

- Bei parametrische Test wird davon ausgegangen, dass das Merkmal (z.B. Erwartungswert) in der oder den Grundgesamtheiten normalverteilt ist („normalverteilte Grundgesamtheiten“).
- Nichtparametrische (verteilungsfreie) Tests setzen keine Verteilungsannahme voraus.

Parametrische Einstichproben tests

Gauß-Test	Vergleich Erwartungswert mit vorgegebenem Wert bei bekannter Varianz oder bei unbekannter Varianz und großen Stichproben
t-Test	Vergleich Erwartungswert mit vorgegebenem Wert bei unbekannter Varianz und kleinen Stichproben
Anteilswerttest	Vergleich Wahrscheinlichkeit (bzw. Anteilswert) mit vorgegebenem Wert (Annäherung mit Normalverteilung)
Varianztest	Vergleich Varianz mit vorgegebenem Wert

Parametrische Zweistichprobentests

Doppelter Gauß-Test	Vergleich zweier Erwartungswerte bei bekannten Varianzen
Doppelter t-Test	Vergleich zweier Erwartungswerte bei unbekanntem, aber gleich großen Varianzen
Anteilswerttest	Vergleich von Wahrscheinlichkeiten (bzw. Anteilswerten) zweier Grundgesamtheiten
F-Test	Vergleich zweier Varianzen (Test auf Gleichheit von Varianzen)

Welch-Test	Vergleich zweier Erwartungswerte bei unbekanntem, aber verschieden großen Varianzen (Näherungslösung); hier nicht behandelt!
------------	---

Nichtparametrische Tests

Chi-Quadrat
Anpassungstest Vergleich einer diskreten Verteilung mit vorgegebener Verteilung

Chi-Quadrat
Unabhängigkeits-
test Test auf Zusammenhänge zwischen 2 Merkmalen

U-Test Test auf Unterschiede zwischen Stichproben; **hier nicht behandelt!**

Übersicht - Lehrveranstaltung

1. Einfache Tests bei Binomialverteilung
 2. Gauß-Test
 3. t-Test
 4. Anteilswerttest
 5. Varianztest
 6. Kontingenztafel, 4-Felder Matrix
 7. Chi-Quadrat-Unabhängigkeitstest
 8. Chi-Quadrat-Anpassungstest
 9. Regressionsanalyse/-test
-
10. p-Wert

1


Hypothesentest (Signifikanztest) Binomialverteilung

Fallbeschreibung

Ein Unternehmen lässt Überraschungseier mit WM-Figuren herstellen und verspricht, dass in mindestens der Hälfte seiner Ü-Eier eine Fußballfigur enthalten ist. Peter, ein Master-Student der Wirtschaftswissenschaften, kauft 50 Ü-Eier und findet in 20 davon WM-Figuren. Er möchte nun die Behauptung des Unternehmens (Signifikanzniveau $\alpha = 5\%$) testen.

Testgröße T: „Anzahl WM-Figuren in Überraschungseiern“
T ist eine $B(50;0,5)$ -verteilte Zufallsgröße

$$H_0: p \geq 0,5$$

$$H_1: p < 0,1$$


Testart: einseitiger (linksseitiger) Signifikanztest; SN $\alpha = 5\%$

Ermittlung kritischer Wert k aus summierter Binomial-Tabelle

1

Hypothesentest (Signifikanztest) Binomialverteilung

Hinweis: k so, dass Wahrscheinlichkeit $\leq \alpha$



k	p 0,01	0,05	0,1	0,2	0,25	0,3	0,4	0,5
0	0,6050	0,0769	0,0052	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,9106	0,2794	0,0338	0,0002	0,0000	0,0000	0,0000	0,0000
...								
15			1,0000	0,9692	0,8369	0,5692	0,0955	0,0033
16				0,9856	0,9017	0,6839	0,1561	0,0077
17				0,9937	0,9449	0,7822	0,2369	0,0164
18				0,9975	0,9713	0,8594	0,3356	0,0325
19				0,9991	0,9861	0,9152	0,4465	0,0595

$$P(X \leq 18) = 0,0325 < \alpha = 0,05$$

Annahmebereich H_0 : [19;50]

Ablehnungsbereich H_0 : [0; 18]

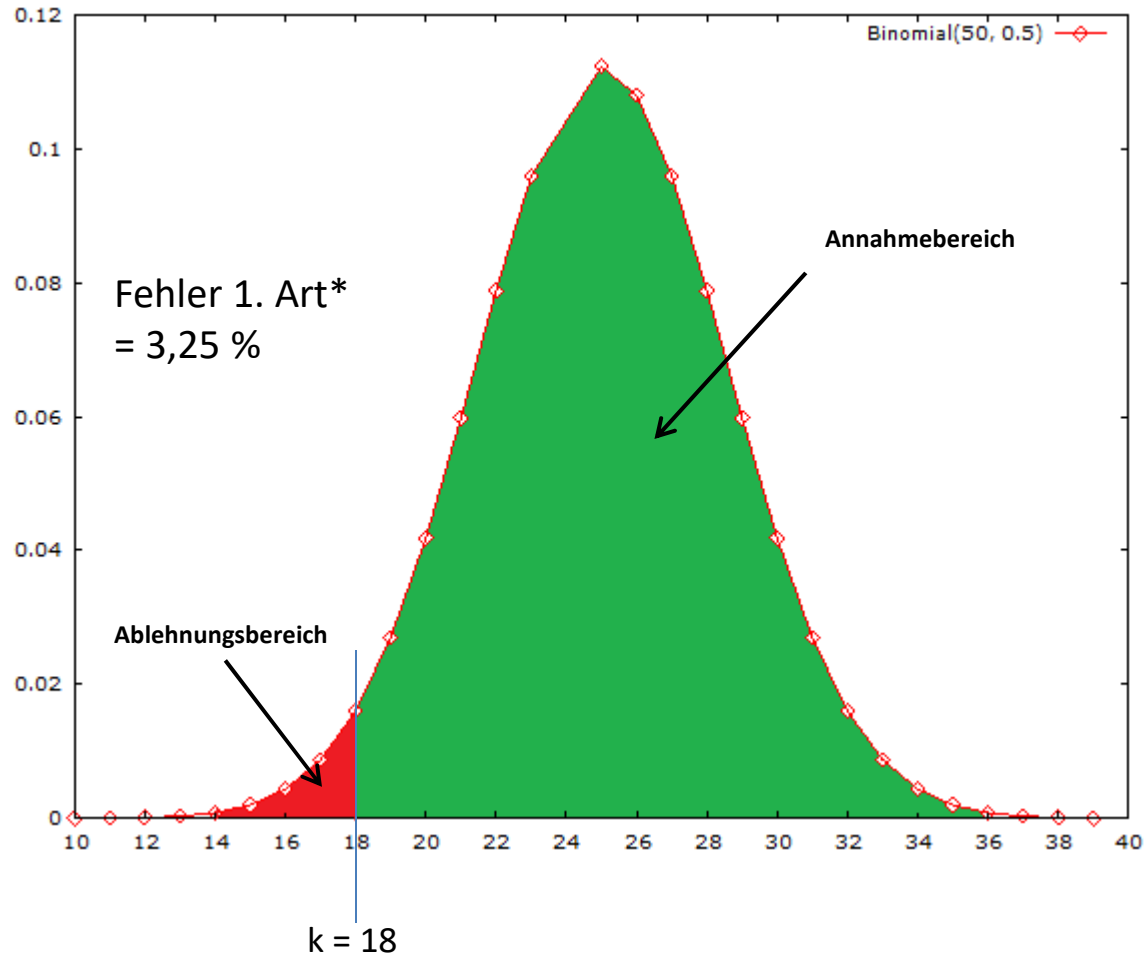
Bei linksseitigem Test gehört k immer zum Ablehnungsbereich

→ H_0 wird angenommen; also Aussage des Herstellers wird angenommen!

→ Die Abweichung ist zufällig und nicht signifikant!

1

Hypothesentest (Signifikanztest) Binomialverteilung



*Fehler 1. Art: H_0 wird abgelehnt, obwohl H_0 richtig ist

1

Hypothesentest (Signifikanztest) Binomialverteilung

Fallbeschreibung

Auf vielfache Nachfrage bietet das Parkrestaurant mehr fleischlose Gerichte als früher an. Durch einen Test ($\alpha = 5\%$) soll herausgefunden werden, ob sich dadurch der Anteil der verkauften fleischlosen Gerichte gegenüber bisher (bis zu 30%) erhöht hat. Hierzu werden die Essensbestellungen von 50 zufällig ausgewählten Gästen ausgewertet; hierunter befinden sich 20 Bestellungen fleischloser Gerichte.

Testgröße T: „Anzahl der Personen, die fleischlose Gerichte bestellen“
T ist eine $B(50;0,3)$ -verteilte Zufallsgröße

$$H_0: p \leq 0,3$$

$$\xrightarrow{\hspace{1cm}} H_1: p > 0,3$$

Testart: einseitiger (rechtsseitiger) Signifikanztest; SN $\alpha = 5\%$

Ermittlung k-Wert aus Tabelle summierte Binomialverteilung

1

Signifikanztest - Binomialverteilung

Hinweis: k so, dass Wahrscheinlichkeit $\geq 1 - \alpha = 0,950$



k	p 0,01	0,05	0,1	0,2	0,25	0,3	0,4	0,5
0	0,6050	0,0769	0,0052	0,0000	0,0000	0,0000	0,0000	0,0000
1	0,9106	0,2794	0,0338	0,0002	0,0000	0,0000	0,0000	0,0000
...								
18				0,9975	0,9713	0,8594	0,3356	0,0325
19				0,9991	0,9861	0,9152	0,4465	0,0595
20				0,9997	0,9937	0,9522	0,5610	0,1013
21				0,9999	0,9974	0,9749	0,6701	0,1611



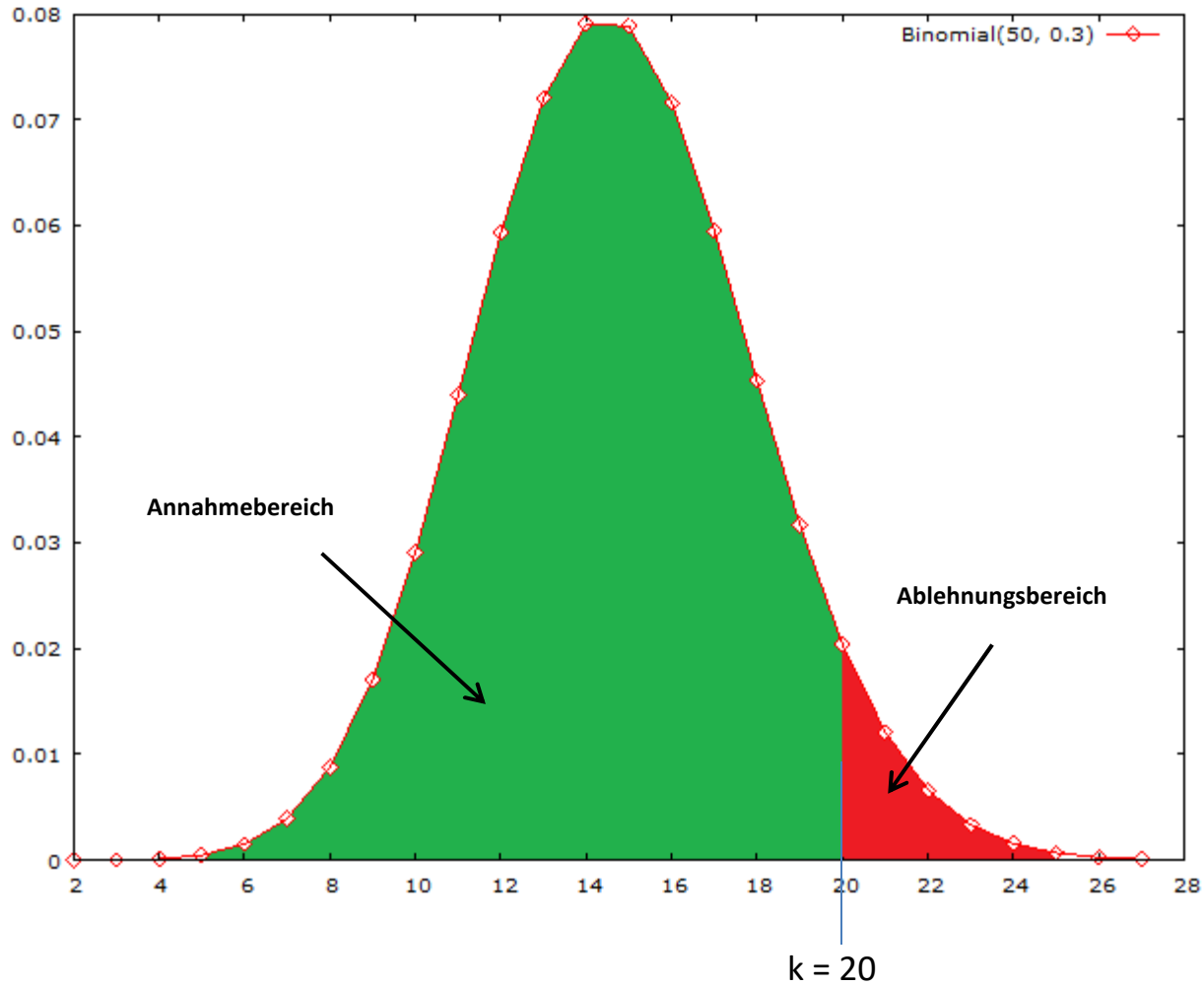
Annahmebereich H_0 : [0;20]
Ablehnungsbereich H_0 : [21; 50]

