III. Schließende Statistik
III.1 Grundbegriffe

Aufgabe der schließenden Statistik:
Aus einer oder mehreren „Zufallsstichproben“ sollen mit Hilfe der Wahrscheinlichkeitsrechnung Rückschlüsse auf die „Grundgesamtheit“ gezogen werden.

Beispiel: Aus einem Lieferlos des Umfangs 10000 Einheiten werden 100 Einheiten zufällig entnommen. Eine Prüfung ergibt, dass 5 Einheiten defekt sind. Welche Aussagen über den Anteil fehlerhafter Einheiten sind damit für das gesamte Lieferlos möglich?

III. Schließende Statistik
III.1 Grundbegriffe

<u>Definition „Grundgesamtheit"</u>
Unter einer Grundgesamtheit verstehen wir die Gesamtheit gleichartiger Objekte oder Elemente, die hinsichtlich eines bestimmten Merkmals untersucht werden sollen. Das interessierende Merkmal beschreiben wir dabei durch eine Zufallsvariable X.

<u>Definition „Zufallsstichprobe" (Stichprobe)</u>
Eine Stichprobe vom Umfang n der Zufallsvariablen X ist die Beobachtung von n unabhängigen, identisch verteilten Zufallsvariablen $X_1, \ldots, X_n$. $X_1, \ldots, X_n$ nennt man Stichprobenvariablen, die beobachteten Werte $x_1, \ldots, x_n$ die Stichprobenwerte.

III. Schließende Statistik
III.1 Grundbegriffe

Weiterverarbeitung des Stichprobenergebnisses zur Schätzung von Parametern der Grundgesamtheit:

- Punktschätzungen
(z.B. Schätzungen für Parameter der Verteilung)

- Intervallschätzungen
(Bereich, der den unbekannte Parameter mit einer vorgegebenen Sicherheit enthält)

- Hypothesentests
(Entscheidung zwischen zwei Vermutungen bzw. Annahmen)

Zur Schätzung eines Parameters θ der Grundgesamtheit wird aus den n Stichprobenvariablen ein Wert gebildet, der i.d.R. von allen X_i abhängt:

Definition „Schätzfunktion" (Schätzer)

Eine Funktion $g(X_1, \ldots, X_n)$ der Stichprobenvariablen heißt Stichprobenfunktion und ist selbst wieder eine zufällige Variable. Wird die Stichprobenfunktion zur Schätzung des Parameters verwendet, so heißt sie Schätzfunktion oder kurz Schätzer für θ und wird mit $\hat{\theta}$ bezeichnet.

Zwei Probleme:

- Wie findet man einen solchen Schätzer?
- Welche Eigenschaften bzw. Qualität hat der Schätzer?

III. Schließende Statistik
III.2 Punktschätzungen

Eine „gute" Schätzfunktion sollte wenigstens „im Mittel" richtig schätzen, d.h. ihr Erwartungswert sollte gleich dem zu schätzenden Parameter sein und die Schätzung sollte mit wachsendem Stichprobenumfang n immer genauer werden.

Eine Schätzfunktion $\hat{\theta}$ für den Parameter θ heißt erwartungstreu oder unverfälscht (engl. unbiased), wenn gilt:

$$\boxed{E(\hat{\theta}) = \theta}$$

Sie heißt konsistent, wenn ihre Varianz mit wachsendem Stichprobenumfang n gegen 0 strebt:

$$\boxed{\lim_{n \to \infty} Var(\hat{\theta}) = 0}$$

III. Schließende Statistik
III.2 Punktschätzungen

Schätzung einer Wahrscheinlichkeit

Sei A ein Ereignis, $p := P(A)$; Betrachte Stichprobe vom Umfang n:

Stichprobenwerte $x_1, \ldots, x_n$ ($x_i = 1$, wenn A eingetreten, $x_i = 0$ sonst)

Eine geeignete Schätzung für p ist die relative Häufigkeit von A:

$$\hat{p} = \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$E(\overline{X}_n) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} p = p \quad \Rightarrow \hat{p} = \overline{X}_n \text{ ist erwartungstreu}$$

$$P(|\overline{X}_n - p| > \varepsilon) \underset{n\to\infty}{\longrightarrow} 0 \ \forall \varepsilon > 0 \quad \text{Schwaches Gesetz großer Zahlen}$$

$$\Rightarrow \hat{p} = \overline{X}_n \text{ ist konsistent}$$

Maß für die Genauigkeit einer erwartungstreuen Schätzung: Varianz des Schätzers, muss hier selber über p geschätzt werden!

$$Var(\overline{X}_n) = \frac{p(1-p)}{n}$$

Schätzung des Erwartungswertes

$$E(\overline{X}_n) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\cdot n\mu = \mu$$

$$Var(\overline{X}_n) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2}\cdot n\sigma^2 = \frac{\sigma^2}{n} \underset{n\to\infty}{\longrightarrow} 0$$

Das arithmetische Mittel ist eine erwartungstreue und konsistente Schätzung des Erwartungswertes von X.

Schätzung der Varianz

$$E(s_n^2) = E\left(\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2\right) = \sigma^2$$

Die Stichprobenvarianz ist eine erwartungstreue und konsistente Schätzung der Varianz von X.

Beweis: s. Präsenzübungen!

Was tun, wenn der Erwartungswert von X bekannt ist und nicht geschätzt werden muss?

$$\boxed{E\left(\frac{1}{n}\sum_{i=1}^{n}\left(X_i - E[X]\right)^2\right) = \sigma^2}$$

Die Maximum-Likelihood-Methode

Strategie:

Den bzw. die gesuchten Parameter der Verteilung bestimmen, indem die Wahrscheinlichkeit für das erhaltene Stichprobenergebnis maximal wird.

III. Schließende Statistik
III.2 Punktschätzungen

<u>Annahme</u>:
X sei eine diskrete Zufallsvariable, deren Wahrscheinlichkeitsfunktion f(x) noch einen unbekannten Parameter θ enthalte, der aus einer Zufallsstichprobe mit n voneinander unabhängigen Stichprobenwerten $x_1, x_2, \ldots, x_n$ geschätzt werden soll:

x_i	x_1	x_2	$\ldots$	x_n
$P(X = x_i) = f(x_i)$	$f(x_1)$	$f(x_2)$	$\ldots$	$f(x_n)$

$$L = f(x_1) \cdot f(x_2) \ldots \cdot f(x_n)$$

Wegen Unabhängigkeit der Stichprobenwerte ist die Wahrscheinlichkeit für diese spezielle Stichprobe durch obiges Produkt gegeben.

L hängt bei festen Stichprobenwerten x_i nur noch vom gesuchten Parameter θ ab:

„Likelihood"-Funktion:

$$L = L(\theta) = f(x_1;\theta) \cdot f(x_2;\theta) \cdot \ldots \cdot f(x_n;\theta)$$

θ wird aus der notwendigen Bedingung für ein Maximum bestimmt (manchmal ohne Differentialrechung!):

$$\frac{\partial L}{\partial \theta} = 0$$

Die hinreichende Bedingung für die zweite Ableitung ist ebenfalls zu überprüfen!

Bei stetigen Zufallsvariablen:
Die Likelihood-Funktion muss mit der entsprechenden Dichtefunktion f(x;θ) gebildet werden!

Falls die Wahrscheinlichkeits- bzw. Dichtefunktion mehrere unbekannte Parameter enthält:

$$L = L(\theta_1;\theta_2;\ldots;\theta_r)$$

$$\frac{\partial L}{\partial \theta_1} = 0; \quad \frac{\partial L}{\partial \theta_2} = 0; \quad \ldots \quad ; \frac{\partial L}{\partial \theta_r} = 0$$

r Gleichungen mit r Unbekannten

III. Schließende Statistik
III.2 Punktschätzungen

Häufig für praktische Berechnung wertvolle Vereinfachung:

$$L^* = \ln(L) = \ln(L(\theta_1;\theta_2;\ldots;\theta_r))$$

$$\frac{\partial}{\partial\theta_1}(\ln(L)) = 0; \quad \frac{\partial}{\partial\theta_2}(\ln(L)) = 0; \quad \ldots \quad ; \frac{\partial}{\partial\theta_r}(\ln(L)) = 0$$

Die Maximum-Likelihood-Methode liefert nicht immer erwartungstreue Schätzer!

FH AACHEN UNIVERSITY OF APPLIED SCIENCES

Eine Punktschätzung eines Parameters ist eine Näherung:
Wie groß ist der Fehler, den man dabei macht?

Betrachte das Konzept bei numerischer Integration:

$$\int_a^b f(x)dx \approx T \quad \text{z.B. Trapezformel}$$

$$\text{Fehlerabschätzung}: R := \left| \int_a^b f(x)dx - T \right| \leq \varepsilon$$

$$\Leftrightarrow T-\varepsilon \leq \int_a^b f(x)dx \leq T+\varepsilon \quad \Leftrightarrow \quad \int_a^b f(x)dx \in [T-\varepsilon; T+\varepsilon] \text{ zu } 100\%$$

Analoges Konzept zur Intervallschätzung in der schließenden Statistik:

Basierend auf einer Stichprobe $X_1, \ldots, X_n$ wird ein Intervall

$$I_n = [c_u ; c_o]$$

berechnet, das den gesuchten Parameter θ mit einer vorgegebenen Sicherheit von x% enthält.

x=100% ist unmöglich, da I_n vom Zufall abhängt!

III. Schließende Statistik
III.3 Intervallschätzungen

Definition :

Sei $X_1,\ldots,X_n$ eine Stichprobe zu einer Verteilung mit Parameter θ und $0<\alpha<1$. Ist dann $I_n := \left[c_u;c_o\right]$ ein Intervall, das von der Stichprobe abhängt und gilt

$P(\theta \in I_n) \geq 1-\alpha$

dann nennt man I_n ein $100\cdot(1-\alpha)\%-$iges Konfidenzintervall für θ und $1-\alpha$ das Konfidenzniveau.

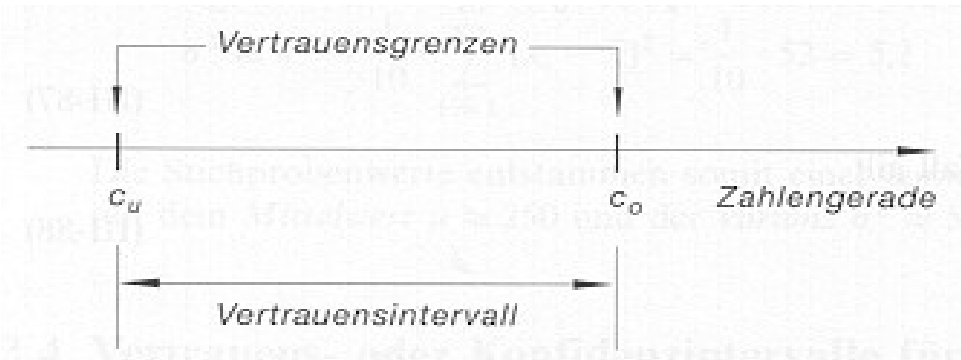
Typische Werte für α sind $\alpha = 0{,}05$ oder $\alpha = 0{,}01$ oder $\alpha = 0{,}001$.

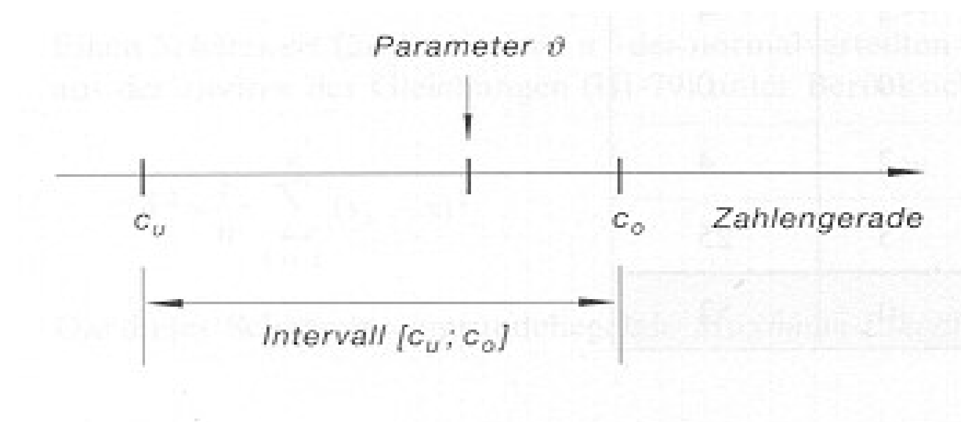
Zweiseitiges Konfidenzintervall: $I_n(x_1,\ldots x_n) = \left[c_u(x_1,\ldots x_n);c_o(x_1,\ldots x_n)\right]$

Einseitiges Konfidenzintervall: $I_n(x_1,\ldots x_n) = \left[c_u^*(x_1,\ldots x_n);\infty\right)$ nach unten beschränkt

$I_n(x_1,\ldots x_n) = \left(-\infty;c_o^*(x_1,\ldots x_n)\right]$ nach oben beschränkt

III. Schließende Statistik
III.3 Intervallschätzungen





Der Parameter liegt entweder innerhalb des berechneten Vertrauensbereiches oder außerhalb!

Nicht: "Der Parameter liegt zu 95% im Konfidenzintervall."!!!

FH AACHEN UNIVERSITY OF APPLIED SCIENCES

III.3.1 Vertrauensintervalle für den unbekannten Mittelwert μ einer Normalverteilung bei bekannter Varianz σ^2

Nur mit der inzwischen bekannten (Standard)-Normalverteilung lässt sich bereits dieses Konfidenzintervall berechnen:

X sei eine normalverteilte Zufallsvariable mit unbekanntem Mittelwert μ und bekannter Varianz σ^2.

Schätzfunktion für den Mittelwert μ: $\overline{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$

$U = \frac{\overline{X} - \mu}{\sigma / \sqrt{n}}$ ist dann standardnormalverteilt mit Mittelwert 0 und Varianz 1.

