Hypothesen

<u>Ausgangssituation und Problemstellung</u>

* Population, μ, σ
* Langjähriger Durchschnitt (Ernte)
* Wahrscheinlichkeit für Ereignisse (Münzen, Roulette)
* Radioaktivität (Hintergrundstrahlung)

<u>Stichprobe, Test oder eine Messung ergibt:</u>

$\Rightarrow$ $\bar{x}$ weicht von μ ab
$\Rightarrow$ diesjährige Ernte über den Durchschnitt
$\Rightarrow$ geänderte (manipulierte?) Wahrscheinlichkeit, (z.B. beim Roulette kommt 10-mal hintereinander eine gerade Zahl)
$\Rightarrow$ Hintergrundstrahlung erhöht oder Kontamination der Luft

Beispiel: Arzt testet ein neues Medikament

Bisher: $p = 0{,}4$; (Wirksamkeitswahrscheinlichkeit, in 40 % der Fälle wirkt das bisher verwendete Mittel)

Wir sagen: $\Rightarrow H_o : p = 0{,}4$ (postuliert den bisherigen Zustand)
$\Rightarrow H_1$ oder H_A: $p > 0{,}4$ (Mittel wirkt besser)

Bezeichnungsweise:

H_o: Nullhypothese	ist meist eine einfache Hypothese ($\mu = \mu_o$) * postuliert den bisherigen Zustand (Ausgangszustand) z.B. gegebene durchschnittliche Größe $p = ½$ für Münzwurf $p = p_o$ für Heilwirkung eines Medikamentes
H_A: Alternativhypothese	meist zusammengesetzt, umfaßt größere Bereiche: $\mu \neq \mu_o$ * postuliert Veränderung gegenüber Status quo

Beispiele: H_o: $\mu = \mu_o$; μ_o =178 cm (eigentlich $\mu \leq \mu_o$, einseitige Hypothese !)
H_1: $\mu > 178$

H_o: $p = ½$
H_1: $p \neq ½$

H_o: $p \geq p_o$, z.B. die durchschnittliche Unfallhäufigkeit
H_1: $p < p_o$, Senkung der durchschn. Unfallhäufigkeit

Weiter zum Beispiel „Medikamententest“

H_o: $p = 0{,}4$ (genau genommen $p \leq 0{,}4$, für $p = 0{,}4$ wird α ein Maximum !)
H_1: $p > 0{,}4$

? welche Kriterien zur Entscheidungsfindung können genannt werden ?

⇨ z.B. experimentelle Verifikation mittels eines Versuches mit n Patienten (= Binomialexperiment !)

Für n = 12 Patienten

X.........Zufallsvariable, bezeichnet die Anzahl der erfolgreich behandelten Patienten
X kann bei 12 Patienten alle Werte aus dem Ereignisraum
$\Omega = \{0,1,.....12\}$ annehmen

Wie ist nun X verteilt bei p = 0.4 ?

Was ist die Bedeutung von p = 0,4? Mit p wird die Wirksamkeitswahrscheinlichkeit des Medikamentes oder einer Behandlung angegeben. Die Wahrscheinlichkeit für Erfolg ergibt sich aus der Binomialverteilung.

Z.B. P(X=12), Wahrscheinlichkeit, daß alle 12 Patienten gesundet sind.

$P(X = 12) = 0.4^{12} = 0.000017$

„Falls die Wirksamkeitswahrscheinlichkeit des Medikamentes p = 0,4 beträgt, werden mit einer Wahrscheinlichkeit von 0,0017% alle 12 Patienten geheilt. Wird bei 12 beobachteten gesunden Patienten die Hypothese H_o: p=0,4 verworfen, dann ist damit eine Irrtumswahrscheinlichkeit von 0,0017% für den Fehler 1. Art verbunden. Die Hypothese wird mit dieser Irrtumswahrscheinlichkeit verworfen, obwohl sie richtig ist! Andererseits wäre bei X = 5 kein Grund vorhanden, die Hypothese H_o zu verwerfen!

Daraus ergeben sich:

=> **Kritischer Bereich** – Bereich in dem H_o verworfen wird
=> **Akzeptanz Bereich** – Bereich in dem H_o akzeptiert wird

k ist der kritische Wert, der sich aus einer festgelegten Irrtumswahrscheinlichkeit von α ergibt !!!

