

# Wirksamkeit, Effizienz

## Definition 3.5 (Wirksamkeit, Effizienz)

Sei  $W$  eine parametrische Verteilungsannahme mit Parameterraum  $\Theta$ .

- ① Seien  $\hat{\theta}$  und  $\tilde{\theta}$  erwartungstreue Schätzfunktionen für  $\theta$ . Dann heißt  $\hat{\theta}$  **mindestens so wirksam** wie  $\tilde{\theta}$ , wenn

$$\text{Var}(\hat{\theta}) \leq \text{Var}(\tilde{\theta}) \text{ für alle } \theta \in \Theta$$

gilt.  $\hat{\theta}$  heißt **wirksamer** als  $\tilde{\theta}$ , wenn *außerdem*  $\text{Var}(\hat{\theta}) < \text{Var}(\tilde{\theta})$  für mindestens ein  $\theta \in \Theta$  gilt.

- ② Ist  $\hat{\theta}$  mindestens so wirksam wie alle (anderen) Schätzfunktionen einer Klasse mit erwartungstreuen Schätzfunktionen für  $\theta$ , so nennt man  $\hat{\theta}$  **effizient** in dieser Klasse erwartungstreuer Schätzfunktionen.

- Die Begriffe „Wirksamkeit“ und „Effizienz“ betrachtet man analog zu Definition 3.5 ebenfalls, wenn Funktionen  $g(\theta)$  von  $\theta$  geschätzt werden.
- $\text{Sd}(\hat{\theta}) = \sqrt{\text{Var}(\hat{\theta})}$  wird auch **Standardfehler** oder **Stichprobenfehler** von  $\hat{\theta}$  genannt.

## Beispiel: Effizienz

- Betrachte Klasse der (linearen) erwartungstreuen Schätzfunktionen

$$\hat{\mu}_{w_1, \dots, w_n} := \sum_{i=1}^n w_i \cdot X_i$$

mit  $\sum_{i=1}^n w_i = 1$  für den Erwartungswert  $\mu := E(Y)$  aus Folie 56.

- Für welche  $w_1, \dots, w_n$  erhält man (bei Vorliegen einer einfachen Stichprobe) die in dieser Klasse **effiziente** Schätzfunktion  $\hat{\mu}_{w_1, \dots, w_n}$ ?
- ↪ Suche nach den Gewichten  $w_1, \dots, w_n$  (mit  $\sum_{i=1}^n w_i = 1$ ), für die  $\text{Var}(\hat{\mu}_{w_1, \dots, w_n})$  möglichst klein wird.
- Man kann zeigen, dass  $\text{Var}(\hat{\mu}_{w_1, \dots, w_n})$  minimal wird, wenn

$$w_i = \frac{1}{n} \text{ für alle } i \in \{1, \dots, n\}$$

gewählt wird.

- Damit ist  $\bar{X}$  also effizient in der Klasse der linearen erwartungstreuen Schätzfunktionen für den Erwartungswert  $\mu$  einer Verteilung!

# Mittlerer quadratischer Fehler (MSE)

- Wenn Erwartungstreue im Vordergrund steht, ist Auswahl nach minimaler Varianz der Schätzfunktion sinnvoll.
- Ist Erwartungstreue nicht das „übergeordnete“ Ziel, verwendet man zur Beurteilung der Qualität von Schätzfunktionen häufig auch den sogenannten mittleren quadratischen Fehler (mean square error, MSE).

## Definition 3.6 (Mittlerer quadratischer Fehler (MSE))

Sei  $W$  eine parametrische Verteilungsannahme mit Parameterraum  $\Theta$ ,  $\hat{\theta}$  eine Schätzfunktion für  $\theta \in \Theta$ . Dann heißt  $\text{MSE}(\hat{\theta}) := E[(\hat{\theta} - \theta)^2]$  der **mittlere quadratische Fehler (mean square error, MSE)** von  $\hat{\theta}$ .

