MEP ASTAT MM H22 (als Probeprüfung)

Modul-End-Prüfung Angewandte Statistik, Herbst 2022

Ergebnisse von Testdurchlauf 1 für Flavio Waser

Evento-ID: Evento:1567881

Testdurchlauf beendet am: 29. Dez 2023, 16:37

Detaillierte Testergebnisse für Testdurchlauf 1

Reihenfolge	Fragen-ID	Fragentitel	Maximale Punktezahl	Erreichte Punkte	Prozent gelöst
1	800293	Streudiagramm Matrix	4	1	25.00 %
2	800284	Korrelation und R ²	4	0	0.00 %
3	800290	Einfache Lineare Regression	2	2	100.00 %
4	800292	Eindimensionaler Datensatz	4	2	50.00 %
5	800285	Stochastische Unabhängigkeit	4	4	100.00 %
6	800294	Multiple lineare Regression	4	0	0.00 %
7	800288	Zentraler Grenzwertsatz	4	0	0.00 %
8	800289	Einfache Lineare Regression	2	0	0.00 %
9	800286	Boxplot	4	0	0.00 %
10	800291	<u>Hypothesentest</u>	4	2	50.00 %
11	800295	Satz von Bayes	4	0	0.00 %
12	800287	Normalverteilung	2	0	0.00 %

Te	Festergebnis in Punkten: 11 von 42 (26.19 %)		
	<u></u>		
D	Detaillierte Testergebnisse für Testdurchlauf 1		

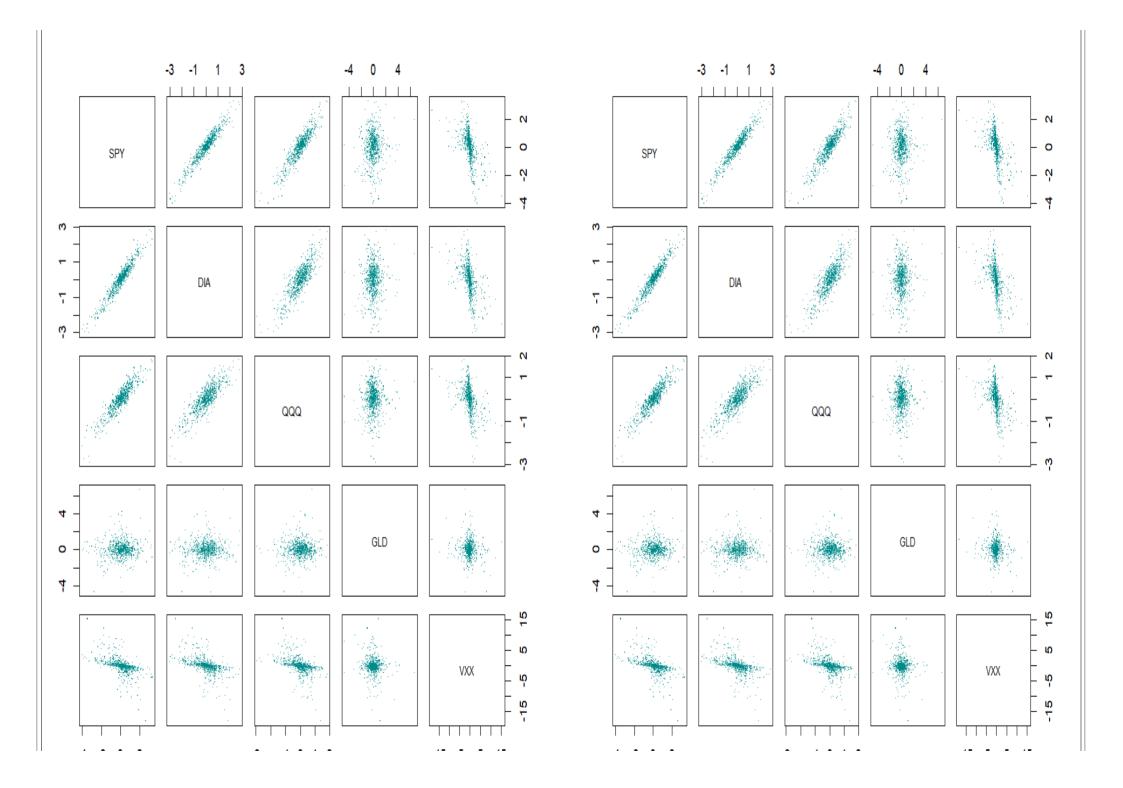
1. Streudiagramm Matrix [ID: 800293]

