Signifikanztests

Was ist ein Signifikanztest?

Typische Fragestellungen

- Sind Bauteile korrekt hersgestellt?
- Sind Nanos besser als Physiker?
- Ist die Übereinstimmung der experimentellen mit der theoretischen Verteilung akzeptabel?

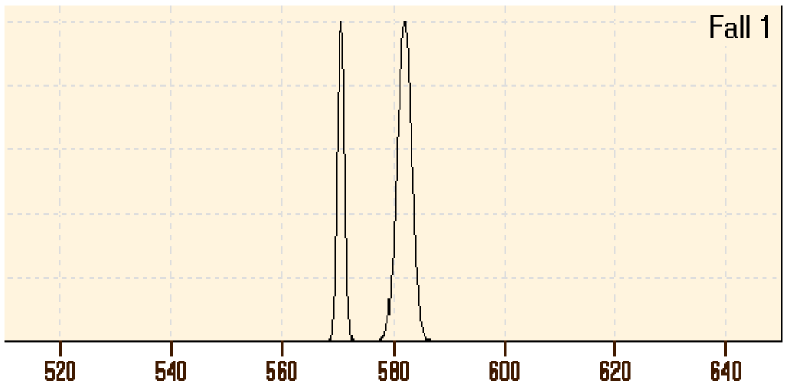
Was heißt signifikant?
Was ist ein Signifikanzniveau?
Was ist eine Hypothese?
Was sind Fehlerarten?
Was sind einseitige (zweiseitige) Tests?
Was ist ein χ^2 – Test?

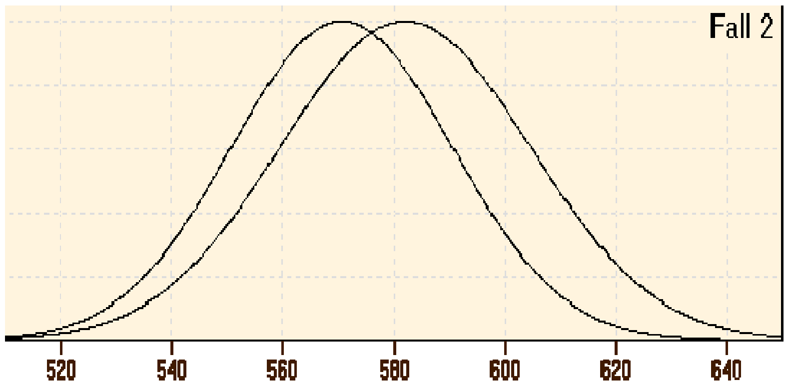
Was heißt signifikant?

Wasserproben: Es wird der Zinkgehalt im Trinkwasser im Norden und im Süden eines Landes gemessen.

	Fall 1: Norden	Fall 1: Süden	Fall 2: Norden	Fall 2: Süden
Einzelwerte:	571.5	581.7	566.3	571.5
	570.8	582.8	549.5	608.1
	570.3	580.4	538.9	544.7
	570.6	583.3	588.9	571.7
	570.4	582.9	592.5	589.7
	571.0	579.5	560.1	588.5
	571.6	582.8	572.9	577.7
	569.4	583.3	575.1	583.7
	570.5	583.0	602.7	561.7
	570.9	581.3	560.1	623.7
Mittelwerte:	570.7	582.1	570.7	582.1

Ist das Wasser im Süden stärker belastet?

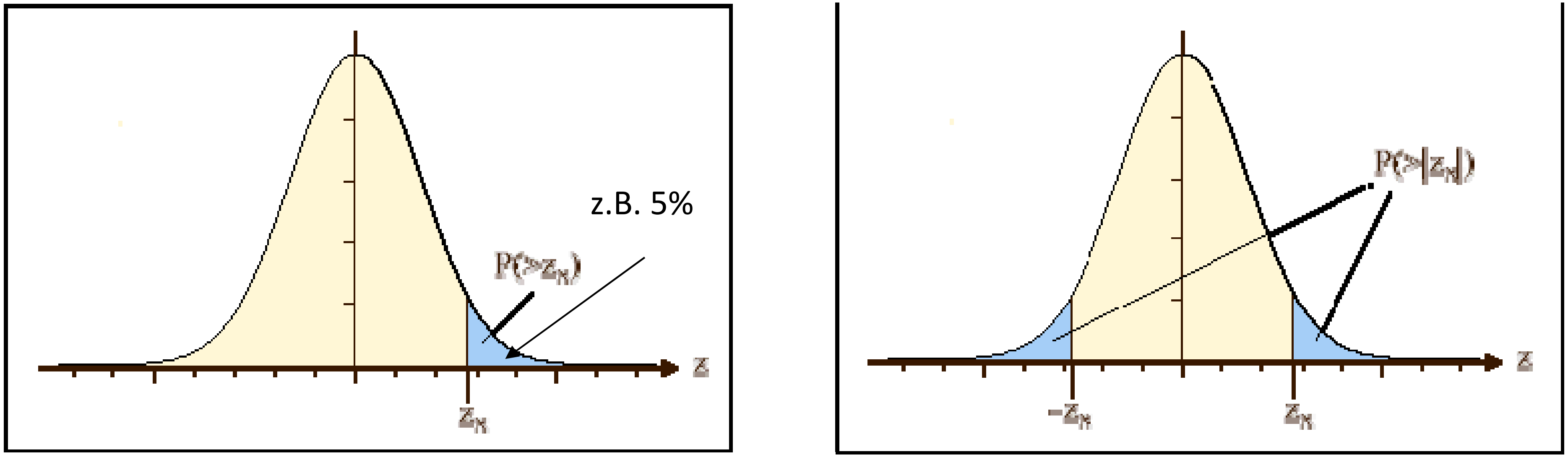




Numerisch sind die Mittelwerte verschieden. Im Fall 1 ist der Unterschied deutlich (lat. significans). Im zweiten Fall ist er nicht so deutlich.

Signifikanzniveau

Rückgriff auf die Normalverteilung:
Bei der Beobachtung von Zufallsprozessen gibt es immer die Wahrscheinlichkeit, dass das Ergebnis einen bestimmten Grenzwert z_x überschreitet.

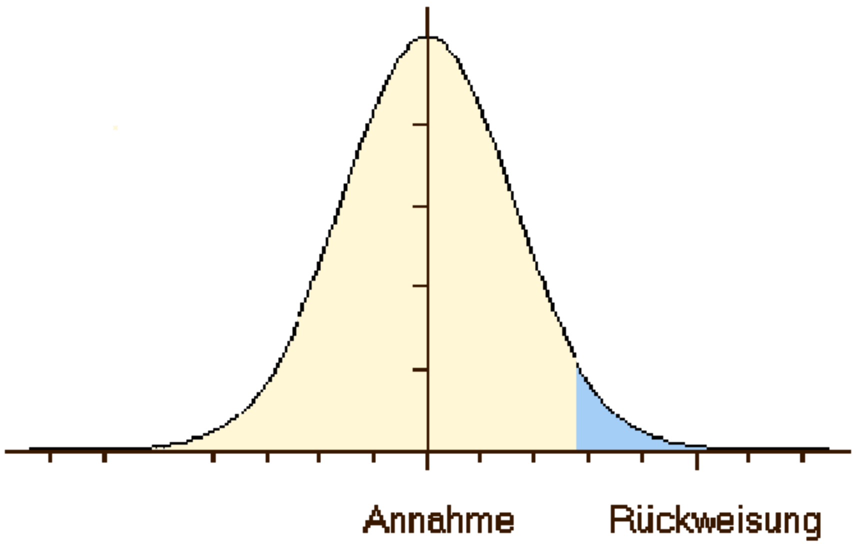


Der eingefärbte Bereich gibt die Wahrscheinlichkeit an, mit der das Ergebnis z einen bestimmten Grenzwert z_x überschreitet (links).

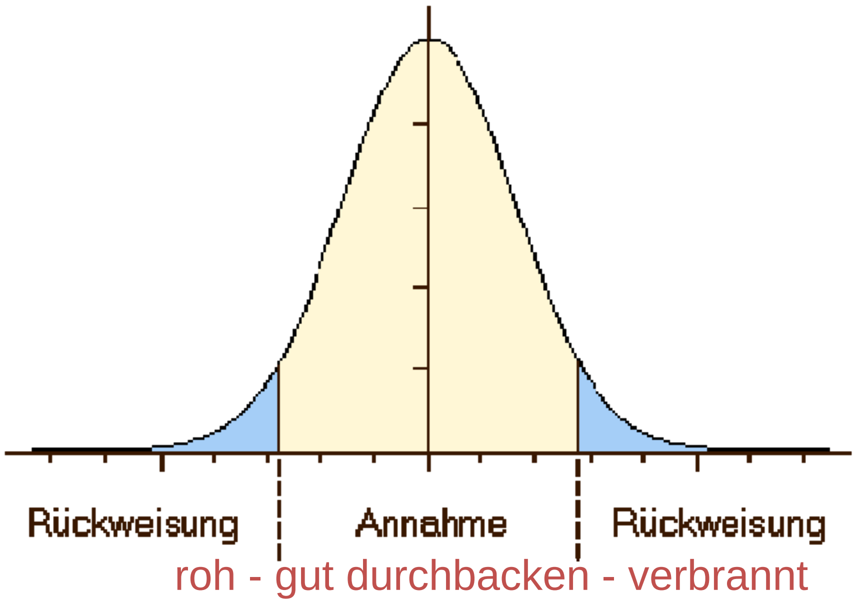
In manchen Fällen werden auch zwei Grenzwerte betrachtet werden müssen (rechtes Beispiel).

