Statistische Methoden der Datenanalyse

Wintersemester 2011/2012

Albert-Ludwigs-Universität Freiburg

Prof. Markus Schumacher und Dr. Stan Lai

Physikalisches Institut Westbau 2 OG Raum 008

Telefonnummer 07621 203 7612

E-Mail: Markus.Schumacher@physik.uni-freiburg.de

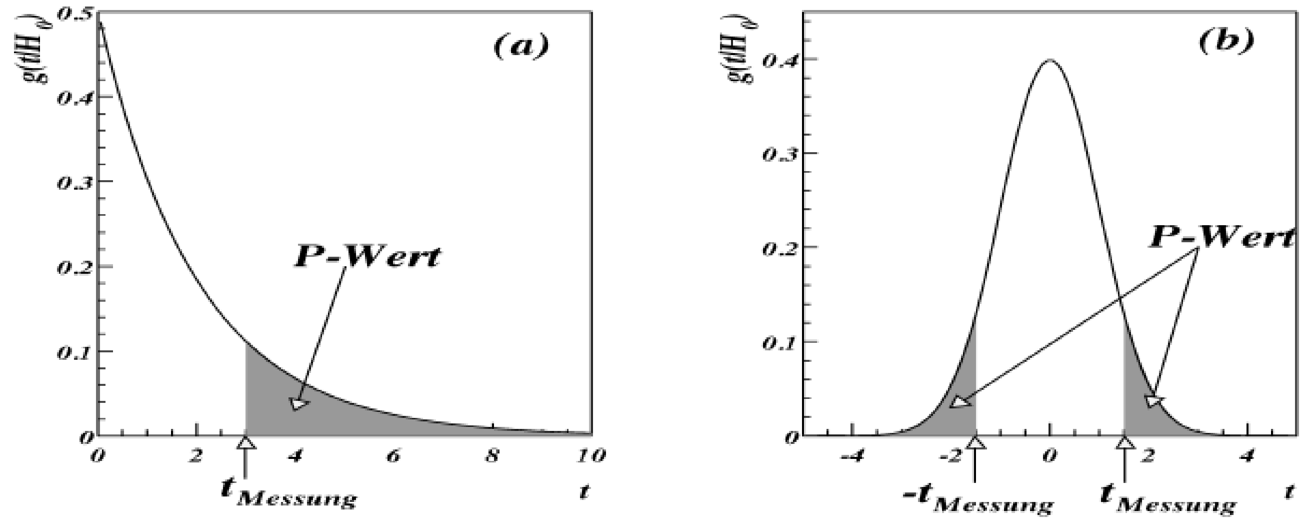
Vorlesung basiert in weiten Teilen auf der von Glen Cowan.

http://terascale.physik.uni-freiburg.de/lehre/ws_1112/statmethoden_ws1112

Kapitel 8:
Statistische Hypothesentests (Teil2)

Grundlegende Begriffe: P-Werte

P-Wert: Wahrscheinlichkeit eine Stichprobe zu beobachten, die genauso verträglich oder weniger verträglich mit der Nullhypothese wie die aktuelle Messung/Beobachtung



t=0 für perfekte Übereinstimmung zwischen Daten und H_0

links: einseitiger P-Wert   rechts: zweiseitiger P-Wert

Grundlegende Begriffe: Bemerkungen zum P-Wert

P-Wert ist Zufallsvariable (vgl.: Signifkanzniveau α vor Messung fixiert)

Wenn P-Wert = Signifkanzniveau α, dann gilt $t_{Messung} = t_{kritisch}$

P-Wert auch beobachtetes Signifkanzniveau genannt

1-P-Wert = Vetrauensniveau des Tests (“confidence level”)
P-Wert = 5% dann ausgeschlossen mit 95% Vetrauensniveau

wenn P-Wert < Signifkanzniveau α, dann Hypothese H_0 verwerfen

Achtung vor Fehlinterpretationen:
P-Wert ist nicht Wkt., dass H_0 falsch ist
1-P-Wert ist nicht Wkt, dass H_0 wahr ist

Ein einfaches Beispiel

Bewertung der Fairness eines Würfels

Stichprobe aus n Münzwürfen

n_K= Anzahl des Auftretens von "Kopf"
$n_Z=n-n_K$ = Anzahl des Auftretens von "Zahl"

H_0: Münze ist fair, d.h. $p_K = 0{,}5$ $p_Z = 0{,}5$
Signifkanzniveau auf $\alpha = 5\%$ fixiert

Teststatisik: $t = n_K$ folgt einer Binomialverteilung

Annahme: n=20 (fix) , Beobachtung $n_K=17$ $E[n_K] = 10$

zweiseitiger P-Wert = $\Sigma_0^3 f_{Bin}(n_K;20) + \Sigma_{17}^{20} f_{Bin}(n_K;20) = 0.26\%$

→ Nullhypothese "fairer Würfel" verwerfen

Eigenschaften von Hypothesentests

Gegeben Nullhypothese H_0 und Signifkanzniveau α

a) bester Test bzgl. Alternativhpothese H_1

maximale Mächtigkeit 1-ß unter H_1

b) gleichmäßig bester Test

maximale Mächtigkeit 1-ß unter allen Alternativhypothesen

c) unverzerrter Test

Mächtigkeit 1-ß > α für alle Alternativhpothesen

Illustratives Beispiel

Test der Hypothese, dass eine Gauss-WDF mit bekannter Varianz σ^2 den Mittelwert $\lambda=\lambda_0$ hat.

Stichprobe vom Umfang n (für Illustration =2): $x_1, x_2, \ldots$

Teststatistik: arithmetischer Mittelwert $\quad \mathsf{x} = \frac{1}{n}(\mathsf{x}_1 + \mathsf{x}_2 + \cdots + \mathsf{x}_n)$

folgt Gauss-Verteilung mit Mittelwert λ und Varianz σ^2/n

$$f(x;\lambda_0) = \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} \exp\left(-\frac{n}{2\sigma^2}(x-\lambda_0)^2\right)$$

Wahl von vier kritischen Regionen mit gleichem Signifikanzniveau α

$$\begin{array}{ll}
U_1 : x < \lambda^{\mathrm{I}} \text{ und } x > \lambda^{\mathrm{II}} & \text{mit } \int_{-\infty}^{\lambda^{\mathrm{I}}} f(x)\,\mathrm{d}x = \int_{\lambda^{\mathrm{II}}}^{\infty} f(x)\,\mathrm{d}x = \frac{1}{2}\alpha \ ; \\
U_2 : x > \lambda^{\mathrm{III}} & \text{mit } \int_{\lambda^{\mathrm{III}}}^{\infty} f(x)\,\mathrm{d}x = \alpha \ ; \\
U_3 : x < \lambda^{\mathrm{IV}} & \text{mit } \int_{-\infty}^{\lambda^{\mathrm{IV}}} f(x)\,\mathrm{d}x = \alpha \ ; \\
U_4 : \lambda^{\mathrm{V}} \leq x < \lambda^{\mathrm{VI}} & \text{mit } \int_{\lambda^{\mathrm{V}}}^{\lambda_0} f(x)\,\mathrm{d}x = \int_{\lambda_0}^{\lambda^{\mathrm{VI}}} f(x)\,\mathrm{d}x = \frac{1}{2}\alpha \ .
\end{array}$$