Ihre Antwort:

Im Datensatz EFTs sind Tagesrenditen von den fünf börsengehandelten Indexfonds SPY, DIA, QQQ, GLD und VXX gespeichert. Mit der Funktion pairs() wurden Streudiagramme für die Variablen des Datensatzes EFTs erstellt.

Bestmögliche Lösung:

Im Datensatz EFTs sind Tagesrenditen von den fünf börsengehandelten Indexfonds SPY, DIA, QQQ, GLD und VXX gespeichert. Mit der Funktion pairs() wurden Streudiagramme für die Variablen des Datensatzes EFTs erstellt.



Ordnen Sie die folgenden Werte korrekt zu:

round(cor(EFTs\$QQQ, EFTs\$GLD), 2)

passt zu

0.04

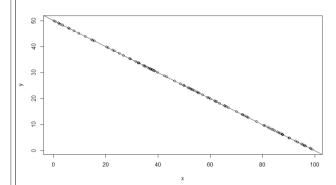
Ordnen Sie die folgenden Werte korrekt zu:

round(cor(EFTs\$SPY, EFTs\$DIA), 2) passt zu 0.95 round(cor(EFTs\$QQQ, EFTs\$GLD), 2) passt zu 0.04 round(cor(EFTs\$SPY, EFTs\$VXX), 2) passt zu -0.55

round(cor(EFTs\$QQQ, EFTs\$DIA), 2) passt zu 0.83

2. Korrelation und \mathbb{R}^2 [ID: 800284]

Ihre Antwort:



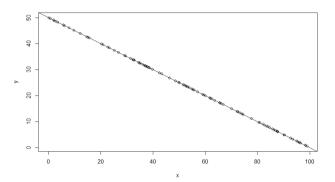
In einem Scatterplot werden die Grössen x und y gegenübergestellt. Alle Punkte liegen auf einer absteigenden Geraden, wobei diese Gerade auch den durch die lineare Regression ermittelten Schätzwerten für y entspricht. Für diese Schätzung wurden \mathbb{R}^2 und RSS berechnet.

Welche Aussagen sind hier wahr?

Für jede Aussage muss entschieden werden: [richtig] oder [falsch]

richtig	falsch	
0	0	der RSS beträgt 1.0.
8		- 2
0	0	der ${ m R}^2$ beträgt -1.0.
8	_	B: K 1 .:
0	0	Die Korrelation zwischen X und Y be-
		trägt -1.
	0	Gemäss dem Regressionsmodell hat
	~	die Grösse x Einfluss auf y.
		•

Bestmögliche Lösung:



In einem Scatterplot werden die Grössen x und y gegenübergestellt. Alle Punkte liegen auf einer absteigenden Geraden, wobei diese Gerade auch den durch die lineare Regression ermittelten Schätzwerten für y entspricht. Für diese Schätzung wurden R^2 und RSS berechnet.

Welche Aussagen sind hier wahr?

richtig	falsch	
0	•	der RSS beträgt 1.0.
0	•	der R^2 beträgt -1.0.
0	0	Die Korrelation zwischen X und Y beträgt -1.
•	0	Gemäss dem Regressionsmodell hat die Grösse x Einfluss auf y.