Bei rechtsseitigem Test gehört k immer zum Annahmebereich

→ H_0 wird angenommen; der Anteil fleischloser Gerichte ist nicht gestiegen!

1

Hypothesentest (Signifikanztest) Binomialverteilung



1

Ermittlung des kritischen Wertes k

B (n;p)	Linksseitiger Test	Rechtsseitiger Test
p bis 50 %	$k: \leq \alpha$	$k: \geq 1-\alpha$
p größer 50 %	$k: \geq 1-\alpha$ (k rechts ablesen)	$k: \leq \alpha$ (k rechts ablesen)

k	p 0,01	0,05	0,1	0,2	0,25	0,3	0,4	0,5	k	n
...										
9		0,9998	0,9755	0,4437	0,1637	0,0402	0,0008	0,0000	40	50
10		1,0000	0,9906	0,5836	0,2622	0,0789	0,0022	0,0000	39	
11			0,9968	0,7107	0,3816	0,1390	0,0057	0,0000	38	
12			0,9990	0,8139	0,5110	0,2229	0,0133	0,0002	37	
13			0,9997	0,8894	0,6370	0,3279	0,0280	0,0005	36	
14			0,9999	0,9393	0,7481	0,4468	0,0540	0,0013	35	
15			1,0000	0,9692	0,8369	0,5692	0,0955	0,0033	34	
16				0,9856	0,9017	0,6839	0,1561	0,0077	33	
17				0,9937	0,9449	0,7822	0,2369	0,0164	32	
18				0,9975	0,9713	0,8594	0,3356	0,0325	31	
19				0,9991	0,9861	0,9152	0,4465	0,0595	30	
20				0,9997	0,9937	0,9522	0,5610	0,1013	29	
21				0,9999	0,9974	0,9749	0,6701	0,1611	28	
...										
25					1,0000	0,9991	0,9427	0,5561	24	
26						0,9997	0,9686	0,6641	23	
27						0,9999	0,9840	0,7601	22	
28						1,0000	0,9924	0,8389	21	
38								1,0000	11	
k	p 0,99	0,95	0,9	0,8	0,75	0,7	0,6	0,5	k	n

z.B.
 $\alpha = 0,05$
 $p = 0,3$
 $p = 0,6$

1

Hypothesentest (Signifikanztest) Binomialverteilung

Fallbeschreibung

Ein Teilnehmer an einem Glücksspiel vermutet, dass der bei dem Spiel verwendete Würfel kein Laplace-Würfel ist. Um dieser Vermutung nachzugehen, würfelt er 100-mal und bestimmt die Anzahl Sechser (12 x Sechs). Ist der Würfel mit einem Signifikanzniveau von 5 % ein Laplace-Würfel?

$$H_0: p = 1/6$$

$$H_1: p \neq 1/6$$

Behauptung
„fairer Würfel“

Vermutung kein
„fairer Würfel“

Testart: beidseitiger Signifikanztest; SN $\alpha = 5\%$

Ermittlung k-Werte aus Tabelle summierte Binomialverteilung

1 Hinweis

Signifikanztest - Binomialverteilung

k_1 (linker Grenzwert) so, dass Wahrscheinlichkeit $\leq \alpha = 0,05$

k_2 (rechts Grenzwert) so, dass Wahrscheinlichkeit $\geq 1 - \alpha/2 = 0,975$



n	k	p 0,01	0,05	0,1	1/6	0,2	0,25	0,3	0,5	k	n
100	...										100
	8		0,9369	0,3209	0,0095	0,0009	0,0000	0,0000	0,0000	91	
	9		0,9718	0,4513	0,0213	0,0023	0,0000	0,0000	0,0000	90	
	...										
	23			1,0000	0,9621	0,8109	0,3711	0,0755	0,0000	76	
	24					0,9783	0,8686	0,4617	0,1136	0,0000	
25					0,9881	0,9125	0,5535	0,1631	0,0000	74	
n	k	p 0,99	0,95	0,9	5/6	0,8	0,75	0,7	0,5	k	n

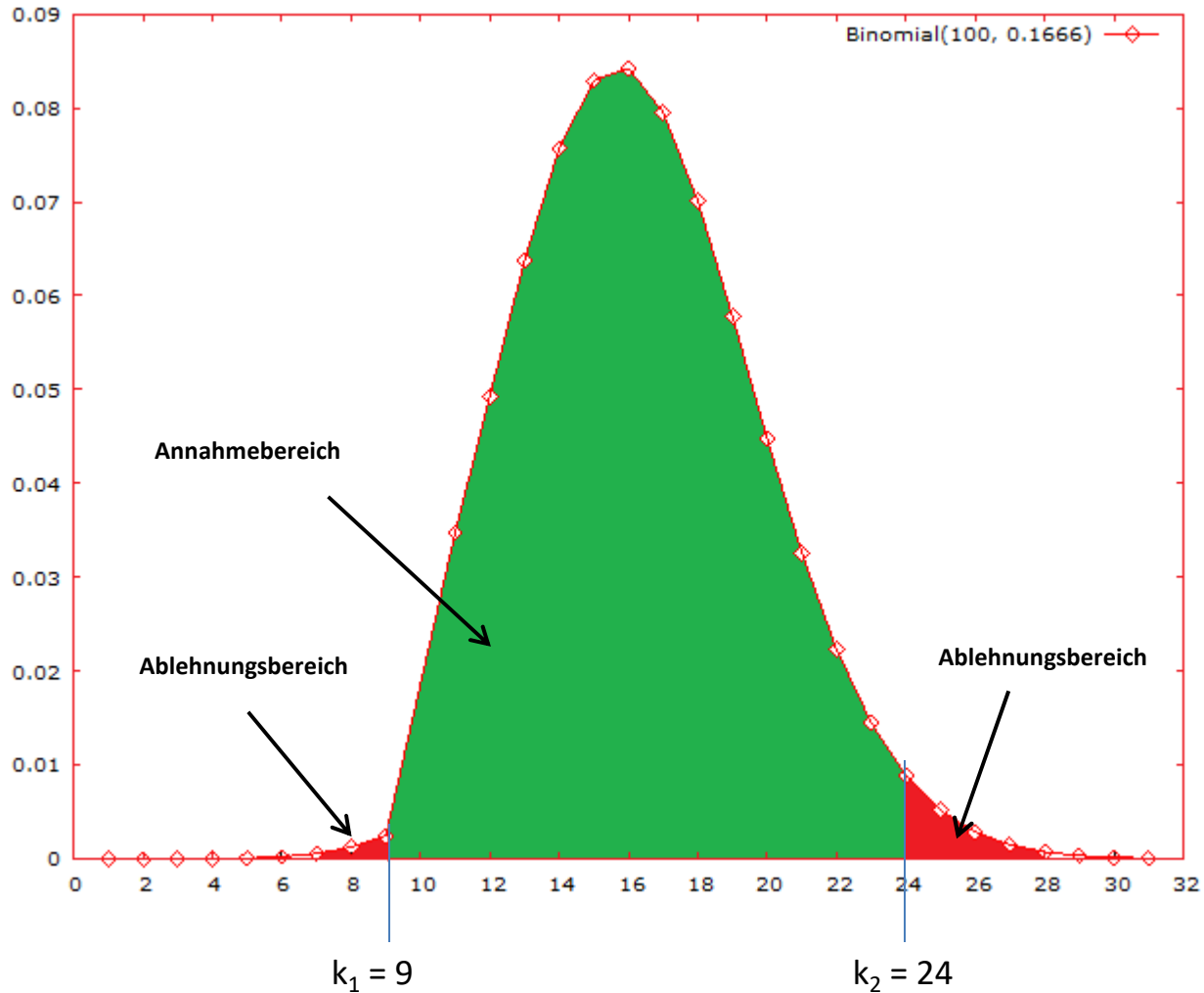
Annahmebereich H_0 : [10;24]

Ablehnungsbereich H_0 : [0; 9] [25; 100]

→ H_0 wird angenommen (12 Sechser bei 100 Würfeln); der Würfel ist fair, d.h. ein Laplace-Würfel

1

Hypothesentest (Signifikanztest) Binomialverteilung



1

Hypothesentest Binomialverteilung

To do

Fallbeschreibung

Die Polizei glaubt durch Schilder mit der Aufschrift „RADARKONTROLLE“ den Anteil der Temposünder vor der Grundschule einer Stadt auf 2 % reduzieren zu können. Die Eltern bezweifeln dies. In einem Test werden 100 Fahrzeuge kontrolliert. Geben Sie die Testgröße sowie die Nullhypothese an und ermitteln Sie deren größtmöglichen Ablehnungsbereich auf einem 5 %-igen Signifikanzniveau.

Fallbeschreibung

Ein Hersteller von Chips für Einkaufswagen garantiert, dass der Anteil an Ausschuss höchstens 10 % beträgt. Ein Käufer findet unter 100 Chips 15 defekte. Kann er hieraus auf einem Signifikanzniveau von 5 % schließen, dass der Anteil an Ausschuss größer als 10 % ist?

Fallbeschreibung

Die Befragung der Studenten einer Hochschule zeigt im vorherigen Jahr, dass höchstens 5 % der befragten Studenten mit der Uni-Bibliothek unzufrieden waren. Man vermutet, dass die Unzufriedenheit zugenommen hat ($n = 50$; $\alpha = 5\%$).

Fallbeschreibung

Eine Münze wird 50-mal geworfen, dabei tritt 30-mal Zahl auf. Kann man mit einer Irrtumswahrscheinlichkeit von 5% schließen, dass die Münze nicht ideal (fair) ist?

