

**INHALT**Nicht verwechseln

Konfidenz, Kontingenz, Konvergenz,  
Regression, Korrelation, Autokorrelation, Interkorrelation.

Nicht verwechseln

normal, Normalverteilung, Standard

normiert z.B.  $v = \frac{s}{\bar{x}}$

Norm z.B. EUKLID-Norm:  $\|x\|_2 = \sqrt{\sum x_i^2}$

Normale, Normalgleichungen

**6.1 Konfidenzintervall**Wir schließen

vom Mittelwert der Stichprobe  $\bar{x}$   
auf den Erwartungswert der Grundgesamtheit  $\mu$ .

Wir beachten

1. stetig, deshalb  $n p q > 9$  nicht zu prüfen
2.  $n < 30$ , dann t-Verteilung statt Normalverteilung anwenden
3.  $n/N > 0,05$ , dann  $\sigma_{\bar{x}}$  mit dem Korrekturfaktor  $\sqrt{\frac{N-n}{N-1}}$  multiplizieren

**6.7 Hypothesentest**Wir schließen

vom behaupteten Erwartungswert der Grundgesamtheit  $\mu$   
auf den Mittelwert der Stichprobe  $\bar{x}$

Wir beachten exakt wie oben

1. stetig, deshalb  $n p q > 9$  nicht zu prüfen
2.  $n < 30$ , dann t-Verteilung statt Normalverteilung anwenden
3.  $n/N > 0,05$ , dann  $\sigma_{\bar{x}}$  mit dem Korrekturfaktor  $\sqrt{\frac{N-n}{N-1}}$  multiplizieren

## > 6.1 SCHLIESSEN VON $\bar{x}$ AUF $\mu$

Wir beschränken uns in diesem Thema auf Zufallsvariable, die

1. stetig oder quasistetig sind
2. einigermaßen normalverteilt sind (das kann man mit dem Anpassungstest aus 5.18 prüfen)

Wir schließen vom Mittelwert der Stichprobe  $\bar{x}$  auf den Erwartungswert der Grundgesamtheit  $\mu$ .

Natürlich schließen wir nicht auf  $\mu$ , sondern auf ein Intervall, in dem  $\mu$  liegt:

$$\bar{x} \Rightarrow \mu \in [\bar{x} - z_c \sigma_{\bar{x}} ; \bar{x} + z_c \sigma_{\bar{x}}] \quad z = \frac{x - \mu}{\sigma} \quad \text{bzw.} \quad x = \mu + z \cdot \sigma \quad \text{oder} \quad \mu = x - z\sigma$$

darin ist  $\sigma_{\bar{x}}$  die Streuung des Stichprobenmittels  $\bar{x}$  und

$z_c$  die kritische **Standardnormalvariable** für ein gegebenes Signifikanzniveau  $1-\alpha$ .

Dieses Intervall nennt man **Konfidenzintervall** (Vertrauensbereich), man vertraut darauf, dass  $\mu$  in diesem Intervall liegt.

Die Wahrscheinlichkeit dafür, dass  $\mu$  im Konfidenzintervall liegt, ist

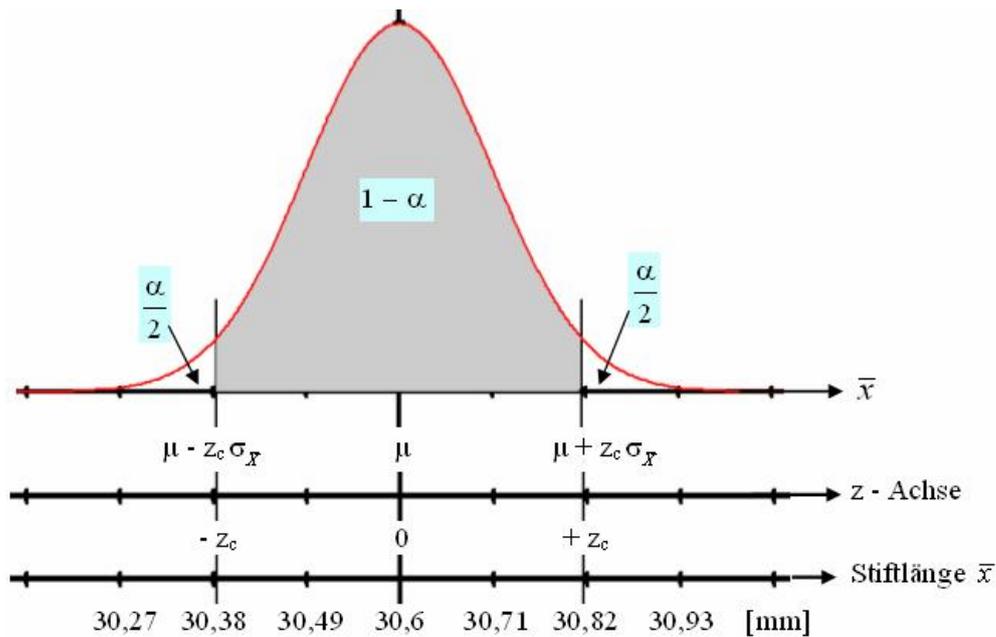
$$W(\bar{x} - z_c \sigma_{\bar{x}} \leq \mu \leq \bar{x} + z_c \sigma_{\bar{x}}) = 1 - \alpha = D(z_c) \quad \text{Tabelle 7.6 } F_{SN}$$

$\mu$  liegt mit einem Sicherheitsgrad von  $1-\alpha$  im Bereich  $[\bar{x} - z_c \sigma_{\bar{x}} ; \bar{x} + z_c \sigma_{\bar{x}}]$ .

$\mu$  liegt mit einem Sicherheitsgrad von 95% im Bereich  $[\bar{x} - 1,96 \sigma_{\bar{x}} ; \bar{x} + 1,96 \sigma_{\bar{x}}]$ .

$\mu$  liegt mit einem Sicherheitsgrad von 95% im Bereich  $\left[ \bar{x} - 1,96 \frac{\sigma}{\sqrt{n}} ; \bar{x} + 1,96 \frac{\sigma}{\sqrt{n}} \right]$ .

## 6.2 KONFIDENZINTERVALL



Man möchte wissen in welchem Konfidenzintervall  $\mu_{\text{Grundgesamtheit}}$  liegt.

$$\mu \in [\bar{x} - z_c \sigma_{\bar{x}} ; \bar{x} + z_c \sigma_{\bar{x}}]$$

Die Irrtums-Wahrscheinlichkeit  $\alpha$  besteht aus zwei gleich großen Teilen,

z.B. bei  $\alpha = 0,05$ , der "left tail"  $\alpha/2 = 2,5\%$  und der "right tail"  $\alpha/2 = 2,5\%$ .

