I Formelsammlung und Wissenswertes

Probeklausur

- **Merkmalstyp/Messniveau**
  - qualitatives Merkmal: Nominalskala (Familienstand, Farbe)
  - Rangmerkmale: Ordinalskala (Rangplätze bei Wettbewerben)
  - quantitatives (metrisches) Merkmal: Intervallskala
    - Nullpunkt NICHT absolut festgelegt
    - nur + und − sinnvoll
  - diskret/stetig: Verhältnisskala
- **dichotone Merkmale**: zwei Ausprägungen

Probeklausur

- **arithmetisches Mittel**

$$\bar{x} = \frac{1}{n}\left(\sum_{j=1}^{k} a_j h(X = a_j)\right) = \sum_{j=1}^{k} a_j r(X = a_j) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\forall\, x_1, \ldots, x_n \in \mathbb{R}\colon \quad \sum_{i=1}^{n}(x_i - \bar{x}) = 0$$

$$z_i = x_i + y_i \implies \bar{z} = \bar{x} + \bar{y}$$

$$z_i = ax_i + b \implies \bar{z} = a\bar{x} + b$$

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 \leq \sum_{i=1}^{n}(x_i - d) \quad \forall\, d \in \mathbb{R}$$

- **gewichtetes arithmetisches Mittel**: $\bar{x}_{gew} = \frac{1}{\sum_{i=1}^{n} g_i} \cdot \sum_{i=1}^{n} g_i x_i$

Probeklausur

- **Stichprobenvarianz**: (Streuungsmaßzahl)

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right) = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2\right)$$
$$= \frac{1}{n-1}\sum_{j=1}^{k}(a_j - \bar{x})^2 \cdot h(X = a_j)$$

- **Standardabweichung**: (Streuungsmaßzahl) $s = \sqrt{s^2}$
- Für $z_i = ax_i + b$ gelten: $s_z^2 = a^2 s_x^2$ und $s_z = |a| \cdot s_x$
- **empirische Verteilungsfunktion**: $F(x) = r(X \leq x) = \sum_{a_j \leq x} r(X = a_j)$

Probeklausur

- **empirischer Median**: (Lagemaßzahl) $\tilde{x} = \begin{cases} x_{\left(\frac{n+1}{2}\right)} & n \text{ ungerade} \rightsquigarrow \text{Zentralwert} \\ \frac{1}{2}\left(x_{\left(\frac{n}{2}\right)} + x_{\left(\frac{n+2}{2}\right)}\right) & n \text{ gerade} \rightsquigarrow \text{zentrales Wertepaar} \end{cases}$

$$\sum_{i=1}^{n}|x_i - \tilde{x}| \leq \sum_{i=1}^{n}|x_i - d| \quad \forall\, d \in \mathbb{R}$$

- **α-Quantil**: $x_\alpha = \begin{cases} \frac{1}{2}\left(x_{(n\alpha)} + x_{(n\alpha+1)}\right) & n\alpha \text{ ganzzahlig} \\ x_{([n\alpha+1])} & \text{sonst} \end{cases}$
- **r-tes Perzentil**: $\frac{r}{100}$-Quantil
- **Quartilabstand**: (Streuungsmaßzahl zum Median) $x_{\frac{3}{4}} - x_{\frac{1}{4}}$

Probeklausur
- **geometrisches Mittel**: $\bar{x}_G = \sqrt[n]{x_1 \cdot \ldots \cdot x_n}$
- **harmonisches Mittel**: $\bar{x}_H = n \cdot \frac{1}{\frac{1}{x_1}+\ldots+\frac{1}{x_n}}$
- **gewichtetes harmonisches Mittel**: $\bar{x}_{H,gew} = \sum_{j=1}^{n} g_j \cdot \frac{1}{\frac{g_1}{x_1}+\ldots+\frac{g_n}{x_n}}$
  g_j sind Zusatzinformationen zu x_j
- **Modalwert $\bar{x}_D$**: (Maßzahl) $h(\bar{x}_D) = \max\left\{h(a) \;\middle|\; a \in A\right\}$

Probeklausur
- **empirische Kovarianz**: $s_{x,y} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n-1}\left(\left(\sum_{i=1}^{n} x_i y_i\right) - (n\bar{x}\cdot\bar{y})\right)$
- **Standardabweichung**: $s_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$

