Übung 10

-

Grundlagen der schliessenden Statistik

Musterlösung

Aktuelle Version: 30. August 2022

Hinweise:

- Übungen sind mit Vorteil alleine zu lösen.
- Benutzen Sie die Musterlösungen nur zur Korrektur.
- Die Übungen sind wichtige Vorbereitungen für die Prüfung. Lösen sie die Übungen sorgfältig und stellen Sie die Lösungswege übersichtlich dar.
- (Ergänzte) Vorlesungsunterlagen und Fachbücher helfen beim Lösen von Übungen und bringen gleichzeitig eine erweiterte Ansicht auf die Problemstellung.
- Wenn Sie die Übungen nicht verstehen, fragen Sie!

Übung 1. *Fragen*

1. Was ist Zweck der schliessenden Statistik?

   Die schliessende Statistik versucht auf der Basis statistischer Modelle und Daten aus Stichproben zu allgemeinen Aussagen über eine Grundgesamtheit zu gelangen.

2. Wie funktioniert schliessende Statistik?

   Es werden Hypothesen über Verteilungen von Merkmalen getestet, indem die angenommenen Verteilungen mit gemessenen Werten (Stichproben) verglichen werden.

3. Wie ist das Stichprobenmittel verteilt?

   Für Stichproben $n > 30$ ist das Stichprobenmittel näherungsweise normalverteilt, sonst Student t-verteilt.

4. Wie ist die zu schätzende Varianz der Grundgesamtheit verteilt?

   Sie ist χ^2-verteilt.

5. Wie ist der Zusammenhang zwischen χ^2-Verteilung und Normalverteilung?

   Die χ^2-Verteilung Verteilung leitet sich aus der Summe von unabhängigen, quadrierten standardnormalverteilten Zufallsgrössen her.

   Zudem gilt auch für die χ^2-Verteilung der zentrale Grenzwertsatz: Sie nähert sich für eine hohe Anzahl an Freiheitsgraden (n>100) der Normalverteilung an.

6. Was ist ein Konfidenzintervall?

   Das Konfidenzintervall, gibt einen Vertrauensbereich an, in dem der wahre Wert mit einer bestimmten Wahrscheinlichkeit liegt.

Übung 2. *Student t-Verteilung*

Eine Zufallsvariable X sei t-verteilt mit $\nu = 10$ Freiheitsgraden.

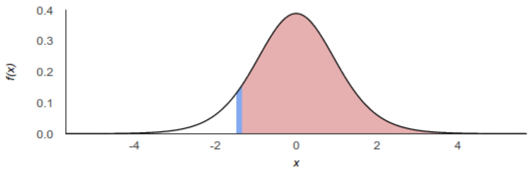
1. Bestimmen Sie die Wahrscheinlichkeit, dass x zwischen -1.4 und 1.8 liegt.

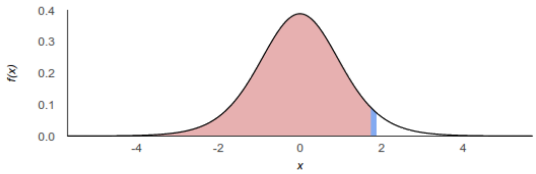
   Aus der Tabelle lesen wir:

   $$F(x < 1.4) \approx 0.9$$
   $$F(x < 1.8) \approx 0.95$$

   Damit ist

   $$\begin{aligned} p(-1.4 < x < 1.8) &= F(x < 1.8) - F(x < -1.4) \\ &= F(x < 1.8) - (1 - F(x < 1.4)) \\ &= 0.95 - (1 - 0.9) = 0.85 \end{aligned}$$





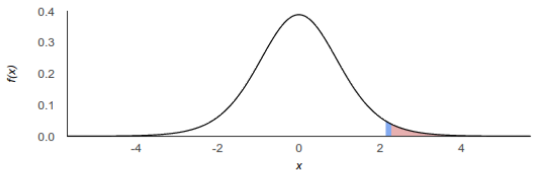
2. Bestimmen Sie die Wahrscheinlichkeit, dass x über 2.2 liegt.

