

SS 2013

# Diskrete Wahrscheinlichkeitstheorie

Javier Esparza

Fakultät für Informatik  
TU München

<http://www7.in.tum.de/um/courses/dwt/ss13>

Sommersemester 2013

Teil V

Induktive Statistik

Das Ziel der **induktiven Statistik** besteht darin,

- aus **gemessenen** Zufallsgrößen (z.B. **Häufigkeiten**)
- auf die zugrunde liegenden Gesetzmäßigkeiten (z.B. **W'keiten**)

zu schließen.

## 20. Schätzvariablen

## Schätzvariablen I: Einführendes Beispiel

Sei  $X$  die Anzahl von Lesezugriffen auf eine Festplatte bis zum ersten Lesefehler.

# Schätzvariablen I: Einführendes Beispiel

Sei  $X$  die Anzahl von Lesezugriffen auf eine Festplatte bis zum ersten Lesefehler.

Wir nehmen  $X \sim \text{Geo}(p)$  an mit unbekannter Erfolgsw'keit  $p$ . Die W'keit  $p$  soll empirisch geschätzt werden.

# Schätzvariablen I: Einführendes Beispiel

Sei  $X$  die Anzahl von Lesezugriffen auf eine Festplatte bis zum ersten Lesefehler.

Wir nehmen  $X \sim \text{Geo}(p)$  an mit unbekannter Erfolgsw'keit  $p$ . Die W'keit  $p$  soll empirisch geschätzt werden.

Wegen  $p = \frac{1}{\mathbb{E}[X]}$  schätzen wir  $\mathbb{E}[X]$ .

# Schätzvariablen I: Einführendes Beispiel

Sei  $X$  die Anzahl von Lesezugriffen auf eine Festplatte bis zum ersten Lesefehler.

Wir nehmen  $X \sim \text{Geo}(p)$  an mit unbekannter Erfolgsw'keit  $p$ . Die W'keit  $p$  soll empirisch geschätzt werden.

Wegen  $p = \frac{1}{\mathbb{E}[X]}$  schätzen wir  $\mathbb{E}[X]$ .

Wir führen  $n$  Messungen der Anzahl von Zugriffen bis zum ersten Lesefehler. Sei  $X_i$  die Zufallsvariable für das Ergebnis der  $i$ -ten Messung.

$X_1, \dots, X_n$  sind unabhängig und besitzen jeweils dieselbe Verteilung wie  $X$ .

Die Variablen  $X_1, \dots, X_n$  nennt man **Stichprobenvariablen**.

# Schätzvariablen I: Einführendes Beispiel

Sei  $X$  die Anzahl von Lesezugriffen auf eine Festplatte bis zum ersten Lesefehler.

Wir nehmen  $X \sim \text{Geo}(p)$  an mit unbekannter Erfolgsw'keit  $p$ . Die W'keit  $p$  soll empirisch geschätzt werden.

Wegen  $p = \frac{1}{\mathbb{E}[X]}$  schätzen wir  $\mathbb{E}[X]$ .

Wir führen  $n$  Messungen der Anzahl von Zugriffen bis zum ersten Lesefehler. Sei  $X_i$  die Zufallsvariable für das Ergebnis der  $i$ -ten Messung.

$X_1, \dots, X_n$  sind unabhängig und besitzen jeweils dieselbe Verteilung wie  $X$ .

Die Variablen  $X_1, \dots, X_n$  nennt man **Stichprobenvariablen**.

## Schätzvariablen II: Einführendes Beispiel

Sei  $\bar{X}$  das arithmetische Mittel der  $X_i$ , d.h.  $\bar{X} := \frac{1}{n} \sum_{i=1}^n X_i$ .

## Schätzvariablen II: Einführendes Beispiel

Sei  $\bar{X}$  das arithmetische Mittel der  $X_i$ , d.h.  $\bar{X} := \frac{1}{n} \sum_{i=1}^n X_i$ .

Wir schätzen  $\mathbb{E}[X]$  durch  $\bar{X}$ .

## Schätzvariablen II: Einführendes Beispiel

Sei  $\bar{X}$  das arithmetische Mittel der  $X_i$ , d.h.  $\bar{X} := \frac{1}{n} \sum_{i=1}^n X_i$ .

Wir schätzen  $\mathbb{E}[X]$  durch  $\bar{X}$ .

Wir nennen  $\bar{X}$  einen Schätzer für den Erwartungswert  $\mathbb{E}[X]$ .

## Definition 150

Sei  $X$  eine Zufallsvariable mit Dichte  $f_X(x) = f(x; \theta)$ .

## Definition 150

Sei  $X$  eine Zufallsvariable mit Dichte  $f_X(x) = f(x; \theta)$ .

Eine **Schätzvariable** oder kurz **Schätzer** für  $\theta$  ist eine Zufallsvariable  $U$  der Gestalt  $U = f(x_1, \dots, X_n)$ , wobei die  $X_1, \dots, X_n$  unabhängige Stichprobenvariablen sind mit derselben Verteilung wie  $X$ .

## Definition 150

Sei  $X$  eine Zufallsvariable mit Dichte  $f_X(x) = f(x; \theta)$ .

Eine **Schätzvariable** oder kurz **Schätzer** für  $\theta$  ist eine Zufallsvariable  $U$  der Gestalt  $U = f(x_1, \dots, X_n)$ , wobei die  $X_1, \dots, X_n$  unabhängige Stichprobenvariablen sind mit derselben Verteilung wie  $X$ .

**Frage:** Welche Eigenschaften soll eine Schätzvariable erfüllen?

## Definition 151

Ein Schätzer  $U$  eines Parameters  $\theta$  heißt **erwartungstreu**, wenn  $\mathbb{E}[U] = \theta$  gilt.

Die Größe  $\mathbb{E}[U - \theta]$  nennt man **Bias** der Schätzvariablen  $U$ .

## Definition 151

Ein Schätzer  $U$  eines Parameters  $\theta$  heißt **erwartungstreu**, wenn  $\mathbb{E}[U] = \theta$  gilt.

Die Größe  $\mathbb{E}[U - \theta]$  nennt man **Bias** der Schätzvariablen  $U$ .

## Satz 152

Sei  $X$  eine beliebige Zufallsvariable.  $\bar{X}$  ist ein erwartungstreuer Schätzer für  $\mathbb{E}[X]$ .

Beweis:

$$\mathbb{E}[\bar{X}] = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i] = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X] = \mathbb{E}[X].$$



## Definition 153

Sei  $U$  ein Schätzer eines Parameters  $\theta$ . Die **mittlere quadratische Abweichung** von  $U$  ist

$$MSE := \mathbb{E}[(U - \theta)^2].$$

Ist  $U$  erwartungstreu, dann  $MSE = \mathbb{E}[(U - \mathbb{E}[U])^2] = \text{Var}[U]$ .

## Definition 153

Sei  $U$  ein Schätzer eines Parameters  $\theta$ . Die **mittlere quadratische Abweichung** von  $U$  ist

$$MSE := \mathbb{E}[(U - \theta)^2].$$

Ist  $U$  erwartungstreu, dann  $MSE = \mathbb{E}[(U - \mathbb{E}[U])^2] = \text{Var}[U]$ .

Eine Schätzvariable heißt **konsistent im quadratischen Mittel**, wenn  $MSE \rightarrow 0$  für  $n \rightarrow \infty$  gilt (mit  $n$  Umfang der Stichprobe).

## Satz 154

*Sei  $X$  eine Zufallsvariable mit endlicher Varianz. Der Schätzer  $\bar{X}$  für  $\mathbb{E}[X]$  ist konsistent im quadratischen Mittel.*

# Schätzvariablen V: Mittlere quadratische Abweichung

## Satz 154

Sei  $X$  eine Zufallsvariable mit endlicher Varianz. Der Schätzer  $\bar{X}$  für  $\mathbb{E}[X]$  ist konsistent im quadratischen Mittel.

Beweis:

Wegen der Unabhängigkeit von  $X_1, \dots, X_n$  gilt

$$MSE = \text{Var} \left[ \frac{1}{n} \sum_{i=1}^n X_i \right] = \frac{1}{n^2} \sum_{i=1}^n \text{Var}[X_i] = \frac{1}{n} \text{Var}[X].$$

und damit  $MSE \rightarrow 0$  für  $n \rightarrow \infty$ . □

# Schätzvariablen VI: Schwache Konsistenz

## Definition 155

Ein Schätzer  $U$  eines Parameters  $\theta$  ist **schwach konsistent** wenn

$$\lim_{n \rightarrow \infty} \Pr[|U - \theta| \geq \varepsilon] = 0$$

# Schätzvariablen VI: Schwache Konsistenz

## Definition 155

Ein Schätzer  $U$  eines Parameters  $\theta$  ist **schwach konsistent** wenn

$$\lim_{n \rightarrow \infty} \Pr[|U - \theta| \geq \varepsilon] = 0$$

## Satz 156

*Ist ein Schätzer konsistent im quadratischen Mittel, dann ist er auch schwach konsistent. Insbesondere ist  $\bar{X}$  ein schwach konsistenter Schätzer für  $X$ .*

## Definition 155

Ein Schätzer  $U$  eines Parameters  $\theta$  ist **schwach konsistent** wenn

$$\lim_{n \rightarrow \infty} \Pr[|U - \theta| \geq \varepsilon] = 0$$

## Satz 156

*Ist ein Schätzer konsistent im quadratischen Mittel, dann ist er auch schwach konsistent. Insbesondere ist  $\bar{X}$  ein schwach konsistenter Schätzer für  $X$ .*

### Beweis:

Mit der Ungleichung von Chebyshev gilt für ein beliebiges, festes  $\varepsilon > 0$ :

$$\Pr[|\bar{X} - \theta| \geq \varepsilon] = \Pr[|\bar{X} - \mathbb{E}[X]| \geq \varepsilon]$$



## Definition 155

Ein Schätzer  $U$  eines Parameters  $\theta$  ist **schwach konsistent** wenn

$$\lim_{n \rightarrow \infty} \Pr[|U - \theta| \geq \varepsilon] = 0$$

## Satz 156

*Ist ein Schätzer konsistent im quadratischen Mittel, dann ist er auch schwach konsistent. Insbesondere ist  $\bar{X}$  ein schwach konsistenter Schätzer für  $X$ .*

### Beweis:

Mit der Ungleichung von Chebyshev gilt für ein beliebiges, festes  $\varepsilon > 0$ :

$$\Pr[|\bar{X} - \theta| \geq \varepsilon] = \Pr[|\bar{X} - \mathbb{E}[X]| \geq \varepsilon] \leq \frac{\text{Var}[\bar{X}]}{\varepsilon^2}$$



## Definition 155

Ein Schätzer  $U$  eines Parameters  $\theta$  ist **schwach konsistent** wenn

$$\lim_{n \rightarrow \infty} \Pr[|U - \theta| \geq \varepsilon] = 0$$

## Satz 156

*Ist ein Schätzer konsistent im quadratischen Mittel, dann ist er auch schwach konsistent. Insbesondere ist  $\bar{X}$  ein schwach konsistenter Schätzer für  $X$ .*

### Beweis:

Mit der Ungleichung von Chebyshev gilt für ein beliebiges, festes  $\varepsilon > 0$ :

$$\Pr[|\bar{X} - \theta| \geq \varepsilon] = \Pr[|\bar{X} - \mathbb{E}[X]| \geq \varepsilon] \leq \frac{\text{Var}[\bar{X}]}{\varepsilon^2} \rightarrow 0 \text{ für } n \rightarrow \infty$$



## Schätzvariablen VII: Schätzung der Varianz

Das Ergebnis einer wiederholten Messung wird präsentiert als

$$\bar{X} \pm S$$

wobei  $S^2$  ein Schätzer für die Varianz von  $X$  darstellt.

## Schätzvariablen VII: Schätzung der Varianz

Das Ergebnis einer wiederholten Messung wird präsentiert als

$$\bar{X} \pm S$$

wobei  $S^2$  ein Schätzer für die Varianz von  $X$  darstellt.

Es liegt nah,

$$S_n^2 := \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

als Schätzer zu verwenden.

## Schätzvariablen VII: Schätzung der Varianz

Das Ergebnis einer wiederholten Messung wird präsentiert als

$$\bar{X} \pm S$$

wobei  $S^2$  ein Schätzer für die Varianz von  $X$  darstellt.

Es liegt nah,

$$S_n^2 := \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

als Schätzer zu verwenden.

$S_n^2$  ist jedoch **keinen** erwartungstreuen Schätzer für die Varianz!

## Satz 157

Sei  $X$  eine Zufallsvariable und sei

$$S_{n-1}^2 := \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

Die Variable  $S_{n-1}^2$  ist ein erwartungstreuer Schätzer für  $\text{Var}[X]$ .