Probeklausur
- **Pearsonscher Korrelationskoeffizient**: (Maßzahl) $r = r_{x,y} = \frac{s_{x,y}}{s_x \cdot s_y}$
  - $s_{x,y}$ enthält Informationen über evtl. Zusammenhang
  - s_x und s_y Normierungsfaktoren
  - $-1 \leq r \leq 1$
  - $r = 1 \iff y_i = ax_i + b$

Probeklausur
- **Regressionskoeffizient**: (Geradensteigung) $a = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \overset{n\geq 2}{=} \frac{s_{x,y}}{s_x^2}$

Probeklausur
- **Regressionskonstante**: (Schnittpunkt mit y-Achse) $b = \bar{y} - a\bar{x}$
  Bemerkung: $(\bar{x}, \bar{y})$ immer auf der Regressionsgerade

Probeklausur
- **Linearisierung durch Transformation**

$$y = b \cdot a^x \implies \underbrace{\ln(y)}_{=:y'} = \underbrace{\ln(b)}_{=:b'} + \underbrace{\ln(a)}_{=:a'} \cdot x \implies y' = b' + a'x$$

$$y = b \cdot x^a \implies \underbrace{\ln(y)}_{=:y'} = \underbrace{\ln(b)}_{=:b'} + \underbrace{\ln(x)}_{=:x'} \cdot a \implies y' = b' + ax'$$

- **Kontingenz**: (Zusammenhang nominalskalierter Merkmale) $E_{ij} = \frac{\overbrace{Z_i}^{\text{Zeilensumme}} \cdot \overbrace{S_j}^{\text{Spaltensumme}}}{n}$

Probeklausur

- **Kontingenzkoeffizient C**: ($\rightsquigarrow$ Nominalskala) $\chi^2 = \sum_{i=1}^{c_1} \sum_{j=1}^{c_2} \frac{(B_{ij}-E_{ij})^2}{E_{ij}}$

  c_1: Anzahl der Zeilen; c_2: Anzahl der Spalten

  $$C := \sqrt{\frac{\chi^2}{\chi^2+n}} \quad \rightsquigarrow \quad 0 \leq C < 1$$
  $$C_{\max} = \sqrt{\frac{m-1}{m}}$$
  $$m := \min(c_1, c_2)$$

- **korrigierter Kontingenzkoeffizient**: $C_k = \frac{C}{C_{\max}} = \sqrt{\frac{\chi^2 \cdot m}{(\chi^2+n)(m-1)}}$ (n: Anzahl der Fälle)

- **Wahrscheinlichkeitsverteilung als Abbildung**: $P \colon \mathcal{A} \longrightarrow \mathbb{R}\ ,\ A \longmapsto P(A)$

Probeklausur

Axiome: $\mathcal{A} \subset \mathcal{P}(\Omega)$

(1) $0 \leq P(A) \leq 1 \quad \forall\ A \in \mathcal{A}$

(2) $P(\Omega) = 1$

(3) $P(A \cup B) = P(A) + P(B) \quad$ falls $A \cap B = \emptyset$, $A, B \in \mathcal{A}$

weitere Eigenschaften

- $P(\emptyset) = 0$
- $P(A_1 \cup \ldots \cup A_r) = P(A_1) + \ldots + P(A_r) \quad$, falls $A_i \cap A_j = \emptyset$ für $i \neq j$
- $P(\bar{A}) = 1 - P(A) \quad$, wobei $\bar{A} = \Omega \setminus A$
- $A_1 \subset A_2 \implies P(A_1) \leq P(A_2)$
- $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$
- $P(A_1 \cup \ldots \cup A_r) \leq P(A_1) + \ldots + P(A_r)$

- **Gleichverteilung/Laplace-Verteilung**: $P(\omega) = \frac{1}{n} \quad \forall\ \omega \in \Omega$

- **Urnenmodelle**

  - geordnet, mit Zurücklegen

    $$\Omega_1 = \left\{ (a_1, \ldots, a_k) \in M^k \;\middle|\; 1 \leq a_i \leq n \right\}$$
    $$\#\Omega_1 = n^k$$

Probeklausur

  - geordnet, ohne Zurücklegen

    $$\Omega_2 = \left\{ (a_1, \ldots, a_k) \in M^k \;\middle|\; a_i \neq a_j \text{ für } i \neq j \right\}$$
    $$\#\Omega_2 = \frac{n!}{(n-k)!} \qquad \text{(Taschenrechner: } \texttt{nPr}\text{)}$$