   Aus der Tabelle lesen wir

$$F(x < 2.2) \approx 0.975$$

   Damit ist

$$p(x > 2.2) = 1 - F(x < 2.2) = 1 - 0.975 = 0.025$$



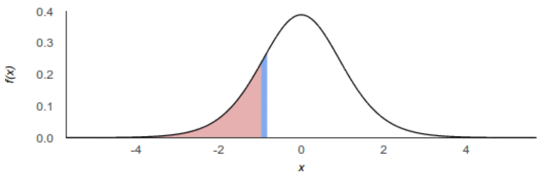
3. Bestimmen Sie die Wahrscheinlichkeit, dass x kleiner als -0.9 ist.

   Aus der Tabelle lesen wir

$$F(x < 0.9) \approx 0.8$$

   Damit ist

$$p(x < -0.9) = 1 - F(x < 0.9) = 1 - 0.8 = 0.2$$



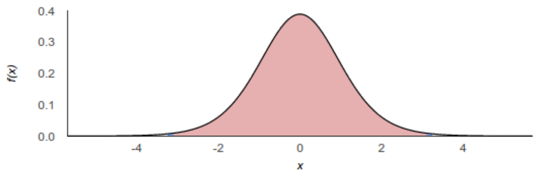
4. In welchem mittleren Bereich liegen ihre Realisationen mit einer Wahrscheinlichkeit von 99%?

   Aus der Tabelle lesen wir

$$F(x > 3.17) = 0.995$$

   Damit ist

$$x = [-3.17, 3.17]$$



5. Wie unterscheiden sich die berechneten Werte, wenn sie statt von einer t-Verteilung von einer Normalverteilung ausgehen?

$$
\begin{aligned}
p(-1.4 < x < 1.8) &= 0.88 \\
p(x > 2.2) &= 0.014 \\
p(x < -0.9) &= 0.18 \\
x &= [-2.58, 2.58]
\end{aligned}
$$

   Die Werte sind für den mittleren Bereich der Kurve ähnlich, unterscheiden sich aber insbesondere am Rand (für kleine ν). Für kleine Stichproben ist die Wahrscheinlichkeit also grösser, Mittelwerte zu erhalten, welche stark vom Populationsmittelwert abweichen.

Übung 3. χ^2-*Verteilung*
Eine Zufallsvariable X sei χ^2-verteilt mit $\nu = 10$ Freiheitsgraden.

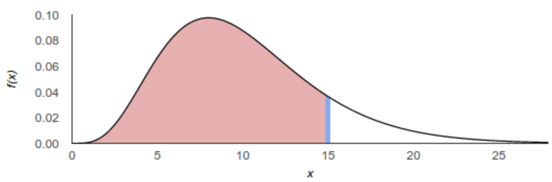
1. Bestimmen Sie die Wahrscheinlichkeit, dass x zwischen 15 und 20 liegt.

   Aus der Tabelle lesen wir:

$$
\begin{aligned}
F(x < 15) &\approx 0.9 \\
F(x < 20) &\approx 0.975
\end{aligned}
$$

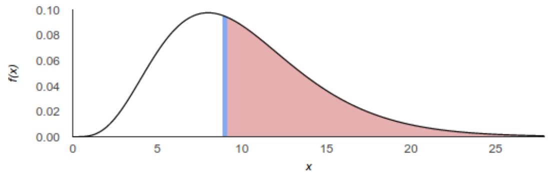
   Damit ist

$$
\begin{aligned}
p(15 < x < 20) &= F(x < 20) - F(x < 15) \\
&= 0.975 - 0.9 = 0.075
\end{aligned}
$$



2. Bestimmen Sie die Wahrscheinlichkeit, dass x über 9 liegt.

   Aus der Tabelle lesen wir
   $$F(x < 9) \approx 0.5$$
   Damit ist
   $$p(x > 9) = 1 - F(x < 9) = 1 - 5 = 0.5$$



3. Wie ändert sich die Verteilung bei höherem ν?

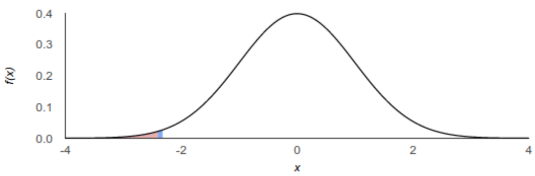
   Ab $\nu \approx 100$ geht die χ^2-Verteilung in eine Normalverteilung über mit $\sigma^2 = 2\nu$.