## Schätzvariablen IX: Schätzung der Varianz

Beweis: Sei  $\mu := \mathbb{E}[X] = \mathbb{E}[X_i] = \mathbb{E}[\bar{X}]$ .

$$= \frac{n-2}{n}(X_i - \mu)^2 + (\mu - \bar{X})^2 - \frac{2}{n} \sum_{j \neq i} (X_i - \mu)(X_j - \mu).$$

Für je zwei unabhängige Zufallsvariablen  $X_i, X_j$  mit  $i \neq j$  gilt

$$\mathbb{E}[(X_i - \mu)(X_j - \mu)] = \mathbb{E}[X_i - \mu] \cdot \mathbb{E}[X_j - \mu] = 0 \cdot 0 = 0.$$

## Schätzvariablen IX: Schätzung der Varianz

Beweis: Sei  $\mu := \mathbb{E}[X] = \mathbb{E}[X_i] = \mathbb{E}[\bar{X}]$ .

$$(X_i - \bar{X})^2$$

Für je zwei unabhängige Zufallsvariablen  $X_i, X_j$  mit  $i \neq j$  gilt

$$\mathbb{E}[(X_i - \mu)(X_j - \mu)] = \mathbb{E}[X_i - \mu] \cdot \mathbb{E}[X_j - \mu] = 0 \cdot 0 = 0.$$

## Schätzvariablen IX: Schätzung der Varianz

Beweis: Sei  $\mu := \mathbb{E}[X] = \mathbb{E}[X_i] = \mathbb{E}[\bar{X}]$ .

$$(X_i - \bar{X})^2 = (X_i - \mu + \mu - \bar{X})^2$$

$$= \mathbb{E}[X_i - \mu] \cdot \mathbb{E}[X_j - \mu] = 0 \cdot 0 = 0.$$

## Schätzvariablen IX: Schätzung der Varianz

Beweis: Sei  $\mu := \mathbb{E}[X] = \mathbb{E}[X_i] = \mathbb{E}[\bar{X}]$ .

$$\begin{aligned}(X_i - \bar{X})^2 &= (X_i - \mu + \mu - \bar{X})^2 \\ &= (X_i - \mu)^2 + (\mu - \bar{X})^2 + 2(X_i - \mu)(\mu - \bar{X})\end{aligned}$$

$$= 0 \cdot 0 = 0.$$

## Schätzvariablen IX: Schätzung der Varianz

Beweis: Sei  $\mu := \mathbb{E}[X] = \mathbb{E}[X_i] = \mathbb{E}[\bar{X}]$ .

$$\begin{aligned}(X_i - \bar{X})^2 &= (X_i - \mu + \mu - \bar{X})^2 \\ &= (X_i - \mu)^2 + (\mu - \bar{X})^2 + 2(X_i - \mu)(\mu - \bar{X}) \\ &= (X_i - \mu)^2 + (\mu - \bar{X})^2 - \frac{2}{n} \sum_{j=1}^n (X_i - \mu)(X_j - \mu)\end{aligned}$$

## Schätzvariablen IX: Schätzung der Varianz

Beweis: Sei  $\mu := \mathbb{E}[X] = \mathbb{E}[X_i] = \mathbb{E}[\bar{X}]$ .

$$\begin{aligned}(X_i - \bar{X})^2 &= (X_i - \mu + \mu - \bar{X})^2 \\ &= (X_i - \mu)^2 + (\mu - \bar{X})^2 + 2(X_i - \mu)(\mu - \bar{X}) \\ &= (X_i - \mu)^2 + (\mu - \bar{X})^2 - \frac{2}{n} \sum_{j=1}^n (X_i - \mu)(X_j - \mu) \\ &= \frac{n-2}{n} (X_i - \mu)^2 + (\mu - \bar{X})^2 - \frac{2}{n} \sum_{j \neq i} (X_i - \mu)(X_j - \mu).\end{aligned}$$

## Schätzvariablen IX: Schätzung der Varianz

Beweis: Sei  $\mu := \mathbb{E}[X] = \mathbb{E}[X_i] = \mathbb{E}[\bar{X}]$ .

$$\begin{aligned}(X_i - \bar{X})^2 &= (X_i - \mu + \mu - \bar{X})^2 \\ &= (X_i - \mu)^2 + (\mu - \bar{X})^2 + 2(X_i - \mu)(\mu - \bar{X}) \\ &= (X_i - \mu)^2 + (\mu - \bar{X})^2 - \frac{2}{n} \sum_{j=1}^n (X_i - \mu)(X_j - \mu) \\ &= \frac{n-2}{n} (X_i - \mu)^2 + (\mu - \bar{X})^2 - \frac{2}{n} \sum_{j \neq i} (X_i - \mu)(X_j - \mu).\end{aligned}$$

Für je zwei unabhängige Zufallsvariablen  $X_i, X_j$  mit  $i \neq j$  gilt

$$\mathbb{E}[(X_i - \mu)(X_j - \mu)]$$

## Schätzvariablen IX: Schätzung der Varianz

Beweis: Sei  $\mu := \mathbb{E}[X] = \mathbb{E}[X_i] = \mathbb{E}[\bar{X}]$ .

$$\begin{aligned}(X_i - \bar{X})^2 &= (X_i - \mu + \mu - \bar{X})^2 \\ &= (X_i - \mu)^2 + (\mu - \bar{X})^2 + 2(X_i - \mu)(\mu - \bar{X}) \\ &= (X_i - \mu)^2 + (\mu - \bar{X})^2 - \frac{2}{n} \sum_{j=1}^n (X_i - \mu)(X_j - \mu) \\ &= \frac{n-2}{n} (X_i - \mu)^2 + (\mu - \bar{X})^2 - \frac{2}{n} \sum_{j \neq i} (X_i - \mu)(X_j - \mu).\end{aligned}$$

Für je zwei unabhängige Zufallsvariablen  $X_i, X_j$  mit  $i \neq j$  gilt

$$\mathbb{E}[(X_i - \mu)(X_j - \mu)] = \mathbb{E}[X_i - \mu] \cdot \mathbb{E}[X_j - \mu]$$

## Schätzvariablen IX: Schätzung der Varianz

Beweis: Sei  $\mu := \mathbb{E}[X] = \mathbb{E}[X_i] = \mathbb{E}[\bar{X}]$ .

$$\begin{aligned}(X_i - \bar{X})^2 &= (X_i - \mu + \mu - \bar{X})^2 \\ &= (X_i - \mu)^2 + (\mu - \bar{X})^2 + 2(X_i - \mu)(\mu - \bar{X}) \\ &= (X_i - \mu)^2 + (\mu - \bar{X})^2 - \frac{2}{n} \sum_{j=1}^n (X_i - \mu)(X_j - \mu) \\ &= \frac{n-2}{n} (X_i - \mu)^2 + (\mu - \bar{X})^2 - \frac{2}{n} \sum_{j \neq i} (X_i - \mu)(X_j - \mu).\end{aligned}$$

Für je zwei unabhängige Zufallsvariablen  $X_i, X_j$  mit  $i \neq j$  gilt

$$\mathbb{E}[(X_i - \mu)(X_j - \mu)] = \mathbb{E}[X_i - \mu] \cdot \mathbb{E}[X_j - \mu] = 0 \cdot 0 = 0.$$

# Schätzvariablen X: Schätzung der Varianz

Daraus folgt

$$\mathbb{E}[(X_i - \bar{X})^2] = \frac{n-2}{n} \cdot \mathbb{E}[(X_i - \mu)^2] + \mathbb{E}[(\mu - \bar{X})^2]$$

$$\begin{aligned}\mathbb{E}[S_{n-1}^2] &= \frac{1}{n-1} \sum_{i=1}^n \mathbb{E}[(X_i - \bar{X})^2] \\ &= \frac{1}{n-1} \cdot n \cdot \frac{n-2}{n} \cdot \text{Var}[X] = \text{Var}[X].\end{aligned}$$

# Schätzvariablen X: Schätzung der Varianz

Daraus folgt

$$\begin{aligned}\mathbb{E}[(X_i - \bar{X})^2] &= \frac{n-2}{n} \cdot \mathbb{E}[(X_i - \mu)^2] + \mathbb{E}[(\mu - \bar{X})^2] \\ &= \frac{n-2}{n} \cdot \text{Var}[X_i] + \text{Var}[\bar{X}]\end{aligned}$$

$$\begin{aligned}&= \frac{1}{n-1} \sum_{i=1}^n \mathbb{E}[(X_i - \bar{X})^2] \\ &= \frac{1}{n-1} \cdot n \cdot \frac{n-1}{n} \cdot \text{Var}[X] = \text{Var}[X].\end{aligned}$$

# Schätzvariablen X: Schätzung der Varianz

Daraus folgt

$$\begin{aligned}\mathbb{E}[(X_i - \bar{X})^2] &= \frac{n-2}{n} \cdot \mathbb{E}[(X_i - \mu)^2] + \mathbb{E}[(\mu - \bar{X})^2] \\ &= \frac{n-2}{n} \cdot \text{Var}[X_i] + \text{Var}[\bar{X}] \\ &= \frac{n-2}{n} \cdot \frac{1}{n} \cdot \text{Var}[X] + \text{Var}[\bar{X}]\end{aligned}$$

$$= \frac{1}{n-1} \cdot n \cdot \frac{n-1}{n} \cdot \text{Var}[X] = \text{Var}[X].$$

# Schätzvariablen X: Schätzung der Varianz

Daraus folgt

$$\begin{aligned}\mathbb{E}[(X_i - \bar{X})^2] &= \frac{n-2}{n} \cdot \mathbb{E}[(X_i - \mu)^2] + \mathbb{E}[(\mu - \bar{X})^2] \\ &= \frac{n-2}{n} \cdot \text{Var}[X_i] + \text{Var}[\bar{X}] \\ &= \frac{n-2}{n} \cdot \frac{1}{n} \cdot \text{Var}[X] + \text{Var}[\bar{X}] = \frac{n-1}{n} \cdot \text{Var}[X] \\ &= \text{Var}[X].\end{aligned}$$

# Schätzvariablen X: Schätzung der Varianz

Daraus folgt

$$\begin{aligned}\mathbb{E}[(X_i - \bar{X})^2] &= \frac{n-2}{n} \cdot \mathbb{E}[(X_i - \mu)^2] + \mathbb{E}[(\mu - \bar{X})^2] \\ &= \frac{n-2}{n} \cdot \text{Var}[X_i] + \text{Var}[\bar{X}] \\ &= \frac{n-2}{n} \cdot \frac{1}{n} \cdot \text{Var}[X] + \text{Var}[\bar{X}] = \frac{n-1}{n} \cdot \text{Var}[X]\end{aligned}$$

und somit gilt für  $S_{n-1}^2$

# Schätzvariablen X: Schätzung der Varianz

Daraus folgt

$$\begin{aligned}\mathbb{E}[(X_i - \bar{X})^2] &= \frac{n-2}{n} \cdot \mathbb{E}[(X_i - \mu)^2] + \mathbb{E}[(\mu - \bar{X})^2] \\ &= \frac{n-2}{n} \cdot \text{Var}[X_i] + \text{Var}[\bar{X}] \\ &= \frac{n-2}{n} \cdot \frac{1}{n} \cdot \text{Var}[X] + \text{Var}[\bar{X}] = \frac{n-1}{n} \cdot \text{Var}[X]\end{aligned}$$

und somit gilt für  $S_{n-1}^2$

$$\mathbb{E}[S_{n-1}^2]$$

# Schätzvariablen X: Schätzung der Varianz

Daraus folgt

$$\begin{aligned}\mathbb{E}[(X_i - \bar{X})^2] &= \frac{n-2}{n} \cdot \mathbb{E}[(X_i - \mu)^2] + \mathbb{E}[(\mu - \bar{X})^2] \\ &= \frac{n-2}{n} \cdot \text{Var}[X_i] + \text{Var}[\bar{X}] \\ &= \frac{n-2}{n} \cdot \frac{1}{n} \cdot \text{Var}[X] + \text{Var}[\bar{X}] = \frac{n-1}{n} \cdot \text{Var}[X]\end{aligned}$$

und somit gilt für  $S_{n-1}^2$

$$\mathbb{E}[S_{n-1}^2] = \frac{1}{n-1} \sum_{i=1}^n \mathbb{E}[(X_i - \bar{X})^2]$$

# Schätzvariablen X: Schätzung der Varianz

Daraus folgt

$$\begin{aligned}\mathbb{E}[(X_i - \bar{X})^2] &= \frac{n-2}{n} \cdot \mathbb{E}[(X_i - \mu)^2] + \mathbb{E}[(\mu - \bar{X})^2] \\ &= \frac{n-2}{n} \cdot \text{Var}[X_i] + \text{Var}[\bar{X}] \\ &= \frac{n-2}{n} \cdot \frac{1}{n} \cdot \text{Var}[X] + \text{Var}[\bar{X}] = \frac{n-1}{n} \cdot \text{Var}[X]\end{aligned}$$

und somit gilt für  $S_{n-1}^2$

$$\begin{aligned}\mathbb{E}[S_{n-1}^2] &= \frac{1}{n-1} \sum_{i=1}^n \mathbb{E}[(X_i - \bar{X})^2] \\ &= \frac{1}{n-1} \cdot n \cdot \frac{n-1}{n} \cdot \text{Var}[X]\end{aligned}$$

# Schätzvariablen X: Schätzung der Varianz

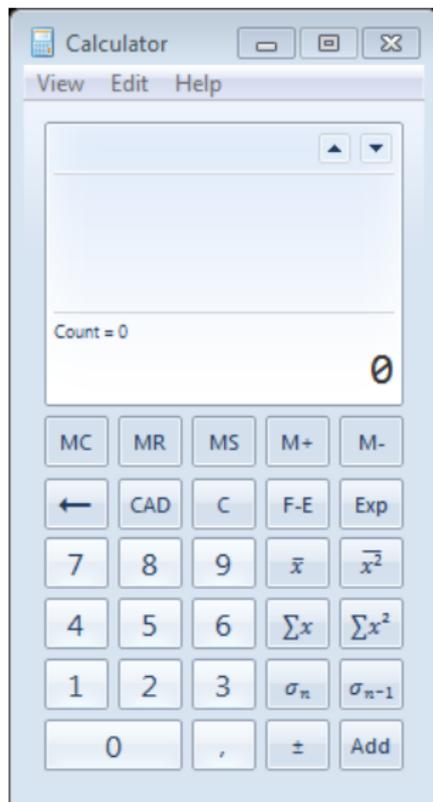
Daraus folgt

$$\begin{aligned}\mathbb{E}[(X_i - \bar{X})^2] &= \frac{n-2}{n} \cdot \mathbb{E}[(X_i - \mu)^2] + \mathbb{E}[(\mu - \bar{X})^2] \\ &= \frac{n-2}{n} \cdot \text{Var}[X_i] + \text{Var}[\bar{X}] \\ &= \frac{n-2}{n} \cdot \frac{1}{n} \cdot \text{Var}[X] + \text{Var}[\bar{X}] = \frac{n-1}{n} \cdot \text{Var}[X]\end{aligned}$$

und somit gilt für  $S_{n-1}^2$

$$\begin{aligned}\mathbb{E}[S_{n-1}^2] &= \frac{1}{n-1} \sum_{i=1}^n \mathbb{E}[(X_i - \bar{X})^2] \\ &= \frac{1}{n-1} \cdot n \cdot \frac{n-1}{n} \cdot \text{Var}[X] = \text{Var}[X].\end{aligned}$$

# Schätzvariablen XI: Schätzung der Varianz



## Definition 158

Die Zufallsvariablen

$$\bar{X} := \frac{1}{n} \sum_{i=1}^n X_i \quad \text{und} \quad S_{n-1}^2 := \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

heißen **Stichprobenmittel** bzw. **Stichprobenvarianz** der Stichprobe  $X_1, \dots, X_n$ .

$\bar{X}$  und  $S_{n-1}^2$  sind erwartungstreue Schätzer für den Erwartungswert bzw. die Varianz von  $X$ .

Oft wird  $S^2$  statt  $S_{n-1}^2$  geschrieben.

## Beispiel 159

Ein Program läuft 10 Mal mit demselben Input mit Laufzeiten (ms)

2582	2580	2568	2576	2582
2581	2623	2616	2622	2617

## Beispiel 159

Ein Program läuft 10 Mal mit demselben Input mit Laufzeiten (ms)

2582	2580	2568	2576	2582
2581	2623	2616	2622	2617

Damit gilt für (die Realisierungen von)  $\bar{X}$  und  $S_{n-1}$

$$\bar{X} \approx 2594,7 \quad S_{n-1} \approx 21,8$$

und wir schreiben als Ergebnis:  $2594.7 \pm 21.8$ .

