## 1.2 Skalentyp [S. 6]

| | | | |
|---|---|---|---|
| Nominalskala | (qualitative Daten) | $=, \neq$ | z.B. Geschlecht, Haarfarbe |
| Ordinalskala | (quantitative Daten) | $=, \neq, <, >$ | z.B. Rangplatz auf der Schachweltbestenliste, Windstärkeskala |
| Intervallskala | (quantitative Daten) | $=, \neq, <, >, +, -$ | z.B. Temperatur in Celisiusgrad |
| Verhältnisskala | (quantitative Daten) | $=, \neq, <, >, +, -, \cdot, \div$ | z.B. Länge (in m), Temperatur in Kelvingrad |

## 2.1 Häufigkeitsverteilung [A2 S.10 L121]

- $x_i$ Messwert
- $N$ Anzahl aller Messwerte $x_i$
- $h_i$: absolute Häufigkeit des $i$-ten Messwertes [S.8]
- $S_i$ kumulierte oder Summenhäufigkeit [S.8]
- $r_i$ relative Häufigkeit $r_i = \frac{h_i}{N}$ [S.9]
- relative Summenhäufigkeit [S.9]

## 2.2 Graphische Darstellung [S.9 ff] [A3 S.11 L121]

- Histogramm (Balkendiagram)
- Polygonzug (Punkte mit Linien verbunden)
  gut um zwei verschiedene Messungen zu vergleichen
- (relative) Summenhäufigkeitskurve
- Streifendiagram
- Kreisdiagram
- Histogramm (Balkendiagram)

## 2.3.1 Maße der zentralen Tendenz

- $\bar{x}$ Arithmetisches Mittel oder Mittelwert (Ausreißer haben mittlerer Wirkung) $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$ (2.1) [S.13]
- $\tilde{x}$ Median: Der Wert, unter dem genau so viele Werte wie darüber liegen (immun gegen Ausreißer) [S.14]
- Modalwert: Der Wert, der am häufigsten vorkommt. (immun gegen Ausreißer) [S.14]

## 2.3.2 Dispersionsmaße

- $s^2$ Varianz (Ausreißer haben ein wenig Wirkung) $s^2 = \dfrac{\sum_{i=1}^{N} x_i^2 - \frac{\left(\sum_{i=1}^{N} x_i\right)^2}{N}}{N-1}$ (2.5) [S.15]
- $s$ Standartabweichung (Ausreißer haben ein wenig Wirkung) [S.15]
- $v_p$ Perzentile: Der Wert, für den p Prozent aller Werte kleiner sind. (immun gegen Ausreißer) $v_p = \frac{p \cdot (N+1)}{100}$ für $0<p<100$ (2.7) [S.16] [A11 S.20 L123]
- $Q$ Mittlerer Quartilabstand: Streuungsmaß (immun gegen Ausreißer) $Q = \frac{C_{75} - C_{25}}{2}$ (2.8) [S.16]
- $R$ Range: Spannweite oder Varianzbreite (Ausreißer wirken sich zu 100% aus!) $R = \max\{x_1, \ldots, x_N\} - \min\{x_1, \ldots, x_N\}$ (2.10) [S.17]

## 2.3.4 Transformation

- z-Transformation: $x_i \mapsto z_i := \frac{1}{s}(x_i - \bar{x})$ D.h.: $\bar{z} = 0$ und $s_z = 1$ [S.19]
- IQ-Transformation: $x_i \mapsto q_i := 100 + 15 z_i$ D.h.: $\bar{q} = 100$ und $s_q = 15$ [S.19]
- und T-Transformation: $x_i \mapsto t_i := 50 + 10 z_i$ D.h.: $\bar{t} = 50$ und $s_t = 10$ [S.19]

## 2.4 Zusammenhangsmaße

- $s_{xy}$ Kovarianz: $s_{xy} = \dfrac{\sum_{i=1}^{N}\left((x_i - \bar{x})(y_i - \bar{y})\right)}{N-1}$ (2.12) [S.22]
- $r$ Produkt-Moment-Korrelation (**<u>invariant</u>** gegenüber zulässigen Transformationen): $r = \dfrac{\sum_{i=1}^{N}(x_i y_i) - \frac{1}{N}\sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i}{\sqrt{\left(\sum_{i=1}^{N} x_i^2 - \frac{1}{N}\left(\sum_{i=1}^{N} x_i\right)^2\right)\left(\sum_{i=1}^{N} y_i^2 - \frac{1}{N}\left(\sum_{i=1}^{N} y_i\right)^2\right)}}$