Übung 4. *Stichprobenmittelwerte*

Sie analysieren eine Produktionseinrichtung. Vom System wissen sie, dass diese Maschine im normalverteilten Mittel 10 Stück pro Sekunde produziert und die Standardabweichung von 1.5 Stück pro Sekunde besitzt. Sie führen 50 Experimente durch.

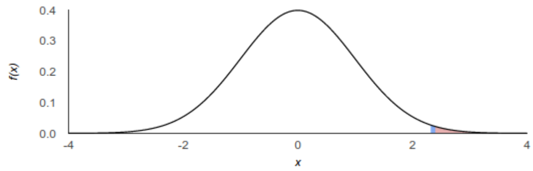
1. Mit welcher Wahrscheinlichkeit wird der mittlere Ausstoss unter 9.5 Stück pro Sekunde liegen?

   Wir suchen die folgende Fläche unter der Kurve mit dem Z-Wert:
   $$Z = \frac{X - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{9.5 - 10}{\frac{1.5}{\sqrt{50}}} = -2.36$$
   Aus der Tabelle lesen wir eine entsprechende Wahrscheinlichkeit von $P(x < 9.5) = F(-2.36) = 0.0091$, d.h. der Mittelwert $\bar{x}$ wird mit einer Wahrscheinlichkeit von 0.91% im Intervall $[\infty, 9.5]$ liegen.



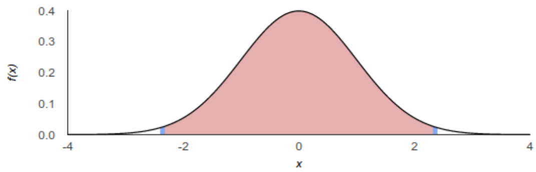
2. Mit welcher Wahrscheinlichkeit wird der Ausstoss über 10.5 Stück pro Sekunde liegen?

   Da die Normalverteilung symmetrisch um 10 Stück pro Sekunde ist, können wir denselben Z-Wert und die Wahrscheinlichkeit verwenden: $P(x > 10.5) = 1 - F(2.36) = F(-2.36) = 0.0091$, d.h. der Mittelwert $\bar{x}$ wird mit einer Wahrscheinlichkeit von 0.91% im Intervall $[10.5, \infty]$ liegen.



3. Mit welcher Wahrscheinlichkeit wird der mittlere Ausstoss zwischen 9.5 und 10.5 Stück pro Sekunde liegen?

   Wir kombinieren die beiden Wahrscheinlichkeiten und erhalten $P(9.5 > x > 10.5) = 1 - 2 \cdot F(-2.36) = 0.9818$, d.h. der Mittelwert $\bar{x}$ wird mit einer Wahrscheinlichkeit von 98.2% im Intervall [9.5, 10.5] liegen.



4. In welchem symmetrischen Intervall liegt der mittlere Ausstoss zu 99%?

   Wir suchen das 99% Konfidenzintervall:

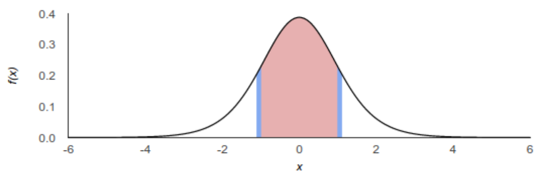
$$\begin{aligned} Z_{(1-\alpha/2)} &= Z_{(0.995)} = 2.58 \\ \bar{x} &= \left[\mu - z\frac{\sigma}{\sqrt{n}}, \mu + z\frac{\sigma}{\sqrt{n}}\right] \\ &= \left[10 - 2.58\frac{1.5}{\sqrt{50}}, 10 + 2.58\frac{1.5}{\sqrt{50}}\right] \\ &= [9.45, 10.54] \end{aligned}$$

5. Wie verändert sich das Ergebnis, wenn Sie stattdessen nur 10 Experimente durchführen?

   Da $n < 30$ ist, müssen wir nun auf die Student t-Verteilung mit $\nu = n-1 = 9$ zurückgreifen. Wir berechnen den t-Wert:

$$t = \frac{X - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{9.5 - 10}{\frac{1.5}{\sqrt{10}}} = -1.05$$