## Beispiel 159

Ein Program läuft 10 Mal mit demselben Input mit Laufzeiten (ms)

2582	2580	2568	2576	2582
2581	2623	2616	2622	2617

Damit gilt für (die Realisierungen von)  $\bar{X}$  und  $S_{n-1}$

$$\bar{X} \approx 2594,7 \quad S_{n-1} \approx 21,8$$

und wir schreiben als Ergebnis:  $2594.7 \pm 21.8$ .

Unter den Annahmen der Normalverteilung beträgt die W'keit, dass die Laufzeit einer Ausführung ausserhalb des Intervalls

$$[2594.7 - 2 \cdot 21.8, 2594.7 + 2 \cdot 21.8] = [2551.1, 2638.3]$$

liegt, höchstens 5% .

## **21. Das Maximum-Likelihood-Prinzip**

# Maximum-Likelihood-Prinzip I

Das Maximum-Likelihood-Prinzip ist ein allgemeines Prinzip zur Bestimmung des Parameters einer  $W$ 'eitsverteilung.

*Wähle den Wert des Parameters, für den die beobachtete Stichprobe maximale  $W$ 'keit hat.*

# Maximum-Likelihood-Prinzip I

Das Maximum-Likelihood-Prinzip ist ein allgemeines Prinzip zur Bestimmung des Parameters einer  $W$ 'eitsverteilung.

*Wähle den Wert des Parameters, für den die beobachtete Stichprobe maximale  $W$ 'keit hat.*

## Beispiel 160

Folgendes Experiment wird 10 mal wiederholt: Eine Münze wird bis zum ersten „Kopf“ geworfen.

Die Münze wird jeweils 4, 4, 7, 2, 4, 6, 5, 3, 5, 2 Mal geworfen.

**Frage:** Welche ist die Erfolgsw'keit der Münze ?

# Maximum-Likelihood-Prinzip I

Das Maximum-Likelihood-Prinzip ist ein allgemeines Prinzip zur Bestimmung des Parameters einer  $W$ 'eitsverteilung.

*Wähle den Wert des Parameters, für den die beobachtete Stichprobe maximale  $W$ 'keit hat.*

## Beispiel 160

Folgendes Experiment wird 10 mal wiederholt: Eine Münze wird bis zum ersten „Kopf“ geworfen.

Die Münze wird jeweils 4, 4, 7, 2, 4, 6, 5, 3, 5, 2 Mal geworfen.

**Frage:** Welche ist die Erfolgsw'keit der Münze ?

Sei  $X$  die Anzahl der Würfe. Es gilt  $X \sim Geo(p)$ .

Nach dem Prinzip sollen wir  $p$  so wählen, das die  $W$ 'keit einer Beobachtung 4, 4, 7, 2, 4, 6, 5, 3, 5, 2 maximiert wird.

## Maximum-Likelihood-Prinzip II: Formalisierung

Sei  $X$  eine Zufallsvariable mit  $\Pr[X = x] = f(x; \theta)$  für eine **bekante** Funktion  $f$  mit einem **unbekannten** Parameter  $\theta$ .

## Maximum-Likelihood-Prinzip II: Formalisierung

Sei  $X$  eine Zufallsvariable mit  $\Pr[X = x] = f(x; \theta)$  für eine **bekante** Funktion  $f$  mit einem **unbekannten** Parameter  $\theta$ .

Sei  $X_1, \dots, X_n$  Stichprobenvariablen mit Dichte  $f(x; \theta)$ .

# Maximum-Likelihood-Prinzip II: Formalisierung

Sei  $X$  eine Zufallsvariable mit  $\Pr[X = x] = f(x; \theta)$  für eine **bekannte** Funktion  $f$  mit einem **unbekannten** Parameter  $\theta$ .

Sei  $X_1, \dots, X_n$  Stichprobenvariablen mit Dichte  $f(x; \theta)$ .

Eine Stichprobe liefert für jede  $X_i$  einen Wert  $x_i$  und wir schreiben  $\vec{x} = (x_1, \dots, x_n)$ .

Für eine **festen** Stichprobe  $\vec{x}$  gibt die **Likelihood-Funktion**

$$L(\vec{x}; \theta) := \prod_{i=1}^n f(x_i; \theta) = \prod_{i=1}^n \Pr_{\theta}[X_i = x_i]$$
$$\stackrel{\text{unabh.}}{=} \Pr_{\theta}[X_1 = x_1, \dots, X_n = x_n]$$

die Wahrscheinlichkeit an, dass man die Stichprobe  $\vec{x}$  erhält, wenn der Parameter den Wert  $\theta$  hat.

## Definition 161

Ein Schätzwert  $\hat{\theta}$  für den Parameter einer Verteilung  $f(x; \theta)$  heißt **Maximum-Likelihood-Schätzwert (ML-Schätzwert)** für eine Stichprobe  $\vec{x}$ , wenn gilt

$$L(\vec{x}; \theta) \leq L(\vec{x}; \hat{\theta}) \text{ für alle } \theta.$$

## Definition 161

Ein Schätzwert  $\hat{\theta}$  für den Parameter einer Verteilung  $f(x; \theta)$  heißt **Maximum-Likelihood-Schätzwert (ML-Schätzwert)** für eine Stichprobe  $\vec{x}$ , wenn gilt

$$L(\vec{x}; \theta) \leq L(\vec{x}; \hat{\theta}) \text{ für alle } \theta.$$

Wenn  $L(\vec{x}; \theta)$  differenzierbar ist, dann kann ein Maximum von  $L(\vec{x}; \theta)$  mit Hilfe der üblichen Methode berechnet werden:

- Berechne  $L'(\vec{x}; \theta) := \frac{dL(\vec{x}; \theta)}{d\theta}$ .
- Finde eine Lösung  $\theta_0$  der Gleichung  $L'(\vec{x}; \theta) = 0$  mit  $L''(\vec{x}; \theta_0) > 0$ .

## Beispiel 162

Wir konstruieren mit der ML-Methode einen Schätzer für den Parameter  $p$  der Bernoulli-Verteilung.

## Beispiel 162

Wir konstruieren mit der ML-Methode einen Schätzer für den Parameter  $p$  der Bernoulli-Verteilung.

Mit  $\Pr_p[X_i = 1] = p$  und  $\Pr_p[X_i = 0] = 1 - p$  gilt

$$\Pr_p[X_i = x_i] = p^{x_i}(1 - p)^{1-x_i} .$$

## Beispiel 162

Wir konstruieren mit der ML-Methode einen Schätzer für den Parameter  $p$  der Bernoulli-Verteilung.

Mit  $\Pr_p[X_i = 1] = p$  und  $\Pr_p[X_i = 0] = 1 - p$  gilt

$$\Pr_p[X_i = x_i] = p^{x_i}(1 - p)^{1-x_i} .$$

Die Likelihood-Funktion ist

## Beispiel 162

Wir konstruieren mit der ML-Methode einen Schätzer für den Parameter  $p$  der Bernoulli-Verteilung.

Mit  $\Pr_p[X_i = 1] = p$  und  $\Pr_p[X_i = 0] = 1 - p$  gilt

$$\Pr_p[X_i = x_i] = p^{x_i}(1 - p)^{1-x_i} .$$

Die Likelihood-Funktion ist

$$L(\vec{x}; p) = \prod_{i=1}^n p^{x_i} \cdot (1 - p)^{1-x_i}$$

# Maximum-Likelihood-Prinzip VI: Beispiele

Statt  $L$  maximieren wir  $\ln L$  (einfachere Berechnung).

$$\begin{aligned}\ln L(\vec{x}; p) &= \sum_{i=1}^n (x_i \cdot \ln p + (1 - x_i) \cdot \ln(1 - p)) \\ &= n\bar{x} \cdot \ln p + (n - n\bar{x}) \cdot \ln(1 - p).\end{aligned}$$

mit  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ .

# Maximum-Likelihood-Prinzip VI: Beispiele

Statt  $L$  maximieren wir  $\ln L$  (einfachere Berechnung).

$$\begin{aligned}\ln L(\vec{x}; p) &= \sum_{i=1}^n (x_i \cdot \ln p + (1 - x_i) \cdot \ln(1 - p)) \\ &= n\bar{x} \cdot \ln p + (n - n\bar{x}) \cdot \ln(1 - p).\end{aligned}$$

$$\text{mit } \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i.$$

Nullsetzen der Ableitung ergibt:

$$\frac{d \ln L(\vec{x}; p)}{d p} = \frac{n\bar{x}}{p} - \frac{n - n\bar{x}}{1 - p} = 0.$$

mit Lösung  $p = \bar{x}$ , ein Maximum.

## Maximum-Likelihood-Prinzip VII: Beispiele

Sei  $X \sim \mathcal{N}(\mu, \sigma^2)$ . Wir bestimmen Schätzer für  $\mu$  und  $\sigma^2$ .

## Maximum-Likelihood-Prinzip VII: Beispiele

Sei  $X \sim \mathcal{N}(\mu, \sigma^2)$ . Wir bestimmen Schätzer für  $\mu$  und  $\sigma^2$ . Es gilt

$$L(\vec{x}; \mu, \sigma^2) = \left( \frac{1}{\sqrt{2\pi}\sigma} \right)^n \cdot \prod_{i=1}^n \exp \left( -\frac{(x_i - \mu)^2}{2\sigma^2} \right).$$

# Maximum-Likelihood-Prinzip VII: Beispiele

Sei  $X \sim \mathcal{N}(\mu, \sigma^2)$ . Wir bestimmen Schätzer für  $\mu$  und  $\sigma^2$ . Es gilt

$$L(\vec{x}; \mu, \sigma^2) = \left( \frac{1}{\sqrt{2\pi}\sigma} \right)^n \cdot \prod_{i=1}^n \exp \left( -\frac{(x_i - \mu)^2}{2\sigma^2} \right).$$

Durch Logarithmieren erhalten wir

$$\ln L(\vec{x}; \mu, \sigma^2) = -n(\ln \sqrt{2\pi} + \ln \sigma) + \sum_{i=1}^n \left( -\frac{(x_i - \mu)^2}{2\sigma^2} \right).$$

# Maximum-Likelihood-Prinzip VII: Beispiele

Nullsetzen der Ableitungen ergibt

$$\frac{\partial \ln L}{\partial \mu} = \sum_{i=1}^n \frac{x_i - \mu}{\sigma^2} \stackrel{!}{=} 0$$

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \sum_{i=1}^n \frac{(x_i - \mu)^2}{\sigma^3} \stackrel{!}{=} 0$$

# Maximum-Likelihood-Prinzip VII: Beispiele

Nullsetzen der Ableitungen ergibt

$$\frac{\partial \ln L}{\partial \mu} = \sum_{i=1}^n \frac{x_i - \mu}{\sigma^2} \stackrel{!}{=} 0$$

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \sum_{i=1}^n \frac{(x_i - \mu)^2}{\sigma^3} \stackrel{!}{=} 0$$

und wir erhalten

$$\mu = \bar{x} \quad \text{und} \quad \sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2.$$

# Maximum-Likelihood-Prinzip VII: Beispiele

Nullsetzen der Ableitungen ergibt

$$\frac{\partial \ln L}{\partial \mu} = \sum_{i=1}^n \frac{x_i - \mu}{\sigma^2} \stackrel{!}{=} 0$$

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \sum_{i=1}^n \frac{(x_i - \mu)^2}{\sigma^3} \stackrel{!}{=} 0$$

und wir erhalten

$$\mu = \bar{x} \quad \text{und} \quad \sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2.$$

# Maximum-Likelihood-Prinzip VII: Beispiele

Nullsetzen der Ableitungen ergibt

$$\frac{\partial \ln L}{\partial \mu} = \sum_{i=1}^n \frac{x_i - \mu}{\sigma^2} \stackrel{!}{=} 0$$

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \sum_{i=1}^n \frac{(x_i - \mu)^2}{\sigma^3} \stackrel{!}{=} 0$$

und wir erhalten

$$\mu = \bar{x} \quad \text{und} \quad \sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2.$$

# Maximum-Likelihood-Prinzip VII: Beispiele

Nullsetzen der Ableitungen ergibt

$$\frac{\partial \ln L}{\partial \mu} = \sum_{i=1}^n \frac{x_i - \mu}{\sigma^2} \stackrel{!}{=} 0$$

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{n}{\sigma} + \sum_{i=1}^n \frac{(x_i - \mu)^2}{\sigma^3} \stackrel{!}{=} 0$$

und wir erhalten

$$\mu = \bar{x} \quad \text{und} \quad \sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2.$$

Der zweite Wert ist **fast** die Stichprobenvarianz  $S_{n-1}^2$  aber mit Vorfaktor  $\frac{1}{n}$  statt  $\frac{1}{n-1}$ .

Die ML-Schätzvariable für die Varianz ist somit **nicht** erwartungstreu.

## **22. Konfidenzintervalle**

**Problem:** Wie gut kann man sich auf einen Schätzwert verlassen?

# Konfidenzintervalle I

**Problem:** Wie gut kann man sich auf einen Schätzwert verlassen?

**Lösung:** Berechne statt einem Schätzer  $U$  zwei Schätzer  $U_1$  und  $U_2$  mit

$$\Pr[U_1 \leq \theta \leq U_2] \geq 1 - \alpha.$$

Die W'keit  $1 - \alpha$  heißt **Konfidenzniveau** und kann dem „Sicherheitsbedürfnis“ angepasst werden.

# Konfidenzintervalle I

**Problem:** Wie gut kann man sich auf einen Schätzwert verlassen?

**Lösung:** Berechne statt einem Schätzer  $U$  zwei Schätzer  $U_1$  und  $U_2$  mit

$$\Pr[U_1 \leq \theta \leq U_2] \geq 1 - \alpha.$$

Die W'keit  $1 - \alpha$  heißt **Konfidenzniveau** und kann dem „Sicherheitsbedürfnis“ angepasst werden.

Informelle Bedeutung:

*Wenn wir für eine Stichprobe die Schätzer  $U_1$  und  $U_2$  berechnen und  $\theta \in [U_1, U_2]$  schliessen, dann irren wir höchstens mit W'keit  $\alpha$ .*

# Konfidenzintervalle I

**Problem:** Wie gut kann man sich auf einen Schätzwert verlassen?

**Lösung:** Berechne statt einem Schätzer  $U$  zwei Schätzer  $U_1$  und  $U_2$  mit

$$\Pr[U_1 \leq \theta \leq U_2] \geq 1 - \alpha.$$

Die W'keit  $1 - \alpha$  heißt **Konfidenzniveau** und kann dem „Sicherheitsbedürfnis“ angepasst werden.