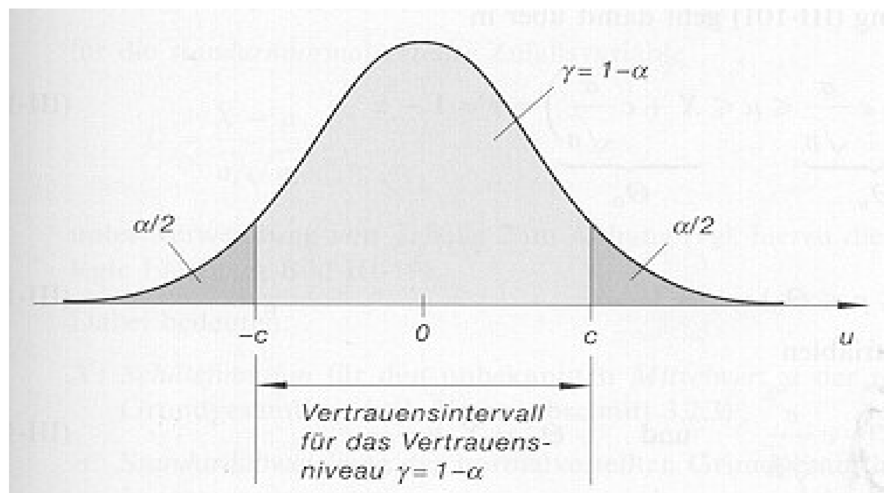
Für U lässt sich schrittweise ein Vertrauensintervall konstruieren:

1.) Wähle ein bestimmtes Vertrauensniveau $\gamma = 1 - \alpha$

(Typisch : 95%, 99% oder 99,9%)

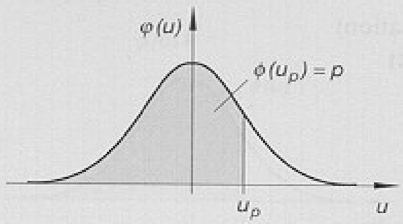
2.) Die Zufallsvariable U soll dann mit der gewählten Wahrscheinlichkeit γ einen Wert in dem symmetrischen Intervall $-c \le U \le c$ annehmen, d.h.

$$P(-c \le U \le c) = \gamma = 1 - \alpha$$



Der Wert für c lässt sich bei gegebenem Vertrauensniveau der Tabelle „Quantile der Standardnormalverteilung“ entnehmen:

Tabelle 2: Quantile der Standardnormalverteilung



p: Vorgegebene Wahrscheinlichkeit ($0 < p > 1$)

u_p: Zur Wahrscheinlichkeit p gehöriges Quantil (*obere* Schranke)

Die Tabelle enthält für spezielle Werte von p das jeweils zugehörige Quantil u_p (*einseitige* Abgrenzung nach *oben*).

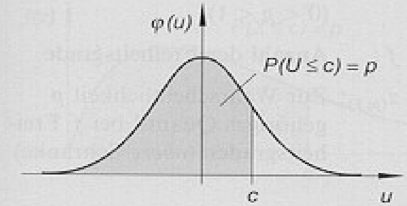
p	u_p	p	u_p
0,90	1,282	0,1	− 1,282
0,95	1,645	0,05	− 1,645
0,975	1,960	0,025	− 1,960
0,99	2,326	0,01	− 2,326
0,995	2,576	0,005	− 2,576
0,999	3,090	0,001	− 3,090

Formeln:

$$u_{1-p} = -u_p$$
$$u_p = -u_{1-p}$$

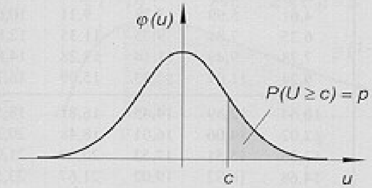
Formeln zur Berechnung von Quantilen

(1) *Einseitige* Abgrenzung nach *oben*



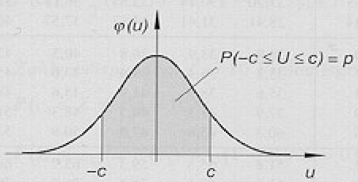
$$P(U \leqslant c) = \phi(c) = p$$
$$\phi(c) = p \;\rightarrow\; c = u_p$$

(2) *Einseitige* Abgrenzung nach *unten*



$$P(U \geqslant c) = 1 - P(U \leqslant c) = 1 - \phi(c) = p$$
$$\phi(c) = 1 - p \;\rightarrow\; c = u_{1-p}$$

(3) *Zweiseitige* (symmetrische) Abgrenzung



$$P(-c \leqslant U \leqslant c) = 2 \cdot \phi(c) - 1 = p$$
$$\phi(c) = \frac{1}{2}(1 + p) \;\rightarrow\; c = u_{(1+p)/2}$$

III. Schließende Statistik
III.3 Intervallschätzungen

3.) $-u_{1-\frac{\alpha}{2}} \leq U \leq u_{1-\frac{\alpha}{2}} \quad \Leftrightarrow \quad \overline{X} - u_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + u_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$

$$P\left(\overline{X} - u_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + u_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right) = \gamma = 1 - \alpha$$

4.) Die Berechnung der Vertrauensgrenzen erfolgt dann anhand einer konkreten Stichprobe: Das Vertrauensintervall

$$\overline{x} - u_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{x} + u_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

enthält den unbekannten Mittelwert μ der normalverteilten Grundgesamtheit mit einem Vertrauen von $\gamma \cdot 100\%$

III.3.2 Vertrauensintervalle für die unbekannte Varianz σ^2 einer Normalverteilung

Bei der Berechnung von Konfidenzintervallen werden i. a. neben der Gaußschen Normalverteilung weitere „Prüf- oder Testverteilungen" benötigt, die zunächst bereitgestellt werden müssen:

Chi-Quadrat-Verteilung (χ^2 – Verteilung)

$X_1, X_2, \ldots, X_n$ seien n mit N(0;1) verteilte, stochastisch unabhängige Zufallsvariable. Betrachte

$$Z = \chi^2 = X_1^2 + X_2^2 + \ldots + X_n^2$$

stetige Zufallsvariable mit Wertebereich $z \geq 0$

Dichtefunktion dieser Zufallsvariablen :

$$f(z) = \begin{cases} A_n \cdot z^{\left(\frac{n-2}{2}\right)} \cdot e^{-\frac{z}{2}} & \text{für} \quad z > 0 \\ 0 & \text{für} \quad z \leq 0 \end{cases} \qquad \chi^2 - Verteilung$$

Der Parameter $n \in N$ bestimmt die Anzahl f = n

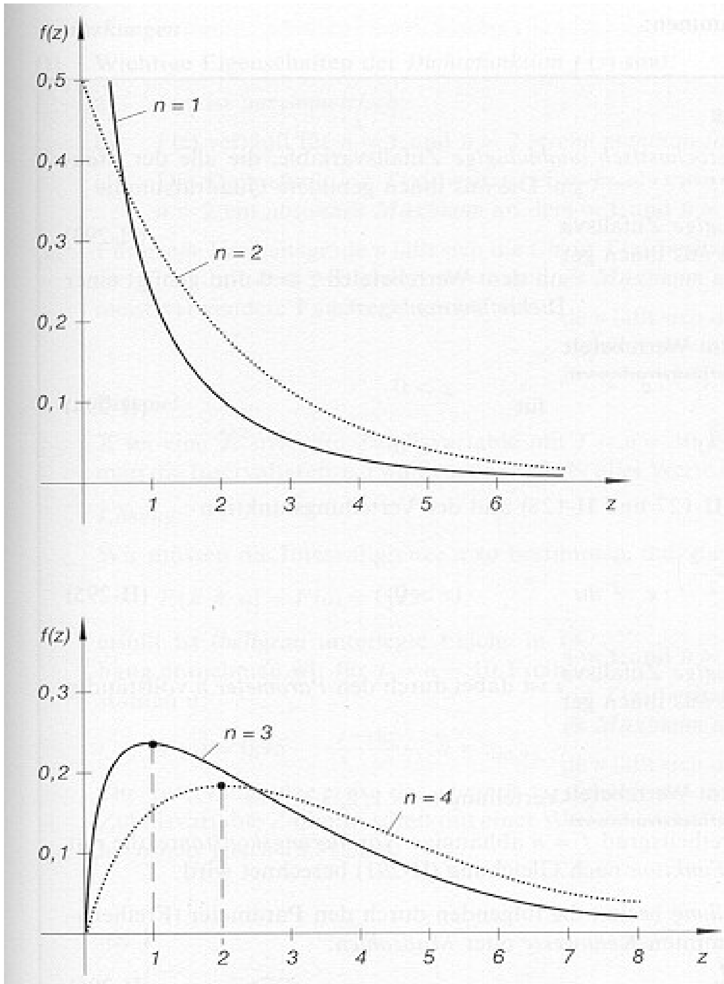
der "Freiheitsgrade" der Verteilung.

Die Konstante A_n wird durch Normierung der Dichtefunktion

bestimmt :

$$\int_0^\infty f(z)\,dz = A_n \cdot \int_0^\infty z^{\left(\frac{n-2}{2}\right)} \cdot e^{-\frac{z}{2}}\,dz = 1 \Rightarrow A_n = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}$$

III. Schließende Statistik
III.3 Intervallschätzungen



Verteilungsfunktion der χ^2 – Verteilung:

$$F(z) = A_n \cdot \int_0^z u^{\left(\frac{n-2}{2}\right)} \cdot e^{-\frac{u}{2}}\, du$$

Mittelwert: $\mu = n$

Varianz: $\sigma^2 = 2n$

<u>Eigenschaften</u>:

- f(z) ist asymmetrisch
- f(z) ist für n=1 und n=2 streng monoton fallend
- f(z) besitzt für n>2 ein absolutes Maximum an der Stelle $z_{max}=n-2$
- Für „große“ Freiheitsgrade (n>100) lässt sich die Chi-Quadrat-Verteilung durch eine Normalverteilung mit $\mu=n$ und Varianz $\sigma^2=2n$ annähern

III. Schließende Statistik
III.3 Intervallschätzungen

X sei eine normalverteilte Zufallsvariable mit

unbekanntem Mittelwert μ und unbekannter Varianz σ^2.

Schätzfunktion für die Varianz σ^2 : $S^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2$

$$Z = (n-1) \cdot \frac{S^2}{\sigma^2} = \frac{1}{\sigma^2} \cdot \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2 = \sum_{i=1}^{n} \left(\frac{X_i - \overline{X}}{\sigma}\right)^2 \text{ genügt}$$

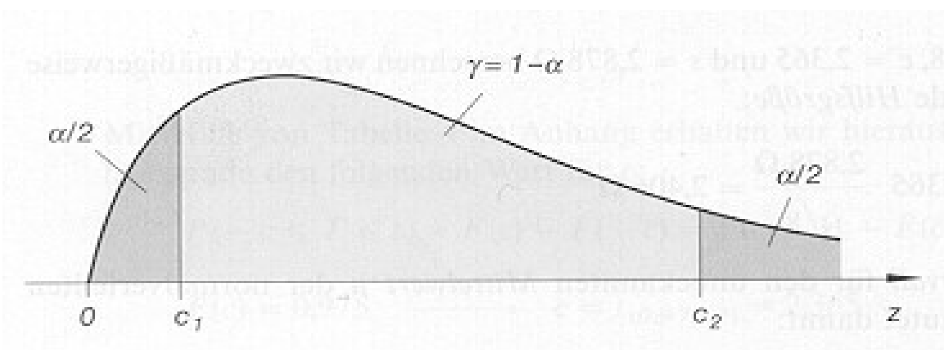
dann einer χ^2 - Verteilung mit $\mathrm{f = n - 1}$ Freiheitsgraden.

Beachte: n-1 Freiheitsgrade statt bisher n, da hier µ über $\overline{X}$ der Stichprobe geschätzt werden muss. Dabei wird ein Freiheitsgrad "verbraucht"!

Für Z lässt sich schrittweise ein Vertrauensintervall konstruieren :

1.) Wähle ein bestimmtes Vertrauensniveau $\gamma = 1 - \alpha$
(Typisch : 95%, 99% oder 99,9%)

2.) Die Zufallsvariable Z soll dann mit der gewählten Wahrscheinlichkeit $\gamma = 1 - \alpha$ einen Wert im Intervall $c_1 \le Z \le c_2$ annehmen, d.h. $P(c_1 \le Z \le c_2) = \gamma = 1 - \alpha$



$$c_1 = \chi^2_{n-1;\alpha/2};$$
$$c_2 = \chi^2_{n-1;1-\alpha/2}$$

III. Schließende Statistik
III.3 Intervallschätzungen

f	0,005	0,01	0,025	0,05	0,10	0,90	0,95	0,975	0,99	0,995
	p									
1	0,000	0,000	0,001	0,004	0,016	2,71	3,84	5,02	6,63	7,88
2	0,01	0,020	0,051	0,103	0,211	4,61	5,99	7,38	9,21	10,60
3	0,07	0,115	0,216	0,352	0,584	6,25	7,81	9,35	11,35	12,84
4	0,21	0,297	0,484	0,711	1,064	7,78	9,49	11,14	13,28	14,86
5	0,41	0,554	0,831	1,15	1,16	9,24	11,07	12,83	15,09	16,75
6	0,68	0,872	1,24	1,64	2,20	10,64	12,59	14,45	16,81	18,55
7	0,99	1,24	1,69	2,17	2,83	12,02	14,06	16,01	18,48	20,28
8	1,34	1,65	2,18	2,73	3,49	13,36	15,51	17,53	20,09	21,96
9	1,73	2,09	2,70	3,33	4,17	14,68	16,92	19,02	21,67	23,59
10	2,16	2,56	3,25	3,94	4,87	15,99	18,31	20,48	23,21	25,19
11	2,60	3,05	3,82	4,57	5,58	17,28	19,67	21,92	24,73	26,76
12	3,07	3,57	4,40	5,23	6,30	18,55	21,03	23,34	26,22	28,30
13	3,57	4,11	5,01	5,89	7,04	19,81	22,36	24,74	27,69	29,82
14	4,07	4,66	5,63	6,57	7,79	21,06	23,68	26,12	29,14	31,32
15	4,60	5,23	6,26	7,26	8,55	22,31	25,00	27,49	30,58	32,80
16	5,14	5,81	6,91	7,96	9,31	23,54	26,30	28,85	32,00	34,27
17	5,70	6,41	7,56	8,67	10,09	24,77	27,59	30,19	33,41	35,72
18	6,26	7,01	8,23	9,39	10,86	25,99	28,87	31,53	34,81	37,16
19	6,84	7,63	8,91	10,12	11,65	27,20	30,14	32,85	36,19	38,58
20	7,43	8,26	9,59	10,85	12,44	28,41	31,41	34,17	37,57	40,00
22	8,6	9,5	11,0	12,3	14,0	30,8	33,9	36,8	40,3	42,8
24	9,9	10,9	12,4	13,8	15,7	33,2	36,4	39,4	43,0	45,6
26	11,2	12,2	13,8	15,4	17,3	35,6	38,9	41,9	45,6	48,3
28	12,5	13,6	15,3	16,9	18,9	37,9	41,3	44,5	48,3	51,0
30	13,8	15,0	16,8	18,5	20,6	40,3	43,8	47,0	50,9	53,7
40	20,7	22,2	24,4	26,5	29,1	51,8	55,8	59,3	63,7	66,8
50	28,0	29,7	32,4	34,8	37,7	63,2	67,5	71,4	76,2	79,5
60	35,5	37,5	40,5	43,2	46,5	74,4	79,1	83,3	88,4	92,0
70	43,3	45,4	48,8	51,7	55,3	85,5	90,5	95,0	100,4	104,2
80	51,2	53,5	57,2	60,4	64,3	96,6	101,9	106,6	112,3	116,3
90	59,2	61,8	65,6	69,1	73,3	107,6	113,1	118,1	124,1	128,3
100	67,3	70,1	74,2	77,9	82,4	118,5	124,3	129,6	135,8	140,2

Die Werte für c_1 und c_2 lassen sich bei gegebenem Vertrauensniveau der Tabelle „Quantile der Chi-Quadrat-Verteilung" entnehmen.