  (2.14) [S.23] [A16 S.29 L125]

- $r_s$ Rangkorrelation nach Spearman: $r_i$ ist der Rangplatz von $x_i$, $s_i$ ist der Rangplatz von $y_i$. $r_s = 1 - \frac{6\sum_{i=1}^{N}(r_i - s_i)^2}{N(N^2-1)}$ (2.15) [S.24]

  ↕ [A18 S.30 L127]

- $r_\tau$ Rangkorrelation nach Kendall: K ist die Anzahl der konkordanten, D die der diskordanten Paare $r_\tau = \frac{2(K-D)}{N(N-1)}$ (2.16) [S.25]

- $r_\varphi$ φ-Koeffizient: Vierfeldertafel $r_\varphi = \frac{|ad-bc|}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$ (2.17) [S.25] [A19 S.30 L128]

| | $y_1$ | $y_2$ | **Randsumme** |
|---|---|---|---|
| $x_1$ | a | b | **a+b** |
| $x_2$ | c | d | **c+d** |
| **Randsumme** | **a+c** | **b+d** | **a+b+c+d** |

- *C* Kontingenzkoeffizient: Kontingenztafel [S.26]

## 2.5 Regression

Regressionsgeraden: $y = b_{yx}x + a_{yx}$ (2.21) mit $b_{yx} = r\frac{s_y}{s_x}$ (2.22) und $a_{yx} = \bar{y} - b_{yx}\bar{x}$ (2.23)

$x = b_{xy}y + a_{xy}$ (2.24) mit $b_{xy} = r\frac{s_x}{s_y}$ (2.25) und $a_{xy} = \bar{x} - b_{xy}\bar{y}$ (2.26)

Die Regressionsgeraden schneiden sich im Punkt $(\bar{x}, \bar{y})$. [S.34] [A21,22,23 S.37 L129]
*bedingte* Verteilung: Man teilt die *Y*-Werte in eine Klasse ein, die zu einem festen *X*-Wert gehören. [S.35]
*Linearität* ist erfüllt, wenn die Mittelwerte der bedingten Verteilungen (der Klassen) auf einer Geraden liegen. [S.36]
*Homoskedastizität* ist erfüllt, wenn die Varianzen der bedingten Verteilungen gleich sind. [S.36]
$\Rightarrow$ *Standartschätzfehler*: (konstante Standartabweichung) $s_{y/x} = s_y\sqrt{1-r^2}$ (2.27) und $s_{x/y} = s_x\sqrt{1-r^2}$ (2.28) [S.36]
Regressionseffekt: Tendenz zur Mitte [S.36] [A24 S.37 L129]

## 3.1 Der Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ [S.40]

- $\Omega$ Zufallsergebnis: Menge aller möglichen Ergebnisse oder Ausgänge eines Vorgangs. [S.39]
- $\omega$ Elementarereignis: $\omega \in \Omega$ [S.39]
- $\mathcal{A}$ Zufallsereignis: $\mathcal{A} \subseteq 2^\Omega$ ($2^\Omega$ ist die Menge aller Teilmengen von $\Omega$) $\emptyset \in \mathcal{A}$: unmögliches Ergebnis $\Omega \in \mathcal{A}$: sicheres Ergebnis [S.39]
- *A* Ereignis: $A \subseteq \Omega$ — $\bar{A}$ Gegenereignis: $\bar{A} = \Omega \neg A$ ($\bar{A}$ gleich $\Omega$ ohne A)
- *P(A)* Wahrscheinlichkeit des Ereignisses *A* — $P(\bar{A})$ Wahrscheinlichkeit des Gegenereignisses $\bar{A}$: $P(\bar{A}) = 1 - P(A)$ [S.41]
- **Laplace-Experiment:** $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$, $P(\omega_i) = \frac{1}{n}$ (*alle Elementarereignisse* $\omega \in \Omega$ *sind gleichwahrscheinlich!*)

