Seminar: Simulation und Bildanalyse mit Java

Einführung in die statistische Testtheorie II

Guntram Seitz

Sommersemester 2004

1 Wiederholung

Grundprinzip:

- Annahme: Beobachtungen bzw. Daten genügen einem gewissen parametrischen Modell
- Somit werden gegebene Beobachtungen $(x_1, \ldots, x_n)$ als Realisierungen von iid. Zufallsvariablen $(X_1, \ldots, X_n)$ aufgefasst
- Hypothese über die genaue Art des Modells (den Parameter) aufstellen

1.1 Die statistische Hypothese

Parametrisches Modell: $(\Omega, \mathcal{F}, P_\theta)$ mit $\Omega \subset \mathbb{R}^n$
Sei $P_\theta \in \mathcal{P} = \{P_\theta \mid \theta \in \Theta\}$ aus einer parametrischen Familie von Verteilungen.
Der Parameterraum Θ wird zerlegt: $\Theta = \Theta_0 \cup \Theta_1$

Null-Hypothese: $H_0 : \theta \in \Theta_0$ gegen die

Alternativ-Hypothese: $H_1 : \theta \notin \Theta_0$

1.1.1 Mögliche Fehler

Fehler 1. Art: H_0 wird irrtümlicherweise verworfen

Fehler 2. Art: H_0 wird irrtümlicherweise nicht verworfen

	H_0 ist korrekt	H_0 ist nicht korrekt
H_0 wird nicht verworfen	korrekt	Fehler 2. Art
H_0 wird verworfen	Fehler 1. Art	korrekt

1.2 Der statistische Test

Zerlege den Stichprobenraum $\Omega = \mathbb{R}^n = K \cup A$, wobei K der kritische Bereich, bzw. A der Akzeptanzbereich genannt wird. Wir betrachten nun die Entscheidungsregel:

$$\varphi(x_1, \ldots, x_n) = \begin{cases} 1 & \text{falls } (x_1, \ldots, x_n) \in K \\ 0 & \text{falls } (x_1, \ldots, x_n) \in \mathbb{R}^n \setminus K \end{cases}$$

Definition: Test Die Stichprobenfunktion $\varphi : \mathbb{R}^n \to \{0,1\}$ heißt *(nichtrandomisierter) Test* zum Niveau α falls:

$$P_\theta(\varphi(X_1,\dots,X_n) = 1) \le \alpha \quad \forall\ \theta \in \Theta_0$$

wobei $(X_1,\dots,X_n)$ die Zufallsstichprobe bezeichnet.

1.2.1 Fehlerwahrscheinlichkeiten

Wahrscheinlichkeit für Fehler 1. Art:

$$\alpha(\theta) = P_\theta(\varphi(X_1, \ldots, X_n) = 1) \text{ für } \theta \in \Theta_0$$

Bei einfachen Hypothesen ist $\alpha = \alpha(\theta)$.

Wahrscheinlichkeit für Fehler 2. Art:

$$\beta(\theta) = P_\theta(\varphi(X_1, \ldots, X_n) = 0) \text{ für } \theta \in \Theta_1$$

1.2.2 Durchführung eines Tests

- Testniveau: Vorgabe eines Testniveaus α (max. Wkt. für Fehler 1. Art)
- Testfunktion: Wahl einer geeigneten Teststatistik $T(X_1, \ldots, X_n)$, deren Verteilung vom zu testenden Parameter θ abhängt
- Ablehnungsbereich: Bestimmung eines Ablehnungsbereichs $K' \subset \mathbb{R}$ mit $P_\theta(T \in K') \leq \alpha$ für $\theta \in \Theta_0$
- Entscheidung: Falls die Realisierung der Testfunktion im kritischen Bereich liegt, wird H_0 verworfen

2 Methoden zur Konstruktion von Tests

Wir betrachten nun zwei verschiedene Methoden zur Konstruktion von statistischen Tests:

