

Teil I: Grundlagen

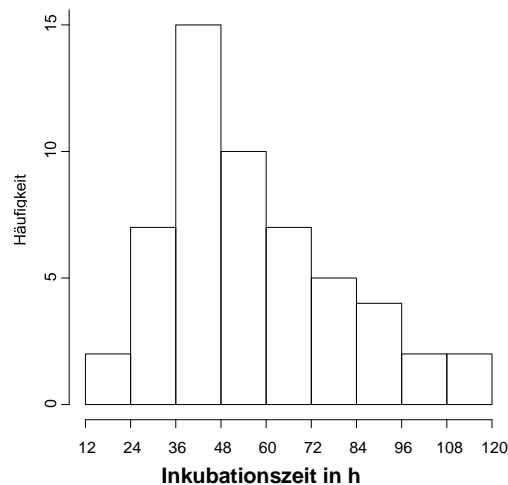
2 Grundlagen der Statistik: Lösungen

Annette Kopp-Schneider und Wiebke Werft

Lösung zu 2.1 Deskriptive Statistik

- Ein mögliches Histogramm sieht wie folgt aus:

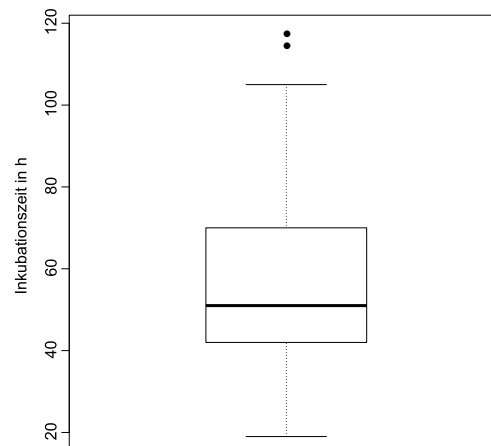
Histogramm der Inkubationszeiten von 54 Angina-Patienten



Es gibt verschiedene Möglichkeiten für die Wahl der Histogrammklassen. Die Einteilung in Klassen als Vielfache von 12 ermöglicht die Interpretation der Inkubationszeit in halben und ganzen Tagen.

- Der arithmetische Mittelwert ergibt sich zu $\bar{x} = 57,46$ h, der Median zu $x_{0,5} = 51$ h, die Spannweite zu $R = 99$ h und die Standardabweichung zu $s = 23,23$ h.
- Ein möglicher Boxplot sieht wie folgt aus:

Boxplot der Inkubationszeiten von 54 Angina-Patienten



Es wurde eine Boxplot-Darstellung gewählt, bei der die Whiskers bis zum kleinsten Wert oberhalb von $Q_1 - 1,5 \cdot IQR$ bzw. bis zum größten Wert unterhalb von $Q_3 + 1,5 \cdot IQR$ reichen. Datenpunkte, die außerhalb der Whiskers liegen, sind durch einzelne Punkte kenntlich gemacht. Wie auch schon im Histogramm zu sehen, ist die Verteilung der Inkubationszeiten schief, und aus dem Boxplot wird klar, dass es zwei auffallend hohe Werte gibt.

Lösung zu 2.2 z-Transformation

Die Größe von erwachsenen Frauen sei $X \sim N(165,49)$ verteilt.

- Der Anteil an Frauen, die kleiner als $x = 158$ cm sind, wird mit der z -Transformation wie folgt bestimmt:

$$z = \frac{x - \mu}{\sigma} = \frac{158 - 165}{7} = -1.$$

Wegen der Symmetrie der Standardnormalverteilung um die 0 ist die Wahrscheinlichkeit, den Wert von $z = -1$ zu unterschreiten, gleich der Wahrscheinlichkeit, den Wert von $z = 1$ zu überschreiten. Die Wahrscheinlichkeit, den Wert von $z = 1$ zu unterschreiten, kann aus Tabelle 2.2 zu 0,841 abgelesen werden. Damit ist der Anteil der Frauen, die kleiner als 158 cm sind, $P(X \leq 158 \text{ cm}) = 1 - 0,841 = 0,159$.