- Mit dem (umgestellten) Varianzzerlegungssatz erhält man direkt

$$E[(\hat{\theta} - \theta)^2] = \underbrace{\text{Var}(\hat{\theta} - \theta)}_{=\text{Var}(\hat{\theta})} + \underbrace{\left[E(\hat{\theta} - \theta)\right]^2}_{=(\text{Bias}(\hat{\theta}))^2},$$

für erwartungstreue Schätzfunktionen stimmt der MSE einer Schätzfunktion also gerade mit der Varianz überein!

# Konsistenz im quadratischen Mittel

- Basierend auf dem MSE ist ein „minimales“ Qualitätskriterium für Schätzfunktionen etabliert.
- Das Kriterium fordert (im Prinzip), dass man den MSE durch Vergrößerung des Stichprobenumfangs beliebig klein bekommen muss.
- Zur Formulierung des Kriteriums müssen Schätzfunktionen  $\hat{\theta}_n$  für „variable“ Stichprobengrößen  $n \in \mathbb{N}$  betrachtet werden.

## Definition 3.7 (Konsistenz im quadratischen Mittel)

Seien  $W$  eine parametrische Verteilungsannahme mit Parameterraum  $\Theta$ ,  $\hat{\theta}_n$  eine Schätzfunktion für  $\theta \in \Theta$  zum Stichprobenumfang  $n \in \mathbb{N}$ .

Dann heißt die (Familie von) Schätzfunktion(en)  $\hat{\theta}_n$  **konsistent im quadratischen Mittel für  $\theta$** , falls

$$\lim_{n \rightarrow \infty} \text{MSE}(\hat{\theta}_n) = \lim_{n \rightarrow \infty} \text{E} \left[ (\hat{\theta}_n - \theta)^2 \right] = 0$$

für alle  $\theta \in \Theta$  gilt.

- Mit der (additiven) Zerlegung des MSE in Varianz und quadrierten Bias aus Folie 59 erhält man sofort:

### Satz 3.8

Seien  $W$  eine parametrische Verteilungsannahme mit Parameterraum  $\Theta$ ,  $\hat{\theta}_n$  eine Schätzfunktion für  $\theta \in \Theta$  zum Stichprobenumfang  $n \in \mathbb{N}$ . Dann ist die Familie  $\hat{\theta}_n$  von Schätzfunktionen genau dann konsistent im quadratischen Mittel für  $\theta$ , wenn sowohl

$$\textcircled{1} \quad \lim_{n \rightarrow \infty} E(\hat{\theta}_n - \theta) = 0 \quad \text{bzw.} \quad \lim_{n \rightarrow \infty} E(\hat{\theta}_n) = \theta \quad \text{als auch}$$

$$\textcircled{2} \quad \lim_{n \rightarrow \infty} \text{Var}(\hat{\theta}_n) = 0$$

für alle  $\theta \in \Theta$  gilt.

- Eigenschaft  $\textcircled{1}$  aus Satz 3.8 wird auch **asymptotische Erwartungstreue** genannt; asymptotische Erwartungstreue ist offensichtlich schwächer als Erwartungstreue.
- Es gibt also auch (Familien von) Schätzfunktionen, die für einen Parameter  $\theta$  zwar konsistent im quadratischen Mittel sind, aber nicht erwartungstreu.

