi) \quad \text{für alle } \theta \in K_1.$$

(b) φ^* ist ein unverfälschter Test zum Niveau α für H gegen K, d.h. $\varphi^* \in \Phi^u_\alpha$.

Bemerkung 29.7
Teilmengen von K_1 sind zum Beispiel

$$\begin{aligned} M_1 &= \{(\mathcal{N}(\mu + \xi, \sigma^2), \mathcal{N}(\mu, \sigma^2)) : (\mu, \xi, \sigma) \in \mathbb{R} \times (0, \infty)^2\} \\ M_2 &= \{(\Gamma(\alpha, \beta - \xi), \Gamma(\alpha, \beta)) : (\alpha, \beta) \in (0, \infty)^2, \xi \in (0, \beta)\} \\ M_3 &= \{(\mathit{Exp}(\beta - \xi, c), \mathit{Exp}(\beta, c)) : (\beta, c) \in (0, \infty)^2, \xi \in (0, \beta)\} \end{aligned}$$

wobei $\mathit{Exp}(\beta, c)$ die auf $(0,c)$ gestutzte Exponentialverteilung mit $\lambda\!\!\lambda$-Dichte

$$f(x) = \frac{\beta e^{-\beta x}}{1 - e^{-\beta c}} \mathbf{1}_{(0,c)}(x)$$

bezeichnet.

Beweis von Satz 29.6(a).

Beweisidee: Konstruktion des gleichmäßig besten Test für H gegen K_1, welcher letztendlich φ^* sein wird. Hierzu gliedern wir den Beweis in drei Schritte.

(1) Umformung des Problems:

Sei $\vartheta \in K_1$. Gesucht ist ein Test $\hat{\varphi}$ mit

$$\hat{\varphi} \in \Phi_{J,\alpha} \text{ und } \mathbb{E}_\vartheta(\hat{\varphi}) = \max_{\varphi \in \Phi_{J,\alpha}} \mathbb{E}_\vartheta(\varphi). \tag{29.3}$$

Wir wissen bereits, dass die Statistik V

$$V = T_{(\cdot)} : (x_{11}, \dots, x_{2n_2}) \mapsto (x_{(11)}, \dots, x_{(2n_2)}) := x_{(\cdot)}$$

vollständig und suffizient für $(\mathbb{P}_\theta^X)_{\theta \in J}$, sowie die Statistik T

$$T = (T_{1,(\cdot)}, T_{2,(\cdot)}) : (x_{11}, \dots, x_{2n_2}) \mapsto (x_{1(1)}, \dots, x_{1(n_1)}, x_{2(1)}, \dots, x_{1(n_2)})$$

suffizient für $\mathcal{E} = (\mathbb{R}^n, \mathcal{B}^n, (\mathbb{Q}_1^{n_1} \otimes \mathbb{Q}_2^{n_2})_{(\mathbb{Q}_1,\mathbb{Q}_2)\in\Theta})$ ist. Insbesondere gilt somit, dass die Funktion $\theta \mapsto \mathbb{E}_\theta(\varphi|V)$ auf J konstant ist und wir verwenden die Notation $\mathbb{E}_J(\varphi|V)$ für $\{\mathbb{E}_\theta(\varphi|V)|\ \theta \in J\}$. Mithilfe dieser beiden Statistiken folgt unter Anwendung von Lemma A.13 und Lemma A.14 (Kapital 20: bedingte Tests), dass ein Test $\hat{\varphi}$ die obigen Forderungen (29.3) genau dann erfüllt, wenn

$$\hat{\varphi} \in \widetilde{\Phi}_{J,\alpha} = \left\{\varphi : \mathbb{E}_J(\varphi|V) = \alpha \quad \mathbb{P}_\theta^X\text{-f.s. für alle } \theta \in J\right\}$$

und

$$\mathbb{E}_\vartheta(\hat{\varphi}|V) \geq \mathbb{E}_\vartheta(\varphi|V) \quad \mathbb{P}_\vartheta^X\text{-f.s. für alle } \varphi \in \widetilde{\Phi}_{J,\alpha}.$$

Durch Übergang zu bedingten Verteilungen nehmen diese beiden Forderungen die Form

$$\hat{\varphi} \in \widetilde{\Phi}_{J,\alpha} = \left\{\varphi : \int \varphi \, d\mathbb{P}_J^{X|V(X)=\nu} = \alpha \ \mathbb{P}_\theta^{V(X)}\text{-f.s. für alle } \theta \in J\right\}, \tag{29.4}$$

$$\int \hat{\varphi} \, d\mathbb{P}_\vartheta^{X|V(X)=\nu} \geq \int \varphi \, d\mathbb{P}_\vartheta^{X|V(X)=\nu} \quad \mathbb{P}_\vartheta^{V(X)}\text{-f.s. für alle } \varphi \in \widetilde{\Phi}_{J,\alpha} \tag{29.5}$$

an. Die Idee besteht nun darin, für jedes $\vartheta \in K_1$ einen optimalen Test $\hat{\varphi}_\vartheta$, also einen Test, der (29.4) und (29.5) erfüllt, zu konstruieren. Hierzu benötigen wir zuerst die konkrete Form der regulär bedingten Verteilung von X gegeben $V(X) = \nu$.

(2) Herleitung der bedingten Verteilungen $\mathbb{P}_\theta^{X|V(X)=\nu}$ für $\theta \in \Theta$:

Für $\theta \in \Theta$ gilt $\mathbb{P}_\theta^X(\mathbb{R}^n_{\neq}) = 1$, also offensichtlich auch $\mathbb{P}_\theta^{V(X)}(\mathbb{R}^n_<) = 1$.
Für jedes $\nu \in \mathbb{R}^n_<$ ist $\mathbb{P}_\theta^{X|V(X)=\nu}$ eine diskrete Verteilung auf der Menge

$$\begin{aligned}\{\mathbb{K}_\pi(\nu) : \pi \in S_n\} &= \{(\nu_{\pi(1)}, \dots, \nu_{\pi(n)}) : \pi \in S_n\} \\ &= \{x \in \mathbb{R}^n : V(x) = \nu\} \qquad (29.6)\\ &:= \mathcal{X}_\nu \qquad (29.7)\end{aligned}$$

Gegeben ein $\theta = (\mathbb{Q}_1, \mathbb{Q}_2) \in \Theta$ sei $\mathbb{Q}$ ein dominierendes σ-endliches Maß für $\mathbb{Q}_1, \mathbb{Q}_2$ (z.B. $\mathbb{Q} = \mathbb{Q}_1 + \mathbb{Q}_2$) und $f_i := \frac{d\mathbb{Q}_i}{d\mathbb{Q}}$ für $i = 1, 2$. Bezeichnen wir nun mit h_θ die $\mathbb{Q}^n$-Dichte von $\mathbb{P}_\theta^X = \mathbb{Q}_1^{n_1} \otimes \mathbb{Q}_2^{n_2}$, also

$$h_\theta(x) = \frac{d(\mathbb{Q}_1^{n_1} \otimes \mathbb{Q}_2^{n_2})}{d\mathbb{Q}^n}(x) = \prod_{i=1}^{2} \prod_{j=1}^{n_i} f_i(x_{ij}),$$

und setzen weiter

$$\begin{aligned}\overline{h}_\theta(x) &:= \sum_{\pi \in S_n} h_\theta \circ \mathbb{K}_\pi(x), \\ &= \sum_{\pi \in S_n} h_\theta((x_{\pi(1)}, \dots, x_{\pi(n)})),\end{aligned}$$

so gilt

$$\mathbb{P}_\theta^{X|V(X)=\nu} = \begin{cases} \sum_{x \in \mathcal{X}_\nu} \frac{h_\theta(x)}{\overline{h}_\theta(\nu)} \delta_x & \text{für } \theta \in \Theta \text{ mit } \overline{h}_\theta(\nu) > 0, \\ \frac{1}{42} & \text{für } \theta \in \Theta \text{ mit } \overline{h}_\theta(\nu) = 0. \end{cases} \qquad (29.8)$$