3. Einfache Lineare Regression [ID: 800290]

Ihre Antwort:

Eine einfache lineare Regression ergibt die Gerade $\hat{y}=32+0.4x$. Eines der Subjekte, Elisabet, hat x=60 und y=52 . Berechne das Residuum von Elisabeth.



Bestmögliche Lösung:

Eine einfache lineare Regression ergibt die Gerade $\hat{y}=32+0.4x$. Eines der Subjekte, Elisabet, hat x=60 und y=52 . Berechne das Residuum von Elisabeth.

Der Wert muss zwischen -4.01 und -3.99 liegen

4. Eindimensionaler Datensatz [ID: 800292]

Ihre Antwort:

In der Abbildung sehen wir die numerischen Werte eines eindimensionalen Datensatzes als dünne Striche auf der Zahlenachse dargestellt. Die Daten wurden mit einem Gerät gemessen, das Werte auf Millimeter rundet.

Bestmögliche Lösung:

In der Abbildung sehen wir die numerischen Werte eines eindimensionalen Datensatzes als dünne Striche auf der Zahlenachse dargestellt. Die Daten wurden mit einem Gerät gemessen, das Werte auf Millimeter rundet.





Für jede Aussage muss entschieden werden: [richtig] oder [falsch]

richtig	falsch	
⊙	0	Aus der graphischen Darstellung ist immer klar, wie viele Messungen der Datensatz enthält.
⊗		
0	⊚	Die Quartilsdifferenz der Daten ändert sich nicht, wenn wir einen Wert im Datensatz (egal welchen) abändern.
\odot		
0	0	Der Mittelwert der Daten ändert sich, wenn wir einen Wert im Daten- satz (egal welchen) abändern.
\odot		
⊙	0	Die Daten sind nicht normalverteilt.
	⊗⊘⊘⊘⊘⊘	

richtig	falsch	
0	•	Aus der graphischen Darstellung ist immer klar, wie viele Messungen der Datensatz enthält.
0	⊚	Die Quartilsdifferenz der Daten ändert sich nicht, wenn wir einen Wert im Datensatz (egal welchen) abändern.
•	0	Der Mittelwert der Daten ändert sich, wenn wir einen Wert im Da- tensatz (egal welchen) abändern.
0	0	Die Daten sind nicht normalverteilt.

5. Stochastische Unabhängigkeit [ID: 800285]

Ihre Antwort:

Angenommen die Wahrscheinlichkeit, dass es an einem beliebigen Tag regnet (Ereignis R) sei 0.05.

Die Wahrscheinlichkeit, dass der Rasen an einem beliebigen Tag gesprengt wird (Ereignis G) sei 0.1, wobei der Rasen nie gesprengt wird, wenn es an diesem Tag auch regnet.

Welche Aussagen sind wahr?

Für jede Aussage muss entschieden werden: [richtig] oder [falsch]

richtig	falsch	
0	•	R und G sind stochastisch unabhängig.
⊙	0	$P(R \cup G) = 0.15$
	0	P(R G) = 0
⊙	0	P(G R) = 0
⊘		

Bestmögliche Lösung:

Angenommen die Wahrscheinlichkeit, dass es an einem beliebigen Tag regnet (Ereignis R) sei 0.05 .

Die Wahrscheinlichkeit, dass der Rasen an einem beliebigen Tag gesprengt wird (Ereignis G) sei 0.1, wobei der Rasen nie gesprengt wird, wenn es an diesem Tag auch regnet.