1. Maximum-Likelihood-Quotienten-Tests
2. Union-Intersection- bzw. Intersection-Union-Tests

2.1 Maximum-Likelihood-Quotienten-Tests

2.1.1 Maximum-Likelihood-Schätzer

Ziel ist es die Modellverteilung möglichst gut an die beobachteten Daten $x = (x_1, \ldots, x_n)$ anzupassen. Die Likelihood-Funktion ist folgendermaßen definiert:

$$L : \mathbb{R}^n \times \Theta \to \mathbb{R} \text{ mit } L(x_1, \ldots, x_n; \theta) = f(x_1, \ldots, x_n; \theta)$$

Hier betrachten wir nur den absolutstetigen Fall, somit ist $f(\cdot, \theta)$ die Dichte von P_θ und $f(x_1, \ldots, x_n; \theta) = f(x_1; \theta) \cdots f(x_n; \theta)$.
Maximum-Likelihood-Schätzer $\hat{\theta}$ für θ:

$$L(x; \hat{\theta}) = \sup_{\theta \in \Theta} L(x; \theta) \quad \forall\, x \in \Omega$$

2.1.2 Der Maximum-Likelihood-Quotienten-Test

H_0 wird genau dann verworfen, wenn $T(x_1, \ldots, x_n) > c$ für eine Konstante c und

$$T(x_1, \ldots, x_n) = \frac{\sup_{\theta \in \Theta} f(x_1, \ldots, x_n; \theta)}{\sup_{\theta \in \Theta_0} f(x_1, \ldots, x_n; \theta)} = \frac{f(x_1, \ldots, x_n; \hat{\theta})}{f(x_1, \ldots, x_n; \tilde{\theta})}$$

wobei $\hat{\theta}$ ein unrestringierter ML-Schätzer und $\tilde{\theta}$ ein ML-Schätzer unter der Bedingung $\theta \in \Theta_0$ ist.

Bestimmung der Konstanten c: Die Konstante c bestimmt sich durch die Bedingung

$$sup_{\theta\in\Theta_0} P_\theta(T(X_1,\ldots,X_n) > c) \leq \alpha$$

Hierfür ist nützlich: Unter H_0 und gewissen Regularitätsbedingungen ist $2\ln T(X_1,\ldots,X_n)$ für $n \to \infty$ asymptotisch χ^2_d-verteilt, wobei d die Differenz der Dimension von Θ und der Anzahl der betrachteten Parameter in Θ_0 ist.

2.1.3 Beispiel: t-Test als Likelihood-Quotienten Test

Gegeben sei eine Zufallsstichprobe $(X_1, \dots, X_n)$ mit $X_i \sim N(\mu, \sigma^2)$ unabhängig und identisch verteilt. Wir testen nun

$$H_0 : \mu = \mu_0 \text{ gegen } H_1 : \mu \neq \mu_0$$

Hier ist dann

$$\Theta_0 = \{(\mu_0, \sigma^2) : \sigma^2 > 0\}$$

und

$$\Theta_1 = \{(\mu, \sigma^2) : \mu \neq \mu_0, \sigma^2 > 0\}.$$

Die Likelihood-Funktion ist gegeben durch:

$$L(x_1, \dots, x_n; \mu, \sigma^2) = \frac{1}{(\sigma\sqrt{2\pi})^n} \exp\left\{-\frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2}\right\}$$

Die unrestringierten ML-Schätzer:

$$\hat{\mu} = \bar{x} \quad \hat{\sigma}^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n}$$

wobei $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ das Stichprobenmittel ist.