Informelle Bedeutung:

*Wenn wir für eine Stichprobe die Schätzer  $U_1$  und  $U_2$  berechnen und  $\theta \in [U_1, U_2]$  schliessen, dann irren wir höchstens mit W'keit  $\alpha$ .*

$[U_1, U_2]$  heißt **Konfidenzintervall**.

# Konfidenzintervalle I

**Problem:** Wie gut kann man sich auf einen Schätzwert verlassen?

**Lösung:** Berechne statt einem Schätzer  $U$  zwei Schätzer  $U_1$  und  $U_2$  mit

$$\Pr[U_1 \leq \theta \leq U_2] \geq 1 - \alpha.$$

Die W'keit  $1 - \alpha$  heißt **Konfidenzniveau** und kann dem „Sicherheitsbedürfnis“ angepasst werden.

Informelle Bedeutung:

*Wenn wir für eine Stichprobe die Schätzer  $U_1$  und  $U_2$  berechnen und  $\theta \in [U_1, U_2]$  schliessen, dann irren wir höchstens mit W'keit  $\alpha$ .*

$[U_1, U_2]$  heißt **Konfidenzintervall**.

Oft setzt man  $U_1 := U - \delta$  und  $U_2 := U + \delta$  für einen Schätzer  $U$  (**symmetrisches Konfidenzintervall**).

## Beispiel 163

Sei  $X \sim \mathcal{N}(\mu, \sigma^2)$  und seien  $X_1, \dots, X_n$  zugehörige Stichprobenvariablen.

## Beispiel 163

Sei  $X \sim \mathcal{N}(\mu, \sigma^2)$  und seien  $X_1, \dots, X_n$  zugehörige Stichprobenvariablen.

Wir Schätzen  $\mu$  durch  $\bar{X}$ .

## Beispiel 163

Sei  $X \sim \mathcal{N}(\mu, \sigma^2)$  und seien  $X_1, \dots, X_n$  zugehörige Stichprobenvariablen.

Wir Schätzen  $\mu$  durch  $\bar{X}$ .

Wir suchen ein symmetrisches Konfidenzintervall für  $\bar{X}$ , d.h. einen Wert  $\delta$  mit

$$\Pr[\bar{X} - \delta \leq \theta \leq \bar{X} + \delta] \geq 1 - \alpha.$$

## Beispiel 163

Sei  $X \sim \mathcal{N}(\mu, \sigma^2)$  und seien  $X_1, \dots, X_n$  zugehörige Stichprobenvariablen.

Wir schätzen  $\mu$  durch  $\bar{X}$ .

Wir suchen ein symmetrisches Konfidenzintervall für  $\bar{X}$ , d.h. einen Wert  $\delta$  mit

$$\Pr[\bar{X} - \delta \leq \theta \leq \bar{X} + \delta] \geq 1 - \alpha.$$

Es gilt (Satz 133)  $\bar{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$ . Nach Lemma 127 ist

$$Z := \sqrt{n} \cdot \frac{\bar{X} - \mu}{\sigma}$$

standardnormalverteilt.

## Konfidenzintervalle III: Normalverteilung

Für  $Z$  suchen wir nach einem Konfidenzintervall  $[-c, c]$  mit

$$\Pr[-c \leq Z \leq c] \stackrel{!}{=} 1 - \alpha.$$

## Konfidenzintervalle III: Normalverteilung

Für  $Z$  suchen wir nach einem Konfidenzintervall  $[-c, c]$  mit

$$\Pr[-c \leq Z \leq c] \stackrel{!}{=} 1 - \alpha.$$

Auflösen nach  $\mu$  ergibt

$$\Pr \left[ \bar{X} - \frac{c\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + \frac{c\sigma}{\sqrt{n}} \right] \stackrel{!}{=} 1 - \alpha.$$

## Konfidenzintervalle III: Normalverteilung

Für  $Z$  suchen wir nach einem Konfidenzintervall  $[-c, c]$  mit

$$\Pr[-c \leq Z \leq c] \stackrel{!}{=} 1 - \alpha.$$

Auflösen nach  $\mu$  ergibt

$$\Pr \left[ \bar{X} - \frac{c\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + \frac{c\sigma}{\sqrt{n}} \right] \stackrel{!}{=} 1 - \alpha.$$

Für  $c$  muss also gelten:

$$\Pr[-c \leq Z \leq c] = \Phi(c) - \Phi(-c) \stackrel{!}{=} 1 - \alpha.$$

## Konfidenzintervalle III: Normalverteilung

Für  $Z$  suchen wir nach einem Konfidenzintervall  $[-c, c]$  mit

$$\Pr[-c \leq Z \leq c] \stackrel{!}{=} 1 - \alpha.$$

Auflösen nach  $\mu$  ergibt

$$\Pr \left[ \bar{X} - \frac{c\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + \frac{c\sigma}{\sqrt{n}} \right] \stackrel{!}{=} 1 - \alpha.$$

Für  $c$  muss also gelten:

$$\Pr[-c \leq Z \leq c] = \Phi(c) - \Phi(-c) \stackrel{!}{=} 1 - \alpha.$$

Mit  $\Phi(-c) = 1 - \Phi(c)$  erhalten wir

$$2 \cdot \Phi(c) - 1 \stackrel{!}{=} 1 - \alpha$$

## Konfidenzintervalle III: Normalverteilung

Für  $Z$  suchen wir nach einem Konfidenzintervall  $[-c, c]$  mit

$$\Pr[-c \leq Z \leq c] \stackrel{!}{=} 1 - \alpha.$$

Auflösen nach  $\mu$  ergibt

$$\Pr \left[ \bar{X} - \frac{c\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + \frac{c\sigma}{\sqrt{n}} \right] \stackrel{!}{=} 1 - \alpha.$$

Für  $c$  muss also gelten:

$$\Pr[-c \leq Z \leq c] = \Phi(c) - \Phi(-c) \stackrel{!}{=} 1 - \alpha.$$

Mit  $\Phi(-c) = 1 - \Phi(c)$  erhalten wir

$$2 \cdot \Phi(c) - 1 \stackrel{!}{=} 1 - \alpha$$

also

$$c = \Phi^{-1} \left( 1 - \frac{\alpha}{2} \right).$$

## Definition 164

$X$  sei eine stetige Zufallsvariable mit Verteilung  $F_X$ .

Die Zahl  $x_\gamma$  mit

$$F_X(x_\gamma) = \gamma$$

heißt  $\gamma$ -Quantil der Verteilung  $F_X$ .

Für die Standardnormalverteilung bezeichnet  $z_\gamma$  das  $\gamma$ -Quantil.

## Definition 164

$X$  sei eine stetige Zufallsvariable mit Verteilung  $F_X$ .

Die Zahl  $x_\gamma$  mit

$$F_X(x_\gamma) = \gamma$$

heißt  $\gamma$ -Quantil der Verteilung  $F_X$ .

Für die Standardnormalverteilung bezeichnet  $z_\gamma$  das  $\gamma$ -Quantil.

Damit kann das gesuchte Konfidenzintervall durch

$$\left[ \bar{X} - \frac{z_{(1-\frac{\alpha}{2})}\sigma}{\sqrt{n}}, \bar{X} + \frac{z_{(1-\frac{\alpha}{2})}\sigma}{\sqrt{n}} \right]$$

angegeben werden.

## **23. Testen von Hypothesen**

# Testen von Hypothesen I: Einführung

Bislang haben wir versucht, Parameter von Verteilungen zu schätzen, z.B. die Erfolgsw'keit  $p$  einer Bernoulli-verteilte Variable.  $X$ .

# Testen von Hypothesen I: Einführung

Bislang haben wir versucht, Parameter von Verteilungen zu schätzen, z.B. die Erfolgsw'keit  $p$  einer Bernoulli-verteilte Variable  $X$ .

In der Praxis ist man jedoch nur an einer Eigenschaft des Parameters interessiert, z.B. ob  $p \geq 1/3$ ,  $p \leq 0.8$ , oder  $p = 0.5$  gilt.

# Testen von Hypothesen I: Einführung

Bislang haben wir versucht, Parameter von Verteilungen zu schätzen, z.B. die Erfolgsw'keit  $p$  einer Bernoulli-verteilte Variable  $X$ .

In der Praxis ist man jedoch nur an einer Eigenschaft des Parameters interessiert, z.B. ob  $p \geq 1/3$ ,  $p \leq 0.8$ , oder  $p = 0.5$  gilt.

**Statistische Tests** werden verwendet, um mit einer gewissen Fehlerw'keit zu entscheiden, ob die Eigenschaft gilt oder nicht.

## Definition 165

Die zu überprüfende Eigenschaft bezeichnet man mit  $H_0$  und wird **Nullhypothese** genannt.

## Definition 165

Die zu überprüfende Eigenschaft bezeichnet man mit  $H_0$  und wird **Nullhypothese** genannt.

Man nimmt an, dass entweder die Nullhypothese oder die mit  $H_1$  bezeichnete **Alternative** gilt.

## Definition 165

Die zu überprüfende Eigenschaft bezeichnet man mit  $H_0$  und wird **Nullhypothese** genannt.

Man nimmt an, dass entweder die Nullhypothese oder die mit  $H_1$  bezeichnete **Alternative** gilt.

Bei den meisten Tests gilt  $H_1 := \neg H_0$  (**triviale Alternative**).

## Definition 165

Die zu überprüfende Eigenschaft bezeichnet man mit  $H_0$  und wird **Nullhypothese** genannt.

Man nimmt an, dass entweder die Nullhypothese oder die mit  $H_1$  bezeichnete **Alternative** gilt.

Bei den meisten Tests gilt  $H_1 := \neg H_0$  (**triviale Alternative**).

Der Test entscheidet, welche von den beiden Eigenschaften **abgelehnt** wird.

# Testen von Hypothesen III: Terminologie

Ein Beispiel, in dem eine nicht triviale Alternative Sinn macht:

## Beispiel 166

Von einer Festplatte ist bekannt, dass sie zu einer von zwei Baureihen gehört. Die mittleren Zugriffszeiten dieser Baureihen betragen 9ms bzw. 12ms.

Es soll mit einem statistischen Test entschieden werden, zu welchem Typ die betrachtete Festplatte gehört.

In diesem Test  $H_0 : \mu \leq 9$  und  $H_1 : \mu \geq 12$  mit  $\mu$  die mittlere Zugriffszeit.

# Testen von Hypothesen IV: Terminologie

Ein Test wiederholt  $n$ -Mal das zur Variablen  $X$  zugehörigen Zufallsexperiment. Seien  $X_1, \dots, X_n$  unabhängige Stichprobenvariablen mit derselben Verteilung wie  $X$ .

## Definition 167

Die **Testgröße** eines Tests ist eine Zufallsvariable der Gestalt  $T := f(X_1, \dots, X_n)$ .

# Testen von Hypothesen IV: Terminologie

Ein Test wiederholt  $n$ -Mal das zur Variablen  $X$  zugehörigen Zufallsexperiment. Seien  $X_1, \dots, X_n$  unabhängige Stichprobenvariablen mit derselben Verteilung wie  $X$ .

## Definition 167

Die **Testgröße** eines Tests ist eine Zufallsvariable der Gestalt  $T := f(X_1, \dots, X_n)$ .

Die **Ablehnungsbedingung** ist eine Bedingung der Gestalt

$$T < c \text{ oder } T > c \quad (\text{einseitiger Test})$$

$$T < c_1 \vee T > c_2 \quad (\text{zweiseitiger Test})$$

# Testen von Hypothesen IV: Terminologie

Ein Test wiederholt  $n$ -Mal das zur Variablen  $X$  zugehörigen Zufallsexperiment. Seien  $X_1, \dots, X_n$  unabhängige Stichprobenvariablen mit derselben Verteilung wie  $X$ .

## Definition 167

Die **Testgröße** eines Tests ist eine Zufallsvariable der Gestalt  $T := f(X_1, \dots, X_n)$ .

Die **Ablehnungsbedingung** ist eine Bedingung der Gestalt

$$T < c \text{ oder } T > c \quad (\text{einseitiger Test})$$

$$T < c_1 \vee T > c_2 \quad (\text{zweiseitiger Test})$$

Die Nullhypothese wird abgelehnt, wenn der Stichprobenvektor  $\vec{x} = (x_1, \dots, x_n)$  die Ablehnungsbedingung erfüllt (d.h.,  $f(\vec{x}) < c$ ,  $f(\vec{x}) > c$ , oder  $f(\vec{x}) < c_1 \vee f(\vec{x}) > c_2$  gilt).

# Testen von Hypothesen IV: Terminologie

Ein Test wiederholt  $n$ -Mal das zur Variablen  $X$  zugehörigen Zufallsexperiment. Seien  $X_1, \dots, X_n$  unabhängige Stichprobenvariablen mit derselben Verteilung wie  $X$ .

## Definition 167

Die **Testgröße** eines Tests ist eine Zufallsvariable der Gestalt  $T := f(X_1, \dots, X_n)$ .

Die **Ablehnungsbedingung** ist eine Bedingung der Gestalt

$$T < c \text{ oder } T > c \quad (\text{einseitiger Test})$$

$$T < c_1 \vee T > c_2 \quad (\text{zweiseitiger Test})$$

Die Nullhypothese wird abgelehnt, wenn der Stichprobenvektor  $\vec{x} = (x_1, \dots, x_n)$  die Ablehnungsbedingung erfüllt (d.h.,  $f(\vec{x}) < c$ ,  $f(\vec{x}) > c$ , oder  $f(\vec{x}) < c_1 \vee f(\vec{x}) > c_2$  gilt).

Der **Ablehnungsbereich** ist die Menge der Stichprobenvektoren  $\vec{x} = (x_1, \dots, x_n)$ , für die die Nullhypothese abgelehnt wird.

## Definition 168

Ein Fehler 1. Art tritt ein, wenn  $H_0$  irrtümlich abgelehnt wird.  
(D.h.,  $H_0$  gilt aber die Stichprobe  $\vec{x}$  liegt im Ablehnungsbereich.)

## Definition 168

Ein **Fehler 1. Art** tritt ein, wenn  $H_0$  **irrtümlich abgelehnt** wird.  
(D.h.,  $H_0$  gilt aber die Stichprobe  $\vec{x}$  liegt im Ablehnungsbereich.)

Ein **Fehler 2. Art** tritt ein, wenn  $H_0$  **irrtümlich angenommen** wird.  
(D.h.,  $H_0$  gilt nicht aber die Stichprobe  $\vec{x}$  liegt nicht im Ablehnungsbereich.)

## Definition 168

Ein **Fehler 1. Art** tritt ein, wenn  $H_0$  **irrtümlich abgelehnt** wird.  
(D.h.,  $H_0$  gilt aber die Stichprobe  $\vec{x}$  liegt im Ablehnungsbereich.)

