

# Grundlagen der Mathematischen Statistik

## Punktschätzungen

Uwe Menzel, 2018

uwe.menzel@matstat.org

[www.matstat.org](http://www.matstat.org)

1

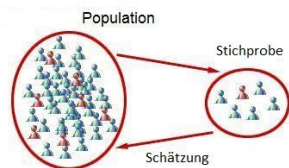
## Wahrscheinlichkeitstheorie und Statistiktheorie

### Wahrscheinlichkeitstheorie:

- was bisher in der Vorlesung behandelt wurde
- Es lag eine bekannte Verteilung vor, z. B.  $X \sim N(\mu, \sigma)$ , mit bekanntem  $\mu$  und  $\sigma$
- davon ausgehend konnten z. B. Erwartungswert, Varianz und Wahrscheinlichkeiten berechnet werden, z. B.  $P(a < X \leq b)$

### Statistiktheorie:

- eher realistisch
- **Verteilungsparameter  $\mu$  und  $\sigma$  unbekannt** (aber möglicherweise ist der Typ der Verteilung bekannt, also z. B.  $X \sim N$ )
- **Die Verteilungsparameter  $\mu$  und  $\sigma$  müssen auf der Basis einer Stichprobe geschätzt werden** → **Punktschätzung, Intervallschätzung (Inferenz)**



In einer **zufälligen** Stichprobe müssen alle Messwerte unabhängig sein.

[www.matstat.org](http://www.matstat.org)

2

## Intuitive Schätzer

### Normalverteilung:

Schätzung von  $\mu$  und  $\sigma$  für die **Normalverteilung** (es muss also bekannt sein, dass Normalverteilung vorliegt).

$X \sim N(\mu, \sigma)$  aber  $\mu$  und/oder  $\sigma$  unbekannt

**Stichprobe:**  $x = (x_1, x_2, \dots, x_n)$

$$\mu^* = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}$$

Schätzung für den **Erwartungswert**  $\mu = E(X)$   
einer normalverteilten Zufallsvariable  
 $X \sim N(\mu, \sigma)$

**Anmerkung:** eine Schätzung wird oft mit einem Stern gekennzeichnet, z. B.  $\mu^*$

$$(\sigma^2)^* = S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

Schätzung für die **Varianz**  $\sigma^2 = V(X)$   
einer normalverteilten  
Zufallsvariable  $X \sim N(\mu, \sigma)$

www.matstat.org

3

## Intuitive Schätzer

### Binomialverteilung:

Schätzung des Parameters  $p$  für eine **Binomialverteilung**. Es muss also bekannt sein dass Binomialverteilung vorliegt (z. B. Münzwurf).

$X \sim \text{Bin}(n, p)$  ... aber  $p$  unbekannt

**Stichprobe:**  $x$  = Anzahl "geglückter" Versuche;  
 $n$  = totale Anzahl der (Bernoulli-) Versuche

$$p^* = \frac{x}{n}$$

Schätzung für die "Erfolgs"-Wahrscheinlichkeit  
 $p$  für  $\text{Bin}(n, p)$

**Beispiel 1:** 10 Münzwürfe ( $n = 10$ ), davon 4 "erfolgreiche" ( $x = 4$ )



$$p^* = \frac{x}{n} = \frac{4}{10} = 0.4 \quad \text{Achtung!}: \text{vielleicht war } n \text{ zu klein?}$$

**Beispiel 2.** Meinungsumfrage: 1000 Personen werden gefragt, ob sie die Partei "Die Besten" wählen würden. 3011 antworten mit "Ja". Wir können den Anteil der "Die Besten"-Wähler in der Gesamtpopulation(!) auf ungefähr 30% schätzen.

www.matstat.org

4

## Schätzer als Zufallsvariablen

Stichprobe 1:  $x^1 = (x_1^1, x_2^1, \dots, x_n^1)$   
 Stichprobe 2:  $x^2 = (x_1^2, x_2^2, \dots, x_n^2)$   
 Stichprobe 3:  $x^3 = (x_1^3, x_2^3, \dots, x_n^3)$

alle Stichproben ergeben verschiedene  $\bar{x}$  und  $S^2 \rightarrow$  **Schätzer müssen als Zufallsvariablen behandelt werden** (meistens mit großen Buchstaben bezeichnet)!

