Statistik

für Betriebswirtschaft, Internationales Management, Wirtschaftsinformatik und Informatik

Sommersemester 2017

HSA Statistik SS 2017 Sessionlist		
Datum	Statistik für BW/IM/I/Winf	Nr.
15.03.2017	Einführung Statistik	1
22.03.2017	Differentialrechnung, 2-dim Diff.Rechnung	2
29.03.2017	univ. deskr. Stat., Quantile, Plots	3
05.04.2017	Streuung, Konzentrationsmaße	4
12.04.2017	Kontingenztabellen, Mosaikplots, Korrelation	5
19.04.2017	Preisindizes, lineare Regression	6
26.04.2017	Wahrscheinlichkeitsbegriff	7
03.05.2017	Bedingte Wahrscheinlichkeit, Bayes	8
10.05.2017	diskrete Zufallsvariablen	9
17.05.2017	Stetige ZV, Gleichverteilung	10
24.05.2017	Pyramid	
31.05.2017	Normalverteilung, Verteilungsparameter	11
07.06.2017	Schätzfunktionen und Punktschätzer	12
14.06.2017	Konfidenzintervalle	13
21.06.2017	Wiederholung, Besprechung Probeklausur	14
28.06.2017	Prüfungswoche	15

- Übungsgruppe Jansen entfällt am 8.6. (Donnerstag)
- Offener Mathe/Statistikraum entfällt am 9.6. (Freitag)
- Mi 11.30 Uhr (W4.03) findet statt
- Di 9.50 W1.06 findet statt
- Di 13.40 Uhr fällt aus
- Do 15.6. (Jansen) auch

Hausaufgabe:
A 82-86

Stefan Etschberger

Rechenregeln für den Erwartungswert

- Ist f symmetrisch bzgl. a, so gilt $E(X) = a$
  Beispiel: f der Gleichverteilung symmetrisch bzgl. $\frac{a+b}{2} \Rightarrow E(X) = \frac{a+b}{2}$
- Lineare Transformation:

$$E(a + bX) = a + b \cdot E(X)$$

- Summenbildung:

$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$$

Beispiel: X gleichverteilt in $[0;10]$, $Y \sim N(1;1)$; $Z = X + 5Y$

$$E(Z) = E(X+5Y) = E(X) + E(5Y) = E(X) + 5 \cdot E(Y) = \frac{10+0}{2} + 5 \cdot 1 = 10$$

- Unabhängigkeit:

$$X,\ Y \text{ unabhängig} \Rightarrow E(X \cdot Y) = E(X) \cdot E(Y)$$

Streuungsparameter

- Varianz $\text{Var}(X)$ bzw. σ^2:

$$\text{Var}(X) = E([X - E(X)]^2) = \begin{cases} \sum\limits_i [x_i - E(X)]^2 f(x_i), & \text{wenn } X \text{ diskret} \\ \int_{-\infty}^{\infty} [x - E(X)]^2 f(x)\,dx, & \text{wenn } X \text{ stetig} \end{cases}$$

- Standardabweichung $\text{Sta}(X)$ bzw. σ:

$$\text{Sta}(X) = \sqrt{\text{Var}(X)}$$

Beispiel: Diskrete Verteilung

x	0	1	2
$f(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

:

$$\text{Var}(X) = (0-1)^2 \cdot \frac{1}{4} + (1-1)^2 \cdot \frac{1}{2} + (2-1)^2 \cdot \frac{1}{4} = \frac{1}{2}$$

Beispiel: Für eine exponentialverteilte Zufallsvariable X (Dichte siehe Erwartungswert) folgt

$$\begin{aligned} \text{Var}(X) &= \int_{-\infty}^{\infty} (x - E(X)) f(x)\,dx = \lambda \int_0^{\infty} \left(x - \tfrac{1}{\lambda}\right)^2 \cdot e^{-\lambda x}\,dx \\ &= e^{-\lambda x}\left(-x^2 + \frac{2x}{\lambda} - \left(\tfrac{1}{\lambda}\right)^2 - \frac{2}{\lambda^2} - \frac{2x}{\lambda} + \frac{2}{\lambda^2}\right)\Bigg|_0^{\infty} \\ &= 0 - \left(-0^2 - \left(\tfrac{1}{\lambda}\right)^2\right) = \frac{1}{\lambda^2} \end{aligned}$$

Rechenregeln für die Varianz

- **Verschiebungssatz:**

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2$$

Beispiel: Diskrete Verteilung

x	0	1	2
$f(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

:

$$\begin{aligned} \mathrm{E}(X^2) &= 0^2 \cdot \tfrac{1}{4} + 1^2 \cdot \tfrac{1}{2} + 2^2 \cdot \tfrac{1}{4} \\ &= \tfrac{3}{2} \\ \Rightarrow \quad \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 &= \tfrac{3}{2} - 1^2 = \tfrac{1}{2} = \mathrm{Var}(X) \end{aligned}$$

- **Lineare Transformation:**

$$\mathrm{Var}(a + bX) = b^2\,\mathrm{Var}(X)$$

- **Summenbildung** gilt nur, wenn die X_i unabhängig! Dann:

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$$

Erwartungswerte und Varianzen wichtiger Verteilungen

Verteilung von X	$E(X)$	$Var(X)$
Binomialverteilung $B(n;p)$	np	$np(1-p)$
Hypergeometrische Verteilung mit den Parametern N, M, n	$n\frac{M}{N}$	$n\frac{M}{N}\frac{N-M}{N}\frac{N-n}{N-1}$
Poisson-Verteilung $P(\lambda)$	λ	λ
Gleichverteilung in $[a;b]$ mit $a<b$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
Normalverteilung $N(\mu;\sigma)$	μ	σ^2

Anwendung: Ungleichung von Tschebyschow

- Für beliebige Zufallsvariablen X und $\varepsilon > 0$ gilt die Ungleichung von Tschebyschow:

$$P\big(|X - E[X]| \geqslant \varepsilon\big) \leqslant \frac{\mathrm{Var}[X]}{\varepsilon^2}$$

Beispiele:

- X ist gleichverteilt mit Parametern a, b und $\varepsilon = \frac{1}{3}(a-b)$, also $E[X] = \frac{1}{2}(a+b)$ und $\mathrm{Var}[X] = \frac{1}{12}(a-b)^2$

$$\Rightarrow P\Big(\big|X - \tfrac{1}{2}(a+b)\big| \geqslant \tfrac{1}{3}(a-b)\Big) \leqslant \frac{(a-b)^2}{12} \cdot \frac{3^2}{(a-b)^2} = 3/4$$

- $X \sim B(100; 0{,}2)$ und $\varepsilon = 10$
damit: $E[X] = 100 \cdot 0{,}2 = 20$ und $\mathrm{Var}[X] = 100 \cdot 0{,}2 \cdot (1 - 0{,}2) = 16$

$$\Rightarrow P\big(|X - 20| \geqslant 10\big) \leqslant \frac{16}{10^2} = 0{,}16$$

diskr. Stat.: Bravais-Pearson-Korr.-Koeff.