– ungeordnet, mit Zurücklegen

$$\Omega_3 = \left\{(a_1, \ldots, a_k) \in M^k \;\middle|\; 1 \le a_1 \le a_2 \le \ldots \le a_k \le n\right\}$$
$$\#\Omega_3 = \binom{n+k-1}{k}$$

– ungeordnet, ohne Zurücklegen ($k \le n$)

$$\Omega_4 = \left\{(a_1, \ldots, a_k) \in M^k \;\middle|\; 1 \le a_1 < a_2 < \ldots < a_k \le n\right\}$$
$$\#\Omega_4 = \binom{n}{k} = \frac{n!}{(n-k)!k!} \qquad \text{(Taschenrechner: } \texttt{nCr}\text{)}$$

Probeklausur

- **bedingte Wahrscheinlichkeit**

$$P(A|B) = P_B(A) = \frac{P(A \cap B)}{P(B)}$$
$$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A) = P(B \cap A)$$

- **stochastisch unabhängig**

$$\begin{aligned} A \text{ stochastisch unabhängig von } B &\iff P(A|B) = P(A|\bar{B}) \\ &\iff P(A|B) = P(A) \\ &\iff P(A \cap B) = P(A) \cdot P(B) \end{aligned}$$

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- **Bayessche Formel**: $P(A_i|B) = \frac{P(B|A_i) \cdot P(A_i)}{P(B)} = \frac{P(B|A_i) \cdot P(A_i)}{\sum_{i=1}^{n} P(B|A_i) \cdot P(A_i)}$

  A_i Disjunktionen von Ω

Probeklausur

- **binomialverteilte Zufallsvariable**: $b_{n;p}(k) = B(n;p;k) = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$
  - $P(l \le X \le m) = \sum_{k=l}^{m} P(X = k) = \sum_{k=l}^{m} b_{n;p}(k)$
  - $P(0 \le X \le n) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \overset{\text{BLS}}{=} \big(p + (1-p)\big)^n = 1$
  - $b_{n;0.5}(k) = \frac{1}{2^n}\binom{n}{k}$
  - $P(X \ge k) = 1 - P(X < k) = 1 - \sum_{k=0}^{k-1} b_{n;p}(k)$

ÜB A34: Kiste mit Werkstücken

- **hypergeometrisch verteilte Zufallsvariable**: $P(X = k) = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}}$

  $\max(0, n - R + r) \le l \le \min(n, r)$

  Notation: $H(n; N; r)$-verteilt

- **stetig verteilte Zufallsvariable**: Verteilungsfunktion $F(x) = \int_{-\infty}^{\infty} f(t)\,\mathrm{d}t \quad \forall\, x \in \mathbb{R}$

- **normalverteilte Zufallsvariable**: $N(\mu,\sigma)$-verteilt, d. h. X stetig verteilt mit Dichte f

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) \quad \forall\, t \in \mathbb{R}$$

wobei μ Erwartungswert und σ Standardabweichung von X.

$N(0,1)$ heißt **standard-normalverteilt**. Dann gilt für die Verteilungsfunktion

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{x} \exp\left(-\frac{1}{2}t^2\right) \mathrm{d}t$$

Probeklausur

und für die Dichtefunktion

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2}x^2\right) \quad \forall\, x \in \mathbb{R}$$

$\varphi(-x) = \varphi(x)$

Probeklausur

$\Phi(-x) = 1 - \Phi(x)$

X ist $N(\mu,\sigma)$-verteilte Zufallsvariable. Dann ist

$$U = \frac{X-\mu}{\sigma}$$

eine $N(0,1)$-verteilte Zufallsvariable.

Probeklausur

X ist $N(\mu,\sigma)$ verteilt. Dann gilt:

$$P(X < b) = P(X \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right)$$

$B(n;p;k) \approx \frac{1}{\sigma} \cdot \varphi\left(\frac{k-\mu}{\sigma}\right)$ mit $\mu = n \cdot p$ und $\sigma = \sqrt{np(1-p)}$

- **Approximation der Binomialverteilung durch die Normalverteilung**
Für $p \in]0,1[$ und „große n“ (mindestens $n \cdot p \cdot (1-p) > 9$) ist

$$B(n;p;k) \approx \frac{1}{\sigma} \cdot \varphi\left(\frac{k-\mu}{\sigma}\right)$$

wobei $\mu = n \cdot p$, $\sigma = \sqrt{np(1-p)}$.