Einseitige und zweiseitige Tests

Übersteigt die Zinkkonzentration einer Wasserprobe einen bestimmten Grenzwert?



Ist die bestellte Pizza gut durchgebacken?



Einseitige Test haben nur einen Ablehnungsbereich (links).

Zweiseitige Tests werden angewendet, wenn ein Parameter auf Gleichwertigkeit mit einem bestimmten Wert überprüft werden soll. Abweichungen in beide Richtungen werden verworfen (rechts).

Konfidenzintervalle

Die Mittelwerte $\overline{x}_i$ sind normalverteilt um den wahren Mittelwert μ

der Grundgesamtheit mit einer Standardabweichung von $\frac{\sigma}{\sqrt{n}}$,

wenn n groß ist.

Das heißt, alle Mittelwerte einer Stichprobe einer Normalverteilung liegen mit einer 68.27 % Wahrscheinlichkeit im Intervall

$$\left(\mu - \frac{\sigma}{\sqrt{n}} \; ; \; \mu + \frac{\sigma}{\sqrt{n}} \right)$$

Aufgabe ist es jetzt, die Länge eines Intervalls so festzulegen,

dass es den wahren Mittelwert μ

mit einer runden Vertrauenswahrscheinlichkeit β überdeckt,

also 90% oder 95%

Konfidenzintervalle

In der Praxis hat β die Werte $\beta = 0{,}90$, $\beta = 0{,}95$, $\beta = 0{,}99$
und gelegentlich $\beta = 0{,}999$

Es ist klar, dass es einen Parameter λ_β geben muss, so dass das Intervall

$$\left(\overline{x}_i - \lambda_\beta \frac{\sigma}{\sqrt{\mathrm{n}}} \; ; \; \overline{x}_i + \lambda_\beta \frac{\sigma}{\sqrt{\mathrm{n}}} \right)$$

ein Konfidenzintervall für zur Vertrauenswahrscheinlichkeit β ist.

$$\Phi\left(\overline{x}_i - \lambda_\beta \frac{\sigma}{\sqrt{n}} \;\langle\; \mu \;\langle\; \overline{x}_i + \lambda_\beta \frac{\sigma}{\sqrt{n}} \right) = \beta$$

Konfidenzintervalle

$$\Phi\left(\overline{x}_i - \lambda_\beta \frac{\sigma}{\sqrt{n}} \;\langle\; \mu \;\langle\; \overline{x}_i + \lambda_\beta \frac{\sigma}{\sqrt{n}} \right) = \beta$$

Diese Ungleichung lässt sich umformen zu

$$\Phi\left(-\lambda_\beta \;\langle\; \frac{\mu - \overline{x}_i}{\sigma}\sqrt{n} \;\langle\; +\lambda_\beta \right) = \beta$$

Der Ausdruck $\frac{\mu - \overline{x}_i}{\sigma}\sqrt{n} = z$ wird im Laufe der Vorlesung wiederholt gebraucht.

Häufig benutzte Werte für λ :

$\lambda_{90\%} = 1{,}645$ $\lambda_{68,27\%} = 1{,}000$

$\lambda_{95\%} = 1{,}960$ $\lambda_{95,44\%} = 2{,}000$

$\lambda_{99\%} = 2{,}576$ $\lambda_{99,73\%} = 3{,}000$

Diese λ sind zweiseitige Schranken der Normalverteilung.

Konfidenzintervalle: Beispiel

Schokoladentafeln: Das wahre Gewicht (Mittelwert der Grundgesamtheit) μ sei 100g.

Wir machen eine Zufallsstichprobe (n>30). $\overline{x}_i = \overline{x} = 93{,}0\,\mathrm{g}$ bei einem $\sigma_{\overline{x}} = 4{,}0\,\mathrm{g}$

Mit 95,5% Sicherheit (Konfidenz) liegt der wahre Mittelwert im Intervall.

$$\overline{x} \pm 2 \cdot \sigma_{\overline{x}}$$

Mit 95,0% Sicherheit (Konfidenz) liegt der wahre Mittelwert im Intervall.

$$\overline{x} \pm 1{,}96 \cdot \sigma_{\overline{x}}$$

Das heißt, alle μ zwischen 100,8 g und 85,2 g können diese Stichprobe mit 95,0% Konfidenz erzeugt haben.

Konfidenzintervalle: Beispiel

Wenn alle μ zwischen 100,8 g und 85,2 g diese Stichprobe mit einer 95,0% Konfidenz erzeugt haben, kann man sich fragen, wie sehen andere Konfidenzintervalle aus?

Zum Beispiel können alle μ zwischen 103,3 g und 82,7 g diese Stichprobe mit 99,0% Konfidenz erzeugt haben.

Zum Beispiel können alle μ zwischen 99,6 g und 86,4 g diese Stichprobe mit 90,0% Konfidenz erzeugt haben.

Das Intervall der möglichen Erzeuger für diesen Stichprobenmittelwert wird umso größer, je kleiner ich die Restwahrscheinlichkeit (Irrtumswahrscheinlichkeit) mache.

Ein Konfidenzintervall kennzeichnet denjenigen Bereich eines Merkmals, in dem sich 95% (99%) aller möglichen Populationsparameter befinden, die den empirisch ermittelten Stichprobenkennwert erzeugt haben können.

VORSICHT!!!! Dies gilt nur bei großen Stichproben n>30.

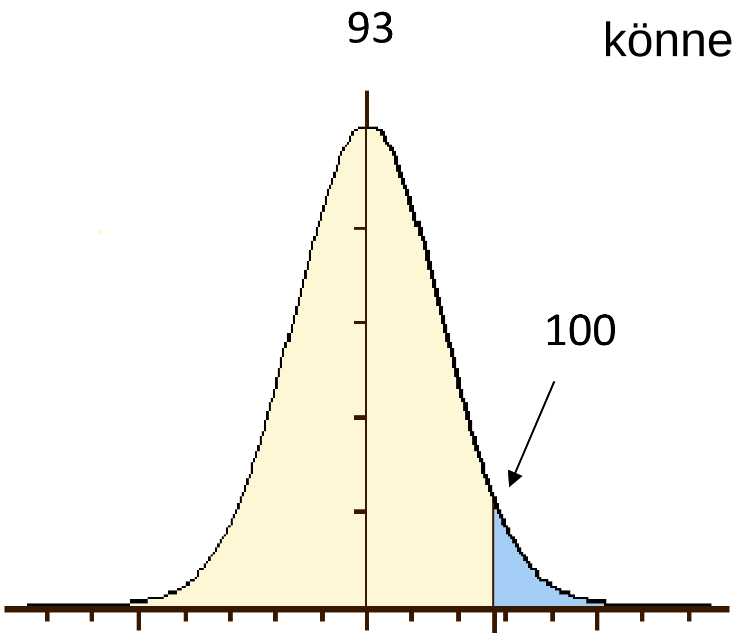
Konfidenzintervalle: Beispiel

Alternative Betrachtung:

$$\bar{x} = 93{,}0\,\mathrm{g} \quad \text{bei einem} \quad \sigma_{\bar{x}} = 4\,\mathrm{g}$$

$$\frac{|\bar{x} - \mu|}{\sigma}\sqrt{n} = z$$

$$\frac{|93 - 100|}{\sigma_{\bar{x}}} = \frac{7}{4} = 1{,}75$$



So kleine Werte, wie den Stichprobenmittelwert 93,0 g, können in 4,01 % der Fälle von Werten größer als 100 g erzeugt werden.

Ist das nun eine signifikante Abweichung?

Oder kann ich behaupten, der wahre Erzeuger für diese Stichprobe hat nicht den Mittelwert 100g?

Testen von Hypothesen

Es gibt zwei verschiedene Fehlerarten:

Fehler erster (α-Fehler) und Fehler zweiter Art (β-Fehler).

Wir kaufen Schokolade.

Die Schokolade soll ein Verpackungsgewicht von 100g aufweisen.

Aufstellen einer Hypothese:

Wir misstrauen dem Lieferanten.
Wir nehmen an, dass der Lieferant unsere Bedingung nicht erfüllt,
d.h. die Tafeln wiegen tatsächlich weniger als 100g

Dies ist unsere Nullhypothese, also die Annahme, die wir widerlegen möchten.

H_0^* oder Nullhypothese:

$$m_{Schokolade} \leq m_{Limit} = 100g$$

Nullhypothese - Alternativhypothese

H_0^* oder Nullhypothese: $$m_{Schokolade} \leq m_{Limit}$$

Zu jeder Nullhypothese gibt es eine Alternativhypothese H_1^*.