III. Schließende Statistik
III.3 Intervallschätzungen

3.) $\chi^2_{n-1;\alpha/2} \leq Z \leq \chi^2_{n-1;1-\alpha/2} \quad \Leftrightarrow \quad (n-1)\cdot\frac{S^2}{\chi^2_{n-1;1-\alpha/2}} \leq \sigma^2 \leq (n-1)\cdot\frac{S^2}{\chi^2_{n-1;\alpha/2}}$

$$P\left((n-1)\cdot\frac{S^2}{\chi^2_{n-1;1-\alpha/2}} \leq \sigma^2 \leq (n-1)\cdot\frac{S^2}{\chi^2_{n-1;\alpha/2}} \right) = \gamma = 1-\alpha$$

4.) Die Berechnung der Vertrauensgrenzen erfolgt dann anhand einer konkreten Stichprobe: Das Vertrauensintervall

$$(n-1)\cdot\frac{s^2}{\chi^2_{n-1;1-\alpha/2}} \leq \sigma^2 \leq (n-1)\cdot\frac{s^2}{\chi^2_{n-1;\alpha/2}}$$

enthält die unbekannte Varianz σ^2 der normalverteilten Grundgesamtheit mit einem Vertrauen von $\gamma \cdot 100\%$. Durch Radizieren erhält man ein entsprechendes Vertrauensintervall für σ.

III.3.3 Vertrauensintervalle für den unbekannten Mittelwert μ einer Normalverteilung bei unbekannter Varianz σ^2

t-Verteilung von Student

X und Y seien zwei stochastisch unabhängige Zufallsvariable. X genügt der Standardnormalverteilung N(0;1); Y genügt der χ^2 - Verteilung mit Freiheitsgrad n. Betrachte die stetige Zufallsvariable

$$T = \frac{X}{\sqrt{\frac{Y}{n}}}$$

III. Schließende Statistik
III.3 Intervallschätzungen

$T = \dfrac{X}{\sqrt{\dfrac{Y}{n}}}$ ist eine stetige Zufallsvariable, die der "t - Verteilung" mit der

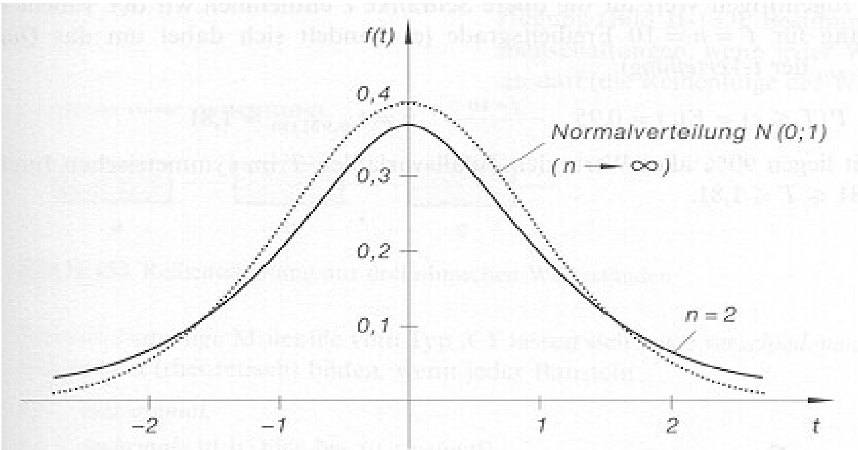
Dichtefunktion $f(t) = A_n \cdot \dfrac{1}{\left(1 + \dfrac{t^2}{n}\right)^{(n+1)/2}} \qquad -\infty < t < \infty$ genügt

Der Parameter $n \in N$ bestimmt die Anzahl f = n der "Freiheitsgrade".

Die Konstante A_n wird durch Normierung der Dichtefunktion bestimmt :

$$\int_{-\infty}^{\infty} f(t)\, dt = A_n \cdot \int_{-\infty}^{+\infty} \frac{dt}{\left(1 + \dfrac{t^2}{n}\right)^{(n+1)/2}} = 1 \Rightarrow A_n = \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\sqrt{n\pi} \cdot \Gamma\left(\dfrac{n}{2}\right)}$$

III. Schließende Statistik
III.3 Intervallschätzungen

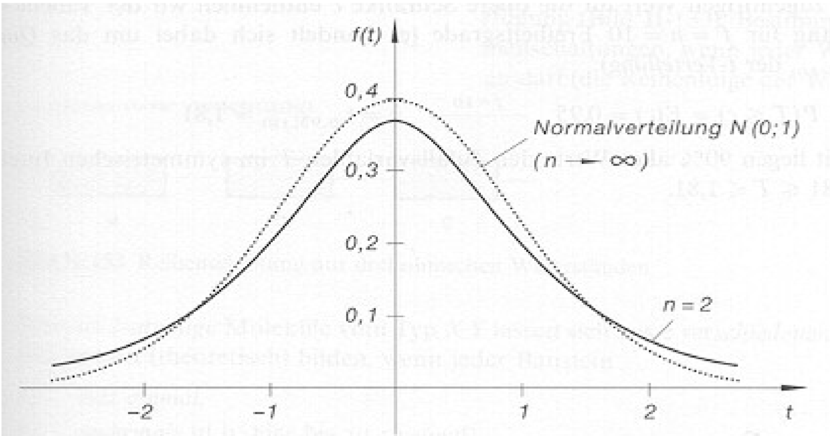


$$\text{Verteilungsfunktion der t-Verteilung: } F(t) = A_n \cdot \int_{-\infty}^{t} \frac{du}{\left(1 + \frac{u^2}{n}\right)^{(n+1)/2}}$$

$$\text{Mittelwert: } \mu = 0 \quad \text{für } n \geq 2 \ \text{(für n = 1 existiert der Mittelwert nicht)}$$

$$\text{Varianz: } \sigma^2 = \frac{n}{n-2} \quad \text{für } n \geq 3 \ \text{(für n = 1 und n = 2 existiert die Varianz nicht)}$$

III. Schließende Statistik
III.3 Intervallschätzungen



<u>Eigenschaften</u>:

- f(t) ist eine gerade Funktion (achsensymmetrisch)
- f(t) besitzt an der Stelle $t_{max}=0$ ein absolutes Maximum
- f(t) nähert sich für $t \to \pm\infty$ asymptotisch der t-Achse
- Für große Freiheitsgrade (Faustregel n>30) lässt sich die t-Verteilung durch die Standardnormal-verteilung mit $\mu=0$ und Varianz $\sigma^2=1$ annähern

III. Schließende Statistik
III.3 Intervallschätzungen

X sei eine normalverteilte Zufallsvariable mit unbekanntem Mittelwert μ und ebenfalls unbekannter Varianz σ^2.

Schätzfunktion für den Mittelwert μ: $\overline{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i$

Schätzfunktion für die Standardabweichung σ:

$$S = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2}$$

$T = \frac{\overline{X} - \mu}{S/\sqrt{n}}$ genügt dann der t - Verteilung von Student mit $\mathrm{f} = \mathrm{n} - 1$ Freiheitsgraden.

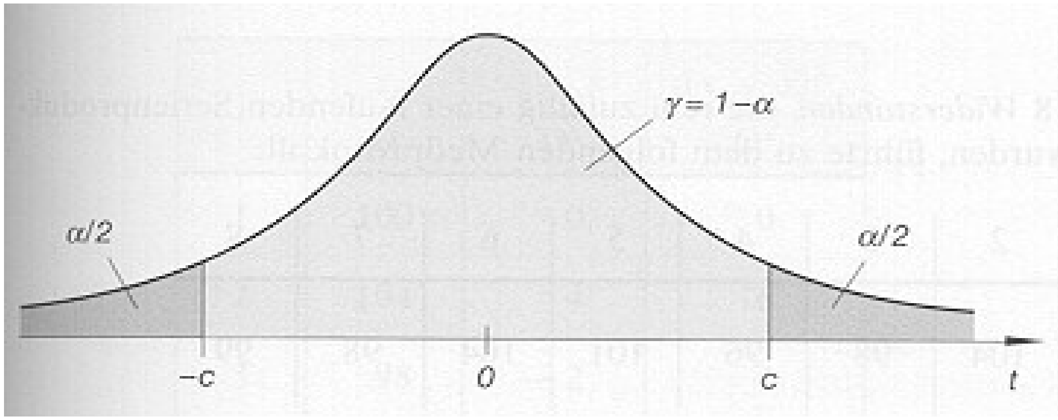
Beachte: Auch hier n-1 Freiheitsgrade statt bisher n.

Für T lässt sich schrittweise ein Vertrauensintervall konstruieren :

1.) Wähle ein bestimmtes Vertrauensniveau $\gamma = 1 - \alpha$

(Typisch : 95%, 99% oder 99,9%)

2.) Die Zufallsvariable T soll dann mit der gewählten Wahrscheinlichkeit γ einen Wert in dem symmetrischen Intervall $-c \leq T \leq c$ annehmen, d.h. $P(-c \leq T \leq c) = \gamma = 1 - \alpha$



III. Schließende Statistik
III.3 Intervallschätzungen

f	p				
	0,90	0,95	0,975	0,99	0,995
1	3,078	6,314	12,707	31,820	63,654
2	1,886	2,920	4,303	6,965	9,925
3	1,638	2,353	3,182	4,541	5,841
4	1,533	2,132	2,776	3,747	4,604
5	1,476	2,015	2,571	3,365	4,032
6	1,440	1,943	2,447	3,143	3,707
7	1,415	1,895	2,365	2,998	3,499
8	1,397	1,860	2,306	2,896	3,355
9	1,383	1,833	2,262	2,821	3,250
10	1,372	1,812	2,228	2,764	3,169
11	1,363	1,796	2,201	2,718	3,106
12	1,356	1,782	2,179	2,681	3,055
13	1,350	1,771	2,160	2,650	3,012
14	1,345	1,761	2,145	2,624	2,977
15	1,341	1,753	2,131	2,602	2,947
16	1,337	1,746	2,120	2,583	2,921
17	1,333	1,740	2,110	2,567	2,898
18	1,330	1,734	2,101	2,552	2,878
19	1,328	1,729	2,093	2,539	2,861
20	1,325	1,725	2,086	2,528	2,845
22	1,321	1,717	2,074	2,508	2,819
24	1,318	1,711	2,064	2,492	2,797
26	1,315	1,706	2,056	2,479	2,779
28	1,313	1,701	2,048	2,467	2,763
30	1,310	1,697	2,042	2,457	2,750
40	1,303	1,684	2,021	2,423	2,704
50	1,299	1,676	2,009	2,403	2,678
60	1,296	1,671	2,000	2,390	2,660
100	1,290	1,660	1,984	2,364	2,626
200	1,286	1,653	1,972	2,345	2,601
500	1,283	1,648	1,965	2,334	2,586
⋮	⋮	⋮	⋮	⋮	⋮
∞	1,282	1,645	1,960	2,326	2,576

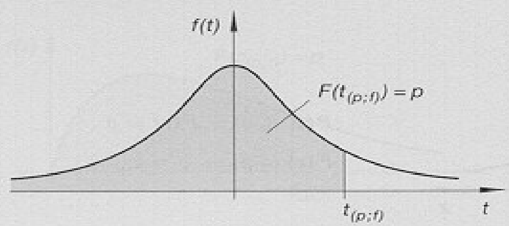
Der Wert für c lässt sich bei gegebenem Vertrauens-niveau der Tabelle „Quantile der t-Verteilung von Student" entnehmen.

Formeln:

$$t_{(1-p;\,f)} = -\,t_{(p;\,f)}$$

$$t_{(p;\,f)} = -\,t_{(1-p;\,f)}$$

Tabelle 4: Quantile der t-Verteilung von „Student"



p: Vorgegebene Wahrscheinlichkeit $(0 < p < 1)$

f: Anzahl der Freiheitsgrade

$t_{(p;\,f)}$: Zur Wahrscheinlichkeit p gehöriges Quantil bei f Freiheitsgraden (*obere* Schranke)

Die Tabelle enthält für spezielle Werte von p das jeweils zugehörige Quantil $t_{(p;\,f)}$ in Abhängigkeit vom Freiheitsgrad f (*einseitige* Abgrenzung nach *oben*).

III. Schließende Statistik
III.3 Intervallschätzungen

3.) $-t_{n-1;1-\alpha/2} \leq T \leq t_{n-1;1-\alpha/2} \quad \Leftrightarrow \quad \overline{X} - t_{n-1;1-\alpha/2} \cdot \frac{S}{\sqrt{n}} \leq \mu \leq \overline{X} + t_{n-1;1-\alpha/2} \cdot \frac{S}{\sqrt{n}}$

$$P\left(\overline{X} - t_{n-1;1-\alpha/2} \cdot \frac{S}{\sqrt{n}} \leq \mu \leq \overline{X} + t_{n-1;1-\alpha/2} \cdot \frac{S}{\sqrt{n}} \right) = \gamma = 1 - \alpha$$

4.) Die Berechnung der Vertrauensgrenzen erfolgt dann anhand einer konkreten Stichprobe:

Das Vertrauensintervall $\overline{x} - t_{n-1;1-\alpha/2} \cdot \frac{s}{\sqrt{n}} \leq \mu \leq \overline{x} + t_{n-1;1-\alpha/2} \cdot \frac{s}{\sqrt{n}}$

enthält den unbekannten Mittelwert μ der normalverteilten Grundgesamtheit mit einem Vertrauen von $\gamma \cdot 100\%$.

III.3.4 Vertrauensintervall für den unbekannten Anteilswert p (Parameter p einer Binomialverteilung)

p=P(A):
Wahrscheinlichkeit für das Ereignis A in einem Bernoulli-Experiment

X=Anzahl des Eintretens von A bei einer n-fachen Ausführung des Bernoulli-Experiments ist eine binomialverteilte Zufallsvariable, die bei „umfang-reichen" Stichproben näherungsweise normalverteilt ist (Grenzwertsatz von Moivre und Laplace) mit

$$E(X) = \mu = np \quad und \quad Var(X) = \sigma^2 = np(1-p)$$

$$\Rightarrow U = \frac{X-\mu}{\sigma} = \frac{X-np}{\sqrt{np(1-p)}} \text{ ist näherungsweise standardnormalverteilt}$$

Erwartungstreue Schätzfunktion für p:

$$\hat{P} = \frac{X}{n} \Rightarrow X = n \cdot \hat{P}$$

$$U = \frac{n\hat{P}-np}{\sqrt{np(1-p)}}; \quad P(-c \leq U \leq c) = \gamma = 1-\alpha; c = u_{1-\alpha/2} \text{ da } U \sim N(0;1)$$

$$\hat{P} - \frac{u_{1-\alpha/2}}{n}\sqrt{np(1-p)} \leq p \leq \hat{P} + \frac{u_{1-\alpha/2}}{n}\sqrt{np(1-p)}$$

III. Schließende Statistik
III.3 Intervallschätzungen

Konkrete Stichprobe: $\hat{p} = \frac{k}{n}$

$$\boxed{\hat{p} - \frac{u_{1-\alpha/2}}{n}\sqrt{n\hat{p}(1-\hat{p})} \le p \le \hat{p} + \frac{u_{1-\alpha/2}}{n}\sqrt{n\hat{p}(1-\hat{p})}}$$

Voraussetzung: $n\hat{p}(1-\hat{p}) > 9$

III. Schließende Statistik
III.4 Statistische Testverfahren

„Das kann kein Zufall sein", sagte sich der Londoner Arzt Arbuthnott vor 250 Jahren, als er in den Geburtsregistern von 80 Jahrgängen ausnahmslos die Knabengeburten häufiger vertreten fand als die Mädchengeburten. Dieser Stichprobenumfang bot ihm eine ausreichende Sicherheit für seinen Schluss."
Nach Lothar Sachs „Angewandte Statistik", 1. Auflage 1967

Ziel:
Durch ein Testverfahren soll festgestellt werden, ob eine statistische Hypothese, d.h. eine Annahme oder Vermutung über die Wahrscheinlichkeitsverteilung einer Zufallsvariablen und ihre Parameter, nicht abgelehnt werden kann oder abgelehnt werden muss.