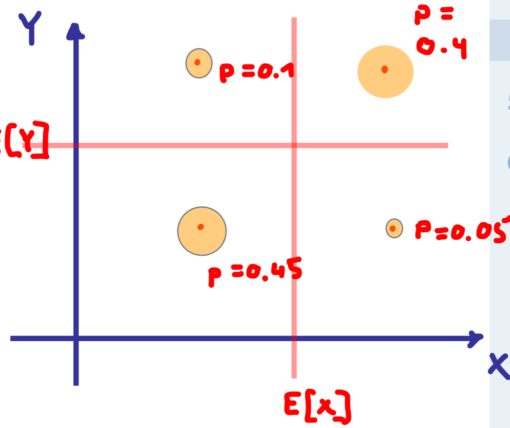
$$r = \frac{\frac{1}{n}\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n}\sum (x_i - \bar{x})^2 \frac{1}{n}\sum (y_i - \bar{y})^2}} \longrightarrow \text{Kovarianz}$$

- **Kovarianz:**

$$\begin{aligned} \mathrm{Cov}(X,Y) &= E[(X - E(X))(Y - E(Y))] \\ &= E(X \cdot Y) - E(X) \cdot E(Y) \end{aligned}$$
(Verschiebungssatz)

- **Korrelationskoeffizient:**

$$\rho(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X) \cdot \mathrm{Var}(Y)}}$$



- **Bemerkungen:**
  - ρ ist r nachgebildet $\Rightarrow \rho \in [-1;1]$
  - $|\rho| = 1 \iff Y = a + bX$ (mit $b \neq 0$)
  - $\rho = 0 \iff X, Y$ **unkorreliert**
- Varianz einer Summe zweier ZV:

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

Beispiel: 2 Zufallsvariablen

$y \backslash x$	-1	0	2	
1	0.5	0	0.05	0.55
2	0	0.1	0.35	0.45
	0.5	0.1	0.4	

$$Cov[X,Y] = E[X \cdot Y] - E[X] \cdot E[Y]$$

$$E[X] = -1 \cdot 0.5 + 0 \cdot 0.1 + 2 \cdot 0.4 = 0.3$$
$$E[Y] = 1 \cdot 0.55 + 2 \cdot 0.45 = 1.45$$

$$E[X \cdot Y] = (-1) \cdot 1 \cdot 0.5 + (-1) \cdot 2 \cdot 0 + 0 \cdot 1 \cdot 0 + 0 \cdot 2 \cdot 0.1$$
$$+ 2 \cdot 1 \cdot 0.05 + 2 \cdot 2 \cdot 0.35$$
$$= 1.0$$

$$Cov[X,Y] = 1.0 - 0.3 \cdot 1.45 = 0.565$$

$$Var[X] = E[X^2] - E^2[X] = (-1)^2 \cdot 0.5 + 0^2 \cdot 0.1 + 2^2 \cdot 0.4 - 0.3^2$$
$$= 2.01$$

$$Var[Y] = E[Y^2] - E^2[Y] = 1^2 \cdot 0.55 + 2^2 \cdot 0.45 - 1.45^2$$
$$= 0.2475$$

$$\Rightarrow \quad \rho[X,Y] = \frac{Cov[X,Y]}{\sqrt{Var[X] \cdot Var[Y]}} = \frac{0.565}{\sqrt{2.01 \cdot 0.2475}}$$

$$\approx 0.8011$$

Aufgabe 5 — 16 Punkte

Für die zweidimensionale Zufallsvariable (X,Y) sei folgendes bekannt:

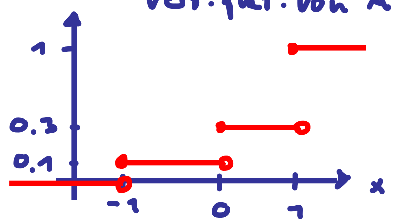
- X hat den Wertebereich $\{-2; 0; 2\}$
- Y hat den Wertebereich $\{0; 1\}$
- Es gilt: $P(X = -2) = P(X = 0) = 0.3$, $P(X = -2, Y = 0) = 0.3$, $P(X = 2, Y = 0) = 0.1$, $P(Y = 0) = 0.6$

a) Berechnen Sie die fehlenden gemeinsamen Wahrscheinlichkeiten und Randwahrscheinlichkeiten von (X,Y) und tragen Sie diese in eine passende Tabelle ein.

Gehen Sie für die Teilaufgaben b) ... d) von der zweidimensionalen Zufallsvariablen (A,B) mit

$\downarrow A \backslash B \rightarrow$	0	3	
−1	0.05	0.05	0.10
0	0.20	0.00	0.20
1	0.40	0.30	0.70
	0.65	0.35	1

Vert. fkt. von A



$$F(x) = \begin{cases} 0 & \text{f.} \quad x < -1 \\ 0.1 & \text{f.} \quad -1 \le x < 0 \\ 0.3 & \text{f.} \quad 0 \le x < 1 \\ 1 & \text{f.} \quad x \ge 1 \end{cases}$$

aus und berechnen Sie bitte folgende Größen:

b) Den Erwartungswert und die Varianz von A.

c) Den Erwartungswert und die Varianz von B.

d) Den Erwartungswert der Zufallsvariablen $C = A \cdot B$ sowie $\mathrm{Cov}[A,B]$.

Lösungshinweis:

a) Lösungstabelle:

$\downarrow X \backslash Y \rightarrow$	0	1	
−2	0.3	0.0	0.3
0	0.2	0.1	0.3
2	0.1	0.3	0.4
	0.6	0.4	1

b) $\mathrm{E}(A) = -1 \cdot 0.1 + 0 \cdot 0.2 + 1 \cdot 0.7 = 0.6$
$\mathrm{E}(A^2) = (-1)^2 \cdot 0.1 + 1^2 \cdot 0.7 = 0.8$
$\mathrm{Var}(A) = \mathrm{E}(A^2) - [\mathrm{E}(A)]^2 = 0.8 - (0.6)^2 = 0.44$

c) $\mathrm{E}(B) = 0 \cdot 0.65 + 3 \cdot 0.35 = 1.05$
$\mathrm{Var}(B) = 3^2 \cdot 0.35 - (1.05)^2 = 3.15 - 1.1025 = 2.0475$