H_1^* oder Alternativhypothese: $$m_{Schokolade} > m_{Limit}$$

Es gibt vier mögliche Ergebnisse für unseren Test.

		Wirklichkeit		
		H_0 = wahr	H_0 = falsch	
Schluss-folgerung	H_0 = wahr	OK	Fehler vom Typ II	Wir kaufen nicht.
	H_0 = falsch	Fehler vom Typ I	OK	Wir kaufen.
		$m_{Schokolade} \leq m_{Limit}$	$m_{Schokolade} > m_{Limit}$	

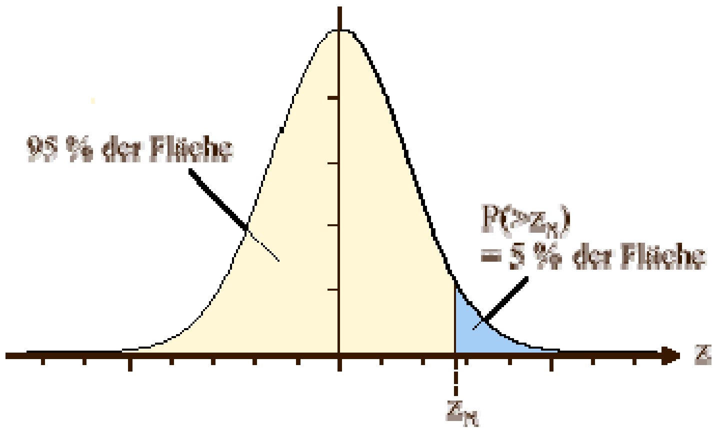
In diesem Beispiel ist der Fehler vom Typ II weniger schädlich.

α - Fehler – β - Fehler

Fehler erster Art oder α-Fehler: Eine richtige Nullhypothese wird zu Gunsten der Alternativhypothese abgelehnt.

Fehler zweiter Art oder β-Fehler: Eine richtige Alternativhypothese wird zu Gunsten der Nullhypothese abgelehnt.

Irrtumswahrscheinlichkeit: Auftreten für die Wahrscheinlichkeit eines α-Fehlers.

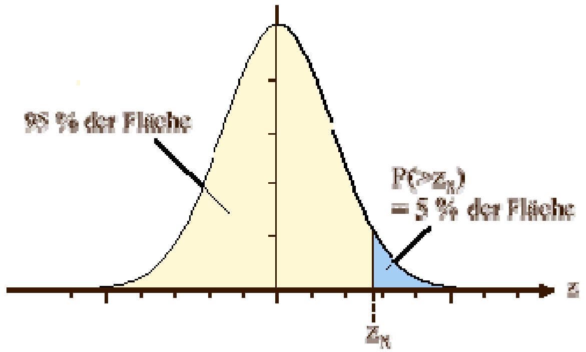


Es gibt immer die Wahrscheinlichkeit, dass eine Stichprobe zu einem Wert führt, der größer ist als der Grenzwert z_k.

Siehe quantitatives Beispiel mit den Schokoladentafeln aus dem Automaten.

Testen von Hypothesen

$$\bar{x} = 93\,\mathrm{g} \quad \text{bei einem} \quad \sigma_{\bar{x}} = 5\,\mathrm{g}$$



Gewicht der Schokoladentafeln

Diskussion mit einer 5% Irrtumswahrscheinlichkeit.

Suche des Grenzwertes z_k

Bei einer Normalverteilung liegen 5% der Ergebnisse oberhalb von 1,645 s.

$z_k = 101{,}2$ g

101,2g sind größer als mein Limit von 100,0g.

Daraus folgt, die Nullhypothese muss verworfen werden.

Diskussion mit einer 10% Irrtumswahrscheinlichkeit.

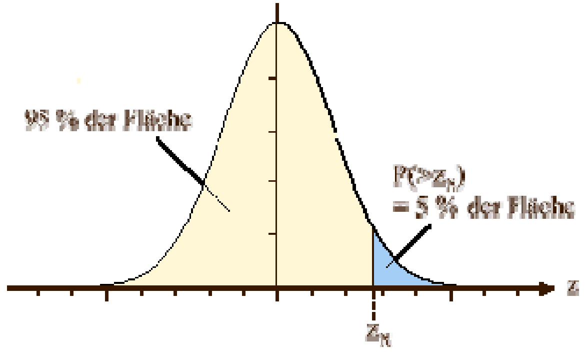
Oberhalb von 1,282 s liegen 10% der Ergebnisse.

$z_k = 99{,}4$ g

Daraus folgt, die Nullhypothese ist gültig.

Berechnung der Irrtumswahrscheinlichkeit

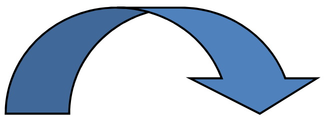
$\overline{x} = 93\text{g}$ bei einem $\sigma_{\overline{x}} = 5\text{g}$



Gewicht der Schokoladentafeln

Der Grenzwert unserer Nullhypothese
z_k = 100,0 g

Bei einer Normalverteilung liegt der Grenzwert von 100,0g etwa 1,40 σ oberhalb vom Stichprobenmittelwert von 93,0g.



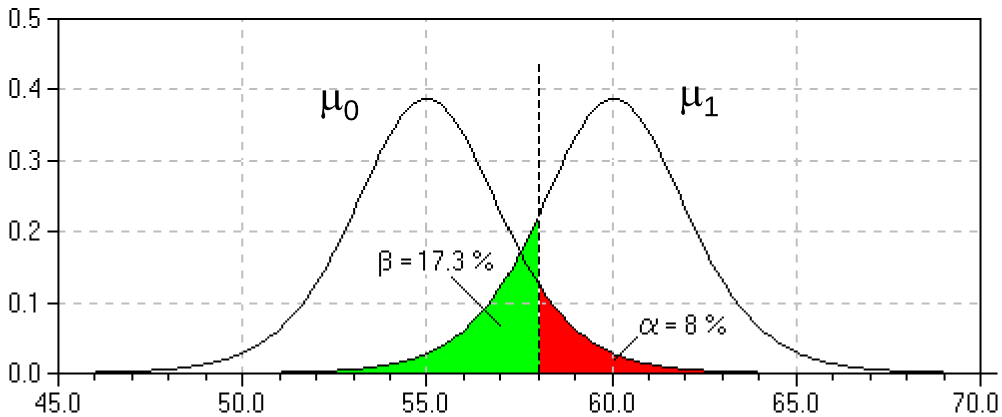
Oberhalb von 1,40 σ liegen 8,08% aller Ereignisse.
Wenn wir H_0^* annehmen, haben wir eine Irrtumswahrscheinlichkeit von 8,08 %.

Dies alles gilt nur, wenn die Stichprobe mehr als 30 Messungen enthält !

Teststärke

Die Teststärke gibt an, mit welcher Wahrscheinlichkeit ein Signifikanztest zu Gunsten einer richtigen Alternativhypothese H_1^* entscheidet.

Die Teststärke hat den Wert (1-β).



Die Teststärke (1-β) wird größer:

- mit wachsender Differenz zwischen Teilpopulationen
- mit kleiner werdender Merkmalsstreuung σ
- mit größer werdendem Stichprobenumfang
- (mit größer werdendem Signifikanzniveau α, wenn β nicht festgelegt ist)

Wie kann ich überprüfen, welche Verteilung meinen Daten zu Grunde liegt?

Chi-Quadrat-Test

Chi-Quadrat-Test

Es gibt verschiedene Arten von Signifikanztests

Neben Signifikanztests, die sich mit dem **Mittelwert** beschäftigen,

gibt es auch Testverfahren für **Verteilungen**.

Der Chi-Quadrat (χ^2) Test stellt eine solche Gruppe von Hypothesentests dar.

Damit können überprüft werden:

Sind vorliegende Daten auf eine bestimmte Art verteilt?

Sind verschiedene Merkmale voneinander unabhängig?

Folgen verschiedene Stichproben der selben Verteilung?

Bei Verteilungen hat man in der Regel die Beantwortung der Frage,

ob eine gemessene Verteilung

Gauß- oder Poisson-verteilt ist oder nicht ?

Chi-Quadrat-Test

$$\chi^2 = \sum_{k=1}^{n} \frac{\left(E_k - T_k\right)^2}{T_k}$$

Die E_k und T_k sind Häufigkeiten und nicht Wahrscheinlichkeiten

Nehmen wir zufällige Schwankungen in den experimentellen Verteilungen an, so erhalten wir für χ^2 wieder eine Verteilung:

$$f\left(\chi^2\right) = \frac{1}{2^{\frac{\nu}{2}}\left(\frac{\nu}{2} - 1\right)!}\left(\chi^2\right)^{\frac{\nu}{2}-1} \exp\left(-\frac{\chi^2}{2}\right)$$

Es wird die Wahrscheinlichkeit berechnet mit der χ^2 Werte auftreten

Chi-Quadrat-Test

Der Begriff der Freiheitsgrade

im allgemeinen ist die Anzahl der Freiheitsgrade ν in einer statistischen Rechnung definiert als
die Anzahl der beobachteten Klassen n minus der Anzahl der aus den Daten berechneten und/oder in der Rechnung verwendeten Parameter c.

