Beurteilende (induktive) Statistik

Die Aufgabe der Beurteilenden Statistik besteht darin, mit Hilfe der Wahrscheinlichkeitsrechnung aus einer Zufallsstichprobe Rückschlüsse auf die Grundgesamtheit zu ermöglichen. Dazu gehören u.a. die Fragestellungen:

Das Schätzproblem (Punktschätzung, Vertrauensintervalle)
Das Testen von Hypothesen

1. Das Schätzproblem

Punktschätzung

Die Wahrscheinlichkeitsverteilung einer Stichprobe ist vom Typ her in vielen Fällen bekannt, sie enthält aber noch unbekannte Parameter. Wie können aus einer konkreten Stichprobe diese Parameter (Wahrscheinlichkeit, Erwartungswert oder Standardabweichung) geschätzt werden?

Beispiele:
Wir nehmen an, dass jeder Chip eines bestimmten Herstellers mit Wahrscheinlichkeit p in Ordnung ist. Findet man in einer Stichprobe von 1000 Chips 921 akzeptable, dann ist plausibelste Schätzung (die sogenannte Maximum-Likelihood-Schätzung) für p gerade die relative Häufigkeit 0.921.

allg.
Ergeben sich in einer Stichprobe vom Umfang n genau g Erfolge, dann wird die unbekannte Erfolgswahrscheinlichkeit p geschätzt durch die relative Häufigkeit

$$p \approx \hat{p} = \frac{g}{n}$$

Am Anfang Kapitels Wahrscheinlichkeitsrechnung wurden nach diesem plausiblen Verfahren die Wahrscheinlichkeiten geschätzt (vgl. etwas Reisnagelbeispiel).

Schätzer für Erwartungswert und Standardabweichung einer normalverteilten Gesamtheit
Schätzer für den Erwartungswert: $\hat{\mu} \approx \frac{1}{n} \cdot \sum_i x_i$ empirischer Mittelwert

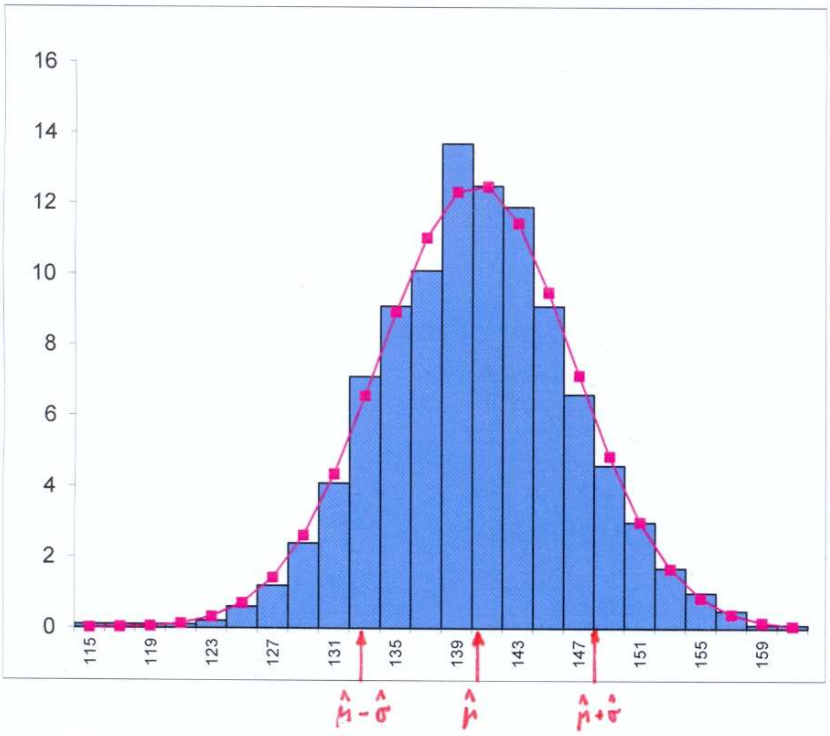
Schätzer für die Varianz bzw. Standardabweichung: $\hat{\sigma}^2 = \frac{1}{n-1} \cdot \sum_i (x_i - \overline{x})^2$

Nach diesen Verfahren können in den folgenden Beispielen „Körpergröße“ und „Widerstand“ die Parameter Mittelwert und Varianz bzw. Standardabweichung geschätzt werden.

