Kapitel II

Schätztheorie

<u>D. 2. 1</u> (*Schätzung*)
Unter *Schätzung* versteht man die Bestimmung eines Parameters der Grundgesamtheit durch die Statistik einer Stichprobe.

<u>D. 2. 2.</u> (*Punktschätzung, Intervallschätzung*)
Bei der *Punktschätzung* handelt es sich um die Ermittlung eines einzelnen Wertes (*Punktschätzer*) für einen Parameter der Grundgesamtheit. Bei der *Intervallschätzung* wird ein sog. *Konfidenzintervall* bestimmt, in dem ein Parameter der Grundgesamtheit liegen soll (*Intervallschätzer*).

<u>D. 2. 3.</u> (*Konfidenzintervall*)
Ein Konfidenzintervall hat folgende Struktur:

$$\left[\text{Punktschätzer-Fehlerspanne Punktschätzer+Fehlerspanne}\right]$$

<u>B. 2. 1.</u> (*Intervallschätzung des Mittelwertes der Grundgesamtheit*)

Zusammenfassung

Grundgesamtheit	Stichproben-umfang	σ bekannt	σ unbekannt
Normalverteilt	Groß $(n \geq 30)$	$\bar{x} \pm z\sigma_{\bar{x}}$	$\bar{x} \pm ts_{\bar{x}}$ oder $\bar{x} \pm zs_{\bar{x}}$
	Klein $(n < 30)$	$\bar{x} \pm z\sigma_{\bar{x}}$	$\bar{x} \pm ts_{\bar{x}}$
Nicht normalverteilt	Groß $(n \geq 30)$	$\bar{x} \pm z\sigma_{\bar{x}}$	$\bar{x} \pm ts_{\bar{x}}$ oder $\bar{x} \pm zs_{\bar{x}}$
	Klein $(n < 30)$	Nichtparametrische Methoden werden benutzt.	

Hier sind:

n:	Stichprobenumfang
σ:	Standardabweichung der Grundgesamtheit
$\bar{x}$:	Mittelwert der Stichprobe
s:	Standardabweichung der Stichprobe
z:	Kritischer $z-$Wert
t:	Kritischer $t-$Wert

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ bzw. $s_{\bar{x}} = \frac{s}{\sqrt{n}}$: Standardabweichung der Stichprobenverteilung

$z\sigma_{\bar{x}} =: E$ bzw. $ts_{\bar{x}} =: E$: Fehlerspanne

B. 2. 2.

Der $z-$Wert ist abhängig vom *Konfidenzniveau* α und kann in der letzten Zeile der Tabelle der $t-$Verteilung (zweiseitig) gefunden werden

Der $t-$Wert ist ebenfalls abhängig vom *Konfidenzniveau* α und kann entsprechend des *Freiheitsgerades* $n-1$ in einer Zeile der Tabelle der $t-$Verteilung (zweiseitig) gefunden werden.

BS. 2. 1.

Es soll das durchschnittliche Gehalt der Universitätsdozenten in den USA geschätzt werden. Stichproben vom Umfang

a) 36
b) 64
c) 144

ergaben einen Mittelwert von $53145. Die Standardabweichung aller Gehälter der Dozenten wird als $16541 geschätzt.
Die Dozentengehälter sind normalverteilt.

1. Geben Sie ein 95%-Konfidenzintervall jeweils für die Fälle a, b und c an.
2. Wie verhält sich die Fehlerspanne mit zunehmendem Stichprobenumfang?
3. Wie verhält sich die Breite des Konfidenzintervalls mit zunehmendem Stichprobenumfang?

Lösung:

1.

a)

$$E = 1.96 \cdot \frac{16541}{\sqrt{36}} \approx 5403.39$$

$$\mu \in [53145 - 5403.39,\ 53145 + 5403.39] = [47741.60\ 58548.39]$$

Intervalllänge: 10806.79.

b)

$$E = 1.96 \cdot \frac{16541}{\sqrt{64}} \approx 4052.55$$

$$\mu \in [53145 - 4052.55,\ 53145 + 4052.55] = [49092.46,\ 57197.55]$$

Intervalllänge: 8105.09.

c)

$$E = 1.96 \cdot \frac{16541}{\sqrt{144}} \approx 2701.70$$

$$\mu \in [53145 - 2701.70,\ 53145 + 2701.70] = [50443.30,\ 55846.70]$$

Intervalllänge: 5403.40.

2.
Die Fehlerspanne nimmt ab.

3.
Die Intervalllänge nimmt ab.

<u>BS. 2. 2.</u>
Die Zeit, in der eine Kassiererin eines Supermarktes ihre Kunden bedient, ist normalverteilt mit einer Standardabweichung von 0.5 Minuten. Eine Stichprobe von 25 Kunden ergab eine mittlere Bedienzeit von 5.2 Minuten.
Geben Sie ein 99%-Konfidenzintervall für die Bedienzeit aller Kunden.

