Mathematik für Biologen

Prof. Dr. Rüdiger W. Braun

Heinrich-Heine-Universität Düsseldorf

15. Dezember 2010

Versuchsplanung

- Die Länge des Konfidenzintervalls wird vorgeschrieben. Wie viele Versuche sind nötig?
- Dazu muss σ bekannt sein
- Das Konfidenzintervall

$$\left[\overline{X} - \frac{\sigma \cdot q_{1-\alpha/2}}{\sqrt{n}}, \overline{X} + \frac{\sigma \cdot q_{1-\alpha/2}}{\sqrt{n}}\right]$$

- hat die Länge

$$\frac{2\sigma \cdot q_{1-\alpha/2}}{\sqrt{n}}$$

- Dieser Wert soll gleich d sein. Das bedeutet

$$n = \left(\frac{2\sigma \cdot q_{1-\alpha/2}}{d}\right)^2$$

Versuchsplanung, Zahlenbeispiel

- Bei unseren Roggenpflanzen streut die Halmlänge um $7cm$
- Wie viele Pflanzen müssen gemessen werden, damit das Konfidenzintervall zum Konfidenzniveau 95% die Länge $5cm$ hat?
- $q_{0.975} = 1.96$, also $n = \left(\frac{2 \cdot 7 \cdot 1.96}{5}\right)^2 = 30.11$
- 30 Halme genügen also
- Wenn das Konfidenzintervall zum Konfidenzniveau 95% nur $1cm$ lang sein darf, dann benötigt man 25-mal so viele Halme, also 753
- Wenn das Konfidenzintervall $5cm$ lang sein soll, aber das Konfidenzniveau 99%, dann $n = \left(\frac{2 \cdot 7 \cdot 2.56}{5}\right)^2 = 51.38$

Schätzung der Varianz

- Bei den bisherigen Schätzern für Konfidenzintervalle musste die Varianz von vornerein bekannt sein. Das ist unrealistisch
- Ausweg: Wir schätzen die Varianz durch die empirische Varianz

$$S^2 = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X})^2$$

  Hierbei ist $\overline{X}$ das arithmetische Mittel
- Dadurch, dass zweimal geschätzt wird (Erwartungswert und Varianz), sinkt die Genauigkeit
- Die Konfidenzintervalle werden also größer
- Statt der Quantile von Φ muss man die Quantile einer anderen Verteilung verwenden. Diese Verteilung hängt von der Anzahl n der Versuche ab

Die t-Verteilung

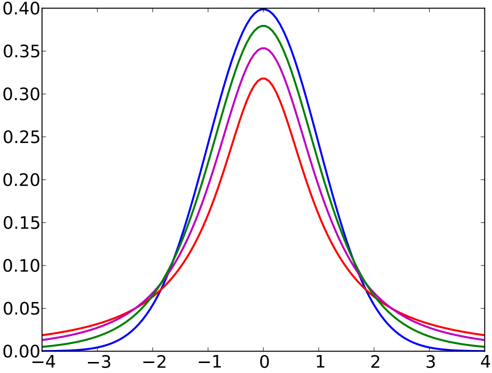
Für $n = 2, 3, 4, \ldots$ ist die t-Verteilung (oder Studentsche Verteilung) mit $n-1$ Freiheitsgraden gegeben durch die Dichte

$$f_{n-1} = c_{n-1} \left(1 + \frac{x^2}{n-1}\right)^{-n/2}$$

Dabei wird c_{n-1} bestimmt durch das Erfordernis, dass $\int_{-\infty}^{\infty} f_{n-1}(x)\, dx = 1$

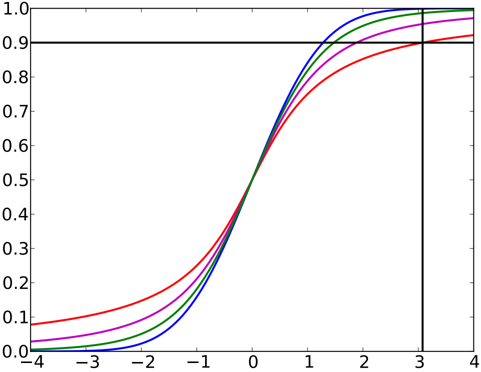
Die t-Verteilungen sind tabelliert

Graphen von Dichten von t-Verteilungen



Dichten der Standard-Normalverteilung (blau) und der t-Verteilungen mit 1 (rot), 2 (violett) bzw. 5 (grün) Freiheitsgraden