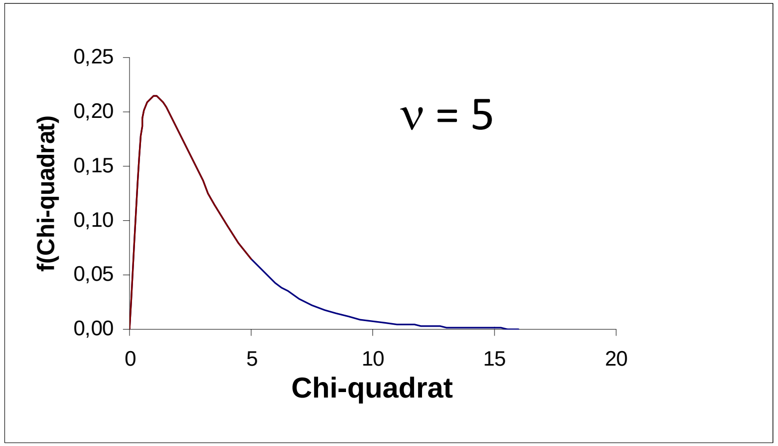
Die Anzahl der Freiheitsgrade ν ist:

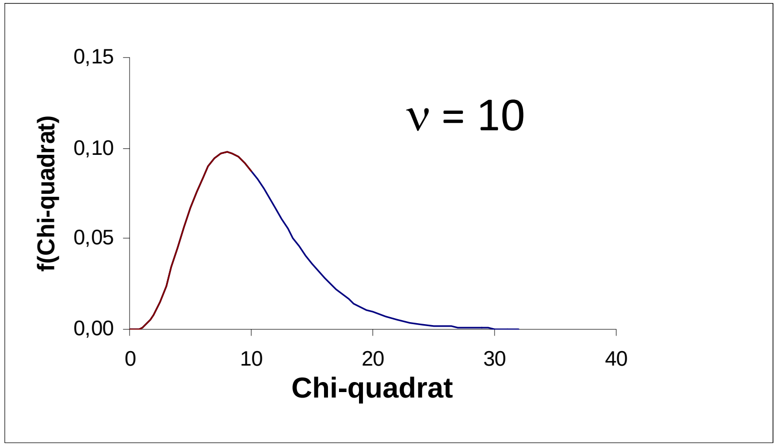
$$\nu = n - c,$$

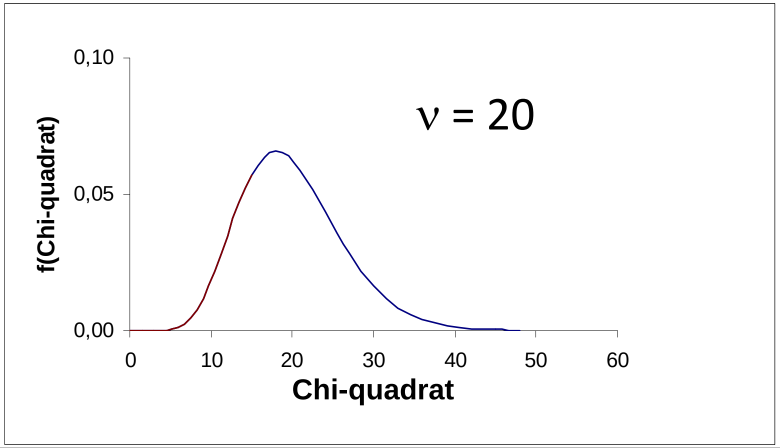
wobei c die Anzahl der Parameter ist, die aus den Daten berechnet werden mussten um die erwarteten Anzahlen T_k zu berechnen.

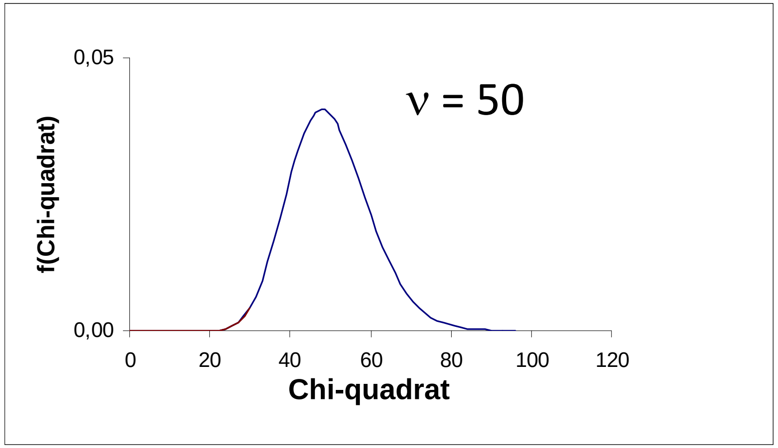
Chi-Quadrat-Test

$$f\left(\chi^2\right) = \frac{1}{2^{\frac{\nu}{2}}\left(\frac{\nu}{2}-1\right)!}\left(\chi^2\right)^{\frac{\nu}{2}-1}\exp\left(-\frac{\chi^2}{2}\right)$$

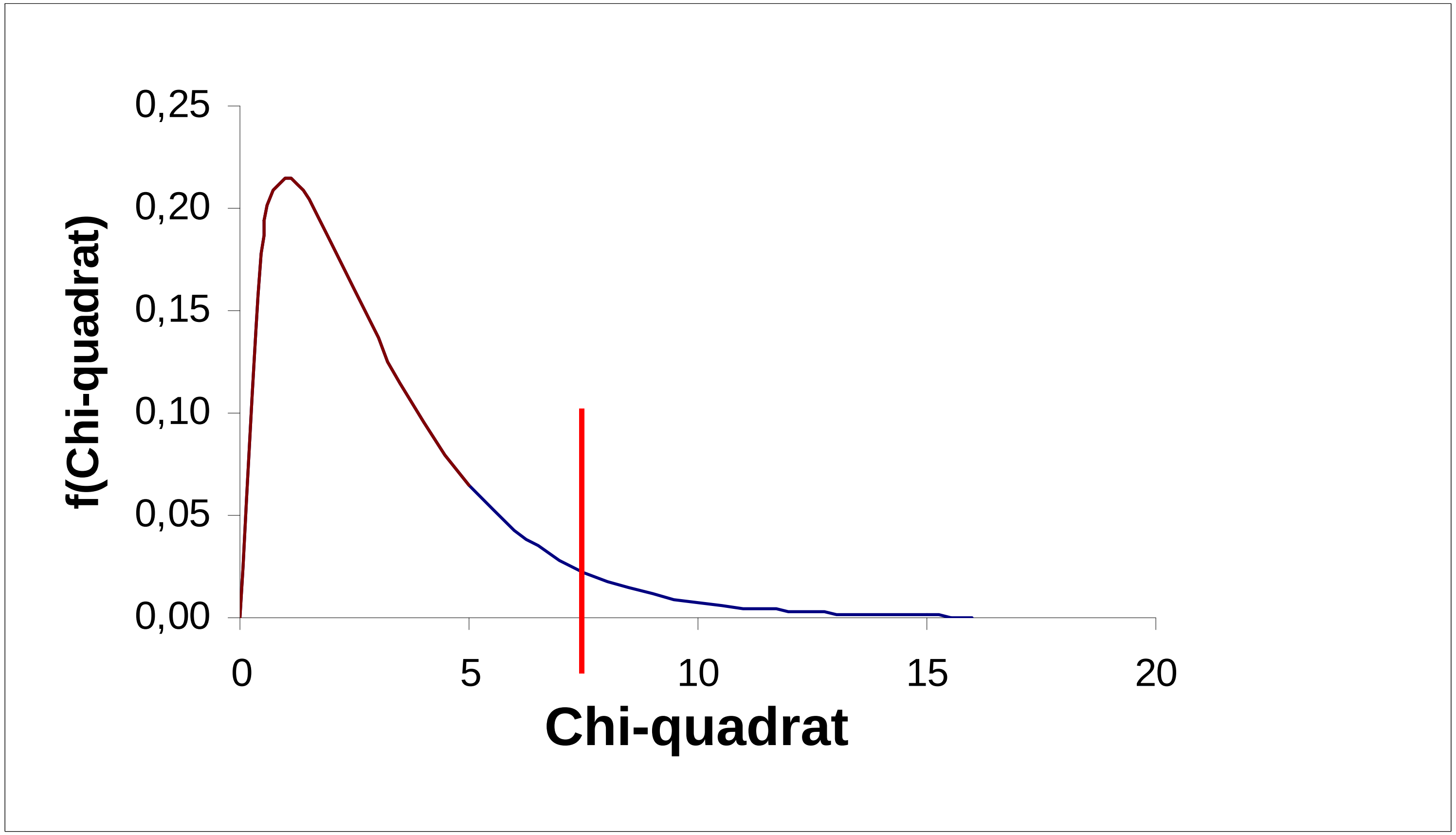








Chi-Quadrat-Test



Wie groß ist die Wahrscheinlichkeit χ^2 – Werte zu finden, die größer sind als ein bestimmter Wert?