Damit ergibt sich für $\sup_{(\mu,\sigma^2)\in\Theta} L(x;\mu,\sigma^2)$:

$$\sup_{(\mu,\sigma^2)\in\Theta} L(x;\mu,\sigma^2) = L(x;\hat{\mu},\hat{\sigma}^2) = \frac{e^{-n/2}}{(2\pi/n)^{n/2}\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\right)^{n/2}}$$

Unter H_0 ist ein ML-Schätzer fur σ^2 gegeben durch:

$$\tilde{\sigma}^2 = \sum_{i=1}^{n} \frac{(x_i - \mu_0)^2}{n}$$

Somit gilt für $\sup_{(\mu,\sigma^2)\in\Theta_0} L(x;\mu,\sigma^2)$:

$$\sup_{(\mu,\sigma^2)\in\Theta_0} L(x;\mu,\sigma^2) = L(x;\mu_0,\tilde{\sigma}^2) = \frac{e^{-n/2}}{(2\pi/n)^{n/2}\left(\sum_{i=1}^n (x_i-\mu_0)^2\right)^{n/2}}$$

Testgröße: Damit ergibt sich die Testgröße $T(x) = \frac{L(x;\mu_0,\tilde{\sigma}^2)}{L(x;\hat{\mu},\hat{\sigma}^2)}$ als:

$$T(x_1,\ldots,x_n) = \left(\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \mu_0)^2}\right)^{n/2}$$

Formulierung des Tests: Es gilt:

$$T(x) < c \Leftrightarrow \frac{\sqrt{n}|\bar{x} - \mu_0|}{\left((n-1)s^2\right)^{1/2}} > c'$$

wobei s^2 die Stichprobenvarianz der konkreten Stichprobe x ist.
Dies können wir nun schreiben als:

$$\frac{\sqrt{n}|\bar{x} - \mu_0|}{s} > c'(n-1)^{1/2} =: c''$$

Verteilung dieser Testgröße: Unter H_0 ist die Teststatistik $\frac{\sqrt{n}(\bar{X}-\mu_0)}{S}$ mit der Zufallsstichprobe X und der Stichprobenvarianz S^2 t-verteilt mit $(n-1)$ Freiheitsgraden.

$$\frac{\sqrt{n}(\bar{X}-\mu_0)}{S} \sim t_{n-1}$$

Bei gegebenem α wählen wir das passende Quantil $c'' = t_{n-1,\alpha/2}$ als Schwellenwert für unseren Test.

2.2 Union-Intersection-Tests und Intersection-Union-Tests

Manchmal ist es sinnvoll, Tests für komplizierte Nullhypothesen aus Tests einfacherer Hypothesen zu konstruieren. Wir betrachten nun zwei (miteinander verwandte) Methoden, die dies leisten: *Vereinigungs- und Schnittmengen-Tests.*

2.2.1 Union-Intersection-Tests

Angenommen H_0 lässt sich als Durchschnitt darstellen:

$$H_0 : \theta \in \bigcap_{\gamma \in \Gamma} \Theta_\gamma$$

Wobei Γ eine beliebige endliche oder unendliche Indexmenge ist.

Die kritischen Bereiche: Seien nun für jede Teilhypothese

$$H_{0\gamma} : \theta \in \Theta_\gamma \text{ gegen } H_{1\gamma} : \theta \in \Theta_\gamma^c$$

entsprechende Tests verfügbar. Dann ist der Ablehnungsbereich für $H_{0\gamma}$

$$K_\gamma = \{x : T_\gamma(x) \in K'_\gamma\}.$$

Der kritische Bereich K für den Union-Intersection-Test ist dann:

$$K = \bigcup_{\gamma \in \Gamma} K_\gamma = \bigcup_{\gamma \in \Gamma} \{x : T_\gamma(x) \in K'_\gamma\}$$

Einfache Konsequenz: Falls eine der Hypothesen $H_{0\gamma}$ verworfen wird, dann muss auch H_0 verworfen werden. Denn nur wenn jede Hypothese $H_{0\gamma}$ nicht verworfen wird, wird auch H_0 nicht verworfen.

Zur Konstruktion des kritischen Bereichs: Üblicherweise hat der Ablehnungsbereich zu $\gamma \in \Gamma$ die Form $\{x : T_\gamma(x) > c\}$, wobei c nicht von γ abhängt. Somit ergibt sich:

$$\bigcup_{\gamma \in \Gamma} \{x : T_\gamma(x) \in K'_\gamma\} = \bigcup_{\gamma \in \Gamma} \{x : T_\gamma(x) > c\} = \{x : \sup_{\gamma \in \Gamma} T_\gamma(x) > c\}$$

Wir erhalten als Teststatisik, H_0 zu testen: $T(X) = \sup_{\gamma \in \Gamma} T_\gamma(X)$.