Graphen von Verteilungsfunktionen von t-Verteilungen



Verteilungsfunktionen der Standard-Normalverteilung (blau) und der t-Verteilungen mit 1 (rot), 2 (violett) bzw. 5 (grün) Freiheitsgraden

Quantile der t-Verteilung

f	90%	95%	97.5%	99%	99.5%	99.9%
1	3.078	6.314	12.706	31.821	63.657	318.309
2	1.886	2.920	4.303	6.965	9.925	22.327
3	1.638	2.353	3.182	4.541	5.841	10.215
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.893
6	1.440	1.943	2.447	3.143	3.707	5.208
7	1.415	1.895	2.365	2.998	3.499	4.785
8	1.397	1.860	2.306	2.896	3.355	4.501
9	1.383	1.833	2.262	2.821	3.250	4.297
10	1.372	1.812	2.228	2.764	3.169	4.144
11	1.363	1.796	2.201	2.718	3.106	4.025
12	1.356	1.782	2.179	2.681	3.055	3.930
13	1.350	1.771	2.160	2.650	3.012	3.852
14	1.345	1.761	2.145	2.624	2.977	3.787
15	1.341	1.753	2.131	2.602	2.947	3.733
16	1.337	1.746	2.120	2.583	2.921	3.686
17	1.333	1.740	2.110	2.567	2.898	3.646
18	1.330	1.734	2.101	2.552	2.878	3.610
19	1.328	1.729	2.093	2.539	2.861	3.579
20	1.325	1.725	2.086	2.528	2.845	3.552
∞	1.282	1.645	1.960	2.326	2.576	3.090

Freiheitsgrade

Heuristisch:

- n Versuche, um einen Parameter θ zu schätzen
- Jeder andere Parameter, der hilfsweise geschätzt werden muss, verringert die Zahl der Freiheitsgrade um 1
- Bei der Schätzung eines Konfidenzintervalls bei unbekannter Varianz verliert man einen Freiheitsgrad durch die Schätzung der Varianz

Schätzung eines Konfidenzintervalls für den Erwartungswert bei unbekannter Varianz

- $X_1, \ldots, X_n$ seien normalverteilt gemäß $N(\mu, \sigma^2)$ für **un**bekanntes σ und unbekanntes μ
- Für μ soll ein zweiseitiges Konfidenzintervall zum Konfidenzniveau $1 - \alpha$ geschätzt werden
- Verwende dazu das arithmetische Mittel

$$\overline{X} = \frac{1}{n}(X_1 + \cdots + X_n)$$

und die Wurzel aus der empirischen Varianz

$$S = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X})^2}$$

Schätzung eines Konfidenzintervalls, Fortsetzung

- Benötigt wird das $(1-\frac{\alpha}{2})$-Quantil der t-Verteilung mit $(n-1)$ Freiheitsgraden; wir bezeichnen es mit $t_{n-1,\,1-\alpha/2}$
- Dann ist das Konfidenzintervall zum Konfidenzniveau $1-\alpha$

$$\left[\overline{X} - \frac{S}{\sqrt{n}} t_{n-1,\,1-\alpha/2},\ \overline{X} + \frac{S}{\sqrt{n}} t_{n-1,\,1-\alpha/2}\right]$$

Beispiel: Reaktionszeit

- In einem Versuch wird die Reaktionszeit eines Patienten auf einen Luftstoß am Auge gemessen. Der Versuch wird 20-mal wiederholt
- Ein Konfidenzintervall zum Konfidenzniveau 95% ist zu schätzen
- Beispielhafter Datensatz (Zeiten im *ms*)

43.14	25.60	35.14	65.50	38.29
23.37	11.00	72.65	47.69	33.12
42.18	49.18	27.35	46.40	32.69
61.52	52.54	68.52	58.92	75.74

- Arithmetisches Mittel und Stichprobenstreuung dieses Datensatzes sind

$$\overline{x} = 45.53 \qquad s = 17.75$$

Reaktionszeit, Fortsetzung

- 20 Versuche bedeuten 19 Freiheitsgrade
- $1-\alpha = 0.95$, also $1-\frac{\alpha}{2} = 0.975$. Wir benötigen das 0.975-Quantil der t-Verteilung mit 19 Freiheitsgraden.