## Beispiel: Konsistenz im quadratischen Mittel

- Voraussetzung (wie üblich):  $X_1, \dots, X_n$  einfache Stichprobe zu  $Y$ .
- Bekannt: Ist  $\mu := E(Y)$  der unbekannte Erwartungswert der interessierenden Zufallsvariable  $Y$ , so ist  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  für alle  $n \in \mathbb{N}$  erwartungstreu.
- Ist  $\sigma^2 := \text{Var}(Y)$  die Varianz von  $Y$ , so erhält man für die Varianz von  $\bar{X}_n$  (vgl. Beweis der Effizienz von  $\bar{X}$  unter allen linearen erwartungstreuen Schätzfunktionen für  $\mu$ ):

$$\text{Var}(\bar{X}_n) = \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n \underbrace{\text{Var}(X_i)}_{=\sigma^2} = \frac{\sigma^2}{n}$$

- Es gilt also  $\lim_{n \rightarrow \infty} \text{Var}(\bar{X}_n) = \lim_{n \rightarrow \infty} \frac{\sigma^2}{n} = 0$ , damit folgt zusammen mit der Erwartungstreue, dass  $\bar{X}_n$  konsistent im quadratischen Mittel für  $\mu$  ist.

# Verteilung des Stichprobenmittels $\bar{X}$

- **Bisher:** Interesse meist an einigen *Momenten* (Erwartungswert und Varianz) von Schätzfunktionen, insbesondere des Stichprobenmittels  $\bar{X}$ .
- Bereits bekannt: Ist  $\mu := E(Y)$ ,  $\sigma^2 := \text{Var}(Y)$  und  $X_1, \dots, X_n$  eine einfache Stichprobe zu  $Y$ , so gilt

$$E(\bar{X}) = \mu \quad \text{sowie} \quad \text{Var}(\bar{X}) = \frac{\sigma^2}{n} .$$

- Damit Aussagen über Erwartungstreue, Wirksamkeit, Konsistenz möglich.
- **Jetzt:** Interesse an ganzer **Verteilung** von Schätzfunktionen, insbesondere  $\bar{X}$ .
- Verteilungsaussagen entweder
  - ▶ auf Grundlage des Verteilungstyps von  $Y$  aus der Verteilungsannahme in speziellen Situationen **exakt** möglich oder
  - ▶ auf Grundlage des zentralen Grenzwertsatzes (bei genügend großem Stichprobenumfang!) allgemeiner **näherungsweise (approximativ)** möglich.
- Wir unterscheiden im Folgenden nur zwischen:
  - ▶  $Y$  normalverteilt  $\rightsquigarrow$  Verwendung der exakten Verteilung von  $\bar{X}$ .
  - ▶  $Y$  nicht normalverteilt  $\rightsquigarrow$  Verwendung der Näherung der Verteilung von  $\bar{X}$  aus dem zentralen Grenzwertsatz.

## Aus „Deskriptive Statistik und Wahrscheinlichkeitsrechnung“:

- ① Gilt  $Y \sim N(\mu, \sigma^2)$ , so ist  $\bar{X}$  **exakt** normalverteilt mit Erwartungswert  $\mu$  und Varianz  $\frac{\sigma^2}{n}$ , es gilt also

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

- ② Ist  $Y$  beliebig verteilt mit  $E(Y) =: \mu$  und  $\text{Var}(Y) =: \sigma^2$ , so rechtfertigt der zentrale Grenzwertsatz **für ausreichend große Stichprobenumfänge**  $n$  die Näherung der tatsächlichen Verteilung von  $\bar{X}$  durch eine Normalverteilung mit Erwartungswert  $\mu$  und Varianz  $\frac{\sigma^2}{n}$  (wie oben!), man schreibt dann auch

$$\bar{X} \dot{\sim} N\left(\mu, \frac{\sigma^2}{n}\right)$$

und sagt „ $\bar{X}$  ist approximativ (näherungsweise)  $N\left(\mu, \frac{\sigma^2}{n}\right)$ -verteilt“.

Der Standardabweichung  $\text{Sd}(\bar{X}) = \sqrt{\text{Var}(\bar{X})}$  von  $\bar{X}$  (also der Standardfehler der Schätzfunktion  $\bar{X}$  für  $\mu$ ) wird häufig mit  $\sigma_{\bar{X}} := \frac{\sigma}{\sqrt{n}}$  abgekürzt.