### Beispiel 6.1

In einem Betrieb werden Messingstifte mit einer Standardabweichung von  $\sigma=1$  [mm] hergestellt. Erfahrungsgemäß ist die Standardabweichung unabhängig von der eingestellten Länge.

a) Bei einer Stichprobe vom Umfang  $n=80$  ergab sich eine mittlere Stiftlänge von  $\bar{x}=30,60$  [mm].

Das Konfidenzintervall bei einem Sicherheitsgrad von  $1-\alpha=96\%$  ist zu bestimmen.

Mit Konfidenzintervall ist immer das Intervall  $[\bar{x} - z_c \sigma_{\bar{x}} ; \bar{x} + z_c \sigma_{\bar{x}}]$  gemeint.

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1}{\sqrt{80}} = 0,11 \text{ mm. Tabelle 7.6: } D(2,05) \approx 0,96.$$

$$[\bar{x} - z_c \sigma_{\bar{x}} ; \bar{x} + z_c \sigma_{\bar{x}}] = \left[ 30,6 - 2,05 \cdot \frac{1}{\sqrt{80}} ; 30,6 + 2,05 \cdot \frac{1}{\sqrt{80}} \right] =$$

$$[30,6 - 0,23 ; 30,6 + 0,23] \Rightarrow \mu \in [30,37 ; 30,83]$$

b) In diesem Betrieb wird mit einem Konfidenzintervall  $\mu \in [30,4 \text{ mm} ; 30,8 \text{ mm}]$  geprüft, bei Stichprobenumfängen von 60 Stück. Mit welcher Irrtumswahrscheinlichkeit  $\alpha$  arbeitet dieser Betrieb? (oder "auf welchem Signifikanzniveau  $1-\alpha$  ")

$$30,8 = 30,6 + z \cdot 1 / \sqrt{60} \Rightarrow 0,2 \dots \Rightarrow z = 1,55 \Rightarrow D(1,55) = 0,8789 = 1 - \alpha$$

$$\sigma_{\bar{x}} = 1 / \sqrt{60} = 0,129. \quad 30,8 = \bar{x} + z_c \sigma_{\bar{x}} \Rightarrow 30,8 = 30,6 + z_c \cdot 0,129 \Rightarrow 0,2 = 0,129 z_c$$

$$\Rightarrow z_c = 1,55 \Rightarrow D(1,55) = 0,8789 = 1 - \alpha \Rightarrow \alpha = 0,12 = 12\%$$

c) Die Ablängvorrichtung soll auf ihre Einstellung geprüft werden. Wie groß muss der Stichprobenumfang mindestens  $n$  sein, damit bei einem Sicherheitsgrad von 95% die Länge des Konfidenzintervalles höchstens 0,3 mm ist?  $\Rightarrow$  **Abweichung = 0,15 mm**

Abweichung  $\varepsilon = \bar{x} - \mu =$  "halbes Konfidenzintervall".  $30,6 + 1,96 \cdot 1 / \sqrt{n} = 30,75 \dots$

$$n = \left( \frac{z_c \sigma}{\bar{x} - \mu} \right)^2 = \left( \frac{1,96 \cdot 1}{0,15} \right)^2 = 170,74 \Rightarrow n = 171 \text{ Stück} \quad \text{ceil}(n) \text{ aufrunden!}$$

Formel-Herleitung siehe 6.3

### 6.3 UMRECHNUNGEN

Das Konfidenzintervall von  $\mu$ :

$$\bar{x} - z_c \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_c \frac{\sigma}{\sqrt{n}}$$

andere Schreibweise:

$$\mu \in \left[ \bar{x} - z_c \frac{\sigma}{\sqrt{n}} ; \bar{x} + z_c \frac{\sigma}{\sqrt{n}} \right]$$

mit  $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

$$\mu \in \left[ \bar{x} - z_c \sigma_{\bar{x}} ; \bar{x} + z_c \sigma_{\bar{x}} \right]$$

mit der Abweichung  $\varepsilon$ :

$$\mu \in \left[ \bar{x} - \varepsilon ; \bar{x} + \varepsilon \right]$$

mit den Grenzen  $g$

$$\mu \in \left[ \mu_{\text{unten}} ; \mu_{\text{oben}} \right]$$

Länge des Intervalls  $2\varepsilon$

$$\mu_{\text{oben}} - \mu_{\text{unten}} = 2\varepsilon = 2z_c \frac{\sigma}{\sqrt{n}}$$

$\bar{x}$ -Strecke unter  
der D-Fläche

Die Wahrscheinlichkeit dafür, dass  $\mu$  in diesem Intervall liegt,  
ist der Sicherheitsgrad  $1 - \alpha$ .  $1 - \alpha$  ist der Tabellen-Wert  $D(z_c)$ .

An der rechten Grenze gilt:

$$\mu_{\text{oben}} = \mu_{\text{grenz}} = \bar{x} + z_c \frac{\sigma}{\sqrt{n}}$$

darin sind alle 5 Größen.

nach der gesuchten Größe auflösen:  $\varepsilon = |\bar{x} - \mu| = z_c \frac{\sigma}{\sqrt{n}}$  die Abweichung von  $\mu$

$\varepsilon$  ist die halbe Länge des Konfidenzintervalls.

$$\boxed{\text{durch } \sigma, \text{ mal } \sqrt{n}} \Rightarrow \frac{\bar{x} - \mu}{\sigma} \sqrt{n} = z_c \quad \text{die kritische Grenze } z_c$$

$$\boxed{\text{mal } \sigma, \text{ durch } \bar{x} - \mu} \Rightarrow \sqrt{n} = \frac{z_c \sigma}{\bar{x} - \mu} \Rightarrow n = \left( \frac{z_c \sigma}{\bar{x} - \mu} \right)^2 \text{ der Stichprobenumfang}$$

Zusammenhang zwischen  $D(z_c)$ ,  $z_c$ ,  $\alpha$ ,  $1 - \alpha$

gegeben: Sicherheitsgrad 99%  $\Rightarrow \alpha = 0,01$ ,  $1 - \alpha = 0,99$ ,  $D(\dots) = 0,99 \Rightarrow z_c = 2,33$

gegeben: Irrtumswahrscheinlichkeit  $\alpha = 5\% \Rightarrow 1 - \alpha = 0,95$ ,  $D(\dots) = 0,95 \Rightarrow z_c = 1,96$

gegeben:  $\bar{x}$ ,  $\mu$ ,  $\sigma$ ,  $n$ . z.B.  $z_c = \frac{\bar{x} - \mu}{\sigma} \sqrt{n} = 1,75 \Rightarrow D(1,75) = 0,92 = 1 - \alpha$ ,  $\alpha_{\text{rechts}} = 0,04 = 4\%$

### 6.3A APPROXIMATION

Ein typisches Beispiel aus der Qualitätskontrolle:

Ein Betrieb fertigt Bauteile vom Typ X125 in hoher Stückzahl. Erfahrungsgemäß sind 10% dieser Bauteile defekt, d.h. die Treffer-Wahrscheinlichkeit ist  $p = 0,1$ .