Illustratives Beispiel

Reihen: 4 verschiedene kritische Regionen

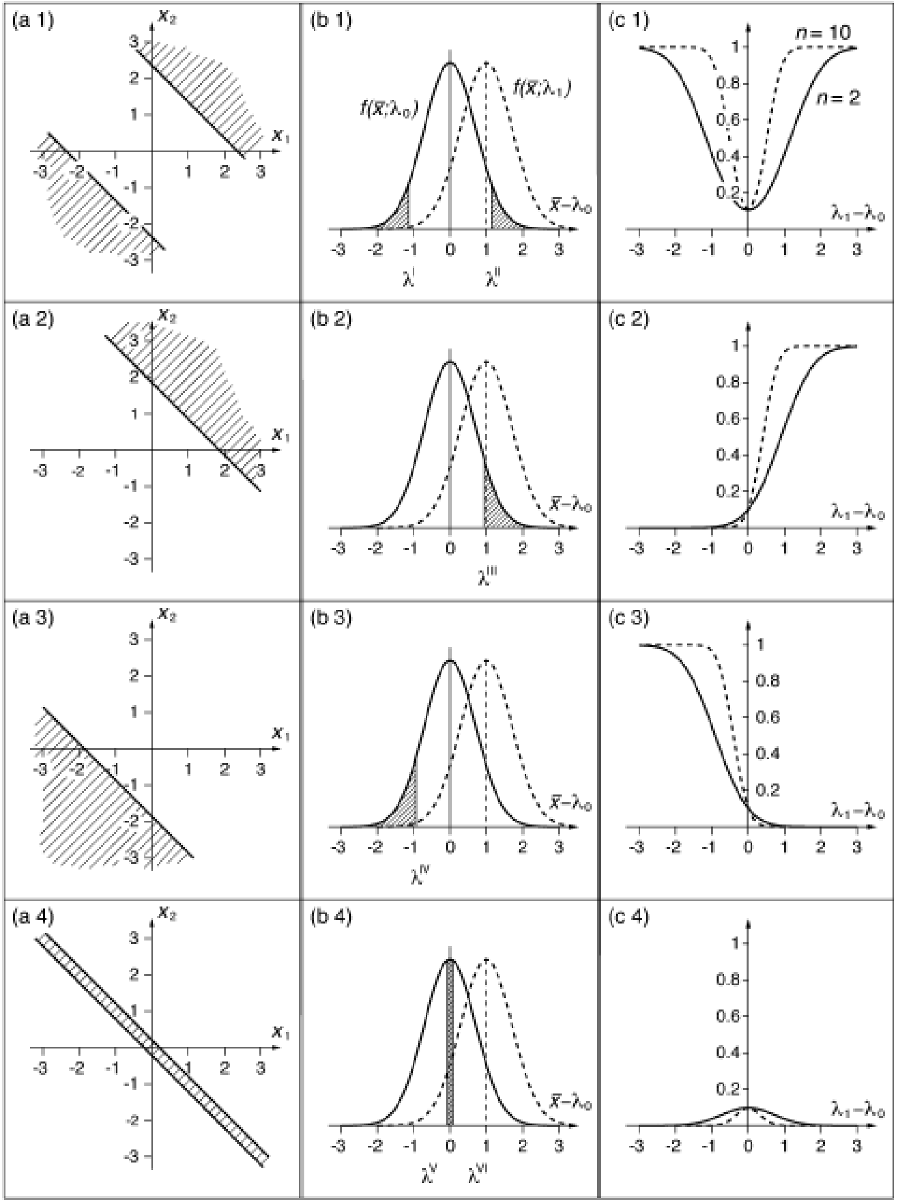
linke Spalte:
kritische Regionen für n=2 im Stichprobenraum

mittlere Spalten:
WDFs für Teststatistik für H_0 und H_1

$$\lambda = \lambda_1 = \lambda_0 + 1$$

+ kritische Regionen

rechte Spalte:
Mächtigkeit für n=2 und n=10



Illustratives Beispiel

$$U_1: x < \lambda^{\mathrm{I}} \text{ und } x > \lambda^{\mathrm{II}} \quad \text{mit } \int_{-\infty}^{\lambda^{\mathrm{I}}} f(x)\,\mathrm{d}x = \int_{\lambda^{\mathrm{II}}}^{\infty} f(x)\,\mathrm{d}x = \tfrac{1}{2}\alpha\,;$$
$$U_2: x > \lambda^{\mathrm{III}} \quad \text{mit } \int_{\lambda^{\mathrm{III}}}^{\infty} f(x)\,\mathrm{d}x = \alpha\,;$$
$$U_3: x < \lambda^{\mathrm{IV}} \quad \text{mit } \int_{-\infty}^{\lambda^{\mathrm{IV}}} f(x)\,\mathrm{d}x = \alpha\,;$$
$$U_4: \lambda^{\mathrm{V}} \le x < \lambda^{\mathrm{VI}} \quad \text{mit } \int_{\lambda^{\mathrm{V}}}^{\lambda_0} f(x)\,\mathrm{d}x = \int_{\lambda_0}^{\lambda^{\mathrm{VI}}} f(x)\,\mathrm{d}x = \tfrac{1}{2}\alpha\,.$$