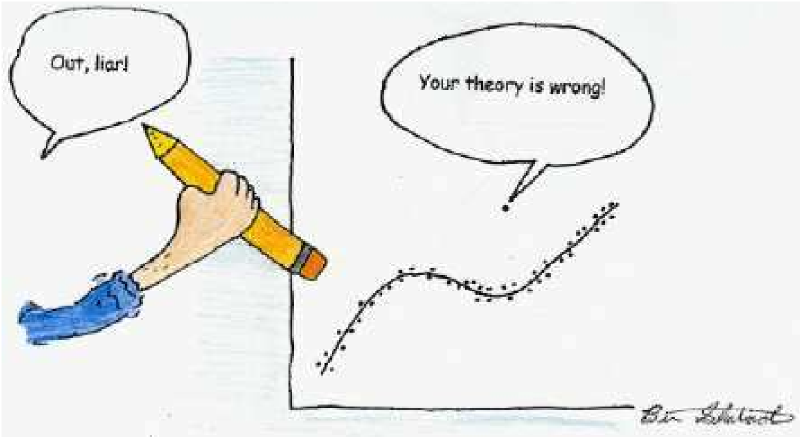
d) $\mathrm{E}(C) = \mathrm{E}(A \cdot B)$. Mit

$A \cdot B$	0	−3	3
$f(A \cdot B)$	0.65	0.05	0.3

folgt: $\mathrm{E}(A \cdot B) = -3 \cdot 0.05 + 3 \cdot 0.3 = 0.75$ und damit
$\mathrm{Cov}[A, B] = \mathrm{E}(A \cdot B) - \mathrm{E}(A) \cdot \mathrm{E}(B) = 0.75 - 0.6 \cdot 1.05 = 0{,}12$

Gliederung

1. Statistik: Einführung
2. Differenzieren 2
3. Deskriptive Statistik
4. Wahrscheinlichkeitstheorie
5. Induktive Statistik



5. Induktive Statistik
   - Grundlagen
   - Punkt-Schätzung
   - Intervall-Schätzung
   - Signifikanztests

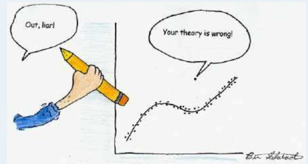
Grundlagen der induktiven Statistik

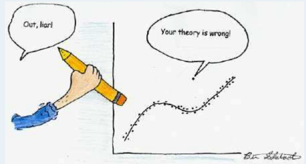
- Vollerhebung of unmöglich,
- Deshalb: Beobachte Teilgesamtheit und schließe auf Grundgesamtheit

Beispiel

Warensendung von 1000 Stück; darunter M Stück Ausschuss.
M ist unbekannt.
$\rightarrow$ Zufällige Entnahme von $n = 30$ Stück („Stichprobe“).
Darunter 2 Stück Ausschuss.
Denkbare Zielsetzungen:

- Schätze M durch eine Zahl (z.B. $\frac{2}{30} \cdot 1000 = 66{,}67$)
- Schätze ein Intervall für M (z.B. $M \in [58; 84]$)
- Teste die Hypothese, dass $M > 50$ ist.





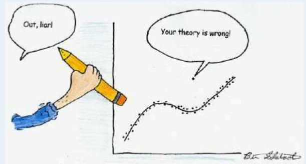
Grundbegriffe

- Grundgesamtheit (G): Menge aller relevanten Merkmalsträger.
- Verteilung von G: $F(x) = P(X \leqq x) =$ Wahrscheinlichkeit, dass ein Merkmalsträger ausgewählt wird, der beim untersuchten Merkmal maximal die Ausprägung x aufweist.
- Uneingeschränkte (reine) Zufallsauswahl:
  Jedes Element von G hat die selbe Chance, ausgewählt zu werden.
- Stichprobenumfang (n): Anzahl der Merkmalsträger in der Stichprobe.
- Einfache Stichprobe:
  Uneingeschränkte Zufallsauswahl und unabhängige Ziehung.
  $\rightarrow$ Alle Stichprobenvariablen $X_1, \ldots, X_n$ sind iid.
  identical independ. distributed
- Stichprobenergebnis:
  n-Tupel der Realisationen der Stichprobenvariablen, $(x_1, \ldots, x_n)$.

Wichtige Stichprobenfunktionen

- Gegeben: Einfache Stichprobe $X_1, \dots, X_n$, Beliebige Verteilung, mit $E(X_i) = \mu$, $Var(X_i) = \sigma^2$

Stichprobenfunktion V	Bezeichnung	$E(V)$	$Var(V)$
$\sum_{i=1}^{n} X_i$	Merkmalssumme	$n\mu$	$n\sigma^2$
$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$	Stichprobenmittel	μ	$\frac{\sigma^2}{n}$
$\frac{\bar{X} - \mu}{\sigma} \sqrt{n}$	Gauß-Statistik	0	1
$\frac{1}{n} \sum_{i=1}^{n} (X_i - \mu)^2$	mittlere quadratische Abweichung bezüglich μ	σ^2	
$\frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$	mittlere quadratische Abweichung	$\frac{n-1}{n} \sigma^2$	
$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$	Stichprobenvarianz	σ^2	
$S = \sqrt{S^2}$	Stichproben-Standardabweichung		
$\frac{\bar{X} - \mu}{S} \sqrt{n}$	t-Statistik		



Beispiel: Stichprobenfunktion $\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ (Stichprobenmittel)

$$E[\bar{X}] = E\left[\frac{1}{n} \cdot \sum_{i=1}^{n} X_i\right] = \frac{1}{n} \cdot \sum_{i=1}^{n} \underbrace{E[X_i]}_{\mu}$$

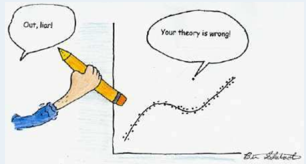
$$= \frac{1}{n} \cdot n \cdot \mu = \mu$$

Varianz: $\mathrm{Var}[\bar{X}] = \mathrm{Var}\left[\frac{1}{n} \cdot \sum_{i=1}^{n} X_i\right] \underset{\text{weil iid-Stichprobe}}{=} \left(\frac{1}{n}\right)^2 \cdot \sum_{i=1}^{n} \underbrace{\mathrm{Var}[X_i]}_{\sigma^2}$