Auf J ergibt sich insbesondere

$$\mathbb{P}_J^{X|V(X)=\nu} = \frac{1}{n!} \sum_{x \in \mathcal{X}_\nu} \delta_x, \qquad (29.9)$$

wobei δ_x das Dirac-Maß in x bildet. Bevor die Formel (29.8) bewiesen wird, noch eine kurze Bemerkung zur Herkunft der Formel:
Für $\nu \in \mathbb{R}^n_<$ mit $\overline{h}_\theta(\nu) > 0$ ist nach (29.6)

$$\overline{h}_\theta(\nu) = \sum_{x \in \mathcal{X}_\nu} h_\theta(x)$$

und somit

$$\mathbb{P}_\theta^{X|V(X)=\nu} = \frac{1}{\overline{h}_\theta(\nu)} \sum_{x\in\mathcal{X}_\nu} h_\theta(x)\delta_x = \frac{1}{\sum\limits_{x:V(x)=\nu} h_\theta(x)} \sum_{x\in\mathcal{X}_\nu} h_\theta(x)\delta_x.$$

Interpretieren kann man dies unter der Voraussetzung $\mathbb{P}_\theta(\{V(X)=\nu\}) > 0$ als

$$\mathbb{P}_\theta(X \in \cdot|V(X)=\nu) = \frac{\mathbb{P}_\theta(\{X\in\cdot\}\cap\{V(X)=\nu\})}{\mathbb{P}_\theta(\{V(X)=\nu\})}.$$

Beweis der Gleichung (29.8):
Für die regulär bedingte Verteilung $\mathbb{P}_\theta^{X|V(X)=\nu}$ gilt für alle $B \in \mathcal{B}^n \cap \mathbb{R}^n_{\neq}, C \in \mathcal{B}^n \cap \mathbb{R}^n_{<}$

$$\int_C \mathbb{P}_\theta^{X|V(X)=\nu}(B)\ \mathbb{P}_\theta^{V(X)}(d\nu) = \mathbb{P}_\theta\Big(X \in B, V(X) \in C\Big)$$

und sie ist $\mathbb{P}_\theta^{V(X)}$-f.s. eindeutig. Damit genügt es zu zeigen, dass für alle $B \in \mathcal{B}^n \cap \mathbb{R}^n_{\neq}$ und $C \in \mathcal{B}^n \cap \mathbb{R}^n_{<}$

$$\int_C \left(\sum_{x\in\mathcal{X}_\nu} \frac{h_\theta(x)}{\overline{h}_\theta(\nu)} \mathbf{1}_B(x) \right) \mathbb{P}_\theta^{V(X)}(d\nu) = \mathbb{P}_\theta\Big(X \in B, V(X) \in C\Big)$$

gilt.

Für beliebiges $C \in \mathcal{B}^n \cap \mathbb{R}^n_{<}$ betrachten wir $\mathbb{P}_\theta^{V(X)}(C) = \mathbb{P}_\theta^X(V^{-1}(C))$.
Es gilt

$$V^{-1}(C) = \{x : V(x) \in C\} = \bigcup_{\nu\in C} \{x : V(x) = \nu\}$$

und mit $C \subseteq \mathbb{R}^n_{<}$ folgt nach (29.6)

$$\begin{aligned} V^{-1}(C) &= \bigcup_{\nu\in C} \{\mathbb{K}_\pi(\nu) : \pi \in S_n\} \\ &= \bigcup_{\nu\in C} \bigcup_{\pi\in S_n} \{\mathbb{K}_\pi(\nu)\} \\ &= \bigcup_{\pi\in S_n} \bigcup_{\nu\in C} \{\mathbb{K}_\pi(\nu)\} \\ &:= \bigcup_{\pi\in S_n} \mathbb{K}_\pi(C). \end{aligned}$$

Behauptung: Die $(\mathbb{K}_\pi(C))_{\pi\in S_n}$ sind paarweise disjunkt.
Beweis: Angenommen, es gilt $x \in \mathbb{K}_{\overline{\pi}}(C) \cap \mathbb{K}_{\hat{\pi}}(C)$ (insbesondere ist dann $x \in \mathbb{R}^n_{\neq}$).
Dann existieren $\overline{\nu}, \hat{\nu} \in C$ mit

$$x = \mathbb{K}_{\overline{\pi}}(\overline{\nu}) \text{ und } x = \mathbb{K}_{\hat{\pi}}(\hat{\nu}), \tag{29.10}$$

also gilt

$$V(x) = V(\mathbb{K}_{\overline{\pi}}(\overline{\nu})) = \overline{\nu} \text{ und } V(x) = V(\mathbb{K}_{\hat{\pi}}(\hat{\nu})) = \hat{\nu}, \text{ da } \overline{\nu}, \hat{\nu} \in \mathbb{R}^n_<,$$

und somit $\overline{\nu} = \hat{\nu}$. Nach (29.10) bedeutet dies

$$\mathbb{K}_{\overline{\pi}}(\overline{\nu}) = \mathbb{K}_{\hat{\pi}}(\overline{\nu}) = x \in \mathbb{R}^n_{\neq},$$

also folgt $\overline{\pi} = \hat{\pi}$.

$V^{-1}(C)$ ist somit Vereinigung der disjunkter Mengen $\mathbb{K}_\pi(C)$, $\pi \in S_n$, und es ergibt sich mit $\mathbb{K}_\pi^{-1} = \mathbb{K}_{\pi^{-1}}(\star)$

$$\begin{aligned}
\mathbb{P}_\theta^{V(X)}(C) &= \mathbb{P}_\theta^X(V^{-1}(C)) \\
&= \mathbb{P}_\theta^X\Big(\bigcup_{\pi\in S_n} \mathbb{K}_\pi(C)\Big) \\
&= \sum_{\pi\in S_n} \mathbb{P}_\theta^X(\mathbb{K}_\pi(C)) \\
&= \sum_{\pi\in S_n} \int_{\mathbb{K}_\pi(C)} h_\theta \; d\mathbb{Q}^n \\
&\overset{\star}{=} \sum_{\pi\in S_n} \int_{\mathbb{K}^{-1}_{\pi^{-1}}(C)} h_\theta(\mathbb{K}_\pi \circ \mathbb{K}_{\pi^{-1}}(\nu)) \; \mathbb{Q}^n(d\nu) \\
&= \sum_{\pi\in S_n} \int_C h_\theta(\mathbb{K}_\pi(\nu)) \; (\mathbb{Q}^n)^{\mathbb{K}_{\pi^{-1}}}(d\nu) \\
&= \sum_{\pi\in S_n} \int_C h_\theta(\mathbb{K}_\pi(\nu)) \; \mathbb{Q}^n(d\nu) \\
&= \int_C \sum_{\pi\in S_n} h_\theta(\mathbb{K}_\pi(\nu)) \; \mathbb{Q}^n(d\nu) \\
&= \int_C \overline{h}_\theta(\nu) \; \mathbb{Q}^n(d\nu).
\end{aligned}$$

