Karlsruher Institut für Technologie

Institut für Volkswirtschaftslehre (ECON)

Lehrstuhl für Ökonometrie und Statistik

Kapitel XII - Gleichmäßig beste unverfälschte Tests und Tests zur Normalverteilung

Induktive Statistik

Prof. Dr. W.-D. Heller
Hartwig Senska
Carlo Siebenschuh

Aus Kapitel XI:

1. Bei einem Test zum Niveau α ist bei stetiger Gütefunktion das Supremum der Fehlerwahrscheinlichkeit 2.Art

$$1 - \alpha$$

2. Es gibt Testprobleme, für die es zu $0 < \alpha < 1$ keinen gleichmäßig besten Test zum Niveau α gibt.

Bei einem Testproblem, zu dem es keinen gleichmäßig besten Test zu gegebenem Niveau α mit $0 < \alpha < 1$ gibt, ist es eine sinnvolle Forderung, nur solche Tests zuzulassen, deren Fehlerwahrscheinlichkeit 2.Art nirgendwo größer als $1 - \alpha$ ist.

Definition:

Ein Test $\delta : \chi \rightarrow \{d_0, d_1\}$ zum Niveau α für $H_0 : \gamma \in \Gamma_0$ gegen $H_1 : \gamma \in \Gamma_1$ heißt **unverfälscht**, wenn

$$P_{II}(\delta, \gamma) \leq 1 - \alpha \ \text{ für alle } \gamma \in \Gamma_1$$

gilt.

Bemerkung:

Dies bedeutet, dass die Wahrscheinlichkeit, die Nullhypothese abzulehnen, wenn sie falsch ist, mindestens so groß ist, wie die Wahrscheinlichkeit, die Nullhypothese abzulehnen, wenn sie richtig ist:

$$1 - P_{II}(\delta, \gamma) \geq \alpha.$$

Gleichmäßig beste Tests existieren häufig nicht

a) bei zweiseitigen Tests (s. Beispiel aus Kapitel XI)
b) bei Verteilungen mit mehreren Parametern, wenn sich das Testproblem auf einen Parameter bezieht und der/die andere(n) unbekannt ist/sind.

Beispiel zu b):

Test zum Mittelwert der Normalverteilung bei unbekannter Varianz,
Parameterraum der Normalverteilung

$$\Gamma = \{(\mu, \sigma^2) | \ \mu \in \mathbb{R}, \ \sigma^2 \ \in \ \mathbb{R}_{++}\}$$
$$\mathbb{R}_{++} \text{ Menge der positiven reellen Zahlen}$$

Das Testproblem

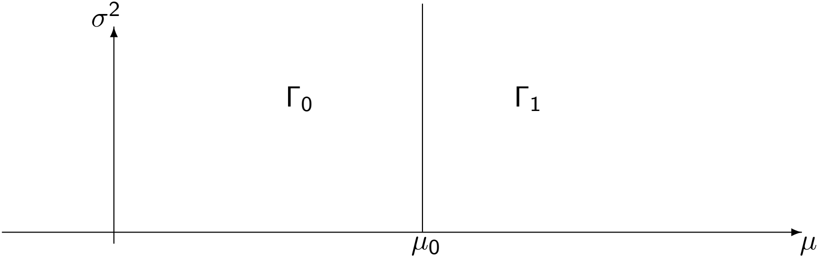
$$H_0: \ \mu \leq \mu_0 \qquad \text{gegen} \qquad H_1: \ \mu > \mu_0$$

entspricht der Aufteilung des Parameterraums in

$$\Gamma_0 = \{(\mu, \sigma^2) | \ \mu \in \mathbb{R} \text{ mit } \mu \leq \mu_0, \ \sigma^2 \in \mathbb{R}_{++}\}$$

und

$$\Gamma_1 = \{(\mu, \sigma^2) | \ \mu \in \mathbb{R} \text{ mit } \mu > \mu_0, \ \sigma^2 \in \mathbb{R}_{++}\}$$



Aufteilung des Parameterraums bei $H_0: \ \mu \leq \mu_0$ gegen $H_1: \ \mu > \mu_0$