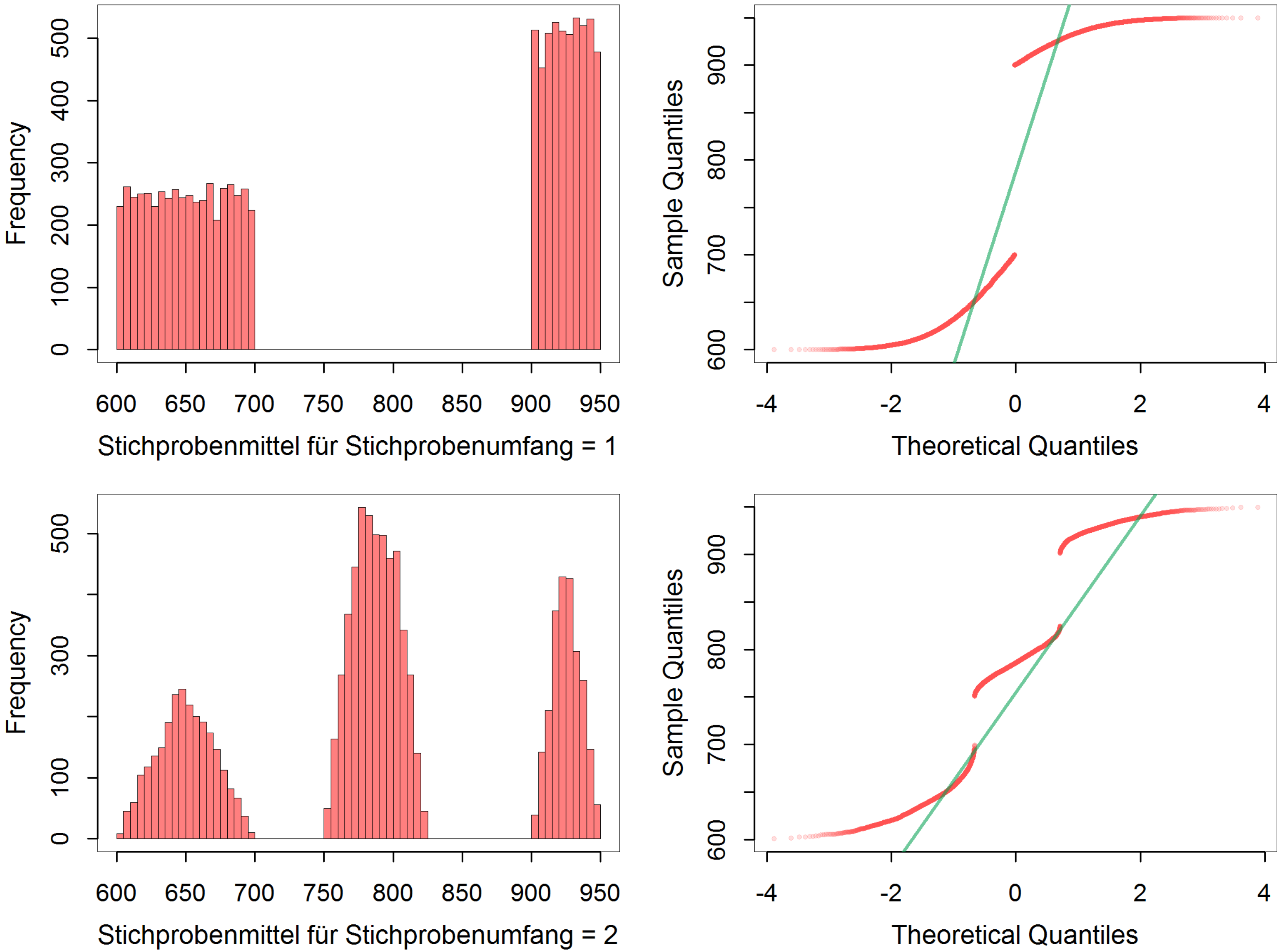
$$= \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n}$$

$$\Rightarrow \mathrm{Std}[\bar{X}] = \frac{\sigma}{\sqrt{n}}$$

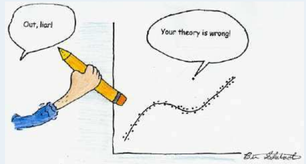
Auswirkungen der Stichprobengröße

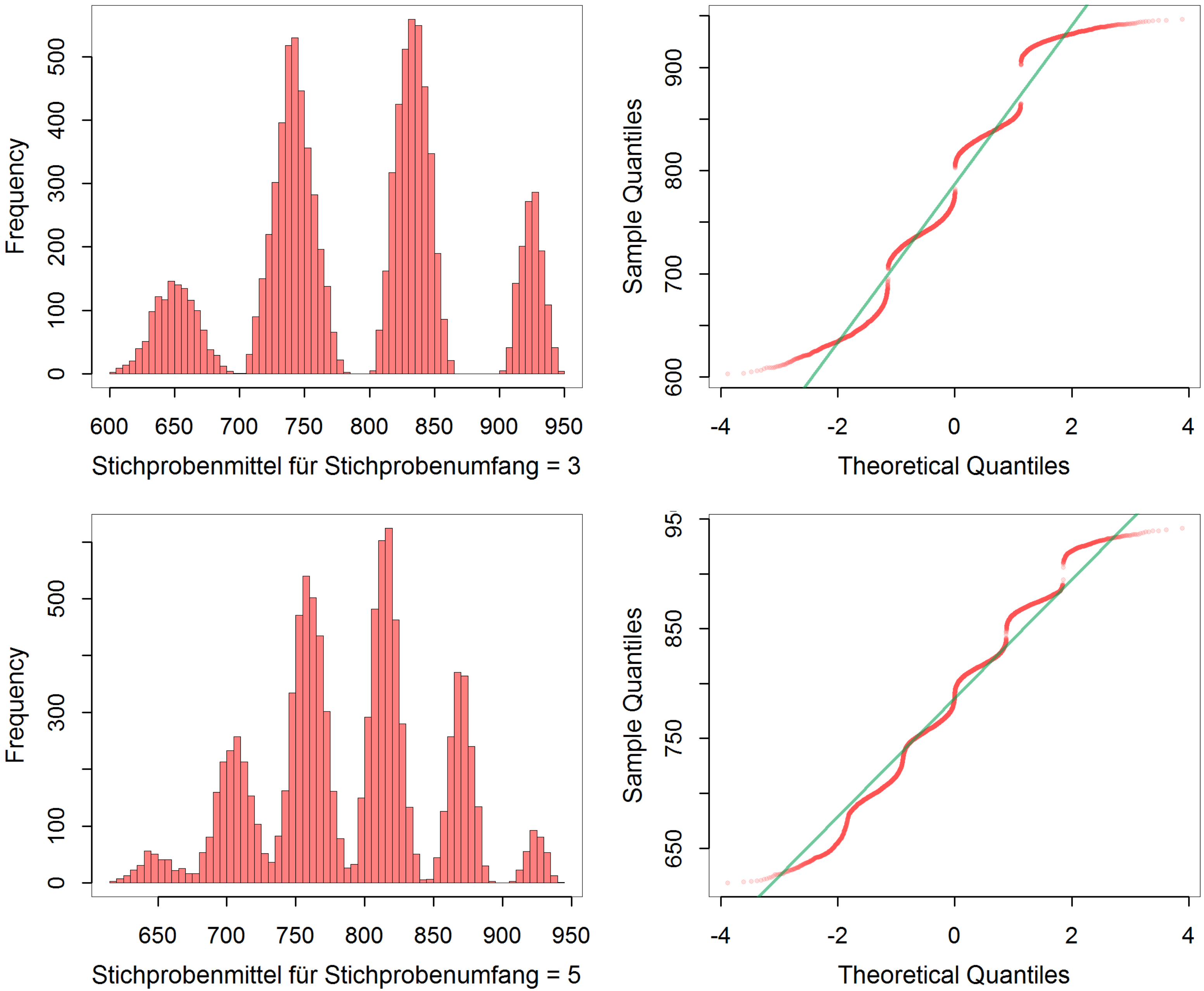


Ziehen von 10.000 Stichproben (jeweils vom Umfang n) und Berechnung der Stichprobenmittel (Verteilung: zwei überlagerte Gleichverteilungen):