Es gilt also auf $\mathcal{B}^n \cap \mathbb{R}^n_<$

$$\overline{h}_\theta = \frac{d\mathbb{P}_\theta^{V(X)}}{d\mathbb{Q}^n}.$$

Damit folgt aber für alle $B \in \mathcal{B}^n \cap \mathbb{R}^n_{\neq}$ und $C \in \mathcal{B}^n \cap \mathbb{R}^n_{<}$

$$\begin{aligned}
&\int_C \left(\sum_{x \in \mathcal{X}_\nu} \frac{h_\theta(x)}{\overline{h}_\theta(\nu)} \mathbf{1}_B(x) \right) \mathbb{P}_\theta^{V(X)}(d\nu) \\
&= \int_{C \cap \{\overline{h}_\theta > 0\}} \left(\sum_{x \in \mathcal{X}_\nu} \frac{h_\theta(x)}{\overline{h}_\theta(\nu)} \mathbf{1}_B(x) \right) \overline{h}_\theta(\nu)\ \mathbb{Q}^n(d\nu) \\
&= \int_{C \cap \{\overline{h}_\theta > 0\}} \left(\sum_{x \in \mathcal{X}_\nu} h_\theta(x) \mathbf{1}_B(x) \right) \mathbb{Q}^n(d\nu)
\end{aligned}$$

und mit $\{\mathbb{K}_\pi(\nu) : \pi \in S_n\} = \{x \in \mathbb{R}^n : V(x) = \nu\}$ nach (29.6) folgt

$$\begin{aligned}
&= \int_{C \cap \{\overline{h}_\theta > 0\}} \sum_{\pi \in S_n} h_\theta(\mathbb{K}_\pi(\nu)) \mathbf{1}_B(\mathbb{K}_\pi(\nu))\ \mathbb{Q}^n(d\nu) \\
&= \int_{\mathbb{K}_{\pi^{-1}}(C \cap \{\overline{h}_\theta > 0\})} \sum_{\pi \in S_n} h_\theta(\nu) \mathbf{1}_B(\nu)\ \mathbb{Q}^n(d\nu) \\
&= \sum_{\pi \in S_n} \int_{B \cap \mathbb{K}_{\pi^{-1}}(C \cap \{\overline{h}_\theta > 0\})} h_\theta(\nu)\ \mathbb{Q}^n(d\nu) \\
&= \sum_{\pi \in S_n} \mathbb{P}_\theta^X \Big(B \cap \mathbb{K}_{\pi^{-1}}(C \cap \{\overline{h}_\theta > 0\}) \Big).
\end{aligned}$$

Die Mengengleichheit $\{\pi : \pi \in S_n\} = \{\pi^{-1} : \pi \in S_n\}$ (da $(S_n, \circ)$ eine Gruppe ist) liefert

$$\begin{aligned}
&= \sum_{\pi \in S_n} \mathbb{P}_\theta^X \Big(B \cap \mathbb{K}_\pi(C \cap \{\overline{h}_\theta > 0\}) \Big) \\
&= \mathbb{P}_\theta^X \Big(B \cap \Big(\bigcup_{\pi \in S_n} \mathbb{K}_\pi(C \cap \{\overline{h}_\theta > 0\}) \Big) \Big) \\
&= \mathbb{P}_\theta \Big(X \in B, X \in \bigcup_{\pi \in S_n} \mathbb{K}_\pi(C \cap \{\overline{h}_\theta > 0\}) \Big) \\
&= \mathbb{P}_\theta \Big(X \in B, X \in (V^{-1}(C \cap \{\overline{h}_\theta > 0\}) \Big) \\
&= \mathbb{P}_\theta \Big(X \in B, V(X) \in (C \cap \{\overline{h}_\theta > 0\}) \Big) \\
&= \mathbb{P}_\theta \Big(X \in B, V(X) \in C \Big),
\end{aligned}$$

wobei im letzen Schritt $\mathbb{P}_\theta(V(X) \in \{\overline{h}_\theta = 0\}) = \int_{\{\overline{h}_\theta = 0\}} \overline{h}_\theta\ d\mathbb{Q}^n = 0$ verwendet wird.

Insgesamt gilt also

$$\int_C \left(\sum_{x\in\mathcal{X}_\nu} \frac{h_\theta(x)}{\overline{h}_\theta(\nu)} \mathbf{1}_B(x) \right) \mathbb{P}_\theta^{V(X)}(d\nu) = \mathbb{P}_\theta\Big(X \in B, V(X) \in C\Big).$$

Da $h_\theta(\nu) > 0$ $\mathbb{P}_\theta^{V(X)}$-f.s., besitzt die bedingte Verteilung $\mathbb{P}_\theta^{X|V(X)=\nu}$ die Form in (29.8).

Für $\theta = (\mathbb{Q}_1, \mathbb{Q}_1) \in J = \Theta_=$ ergibt sich mit $\mathbb{Q} = \mathbb{Q}_1$ offenkundig

$$h_\theta = 1 \text{ sowie } \overline{h}_\theta = n!,$$

was zu

$$\mathbb{P}_J^{X|V(X)=\nu} = \frac{1}{n!} \sum_{x\in\mathcal{X}_\nu} \delta_x$$

führt.

(3) Anwendung des Neyman-Pearson-Lemma:

Mit der in Abschnitt 20 (bedingte Tests) entwickelten Theorie ergibt sich das weitere Vorgehen. Auf jeder der disjunkten Mengen $\mathcal{X}_\nu = \{x \in \mathbb{R}^n_{\neq} : V(x) = \nu\}, \nu \in \mathbb{R}^n_<$, wird ein bester Test zum Niveau α für $\mathbb{P}_J^{X|V(X)=\nu}$ gegen $\mathbb{P}_\vartheta^{X|V(X)=\nu}$ konstruiert. Ist φ^* ein Test, der für alle $\nu \in \mathbb{R}^n_<$ den Bedingungen

$$\int \varphi^* \, d\mathbb{P}_J^{X|V(X)=\nu} = \alpha \tag{29.11}$$

$$\int \varphi^* \, d\mathbb{P}_\vartheta^{X|V(X)=\nu} = \max_{\varphi : \int \varphi \, d\mathbb{P}_J^{X|V(X)=\nu} = \alpha} \int \varphi \, d\mathbb{P}_\vartheta^{X|V(X)=\nu} \tag{29.12}$$

genügt, so folgt $\mathbb{E}_\vartheta \varphi^* \geq \mathbb{E}_\vartheta \varphi$ für alle $\varphi \in \Phi_{J,\alpha}$.
Begründung: (siehe auch Satz 20.4 im Alsmeyer-Skript)
Für jedes $\varphi \in \Phi_{J,\alpha}$ bildet $N = \{\nu \in \mathbb{R}^n : \int \varphi \, d\mathbb{P}_J^{X|V(X)=\nu} \neq \alpha\}$ eine $(\mathbb{P}_\theta^{V(X)})_{\theta\in J}$-Nullmenge und damit auch eine $\mathbb{P}_\vartheta^{V(X)}$-Nullmenge für alle $\vartheta \in K$, denn $(\mathbb{P}_\theta^{V(X)})_{\theta\in J}$ dominiert $(\mathbb{P}_\theta^{V(X)})_{\theta\in\Theta}$. Da aber nach (29.12) für $\nu \in N^c$

$$\int \varphi^* \, d\mathbb{P}_\vartheta^{X|V(X)=\nu} \geq \int \varphi \, d\mathbb{P}_\vartheta^{X|V(X)=\nu},$$

gilt, ergibt sich die Behauptung mittels

$$\mathbb{E}_\vartheta(\psi) = \int \int \psi \, d\mathbb{P}_\vartheta^{X|V(X)=\nu} \, d\mathbb{P}_\vartheta^{V(X)}.$$

Gesucht ist also ein Test φ^* mit den Eigenschaften (29.11) und (29.12), der zusätzlich *nicht* von $\vartheta \in K$ abhängt. Da dies nicht für alle $\vartheta \in K$ möglich ist, wählen wir die Menge K_1 der Verteilungspaare mit Exponentialdichten bezüglich eines $\lambda\!\!\lambda$-stetigen Maßes $g\lambda\!\!\lambda$. Sei also $\vartheta = (\mathbb{Q}_1, \mathbb{Q}_2) \in K_1$. Dann gilt mit $\mathbb{Q} = \lambda\!\!\lambda$