Körpergrössen von 4922 Jugendlichen im Alter zwischen 10 und 11 Jahren

Die Daten stammen von Jugendgesundheitsuntersuchungen an Volksschulen Bayerns im Jahre 1965/66

Körpergrösse in cm	rel. Häufigkeit in %	Dichte der Normalverteilung
115	0.1	0.0
117	0.1	0.0
119	0.1	0.0
121	0.1	0.1
123	0.2	0.3
125	0.6	0.7
127	1.2	1.4
129	2.4	2.6
131	4.1	4.4
133	7.1	6.6
135	9.1	8.9
137	10.1	11.0
139	13.7	12.3
141	12.5	12.5
143	11.9	11.4
145	9.1	9.5
147	6.6	7.1
149	4.6	4.9
151	3.0	3.0
153	1.7	1.7
155	1	0.8
157	0.5	0.4
159	0.1	0.2
161	0.1	0.1
		100.0



emp. Mittelwert $\hat{\mu}$ = **140.246**

emp. Varianz 40.31

emp. Standardabweichung $\hat{\sigma}$ = **6.35**

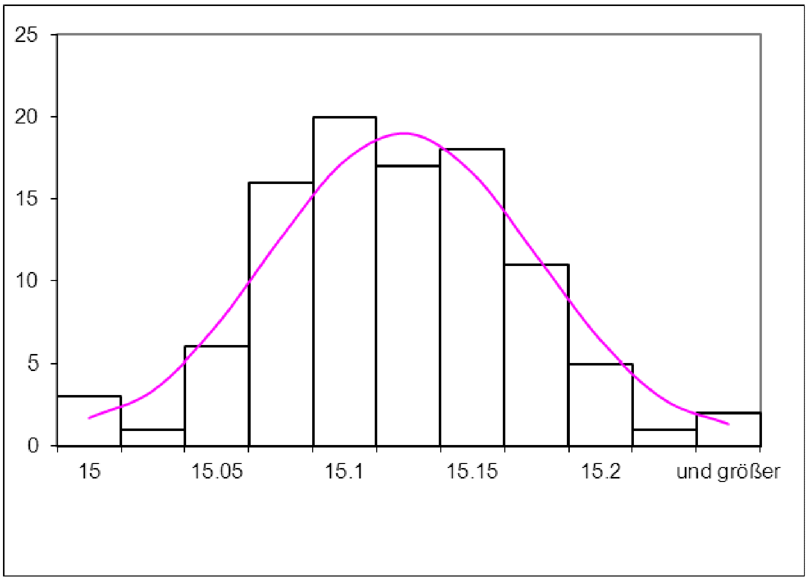
Stichprobe von n=100 Widerständen spezifiziert als 15 kOhm +/- 2% (bei 20°C)

15.11 Erwartungswert
(Schätzung durch emp.Mittelwert der Daten)
0.052 Standardabweichung
(Schätzung durch emp. Standardabweichung)

	KiloOhm
	14.950
21	**14.980**
80	15.000
61	15.010
45	15.040
39	**15.040**
64	15.050
19	15.050
26	15.050
53	15.050
85	15.060
7	15.060
15	15.060
16	15.060
49	15.060
51	15.060
56	15.060
76	15.060
81	15.070
18	15.070
33	15.070
35	15.070
54	15.070
58	15.070
68	15.070
75	15.070
82	15.080
10	15.080
22	15.080
28	15.080
36	15.080
40	15.080
42	15.080
52	15.090
3	15.090
29	15.090
41	15.090
43	15.090
74	15.090
89	15.100
27	15.100

Messdaten

Klasse	*Häufigkeit*
15	**3**
15.025	**1**
15.05	**6**
15.075	**16**
15.1	**20**
15.125	**17**
15.15	**18**
15.175	**11**
15.2	**5**
15.225	**1**
und größer	**2**
Summe	100



Vertrauenintervalle

Gibt man statt einer Zahl ein Intervall an, in dem der Parameter mit grosser Wahrscheinlichkeit liegen wird, so spricht man von einer Intervallschätzung bzw. einem Vertrauensintervall (Konfidenzintervall).

Einführende Aufgabe:
Vertrauensintervall für die Wahrscheinlichkeit einer Mädchengeburt (nach Ineichen)
Von den n = 2 644 757 Geburten in der Schweiz (1871 – 1900) waren k = 1 285 086 Mädchengeburten. Gesucht ist ein 95%-Vertrauensintervall für die (unbekannte) Wahrscheinlichkeit einer Mädchengeburt.
Die Anzahl der Mädchengeburten k ist der Wert einer binomialverteilten Zufallsvariablen, die durch die Normalverteilung angenähert werden kann. K erfüllt mit der Wahrscheinlichkeit 1 - α die Ungleichungen

$-u_{1-\alpha/2} < \frac{k-np}{\sqrt{npq}} < u_{1-\alpha/2}$ oder nach p aufgelöst

$$-u_{1-\alpha/2} \cdot \sqrt{npq} < k - np < u_{1-\alpha/2} \cdot \sqrt{npq}$$

$$-u_{1-\alpha/2} \cdot \sqrt{\frac{pq}{n}} < \frac{k}{n} - p < u_{1-\alpha/2} \cdot \sqrt{\frac{pq}{n}}$$

$$-\frac{k}{n} - u_{1-\alpha/2} \cdot \sqrt{\frac{pq}{n}} < -p < -\frac{k}{n} + u_{1-\alpha/2} \cdot \sqrt{\frac{pq}{n}}$$