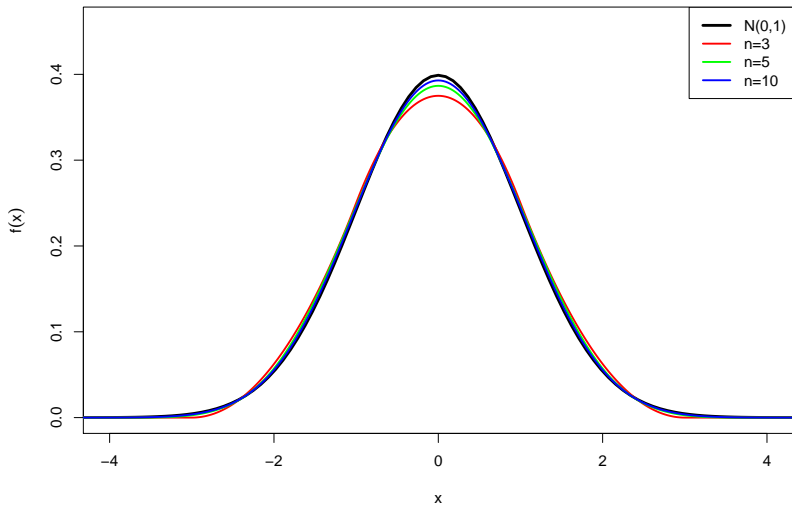
- Die Qualität der Näherung der Verteilung im Fall ② wird mit zunehmendem Stichprobenumfang höher, hängt aber **ganz entscheidend** vom Verteilungstyp (und sogar der konkreten Verteilung) von  $Y$  ab!
- Pauschale Kriterien an den Stichprobenumfang  $n$  („Daumenregeln“, z.B.  $n \geq 30$ ) finden sich häufig in der Literatur, sind aber nicht ganz unkritisch.
- Verteilungseigenschaft  $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$  bzw.  $\bar{X} \overset{\circ}{\sim} N\left(\mu, \frac{\sigma^2}{n}\right)$  wird meistens (äquivalent!) in der (auch aus dem zentralen Grenzwertsatz bekannten) Gestalt

$$\frac{\bar{X} - \mu}{\sigma} \sqrt{n} \sim N(0, 1) \quad \text{bzw.} \quad \frac{\bar{X} - \mu}{\sigma} \sqrt{n} \overset{\circ}{\sim} N(0, 1)$$

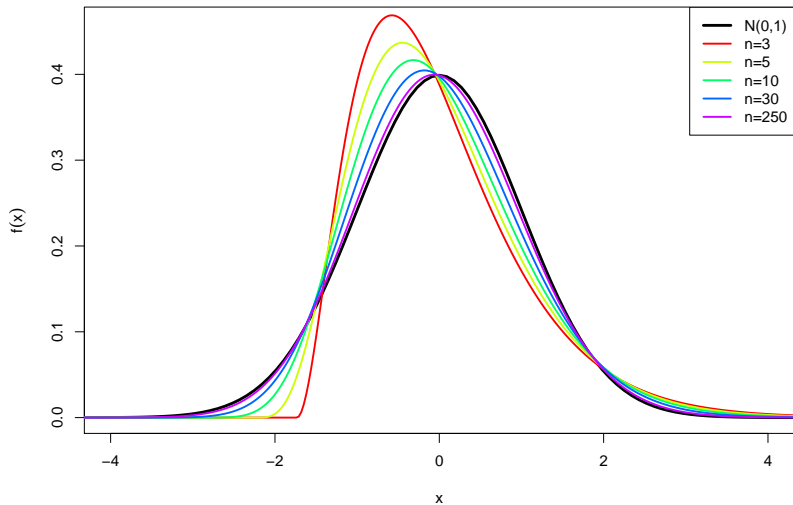
verwendet, da dann Verwendung von Tabellen zur Standardnormalverteilung möglich.