Bei bekanntem σ^2 ist

$$\delta(x) = \begin{cases} d_1 & \overline{x} > \mu_0 + \frac{\sigma}{\sqrt{n}} \cdot u_{1-\alpha} \\ d_0 & \overline{x} \leq \mu_0 + \frac{\sigma}{\sqrt{n}} \cdot u_{1-\alpha} \end{cases}$$

gleichmäßig bester Test zum Niveau α ($u_{1-\alpha}$ ist $(1-\alpha)$-Quantil der Standardnormalverteilung). Die Konstruktion des Tests beruht darauf, dass bei Mittelwert μ

$$\frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad \text{standardnormalverteilt ist.}$$

Bei unbekanntem σ^2 kann σ durch einen Schätzwert ersetzt werden.
Es gilt

$$\frac{\overline{X} - \mu}{\frac{S^*(X)}{\sqrt{n}}} \quad \text{ist } t(n-1)\text{-verteilt ("}t\text{-verteilt" mit } n-1 \text{ Freiheitsgraden)}$$

mit $S^*(X) = \sqrt{\frac{1}{n-1} \sum\limits_{i=1}^{n} (x_i - \overline{x})^2}$ korrigierte Stichprobenstandardabweichung

(**Achtung:** $S^*(X)$ ist nicht erwartungstreu für σ, obwohl $S^{*2}(X)$ erwartungstreu für σ^2 ist)

Damit ersetzen wir σ in δ durch $S^*(X)$ und $u_{1-\alpha}$ durch das $(1-\alpha)$-Quantil der $t(n-1)$-Verteilung $t(n-1)_{1-\alpha}$.

Der resultierende Test

$$\delta(x) = \begin{cases} d_1 & \overline{x} > \mu_0 + \frac{S*(x)}{\sqrt{n}} t(n-1)_{1-\alpha} \\ d_0 & \overline{x} \leq \mu_0 + \frac{S*(x)}{\sqrt{n}} t(n-1)_{1-\alpha} \end{cases}$$

ist <u>kein</u> gleichmäßig bester Test zum Niveau α, aber ein gleichmäßig bester unverfälschter Test zu Niveau α.

Analog ist für

$$H_0 : \mu \geq \mu_0 \qquad \text{gegen} \qquad H_1 : \mu < \mu_0$$

bei unbekannter Varianz

$$\delta(x) = \begin{cases} d_1 & \overline{x} < \mu_0 - \frac{s*(x)}{\sqrt{n}} t(n-1)_{1-\alpha} \\ d_0 & \overline{x} \geq \mu_0 - \frac{s*(x)}{\sqrt{n}} t(n-1)_{1-\alpha} \end{cases}$$

gleichmäßig bester unverfälschter Test zum Niveau α, aber <u>kein</u> gleichmäßig bester Test zum Niveau α.

Beispiel zu a) Zweiseitige Tests zum Mittelwert einer Normalverteilung:

Testsituation:

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

Bemerkung:

Die umgekehrte Testsituation

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

sollte nicht angesetzt werden, da dann bei stetiger Gütefunktion die Fehlerwahrscheinlichkeit 2. Art (Γ_1 hat nur das Element μ_0) bei einem Test zum Niveau α mindestens $1 - \alpha$ ist.

Naheliegend ist, die Nullhypothese abzulehnen wenn der Stichprobenmittelwert stark von μ_0 abweicht. Dies führt zu

$$\delta(x) = \begin{cases} d_1 & |\overline{x} - \mu_0| > c \\ d_0 & |\overline{x} - \mu_0| \leq c \end{cases}$$

Gesucht ist c zu vorgegebenem Niveau α. Bei Gültigkeit von H_0 ist $\mu = \mu_0$.
Gesucht ist also c mit

$$P_{\mu_0}(|\overline{X} - \mu_0| > c) = \alpha$$

i) σ^2 bekannt:

Bei Mittelwert μ_0 ist

$$\frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \qquad \text{standardnormalverteilt}(N(0,1)).$$