$\frac{k}{n} + u_{1-\alpha/2} \cdot \sqrt{\frac{pq}{n}} > p > \frac{k}{n} - u_{1-\alpha/2} \cdot \sqrt{\frac{pq}{n}}$ oder

$$\frac{k}{n} - u_{1-\alpha/2} \cdot \sqrt{\frac{pq}{n}} < p < \frac{k}{n} + u_{1-\alpha/2} \cdot \sqrt{\frac{pq}{n}}$$

Die Aussage bedeutet, dass dieses Intervall das unbekannte p mit der Wahrscheinlichkeit 1 - α überdeckt. Der unbekannte Wert für p wird mit k/n geschätzt: d.h. es gilt ungefähr:

$$\sqrt{\frac{pq}{n}} \approx \sqrt{\frac{\frac{k}{n} \cdot (1 - \frac{k}{n})}{n}}$$

Ausserdem wird das Intervall noch etwas vergrößert.

Führt man zur Abkürzung die Größe d ein mit $d := u_{1-\alpha/2} \cdot \sqrt{\frac{\frac{k}{n} \cdot (1 - \frac{k}{n})}{n}}$ so erhält man das folgende Ergebnis

Das Intervall $[\frac{k-\frac{1}{2}}{n} - d, \frac{k+\frac{1}{2}}{n} + d]$ überdeckt das unbekannte p mit der Wahrscheinlich-keit 1 - α.
Ist α = 5% und damit 1 - α = 95% so spricht man von einem 95%-Vertrauensintervall.
Im Geburtenbeispiel ergibt sich als 95%-Vertrauensintervall für die Wahrscheinlichkeit einer Mädchengeburt
$[0.4853, 0.4865]$

Aufgabe:
Eine Urne enthält schwarze und weisse Kugeln, deren Anteile unbekannt sind. Um den Anteil p an schwarzen Kugeln zu schätzen, werden der Urne 100 Kugeln mit Zurücklegen gezogen. Gesucht ist ein 95%-Vertrauensintervall für p, wenn beim Versuch 68 schwarze und 32 weisse gezogen werden.

Urne

Farbe	schwarz	weiss	n
Anzahl gezogen	68	32	100
Vertrauensintervall		95%	
u		1.96	
x/n	0.68		
1-x/n	0.32		
d	0.0914		
95%Vertr.intervall		**0.589**	**0.771**

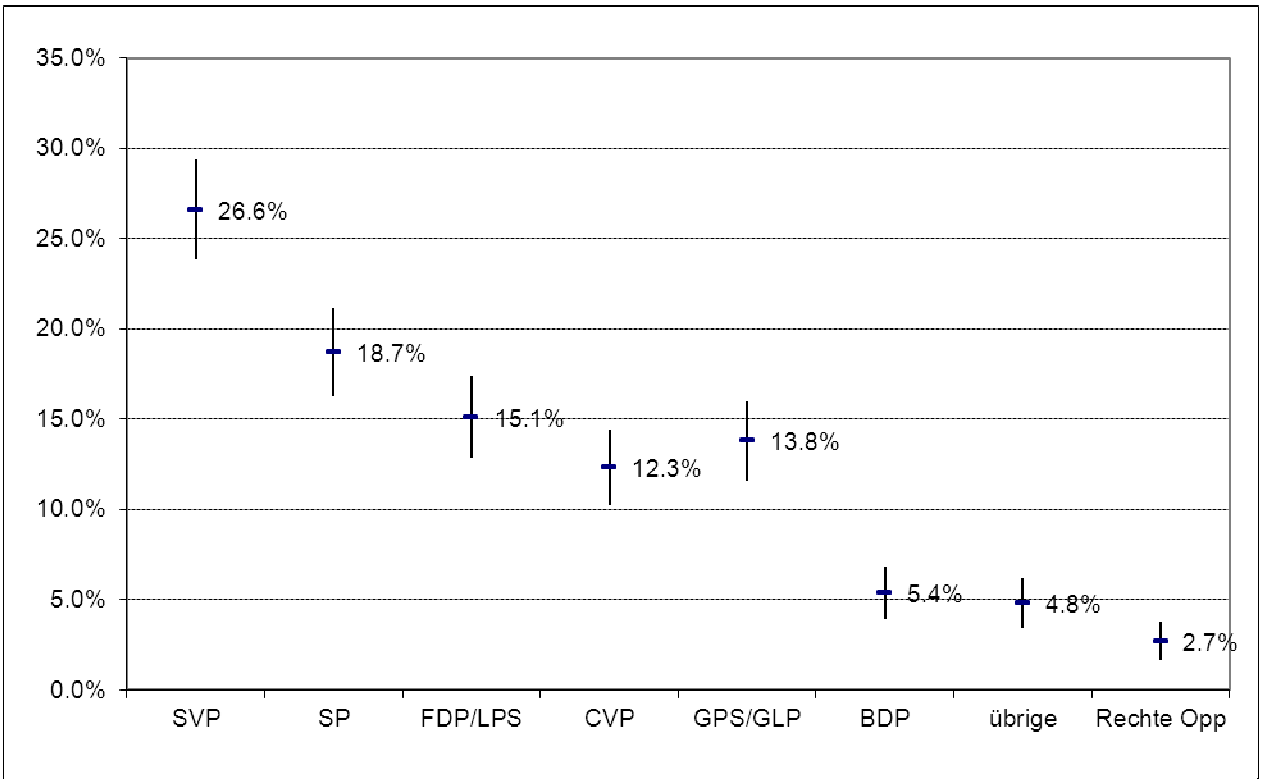
Zur Interpretation des Vertrauensintervalls
Wird das Experiment wiederholt, so würde sich ein anderes Vertrauensintervall ergeben. Das Vertrauensintervall ergibt sich also zufällig. Der unbekannte wahre Wert p liegt also höchstens mit einer Wahrscheinlichkeit von 5% nicht im Vertrauensintervall oder kurz:
Das zufällige Vertrauensintervall enthält die unbekannte Wahrscheinlichkeit p in etwa 19 von 20 Fällen.