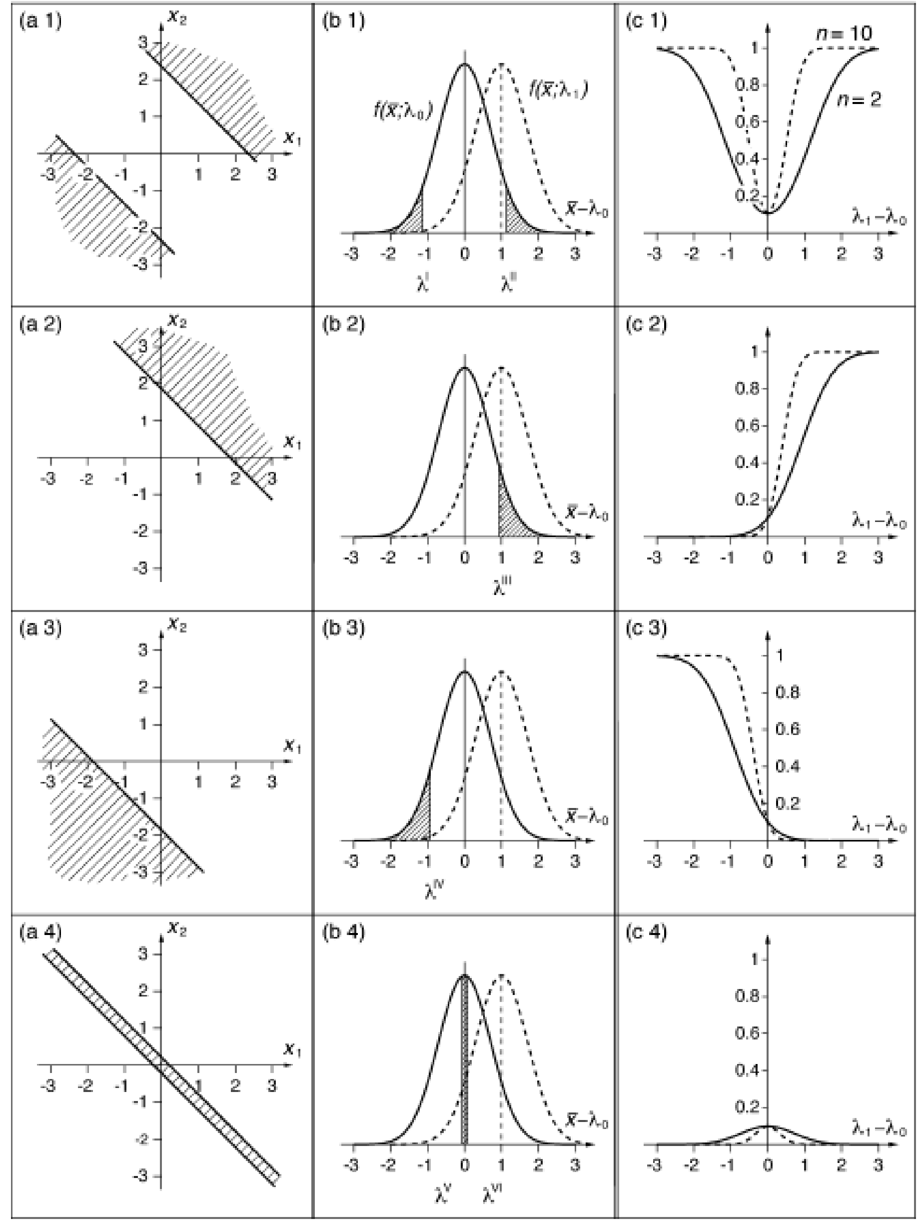
U1 ist unverzerrter Test

U2: mächtiger für $\lambda_1>\lambda_0$

U3: mächtiger für $\lambda_1<\lambda_0$

U4: kein guter Test
maximale Mächtigkeit für
$\lambda_1=\lambda_0$

keiner von U1 bis U3
ist ein gleichmäßig bester Test



Das Neyman-Person-Lemma

bester Test: für gegebenes Signifikanzniveau α maximale Mächtigkeit $1-\beta$

Fragen: welche Teststatistik t? welche Wahl der kritischen Region S_{krit}?

Antwort durch Neyman-Pearson-Lemma:
Ein Test der einfachen Nullhypothesen H_0 bzgl der einfachen Alternativhypothese H_1 ist ein bester Test, wenn die kritische Region S_{krit} im Stichprobenraum E so gewählt wird, dass gilt

$$\frac{P(\mathbf{x}|H_1)}{P(\mathbf{x}|H_0)} > c$$

(≤ c außerhalb kritischer Region)
c ist Konstante die von Signifkanzniveau abhängt.

äquivalente Aussage: die optimale Teststatistik ist $$\boxed{t(\mathbf{x}) = \frac{P(\mathbf{x}|H_1)}{P(\mathbf{x}|H_0)}}$$

Achtung: oft wird auch der Kehrwert des Likelihoodverhältnisses verwendet.
Jede monotone Funktion von t(x) ist ebenso optimal wie t selbst z.B. t/(1+t)

Beispiel 1 für Neyman-Pearson

Test der Hypothese, dass eine Gauss-WDF mit bekannter Varianz σ^2 den Mittelwert $\lambda=\lambda_0$ hat.

Stichprobe vom Umfang n: $x_1, x_{2,} \ldots x_n$

Likelihoodfunktionen bzw. gemeinsame WDF für die Stichprobe unter den Hypothesen H_0: $\lambda = \lambda_0$ und H_1: $\lambda=\lambda_1$

$$f(X|H_0) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^N \exp\left[-\frac{1}{2\sigma^2}\sum_{j=1}^{N}(\mathsf{x}^{(j)} - \lambda_0)^2\right]$$

$$f(X|H_1) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^N \exp\left[-\frac{1}{2\sigma^2}\sum_{j=1}^{N}(\mathsf{x}^{(j)} - \lambda_1)^2\right]$$

Der Likelihoodquotient (inverse von oben def. t(x)) ist dann:

$$Q = \frac{f(X|H_0)}{f(X|H_1)} = \exp\left[-\frac{1}{2\sigma^2}\left\{\sum_{j=1}^{N}(\mathsf{x}^{(j)} - \lambda_0)^2 - \sum_{j=1}^{N}(\mathsf{x}^{(j)} - \lambda_1)^2\right\}\right]$$

$$= \exp\left[-\frac{1}{2\sigma^2}\left\{N(\lambda_0^2 - \lambda_1^2) - 2(\lambda_0 - \lambda_1)\sum_{j=1}^{N}\mathsf{x}^{(j)}\right\}\right]$$

Beispiel 1 für Neyman-Pearson (2)

$$\exp\left[-\frac{N}{2\sigma^2}(\lambda_0^2-\lambda_1^2)\right]=k\geq 0$$ ist nicht negative Konstante

Die kritischen Regionen gemäß Neyman-Pearson-Lemma ergeben sich zu:

$$k\exp\left[\frac{\lambda_0-\lambda_1}{\sigma^2}\sum_{j=1}^{N}\mathbf{x}^{(j)}\right]\begin{cases}\leq c, & X\in S_c\\ \geq c, & X\notin S_c\end{cases}$$ bzw. $$(\lambda_0-\lambda_1)\mathbf{x}\begin{cases}\leq c', & X\in S_c\\ \geq c', & X\notin S_c\end{cases}$$

c und c' sind Konstanten, die abhängig vom Signifkanzniveau α sind

Arithmetischer Mittelwert ist optimale Teststatistik

Klare Vorschrift für Konstruktion der kritischen Regionen

a) $\lambda_1<\lambda_0$: linkseitiger Test b) $\lambda_1>\lambda_0$: rechtseitiger Test (siehe U2/3 in ill. Bsp.)

$$\mathbf{x}\begin{cases}\leq c'', & X\in S_c\\ \geq c'', & X\notin S_c\end{cases}$$ $$\mathbf{x}\geq c'''$$