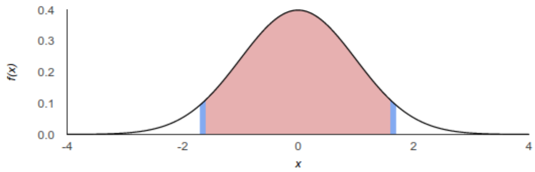
   Aus der Tabelle lesen wir eine entsprechende Wahrscheinlichkeit von $P(x < 9.5) = F(-1.05, 9) = 0.161$, d.h. der Mittelwert $\bar{x}$ wird nur mit einer Wahrscheinlichkeit von 67.8% im Intervall [9.5, 10.5] liegen.



6. Wie viele Experimente müssen Sie durchführen, um zu 90% sicher zu sein, dass der mittlere Ausstoss im Intervall [9.5, 10.5] liegt?

   Wir gehen in einem ersten Schritt von einer beidseitigen Normalverteilung aus:

$$Z_{(1-\alpha/2)} = Z_{(0.95)} = 1.64$$
$$Z = \frac{x-\mu}{\frac{\sigma}{\sqrt{n}}} = \sqrt{n}\frac{x-\mu}{\sigma}$$
$$n = \left(\frac{Z\sigma}{(x-\mu)}\right)^2 = \left(\frac{1.64 \cdot 1.5}{0.5}\right)^2 = 24.2$$



   Es müssen also mindestens 25 Experimente durchgeführt werden.

   Auch mit der Student t-Verteilung erhalten wir in etwa dieselbe Anzahl Experimente. In der Tabelle suchen wir den Wert, welcher die Bedingung erfüllt:

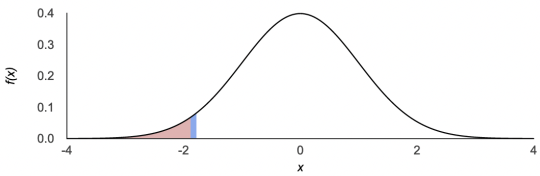
$$t_{(0.95,n)} = \sqrt{n}\frac{x-\mu}{\sigma} = \sqrt{n}\frac{0.5}{1.5} = \frac{\sqrt{n}}{3}$$

   Wir finden $t_{(0.95,26)} \approx 1.7$. Es müssen also mindestens 26 Experimente durchgeführt werden.

Übung 5. *Stichprobenmittelwerte*
Es seien die Ergebnisse bei Prüfungen an einer Hochschule normalverteilt mit einer Standardabweichung von $\sigma = 1.5$ und einem Durchschnitt von $\mu = 4.5$ Notenpunkten. In der aktuellen Prüfungssession finden $n = 30$ Prüfungen statt.

1. Wie hoch ist die Wahrscheinlichkeit, dass der Durchschnitt der Prüfungen kleiner als 4 ist?

$$Z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{4-4.5}{\frac{1.5}{\sqrt{30}}} = -1.83$$
$$P(\bar{x} < 4) = F(-1.83) = 3.36\%$$



Alternativ können wir auch eine t-Verteilung mit $\nu = n - 1 = 29$ Freiheitgraden verwenden und erhalten mit dem bereits berechneten z-Wert $P(\bar{x} < 4) = F(-1.83) = 3.88\%$.

2. Im welchem Intervall liegt der Durchschnitt zu 99%?

$$Z_{(1-\alpha/2)} = Z_{(0.995)} = 2.58$$
$$\bar{x} = \mu \pm z\frac{\sigma}{\sqrt{n}} = 4.5 \pm 2.58\frac{1.5}{\sqrt{30}} = [3.79, 5.21]$$

Alternativ können wir auch eine t-Verteilung mit $\nu = n - 1 = 29$ Freiheitgraden verwenden:

$$Z_{(1-\alpha/2)} = Z_{(0.995)} = 2.76$$
$$\bar{x} = \mu \pm z\frac{\sigma}{\sqrt{n}} = 4.5 \pm 2.76\frac{1.5}{\sqrt{30}} = [3.74, 5.26]$$