Welche Aussagen sind wahr?

richtig	falsch	
0	•	R und G sind stochastisch
		unabhängig.
•	0	$P(R \cup G) = 0.15$
•	0	P(R G) = 0
•	0	P(G R) = 0

6. Multiple lineare Regression [ID: 800294]

Ihre Antwort:

Ein Datensatz zu Prüfungsergebnissen der Berufsmaturität wird untersucht. Es wurde ein multiples lineares Modell mit der Zielvariablen BMSchnitt (Abschlussnote von 1 bis 6) und den Prädiktoren IQ, Motivation, Grösse und Geschlecht angepasst. Der R Output sieht wie folgt aus:

call:

lm(formula = BMSchnitt ~ IQ + Motivation + Größe + Geschlecht, data = BM)

Residuals:

Min 1Q Median 3Q Max -0.70065 -0.13915 0.01281 0.16483 0.49899

Coefficients:

Estimate Std. Error t value Pr(>|t|)(Intercept) 0.305073 0.714277 0.427 0.671 8.533 4.91e-11 *** 0.037845 0.004435 ΙQ 6.119 1.92e-07 *** Motivation 0.150587 0.024609 Größe -0.434140 0.336441 -1.290 0.203 GeschlechtMann -0.068264 0.090344 -0.756 0.454 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '1

Residual standard error: 0.2743 on 46 degrees of freedom Multiple R-squared: 0.9056, Adjusted R-squared: 0.8974 F-statistic: 110.4 on 4 and 46 DF, p-value: < 2.2e-16

Für jede Aussage muss entschieden werden: [richtig] oder [falsch]

richtig falsch

Gemäss den Angaben (Signifikanz ausser acht gelassen) schliessen

Bestmögliche Lösung:

Ein Datensatz zu Prüfungsergebnissen der Berufsmaturität wird untersucht. Es wurde ein multiples lineares Modell mit der Zielvariablen BMSchnitt (Abschlussnote von 1 bis 6) und den Prädiktoren IQ, Motivation, Grösse und Geschlecht angepasst. Der R Output sieht wie folgt aus:

```
Call:
```

lm(formula = BMSchnitt ~ IQ + Motivation + Größe + Geschlecht, data = BM)

Residuals:

