

ZUSAMMENFASSUNG STATISTIK 1

MARKUS KEPPELER

1. GRUNDLAGEN

1.1. Grundlegende Definitionen.

Definition 1. Ein statistisches Experiment ist ein Tripel

$$\xi = (\mathfrak{X}, \mathcal{A}, (\mathbb{P}_\vartheta)_{\vartheta \in \Theta}),$$

wobei \mathfrak{X} der Stichprobenraum ist, \mathcal{A} eine σ -Algebra über \mathfrak{X} und (\mathbb{P}_ϑ) eine Familie von Wahrscheinlichkeiten, die über Θ parametrisiert sind.

Definition 2. Eine messbare Funktion T auf $(\mathfrak{X}, \mathcal{A})$ heisst Statistik.

Definition 3. D sei die (nichtleere) Menge der möglichen Entscheidungen, \mathcal{D} eine σ -Algebra über D . Dann heisst (D, \mathcal{D}) Entscheidungsraum, und jede messbare Abbildung $\delta : \mathfrak{X} \rightarrow D$ heisst Entscheidungsfunktion.

Interpretation der Entscheidungsfunktion: Bei Vorliegen einer Stichprobe $x \in \mathfrak{X}$ liefert $\delta(x)$ die Entscheidung, die man bei Vorliegen von x trifft.

Definition 4. Eine Abbildung $L : \Theta \times D \rightarrow [0, \infty]$ heisst Verlustfunktion, wenn $L(\vartheta, \cdot) : d \mapsto L(\vartheta, d)$ für alle $\vartheta \in \Theta$ messbar ist.

Interpretation: L ordnet einer Entscheidung $d \in D$ den Verlust bei Vorliegen des „richtigen“ Parameters $\vartheta \in \Theta$ zu. Wenn also ϑ der richtige Parameter ist, so ist $L(\vartheta, d)$ der Verlust bei der Entscheidung für d .

Definition 5. Ein statistisches Modell ist ein Tripel

$$\mathcal{S} = (\xi, (D, \mathcal{D}), L)$$

aus einem statistischen Experiment, einem Entscheidungsraum und einer Verlustfunktion.

Definition 6. Wenn \mathcal{F} die Menge der zu einem statistischen Modell gehörigen Entscheidungsfunktionen sei, so heisst $R : \Theta \times \mathcal{F} \rightarrow [0, \infty]$ definiert durch

$$R(\vartheta, \delta) := \int_{\mathfrak{X}} L(\vartheta, \delta(x)) \mathbb{P}_\vartheta(dx) = \mathbb{E}_\vartheta L(\vartheta, \delta(X))$$

Risikofunktion. Die Risikofunktion ist somit der erwartete Verlust von δ bei vorliegendem ϑ .

2. SCHÄTZTHEORIE

2.1. Zulässigkeit.

Definition 7. Auf \mathcal{F} (die Menge der zu einem Experiment gehörigen Entscheidungsfunktionen) lässt sich folgende Ordnungsstruktur einführen:

$$\delta_1 \preceq \delta_2 \Leftrightarrow R(\vartheta, \delta_1) \leq R(\vartheta, \delta_2) \quad \forall \vartheta \in \Theta,$$

d.h. „ δ_1 ist nicht schlechter als δ_2 “. Weiter wird eingeführt:

$$\delta_1 \prec \delta_2 \Leftrightarrow \delta_1 \preceq \delta_2 \text{ und } R(\vartheta, \delta_1) < R(\vartheta, \delta_2) \text{ für mind. ein } \vartheta \in \Theta,$$

d.h. „ δ_1 ist besser als δ_2 “.

Definition 8. Für $\mathcal{K} \subset \mathcal{F}$ heisst eine Entscheidungsfunktion δ_0 gleichmäßig beste Entscheidungsfunktion in \mathcal{K} , falls gilt

$$\delta_0 \in \mathcal{K} \text{ und } \delta_0 \preceq \delta \quad \forall \delta \in \mathcal{K}$$

Im allgemeinen Fall existiert nicht immer eine gleichmäßig beste Entscheidungsfunktion.