- verschiedene numerische Werte für verschiedene Stichproben  $\rightarrow$  Schätzungen sind zufallsbehaftet
- die (unabhängigen) Observations  $x_1, x_2, \dots, x_n$  werden als Realisationen unabhängiger Zufallsvariablen  $X_1, X_2, \dots, X_n$  angesehen, die alle von der gleichen Verteilung erzeugt werden
- ebenso wird der konkrete numerische Wert  $\bar{x}$  für eine einzelne Schätzung als **Realisation einer Zufallsvariable  $\bar{X}$**  betrachtet

$$\underbrace{\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i}_{\text{Stichprobenwert}} \quad \Rightarrow \quad \underbrace{\bar{X} = \frac{1}{n} \cdot \sum_{i=1}^n X_i}_{\text{Zufallsvariable}}$$

www.matstat.org

5

## Schätzer als Zufallsvariablen

Stichprobe 1:  $x^1 = (x_1^1, x_2^1, \dots, x_n^1)$

Stichprobe 2:  $x^2 = (x_1^2, x_2^2, \dots, x_n^2)$

Stichprobe 3:  $x^3 = (x_1^3, x_2^3, \dots, x_n^3)$

$$\underbrace{\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i}_{\text{Stichprobenwert}} \quad \Rightarrow \quad \underbrace{\bar{X} = \frac{1}{n} \cdot \sum_{i=1}^n X_i}_{\text{Zufallsvariable}}$$

Beispiel Bin:  $p^* = \frac{x}{n} \quad \Rightarrow \quad p^* = \frac{X}{n} \quad X \sim \text{Bin}(n, p)$

$E(p^*), V(p^*), \dots$  können berechnet werden

www.matstat.org

6

## Schätzungen

Observationen  $x_1, x_2, \dots, x_3 \rightarrow$  Zufallsvariablen  $X_1, X_2, \dots, X_3$

Stichprobenresultat	Zufallsvariable	Verteilung
$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$	$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$	$X_i \sim N(\mu, \sigma)$
$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$	$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$	$X_i \sim N(\mu, \sigma)$
$p^* = \frac{x}{n}$	$p^* = \frac{X}{n}$	$X \sim Bin(n, p)$

www.matstat.org

7

## Eigenschaften von Schätzungen

Ein guter Schätzer sollte verschiedene Bedingungen erfüllen :

- Erwartungstreue
- Konsistenz
- Effektivität

Mit Hilfe dieser Eigenschaften kann man die Qualität eines Schätzers beurteilen.

www.matstat.org

8

## Erwartungstreue

$\theta^*$  = Schätzung für  $\theta$       $\theta$  steht für einen Verteilungsparameter ( $\mu, \sigma, p, \dots$ )

Wenn Erwartungstreue vorliegt, muss  $E(\theta^*) = \theta$

d.h. der Erwartungswert des Schätzers muss mit dem zu schätzenden (formalen) Parameter der zugeordneten Zufallsvariable übereinstimmen.