- **Erwartungswert**: $E(X) = \sum\limits_{\omega\in\Omega} X(\omega) \cdot P(\omega) = \sum\limits_{k=1}^{r} a_k \cdot P(X = a_k) = \int\limits_{-\infty}^{\infty} t \cdot f(t)\,\mathrm{d}t$

- **Varianz**: $\mathrm{Var}(X) = \begin{cases} \sum\limits_{\omega\in\Omega} \big(X(\omega)-\mu\big)^2 \cdot P(\omega) & \text{endlich viele } \omega \in \Omega \\ \int\limits_{-\infty}^{\infty} (t-\mu)^2 f(t)\,\mathrm{d}t & \text{stetige Zufallsvariable mit Dichte } f \end{cases}$

Ω oder $X(\Omega)$ endlich $\Longrightarrow$ $\mathrm{Var}(X) = \sum\limits_{k=1}^{r} (a_k-\mu)^2 \cdot P(X = a_k)$

Probeklausur

- **Standardabweichung**: $r_x = \sqrt{\mathrm{Var}(X)}$

$$\mathrm{Var}(X) = E\big((X-\mu)^2\big) = E(X^2) - \mu^2 = E(X^2) - \big(E(X)\big)^2$$
$$\mathrm{Var}(\lambda X) = \lambda^2 \cdot \mathrm{Var}(X)$$
$$\mathrm{Var}(X+\lambda) = \mathrm{Var}(X)$$

- **Standardisierung**: $X^{\varphi} = \frac{X-\mu}{\sigma_x}$ mit $\mu = E(X)$ und σ_x Standardabweichung

Probeklausur

X standardisiert $\iff E(X) = 0 \ \wedge \ \mathrm{Var}(X) = 1$

- **Covarianz**: $\sigma_{x,y} = \mathrm{Cov}(X,Y) = \sum\limits_{\omega\in\Omega} \big(X(\omega)-\mu\big)\big(Y(\omega)-\nu\big) = E\big((X-\mu)(Y-\nu)\big)$
mit $\mu = E(X)$ und $\nu = E(Y)$
- **Korrelationskoeffizient**: $\varrho = \varrho_{x,y} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\cdot\mathrm{Var}(Y)}} = \frac{\mathrm{Cov}(X,Y)}{\sigma_x\cdot\sigma_y}$

$$\mathrm{Cov}(X,X) = \mathrm{Var}(X)$$
$$E(X\cdot Y) = E(X)\cdot E(Y) + \mathrm{Cov}(X,Y)$$
$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\big(\mathrm{Cov}(X,Y)\big)$$

X, Y stochastisch unabhängig $\implies E(X\cdot Y) = E(X)\cdot E(Y) \ \wedge \ \mathrm{Var}(X+Y) = \mathrm{Var}(X)+\mathrm{Var}(Y)$

$X_1, \ldots, X_n$ stochastisch unabhängig $\implies \mathrm{Var}\,(X_1, \ldots, X_n) = \mathrm{Var}(X_1) + \ldots + \mathrm{Var}(X_n)$

- **Erwartungstreue**: $X_1, \ldots, X_n$ paarweise unabhängig, $E(X_i) = \mu$, $\mathrm{Var}(X_i) = \sigma^2$, $\bar{x} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$, $S^2 = \frac{1}{n-1}\sum\limits_{i=1}^{n}(X_i - \bar{X})^2 \implies E(S^2) = \sigma^2$, d. h. S^2 ist erwartungstreu

ÜB A31: Schließfach Bahnhof, Meyer

- **Anzahl der Permutationen**: Von n Objekten seien jeweils $n_1, \ldots, n_r$ Objekte gleich, wobei $n_1 + \ldots + n_r = n$ $(r \le n)$. Dann lassen sich n Objekte auf $\frac{n!}{n_1!\cdot\ldots\cdot n_r!}$ verschiedene Arten anordnen.
- **Signifikanztest** (Beispiel)
  (1) Zweifel an gleiche Wahrscheinlichkeiten beim Werfen einer Münze
  W: „Wappen oben“; Z: „Zahl oben“
  (2) $P(W)$ und $P(Z)$ unbekannt; H_0: $P(W) = P(Z) = \frac{1}{2}$; $\alpha = 0.05$
  (3) 200 mal werfen der Münze $\implies X_{200} \stackrel{\wedge}{=}$ Anzahl von „Wappen oben“
  (4) X_{200} binomialverteilt: $P_{H_0}(X_{200} = k) = \binom{200}{k}\left(\frac{1}{2}\right)^{200}$
  (5) gesucht: $c \in \mathbb{N}$, $c \in [0, 200]$, mit der Eigenschaft

$$P_{H_0}(|X_{200} - 100| > c) = P_{H_0}(X < 100 - c) + P_{H_0}(X > 100 + c) \le \alpha$$

