Test auf den Erwartungswert

- Wir interessieren uns für den Erwartungswert μ einer metrischen Zufallsgröße.
  Beispiele: Alter, Einkommen, Körpergröße, Scorewert ...
- Wir können einseitige oder zweiseitige Hypothesen formulieren.
- Beispiele
  - "Vor 10 Jahren betrug die Durchschnittsgröße von Studienanfängern und -anfängerinnen 167 cm. Heute sind sie im Schnitt größer als 167 cm."
  - Die Differenz zwischen Gewicht vor und nach einer Diät ist 0.

Zweiseitiger Gauss-Test auf den Erwartungswert μ

- X Zufallsgröße mit Erwartungwert μ.
- Zweiseitige Hypothese über μ:

$$\begin{aligned} H_0 &: \quad \mu = \mu_0 \\ H_1 &: \quad \mu \neq \mu_0 \end{aligned}$$

- Testgröße: Mittelwert in der Stichprobe $X_1, \dots, X_n$:

$$\overline{X} \quad = \quad \frac{1}{n} \sum_{i=1}^{n} X_i$$

- Stichprobenumfang n ist genügend groß
(Faustregel $n > 30$)

Zweiseitiger Gauss-Test auf den Erwartungswert μ

Hypothesen: $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$

Testentscheidung zum Signifikanzniveau α

Annahmebereich

$$\mu_0 \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{S^2}{n}}$$

mit

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2$$

H_0 wird abgelehnt, falls

$$\overline{X} < \mu_0 - z_{1-\frac{\alpha}{2}} \sqrt{\frac{S^2}{n}} \quad \text{oder} \quad \overline{X} > \mu_0 + z_{1-\frac{\alpha}{2}} \sqrt{\frac{S^2}{n}}$$

$z_{1-\frac{\alpha}{2}}$ ist das $(1-\alpha/2)$-Quantil der Standardnormalverteilung.

Zweiseitiger Gauss-Test: p–Wert

Der p-Wert ergibt sich zu :

$$2 \cdot \left[1 - \Phi \left(|\overline{X} - \mu_0| / \sqrt{\frac{S^2}{n}} \right) \right]$$

Φ ist die Verteilungsfunktion der Standard Normalverteilung

$$(\overline{X} - \mu_0) / \sqrt{\frac{S^2}{n}}$$

wird als t-Wert (oder z-Wert) bezeichnet.

Beispiel:
Verändert sich der Blutdruck nach einer Intervention

Beispiel zweiseitiger Test

- Nullhypothese: Die Blutdruckdifferenz ist 0.

$$H_0 \quad : \quad \mu = 0$$
$$H_1 \quad : \quad \mu \neq 0$$

- Testgröße: Durchschnittlicher Bluddruck
- Faustregel $n = 62 > 30$ ist erfüllt
  $\longrightarrow$ zweiseitiger Gauß-Test
- Falls wir annehmen, dass die Blutdruckdifferenz normalverteilt ist, können wir auch den zweiseitigen t-Test anwenden. (Standardeinstellung in SPSS)

Dualität Test und Konfidenzintervall

- Annahmebereich: Wir behalten H_0 bei, falls die Testgröße T in der Nähe von μ_0 liegt:
- Äquivalente Formulierung über ein Konfidenzintervall: Wir behalten H_0 bei, falls μ_0 in der Nähe der Testgröße liegt
- Wir behalten H_0 bei, falls μ_0 im Konfidenzintervall für die Differenz liegt
- Dabei hängen das Konfindenzniveau γ und das Signifikanzniveau α wie folgt zusammen:
  $1 - \alpha = \gamma$
- Dies gilt sehr allgemein für zweiseitige Test und Konfidenzintervalle
- Dies Prinzip kann zur Konstruktion von Konfidenzintervallen verwendet werden

Einseitiger Gauss-Test auf den Erwartungswert μ

- X Zufallsgröße mit Erwartungwert μ.
- Einseitige Hypothese über μ:

$$\begin{aligned} H_0 &: \quad \mu \leq \mu_0 \\ H_1 &: \quad \mu > \mu_0 \end{aligned}$$