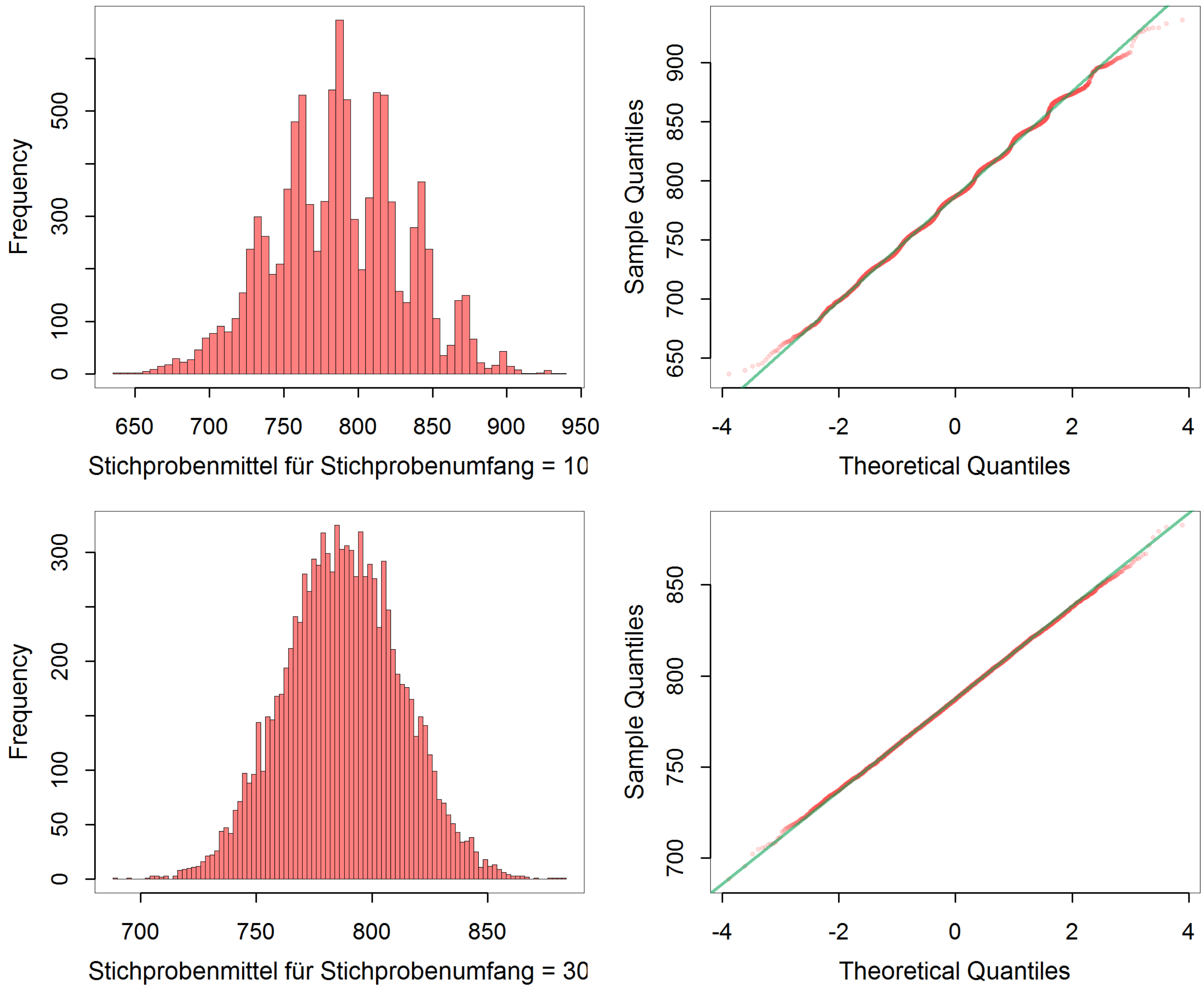


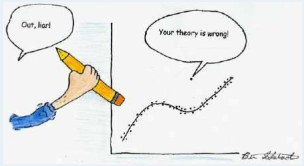
Auswirkungen der Stichprobengröße



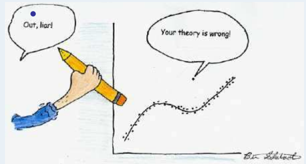


Auswirkungen der Stichprobengröße





Punkt-Schätzung



(kleiner) theta

- Ein unbekannter Parameter ϑ der Verteilung von G soll auf Basis einer Stichprobe geschätzt werden.
- Zum Beispiel: σ von $N(10; \sigma)$
- Schätzwert: $\hat{\vartheta}$
- Vorgehen: Verwendung einer Schätzfunktion

$$\hat{\Theta} = g(X_1, \ldots, X_n)$$

(großes) Theta

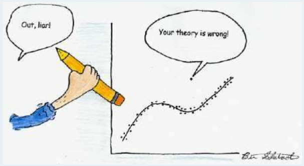
Beachte: Der Schätzwert $\hat{\vartheta}$ ist die Realisierung der ZV (!) $\hat{\Theta}$.

- Frage: Welche Stichprobenfunktion ist zur Schätzung geeignet?

➡ Kriterien für die Beurteilung/Konstruktion von Schätzfunktionen!