Definition 9. Eine Entscheidungsfunktion $\delta_0 \in \mathcal{F}$ heisst zulässig, falls kein $\delta \in \mathcal{F}$ existiert mit $\delta \prec \delta_0$.

Definition 10. Ein $\delta_0 \in \mathcal{F}$ heisst Minimaxverfahren, falls gilt:

$$\sup_{\vartheta \in \Theta} R(\vartheta, \delta_0) \leq \sup_{\vartheta \in \Theta} R(\vartheta, \delta) \text{ für alle } \delta \in \mathcal{F}.$$

Dies lässt sich (unter der Annahme das alle auftretenden Suprema auch Maxima sind) alternativ wie folgt beschreiben:

$$\max_{\vartheta \in \Theta} R(\vartheta, \delta_0) = \min_{\delta \in \mathcal{F}} \max_{\vartheta \in \Theta} R(\vartheta, \delta),$$

was die Bezeichnung „Minimax“ erklärt.

Definition 11. Auf Θ seien eine σ -Algebra und ein W -Maß ξ gegeben. ξ wird als a-priori-Verteilung bezeichnet. Dann heisst $\delta_0 \in \mathcal{F}$ ein Bayes-Verfahren zu ξ , falls gilt

$$\int_{\Theta} R(\vartheta, \delta_0) \xi(d\vartheta) \leq \int_{\Theta} R(\vartheta, \delta) \xi(d\vartheta) \text{ für alle } \delta \in \mathcal{F}.$$

Das W -Maß ξ drückt hierbei eine gewisse Vorbewertung der Parameter aus.

2.2. Exponentialfamilien.

Definition 12. Eine durch ν dominierte Familie von W -Maßen $(\mathbb{P}_{\vartheta})_{\vartheta \in \Theta}$ heisst Exponentialfamilie, falls die ν -Dichten folgende Gestalt besitzen: Es existiert ein $k \in \mathbb{N}$ und Abbildungen

$$C, Q_1, \dots, Q_k : \Theta \rightarrow \mathbb{R}$$

und messbare Abbildungen

$$h, T_1, \dots, T_k : \mathfrak{X} \rightarrow \mathbb{R},$$

so dass für jedes $\vartheta \in \Theta$ gilt:

$$f_{\vartheta} = C(\vartheta) \exp\left(\sum_{i=1}^k Q_i(\vartheta) T_i\right) h \quad \nu\text{-f.s.}$$

Satz 1. Sind X_1, \dots, X_n unabhängige Zufallsvariablen, deren Verteilungen jeweils eine Exponentialfamilie bilden, so ist auch die Verteilung von

$$X = (X_1, \dots, X_n)$$

eine Exponentialfamilie.

3. FINDEN VON SCHÄTZERN

3.1. Einleitende Definitionen.

Definition 13. Ein Schätzer g für $\gamma(\vartheta)$ heisst erwartungstreu, falls $\mathbb{E}_{\vartheta} g = \gamma(\vartheta)$ gilt.

Im Folgenden werden die wichtigsten drei Methoden zum Finden von Schätzern vorgestellt:

3.2. Momentenmethode.

Definition 14. Seien X_1, \dots, X_n unabhängige, identisch verteilte Zufallsvariablen zum schätzen von $\vartheta = (\vartheta_1, \dots, \vartheta_d)$. Eine Lösung $\hat{\vartheta} = \hat{\vartheta}(x_1, \dots, x_n)$ für das Gleichungssystem

$$\begin{aligned}\bar{x}_n^{(1)} &= \frac{x_1 + \dots + x_n}{n} = \mathbb{E}_\vartheta X_1 \\ \bar{x}_n^{(2)} &= \frac{x_1^2 + \dots + x_n^2}{n} = \mathbb{E}_\vartheta X_1^2 \\ &\vdots \\ \bar{x}_n^{(d)} &= \frac{x_1^d + \dots + x_n^d}{n} = \mathbb{E}_\vartheta X_1^d\end{aligned}$$

heißt Momentenmethode-Schätzer für ϑ .