III.4.1 Planung und Durchführung eines Parametertests

Die Wahrscheinlichkeitsverteilung einer Zufalls-variablen X sei von der Art her bekannt, enthalte jedoch noch einen oder mehrere unbekannte Parameter.

1.) Formulieren von "Nullhypothese" H_0 und "Alternativhypothese" H_1:

$H_0: \quad \vartheta = \vartheta_0$ ("der Parameter besitzt den Wert ϑ_0")

$H_1: \quad \vartheta \neq \vartheta_0$ ("der Parameter ist von ϑ_0 verschieden")

(zweiseitiger Test, auch einseitig möglich:

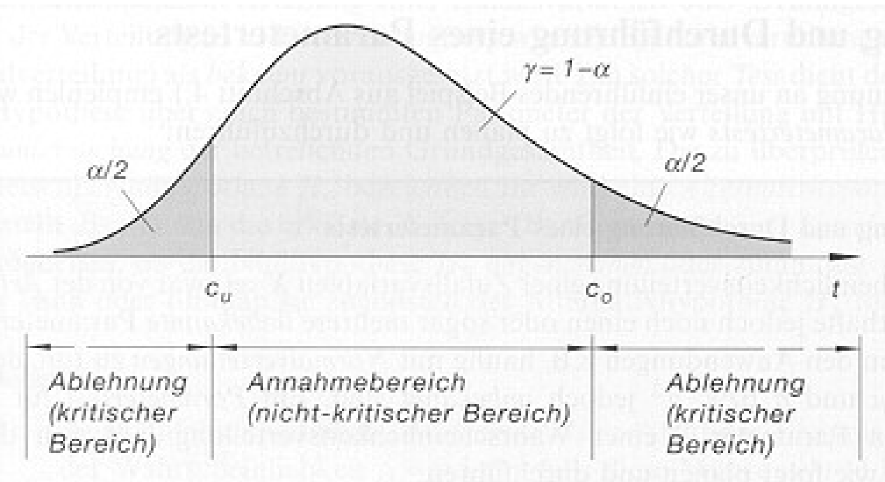
$H_1: \quad \vartheta > \vartheta_0$ oder $H_1: \quad \vartheta < \vartheta_0$)

III. Schließende Statistik
III.4 Statistische Testverfahren

2.) Festlegung des Signifikanzniveaus α. α gibt die Wahrscheinlichkeit an, die Nullhypothese abzulehnen, obwohl sie zutrifft ("Fehler 1.Art", s. später)
Typisch: $\alpha = 0{,}05$; $\alpha = 0{,}01$; $\alpha = 0{,}001$

3.) Bestimmung einer Test- oder Prüfvariablen $T = g(X_1; X_2; \ldots; X_n)$
(Stichprobenfunktion, deren Wahrscheinlichkeitsverteilung als bekannt vorausgesetzt werden muss.)

4.) Auf der Basis des Signifikanzniveaus α werden zwei kritische Grenzen c_u und c_o berechnet, so dass die Testvariable T mit der Wahrscheinlichkeit $\gamma = 1 - \alpha$ Werte aus dem Intervall $c_u \le T \le c_o$ annimmt:
$$P(c_u \le T \le c_o)_{H_0} = \gamma = 1 - \alpha$$



Mögliches Aussehen der Wahrscheinlichkeits-dichte der Testvariablen T

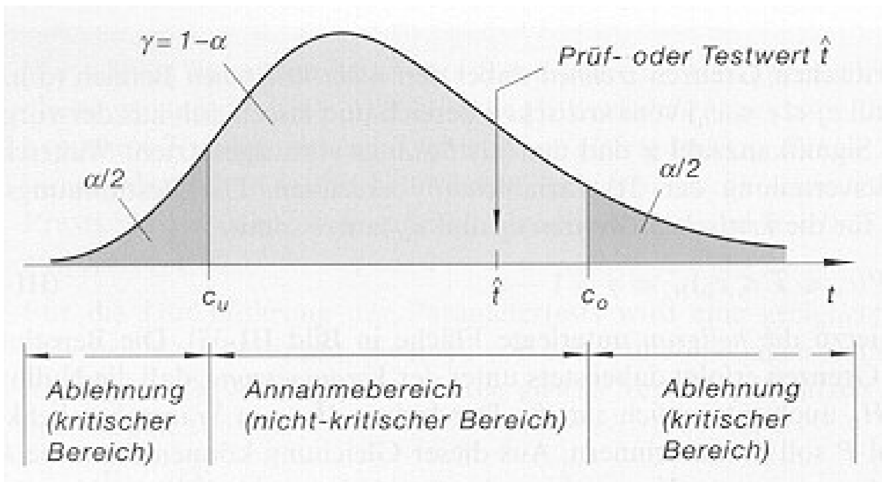
5.) Berechnung eines Wertes der Test variablen aus einer konkreten Stichprobe: $x_1; x_2; \ldots; x_n \Rightarrow \hat{t} = g(x_1; x_2; \ldots; x_n)$ als Test - oder Prüfwert von T

6.) Testentscheidung anhand des Wertes von $\hat{t}$:

III. Schließende Statistik
III.4 Statistische Testverfahren

1. Fall : $\hat{t}$ fällt in den nichtkritischen Bereich der Testvariablen T :

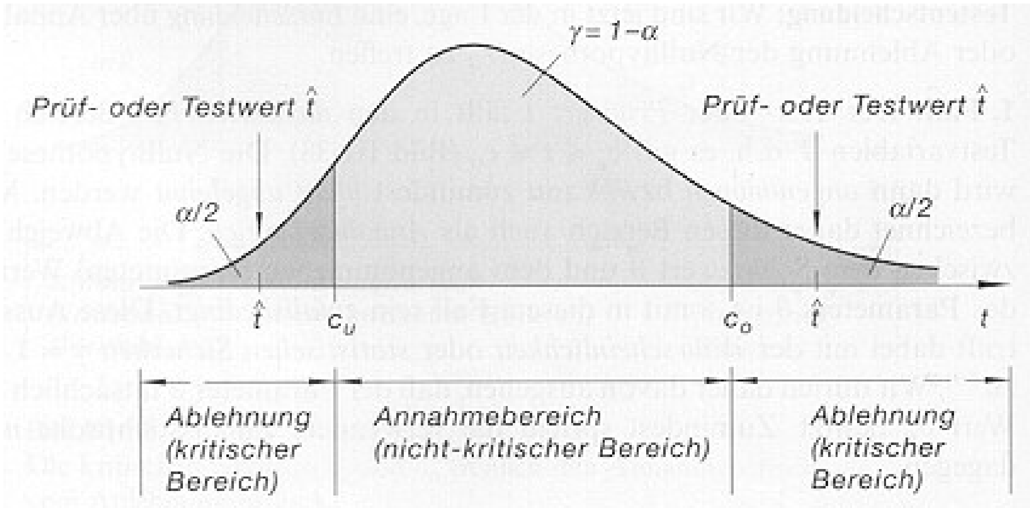
$$c_u \le \hat{t} \le c_o$$



H_0 wird nicht verworfen; der Unterschied zwischen dem Schätzwert $\hat{\vartheta}$ und dem angenommenen Wert des Parameters ϑ_0 ist zufallsbedingt.

$$2.\ \text{Fall}: \hat{t}\ \text{fällt in den kritischen Bereich der Testvariablen T}:$$

$$\hat{t} < c_u \ \text{oder}\ \hat{t} > c_o$$

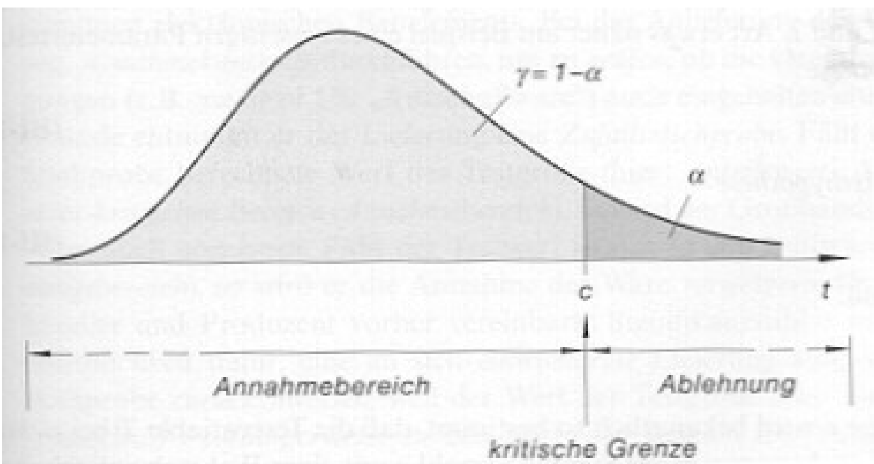


H_0 wird zugunsten der Alternativhypothese H_1 verworfen; der Schätzwert $\hat{\vartheta}$ weicht signifikant vom angenommenen Wert ϑ_0 des Parameters ab.

Vorgehen bei einem einseitigen Test:

Nullhypothese: $H_0 : \vartheta \leq \vartheta_0$; Alternativhypothese: $H_1 : \vartheta > \vartheta_0$

Bei einem einseitigen Test gibt es nur eine kritische Grenze c:



Bestimmungsgleichung:

$$P(T \leq c)_{H_0} = \gamma = 1 - \alpha$$

Merke:

- H_0 enthält immer das Gleichheitszeichen
- H_0 nicht verwerfbar, wenn durch Stichprobe gestützt

Mögliche Fehlentscheidungen:

- Die Testgröße fällt in den kritischen Bereich, obwohl die Nullhypothese zutrifft:
„Verwerfen der Nullhypothese, obwohl sie zutrifft."

- Die Testgröße fällt in den nicht-kritischen Bereich, obwohl die Nullhypothese nicht zutrifft:
„Nicht-Verwerfen der Nullhypothese, obwohl sie nicht zutrifft."

Bei jeder Testentscheidung besteht somit eine bestimmte Wahrscheinlichkeit dafür, dass die getroffene Entscheidung falsch ist:

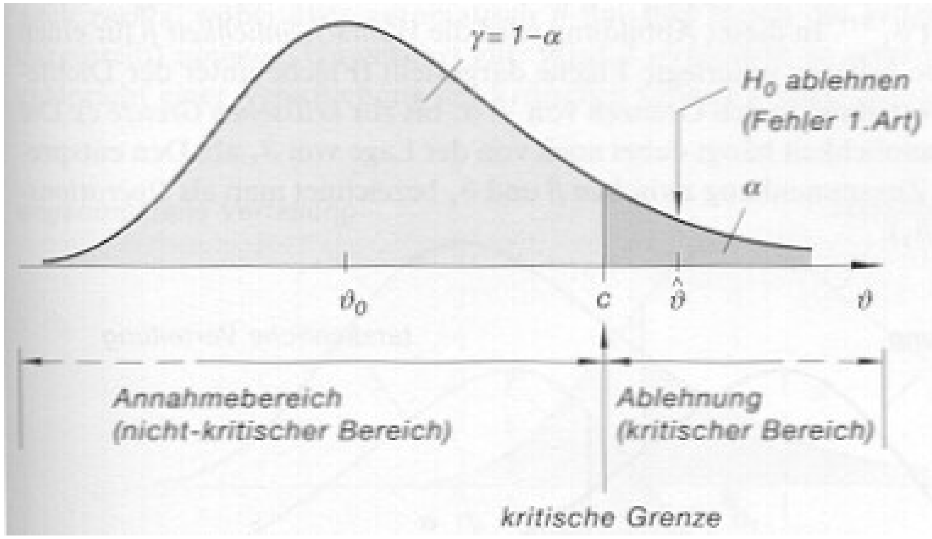
- **Fehler 1. Art:**

Eine an sich richtige Nullhypothese H_0 wird abgelehnt. Die Wahrscheinlichkeit für einen Fehler 1. Art wird mit α bezeichnet.

- **Fehler 2. Art:**

Eine an sich falsche Nullhypothese H_0 wird nicht abgelehnt. Die Wahrscheinlichkeit für einen Fehler 2. Art wird mit β bezeichnet.

Welche Wahrscheinlichkeit geben wir vor der Testdurchführung vor?

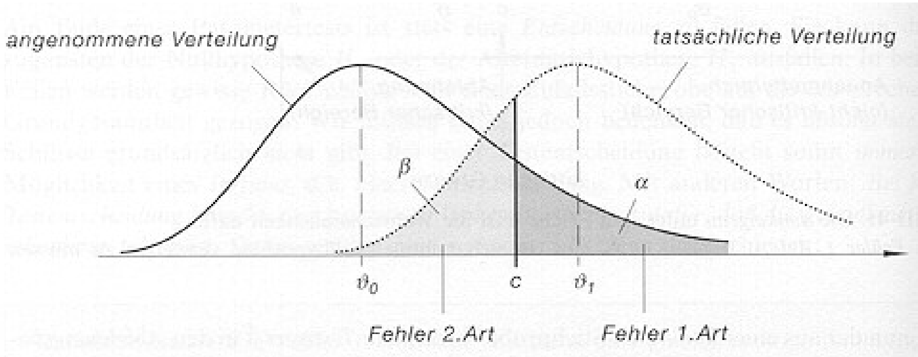


Beispiel anhand eines einseitigen Parametertests

Die Nullhypothese wird verworfen, wenn die Testgröße in den kritischen Bereich fällt. Dies ist mit der Irrtumswahrscheinlichkeit (Signifikanzniveau) α der Fall.

Über die Entscheidung für ein Signifikanzniveau wird somit die Wahrscheinlichkeit für das Auftreten des Fehlers 1. Art festgelegt!