Ex1: Schätzung von Erwartungswert und Varianz für eine Normalverteilung:

$$E(\mu^*) = E(\bar{X}) = E\left[\frac{1}{n} \sum_{i=1}^n X_i\right] = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{n} \cdot n \cdot \mu = \mu \quad \checkmark$$

Der wahre Werte von  $\mu$  ist unbekannt, wir wissen jedoch, dass der Erwartungswert des Schätzers mit dem wahren  $\mu$  übereinstimmt!

$$E(S^2) = \dots = \sigma^2 \quad \checkmark \quad (\text{genau deshalb müssen wir } n - 1 \text{ im Nenner von } S^2 \text{ haben!})$$

Ex2: Schätzung der "Erfolgs"- Wahrscheinlichkeit  $p$  für Binomialverteilung:

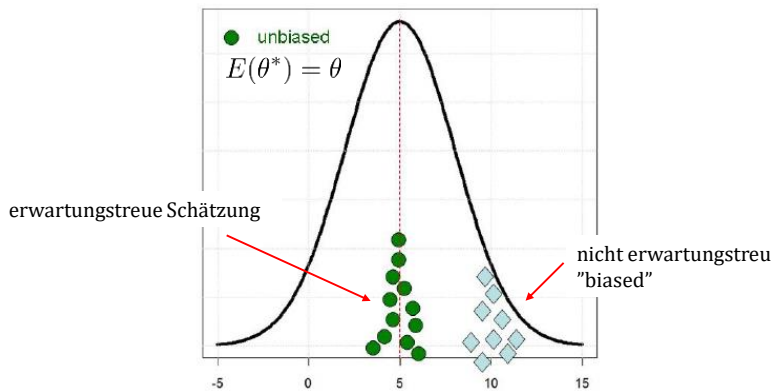
$$E(p^*) = E\left(\frac{X}{n}\right) = \frac{1}{n} \cdot E(X) = \frac{1}{n} \cdot n \cdot p = p \quad \checkmark$$

$\leftarrow E(X)$  für  $\text{Bin}(n, p)$

[www.matstat.org](http://www.matstat.org)

9

## Erwartungstreue



- Eine erwartungstreue Schätzung  $\theta^*$  ist um den wahren Wert  $\theta$  herum konzentriert, der Erwartungswert aller Schätzungen liegt beim wahren Wert  $\theta$ .
- Der Mittelwert einer erwartungstreuen Schätzung ergibt den wahren Wert mit desto größerer Genauigkeit, je mehr Messwerte ermittelt werden.

[www.matstat.org](http://www.matstat.org)

10

10

## Konsistenz

$\theta_n^*$  = Schätzung aufgrund von  $n$  Observationen (Messwerten)

$$\lim_{n \rightarrow \infty} V(\theta_n^*) = 0$$

Je größer  $n$ , desto kleiner soll die Varianz des Schätzers werden. Für  $n \rightarrow \infty$  soll die Varianz gegen Null gehen.

Wenn eine Schätzung konsistent ist, lohnt es sich, mehr Observationen zu machen, denn dafür wird die Varianz des Schätzers – also die Unsicherheit der Schätzung – kleiner. Die Schätzung liegt dann mit größerer Wahrscheinlichkeit in der Nähe des wahren Wertes.

Ex 1: Schätzung für Erwartungswert und Varianz für die Normalverteilung:

$$V(\mu^*) = V(\bar{X}) = V\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n V(X_i) = \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n}$$

Klingt das komisch?  $\mu^*$  ist eine Zufallsvariable, hat also eine Varianz!  $V(\mu^*)$  ist die Varianz des Schätzers für den Mittelwert.

$$\lim_{n \rightarrow \infty} V(\mu^*) = \lim_{n \rightarrow \infty} \frac{\sigma^2}{n} = 0 \quad \checkmark$$

$$\lim_{n \rightarrow \infty} V(S^2) = \dots = 0 \quad \checkmark$$

www.matstat.org

11

## Konsistenz

Ex 2: Schätzung für die "Erfolgs"-Wahrscheinlichkeit  $p$  in  $B(n, p)$ :

$$p^* = \frac{X}{n} \quad \text{Schätzung für } p \text{ in } \text{Bin}(n, p) \quad X \sim \text{Bin}(n, p)$$

$$V(p^*) = V\left(\frac{X}{n}\right) = \frac{1}{n^2} V(X) = \frac{1}{n^2} \cdot n \cdot p \cdot (1-p) = \frac{p \cdot (1-p)}{n}$$

$$\lim_{n \rightarrow \infty} V(p^*) = \lim_{n \rightarrow \infty} \frac{p \cdot (1-p)}{n} = 0 \quad \checkmark$$