2

Parametrischer Test - Einstichprobentest (hier: Gauß-Test über Mittelwert)

Voraussetzungen/Merkmale beim Gauß-Test

- Stichprobengröße $n \leq 30$
 - bekannte Varianz in der Grundgesamtheit
 - Normalverteilung des Merkmals in der Grundgesamtheit
- Stichprobengröße $n > 30$
 - Varianz in der Grundgesamtheit muss nicht bekannt sein
 - Merkmals in der Grundgesamtheit braucht nicht normalverteilt sein

→ Siehe folgende Arbeitsschritte zur Durchführung des Gauß-Tests

2

Parametrischer Test - Einstichprobentest (Arbeitsschritte Gauß-Test über Mittelwert)

1. Signifikanzniveau und Nullhypothese

Signifikanzniveau $\alpha = ?$

Nullhypothese:

- einseitig: $H_0: \mu \leq (\geq) \mu_0$
- zweiseitig: $H_0: \mu = \mu_0$

2. Prüfgröße und ihre Verteilung

Prüfgröße (empirischer Wert):

- bekannte Varianz:

$$Z_0 = \frac{\bar{X} - \mu_0}{\sigma_{\bar{x}}} \text{ mit } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

- unbekannte Varianz und $n > 30$

$$Z_0 = \frac{\bar{X} - \mu_0}{S_{\bar{x}}} \text{ mit } S_{\bar{x}} = \frac{S}{\sqrt{n}}$$

Z_0 folgt einer $N(0;1)$ -Verteilung

3. Kritischer Wert

Kritischer Wert beim:

- einseitigen Test: $z_{1-\alpha}$
- zweiseitigen Test: $z_{1-\alpha/2}$

4. Testentscheidung

- einseitiger Test:

$|Z_0| \leq z_{1-\alpha} : H_0$ annehmen

$|Z_0| > z_{1-\alpha} : H_0$ ablehnen

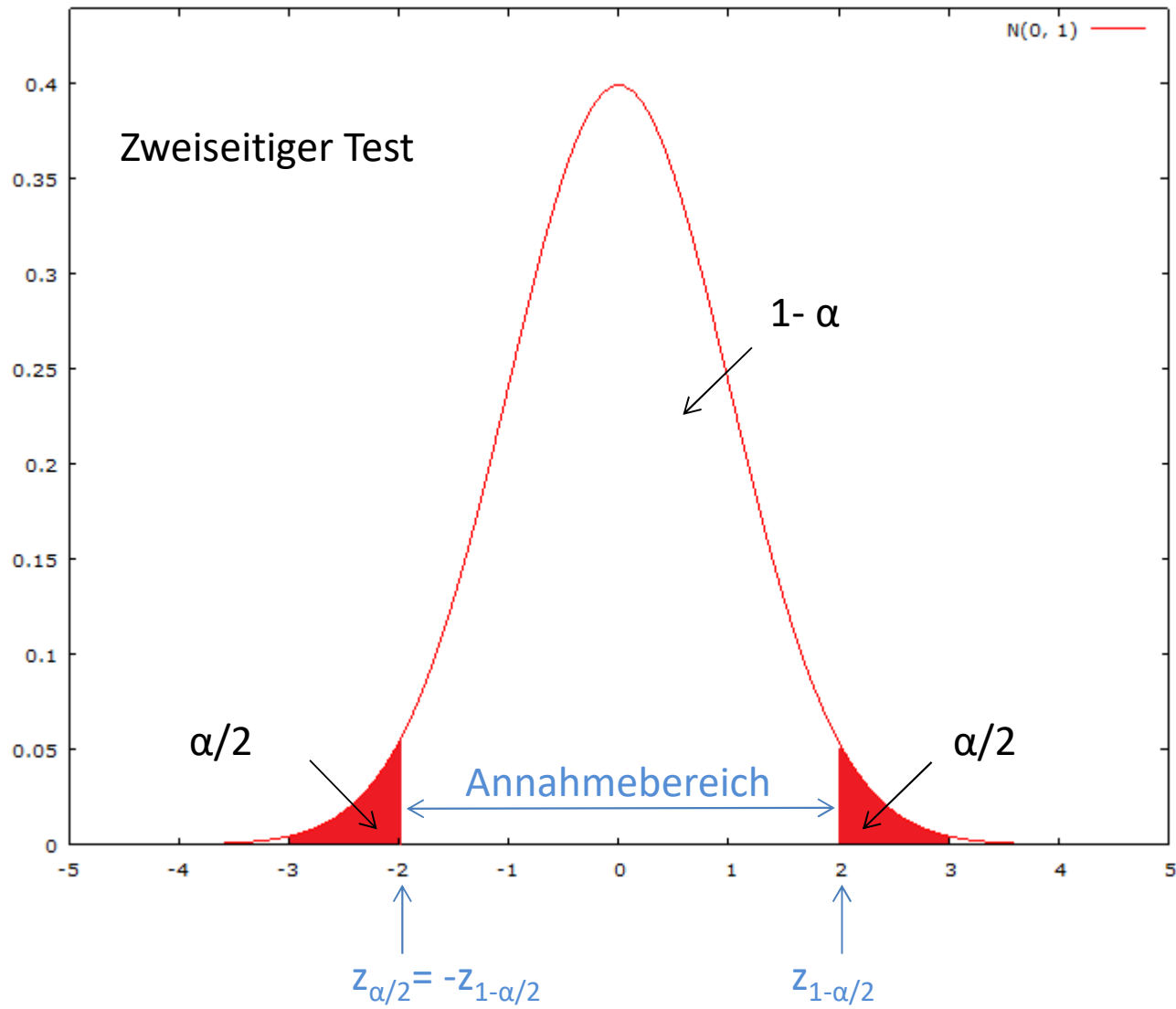
- beidseitiger Test:

$|Z_0| \leq z_{1-\alpha/2} : H_0$ annehmen

$|Z_0| > z_{1-\alpha/2} : H_0$ ablehnen

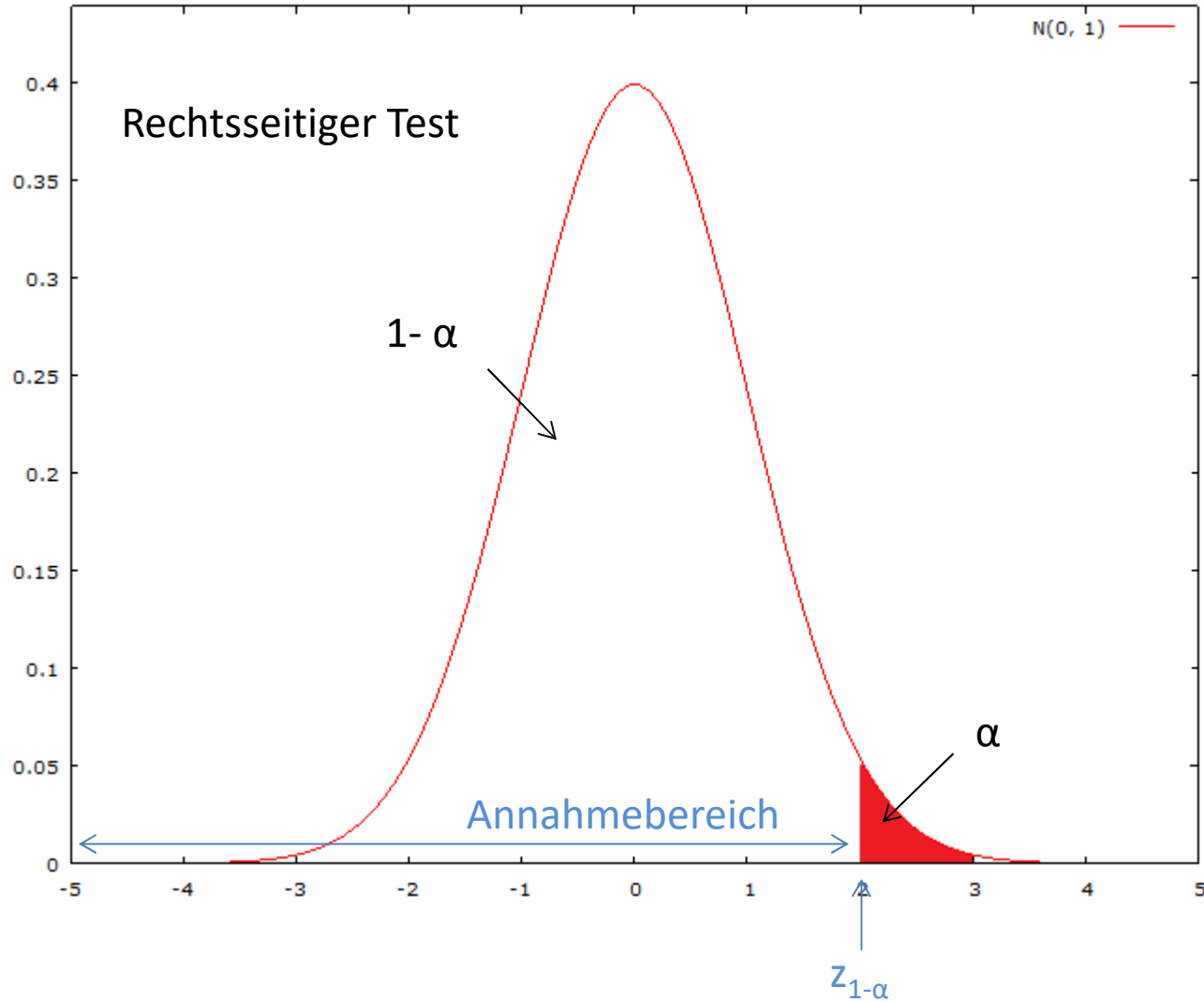
2

Parametrischer Test - Einstichprobentest (hier: Gauß-Test über Mittelwert)



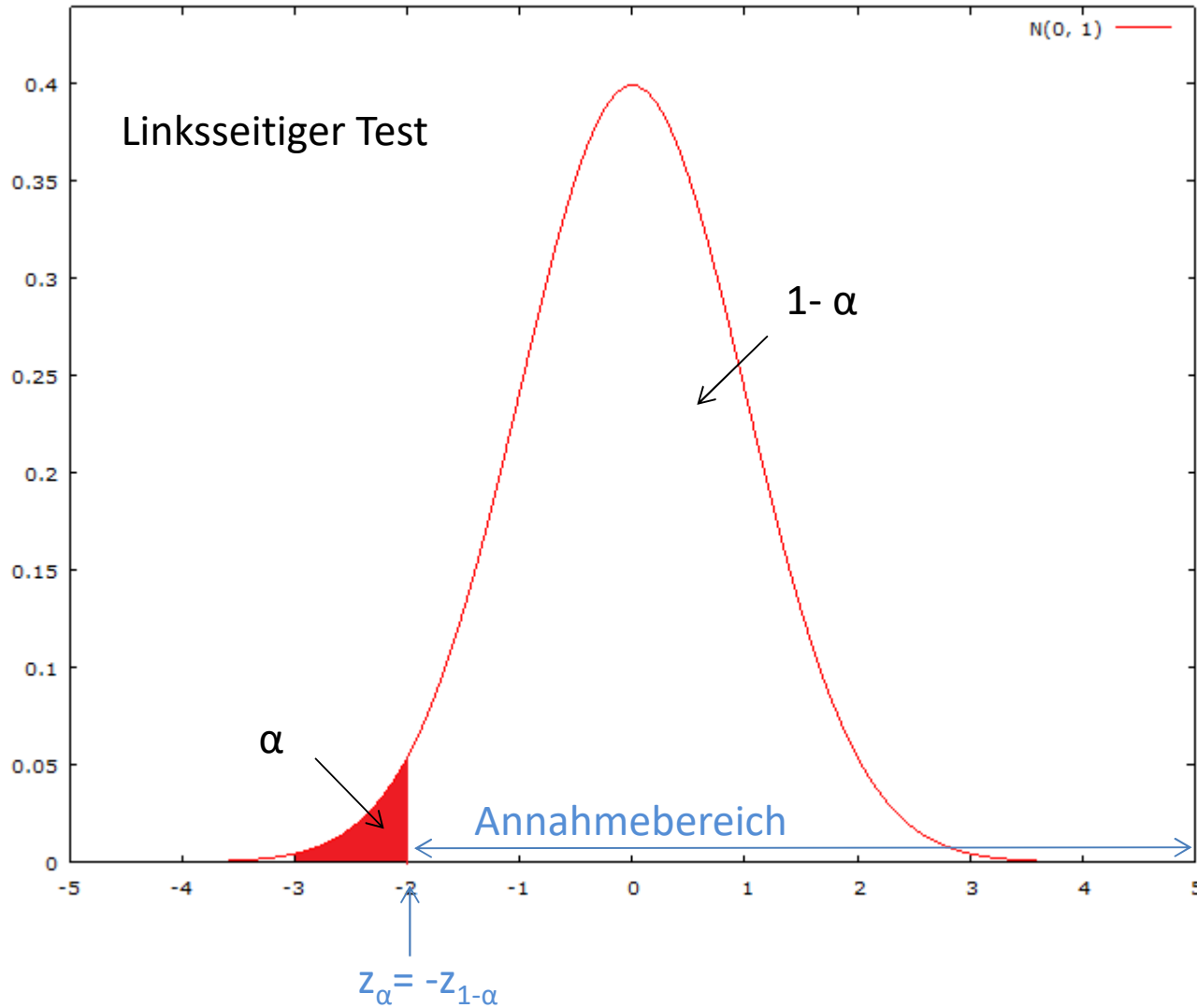
2

Parametrischer Test - Einstichprobentest (hier: Gauß-Test über Mittelwert)



2

Parametrischer Test - Einstichprobentest (hier: Gauß-Test über Mittelwert)



2

Parametrischer Test - Einstichprobentest (hier: Gauß-Test über Mittelwert)

Allgemeiner Hinweis zum Verständnis

Ein Fehler 1. Art beinhaltet ein fälschliches Verwerfen der Nullhypothese (Nullhypothese wird abgelehnt, obwohl sie richtig ist)

Die Wahrscheinlichkeit für einen Fehler 1. Art, die maximal toleriert werden soll, heißt **Signifikanzniveau** oder **Irrtumswahrscheinlichkeit** und wird mit α bezeichnet.