- Die Größe einer Frau, die zu den 5% größten Frauen gehört, ist gerade das 0,95-Quantil dieser Größenverteilung, also von $\mathfrak{N}(165,49)$. Durch Inversion der z -Transformation und Benutzung des 0,95-Quantils der Standardnormalverteilung $z_{0,95} = 1,64$ ergibt sich

$$x = \mu + z_{0,95} \cdot \sigma = 165 \text{ cm} + 1,64 \cdot 7 \text{ cm} = 176,5 \text{ cm}.$$

Somit muss eine Frau 176,5 cm groß sein, damit sie zu den 5% größten Frauen gehört.



Lösung zu 2.3 Konfidenzintervall für den Erwartungswert einer Normalverteilung

- In der Aufgabenstellung fehlt die Angabe von α , daher verwenden wir den Standardwert $\alpha = 5\%$. Aus den Angaben $n = 21$, $\bar{x} = 0,9^\circ\text{C}$, $s = 0,8^\circ\text{C}$ und dem 97,5%-Quantil der t -Verteilung mit 20 Freiheitsgraden von $t_{n-1;1-\alpha/2} = t_{20;0,975} = 2,086$ berechnet sich das 95%-Konfidenzintervall mit der Formel (2.14):

$$\begin{aligned} & \left[\bar{x} - t_{n-1;1-\alpha/2} \frac{s}{\sqrt{n}} ; \bar{x} + t_{n-1;1-\alpha/2} \frac{s}{\sqrt{n}} \right] \\ &= \left[0,9^\circ\text{C} - 2,086 \cdot \frac{0,8^\circ\text{C}}{\sqrt{21}} ; 0,9^\circ\text{C} + 2,086 \cdot \frac{0,8^\circ\text{C}}{\sqrt{21}} \right] \\ &= [0,54^\circ\text{C} ; 1,26^\circ\text{C}]. \end{aligned}$$

- Die wahre Temperaturdifferenz eine Stunde nach Gabe des fiebersenkenden Medikaments liegt mit Wahrscheinlichkeit 95% im (zufälligen, weil von dieser speziellen Stichprobe abhängigen) Intervall $[0,54^\circ\text{C} ; 1,26^\circ\text{C}]$.
- Hier ist der Einstichproben- t -Test durchzuführen. Es wird die Nullhypothese $H_0 : \mu = 0^\circ\text{C}$ gegen die Alternative $H_1 : \mu \neq 0^\circ\text{C}$ getestet, wobei μ den Erwartungswert der Temperaturdifferenz bezeichnet. Die Teststatistik berechnet sich zu

$$t = \frac{\bar{x}}{s/\sqrt{n}} = \frac{0,9^\circ\text{C}}{0,8^\circ\text{C}/\sqrt{21}} = 5,16$$

Da der Wert der Teststatistik 5,16 größer als das 97,5%-Quantil der t -Verteilung mit 20 Freiheitsgraden, $t_{n-1;1-\alpha/2} = t_{20;0,975} = 2,086$, ist, kann die Nullhypothese auf dem 5%-Niveau abgelehnt werden. Es gibt also eine auf dem 5%-Niveau signifikante Temperaturreduktion.



Lösung zu 2.4 Interpretation von p-Werten

- Wenn ein Test ein zum 5%-Niveau signifikantes Ergebnis liefert, ist das Ergebnis nicht unbedingt zum 1%-Niveau signifikant, da der Wert der Teststatistik zwischen dem kritischen Wert für das 5% Niveau und dem für das 1%-Niveau liegen könnte und somit zwar zum 5%-Niveau signifikant wäre, aber nicht zum 1%-Niveau.
- Wenn ein Test ein zum 1%-Niveau signifikantes Ergebnis liefert, ist das Ergebnis auch zum 5%-Niveau signifikant, da der Wert der Teststatistik extremer ist als der kritische Wert für das 1% Niveau, also auch extremer als der kritische Wert für das 5% Niveau.
- Hier vergleicht man den p -Wert mit dem vorher spezifizierten Testniveau: da $p = 0,02 < \alpha = 0,05$ ist, kann die Nullhypothese abgelehnt werden. Es wurde also ein auf dem 5%-Niveau signifikanter Effekt gefunden. Da die Nullhypothese abgelehnt wird, könnte ein Fehler 1. Art aufgetreten sein. Die Wahrscheinlichkeit hierfür beträgt 0,02.
- Auch hier vergleicht man den p -Wert mit dem vorher spezifizierten Testniveau: da $p = 0,2 > \alpha = 0,05$ ist, kann die Nullhypothese nicht abgelehnt werden. Es wurde also kein auf dem 5%-Niveau signifikanter Effekt gefunden. Da die Nullhypothese nicht abgelehnt werden kann, könnte ein Fehler 2. Art aufgetreten sein.