Es gibt keinen gleichmäßig besten Test, wegen Vorzeichenwechsels von λ_1-λ_0

Beispiel 2 für Neyman-Pearson

Gegeben: Stichprobe vom Umfang n (x_i; i=1,n) aus Exponentialverteilung

$$f(t;\tau) = \frac{1}{\tau}\exp(-t/\tau)$$

Hypothese H_0: Zerfallszeit ist t1=1 Hypothese H_1: Zerfallszeit ist t2=2

Das Likelihoodverhältnis ist dann gegeben durch:

$$\frac{L(\underline{t}\,|\,\tau=2)}{L(\underline{t}\,|\,\tau=1)} = \frac{\prod_{i=1}^{n} \frac{1}{2}\exp(-t_i/2)}{\prod_{i=1}^{n} \exp(-t_i)} = \left(\tfrac{1}{2}\right)^n \exp\left(\tfrac{1}{2}\sum_{i=1}^{n} t_i\right) > k$$

Ausgedrückt durch den Mittelwert $\bar{t} = \frac{1}{n}\sum_{i=1}^{n} t_i$ kann die Ungleichung geschrieben werden als:

$$\bar{t} > 2\left(\frac{1}{n}\ln k + \ln 2\right) \equiv T_n$$

Arithmetischer Mittelwert ist optimale Teststatistik

T_n ist eine Konstante, die für jedes vorgegebene Signifkanzniveau α bestimmt werden muss und kann

Beispiel 2 für Neyman-Pearson (2)

Um den kritischen Wert für gebebnes α zu bestimmen, muss WDF für arithmetischen Mittwelert bekannt sein.

Hier nur zwei Grenzfälle n=1 und n→unendlich

a) n=1 WDF trivial $$f_1(\bar{t}) = \frac{1}{\tau} \exp(-\bar{t}/\tau)$$

für Signifkanzniveau α = 5% zu lösen: $$0.05 = \alpha = \int_{T_1}^{\infty} e^{-\bar{t}} d\bar{t} .$$

Entscheidungsgrenze ergibt sich zu: $$T_1 = -\ln \alpha \simeq 3.00$$

Die Mächtigkeit zu: $$1-\beta = \int_{T_1}^{\infty} \tfrac{1}{2}\, e^{-\bar{t}/2} d\bar{t} = \sqrt{\alpha} \simeq 0.22$$

d.h. für Mittelwert von t > 3 verwerfe Nullhypothese

Beispiel 2 für Neyman-Pearson (3)

b) n→ unendlich: WDF geht gegen Normalverteilung

$$f_n(\bar{t}) = N(\tau, \tau^2/n) = \frac{1}{\sqrt{2\pi}\ \tau/\sqrt{n}} \exp\left(-\frac{1}{2}\ \frac{(\bar{t}-\tau)^2}{\tau^2/n}\right)$$

für Signifkanzniveau α = 5% zu lösen:

$$0.05 = \alpha = \int_{T_n}^{\infty} N(1, \frac{1}{n})\, d\bar{t} = 1 - \int_{-\infty}^{T_n} N(1, \frac{1}{n})\, d\bar{t} = 1 - G\left(\frac{T_n - 1}{1/\sqrt{n}}\right)$$

Entscheidungsgrenze ergibt sich zu: $T_n = 1 + \frac{1.645}{\sqrt{n}}$

Die Mächtigkeit zu: $1-\beta = \int_{T_n}^{\infty} N(2, \frac{4}{n})\, d\bar{t} = 1 - \int_{-\infty}^{T_n} N(2, \frac{4}{n})\, d\bar{t} = 1 - G\left(\frac{T_n - 2}{2/\sqrt{n}}\right)$

Beispiel 2 für Neyman-Pearson (4)

kritischer Wert T_n und Mächtigkeit hängen vom Stichprobenumfang n ab.

Für n=100 findet man: $T_{100}=1.16$ und $1-\beta=0.99999$

Für fixes Signifikanzniveau wächst Mächtigkeit mit n stark an.

Hätten wir das Signifikanzniveau zu $\alpha = 1\%$ gewählt
dann wären die Mächtigkeiten für
$\alpha=0.01$ n=1: $1-\beta=0.10$ und n=100: $1-\beta=0.9994$

Vgl. $\alpha=0.05$ N=1: $1-\beta=0.28$ und n=100: 1-ß=0.99999

Die Mächtigkeit 1-ß des Test wird kleiner,
wenn das Signifikanzniveau α verringert wird.

Dies entspricht unserer Erwartung.

Likelihood-Quotientenmethode

Neyman-Pearson-Lemma gilt nur für einfache Hypothesen
z.B.. kein gleichmäßig bester Test für Mittelwert der Gaussverteilung I, wenn λ unter H1 unbekannt.

Betrachte folgendes Szenario:

p Parameter λ_1, .. λ_p im p-dimensionalen Ω-Parameterraum
Stichprobe vom Umfang n: x_1, … x_n
sei ω ein p-dimensionaler Unterraum von Ω, in dem die Parameter der Nullhypothese H_0 entsprechen

→ allgemeinste Alternativhypothese H_1: Parameter liegen im Komplement zu ω, also in $\Omega-\omega$

Likelihood-Quotientenmethode (2)

Die Likelihoodfunktion der Stichprobe sei $L = \prod_{i=1}^{n} f(x_i | \underline{\theta})$

Als Teststatistik wählt man den Quotienten der maximierten Likelihoods

$$\lambda \equiv \frac{L(\hat{\omega})}{L(\hat{\Omega})}$$

λ aus [0,1]

$L(\hat{\Omega})$ maximale Likelihood auf Gesamtraum Ω (H_0+H_1)

$\hat{\Omega}$ ML-Schätzer auf Gesamtraum Ω (H_0+H_1)

$L(\hat{\omega})$ maximale Likelihood auf Unterraum ω (H_0)

$\hat{\omega}$ ML-Schätzer auf Unterraum ω (H_0)

Likelihood-Quotientenmethode (3)

$$\lambda \equiv \frac{L(\hat{\omega})}{L(\hat{\Omega})}$$ λ aus [0,1]

Teilweise wird auch Kehrwert verwendet → aufpassen

Für λ nahe an 1, dann liefert $L(\hat{\omega})$ einen Wert nahe am globalen maximum $L(\hat{\Omega})$. Dann besitzt H_0 ein große Wkt. die Stichprobe erzeugt zu haben.

Für λ nahe an 0, besitzt H_0 eine kleine Wkt.die Stichprobe erzeigt zu haben.
→ λ scheint eine sinnvolle Teststatistik zu sein.

Verwerfe Nulhypothese für $0 < \lambda < \lambda_\alpha$ mit kritsichem Wert λ_α aus

$$\alpha = \int_0^{\lambda_\alpha} g(\lambda | H_0)\, d\lambda$$ g ist WDF für λ unter H_0, α ist Siginifikanzniveau

Likelihood-Quotientenmethode (4)

Problem: Bestimmung der WDFS für λ unter H_0 (H_1)

genereell durch MC-Methode, Simulation von vielen Stichproben

für n groß kann Satz von Wilks angewendet werden.