Mit dem $\frac{\alpha}{2}$ - Quantil $u_{\frac{\alpha}{2}}$ und dem $(1 - \frac{\alpha}{2})$ - Quantil $u_{1-\frac{\alpha}{2}}$ gilt damit

$$P_{\mu_0}\left(u_{\frac{\alpha}{2}} \leq \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \leq u_{1-\frac{\alpha}{2}} \right) = 1 - \alpha$$

Wegen $u_{\frac{\alpha}{2}} = -u_{1-\frac{\alpha}{2}}$ ist dann

$$P\left(|\overline{X} - \mu_0| > \underbrace{\frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}}}_{c}\right) = \alpha$$

Somit ist

$$\delta(x) = \begin{cases} d_1 & |\overline{x} - \mu_0| > \frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}} \\ \\ d_0 & |\overline{x} - \mu_0| \leq \frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}} \end{cases}$$

gleichmäßig bester unverfälschter Test zum Niveau α.

ii) σ unbekannt:
Bei Mittelwert μ_0 ist

$$\frac{\overline{X} - \mu_0}{\frac{S^*(X)}{\sqrt{n}}} \qquad t(n-1) - \text{verteilt.}$$

Mit dem $\frac{\alpha}{2}$ - Quantil $t(n-1)_{\frac{\alpha}{2}}$ und dem $(1-\frac{\alpha}{2})$ - Quantil $t(n-1)_{1-\frac{\alpha}{2}}$ gilt damit

$$P_{\mu_0}\left(t(n-1)_{\frac{\alpha}{2}} \leq \frac{\overline{X} - \mu_0}{\frac{S^*(X)}{\sqrt{n}}} \leq t(n-1)_{1-\frac{\alpha}{2}} \right) = 1 - \alpha$$

Wegen $t(n-1)_{\frac{\alpha}{2}} = -t(n-1)_{1-\frac{\alpha}{2}}$ ist dann

$$P\left(|\overline{X}-\mu_0| > \underbrace{\frac{S*(X)}{\sqrt{n}}t(n-1)_{1-\frac{\alpha}{2}}}_{c}\right) = \alpha.$$

Somit ist

$$\delta(x) = \begin{cases} d_1 & |\overline{x}-\mu_0| > \frac{S^*(X)}{\sqrt{n}}t(n-1)_{1-\frac{\alpha}{2}} \\ \\ d_0 & |\overline{x}-\mu_0| \leq \frac{S^*(X)}{\sqrt{n}}t(n-1)_{1-\frac{\alpha}{2}} \end{cases}$$

ist gleichmäßig bester unverfälschter Test zum Niveau α.

Einseitige Tests zur Varianz der Normalverteilung

i) Bei bekanntem Mittelwert μ
Wie gesehen ist für das Testproblem

$$H_0 : \sigma^2 \leq \sigma_0^2 \qquad \text{gegen} \qquad H_1 : \sigma^2 > \sigma_0^2$$

$$\delta(x) = \begin{cases} d_1 & \sum_{i=0}^{n} (x_i - \mu)^2 > \chi^2(n)_{1-\alpha} \cdot \sigma_0^2 \\ \\ d_0 & \sum_{i=0}^{n} (x_i - \mu)^2 \leq \chi^2(n)_{1-\alpha} \cdot \sigma_0^2 \end{cases}$$

gleichmäßig bester Test zum Niveau α.

Ebenso ist für

$$H_0 : \sigma^2 \geq \sigma_0^2 \qquad \text{gegen} \qquad H_1 : \sigma^2 < \sigma_0^2$$

$$\delta(x) = \begin{cases} d_1 & \sum_{i=0}^{n} (x_i - \mu)^2 < \chi^2(n)_\alpha \cdot \sigma_0^2 \\ & \\ d_0 & \sum_{i=0}^{n} (x_i - \mu)^2 \geq \chi^2(n)_\alpha \cdot \sigma_0^2 \end{cases}$$

gleichmäßig bester Test zum Niveau α.