Lösung zu 2.5 Wahl des Hypothesentests

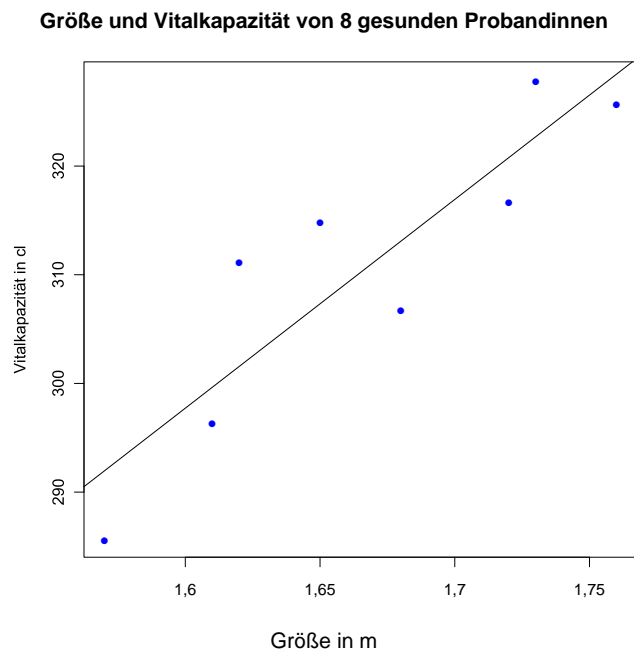
- Hier ist der χ^2 -Test für die Auswertung zu wählen, sodass es sich um einen Vergleich von Anteilen handelt: Es wird die Nullhypothese $H_0: \pi_1 = \pi_2$ gegen die Alternativhypothese $H_1: \pi_1 \neq \pi_2$ getestet, wobei π_i den Anteil der 0“ beim Roulettetisch i bezeichnet.
- Die Vitalkapazität ist eine quantitative Größe, die annähernd normalverteilt ist. Da hier zwei unabhängige Stichproben verglichen werden, ist der Zweistichproben t-Test für die Auswertung zu wählen.
- Die Rötung wird mit einem quantitativen Maß gemessen, von dem wir annehmen, dass es normalverteilt ist. Hier handelt es sich um verbundene Stichproben (linker und rechter Arm desselben Probanden), daher ist der t -Test für verbundene Stichproben zu wählen. Faktisch wird für jede Testperson die Differenz der Rötungswerte zwischen dem linken und dem rechten Arm berechnet und es wird der Einstichproben t-Test mit diesen Differenzen durchgeführt.



- Hier handelt es sich um eine Auswertung von Zeitangaben. Zeitangaben sind typischerweise schief verteilt, mit vielen kleinen Werten und ein paar langen Zeitdauern (vgl. z.B. auch Aufgabe 2.1). Daher kann man nicht von einer normalverteilten Variablen ausgehen. Da hier die Dauer von Reparaturen von zwei verschiedenen Werkstätten, also zwei unabhängige Stichproben verglichen werden sollen, ist ein nicht-parametrischer Test, speziell der Wilcoxon-Rangsummentest, für die Auswertung zu wählen.

Lösung zu 2.6 Korrelation und Regression

- Die Daten werden als Punktwolke grafisch dargestellt. Dabei ist zu beachten, dass in der Aufgabenstellung die Einheit für die Vitalkapazität fälschlicherweise als [l] statt als [cl] angegeben ist:



- Der Korrelationskoeffizient nach Pearson ergibt sich zu $r = 0,893$. Der Wert der Teststatistik berechnet sich zu

$$t = r \sqrt{\frac{n-2}{1-r^2}} = 0,893 \cdot \sqrt{\frac{8-2}{1-(0,893)^2}} = 4,849$$

Das 0,975-Quantil der t -Verteilung mit $8 - 2 = 6$ Freiheitsgraden ist $t_{n-2; 1-\frac{\alpha}{2}} =$

$t_{6; 0,975} = 2,447$. Da der Absolutwert der Teststatistik das Quantil überschreitet, gibt es eine auf dem 5%-Niveau signifikante Korrelation zwischen den beiden Variablen.