3.3. Maximum-Likelihood-Methode.

Definition 15. Sei ξ ein durch ν dominiertes Experiment mit Dichten $f_\vartheta = \frac{d\mathbb{P}_\vartheta}{d\nu}$. Für gegebenen Beobachtungswert $X = x \in \mathfrak{X}$ wird $L(\cdot|x) : \Theta \rightarrow [0, \infty)$ definiert durch

$$L(\vartheta|x) = f_\vartheta(x)$$

die Likelihood-Funktion von X (gegeben x) genannt. Ausserdem heisst $l(\vartheta|x) = \log L(\vartheta|x)$ die Log-Likelihood-Funktion von X (gegeben x).

Definition 16. Existiert ein $\hat{\vartheta}(x) \in \Theta$, in dem $L(\cdot|x)$ ein Maximum annimmt, so heisst $\hat{\vartheta}(x)$ Maximum-Likelihood-Schätzer für ϑ auf der Basis von x .

Die Idee hinter dem MLS ist die Verteilung zu wählen, unter der die Wahrscheinlichkeit der gegebenen Beobachtung maximal ist.

Satz 2. Sei γ eine Parameterfunktion und $\hat{\vartheta}(x)$ ein MLS für ϑ bei gegebenem $x \in \mathfrak{X}$. Dann ist $\hat{\nu}(x) = \gamma(\hat{\vartheta}(x))$ ein MLS für $\gamma(\vartheta)$.

3.4. Bayes-Methode.

Satz 3. Wenn g eine Dichte der a-priori-Verteilung ξ ist, und f_ϑ die Verteilung der Zufallsvariable, so bildet

$$\frac{f_\vartheta(x)g(\vartheta)}{f_\xi(x)}\mathbf{1}_{\{f_\xi > 0\}} + g(\vartheta)\mathbf{1}_{\{f_\xi = 0\}}$$

eine Dichte der a-posteriori-Verteilung $f^{\Lambda|X=x}$. Hierbei ist

$$f_\xi(x) = \int_{\Theta} f_\vartheta(x)d\xi(\vartheta).$$

Satz 4. Für den Bayes-Schätzer $\hat{\delta}(x)$ gilt:

$$\hat{\delta}(x) = \mathbb{E}(\Lambda|X = x)$$

Achtung: In obigen Sätzen zur Bayes-Methode wurden einige Voraussetzungen weggelassen, da sie der Übersicht schaden.

4. DATENREDUKTION UND SUFFIZIENZ

4.1. Einleitende Definitionen.

Definition 17. Es sei $\{\mathbb{P}_\vartheta : \vartheta \in \Theta\}$ eine Menge von W -Maßen auf $(\mathfrak{X}, \mathcal{A})$. Eine Statistik $T : \mathfrak{X} \rightarrow \mathfrak{Y}$ heisst suffizient für (\mathbb{P}_ϑ) , falls

$$\mathbb{P}_\vartheta^{X|T=t} = \mathbb{P}_\vartheta[X \in \cdot | T = t] \quad \forall t$$

nicht mehr von ϑ abhängt.

Definition 18. Eine suffiziente Statistik T^* heisst minimalsuffizient, wenn sie über jeder weiteren suffizienten Statistik T faktorisiert, d.h. wenn

$$T^* = h \circ T \quad \text{f.s.}$$

für eine geeignete messbare Funktion h gilt.