- Im Folgenden: Einige Beispiele für Qualität von Näherungen durch Vergleich der Dichtefunktion der Standardnormalverteilungsapproximation mit der tatsächlichen Verteilung von  $\frac{\bar{X} - \mu}{\sigma} \sqrt{n}$  für unterschiedliche Stichprobenumfänge  $n$ .

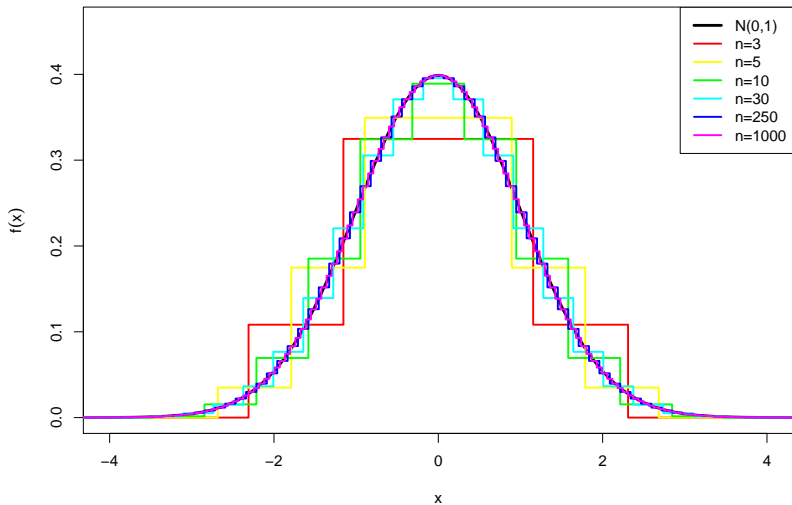
Beispiel: Näherung, falls  $Y \sim \text{Unif}(20, 50)$



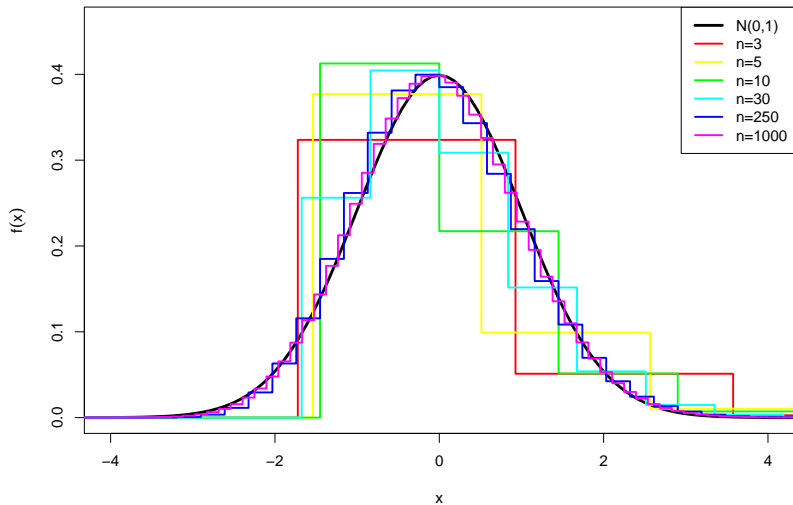
# Beispiel: Näherung, falls $Y \sim \text{Exp}(2)$



# Beispiel: Näherung, falls $Y \sim B(1, 0.5)$



Beispiel: Näherung, falls  $Y \sim B(1, 0.05)$



# Schwankungsintervalle für $\bar{X}$

- Eine Verwendungsmöglichkeit für Verteilung von  $\bar{X}$ :

Berechnung von (festen) Intervallen mit der Eigenschaft, dass die Stichprobenziehung mit einer vorgegebenen Wahrscheinlichkeit zu einer Realisation von  $\bar{X}$  führt, die in dieses berechnete Intervall fällt.

Solche Intervalle heißen **Schwankungsintervalle**.