  $P(A) = \frac{Anzahl \cdot der \cdot günstigen \cdot Ergebnisse}{Anzahl \cdot der \cdot möglichen \cdot Ergebnisse} = \frac{|A|}{n}$ [S.41]

## 3.2 Bedingte Wahrscheinlichkeit und Unabhängigkeit

- Die Wahrscheinlichkeit von *A* unter der Bedingung *B*: $P(A|B) = \frac{P(A \cap B)}{P(B)}$ (3.11) [S.43]
- **Satz von Bayes:** $P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|\bar{A})P(\bar{A})}$ (3.12-13) [S.43] [A26 S.45 L131]
- Anwendung bedingter Wahrscheinlichkeiten: [S.43 f.] [A27 S.46 L131]

| | $A$ | $\bar{A}$ | Randsumme |
|---|---|---|---|
| $B$ | $P(A \cap B)$ | $P(\bar{A} \cap B)$ | $P(B)$ |
| $\bar{B}$ | $P(A \cap \bar{B})$ | $P(\bar{A} \cap \bar{B})$ | $P(\bar{B})$ |
| Randsumme | $P(A)$ | $P(\bar{A})$ | |

- *A* ist unabhängig von *B*: $P(A|B) = \frac{P(A \cap B)}{P(B)} = P(A)$ (3.14) $\Leftrightarrow$ $P(A \cap B) = P(A)P(B)$ (3.15) [S.45]

## 3.4 Verteilungen von Zufallsvariablen

- $X$ Zufallsvariable: Eine Abbildung $X : \Omega \to R$ die jedem Elementarereignis $\omega$ eine Zahl zuordnet.

***Diskrete Verteilung*** [S.49] [A32 S.56 L135]

$f_X$ Wahrscheinlichkeitsfunktion von $X$: $f_X(x) = P(X = x)$

$F_X$ Verteilungsfunktion von $X$: $F_X(x) = P(X \le x) = \sum_{x_i \le x} f(x_i)$

***Stetige Verteilung*** [S.49]

$f_X$ Dichtefunktion

$F_X$ Verteilungsfunktion von $X$: $F_X(x) = P(X \le x) = \int_{-\infty}^{x} f(t)dt$

- Die Verteilungsfunktion $F_X$ ist für diskrete und für stetige $X$ monoton steigend.

## 3.4.3 Parameter von Zufallsvariablen

***Formeln für diskrete X*** | ***Formeln für Laplace-Experiment***

- $\mu = EX$ Erwartungswert von $X$: $\mu = EX = \sum_{i=1}^{n}(x_i P(X = x_i)) = \sum_{i=1}^{n}(x_i f_X(x_i))$ (3.27) [S.53] | $\mu = EX = \frac{1}{n}\sum_{i=1}^{n} x_i$ [S.53]
- $\sigma^2 = VarX$ Varianz von X: $\sigma^2 = VarX = \sum_{i=1}^{n}((x_i - EX)^2 P(X = x_i))$ (3.28) [S.53] | $\sigma^2 = VarX = \frac{1}{n}\sum_{i=1}^{n}(x_i - EX)^2$ [S.54]
- **Tschebyscheffsche Ungleichung:** $P(|X - \mu| \ge k\sigma) \le \frac{1}{k^2}$ für $k \in R_+^*$ (3.30) [S.54]
- $Cov(X,Y)$ **Kovarianz:** Maß für Zusammenhang von zwei Zufallsvariablen $Cov(X,Y) = E(X \cdot Y) - EX \cdot EY$ (3.40) [S.55]
- $p$ **Korellatioskoeffizient**: $p = p_{XY} = \frac{Cov(X,Y)}{\sqrt{VarX \cdot VarY}}$ (3.41) [S.56]

  allgemein gilt: Unabhängigkeit $\Rightarrow$ Unkorreliertheit    für Normalverteilung gilt: Unabhängigkeit $\Leftrightarrow$ Unkorreliertheit