Lösung:

$$\mu \in \left[5.2 - 2.578 \cdot \frac{0.5}{\sqrt{25}},\ 5.2 + 2.578 \cdot \frac{0.5}{\sqrt{25}}\right] = [4.9422,\ 5.4578].$$

<u>BS. 2. 3.</u>
Die Lebensdauer einer TV-Bildröhre ist nicht normalverteilt. Ihre Standardabweichung beträgt 500 Stunden. Eine Stichprobe von 35 Röhren ergab einen Mittelwert von 8900 Stunden.
Geben Sie ein 95%-Konfidenzintervall für die mittlere Lebensdauer aller Bildröhren an.

Lösung:

$$\mu \in \left[8900 - 1.96 \cdot \frac{500}{\sqrt{35}},\ 8900+1.96 \cdot \frac{500}{\sqrt{35}}\right] = [8734.35,\ 9065.65].$$

<u>BS. 2. 4.</u>
Ein Wissenschaftler untersucht die Menge eines Bestanteils in Lebensmitteln. Eine Stichprobe von 50 Verbrauchern ergab einen Mittelwert von 756 g und eine Standardabweichung von 35 g. Die Grundgesamtheit ist normalverteilt.
Geben Sie ein 95%-Konfidenzintervall für den Mittelwert der Grundgesamtheit an.

$$\mu \in \left[756 - 2.001 \cdot \frac{35}{\sqrt{50}},\ 756 + 2.001 \cdot \frac{35}{\sqrt{50}}\right] = [746.10,\ 765.90].$$

(Approximation durch die Normalverteilung:

$$\mu \in \left[756 - 1.96 \cdot \frac{35}{\sqrt{50}},\ 756 + 1.96 \cdot \frac{35}{\sqrt{50}}\right] = [746.30,\ 765.70].\ )$$

BS. 2. 5.
Der Manager eines Restaurants möchte die durchschnittliche Einnahme des Restaurants in der Woche schätzen. Dazu werden 115 Kunden berücksichtigt. Sie zahlten durchschnittlich 9.74 € mit einer Standardabweichung von 2.93 €. Die Grundgesamtheit ist nicht normal.

1. Bestimmen Sie ein 95%-Konfidenzintervall für die durchschnittliche Ausgabe der Kunden.
2. Konstruieren Sie ein 95%-Konfidenzintervall für die Einnahmen des Restaurants an einem Tag mit 115 Kunden.

Lösung:

1.

$$\mu \in \left[9.74 - 1.984 \cdot \frac{2.93}{\sqrt{115}},\ 9.74 + 1.984 \cdot \frac{2.93}{\sqrt{115}} \right] = [9.1979,\ 10.2820].$$

(Approximation durch die Normalverteilung:

$$\mu \in \left[9.74 - 1.96 \cdot \frac{2.93}{\sqrt{115}},\ 9.74 + 1.96 \cdot \frac{2.93}{\sqrt{115}} \right] = [9.2045,\ 10.2755].$$

2.

$$[9.20 \cdot 115,\ 10.28 \cdot 115] = [1058.00,\ 1182.00])$$

B. 2. 3. (***Bestimmung des Mindestumfangs der Stichprobe***)
Unter der Voraussetzung, dass die Standardabweichung der Grundgesamtheit bekannt ist, gilt:

$$n = \left(\frac{z \cdot \sigma}{E} \right)^2.$$

BS. 2. 6.
Bestimmen Sie die Mindesttrainingszeit für eine bestimmte Operation mit einer Fehlerspanne von 3 Tagen und einer Standardabweichung von 20 Tagen mit 90%-Konfidenz.

Lösung:

$$n = \left(\frac{z \cdot \sigma}{E} \right)^2 = \left(\frac{1.645 \cdot 20}{3} \right)^2 \approx 120.27.$$

Also mindestens 121Tage.

B. 2. 4. (***Die t-Verteilung***)
Die *t-Verteilung* ist eine Wahrscheinlichkeitsfunktion, die 1908 von William Sealy Gosset entwickelt und nach seinem Pseudonym *Student* benannt wurde.

Eine Stetige Zufallsvariable X genügt der *t-Verteilung* mit $n > 0$ *Freiheitgraden*, wenn Sie folgende Wahrscheinlichkeitsdichte besitzt:

$$f_n(x) := \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \qquad -\infty < x < +\infty$$