III. Schließende Statistik
III.4 Statistische Testverfahren



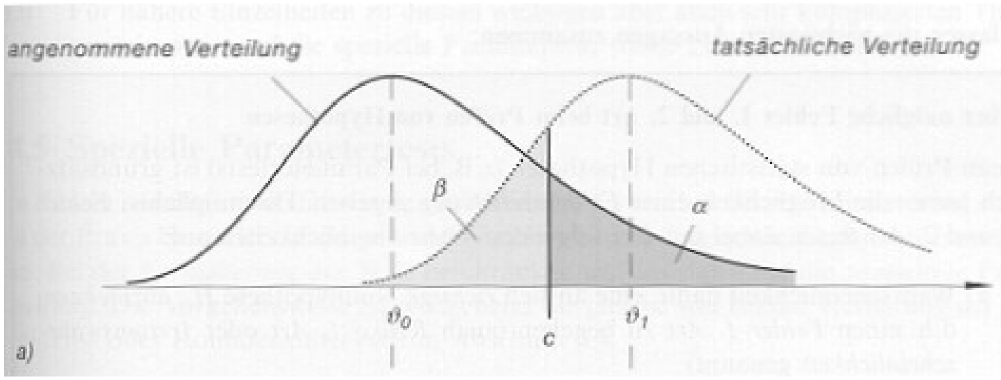
Weicht der tatsächlich vorhandene Parameter ϑ_1 von dem in H_0 angenommenen ϑ_0 ab, so ergibt sich dennoch eine endliche Wahrscheinlichkeit β, einen Wert der Prüfgröße im Annahmebereich für H_0 zu erhalten.

Die Wahrscheinlichkeit für den Fehler 2. Art hängt dabei noch von der Lage des wahren Parameterwertes ϑ_1 ab: $\beta = \beta(\vartheta_1)$

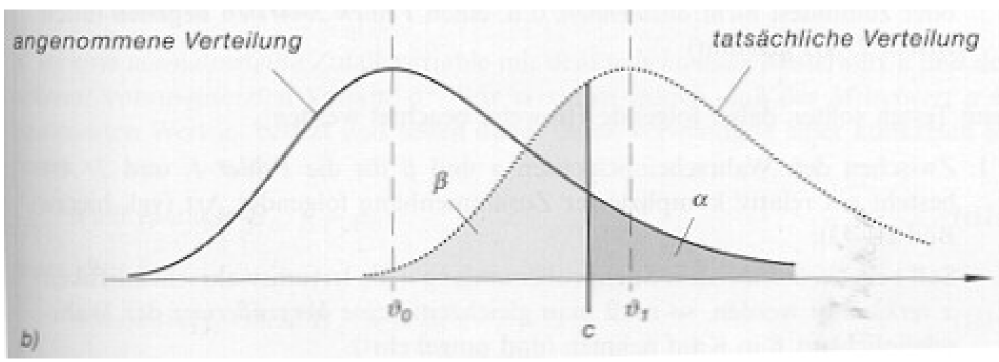
„Operationscharakteristik"

III. Schließende Statistik
III.4 Statistische Testverfahren

Zusammenhang zwischen den Fehlern 1. und 2. Art:



„Hohes" Signifikanzniveau α



„Niedriges" Signifikanzniveau α

Eine Verkleinerung von α zieht automatisch eine Vergrößerung von β nach sich und umgekehrt (bei konstantem Stichprobenumfang)!

$\beta = \beta(\vartheta_1)$ „Operationscharakteristik"
(Annahmewahrscheinlichkeit für H_0)
- Wahrscheinlichkeit für Fehler 2. Art, wenn H_0 eigentlich abzulehnen wäre
- Wahrscheinlichkeit einer richtigen Entscheidung, wenn H_0 zutrifft

$g(\vartheta_1) = 1 - \beta(\vartheta_1)$ „Gütefunktion", auch „Machtfunktion" oder „Power function"
(Ablehnwahrscheinlichkeit für H_0)
- Wahrscheinlichkeit für Fehler 1. Art, wenn H_0 eigentlich zutrifft
- Wahrscheinlichkeit einer richtigen Entscheidung, wenn H_1 zutrifft

III.4.2 Test für eine Wahrscheinlichkeit

In welcher Verteilung für eine diskrete ZV ist die „Wahrscheinlichkeit" ein Parameter?

Typische Fragestellungen:

- Ist die Wahrscheinlichkeit, eine „6" beim einfachen Würfelwurf mit diesem Würfel zu erhalten, 1/6?
- Ist der Anteil fehlerhafter ICs in einer Lieferung höchstens 1%?
- Mit welchem Stichprobenverfahren kann man zwischen Lieferungen 1. Wahl (1% Ausschuss) und 2. Wahl (10% Ausschuss) „sicher" unterscheiden?

III.4.2.1 Einfacher Alternativtest

H_0: $p=p_0$ testen gegen H_1: $p=p_1$ (o.b.d.A. $p_1>p_0$)

Beispiel: Anteil fehlerhafter Schrauben in Verpackungen
$p_0=0{,}15$; $p_1=0{,}40$

Stichprobe des Umfangs n (z.B. n=10)

Ergebnis: N defekte Schrauben in der Stichprobe

Große Werte von N sprechen für H_1, kleine für H_0

$$N \begin{cases} > k \Rightarrow \text{Entscheidung für } H_1 \\ \leq k \Rightarrow \text{Entscheidung für } H_0 \end{cases}$$

k=?

III.4.2.1 Einfacher Alternativtest

$$N \begin{cases} > k \Rightarrow \text{Entscheidung für } H_1 \\ \leq k \Rightarrow \text{Entscheidung für } H_0 \end{cases}$$

k=?

Wahrscheinlichkeit für Fehler 1. Art:

$$\alpha = \alpha(k) = P(N > k | p = p_o) = \sum_{l=k+1}^{n} \binom{n}{l} {p_0}^l (1 - p_0)^{n-l}$$

Wahrscheinlichkeit für Fehler 2. Art:

$$\beta = \beta(k) = P(N \leq k | p = p_1) = \sum_{l=0}^{k} \binom{n}{l} {p_1}^l (1 - p_1)^{n-l}$$

III.4.2.2 Einseitiger Hypothesentest für eine Wahrscheinlichkeit

Verteilungsannahme für die den Beobachtungen zugrunde liegende Verteilung: Binomialverteilung

Testgröße: In einer Stichprobe des Umfangs n gefundene Zahl N ein bestimmtes Merkmal tragende Einheiten, z.B. defekte Schrauben

$$H_0\colon p \leq p_0 \;;\; H_1\colon p > p_0$$

$$N \begin{cases} > k & \Rightarrow H_0 \text{ wird abgelehnt}, H_1 \text{ wird bestätigt} \\ \leq k & \Rightarrow H_0 \text{ wird nicht abgelehnt (aber auch nicht bestätigt)} \end{cases}$$

III.4.2.2 Einseitiger Hypothesentest für eine Wahrscheinlichkeit

$$N \begin{cases} > k & \Longrightarrow H_0 \, wird \, abgelehnt, H_1 \, wird \, bestätigt \\ \leq k \Longrightarrow H_0 \, wird \, nicht \, abgelehnt \, (aber \, auch \, nicht \, bestätigt) & \end{cases}$$

Wie sollte man k („Signifikanzschranke") wählen, um möglichst selten eine falsche Entscheidung zu treffen?

Vorgehensweise: Festlegung des kritischen Wertes k durch Vorgabe einer oberen Schranke für die Fehlerwahrscheinlichkeit 1.Art α:

$\alpha \leq \alpha_0$:Signifikanzniveau des Tests

III.4.2.2 Einseitiger Hypothesentest für eine Wahrscheinlichkeit

$\alpha \leq \alpha_0$:Signifikanzniveau des Tests

$$\alpha(p) = P(N > k|p) = \sum_{l=k+1}^{n} \binom{n}{l} p^l (1-p)^{n-l} \leq \alpha_0 \quad \forall p \leq p_0$$

$\alpha(p)$ist streng monoton wachsend in p

$\Rightarrow \alpha(p) \leq \alpha(p_0) \ \forall p \leq p_0$

Suche kleinstes k, das die folgende Ungleichung erfüllt:

$$\sum_{l=k+1}^{n} \binom{n}{l} {p_0}^l (1-p_0)^{n-l} \leq \alpha_0$$

III.4.2.3 Zweiseitiger Hypothesentest für eine Wahrscheinlichkeit

H_0: $p=p_0$; H_1: $p\neq p_0$

$N > k_2 \vee N < k_1 \Rightarrow H_0$ wird abgelehnt, H_1 bestätigt.
$k_1 \leq N \leq k_2 \Rightarrow H_0$ wird nicht abgelehnt.

Vorgehensweise: Festlegung der kritischen Werte k_1 und k_2 durch Vorgabe einer oberen Schranke für die Fehlerwahrscheinlichkeit 1.Art α:

$$\alpha(p_0) = P(N > k_2 \vee N < k_1 \,|p = p_0) =$$
$$\sum_{l=0}^{k_1-1} \binom{n}{l} {p_0}^l (1-p_0)^{n-l} + \sum_{l=k_2+1}^{n} \binom{n}{l} {p_0}^l (1-p_0)^{n-l} \leq \alpha_0$$

III.4.2.3 Zweiseitiger Hypothesentest für eine Wahrscheinlichkeit

$$\alpha(p_0) = P(N > k_2 \vee N < k_1 \,|p = p_0) =$$

$$\sum_{l=0}^{k_1-1} \binom{n}{l} {p_0}^l (1-p_0)^{n-l} + \sum_{l=k_2+1}^{n} \binom{n}{l} {p_0}^l (1-p_0)^{n-l} \le \alpha_0$$

Verteile die Wahrscheinlichkeit für den Fehler 1. Art symmetrisch auf beide Seiten:

$$\sum_{l=0}^{k_1-1} \binom{n}{l} {p_0}^l (1-p_0)^{n-l} \le \frac{\alpha_0}{2} \text{ mit größtem } \mathrm{k}_1$$

$$\sum_{l=k_2+1}^{n} \binom{n}{l} {p_0}^l (1-p_0)^{n-l} \le \frac{\alpha_0}{2} \text{ mit kleinstem } \mathrm{k}_2$$

III.4.3 Test für den unbekannten Mittelwert μ einer Normalverteilung bei bekannter Varianz σ²

Zweiseitiger Test

$$H_0 : \mu = \mu_0 \quad gegen \quad H_1 : \mu \neq \mu_0$$

$$\text{Test} : |\overline{X} - \mu_0| \begin{cases} > c \Rightarrow H_0 \text{ wird zugunsten von } H_1 \text{ verworfen} \\ \leq c \qquad \Rightarrow H_0 \text{ wird nicht verworfen} \end{cases}$$

1.) Vorgabe des Signifikanzniveaus α

(Wahrscheinlichkeit für Fehler 1.Art)

2.) Berechnung der kritischen Grenzen c unter der Annahme,

dass H_0 gültig ist

$$\alpha = P(\overline{X} - \mu_0 > c) + P(\overline{X} - \mu_0 < -c)$$

$$= P\left(\frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} > \frac{c}{\sigma / \sqrt{n}}\right) + P\left(\frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} < \frac{-c}{\sigma / \sqrt{n}}\right)$$

$u = \dfrac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$ ist $N(0;1)$ – verteilt (Dies ist die geeignete Prüfvariable, s.u.!)

$$\alpha = 1 - \phi\left(\frac{c}{\sigma/\sqrt{n}}\right) + \phi\left(\frac{-c}{\sigma/\sqrt{n}}\right) = 2 - 2\phi\left(\frac{c}{\sigma/\sqrt{n}}\right) \Rightarrow \phi\left(\frac{c}{\sigma/\sqrt{n}}\right) = 1 - \frac{\alpha}{2}$$

$$\Rightarrow \frac{c}{\sigma/\sqrt{n}} = u_{1-\alpha/2} \qquad \boxed{c = u_{1-\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}}$$

3.) Testentscheidung

$|\bar{x} - \mu_0| \le u_{1-\alpha/2} \cdot \dfrac{\sigma}{\sqrt{n}} \Rightarrow H_0 : \mu = \mu_0$ wird nicht verworfen

$|\bar{x} - \mu_0| > u_{1-\alpha/2} \cdot \dfrac{\sigma}{\sqrt{n}} \Rightarrow H_0$ wird zugunsten $H_1 : \mu \ne \mu_0$ verworfen

Andere Formulierung mit einer "Prüfgröße": $\hat{u} = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

$-u_{1-\alpha/2} \leq \hat{u} \leq u_{1-\alpha/2}$ $\Rightarrow H_0 : \mu = \mu_0$ wird nicht verworfen

$|\hat{u}| > u_{1-\alpha/2}$ $\Rightarrow H_0$ wird zugunsten $H_1 : \mu \neq \mu_0$ verworfen

4.) Fehler 2. Art

<u>Definition der „Gütefunktion“ des Tests</u>:
g(μ) ist die Wahrscheinlichkeit, die Nullhypothese H_0 abzulehnen, wenn μ der wahre Parameter der Grundgesamtheit ist.

$$g(\mu) = P(\bar{X} - \mu_0 > c) + P(\bar{X} - \mu_0 < -c)$$
$$= P(\bar{X} > c + \mu_0) + P(\bar{X} < -c + \mu_0)$$

III. Schließende Statistik
III.4 Statistische Testverfahren

$$g(\mu) = P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} > \frac{c+\mu_0-\mu}{\sigma/\sqrt{n}}\right) + P\left(\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < \frac{-c+\mu_0-\mu}{\sigma/\sqrt{n}}\right)$$

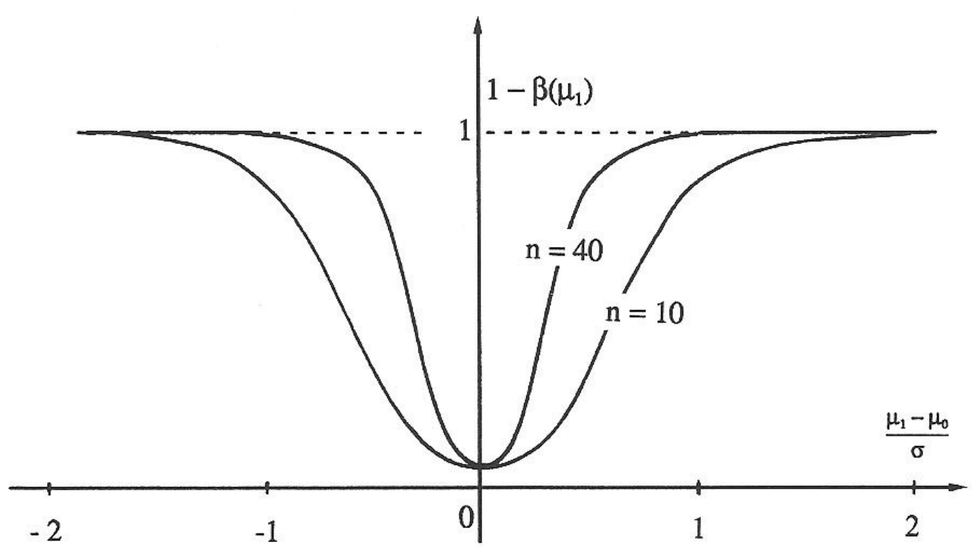
$$= 1-\phi\left(\frac{c+\mu_0-\mu}{\sigma/\sqrt{n}}\right) + \phi\left(\frac{-c+\mu_0-\mu}{\sigma/\sqrt{n}}\right)$$

$$= 1-\phi\left(\frac{c}{\sigma/\sqrt{n}} + \frac{\sqrt{n}}{\sigma}\cdot(\mu_0-\mu)\right) + \phi\left(\frac{-c}{\sigma/\sqrt{n}} + \frac{\sqrt{n}}{\sigma}\cdot(\mu_0-\mu)\right)$$