$$\begin{aligned}
h_\theta(x) &= \frac{d(\mathbb{Q}_1^{n_1} \otimes \mathbb{Q}_2^{n_2})}{d\mathbb{Q}^n}(x) \\
&= \frac{d(\mathbb{Q}_1^{n_1} \otimes \mathbb{Q}_2^{n_2})}{d\lambda\!\!\lambda^n}(x) \\
&= \prod_{i=1}^{2}\prod_{j=1}^{n_i} f_i(x_{ij}) \\
&= \prod_{i=1}^{2}\prod_{j=1}^{n_i} C_{\beta_i} e^{\beta_i x_{ij}} g(x_{ij}) \\
&= C_{\beta_1}^{n_1} C_{\beta_2}^{n_2} \exp\left(\sum_{i=1}^{2} \beta_i \sum_{j=1}^{n_i} x_{ij}\right) \prod_{i=1}^{2}\prod_{j=1}^{n_i} g(x_{ij}) \\
&= C_{\beta_1}^{n_1} C_{\beta_2}^{n_2} \exp\left(\beta_1 \sum_{j=1}^{n_1} x_{1j} + \beta_2 \sum_{j=1}^{n_2} x_{2j}\right) \prod_{i=1}^{2}\prod_{j=1}^{n_i} g(x_{ij}) \\
&= C_{\beta_1}^{n_1} C_{\beta_2}^{n_2} \exp\left((\beta_1 - \beta_2) \sum_{j=1}^{n_1} x_{1j} + \beta_2 (\sum_{j=1}^{n_1} x_{1j} + \sum_{j=1}^{n_2} x_{2j})\right) \prod_{i=1}^{2}\prod_{j=1}^{n_i} g(x_{ij}) \\
&= C_{\beta_1}^{n_1} C_{\beta_2}^{n_2} \exp\left((\beta_1 - \beta_2) \sum_{j=1}^{n_1} x_{1j} + \beta_2 \sum_{j=1}^{n} \nu_j\right) \prod_{j=1}^{n} g(\nu_j)
\end{aligned}$$

mit $\nu = (\nu_1, \dots, \nu_n) := V(x)$. Ein bester Test φ_ν auf $\mathcal{X}_\nu$ für $\mathbb{P}_J^{X|V(X)=\nu}$ gegen $\mathbb{P}_\vartheta^{X|V(X)=\nu}$ ist nach dem Neyman-Pearson-Lemma gegeben durch

$$\varphi_\nu(x) = \begin{cases} 1 & & > \\ \gamma_\nu(\vartheta) & \text{falls } \frac{\mathbb{P}_\vartheta^{X|V(X)=\nu}(\{x\})}{\mathbb{P}_J^{X|V(X)=\nu}(\{x\})} & = & k_\nu(\vartheta), \\ 0 & & < \end{cases}$$

wobei $k_\nu(\vartheta), \gamma_\nu(\vartheta)$ durch die Gleichung $\int \varphi_\nu \; d\mathbb{P}_J^{X|V(X)=\nu} = \alpha$ festgelegt werden. Nun gilt auf $\mathcal{X}_\nu$ nach Schritt 2

$$\begin{aligned}
\frac{\mathbb{P}_\vartheta^{X|V(X)=\nu}(\{x\})}{\mathbb{P}_J^{X|V(X)=\nu}(\{x\})} &= \frac{\sum_{y \in \mathcal{X}_\nu} \frac{h_\theta(y)}{\overline{h}_\theta(\nu)} \delta_y(x)}{\frac{1}{n!} \sum_{y \in \mathcal{X}_\nu} \delta_y(x)} \\
&= \frac{h_\theta(x)}{\overline{h}_\theta(\nu)} \cdot n!
\end{aligned}$$

$$= \frac{C_{\beta_1}^{n_1} C_{\beta_2}^{n_2} \exp\left((\beta_1 - \beta_2) \sum_{j=1}^{n_1} x_{1j} + \beta_2 \sum_{j=1}^{n} \nu_j \right) \prod_{j=1}^{n} g(\nu_j)}{\overline{h}_\theta(\nu)} \cdot n!,$$

welches offensichtlich monoton wachsend in $\sum_{j=1}^{n_1} x_{1j}$ ist ($\beta_1 > \beta_2$). Dies impliziert aber die Gleichheit der Mengen

$$\left\{ \frac{\mathbb{P}_\vartheta^{X|V(X)=\nu}(\{x\})}{\mathbb{P}_J^{X|V(X)=\nu}(\{x\})} \gtreqless k_\nu(\vartheta) \right\} = \left\{ \sum_{j=1}^{n_1} x_{1j} \gtreqless c_\nu(\vartheta) \right\} \tag{29.13}$$

mit $c_\nu(\vartheta)$ eindeutig festgelegt durch $k_\nu(\vartheta)$. Aus der Forderung

$$\int \varphi_\nu \, d\mathbb{P}_J^{X|V(X)=\nu} = \alpha$$

ergibt sich mittels der regulär bedingten Verteilung in (29.9) und der Mengengleichheit (29.13)

$$\begin{aligned}
\alpha = \int \varphi_\nu \, d\mathbb{P}_J^{X|V(X)=\nu} &= \sum_{x \in \mathcal{X}_\nu} \varphi_\nu(x) \cdot \mathbb{P}_J^{X|V(X)=\nu}(x) \\
&= \sum_{x \in \mathcal{X}_\nu} \varphi_\nu(x) \cdot \frac{1}{n!} \\
&= \frac{1}{n!} \sum_{x \in \mathcal{X}_\nu} \left(\mathbf{1}_{\{\sum_{j=1}^{n_1} x_{1j} > c_\nu(\vartheta)\}} + \gamma_\nu(\vartheta) \mathbf{1}_{\{\sum_{j=1}^{n_1} x_{1j} = c_\nu(\vartheta)\}} \right) \\
&= \frac{1}{n!} \sum_{x \in \mathcal{X}_\nu} \mathbf{1}_{\{\sum_{j=1}^{n_1} x_{1j} > c_\nu(\vartheta)\}} + \frac{\gamma_\nu(\vartheta)}{n!} \sum_{x \in \mathcal{X}_\nu} \mathbf{1}_{\{\sum_{j=1}^{n_1} x_{1j} = c_\nu(\vartheta)\}} \\
&= \frac{1}{n!} \left| \left\{ x \in \mathbb{R}_{\neq}^n : x_{(\cdot)} = \nu, \sum_{j=1}^{n_1} x_{1j} > c_\nu(\vartheta) \right\} \right| \\
&\quad + \frac{\gamma_\nu(\vartheta)}{n!} \left| \left\{ x \in \mathbb{R}_{\neq}^n : x_{(\cdot)} = \nu, \sum_{j=1}^{n_1} x_{1j} = c_\nu(\vartheta) \right\} \right|
\end{aligned}$$

Dies bedeutet aber, dass $c_\nu(\vartheta), \gamma_\nu(\vartheta)$ nicht von ϑ abhängen (wohl aber von ν bzw. x). Mit $c_\nu^* := c_\nu(\vartheta)$ und $\gamma_\nu^* := \gamma_\nu(\vartheta)$ ergibt sich für jedes $\vartheta \in K_1$, dass

$$\varphi_\nu(x) = \begin{cases} 1 & & > \\ \gamma_\nu^* & \text{falls } \sum_{j=1}^{n_1} x_{1j} & = & c_\nu^* \\ 0 & & < \end{cases}$$

der gleichmäßig beste Test für H gegen $\{\vartheta\}$ unter allen J-ähnlichen Tests zum Niveau α ist. Durch $\varphi^*(x) = \varphi_{V(x)}(x)$ erhalten wir dann den gleichmäßig besten Test für H gegen K_1 unter allen J-ähnlichen Tests zum Niveau α ist. □

Beweis von Satz 29.6(b).
Wir verweisen hier auf Alsmeyer, Mathematische Statistik, Seite 203ff. □

Bemerkung 29.8 (Durchführung des Tests)
Für die Auswertung eines Tests wird eine Vertafelung der kritischen Werte c_ν^* benötigt. In Gegensatz zu der Statistik-Vorlesung, wo zum Beispiel im Kapital "Tests im Zusammenhang mit der Normalverteilung" der kritische Wert immer ein durch die Wahl von α eindeutig bestimmtes Quantil (Fraktil) einer bekannten Verteilungsfunktion darstellte, ist c_ν^* aufgrund von $\nu = V(x) = x_{(\cdot)}$ von der Realisation abhängig und eine Vertafelung daher nicht möglich.
Im Folgenden möchte ich eine einfache Vorgehensweise zur Bestimmung des kritischen Wertes c_ν^* vorführen:
Sei

$$v_I := \{\nu_i : i \in I\},\ I \subset \{1, \dots, n\}, |I| = n_1.$$

Es gibt $\binom{n}{n_1}$ Teilmengen von $\{1, \dots, n\}$ und zu diesen $I_1, \dots, I_{\binom{n}{n_1}}$ definieren wir die Summen $s_j = \sum_{i \in I_j} \nu_i, 1 \le j \le \binom{n}{n_1}$, derart dass $s_1 \ge s_2 \ge \dots \ge s_{\binom{n}{n_1}}$. Dann können wir γ_ν^* darstellen durch

$$\gamma_\nu^* = \frac{\alpha\binom{n}{n_1} - \left|\left\{\nu_I : \sum_{\nu_i \in \nu_I} \nu_i > c_\nu^*\right\}\right|}{\left|\left\{\nu_I : \sum_{\nu_i \in \nu_I} \nu_i = c_\nu^*\right\}\right|}$$