Dabei ist

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$$

die $\Gamma-$ *Funktion.*

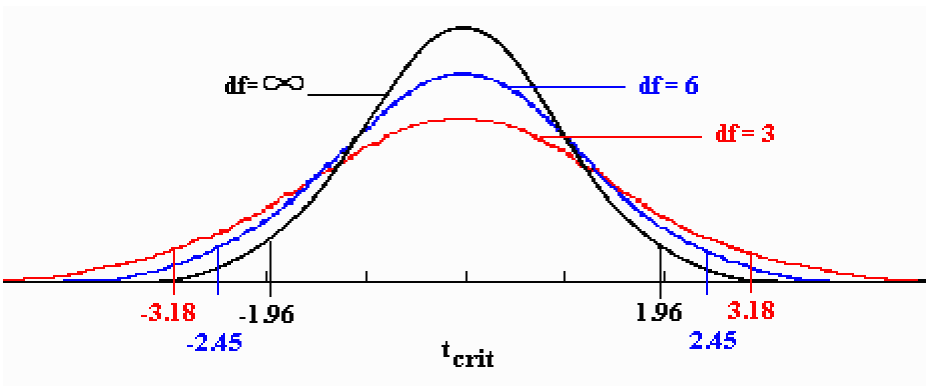
Die *t-Verteilung* hat nur einen Parameter, namens *Freiheitsgrad*

$$df := n-1.$$

Für die *t-Verteilung* gilt:

$$E(X) = 0, \qquad D(X) = \sqrt{df/(df-2)} > 1$$

Für $df \to +\infty$ geht die $t-$*Verteilung* in die *Normalverteilung* über. daher findet man in der letzten Tabelle der *t-Verteilung* die kritischen Werte der standardisierten Normalverteilung:



BS. 2. 7.
In einer Gemeinde mit 10000 Erwerbstätigen wurde durch eine Stichprobe von 100 Personen für die Verteilung der Wochenendverdiente ein Mittelwert von 400 € mit einer Standardabweichung von 50 € festgestellt.
Der Wochenendverdienst aller Erwerbstätigen ist normalverteilt

1. Bestimmen Sie für den Mittelwert der Grundgesamtheit mit einem Konfidenzniveau von

i) 95%

ii) 99%

2. Wie würde sich das 95%-Intervall ändern, wenn eine Stichprobe mit 2000 Erwerbstätigen gezogen worden wäre?

Lösung:
1.

i)
Es ist

$$\bar{x} = 400, \quad s = 50, \quad n = 100, \quad df = 99, \quad t_{99,0.05} = 1.98421695 \approx 1.984$$

$$\mu \in \left[400 - 1.984 \cdot \frac{50}{\sqrt{100}},\ 400 + 1.984 \cdot \frac{50}{\sqrt{100}}\right] \approx [390.08,\ 409.92]$$

ii)
Es ist

$$\bar{x} = 400, \quad s = 50, \quad n = 100, \quad df = 99, \quad t_{99,0.01} = 2.62640546 \approx 2.626$$

$$\mu \in \left[400 - 2.626 \cdot \frac{50}{\sqrt{100}},\ 400 + 2.626 \cdot \frac{50}{\sqrt{100}}\right] \approx [386.87,\ 413,13]$$

2.

Es ist

$$\bar{x} = 400, \quad s = 50, \quad n = 2000, \quad df = 1999, \quad t_{1999,0.05} = 1.96115139 \approx 1.961$$

$$\mu \in \left[400 - 1.961 \cdot \frac{50}{\sqrt{2000}},\ 400 + 1.961 \cdot \frac{50}{\sqrt{2000}}\right] \approx [397.808,\ 402.193]$$

Die Intervalllänge hat sich von 19.84 auf 4.385 verringert.

BS. 2. 8.

Eine Abfüllanlage für Bier soll daraufhin überprüft werden, ob sie korrekt für eine Sollmenge von 1 hl (= 100 Liter) pro Fass eingestellt ist. Eine Stichprobe von 25 Fässern ergab eine durchschnittliche Abfüllmenge von 100.2 Liter bei einer Standardabweichung von 0.28 Liter. Die Abfüllmenge ist normal verteilt.
Wie lautet das 95%-Konfidenzintervall für die durchschnittliche Abfüllmenge?

Lösung:
Es ist

$$\bar{x} = 100.2, \quad s = 0.28, \quad n = 25, \quad df = 24, \quad t_{24,0.05} = 2.06389857 \approx 2.064$$

$$\mu \in \left[100.2 - 2.064 \cdot \frac{0.28}{\sqrt{25}},\ 100.2 + 2.064 \cdot \frac{0.28}{\sqrt{25}}\right] \approx [100.08,\ 100.32].$$

<u>BS. 2. 9.</u>
Eine Stichprobe von 115 Personen zahlten an einem Tag durchschnittlich 9.74 € für das Mittagessen in einem Restaurant mit einer Standardabweichung von 2.93 €.
Ermitteln Sie ein 95%-Konfidenzintervall für das Geld, dass die Personen an einem Tag in diesem Restaurant zahlen.

Lösung:

$$s = 2.93, \quad \bar{x} = 9.74, \quad n = 115, \quad df = 114, \quad t_{114,0.05} = 1.98099229 \approx 1.981$$

$$\mu \in \left[9.74 - 1.981 \cdot \frac{2.93}{\sqrt{115}},\ 9.74 + 1.981 \cdot \frac{2.93}{\sqrt{115}}\right] \approx [9.20,\ 10.28].$$

(Letzte Aktualisierung: 08.06.2017)