Die Gegen-Wahrscheinlichkeit ist dann  $q = 1 - p = 0,9$ .

Ein Kunde prüft eine Stichprobe vom Umfang  $n = 80$  Stück auf Funktionstüchtigkeit.

Will man einen solchen Sachverhalt mit Hilfe der Standard-Normalverteilung prüfen hat man drei Approximationsprobleme:

1. Für diskrete Zufallsvariable kann man die Normalverteilung benutzen, wenn  $n p q > 9$ .

□ Hier ist  $n p q = 80 \cdot 0,1 \cdot 0,9 = 7,2$ .

Die Stichprobe müsste also mindestens 101 Bauteile umfassen.

Wir betrachten **hier nur stetige Zufallsvariablen**, also können wir bei jeder Aufgabe angeben:

**" $n p q > 9$  ist nicht zu prüfen, weil eine stetige Zufallsvariable vorliegt".**

2. Versuchsreihen werden in der Regel ohne Zurücklegen durchgeführt.

Man wird das gerade geprüfte Bauteil nicht wieder zurücklegen, mischen, dann neu "ziehen".

Wenn man die geprüften Bauteile nicht zurücklegt, ändert sich nach jeder Prüfung die Wahrscheinlichkeit  $p$ , ein defektes Bauteil zu "ziehen".

Wenn für den Auswahlssatz gilt  **$n/N \leq 0,05$**  kann man trotzdem die Normalverteilung verwenden. andernfalls muss man die Standardabweichung  $s_{\bar{x}}$  bzw.  $\sigma_{\bar{x}}$  mit dem Korrekturfaktor

$$\sqrt{\frac{N-n}{N-1}}$$

multiplizieren.

□ Hier ist  $n = 80$ .  $\frac{n}{N} = \frac{80}{1600} = 0,05$ .

Die gesamte Fertigung müsste also mindestens 1600 Stück umfassen.

Aus der Formulierung "hohe Stückzahl" folgert man hier, dass  $N > 1600$  Stück.

Bei jeder Aufgabe ist also anzugeben:

**"Korrekturfaktor (nicht) erforderlich, weil  $n/N = \boxed{80/1600} \leq 0,05$  "**

3. Im Konfidenzintervall  $[\bar{x} - z_c \sigma_{\bar{x}} ; \bar{x} + z_c \sigma_{\bar{x}}]$  benötigt man die erwartete Streuung  $\sigma_{\bar{x}}$ .

Die Streuung  $\sigma_{\bar{x}}$  in der Grundgesamtheit ist normalerweise unbekannt.

Bei der Auswertung einer Stichprobe ergeben sich  $\bar{x}$  und  $s_{\bar{x}}$ .

Man kann näherungsweise  $s_{\bar{x}}$  statt  $\sigma_{\bar{x}}$  verwenden, wenn  $n \geq 30$ .

andernfalls muss man die STUDENT-t-Verteilung statt der Normalverteilung anwenden.

(Bei nicht-normalverteilten Zufallsvariablen, wenn  $n \geq 50$ )

□ Hier ist  $n = 80 > 30$

Bei jeder Aufgabe ist also anzugeben:

**"STUDENT-t-Verteilung (nicht) erforderlich, weil  $n = \boxed{80} \geq 30$  "**

In allen Aufgaben zum Konfidenzintervall sind diese drei Bedingungen anzugeben:

1.  $n p q > 9$  nicht zu prüfen, weil Zufallsvariable stetig ist

2.  $n < 30$  also t-Verteilung benutzen oder  $n \geq 30$  also Normalverteilung benutzen

3.  $n/N > 0,05$  Korrekturfaktor anwenden oder  $n/N \leq 0,05$  kein Korrekturfaktor

## 6.4 BREMSBELÄGE

### Beispiel 6.2

Ein Betrieb fertigt Bremsbeläge in einer typischen Produktionsserie von  $N=1500$  Stück. Die Nutzungsdauer einer bestimmten Sorte Bremsbeläge ist erfahrungsgemäß normalverteilt. In einer Stichprobe vom Umfang  $n=64$  [Stück] wird eine mittlere Nutzungsdauer von  $\bar{x}=41.000$  [km] und eine Standardabweichung von  $s=5000$  [km] ermittelt.

a) Es ist genau zu erläutern, ob hier die Normalverteilung verwendet werden kann.

(1) Zufallsvariable stetig, also ist  $npq > 9$  nicht zu prüfen.

(2)  $s = 5000$  näherungsweise für  $\sigma$ :  $n > 30$ , also ist die  $t$ -Verteilung nicht erforderlich.

(3) ohne Zurücklegen  $n/N = 64 / 1500 = 0,043 < 0,05$  also den Korrekturfaktor weglassen

b) Bestimmen Sie den Vertrauensbereich für  $\mu$  bei einer Irrtumswahrscheinlichkeit  $\alpha = 0,02$ .

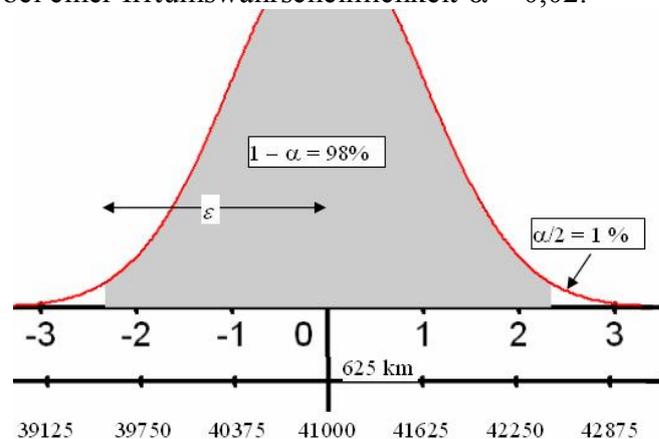
$$1 - \alpha = 0,98 = D(2,33). \quad z_c = 2,33$$

$$[\bar{x} - z_c \sigma_{\bar{x}} ; \bar{x} + z_c \sigma_{\bar{x}}]$$

$$= \left[ 41000 - 2,33 \cdot \frac{5000}{\sqrt{64}} ; 41000 + 2,33 \cdot \frac{5000}{\sqrt{64}} \right]$$

$$= [41.000 - 1.456 ; 41.000 + 1.456]$$

$$= [39.544 ; 42.456] \text{ km}$$



c) Bestimmen Sie den Vertrauensbereich für  $\mu$  bei einem Sicherheitsgrad von 95%.