4.2. Neyman-Kriterium für Suffizienz.

Satz 5. Sei $(\mathbb{P}_\vartheta)_{\vartheta \in \Theta}$ eine durch μ dominierte Familie von Wahrscheinlichkeitsmaßen. Genau dann ist eine Statistik $T : (\mathfrak{X}, \mathcal{A}) \rightarrow (\mathfrak{X}', \mathcal{A}')$ suffizient, wenn für jedes $\vartheta \in \Theta$ eine messbare Funktion g_ϑ und eine messbare Funktion h existieren, so dass $(g_\vartheta \circ T)h$ eine μ -Dichte von \mathbb{P}_ϑ bildet.

Bemerkung 1. Für eine Exponentialfamilie in T und Q (vollen Ranges) ist T suffizient.

Satz 6. Falls T eine suffiziente Statistik ist, so ist jede weitere Statistik S mit $T = k \circ S$ für eine messbare Abbildung k ebenfalls suffizient.

5. VOLLSTÄNDIGKEIT, VERTEILUNGSFREIHEIT UND DER SATZ VON BASU

5.1. Verteilungsfreiheit.

Definition 19. Eine Statistik T heisst verteilungsfrei, falls $\mathbb{P}_\vartheta^{T(X)}$ nicht mehr von ϑ abhängt.

Definition 20. Eine Statistik T heisst verteilungsfrei 1. Ordnung, falls $\mathbb{E}_\vartheta T(X)$ nicht von ϑ abhängt.

5.2. Vollständigkeit.

Definition 21. Eine Statistik T heisst vollständig, falls gilt

$$\mathbb{E}_\vartheta f(T(X)) = 0 \quad \forall \vartheta \in \Theta \Rightarrow f = 0 \quad \mathbb{P}_\vartheta^{T(X)}\text{-f.s.} \quad \forall \vartheta \in \Theta$$

Satz 7. Ist T vollständig und suffizient, so ist T minimalsuffizient.

Satz 8. Ist T eine vollständige Statistik, so ist auch $f \circ T$ vollständig für jede messbare Funktion f .

Bemerkung 2. Ist $(\mathbb{P}_\vartheta)_{\vartheta \in \Theta}$ eine k -parametrische Exponentialfamilie und besitzt $\mathcal{Q} := \{Q_1(\Theta), \dots, Q_k(\Theta)\}$ innere Punkte, so ist $T = (T_1, \dots, T_k)$ vollständig.

5.3. Basu.

Satz 9 (Satz von Basu). Sei T eine vollständige und suffiziente Statistik. Dann ist jede verteilungsfreie Statistik unabhängig von T .

6. GLEICHMÄSSIG BESTE ERWARTUNGSTREUE SCHÄTZER

6.1. Einleitende Definitionen. Im Folgenden sei $(D, \mathcal{D}) = (\mathbb{R}, \mathcal{B})$, sowie $L(\vartheta, d) = (d - \gamma(\vartheta))^2$. Die Risikofunktion eines Schätzers $g : \mathfrak{X} \rightarrow \mathbb{R}$ hat somit die Form

$$R(\vartheta, g) = \int (g(x) - \gamma(\vartheta))^2 \mathbb{P}_\vartheta(dx) = \mathbb{E}_\vartheta(g(X) - \gamma(\vartheta))^2$$

Wir beschränken uns im Folgenden auf erwartungstreue Schätzer (für $\gamma(\vartheta)$), d.h. Schätzer, für die gilt

$$\mathbb{E}_\vartheta g = \int g(x) \mathbb{P}_\vartheta(dx) = \gamma(\vartheta).$$

Für solche Schätzer gilt für das Risiko:

$$R(\vartheta, g) = E_\vartheta(g(X) - \mathbb{E}_\vartheta g(X))^2 = \mathbb{V}_\vartheta g(X)$$

für alle $\vartheta \in \Theta$.