- Testgröße: Mittelwert in der Stichprobe $X_1, \ldots, X_n$:

$$\overline{X} \quad = \quad \frac{1}{n} \sum_{i=1}^{n} X_i$$

- Stichprobenumfang n ist genügend groß
  (Faustregel $n > 30$)

Einseitiger Gauss-Test auf den Erwartungswert μ

Hypothesen: $H_0 : \mu \leq \mu_0$ vs. $H_1 : \mu > \mu_0$

Testentscheidung zum Signifikanzniveau α

Annahmebereich

$$\overline{X} \quad \leq \quad \mu_0 + z_{1-\alpha}\sqrt{\frac{S^2}{n}}$$

H_0 wird abgelehnt, falls

$$\overline{X} \quad > \quad \mu_0 + z_{1-\alpha}\sqrt{\frac{S^2}{n}}$$

$z_{1-\alpha}$ ist das $(1-\alpha)$-Quantil der Standardnormalverteilung.

Ein- bzw. zweiseitiger t-Test auf den Erwartungswert

Was tun wir, falls die Faustregel $n > 30$ nicht erfüllt ist?

Zusätzliche Voraussetzung

- Zufallsgröße X ist normalverteilt.

Wir stellen keine Bedingung an die Stichprobengröße n.

t-Test

- Der einseitige bzw. der zweiseitige t-Test auf den Erwartungswert μ hat die gleiche Form wie der einseitige bzw. zweiseitige Gauss-Test.
- Der t-Test unterscheidet sich vom Gauss-Test dadurch, dass wir das Quantil z der Standardnormalverteilung durch das Quantil t der t-Verteilung mit $n-1$ Freiheitsgraden ersetzen.

Verbundene und unverbundene Stichprobe

Verbundene, abhängige Stichprobe

Zwei Stichproben heißen verbunden, falls an einem Merkmalsträger (z.B. einer Person) zwei vergleichbare Merkmale erhoben werden.
Man nennt verbundene Stichproben oft auch abhängige Stichproben.

Beispiel: Das Ziel einer medizinischen Studie ist es, die Wirkung eines cholesterin-senkenden Medikaments zu überprüfen.

- Unterteilung der Probanden und Probandinnen in 2 Gruppen:
  1 Gruppe erhält das Medikament, 1 Gruppe erhält ein Placebo.
  unverbundene Stichprobe
- Alle Probanden und Probandinnen erhalten das Medikament. Von allen Personen wird der Cholesterinspiegel am Anfang und am Ende der Studie erhoben.
  verbundene Stichprobe

Approximativer Test auf Erwartungswert-Differenz bei unabhängigen Stichproben

Voraussetzungen:

- X und Y sind zwei Größen mit Erwartungswerten μ_X und μ_Y
- $X_1, \ldots, X_m$ und $Y_1, \ldots, Y_n$ **unabhängige** Stichproben
- Testgröße: standardisierte Differenz der Mittelwerte

$$\begin{aligned} T &= \overline{X} - \overline{Y} \\ &= \frac{1}{m}\sum_{i=1}^{m} X_i - \frac{1}{n}\sum_{i=1}^{n} Y_i \end{aligned}$$

- Faustregel: Stichprobenumfänge $m, n > 30$

Differenz von Erwartungswerten bei unabhängigen Stichproben

Annahmebereich

Für die beiden einseitigen Tests und den zweiseitigen Test auf die Differenz

$$H_0 : \mu_X - \mu_Y = d_0 \qquad H_0 : \mu_X - \mu_Y \leq d_0 \qquad H_0 : \mu_X - \mu_Y \geq d_0$$
$$H_1 : \mu_X - \mu_Y \neq d_0 \qquad H_1 : \mu_X - \mu_Y > d_0 \qquad H_1 : \mu_X - \mu_Y < d_0$$

ist der Annahmebereich

$$d_0 \pm z_{1-\frac{\alpha}{2}} \cdot s \qquad]-\infty, d_0 + z_{1-\alpha} \cdot s] \qquad [d_0 - z_{1-\alpha} \cdot s, \infty[$$

mit

$$s = \sqrt{\frac{s_X^2}{m} + \frac{s_Y^2}{n}}$$

Beispiel: Radio-Hördauer Ost-West

- Hören Personen in den alten Bundesländern im Schnitt mehr Radio?
  X : Hördauer im den alten Bundesländern,
  Y: Radiodauer in den neuen Bundesländern

$$
\begin{aligned}
H_0 &: \quad \mu_X - \mu_Y \leq 0 \\
H_1 &: \quad \mu_X - \mu_Y > 0
\end{aligned}
$$