Vgl. dazu in der DMK-Formelsammlung die Vertrauensintervalle für die Binomialverteilung sofern $n < 100$.

Wahlprognose

Bei den Nationalratswahlen 2011 haben die grössten Parteien die in der Tabelle angegebenen Wähleranteile erreicht. Wir nehmen an, vor den nächsten Wahlen werden 1000 Personen zufällig ausgewählt und gefragt, welche Partei sie unterstützen. Gesucht sind 95%-Vertrauensintervalle für die Wähleranteile der verschiedenen Parteien.

Vertrauensintervalle		**Wähleranteile**		**Nationalratswahlen 2011**		
n	1000	x/n	1-x/n	d	95%-Vertrauensintervall	
SVP	266	0.266	0.734	0.0274	**23.8%**	**29.4%**
SP	187	0.187	0.813	0.0242	**16.2%**	**21.2%**
FDP/LPS	151	0.151	0.849	0.0222	**12.8%**	**17.4%**
CVP	123	0.123	0.877	0.0204	**10.2%**	**14.4%**
GPS/GLP	138	0.138	0.862	0.0214	**11.6%**	**16.0%**
BDP	54	0.054	0.946	0.0140	**3.9%**	**6.9%**
übrige	48	0.048	0.952	0.0132	**3.4%**	**6.2%**
Rechte Opp	27	0.027	0.973	0.0100	**1.6%**	**3.8%**



Liegen die Ergebnisse der Umfrage ausserhalb der 95%-Intervalle, dann sind zwei Interpretationen möglich:

a) die Umfrage hat bei dieser Partei ein ungewöhnliches (signifikantes) Ergebnis gezeigt

b) der Anteil der Partei hat sich verändert.

Ergänzung:
Um die unbekannte Wahrscheinlichkeit mit einer bestimmten Genauigkeit ε zu schätzen ist ein Stichprobenumfang n nötig, wobei gilt:

$$n \geq \frac{\hat{p} \cdot \hat{q} \cdot u^2_{1-\alpha/2}}{\varepsilon^2}$$

Wegen $\hat{p} \cdot \hat{q} = \hat{p} \cdot (1-\hat{p}) \leq \frac{1}{4}$ gilt die ungenauere Abschätzung $n \geq \frac{u^2_{1-\alpha/2}}{4\varepsilon^2}$

zulässiger Fehler ε	0.05	0.02	0.01
notwendiger Stichprobenumfang	400	2500	10000

Übungsaufgaben:

Die FDP hat bei den NR-Wahlen 2007 17.6 % erreicht. Wie gross ist der Stichprobenumfang zu wählen, damit der Wähleranteil mit 95% Sicherheit auf 1.5% genau geschätzt werden kann?

2. Testen von Hypothesen

Signifikanztest (Normalverteilung)

Einführendes Beispiel:
Einer Kandidatin werden n Fragen mit zwei Auswahlantworten vorgelegt, von denen genau eine richtig ist (einfacher Multiple choice-Test). Es soll geprüft werden, ob die Kandidatin eine Könnerin ist ($p > 0.5$)oder nur rät ($p = 0.5$).