Min 1Q Median 3Q Max -0.70065 -0.13915 0.01281 0.16483 0.49899

Coefficients:

```
Estimate Std. Error t value Pr(>|t|)
(Intercept)
               0.305073
                         0.714277
                                            0.671
               0.037845
                         0.004435
                                   8.533 4.91e-11 ***
ΙQ
Motivation
               0.150587
                         0.024609
                                   6.119 1.92e-07 ***
Größe
              -0.434140
                         0.336441 -1.290
                                            0.203
GeschlechtMann -0.068264
                         0.090344 -0.756
                                            0.454
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 0.2743 on 46 degrees of freedom Multiple R-squared: 0.9056, Adjusted R-squared: 0.8974 F-statistic: 110.4 on 4 and 46 DF, p-value: < 2.2e-16

Für jede Aussage muss entschieden werden: [richtig] oder [falsch]

richtig falsch

Frauen schle	echter ab als Männer. O	Wenn der IQ um eine Einheit zu- nimmt, steigt der BMSchnitt um	ausser acht	n Angaben (Signifikanz gelassen) schliessen lechter ab als Männer.	
8	0	0.037845 Prozent. Die Prädiktorvariable Motivation	0	•	Wenn der IQ um eine Einheit zu- nimmt, steigt der BMSchnitt um 0.037845 Prozent.
8		liegt ausserhalb des Bereichs, bei dem von einer zufälligen Verteilung zwischen Prädiktorvariable und Ziel- variable ausgegangen werden kann.	•	0	Die Prädiktorvariable Motivation liegt ausserhalb des Bereichs, bei dem von einer zufälligen Verteilung zwischen Prädiktorvariable und Zielvariable ausgegangen werden
0	0	Der Anteil der Varianz in den Daten, der durch dieses Modell erklärt wird, beträgt ca. 90 Prozent.	0	0	kann. Der Anteil der Varianz in den Daten, der durch dieses Modell erklärt wird, beträgt ca. 90 Prozent.

7. Zentraler Grenzwertsatz [ID: 800288]

Ihre Antwort:

Die Einkommensverteilung in einigen Ländern der Dritten Welt gilt als keilförmig (viele sehr arme Menschen, sehr wenige Menschen mit mittlerem Einkommen und noch weniger wohlhabende Menschen). Angenommen, wir wählen ein Land mit einer keilförmigen Verteilung aus. Lassen Sie das durchschnittliche Gehalt 2000 \$ pro Jahr mit einer Standardabweichung von 8000 \$ betragen. Wir befragen zufällig 1000 Einwohner dieses Landes und erhalten einen Stichprobenmittelwert \bar{X} . Berechnen Sie die Wahrscheinlichkeit

 $P(2100 < \bar{X} < 2200)$

auf mindestens vier Stellen nach dem Komma genau (maximal 15 Zeichen erlaubt).



Bestmögliche Lösung:

Die Einkommensverteilung in einigen Ländern der Dritten Welt gilt als keilförmig (viele sehr arme Menschen, sehr wenige Menschen mit mittlerem Einkommen und noch weniger wohlhabende Menschen). Angenommen, wir wählen ein Land mit einer keilförmigen Verteilung aus. Lassen Sie das durchschnittliche Gehalt 2000 $\$ pro Jahr mit einer Standardabweichung von $\$ 8000 $\$ betragen. Wir befragen zufällig $\$ 1000 Einwohner dieses Landes und erhalten einen Stichprobenmittelwert $\$ \bar{X}. Berechnen Sie die Wahrscheinlichkeit

 $P(2100 < \bar{X} < 2200)$

auf mindestens vier Stellen nach dem Komma genau (maximal 15 Zeichen erlaubt).

Der Wert muss zwischen 0.13155 und 0.13185 liegen

8. Einfache Lineare Regression [ID: 800289]

Ihre Antwort:

Wir führen eine einfache, lineare Regression mit R aus, wobei wir $H_0:\beta_1=0$ gegen $H_A:\beta_1\neq 0$ testen. Die R Ausgabe gibt einen p— Wert von 0.015 an. Wir schliessen daraus dass