$$1 - \alpha = 0,95 = D(1,96). \quad z_c = 1,96$$

$$[41000 - 1,96 \cdot 5000/8 ; 41000 + 1,96 \cdot 5000/8 ] = [39.775 ; 42.225] \text{ km.}$$

d) In diesem Betrieb wird bei üblichen Stichproben von 50 Bremsbelägen mit dem Konfidenzintervall [40.000 ; 42.000] km gerechnet. Mit welcher Irrtumswahrscheinlichkeit arbeitet der Betrieb?

$$\bar{x} - \mu = 1000 = z_c \frac{\sigma}{\sqrt{n}}. \quad 1000 = z_c \frac{5000}{\sqrt{50}} \Rightarrow z_c = 1,41 \Rightarrow D(1,41) = 0,8415$$

$$1 - \alpha = 0,8415. \quad \underline{\alpha = 15,85 \%}$$

Ergebnis mit Skizze vergleichen:  $z_c = 1,41$  entspricht 42.000 km

e) Wie groß muss der Stichprobenumfang mindestens sein, damit bei einem Sicherheitsgrad von 99% die Abweichung  $\epsilon$  höchstens 1000 km beträgt?

$$|\bar{x} - \mu| = z_c \frac{\sigma}{\sqrt{n}} \Rightarrow n \geq \left( \frac{z_c \sigma}{\bar{x} - \mu} \right)^2 = \left( \frac{z_c \sigma}{\epsilon} \right)^2 = \frac{z_c^2 \sigma^2}{(\bar{x} - \mu)^2}. \quad 1000 = 2,58 \cdot \frac{5000}{\sqrt{n}} \Rightarrow n = 167 \text{ Stück}$$

## 6.5 KONFIDENZINTERVALL - VARIANTEN

### 1. Konfidenzintervall bei unbekannter Standardabweichung $\sigma$

Bei der Bestimmung des Konfidenzintervalles für den Mittelwert der Grundgesamtheit  $\mu$  wurde die Standardabweichung der Grundgesamtheit  $\sigma$  als bekannt vorausgesetzt.

$$\mu \in \left[ \bar{x} - z_c \frac{\sigma}{\sqrt{n}} ; \bar{x} + z_c \frac{\sigma}{\sqrt{n}} \right]$$

Wenn  $\sigma$  nicht bekannt ist, dann kann man  $\sigma$  näherungsweise durch die Standardabweichung

$$\text{der Stichprobe } s = +\sqrt{s^2} \quad \text{mit} \quad s^2 = \frac{1}{n-1} \left( \sum_{i=1}^k x_i^2 n_i - n \bar{x}^2 \right) \text{ ersetzen.}$$

Die Aussage über das Konfidenzintervall von  $\mu$  wird dadurch ungenauer.

Bei kleinen Stichproben kann man nicht ohne weiteres  $\sigma_{\text{Grundgesamtheit}}$  durch  $s_{\text{Stichprobe}}$  ersetzen.

Bei normalverteilten Zufallsvariablen benutzt man deshalb für  $n < 30$  die Student-t-Verteilung;

$$\text{t-Werte statt z-Werte: } z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \Rightarrow t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} \Rightarrow \mu \in \left[ \bar{x} - t_c \frac{s}{\sqrt{n}} ; \bar{x} + t_c \frac{s}{\sqrt{n}} \right]$$

Die Werte der Verteilungsfunktion  $F_v(t)$  und für die Mittige Fläche  $1 - \alpha = D_v(t)$  sind für wichtige Sicherheitsgrade tabelliert.  $\rightarrow$  Tabelle 7.5b

#### Beispiel 6.3

Eine Stichprobe von  $n=25$  aus einer normalverteilten Grundgesamtheit ergab den Mittelwert  $\bar{x}=105$  und die Varianz  $s^2=40$ . Zum Signifikanzniveau  $\alpha=0,05$  soll ein Konfidenzintervall für den Erwartungswert  $\mu$  der Grundgesamtheit bestimmt werden.

*Konfidenzintervall von  $\mu$  gesucht,  $\alpha = 0,05$ .  $\sigma$  unbekannt,*

*$n = 25 < 30$  also Student-t-Verteilung benutzen. Freiheitsgrade  $\nu = 25 - 1 = 24$ .*

*Tabelle 7.5b liefert  $0,95 = D_{24}(t_c) = D_{24}(2,064)$*

$$\Rightarrow \left[ \bar{x} - t_c \frac{s}{\sqrt{n}} ; \bar{x} + t_c \frac{s}{\sqrt{n}} \right] = \left[ 105 - 2,064 \cdot \frac{\sqrt{40}}{5} ; 105 + 2,064 \cdot \frac{\sqrt{40}}{5} \right] = [102,39 ; 107,61]$$

### 2. Konfidenzintervall bei Stichproben ohne Zurücklegen

Der Ansatz  $\mu \in \left[ \bar{x} - z_c \frac{\sigma}{\sqrt{n}} ; \bar{x} + z_c \frac{\sigma}{\sqrt{n}} \right]$  gilt für Stichproben mit Zurücklegen.

Für  $n/N \leq 0,05$  kann man diesen Ansatz auch für Stichproben ohne Zurücklegen verwenden.

Für Auswahlsätze  $n/N > 0,05$  ist ein Korrekturfaktor erforderlich:  $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$

#### Beispiel 6.4

Aus einer Grundgesamtheit von  $N=1000$  Rechnungen eines Einzelhandelsgeschäfts wird eine Stichprobe vom Umfang  $n=100$  ohne Zurücklegen gezogen. In der Stichprobe haben die Rechnungsbeträge einen Mittelwert von  $\bar{x}=15,00$  € und eine Standardabweichung von  $s=4,10$  €. Es soll das Konfidenzintervall für die Rechnungsbeträge beim Sicherheitsgrad von 96% bestimmt werden.