Satz von Wilks:
Sei WDF für ZV x durch stetige $f(x;\lambda_1, ..\lambda_p)$ gegeben
und durch H_0 r von p Parametern festgelegt, dann folgt
transformierte Teststatistik $-2 \ln \lambda$ einer Chi2-Verteilung
mit Anzahl der Freiheitsgrade r für Stichprobenumfang → unendlich.

Kein generelle Aussage über Optimalität von λ als Teststatistik.

Aber in der Praxis ist es meist schwierig eine bessere zu finden.

Likelihood-Quotientenmethode Beispiel

Wiederum Test der Hypothese über Mitttelwert λ einer Gauss-WDF mit bekannter Varianz σ^2

H_0: $\lambda=\lambda_0$ (einfach) H1: $\lambda\neq\lambda_0$ (zusammengesetzt)

Wert von λ unter H_1 ist fixiert: $\tilde{\lambda}^{(\omega)} = \lambda_0$
ML-Schätzer für λ unter H_1 ist arithmetischer Mittelwert $\bar{x}$.
gemeinsame WDF für Stichprobe bzw. Likelihoodfunktion:

$$f(x^{(1)},x^{(2)},\dots,x^{(N)}) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^N \exp\left[-\frac{1}{2\sigma^2}\sum_{j=1}^{N}(x^{(j)}-\lambda)^2\right]$$

Der Likelihoodquotient (Achtung: Kehrwert von früherer Definition) ist:

$$T = \frac{f(\mathrm{x}^{(1)},\mathrm{x}^{(2)},\dots,\mathrm{x}^{(N)};\bar{\mathrm{x}})}{f(\mathrm{x}^{(1)},\mathrm{x}^{(2)},\dots,\mathrm{x}^{(N)};\lambda_0)}$$

$$\begin{aligned} T &= \exp\left[\frac{1}{2\sigma^2}\left\{-\sum_{j=1}^{N}(\mathrm{x}^{(j)}-\bar{\mathrm{x}})^2+\sum_{j=1}^{N}(\mathrm{x}^{(j)}-\lambda_0)^2\right\}\right] \\ &= \exp\left[\frac{1}{2\sigma^2}\sum_{j=1}^{N}(\bar{\mathrm{x}}-\lambda_0)^2\right] = \exp\left[\frac{N}{2\sigma^2}(\bar{\mathrm{x}}-\lambda_0)^2\right]. \end{aligned}$$

Likelihood-Quotientenmethode Beispiel (2)

Wir müssen nun WDF für T und kritischen Wert (hier $T_{1-\alpha}$ genannt) bestimmen

Verwende monotone Transformation für neue Teststatistik T' $T' = 2\ln T = \frac{N}{\sigma^2}(\mathsf{x} - \lambda_0)^2$

Suche kritischen Wert von T' mit: $T' > T'_{1-\alpha}$ $\int_{T'_{1-\alpha}}^{\infty} h(T'|H_0)\,\mathrm{d}T' = \alpha$

Aufgabe: Bestimmung von WDF $h(T'|H_0)$

Ausgangspunkt: WDF für arithmetischen Mittelwert $f(x|H_0) = \sqrt{\frac{N}{2\pi\sigma^2}} \exp\left(-\frac{N}{2\sigma^2}(x - \lambda_0)^2\right)$

Variablentransformation: $\left|\frac{\mathrm{d}x}{\mathrm{d}T'}\right| = \frac{1}{2}\sqrt{\frac{\sigma^2}{N}} T'^{-1/2}$ $h(T'|H_0) = \left|\frac{\mathrm{d}x}{\mathrm{d}T'}\right| f(x|H_0) = \frac{1}{\sqrt{2\pi}} T'^{-1/2} \mathrm{e}^{-T'/2}$

Dies ist Chi-Quadrat-WDF für 1 Freiheitsgrad.
Hier gilt Satz von Wilks exakt für beliebige Stichprobenumfänge n

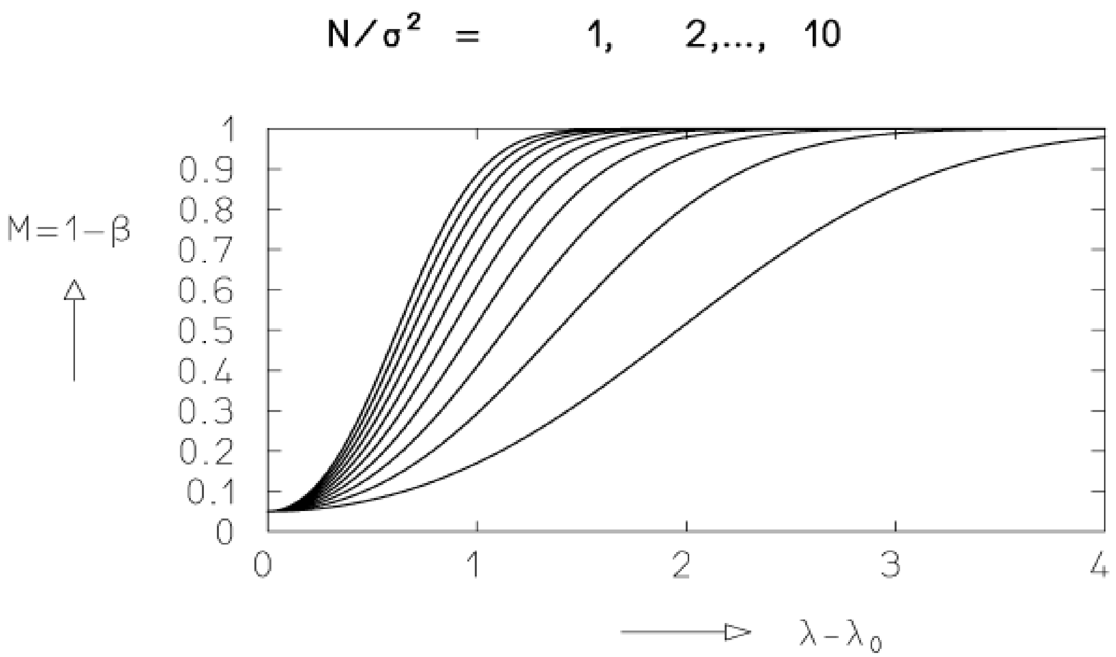
Der Likelihood-Quotiententest liefert als Entscheidungsgrenze:

$$T' = \frac{N}{\sigma^2}(x-\lambda_0)^2 > T'_{1-\alpha}$$ dies ist zweiseitiger Test aus Beispiel

Die Entscheidungsgrenze ist äquivalent zu:

$$\left(\frac{N}{\sigma^2}\right)^{1/2}|x-\lambda_0| < \lambda', \quad \left(\frac{N}{\sigma^2}\right)^{1/2}|x-\lambda_0| > \lambda'' \qquad -\lambda' = \lambda'' = (T'_{1-\alpha})^{1/2} = (\chi^2_{1-\alpha})^{1/2} = \chi_{1-\alpha}$$