2.2.2 Beispiel: Union-Intersection Test bei Normalverteilung

Sei $(X_1, \ldots, X_n)$ eine Zufallsstichprobe mit $X_i \sim N(\mu, \sigma^2)$ $\forall i$. Wir wollen nun die Hypothese $H_0 : \mu = \mu_0$ gegen $H_1 : \mu \neq \mu_0$ testen. Dafür schreiben wir H_0 als Durchschnitt:

$$H_0 : \{\mu : \mu \leq \mu_0\} \cap \{\mu : \mu \geq \mu_0\}$$

Also ist hier $\Gamma = \{L, U\}$ und:

$H_{0L} : \mu \leq \mu_0$ gegen $H_{1L} : \mu > \mu_0$ wird verworfen, falls

$$\frac{\bar{X} - \mu_0}{S/\sqrt{n}} \geq t_L$$

$H_{0U} : \mu \geq \mu_0$ gegen $H_{1L} : \mu < \mu_0$ wird verworfen, falls

$$\frac{\bar{X} - \mu_0}{S/\sqrt{n}} \leq t_U$$

Für $t_L = -t_U \geq 0$ ist dies der zweiseitige t-Test:

$$\frac{|\bar{X} - \mu_0|}{S/\sqrt{n}} \geq t_L \Leftrightarrow H_0 \text{ wird verworfen}$$

Mit $t_L = t_{n-1,\alpha/2}$ erhält man einen Test mit α als Wahrscheinlichkeit für den Fehler 1. Art.

2.2.3 Intersection-Union-Tests

Nun nehmen wir an, dass sich die Nullhypothese durch eine Vereinigungsmenge ausdrücken lässt.

$$H_0 : \theta \in \bigcup_{\gamma \in \Gamma} \Theta_\gamma$$

Falls für jedes $\gamma \in \Gamma$, der Ablehnungsbereich für $H_{0\gamma}$ durch $\{x : T_\gamma(x) \in K'_\gamma\}$ gegeben ist, dann ist der Ablehnungsbereich für den Intersection-Union-Test H_0 gegen H_1 gegeben durch:

$$K = \bigcap_{\gamma \in \Gamma} \{x : T_\gamma(x) \in K'_\gamma\}$$

Einfache Konsequenz: H_0 wird genau dann verworfen, wenn jede einzelne der Hypothesen $H_{0\gamma}$ verworfen wird.

Zur Konstruktion des kritischen Bereichs: Der Ablehnungsbereich lässt sich einfacher formulieren, falls die Ablehnungsbereiche der Teilhypothesen alle von der Form $\{x : T_\gamma(x) \geq c\}$ sind. Auch hier sei c unabhängig von γ.

$$\bigcap_{\gamma \in \Gamma} \{x : T_\gamma(x) \in K'_\gamma\} = \bigcap_{\gamma \in \Gamma} \{x : T_\gamma(x) \geq c\} = \{x : \inf_{\gamma \in \Gamma} T_\gamma(x) \geq c\}$$

So ist

$$T(X) = \inf_{\gamma \in \Gamma} T_\gamma(X)$$

die Teststatistik für den Test H_0 gegen H_1.