*$n \geq 50$  d.h. man kann  $s_{\text{Stichprobe}}$  verwenden statt  $\sigma_{\text{Grundgesamtheit}}$ . Keine t-Verteilung.*

$$\frac{n}{N} = \frac{100}{1000} = 0,1 > 0,05 \Rightarrow \sigma_{\bar{x}, \text{ohne Zurücklegen}} = \frac{4,1}{\sqrt{100}} \cdot \sqrt{\frac{1000-100}{1000-1}} = 0,41 \cdot 0,949 = 0,389$$

$$D(z_c) = 0,96 \approx D(2,05) \Rightarrow \mu \in [15 - 2,05 \cdot 0,389 ; 15 + 2,05 \cdot 0,389] = [14,20 ; 15,80]$$

## 6.6 AUFGABE KONFIDENZINTERVALL

### Aufgabe Konfidenzintervall vgl. Übung4

gegeben: Sachverhalt wie in den Beispielen 6.1 – 6.4, stetige Zufallsvariable,  $\sigma$  oder  $s$  gegeben, mit oder ohne Zurücklegen,  $\alpha$  oder  $1 - \alpha$  oder Intervall oder Abweichung  $\varepsilon$ .  
Bei Qualitätskontrollen ist immer "ohne Zurücklegen" anzunehmen, auch wenn sonst kein Hinweis vorliegt.

gesucht / Schritte:

1. Welche Verteilung wird verwendet?

- (1) stetige Zufallsvariable:  $npq > 9$  ist nicht zu prüfen
- (2) prüfen  $n < 30$  bzw.  $n < 50$ , dann  $t$ -Verteilung benutzen
- (3) prüfen "ohne Zurücklegen" und  $n/N > 0,05$ , dann Korrekturfaktor

2. Konfidenzintervall

- (1) die gegebenen Größen zusammenstellen,
- (2)  $\sigma_{\bar{x}}$  oder  $\sigma_{\bar{x}, \text{ohne Zurücklegen}}$  berechnen
- (3)  $D(z_c)$  bzw.  $D_v(t_c)$  in der Tabelle ablesen
- (4) einsetzen in  $[\bar{x} - z_c \sigma_{\bar{x}} ; \bar{x} + z_c \sigma_{\bar{x}}]$  bzw.  $[\bar{x} - t_c \sigma_{\bar{x}} ; \bar{x} + t_c \sigma_{\bar{x}}]$

3. Irrtums-Wahrscheinlichkeit  $\alpha$

- (1) prüfen wie bei 1., ob  $t$ -Verteilung bzw. Korrekturfaktor erforderlich gegebenenfalls  $D_v(t_c)$  benutzen oder/und  $\sigma_{\bar{x}}$  korrigieren
- (2) gegebene Werte in  $\bar{x} - \mu = z_c \frac{\sigma}{\sqrt{n}}$  einsetzen, nach  $z_c$  auflösen.
- (3)  $\alpha = 1 - D(z_c)$

4. Stichprobenumfang

- (1) prüfen wie bei 1., ob  $t$ -Verteilung bzw. Korrekturfaktor erforderlich gegebenenfalls  $D_v(t_c)$  benutzen oder/und  $\sigma_{\bar{x}}$  korrigieren
- (2) je nachdem, ob der Korrekturfaktor erforderlich ist,  $n$  berechnen

$$\text{aus } \varepsilon = |\bar{x} - \mu| = z_c \frac{\sigma}{\sqrt{n}} \quad \text{bzw. aus } \varepsilon = |\bar{x} - \mu| = z_c \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

$$\text{oder mit den Formeln } \left( \frac{z_c \sigma}{\varepsilon} \right)^2 \quad \text{bzw. } n \geq \frac{N}{1 + (N-1) \left( \frac{\varepsilon}{z_c \sigma} \right)^2} \quad *)$$

(3)  $n$  auf nächste ganze Zahl aufrunden, Antwort: "n mindestens ... Stück."

$$*) |\bar{x} - \mu| = \varepsilon = z_c \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \Rightarrow \varepsilon^2 = z_c^2 \frac{\sigma^2}{n} \frac{N-n}{N-1} \Rightarrow n \frac{\varepsilon^2}{z_c^2 \sigma^2} = \frac{N-n}{N-1}$$

$$\Rightarrow (N-1)n \frac{\varepsilon^2}{z_c^2 \sigma^2} = N-n \Rightarrow n + n(N-1) \frac{\varepsilon^2}{z_c^2 \sigma^2} = N$$

$$\Rightarrow n \left( 1 + (N-1) \frac{\varepsilon^2}{z_c^2 \sigma^2} \right) = N \Rightarrow n \geq \frac{N}{1 + (N-1) \left( \frac{\varepsilon}{z_c \sigma} \right)^2}$$

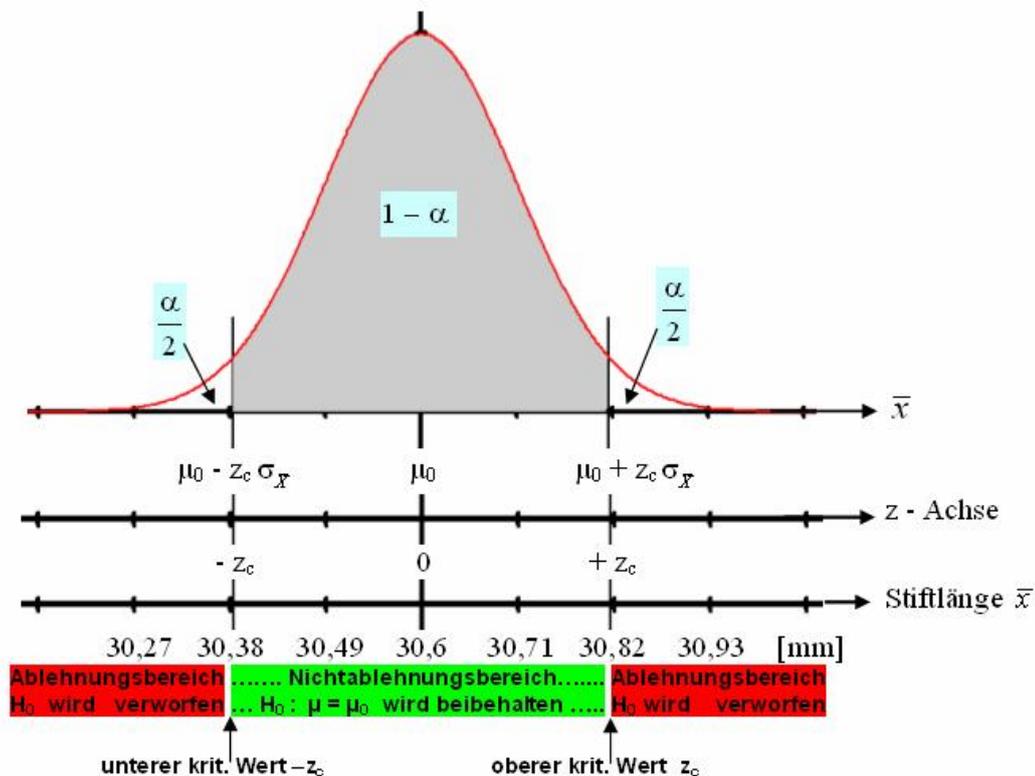
## > 6.7 SCHLIESSEN VON $\mu$ AUF $\bar{x}$