Hypothese H_1:	Die Kandidatin ist eine Könnerin	$p > 0.5$ (Test einseitig)
Nullhypothese H_0:	Die Kandidatin rät nur	$p = 0.5$

Planung eines Experiments:
Es werden $n = 100$ Fragen gestellt
Testgrösse X: Anzahl der richtigen Antworten.
Die Zufallsvariable X ist binomialverteilt mit dem Erwartungswert $\mu = 50$ und der Standardabweichung $\sigma = 5$. Da die Faustregel erfüllt ist, kann die Binomialverteilung durch die Normalverteilung approximiert werden.

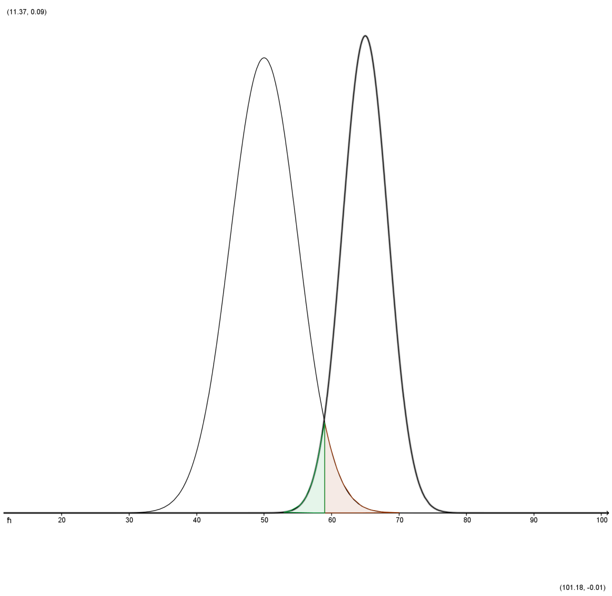
Beim gebräuchlichen Signifikanzniveau $\alpha = 5\%$, ist u so zu wählen, dass $\Phi(u) = 0.95$ gilt. Für die kritische Grenze erhält man damit: $1.645 \cdot \sigma = 1.645\ 5 = 8.225$

Entscheidungsregel (einseitiger Test):
Beantwortet die Kandidatin mindestens $\mu + 1.645 \cdot \sigma \approx 50 + 8.2 \approx 59$ Fragen richtig, so stuft man sie als Könnerin ein.

Fehlermöglichkeiten:
Fehler 1. Art:
Die Nullhypothese wird abgelehnt, obwohl sie richtig ist.
Im Beispiel bedeutet dies: Eine ratende Kandidatin beantwortet, obwohl sie nur rät, mindestens 59 Fragen richtig und wird damit fälschlicherweise als Könnerin eingestuft.
Fehler 2. Art:
Die Nullhypothese ist falsch, wird aber fälschlicherweise beibehalten.
Im Beispiel bedeutet dies: Eine Könnerin (z.B. p = 0.65) hat Pech, beantwortet zu wenig Fragen richtig und wird deshalb als Raterin eingestuft.

In der Abbildung sind die beiden Wahrscheinlichkeiten für den Fehler 1. und 2. Art (unter der Annahme p = 0.65) zu erkennen.

Die Wahrscheinlichkeit, dass eine ratende Kandidatin die Prüfung besteht (d.h. mindestens 59 Fragen richtig beantwortet ist höchstens 5%.
Die Wahrscheinlichkeit, dass eine Kandidatin mit p = 0.65 die Prüfung nicht besteht (d.h. weniger als 59 Fragen richtig beantwortet) ist 8.8%.



Die beiden Fehler in zwei anderen Kontexten:

Fehler erster und zweiter Art

	Schuldig	unschuldig
Schuld-spruch	Richtig	Justizirrtum zu Ungunsten des Angeklagten
Freispruch	Justizirrtum zu Gunsten des Angeklagten	Richtig

(Bohley, 1992, S. 568 zitiert nach Conzelmann, 1999, S. 251)

Fehler erster und zweiter Art

	H0 wahr	H0 falsch
H0 abgelehnt	Fehler 1. Art (α-Fehler, Irrtums-wahrscheinlichkeit) → **Trainingsverfahren überschätzt**	richtige Entscheidung (1 – β, *Power*) → **Beleg für besseres Trainingsverfahren**
H0 beibehalten	richtige Entscheidung 1 - α → **Beleg, dass Trainingsverfahren *nicht* besser**	Fehler 2. Art (β-Fehler) → **Trainingsverfahren unterschätzt**

Je nach den möglichen Auswirkungen des Fehlers wird man versuchen den Fehler 1. oder 2. Art oder beide zu minimieren (Stichwort: Pilzsammeln).