Quantile der t-Verteilung

f	90%	95%	97.5%	99%	99.5%	99.9%
1	3.078	6.314	12.706	31.821	63.657	318.309
2	1.886	2.920	4.303	6.965	9.925	22.327
3	1.638	2.353	3.182	4.541	5.841	10.215
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.893
6	1.440	1.943	2.447	3.143	3.707	5.208
7	1.415	1.895	2.365	2.998	3.499	4.785
8	1.397	1.860	2.306	2.896	3.355	4.501
9	1.383	1.833	2.262	2.821	3.250	4.297
10	1.372	1.812	2.228	2.764	3.169	4.144
11	1.363	1.796	2.201	2.718	3.106	4.025
12	1.356	1.782	2.179	2.681	3.055	3.930
13	1.350	1.771	2.160	2.650	3.012	3.852
14	1.345	1.761	2.145	2.624	2.977	3.787
15	1.341	1.753	2.131	2.602	2.947	3.733
16	1.337	1.746	2.120	2.583	2.921	3.686
17	1.333	1.740	2.110	2.567	2.898	3.646
18	1.330	1.734	2.101	2.552	2.878	3.610
19	1.328	1.729	**2.093**	2.539	2.861	3.579
20	1.325	1.725	2.086	2.528	2.845	3.552
∞	1.282	1.645	1.960	2.326	2.576	3.090

Reaktionszeit, Fortsetzung

- Erinnerung

$$\overline{x} = 45.53 \qquad s = 17.75 \qquad t_{19,\,0.975} = 2.093$$

- Die Realisierung der oberen Vertrauensgrenze ist

$$\begin{aligned} G_o &= \overline{x} + \frac{s}{\sqrt{n}} \cdot t_{n-1,\,1-\alpha/2} \\ &= 45.53 + \frac{17.75}{\sqrt{20}} \cdot 2.093 \\ &= 45.53 + 8.30 \\ &= 53.83 \end{aligned}$$

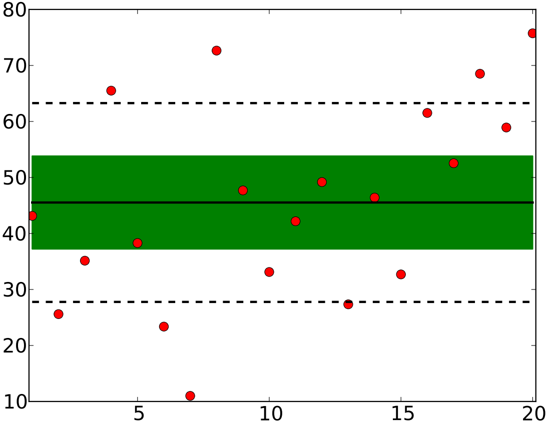
- Entsprechend ist $G_u = 45.53 - 8.30 = 37.23$ die Realisierung der unteren Vertrauensgrenze.

Reaktionszeit, Zusammenfassung

Zum Konfidenzniveau 95% kann gesagt werden:

> *Die Reaktionszeit des Probanden auf Luftstöße am Auge liegt zwischen* $37.23\,ms$ *und* $53.83\,ms$

Reaktionszeit, Skizze



Der grüne Bereich ist das Konfidenzintervall zum Konfidenzniveau 95%

Tabelle: Konfidenzintervall für den Erwartungswert bei unbekannter Varianz

- Das Konfidenzniveau sei $1 - \alpha$
- Bilde arithmetisches Mittel

$$\overline{X} = \frac{1}{n} \sum_{k=1}^{n} X_k$$

und Stichprobenstreuung

$$S = \sqrt{\frac{1}{n-1} \sum_{k=1}^{n} (x - \overline{x})^2}$$

- Bei einseitigem Konfidenzintervall: Bestimme das Quantil $t_{n-1,\,1-\alpha}$
- Bei zweiseitigem Konfidenzintervall: Bestimme das Quantil $t_{n-1,\,1-\alpha/2}$