### Schwankungsbereich der Zufallsvariablen $\bar{X}$ = Wiederholung aus 6.1

Die Werte  $\bar{x}$  eines Stichprobenmittels  $\bar{X}$  liegen symmetrisch um den Mittelwert  $\mu$ .  
Mit Hilfe der Standard-Normalverteilung kann man die Wahrscheinlichkeit berechnen, dafür

$$\text{dass } \bar{x} \in [\mu - z \sigma_{\bar{x}}, \mu + z \sigma_{\bar{x}}] = [\bar{x}_{\text{unten}}, \bar{x}_{\text{oben}}] \quad \text{mit } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Für jeden Sicherheitsgrad  $1 - \alpha = D(z)$  gibt es einen solchen Schwankungsbereich.



Speziell für den Sicherheitsgrad  $1 - \alpha = 0,95$ :

liegen die Werte  $xy$  im Bereich  $[\mu - 1,96 \sigma_{\bar{x}}; \mu + 1,96 \sigma_{\bar{x}}]$ .  $\alpha = 0,05$ ,  $z_c = 1,96$ .  $D(1,96) = 0,95$

Wenn sich bei der Auswertung der Stichprobe ergibt  $\bar{x} \notin [\mu - 1,96 \sigma_{\bar{x}}; \mu + 1,96 \sigma_{\bar{x}}]$

dann besteht ein **signifikanter Unterschied** zwischen der behaupteten Nullhypothese  $\mu_0 = \mu$  und den Ergebnissen der Stichprobe.

Die Wahrscheinlichkeit  $\alpha$  dafür, dass ein Stichprobenmittelwert  $\bar{x}$  nicht im Schwankungsbereich liegt, ist  $\alpha = 0,05 = 5\%$  (Signifikanzniveau, Irrtumswahrscheinlichkeit).

$\alpha$  oder  $1 - \alpha$  werden für den Test auf signifikanten Unterschied (oder auf statistisch gesicherte Abweichung) vorgegeben.

$-z_c$  und  $+z_c$  sind die kritischen unteren bzw. Grenzen aus der Tabelle  $D(z)$ .

Hypothesentest im engeren Sinn:  $\mu$ -Bereich gegeben, testen, ob  $\bar{x}$  im Bereich liegt

$\bar{x} \in [\mu_0 - z_c \sigma_{\bar{x}}, \mu_0 + z_c \sigma_{\bar{x}}]$  bedeutet  $H_0$  wird bestätigt:  $\mu = \mu_0$

## 6.8 TYPISCHE HYPOTHESENTESTS

### Beispiel 6.5 (Fortsetzung von Beispiel 6.1)

In einem Betrieb werden Messingstifte mit einer Soll-Länge von  $\mu=30,60$  mm hergestellt. Aus früheren Untersuchungen ist bekannt, dass die Stiftlänge normalverteilt ist und eine Standardabweichung von  $\sigma=1$  [mm] aufweist. Grafik in Abschnitt 6.7.

a) Prüfen Sie, welche Verteilung man hier benutzen kann.

$\bar{X}$  stetig, also  $npq > 9$  nicht prüfen,

$\sigma$  gegeben, STUDENT-t-Verteilung nicht erforderlich.

wir gehen davon aus, dass  $N$  groß ist und deshalb  $n/N \leq 0,05$ , also kein Korrekturfaktor.

b) Bei einer Stichprobe vom Umfang 80 ergab sich eine mittlere Stiftlänge von  $\bar{x}=30,35$  [mm]. Kann die Nullhypothese bei einer Irrtumswahrscheinlichkeit von 5% beibehalten werden?

$$\mu_0 = 30,6 \text{ mm} \quad \sigma = 1 \text{ mm}, \quad \alpha = 0,05, \quad 1-\alpha = 0,95$$

$$D(z_c) = 0,95 = D(1,96) \text{ aus Tabelle 7.6}$$

$$z_{\text{empirisch}} = \frac{\bar{x} - \mu}{\sigma} \sqrt{n} = \frac{30,35 - 30,6}{1} \sqrt{80} = -2,23. \quad |z| = 2,23 > z_c = 1,96$$

$H_0$  wird verworfen, weil  $z_{\text{empirisch}} > z_{\text{kritisch}}$

oder auch über  $[\mu - 1,96 \sigma_{\bar{x}}; \mu + 1,96 \sigma_{\bar{x}}]$

$$\bar{x} = 30,35 \notin [30,6 - 1,96 \cdot 0,112; 30,6 + 1,96 \cdot 0,112] = [30,38; 30,82]$$

Es besteht ein signifikanter Unterschied zwischen der Behauptung und der Messung.

c) Kann die Nullhypothese bei einer Irrtumswahrscheinlichkeit von 1% beibehalten werden?

$$z_{\text{empirisch}} = \frac{30,35 - 30,6}{0,112} = -2,23 \text{ wie oben} \quad |z_{\text{empir}}| = 2,23 < z_c = 2,58$$

$$\text{oder auch: } \bar{x} = 30,35 \in [30,6 - 2,58 \cdot 0,112; 30,6 + 2,58 \cdot 0,112] = [30,32; 30,88]$$

$\Rightarrow$  Aus dem Signifikanzniveau von 1% wird  $H_0$  beibehalten, weil  $z_{\text{empirisch}} < z_{\text{kritisch}}$

Der Unterschied ist nicht hochsignifikant.

d) Formulieren Sie Null- und Alternativ-Hypothese.

$$H_0: \mu = 30,60 \text{ mm gegen die Alternativhypothese } H_1: \mu \neq 30,60 \text{ mm.}$$

### Beispiel 6.6

In einem Betrieb werden Druckfedern mit einer mittleren Verdrehspannung von  $\mu=600$  N mm<sup>-2</sup> gefertigt. Die Verdrehspannung ist normalverteilt, die Varianz liegt erfahrungsgemäß bei  $\sigma^2=900$ . In einer Stichprobe von  $n=60$  Druckfedern wurde eine mittlere Verdrehspannung von  $\bar{x}=592$  N mm<sup>-2</sup> gemessen. Kann die Nullhypothese auf dem Signifikanzniveau von  $\alpha=0,02$  verworfen werden?