Definition 22. Sei \mathcal{U}_γ die Menge aller erwartungstreuen Schätzer für $\gamma(\vartheta)$. Ein Schätzer $g^* \in \mathcal{U}_\gamma$ heisst gleichmässig bester erwartungstreuer Schätzer für $\gamma(\vartheta)$, wenn für alle $\vartheta \in \Theta$ und $g \in \mathcal{U}_\gamma$ gilt:

$$R(\vartheta, g^*) = \mathbb{V}_\vartheta g^*(X) \leq \mathbb{V}_\vartheta g(X) = R(\vartheta, g).$$

6.2. Finden eines GBES.

Satz 10 (Rao und Blackwell). *Es sei g ein erwartungstreuer Schätzer für $\gamma(\vartheta)$ und T eine suffiziente Statistik. Definiere g^* durch*

$$g^*(t) = \mathbb{P}(g|T = t) = \mathbb{E}(g(X)|T(X) = t)$$

und man erhält eine neue Schätzfunktion $g^* \circ T$, für die gilt:

- (1) $g^* \circ T$ ist ebenfalls erwartungstreu.
- (2) $\mathbb{V}_\vartheta g^* \circ T(X) \leq \mathbb{V}_\vartheta g(X)$ für alle $\vartheta \in \Theta$.

Der Satz sagt also, dass zu einem gegebenen erwartungstreuen Schätzer bei Vorliegen einer suffizienten Statistik durch Bedingen stets ein mindestens ebensoguter Schätzer ($g^* \circ T$) konstruiert werden kann.

Satz 11 (Lehmann und Scheffe). *Ist in der Situation von Rao/Blackwell die Statistik T sogar vollständig, so ist $g^* \circ T$ sogar ein GBES.*

Bemerkung 3. *Existiert ein GBES für $\gamma(\vartheta)$, so ist dieser bereits (f.s.) eindeutig bestimmt.*

Bemerkung 4. *Ein gleichmäßig bester erwartungstreuer Schätzer muss nicht zulässig sein!*

7. INFORMATIONS-UNGLEICHUNG

7.1. Regularitätsbedingungen. Es gelten im Folgenden folgende Regularitäten:

- Θ ist offene Teilmenge von \mathbb{R} .
- Es existieren Versionen $f(\vartheta, \cdot) = \frac{d\mathbb{P}_\vartheta}{d\nu}$, so dass $\mathfrak{X}^* \stackrel{\text{def}}{=} \{x \in \mathfrak{X} : f_\vartheta(x) > 0\}$ nicht von ϑ abhängt (insbesondere sind die Maße paarweise äquivalent).
- Für alle $\vartheta \in \Theta$ und $x \in \mathfrak{X}^*$ existiert $\frac{d}{d\vartheta} f(\vartheta, x)$ und ist stetig in ϑ .

7.2. Cramer-Rao-Schranke.

Definition 23. Die Fisher-Information $I(\vartheta)$ ist definiert durch

$$I(\vartheta) \stackrel{\text{def}}{=} \mathbb{E}_{\vartheta} \left(\frac{d}{d\vartheta} \log f_{\vartheta}(X) \right)^2 = \mathbb{V}_{\vartheta} \left(\frac{d}{d\vartheta} \log f_{\vartheta}(X) \right).$$

Satz 12. Falls eine Verteilungsfamilie (\mathbb{P}_{ϑ}) die Fisher-Information $I(\vartheta)$ besitzt, so besitzt die Familie $(\otimes_{i=1}^n \mathbb{P}_{\vartheta})$ der n -fachen Produktverteilungen die Fisher-Information $nI(\vartheta)$.

Satz 13 (Cramer-Rao-Schranke). Sei T eine Statistik mit $\mathbb{V}_{\vartheta}T < \infty$. Dann gilt

$$\mathbb{V}_{\vartheta}T(X) \geq \frac{(\mathbb{E}'_{\vartheta}T(X))^2}{I(\vartheta)}$$

für alle $\vartheta \in \Theta$.