Begründung: *Für $x = (x_{11}, \dots, x_{1n_1}, x_{21}, \dots, x_{2n_2}) \in \mathbb{R}^n_{\neq}$ hat die Permutation der forderen n_1 und die der hinteren n_2 Einträge keinen Einfluss auf die Summe $\sum\limits_{j=1}^{n_1} x_{1j}$. Daher gilt*

$$\left|\left\{x \in \mathbb{R}^n_{\neq} : x_{(\cdot)} = v, \sum_{j=1}^{n_1} x_{1j} > c_v^*\right\}\right| = \left|\left\{\nu_I : \sum_{\nu_i \in \nu_I} \nu_i > c_\nu^*\right\}\right| \cdot n_1! \cdot n_2!$$

sowie

$$\left|\left\{x \in \mathbb{R}^n_{\neq} : x_{(\cdot)} = v, \sum_{j=1}^{n_1} x_{1j} = c_v^*\right\}\right| = \left|\left\{\nu_I : \sum_{\nu_i \in \nu_I} \nu_i = c_\nu^*\right\}\right| \cdot n_1! \cdot n_2!.$$

Hieraus ergibt sich

$$\begin{aligned}
&\frac{1}{n!}\cdot\left|\left\{x\in\mathbb{R}^n_{\neq}: x_{(\cdot)}=v,\sum_{j=1}^{n_1}x_{1j}>c_v^*\right\}\right| + \frac{\gamma_v^*}{n!}\cdot\left|\left\{x\in\mathbb{R}^n_{\neq}: x_{(\cdot)}=v,\sum_{j=1}^{n_1}x_{1j}=c_v^*\right\}\right| = \alpha\\
\Leftrightarrow\ &\frac{1}{n!}\cdot\left|\left\{\{\nu_I:\sum_{\nu_i\in\nu_I}\nu_i>c_\nu^*\right\}\right|\cdot n_1!\cdot n_2! \quad + \frac{\gamma_v^*}{n!}\cdot\left|\left\{\nu_I:\sum_{\nu_i\in\nu_I}\nu_i=c_\nu^*\right\}\right|\cdot n_1!\cdot n_2! \quad = \alpha\\
\Leftrightarrow\ &\frac{1}{\binom{n}{n_1}}\cdot\left|\left\{\nu_I:\sum_{\nu_i\in\nu_I}\nu_i>c_\nu^*\right\}\right| \quad + \frac{\gamma_v^*}{\binom{n}{n_1}}\cdot\left|\left\{\nu_I:\sum_{\nu_i\in\nu_I}\nu_i=c_\nu^*\right\}\right| \quad = \alpha
\end{aligned}$$

und somit die Formel für γ_ν^*.
Aus der Annahme, dass allen Beobachtungen eine stetige Verteilung zugrundeliegt, lässt sich aber weiter leicht folgern, dass für $(\mathbb{P}_\theta^{V(X)})_{\theta\in\Theta}$*-fast alle* ν *sogar* $s_1>\ldots>s_{\binom{n}{n_1}}$ *gilt. Mit* $k^*:=\lfloor\alpha\binom{n}{n_1}\rfloor$ *und* $c_\nu^*=s_{k^*+1}$ *folgt somit*

$$\gamma_\nu^*=\alpha\binom{n}{n_1}-\left\lfloor\alpha\binom{n}{n_1}\right\rfloor\in[0,1).$$

Beispiel 29.9 (Anwendung des einseitigen Pitman-Zweistichprobentest)
Wir können nun die Behauptung Methode I ist besser als Methode aus dem einleitenden Beispiel zum Niveau $\alpha\in(0,1)$ *testen. Wir beitzen also folgende Daten:*

- $\alpha=0,05$,
- $n_1=5, n_2=4$ *(also* $n=9$*) sowie*
- *die Meßwerte bei Anwendung von Verfahren I bzw. II:*

I	*6,8*	*10,8*	*5,9*	*7,2*	*5,3*
II	*12,6*	*6,9*	*10,3*	*13,8*	

Es gilt somit

- $\binom{n}{n_1}=\binom{9}{5}=126$,
- $k^*=\lfloor\alpha\cdot\binom{n}{n_1}\rfloor=\lfloor 0,05\cdot 126\rfloor=6$ *sowie*
- $\gamma_\nu^*=\alpha\binom{n}{n_1}-\lfloor\alpha\binom{n}{n_1}\rfloor=0,3$

Hieraus ergibt sich $c_\nu^*:=s_{k^*+1}=s_7$, *dies bedeutet, wir müssen die 7 größten Summen aus jeweils* $n_1=5$ *Meßwerten bilden:*

5,3	*5,8*	*6,8*	*6,9*	*7,2*	*10,3*	*10,8*	*12,6*	*13,8*	$\sum_{i\in I} \nu_i$
				$*$	$*$	$*$	$*$	$*$	$54,7$
			$*$		$*$	$*$	$*$	$*$	$54,4$
		$*$			$*$	$*$	$*$	$*$	$54,3$
	$*$				$*$	$*$	$*$	$*$	$53,3$
$*$					$*$	$*$	$*$	$*$	$52,8$
			$*$	$*$		$*$	$*$	$*$	$51,3$
		$*$		$*$		$*$	$*$	$*$	$51,2$

Der einseitige Pitman-Zweistichprobentest ist also gegeben durch

$$\varphi^*(x) = \begin{cases} 1 & & > \\ 0,3 & \textit{falls } \sum_{j=1}^{n_1} x_{1j} & = \ 51,2 \\ 0 & & < \end{cases} \tag{29.14}$$

Da $\sum_{j=1}^{n_1} x_{1j} = 6,8 + 10,8 + 5,9 + 7,2 + 5,3 = 36 < 51,2$ *ist* $\varphi^*(x) = 0$. *Der Test entscheidet sich also für die Hypothese "Verfahren I ist nicht besser als Verfahren II", wobei dies bedeutet, dass die Alternative zum Niveau* $\alpha = 0,05$ *nicht gesichert angenommen werden kann (die Hypothese kann bei Tests zum Niveau* α *nie gesichert angenommen werden, da nicht bekannt ist, wie groß der minimierte Fehler 2.Art ist).*

Bemerkung 29.10

Das Resultat im obigen Beispiel legt nahe, die Hypothese und Alternative zu vertauschen, um "Verfahren II ist besser als Verfahren I" zum Niveau α *zu verifizieren. D.h. die Rollen der beiden Verfahren werden vertauscht (also* $\widetilde{I}$*=II und* $\widetilde{II}$*=I) und wir testen "Verfahren* $\widetilde{I}$ *ist besser als* $\widetilde{II}$*".*

Aufgrund der Gleichheit $\binom{n}{n_1} = \binom{n}{n_2}$ *gilt auch in diesem Test* $c^*_\nu = s_7$*, wobei nun die 7 größsten Summen aus jeweils 4 Meßwerten gebildet werden müssen:*

5,3	*5,8*	*6,8*	*6,9*	*7,2*	*10,3*	*10,8*	*12,6*	*13,8*	$\sum_{i\in I} \nu_i$
					$*$	$*$	$*$	$*$	$47,5$
				$*$		$*$	$*$	$*$	$44,4$
			$*$			$*$	$*$	$*$	$44,1$
		$*$				$*$	$*$	$*$	$44,0$
				$*$	$*$		$*$	$*$	$43,9$
			$*$		$*$		$*$	$*$	$43,6$
		$*$			$*$		$*$	$*$	$43,5$