$$1 - 0,02 = 0,98 = D(2,33) \Rightarrow z_{\text{crit.}} = 2,33.$$

$$z_{\text{empirisch}} = \frac{\bar{x} - \mu}{\sigma} \sqrt{n} = \frac{592 - 600}{\sqrt{900}} \sqrt{60} = -2,066 < 2,33 \quad |-2,066| < 2,33$$

$$\text{oder auch: } \bar{x} = 592 \in [600 - 2,33 \cdot 3,873; 600 + 2,33 \cdot 3,873] = [591; 609]$$

$\Rightarrow H_0$  wird beibehalten, weil  $z_{\text{empirisch}} < z_{\text{kritisch}}$

## 6.9 HYPOTHESENTEST – VARIANTEN

### 1. Unbekannte Standardabweichung $\sigma$

#### Beispiel 6.7

Nach Angaben eines Autoherstellers soll der Benzinverbrauch eines bestimmten Autos bei einer Geschwindigkeit von 120 km/h bei  $\mu$  9 Liter/100 km liegen. Die Zufallsvariable "Benzinverbrauch" ist näherungsweise normalverteilt. Bei 26 Testfahrten wurde ein Mittelwert von  $\bar{x}$  8,83 l bei einer Standardabweichung von  $s=0,7$  l ermittelt.

- a) Zunächst ist die Anwendbarkeit der Normalverteilung zu prüfen.
- (1) die Zufallsvariable ist stetig, also  $npq > 9$  nicht zu prüfen.
  - (2)  $\sigma$  ist unbekannt, wir benutzen ersatzweise  $s$  aus der Stichprobe.  
 $n < 30$ , also müssen wir die **STUDENT-t-Verteilung** anwenden.
  - (3) es ist eine Stichprobe ohne Zurücklegen,  $N$  ist groß (es werden mit dem Auto noch viele Fahrten durchgeführt),  $n/N < 0,05$ , kein Korrekturfaktor notwendig.

- b) Prüfen Sie, ob sich die Behauptung des Herstellers bei der Irrtums-Wahrscheinlichkeit von 5% aufrechterhalten lässt.  $\alpha = 0,05$   $D(z) = 0,95$   $t_{crit} = 2,06$

$$n=26 \text{ Stück } s=0,7 \text{ l, } \bar{x} = 8,83 \text{ l, Freiheitsgrade } v = 26 - 1 = 25. \sigma_{\bar{x}} = 0,137$$

$$1 - 0,05 = 0,95 = D_{25}(2,06) \Rightarrow t_{crit.} = 2,06.$$

$$t_{\text{empirisch}} = \frac{\bar{x} - \mu}{s} \sqrt{n} = \frac{8,83 - 9}{0,7} \sqrt{26} = -1,238. \quad |-1,238| < 2,06$$

$$\text{oder auch: } \bar{x} = 8,83 \in [9 - 2,06 \cdot 0,137; 9 + 2,06 \cdot 0,137] = [8,72; 9,28]$$

$\Rightarrow$  ja, die Behauptung  $H_0$  wird beibehalten, weil  $t_{\text{empirisch}} < t_{\text{kritisch}}$

Hätte man hier die Standardnormalverteilung benutzt, dann wäre  $z_c = 1,96$

Beim gleichen Sicherheitsgrad gilt:  $z_c < t_c$ .

Die t-Verteilung schützt die Nullhypothese stärker als die Normalverteilung.

- c) Ab welchem Benzinverbrauch wird man auf dem Signifikanzniveau von 0,01 behaupten können, dass der tatsächliche Benzinverbrauch vom behaupteten abweicht?

$$D_{25}(2,787) = 0,99 \Rightarrow t_c = 2,787.$$

$$[9 - 2,787 \cdot 0,137; 9 + 2,787 \cdot 0,137] = [8,62; 9,38]$$

wenn der Benzinverbrauch unter 8,62 l/100 km oder über 9,38 l/km liegt, dann besteht ein signifikanter Unterschied zur Behauptung vor.

### 2. Stichproben ohne Zurücklegen

Normalerweise werden Stichproben ohne Zurücklegen durchgeführt.

Für Auswahlsätze  $n/N > 0,05$  ist ein Korrekturfaktor erforderlich.

Gegenüber den bisherigen Rechnungen ändert sich nur  $\sigma_{\bar{x}}$ :  $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$

## 6.10 AUFGABE HYPOTHESENTEST

### Aufgabe Hypothesentest

gegeben: Sachverhalt wie in den Beispielen 6.5 – 6.7, stetige Zufallsvariable,  $\sigma$  oder  $s$  gegeben, mit oder ohne Zurücklegen,  $\alpha$  oder  $1 - \alpha$  oder Intervall oder Abweichung  $\varepsilon$ .  
Bei Qualitätskontrollen ist immer "ohne Zurücklegen" anzunehmen, auch wenn sonst kein Hinweis vorliegt.

gesucht / Schritte:

1. Welche Verteilung wird verwendet?
  - (1) *stetige Zufallsvariable:*  
 *$npq > 9$  und Stetigkeitskorrektur nicht erforderlich*
  - (2) *prüfen  $n < 30$  bzw.  $n < 50$ , dann  $t$ -Verteilung benutzen*
  - (3) *prüfen "ohne Zurücklegen" und  $n/N > 0,05$ , dann Korrekturfaktor*
2. Nullhypothese bestätigen oder verwerfen?
  - (1)  *$z_{\text{empirisch}} > z_{\text{crit.}}$  ? bzw.  $t_{\text{empirisch}} > t_{\text{crit.}}$  ?*
  - (2) *die gegebenen Größen zusammenstellen,*
  - (3)  *$\sigma_{\bar{x}}$  oder  $\sigma_{\bar{x}, \text{ohne Zurücklegen}}$  berechnen*
  - (3)  *$D(z_c)$  bzw.  $D_v(t_c)$  in der Tabelle ablesen*
  - (4)  *$z_{\text{empirisch}} = \frac{\bar{x} - \mu}{\sigma} \sqrt{n}$  bzw.  $t_{\text{empirisch}} = \frac{\bar{x} - \mu}{s} \sqrt{n}$  berechnen*  
*oder*  
 *$[\mu - z_c \sigma_{\bar{x}} ; \mu + z_c \sigma_{\bar{x}}]$  bzw.  $[\mu - t_c \sigma_{\bar{x}} ; \mu + t_c \sigma_{\bar{x}}]$*   
*prüfen, ob  $\bar{x}$  in diesem Intervall liegt oder nicht.*
  - (5) *Entscheidung: passende Antwort verbal formulieren.*
3. Nullhypothese bestätigen oder verwerfen, bei veränderten Werten  
*Gleiche Vorgehensweise wie bei Nr. 2*
4. Grenzen, ab denen man die Nullhypothese verwerfen kann  
oder Grenzen, bis zur denen man die Nullhypothese beibehalten muss.  
 *$z_c$  bzw.  $t_c$  über  $1 - \alpha = D(z)$  bzw.  $1 - \alpha = D_v(t)$  berechnen,*  
 *$\sigma_{\bar{x}}$  oder  $\sigma_{\bar{x}, \text{ohne Zurücklegen}}$  berechnen*  
 *$[\mu - z_c \sigma_{\bar{x}} ; \mu + z_c \sigma_{\bar{x}}]$  bzw.  $[\mu - t_c \sigma_{\bar{x}} ; \mu + t_c \sigma_{\bar{x}}]$*   
Mit den Grenzen dieses Intervalls passende verbale Antwort geben.
5. Irrtums-Wahrscheinlichkeit  $\alpha$   
*Man bestimmt  $z_c$  bzw.  $t_c$ . Daraus bestimmt man  $\alpha$  über*  
 *$1 - \alpha = D(z)$  bzw.  $1 - \alpha = D_v(t)$ .*