2.2.4 Anwendung am Beispiel eines Sitzpolster-Herstellers

Ein Hersteller produziert Sitzpolster. Diese sollen auf ihre Qualität getestet werden.
Qualitätsmerkmale:

- Belastbarkeit (in Kilogramm)
- Entflammbarkeit (Wahrscheinlichkeit)

Qualitätsansprüche:

- Belastbarkeit größer als 50 Kilogramm
- Wahrscheinlichkeit, einen Entflammbarkeitstest zu bestehen, über 0.95

Statistische Modellierung: Wir nehmen an, die Belastbarkeit eines Polsters sei die Realisierung einer Zufallsvariablen $X_i \sim N(\theta_1, \sigma^2)$ für $i = 1, \ldots, n$. Der Entflammbarkeitstest sei ein Bernoulli-Experiment $Y_i \sim Bernoulli(\theta_2)$ für $i = 1, \ldots, m$. Unsere Hypothesen gemäß dem Intersection-Union-Ansatz sehen wie folgt aus:

$$\underbrace{H_0 : \{\theta_1 \leq 50\} \cup \{\theta_2 \leq 0.95\}}_{Negationsprinzip} \text{ gegen } H_1 : \{\theta_1 > 50\} \cap \{\theta_2 > 0.95\}$$

Ein Sitzpolster wird also genau dann als akzeptabel angesehen, falls H_0 verworfen wird.

Die Ablehnungsbereiche: Für die beiden Teilhypothesen gilt:

$$H_{01}: \theta_1 \leq 50 \text{ wird verworfen, falls } \tfrac{\bar{X}-50}{S/\sqrt{n}} > t$$
$$H_{02}: \theta_2 \leq 0.95 \text{ wird verworfen, falls } \textstyle\sum_{i=1}^{m} Y_i > b$$

Somit ist der Ablehnungsbereich für den Intersection-Union-Test wie folgt gegeben:

$$\left\{(x, y): \frac{\bar{x}-50}{s/\sqrt{n}} > t \text{ und } \sum_{i=1}^{m} y_i > b\right\}$$

Auf diese Weise lassen sich umfassende Tests zur Produktqualität, auch mit mehreren Parametern, konstruieren.

3 Invariante Tests

Wir werden nun noch Tests kennen lernen, die bezüglich einer Gruppe von Transformationen im Stichprobenraum invariant sind, d.h. immer noch die selben Ergebnisse liefern, die genauso gedeutet werden können.
Hierfür definieren wir zunächst eine *Gruppe von Transformationen* und betrachten dann das Testen von statistischen Hypothesen aus entscheidungstheoretischer Sicht.

3.1 Transformationen im Stichprobenraum

3.1.1 Definition

Eine Menge G von Transformation heißt *Gruppe* von Transformationen auf dem Raum Ω , falls

(i) Existenz der Identität: $\mathrm{id} \in G : \mathrm{id}(x) = x$

(ii) Existenz der Inversen: $g \in G \Rightarrow \exists\ g^* \in G$ mit $g^*(g(x)) = x$

(iii) Abgeschlossenheit bzgl. der Komposition:
$g_1 \in G, g_2 \in G \Rightarrow g_1 \circ g_2 \in G$

3.1.2 Beispiel: Skalierung und Verschiebung

Sei $\Omega = \mathbb{R}$ und $G = \{g_{a,b} : g_{a,b}(x) = ax + b \text{ mit } a, b \in \mathbb{R} \text{ und } a \neq 0\}$ für $x \in \mathbb{R}$.

- Die Identität ist: $g_{1,0}$
- Das Inverse zu $g_{a,b}$ ist: $g_{\frac{1}{a},-\frac{b}{a}}$
- Komposition: $g_{a_2,b_2} \circ g_{a_1,b_1} = g_{a_1 a_2, a_2 b_1 + b_2} \in G$

Damit ist G eine Gruppe von Transformationen auf Ω.

3.1.3 Invariante Verteilungsfamilien

Sei G eine Gruppe von Transformationen auf Ω und
$\mathcal{P} = \{P_\theta \mid \theta \in \Theta\}$ sei eine Familie von Verteilungen auf Ω.
Dann heißt $\mathcal{P}$ *invariant*, falls es für alle $\theta \in \Theta$ und für alle $g \in G$, ein θ' gibt, so dass:

$$X \sim P_\theta \Rightarrow g(X) \sim P_{\theta'}$$

Wir schreiben $\theta' = g'(\theta)$, und $G' = \{g'\}$ ist eine Gruppe von Transformationen auf auf Θ.