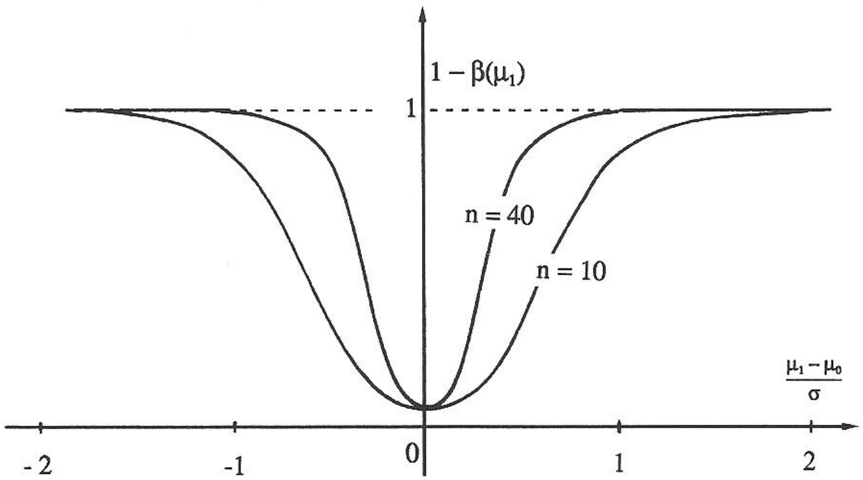
$$= 1-\phi\left(u_{1-\alpha/2} + \frac{\sqrt{n}}{\sigma}\cdot(\mu_0-\mu)\right) + \phi\left(-u_{1-\alpha/2} + \frac{\sqrt{n}}{\sigma}\cdot(\mu_0-\mu)\right)$$

Bei festem Signifikanzniveau α ist g eine Funktion des Abstandes des wahren Mittelwertes μ vom vermuteten μ_0 in “Einheiten” der Standardabweichung σ. Der Stichprobenumfang n ist ein Parameter:

Gütefunktion g(μ) beim zweiseitigen Test für den unbekannten Mittelwert einer Normalverteilung bei bekannter Varianz; Signifikanzniveau α=0,05



„Trennschärfe" wächst mit wachsendem Stichprobenumfang n.



$$g(\mu) \underset{n\to\infty}{\overset{\mu\neq\mu_0}{\longrightarrow}} 1$$

$$g(\mu_0) = 1 - \phi(u_{1-\alpha/2}) + \phi(-u_{1-\alpha/2}) = 1 - (1 - \frac{\alpha}{2}) + 1 - (1 - \frac{\alpha}{2}) = \alpha$$ Fehler 1.Art

Bei ausreichend großen Stichprobenumfängen wird H_0 immer abgelehnt, falls dies berechtigt ist! Mit wachsendem Stichprobenumfang wächst also die Wahrscheinlichkeit, einen Fehler 2. Art zu vermeiden: Der Parametertest ist „konsistent".

<u>Typische Forderung bei der Planung eines Tests</u>:
Berechne den Mindeststichprobenumfang n_0 so, dass damit bei vorgegebenem α für einen Wert $\mu=\mu_1$ die Wahrscheinlichkeit für den Fehler 2. Art höchstens $\beta_0= \beta(\mu_1)$ ist.

Typisch (aber nicht zwingend!) bei technischen Anwendungen: $\alpha = 0{,}01; \quad \beta_0 = 0{,}1$

$$\boxed{n_0 = (u_{1-\alpha/2} + u_{1-\beta_0})^2 \cdot \frac{\sigma^2}{(\mu_1 - \mu_0)^2}} \quad \boxed{n = 14{,}88 \cdot \frac{\sigma^2}{\Delta\mu^2} \approx 15 \cdot \frac{\sigma^2}{\Delta\mu^2}}$$

FH AACHEN UNIVERSITY OF APPLIED SCIENCES

Einseitiger Test: Abgrenzung nach oben

$$H_0 : \mu \le \mu_0 \quad gegen \quad H_1 : \mu > \mu_0$$

$$\text{Test}: \ \overline{X} - \mu_0 \begin{cases} > c \Rightarrow H_0 \text{ wird zugunsten von } H_1 \text{ verworfen} \\ \le c \qquad \Rightarrow H_0 \text{ wird nicht verworfen} \end{cases}$$

1.) Vorgabe des Signifikanzniveaus α

(Wahrscheinlichkeit für Fehler 1. Art)

2.) Berechnung der kritischen Grenze c unter der Annahme, dass H_0 gültig ist

$$1 - \alpha = P(\overline{X} - \mu_0 \le c) = P\left(\frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \le \frac{c}{\sigma / \sqrt{n}} \right); \quad u = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \text{ ist } N(0;1) - \text{verteilt}$$

$$1 - \alpha = \phi\left(\frac{c}{\sigma / \sqrt{n}} \right); \quad \Rightarrow \frac{c}{\sigma / \sqrt{n}} = u_{1-\alpha}$$

$$\boxed{c = u_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}}$$

3.) Testentscheidung

$$\bar{x} - \mu_0 \le u_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}} \Rightarrow H_0 : \mu \le \mu_0$$ wird nicht verworfen

$$\bar{x} - \mu_0 > u_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}} \Rightarrow H_0$$ wird zugunsten $H_1 : \mu > \mu_0$ verworfen

Andere Formulierung mit einer "Prüfgröße": $\hat{u} = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

$\hat{u} \le u_{1-\alpha} \quad \Rightarrow H_0 : \mu \le \mu_0$ wird nicht verworfen

$\hat{u} > u_{1-\alpha} \quad \Rightarrow H_0$ wird zugunsten $H_1 : \mu > \mu_0$ verworfen

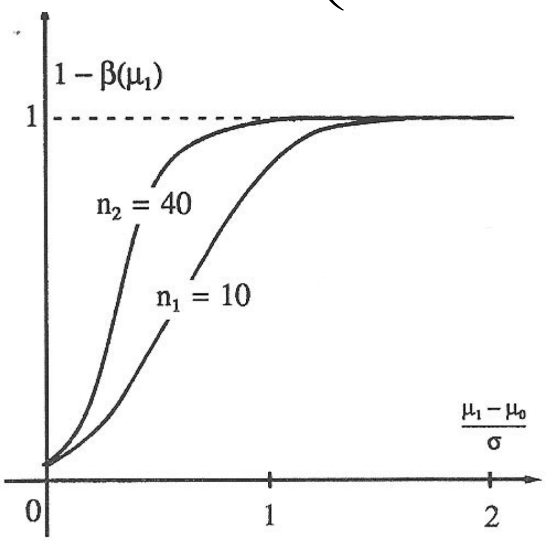
„Gütefunktion“ des Tests:
g(μ) ist die Wahrscheinlichkeit, die Nullhypothese H_0 abzulehnen, wenn μ der wahre Parameter der Grundgesamtheit ist.

III. Schließende Statistik
III.4 Statistische Testverfahren

$$g(\mu) = P(\overline{X} - \mu_0 > c) = P(\overline{X} > c + \mu_0)$$

$$g(\mu) = P\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} > \frac{c + \mu_0 - \mu}{\sigma/\sqrt{n}}\right) = 1 - \phi\left(\frac{c + \mu_0 - \mu}{\sigma/\sqrt{n}}\right)$$

$$= 1 - \phi\left(\frac{c}{\sigma/\sqrt{n}} + \frac{\sqrt{n}}{\sigma}\cdot(\mu_0 - \mu)\right) = 1 - \phi\left(u_{1-\alpha} + \frac{\sqrt{n}}{\sigma}\cdot(\mu_0 - \mu)\right)$$



Gütefunktion g(μ) beim einseitigen Test für den unbekannten Mittelwert einer Normalverteilung bei bekannter Varianz; Signifikanzniveau α=0,05; „Trennschärfe" wächst mit wachsendem Stichprobenumfang n.

III. Schließende Statistik
III.4 Statistische Testverfahren

Bei gleichem Signifikanzniveau und Stichprobenumfang ist die „Macht" (Wert der Gütefunktion) des einseitigen Tests höher als die des zweiseitigen Tests.

Typische Forderung bei der Planung eines Tests:
Berechne den Mindeststichprobenumfang n_0 so, dass damit bei vorgegebenem α für einen Wert $\mu=\mu_1$ die Wahrscheinlichkeit für den Fehler 2. Art höchstens $\beta_0 = \beta(\mu_1)$ ist.

$$\boxed{n_0 = (u_{1-\alpha} + u_{1-\beta_0})^2 \cdot \frac{\sigma^2}{(\mu_1 - \mu_0)^2}}$$

$$\boxed{n = 13{,}02 \cdot \frac{\sigma^2}{\Delta\mu^2} \approx 15 \cdot \frac{\sigma^2}{\Delta\mu^2}}$$

Typisch (aber nicht zwingend!) bei technischen Anwendungen:

$$\alpha = 0{,}01; \quad \beta_0 = 0{,}1$$

Einseitiger Test: Abgrenzung nach unten

Analoge Berechnung wie bei Abgrenzung nach oben

$$H_0 : \mu \geq \mu_0 \quad gegen \quad H_1 : \mu < \mu_0$$

$$\text{Test}: \ \overline{X} - \mu_0 \begin{cases} < c \Rightarrow H_0 \text{ wird zugunsten von } H_1 \text{ verworfen} \\ \geq c \qquad \Rightarrow H_0 \text{ wird nicht verworfen} \end{cases}$$

1.) Vorgabe des Signifikanzniveaus α (Wahrscheinlichkeit für Fehler 1.Art)

2.) Berechnung der kritischen Grenze c unter der Annahme, dass H_0 gültig ist

$$\alpha = P(\overline{X} - \mu_0 \leq c) = P\left(\frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \leq \frac{c}{\sigma / \sqrt{n}}\right); \quad u = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \text{ ist } N(0;1) - \text{verteilt}$$

$$\alpha = \phi\left(\frac{c}{\sigma / \sqrt{n}}\right); \quad \Rightarrow \frac{c}{\sigma / \sqrt{n}} = u_\alpha$$

$$\boxed{c = u_\alpha \cdot \frac{\sigma}{\sqrt{n}}}$$

FH AACHEN UNIVERSITY OF APPLIED SCIENCES

3.) Testentscheidung

$$\bar{x} - \mu_0 \geq u_\alpha \cdot \frac{\sigma}{\sqrt{n}} \Rightarrow H_0 : \mu \geq \mu_0 \text{ wird nicht verworfen}$$

$$\bar{x} - \mu_0 < u_\alpha \cdot \frac{\sigma}{\sqrt{n}} \Rightarrow H_0 \text{ wird zugunsten } H_1 : \mu < \mu_0 \text{ verworfen}$$

Andere Formulierung mit einer "Prüfgröße": $\hat{u} = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

$\hat{u} \geq u_\alpha \quad \Rightarrow H_0 : \mu \geq \mu_0$ wird nicht verworfen

$\hat{u} < u_\alpha \quad \Rightarrow H_0$ wird zugunsten $H_1 : \mu < \mu_0$ verworfen

Zusammenfassung:

H_0	H_1	Kritischer Bereich K bei einem Signifikanzniveau α
$\mu = \mu_0$	$\mu \neq \mu_0$	$\frac{\lvert \bar{x} - \mu_0 \rvert}{\sigma}\sqrt{n} > u_{1-\frac{\alpha}{2}}$
$\mu \geq \mu_0$ $\mu \leq \mu_0$	$\mu < \mu_0$ $\mu > \mu_0$	$\frac{\lvert \bar{x} - \mu_0 \rvert}{\sigma}\sqrt{n} > u_{1-\alpha}$

III.4.4 Test für den unbekannten Mittelwert μ einer Normalverteilung bei unbekannter Varianz σ²

$$H_0 : \mu = \mu_0 \quad gegen \quad H_1 : \mu \neq \mu_0$$

1.) Vorgabe des Signifikanzniveaus α (Wahrscheinlichkeit für Fehler 1.Art)

2.) Berechnung der Test - oder Prüfvariablen

$$T = \frac{\overline{X} - \mu_0}{S / \sqrt{n}} \text{ ist t} - \text{verteilt nach Student mit f} = \text{n - 1 Freiheitsgraden}$$

$$1 - \alpha = P(-c \leq T \leq c) \;\Rightarrow\; c = t_{1-\frac{\alpha}{2};f}$$

3.) Testentscheidung anhand eines konkreten Stichprobenergebnisses $\hat{t} = \dfrac{\overline{x} - \mu_0}{s / \sqrt{n}}$

$$|\,\hat{t}\,| \leq t_{1-\frac{\alpha}{2};f} \Rightarrow H_0 : \mu = \mu_0 \text{ wird nicht verworfen}$$

$$|\,\hat{t}\,| \geq t_{1-\frac{\alpha}{2};f} \Rightarrow H_0 \text{ wird zugunsten } H_1 : \mu \neq \mu_0 \text{ verworfen}$$

Zusammenfassung:
Einseitige Tests analog zu Tests bei bekannter Varianz, ersetze dort u durch t!