Satz 14. Gilt im Satz 13 zusätzlich, dass T erwartungstreu ist, so gilt

$$\mathbb{V}_{\vartheta}T(X) \geq \frac{1}{I(\vartheta)}$$

Satz 15. Bildet die Verteilungsfamilie eine einparametrische Exponentialfamilie in T und Q , so nimmt $T(x)$ in jedem ϑ die Schranke der Informationsungleichung an und ist damit GBES für $\mathbb{E}_{\vartheta}T(X)$.

8. LINEARE MODELLE / KQS

8.1. Lineare Modelle.

Definition 24. Ein lineares Modell liegt vor, wenn der Beobachtungsvektor $X = (X_1, \dots, X_n)^t$ die Regressionsgleichung

$$X = A\vartheta + \epsilon$$

erfüllt, mit $\vartheta = (\vartheta_1, \dots, \vartheta_p)^t$ und $\epsilon = (\epsilon_1, \dots, \epsilon_n)^t$.

8.2. Kleinste-Quadrate-Schätzer.

Definition 25. Ein Schätzer $\hat{\vartheta} = (\hat{\vartheta}_1, \dots, \hat{\vartheta}_p)^t$ heisst Kleinste-Quadrate-Schätzer für ϑ , wenn

$$(x - A\hat{\vartheta}(x))^t(x - A\hat{\vartheta}(x)) = \min_{\vartheta \in \Theta} (x - A\vartheta)^t(x - A\vartheta)$$

für alle $x \in \mathbb{R}^n$ gilt.

Definition 26. Die Normalengleichung bezeichnet die Gleichheit

$$A^t A \hat{\vartheta}(x) = A^t x.$$

$\hat{\vartheta}$ ist genau dann KQS, wenn er die Normalengleichung erfüllt.

Satz 16. In einem linearen Modell vollen Ranges gilt für den KQS $\hat{\vartheta}$:

$$\hat{\vartheta}(x) = (A^t A)^{-1} A^t x$$

8.3. Lineare Schätzer.

Definition 27. Ein Schätzer g_b heisst linearer Schätzer, falls er von der Gestalt

$$g_b(x) = b^t x$$

für ein $b \in \mathbb{R}^n$ ist. g^* heisst BLUE (best linear unbiased estimator), falls er linear und erwartungstreu ist, und seine Varianz minimal ist, d.h. wenn

$$\mathbb{V}_\vartheta g^* = \min_{b \in \mathbb{R}^n: \mathbb{E}_\vartheta g_b(X) = \gamma(\vartheta)} \mathbb{V}_\vartheta g_b(X)$$

gilt.

Satz 17 (Gauß-Markov). Wenn $\hat{\vartheta}$ der KQS für ϑ ist, so ist $\beta^t \hat{\vartheta}$ der BLUE für $\gamma(\vartheta) = \beta^t \vartheta$.

Satz 18. Wenn in dem linearen Modell die ϵ_i normalverteilt sind, dann ist $\beta^t \hat{\vartheta}$ sogar der Beste unter allen erwartungstreuen Schätzern für $\gamma(\vartheta) = \beta^t \vartheta$, also der GBES (statt nur BLUE).

9. TESTTHEORIE

9.1. Einleitende Definitionen. Getestet werden soll eine Hypothese H gegen eine Alternative K .

Definition 28. Eine Test ist eine Funktion

$$\varphi : \mathfrak{X} \rightarrow [0, 1].$$

φ heisst nicht-randomisierter Test, falls $\varphi(x) \in \{0, 1\}$ gilt. Ansonsten heisst φ randomisierter Test.

$\varphi(x) = \gamma \in [0, 1]$ gibt die Wahrscheinlichkeit an, sich bei Vorliegen von x für die Alternative K zu entscheiden.