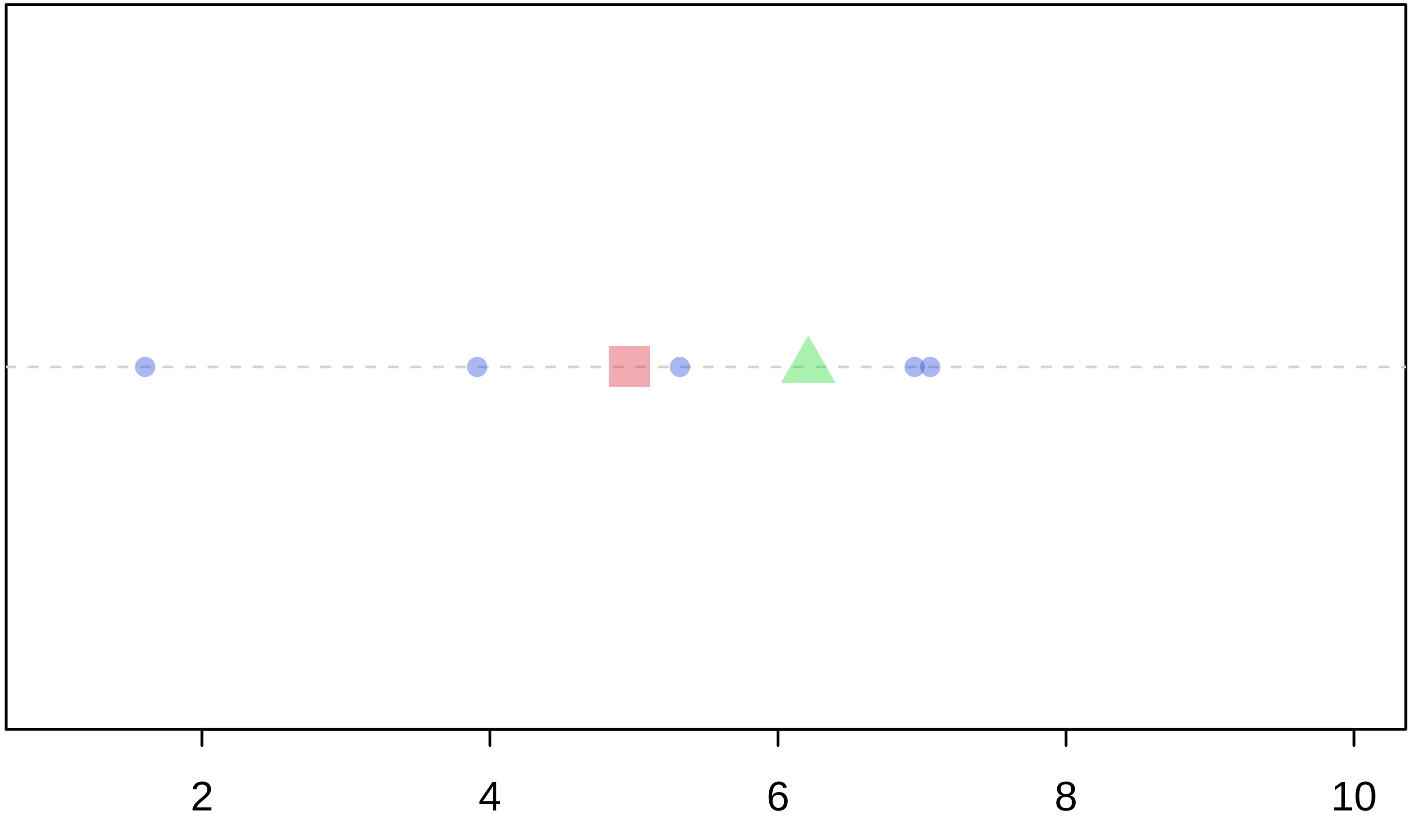
- Im Folgenden: Vorliegen einer einfachen Stichprobe, d.h. $X_1, \ldots, X_n$ iid.

Beispiel

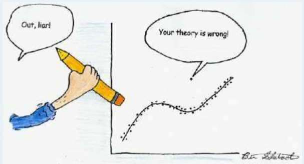


- Schätzen des Mittelwertes einer Grundgesamtheit
- dazu: Einfache Stichprobe vom Umfang 5
- und den beiden Stichprobenfunktionen

$$\boxed{\hat{\Theta}_1} = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad \hat{\Theta}_2 = \frac{1}{n-1}\sum_{i=1}^{n} X_i$$

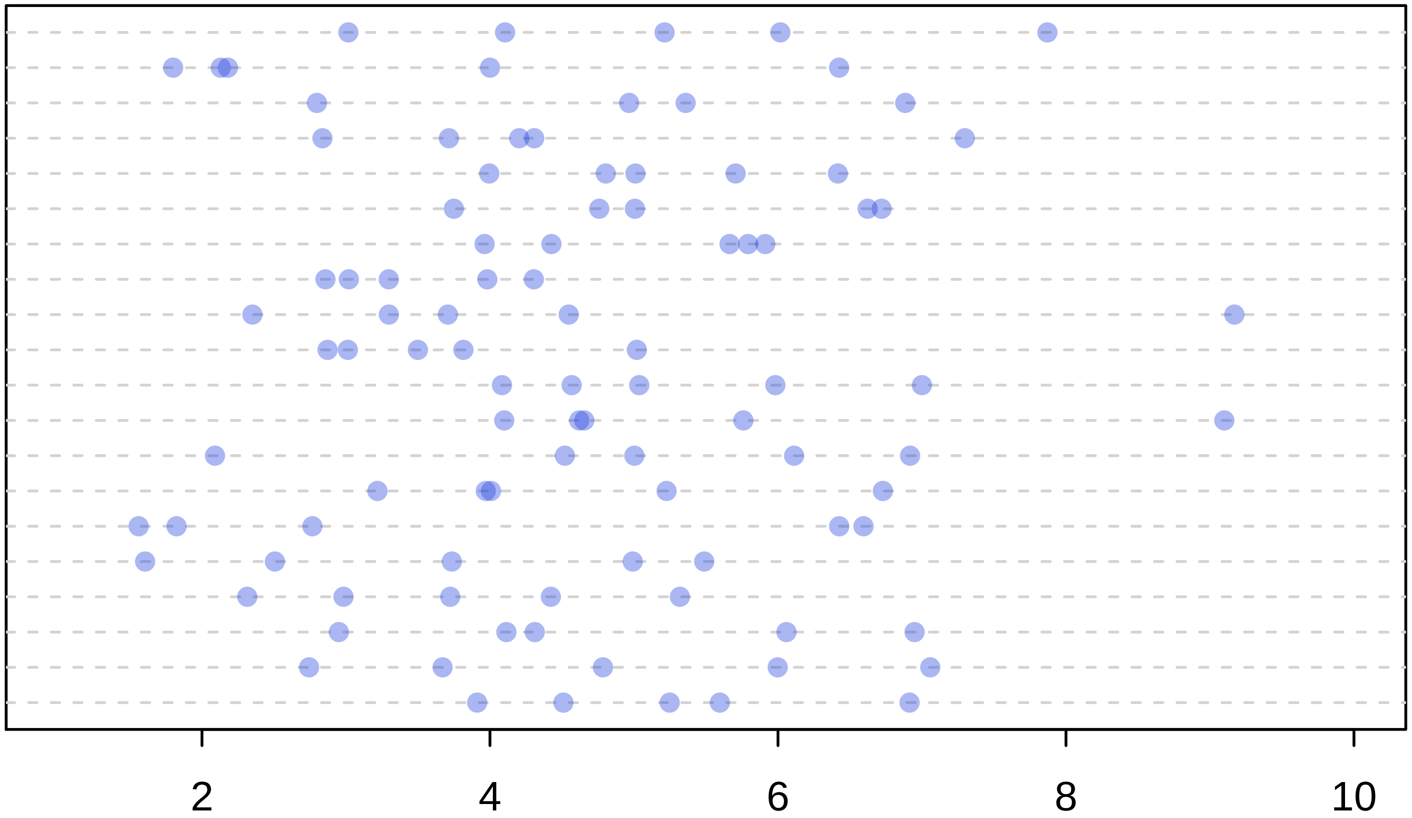


Beispiel

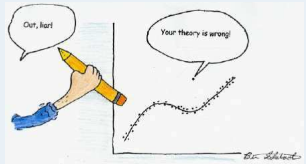


- Schätzen des Mittelwertes einer Grundgesamtheit
- dazu: Einfache Stichprobe vom Umfang 5
- und den beiden Stichprobenfunktionen

$$\hat{\Theta}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad \hat{\Theta}_2 = \frac{1}{n-1}\sum_{i=1}^{n} X_i$$