Tabelle, Fortsetzung

- zweiseitig zum Konfidenzniveau $1-\alpha$

$$\overline{X} - \frac{S \cdot t_{n-1,\,1-\alpha/2}}{\sqrt{n}} \leq \mu \leq \overline{X} + \frac{S \cdot t_{n-1,\,1-\alpha/2}}{\sqrt{n}}$$

- einseitig, obere Vertrauensgrenze zum Konfidenzniveau $1-\alpha$

$$\mu \leq \overline{X} + \frac{S \cdot t_{n-1,\,1-\alpha}}{\sqrt{n}}$$

- einseitig untere Vertrauensgrenze zum Konfidenzniveau $1-\alpha$

$$\overline{X} - \frac{S \cdot t_{n-1,\,1-\alpha}}{\sqrt{n}} \leq \mu$$

Schätzung eines Konfidenzintervalls für die Erfolgswahrscheinlichkeit

- Ein ja/nein-Experiment wird n-mal wiederholt
- Es soll ein Konfidenzintervall für die Erfolgswahrscheinlichkeit im Einzelfall geschätzt werden.
- Es seien k Erfolge beobachtet worden
- Wir wissen bereits, dass

  $$\hat{p} = \frac{k}{n}$$

  der ML-Schätzer für die Erfolgswahrscheinlichkeit ist
- Wenn

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

  kann die Normalapproximation angewandt werden

Schätzung eines Konfidenzintervall für die Erfolgswahrscheinlichkeit, Fortsetzung

$$Y = \sqrt{n}\frac{\hat{p} - p}{\sqrt{p \cdot (1-p)}}$$

ist näherungsweise standardnormalverteilt

- $q_{1-\alpha/2}$ sei das Quantil der Standardnormalverteilung. Wir kürzen es mit q ab

$$P(-q \leq Y \leq q) \cong 1 - \alpha$$

- Rückeinsetzen von Y gibt

$$-q \leq \sqrt{n}\frac{\hat{p} - p}{\sqrt{p \cdot (1-p)}} \leq q$$

- Das muss nach p aufgelöst werden. Das größere p ist G_o, das kleinere G_u

Schätzung eines Konfidenzintervall für die Erfolgswahrscheinlichkeit, Formeln

$$G_o = \frac{1}{1 + q^2/n} \left(\hat{p} + \frac{q^2}{2n} + \frac{q}{\sqrt{n}} \cdot \sqrt{\hat{p} \cdot (1 - \hat{p}) + \frac{q^2}{4n}} \right)$$

$$G_u = \frac{1}{1 + q^2/n} \left(\hat{p} + \frac{q^2}{2n} - \frac{q}{\sqrt{n}} \cdot \sqrt{\hat{p} \cdot (1 - \hat{p}) + \frac{q^2}{4n}} \right)$$

Konfidenzintervall für das Rückfangexperiment

- Erinnerung: N Mücken in der Halle, davon 160 markierte.
- 100 Mücken gefangen, darunter waren 19 markierte.
- Realisierung des Punktschätzers für die Zahl der Mücken:

$$\widehat{N} = 842$$

- Frage: Wie genau ist das?

Rückfangexperiment, Fortsetzung

- Die Zahl der markierten unter den Mücken des Rückfangs ist verteilt gemäß $B_{100,\,p}$ mit unbekanntem p
- Wir bestimmen ein Konfidenzintervall zum Konfidenzniveau 95% für dieses p
- Der Zusammenhang zum Schätzer für die Anzahl aller Mücken ist
$$\hat{N} = \frac{160}{\hat{p}}$$
- Wenn G_u eine untere Vertrauensgrenze für p ist, dann ist
$$\frac{160}{G_u}$$
eine obere Vertrauensgrenze für N

Rückfangexperiment, Fortsetzung

- $\hat{p} = 0.19$ und $1 - \frac{\alpha}{2} = 0.975$
- $q_{0.975} = 1.960$ aus der Tabelle
- $G_u = 0.1251$
- $\dfrac{160}{G_u} = 1278$
- genauso $G_o = 0.2778$ und $\dfrac{160}{G_o} = 576$
- Zum Konfidenzniveau 95% liegt die Anzahl der Mücken in der Halle zwischen 576 und 1278