Definition 29. Als Verlustfunktion wird die Neymann-Pearsonsche Funktion gewählt:

$$L(\vartheta, \gamma) = \begin{cases} \gamma & \vartheta \in H \\ 1 - \gamma & \vartheta \in K \end{cases}$$

Die Verlustfunktion gibt jeweils die Wahrscheinlichkeit an sich falsch zu entscheiden. Das zugehörige Risiko ist dann

$$R(\vartheta, \varphi) = \begin{cases} \int \varphi d\mathbb{P}_\vartheta & \vartheta \in H \\ 1 - \int \varphi d\mathbb{P}_\vartheta & \vartheta \in K \end{cases}$$

Definition 30. Als Gütefunktion eines Tests φ bezeichnen wir die Funktion β_φ mit

$$\beta_\varphi : \vartheta \mapsto \int \varphi d\mathbb{P}_\vartheta$$

Definition 31. Ein Test zum Niveau α ist ein Test φ für den $\beta_\varphi(\vartheta) \leq \alpha$ für alle $\vartheta \in H$ gilt. Die Menge aller Tests zum Niveau α wird mit Φ_α bezeichnet, Φ ist die Menge aller Tests.

9.2. Das Neymann-Pearson-Lemma.

Satz 19 (Neymann-Pearson). *Es seien \mathbb{P}_0 und \mathbb{P}_1 Wahrscheinlichkeitsmaße mit Dichten f_0 und f_1 bezüglich eines dominierenden Maßes. Zu testen sei $H = \{\vartheta_0\}$ gegen $K = \{\vartheta_1\}$. Dann gilt:*

- Gilt $\psi \in \Phi_\alpha$ mit $\int \psi d\mathbb{P}_0 = \alpha$, und ist ψ von der Gestalt

$$(1) \quad \psi(x) = \begin{cases} 1 & f_1(x) > k f_0(x) \\ 0 & f_1(x) < k f_0(x) \end{cases}$$

für ein konstantes k , so gilt

$$(2) \quad \beta_\psi(\vartheta_1) = \max_{\varphi \in \Phi_\alpha} \beta_\varphi(x).$$

- Für jedes $\alpha \in (0, 1)$ gibt es einen Test wie in (1).
- Gibt es einen Test $\psi \in \Phi_\alpha$, der (2) erfüllt, so hat ψ die Gestalt (1). Gilt zudem $\int \psi d\mathbb{P}_1 < 1$, so folgt schon $\int \psi d\mathbb{P}_0 = \alpha$. Dies bedeutet, dass der beste Test zum Niveau α das Niveau auch voll ausschöpft.

Satz 20. *In der Situation von Satz 19 gilt für einen besten Test $\psi: \mathbb{E}_{\vartheta_1} \psi \geq \alpha$.*

9.3. Einseitige Test bei monotonem Dichtequotienten. Bisher haben wir Hypothesen der Form $H = \{\vartheta_0\}$ gegen Alternativen der Form $K = \{\vartheta_1\}$ getestet. Jetzt möchten wir Hypothesen der Form $H = \{\vartheta : \vartheta \leq \vartheta_0\}$ gegen Alternativen der Gestalt $K = \{\vartheta : \vartheta > \vartheta_0\}$ testen (oder andersherum). Da der oben konstruierte beste Test φ^* unabhängig von dem $\vartheta_1 > \vartheta_0$ ist, ist er auch schon gleichmäßig bester Test für $H = \{\vartheta_0\}$ gegen $K = \{\vartheta : \vartheta > \vartheta_0\}$. Um dies auf $H = \{\vartheta : \vartheta \leq \vartheta_0\}$ zu verallgemeinern, werden wir Forderungen an den Dichtequotienten stellen müssen.

Definition 32. *Ein einseitiges Testproblem für den eindimensionalen Parameter ϑ ist ein Testproblem der Form*

$$H = \{\vartheta : \vartheta \leq \vartheta_0\} \text{ gegen } K = \{\vartheta : \vartheta > \vartheta_0\},$$

wobei H und K jeweils nicht leer sein sollen.