Es ergibt sich $c^*_\nu = 43,5 < 43,6 = 12,6 + 6,9 + 10,3 + 13,8$ *und* $\varphi^*(x) = 1$.
Damit kann zum Niveau $\alpha = 0,05$ *"Verfahren II ist besser als Verfahren I" verifiziert werden.*

29.1.2. Verbundene Stichproben

Erinnern wir uns an die Annahmen in der Problemstellung 29.1, so haben wir deren Gültigkeit in der Realität kritisch geäugt. Dies lag vor allem an der Tatsache, dass das Behandlungsergebnis häufig vom sonstigen Zustand des Patienten (Alter, Sportlichkeit, Ernährungsgewohnheiten) abhängt. In unserem überspitzten Beispiel mit Sarah und Harald wurde deutlich, dass die Behandlungsergebnisse der Methoden I und II nur sinnvoll interpretierbar sind, wenn die Patienten identische Voraussetzungen mitbringen. Optimal wäre daher ein und dieselbe Person zuerst mit Methode I und beim nächsten Krankheitsausbruch mit Methode II zu behandeln. Die verbundene Stichprobe nimmt diesen Ansatz auf, indem sie Paare von Patienten bildet, deren Eigenschaften in möglichst vielen heilungsrelevanten Komponenten übereinstimmen.

Problemstellung 29.11 (Verbundene Stichprobe)
Modifikation des Zweistichprobenvergleichs

Zum Vergleich zweier Behandlungsmethoden I und II werden dieses Mal zunächst Paare von Versuchseinheiten mit gleichartigen Komponenten (z.B. eineiige Zwillinge, Augenpaare) gebildet und anschließend auf ein Mitglied eines Paares Methode I, auf das andere Mitglied Methode II angewendet. Die Differenz der Meßwerte interpretiert man als Auswirkung der Unterschiede zwischen I und II, wobei sich größere Güte wie zuvor durch höhere Meßwerte ausdrückt.
Wir bezeichnen die Meßwerte im i-ten Paar, $i = 1, \dots, n$, *mit* X_{1i} *und* X_{2i}, *wobei* $(X_{11}, \dots, X_{1n})$ *die Ergebnisse bei Verwendung von Methode I und* $(X_{21}, \dots, X_{2n})$ *die Ergebnisse bei Verwendung von Methode II repräsentieren.*

Modell 29.12 (Verbundene Stichprobe)
Wir machen folgende Annahmen

- *Die Vektoren* $(X_{11}, X_{21}), \dots, (X_{1n}, X_{2n})$ *sind stochastisch unabhängig und identisch verteilt*

- *Die Stichprobendifferenz*

$$\begin{aligned} X_i &:= X_{1i} - X_{2i}, \quad i = 1, \dots, n \\ -X_i &:= X_{2i} - X_{1i}, \quad i = 1, \dots, n \end{aligned}$$

  besitzt eine stetige Verteilung $\mathbb{P}$ *bzw.* $\mathbb{P}^-$.

- F *bezeichne die Verteilungsfunktion der* X_i *und* F^- *die der* $-X_i$, *d.h. es gilt*

$$F^-(t) = 1 - F(-t) \quad \textit{für alle } t \in \mathbb{R}.$$

- $\mathbb{P} \in \mathfrak{F}$ *mit Parameterraum*

$$\mathfrak{F} := \{\mathbb{Q} : \mathbb{Q} \textit{ stetiges W'Maß auf } (\mathbb{R}, \mathcal{B}) \textit{ mit } \mathbb{Q} \succeq \mathbb{Q}^- \textit{ oder } \mathbb{Q} \preceq \mathbb{Q}^-\}$$

Wir wollen überprüfen, ob Methode I besser ist als Methode II. Mathematisch bedeutet dies, dass die Verteilung der X_i die der $-X_i$ stochastisch majorisiert, d.h. für alle $t \in \mathbb{R}$ gilt

$$\mathbb{P}((t, \infty)) \geq \mathbb{P}^-((t, \infty))$$

mit strikter Ungleichheit für mindestens ein t. Das Testproblem lässt sich also in der Form

$$\begin{array}{ccc} H = \{\mathbb{Q} \in \mathfrak{F} : \mathbb{Q} \preceq \mathbb{Q}^-\} & \text{gegen} & K = \{\mathbb{Q} \in \mathfrak{F} : \mathbb{Q} \succ \mathbb{Q}^-\} \\ \hat{=} \text{ I ist nicht besser als II} & & \hat{=} \text{ I ist besser als II} \end{array}$$

bei gegebenem statistischem Experiment

$$\mathcal{E} = (\mathbb{R}^n, \mathcal{B}^n, (\mathbb{Q}^n)_{\mathbb{Q} \in \mathfrak{F}})$$

präzisieren.

Lemma 29.13
In der Problemstellung 29.11 gilt

- $\mathbb{E}_{\mathbb{Q}}(\varphi) = \int_{\mathbb{R}^n} \varphi \, d\mathbb{Q}^n = \alpha$ *für alle* $\mathbb{Q} \in J = \overline{H} \cap \overline{K}$ *und* $\varphi \in \Phi_\alpha^u$

- $J = \mathfrak{F}_s := \{\mathbb{Q} \in \mathfrak{F} : \mathbb{Q} = \mathbb{Q}^-\}$, *d.h.* J *ist die Menge der stetigen, in 0 symmetrischen Verteilungen auf* $(\mathbb{R}, \mathcal{B})$.

Bemerkung 29.14
Die Statistik V

$$V : x = (x_1 \dots , x_n) \mapsto |x|_{(\cdot)} := T_{(\cdot)}(|x_1|, \dots , |x_n|)$$

ist vollständig und suffizient für $(\mathbb{Q}^n)_{\mathbb{Q} \in \mathfrak{F}_s}$.

Auch für die Modifikation des Zweistichprobenvergleichs sind bisher weder gleichmäßig beste unverfälschte noch gleichmäßig beste invariante Tests zum Niveau $\alpha \in (0, 1)$ bekannt. Daher begnügen wir uns erneut mit der Angabe eines unverfälschten Tests zum Niveau α, der lediglich auf einer *interessanten* Teilmenge K_1 der Alternative K den Fehler 2. Art gleichmäßig minimiert.

Satz 29.15 (Einseitiger Pitman-Symmetrietest)
Sei

$$K_1 := \left\{ \mathbb{P} \in K : \frac{d\mathbb{P}}{d\lambda\!\!\lambda}(x) = C_\beta e^{\beta x} g(|x|) \text{ mit } \beta > 0 \text{ und } g \geq 0 \right\}.$$

Zum Niveau $\alpha \in (0, 1)$ *wird der Pitman-Symmetrietest durch*

$$\varphi^*(x) = \begin{cases} 1 & > \\ \gamma_v^*, \quad \text{falls } \sum_{j=1}^n x_j & = c_v^* \\ 0 & < \end{cases} \tag{29.15}$$

definiert, $v = V(x) = |x|_{(\cdot)} := T_{(\cdot)}(|x_1|, \dots |x_n|)$ *für* $x \in \mathbb{R}^n_{|\neq|} := \{y \in \mathbb{R}^n : |y_i| \neq |y_j| \text{ für } i \neq j\}$*, wobei sich* $\gamma_v^* \in [0, 1)$ *und* c_v^* *aus der Gleichung*

$$\frac{1}{2^n n!} \left| \left\{ x \in \mathbb{R}^n_{|\neq|} : |x|_{(\cdot)} = v, \sum_{j=1}^n x_j > c_v^* \right\} \right| + \frac{\gamma_v^*}{2^n n!} \left| \left\{ x \in \mathbb{R}^n_{|\neq|} : |x|_{(\cdot)} = v, \sum_{j=1}^n x_j = c_v^* \right\} \right| = \alpha$$

bestimmen. Dann gilt:

(a) φ^* *ist gleichmäßig bester Test für* H *gegen* K_1 *unter allen* J*-ähnlichen Tests zum Niveau* α $(J = \mathfrak{F}_s)$*, d.h.*

$$\varphi^* \in \Phi_{J,\alpha} \quad \text{und} \quad \mathbb{E}_\theta(\varphi^*) = \max_{\varphi \in \Phi_{J,\alpha}} \mathbb{E}_\theta(\varphi) \quad \text{für alle } \theta \in K_1.$$

(b) φ^* *ist ein unverfälschter Test zum Niveau* α *für* H *gegen* K*, d.h.* $\varphi^* \in \Phi_\alpha^u$.