Ein **Fehler 2. Art** tritt ein, wenn  $H_0$  **irrtümlich angenommen** wird.  
(D.h.,  $H_0$  gilt nicht aber die Stichprobe  $\vec{x}$  liegt nicht im Ablehnungsbereich.)

Sei  $\alpha$  eine obere Schranke der W'keit für den Fehler 1. Art. Wir sagen, dass der Test  **$\alpha$ -Fehler** oder **Signifikanzniveau** von  $\alpha$  hat.

## Definition 168

Ein **Fehler 1. Art** tritt ein, wenn  $H_0$  **irrtümlich abgelehnt** wird.  
(D.h.,  $H_0$  gilt aber die Stichprobe  $\vec{x}$  liegt im Ablehnungsbereich.)

Ein **Fehler 2. Art** tritt ein, wenn  $H_0$  **irrtümlich angenommen** wird.  
(D.h.,  $H_0$  gilt nicht aber die Stichprobe  $\vec{x}$  liegt nicht im Ablehnungsbereich.)

Sei  $\alpha$  eine obere Schranke der W'keit für den Fehler 1. Art. Wir sagen, dass der Test  **$\alpha$ -Fehler** oder **Signifikanzniveau** von  $\alpha$  hat.

In der Praxis gibt man sich einen Wert für  $\alpha$  vor und bestimmt man den Ablehnungsbereich so, dass der Test Signifikanzniveau  $\alpha$  hat.

Übliche Werte für  $\alpha$  sind 0,05, 0,01 oder 0,001.

# Testen von Hypothesen VI: Terminologie

Die Minimierung des Fehlers 1. Art und des Fehlers 2. Art sind **gegenläufige Ziele**, z.B.:

- Lehnt man die Nullhypothese **nie** ab, hat ein Fehler 1. Art W'keit **0**. Allerdings tritt ein Fehler 2. Art immer ein, wenn  $H_0$  nicht gilt.
- Lehnt man die Nullhypothese **immer** ab, hat ein Fehler 2. Art W'keit **0**. Allerdings tritt ein Fehler 1. Art immer ein, wenn  $H_0$  gilt.

# Testen von Hypothesen VI: Terminologie

Die Minimierung des Fehlers 1. Art und des Fehlers 2. Art sind **gegenläufige Ziele**, z.B.:

- Lehnt man die Nullhypothese **nie** ab, hat ein Fehler 1. Art W'keit **0**. Allerdings tritt ein Fehler 2. Art immer ein, wenn  $H_0$  nicht gilt.
- Lehnt man die Nullhypothese **immer** ab, hat ein Fehler 2. Art W'keit **0**. Allerdings tritt ein Fehler 1. Art immer ein, wenn  $H_0$  gilt.

Ziel der meisten statistischen Tests ist eine kleine W'keit **für den Fehler 1. Art** zu garantieren. Die W'keit des Fehlers 2. Art kann groß sein!

# Testen von Hypothesen VI: Terminologie

Die Minimierung des Fehlers 1. Art und des Fehlers 2. Art sind **gegenläufige Ziele**, z.B.:

- Lehnt man die Nullhypothese **nie** ab, hat ein Fehler 1. Art W'keit 0. Allerdings tritt ein Fehler 2. Art immer ein, wenn  $H_0$  nicht gilt.
- Lehnt man die Nullhypothese **immer** ab, hat ein Fehler 2. Art W'keit 0. Allerdings tritt ein Fehler 1. Art immer ein, wenn  $H_0$  gilt.

Ziel der meisten statistischen Tests ist eine kleine W'keit **für den Fehler 1. Art** zu garantieren. Die W'keit des Fehlers 2. Art kann groß sein!

Bei der Wahl der Nullhypothese (setzt man  $H_0 = H$  oder  $H_0 = \neg H$  ?) muss das berücksichtigt werden.

# Testen von Hypothesen VII: Entwurf eines Tests

Wir entwerfen einen Test für den Parameter  $p$  einer Bernoulli-verteilten Zufallsvariablen  $X$ .

Der Wert  $k \in \mathbb{R}$  muss in Abhängigkeit des Signifikanzniveaus  $\alpha$  festgelegt werden.

# Testen von Hypothesen VII: Entwurf eines Tests

Wir entwerfen einen Test für den Parameter  $p$  einer Bernoulli-verteilten Zufallsvariablen  $X$ .

Wir setzen  $H_0 : p \geq p_0$  und  $H_1 : p < p_0$ .

# Testen von Hypothesen VII: Entwurf eines Tests

Wir entwerfen einen Test für den Parameter  $p$  einer Bernoulli-verteilten Zufallsvariablen  $X$ .

Wir setzen  $H_0 : p \geq p_0$  und  $H_1 : p < p_0$ .

Als Testgröße verwenden wir  $T := X_1 + \dots + X_n$ .  
(Anzahl der Köpfe.)

## Testen von Hypothesen VII: Entwurf eines Tests

Wir entwerfen einen Test für den Parameter  $p$  einer Bernoulli-verteilten Zufallsvariablen  $X$ .

Wir setzen  $H_0 : p \geq p_0$  und  $H_1 : p < p_0$ .

Als Testgröße verwenden wir  $T := X_1 + \dots + X_n$ .  
(Anzahl der Köpfe.)

Wir lehnen  $H_0$  ab wenn der Wert von  $T$  “zu klein“ ist.  
Als Gestalt des Ablehnungsbereichs wählen wir also  $T < k$ .  
(Einseitiger Test.)

# Testen von Hypothesen VII: Entwurf eines Tests

Wir entwerfen einen Test für den Parameter  $p$  einer Bernoulli-verteilten Zufallsvariablen  $X$ .

Wir setzen  $H_0 : p \geq p_0$  und  $H_1 : p < p_0$ .

Als Testgröße verwenden wir  $T := X_1 + \dots + X_n$ .  
(Anzahl der Köpfe.)

Wir lehnen  $H_0$  ab wenn der Wert von  $T$  “zu klein“ ist.

Als Gestalt des Ablehnungsbereichs wählen wir also  $T < k$ .  
(Einseitiger Test.)

Für eine Nullhypothese  $H_0 : p = p_0$  würden wir  $T < k_1 \vee T > k_2$  wählen.

# Testen von Hypothesen VII: Entwurf eines Tests

Wir entwerfen einen Test für den Parameter  $p$  einer Bernoulli-verteilten Zufallsvariablen  $X$ .

Wir setzen  $H_0 : p \geq p_0$  und  $H_1 : p < p_0$ .

Als Testgröße verwenden wir  $T := X_1 + \dots + X_n$ .  
(Anzahl der Köpfe.)

Wir lehnen  $H_0$  ab wenn der Wert von  $T$  “zu klein“ ist.

Als Gestalt des Ablehnungsbereichs wählen wir also  $T < k$ .  
(Einseitiger Test.)

Für eine Nullhypothese  $H_0 : p = p_0$  würden wir  $T < k_1 \vee T > k_2$  wählen.

Der Wert  $k \in \mathbb{R}$  muss in Abhängigkeit des Signifikanzniveaus  $\alpha$  festgelegt werden.

# Testen von Hypothesen VIII: Entwurf eines Tests

Wir bestimmen  $k$ .

# Testen von Hypothesen VIII: Entwurf eines Tests

Wir bestimmen  $k$ .

Es gilt  $T \sim \text{Bin}(n, p)$ . Bei großer Stichprobenumfang  $n$  ist die Variable

$$\tilde{T} := \frac{T - np}{\sqrt{np(1-p)}}.$$

annähernd standardnormalverteilt (siehe Korollar 139).

# Testen von Hypothesen VIII: Entwurf eines Tests

Wir bestimmen  $k$ .

Es gilt  $T \sim \text{Bin}(n, p)$ . Bei großer Stichprobenumfang  $n$  ist die Variable

$$\tilde{T} := \frac{T - np}{\sqrt{np(1-p)}}.$$

annähernd standardnormalverteilt (siehe Korollar 139).

Wir berechnen für jeden Wert von  $k$  die maximale Fehlerw'keit über die möglichen Werten von  $p$ :

$$\text{Fehlerw'keit 1. Art} \leq \sup_{p \text{ erfüllt } H_0} \Pr_p[„H_0 \text{ wird abgelehnt“}]$$

$$= \sup_{p < p_0} \Pr_p[T \geq k] = \Pr_{p=p_0}[T \geq k]$$

# Testen von Hypothesen VIII: Entwurf eines Tests

Wir bestimmen  $k$ .

Es gilt  $T \sim \text{Bin}(n, p)$ . Bei großer Stichprobenumfang  $n$  ist die Variable

$$\tilde{T} := \frac{T - np}{\sqrt{np(1-p)}}.$$

annähernd standardnormalverteilt (siehe Korollar 139).

Wir berechnen für jeden Wert von  $k$  die maximale Fehlerw'keit über die möglichen Werten von  $p$ :

$$\begin{aligned} \text{Fehlerw'keit 1. Art} &\leq \sup_{p \text{ erfüllt } H_0} \Pr_p[\text{„}H_0 \text{ wird abgelehnt“}] \\ &= \max_{p \geq p_0} \Pr_p[T < k] \end{aligned}$$

$$= \Pr_{p=p_0}[T \geq k]$$

# Testen von Hypothesen VIII: Entwurf eines Tests

Wir bestimmen  $k$ .

Es gilt  $T \sim \text{Bin}(n, p)$ . Bei großer Stichprobenumfang  $n$  ist die Variable

$$\tilde{T} := \frac{T - np}{\sqrt{np(1-p)}}.$$

annähernd standardnormalverteilt (siehe Korollar 139).

Wir berechnen für jeden Wert von  $k$  die maximale Fehlerw'keit über die möglichen Werten von  $p$ :

$$\begin{aligned} \text{Fehlerw'keit 1. Art} &\leq \sup_{p \text{ erfüllt } H_0} \Pr_p[„H_0 \text{ wird abgelehnt“}] \\ &= \max_{p \geq p_0} \Pr_p[T < k] = \Pr_{p=p_0}[T < k] \end{aligned}$$

# Testen von Hypothesen VIII: Entwurf eines Tests

Wir bestimmen  $k$ .

Es gilt  $T \sim \text{Bin}(n, p)$ . Bei großer Stichprobenumfang  $n$  ist die Variable

$$\tilde{T} := \frac{T - np}{\sqrt{np(1-p)}}.$$

annähernd standardnormalverteilt (siehe Korollar 139).

Wir berechnen für jeden Wert von  $k$  die maximale Fehlerw'keit über die möglichen Werten von  $p$ :

$$\begin{aligned} \text{Fehlerw'keit 1. Art} &\leq \sup_{p \text{ erfüllt } H_0} \Pr_p[„H_0 \text{ wird abgelehnt“}] \\ &= \max_{p \geq p_0} \Pr_p[T < k] = \Pr_{p=p_0}[T < k] \end{aligned}$$

Fehlerw'keit 2. Art

# Testen von Hypothesen VIII: Entwurf eines Tests

Wir bestimmen  $k$ .

Es gilt  $T \sim \text{Bin}(n, p)$ . Bei großer Stichprobenumfang  $n$  ist die Variable

$$\tilde{T} := \frac{T - np}{\sqrt{np(1-p)}}.$$

annähernd standardnormalverteilt (siehe Korollar 139).

Wir berechnen für jeden Wert von  $k$  die maximale Fehlerw'keit über die möglichen Werten von  $p$ :

$$\begin{aligned} \text{Fehlerw'keit 1. Art} &\leq \sup_{p \text{ erfüllt } H_0} \Pr_p[\text{„}H_0 \text{ wird abgelehnt“}] \\ &= \max_{p \geq p_0} \Pr_p[T < k] = \Pr_{p=p_0}[T < k] \end{aligned}$$

$$\text{Fehlerw'keit 2. Art} \leq \sup_{p \text{ erfüllt } H_1} \Pr_p[\text{„}H_1 \text{ wird abgelehnt“}]$$

# Testen von Hypothesen VIII: Entwurf eines Tests

Wir bestimmen  $k$ .

Es gilt  $T \sim \text{Bin}(n, p)$ . Bei großer Stichprobenumfang  $n$  ist die Variable

$$\tilde{T} := \frac{T - np}{\sqrt{np(1-p)}}.$$

annähernd standardnormalverteilt (siehe Korollar 139).

Wir berechnen für jeden Wert von  $k$  die maximale Fehlerw'keit über die möglichen Werten von  $p$ :

$$\begin{aligned} \text{Fehlerw'keit 1. Art} &\leq \sup_{p \text{ erfüllt } H_0} \Pr_p[„H_0 \text{ wird abgelehnt“}] \\ &= \max_{p \geq p_0} \Pr_p[T < k] = \Pr_{p=p_0}[T < k] \end{aligned}$$

$$\begin{aligned} \text{Fehlerw'keit 2. Art} &\leq \sup_{p \text{ erfüllt } H_1} \Pr_p[„H_1 \text{ wird abgelehnt“}] \\ &= \sup_{p < p_0} \Pr_p[T \geq k] \end{aligned}$$

# Testen von Hypothesen VIII: Entwurf eines Tests

Wir bestimmen  $k$ .

Es gilt  $T \sim \text{Bin}(n, p)$ . Bei großer Stichprobenumfang  $n$  ist die Variable

$$\tilde{T} := \frac{T - np}{\sqrt{np(1-p)}}.$$

annähernd standardnormalverteilt (siehe Korollar 139).

Wir berechnen für jeden Wert von  $k$  die maximale Fehlerw'keit über die möglichen Werten von  $p$ :

$$\begin{aligned} \text{Fehlerw'keit 1. Art} &\leq \sup_{p \text{ erfüllt } H_0} \Pr_p[„H_0 \text{ wird abgelehnt“}] \\ &= \max_{p \geq p_0} \Pr_p[T < k] = \Pr_{p=p_0}[T < k] \end{aligned}$$

$$\begin{aligned} \text{Fehlerw'keit 2. Art} &\leq \sup_{p \text{ erfüllt } H_1} \Pr_p[„H_1 \text{ wird abgelehnt“}] \\ &= \sup_{p < p_0} \Pr_p[T \geq k] = \Pr_{p=p_0}[T \geq k] \end{aligned}$$

# Testen von Hypothesen IX: Entwurf eines Tests

Für das gewünschte Signifikanzniveau  $\alpha$  erhalten wir

$$\alpha = \Pr_{p=p_0}[T < k]$$

# Testen von Hypothesen IX: Entwurf eines Tests

Für das gewünschte Signifikanzniveau  $\alpha$  erhalten wir

$$\begin{aligned}\alpha &= \Pr_{p=p_0}[T < k] \\ &= \Pr_{p=p_0}\left[\tilde{T} < \frac{k - np}{\sqrt{np(1-p)}}\right]\end{aligned}$$

# Testen von Hypothesen IX: Entwurf eines Tests

Für das gewünschte Signifikanzniveau  $\alpha$  erhalten wir

$$\begin{aligned}\alpha &= \Pr_{p=p_0}[T < k] \\ &= \Pr_{p=p_0} \left[ \tilde{T} < \frac{k - np}{\sqrt{np(1-p)}} \right] \\ &= \Pr \left[ \tilde{T} < \frac{k - np_0}{\sqrt{np_0(1-p_0)}} \right] \approx \Phi \left( \frac{k - np_0}{\sqrt{np_0(1-p_0)}} \right).\end{aligned}$$

# Testen von Hypothesen X: Entwurf eines Tests

Damit gilt: Für jedes

$$k \leq z_\alpha \sqrt{np_0(1-p_0)} + np_0$$

kann die Nullhypothese  $p \geq p_0$  abgelehnt werden mit Signifikanzniveau  $\alpha$ .