  (6) Berechnung der Wahrscheinlichkeiten: $n\cdot p(1-p) = 200\cdot\frac{1}{2}\cdot\frac{1}{2} = 50 > 9$ ✓
  $\implies$ Verwendung der Normalverteilung

$$\mu = n\cdot p = 100 \quad \text{und} \quad \sigma = \sqrt{n\cdot p(1-p)} = \sqrt{50}$$

Mit (5) folgt

$$P_{H_0}(|X_{200}-100| > c) \leq \alpha \iff 1-P_{H_0}(|X_{200}-100| > c) = P_{H_0}(|X_{200}-100| \leq c) \geq 1-\alpha$$

$$\begin{aligned} P_{H_0}(|X_{200}-100| \leq c) &= P_{H_0}(100-c \leq X_{200} \leq 100+c) \\ &\approx \Phi\left(\frac{100+c-100}{\sqrt{50}}\right) - \Phi\left(\frac{100-c-100}{\sqrt{50}}\right) \\ &= \Phi\left(\frac{c}{\sqrt{50}}\right) - \Phi\left(\frac{-c}{\sqrt{50}}\right) = 2\Phi\left(\frac{c}{\sqrt{50}}\right) - 1 \end{aligned}$$

Also

$$P_{H_0}(|X_{200}-100| > c) \leq \alpha \iff 2\Phi\left(\frac{c}{\sqrt{50}}\right) - 1 \geq 1-\alpha \iff \Phi\left(\frac{c}{\sqrt{50}}\right) \geq 1-\frac{\alpha}{2}$$

gesucht: kleinstes $x \in [0,200]$, $c \in \mathbb{N}$, mit der Eigenschaft $\Phi\left(\frac{c}{\sqrt{50}}\right) \geq 1-\frac{\alpha}{2}$

(7) Bestimmung von c: $\Phi\left(\frac{c}{\sqrt{50}}\right) \approx 1-\frac{\alpha}{2} = 0.975$
$\xrightarrow{\text{Tabelle } N(0,1)\text{-Verteilung}} \Phi(X) = 0.975$ für $X \approx 1.96$
$\implies \frac{c}{\sqrt{50}} \approx 1.96 \implies c \approx 13.8 \longrightarrow$ wähle $c = 14$

(8) Entscheidungsregel: H_0 abgelehnt für $X_{200} < 86$ oder $X_{200} > 114$. Für $X_{200} \in [86,114]$ wird H_0 NICHT abgelehnt. Das heißt NICHT, dass für $X_{200} \in [86,114]$ H_0 beseitigt ist.

II Wichtige Excel-Befehle

Ein `$` Zeichen vor der Spalte bzw. Zeile stellt einen absoluten Bezug dar.

Summe der Zellen B2 bis B81	`=SUMME(B2:B81)`
Anzahl der beschriebenen Zellen zwischen B2 und B81	`=ANZAHL(B2:B81)`
Mittelwert der Zellen B2 bis B81	`=MITTELWERT(B2:B81)`
Median der Zellen B2 bis B81	`=MEDIAN(B2:B81)` (Probeklausur)
Maximum der Zellen B2 bis B81	`=MAX(B2:B81)`
Minimum der Zellen B2 bis B81	`=MIN(B2:B81)`
Absolute Häufigkeit des Wertes 1 in den Zellen B2 bis B81	`=ZÄHLENWENN($B$2:$B$81;1)`
Relative Häufigkeit des Wertes 1 in den Zellen B2 bis B81	`=ZÄHLENWENN($B$2:$B$81;1)/ANZAHL($B$2:$B$81)`
Varianz der Zellen B2 bis B81	`=VARIANZ(B2:B81)` (Probeklausur)
Standardabweichung der Zellen B2 bis B81	`=STABW(B2:B81)`
Wurzel des Wertes in Zelle B5	`=WURZEL(B5)`
unteres Quartil der Zellen B2 bis B81	`=QUARTILE(B2:B81;1)`
oberes Quartil der Zellen B2 bis B81	`=QUARTILE(B2:B81;3)`
a-tes Perzentil der Zellen B2 bis B81	`=QUANTIL(B2:B81;a/100)`
b-Quantil der Zellen B2 bis B81	`=QUANTIL(B2:B81;b)`
Kovarianz (x-Werte in A2 bis A8, y-Werte in B2 bis B8)	`=KOVAR(A2:A8;B2:B8)`
Korrelationskoeffizient	`=KORREL(A2:A8;B2:B8)`
Steigung der Regressionsgeraden	`=STEIGUNG(B2:B8;A2:A8)` (Probeklausur)
y-Achsenabschnitt der Regressionsgeraden	`=ACHSENABSCHNITT(B2:B8;A2:A8)`