- A. $H_0: \beta_1 = 0$ ist wahr mit Wahrscheinlichkeit 0.015
- B. $H_0: \beta_1 = 0$ wird verworfen mit $\alpha = 0.05$
- **C.** $H_A: \beta_1 \neq 0$ ist wahr mit Wahrscheinlichkeit 0.015
- **D.** $H_A: \beta_1 \neq 0$ wird verworfen mit $\alpha = 0.05$
- E. Beides, A und B.
- O A 😵
- о в **⊗** о с **⊗**
- ⊙ D 🐼
- O E 🕃

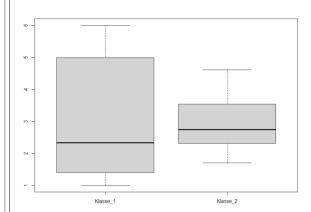
Bestmögliche Lösung:

Wir führen eine einfache, lineare Regression mit R aus, wobei wir $H_0: \beta_1=0 \text{ gegen } H_A: \beta_1\neq 0 \text{ testen. Die R Ausgabe gibt einen } p\text{--Wert von 0.015 an. Wir schliessen daraus dass}$

- A. $H_0: \beta_1 = 0$ ist wahr mit Wahrscheinlichkeit 0.015
- B. $H_0: \beta_1 = 0$ wird verworfen mit $\alpha = 0.05$
- **C.** $H_A: \beta_1 \neq 0$ ist wahr mit Wahrscheinlichkeit 0.015
- D. $H_A: \beta_1 \neq 0$ wird verworfen mit $\alpha = 0.05$
- E. Beides, A und B.
- O A
- B
- O C
- O D
- O E

9. Boxplot [ID: 800286]

Ihre Antwort:

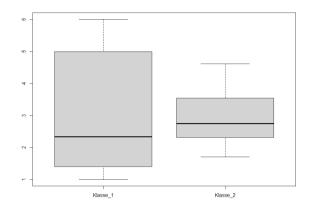


Betrachten Sie den Boxplot oben für die Noten zweier Schulklassen. Beantworten Sie, ob die unteren Aussagen wahr oder falsch sind.

Für jede Aussage muss entschieden werden: [richtig] oder [falsch]

richtig	falsch	
0	0	Die Median-Note von Klasse 2 ist höher als die von Klasse 1.
⊘		
0	•	Die Quartilsdifferenz von Klasse 1 ist kleiner 2.
⊘		
0	0	Beide Datensätze sind normalverteilt.
⊗		
0	•	Beide Datensätze enthalten keine Ausreisser.
6		

Bestmögliche Lösung:



Betrachten Sie den Boxplot oben für die Noten zweier Schulklassen. Beantworten Sie, ob die unteren Aussagen wahr oder falsch sind.

richtig	falsch	
⊙	0	Die Median-Note von Klasse 2 ist
		höher als die von Klasse 1.
0	•	Die Quartilsdifferenz von Klasse 1
		ist kleiner 2.
0	•	Beide Datensätze sind
		normalverteilt.
•	0	Beide Datensätze enthalten keine
		Ausreisser.

10. Hypothesentest [ID: 800291]

Ihre Antwort:

Die National Collegiate Athletic Assocation hat ein neues Trainingsprogram entwickelt, das die Sprunghöhe von Basketballspielern erhöhen soll. Um die Wirksamkeit des Programs zu testing, wurden zufällig 12 College Basketballspieler rekrutiert, und deren Sprunghöhe vor und nach dem einmonatigen Training gemessen.

Die folgenden Daten zeigen für jeden Spieler die maximale Sprunghöhe (in Zoll) vor und nach dem Training:

Vorher: 22, 24, 20, 19, 19, 20, 22, 25, 24, 23, 22, 21

Nachher: 23, 25, 20, 24, 18, 22, 23, 28, 24, 25, 24, 20

Der folgende Code zeigt wie der Hypothesentest in R ausgefürt wurde:

Bestmögliche Lösung:

Die National Collegiate Athletic Assocation hat ein neues Trainingsprogram entwickelt, das die Sprunghöhe von Basketballspielern erhöhen soll. Um die Wirksamkeit des Programs zu testing, wurden zufällig 12 College Basketballspieler rekrutiert, und deren Sprunghöhe vor und nach dem einmonatigen Training gemessen.

Die folgenden Daten zeigen für jeden Spieler die maximale Sprunghöhe (in Zoll) vor und nach dem Training:

Vorher: 22, 24, 20, 19, 19, 20, 22, 25, 24, 23, 22, 21

Nachher: 23, 25, 20, 24, 18, 22, 23, 28, 24, 25, 24, 20

Der folgende Code zeigt wie der Hypothesentest in R ausgefürt wurde:

Beantworten Sie die folgenden Fragen als Richtig oder Falsch:

Für jede Aussage muss entschieden werden: [richtig] oder [falsch]

falsch

0

richtig

ricitas	laiseil	
0	0	Die Nullhypothese sagt aus, dass das Program wirksam ist.
⊗		
0	•	Da der Vertrauensintervall als 95% angegeben ist, wird alpha = 0.025 verwendet.
⊘		
0	•	Der p-Wert liegt über alpha und die Nullhypothese wird verworfen.
⊘		