2

Parametrischer Test - Einstichprobentest (hier: Gauß-Test über Mittelwert)

Ablesen des kritischen z-Wertes aus der Tabelle der Normalverteilung

Für z.B. ein Signifikanzniveau $\alpha = 0,01$ ist beim einseitigen Test der $1 - \alpha$ %-Punkt abzulesen

	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
2,00	0,977250	0,977784	0,978308	0,978822	0,979325	0,979818	0,980301	0,980774	0,981237	0,981691
2,10	0,982136	0,982571	0,982997	0,983414	0,983823	0,984222	0,984614	0,984997	0,985371	0,985738
2,20	0,986097	0,986447	0,986791	0,987126	0,987455	0,987776	0,988089	0,988396	0,988696	0,988989
2,30	0,989276	0,989556	0,989830	0,990097	0,990358	0,990613	0,990863	0,991106	0,991344	0,991576
2,40	0,991802	0,992024	0,992240	0,992451	0,992656	0,992857	0,993053	0,993244	0,993431	0,993613

$$\rightarrow z_{1-\alpha} = z_{1-0,01} = z_{0,99} = 2,33$$

Oder alternativ aus der Tabelle der Quantile der Normalverteilung!

$1-\alpha$	$z_{1-\alpha}$	$1-\alpha$	$z_{1-\alpha}$	$1-\alpha$	$z_{1-\alpha}$	$1-\alpha$	$z_{1-\alpha}$
0.9994	3.2389	0.9930	2.4573	0.950	1.6449	0.730	0.6128
0.9993	3.1946	0.9925	2.4324	0.945	1.5982	0.720	0.5828
0.9992	3.1559	0.9920	2.4089	0.940	1.5548	0.710	0.5534
0.9991	3.1214	0.9915	2.3867	0.935	1.5141	0.700	0.5244
0.9990	3.0902	0.9910	2.3656	0.930	1.4758	0.690	0.4959
0.9989	3.0618	0.9905	2.3455	0.925	1.4395	0.680	0.4677
0.9988	3.0357	0.9900	2.3263	0.920	1.4051	0.670	0.4399
0.9987	3.0115	0.9895	2.3080	0.915	1.3722	0.660	0.4125

2

Parametrischer Test - Einstichprobentest (hier: Gauß-Test über Mittelwert)

To do

Fallbeschreibung

Ein Automobilwerk behauptete, dass ein von ihm gefertigter Pkw einen durchschnittlichen Spritverbrauch von höchstens 7,6 l/100 km habe. Die Überprüfung einer Testzeitschrift anhand von 40 zufällig ausgewählten Autos ergab einen Durchschnittsverbrauch von 7,8 l/100 km. Die Herstellerangabe für die Standardabweichung in der Grundgesamtheit $\sigma = 0,6$ wird als richtig vorausgesetzt. Testen Sie die Behauptung des Automobilwerks ($\alpha = 5\%$).

2

Parametrischer Test - Einstichprobentest (hier: Gauß-Test über Mittelwert)

To do

Fallbeschreibung

Psychologische Studien haben einen durchschnittlichen Intelligenzquotienten von 100 Punkten bei einer Standardabweichung von 15 Punkten ergeben. Ein Master-Absolvent vermutet, dass sich der Durchschnittswert aufgrund des Medienkonsums verändert hat. Eine Stichprobe vom Umfang 64 liefert einen Durchschnittswert von 102 Punkten ($\alpha = 5\%$).

2

Parametrischer Test - Einstichprobentest (hier: Gauß-Test über Mittelwert)

To do

Fallbeschreibung

Bei der Überprüfung der ausgeschenkten Biermenge in 50 0,3 l-Gläsern beim Gastwirt Johann Merse ergab sich eine durchschnittliche Füllung von 0,194 l bei einer Standardabweichung von 0,1. Schenkt Gastwirt Merse systematisch zu wenig Bier aus? Ist also davon auszugehen, dass auch in der Grundgesamtheit aller verkauften Biergläser ein zu geringer Inhalt vorhanden ist? Als Signifikanzniveau wird ein α von 5% verwendet.

2

Parametrischer Test - Zweistichprobentest (hier: doppelter Gauß-Test über Mittelwert)

Voraussetzungen/Merkmale beim doppelten Gauß-Test

- Wird zur Untersuchung des Unterschieds zweier arithmetischer Mittel angewendet, wenn die Varianzen der Grundgesamtheit vorliegen
 - Bei Stichprobengrößen $n_1 \leq 30$ und $n_2 \leq 30$ ist zusätzlich die Annahme von normalverteilten Grundgesamtheiten erforderlich
- Siehe folgende Arbeitsschritte zur Durchführung des doppelten Gauß-Tests

2

Parametrischer Test - Zweistichprobentest (Arbeitsschritte doppelter Gauß-Test über Mittelwert)

Standardmäßig beidseitiger Test

1. Signifikanzniveau und Nullhypothese

Signifikanzniveau $\alpha = ?$

Nullhypothese: $H_0: \mu_1 = \mu_2$

2. Prüfgröße und ihre Verteilung

Prüfgröße (empirischer Wert):

- bekannte Varianz:

$$Z_0 = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Z_0 folgt einer $N(0;1)$ -Verteilung

3. Kritischer Wert

zweiseitig ablesen: $z_{1-\alpha/2}$

4. Testentscheidung

$|Z_0| \leq z_{1-\alpha/2} : H_0$ annehmen

$|Z_0| > z_{1-\alpha/2} : H_0$ ablehnen

2

Parametrischer Test - Zweistichprobentest (hier: doppelter Gauß-Test über Mittelwert)

To do

Fallbeschreibung

Ein Marktforschungsinstitut untersucht, ob sich die West- und Ostdeutschen in ihren Fernsehgewohnheiten unterscheiden. Für die 800 westdeutschen Befragten ergibt sich eine durchschnittliche Fernsehdauer von 2 Stunden. Die 600 Befragten aus den neuen Bundesländern wiesen dagegen einen Durchschnittswert von 2,5 Stunden auf. Die aus anderen Untersuchungen bekannten Standardabweichungen (1 Stunde für Westdeutschland und 0,5 Stunden für Ostdeutschland) werden als bekannte Varianzen für die Grundgesamtheiten vorausgesetzt. Testen Sie, ob zwischen den west- und ostdeutschen Befragten bei der durchschnittlichen Fernsehdauer signifikante Differenzen bestehen ($\alpha = 5\%$).

Anmerkungen:

- Der Index 1 wird für Westdeutschland verwendet
- Aufgrund der großen Stichprobenumfänge ist eine Normalverteilung der Merkmale in den Grundgesamtheiten nicht erforderlich
- Wegen den bekannten Varianzen in den Grundgesamtheiten lässt sich der doppelte Gauß-Test anwenden

3

Parametrischer Test - Einstichprobentest (t-Test über Mittelwert)

Voraussetzungen/Merkmale beim t-Test

- unbekannte Varianz in der Grundgesamtheit
- kleine Stichproben ($n \leq 30$, sonst Gauß-Test)
- Merkmal ist in der Grundgesamtheit normalverteilt

→ Siehe folgende Arbeitsschritte zur Durchführung des t-Tests

3

Rückblick: t-Verteilung

- t-Verteilung $t(v)$ mit v Freiheitsgraden ergibt sich als Quotient aus einer normal- und einer Chi-Quadrat- verteilten Zufallsvariablen
- verläuft symmetrisch um den Nullpunkt
- ist flacher als die Standardnormalverteilung, geht aber bei großer Anzahl an Freiheitsgraden in die Standardnormalverteilung über.

Beispiel

Welchen Wert unterschreitet eine t-verteilte ZV ($v = 8$) mit einer Wahrscheinlichkeit von 99 %?

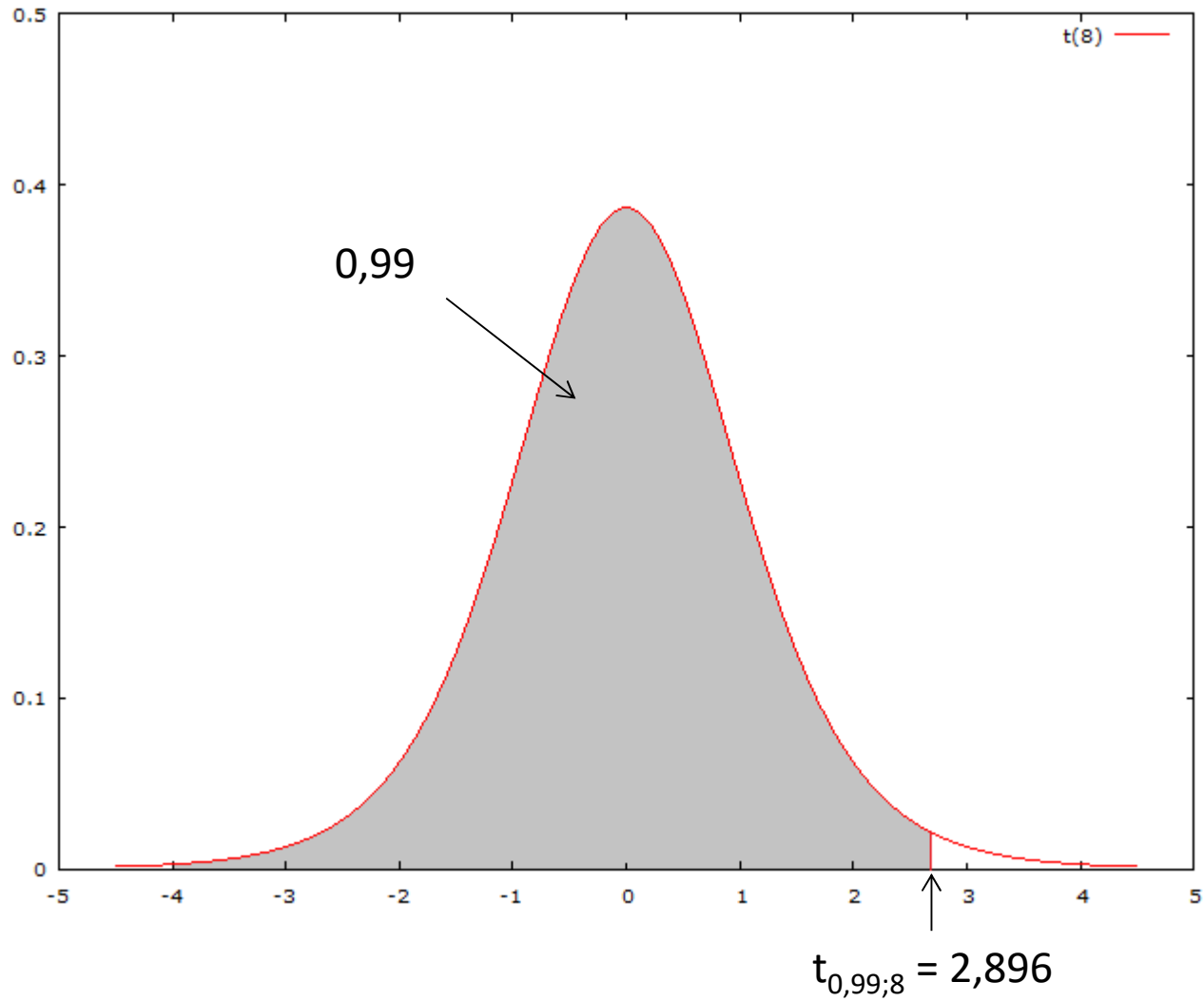
→ gesucht: $P(T \leq t_{1-\alpha;v}) = 1 - \alpha (=0,99)$