# Testen von Hypothesen X: Entwurf eines Tests

Damit gilt: Für jedes

$$k \leq z_\alpha \sqrt{np_0(1-p_0)} + np_0$$

kann die Nullhypothese  $p \geq p_0$  abgelehnt werden mit Signifikanzniveau  $\alpha$ .

Das größte solche  $k$  minimiert den Fehler 2. Art.

# Testen von Hypothesen X: Entwurf eines Tests

Damit gilt: Für jedes

$$k \leq z_\alpha \sqrt{np_0(1 - p_0)} + np_0$$

kann die Nullhypothese  $p \geq p_0$  abgelehnt werden mit Signifikanzniveau  $\alpha$ .

Das größte solche  $k$  minimiert den Fehler 2. Art.

Für diesen Wert gilt:

$$\text{Fehlerw'keit 1. Art} = \Pr_{p=p_0}[T < k]$$

$$\text{Fehlerw'keit 2. Art} = \Pr_{p=p_0}[T \geq k]$$

und damit Fehlerw'keit 2. Art = 1 – Fehlerw'keit 1. Art

# Testen von Hypothesen X: Entwurf eines Tests

Damit gilt: Für jedes

$$k \leq z_\alpha \sqrt{np_0(1 - p_0)} + np_0$$

kann die Nullhypothese  $p \geq p_0$  abgelehnt werden mit Signifikanzniveau  $\alpha$ .

Das größte solche  $k$  minimiert den Fehler 2. Art.

Für diesen Wert gilt:

$$\text{Fehlerw'keit 1. Art} = \Pr_{p=p_0}[T < k]$$

$$\text{Fehlerw'keit 2. Art} = \Pr_{p=p_0}[T \geq k]$$

und damit Fehlerw'keit 2. Art = 1 – Fehlerw'keit 1. Art

Es ist also nicht möglich, beide W'keiten gleichzeitig zu reduzieren.

## Beispiel 169

Ein Test soll mit Signifikanz  $\alpha = 0.05$  bestimmen, ob die W'keit von Kopf bei einer Münze mindestens  $1/3$  beträgt.

Wenn die Münze  $n = 450$  Mal geworfen wird, dann wird die Hypothese  $p \geq 1/3$  abgelehnt werde, wenn die Anzahl der Köpfe unter

$$\begin{aligned}k &= z_{0.05} \cdot \sqrt{450 \cdot \frac{1}{3} \cdot \frac{2}{3}} + 450 \cdot \frac{1}{3} \\ &= z_{0.05} \cdot 10 + 150 \\ &\approx -1.64 \cdot 10 + 150 = 133,6\end{aligned}$$

liegt.

# Testen von Hypothesen XI: Entwurf eines Tests

Bei echten **Alternativtests** werden für hinreichend große Stichproben beide Testfehler klein.

## Beispiel 170

Bei zwei Sorten von gezinkten Münzen ist die Erfolgs'wkeit jeweils  $p \geq 3/5$  und  $p \leq 2/5$ . Mit Hilfe eines Tests soll die Sorte einer gegebenen Münze bestimmt werden.

# Testen von Hypothesen XI: Entwurf eines Tests

Bei echten **Alternativtests** werden für hinreichend große Stichproben beide Testfehler klein.

## Beispiel 170

Bei zwei Sorten von gezinkten Münzen ist die Erfolgs'wkeit jeweils  $p \geq 3/5$  und  $p \leq 2/5$ . Mit Hilfe eines Tests soll die Sorte einer gegebenen Münze bestimmt werden.

Unter dieser Annahme setzt man:

$$H_0 : p \geq 3/5 \quad H_1 : p \leq 2/5.$$

# Testen von Hypothesen XII: Entwurf eines Tests

$$\text{Fehlerw'keit 1. Art} \leq \sup_{p \text{ erfüllt } H_0} \Pr_p[„H_0 \text{ wird abgelehnt“}]$$

$$\begin{aligned} &\leq \sup_{p \text{ erfüllt } H_1} \Pr_p[„H_1 \text{ wird abgelehnt“}] \\ &= \sup_{p \leq 2/5} \Pr_p[T \geq k] = \Pr_{p=2/5}[T \geq k] \\ &\approx 1 - \Phi\left(\frac{k - n \cdot 2/5}{\sqrt{n \cdot 6/25}}\right) \end{aligned}$$

# Testen von Hypothesen XII: Entwurf eines Tests

$$\begin{aligned}\text{Fehlerw'keit 1. Art} &\leq \sup_{p \text{ erfüllt } H_0} \Pr_p[„H_0 \text{ wird abgelehnt“}] \\ &= \max_{p \geq 3/5} \Pr_p[T < k]\end{aligned}$$

$$\begin{aligned}&= \sup_{p \leq 2/5} \Pr_p[T \geq k] = \Pr_{p=2/5}[T \geq k] \\ &\approx 1 - \Phi\left(\frac{k - n \cdot 2/5}{\sqrt{n \cdot 6/25}}\right)\end{aligned}$$

# Testen von Hypothesen XII: Entwurf eines Tests

$$\begin{aligned}\text{Fehlerw'keit 1. Art} &\leq \sup_{p \text{ erfüllt } H_0} \Pr_p[„H_0 \text{ wird abgelehnt“}] \\ &= \max_{p \geq 3/5} \Pr_p[T < k] = \Pr_{p=3/5}[T < k]\end{aligned}$$

$$= \Pr_{p=2/5}[T \geq k]$$

$$\approx 1 - \Phi\left(\frac{k - n \cdot 2/5}{\sqrt{n \cdot 6/25}}\right)$$

# Testen von Hypothesen XII: Entwurf eines Tests

$$\begin{aligned}\text{Fehlerw'keit 1. Art} &\leq \sup_{p \text{ erfüllt } H_0} \Pr_p[„H_0 \text{ wird abgelehnt“}] \\ &= \max_{p \geq 3/5} \Pr_p[T < k] = \Pr_{p=3/5}[T < k] \\ &\approx \Phi\left(\frac{k - n \cdot 3/5}{\sqrt{n \cdot 6/25}}\right)\end{aligned}$$

$$\approx 1 - \Phi\left(\frac{k - n \cdot 2/5}{\sqrt{n \cdot 6/25}}\right)$$

# Testen von Hypothesen XII: Entwurf eines Tests

$$\begin{aligned}\text{Fehlerw'keit 1. Art} &\leq \sup_{p \text{ erfüllt } H_0} \Pr_p[„H_0 \text{ wird abgelehnt“}] \\ &= \max_{p \geq 3/5} \Pr_p[T < k] = \Pr_{p=3/5}[T < k] \\ &\approx \Phi\left(\frac{k - n \cdot 3/5}{\sqrt{n \cdot 6/25}}\right)\end{aligned}$$

Fehlerw'keit 2. Art

# Testen von Hypothesen XII: Entwurf eines Tests

$$\begin{aligned}\text{Fehlerw'keit 1. Art} &\leq \sup_{p \text{ erfüllt } H_0} \Pr_p[„H_0 \text{ wird abgelehnt“}] \\ &= \max_{p \geq 3/5} \Pr_p[T < k] = \Pr_{p=3/5}[T < k] \\ &\approx \Phi\left(\frac{k - n \cdot 3/5}{\sqrt{n \cdot 6/25}}\right)\end{aligned}$$

$$\text{Fehlerw'keit 2. Art} \leq \sup_{p \text{ erfüllt } H_1} \Pr_p[„H_1 \text{ wird abgelehnt“}]$$

# Testen von Hypothesen XII: Entwurf eines Tests

$$\begin{aligned}\text{Fehlerw'keit 1. Art} &\leq \sup_{p \text{ erfüllt } H_0} \Pr_p[„H_0 \text{ wird abgelehnt“}] \\ &= \max_{p \geq 3/5} \Pr_p[T < k] = \Pr_{p=3/5}[T < k] \\ &\approx \Phi\left(\frac{k - n \cdot 3/5}{\sqrt{n \cdot 6/25}}\right)\end{aligned}$$

$$\begin{aligned}\text{Fehlerw'keit 2. Art} &\leq \sup_{p \text{ erfüllt } H_1} \Pr_p[„H_1 \text{ wird abgelehnt“}] \\ &= \sup_{p \leq 2/5} \Pr_p[T \geq k]\end{aligned}$$

# Testen von Hypothesen XII: Entwurf eines Tests

$$\begin{aligned}\text{Fehlerw'keit 1. Art} &\leq \sup_{p \text{ erfüllt } H_0} \Pr_p[„H_0 \text{ wird abgelehnt“}] \\ &= \max_{p \geq 3/5} \Pr_p[T < k] = \Pr_{p=3/5}[T < k] \\ &\approx \Phi\left(\frac{k - n \cdot 3/5}{\sqrt{n \cdot 6/25}}\right)\end{aligned}$$

$$\begin{aligned}\text{Fehlerw'keit 2. Art} &\leq \sup_{p \text{ erfüllt } H_1} \Pr_p[„H_1 \text{ wird abgelehnt“}] \\ &= \sup_{p \leq 2/5} \Pr_p[T \geq k] = \Pr_{p=2/5}[T \geq k]\end{aligned}$$

# Testen von Hypothesen XII: Entwurf eines Tests

$$\begin{aligned}\text{Fehlerw'keit 1. Art} &\leq \sup_{p \text{ erfüllt } H_0} \Pr_p[„H_0 \text{ wird abgelehnt“}] \\ &= \max_{p \geq 3/5} \Pr_p[T < k] = \Pr_{p=3/5}[T < k] \\ &\approx \Phi\left(\frac{k - n \cdot 3/5}{\sqrt{n \cdot 6/25}}\right)\end{aligned}$$

$$\begin{aligned}\text{Fehlerw'keit 2. Art} &\leq \sup_{p \text{ erfüllt } H_1} \Pr_p[„H_1 \text{ wird abgelehnt“}] \\ &= \sup_{p \leq 2/5} \Pr_p[T \geq k] = \Pr_{p=2/5}[T \geq k] \\ &\approx 1 - \Phi\left(\frac{k - n \cdot 2/5}{\sqrt{n \cdot 6/25}}\right)\end{aligned}$$

## Testen von Hypothesen XIII: Entwurf eines Tests

Mit z.B.  $n = 100$  und  $\alpha = 0.05$  ergibt sich  $k \approx 51.94$ . Für diesen Wert erhalten wir

$$\begin{aligned}\text{Fehlerw'keit 2. Art} &\approx 1 - \Phi\left(\frac{51.94 - 40}{2\sqrt{6}}\right) \\ &\approx 1 - \Phi(2.437) \approx 1 - 0.9926 \\ &= 0.0074\end{aligned}$$

## **24. Praktische Anwendung statistischer Tests**

# Praktische Anwendung statistischer Tests

Das im vorhergehenden Abschnitt konstruierte Testverfahren taucht in der Literatur unter dem Namen **approximativer Binomialtest** auf.

# Praktische Anwendung statistischer Tests

Das im vorhergehenden Abschnitt konstruierte Testverfahren taucht in der Literatur unter dem Namen **approximativer Binomialtest** auf.

Die folgende Tabelle 1 gibt einen Überblick über die Eckdaten dieses Tests.

# Praktische Anwendung statistischer Tests

## Tabelle: Approximativer Binomialtest

### Annahmen:

$X_1, \dots, X_n$  seien unabhängig und identisch verteilt mit  $\Pr[X_i = 1] = p$  und  $\Pr[X_i = 0] = 1 - p$ , wobei  $p$  unbekannt sei.

$n$  sei hinreichend groß, so dass die Approximation aus Korollar 139 brauchbare Ergebnisse liefert.

### Hypothesen:

- a)  $H_0 : p = p_0$  gegen  $H_1 : p \neq p_0$ ,
- b)  $H_0 : p \geq p_0$  gegen  $H_1 : p < p_0$ ,
- c)  $H_0 : p \leq p_0$  gegen  $H_1 : p > p_0$ .

### Testgröße:

$$Z := \frac{h - np_0}{\sqrt{np_0(1 - p_0)}},$$

wobei  $h := X_1 + \dots + X_n$  die Häufigkeit bezeichnet, mit der die Ereignisse  $X_i = 1$  aufgetreten sind.

### Ablehnungskriterium für $H_0$ bei Signifikanzniveau $\alpha$ :

- a)  $|Z| > z_{1-\alpha/2}$ ,
- b)  $Z < z_\alpha$ ,
- c)  $Z > z_{1-\alpha}$ .

# Allgemeines Vorgehen bei statistischen Tests

## 1. Schritt: Formulierung von Annahmen.

Ganz ohne Annahmen kommt man meist nicht aus.

Übliche Annahmen betreffen die Verteilung der Stichprobenvariablen und deren Unabhängigkeit.

# Allgemeines Vorgehen bei statistischen Tests

## 1. Schritt: Formulierung von Annahmen.

Ganz ohne Annahmen kommt man meist nicht aus.

Übliche Annahmen betreffen die Verteilung der Stichprobenvariablen und deren Unabhängigkeit.

## 2. Schritt: Formulierung der Nullhypothese.

# Allgemeines Vorgehen bei statistischen Tests

1. **Schritt:** Formulierung von Annahmen.

Ganz ohne Annahmen kommt man meist nicht aus.

Übliche Annahmen betreffen die Verteilung der Stichprobenvariablen und deren Unabhängigkeit.

2. **Schritt:** Formulierung der Nullhypothese.

3. **Schritt:** Auswahl des Testverfahrens.

# Allgemeines Vorgehen bei statistischen Tests

1. **Schritt:** Formulierung von Annahmen.

Ganz ohne Annahmen kommt man meist nicht aus.

Übliche Annahmen betreffen die Verteilung der Stichprobenvariablen und deren Unabhängigkeit.

2. **Schritt:** Formulierung der Nullhypothese.