Definition 33. *Sei $\mathcal{P} = \{\mathbb{P}_\vartheta : \vartheta \in \vartheta\}$ eine durch ein Maß μ dominierte Familie von Wahrscheinlichkeitsmaßen und $T: \mathcal{X} \rightarrow \mathbb{R}$ sei eine Statistik.*

Wir sagen \mathcal{P} hat (streng) isotone¹ Dichtequotienten in T , falls es für jedes Paar $\vartheta_0 < \vartheta_1$ eine (streng) isotone Funktion $H_{\vartheta_0, \vartheta_1}$ gibt, so dass

$$\frac{f_1}{f_2} = \frac{d\mathbb{P}_{\vartheta_1}}{d\mu} / \frac{d\mathbb{P}_{\vartheta_0}}{d\mu}(x) = H_{\vartheta_0, \vartheta_1} \circ T(x) \quad \mu\text{-f.s.}$$

gilt. $\frac{f_1}{f_2}(x)$ heisst Dichtequotient oder auch Likelihoodquotient.

Satz 21. *Sei \mathcal{P} eine Verteilungsfamilie mit streng isotonen Dichtequotienten in einer Statistik T , und sei $\alpha \in (0, 1)$ ein Niveau. Weiterhin seien $\vartheta_0 < \vartheta_1 \in \vartheta$ und $H = \{\vartheta : \vartheta \leq \vartheta_0\}$ und $K = \{\vartheta : \vartheta > \vartheta_0\}$ und*

$$\varphi^*(x) = \begin{cases} 1 & T(x) > k \\ \gamma & T(x) = k \\ 0 & T(x) < k, \end{cases}$$

wobei wir k und γ so wählen, dass

$$\mathbb{P}_{\vartheta_0}(T(x) > k) + \gamma \mathbb{P}_{\vartheta_0}(T(x) = k) = \alpha$$

ist. Dann gilt:

¹isoton = monoton wachsend

- (1) φ^* minimiert gleichmäßig die Fehlerwahrscheinlichkeiten erster und zweiter Art unter allen Test φ von H gegen K mit $\mathbb{E}_{\vartheta_0}\varphi = \alpha$.
- (2) φ^* ist gleichmäßig bester Test zum Niveau α von H gegen K .
- (3) $\vartheta \mapsto \mathbb{E}_{\vartheta}\varphi^*$ ist streng isoton auf $\{\vartheta : 0 < \mathbb{E}_{\vartheta}\varphi^* < 1\}$.

Satz 22. In jedem einseitigem Testproblem mit monotonen Dichtequotienten in T haben die gleichmäßig besten Tests zu einem Niveau α die Form

$$\varphi^*(x) = \mathbf{1}_{(k^*, \infty)}(T(x)) + \gamma \mathbf{1}_{\{k^*\}}(T(x))$$

bzw.

$$\varphi^*(x) = \mathbf{1}_{(-\infty, k^*)}(T(x)) + \gamma \mathbf{1}_{\{k^*\}}(T(x))$$

Satz 23. In der Situation von Satz 21 sei $h : \mathbb{R} \rightarrow \mathbb{R}$ eine weitere streng isotone Funktion und $\tilde{T} = h \circ T$. Es sei

$$\tilde{\varphi}(x) = \begin{cases} 1 & \tilde{T}(x) > \tilde{k} \\ \tilde{\gamma} & \tilde{T}(x) = \tilde{k} \\ 0 & \tilde{T}(x) < \tilde{k} \end{cases}$$

Dabei seien \tilde{k} und $\tilde{\gamma}$ so gewählt, dass $\mathbb{E}_{\vartheta_0}\tilde{\varphi} = \alpha$ gilt. Dann gilt $\varphi^* = \tilde{\varphi}$ (fast sicher).