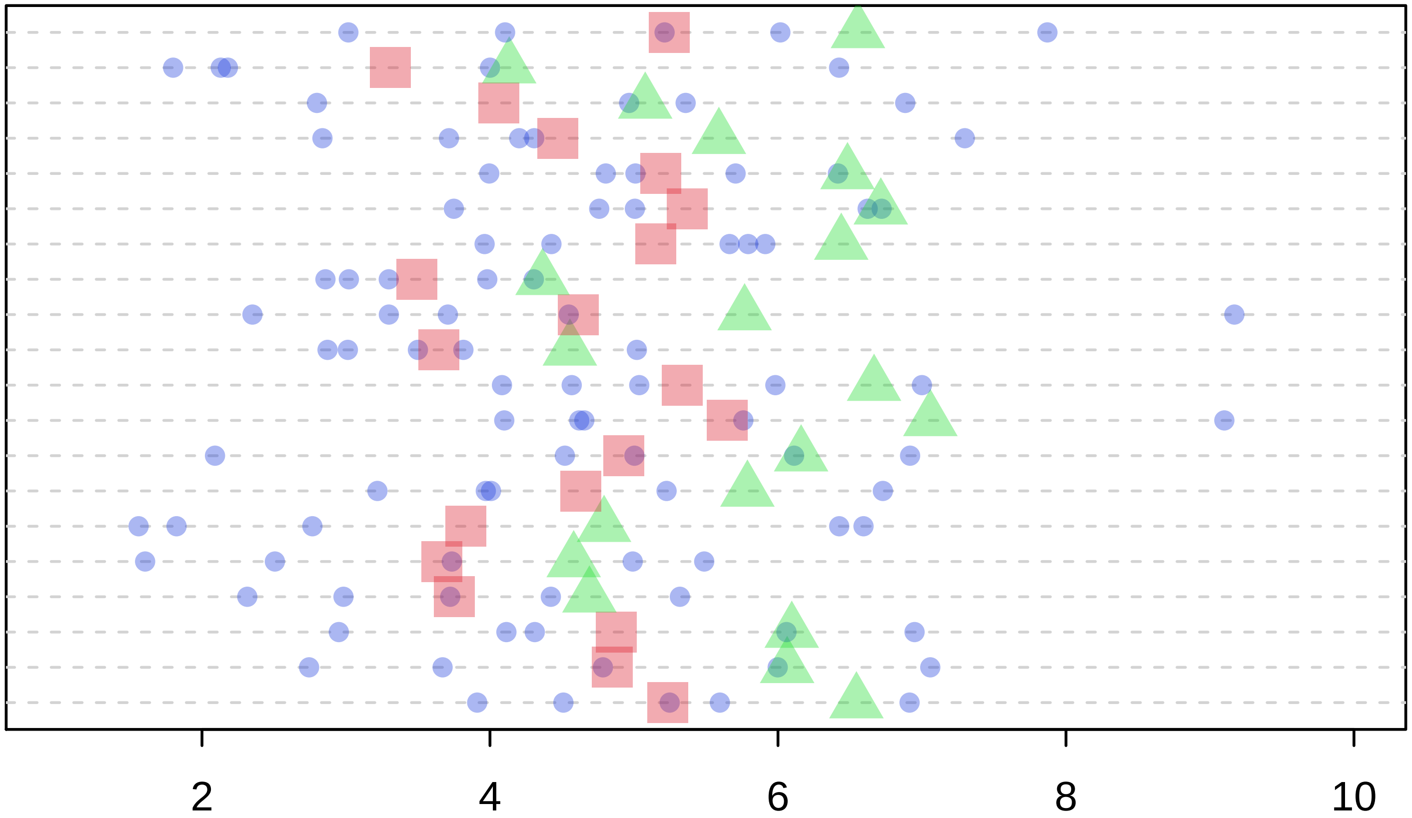


Beispiel

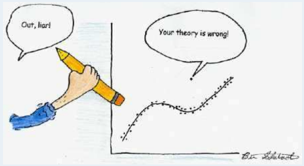


- Schätzen des Mittelwertes einer Grundgesamtheit
- dazu: Einfache Stichprobe vom Umfang 5
- und den beiden Stichprobenfunktionen

$$\hat{\Theta}_1 = \frac{1}{n} \sum_{i=1}^{n} X_i, \qquad \hat{\Theta}_2 = \frac{1}{n-1} \sum_{i=1}^{n} X_i$$