3.1.4 Beispiel: Familie von Normalverteilungen

Seien $X_1, \ldots, X_n \sim N(\theta, 1)$ iid. Und es sei $G = \{g_{1,b}\}$ wie oben definiert. Dann hat $g_{1,b}(X) = (X_1 + b, \ldots, X_n + b)$ iid.
$N(\theta + b, 1)$-verteilte Komponenten.
Somit existiert $\theta' = g'_{1,b}(\theta) = \theta + b$ und die betrachtete Familie von Normalverteilungen ist invariant bzgl. Verschiebung.

3.2 Entscheidungstheorie

Die Entscheidungstheorie legt einen weiteren Rahmen, um das Testen von Hypothesen, ebenso wie um das Schätzen von Parametern.

3.2.1 Das Entscheidungsproblem

Die Idee ist einfach: Zu unseren Räumen Ω und Θ, haben wir nun noch den Raum der möglichen Aktionen $\mathcal{A}$. D.h. entsprechend einer Stichprobe $x \in \Omega$, wählen wir eine Aktion $a \in \mathcal{A}$.

3.2.2 Die Verlust-Funktion

Eine Abbildung

$$L : \Theta \times \mathcal{A} \to \mathbb{R}$$

mit

$$L(\theta, a) = \text{ Kosten der Aktion } a$$

heißt Verlust-Funktion (*Loss*).

3.2.3 Die Entscheidungsregel

Eine Abbildung

$$d : \Omega \to \mathcal{A}$$

mit

$$d(x) = \text{ Aktion, falls } x \text{ beobachtet wurde}$$

heißt Entscheidungsregel.

3.2.4 Das Risiko

Als Maß für die Güte einer Entscheidungsregel definieren wir noch das Risiko als den erwarteten Verlust eines Entscheidungsproblems:

$$R(\theta, d) = \mathbb{E}_\theta(L(\theta, d(X)))$$

3.3 Invariante Entscheidungsregeln

3.3.1 Das invariante Entscheidungsproblem

Gegeben sei eine invariante Familie von Verteilungen. Dann ist ein *Entscheidungsproblem invariant* bzgl. einer Gruppe G, falls für jedes $g \in G$ (bzw. $g' \in G'$) und jedes $a \in \mathcal{A}$ genau ein $a^* \in \mathcal{A}$ existiert, so daß $L(\theta, a) = L(g'(\theta), a^*)$ für alle $\theta \in \Theta$.

Wir schreiben $a^* = g^*(a)$. Somit ist eine Gruppe von Transformationen auf $\mathcal{A}$ gegeben.

3.3.2 Definition: invariante Entscheidungsregel

Eine Entscheidungsregel φ heißt *invariant* bzgl. G, falls für jedes $g \in G$ und $x \in \Omega$ gilt:

$$\varphi(g(x)) = g^*(d(x))$$

Idee: Invarianz als Kriterium für die Auswahl einer Entscheidungsregel. $\Rightarrow$ Möglichkeit der Konstruktion einer Entscheidungsregel, die für eine ganze *Klasse von Daten* optimal ist.

3.4 Invariantes Testen von Hypothesen

Wir testen $H_0 : \theta \in \Theta_0$ gegen $H_1 : \theta \in \Theta_1$. Für einen invarianten Test benötigen wir:

- invariantes Entscheidungsproblem
  - invariante Verteilungsfamilie
  - invarianter Verlust
  - invariante Parameterräume Θ_0 und Θ_1
- invariante Entscheidungsregel

3.4.1 Die Verlustfunktion im Testfall

$$\begin{array}{lll} L(a_0,\theta)=0 & L(a_1,\theta)=c_1 & \text{falls } \theta\in\Theta_0 \\ L(a_0,\theta)=c_0 & L(a_1,\theta)=0 & \text{falls } \theta\in\Theta_1 \end{array}$$

Wobei a_i bedeutet, dass die Hypothese H_i nicht verworfen wird.