→ siehe Quantil-Tabelle: $t_{0,99;8} = 2,896$

n=df\1- α	0,9	0,95	0,975	0,99	0,995
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

3

Rückblick: t-Verteilung



3

Parametrischer Test - Einstichprobentest (Arbeitsschritte t-Test über Mittelwert)

1. Signifikanzniveau und Nullhypothese

Signifikanzniveau $\alpha = ?$

Nullhypothese:

- einseitig: $H_0: \mu \leq (\geq) \mu_0$

- zweiseitig: $H_0: \mu = \mu_0$

2. Prüfgröße und ihre Verteilung

Prüfgröße (empirischer Wert):

$$T_0 = \frac{\bar{X} - \mu_0}{S_{\bar{x}}} \quad \text{mit} \quad S_{\bar{x}} = \frac{S}{\sqrt{n}}$$

T_0 folgt einer $t(n-1)$ -Verteilung;

da durch die Schätzung des arithmetischen Mittels ein Freiheitsgrad „verloren“ geht

3. Kritischer Wert

Kritischer Wert beim:

- einseitigen Test: $t_{1-\alpha;n-1}$

- zweiseitigen Test: $t_{1-\alpha/2;n-1}$

4. Testentscheidung

- einseitiger Test:

$|T_0| \leq t_{1-\alpha;n-1} : H_0$ annehmen

$|T_0| > t_{1-\alpha;n-1} : H_0$ ablehnen

- beidseitiger Test:

$|T_0| \leq t_{1-\alpha/2;n-1} : H_0$ annehmen

$|T_0| > t_{1-\alpha/2;n-1} : H_0$ ablehnen

3

Parametrischer Test - Einstichprobentest (t-Test über Mittelwert)

Ablezen des kritischen Wertes aus der Tabelle der t-Verteilung

Für z.B. $n = 6$ und einem einseitigen Signifikanzniveau $\alpha = 0,05$ ist beim einseitigen Test der 95%-Punkt der t-Verteilung mit 5 Freiheitsgraden (df) (Achtung: in der Tabelle häufig irritierend auch mit n bezeichnet) abzulesen.

n=df\1- α	0,9	0,95	0,975	0,99	0,995
1	3.078	6.314	12.706	31.821	63.657
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

$$\rightarrow t_{1-\alpha;n-1} = t_{0,95;6-1} = 2,015$$

3

Parametrischer Test - Einstichprobentest (t-Test über Mittelwert)

To do

Fallbeschreibung

Ein Pharmaunternehmer behauptet, dass seine Tabletten durchschnittlich 60 mg eines Spurenelements enthalten. Eine Stichprobenuntersuchung von $n = 20$ ergibt einen Inhalt von durchschnittlich 60,02 mg bei einer Varianz von $0,002 \text{ mg}^2$. Eine zu hohe Dosis wirkt sich ebenso wie eine zu geringe Dosis negativ auf die Gesundheit aus. Ist davon auszugehen, dass die Angaben des Pharmaunternehmers falsch sind ($\alpha = 0,10$)?

3

Parametrischer Test - Zweistichprobentest (doppelter t-Test über Mittelwert)

Voraussetzungen/Merkmale beim doppelten Gauß-Test

- Normalverteilung des Merkmals in den Grundgesamtheiten (normalverteilte Grundgesamtheiten)
- Bei Vorliegen kleiner Stichprobengrößen $n_1 \leq 30$ oder $n_2 \leq 30$
- Bei stark unterschiedlichen Stichprobengrößen u.U. problematisch (dann ggf. U-Test anwenden; hier nicht behandelt)
- Wird dann angewendet, wenn die Varianzen in den Grundgesamtheiten unbekannt, aber gleich groß sind
- Die Annahme gleicher Varianzen kann mit dem F-Test überprüft werden

→ Siehe folgende Arbeitsschritte zur Durchführung des doppelten t-Tests

3

Parametrischer Test - Zweistichprobentest (Arbeitsschritte doppelter t-Test über Mittelwert)

Standardmäßig beidseitiger Test

1. Signifikanzniveau und Nullhypothese

Signifikanzniveau $\alpha = ?$

Nullhypothese: $H_0: \mu_1 = \mu_2$

2. Prüfgröße und ihre Verteilung

Prüfgröße (empirischer Wert):

$$T_0 = \frac{\bar{X}_1 - \bar{X}_2}{S \cdot \sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}}} \quad \text{wobei}$$

$$S^2 = \frac{(n_1 - 1) \cdot S_1^2 + (n_2 - 1) \cdot S_2^2}{n_1 + n_2 - 2}$$

S^2 = „gepoolte“ Varianz; eine Art Mittelwert der Varianzen der Stichproben

T_0 folgt einer $t(n-2)$ -Verteilung;

da durch die Schätzung der beiden arithmetischen Mittel zwei Freiheitsgrade „verloren“ gehen

3. Kritischer Wert

zweiseitig ablesen:

$$t_{1-\alpha/2; n_1+n_2-2}$$

oder auch :

$$z_{1-\alpha/2}$$

wenn $n_1 + n_2 > 30$ gilt, da die $t(n-2)$ -Verteilung dann durch eine $N(0,1)$ -Verteilung approximiert werden kann!

4. Testentscheidung

$$|T_0| \leq t_{1-\alpha/2; n_1+n_2-2} \quad \text{bzw.} \quad |Z_0| \leq z_{1-\alpha/2}$$

$\rightarrow H_0$ annehmen

$$|T_0| > t_{1-\alpha/2; n_1+n_2-2} \quad \text{bzw.} \quad |Z_0| > z_{1-\alpha/2}$$

$\rightarrow H_0$ ablehnen

3

Parametrischer Test - Zweistichprobentest (doppelter t-Test über Mittelwert)

To do

Fallbeschreibung

Von Merseburger und Göttinger Studenten wurde das Einkommen pro Monat erhoben. Bei 10 zufällig ausgewählten Merseburger Studierenden ergeben sich

$$\bar{x}_1 = 551[\text{€}] \text{ und } s_1^2 = 22.915,111[\text{€}^2]$$

Für 18 Göttinger Studenten betragen

$$\bar{x}_2 = 606[\text{€}] \text{ und } s_2^2 = 20.836,706[\text{€}^2]$$

Stehen diese Stichprobenergebnisse in Übereinstimmung mit der Hypothese, dass alle Merseburger und alle Göttinger Studierenden ein gleiches Durchschnittseinkommen besitzen ($\alpha = 5\%$)?

Der F-Test (siehe später) habe ergeben, dass die Varianzen in beiden Grundgesamtheiten gleich groß sind.

4

Parametrischer Test - Einstichprobentest (hier: Anteilswert-Test)

Voraussetzungen/Merkmale beim Anteilswert-Test

- Vergleich einer Wahrscheinlichkeit (Anteilswert) mit einem vorgegebenen Wert

→ Siehe folgende Arbeitsschritte zur Durchführung des Anteilswert-Tests

4

Parametrischer Test - Einstichprobentest (Arbeitsschritte Anteilswert-Test)

1. Signifikanzniveau und Nullhypothese

Signifikanzniveau $\alpha = ?$

Nullhypothese:

- einseitig: $H_0: p \leq (\geq) p_0$

- zweiseitig: $H_0: p = p_0$

2. Prüfgröße und ihre Verteilung

Prüfgröße (empirischer Wert):

$$Z_0 = \frac{\bar{P} - p_0}{\sigma_{\bar{P}}} \quad \text{mit} \quad \bar{P} = \frac{n_A}{n}$$

$$\text{und } \sigma_{\bar{P}} = \sqrt{\frac{p_0 \cdot (1 - p_0)}{n}}$$

Die Binomialverteilung kann bei

$n > \frac{9}{p_0 \cdot (1 - p_0)}$ durch eine
Normalverteilung
approximiert werden

3. Kritischer Wert

Kritischer Wert beim:

- einseitigen Test: $z_{1-\alpha}$

- zweiseitigen Test: $z_{1-\alpha/2}$

4. Testentscheidung

- einseitiger Test:

$|Z_0| \leq z_{1-\alpha} : H_0$ annehmen

$|Z_0| > z_{1-\alpha} : H_0$ ablehnen

- beidseitiger Test:

$|Z_0| \leq z_{1-\alpha/2} : H_0$ annehmen

$|Z_0| > z_{1-\alpha/2} : H_0$ ablehnen

4

Parametrischer Test - Einstichprobentest (hier: Anteilswerts-Test)

To do

Fallbeschreibung

Ein Hersteller von Werkzeugmaschinen behauptete, dass seine Produkte nur zu höchstens 10 % Mängel aufweisen. Eine Stichprobe vom Umfang $n = 100$ ergab 21 defekte Maschinen. Kann bei $\alpha = 1\%$ die Behauptung des Herstellers abgelehnt werden?

4

Parametrischer Test - Einstichprobentest (hier: Anteilswerts-Test)

To do

Fallbeschreibung

Ein Lebensmittelhersteller hat seinen Umsatz und seine Mitarbeiterzahlen auf einen Marktanteil von 25 % abgestimmt. Ein geringerer Marktanteil würde einen geringeren Umsatz und damit geringeren Gewinn nach sich ziehen. Ein höherer Marktanteil wäre mit einem Mitarbeitermangel und unzufriedenen Kunden verbunden. Eine Erhebung in 100 Supermärkten hat einen Marktanteil von 30 % ergeben. Ist davon auszugehen, dass der Marktanteil von 25 % abweicht ($\alpha = 0,05$)?

4

Parametrischer Test - Zweistichprobentest (hier: Anteilswert-Test)

Voraussetzungen/Merkmale beim Anteilswert-Test

- Getestet werden soll, ob sich die Anteilswerte zweier Grundgesamtheiten signifikant voneinander unterscheiden

→ Siehe folgende Arbeitsschritte zur Durchführung des Anteilswert-Tests

4

Parametrischer Test - Zweistichprobentest (Arbeitsschritte Anteilswert-Test)

1. Signifikanzniveau und Nullhypothese

Signifikanzniveau $\alpha = ?$

Nullhypothese: $H_0: p_1 = p_2$

2. Prüfgröße und ihre Verteilung

Prüfgröße (empirischer Wert):

$$Z_0 = \frac{\bar{P}_1 - \bar{P}_2}{\sqrt{\bar{P} \cdot (1 - \bar{P}) \cdot \left(\frac{n_1 + n_2}{n_1 \cdot n_2}\right)}} \quad \text{mit}$$

$$\bar{P} = \frac{n_1 \cdot \bar{P}_1 + n_2 \cdot \bar{P}_2}{n_1 + n_2}$$

Die Binomialverteilung lässt sich bei

$$n_1 > \frac{9}{p_1 \cdot (1 - p_1)} \quad \text{und} \quad n_2 > \frac{9}{p_2 \cdot (1 - p_2)}$$

durch eine Normalverteilung approximieren

3. Kritischer Wert

Standardmäßig beidseitiger Test

zweiseitig ablesen: $z_{1-\alpha/2}$

4. Testentscheidung

$|Z_0| \leq z_{1-\alpha/2} : H_0$ annehmen

$|Z_0| > z_{1-\alpha/2} : H_0$ ablehnen

4

Parametrischer Test - Zweistichprobentest (hier: Anteilswerttest)

To do

Fallbeschreibung

Ein Institut hat 1.500 Wahlberechtigten nach ihrer Wahlabsicht befragt. Von den 500 Männern gaben 230 an, CDU/CSU wählen zu wollen. Bei den Frauen lag der Anteil von CDU/CSU-Sympathisanten bei $430/1.000 = 0,43$. Unterscheiden sich beide Anteilswerte für die männlichen und weiblichen Befragten signifikant voneinander ($\alpha = 1\%$)?