3. **Schritt:** Auswahl des Testverfahrens.

4. **Schritt:** Durchführung des Tests und Entscheidung.

# Allgemeines Vorgehen bei statistischen Tests

1. **Schritt:** Formulierung von Annahmen.

Ganz ohne Annahmen kommt man meist nicht aus.

Übliche Annahmen betreffen die Verteilung der Stichprobenvariablen und deren Unabhängigkeit.

2. **Schritt:** Formulierung der Nullhypothese.

3. **Schritt:** Auswahl des Testverfahrens.

4. **Schritt:** Durchführung des Tests und Entscheidung.

# Wie findet man das richtige Testverfahren? I

Statistische Tests kann man nach mehreren Kriterien in Klassen einteilen.

- **Anzahl der beteiligten Zufallsgrößen**

- **Ein-Stichproben-Test.**

Es wird nur eine Zufallsgröße untersucht, für die eine Stichprobe erzeugt wird.

**Beispiel:** Beträgt die mittlere Zugriffszeit auf einen Datenbankserver im Mittel höchstens 10ms?

# Wie findet man das richtige Testverfahren? I

Statistische Tests kann man nach mehreren Kriterien in Klassen einteilen.

- **Anzahl der beteiligten Zufallsgrößen**

- **Ein-Stichproben-Test.**

Es wird nur eine Zufallsgröße untersucht, für die eine Stichprobe erzeugt wird.

**Beispiel:** Beträgt die mittlere Zugriffszeit auf einen Datenbankserver im Mittel höchstens 10ms?

- **Zwei-Stichproben-Test.**

Zwei Zufallsgrößen, für die jeweils eine Stichprobe erzeugt wird, werden verglichen.

**Beispiel:** Hat Datenbankserver A eine kürzere mittlere Zugriffszeit als Datenbankserver B?

- **Art der Nullhypothese**

- Tests auf Lageparameter.

Die Nullhypothese ist eine Aussage über **Lageparameter** der Verteilung wie den Erwartungswert oder die Varianz.

- **Art der Nullhypothese**

- Tests auf Lageparameter.

Die Nullhypothese ist eine Aussage über **Lageparameter** der Verteilung wie den Erwartungswert oder die Varianz.

- Tests auf eine vorgegebene Verteilung.

Die Nullhypothese ist eine Aussage über den Verteilungstyp, z.B. dass die Zufallsgröße Normalverteilt ist.

- **Art der Nullhypothese**

- Tests auf Lageparameter.

Die Nullhypothese ist eine Aussage über **Lageparameter** der Verteilung wie den Erwartungswert oder die Varianz.

- Tests auf eine vorgegebene Verteilung.

Die Nullhypothese ist eine Aussage über den Verteilungstyp, z.B. dass die Zufallsgröße Normalverteilt ist.

- Tests auf ein Maß für die Abhängigkeit verschiedener Zufallsgröße.

Z.B. sagt die Nullhypothese, dass zwei Zufallsvariablen unabhängig sind.

# Wie findet man das richtige Testverfahren? III

- **Annahmen über die Zufallsgrößen**

# Wie findet man das richtige Testverfahren? III

- **Annahmen über die Zufallsgrößen**
  - Bekannter/unbekannter Verteilungstyp.

- **Annahmen über die Zufallsgrößen**
  - Bekannter/unbekannter Verteilungstyp.
  - Bekannter/unbekannter Erwartungswert.

- **Annahmen über die Zufallsgrößen**
  - Bekannter/unbekannter Verteilungstyp.
  - Bekannter/unbekannter Erwartungswert.
  - Bekannte/unbekannte Varianz.

# Ein-Stichproben-Tests für Lageparameter I: Gaußtest

## Tabelle: Gaußtest

### Annahmen:

$X_1, \dots, X_n$  seien unabhängig und identisch verteilt mit  $X_i \sim \mathcal{N}(\mu, \sigma^2)$ ,  
wobei  $\sigma^2$  bekannt ist.

Alternativ gelte  $\mathbb{E}[X_i] = \mu$  und  $\text{Var}[X_i] = \sigma^2$ , und  $n$  sei groß genug.

### Hypothesen:

- a)  $H_0 : \mu = \mu_0$  gegen  $H_1 : \mu \neq \mu_0$ ,
- b)  $H_0 : \mu \geq \mu_0$  gegen  $H_1 : \mu < \mu_0$ ,
- c)  $H_0 : \mu \leq \mu_0$  gegen  $H_1 : \mu > \mu_0$ .

### Testgröße:

$$Z := \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n}.$$

### Ablehnungskriterium für $H_0$ bei Signifikanzniveau $\alpha$ :

- a)  $|Z| > z_{1-\alpha/2}$ ,
- b)  $Z < z_\alpha$ ,
- c)  $Z > z_{1-\alpha}$ .

# Ein-Stichproben-Tests für Lageparameter II: Gaußtest

Der Gaußtest hat den Nachteil, dass man die Varianz  $\sigma^2$  der Variablen  $X_1, \dots, X_n$  kennen muss.

# Ein-Stichproben-Tests für Lageparameter II: Gaußtest

Der Gaußtest hat den Nachteil, dass man die Varianz  $\sigma^2$  der Variablen  $X_1, \dots, X_n$  kennen muss.

Wenn diese unbekannt ist kann die Varianz durch die Stichprobenvarianz  $S^2$  (siehe Definition 158) ersetzt werden.

Dies führt auf den so genannten  $t$ -Test.

# Ein-Stichproben-Tests für Lageparameter III: $t$ -Test

## Tabelle: $t$ -Test

### Annahmen:

$X_1, \dots, X_n$  seien unabhängig und identisch verteilt mit  $X_i \sim \mathcal{N}(\mu, \sigma^2)$ .  
Alternativ gelte  $\mathbb{E}[X_i] = \mu$  und  $\text{Var}[X_i] = \sigma^2$ , und  $n$  sei groß genug.

### Hypothesen:

- a)  $H_0 : \mu = \mu_0$  gegen  $H_1 : \mu \neq \mu_0$ ,
- b)  $H_0 : \mu \geq \mu_0$  gegen  $H_1 : \mu < \mu_0$ ,
- c)  $H_0 : \mu \leq \mu_0$  gegen  $H_1 : \mu > \mu_0$ .

### Testgröße:

$$T := \frac{\bar{X} - \mu_0}{S} \sqrt{n}.$$

### Ablehnungskriterium für $H_0$ bei Signifikanzniveau $\alpha$ :

- a)  $|T| > t_{n-1, 1-\alpha/2}$ ,
- b)  $T < t_{n-1, \alpha}$ ,
- c)  $T > t_{n-1, 1-\alpha}$ .

# Ein-Stichproben-Tests für Lageparameter IV: $t$ -Test

Hierbei bezeichnet  $t_{n-1,1-\alpha}$  das

$(1 - \alpha)$ -Quantil der  $t$ -Verteilung mit  $n - 1$  Freiheitsgraden

an.

# Ein-Stichproben-Tests für Lageparameter IV: $t$ -Test

Hierbei bezeichnet  $t_{n-1,1-\alpha}$  das

$(1 - \alpha)$ -Quantil der  $t$ -Verteilung mit  $n - 1$  Freiheitsgraden

an.

Die  $t$ -Verteilung nennt man auch **Student-Verteilung**.

# Ein-Stichproben-Tests für Lageparameter IV: $t$ -Test

Hierbei bezeichnet  $t_{n-1,1-\alpha}$  das

$(1 - \alpha)$ -Quantil der  $t$ -Verteilung mit  $n - 1$  Freiheitsgraden

an.

Die  $t$ -Verteilung nennt man auch **Student-Verteilung**.

Es handelt sich eigentlich um eine Familie von Verteilungen, eine für jede Anzahl von Freiheitsgraden.

# Ein-Stichproben-Tests für Lageparameter IV: $t$ -Test

Hierbei bezeichnet  $t_{n-1,1-\alpha}$  das

$(1 - \alpha)$ -Quantil der  $t$ -Verteilung mit  $n - 1$  Freiheitsgraden

an.

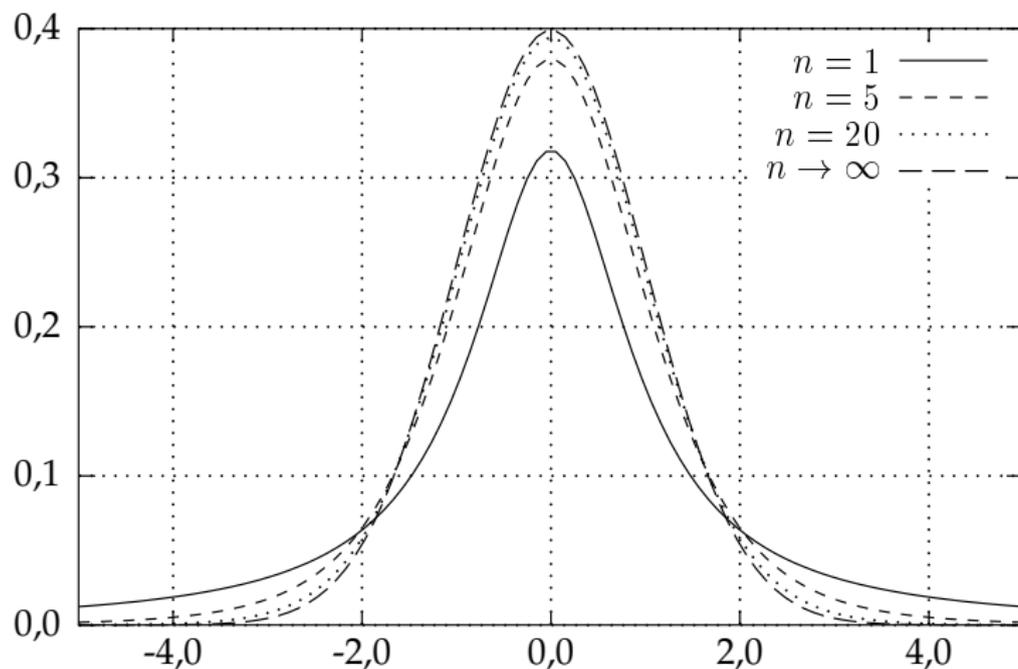
Die  $t$ -Verteilung nennt man auch **Student-Verteilung**.

Es handelt sich eigentlich um eine Familie von Verteilungen, eine für jede Anzahl von Freiheitsgraden.

Die Dichte ähnelt der Dichte der Normalverteilung.

Für große  $n$  (Faustregel:  $n \geq 30$ ) wird in der Praxis die  $t$ -Verteilung durch die Normalverteilung angenähert.

# Ein-Stichproben-Tests für Lageparameter V: $t$ -Test



Dichte der  $t$ -Verteilung mit  $n$  Freiheitsgraden

# Zwei-Stichproben-Tests für Lageparameter I: $t$ -Test

## Tabelle: Zwei-Stichproben- $t$ -Test

### Annahmen:

$X_1, \dots, X_m$  und  $Y_1, \dots, Y_n$  seien unabhängig und jeweils identisch verteilt, wobei  $X_i \sim \mathcal{N}(\mu_X, \sigma_X^2)$  und  $Y_i \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$  gelte.  
Die Varianzen seien identisch, also  $\sigma_X^2 = \sigma_Y^2$ .

### Hypothesen:

- a)  $H_0 : \mu_X = \mu_Y$  gegen  $H_1 : \mu_X \neq \mu_Y$ ,  
b)  $H_0 : \mu_X \geq \mu_Y$  gegen  $H_1 : \mu_X < \mu_Y$ ,  
c)  $H_0 : \mu_X \leq \mu_Y$  gegen  $H_1 : \mu_X > \mu_Y$ .

### Testgröße:

$$T := \sqrt{\frac{n+m-2}{\frac{1}{n} + \frac{1}{m}}} \cdot \frac{\bar{X} - \bar{Y}}{\sqrt{(m-1) \cdot S_X^2 + (n-1) \cdot S_Y^2}}.$$

### Ablehnungskriterium für $H_0$ bei Signifikanzniveau $\alpha$ :

- a)  $|T| > t_{m+n-2, 1-\alpha/2}$ ,  
b)  $T < t_{m+n-2, \alpha}$ ,  
c)  $T > t_{m+n-2, 1-\alpha}$ .

# Zwei-Stichproben-Tests für Lageparameter II

## Beispiel 171

Im SS 2012 nahmen 48 Studierende an die Evaluierung der DWT-Vorlesung teil. Sie beurteilten die Vorlesung 21x mit der Note 1, 21x mit der Note 2 und 6x mit der Note 3.

Damit ergibt sich für die Zufallsvariable  $X$ , die die Note darstellt,  $\bar{X} \approx 1.69$  und  $\overline{S}_X \approx 0.68$ .

# Zwei-Stichproben-Tests für Lageparameter II

## Beispiel 171

Im SS 2012 nahmen 48 Studierende an die Evaluierung der DWT-Vorlesung teil. Sie beurteilten die Vorlesung 21x mit der Note 1, 21x mit der Note 2 und 6x mit der Note 3.

Damit ergibt sich für die Zufallsvariable  $X$ , die die Note darstellt,  $\bar{X} \approx 1.69$  und  $\bar{S}_X \approx 0.68$ .

Im SS 2013 nahmen 61 Studierende teil. Sie beurteilten die Vorlesung 20x mit der Note 1, 33x mit der Note 2, 7x Mal mit der Note 3 und 1x mit der Note 4.

Damit ergibt sich für die Zufallsvariable  $Y$ , die die Note darstellt,  $\bar{Y} \approx 1.82$  und  $\bar{S}_Y \approx 0.69$ .

# Zwei-Stichproben-Tests für Lageparameter II

## Beispiel 171

Im SS 2012 nahmen 48 Studierende an die Evaluierung der DWT-Vorlesung teil. Sie beurteilten die Vorlesung 21x mit der Note 1, 21x mit der Note 2 und 6x mit der Note 3.

Damit ergibt sich für die Zufallsvariable  $X$ , die die Note darstellt,  $\bar{X} \approx 1.69$  und  $\bar{S}_X \approx 0.68$ .

Im SS 2013 nahmen 61 Studierende teil. Sie beurteilten die Vorlesung 20x mit der Note 1, 33x mit der Note 2, 7x Mal mit der Note 3 und 1x mit der Note 4.

Damit ergibt sich für die Zufallsvariable  $Y$ , die die Note darstellt,  $\bar{Y} \approx 1.82$  und  $\bar{S}_Y \approx 0.69$ .

Prof. Evilsparza ist vom schlechteren Ergebnis im 2013 schwer beleidigt. Er erwägt, alle Teilnehmer der Klausur eine 5 zu verpassen. Dr. Luttenberger redet ihn jedoch ein, dass er erst die statistische Signifikanz der Ergebnisse prüfen soll.