Hypothesentests

- Medikament 1 wirkt gegen eine bestimmte Krankheit mit einer Heilungswahrscheinlichkeit $p = 0.2$
- Das neue Medikament 2 wirkt gegen dieselbe Krankheit, und zwar nach Angaben des Herstellers mit einer größeren Heilungswahrscheinlichkeit als 0.2
- Ein Test an 24 Patienten soll klären, ob das Medikament 2 dem Medikament 1 vorzuziehen ist
- Es muss eine Entscheidung zwischen zwei Optionen gefällt werden. Diese bezeichnet man als Nullhypothese und Alternativhypothese

Nullhypothese und Alternativhypothese

- **Nullhypothese H_0:** Das ist diejenige Hypothese, deren fälschliche Ablehnung man nach Möglichkeit vermeiden will

  Im Fall des Medikamentenvergleichs ist die Nullhypothese die Annahme, dass das alte Medikament mindestens so gut ist wie das neue

- **Alternativhypothese H_1:** Das ist die Alternative zur Nullhypothese

  Im Falle des Medikamentenvergleichs also die Annahme, dass das neue Medikament besser ist als das alte

Nullhypothese und Alternativhypothese, Fortsetzung

- Wissenschaft ist konservativ. Wer mit einer neuen Idee kommt, muss zeigen, dass sie besser ist als die alte
- Daher ist die Nullhypothese in der Regel die Annahme, dass die bestehende Theorie bzw. das vorhandene Medikament mindestens so gut ist wie die Neuerungen

Fehler erster und zweiter Art

- Die Zufallsvariable X gibt die Anzahl der Heilerfolge des zweiten Medikaments an.
- Wir suchen eine Zahl c derart, dass die Nullhypothese für $X \leq c$ angenommen und für $X > c$ abgelehnt wird.

	$X \leq c$	$X > c$
$p \leq 0.2$	richtige Entscheidung	Fehler 1. Art
$p > 0.2$	Fehler 2. Art	richtige Entscheidung

- Der Fehler 1. Art ist die fälschliche Ablehnung der Nullhypothese. Schlimm!
- Der Fehler 2. Art ist die fälschliche Annahme der Nulhypothese

Medikamentenvergleich, Fehler 1. Art

- Je größer c gewählt wird, umso unwahrscheinlicher ist der Fehler 1. Art. Dafür wird der Fehler zweiter Art immer wahrscheinlicher
- Im Medikamentenvergleich hatten wir $n = 24$ und

$$\begin{aligned} H_0 &: \quad p \leq 0.2 \\ H_1 &: \quad p > 0.2 \end{aligned}$$

- Naiv würde man setzen $c = 0.2 \cdot 24 = 4.8$, also gerundet $c = 5$

Medikamentenvergleich, Fehler 1. Art

- Berechne den Fehler 1. Art für $c = 5$ und $p = 0.2$
- Die Anzahl der geheilten Patienten ist $B_{n,p}$-verteilt mit $n = 24$ und $p = 0.2$, also

$$
\begin{aligned}
P(X > 5) &= 1 - P(X \leq 5) \\
&= 1 - \sum_{k=0}^{5} B_{24,\,0.2}(k) = 1 - 0.65589 = 0.34411
\end{aligned}
$$

- Wahrscheinlichkeit von 34% für den Fehler erster Art ist nicht akzeptabel

Tabelle von $\sum_{k=0}^{r} B_{24,\,p}(k)$

r	p	0.20	0.21	0.22	0.23	0.24
0	0.	00472	00349	00257	00189	00138
1		03306	02577	01998	01541	01183
2		11452	09387	07645	06188	04978
3		26386	22662	19326	16367	13767
4		45988	41188	36622	32330	28338
5		65589	60887	56136	51401	46743
6		81107	77469	73565	69441	65149
7		91083	88804	86206	83297	80094
8		96383	95206	93782	92092	90124
9		98738	98232	97581	96763	95754
10		99621	99438	99188	98855	98421
11		99902	99846	99765	99651	99493
12		99978	99964	99942	99908	99860
13		99996	99993	99988	99979	99967
14		99999	99999	99998	99996	99993
15					99999	99999