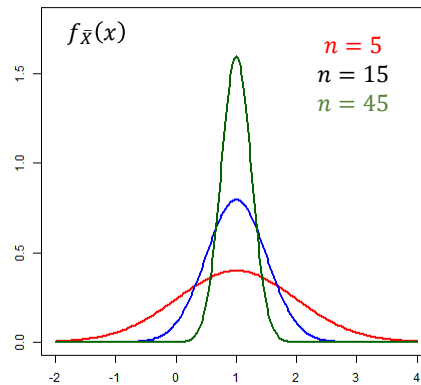
**Zusammenfassung:** Die intuitiven Schätzer für  $\mu, \sigma$  in  $N(\mu, \sigma)$  bzw. für  $p$  in  $\text{Bin}(n, p)$  sind erwartungstreu und konsistent.

www.matstat.org

12

## Konsistenz

Dichtefunktion für den Schätzer  $\bar{X}$



Die Schätzungen konzentrieren sich mehr und mehr um den wahren Wert wenn die Stichprobe größer und größer wird, d.h. wenn  $n$  wächst.

www.matstat.org

13

## Effektivität

Ein Schätzer sollte eine möglichst kleine Varianz haben, denn größere Varianz bedeutet größere Unsicherheit. Wenn man zwischen verschiedenen Schätzern wählen kann, sollte man denjenigen mit der kleinsten Varianz anwenden, also den am **effektivsten**.

Beispiel: Stichprobe  $x = (x_1, x_2, \dots, x_{10})$  mit  $X \sim N(\mu, \sigma)$

$\mu_1^* = \bar{X}_{10}$  **Schätzer Nr. 1:** arithmetischer Mittelwert aller 10 Messwerte.

$$V(\mu_1^*) = V(\bar{X}_{10}) = \frac{\sigma^2}{10}$$

$\mu_2^* = \frac{X_1 + X_{10}}{2}$  **Schätzer Nr. 2:** Mittelwert des größten und kleinsten Wertes

$$V(\mu_2^*) = V\left(\frac{X_1 + X_{10}}{2}\right) = \frac{1}{4} [V(X_1) + V(X_{10})] = \frac{1}{4} \cdot 2\sigma^2 = \frac{\sigma^2}{2}$$

$V(\mu_1^*) < V(\mu_2^*) \Rightarrow \mu_1^*$  ist effektiver

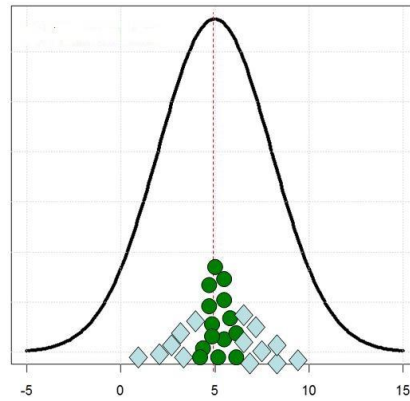
Am besten: **MVUE** = Minimum Variance Unbiased Estimator

www.matstat.org

14

## Effektivität

$\theta^*$  ist effektiver als  $\hat{\theta}$  wenn  $V(\theta^*) < V(\hat{\theta})$



● effektiver

◆ weniger effektiv

www.matstat.org

15

## Methoden zum Finden von Schätzern

( ... also einer auf die Messwerte anzuwendenden Formel)

- a) Momentenmethode
- b) Methode der kleinsten Quadrate
- c) Maximum-Likelihood-Methode
- d) .....