## 5.4 Grundaufgaben der Kombinatorik

- Fundamentalprinzip [S.57]

| ***n ist die Anzahl der Kugeln*** | ***mit Zurücklegen/Wiederholung*** $(k \in N^*)$ | ***ohne Zurücklegen/Wiederholung*** $(k \le n)$ |
|---|---|---|
| ***Geordnete Stichprobe (mit Reihenfolge) mit Ziehen von k Kugeln:*** | $n^k$ [S.59] | $\frac{n!}{(n-k)!}$ [S.59] |
| ***Geordnete Stichprobe (mit Reihenfolge) mit Ziehen von n Kugel:*** | $\frac{n!}{n_1! \cdot n_2! \cdot \ldots \cdot n_r!}$ [S.60] | $n!$ [S.60] (Permutation von $n$) |
| ***Ungeordnete Stichprobe (ohne Reihenfolge) mit Ziehen von k Kugeln:*** | $\binom{n}{k}$ [S.59] | $\binom{n+k-1}{k}$ [S.59] |

[A33,A34 S.61 L137]

## Spezielle diskrete Verteilungen

- **Diskrete Gleichverteilung:** X hat endlich viele Werte $\{x_1, x_2, \cdots, x_n\}$, die alle gleichwahrscheinlich sind.

$$f_X(x) = \begin{cases} \frac{1}{n} \forall x \in \{x_1, x_2, \cdots, x_n\} \\ 0 \forall x \notin \{x_1, x_2, \cdots, x_n\} \end{cases} \quad (3.48) \quad \mu = EX = \frac{1}{n}\sum_{i=1}^{n} x_i \quad (3.49) \quad \sigma^2 = VarX = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \frac{1}{n^2}\left(\sum_{i=1}^{n} x_i\right)^2 \quad (3.50) \text{ [S.62]}$$

- **Bernoulliexperiment und Binomialverteilung** [S.63 f.]

  Ein Zufallsvorgang <u>mit Zurücklegen</u> wird $n$-mal wiederholt: $k$ ist die Anzahl des Ereignis $A$ mit $P(A) = p$ und $P(\overline{A}) = 1 - p$.

  Die Wahrscheinlichkeit, dass $A$ $k$-mal eintritt: **TABELLE A** oder $f_X(x) = P(X = x) = \binom{n}{k} p^k (1-p)^{n-k}$ für $k=0,1,\ldots,n$ (3.51) [S.63]

  $\mu = EX = n \cdot p$ (3.54) $\sigma^2 = VarX = n \cdot p(1-p)$ (3.55) [S.64] [A38 S.67 L139]

- **Hypergeometrische Verteilung** *(wie Binomialverteilung, nur <u>ohne Zurücklegen</u>)* [S.64 f.]

  Es wird $n$-mal aus einer Urne mit N Kugeln - davon M weiße - gezogen: $m$ gibt die Zahl der weißen Kugeln an, die gezogen wurden.

  Wahrscheinlichkeit, dass $m$ weiße Kugeln gezogen wurden: $f_X(x) = P(X = x) = \frac{\binom{M}{n}\binom{N-M}{n-m}}{\binom{N}{n}}$ für $m=0,1,\ldots,n$ (3.56) [S.64]

  $\mu = EX = n \cdot \frac{M}{N}$ (3.57) $\sigma^2 = VarX = n \cdot \frac{M}{N} \cdot \frac{N-M}{N} \cdot \frac{N-n}{N-1}$ (3.58)

- **Poissonverteilung** [S.65 ff.]

## 3.7 Normalverteilung

- $X$ ist **normalverteilt**, wenn die Messwerte eine Glockenkurve bilden: $X$ ist in $\mathcal{N}(\mu,\sigma)$
- $X$ ist **standartnormalverteilt**, wenn $X$ **normalverteilt** ist, $\mu = EX = 0$ und $\sigma^2 = VarX = 1$ **TABELLE B**
- Ist $X$ **normalverteilt**, dann ist $Z$ **standartnormalverteilt:** $Z = \frac{X - EX}{\sqrt{VarX}} = \frac{X-\mu}{\sigma}$ (3.66) [S.68]
- Ist $X$ **normalverteilt**, dann sind $P(|X-\mu| \le \sigma) = 0{,}682$, $P(|X-\mu| \le 2\sigma) = 0{,}955$ und $P(|X-\mu| \le 3\sigma) = 0{,}997$ (3.67-69) [S.69]
- **Zentraler Grenzwertsatz:** für $N>20$ Normalverteilungsannahme [S.69 f.]
- **Gesetz der großen Zahlen:** *„Gambler's fallacy"* [S.70 f.]