- Befragung unter 253 Personen aus den alten Bundesländern und 932 Personen aus den neuen Bundesländern
  - unverbundene Stichproben $X_1, \ldots, X_{253}$ und $Y_1, \ldots, Y_{932}$
  - Stichprobengrößen $m = 253, n = 932 > 30$
- Durchschnittliche Hördauer:
  11.4 h (Standardabweichung 8.4 h) in den alten Bundesländern
  9.5 h (Standardabweichung 8.4 h) in den neuen Bundesländern

Beispiel: Radio-Hördauer Ost-West

- Signifikanzniveau: $\alpha = 0,1$
- Differenz der Radio-Hördauer

$$\overline{X} - \overline{Y} = 11.4 - 9.5 = 1.9$$

- Annahmebereich

$$\begin{aligned} \overline{X} - \overline{Y} &\leq z_{1-\alpha} \cdot \sqrt{\frac{s_X^2}{m} + \frac{s_Y^2}{n}} \\ &= z_{0.9} \cdot \sqrt{\frac{8.4^2}{932} + \frac{8.4^2}{253}} \\ &\approx 1.28 \cdot 0.65 \\ &\approx 0.83 \end{aligned}$$

- H_0 wird abgelehnt, Personen aus den alten Bundesländern hören signifikant länger Radio.

Doppelter t-Test auf die Erwartungswertdifferenz bei unabhängigen Stichproben

Voraussetzungen:

- X und Y sind zwei Größen mit Erwartungswerten μ_X und μ_Y
- $X_1, \ldots, X_m$ und $Y_1, \ldots, Y_n$ **unabhängige** Stichproben
- Testgröße: Differenz der Mittelwerte

$$\begin{aligned} T &= \overline{X} - \overline{Y} \\ &= \frac{1}{m}\sum_{i=1}^{m} X_i - \frac{1}{n}\sum_{i=1}^{n} Y_i \end{aligned}$$

- X und Y sind **normalverteilt**.
- Die Varianzen sind gleich $\sigma_X^2 = \sigma_Y^2$

Doppelter t-Test auf die Erwartungswertdifferenz bei unabhängigen Stichproben

Annahmebereich

Für die beiden einseitigen t-Tests und den zweiseitigen t-Test auf die Differenz

$$H_0 : \mu_X - \mu_Y = d_0 \qquad H_0 : \mu_X - \mu_Y \leq d_0 \qquad H_0 : \mu_X - \mu_Y \geq d_0$$
$$H_1 : \mu_X - \mu_Y \neq d_0 \qquad H_1 : \mu_X - \mu_Y > d_0 \qquad H_1 : \mu_X - \mu_Y < d_0$$

ist der Annahmebereich

$$d_0 \pm t_{(m+n-2;1-\frac{\alpha}{2})} s_d \qquad]-\infty, d_0 + t_{(m+n-2;1-\alpha)} s_d] \qquad [d_0 - t_{(m+n-2;1-\alpha)} s_d, \infty[$$

mit

$$s_d = s_{X,Y} \cdot \sqrt{\frac{1}{m} + \frac{1}{n}}$$
$$s^2_{X,Y} = \frac{1}{m+n-2}\left((m-1)S^2_X + (n-1)S^2_Y\right)$$

$t_{(m+n-2;1-\frac{\alpha}{2})}$ und $t_{(m+n-2;1-\alpha)}$ sind die Quantile der t-Verteilung mit $m+n-2$ Freiheitsgraden.