- Gesucht sind also Intervallgrenzen  $g_u < g_o$  von Intervallen  $[g_u, g_o]$  mit

$$P_{\bar{X}}([g_u, g_o]) = P\{\bar{X} \in [g_u, g_o]\} \stackrel{!}{=} p_S$$

für eine vorgegebene Wahrscheinlichkeit  $p_S \in (0, 1)$ .

- Aus bestimmten Gründen (die später verständlich werden) gibt man nicht  $p_S$  vor, sondern die Gegenwahrscheinlichkeit  $\alpha := 1 - p_S$ , d.h. man fordert

$$P_{\bar{X}}([g_u, g_o]) = P\{\bar{X} \in [g_u, g_o]\} \stackrel{!}{=} 1 - \alpha$$

für ein vorgegebenes  $\alpha \in (0, 1)$ .

$1 - \alpha$  wird dann auch **Sicherheitswahrscheinlichkeit** genannt.

- Eindeutigkeit für die Bestimmung von  $g_u$  und  $g_o$  erreicht man durch die Forderung von **Symmetrie** in dem Sinn, dass die untere bzw. obere Grenze des Intervalls jeweils mit einer Wahrscheinlichkeit von  $\alpha/2$  unter- bzw. überschritten werden soll, d.h. man fordert genauer

$$P\{\bar{X} < g_u\} \stackrel{!}{=} \frac{\alpha}{2} \quad \text{und} \quad P\{\bar{X} > g_o\} \stackrel{!}{=} \frac{\alpha}{2}.$$

- Unter Verwendung der Verteilungseigenschaft

$$\frac{\bar{X} - \mu}{\sigma} \sqrt{n} \sim N(0, 1) \quad \text{bzw.} \quad \frac{\bar{X} - \mu}{\sigma} \sqrt{n} \dot{\sim} N(0, 1)$$

erhält man also exakt bzw. näherungsweise

$$\begin{aligned} P\{\bar{X} < g_u\} &= P\left\{\frac{\bar{X} - \mu}{\sigma} \sqrt{n} < \frac{g_u - \mu}{\sigma} \sqrt{n}\right\} \stackrel{!}{=} \frac{\alpha}{2} \\ \Leftrightarrow \frac{g_u - \mu}{\sigma} \sqrt{n} &= \Phi^{-1}\left(\frac{\alpha}{2}\right) \\ \Rightarrow g_u &= \mu + \frac{\sigma}{\sqrt{n}} \cdot \Phi^{-1}\left(\frac{\alpha}{2}\right) \end{aligned}$$

als untere Intervallgrenze.

- Analog erhält man exakt bzw. näherungsweise

$$\begin{aligned}
 P\{\bar{X} > g_o\} &= P\left\{\frac{\bar{X} - \mu}{\sigma}\sqrt{n} > \frac{g_o - \mu}{\sigma}\sqrt{n}\right\} \stackrel{!}{=} \frac{\alpha}{2} \\
 \Leftrightarrow \frac{g_o - \mu}{\sigma}\sqrt{n} &= \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \\
 \Rightarrow g_o &= \mu + \frac{\sigma}{\sqrt{n}} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) .
 \end{aligned}$$

als die obere Intervallgrenze.

- Als Abkürzung für  $p$ -Quantile der Standardnormalverteilung (also Funktionswerte von  $\Phi^{-1}$  an der Stelle  $p \in (0, 1)$ ) verwenden wir:

$$N_p := \Phi^{-1}(p)$$

- Man erhält also insgesamt als symmetrisches Schwankungsintervall für  $\bar{X}$  exakt bzw. näherungsweise das Intervall

$$\left[ \mu + \frac{\sigma}{\sqrt{n}} \cdot N_{\frac{\alpha}{2}}, \mu + \frac{\sigma}{\sqrt{n}} \cdot N_{1 - \frac{\alpha}{2}} \right] .$$