Berechnung der Mächtigkeit für α=0.05 liefert



Likelihood-Quotientenmethode Beispiel (4)

Wiederum Test der Hypothese über Mittelwert λ einer Gauss-WDF aber, diesmal mit unbekannter Varianz σ^2

H_0: $\lambda=\lambda_0$ (zusammengesetzt) H1: $\lambda\neq\lambda_0$ (zusammengesetzt)

ML-Schätzer unter H_0: $\tilde{\lambda}^{(\omega)} = \lambda_0\,,\quad \tilde{\sigma}^{2(\omega)} = \frac{1}{N}\sum_{j=1}^{N}(\mathrm{x}^{(j)} - \lambda_0)^2$

ML-Schätzer unter H_1 : $\tilde{\lambda}^{(\Omega)} = \bar{\mathrm{x}}\,,\quad \tilde{\sigma}^{2(\Omega)} = \frac{1}{N}\sum_{j=1}^{N}(\mathrm{x}^{(j)} - \bar{\mathrm{x}})^2 = \mathrm{s}'^2$

Der Likelihoodquotient ist:

$$\begin{aligned} T &= \left(\frac{\sum(\mathrm{x}^{(j)}-\lambda_0)^2}{\sum(\mathrm{x}^{(j)}-\bar{\mathrm{x}})^2}\right)^{N/2} \exp\left(-\frac{N}{2}\frac{\sum(\mathrm{x}^{(j)}-\bar{\mathrm{x}})^2}{\sum(\mathrm{x}^{(j)}-\bar{\mathrm{x}})^2} + \frac{N}{2}\frac{\sum(\mathrm{x}^{(j)}-\lambda_0)^2}{\sum(\mathrm{x}^{(j)}-\lambda_0)^2}\right) \\ &= \left(\frac{\sum(\mathrm{x}^{(j)}-\lambda_0)^2}{\sum(\mathrm{x}^{(j)}-\bar{\mathrm{x}})^2}\right)^{N/2}. \end{aligned}$$

Variablentrafo (anders als in Beispiel 1):

$$\begin{aligned} T' &= T^{2/N} = \frac{\sum(\mathrm{x}^{(j)}-\lambda_0)^2}{\sum(\mathrm{x}^{(j)}-\bar{\mathrm{x}})^2} = \frac{\sum(\mathrm{x}^{(j)}-\bar{\mathrm{x}})^2 + N(\bar{\mathrm{x}}-\lambda_0)^2}{\sum(\mathrm{x}^{(j)}-\bar{\mathrm{x}})^2}, \\ T' &= 1 + \frac{t^2}{N-1}. \end{aligned}$$

$$t = \sqrt{N}\frac{\bar{\mathrm{x}}-\lambda_0}{\left(\frac{\sum(\mathrm{x}^{(j)}-\bar{\mathrm{x}})^2}{N-1}\right)^{1/2}} = \sqrt{N}\frac{\bar{\mathrm{x}}-\lambda_0}{\mathrm{s}_{\mathrm{x}}} = \frac{\bar{\mathrm{x}}-\lambda_0}{\mathrm{s}_{\bar{\mathrm{x}}}}$$

t folgt Studentscher t-Verteilung mit n-1Freiheitsgraden

Zweiseitiger Test mit kritischer Region definiert über $|t| > t_{1-\frac{1}{2}\alpha}$

Hypothesentest für ZV aus Gauss-WDF

A) Vergleich von Stichprobe $x_1, \ldots, x_n$ mit Theorie $f_G(x; \mu, \sigma^2)$

Erwartungstreue Schätzwerte für Mittelwert und Varianz

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \qquad s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

i) Test der Mittelwerthypothese $H_0: \mu=\mu_0 \quad H_1: \mu\neq\mu_0$

a) Varianz bekannt

Teststatistik $\dfrac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$ folgt Standardgauss-WDF $N(0,1)$

b) Varianz unbekannt

Teststatistik $\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}}$ folgt Studentscher t-Verteilung mit n-1 Freiheitsgraden $t(n-1)$

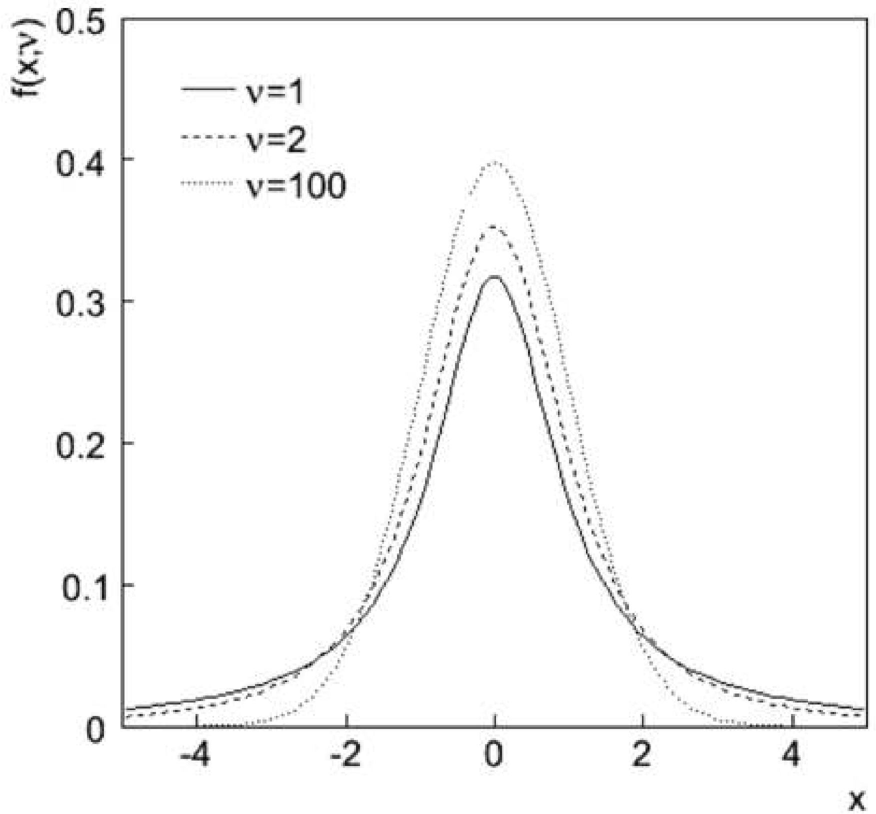
Studentsche t-Verteilung

$$f(x;\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma(\nu/2)}\left(1+\frac{x^2}{\nu}\right)^{-\left(\frac{\nu+1}{2}\right)}$$

ν = Anzahl der Freiheitsgrade (nicht notwendigerweise ganzzahlig)

$$E[x] = 0 \quad (\nu > 1)$$

$$V[x] = \frac{\nu}{\nu-2} \quad (\nu > 2)$$



ν = 1 ergibt Cauchy-WDF

$\nu \to \infty$ ergibt Gauss-WDF

Studentsche t-Verteilung (2)

Wenn
x ZV aus Gauss-WDF mit $\mu = 0$, $\sigma^2 = 1$, und
z ZV aus χ^2-WDF mit n Freiheitsgraden,
dann ist ZV $t = x / (z/n)^{1/2}$ verteilt gemäß
Studentscher *t-WDF* mit $\nu = n$ Freiheitsgraden.