ii) Bei unbekanntem Mittelwert
Ersetzen wir den unbekannten Mittelwert μ durch den Schätzwert $\bar{x}$, so ist bei Varianz σ_0^2

$$\frac{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}{\frac{1}{n-1}\sigma_0^2} = \frac{S^{*^2}(X)}{\frac{\sigma_0^2}{n-1}} = \sum_{i=1}^{n}\frac{(X_i - \overline{X})^2}{\sigma_0^2}$$

$$\chi^2(n-1) - \text{verteilt}$$

Mit dem α-Quantil $\chi^2(n-1)_\alpha$ der $\chi^2(n-1)$-Verteilung gilt damit für σ_0^2

$$P_{\sigma_0^2}\left(\frac{S^{*^2}(X)}{\frac{\sigma_0^2}{n-1}} \leq \chi^2(n-1)_\alpha\right) = \alpha$$

und für das Testproblem

$$H_0 : \sigma \leq \sigma_0^2 \qquad \text{gegen} \qquad H_1 : \sigma^2 > \sigma_0^2$$

ist

$$\delta(x) = \begin{cases} d_1 & {S^*}^2(x) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 > \chi^2(n-1)_{1-\alpha} \sigma_0^2 \frac{1}{n-1} \\ d_0 & {S^*}^2(x) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \leq \chi^2(n-1)_{1-\alpha} \sigma_0^2 \frac{1}{n-1} \end{cases}$$

Test zum Niveau α.
Man kann zeigen: δ ist gleichmäßig bester Test zum Niveau α.

Entsprechend ist bei σ_0^2 mit dem $(1-\alpha)$-Quantil $\chi^2(n-1)_{1-\alpha}$

$$P_{\sigma_0^2}\left(\frac{{S^*}^2(X)}{\frac{\sigma_0^2}{n-1}} \geq \chi^2(n-1)_{1-\alpha}\right) = \alpha$$

und für das Testproblem

$$H_0 : \sigma \geq \sigma_0^2 \qquad \text{gegen} \qquad H_1 : \sigma < \sigma_0^2$$

...

...

ist

$$\delta(x) = \begin{cases} d_1 & S^{*^2}(x) = \frac{1}{n-1} \sum\limits_{i=1}^{n} (x_i - \bar{x})^2 < \chi^2(n-1)_\alpha \frac{\sigma_0^2}{n-1} \\ d_0 & S^{*^2}(x) = \frac{1}{n-1} \sum\limits_{i=1}^{n} (x_i - \bar{x})^2 \geq \chi^2(n-1)_\alpha \frac{\sigma_0^2}{n-1} \end{cases}$$

ein Test zum Niveau α.

Es gilt:

δ ist gleichmäßig bester unverfälschter Test zum Niveau α.

Aber:

δ ist kein gleichmäßig bester Test zum Niveau α.

Zweiseitige Tests zur Varianz der Normalverteilung

Das Testproblem lautet

$$H_0 : \sigma = \sigma_0^2 \qquad \text{gegen} \qquad H_1 : \sigma \neq \sigma_0^2$$

(Wie bei den zweiseitigen Tests zum Mittelwert wird das Testproblem mit vertauschten Hypothesen nicht verwendet.)

Zweiseitige Tests zur Varianz der Normalverteilung

i) Bei bekanntem Mittelwert μ

Naheliegend ist es, wieder die Testgröße

$$\sum_{i=1}^{n}(x_i - \mu)^2 \quad \text{bzw.} \quad S^2(x) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2$$

zu verwenden und H_0 abzulehnen, wenn $S^2(x)$ stark von σ_0^2 abweicht. $S^2(x)$ ist ja GBE-Schätzfunktion für σ^2 bei bekanntem Mittelwert μ_0.

Gesucht sind also Testschranken c_1, c_2, so dass der Test

$$\delta(x) = \begin{cases} d_0 & c_1 \leq S^2(x) \leq c_2 \\ d_1 & S^2(x) < c_1 \text{ oder } c_2 < S^2(x) \end{cases}$$

ein Test zum Niveau α ist.