## 6.11 EINSEITIGE HYPOTHESENTESTS

### 1. Einseitige Tests

Wenn man sich nur dafür interessiert, ob das Stichprobenmittel  $\bar{x}$  signifikant kleiner ist als der gegebene Mittelwert der Grundgesamtheit  $\mu_0$ , dann liegt ein einseitiger Test vor.

Der untere kritischen Wert ist  $-z_c$  und die Alternativ-Hypothese heißt  $H_1: \mu < \mu_0$ .

Es gibt dann nur einen (links-seitigen) Ablehnungsbereich mit den Wahrscheinlichkeiten  $F_{SN}(-z)$ .

Ganz analog kann von Bedeutung sein, ob das Stichprobenmittel  $\bar{x}$  signifikant größer ist als der gegebene Mittelwert der Grundgesamtheit  $\mu_0$ , dann liegt ebenfalls ein einseitiger Test vor.

Der obere kritischen Wert ist  $+z_c$  und der Alternativ-Hypothese heißt  $H_1: \mu > \mu_0$ .

Es gibt dann nur einen (rechts-seitigen) Ablehnungsbereich mit den Wahrscheinlichkeiten  $F_{SN}(z)$

#### Beispiel 6.8 (Fortsetzung von Beispiel 6.7)

Nach Angaben eines Autoherstellers soll der Benzinverbrauch eines bestimmten Autos bei einer Geschwindigkeit von 120 km/h unter 9 Liter/100 km liegen. Die Zufallsvariable "Benzinverbrauch" ist näherungsweise normalverteilt. Bei 35 Testfahrten wurde ein Mittelwert von 8,83 Liter bei einer Standardabweichung von 0,7 Liter ermittelt.

a) Zunächst ist die Anwendbarkeit der Normalverteilung zu prüfen.

(1) die Zufallsvariable ist stetig,  $npq > 9$  ist nicht zu prüfen.

(2)  $\sigma$  ist unbekannt,  $n > 30$ , wir benutzen ersatzweise  $s$  aus der Stichprobe.

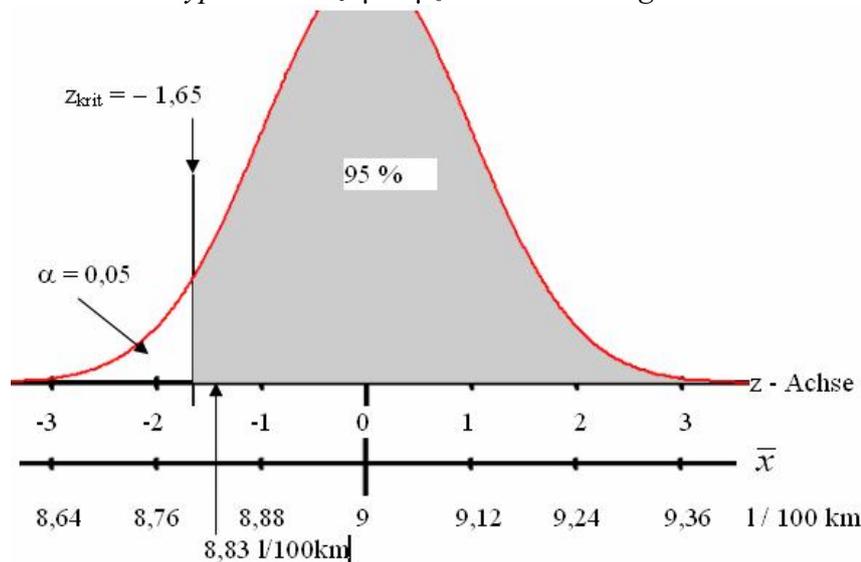
(3) es ist eine Stichprobe ohne Zurücklegen,  $N$  ist groß (es werden mit dem Auto noch viele Fahrten durchgeführt),  $n/N < 0,05$ , kein Korrekturfaktor notwendig.

b) Prüfen Sie, ob sich die Behauptung des Herstellers bei der Irrtumswahrscheinlichkeit von 5% aufrechterhalten lässt.

mit "unter 9 l" ist gemeint "signifikant unter 9 l"

$n=35$   $s=0,7 \Rightarrow \sigma = 0,7$ , weil  $v > 30$ .  $\bar{x} = 8,83$

$\mu_0 = 9$ . Alternativhypothese  $H_1: \mu < \mu_0$ . links-einseitiger Test.



$F_{SN}(-1,64) = 0,505$ ,  $F_{SN}(-1,65) = 0,495$ , interpolierter Zwischenwert  $F_{SN}(-1,645) = 0,05$ .

$\Rightarrow z_c = 1,645$

$$z_{\text{empirisch}} = \frac{\bar{x} - \mu}{\sigma} \sqrt{n} = \frac{8,83 - 9}{0,7} \sqrt{35} = -1,437. \quad |-1,435| < 1,645 \Rightarrow z_{\text{empir.}} < z_{\text{crit.}}$$

oder auch  $8,83 \in [9 - 1,645 \cdot 0,118; 9 + 1,645 \cdot 0,118] = [8,81; 9,19]$

Der Autohersteller kann nicht behaupten, dass das Auto unter 9 l verbraucht.