Auswertung mit Statistikprogrammen ergibt einen p -Wert von $p = 0,003$, der erheblich



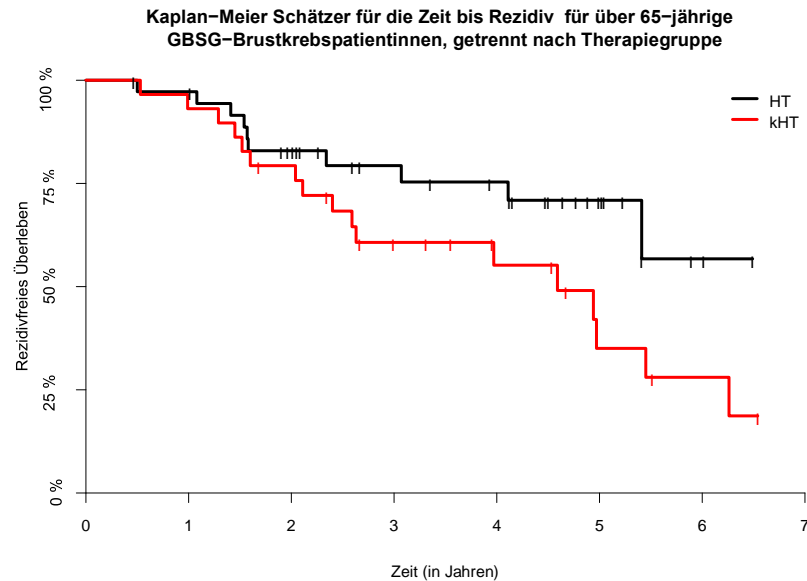
kleiner als 5% ist. Die Nullhypothese, dass die beiden Variablen unkorreliert sind, kann also auf dem 5%-Niveau verworfen werden.

- Es wird eine Ausgleichsgerade $y = \beta_0 + \beta_1 \cdot x$ durch die Daten gelegt. Die Parameter der linearen Regressionsgeraden berechnen sich zu $\beta_0 = -9,6 \text{ cl}$ und $\beta_1 = 192,1 \text{ cl/m}$. Die Ausgleichsgerade ist in die Punktwolke eingezeichnet.
- Für eine Frau von 1,70m Größe ist eine Vitalkapazität von $-9,6 \text{ cl} + 192,1 \frac{\text{cl}}{\text{m}} \cdot 1,7 \text{ m} = 316,9 \text{ cl} = 3,169 \text{ l}$ zu erwarten.



Lösung zu 2.7 Ereignisdaten

- Die grafische Darstellung der Daten beruht auf dem Kaplan-Meier-Schätzer der Zeit bis Rezidiv. Getrennt nach Therapiegruppen ergibt sich folgendes Bild:



- Der Logrank-Test wird benutzt, um zu testen, ob es einen Unterschied in der Zeit bis Rezidiv für Brustkrebspatientinnen, die eine Hormontherapie erhalten haben (HT), und Patientinnen, die keine Hormontherapie erhalten haben (kHT), gibt. Von 37 Patientinnen in der Gruppe mit Hormontherapie erlitten 10 ein Rezidiv, während von 29 Patientinnen in der Gruppe ohne Hormontherapie 17 ein Rezidiv hatten. Der Wert der Logrank-Teststatistik ist 3,2. Die Teststatistik folgt einer χ^2_1 -Verteilung mit 0,95-Quantil 3,84. Daher kann die Nullhypothese, dass es zwischen den beiden Gruppen keinen Unterschied im rezidivfreien Überleben gibt, nicht zum 5%-Niveau abgelehnt werden. Der p -Wert beträgt 0,07. Die grafischen Darstellung und die recht geringe Fallzahl legen nahe, dass hier ein Fehler 2. Art gemacht wird.