Chi-Quadrat-Test

Sicherheitsschwellen:

ν \ P	0.990	0.975	0.950	0.900	0.800	0.700	0.600	0.500	0.400	0.300	0.200	0.100	0.050	0.025	0.010	0.001
1	0.0002	0.001	0.004	0.02	0.06	0.15	0.27	0.45	0.71	1.07	1.64	2.71	3.84	5.02	6.63	10.83
2	0.020	0.051	0.103	0.21	0.45	0.71	1.02	1.39	1.83	2.41	3.22	4.61	5.99	7.38	9.21	13.82
3	0.115	0.216	0.352	0.58	1.01	1.42	1.87	2.37	2.95	3.66	4.64	6.25	7.81	9.35	11.34	16.27
4	0.297	0.484	0.711	1.06	1.65	2.19	2.75	3.36	4.04	4.88	5.99	7.78	9.49	11.14	13.28	18.47
5	0.554	0.831	1.15	1.61	2.34	3.00	3.66	4.35	5.13	6.06	7.29	9.24	11.07	12.83	15.09	20.52
6	0.872	1.24	1.64	2.20	3.07	3.83	4.57	5.35	6.21	7.23	8.56	10.64	12.59	14.45	16.81	22.46
7	1.24	1.69	2.17	2.83	3.82	4.67	5.49	6.35	7.28	8.38	9.80	12.02	14.07	16.01	18.48	24.32
8	1.65	2.18	2.73	3.49	4.59	5.53	6.42	7.34	8.35	9.52	11.03	13.36	15.51	17.53	20.09	26.12
9	2.09	2.70	3.33	4.17	5.38	6.39	7.36	8.34	9.41	10.66	12.24	14.68	16.92	19.02	21.67	27.88
10	2.56	3.25	3.94	4.87	6.18	7.27	8.30	9.34	10.47	11.78	13.44	15.99	18.31	20.48	23.21	29.59
11	3.05	3.82	4.57	5.58	6.99	8.15	9.24	10.34	11.53	12.90	14.63	17.28	19.68	21.92	24.72	31.26
12	3.57	4.40	5.23	6.30	7.81	9.03	10.18	11.34	12.58	14.01	15.81	18.55	21.03	23.34	26.22	32.91
13	4.11	5.01	5.89	7.04	8.63	9.93	11.13	12.34	13.64	15.12	16.98	19.81	22.36	24.74	27.69	34.53
14	4.66	5.63	6.57	7.79	9.47	10.82	12.08	13.34	14.69	16.22	18.15	21.06	23.68	26.12	29.14	36.12
15	5.23	6.26	7.26	8.55	10.31	11.72	13.03	14.34	15.73	17.32	19.31	22.31	25.00	27.49	30.58	37.70
20	8.26	9.59	10.85	12.44	14.58	16.27	17.81	19.34	20.95	22.77	25.04	28.41	31.41	34.17	37.57	45.31
25	11.52	13.12	14.61	16.47	18.94	20.87	22.62	24.34	26.14	28.17	30.68	34.38	37.65	40.65	44.31	52.62
30	14.95	16.79	18.49	20.60	23.36	25.51	27.44	29.34	31.32	33.53	36.25	40.26	43.77	46.98	50.89	59.70
35	18.51	20.57	22.47	24.80	27.84	30.18	32.28	34.34	36.47	38.86	41.78	46.06	49.80	53.20	57.34	66.62
40	22.16	24.43	26.51	29.05	32.34	34.87	37.13	39.34	41.62	44.16	47.27	51.81	55.76	59.34	63.69	73.40
48	28.18	30.75	33.10	35.95	39.62	42.42	44.92	47.34	49.84	52.62	55.99	60.91	65.17	69.02	73.68	84.04
50	29.71	32.36	34.76	37.69	41.45	44.31	46.86	49.33	51.89	54.72	58.16	63.17	67.50	71.42	76.15	86.66

Chi-Quadrat-Test

Sicherheitsschwellen: Konkretes Beispiel

Auf einer vierspurigen Autobahn werden die Fahrzeuge pro Spur gezählt:

Bevorzugen die Fahrer eine Fahrbahn ?

Spur	1	2	3	4
Fahrzeuge	294	276	238	192

Insgesamt 1000 Fahrzeuge

Annahme: Die Fahrer bevorzugen keine Fahrbahn, das heißt $p_k = 1/4$;
somit ist $T_k = 250$ für alle k

$$\chi^2 = \frac{(294-250)^2}{250} + \frac{(276-250)^2}{250} + \frac{(238-250)^2}{250} + \frac{(192-250)^2}{250} = 24,48$$

Chi-Quadrat-Test

Sicherheitsschwellen: Konkretes Beispiel

Wie viele Freiheitsgrade?

Sobald N festliegt, können wir die Gleichung $N = \sum_{k=1}^{n} T_k$ als Zwangsbedingung auffassen.

Wir haben freie Wahl für die Häufigkeiten T_1,... T_{n-1}.

Die letzte Zahl T_n ist jedoch nicht mehr frei wählbar.

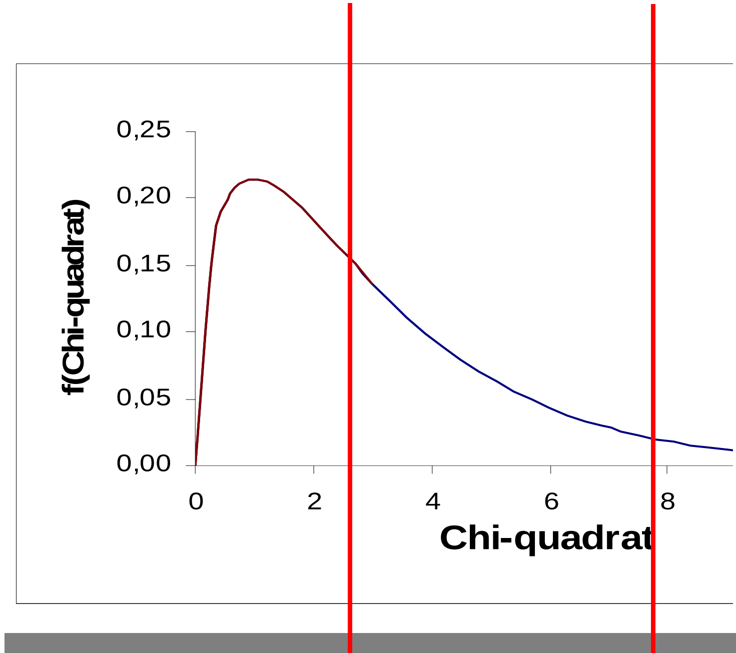
Sie muss so gewählt werden, dass die Summe der T_i die Anzahl N ergibt.

Wir werden sehen, dass die Zahl der Zwangsbedingungen größer als eins werden kann (Gaußverteilung etc.).

Chi-Quadrat-Test

Sicherheitsschwellen: Konkretes Beispiel

ν \ P	0.990	0.975	0.950	0.900	0.800	0.700	0.600	0.500	0.400	0.300	0.200	0.100	0.050	0.025	0.010	0.001
1	0.0002	0.001	0.004	0.02	0.06	0.15	0.27	0.45	0.71	1.07	1.64	2.71	3.84	5.02	6.63	10.83
2	0.020	0.051	0.103	0.21	0.45	0.71	1.02	1.39	1.83	2.41	3.22	4.61	5.99	7.38	9.21	13.82
3	0.115	0.216	0.352	0.58	1.01	1.42	1.87	2.37	2.95	3.66	4.64	6.25	7.81	9.35	11.34	16.27
4	0.297	0.484	0.711	1.06	1.65	2.19	2.75	3.36	4.04	4.88	5.99	7.78	9.49	11.14	13.28	18.47
5	0.554	0.831	1.15	1.61	2.34	3.00	3.66	4.35	5.13	6.06	7.29	9.24	11.07	12.83	15.09	20.52
6	0.872	1.24	1.64	2.20	3.07	3.83	4.57	5.35	6.21	7.23	8.56	10.64	12.59	14.45	16.81	22.46
7	1.24	1.69	2.17	2.83	3.82	4.67	5.49	6.35	7.28	8.38	9.80	12.02	14.07	16.01	18.48	24.32



Beispiel Autobahn:

Vier Fahrspuren: $\nu = 3$

$$\chi^2 = 24{,}48$$

Fazit:

So große Abweichungen kommen in weniger als 0,1% der Fälle vor.