Bei σ_0^2 ist $\frac{S^2(X)}{\frac{\sigma_0^2}{n}}$ $\chi^2(n)$-verteilt

Mit den Quantilen $\chi^2(n)_{\alpha_1}$ und $\chi^2(n)_{1-\alpha_2}$ gilt also

$$P_{\sigma_0^2}\left(\frac{S^2(X)}{\frac{\sigma_0^2}{n}} \leq \chi^2(n)_{\alpha_1}\right) + P_{\sigma_0^2}\left(\chi^2(n)_{1-\alpha_2} < \frac{S^2(X)}{\frac{\sigma_0^2}{n}}\right) = \alpha_1 + \alpha_2$$

Gilt $\alpha_1 + \alpha_2 = \alpha$, so ist

$$\delta(x) = \begin{cases} d_0 & \chi^2(n)_{\alpha_1}\frac{\sigma_0^2}{n} \leq S^2(x) \leq \chi^2(n)_{1-\alpha_2}\frac{\sigma_0^2}{n} \\ d_1 & S^2(x) < \chi^2(n)_{\alpha_1}\frac{\sigma_0^2}{n} \;\text{ oder }\; \chi^2(n)_{1-\alpha_2}\frac{\sigma_0^2}{n} < S^2(x) \end{cases}$$

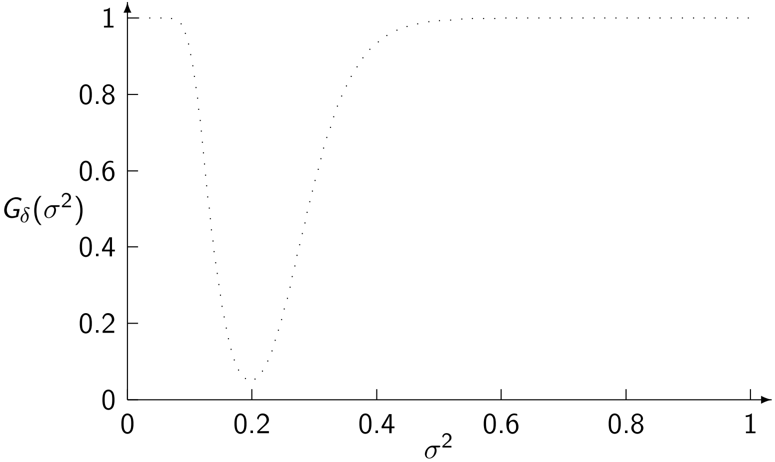
ein Test zum Niveau α.

Bemerkung:

α_1 und α_2 können so bestimmt werden, dass δ ein gleichmäßig bester unverfälschter Test zum Niveau α ist.

Gilt $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$, so heißt δ Equal-Tails-Test. Wegen der Asymmetrie der χ^2-Verteilung ist der Equal-Tails-Test <u>nicht</u> unverfälscht (s. Abb.).

Bei Anwendungen wird dennoch in der Regel der Equal-Tails-Test verwendet.



Gütefunktion des "equal-tails-test" für verschiedene Werte

Damit sind die anschließenden Überlegungen analog zu Fall i) mit unbekanntem Mittelwert.

$$\delta(x) = \begin{cases} d_0 & \chi^2(n-1)_{\frac{\alpha}{2}} \frac{\sigma_0^2}{n-1} \le {S^*}^2(x) \le \\ & \chi^2(n-1)_{1-\frac{\alpha}{2}} \frac{\sigma_0^2}{n-1} \\ d_1 & {S^*}^2(x) < \chi^2(n-1)_{\frac{\alpha}{2}} \ \text{oder} \\ & \chi^2(n-1)_{1-\frac{\alpha}{2}} \frac{\sigma_0^2}{n-1} \\ & < {S^*}^2(x) \end{cases}$$

ist Equal-Tails-Test zum Niveau α. δ ist wiederum nicht unverfälscht.

Bemerkung:

Tests zur Normalverteilung sind von besonderer Bedeutung, weil nach dem zentralen Grenzwertsatz

$$Z = \frac{\frac{1}{n}\sum_{i=1}^{n} X_i - \mu}{\sigma}\sqrt{n} \sim N(0,1)$$

für X_i u.i.v. mit $E(X_i) = \mu$ und $Var(X_i) = \sigma^2$.

$\Longrightarrow$ Für eine Testsituation zum Erwartungswert kann man für beliebige Verteilungen die Tests der Normalverteilung als Approximation heranziehen, falls nur n genügend groß ist (meist $n \geq 30$).

Dies gilt insbesondere auch dann, wenn der Verteilungstyp nicht bekannt ist.