www.matstat.org

16



## Momentenmethode

Das **k-te Moment** einer Zufallsvariable  $X$  ist laut Definition der Erwartungswert der  $k$ -ten Potenz dieser Zufallsvariable:

$$m_k = E(X^k)$$

Für die Normalverteilung gilt z.B.: → Vorlesung **F3**:  $V(X) = E(X^2) - E(X)^2$

$$m_1 = E(X) = \mu \quad m_2 = E(X^2) = V(X) + E(X)^2 = \sigma^2 + \mu^2$$

Die Momente sind somit Parameter, die die Verteilung beschreiben; hat man alle Momente, so ist die Verteilung eindeutig beschrieben. Analog kann eine Stichprobe durch die sog. **empirischen Momente** beschrieben werden:

$$\hat{m}_k = \frac{1}{n} \cdot \sum_{i=1}^n x_i^k$$

Man nimmt nun an, dass die theoretischen Momente einer Zufallsvariable mit Hilfe der entsprechenden Momente einer Stichprobe geschätzt werden können  
→ **Momentenmethode**.

www.matstat.org

17

## Momentenmethode

1. Moment der Verteilung = 1. Moment der Stichprobe

$$E(X; \theta) = \frac{1}{n} \cdot \sum_{i=1}^n x_i$$

1. Moment der Zufallsvariable  
(welches von einem Parameter  $\theta$   
abhängt)

1. Moment der Stichprobe

Beispiel:  $Bin(n, p)$

$$n \cdot p = x_1$$

Parameter  $p$  soll  
geschätzt werden

nur ein Messwert  
(Anzahl der "geglückten"  
Bernoulli-Versuche)

$$\Rightarrow p^* = \frac{x_1}{n} \quad \text{Schätzung für } p \text{ in } Bin(n, p)$$

www.matstat.org

18

## Methode der kleinsten Quadrate

Man bestimmt den Verteilungsparameter  $\theta$  so, dass die Summe der quadratischen Abstände der Messwerte zum Erwartungswert minimiert wird (vergleiche Vorlesung lineare Regression, **F10**)

$$Q(\theta) = \sum_{i=1}^N (x_i - E(X; \theta))^2 \rightarrow \text{Min}$$

$x_i$ : Stichprobe  
 $E(X)$  hängt von einem Verteilungsparameter  $\theta$  ab

Beispiel:  $\text{Bin}(n, p)$

$$X \sim \text{Bin}(n, p) \quad p \text{ soll geschätzt werden} \quad E(X) = n \cdot p$$

nur ein Wert in der Stichprobe:  $x_1$

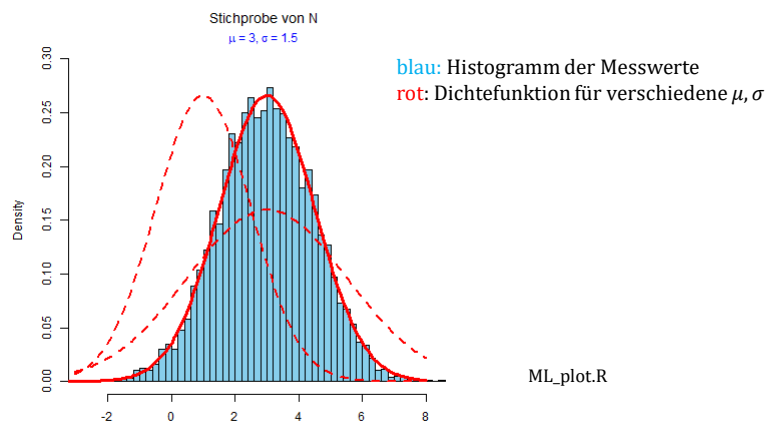
$$Q(p) = (x_1 - n \cdot p)^2 \rightarrow \text{Min}$$

$$\frac{dQ}{dp} = -2n \cdot (x_1 - n \cdot p) = 0 \rightarrow p = \frac{x_1}{n}$$

www.matstat.org

19

## Maximum-Likelihood-Methode



**Idee:** wähle die Parameter  $\mu, \sigma$  so dass die Messwerte möglichst plausibel werden  $\rightarrow$  wähle die Parameter welche die durchgezogene rote Linie ergeben.