## 3.8 Testverteilungen

- $\chi^2(n,\alpha)$ $\chi^2$-Verteilung mit $n$ Freiheitsgeraden: $P(X \le \chi^2_{n,\alpha}) = 1-\alpha$ (3.77) $\mu = EX = n$ (3.78) $\sigma^2 = VarX = 2n$ (3.79) [S.73]
  $n := df$ **TABELLE C**
- $t(n,\alpha)$ $t$-Verteilung mit $n$ Freiheitsgeraden: $P(Y \le t_{n,\alpha}) = 1-\alpha$ (3.80) $\mu = EY = 0$ $\forall n > 1$ (3.81) [S.73]
  $\sigma^2 = VarY = \frac{n}{n-2}$ $\forall n > 2$ (3.82) $P(-t_{n,\alpha/2} \le Y \le t_{n,\alpha/2}) = 1-\alpha$ (3.83) [S. 73 f.] $n := df$ **TABELLE D**
- $F((n,m),\alpha)$ $F$-Verteilung mit $m$ Zählerfreiheitsgeraden und $n$ Nennerfreiheitsgeraden: : $P(X \le F_{m,n;\alpha}) = 1-\alpha$ (3.84)
  $\mu = EX = \frac{n}{n-2}$ $\forall n > 2$ (3.85) $\sigma^2 = VarX = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$ $\forall n > 4$ (3.86) [S.74] $m := df_{Zähler}$ und $n := df_{Nenner}$ **TABELLE E**

## 4.1 Inferenzstatistik: Die Brücke zur Anwendung der Wahrscheinlichkeitsrechnung

| Population | Kennwert | Stichprobe |
|---|---|---|
| $\lvert\Omega\rvert$ | Größe/Anzahl der Messwerte | N |
| $\mu = EX$ | Erwartungswert/Mittelwert | $\bar{x}$ |
| $\tilde{\mu}$ | Median | $\tilde{x}$ |
| $\sigma$ | Standartabweichung | $s$ |
| $\sigma^2 = VarX$ | Varianz | $s^2$ |
| $p$ | Korrelation | $r$ |
| Wahrscheinlichkeitsfunktion/Dichte | $\leftrightarrow$ | relative Häufigkeitsverteilung |
| Verteilungsfunktion | $\leftrightarrow$ | relative Summenhäufigkeit |

[S.78]

- $s_{\bar{x}}$ **Standartfehler des Mittelwerts**: $s_{\bar{x}} = \frac{s}{\sqrt{N}}$ (4.2) S.80 (bei Stichproben ohne Zurücklegen mit $\sqrt{\frac{|\Omega|-N}{|\Omega|-1}}$ multiplizieren [S.81])
- Verschiedene Typen von Fragestellungen: **Parameterschätzung**, **Hypothesenprüfung** und **Anpassungstest** [S.81 ff.]
- **abhängige Stichproben** und **Messwiederholungen:** [A 46 S.87 **L142**]