Beweis von Satz 29.15.
Die Beweisführung ist weitgehend analog zum Beweis des Satzes 29.6. □

Bemerkung 29.16
Als Teilmenge von K_1 sei hier

$$M_1 = \{\mathcal{N}(\mu, \sigma^2) : (\mu, \sigma^2) \in (0, \infty)^2\}$$

erwähnt.

29.2. Der diskrete Fall - Ausblick

Kritikpunkt von Anwendern:
Die Voraussetzung stetig verteilter Beobachtungen, die dafür sorgt, dass (mit Wahrscheinlichkeit 1) keine Bindungen innerhalb der Stichprobe entstehen.

Die Kritiker berufen sich auf die begrenzte Meßgenauigkeit in der Praxis, aufgrund derer man eigentlich immer nur diskrete Messdaten erhält. Daher ist die Annahme, dass nur bindungsfreie Realisationen möglich sind (selbst wenn die Beobachtungen tatsächlich von einer stetigen Verteilung herrühren), nicht gerechtfertigt.

In diesem Subabschnitt wird nun die Parametermenge

$$\mathcal{D} = \mathcal{D}(S) := \{(\mathbb{Q}_1, \mathbb{Q}_2) : \mathbb{Q}_i \text{ W'Maß mit Träger } \subset S \text{ und } \mathbb{Q}_1 \succeq \mathbb{Q}_2 \text{ oder } \mathbb{Q}_1 \preceq \mathbb{Q}_2\},$$

anstelle von Θ bzw.

$$\mathcal{D} = \mathcal{D}(S) := \{\mathbb{Q} : \mathbb{Q} \text{ W'Maß mit Träger } \subset S \text{ und } \mathbb{Q} \succeq \mathbb{Q}^- \text{ oder } \mathbb{Q} \preceq \mathbb{Q}^-\},$$

anstelle von $\mathfrak{F}$ zugrundegelegt. Hierbei ist S eine beliebige abzählbare Teilmenge von $\mathbb{R}$, wobei typischerweise $S = d\mathbb{Z}$ für ein $d > 0$ gilt.

Betrachten wir den Beweis von Aussage (a) des einseitigen Pitman-Zweistichprobentests, so fällt auf, dass nur bei der Berechnung der bedingten Verteilung $\mathbb{P}_\theta^{X|V(X)=\nu}$ die Stetigkeit der auftretenden Verteilungen verwendet wurde. Es verursacht aber keine größeren Probleme die Verteilung $\mathbb{P}_\theta^{X|V(X)=\nu}$ im diskreten Fall (also unter der Annahme, dass ν identische Einträge besitzen kann) zu berechnen. Daher können die Pitmanschen Permutationstests für den diskreten Fall aufgestellt und mit dem analogen Vorgehen wie im stetigen Fall bewiesen werden.

A. Anhang

A.1. Stochastische Majorisierung

Definition A.1 (stochastisch größer)
Seien $\mathbb{Q}$ und $\mathbb{P}$ zwei W'Maße auf $(\mathbb{R}, \mathcal{B})$. Dann heißt $\mathbb{Q}$ stochastisch größer als $\mathbb{P}$ (alternative $\mathbb{Q}$ majorisiert $\mathbb{P}$), falls

$$\mathbb{Q}((x,\infty)) \geq \mathbb{P}((x,\infty)) \quad [F_{\mathbb{Q}}(x) \leq F_{\mathbb{P}}(x)] \text{ für alle } x \in \mathbb{R}.$$

Notation: $\mathbb{Q} \succeq \mathbb{P}$.
Gilt außerdem $\mathbb{P} \neq \mathbb{Q}$, so schreiben wir $\mathbb{Q} \succ \mathbb{P}$.

Die anschauliche Interpretation liegt auf der Hand: $\mathbb{Q}$ ist stochastisch größer als $\mathbb{P}$, wenn $\mathbb{Q}$ für jeden Wert $x \in \mathbb{R}$ mehr Masse rechts von x als $\mathbb{P}$ besitzt. Dies impliziert:

Satz A.2
Seien X, Y Zufallsvariablen sowie $\mathbb{Q}, \mathbb{P}$ W'Maße mit $X \sim \mathbb{Q}, Y \sim \mathbb{P}$. Dann gilt die Implikation

$$\mathbb{Q} \succeq \mathbb{P} \quad \Rightarrow \quad \mathbb{E}X \geq \mathbb{E}Y.$$

A.2. Testtheorie

Setting: Gegeben sei ein statistisches Experiment $\mathcal{E} = (\mathcal{X}, \mathcal{A}, (W_\theta)_{\theta \in \Theta})$, bestehend aus einem messbaren Raum $(\mathcal{X}, \mathcal{A})$ und einer Familie $(W_\theta)_{\theta \in \Theta}$ von W'Maßen auf $(\mathcal{X}, \mathcal{A})$. Zudem sei X : $(\Omega, \mathcal{F}) \to (\mathcal{X}, \mathcal{A})$ eine Zufallsvariable mit $\mathbb{P}_\theta^X = W_\theta$.

Definition A.3 (Statistik/Test)
Sei $(\mathcal{X}', \mathcal{A}')$ ein Messraum. Eine messbare Funktion

$$\mathrm{T}: \ (\mathcal{X}, \mathcal{A}) \to (\mathcal{X}', \mathcal{A}')$$

heißt Statistik. Ein Test oder eine Testfunktion auf $\mathcal{X}$ ist eine Statistik φ mit

$$\varphi: \ \mathcal{X} \to [0, 1].$$

Wir betrachten in den folgenden Abschnitten ein Testproblem H gegen K mit

- $H, K \subset \Theta$
- $H \cup K = \Theta$ und
- $H \cap K = \emptyset$.

A.2.1. Gleichmäßig beste Tests

Definition A.4 (Test zum Niveau α)
Ein Test φ heißt Test zum Niveau $\alpha \in (0, 1)$, falls

$$\mathbb{E}_\theta \varphi(X) \leq \alpha$$

für alle $\theta \in H$. Es bezeichne Φ_α die Gesamtheit aller Tests zum Niveau α.

Definition A.5 (Gleichmäßig bester Test)
Ein Test φ^* heißt gleichmäßig bester Test zum Niveau $\alpha \in (0, 1)$, falls $\varphi^* \in \Phi_\alpha$ und

$$\mathbb{E}_\theta \varphi^*(X) = \max_{\varphi \in \Phi_\alpha} \mathbb{E}_\theta \varphi(X) \quad \text{für alle } \theta \in K.$$

Satz A.6 (Neyman-Pearson-Lemma)
Seien $\mathbb{Q}_0$ und $\mathbb{Q}_1$ W'Maße auf $(\mathcal{X}, \mathcal{A})$ mit Dichten f_0 bzw f_1 bzgl. eines dominierenden Maßes μ (etwa $\mu = \mathbb{Q}_0 + \mathbb{Q}_1$).
Ferner sei $\alpha \in (0,1)$ und $\Phi_\alpha = \{\varphi : \varphi$ Test mit $\int \varphi \, d\mathbb{Q}_0 \leq \alpha\}$. Dann gilt:

- *(Hinreichende Bedingung) Sei ψ ein Test mit*

  1. $\int \psi \, d\mathbb{Q}_0 = \alpha$

  2. *Es existiert ein $k \in [0, \infty)$ derart, dass*

$$\psi(x) = \begin{cases} 1 & f_1(x) > k f_0(x) \\ 0 & f_1(x) < k f_0(x) \end{cases} \quad \mu\text{-f.ü.} \tag{A.1}$$

  Dann folgt

$$\int \psi \, d\mathbb{Q}_1 = \max_{\varphi \in \Phi_\alpha} \int \varphi \, d\mathbb{Q}_1. \tag{A.2}$$

- *(Existenz) Es existiert ein Test ψ, der die Voraussetzungen* (A.1) *und* (A.2) *erfüllt.*

A.2.2. Unverfälschte Tests

Definition A.7 (Unverfälschtheit)
Ein Test φ heißt unverfälscht zum Niveau $\alpha \in (0,1)$, falls

- $\mathbb{E}_\theta \varphi(X) \leq \alpha$ *für alle* $\theta \in H$

- $\mathbb{E}_\theta \varphi(X) \geq \alpha$ *für alle* $\theta \in K$

Es bezeichne Φ_α^u die Gesamtheit aller unverfälschten Tests zum Niveau α.