Signifikanztest (Binomialverteilung)

Das folgende Beispiel stammt von Sir Ronald Aylmer Fisher (1890-1962), der als Begründer der modernen mathematisch orientierten Statistik gilt:

Die englische Lady trinkt ihren Tee stets mit einem Zusatz Milch. Eines Tages verblüfft sie ihre Teerunde mit der Behauptung, sie könne allein am Geschmack entscheiden, ob zuerst die Milch oder zuerst der Tee eingegossen worden sei. Sie behauptet, dass sie im Vergleich zum blinden Raten öfter die richtige Eingiess-Reihenfolge treffen würde. Wie könnte die Behauptung der Lady überprüft werden

Planen eines Zufallsexperiments
Die Lady erhält n = 20-mal je 2 Tassen Tee überreicht, jeweils eine vom Typ Milch vor Tee, die andere vom Typ Tee vor Milch. Die Reihenfolge dieser beiden Tassen wird durch den Wurf einer L-Münze festgelegt

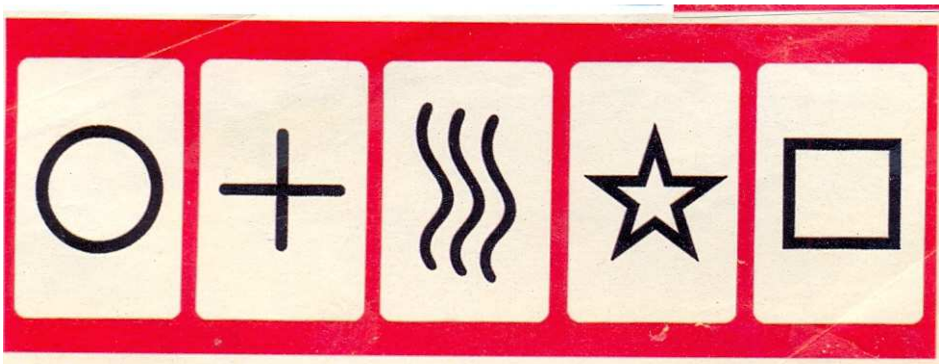
Hypothese H_1: Die Lady hat die behauptete Fähigkeit, d.h. ihre Trefferwahrscheinlichkeit p ist grösser als ½
Nullhypothese H_0.: Die Lady rät nur: $p = ½$
.
Entscheidungsregel: sei X die Anzahl der Treffer
$X \geq 14$ Entscheid für H_1
$X < 14$ Entscheid für H_0

Fehler 1. Art:
H_0 wird verworfen, obwohl sie richtig ist (d.h. man attestiert der Dame eine Fähigkeit, die nicht vorhanden ist (die Dame erzielt also mindestens 14 Treffer durch Raten)

Fehler 2. Art.
H_0 wird beibehalten, obwohl H_1 richtig ist. Die Dame hat tatsächlich die behauptete Begabung, erzielt aber weniger als 14 Treffer.

Die Entscheidungsregel wird oft so gewählt, dass die Wahrscheinlichkeit α des Fehlers erster Art kleiner als 0.05 ist. Ein solcher Test heisst Signifikanztest zum Signifikanzniveau a = 5%. Bei dieser Vorgehensweise wird der Fehler 1. Art im Vergleich zum Fehler 2. Art als schwerwiegender erachtet. Die meisten Tests werden in der Hoffnung auf eine signifikante Ablehnung der Nullhypothese angelegt Die Nullhypothese behauptet die Abwesenheit eines Effekts, den man eigentlich nachweisen möchte!

Ein Beispiel aus der Parapsychologie (ESP: Extra Sensory Perception)
Gibt es Menschen, die über eine Begabung verfügen, die sich naturwissenschaftlich nicht erklären lässt?



ASW-Symbole: Kreis, Kreuz, Welle, Stern, Quadrat

Der amerikanische Parapsychologe B. Rhyne (1895 bis 1980) hat folgendes Experiment vorgeschlagen:
Die Testperson muss versuchen, eine aus 5 verschiedenen Karten ausgewählte Karte zu identifizieren, ohne sie gesehen zu haben.

Hypothese H_1: Die Person verfügt über eine besondere Begabung: $p > 0.2$
Nullhypothese H_0: Die Person hat keine besondere Begabung, rät nur: $p = 0.2$