Tests auf Erwartungswertdifferenz bei abhängigen Stichproben

1. Gegeben ist eine verbundene Stichprobe
   $X_1, \dots, X_n$ und $Y_1, \dots, Y_n$
2. Bilde die Differenz

$$W_i = X_i - Y_i \quad i = 1, \dots, n$$

3. Führe einen Test auf den Erwartungswert von W durch
   - $n > 30 \longrightarrow$ Gauß-Test
   - W normalverteilt $\longrightarrow$ t-Test

Differenz von Anteilen bei unabhängigen Stichproben

- X und Y sind zwei Bernoulli-Größen mit

$$\begin{aligned} p_X &= P(X = 1) \\ p_Y &= P(Y = 1) \end{aligned}$$

- $X_1, \ldots, X_m$ und $Y_1, \ldots, Y_n$ **unabhängige** Stichproben
- Testgröße: Differenz der Anteile

$$\begin{aligned} T &= R_X - R_Y \\ &= \frac{1}{m}\sum_{i=1}^{m} X_i - \frac{1}{n}\sum_{i=1}^{n} Y_i \end{aligned}$$

- Faustregel: Stichprobenumfänge $m, n > 30$

Differenz von Anteilen bei unabhängigen Stichproben

Annahmebereich

Für die beiden einseitigen Tests und den zweiseitigen Test auf die Differenz

$$H_0 : p_X - p_Y = 0 \qquad H_0 : p_X - p_Y \leq 0 \qquad H_0 : p_X - p_Y \geq 0$$
$$H_1 : p_X - p_Y \neq 0 \qquad H_1 : p_X - p_Y > 0 \qquad H_1 : p_X - p_Y < 0$$

ist der Annahmebereich

$$\left[-z_{1-\frac{\alpha}{2}} \cdot s_R \,;\, z_{1-\frac{\alpha}{2}} \cdot s_R\right] \qquad \left]-\infty, z_{1-\alpha} \cdot s_R\right] \qquad \left[-z_{1-\alpha} \cdot s_R, \infty\right[$$

mit

$$s_R = \sqrt{R(1-R)\left(\frac{1}{m} + \frac{1}{n}\right)} \text{ und } R = \frac{m \cdot R_X + n \cdot R_Y}{m+n}$$

Beispiel: Ist Fernsehen informativ?

Weiterführung des Beispiels aus dem Thema ,,Schätzen"

- Beurteilen Personen aus den alten Bundesländern den Informationsgehalt im Fernsehen anders als Personen aus den neuen Bundesländern?
  zweiseitiger Test
  X: Person aus den alten Bundesländern hält Fernsehen für informativ
  Y: Person aus den neuen Bundesländern hält Fernsehen für informativ
- Signifikanzniveau: $\alpha = 0.05$
- Umfrage: 253 Personen aus den alten Bundesländern, 932 Personen aus den neuen Bundesländern:
  ,,Halten Sie Fernsehen für informativ? Ja/Nein"
  - unverbundene Stichproben $X_1, \ldots, X_{253}$ und $Y_1, \ldots, Y_{932}$
  - Stichprobengrößen $m = 253, n = 932 > 30$

Beispiel: Ist Fernsehen informativ?

- alte Bundesländer: 206 Personen halten Fernsehen für informativ
  neue Bundesländer: 747 Personen halten Fernsehen für informativ

- Anteile

$$R_X = \frac{206}{253} \approx 0.81 \quad \text{und} \quad R_Y = \frac{747}{932} \approx 0.80$$

- Standardfehler

$$\begin{aligned} R &= \frac{m \cdot R_X + n \cdot R_Y}{m+n} = \frac{253 \cdot 0.81 + 932 \cdot 0.8}{932 + 253} \\ &= \frac{950.53}{1185} \approx 0.802 \\ s_R &= \sqrt{R(1-R)(\frac{1}{m} + \frac{1}{n})} = \sqrt{0.802 \cdot (1 - 0.802) \left(\frac{1}{932} + \frac{1}{253} \right)} \\ &\approx 0.03 \end{aligned}$$

Beispiel: Ist Fernsehen informativ?