## 4.2 Parameterschätzung

- **Erwartungswert einer Zufallsvariable** [S.88]
  *Vorgehen:* wähle Irrtumswahrscheinlichkeit z.B. $\alpha = 0{,}05$; berechne Stichprobenmittelwert $\bar{x}$, Stichprobenvarianz $s^2$ und $s_{\bar{x}}$ Standartfehler des Mittelwerts; Freiheitsgrade $df = N-1$; $t_{N-1,\alpha/2}$: **TABELLE D** (Bei Stichprobengröße $N > 30$ $z_{\alpha/2}$: **TABELLE B**)
  Mit Wahrscheinlichkeit $1-\alpha$ gilt: $\bar{x} - t_{N-1,1/\alpha} s_{\bar{x}} \le \mu \le \bar{x} + t_{N-1,1/\alpha} s_{\bar{x}}$ (4.3) [S.88] [A 48a S.90 L140]
- **Vertrauensintervall für Parameter *p* einer Binomialverteilung** [S.88 f.]
  *Vorgehen:* wähle Irrtumswahrscheinlichkeit z.B. $\alpha = 0{,}05$; relative Häufigkeit des Eintretens von $A$ $\frac{k}{N}$ berechnen; $z_{\alpha/2}$: **TABELLE B**
  Mit Wahrscheinlichkeit $1-\alpha$ gilt: $\frac{k}{N} - z_{1/\alpha}\sqrt{\frac{\frac{k}{N}\left(1-\frac{k}{N}\right)}{N}} \le p \le \frac{k}{N} + z_{1/\alpha}\sqrt{\frac{\frac{k}{N}\left(1-\frac{k}{N}\right)}{N}}$ (4.4) [S.89] [A 49 S.90 L141]
- **Bestimmung der Stichprobengröße** [S.89 f.]
  *Vorgehen:* Irrtumswahrscheinlichkeit z.B. $\alpha = 0{,}05$, Länge des Vertrauensintervalls $\Delta$ wählen
  - für $\mu$ (irgendeine Information über die Varianz ist notwendig): $N = \frac{4z^2_{\alpha/2}\sigma^2}{\Delta^2} \approx \frac{4z^2_{\alpha/2}s^2}{\Delta^2}$ (4.6) [S.89] [A 48b S.90 L141]
  - für $p$ ($\hat{p}$ ist eine Schätzung für $p$): $N = \frac{4z^2_{\alpha/2}\hat{p}(1-\hat{p})}{\Delta^2}$ (4.7) [S.89]

## 4.3 Testen von Unterschiedshypothesen

| Skalenniveau | $H_0$ Hypothese | unabhängige Stichprobe | abhängig Stichprobe |
|---|---|---|---|
| Nominal | gewisse Häufigkeitsverhältnisse in den Populationen sind gleich | *Vierfelder* $\chi^2$ -Test | *McNemar*-Test |
| Ordinal | Die Wahrscheinlichkeit ist $1/2$, dass Werte aus einer Population größer sind als aus einer andern | *U*-Test | *Wilcoxon*-Test |
| Intervall | Erwartungswerte in beiden Populationen sind gleich | $t_{\text{hom}}$ oder *U*-Test | $t_{cor}$ oder *Wilcoxon*-Test |
| Intervall | Varianzen in beiden Populationen sind gleich | *F*-Test | $t_{cor_s}$ |

[S.104]

- $t_{\text{hom}}$ **Test für unabhängige Stichproben** [S.92]
  Man vergleicht die Erwartungswerte von zwei unabhängigen Stichproben mit den Stichprobengrößen $N_1, N_2$
  *Voraussetzung:* Normalverteilung, Varianzhomogenität (Anpassungstest: F-Test mit großem Signifikanzniveau z.B. $\alpha = 0{,}25$ um den $\beta$ - Fehler klein zu halten. Wenn die Varianzen nicht homogen sind: *U*-Test)
  *Vorgehen:* $H_0: \mu_1 = \mu_2; H_1: \mu_1 \neq \mu_2$; wähle z.B. $\alpha = 0{,}05$; berechne Mittelwerte $\bar{x}_1, \bar{x}_2$ und Varianzen $s_1^2, s_2^2$
  $$s_g^2 = \frac{(N_1-1)s_1^2 + (N_2-1)s_2^2}{N_1+N_2-2} \quad \textit{Teststatistik:}\ t = \frac{\bar{x}_1 - \bar{x}_2}{s_g\sqrt{\frac{1}{N_1}+\frac{1}{N_2}}} \text{ mit } df = N_1+N_2-2 \quad (4.8) \quad t_{crit} := t_{N_1+N_2-2,\alpha/2} \ \textbf{TABELLE D}$$
  *Entscheidung(zweiseitig):* $|t| \geq t_{crit} \Rightarrow H_0$ verwerfen; $|t| < t_{crit} \Rightarrow H_0$ beibehalten [A50 S.96 L144]