Bemerkung: Für den Test φ ist das Risiko dann gerade:

$$R(\theta,\varphi)=\begin{cases} c_1\alpha(\theta) & \text{falls } \theta\in\Theta_0 \\ c_0\beta(\theta) & \text{falls } \theta\in\Theta_1 \end{cases}$$

3.4.2 Invarianz der Parameterräume, des Verlusts und des Tests

Es wird gefordert, daß:

$$\theta \in \Theta_i \Rightarrow g'(\theta) \in \Theta_i \quad i = 0, 1$$

Der Verlust L ist hier zwangsläufig invariant, da
$L(\theta, a_i) = L(g'(\theta), a^*)$ mit $a^* = \mathrm{id}(a_i) = a_i \in \mathcal{A} = \{a_1, a_2\}$. Damit ist $g^* = \mathrm{id}$ und der Test φ ist dann invariant, falls

$$\varphi(g(x)) = \varphi(x) \quad \forall\, x \in \Omega,\ g \in G$$

3.4.3 Beispiel: ein invarianter t-Test

Gegeben seien $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ iid. Wir testen

$$H_0 : \mu \leq 0 =: \mu_0 \text{ gegen } H_1 : \mu > 0.$$

o.B.d.A. sei

$$\mu_0 = 0.$$

Invarianz der Normalverteilungsfamilie: Die Normalverteilungsfamilie ist invariant bzgl. Verschiebung und Skalierung, d.h. wir wählen wieder

$$G = \{g_{a,b} : g_{a,b}(x) = ax + b\}.$$

Dann ist $Z_i = aX_i + b \sim N(a\mu + b, a^2\sigma^2)$. Wir definieren

$$Z = g_{a,b}(X) = aX + b,$$

somit auch

$$g_{a,b}(\bar{X}_n, S^2) = (a\bar{X}_n + b, a^2 S^2) \text{ und } g'_{a,b}(\mu, \sigma^2) = (a\mu + b, a^2\sigma^2).$$

Invarianz des Parameterraumes Für die Invarianz des Parameterraumes benötigen wir $\mu \leq 0 \Rightarrow a\mu + b \leq 0$. Dies kann nur dann der Fall sein, wenn $a > 0$ und $b = 0$. Somit kann unser Testproblem nur bzgl. positiver Skalierung der Daten invariant sein.

Invariante Tests Mit $b = 0$ gilt $g_a(\bar{X}_n, S^2) = (a\bar{X}_n, a^2 S^2)$. Also fordern wir für einen invarianten Test, der auf diesen suffizienten Schätzern basiert:

$$\varphi(\bar{X}_n, S^2) = \varphi(a\bar{X}_n, a^2 S^2) \quad \forall\ a > 0$$

Insbesondere erhalten wir für $a = 1/S$: $\varphi(\bar{X}_n, S^2) = \varphi(\bar{X}_n/S, 1)$. Jeder invariante Test ist also eine Funktion nur von $\bar{X}_n/S$.

Teststatistik des invarianten Tests Da $\bar{X}_n \sim N(\mu, \sigma^2/n)$ und $(n-1)S^2 \sim \sigma^2\chi^2_{n-1}$, und $\bar{X}_n$ und S^2 unabhängig sind, verwerfen wir $H_0 : \mu \leq 0$, falls:

$$T(X) = \sqrt{n}\bar{X}_n/S$$

groß genug ist. Für $\mu = 0$ ist $T \sim t_{n-1}$, wähle dann $t_{n-1,\alpha}$ als Schwelle.

Fazit: Wir erhalten hiermit einen Test, der für eine ganze Klasse von Datensätzen - nämlich die, die sich nur durch einen positiven Skalierungsfaktor unterscheiden - anwendbar ist.

Literatur

[1] Elizabeth Thompson (1998): *Lecture Notes: Inference, Decision Theory*. University of Washington, USA.

[2] George Casella, Roger L. Berger (2002): *Statistical Inference.* Wadsworth Group, Duxbury, CA, USA.

[3] Volker Schmidt (2003): *Vorlesungsskript: Statistik I*. Universität Ulm.