www.matstat.org

20

## Maximum-Likelihood-Methode

**Likelihood:** Funktion eines Verteilungsparameters      Stichprobe:  
 $x = (x_1, x_2, \dots, x_n)$

$$L(\theta) = \begin{cases} p_X(x_1) \cdot p_X(x_2) \cdot \dots \cdot p_X(x_n) & \text{diskret} \\ f_X(x_1) \cdot f_X(x_2) \cdot \dots \cdot f_X(x_n) & \text{kontinuierlich} \end{cases}$$

Beispiel: Normalverteilung       $x_i$  bekannt (Messwerte)

$$L(\mu, \sigma) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i - \mu)^2}{2\sigma^2}} = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\sum (x_i - \mu)^2 / 2\sigma^2}$$

Bestimme nun  $\mu$  und  $\sigma$  so dass  $L(\mu, \sigma)$  maximiert wird:

$$\frac{\partial L}{\partial \mu} = 0 \quad \frac{\partial L}{\partial \sigma} = 0 \quad \text{ergibt } \mu \text{ und } \sigma \text{ als Funktion}$$

der Messwerte  $x_i$ , also den  
**Schätzer**

www.matstat.org

21

## Parametrische Punktschätzer

Verteilung	Parameter	Schätzer für den Parameter	Methode
$X \sim N(\mu, \sigma)$	$\mu$	$\mu^* = \bar{X}$	ML, MK, MM
$X \sim N(\mu, \sigma)$	$\sigma^2$	$S^2 = \frac{1}{n-1} S_{xx}$	MM, ML
$X \sim N(\mu, \sigma)$	$\sigma$	$S = \sqrt{S^2}$	ML
$X \sim Po(\mu)$	$\mu$	$\mu^* = \bar{X}$	MM, MK, ML
$X \sim Bin(n, p)$	$p$	$p^* = \frac{X}{n}$	ML
$X \sim Hyp(N, n, m)$	$p$	$p^* = \frac{X}{n}$	MM, MK, ML
$X \sim Exp(\lambda)$	$\lambda$	$\lambda^* = \frac{1}{\bar{X}}$	ML
$X \sim U(0, \theta)$	$\theta$	$\theta^* = \frac{n+1}{n} \cdot \max(X_i)$	ML

$$\bar{X} = \sum_{i=1}^n X_i$$

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

www.matstat.org

22

## Parametrische Punktschätzer, Forts.

Verteilung	Parameter	Schätzung für den Parameter	Methode
$X \sim N(\mu_x, \sigma)$ $Y \sim N(\mu_y, \sigma)$	$\mu_x$	$\mu_x^* = \bar{X}$	ML
	$\mu_y$	$\mu_y^* = \bar{Y}$	ML
	$\sigma^2$	$S^2 = \frac{S_{xx} + S_{yy}}{(n_x - 1) + (n_y - 1)}$	ML
$X_1 \sim \text{Bin}(n_1, p)$ $X_2 \sim \text{Bin}(n_2, p)$	$p$	$p^* = \frac{X_1 + X_2}{n_1 + n_2}$	ML
$X_1 \sim \text{Po}(\lambda \cdot t_1)$ $X_2 \sim \text{Po}(\lambda \cdot t_2)$	$\lambda$	$\lambda^* = \frac{X_1 + X_2}{t_1 + t_2}$	ML

$$\bar{X} = \sum_{i=1}^n X_i$$

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

www.matstat.org

23

## Standardfehler

Der Standardfehler ist eine **Approximation für die Standardabweichung eines Schätzers** (größerer Standardfehler → größere Unsicherheit).