- $t_{cor}$ **Test für abhängige Stichproben** [S.93]
  Man vergleicht die Erwartungswerte von zwei abhängigen Stichproben mit der Stichprobengröße $N$. Die Daten liegen paarweise vor.
  *Voraussetzung:* Normalverteilung der Differenzen der zusammengehörenden Messwertpaare
  *Vorgehen:* $H_0: \mu_1 = \mu_2; H_1: \mu_1 \neq \mu_2$; wähle z.B. $\alpha = 0{,}05$; berechne Mittelwerte $\bar{x}, \bar{y}$ und Differenzen $d_i = x_i - y_i$ (für *i=1, …, N)*
  $$s_d^2 = \frac{\sum_{i=1}^{N} d_i^2 - \frac{1}{N}\left(\sum_{i=1}^{N} d_i\right)^2}{N(N-1)} \quad \textit{Teststatistik:}\ t = \frac{\bar{x}-\bar{y}}{s_d} \text{ mit } df = N-1 \quad (4.9) \quad t_{crit} := t_{N-1,\alpha/2} \ \textbf{TABELLE D}$$
  *Entscheidung(zweiseitig):* $|t| \geq t_{crit} \Rightarrow H_0$ verwerfen; $|t| < t_{crit} \Rightarrow H_0$ beibehalten [A51 S.96 L145]

- ***F*-Test Varianzvergleich in unabhängigen Stichproben** [S.94]
  Man vergleicht die Varianzen von zwei unabhängigen Stichproben mit den Stichprobengrößen $N_1, N_2$
  *Voraussetzung:* Normalverteilung, Normalverteilung in beiden Populationen
  *Vorgehen:* $H_0: \sigma_1 = \sigma_2; H_1: \sigma_1 \neq \sigma_2$; wähle $\alpha$ (z.B. auf 0,25 bei Voraussetzung für $t_{\text{hom}}$); berechne
  $$\textit{Teststatistik:}\ F = \frac{\max(s_1^2, s_2^2)}{\min(s_1^2, s_2^2)} \quad (4.10) \quad \text{für } s_1^2 \geq s_2^2:\ df_{Z\ddot{a}hler} = N_1 - 1 \ \text{ und } \ df_{Nenner} = N_2 - 1$$
  für $s_1^2 < s_2^2$: $df_{Z\ddot{a}hler} = N_2 - 1$ und $df_{Nenner} = N_1 - 1$ $F_{crit} := Ft_{df_{Z\ddot{a}hler}, df_{Nenner};\alpha}$ **TABELLE E**
  *Entscheidung(zweiseitig):* $F \geq F_{crit} \Rightarrow H_0$ verwerfen; $F < F_{crit} \Rightarrow H_0$ beibehalten [A50 S.96 L144]

- ***U*-Test von Mann-Whitney für unabhängige Stichproben** [S.97 f.]
  *Voraussetzung:* stetige Verteilung in den Populationen, unabhängige Stichproben mit den Stichprobengrößen $N_1, N_2$
  *Vorgehen:* $H_0: P(X \leq Y) = \frac{1}{2}; H_1: P(X \leq Y) \neq \frac{1}{2}$; wähle z.B. $\alpha = 0{,}05$; bringe die Stichproben in eine <u>gemeinsame</u>(!) Rangreihe;
  berechne $R_1$ *:= Summe der Rangplätze der 1. Stichprobe* und $R_2$ *:= Summe der Rangplätze der 2. Stichprobe*
  $$\textit{Teststatistik:}\ U = N_1N_2 + \frac{N_1(N_1+1)}{2} - R_1 \quad (4.13) \quad U' = N_1N_2 + \frac{N_2(N_2+1)}{2} - R_2 \quad (4.14) \quad U + U' = N_1N_2 \quad U_{crit}: \textbf{TABELLE F}$$
  *Entscheidung(zweiseitig):* $\min(U, U') \leq U_{crit} \Rightarrow H_0$ verwerfen; $\min(U, U') > U_{crit} \Rightarrow H_0$ beibehalten
  *wenn* $H_0$ verworfen wurde: $U < U' \Rightarrow P(X \leq Y) < \frac{1}{2}$; $U > U' \Rightarrow P(X \leq Y) > \frac{1}{2}$ [A53 S.100 L146]