H_0	H_1	Kritischer Bereich K bei einem Signifikanzniveau α
$\mu = \mu_0$	$\mu \neq \mu_0$	$\frac{\lvert \bar{x} - \mu_0 \rvert}{s}\sqrt{n} > t_{n-1;1-\frac{\alpha}{2}}$
$\mu \geq \mu_0$ $\mu \leq \mu_0$	$\mu < \mu_0$ $\mu > \mu_0$	$\frac{\lvert \bar{x} - \mu_0 \rvert}{s}\sqrt{n} > t_{n-1;1-\alpha}$

III.4.5 Test für die unbekannte Varianz σ^2 einer Normalverteilung

$$H_0 : \sigma^2 = \sigma_0^2 \quad gegen \quad H_1 : \sigma^2 \neq \sigma_0^2$$

1.) Vorgabe des Signifikanzniveaus α

(Wahrscheinlichkeit für Fehler 1.Art)

2.) Berechnung der Test - oder Prüfvariablen

$$Z = (n-1) \cdot \frac{S^2}{\sigma_0^2} \text{ ist } \chi^2 - \text{verteilt mit f} = \text{n - 1 Freiheitsgraden}$$

$$1-\alpha = P(c_1 \le Z \le c_2) \Rightarrow c_1 = \chi^2_{\frac{\alpha}{2};f} \text{ und } c_2 = \chi^2_{1-\frac{\alpha}{2};f}$$

3.) Testentscheidung anhand eines konkreten Stichprobenergebnisses

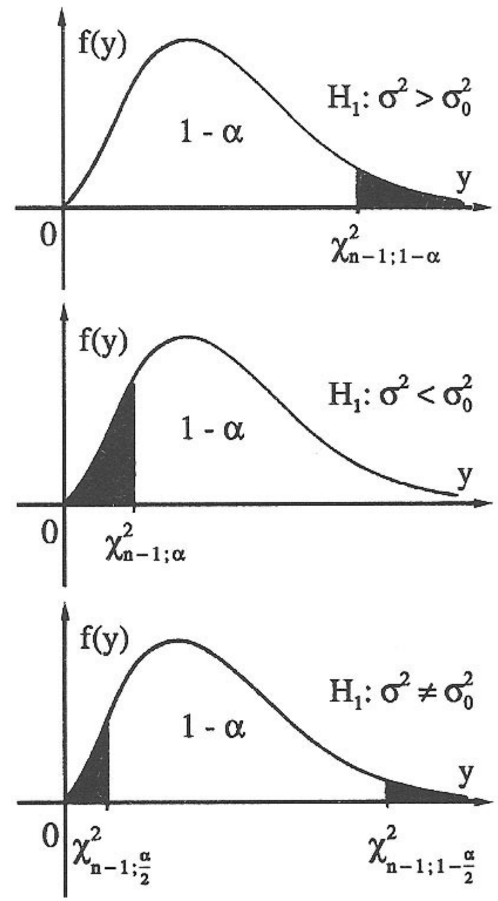
$$\hat{z} = (n-1)\frac{s^2}{\sigma_0^2}$$

$$\chi^2_{\frac{\alpha}{2};f} \le \hat{z} \le \chi^2_{1-\frac{\alpha}{2};f} \qquad \Rightarrow H_0 : \sigma^2 = \sigma_0^2 \text{ wird nicht verworfen}$$

$$\hat{z} < \chi^2_{\frac{\alpha}{2};f} \text{ oder } \hat{z} > \chi^2_{1-\frac{\alpha}{2};f} \Rightarrow H_0 \text{ wird zugunsten } H_1 : \sigma^2 \ne \sigma_0^2 \text{ verworfen}$$

Zusammenfassung:

H_0	H_1	Kritischer Bereich K bei einem Signifikanzniveau α
$\sigma^2 \le \sigma_0^2$	$\sigma^2 > \sigma_0^2$	$\frac{(n-1)s^2}{\sigma_0^2} > \chi^2_{n-1;1-\alpha}$
$\sigma^2 \ge \sigma_0^2$	$\sigma^2 < \sigma_0^2$	$\frac{(n-1)s^2}{\sigma_0^2} < \chi^2_{n-1;\alpha}$
$\sigma^2 = \sigma_0^2$	$\sigma^2 \ne \sigma_0^2$	$\frac{(n-1)s^2}{\sigma_0^2} > \chi^2_{n-1;1-\frac{\alpha}{2}}$ oder $\frac{(n-1)s^2}{\sigma_0^2} < \chi^2_{n-1;\frac{\alpha}{2}}$



III.4.6 Test für einen unbekannten Anteilswert p (Parameter p einer Binomialverteilung)

$$H_0: p = p_0 \quad gegen \quad H_1: p \neq p_0$$

1.) Vorgabe des Signifikanzniveaus α (Wahrscheinlichkeit für Fehler 1. Art)

2.) Schätzfunktion für den unbekannten Parameter p:

$\hat{P} = \dfrac{X}{n}$ Relative Häufigkeit des betrachteten Ereignisses in der Stichprobe

Wegen X binomialverteilt mit $E(X) = np_0$ und $Var(X) = np_0(1-p_0)$:

$$E(\hat{P}) = p_0 \text{ und } Var(\hat{P}) = \frac{p_0(1-p_0)}{n}$$

Für $np_0(1-p_0) > 9$ ist $\hat{P}$ annähernd normalverteilt mit $\mu = p_0$ und $\sigma^2 = \dfrac{p_0(1-p_0)}{n}$

$$U = \frac{\hat{P} - \mu}{\sigma} = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \sqrt{\frac{n}{p_0(1-p_0)}} \cdot (\hat{P} - p_0)$$

ist näherungsweise $N(0;1)$ – verteilt

3.) Testentscheidung

Formulierung mit einer "Prüfgröße": $\hat{u} = \sqrt{\frac{n}{p_0(1-p_0)}} \cdot (\hat{p} - p_0)$

$-u_{1-\alpha/2} \leq \hat{u} \leq u_{1-\alpha/2} \quad \Rightarrow H_0 : p = p_0$ wird nicht verworfen

$|\hat{u}| > u_{1-\alpha/2} \quad \Rightarrow H_0$ wird zugunsten $H_1 : p \neq p_0$ verworfen

Einseitige Tests

Analog zu den Testverfahren für den Mittelwert einer Normalverteilung bei bekannter Varianz!

Bisher: Vergleich eines Stichprobenergebnisses mit einem Verteilungsparameter
Genauso wichtig: Vergleich zweier Stichproben-ergebnisse gegeneinander:

- Vergleich zweier Stichprobenvarianzen: Haben die beiden Grundgesamtheiten, aus denen die Stichproben stammen, dieselbe Varianz?
- Vergleich zweier Stichprobenmittelwerte: Haben die beiden Grundgesamtheiten, aus denen die Stichproben stammen, denselben Mittelwert?
- Vergleich zweier Anteilswerte: Haben die beiden Grundgesamtheiten, aus denen die Stichproben stammen, dieselben Anteilswerte p?

FH AACHEN UNIVERSITY OF APPLIED SCIENCES

III.4.7 Test für die Gleichheit der Varianzen zweier unabhängiger Normalverteilungen

1. Grundgesamtheit: Merkmal X normalverteilt mit $E(X) = \mu_1$ und $Var(X) = \sigma_1^2$	$\rightarrow$	**1. Stichprobe:** Stichprobenumfang $= n_1$ Stichprobenvarianz $= s_1^2$
2. Grundgesamtheit: Merkmal X normalverteilt mit $E(X) = \mu_2$ und $Var(X) = \sigma_2^2$	$\rightarrow$	**2. Stichprobe:** Stichprobenumfang $= n_2$ Stichprobenvarianz $= s_2^2$

$$H_0 : \sigma_1^2 = \sigma_2^2$$

Wann ist der Unterschied zwischen s_1 und s_2 so groß, dass man von unterschiedlichen Werten für σ_1 und σ_2 ausgehen muss?

Die Zufallsgrößen $Z_1 = (n_1 - 1) \cdot \frac{S_1^{\ 2}}{\sigma_1^2}$ und $Z_2 = (n_2 - 1) \cdot \frac{S_2^{\ 2}}{\sigma_2^2}$

sind χ^2 – verteilt mit $f_1 = n_1 - 1$ Freiheitsgraden bzw.

$f_2 = n_2 - 1$ Freiheitsgraden

Gibt es eine Verteilung, die den Zusammenhang zwischen diesen beiden Zufallsvariablen beschreibt?

Die Zufallsvariable $F = \frac{\left(\frac{Z_1}{f_1}\right)}{\left(\frac{Z_2}{f_2}\right)} = \frac{\left(\frac{S_1^2}{\sigma_1^2}\right)}{\left(\frac{S_2^2}{\sigma_2^2}\right)} = \frac{\sigma_2^2 \cdot S_1^2}{\sigma_1^2 \cdot S_2^2}$

genügt einer F - Verteilung mit $f_1 = n_1 - 1$ und $f_2 = n_2 - 1$ Freiheitsgraden

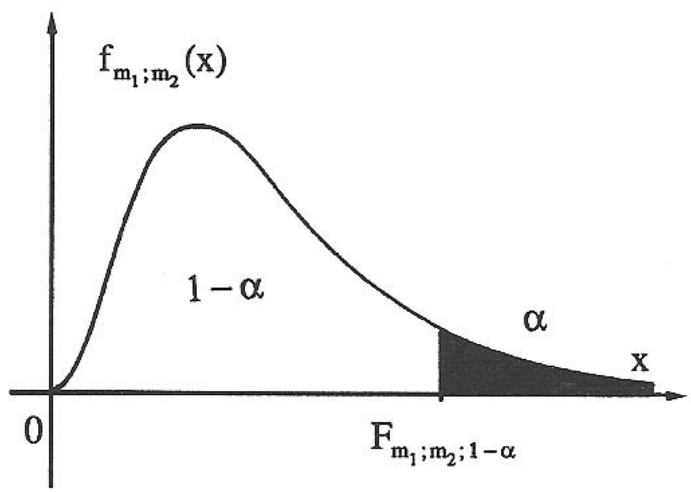
III. Schließende Statistik
III.4 Statistische Testverfahren

Falls $H_0: \sigma_1^2 = \sigma_2^2$ richtig ist, ergibt sich für die neue Zufallsvariable $F = \frac{S_1^2}{S_2^2}$

"Varianzenquotient": Testvariable des Tests auf Gleichheit der Varianzen

F folgt einer "F - Verteilung" mit f_1 und f_2 Freiheitsgraden.

Zur Durchführung des Tests sind wiederum nur die Quantile der F-Verteilung von Interesse (ausführlich tabelliert):

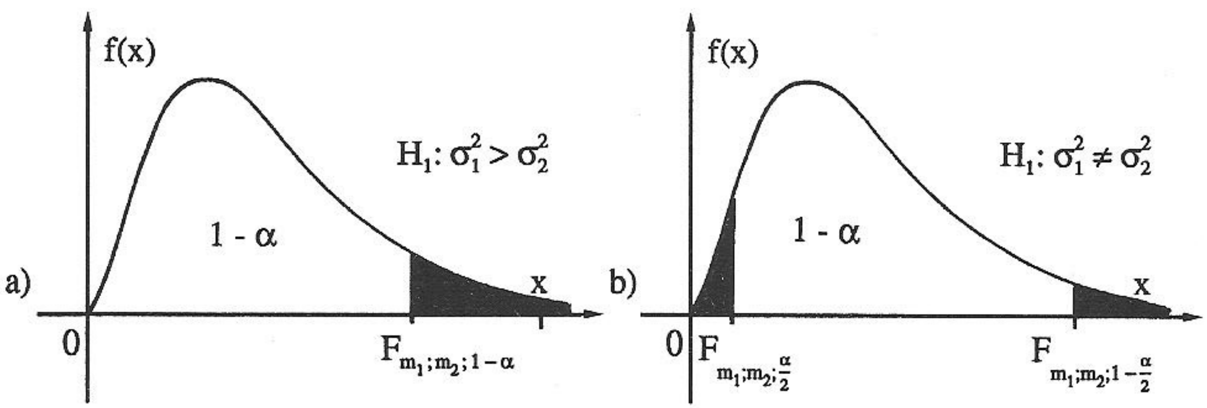


$$F_{f_1;f_2;\alpha} = \frac{1}{F_{f_2;f_1;1-\alpha}}$$

Zusammenfassung:

H_0	H_1	Kritischer Bereich K bei einem Signifikanzniveau α
$\sigma_1^2 \le \sigma_2^2$	$\sigma_1^2 > \sigma_2^2$	$\frac{s_1^2}{s_2^2} > F_{m_1;m_2;1-\alpha}$
$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 \ne \sigma_2^2$	$\frac{s_1^2}{s_2^2} > F_{m_1;m_2;1-\frac{\alpha}{2}}$ oder $\frac{s_1^2}{s_2^2} < F_{m_1;m_2;\frac{\alpha}{2}}$

Wichtig:
Wähle $s_1 > s_2$



Einseitiger
und Zweiseitiger

„F-Test“

III.4.8 Test für die Gleichheit der unbekannten Mittelwerte μ_1 und μ_2 zweier unabhängiger Normal-verteilungen

Stimmen die Mittelwerte zweier normalverteilter Grundgesamtheiten überein oder unterscheiden sie sich signifikant voneinander?

Entnehme jeder Grundgesamtheit eine Zufallsstichprobe (Zufallsvariablen X bzw. Y):

$$x_1; x_2; \ldots x_{n_1} \quad bzw. \quad y_1; y_2; \ldots y_{n_2}$$

(i.a. sind die Stichprobenumfänge n_1 und n_2 ver-schieden!)

Zu testende Hypothesen: $H_0 : \mu_1 = \mu_2; \quad H_1 : \mu_1 \neq \mu_2$

Wichtige Unterscheidung vor dem Test:
„Abhängige" oder „unabhängige" Stichproben?

Definition:
Zwei Stichproben heißen voneinander abhängig, wenn zu jedem Wert der einen Stichprobe genau ein Wert der anderen Stichprobe gehört und umgekehrt:

$$x_i \leftrightarrow y_i \quad (i = 1, 2, \ldots, n)$$

Umkehrbar eindeutige Zuordnungsvorschrift

Zwei abhängige („verbundene") Stichproben können daher nur denselben Umfang n haben!

Differenzentests bei abhängigen Stichproben

Hilfsgröße: $\mu := \mu_1 - \mu_2 \quad \Rightarrow \quad H_0 : \mu = 0 \quad H_1 : \mu \neq 0$

Entsprechende Differenzen der beiden Stichproben:

$z_i = x_i - y_i \; (i = 1, \ldots, n)$

Wenn die Nullhypothese H_0 zutrifft, müssen die z_i aus einer Normalverteilung mit Mittelwert $\mu=0$ kommen!

Analog zum Parametertest für den Mittelwert μ einer normalverteilten Grundgesamtheit!

<u>1. Fall</u>: Die Varianzen der beiden Zufallsvariablen X und Y sind bekannt

<u>1. Fall</u>: Die Varianzen der beiden Zufallsvariablen X und Y sind bekannt

$$Var(X) = \sigma_1^2; \quad Var(Y) = \sigma_2^2$$

$$\Rightarrow \sigma^2 = Var(Z) = Var(X - Y) = Var(X) + Var(Y) = \sigma_1^2 + \sigma_2^2$$

$$\text{Prüfgröße}: \ \hat{u} = \frac{\bar{z} - 0}{\sigma / \sqrt{n}} = \frac{\bar{z}}{\sigma / \sqrt{n}} \quad \text{Test gegen}: \ u_{1-\alpha/2} \ \text{(zweiseitig)}$$

$-u_{1-\alpha/2} \le \hat{u} \le u_{1-\alpha/2}$ $\Rightarrow H_0 : \mu = 0$ wird nicht verworfen

$|\hat{u}| > u_{1-\alpha/2}$ $\Rightarrow H_0$ wird zugunsten $H_1 : \mu \neq 0$ verworfen

Einseitiger Test:
Analog, kritische Schranke für die Prüfgröße wird $u_{1-\alpha}$.

2. Fall: Die Varianzen der beiden Zufallsvariablen X und Y sind unbekannt

$$S^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (Z_i - \overline{Z})^2 \text{ als Schätzung für die Varianz von Z; } f = n-1$$

$$\text{Prüfgröße: } \hat{t} = \frac{\bar{z} - 0}{s/\sqrt{n}} = \frac{\bar{z}}{s/\sqrt{n}} \quad \text{Test gegen } t_{1-\alpha/2;f} \text{ (zweiseitig)}$$

$$|\hat{t}| \leq t_{1-\frac{\alpha}{2};f} \Rightarrow H_0 : \mu = 0 \text{ wird nicht verworfen}$$

$$|\hat{t}| \geq t_{1-\frac{\alpha}{2};f} \Rightarrow H_0 \text{ wird zugunsten } H_1 : \mu \neq 0 \text{ verworfen}$$

Einseitiger Test:
Analog, kritische Schranke für die Prüfgröße wird $t_{1-\alpha}$.