Diese ZV taucht in Problemen auf, in denen man das Verhältnis
Von Stichprobenmittelwert und Stichprobenvarianz betrachtet.

Studentsche t-WDF liefert Glockenkurve mit adjustierbaren Ausläufern
von Gauss-WDF $(\nu \to \infty)$ bis Cauchy-WDF $(\nu = 1)$.

Entwickelt in 1908 von William Gosset, der unter dem Pseudonym
"Student" für die Guinness-Brauerei arbeitete.

Hypothesentest für ZV aus Gauss-WDF (2)

Abweichung von $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ von μ_0 in beiden Richtungen

in beide Richtungen ist schlechte Übereinstimmung mt Nullhypothese.
→ beidseitiger Test mit $\alpha/2$ in jedem Ausläufer.

ii) Test derVarianzhypothese H_0: $\sigma=\sigma_0$ H_1:$\sigma\neq\sigma_0$

a) Mittelwert bekannt

Teststatistik $(n-1)s^2/\sigma_0^2 = \sum_{i=1}^{n} (x_i-\mu)^2/\sigma_0^2$, folgt Chi-2Qudratverteilung mit n Freiheitsgraden $\chi^2(n)$

b) Mittelwert unbekannt

Teststatistik $(n-1)s^2/\sigma_0^2 = \sum_{i=1}^{n} (x_i-\bar{x})^2/\sigma_0^2$ folgt Chi-Quadratverteilung mit n-1 Freiheitsgraden $\chi^2(n-1)$

Hypothesentest für ZV aus Gauss-WDF (3)

B) Vergleich von 2 Stichprobe $x_1, \ldots, x_n$ und $y_1, \ldots y_m$

i) Test der der Gleicheit der Mittelwerte $H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$

beide Varianzen bekannt

$\bar{x}$ and $\bar{y}$ folgen Gauss-WDFs $N(\mu_1, \sigma_1^2/n)$ $N(\mu_2, \sigma_2^2/m)$

Gemäß Additionstheorems für Gauss-WDFs ist die Differenz $\bar{x}-\bar{y}$ gaussverteilt mit Mittelwert $(\mu_1 - \mu_2)$ und Varianz $(\sigma_1^2/n + \sigma_2^2/m)$

Variable $$\frac{(\bar{x}-\bar{y}) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n + \sigma_2^2/m}}$$ folgt Standardgauss-WDF $N(0,1)$

Ebenso Teststatistik: $$\frac{\bar{x} - \bar{y}}{\sqrt{\sigma_1^2/n + \sigma_2^2/m}}$$

Hypothesentest für ZV aus Gauss-WDF (4)

Test der der Geichheit der Mittelwerte H_0: $\mu_1=\mu_2$ H_1:$\mu_1\neq\mu_2$

beide Varianzen unbekannt aber gleich.

Es gilt für die 4 unabhängigen ZV:

$\bar{x}$, which is $N(\mu_1, \sigma_1^2/n)$, $(n-1)s_1^2/\sigma_1^2$, which is $\chi^2(n-1)$,

$\bar{y}$, which is $N(\mu_2, \sigma_2^2/m)$, $(m-1)s_2^2/\sigma_2^2$, which is $\chi^2(m-1)$.

Differenz der Mittelwerte $\dfrac{(\bar{x}-\bar{y}) - (\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n + \sigma_2^2/m}}$ folgt $N(0,1)$

Summe der Varianzen $(n-1)s_1^2/\sigma_1^2 + (m-1)s_2^2/\sigma_2^2$ folgt $\chi^2(n+m-2)$

(aus Additionstheorem für Chi-Quadrat-ZV u. $(n-1)+(m-1)=n+m-2$

Hypothesentest für ZV aus Gauss-WDF (5)

Die Observable

$$\frac{\dfrac{(\bar{x}-\bar{y})-(\mu_1-\mu_2)}{\sqrt{\sigma_1^2/n+\sigma_2^2/m}}}{\sqrt{\{(n-1)s_1^2/\sigma_1^2+(m-1)s_2^2/\sigma_2^2\}/(n+m-2)}}$$

folgt Studentscher t-Verteilung mit n+m-2 Freiheitsgraden

Aber Varianzen unbekannt. Bei Gleichheit vereinfacht sich Ausdruck zu

$$\frac{(\bar{x}-\bar{y})-(\mu_1-\mu_2)}{S\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}} \quad \text{mit} \quad S^2 \equiv \frac{1}{n+m-2}\left((n-1)s_1^2+(m-1)s_2^2\right)=\frac{1}{n+m-2}\left(\sum_{i=1}^{n}(x_i-\bar{x})^2+\sum_{i=1}^{m}(y_i-\bar{y})^2\right)$$

Teststatistik $\dfrac{\bar{x}-\bar{y}}{S\sqrt{\frac{1}{n}+\frac{1}{m}}}$ folgt Studentscher t-Verteilung

mit n+m-2 Freiheitsgraden.

Für ungleiche Varianzen keine geschlossene exakte Lösung.

Hypothesentest für ZV aus Gauss-WDF (6)

i) Test der der Gleicheit der Varianzen H_0: $\sigma_1=\sigma_2$ H_1: $\sigma_1>\sigma_2$

a) beide Mittelwerte unbekannt

Es gilt:

$$\frac{(n-1)s_1^2}{\sigma_1^2} = \frac{1}{\sigma_1^2}\sum_{i=1}^{n}(x_i-\bar{x})^2 \quad \text{is } \chi^2(n-1)$$

$$\frac{(m-1)s_2^2}{\sigma_2^2} = \frac{1}{\sigma_2^2}\sum_{i=1}^{m}(y_i-\bar{y})^2 \quad \text{is } \chi^2(m-1)$$

Der Ausdruck

$$\frac{\frac{(n-1)s_1^2}{\sigma_1^2}/(n-1)}{\frac{(m-1)s_2^2}{\sigma_2^2}/(m-1)} = \frac{s_1^2}{s_2^2}\cdot\frac{\sigma_2^2}{\sigma_1^2}$$

folgt einer Fisherschen F-Verteilung mit n-1, m-1 Freiheitsgraden

Ebenso die Teststatistik:

$$\frac{s_1^2}{s_2^2} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2 \Big/ \frac{1}{m-1}\sum_{i=1}^{m}(y_i-\bar{y})^2$$