# Zwei-Stichproben-Tests für Lageparameter II

## Beispiel 171

Im SS 2012 nahmen 48 Studierende an die Evaluierung der DWT-Vorlesung teil. Sie beurteilten die Vorlesung 21x mit der Note 1, 21x mit der Note 2 und 6x mit der Note 3.

Damit ergibt sich für die Zufallsvariable  $X$ , die die Note darstellt,  $\bar{X} \approx 1.69$  und  $\bar{S}_X \approx 0.68$ .

Im SS 2013 nahmen 61 Studierende teil. Sie beurteilten die Vorlesung 20x mit der Note 1, 33x mit der Note 2, 7x Mal mit der Note 3 und 1x mit der Note 4.

Damit ergibt sich für die Zufallsvariable  $Y$ , die die Note darstellt,  $\bar{Y} \approx 1.82$  und  $\bar{S}_Y \approx 0.69$ .

Prof. Evilsparza ist vom schlechteren Ergebnis im 2013 schwer beleidigt. Er erwägt, alle Teilnehmer der Klausur eine 5 zu verpassen. Dr. Luttenberger redet ihn jedoch ein, dass er erst die statistische Signifikanz der Ergebnisse prüfen soll.

**Frage:** Kann Prof. Evilsparza mit Signifikanzniveau von  $\alpha = 0.05$  behaupten, dass die Studierenden vom SS 2013 eine schlechtere Meinung haben?

## Zwei-Stichproben-Tests für Lageparameter III

Seien  $X_1, \dots, X_{48}$  und  $Y_1, \dots, Y_{61}$  jeweils unabhängig und identisch verteilt. Wir definieren

$$\begin{aligned} X'_1 &= (X_1 + \dots + X_6)/6 & \dots & & X'_8 &= (X_{43} + \dots + X_{48})/6 \\ Y'_1 &= (Y_1 + \dots + Y_6)/6 & \dots & & Y'_{10} &= (Y_{55} + \dots + X_{60})/6 \end{aligned}$$

und nehmen an, dass

$$X'_i \sim \mathcal{N}(\mu_{X'}, \sigma^2) \quad \text{und} \quad \mu_{Y'} \sim \mathcal{N}(\mu_{Y'}, \sigma^2)$$

für unbekannte  $\mu_{X'}, \mu_{Y'}, \sigma^2$  gilt.

## Zwei-Stichproben-Tests für Lageparameter III

Seien  $X_1, \dots, X_{48}$  und  $Y_1, \dots, Y_{61}$  jeweils unabhängig und identisch verteilt. Wir definieren

$$\begin{aligned} X'_1 &= (X_1 + \dots + X_6)/6 & \dots & & X'_8 &= (X_{43} + \dots + X_{48})/6 \\ Y'_1 &= (Y_1 + \dots + Y_6)/6 & \dots & & Y'_{10} &= (Y_{55} + \dots + X_{60})/6 \end{aligned}$$

und nehmen an, dass

$$X'_i \sim \mathcal{N}(\mu_{X'}, \sigma^2) \quad \text{und} \quad \mu_{Y'} \sim \mathcal{N}(\mu_{Y'}, \sigma^2)$$

für unbekanntes  $\mu_{X'}, \mu_{Y'}, \sigma^2$  gilt.

Wir haben

$$\overline{X'} = \overline{X} \quad \overline{Y'} = \overline{Y}$$

und sowie

$$S_{X'} = \frac{1}{\sqrt{6}} S_X \quad S_{Y'} = \frac{1}{\sqrt{6}} S_Y$$

## Zwei-Stichproben-Tests für Lageparameter IV

Wir wählen als Nullhypothese

$$H_0 : \mu_{X'} = \mu_{Y'}$$

(identischen Verteilungen in beiden Semestern).

## Zwei-Stichproben-Tests für Lageparameter IV

Wir wählen als Nullhypothese

$$H_0 : \mu_{X'} = \mu_{Y'}$$

(identischen W'keitsverteilungen in beiden Semestern).

Für die Testgröße ergibt sich

$$T = \sqrt{\frac{8 + 10 - 2}{1/8 + 1/10}} \cdot \frac{1.69 - 1.82}{\sqrt{7 \cdot 0.68^2/6 + 9 \cdot 0.69^2/6}} \approx -0.861$$

## Zwei-Stichproben-Tests für Lageparameter IV

Wir wählen als Nullhypothese

$$H_0 : \mu_{X'} = \mu_{Y'}$$

(identischen Verteilungen in beiden Semestern).

Für die Testgröße ergibt sich

$$T = \sqrt{\frac{8 + 10 - 2}{1/8 + 1/10}} \cdot \frac{1.69 - 1.82}{\sqrt{7 \cdot 0.68^2/6 + 9 \cdot 0.69^2/6}} \approx -0.861$$

Aus einer Tabelle entnehmen wir

$$t_{m+n-2, 1-\alpha/2} = t_{16, 0.975} = 2.120 > 0.861 = |T|$$

Damit kann  $H_0$  nicht abgelehnt werden.

## Zwei-Stichproben-Tests für Lageparameter IV

Wir wählen als Nullhypothese

$$H_0 : \mu_{X'} = \mu_{Y'}$$

(identischen W'keitsverteilungen in beiden Semestern).

Für die Testgröße ergibt sich

$$T = \sqrt{\frac{8 + 10 - 2}{1/8 + 1/10}} \cdot \frac{1.69 - 1.82}{\sqrt{7 \cdot 0.68^2/6 + 9 \cdot 0.69^2/6}} \approx -0.861$$

Aus einer Tabelle entnehmen wir

$$t_{m+n-2, 1-\alpha/2} = t_{16, 0.975} = 2.120 > 0.861 = |T|$$

Damit kann  $H_0$  nicht abgelehnt werden.

Für  $\alpha = 0.4$  gilt

$$t_{m+n-2, 1-\alpha/2} = t_{16, 0.40} = 0.865 > 0.861 = |T|$$

und so kann  $H_0$  wieder nicht abgelehnt werden (aber fast).

# Tests auf Verteilungen I: $\chi^2$ -Anpassungstest

## Tabell: $\chi^2$ -Anpassungstest

### Annahmen:

$X_1, \dots, X_n$  seien unabhängig und identisch verteilt mit  $W_{X_i} = \{1, \dots, k\}$ .

### Hypothesen:

$$H_0 : \Pr[X = i] = p_i \quad \text{für } i = 1, \dots, k,$$

$$H_1 : \Pr[X = i] \neq p_i \quad \text{für mindestens ein } i \in \{1, \dots, k\},$$

### Testgröße:

$$T = \sum_{i=1}^k \frac{(h_i - np_i)^2}{np_i},$$

wobei  $h_i$  die Häufigkeit angibt, mit der  $X_1, \dots, X_n$  den Wert  $i$  angenommen haben.

### Ablehnungskriterium für $H_0$ bei Signifikanzniveau $\alpha$ :

$$T > \chi_{k-1, 1-\alpha}^2;$$

dabei sollte gelten, dass  $np_i \geq 1$  für alle  $i$  und  $np_i \geq 5$  für mindestens 80% der Werte  $i = 1, \dots, k$ .

## Tests auf Verteilungen II: $\chi^2$ -Anpassungstest

Für die Testgröße  $T$  wird näherungsweise eine

*$\chi^2$ -Verteilung mit  $k - 1$  Freiheitsgraden*

angenommen.

Das  $\gamma$ -Quantil einer  $\chi^2$ -Verteilung mit  $k$  Freiheitsgraden bezeichnen wir mit  $\chi_{k,\gamma}^2$ .

## Tests auf Verteilungen II: $\chi^2$ -Anpassungstest

Für die Testgröße  $T$  wird näherungsweise eine

*$\chi^2$ -Verteilung mit  $k - 1$  Freiheitsgraden*

angenommen.

Das  $\gamma$ -Quantil einer  $\chi^2$ -Verteilung mit  $k$  Freiheitsgraden bezeichnen wir mit  $\chi_{k,\gamma}^2$ .

Die Werte dieser Verteilung finden sich in entsprechenden Tabellen in der Literatur.

## Tests auf Verteilungen II: $\chi^2$ -Anpassungstest

Für die Testgröße  $T$  wird näherungsweise eine

*$\chi^2$ -Verteilung mit  $k - 1$  Freiheitsgraden*

angenommen.

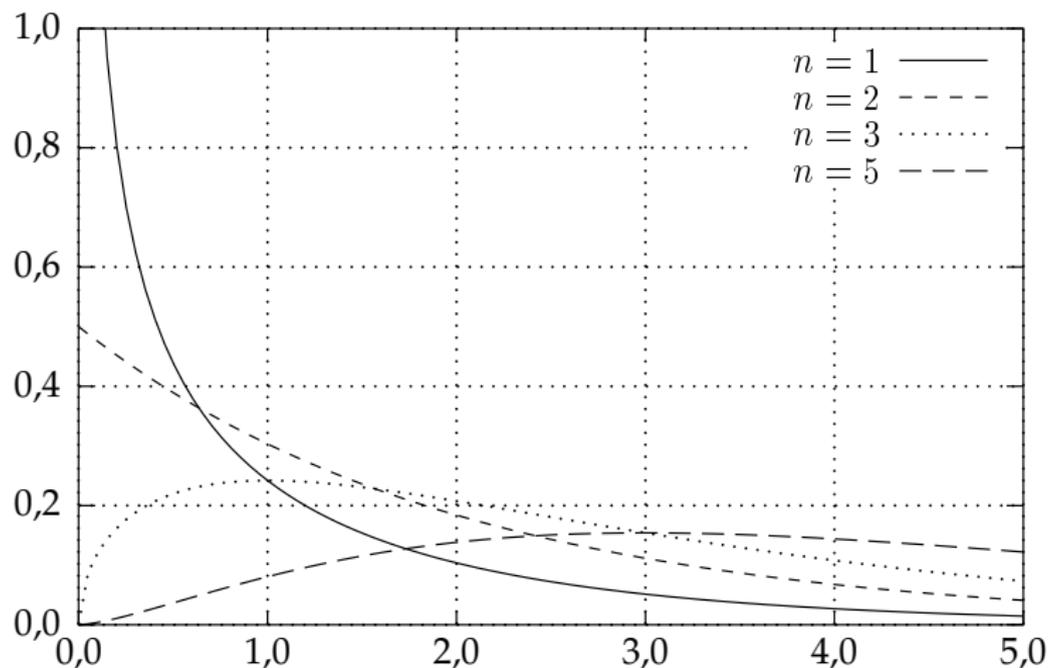
Das  $\gamma$ -Quantil einer  $\chi^2$ -Verteilung mit  $k$  Freiheitsgraden bezeichnen wir mit  $\chi_{k,\gamma}^2$ .

Die Werte dieser Verteilung finden sich in entsprechenden Tabellen in der Literatur.

Damit die Approximation gerechtfertigt ist, sollte gelten (Faustregel)

- $np_i \geq 1$  für alle  $i$  und
- $np_i \geq 5$  für mindestens 80% der Werte  $i = 1, \dots, k$ .

# Tests auf Verteilungen III: $\chi^2$ -Anpassungstest



Dichte der  $\chi^2$ -Verteilung mit  $n$  Freiheitsgraden

# Tests auf Verteilungen IV: $\chi^2$ -Anpassungstest

## Beispiel 172

Wir wollen überprüfen, ob der Zufallszahlengenerator von Maple eine gute Approximation der Gleichverteilung liefert.

# Tests auf Verteilungen IV: $\chi^2$ -Anpassungstest

## Beispiel 172

Wir wollen überprüfen, ob der Zufallszahlengenerator von Maple eine gute Approximation der Gleichverteilung liefert.

Dazu lassen wir Maple  $n = 100000$  Zufallszahlen aus der Menge  $\{1, \dots, 10\}$  generieren.

# Tests auf Verteilungen IV: $\chi^2$ -Anpassungstest

## Beispiel 172

Wir wollen überprüfen, ob der Zufallszahlengenerator von Maple eine gute Approximation der Gleichverteilung liefert.

Dazu lassen wir Maple  $n = 100000$  Zufallszahlen aus der Menge  $\{1, \dots, 10\}$  generieren.

Die Nullhypothese lautet  $p_1 = \dots = p_{10} = 1/10$ , wobei  $p_i$  die W'keit von  $i$  bezeichnet.

# Tests auf Verteilungen IV: $\chi^2$ -Anpassungstest

## Beispiel 172

Wir wollen überprüfen, ob der Zufallszahlengenerator von Maple eine gute Approximation der Gleichverteilung liefert.

Dazu lassen wir Maple  $n = 100000$  Zufallszahlen aus der Menge  $\{1, \dots, 10\}$  generieren.

Die Nullhypothese lautet  $p_1 = \dots = p_{10} = 1/10$ , wobei  $p_i$  die W'keit von  $i$  bezeichnet.

Die Nullhypothese soll mit Signifikanzniveau von  $\alpha = 0,05$  getestet werden.

# Tests auf Verteilungen V: $\chi^2$ -Anpassungstest

Beispiel:

$i$	1	2	3	4	5	6	7	8	9	10
$h_i$	10102	10070	9972	9803	10002	10065	10133	9943	10009	9901

# Tests auf Verteilungen V: $\chi^2$ -Anpassungstest

Beispiel:

$i$	1	2	3	4	5	6	7	8	9	10
$h_i$	10102	10070	9972	9803	10002	10065	10133	9943	10009	9901

Es gilt  $T = 8,9946$  und  $\chi_{9,0,95}^2 \approx 16,919$ .

Die Nullhypothese wird nicht abgelehnt.

# Tests auf Verteilungen V: $\chi^2$ -Anpassungstest

Beispiel:

$i$	1	2	3	4	5	6	7	8	9	10
$h_i$	10102	10070	9972	9803	10002	10065	10133	9943	10009	9901

Es gilt  $T = 8,9946$  und  $\chi_{9,0,95}^2 \approx 16,919$ .

Die Nullhypothese wird nicht abgelehnt.

Was bedeutet das?

# Tests auf Verteilungen V: $\chi^2$ -Anpassungstest

Beispiel:

$i$	1	2	3	4	5	6	7	8	9	10
$h_i$	10102	10070	9972	9803	10002	10065	10133	9943	10009	9901

Es gilt  $T = 8,9946$  und  $\chi_{9,0,95}^2 \approx 16,919$ .

Die Nullhypothese wird nicht abgelehnt.

Was bedeutet das?

Nicht viel! Der Test garantiert nur, dass die W'keit, die Nullhypothese **irrtümlich abzulehnen**, höchstens **5%** beträgt. Die Nullhypothese ist jedoch nicht abgelehnt worden!