Chi-Quadrat-Test

Weiteres Beispiel: Bedeutung der Sicherheitsschwellen

Rechteckverteilungen (Gleichverteilungen) kommen beim Würfeln vor,

hier ist die Wahrscheinlichkeit für das Würfeln einer Zahl $p_k=1/6$

Wir würfeln 240 Mal mit einem Würfel A

Augenzahl	1	2	3	4	5	6
Häufigkeit	35	42	49	37	45	32

$$\chi^2 = 5.20$$

Wir würfeln 360 Mal mit einem Würfel B

Augenzahl	1	2	3	4	5	6
Häufigkeit	50	66	80	50	53	60

$$\chi^2 = 12.15$$

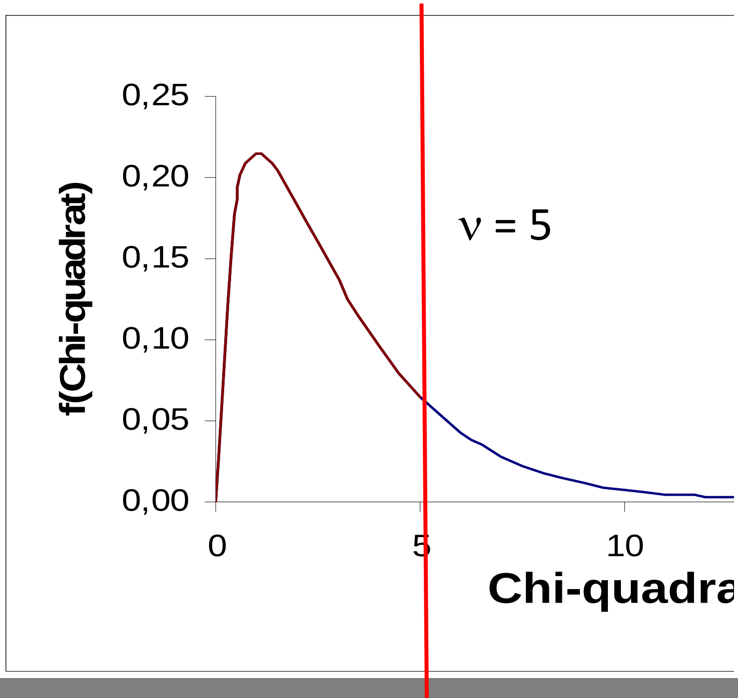
Hypothese: Würfel B ist gezinkt!

Testen der Hypothese mit einer Irrtumswahrscheinlichkeit von 5%, 1% und 0.1%

Chi-Quadrat-Test

Weiteres Beispiel: Bedeutung der Sicherheitsschwellen

P / ν	0.990	0.975	0.950	0.900	0.800	0.700	0.600	0.500	0.400	0.300	0.200	0.100	0.050	0.025	0.010	0.001
1	0.0002	0.001	0.004	0.02	0.06	0.15	0.27	0.45	0.71	1.07	1.64	2.71	3.84	5.02	6.63	10.83
2	0.020	0.051	0.103	0.21	0.45	0.71	1.02	1.39	1.83	2.41	3.22	4.61	5.99	7.38	9.21	13.82
3	0.115	0.216	0.352	0.58	1.01	1.42	1.87	2.37	2.95	3.66	4.64	6.25	7.81	9.35	11.34	16.27
4	0.297	0.484	0.711	1.06	1.65	2.19	2.75	3.36	4.04	4.88	5.99	7.78	9.49	11.14	13.28	18.47
5	0.554	0.831	1.15	1.61	2.34	3.00	3.66	4.35	5.13	6.06	7.29	9.24	11.07	12.83	15.09	20.52
6	0.872	1.24	1.64	2.20	3.07	3.83	4.57	5.35	6.21	7.23	8.56	10.64	12.59	14.45	16.81	22.46
7	1.24	1.69	2.17	2.83	3.82	4.67	5.49	6.35	7.28	8.38	9.80	12.02	14.07	16.01	18.48	24.32



Beispiel Würfel A:

Sechs Seiten: $\nu = 5$

$\chi^2 = 5{,}20$

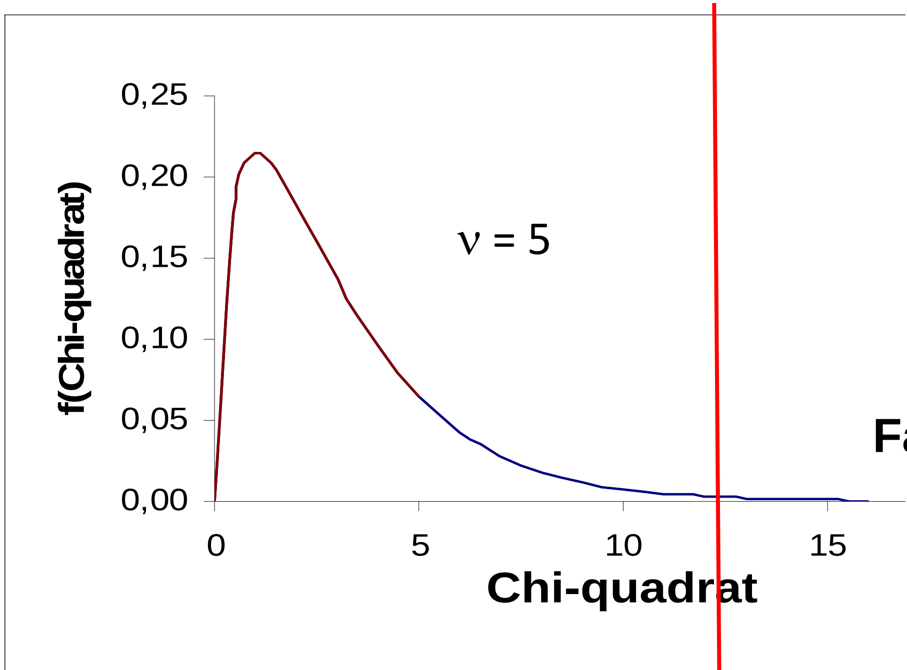
Fazit:

So große Abweichungen kommen in fast 40% der Fälle vor.

Chi-Quadrat-Test

Weiteres Beispiel: Bedeutung der Sicherheitsschwellen

ν \ P	0.990	0.975	0.950	0.900	0.800	0.700	0.600	0.500	0.400	0.300	0.200	0.100	0.050	0.025	0.010	0.001
1	0.0002	0.001	0.004	0.02	0.06	0.15	0.27	0.45	0.71	1.07	1.64	2.71	3.84	5.02	6.63	10.83
2	0.020	0.051	0.103	0.21	0.45	0.71	1.02	1.39	1.83	2.41	3.22	4.61	5.99	7.38	9.21	13.82
3	0.115	0.216	0.352	0.58	1.01	1.42	1.87	2.37	2.95	3.66	4.64	6.25	7.81	9.35	11.34	16.27
4	0.297	0.484	0.711	1.06	1.65	2.19	2.75	3.36	4.04	4.88	5.99	7.78	9.49	11.14	13.28	18.47
5	0.554	0.831	1.15	1.61	2.34	3.00	3.66	4.35	5.13	6.06	7.29	9.24	11.07	12.83	15.09	20.52
6	0.872	1.24	1.64	2.20	3.07	3.83	4.57	5.35	6.21	7.23	8.56	10.64	12.59	14.45	16.81	22.46
7	1.24	1.69	2.17	2.83	3.82	4.67	5.49	6.35	7.28	8.38	9.80	12.02	14.07	16.01	18.48	24.32



Beispiel Würfel gezinkt (?):

Sechs Seiten: $\nu = 5$

$\chi^2 = 12,15$

Fazit:

So große Abweichungen kommen in weniger als 5% der Fälle vor.

Chi-Quadrat-Test

Weiteres Beispiel: Bedeutung der Sicherheitsschwellen

Hypothese: Meine Kumpels sind Ganoven !

Wenn dem so ist, dann erhalte ich bei der Messung einen Wert, der größer ist als der Schwellwert

Würfel A $\chi^2 = 5.20$ Würfel B $\chi^2 = 12.15$

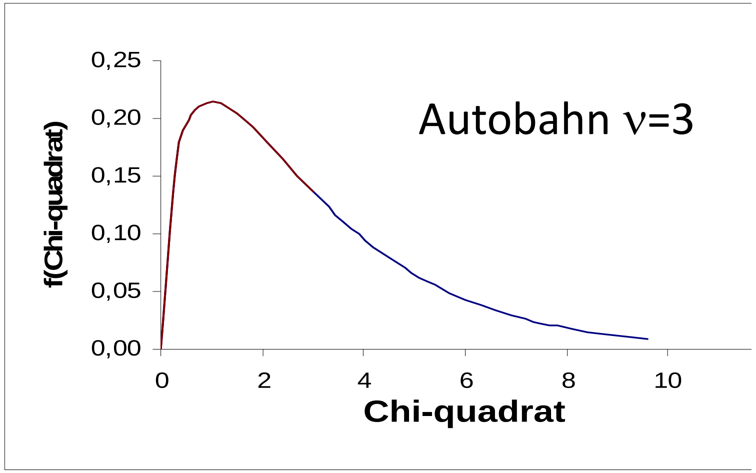
		Würfel A	Würfel B
Irrtums-wahrscheinlichkeit	Schwellwert	Hypothese	Hypothese
5 %	11.07	Verwerfen	Bestätigt
1%	15.09	Verwerfen	Verwerfen
0.1%	20.52	Verwerfen	Verwerfen