5

Parametrischer Test - Einstichprobentest (Varianztest; χ^2 – Test für Varianzen)

Voraussetzungen/Merkmale beim Varianz-Test

- Vergleich der Varianz mit einem vorgegebenem Wert
- Normalverteilte Grundgesamtheit
- Unbekannter Erwartungswert μ der Grundgesamtheit

→ Siehe folgende Arbeitsschritte zur Durchführung des Varianz-Tests

5

Parametrischer Test - Einstichprobentest (Arbeitsschritte χ^2 – Test für Varianzen)

1. Signifikanzniveau und Nullhypothese

Signifikanzniveau $\alpha = ?$

Nullhypothese:

- einseitig: $H_0: \sigma^2 \leq (\geq) \sigma_0^2$
- zweiseitig: $H_0: \sigma^2 = \sigma_0^2$

2. Prüfgröße und ihre Verteilung

Prüfgröße (empirischer Wert):

$$X_0^2 = \frac{n-1 \cdot S^2}{\sigma_0^2} \quad \text{mit} \quad S^2 = \frac{1}{n-1} \cdot \sum_{i=1}^n (X_i - \bar{X})^2$$

χ_0^2 folgt einer χ^2 -Verteilung mit $(n-1)$ Freiheits-Graden. Bei $n > 30$ kann die χ^2 -Verteilung hinreichend genau durch $N(n-1; 2(n-1))$ approximiert werden.

3. Kritischer Wert

Kritischer Wert beim einseitigen Test :

$\chi_{1-\alpha; n-1}^2$ bzw. bei einer Approximation

$$\chi_{1-\alpha; n-1}^2 \approx n-1 + z_{1-\alpha} \cdot \sqrt{2 \cdot (n-1)}$$

Kritischer Wert beim zweiseitigen Test :

$\chi_{1-\alpha/2; n-1}^2$ bzw. bei einer Approximation

$$\chi_{1-\alpha/2; n-1}^2 \approx n-1 + z_{1-\alpha/2} \cdot \sqrt{2 \cdot (n-1)}$$

4. Testentscheidung

- einseitiger Test:

$X_0^2 \leq \chi_{1-\alpha; n-1}^2$: H_0 annehmen

$X_0^2 > \chi_{1-\alpha; n-1}^2$: H_0 ablehnen

- beidseitiger Test:

$X_0^2 \leq \chi_{1-\alpha/2; n-1}^2$: H_0 annehmen

$X_0^2 > \chi_{1-\alpha/2; n-1}^2$: H_0 ablehnen

5

Parametrischer Test - Einstichprobentest (Varianztest; χ^2 – Test für Varianzen)

To do

Fallbeschreibung

Im Rahmen einer Diskussion über die Einkommensverteilung in Deutschland behauptete die Bundesregierung, dass die Varianz des Einkommens höchstens 10.000 [€²]. Bei einer Erhebung an 200 Haushalten ergebe sich ein $s^2 = 14.000$ [€²]. Ist die Aussage der Bundesregierung widerlegt ($\alpha = 5\%$)?

Fallbeschreibung

Das Risiko auf Kapitalmärkten wird häufig über die Streuung der Renditen gemessen. Je weiter die Renditen streuen, desto höher ist die Unsicherheit und damit das Risiko. Ein Aktienhändler behauptete, dass das durchschnittliche Risiko der Aktien auf einem bestimmten Markt, abgebildet durch die Varianz der Renditen, bei 10 [%²]. Eine Stichprobenuntersuchung vom Umfang zehn ergibt eine Varianz von 12 [%²]. Ein risikoaverser Anleger möchte wissen, ob er der Aussage des Aktienhändlers im Hinblick auf das geringe Risiko trotzdem vertrauen kann. Testen Sie die Behauptung des Aktienhändlers bei einem α von 5 %.

5

Parametrischer Test - Zweistichprobentest (F-Test über Varianz)

Voraussetzungen/Merkmale beim F-Test

- Normalverteilte Grundgesamtheiten als Voraussetzung
 - F-Test überprüft, ob zwei normalverteilte Grundgesamtheiten die gleiche Varianz aufweisen
 - F-Test kann zur Kontrolle der Voraussetzungen für den doppelten t-Test verwendet werden
 - Verteilung der Indizes für die Gruppen ist (im Unterschied zu den anderen Zwei-Stichprobentest) nicht beliebig. Die Stichprobe mit der größeren Stichprobenvarianz erhält den Index 1 zugewiesen.
- Siehe folgende Arbeitsschritte zur Durchführung des F-Tests

5

Rückblick: F-Verteilung

F-Verteilung $F(v;w)$ mit v Zähler- und w Nennerfreiheitsgraden ergibt sich als Quotient zweier Chi-Quadrat-verteilter Zufallsvariablen

Unter Anwendung der Tabelle der F-Verteilung lässt sich der Wert $F_{1-\alpha;v;w}$ berechnen, den die F-verteilte ZV mit einer Wahrscheinlichkeit $1-\alpha$ unterschreitet

Beispiel

Eine ZV sei F-verteilt mit $v = 4$ Zähler- und $w = 6$ Nennerfreiheitsgraden. Welchen Wert (auf der Abzisse) wird die ZV einer Wahrscheinlichkeit von 95 % nicht unterschreiten?

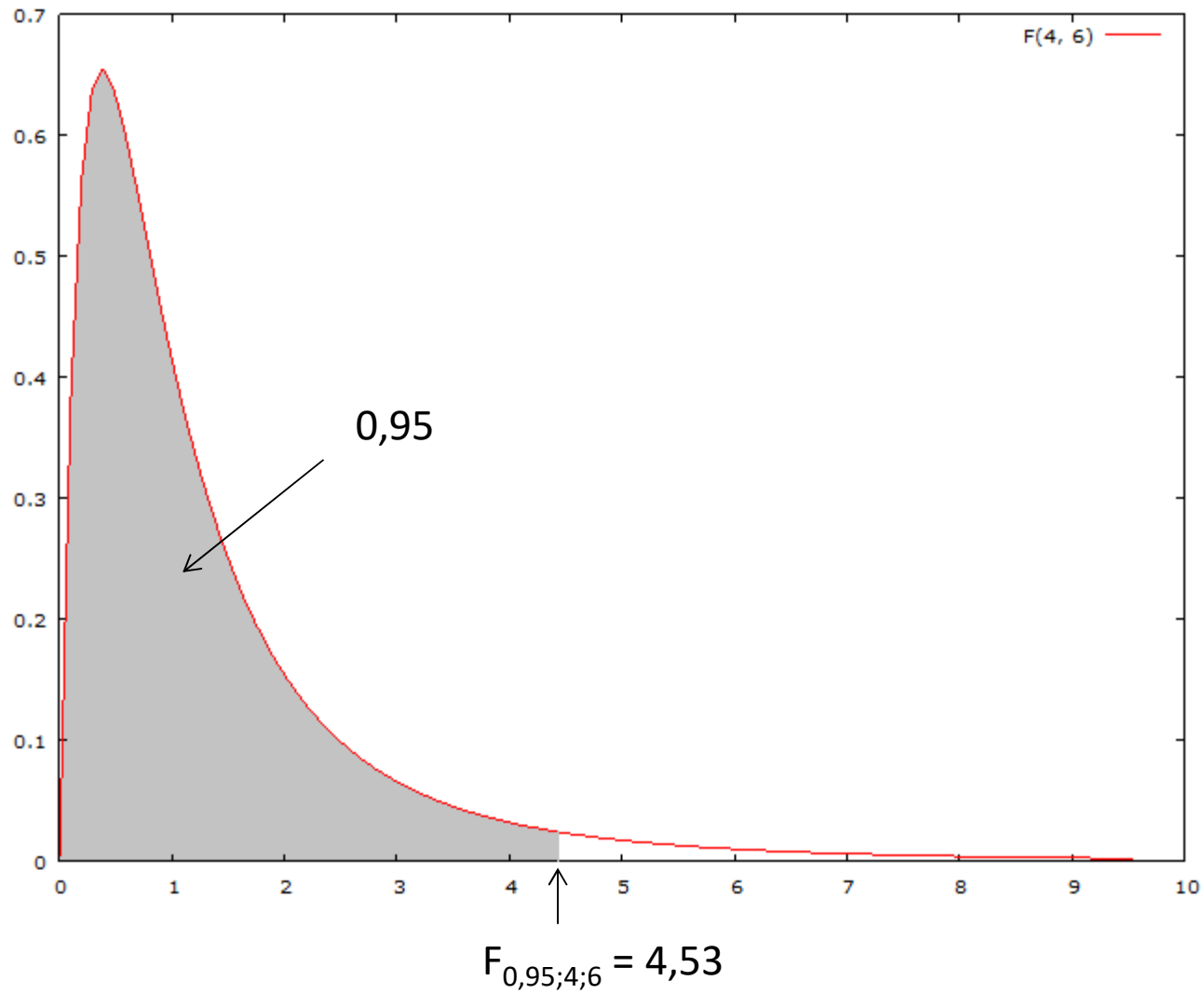
→ gesucht: $P(F \leq F_{1-\alpha;v;w}) = 1-\alpha$

→ siehe Tabelle: $F_{0,95;4;6} = 4,53$

v \ w	1- α	1	2	3	4	5	6	7	8
5	0,950	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82
	0,990	16,3	13,3	12,1	11,4	11,0	10,7	10,5	10,3
6	0,950	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15
	0,990	13,7	10,9	9,78	9,15	8,75	8,47	8,26	8,10
7	0,950	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73
	0,990	12,2	9,55	8,45	7,85	7,46	7,19	6,99	6,84

5

Rückblick: F-Verteilung



5

Parametrischer Test - Zweistichprobentest (Arbeitsschritte F-Test über Varianz)

Standardmäßig beidseitiger Test

1. Signifikanzniveau und Nullhypothese

Signifikanzniveau $\alpha = ?$

Nullhypothese: $H_0: \sigma^2_1 = \sigma^2_2$

2. Prüfgröße und ihre Verteilung

Prüfgröße (empirischer Wert):

In der Prüfgröße ist jeweils die größere Varianz in den Zähler zu setzen:

$$F_0 = \frac{S_1^2}{S_2^2} \text{ mit } S_1^2 > S_2^2$$

F_0 folgt einer F-Verteilung mit $(n_1-1; n_2-1)$ -Freiheitsgraden

3. Kritischer Wert

zweiseitig ablesen: $F_{1-\alpha/2; n_1-1; n_2-1}$

4. Testentscheidung

$F_0 \leq F_{1-\alpha/2; n_1-1; n_2-1}$: H_0 annehmen

$F_0 > F_{1-\alpha/2; n_1-1; n_2-1}$: H_0 ablehnen

5

Parametrischer Test - Zweistichprobentest (F-Test über Varianz)

To do

Fallbeschreibung

Bei einer Erhebung in zwei Regionen 1 und 2 ergaben sich in Bezug auf das monatliche Einkommen folgende Ausprägungen :

$$\bar{x}_1 = 3.314 \text{ €} \quad \bar{x}_2 = 2.625 \text{ €}$$

Die Varianzen sind:

$$s_2^2 = 14.516 \quad s_1^2 = 24.936$$

n_1 ist gleich 61, n_2 ist 121. Es interessiert die Frage, ob die Einkommensunterschiede, gemessen durch die Varianz, signifikant verschieden sind ($\alpha = 10\%$).