Falls: $k \leq X \leq 12$ => verwerfe H_0
$0 \leq X \leq k - 1$ => akzeptiere H_0

Das Kriterium für die Festlegung des kritischen Wertes k ist (vorerst) die Irrtumswahrscheinlichkeit α, die zum Tragen kommt, wenn H_0 verworfen wird, obwohl diese richtig ist. Dabei können folgende Fehler auftreten:

	H_0 wahr (H_1 falsch)	H_0 falsch (H_1 wahr)
Akzeptiere H_0 (verwerfe H_1)	Richtige Entscheidung	Fehler 2. Art (Falscher Nichtnachweis) Ignoranz
Verwerfe H_0 (Akzeptiere H_1)	Fehler 1. Art (falscher Nachweis) („Aberglaube“)	Richtige Entscheidung

k_w	$P(X = k_w)$	Größe des kritischen Bereiches $\alpha = P(k_w \leq X \leq 12)$
4	0,21284	0,7747
5	0,22703	0,5618
6	0,17658	0,3348
7	0,10090	0,1582
8	0,04204	0,0573
9	0,01246	0,0153
10	0,00249	0,0028
11	0,00030	0,0003
12	0,00002	0,00002

Im Beispiel der Untersuchung der Wirksamkeit eines neuen Medikamentes ergibt sich die Größe des kritischen Bereiches aus der Binomialverteilung

α = Fehler 1. Art. Dieser wird ein Maximum, wenn angenommen wird $p = 0{,}4$

Überlegen Sie, wie α sich ändert, wenn $p = 0{,}3$ oder $p = 0{,}2$?

Wird H_o bei $X = 8$ abgelehnt, dann ist $\alpha = 5{,}73\%$; bei $X = 9$ ist $\alpha = 1{,}53\%$. In den Naturwissenschaften hat sich eingebürgert für $\alpha = 5\%$ festzulegen. Damit wären also mindestens 9 geheilte Patienten notwendig, um H_o zu verwerfen

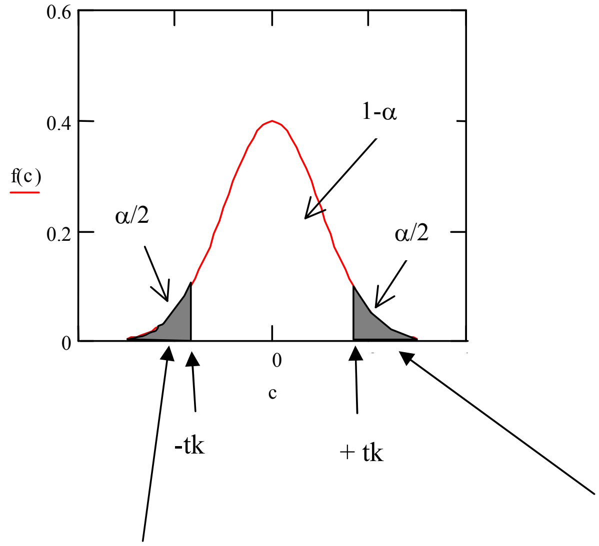
Welche Information gibt es bezüglich des Fehlers 2. Art?

- ⇨ noch keine, weil keine Angaben darüber, in welchem Bereich p^* (H_A: $p = p^*$; $p^* > p$) liegen soll, sondern die Hypothese H_A nur behauptet, daß $p^* > 0{,}4$ ist, aber keinen definierten Wart dafür angibt!
- ⇨ Allgemein sind kleine Unterschiede nur schwer festzustellen, bzw. es können diese nur mit großen Probemengen ermittelt werden

Zur t-Verteilung und Hypothesentests

$T = \dfrac{\overline{X} - \mu}{S/\sqrt{n}}$ ist t-verteilt mit FG ν: $\nu = n - 1$

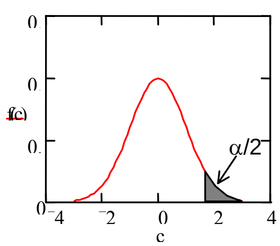
Der Erwartungswert $E(T) = 0$
Die Varianz $V(T) = f(n)$, sie liegt nahe bei 1



kritische Regionen links und rechts mit den kritischen Werten $-t_k$ und $+t_k$

Es gibt nun 3 Testmöglichkeiten :

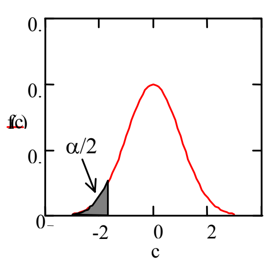
a) Einseitiger Test, rechts-seitig



H_o: $\mu \leq \mu_o$
H_1: $\mu > \mu_o$

$$\alpha = P\left(\frac{(\overline{X} - \mu)}{S/\sqrt{n}} \geq c \Big|_{\mu \leq \mu_o} \right)$$