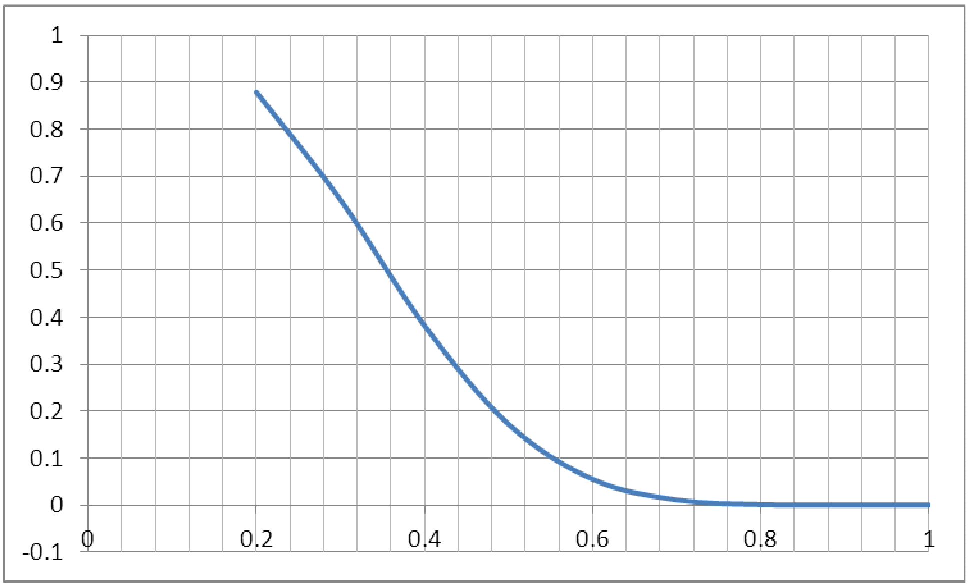
2. Planen eines Experiment:
Es werden 10 Versuche durchgeführt. Die Testgrösse Z: Anzahl der richtig geratenen Karten ist binomialverteilt.

3. Signifikanzniveau:
Eine Person, die nur rät, wird im Mittel bei 10 Versuchen 2 Treffer erzielen. Wir werden einer Person erst dann besondere Begabung zuerkennen, wenn Z deutlich (man sagt auch signifikant) höher liegt.
Entscheidungsregel:
$Z \leq 3 \rightarrow H_0$ beibehalten
$Z > 3 \rightarrow H_0$ ablehnen, Entscheid für H_1

4. Fehler 1. Art:
Wahrscheinlichkeit, dass einer "normalen" Person irrtümlich besondere Begabung zugeschrieben wird ($p = 0.2$, aber $Z > 3$) die zugehörige Wahrscheinlichkeit α ist 3.3%
Fehler 2. Art:
Wahrscheinlichkeit bei einer Person die besondere Begabung nicht anzuerkennen.

Das Risiko 2. Art hängt von Trefferwahrscheinlichkeit der Testperson ab:

Je grösser der Grad aussersinnlicher Begabung ist (das heisst je grösser p ist) umso kleiner ist die Wahrscheinlichkeit, diese besondere Begabung nicht festzustellen.



Verschärft man die Entscheidungsregel ($Z > 5$) so verkleinert sich zwar der Fehler 1. Art, aber der Fehler 2. Art vergrössert sich.
Will man beide Fehler verkleinern, dann bedingt dies eine Erhöhung der Anzahl Versuche.

Zusammenfassung Signifikanztest:

1. Hypothese H_1:
Aufgrund der Daten ergibt sich eine bestimmte Vermutung über den unbekannten Parameter einer Wahrscheinlichkeitsverteilung (z.B. über die Erfolgswahrscheinlichkeit bei einem Bernoulliversuch).
2. Nullhypothese H_0. Gegenhypothese zu H_1:
Die Nullhypothese behauptet Abwesenheit eines Effekts, den man eigentlich nachweisen möchte.
3. Planen eines Zufallsexperiments, Festlegen einer Entscheidungsregel
4. Mögliche Fehlentscheide:
Fehler 1. Art: H_0 wird verworfen, obwohl H_0 richtig ist. Die zugehörige Wahrscheinlichkeit wird mit α bezeichnet, α heisst Signifikanzniveau. Übliches Signifikanzniveau ist $\alpha = 5\%$.
Fehler 2. Art: H_0 wird nicht verworfen, obwohl H_0 falsch ist.

Weitere Beispiele für Signifikanztests:
a)

Ist das neue Medikament besser als das alte?	Nullhypothese: Sie sind beide gleich gut
Ist ein konkreter Würfel gefälscht?	Nullhypothese: Es handelt sich um einen Laplace-Würfel.
Statistische Qualitätskontrolle:	
Der Ausschussanteil ist zu gross?	Die Lieferung ist in Ordnung

3. Beispiel eines Anpassungstests : Chiquadrattest

Die Güte der Übereinstimmung von Daten mit einer bestimmten Wahrscheinlichkeits-verteilung kann mit dem sogenannten Chi-Quadrattest geprüft werden.

Beispiele:
Sind die Augenzahlen eines bestimmten Würfels gleichverteilt?
Ist die Anzahl der Mädchengeburten in Vierkinderfamilien binomialverteilt?
Sind die Widerstände normalverteilt?

Nach Pearson wird folgende Testgrösse verwendet:

$$\chi^2 = \sum_{i=1}^{k} \frac{(b_i - e_i)^2}{e_i}$$

wobei b_i bzw. e_i die beobachteten bzw. erwarteten Häufigkeiten bedeuten. Zu grosse bzw. zu kleine Werte der Testgrösse sind verdächtig vgl. dazu die Tabellen der kritischen Werte in der Formelsammlung.