6

Kontingenztafel (Stochastische Unabhängigkeit von Ereignissen)

Fallbeschreibung

An einer Hochschule werden die Studiengänge Sozialwesen (S), Technik (T) und Wirtschaft (W) angeboten. Eingeschrieben sind an der Hochschule 275 Frauen (F) und 225 Männer (M). 170 Studierende haben den Studiengang S gewählt, davon 143 Frauen. Für den Studiengang T haben sich 22 Frauen und 108 Männer entschieden. Bei einer Zufallsexperiment werden bei einem zufällig herausgegriffenen Studierenden Geschlecht und Studien-gang notiert. Relative Häufigkeiten werden dabei als Wahrscheinlichkeiten interpretiert .

Ereignisse

E1: „Die zufällig ausgewählte Person ist ein Mann“

E2: „Die zufällig ausgewählte Person ist im Studiengang W immatrikuliert“

Frage

Ist die Auswahl des Studiengangs W vom Geschlecht abhängig?

6

Kontingenztabelle (Stochastische Unabhängigkeit von Merkmalen)

To do

Fallbeschreibung

Eine Versicherungsgesellschaft will überprüfen, ob ein Zusammenhang zwischen Geschlecht und Unfallhäufigkeit besteht. Es wurden hierfür folgende Wahrscheinlichkeiten ermittelt:

Y Geschlecht X Unfallhäufigkeit	y_1 (männlich)	y_2 (weiblich)
x_1 (keinmal)	0,20	0,30
x_2 (ein- oder zweimal)	0,15	0,20
x_2 (mehr als zweimal)	0,10	0,05

7

Rückblick: Chi-Quadrat-Verteilung

Die Chi-Quadrat-Verteilung $X^2(v)$ ergibt sich als Summe von v unabhängigen standardnormalverteilten Zufallsvariablen. Der Parameter v bezeichnet die Freiheitsgrade (auch: $df = \text{degrees of freedom}$) .

Die Chi-Quadrat-verteilte Zufallsvariable mit v Freiheitsgraden nimmt mit einer Wahrscheinlichkeit von $1 - \alpha$ einen Wert an, der kleiner oder gleich dem $1 - \alpha$ – Quantil $X^2_{1-\alpha;v}$ ist.

Beispiel

Eine ZV weise eine X^2 -Verteilung mit zehn Freiheitsgraden auf. Welchen Wert nimmt die ZV mit 95 %-iger Wahrscheinlichkeit höchstens an?

→ siehe Tabelle:

7

Rückblick: Chi-Quadrat-Verteilung

Beispiel

Eine ZV weise eine X^2 -Verteilung mit zehn Freiheitsgraden auf. Welchen Wert nimmt die ZV mit 95 %-iger Wahrscheinlichkeit höchstens an?

→ gesucht: $P(X^2 \leq X^2_{1-\alpha;v}) = 1-\alpha (=0,95)$

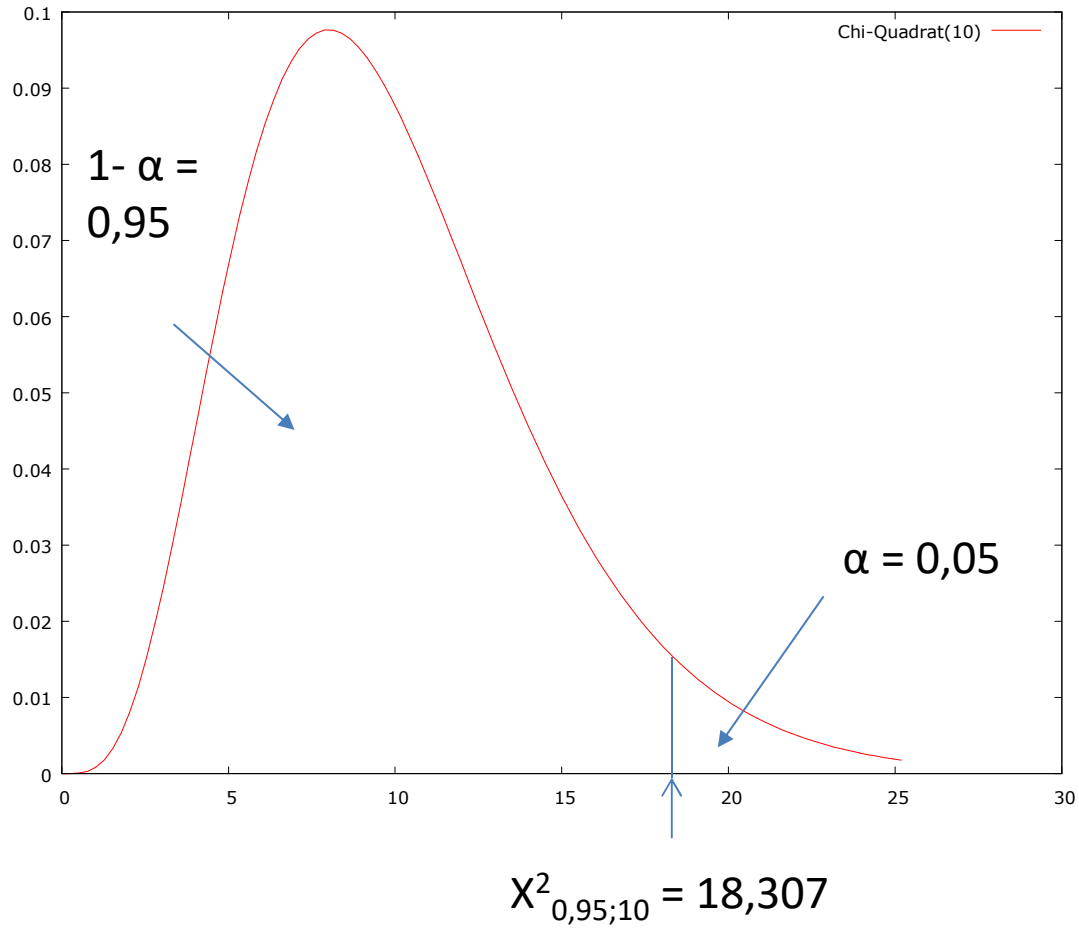
→ siehe Tabelle: $X^2_{0,95;10} = 18,31$



v	(1-a)								
	<u>0,7</u>	<u>0,75</u>	<u>0,8</u>	<u>0,85</u>	<u>0,9</u>	<u>0,95</u>	<u>0,975</u>	<u>0,99</u>	<u>0,995</u>
1	1,07	1,32	1,64	2,07	2,71	3,84	5,02	6,63	7,88
2	2,41	2,77	3,22	3,79	4,61	5,99	7,38	9,21	10,60
3	3,66	4,11	4,64	5,32	6,25	7,81	9,35	11,34	12,84
4	4,88	5,39	5,99	6,74	7,78	9,49	11,14	13,28	14,86
5	6,06	6,63	7,29	8,12	9,24	11,07	12,83	15,09	16,75
6	7,23	7,84	8,56	9,45	10,64	12,59	14,45	16,81	18,55
7	8,38	9,04	9,80	10,75	12,02	14,07	16,01	18,48	20,28
8	9,52	10,22	11,03	12,03	13,36	15,51	17,53	20,09	21,95
9	10,66	11,39	12,24	13,29	14,68	16,92	19,02	21,67	23,59
10	11,78	12,55	13,44	14,53	15,99	18,31	20,48	23,21	25,19

7

Rückblick: Chi-Quadrat-Verteilung



7

Nichtparametrischer Test Chi-Quadrat-Unabhängigkeitstest

Voraussetzungen/Merkmale

- Nichtparametrischer Test, mit dem überprüft werden kann, ob zwei Zufallsvariablen X und Y stochastisch unabhängig sind.

7

Chi-Quadrat-Unabhängigkeitstest

Fallbeschreibung

Im Rahmen einer Marktstudie werden 100 Personen nach der Kaufhäufigkeit sowie dem Geschlecht gefragt.

X Kaufhäufigkeit	Y Geschlecht	y ₁ (männlich)	y ₂ (weiblich)
x ₁ (einmal)		10	15
x ₂ (zweimal)		30	20
x ₃ (dreimal)		5	20

Aufgabe

Den Untersuchungsleiter interessiert, ob zwischen beiden Merkmalen ein Zusammenhang in der Grundgesamtheit besteht. Die Fragestellung soll unter Verwendung des Chi-Quadrat-Unabhängigkeitstests bearbeitet werden ($\alpha = 005$).

7

Chi-Quadrat-Unabhängigkeitstest

To do

Fallbeschreibung

Es ist zu klären, ob die i.R. einer Studie erhobenen Daten (s. Tabelle) auf einen gesicherten Zusammenhang zwischen Geschlecht und Erwerbstätigkeit in der Grundgesamtheit schließen lassen ($\alpha = 10\%$)!

		Geschlecht		Summe
		männlich	weiblich	
Erwerbsbeteiligung	voll erwerbstätig	60	26	86
	teilzeit beschäftigt	2	16	18
	geringfügig beschäftigt	4	8	12
	nicht erwerbstätig	34	50	84
Summe:		100	100	200

H_0 : Es besteht **kein** Zusammenhang zwischen Erwerbstätigkeit und Geschlecht

H_1 : Es besteht **ein** Zusammenhang zwischen Erwerbstätigkeit und Geschlecht

Fallbeschreibung

Der Suchtbeauftragte eines Unternehmens vermutet, dass das Geschlecht einen Einfluss darauf hat, ob ein Mitarbeiter Raucher oder Nichtraucher ist. Um dies zu überprüfen, wählt er aus allen Mitarbeitern zufällig 100 aus und erhebt von diesen Geschlecht sowie Suchtverhalten (Raucher/Nichtraucher). Er erhält folgende Daten.

	Weiblich	Männlich	
Raucher	35	25	beobachtet
Nichtraucher	20	20	

Wenn es keinen Zusammenhang zwischen den Merkmalen geben würde, dann würde die Tabelle wie folgt aussehen:

	Weiblich	Männlich	
Raucher	33	27	erwartet
Nichtraucher	22	18	

H_0 : Geschlecht und Suchtverhalten sind in der Grundgesamtheit unabhängig

H_1 : Das Geschlecht hat einen Einfluss auf das Suchtverhalten

7

Chi-Quadrat-Unabhängigkeitstest

To do

Aufgabe

Es soll mittels χ^2 – Anpassungstests (Signifikanzniveau $\alpha = 10\%$) ermittelt werden, ob ein signifikanter Zusammenhang zwischen dem Suchtverhalten (Raucher/Nichtraucher) eines Mitarbeiters und dessen Geschlecht besteht.

7

Einschränkungen: Chi-Quadrat-Unabhängigkeitstest

- Der X^2 -Wert sagt nichts über die Stärke des Effektes aus
- Ungeeignet für kleine Stichproben; hierfür: Gruppierung der Beobachtungen notwendig (Häufigkeiten pro Klasse müssen > 5 sein)

8

Nichtparametrischer Test Chi-Quadrat-Anpassungstest

Voraussetzungen/Merkmale

- Nichtparametrischer Test, mit dem überprüft werden kann, ob ein Merkmal X in der Grundgesamtheit einer bestimmten Verteilung folgt.
- Es kann z.B. auch getestet werden, ob die Normalverteilungsannahme, die die parametrischen Tests in der Regel voraussetzen, vorliegt.

→ Siehe folgende Arbeitsschritte zur Durchführung des Chi-Quadrat-Anpassungstests

8

Nichtparametrischer Test (Arbeitsschritte Chi-Quadrat-Anpassungstest)

1. Signifikanzniveau und Nullhypothese

Signifikanzniveau $\alpha = ?$

Nullhypothese: Die beobachteten Häufigkeiten stimmen mit den erwarteten in der Grundgesamtheit überein.