- *Wilcoxon*-**Test für unabhängige Stichproben** [S.99]
  *Voraussetzung:* stetige Verteilung in den Populationen, abhängige Stichproben mit der Stichprobengrößen $N$ , Messwertpaare
  *Vorgehen:* $H_0: P(X \le Y) = \frac{1}{2}; H_1: P(X \le Y) \ne \frac{1}{2}$ ; wähle z.B. $\alpha = 0{,}05$ ; berechne Differenzen $d_i = x_i - y_i$ für $i=1, \ldots, N$ ;
  Beträge von $d_i$ in eine Rangreihe bringen; $R_1 :=$ *Summe der Rangplätze der positiven* $d_i$ ; $R_2 :=$ *Summe der Rangplätze negativen* $d_i$
  *Teststatistik:* $R = \min(R_+, R_-)$  $R_{crit}$ : **TABELLE G**
  *Entscheidung(zweiseitig):* $R \le R_{crit} \Rightarrow H_0$ verwerfen; $R > R_{crit} \Rightarrow H_0$ beibehalten
  *wenn* $H_0$ verworfen wurde: $R_- < R_+ \Rightarrow P(X \le Y) < \frac{1}{2}$; $R_+ < R_- \Rightarrow P(X \le Y) > \frac{1}{2}$ [A54 S.100 L146]

- *Vierfelder* $\chi^2$ **-Test für unabhängige Stichproben** [S.101]
  *Voraussetzung:* zwei Klassen $K_1, K_2$ , zwei alternative Reaktionskategorien $R, \overline{R}$ , unabhängige Stichproben, weiteres siehe Anmerkung.
  $p_1 / p_2$ seien die Wahrscheinlichkeiten, mit der in der ersten/zweiten Klasse das Ereignis $R$ eintritt.
  *Vorgehen:* $H_0: p_1 = p_2 ; H_1: p_1 \ne p_2$ ; wähle z.B. $\alpha = 0{,}05$ ; Vierfeldertafel berechnen:

| | $K_1$ | $K_2$ | |
|---|---|---|---|
| $R$ | $a$ | $b$ | $a+b$ |
| $\overline{R}$ | $c$ | $d$ | $c+d$ |
| | $a+c$ | $b+d$ | $N$ |

*Teststatistik:* $\chi^2 = N \frac{(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$ (4.16) [S.101] $\chi_{crit} := \chi^2_{1,\alpha}$ **TABELLE C**

*Entscheidung:* $\chi^2 \ge \chi_{crit} \Rightarrow H_0$ verwerfen; $\chi^2 < \chi_{crit} \Rightarrow H_0$ beibehalten [A55 S.103 L147]

**Anmerkung:** Stichprobengröße sollte $N \ge 20$ (besser $N \ge 40$ ) und folgende vier Werte $\ge 1$ (besser $\ge 5$ ) sein:

| | $K_1$ | $K_2$ |
|---|---|---|
| $R$ | $\frac{(a+b)(a+c)}{N}$ | $\frac{(a+b)(a+c)}{N}$ |
| $\overline{R}$ | $\frac{(a+b)(a+c)}{N}$ | $\frac{(a+b)(a+c)}{N}$ |

- *McNemar*-**Test für unabhängige Stichproben** [S.101]
  *Voraussetzung:* abhängige Stichproben, d.h. in der Klasse $K_1$ und $K_2$ wird von den gleichen Versuchspersonen festgestellt, ob sie zur gleichen Reaktionsklasse $R$ oder $\overline{R}$ gehören; es sollte $\frac{b+c}{2} \ge 5$ gelten

| | | $K_2$: $R$ | $K_2$: $\overline{R}$ | |
|---|---|---|---|---|
| $K_1$ | $R$ | $a$ | $b$ | $a+b$ |
| $K_1$ | $\overline{R}$ | $c$ | $d$ | $c+d$ |
| | | $a+c$ | $b+d$ | $N$ |

*b, c* sind die *Wechsler*, d.h. die Personen, die sich unter $K_1$ und $K_2$ verschieden verhalten.

*p* sei die Wahrscheinlichkeit, dass ein Wechsler zu *b* gehört. *Vorgehen:* $H_0: p = \frac{1}{2}; H_1: p \ne \frac{1}{2}$ ;

*Teststatistik:* $\chi^2 = \frac{(b-c)^2}{b+c}$ (4.17) $\chi_{crit} := \chi^2_{1,\alpha}$ **TABELLE C**

*Entscheidung:* $\chi^2 \ge \chi_{crit} \Rightarrow H_0$ verwerfen; $\chi^2 < \chi_{crit} \Rightarrow H_0$ beibehalten [A56 S.103 L148]