- Annahmebereich

$$\begin{aligned} \pm z_{1-\frac{\alpha}{2}} \cdot s_R &\approx \pm z_{0.975} \cdot 0.03 \\ &\approx \pm 1.96 \cdot 0.03 \\ &\approx \pm 0.059 \end{aligned}$$

- Differenz der Anteile in der Stichprobe

$$R_X - R_Y \approx 0.81 - 0.8 = 0.01$$

- H_0 wird beibehalten, der Unterschied zwischen alten und neuen Bundesländern ist nicht signifikant

Differenz von Anteilen bei abhängigen Stichproben

Voraussetzungen:

- X und Y sind zwei Bernoulli-Größen mit

$$\begin{aligned} p_X &= P(X=1) \\ p_Y &= P(Y=1) \end{aligned}$$

- $(X_1, Y_1), \ldots, (X_n, Y_n)$ **abhängige, verbundene** Stichproben
- Absolute Haufigkeiten werden in einer Kontingenztafel festgehalten

	Y=0	Y=1
X=0	n_{11}	n_{12}
X=1	n_{21}	n_{22}

- Verwende Test von McNemar

Zusammenhang zwischen 2 kategorialen Merkmalen

Sind zwei kategoriale Merkmale unabhängig? **Beispiele**

- Gibt es einen Zusammenhang zwischen besuchter Schule (Hauptschule, Realschule, Gymnasium) und Fernsehkonsum (hoch/niedrig)?
- Gibt es einen Zusammenhang zwischen Geschlecht (m/w) und der Affinität zu Fußball (Fan/kein Fan)?
- ...

→ Empirische Untersuchung mittels Kontingenztafeln und Kennzahlen

Randverteilungen

Die zu den einzelnen Merkmalen gehörigen empirischen Häufigkeitsverteilungen heißen **Randverteilungen** oder **marginale** Verteilungen.

Geschlecht \ Fußball	ja	nein	Summe
m	87	10	97
w	23	31	54
Summe	110	41	151

Randverteilung Geschlecht

m	w	
97	54	← Absolute Häufigkeiten
$\frac{97}{151}$	$\frac{54}{151}$	← Relative Häufigkeiten

Randverteilung Fußballfan

ja	nein	
110	41	← Absolute Häufigkeiten
$\frac{110}{151}$	$\frac{41}{151}$	← Relative Häufigkeiten

Bedingte Verteilung

Unter der **bedingten Verteilung** von X gegeben $Y = B_j$ versteht man die Verteilung von X in der Teilgesamtheit der Untersuchungseinheiten, die die Ausprägung B_j des Merkmals Y aufweisen.

Geschlecht \ Fußball	ja	nein	Summe
m	87	10	97
w	23	31	54
Summe	110	41	151

Verteilung „Fußballfan“
bei Männern

ja	nein	Ges.
87	10	97
90%	10%	100%

Verteilung “Geschlecht“
bei Fußballfans

m	w	Ges.
87	23	110
79%	21%	100%

Empirische Unabhängigkeit

- Falls die beiden Merkmale unabhängig voneinander sind, so sollte

  ,,relative Häufigkeit Fußballfan"
  ≈ ,,relative Häufigkeit Fußballfans unter Frauen"
  ≈,,relative Häufigkeit Fußballfans unter Männern"

- **Formel:**

  $$\text{bedingte relative Häufigkeit } \frac{n_{ij}}{n_{\bullet j}} = \frac{n_{i\bullet}}{n} \text{ relative Randhäufigkeit}$$

- Daraus folgt

  $$e_{ij} := \frac{n_{i\bullet} n_{\bullet j}}{n}$$

  ist die **unter Unabhängigkeit erwartete Anzahl**.

- Falls die Merkmale **unabhängig** sind, sollte gelten:

  $$e_{ij} \approx n_{ij}$$

χ^2-Unabhängigkeitstest

- Zwei Zufallsgrößen X und Y mit k bzw. l Ausprägungen

$$p_{ij} = P(X = i, Y = j)$$
$$p_{i\bullet} = P(X = i) \quad p_{\bullet j} = P(Y = j)$$

- Hypothesen:

$$\begin{aligned} H_0 &: \quad \text{X und Y sind stochastisch unabhängig} \\ & \quad\; p_{ij} = p_{i\bullet} \cdot p_{\bullet j} \text{ für alle } i = 1, \ldots, k, \; j = 1, \ldots, l \\ H_1 &: \quad \text{X und Y sind stochastisch abhängig} \\ & \quad\; p_{ij} \neq p_{i\bullet} \cdot p_{\bullet j} \text{ für mindestens eine ij-Kombination} \end{aligned}$$