Fischers F-Verteilung

$$f(F;\nu_1,\nu_2) = \frac{\Gamma\left(\frac{1}{2}(\nu_1+\nu_2)\right)}{\Gamma(\frac{1}{2}\nu_1)\Gamma(\frac{1}{2}\nu_2)} \left(\frac{\nu_1}{\nu_2}\right)^{\frac{1}{2}\nu_1} \frac{F^{\frac{1}{2}\nu_1 - 1}}{\left(1+\frac{\nu_1 F}{\nu_2}\right)^{\frac{1}{2}(\nu_1+\nu_2)}}$$

$$0 \leq F \leq \infty; \qquad \nu_1, \nu_2 > 0$$

$$F_{mode} = \frac{\nu_1 - 2}{\nu_1} \frac{\nu_2}{\nu_2 + 2} \qquad \nu_1 > 2$$

$$E(F) = \frac{\nu_2}{\nu_2 - 2}, \qquad \nu_2 > 2$$

$$V(F) = \frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)}, \qquad \nu_2 > 4$$

$\nu_1 = 1$ ergibt Student t-WDF

$\nu 1_{/2} \rightarrow \infty$ ergibt Gauss-WDF

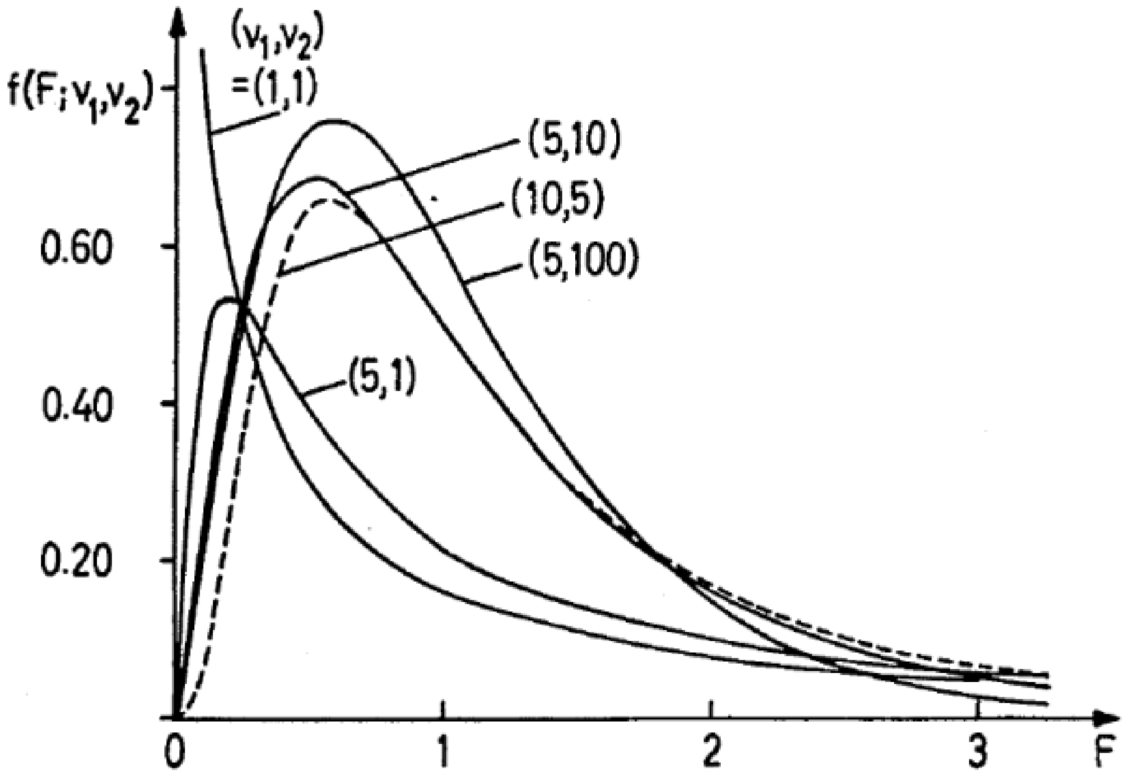


Fig. 5.4. The F-distribution for selected degrees of freedom (ν_1,ν_2).

Hypothesentest für ZV aus Gauss-WDF (7)

für den Fall $H_1{:}\sigma_1>\sigma_2$ kritische Region im oberen Ausläufer der F-Verteilung

für den Fall $H_1{:}\sigma_1<\sigma_2$ kritische Region im unteren Ausläufer der F-Verteilung

b) beide Mittelwerte unbekannt

Teststatistik $$\frac{s_1^2}{s_2^2} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\mu_1)^2 \Big/ \frac{1}{m-1}\sum_{i=1}^{m}(y_i-\mu_2)^2$$

folgt einer Fischerchen F-Verteilung mit n,m Freiheitsgraden

Beispiele und praktische Anwendung in den Übungen

Zusammenfassende Tabelle für Test mit Gauss-WDF

Table 14.1 Tests for the mean and variance of normal distributions

H_0	Conditions	Test statistic	Test distribution
$\mu=\mu_0$	σ^2 known	$\frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$	$N(0,1)$
	σ^2 unknown	$\frac{\bar{x}-\mu_0}{s/\sqrt{n}}$	$t(n-1)$
$\sigma^2=\sigma_0^2$	μ known	$(n-1)s^2/\sigma_0^2=\sum_{i=1}^{n}(x_i-\mu)^2/\sigma_0^2$	$\chi^2(n)$
	μ unknown	$(n-1)s^2/\sigma_0^2=\sum_{i=1}^{n}(x_i-\bar{x})^2/\sigma_0^2$	$\chi^2(n-1)$
$\mu_1-\mu_2=0$	$\sigma_1^2=\sigma_2^2=\sigma^2$ known	$\frac{\bar{x}-\bar{y}}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}}$	$N(0,1)$
	σ_1^2 and σ_2^2 known	$\frac{\bar{x}-\bar{y}}{\sqrt{\sigma_1^2/n+\sigma_2^2/m}}$	
	$\sigma_1^2=\sigma_2^2=\sigma^2$ unknown	$\frac{\bar{x}-\bar{y}}{S\sqrt{\frac{1}{n}+\frac{1}{m}}}$ $S^2\equiv\frac{1}{n+m-2}\left((n-1)s_1^2+(m-1)s_2^2\right)$	$t(n+m-2)$
	$\sigma_1^2\neq\sigma_2^2$ unknown	$\frac{\bar{x}-\bar{y}}{\sqrt{s_1^2/n+s_2^2/m}}$	not exactly known, $\approx N(0,1)$
$\frac{\sigma_1^2}{\sigma_2^2}=1$	μ_1 and μ_2 known	$\frac{s_1^2}{s_2^2}=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\mu_1)^2\Big/\frac{1}{m-1}\sum_{i=1}^{m}(y_i-\mu_2)^2$	$F(n,m)$
	μ_1 and μ_2 unknown	$\frac{s_1^2}{s_2^2}=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2\Big/\frac{1}{m-1}\sum_{i=1}^{m}(y_i-\bar{y})^2$	$F(n-1,m-1)$