Differenzentests bei unabhängigen Stichproben

2 Stichproben: $x_1, x_2, \ldots x_{n_1} \quad bzw. \quad y_1, y_2, \ldots y_{n_2}$

(i.a. sind die Stichprobenumfänge n_1 und n_2 verschieden!)

$$H_0 : \mu_1 = \mu_2 \quad H_1 : \mu_1 \neq \mu_2$$

<u>1. Fall</u>: Die Varianzen der beiden Zufallsvariablen X und Y sind bekannt

$$Var(X) = \sigma_1^2; \quad Var(Y) = \sigma_2^2 \quad \Rightarrow \quad Var(\overline{X}) = \frac{\sigma_1^2}{n_1}; \quad Var(\overline{Y}) = \frac{\sigma_2^2}{n_2}$$

Bei Gültigkeit von H_0 :

Die Zufallsvariable $Z = \overline{X} - \overline{Y}$ ist normalverteilt mit $\mu = 0$

und Varianz $\sigma^2 = \dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}$

Prüfgröße : $\hat{u} = \dfrac{\overline{x} - \overline{y}}{\sigma}$

Test gegen : $u_{1-\alpha/2}$ (zweiseitig) bzw. $u_{1-\alpha}$ (einseitig)

FH AACHEN UNIVERSITY OF APPLIED SCIENCES

<u>2. Fall</u>: Die Varianzen der beiden Zufallsvariablen X und Y sind gleich, aber unbekannt

$$S_1^{\;2} = \frac{1}{n_1 - 1} \cdot \sum_{i=1}^{n_1} (X_i - \overline{X})^2 \text{ als Schätzung für die Varianz } \sigma_1^2 \text{ von X}$$

$$S_2^{\;2} = \frac{1}{n_2 - 1} \cdot \sum_{i=1}^{n_2} (Y_i - \overline{Y})^2 \text{ als Schätzung für die Varianz } \sigma_2^2 \text{ von Y}$$

Wenn beide Varianzen als gleich vorausgesetzt werden können,

Zusammenfassung :

$$S^2 = \frac{(n_1 - 1) \cdot S_1^2 + (n_2 - 1) \cdot S_2^2}{n_1 + n_2 - 2} \quad \text{als Schätzung für } \sigma^2$$

$$Var(\overline{X}) = \frac{\sigma^2}{n_1}; \quad Var(\overline{Y}) = \frac{\sigma^2}{n_2}$$

III. Schließende Statistik
III.4 Statistische Testverfahren

$$\Rightarrow Var(Z) = Var(\overline{X} - \overline{Y}) = \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2} = \sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right) = \sigma^2 \cdot \frac{n_2 + n_1}{n_1 \cdot n_2}$$

Prüfgröße:

$$\hat{t} = \frac{\overline{x} - \overline{y}}{s \cdot \sqrt{\dfrac{n_2 + n_1}{n_1 \cdot n_2}}} = \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} \cdot \frac{\overline{x} - \overline{y}}{\sqrt{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}}$$

Test gegen $t_{1-\alpha/2;f}$ (zweiseitig) bzw. $t_{1-\alpha;f}$ und $t_{\alpha;f}$ (einseitig)

III.4.9 Test für die Gleichheit von Anteilswerten zweier unabhängiger Grundgesamtheiten

1. zweistufige Grundgesamtheit: Anteilswert p_1	→	**Stichprobe vom Umfang n_1:** Anteilswert $\hat{p}_1$
2. zweistufige Grundgesamtheit: Anteilswert p_2	→	**Stichprobe vom Umfang n_2:** Anteilswert $\hat{p}_2$

$$H_0 : p_1 = p_2 = p$$

Wann ist der Unterschied zwischen $\hat{p}_1$ und $\hat{p}_2$ so groß, dass man von unterschiedlichen Werten für p_1 und p_2 ausgehen muss?

III. Schließende Statistik
III.4 Statistische Testverfahren

Falls H_0 gültig:

Für $n_{1/2}p(1-p)>9$ ist $\hat{P}_{1/2}$ annähernd normalverteilt

mit $\mu_{1/2}=p$ und $\sigma_{1/2}{}^2=\dfrac{p(1-p)}{n_{1/2}}$

$$\Rightarrow E(\hat{P}_2-\hat{P}_1)=0 \text{ und } Var(\hat{P}_2-\hat{P}_1)=Var(\hat{P}_1)+Var(\hat{P}_2)$$

Der unbekannte Anteil p wird als "gewogenes" arithmetisches Mittel geschätzt:

$$\hat{p}=\frac{n_1\cdot\hat{p}_1+n_2\cdot\hat{p}_2}{n_1+n_2} \quad \text{Zusammenfassung beider Stichproben}$$

Die Zufallsgröße $$U = \frac{\hat{P}_2 - \hat{P}_1}{\sqrt{\hat{P}(1-\hat{P}) \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

ist dann näherungsweise standardnormalverteilt.

Testentscheidung:

H_0	H_1	Kritischer Bereich K bei einem Signifikanzniveau α
$p_1 = p_2$	$p_1 \neq p_2$	$\frac{\lvert \hat{p}_2 - \hat{p}_1 \rvert}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} > u_{1-\frac{\alpha}{2}}$
$p_1 \geq p_2$ $p_1 \leq p_2$	$p_1 < p_2$ $p_1 > p_2$	$\frac{\lvert \hat{p}_2 - \hat{p}_1 \rvert}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} > u_{1-\alpha}$

Bisher:

Hypothesenprüfungen waren Parametertests für eine vorausgesetzte Verteilung

Nun:

Prüfen von Hypothesen über die Art des Verteilungsgesetzes: „Anpassungstest"

Nullhypothese: Das Merkmal X in der Grundgesamtheit wird durch die Verteilungsfunktion F_0 beschrieben: $H_0 : F(x) = F_0(x)$

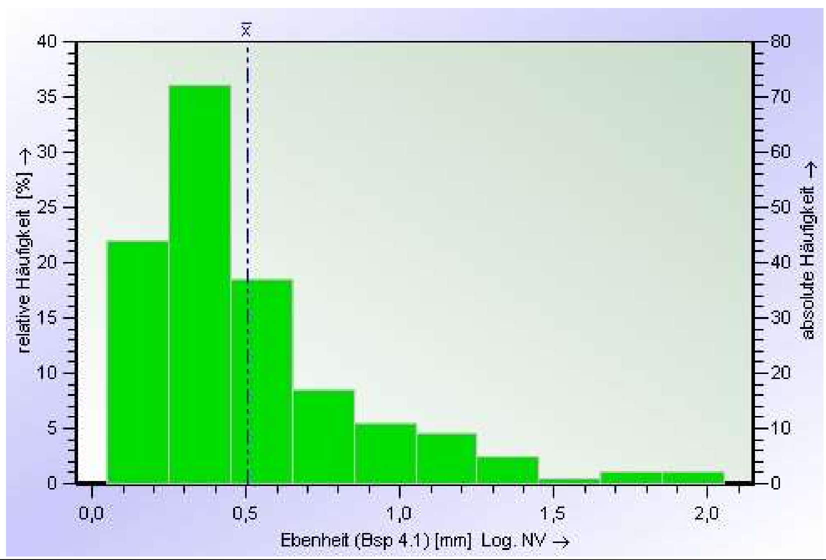
Alternativhypothese: $H_1 : F(x) \neq F_0(x)$

III.4.10 χ^2 – Anpassungs test

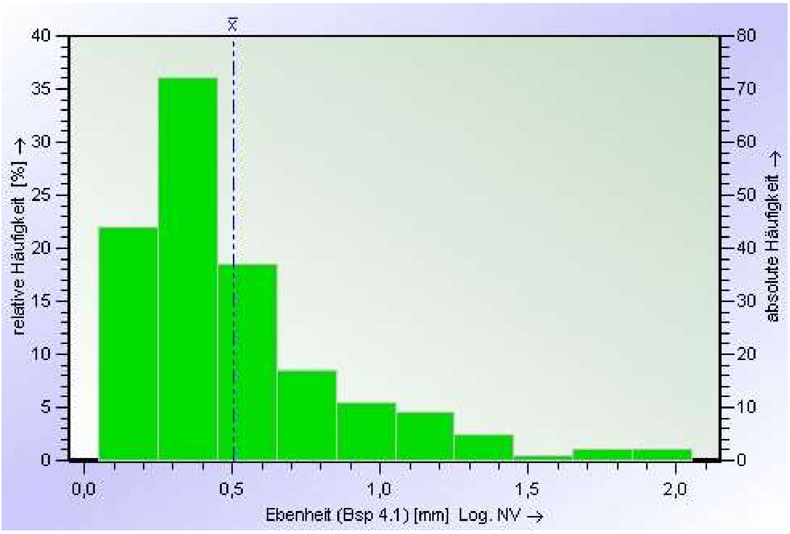
Testsituation :

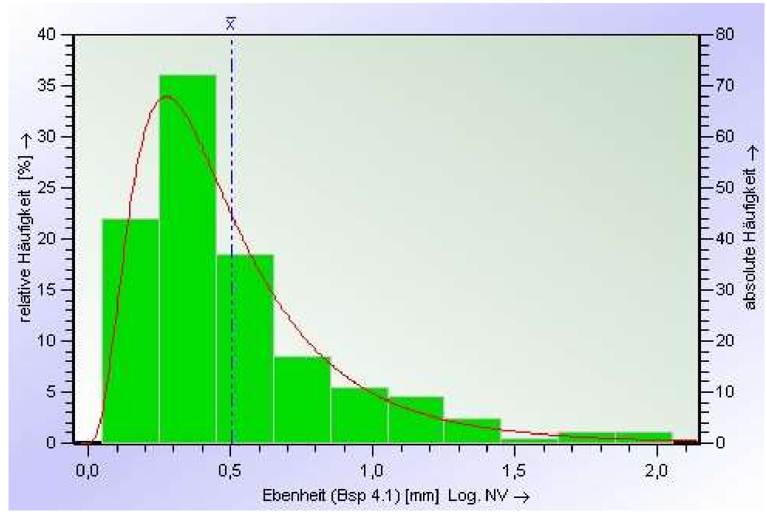
Stichprobe des Umfangs n; unterteilt in $i = 1 \ldots k$ disjunkte Merkmalsklassen A_i; Klassenmitten $\bar{x}_i$; in den Klassen beobachtete absolute Häufigkeiten n_i

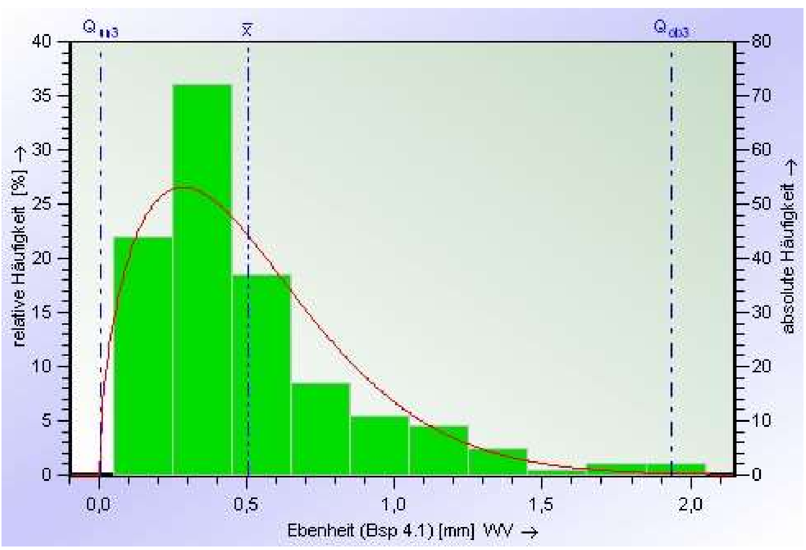
Beispiel:



III. Schließende Statistik
III.4 Statistische Testverfahren







Welche Wahrscheinlichkeitsdichte „passt"?

Die beobachteten Häufigkeiten n_i werden den theoretisch zu erwartenden Häufigkeiten

$$\pi_i = n \cdot p_i$$

gegenübergestellt, wobei $p_i = P(X \in A_i \mid H_0)$, d.h. die Wahrscheinlichkeit, bei richtiger Nullhypothese, einen Merkmalswert in der i-ten Merkmalsklasse zu erhalten.

III. Schließende Statistik
III.4 Statistische Testverfahren

$\bar{x}_i$	n_i	$\pi_i = n \cdot p_i$	$y_i^2 = \frac{(n_i - \pi_i)^2}{\pi_i}$
⋮	⋮	⋮	⋮
	$\sum_{i=1}^{k} n_i = n$	$\sum_{i=1}^{k} \pi_i = n$	$y^2 = \sum_{i=1}^{k} y_i^2$

Prüffunktion:

$$Y^2 = \sum_{i=1}^{k} \frac{(n_i - \pi_i)^2}{\pi_i}$$

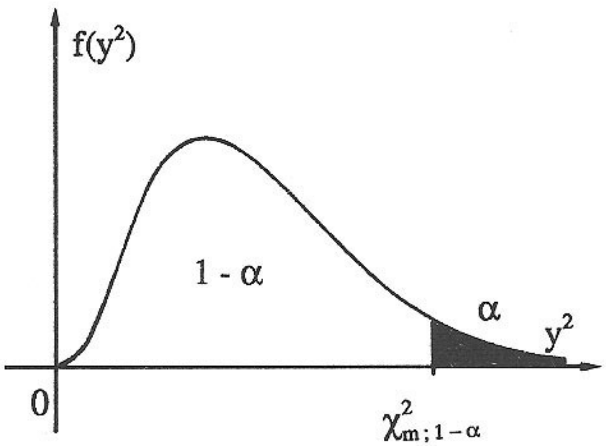
Y^2 ist mit $k - r - 1$ Freiheitsgraden asymptotisch χ^2 - verteilt:

k : Anzahl der Merkmalsklassen

r : Anzahl der aus der Stichprobe zu schätzenden Parameter

Asymptotisch: $\pi_i \geq 5 \; \forall i$ und $n \geq 50$

Die Anpassung an die vermutete Verteilung ist gut, wenn Y^2 nicht zu große Werte annimmt:



Testentscheidung:

H_0	H_1	Kritischer Bereich K bei einem Signifikanzniveau α
$F(x) = F_0(x)$	$F(x) \neq F_0(x)$	$y^2 = \sum_{i=1}^{k} \frac{(n_i - \pi_i)^2}{\pi_i} > \chi^2_{k-r-1;1-\alpha}$

FH Aachen
Fachbereich Medizintechnik und Technomathematik
Prof. Dr. Horst Schäfer
Heinrich-Mußmann-Str. 1
52428 Jülich
T +49. 241. 6009 53927
horst.schaefer@fh-aachen.de
www.fh-aachen.de