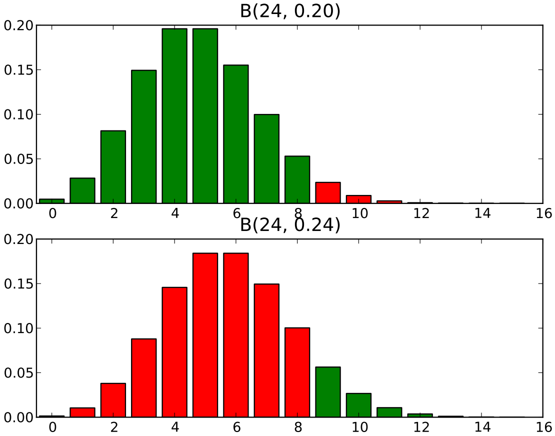
Medikamentenvergleich, Fortsetzung

- Die Tabelle zeigt, dass $c = 8$ gewählt werden muss, wenn die Wahrscheinlichkeit eines Fehlers 1. Art kleiner als 5% sein soll.
- Bestimme nun die Wahrscheinlichkeit eines Fehlers 2. Art falls $p = 0.24$
- Für $p = 0.24$ kommt es bei $k = 0, 1, 2, \ldots, 8$ Erfolgen zum Fehler 2. Art.
- Die Wahrscheinlichkeit hierfür ist

$$\sum_{k=0}^{8} B_{24,\,0.24}(k) = 0.90124$$

- Bei $p = 0.24$ beträgt die Wahrscheinlichkeit eines Fehlers 2. Art 90%

Fehler 1. und 2. Art im Medikamentenvergleich



Testverfahren

- Es sei Θ eine Menge von Parametern. Zu jedem $\theta \in \Theta$ gebe es eine Verteilung P_θ
- Die Zufallsvariablen $X_1, \ldots, X_n$ seien unabhängig und alle nach demselben P_θ verteilt. Dieses θ sei unbekannt
- Die Menge M aller möglichen Werte $(x_1, \ldots, x_n)$ von $(X_1, \ldots, X_n)$ heißt *Stichprobenraum*
- Der Parameterraum sei in zwei Mengen H_0 und H_1 zerlegt. Dabei ist H_0 die Nullhypothese und H_1 die Alternative

Testverfahren, Fortsetzung

- Ein *Test* besteht aus der Zerlegung des Stichprobenraums M in zwei Teilmengen K_0 und K_1 derart, dass die Nullhypothese akzeptiert wird, wenn $(x_1, \ldots, x_n) \in K_0$, und abgelehnt wird, wenn $(x_1, \ldots, x_n) \in K_1$
- Die Menge K_0 heißt *Annahmebereich*, die Menge K_1 heißt *kritischer Bereich*

	$x \in K_0$	$x \in K_1$
$\theta \in H_0$	richtige Entscheidung	Fehler 1. Art
$\theta \in H_1$	Fehler 2. Art	richtige Entscheidung

Interpretation im Beispiel

- Im Beispiel bezeichnen die Zufallsvariablen $X_1, \ldots, X_{24}$ die Heilerfolge der einzelnen Patienten
- Der Parameterraum ist $\Theta = [0, 1]$
- Alle X_j sind $B_{1,p}$ verteilt mit unbekanntem $p \in \Theta$
- Der Stichprobenraum ist $M = \{0, 1\}^n$
- Die Nullhypothese ist $H_0 = \{p \leq 0.2\}$, die Alternative $H_1 = \{p > 0.2\}$
- Ich habe den folgenden Test vorgestellt:

$$K_0 = \left\{ (x_1, \ldots, x_n) \in M \;\middle|\; \sum_{j=1}^{24} x_j \leq 8 \right\}$$

$$K_1 = \left\{ (x_1, \ldots, x_n) \in M \;\middle|\; \sum_{j=1}^{24} x_j > 8 \right\}$$

Signifikanztests

- Für jedes $\theta \in H_0$ bezeichnet man die Wahrscheinlichkeit $P_\theta(K_1)$ als eine *Fehlerwahrscheinlichkeit erster Art*
- Ein Test heißt *Signifikanztest* zum Niveau α, wenn alle Fehlerwahrscheinlichkeiten erster Art $\leq \alpha$ sind
- Im Beispiel hatte ich einen Signifikanztest zum Niveau $\alpha = 0.05$ angegeben. Übliche Niveaus sind 0.1, 0.05 und 0.01