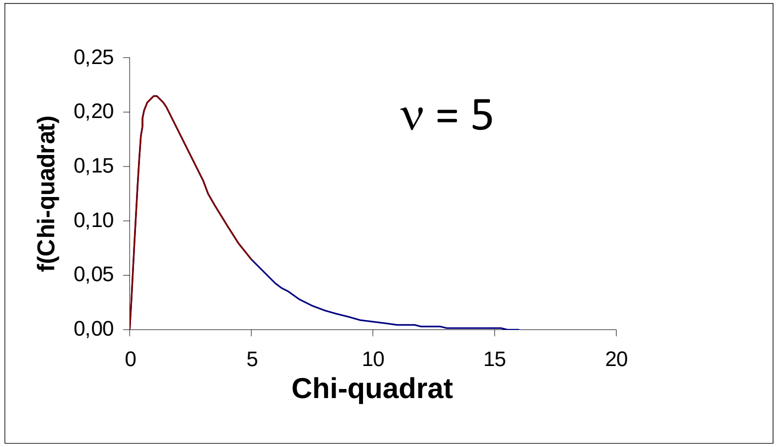
Chi-Quadrat-Test

Weiteres Beispiel: Bedeutung der Sicherheitsschwellen

Sicherheitsschwelle $\neq$ Schwellwert

Beispiel	χ^2	Freiheits-grad	P = 90 %	P = 1%	P = 0.1%	P(χ^2)
Autobahn	24.48	3	0.58	11.34	16.27	< 0.1%
Würfel	5.20	5	1.61	15.09	20.52	0.391





Chi-Quadrat-Test

Berechnen von Zwischenwerten der Wahrscheinlichkeit

Um aus der Tabelle Zwischenwerte für P zu errechnen,
interpoliert man logarithmisch

Beispiel Würfel: $\chi^2_{EX} = 5{,}20$

ν \ P	0.990	0.975	0.950	0.900	0.800	0.700	0.600	0.500	0.400	0.300	0.200	0.100	0.050	0.025	0.010	0.001
1	0.0002	0.001	0.004	0.02	0.06	0.15	0.27	0.45	0.71	1.07	1.64	2.71	3.84	5.02	6.63	10.83
2	0.020	0.051	0.103	0.21	0.45	0.71	1.02	1.39	1.83	2.41	3.22	4.61	5.99	7.38	9.21	13.82
3	0.115	0.216	0.352	0.58	1.01	1.42	1.87	2.37	2.95	3.66	4.64	6.25	7.81	9.35	11.34	16.27
4	0.297	0.484	0.711	1.06	1.65	2.19	2.75	3.36	4.04	4.88	5.99	7.78	9.49	11.14	13.28	18.47
5	0.554	0.831	1.15	1.61	2.34	3.00	3.66	4.35	5.13	6.06	7.29	9.24	11.07	12.83	15.09	20.52
6	0.872	1.24	1.64	2.20	3.07	3.83	4.57	5.35	6.21	7.23	8.56	10.64	12.59	14.45	16.81	22.46
7	1.24	1.69	2.17	2.83	3.82	4.67	5.49	6.35	7.28	8.38	9.80	12.02	14.07	16.01	18.48	24.32

$\chi^2_{P1} = 5{,}13$

$\chi^2_{EX} = 5{,}20$

$\chi^2_{P2} = 6{,}06$

Chi-Quadrat-Test

Berechnen von Zwischenwerten der Wahrscheinlichkeit

Um aus der Tabelle Zwischenwerte für P zu errechnen, interpoliert man logarithmisch

P_1	P	P_2
χ_1^2	χ^2	χ_2^2
5,13	5,20	6,06
0,400	**0,391**	0,300

$$\frac{\ln P - \ln P_1}{\ln P_2 - \ln P_1} = \frac{\chi^2 - \chi_1^2}{\chi_2^2 - \chi_1^2}$$

$$\begin{aligned}
\ln P &= \frac{\left(\chi^2 - \chi_1^2\right)\left(\ln P_2 - \ln P_1\right)}{\chi_2^2 - \chi_1^2} + \ln P_1 \\
&= \frac{\left(5{,}20 - 5{,}13\right)\left(-1{,}20397 + 0{,}91629\right)}{6{,}06 - 5{,}13} - 0{,}91629 \\
&= -0{,}93971
\end{aligned}$$

Chi-Quadrat-Test: Allgemeines Vorgehen

1.) Aufstellen einer Hypothese — Alle Fahrbahnen werden gleichmäßig benutzt

2.) Festlegen einer Irrtumswahrscheinlichkeit — 5 %

3.) Bestimmung der Schwellwertes (Tabelle) — $\chi^2_{SW} = 7.81$

4.) Vergleich des Schwellwertes mit dem experimentellen Wert — $\chi^2_{EX} = 24.48$

Das χ^2 des Schwellwertes ist kleiner als der experimentelle Wert

Das heißt, in 5% aller Fälle gibt es Werte, die größer sind als 7.81.
Der experimentelle Wert ist deutlich größer als die Schwelle,
solche Werte sind also extrem selten

Die Hypothese muss verworfen werden.

Chi-Quadrat-Test: Test auf Poisson Verteilung

Sind vorliegende Daten auf eine bestimmte Art verteilt?

k	$\mathbf{E_k}$
0	54
1	97
2	90
3	55
4	17
5	19
6	3
7	1
8	0

Insgesamt N = 336 Messungen

Mittelwert $\mu = 1.875$

Beide Daten werden benötigt, um die theoretische Häufigkeit zu berechnen.

Festlegen der Irrtumswahrscheinlichkeit: Testen auf 5 %

Chi-Quadrat-Test: Test auf Poisson Verteilung

Sind vorliegende Daten auf eine bestimmte Art verteilt?

k	$\mathbf{E_k}$	$\mathbf{T_k}$
0	54	51.53
1	97	96.61
2	90	90.58
3	55	56.61
4	17	26.54
5	19	9.95
6	3	3.11
7	1	0.83
8	0	0.24

Insgesamt N = 336 Messungen

Mittelwert $\mu = 1.875$

Beide Daten werden benötigt, um die theoretische Häufigkeit zu berechnen.

9 Klassen – 2 Zwangsbedingungen

$$\chi^2 = \sum_{k=1}^{n} \frac{\left(E_k - T_k\right)^2}{T_k} = \sum_{k=1}^{n} Q_k$$

Chi-Quadrat-Test: Test auf Poisson Verteilung

Sind vorliegende Daten auf eine bestimmte Art verteilt?

k	**E_k**	**T_k**	**Q_k**
0	54	51.53	0.12
1	97	96.61	0.00
2	90	90.58	0.00
3	55	56.61	0.05
4	17	26.54	3.43
5	19	9.95	8.23
6	3	3.11	0.00
7	1	0.83	0.03
8	0	0.24	0.24

Chi-Quadrat **12.10**

Freiheitsgrade **7**

Insgesamt N = 336 Messungen

Mittelwert $\mu = 1.875$

Beide Daten werden benötigt, um die theoretische Häufigkeit zu berechnen.

9 Klassen – 2 Zwangsbedingungen

Chi-Quadrat-Test

Sicherheitsschwellen:

ν \ P	0.990	0.975	0.950	0.900	0.800	0.700	0.600	0.500	0.400	0.300	0.200	0.100	0.050	0.025	0.010	0.001
1	0.0002	0.001	0.004	0.02	0.06	0.15	0.27	0.45	0.71	1.07	1.64	2.71	3.84	5.02	6.63	10.83
2	0.020	0.051	0.103	0.21	0.45	0.71	1.02	1.39	1.83	2.41	3.22	4.61	5.99	7.38	9.21	13.82
3	0.115	0.216	0.352	0.58	1.01	1.42	1.87	2.37	2.95	3.66	4.64	6.25	7.81	9.35	11.34	16.27
4	0.297	0.484	0.711	1.06	1.65	2.19	2.75	3.36	4.04	4.88	5.99	7.78	9.49	11.14	13.28	18.47
5	0.554	0.831	1.15	1.61	2.34	3.00	3.66	4.35	5.13	6.06	7.29	9.24	11.07	12.83	15.09	20.52
6	0.872	1.24	1.64	2.20	3.07	3.83	4.57	5.35	6.21	7.23	8.56	10.64	12.59	14.45	16.81	22.46
7	1.24	1.69	2.17	2.83	3.82	4.67	5.49	6.35	7.28	8.38	9.80	12.02	14.07	16.01	18.48	24.32
8	1.65	2.18	2.73	3.49	4.59	5.53	6.42	7.34	8.35	9.52	11.03	13.36	15.51	17.53	20.09	26.12
9	2.09	2.70	3.33	4.17	5.38	6.39	7.36	8.34	9.41	10.66	12.24	14.68	16.92	19.02	21.67	27.88
10	2.56	3.25	3.94	4.87	6.18	7.27	8.30	9.34	10.47	11.78	13.44	15.99	18.31	20.48	23.21	29.59
11	3.05	3.82	4.57	5.58	6.99	8.15	9.24	10.34	11.53	12.90	14.63	17.28	19.68	21.92	24.72	31.26
12	3.57	4.40	5.23	6.30	7.81	9.03	10.18	11.34	12.58	14.01	15.81	18.55	21.03	23.34	26.22	32.91
13	4.11	5.01	5.89	7.04	8.63	9.93	11.13	12.34	13.64	15.12	16.98	19.81	22.36	24.74	27.69	34.53
14	4.66	5.63	6.57	7.79	9.47	10.82	12.08	13.34	14.69	16.22	18.15	21.06	23.68	26.12	29.14	36.12
15	5.23	6.26	7.26	8.55	10.31	11.72	13.03	14.34	15.73	17.32	19.31	22.31	25.00	27.49	30.58	37.70
20	8.26	9.59	10.85	12.44	14.58	16.27	17.81	19.34	20.95	22.77	25.04	28.41	31.41	34.17	37.57	45.31
25	11.52	13.12	14.61	16.47	18.94	20.87	22.62	24.34	26.14	28.17	30.68	34.38	37.65	40.65	44.31	52.62
30	14.95	16.79	18.49	20.60	23.36	25.51	27.44	29.34	31.32	33.53	36.25	40.26	43.77	46.98	50.89	59.70
35	18.51	20.57	22.47	24.80	27.84	30.18	32.28	34.34	36.47	38.86	41.78	46.06	49.80	53.20	57.34	66.62
40	22.16	24.43	26.51	29.05	32.34	34.87	37.13	39.34	41.62	44.16	47.27	51.81	55.76	59.34	63.69	73.40
48	28.18	30.75	33.10	35.95	39.62	42.42	44.92	47.34	49.84	52.62	55.99	60.91	65.17	69.02	73.68	84.04
50	29.71	32.36	34.76	37.69	41.45	44.31	46.86	49.33	51.89	54.72	58.16	63.17	67.50	71.42	76.15	86.66

Chi-Quadrat-Test: Test auf Poisson Verteilung

Sind vorliegende Daten auf eine bestimmte Art verteilt?

k	**E_k**	**T_k**	**Q_k**
0	54	51.53	0.12
1	97	96.61	0.00
2	90	90.58	0.00
3	55	56.61	0.05
4	17	26.54	3.43
5	19	9.95	8.23
6	3	3.11	0.00
7	1	0.83	0.03
8	0	0.24	0.24

Chi-Quadrat **12.10**

Freiheitsgrade **7**
Sicherheitsschwelle **0.097**

Insgesamt N = 336 Messungen

Mittelwert $\mu = 1.875$

Beide Daten werden benötigt, um die theoretische Häufigkeit zu berechnen.

9 Klassen – 2 Zwangsbedingungen

Bei Annahme einer Irrtumswahrscheinlichkeit von 5 % entsprechen die Daten einer Poisson-Verteilung

Chi-Quadrat-Test: Test auf Poisson Verteilung

Poisson-Verteilung

Vergleich zweier fast identischer Messungen (Mittelwert ist identisch)

k	Ek	Tk	Qk
0	54	51.53	0.12
1	97	96.61	0.00
2	90	90.58	0.00
3	55	56.61	0.05
4	17	26.54	3.43
5	19	9.95	8.23
6	3	3.11	0.00
7	1	0.83	0.03
8	0	0.24	0.24

Bei Annahme einer Irrtumswahrscheinlichkeit von 5 % entsprechen die rechten Daten keiner Poisson-Verteilung

????????

k	Ek	Tk	Qk	ΔQk
0	55	51.53	0.23	-0.11
1	96	96.61	0.00	-0.00
2	90	90.58	0.00	
3	55	56.61	0.05	
4	17	26.54	3.43	
5	19	9.95	8.23	
6	3	3.11	0.00	
7	0	0.83	0.83	-0.80
8	1	0.24	2.41	-2.17

Chi-Quadrat	**15.18**
Freiheitsgrade	**7**
Sicherheitsschwelle	**0.034**

Chi-Quadrat-Test: Test auf Poisson Verteilung

Der χ^2 Test darf nur angewandt werden,
wenn die theoretische Häufigkeit einer Klasse $\geq$ 5 ist

k	**E_k**	**T_k**	**Q_k**
0	54	51.53	0.12
1	97	96.61	0.00
2	90	90.58	0.00
3	55	56.61	0.05
4	17	26.54	3.43
5	19	9.95	8.23
6	3	3.11	0.00
7	1	0.83	0.03
8	0	0.24	0.24

3.11 + 0.83 + 0.24 = 4.18 < 5

Zusammenfassung von Klassen, so lange bis die Bedingung erfüllt ist

Chi-Quadrat-Test: Test auf Poisson Verteilung

Der χ^2 Test darf nur angewandt werden,
wenn die theoretische Häufigkeit einer Klasse $\geq$ 5 ist

k	**E_k**	**T_k**	**Q_k**
0	54	51.53	0.12
1	97	96.61	0.00
2	90	90.58	0.00
3	55	56.61	0.05
4	17	26.54	3.43
5	19	9.95	8.23
6	3	3.11	0.00
7	1	0.83	0.03
8	0	0.24	0.24

9.95 + 3.11 + 0.83 + 0.24 = 14.13 > 5

Zusammenfassung von Klassen, so lange bis die Bedingung erfüllt ist

Dies führt zu einer neuen Tabelle

Chi-Quadrat-Test: Test auf Poisson Verteilung

Poisson-Verteilung

Vergleich zweier fast identischer Messungen

k	Ek	Tk	Qk	k	Ek	Tk	Qk	Qk
0	54	51.53	0.12	0	55	51.53	0.23	-0.11
1	97	96.61	0.00	1	96	96.61	0.00	-0.00
2	90	90.58	0.00	2	90	90.58	0.00	
3	55	56.61	0.05	3	55	56.61	0.05	
4	17	26.54	3.43	4	17	26.54	3.43	
5....	23	14.13	5.57	5...	23	14.13	5.57	

Chi-Quadrat	**9.17**	**Chi-Quadrat**	**9.28**
Freiheitsgrade	**4**	**Freiheitsgrade**	**4**
Sicherheitsschwelle	**>0.05**	**Sicherheitsschwelle**	**> 0.05**

Beide Verteilungen sind mit einer Irrtumswahrscheinlichkeit von 5% Poisson-verteilt

Chi-Quadrat-Test: Test auf Normal-Verteilung

Alle bislang diskutierten Messgrößen waren diskret.
Es gab keine Schwierigkeiten bei der Klasseneinteilung.

Die Gauß-Verteilung ist eine kontinuierliche Verteilung.
Es gibt andere häufig vorkommende kontinuierliche Verteilungen
(Lorentz-, Maxwell-, Boltzmann-, Fermi-Verteilung sind in der Physik sehr häufig).

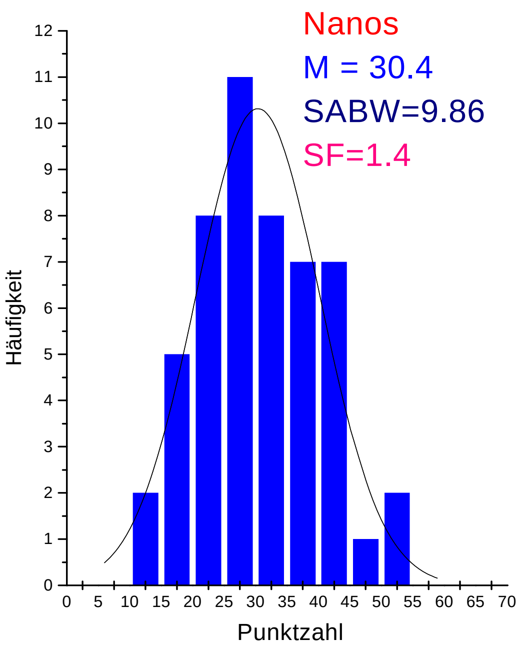
Wie immer die Verteilung aussehen mag, die gesamte Fläche
unter dem Graphen $f(x)$ aufgetragen gegen x, ist gleich 1
und die Wahrscheinlichkeit eines Messwertes zwischen
$x = a$ und $x = b$
ist gleich der Fläche zwischen a und b.

$$P(a < x < b) = \int_a^b f(x)\,dx$$

Chi-Quadrat-Test: Test auf Normal-Verteilung

Wenn also die k-te Klasse von $x = a_k$ bis $x = a_{k+1}$ läuft,
ist die (nach insgesamt N Messungen) erwartete
Anzahl von Messwerten in der k-ten Klasse:

$$T_k = N \cdot P\left(a_k \leq x < a_{k+1}\right) = N \cdot \int_{a_k}^{a_{k+1}} f\left(x\right)\, dx$$



Beispiele:
Klausurergebnisse
Körpergröße von Menschen

Aufwändiges Verfahren