### 1. Standardfehler für Schätzung von $\mu$ in $N(\mu, \sigma)$ :

$$\mu^* = \frac{1}{n} \sum_{i=1}^n x_i = \bar{X} \quad \text{Schätzung für } \mu \text{ in } N(\mu, \sigma)$$

$$V(\bar{X}) = V\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n V(X_i) = \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n} \quad \text{Varianz des Schätzers}$$

$$D(\bar{X}) = \frac{\sigma}{\sqrt{n}} \quad \text{Standardabweichung des Schätzers von } \mu, \text{ aber } \sigma \text{ unbekannt!}$$



$\sigma \rightarrow s$  ( $\sigma$  wird durch seine Schätzung  $s$  ersetzt)

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

$$d(\bar{X}) = \frac{s}{\sqrt{n}} \quad \text{Standardfehler, ein numerischer Wert kann berechnet werden, Approximation für die Standardabweichung (für eine Schätzung von } \mu \text{ in } N(\mu, \sigma) \text{ mit } \bar{X})$$

www.matstat.org

24

## Standardfehler

### 2. Standardfehler für den Schätzer von $p$ in $\text{Bin}(n, p)$ :

$$p^* = \frac{X}{n} \quad \text{schätzt } p \text{ in } \text{Bin}(n, p)$$

$$V(p^*) = V\left(\frac{X}{n}\right) = \frac{1}{n^2} V(X) = \frac{1}{n^2} \cdot n \cdot p \cdot (1-p) = \frac{p \cdot (1-p)}{n}$$

$$D(p^*) = \sqrt{\frac{p \cdot (1-p)}{n}} \quad \text{aber } p \text{ unbekannt (wir wollen ja } p \text{ schätzen!)}$$

$$\Downarrow \quad p \implies p^* = \frac{x}{n} \quad \text{ersetze } p \text{ mit Schätzung für } p$$

$$d(p^*) = \sqrt{\frac{p^* \cdot (1-p^*)}{n}} = \sqrt{\frac{\frac{x}{n} \cdot (1 - \frac{x}{n})}{n}} \quad \text{Standardfehler, ein numerischer Wert kann berechnet werden}$$

Standardfehler für Schätzung von  $p$  in  $\text{Bin}(n, p)$  mit  $p^* = x/n$

www.matstat.org

25

## Standardfehler für den Schätzer einer Proportion

**Stichprobe:**  $n = 1000$  Personen,  $x = 350$  antworteten "Ja" (Observation).  
Es wird nun eine Schätzung für den Anteil der "Ja"-Sager in der Gesamtpopulation gesucht, sowie der zugeordnete Standardfehler.

**Observation:**  $p_{obs}^* = x/n = 35\%$

**Zufallsvariable**  $p^* = X/n$  mit  $X \sim \text{Bin}(n, p)$

### Standardfehler

$$d(\hat{p}) = \sqrt{\frac{p_{obs} \cdot (1 - p_{obs})}{n}} = \sqrt{\frac{0.35 \cdot (1 - 0.35)}{1000}} = 1.5\%$$

**Ergebnis:** Anteil der "Ja"-Sager:  $35 \pm 1.5\%$

www.matstat.org

26

## Standardfehler

Der Standardfehler wird berechnet, indem man unbekannte Parameter in der Formel für die Standardabweichung eines Schätzers mit der Schätzung dieser Parameter ersetzt:

Verteilung	zu schätzen	Schätzer	Standardabw. des Schätzers	ersetze	Standardfehler des Schätzers
$N(\mu, \sigma)$	$\mu$	$\bar{X}$	$\frac{\sigma}{\sqrt{n}}$	$\sigma \rightarrow s$	$d = \frac{s}{\sqrt{n}}$
$Bin(n, p)$	$p$	$\frac{x}{n}$	$\sqrt{\frac{p \cdot (1-p)}{n}}$	$p \rightarrow p^*$	$\sqrt{\frac{p^* \cdot (1-p^*)}{n}}$

www.matstat.org

$$p^* = \frac{x}{n}$$