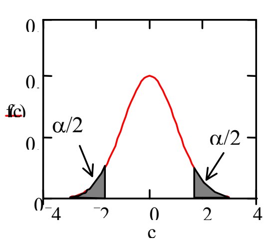
b) Einseitiger Test, links-seitig



H_o : $\mu \geq \mu_o$
H_1 : $\mu < \mu_o$

$$\alpha = P\left(\frac{(\overline{X} - \mu)}{S/\sqrt{n}} \leq c \Big|_{\mu \geq \mu_o} \right)$$

c) beidseitig



H_o : $\mu = \mu_o$
H_1 $\neq \mu_o$

$$\alpha = P\left(\frac{\left|\overline{X} - \mu\right|}{S/\sqrt{n}} \geq c \Big|_{\mu \neq \mu_o} \right)$$

Beispiele

zu a) Versuchsreihe mit $n = 20$
$S^2 = 32$
$\overline{X} = 95$
? Signifikant größer als 92 auf α_{eins} ($\alpha/2$) = 5; 2,5; 1% ?

Berechne die Prüfgröße t_p; $t_p = \dfrac{\overline{X}-\mu}{S/\sqrt{n}}$ = (95-92)/(32/20)^{0,5} = 2,37

H_o: $\mu \leq \mu_o$; $\mu_o = 92$
H_1: $\mu > \mu_o$

α/2	**kritische Werte t_K**	**t_p = 2,37; Vergleich $t_p - t_k$;**
5%	1.7291	$t_p > t_k \Rightarrow H_o$ verwerfen auf $\alpha/2$ = 5%
2,5%	2.0930	$t_p > t_k \Rightarrow H_o$ verwerfen auf $\alpha/2$ = 2,5%
1%	2.5395	$t_p < t_k \Rightarrow H_o$ nicht verwerfen auf $\alpha/2$ = 1%

? Wie würde sich das Ergebnis ändern, wenn statt n = 20 nur n = 10 wäre ?

zu b) $n = 10$
$S^2 = 25$
$\overline{X} = 65$
H_o: $\mu \geq \mu_o$: $\mu_o = 70$
H_1: $\mu < \mu_o$; ? Signifikant kleiner als 70 ?

Berechne die Prüfgröße t_p; $t_p = \dfrac{\overline{X}-\mu}{S/\sqrt{n}}$ = (65-7)/(25/10)^{0,5} = -3,16

H_o wird verworfen: Die Unterschiede sind signifikant auf allen α-Niveaus !!

zu c)

Langjähriger Ø einer Produktion bekannt. z.B. durchschnittliches Gewicht von Fischen nach 2 Monaten Aufzucht = 50 dkg

Stichprobe: 10 Fische werden entnommen: 45, 51, 49, 48, 42, 44, 50, 47, 49, 53
? hat sich das durchschnittliche Gewicht geändert ?

H_o: $\mu = \mu o$; $\mu_o = 50$
H_1: $\mu \neq \mu_o$

$N = 10$, $\nu = 9$
$S^2 = 3{,}36^2$
$\overline{X} = 47{,}8$

Berechne die Prüfgröße t_p; $t_p = \dfrac{|\overline{X} - \mu|}{S/\sqrt{n}} = |47{,}8 - 50|/(3{,}36/10^{0,5}) = 2{,}07$

α	**kritische Werte $t_{K,\nu=9}$**	**$t_p = 2{,}07$; Vergleich $t_p - t_k$;**
10%	1,8331	$t_p > t_k \Rightarrow H_o$ verwerfen auf $\alpha = 10\%$
5%	2,2622	$t_p < t_k \Rightarrow H_o$ akzeptieren auf $\alpha = 5\%$
1%	2.8214	$t_p < t_k \Rightarrow H_o$ akzeptieren auf $\alpha = 1\%$

t-Test und Welch-Test für den Vergleich zweier Mittelwerte unabhängiger Stichproben
(nach L. Sachs: Statistische Methoden)

Voraussetzung:

1. **Unabhängige Stichproben**
2. **Normalverteilung (zumindest angenähert)**

Nullhypothese und Alternativhypothese:

H_0: $\mu 1 = \mu 2$ Den beiden Stichproben liegen Grundgesamtheiten zugrunde, deren Mittelwerte µ1 und µ2 gleich sind

H_1: $\mu 1 \neq \mu 2$ Den beiden Stichproben liegen Grundgesamtheiten zugrunde, deren Mittelwerte µ1 und µ2 nicht gleich sind

H_0 wird auf dem 100α% Niveau abgelehnt, sobald die Prüfgröße $t_p \geq t_{\nu,\alpha}$ ist. t_p ist je nach Gleichheit oder Ungleichheit der Varianzen unterschiedlich zu berechnen.