- Prüfgröße:

$$\chi^2 = \sum_{i=1}^{k} \sum_{j=1}^{l} \frac{(n_{ij} - e_{ij})^2}{e_{ij}}$$

- Faustregel $n_{ij} \geq 5$ für alle i, j

χ^2-Unabhängigkeitstest

Annahmebereich

Für den χ^2-Unabhängigkeitstest ist der Annahmebereich

$$\chi^2 \quad \leq \quad q_{1-\alpha,\,(k-1)(l-1)}$$

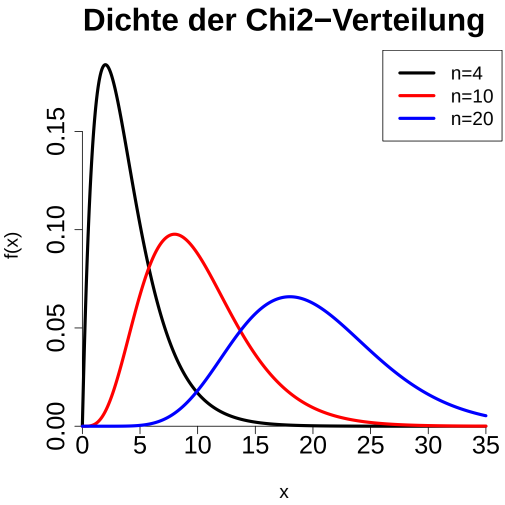
Die Nullhypothese wird abgelehnt, falls

$$\chi^2 \quad > \quad q_{1-\alpha,\,(k-1)(l-1)}$$

Dabei ist $q_{1-\alpha,\,(k-1)(l-1)}$ das

- $(1-\alpha)$-Quantil
- der χ^2-Verteilung
- mit $(k-1)\cdot(l-1)$ Freiheitsgraden.

Dichte der χ^2-Verteilung mit n Freiheitsgraden



Beispiel

$$e_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n}$$

Geschlecht \ Fußball	ja	nein	Summe
m	87	10	97
w	23	31	54
Summe	110	41	151

Erwartete Besetzungszahlen bei Unabhängigkeit

	ja (j=1)	nein (j=2)
m (i=1)	$\frac{97 \cdot 110}{151} \approx 71$	$\frac{97 \cdot 41}{151} \approx 26$
w (i=2)	$\frac{54 \cdot 110}{151} \approx 39$	$\frac{54 \cdot 41}{151} \approx 15$

$$\begin{aligned}
\chi^2 &= \sum_{i=1}^{k} \sum_{j=1}^{l} \frac{(n_{ij} - e_{ij})^2}{e_{ij}} \\
&\approx \frac{(87-71)^2}{71} + \frac{(10-26)^2}{26} + \frac{(23-39)^2}{39} + \frac{(31-15)^2}{15} \\
&\approx 37.09
\end{aligned}$$

Beispiel

- Signifikanzniveau: $\alpha = 0.01$
- Faustregel gültig?
  Besetzungszahlen $n_{ij} \geq 5$
- Bestimmung der Freiheitsgrade: $k = l = 2$

$$\text{Freiheitsgrade} \quad = \quad (k-1)\cdot(l-1) = (2-1)\cdot(2-1) = 1$$

$$q_{1-0.01;\,(2-1)(2-1)} \quad = \quad q_{0.09;\,1} \approx 6,63$$

- H_0 wird abgelehnt

Unabhängigkeit und Differenz von Anteilen

Die beide Fragen:

- Gibt es Unterschiede in den Anteilen von Y=1 zweier Gruppen ?
- Gibt es einen Zusammenhang zwischen Gruppen–Zugehörigkeit und einem binären Binären Merkmal Y ?

sind äquivalent.

Die beiden Testprozeduren Gauss Test für Differenzen von Anteilen und χ^2-Unabhängigkeitstest führen zum gleichen Ergebnis.