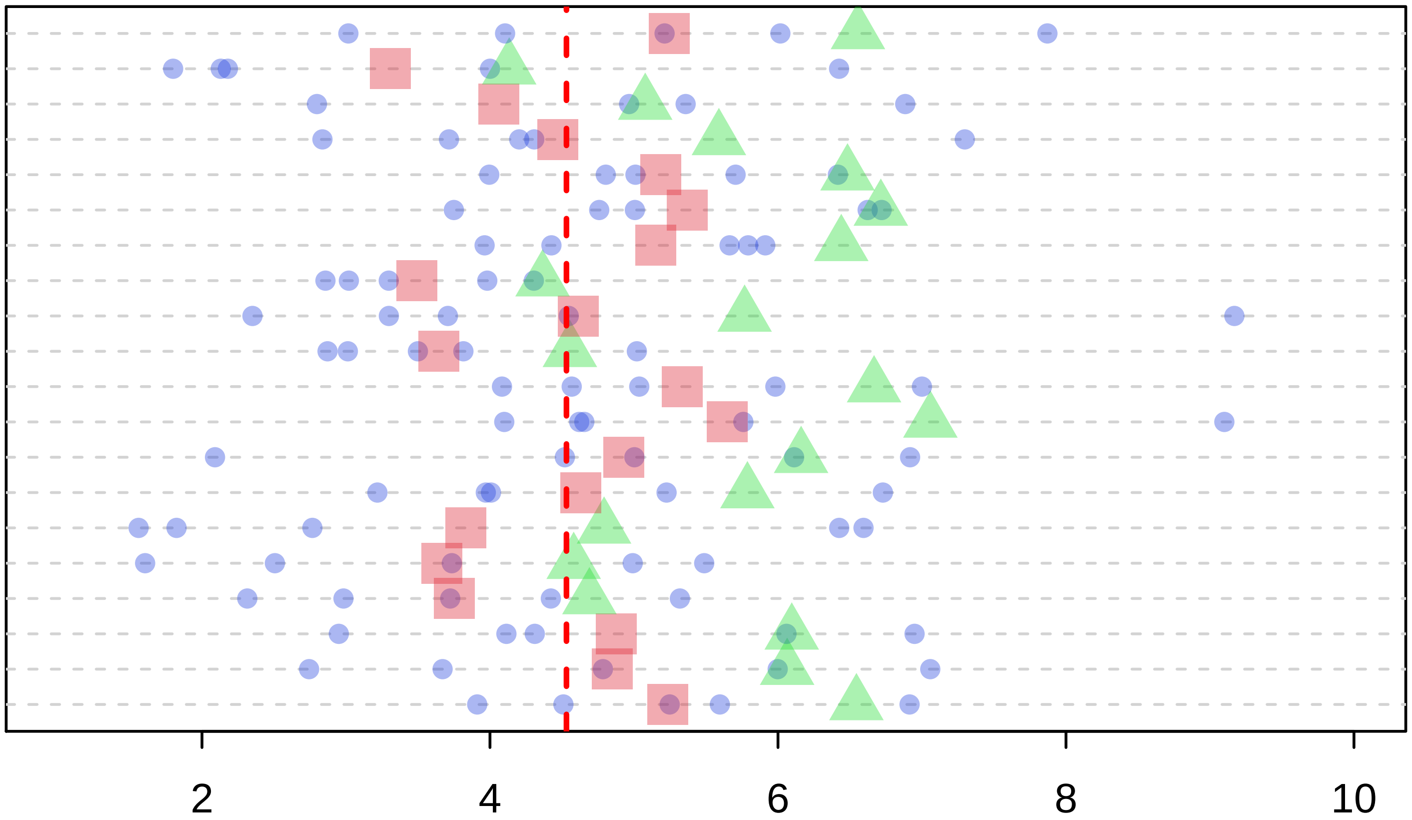


Beispiel

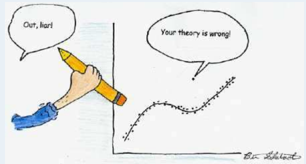


- Schätzen des Mittelwertes einer Grundgesamtheit
- dazu: Einfache Stichprobe vom Umfang 5
- und den beiden Stichprobenfunktionen

$$\hat{\Theta}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad \hat{\Theta}_2 = \frac{1}{n-1}\sum_{i=1}^{n} X_i$$



Mittelwert Grundgesamtheit = 4.53



- Eine Schätzfunktion $\hat{\Theta} = g(X_1, \ldots, X_n)$ heißt erwartungstreu oder unverzerrt für ϑ, wenn unabhängig vom numerischen Wert von ϑ gilt:

valide

$$E(\hat{\Theta}) = \vartheta$$

erwartungstreu?

Beispiel

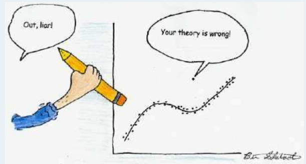
Sind $\hat{\Theta}_1 = \bar{X}$, $\hat{\Theta}_2 = \frac{X_1 + X_n}{2}$, $\hat{\Theta}_3 = \frac{1}{n-1} \sum\limits_{i=1}^{n} X_i$ erwartungstreu für μ?

a) $\hat{\Theta}_1$: $\quad E(\bar{X}) = \mu$
$\Rightarrow \hat{\Theta}_1$ ist erwartungstreu.

b) $\hat{\Theta}_2$: $\quad E\left(\frac{X_1 + X_n}{2}\right) = \frac{1}{2}[\underbrace{E(X_1)}_{\mu} + \underbrace{E(X_n)}_{\mu}] = \frac{1}{2}(\mu + \mu) = \mu$
$\Rightarrow \hat{\Theta}_2$ ist erwartungstreu.

c) $\hat{\Theta}_3$: $E\left(\frac{1}{n-1} \sum\limits_{i=1}^{n} X_i\right) = \frac{1}{n-1} \sum\limits_{i=1}^{n} E(X_i) = \frac{1}{n-1} \sum\limits_{i=1}^{n} \mu = \frac{n}{n-1} \mu \neq \mu$
$\Rightarrow \hat{\Theta}_3$ ist nicht erwartungstreu

Erwartungstreue und Wirksamkeit



- Welche der erwartungstreuen Schätzfunktionen $\widehat{\Theta}_1, \widehat{\Theta}_2$ ist „besser“?
- Von zwei erwartungstreuen Schätzfunktionen $\widehat{\Theta}_1$, $\widehat{\Theta}_2$ für ϑ heißt $\widehat{\Theta}_1$ wirksamer als $\widehat{\Theta}_2$, wenn unabhängig vom numerischen Wert von ϑ gilt:

$$\operatorname{Var}(\widehat{\Theta}_1) < \operatorname{Var}(\widehat{\Theta}_2)$$

Beispiel: ($\widehat{\Theta}_1 = \bar{X}$, $\widehat{\Theta}_2 = \frac{X_1 + X_n}{2}$)
Wegen

$$\left.\begin{aligned} \operatorname{Var}(\widehat{\Theta}_1) &= \operatorname{Var}(\bar{X}) && = \frac{\sigma^2}{n} \\ \operatorname{Var}(\widehat{\Theta}_2) &= \operatorname{Var}\left(\frac{X_1 + X_n}{2}\right) = \frac{1}{4}(\sigma^2 + \sigma^2) && = \frac{\sigma^2}{2} \end{aligned}\right\} \Rightarrow \operatorname{Var}(\widehat{\Theta}_1) < \operatorname{Var}(\widehat{\Theta}_2)$$

(falls $n > 2$) ist $\widehat{\Theta}_1$ wirksamer als $\widehat{\Theta}_2$.

Schätzfunktionen für die Varianz

$$\hat{\Theta}_1 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$\hat{\Theta}_2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

Welches $\hat{\Theta}$ ist erwartungstreu?

[schon bekannt: $Var[\bar{x}] = \frac{\sigma^2}{n}$]

$$E[\hat{\Theta}_1] = \frac{1}{n} \sum_{i=1}^{n} E\left[\left(x_i - \mu + \mu - \bar{x}\right)^2\right]$$

$$= \frac{1}{n} \sum E\left[(x_i-\mu)^2 + 2\cdot(x_i-\mu)(\mu-\bar{x}) + (\mu-\bar{x})^2\right]$$

$$= \frac{1}{n} E\left(\sum \underbrace{E(x_i-\mu)^2}_{=\sigma^2}\right) + 2\cdot(\mu-\bar{x})\cdot \underbrace{\sum(x_i-\mu)}_{n\cdot(\bar{x}-\mu)} + n(\mu-\bar{x})^2 \Big)$$

$$= \frac{1}{n} E\left(\sum (x_i-\mu)^2 - n(\bar{x}-\mu)^2\right)$$

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

$\Rightarrow$ $\hat{\Theta}_1$ ist nicht erwartungstreu

$$\hat{\Theta}_2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$$