Fall 1: Gleiche Varianzen $(\sigma_1^2 = \sigma_2^2)$:
(Die Gleichheit der Varianzen ist gegebenenfalls mit einem F-Test zu prüfen !!!)

$$t_p = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1-1)\,s_1^{\,2} + (n_2-1)\,s_2^{\,2}}{n_1+n_2-2}\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}}$$

$$\nu = n_1 + n_2 - 2$$

Für $n_1 = n_2$ ergibt sich eine wesentlich einfachere Prüfgröße:

$$t_p = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^{\,2}+s_2^{\,2}}{n_1}}}$$

$$\nu = 2n_1 - 2$$

Vertrauensbereiche für die Differenzen zweier Mittelwerte unabhängiger Stichproben

$$(\bar{x}_1 - \bar{x}_2) - t_{\nu,\alpha}\; A \le (\mu_1 - \mu_2) \le (\bar{x}_1 - \bar{x}_2) + t_{\nu,\alpha}\; A$$

$$\text{und } A = \sqrt{\frac{(n_1-1)\,s_1^{\,2} + (n_2-1)\,s_2^{\,2}}{n_1+n_2-2}\left[\frac{1}{n_1}+\frac{1}{n_2}\right]}$$

$$\text{Für } \mathrm{n}_1 = \mathrm{n}_2 \text{ gilt: } A = \sqrt{\frac{s_1^{\,2}+s_2^{\,2}}{n_1}}$$

Für den Fall $\sigma_1^{\,2} = \sigma_2^{\,2}$ <u>mit bekannten Varianzen</u> gilt:

$$(\bar{x}_1 - \bar{x}_2) - z\,\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \le (\mu_1 - \mu_2) \le (\bar{x}_1 - \bar{x}_2) + z\,\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

und für $n_1 = n_2$:

$$(\bar{x}_1 - \bar{x}_2) - z\sigma\sqrt{\frac{2}{n_1}} \le (\mu_1 - \mu_2) \le (\bar{x}_1 - \bar{x}_2) + z\sigma\sqrt{\frac{2}{n_1}}$$

mit z = 1.96 für den 95% VB und z = 2.58 für den 99% VB

Fall 2: <u>Ungleiche Varianzen, t- Test nach Welch oder Welch – Test.</u>

(unbekannte Varianzen, die möglicherweise ungleich sind, knapp $(\sigma_1^{\,2} \neq \sigma_2^{\,2})$.

Die t- Werte sind <u>Approximationen</u>, ν ist zur ganzen Zahl abzurunden. Der zweiseitige Test ist gegenüber Abweichungen von der Normalverteilung robust, sobald (1) $n_1 > 10$, $n_2 > 10$;

(2) ¼ $\le (n_1/n_2) \le 4$. Bei schiefen Verteilungen nicht zu unterschiedlichen Verteilungstyps sollte $n_1 = n_2 \sim\ge 20$ sein.

$$t_p = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1^{\,2}}{n_1} + \dfrac{s_2^{\,2}}{n_2}}}, \qquad n_1, n_2 \geq 6$$

$$\nu = \frac{\left(\dfrac{s_1^{\,2}}{n_1} + \dfrac{s_2^{\,2}}{n_2}\right)^2}{\dfrac{\left(\dfrac{s_1^{\,2}}{n_1}\right)^2}{n_1 - 1} + \dfrac{\left(\dfrac{s_2^{\,2}}{n_2}\right)^2}{n_2 - 1}}$$

Für $n_1 = n_2$

$$t_p = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1^{\,2} + s_2^{\,2}}{n_1}}},$$

$$\nu = \frac{(n_1 - 1)\left(s_1^{\,2} + s_2^{\,2}\right)^2}{\left(s_1^{\,2}\right)^2 + \left(s_2^{\,2}\right)^2}$$

Vertrauensbereiche für die Differenz zweier Mittelwerte unabhängiger Stichproben

$$(\bar{x}_1 - \bar{x}_2) - t_{\nu,\alpha}\, B \leq (\mu_1 - \mu_2) \leq (\bar{x}_1 - \bar{x}_2) + t_{\nu,\alpha}\, B, \ \mathit{mit}$$

$$B = \sqrt{\frac{s_1^{\,2}}{n_1} + \frac{s_2^{\,2}}{n_2}}$$

für $n_1 = n_2$ gilt:

$$B = \sqrt{\frac{s_1^{\,2} + s_2^{\,2}}{n_1}}$$