2. Prüfgröße und ihre Verteilung

Prüfgröße (empirischer Wert):

$$X_0^2 = \sum_{j=1}^m \frac{(n_j - \tilde{n}_j)^2}{\tilde{n}_j}$$

n_j : Beobachtete Häufigkeit

\tilde{n}_j : Erwartete Häufigkeit

X_0^2 : Chi-Quadrat-Verteilung mit $m-k-1$ Freiheitsgraden (k = Anzahl zu schätzender Parameter)

3. Kritischer Wert

Kritischer Wert: $X_{1-\alpha; p-k-1}^2$

4. Testentscheidung

$X_0^2 \leq X_{1-\alpha; p-k-1}^2$: H_0 annehmen

$X_0^2 > X_{1-\alpha; p-k-1}^2$: H_0 ablehnen

8

Chi-Quadrat-Anpassungstest

To do

Fallbeschreibung

A als Teilnehmer an einem Glücksspiel bezweifelt, dass ein beim Spiel benutzter Würfel fair ist. Bei einem fairen Würfel würde man erwarten, dass jede Augenzahl ungefähr gleich häufig gewürfelt würde (Gleichverteilung). A stützt seine Vermutung darauf, dass die beobachtete Häufigkeit davon teilweise erheblich abweicht. Die Stichprobenergebnisse gehen aus folgender Tabelle hervor:

Augenzahl des Würfels	1	2	3	4	5	6
Beobachtete Häufigkeit n_j	279	311	324	263	367	256
Erwartete Häufigkeit \tilde{n}_j	300	300	300	300	300	300

Ist bei einem Signifikanzniveau von $\alpha = 5\%$ davon auszugehen, dass A mit seiner Vermutung richtig liegt?

8

Chi-Quadrat-Anpassungstest

To do

Fallbeschreibung

Bei der Befragung von 100 Merseburger Studierenden ergaben sich folgende Besatzziffern in den gewählten Einkommensklassen:

Einkommens- klasse	<450	450-550	550-650	650-750	750-850	>850
Anzahl Merk- malsträger	18	28	21	18	9	6

Es gelte : $\bar{x}_1 = 551$ und $s^2 = 22.915$

Testen Sie bei einem Signifikanzniveau von $\alpha = 10 \%$, ob das Einkommen einer Normalverteilung folgt?

9

Regressionsanalyse

Beispiel: Charakteristische Linie eines Wertpapiers; gesucht ist eine Schätzgleichung für künftige Wertpapierrenditen

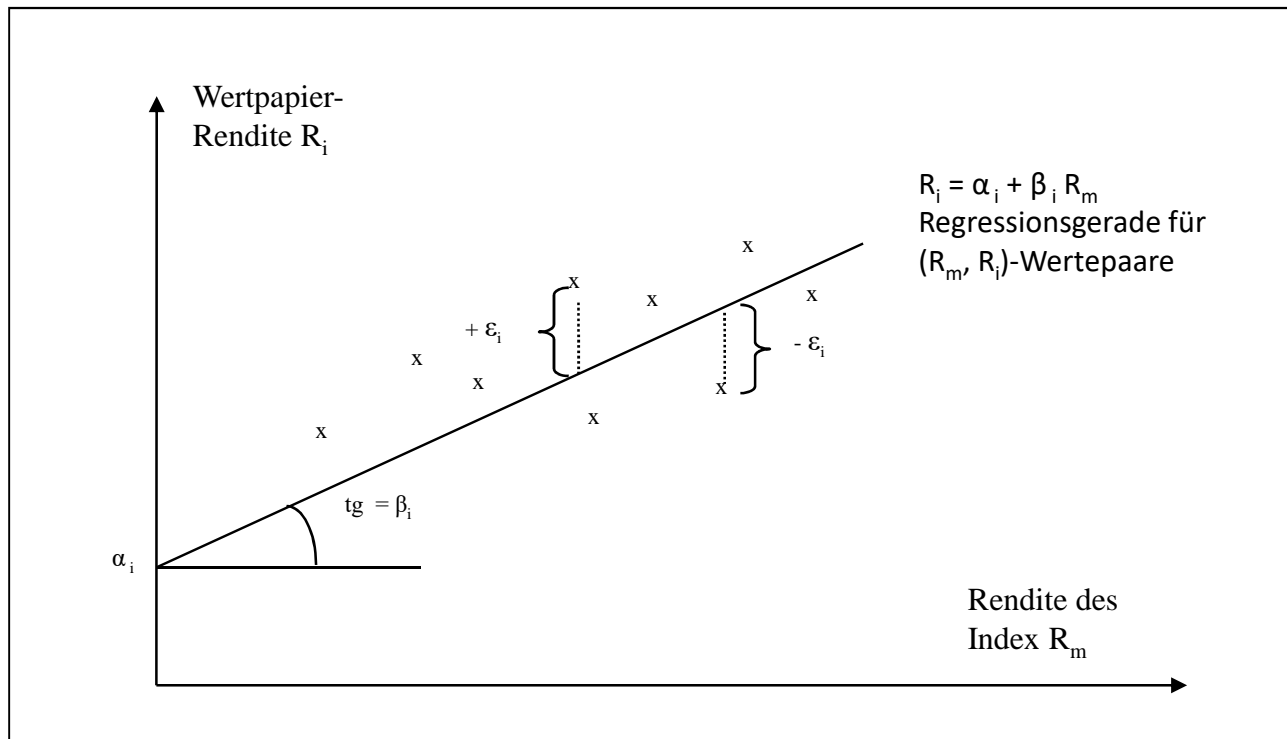
$$R_i = \alpha_i + \beta_i \cdot R_m$$

R_i = Rendite des Wertpapiers i

R_m = Marktrendite

α_i = von der Marktrendite unabhängig

β_i = Beta – Faktor



9

Regressionsanalyse

Bestimmungsgleichung der Regressionskoeffizienten

$$\alpha_i = \frac{\sum_{t=1}^T R_{m,t}^2 \cdot \sum_{t=1}^T R_{i,t} - \sum_{t=1}^T R_{m,t} \cdot \sum_{t=1}^T R_{m,t} \cdot R_{i,t}}{T \cdot \sum_{t=1}^T R_{m,t}^2 - \left(\sum_{t=1}^T R_{m,t} \right)^2}$$

$$\beta_i = \frac{T \cdot \sum_{t=1}^T R_{m,t} \cdot R_{i,t} - \sum_{t=1}^T R_{m,t} \cdot \sum_{t=1}^T R_{i,t}}{T \cdot \sum_{t=1}^T R_{m,t}^2 - \left(\sum_{t=1}^T R_{m,t} \right)^2}$$

Schema/Arbeitstabelle

Jahre t	R_m	R_i	R_m^2	$R_m \cdot R_i$
1				
2				
...				
T				
	$\sum_{t=1}^T R_{m,t}$	$\sum_{t=1}^T R_{i,t}$	$\sum_{t=1}^T R_{m,t}^2$	$\sum_{t=1}^T R_{m,t} \cdot R_{i,t}$
	$\left(\sum_{t=1}^T R_{m,t} \right)^2$			

9

Regressionsanalyse Zahlenbeispiel Wertpapierrendite

Jahre t	„Markt“- Rendite R_m	Aktienrendite R_i
1	6 %	5 %
2	6 %	6 %
3	4 %	1 %
4	7 %	9 %
5	9 %	5 %
6	6 %	10 %
7	6 %	5 %
8	7 %	8 %
9	8 %	10 %
10	8 %	9 %

Jahre t	R_m	R_i	R_m^2	$R_m \cdot R_i$
1	6	5	36	30
2	6	6	36	36
3	4	1	16	4
4	7	9	49	63
5	9	5	81	45
6	6	10	36	60
7	6	5	36	30
8	7	8	49	56
9	8	10	64	80
10	8	9	64	72
Summe	67	68	467	476
$\left(\sum_{t=1}^T R_{m,t}\right)^2$	4.489			

$$\alpha = \frac{467 \cdot 68 - 67 \cdot 476}{10 \cdot 467 - 4.489} = -0,751$$

$$\beta = \frac{10 \cdot 476 - 67 \cdot 68}{10 \cdot 467 - 4.489} = 1,127$$

Regressionsgerade: $R_i = \alpha_i + \beta_i \cdot R_m = -0,751 + 1,127 \cdot R_m$

9

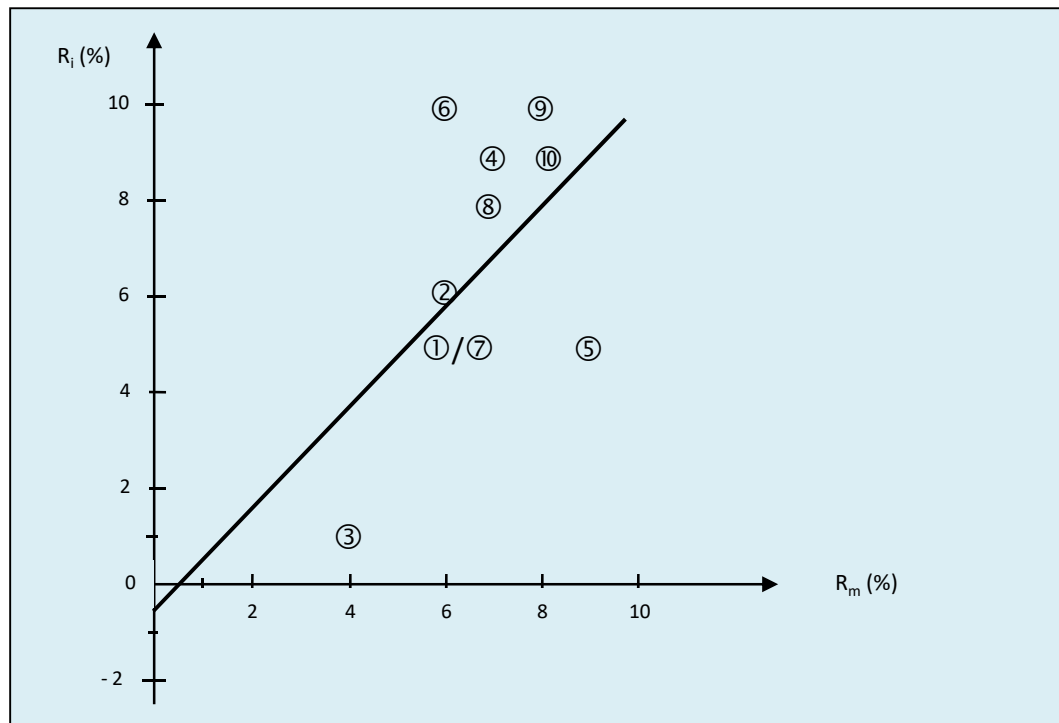
Regressionsanalyse Zahlenbeispiel Wertpapierrendite

→ Alpha und Beta : $\alpha = \frac{467 \cdot 68 - 67 \cdot 476}{10 \cdot 467 - 4 \cdot 489} = -0,751$ $\beta = \frac{10 \cdot 476 - 67 \cdot 68}{10 \cdot 467 - 4 \cdot 489} = 1,127$

→ Regressionsgerade :

$$R_i = \alpha_i + \beta_i \cdot R_m = -0,751 + 1,127 \cdot R_m$$

→ Charakteristische Linie :



9 Regressionsanalyse und Test (Excel-Beispiel)

To do

Jahre t	„Markt“- Rendite R_m	Aktienrendite R_i
1	7 %	7 %
2	6 %	8 %
3	5 %	4 %
4	9 %	8 %
5	11 %	10 %
6	14 %	12 %
7	9 %	10 %
8	8 %	8 %
9	15 %	12 %
10	8 %	9 %

Jahre t	R_m	R_i	R_m^2	$R_m \cdot R_i$
1	7	7	49	49
2	6	8	36	48
3	5	4	25	20
4	9	8	81	72
5	11	10	121	110
6	14	12	196	168
7	9	10	81	90
8	8	8	64	64
9	15	12	225	180
10	8	9	64	72
Summe	92	88	942	873
$\left(\sum_{t=1}^T R_{m,t}\right)^2$	8464			

$$\alpha = \frac{942 \cdot 88 - 92 \cdot 873}{10 \cdot 942 - 8.464} = 2,699$$

$$\beta = \frac{10 \cdot 873 - 92 \cdot 88}{10 \cdot 942 - 8.464} = 0,663$$

Regressionsgerade: $R_i = \alpha_i + \beta_i \cdot R_m = 2,699 + 0,663 \cdot R_m$