Beispiel:
Ist ein bestimmter Würfel ein Laplacewürfel?

Der Würfel wird n = 542-mal geworfen:

Augenzahl X	1	2	3	4	5	6
beob. abs. Häufigkeit bi	78	91	86	92	90	105
erwartete Häufigkeit ei	$\frac{542}{6}$	$\frac{542}{6}$	$\frac{542}{6}$	$\frac{542}{6}$	$\frac{542}{6}$	$\frac{542}{6} \approx 90.3$

Der gerechnete Wert von $\chi^2 = 4.3$ ist mit dem kritischen Wert zu vergleichen. Dieser hängt von der Anzahl k der Klassen (k = 6) bzw. von $\nu = k - 1 = 5$ ab, ν heisst Freiheitsgrad. In der Tabelle liest man als kritischen Wert 11.1 bei einem Signifikanzniveau von 5% ab. Da der gerechnete Wert $\chi^2 = 4.3 < 11.1$ ist, kann der Laplacewürfel als unverdächtig eingestuft werden.

Zur Beurteilung können auch Erwartungswert und Standardabweichung der Chiquadrat-verteilung zugezogen werden. Es gilt nämlich
für den Erwartungswert: $\mu = k - 1 = \nu$ Freiheitsgrad
für die Varianz $\sigma \approx 2 \cdot (k-1) = 2\nu$
im Beispiel also $\mu = 5$ bzw. $\sigma = 10$

Entscheidungsregel:
Weicht der Wert von Chiquadrat um mehr als 2 Standardabweichungen vom Erwartungswert ab, dann gilt die Abweichung als auffällig und man würde in diesem Fall annehmen, dass der Würfel kein Laplacewürfel ist.
Die kritischen Werte findet man in der Formelsammlung tabelliert.

Wie schon erwähnt sind auch zu kleine Werte von Chiquadrat auffällig

Beispiel:
Hat Mendel Daten manipuliert?
Mendels Theorie sagt aus: Beim Kreuzen von gelben runden mit grünen eckigen Erbsen gibt es vier Nachkommen im Verhältnis

	9 :	3 :	3 :	1
	gelb/rund	gelb/eckig	grün/rund	grün eckig
beobachtete Häufigkeit b_i:	315	101	108	32
erwartete Häufigkeit e_i:	$\frac{9}{16}\cdot 556$	$\frac{3}{16}\cdot 556$	$\frac{3}{16}\cdot 556$	$\frac{1}{16}\cdot 556$
	312.75	104.25	104.25	34.75

$$\chi^2 = \frac{(315-312.75)^2}{312.75} = \frac{(108-104.25)^2}{104.25} = \frac{(101-104.25)^2}{104.25} = \frac{(32-34.75)^2}{34.75} = 0.470$$

Freiheitsgrad $\nu = k - 1 = 3$
Der kritische Wert 7.81 bei einem Signifikanzniveau von 5% ist weit unterschritten. Es handelt sich um einen verdächtig kleinen Wert, denn aus der Tabelle schätzt man ab, dass in etwa 93% aller Fälle grössere Werte von Chiquadrat auftreten.
Die neueste Deutung der zu guten Werte: Mendel hat jedes Mal so lange Material gesammelt, bis seine Werte nahe bei den Erwartungswerten lagen!

Mit dem Chiquadrattest können die Daten bei der Untersuchung der Widerstände im ersten Abschnitt überprüft werden:

Klasse	Messdaten Häufigkeit	Werte der NV mit Erw.wert 15.11 Standabw. 0.052	Chiquadrat
15	**3**	**2**	0.95
15.025	**1**	**3**	1.68
15.05	**6**	**7**	0.24
15.075	**16**	**13**	0.91
15.1	**20**	**17**	0.41
15.125	**17**	**19**	0.21
15.15	**18**	**17**	0.12
15.175	**11**	**12**	0.02
15.2	**5**	**6**	0.30
15.225	**1**	**3**	1.18
und größer	**2**	**1**	0.31
Summe	100	100	
			6.34

15.11 Erwartungswert
(Schätzung durch emp.Mittelwert der Daten)
0.052 Standardabweichung
(Schätzung durch emp. Standardabweichung)

Anpassungstest Chiquadrat

Chiquadrat aus den Daten		6.34	
kritisches Chiquadrat		Signifikanzniveau	5%
Freiheitsgrad	10	(11 Klassen)	
aus Tabelle	18.3		

Da der Chiquadratwert 6.34 kleiner als der kritische Wert 18.3 ist, kann die Hypothese, es liege eine Normalverteilung vor beibehalten werden.
Die Übereinstimmung einer Verteilung mit der Normalverteilung kann auch graphisch überprüft werden.