Erstellen eines Punktdiagramms mit Regressionsgerade

(1) Tabelle mit Werten markieren $\longrightarrow$ Einfügen $\longrightarrow$ Diagramm

(2) Diagrammtyp: Punkt(XY) $\longrightarrow$ weiter $\longrightarrow$ weiter $\longrightarrow$ Fertig stellen

(3) Rechtsklick auf Punkt im Diagramm $\longrightarrow$ Trendlinie hinzufügen $\longrightarrow$ Regressionstyp linear

(4) ggf. Haken setzen bei Gleichung anzeigen

(5) Diagramm nach Belieben beschriften, skalieren und formatieren

III Wichtige R-Befehle

Wurzel von a	`sqrt(a)`
7. Spalte von dat wird GRB zugeordnet	`GRB<-dat[,7]`
Maximum von GRB	`max(GRB)`
Minimum von GRB	`min(GRB)`
Summe von GRB	`sum(GRB)`
arithmetisches Mittel von GRB	`mean(GRB)`
Median von GRB	`median(GRB)` (Probeklausur)
Standardabweichung von GRB	`sd(GRB)`
Varianz von GRB	`var(GRB)`
Sortierung von GRB	`sort(GRB)`
absolute Häufigkeitstabelle von GRB	`table(GRB)`
relative Häufigkeitstabelle von GRB	`table(GRB/length(GRB))`
Länge von GRB	`length(GRB)`
zusammenfassende Statistiken von GRB	`summary(GRB)`
Quantile von GRB	`quantile(GRB)`
a-tes Perzentil von GRB	`quantile(GRB, a/100, type=1)`
b-Quantil von GRB	`quantile(GRB, b, type=1)`
Datenreihe dat manuell mit Daten füllen	`dat<-c(149, 147, 158)` (Probeklausur)
empirische Verteilungsfunktion aus dat erstellen und Fn zuordnen	`Fn<-ecdf(dat)`
Fn zeichnen	`plot(Fn, main="Überschrift", xlab="Beschriftung x-Achse", ylab="Beschriftung y-Achse")`
Kovarianz nach Pearson der Datenreihen x und y	`cov(x,y)`
Korrelationskoeffizient nach Pearson	`cor(x,y)`
lineare Regression	`reg<-lm(y~x)`
Punktediagramm plotten (`lwd` ist Linienstärke)	`plot(y~x, main="Punktediagramm", col="blue", lwd=4)`
Regressionsgerade plotten	`abline(ref, lwd=3)`
kumulative Summe der Datenreihe x	`cumsum(x)`
Boxplot der Datenreihe x	`boxplot(x, main="Überschrift")` (Probeklausur)

- VorlDreiWuerfel.txt (im selben Ordner wie R-Datei gespeichert) in R einlesen und `dat` zuordnen funktioniert mit `dat<-read.table(file="VorlDreiWuerfel.txt",header=TRUE)`

- Histogramm wird geplottet mit `hist(GRB, xlab="Beschriftung x-Achse", ylab="Beschriftung y-Achse", main="Überschrift", cex.main=1.9, col="gray75", border="black")`
  `cex.main` skaliert dabei die Größe der Überschrift, `col` definiert die Füllfarbe des Plots und `border` legt die Farbe des Plots fest

- Datenmatrix `dat` manuell mit Daten füllen; findet Spaltenweise statt, `ncol` definiert die Anzahl der Spalten, `nrow` die Anzahl der Zeilen: `dat<-matrix(c(1,2,3,4,5,6,7,8,9,7,4,1,2,5,8), ncol=5, nrow=3))`
  Matrix sieht dann wie folgt aus:

1	4	7	7	2
2	5	8	4	5
3	6	9	1	8