Definition A.8 (Gleichmäßig bester unverfälschter Test)
Ein Test φ^* heißt gleichmäßig bester unverfälschter Test zum Niveau $\alpha \in (0,1)$, falls $\varphi^* \in \Phi_\alpha^u$ und

$$\mathbb{E}_\theta \varphi^*(X) = \max_{\varphi \in \Phi_\alpha^u} \mathbb{E}_\theta \varphi(X) \quad \text{für alle } \theta \in K.$$

A.2.3. J-ähnliche Tests

Definition A.9 (J-Ähnlichkeit)
Sei J eine beliebige Teilmenge von Θ. Ein Test $\varphi: \mathcal{X} \to [0,1]$ heißt ähnlich auf J zum Niveau α oder auch J-ähnlich zum Niveau α, falls

$$\mathbb{E}_\theta(\varphi(X)) = \alpha \quad \textit{für alle } \theta \in J.$$

Die Menge aller solcher Tests bezeichnen wir mit $\Phi_{J,\alpha}$.

Satz A.10
Bezeichne $\mathcal{W} = \mathcal{W}(\mathcal{X}, \mathcal{A})$ die Menge der W'Maße auf $(\mathcal{X}, \mathcal{A})$. Sei φ ein Test auf $(\mathcal{X}, \mathcal{A})$ und $\overline{H}, \overline{K}$ der topologische Abschluss von H, K bzgl. d_V ($d_V(\mathbb{P}, \mathbb{Q}) := \sup_{A \in \mathcal{A}} |\mathbb{P}(A) - \mathbb{Q}(A)|$). Dann gilt

1. *die Gütefunktion $\mathbb{P} \mapsto \mathbb{E}_\mathbb{P}\varphi$ ist gleichmäßig stetig auf $\mathcal{W}(\mathcal{X}, \mathcal{A})$ bezüglich d_V.*
2. *Ist φ ein unverfälschter Test zum Niveau α für H gegen K, wobei $H, K \subseteq \mathcal{W}$ beliebig mit $J = \overline{H} \cap \overline{K} \neq \emptyset$, so ist φ J-ähnlich zum Niveau α.*

A.2.4. Suffienz und Vollständigkeit von Statistiken

Definition A.11 (Suffizienz)
Eine Statistik $\mathrm{T} : (\mathcal{X}, \mathcal{A}) \to (\mathcal{X}', \mathcal{A}')$ heißt suffizient für $\mathcal{E}$, falls für jedes $A \in \mathcal{A}$ eine meßbare Abbildung $\mathrm{W(A|T = \cdot)} : \mathcal{X}' \to [0,1]$ existiert, so dass $W(A|T = \cdot)$ für jedes $\theta \in \Theta$ eine Version der faktorisierten bedingten Wahrscheinlichkeit $W_\theta(A|T = \cdot)$ bildet. Dies bedeutet, dass

$$W_\theta(A \cap T^{-1}(A')) = \int_{A'} W(A|T = t)\ W_\theta^T(dt)$$

für alle $A \in \mathcal{A}, A' \in \mathcal{A}'$ und $\theta \in \Theta$ gelten soll.

Eine Statistik T heißt also suffizient für $\mathcal{E}$, falsch die bedingte Verteilung $\mathbb{P}_\theta^{X|T(X)=t}$ unter jedem $\mathbb{P}_\theta$ dieselbe ist, d.h. von θ nicht mehr abhängt.

Definition A.12 (Vollständigkeit)
Eine Statistik $\mathrm{T} : (\mathcal{X}, \mathcal{A}) \to (\mathcal{X}', \mathcal{A}')$ heißt vollständig für $\mathcal{E}$, falls die Implikation gilt:

$$\mathbb{E}_\theta f(T(X)) = 0 \textit{ f.a. } \theta \in \Theta \Rightarrow f \equiv 0 \ \mathbb{P}_\theta^{T(X)}\textit{-f.s. f.a. } \theta \in \Theta$$

A.2.5. Bedingte Tests (Kapital 20 Alsmeyer)

Nun sei $\mathcal{E} = (\mathcal{X}, \mathcal{A}, (W_\theta)_{\theta\in\Theta})$ ein statistisches Experiment mit $\Theta \subset \mathbb{R}^d, d \geq 2$. Zudem bezeichne $T : (\mathcal{X}, \mathcal{A}) \to (\mathcal{X}', \mathcal{A}')$ eine suffiziente Statistik für $\mathcal{E}$ und wir definieren $\mathbb{E}_{\vartheta,T}(\varphi) := \mathbb{E}_\vartheta(\varphi \circ T(X))$.

Wegen der Suffizienz von T existiert zu jedem Text φ ein ebenso guter Test ϕ der Form $\psi \circ T$, und zwar $\phi = W(\varphi|T)$. Aus diesem Grund betrachten nur noch das durch T reduzierte Experiment $\mathcal{E}^T = (\mathcal{X}', \mathcal{A}', (W_\theta^T)_{\theta\in\Theta})$, wobei jetzt $\Phi_{J,\alpha}$ die Menge aller J-ähnlichen Tests $\psi : \mathcal{X}' \to [0,1]$ bezeichne.

Lemma A.13 (Lemma 20.2 im Alsmeyer)
Sei $V : (\mathcal{X}', \mathcal{A}') \to (\mathcal{X}'', \mathcal{A}'')$ eine vollständige und suffiziente Statistik für $(W_\theta^T)_{\theta\in J}$. Dann gilt:

$$\psi \in \Phi_{J,\alpha} \quad \Leftrightarrow \quad \mathbb{E}_{\cdot,T}(\psi|V) = \alpha \quad W_\theta^T\text{-f.s. für alle } \theta \in J.$$

Lemma A.14 (Lemma 20.3 im Alsmeyer)
Sei $\vartheta \in \Theta \setminus J$ und V wie in Lemma A.13. Dann gilt für $\psi^* \in \Phi_{J,\alpha}$:

$$\mathbb{E}_{\vartheta,T}(\psi^*) = \max_{\psi\in\Phi_{J,\alpha}} \mathbb{E}_{\vartheta,T}(\psi) \quad \Leftrightarrow \quad \mathbb{E}_{\vartheta,T}(\psi^*|V) \geq \mathbb{E}_{\vartheta,T}(\psi|V)\ W_\vartheta^T\text{-f.s. für alle } \psi \in \Phi_{J,\alpha}.$$

Zusammenfassend besagen A.13 und A.14, dass bei Vorliegen einer vollständigen und suffizienten Statistik V für $W_\theta^T)_{\theta\in J}$ die folgenden Problemstellungen äquivalent sind:

1. Finde einen Test $\psi^* :\ \mathcal{X}' \to [0,1]$ mit
   - $\mathbb{E}_{\theta,T}(\psi^*) = \alpha$ für alle $\theta \in J$
   - $\mathbb{E}_{\vartheta,T}(\psi^*) = \max_{\psi\in\Phi_{J,\alpha}} \mathbb{E}_{\vartheta,T}(\psi)$
2. Finde einen Test $\psi^* :\ \mathcal{X}' \to [0,1]$ mit
   - $\mathbb{E}_{\cdot,T}(\psi|V) = \alpha \quad W_\theta^T$-f.s. für alle $\theta \in J$
   - $\mathbb{E}_{\vartheta,T}(\psi^*|V) \geq \mathbb{E}_{\vartheta,T}(\psi|V) \quad W_\vartheta^T$-f.s. für alle $\psi \in \Phi_{J,\alpha}$