```
#define before and after max jump heights
before <- c(22, 24, 20, 19, 19, 20, 22, 25, 24, 23, 22, 21)
after <- c(23, 25, 20, 24, 18, 22, 23, 28, 24, 25, 24, 20)

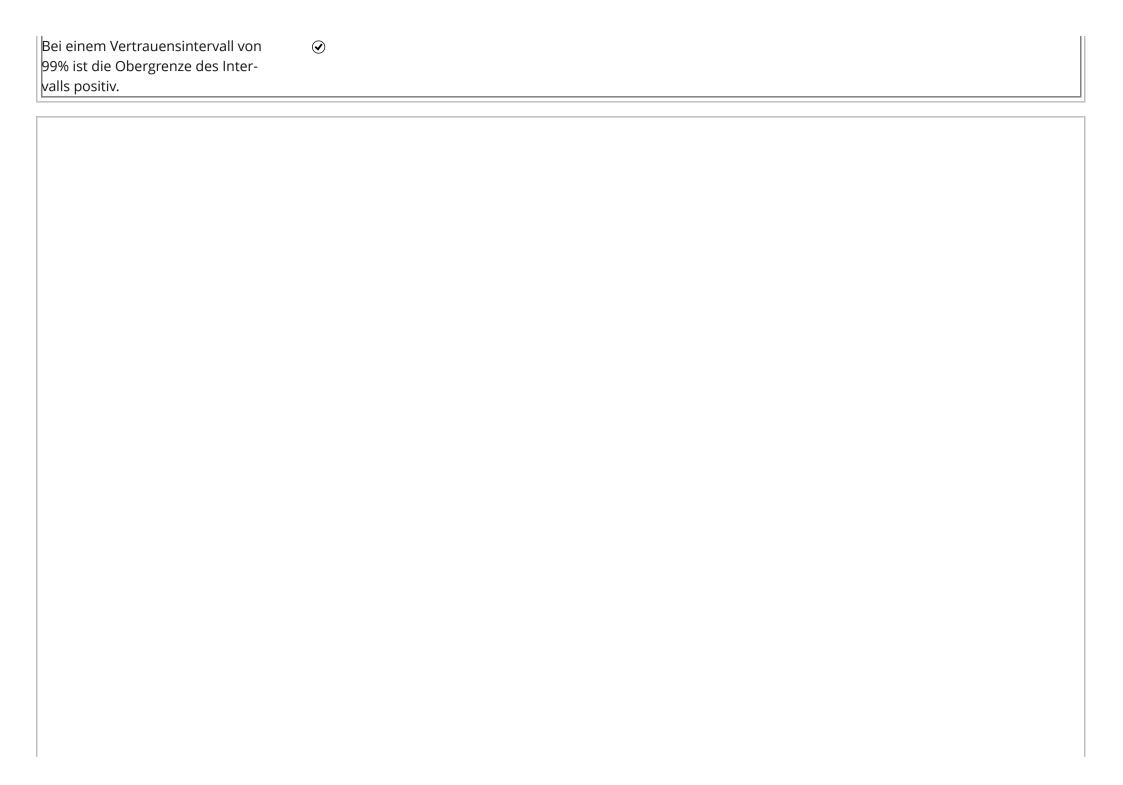
#perform paired samples t-test
t.test(x = before, y = after, paired = TRUE)

Paired t-test

data: before and after
t = -2.5289, df = 11, p-value = 0.02803
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-2.3379151 -0.1620849
sample estimates:
mean of the differences
-1.25</pre>
```

Beantworten Sie die folgenden Fragen als Richtig oder Falsch:

richtig	falsch	
0	•	Die Nullhypothese sagt aus, dass
		das Program wirksam ist.
0	•	Da der Vertrauensintervall als 95%
		angegeben ist, wird alpha = 0.025
		verwendet.
0	•	Der p-Wert liegt über alpha und die
		Nullhypothese wird verworfen.
•	0	Bei einem Vertrauensintervall von
		99% ist die Obergrenze des Inter-
		valls positiv.



11. Satz von Bayes [ID: 800295]

Ihre Antwort:

Aus einer Zusammenstellung des Erfolgs der Wetterprognosen ist bekannt, dass der Wetterbericht

- mit Wahrscheinlichkeit $\frac{1}{4}$ schönes Wetter fälschlicherweise als schlecht voraussagt und
- mit Wahrscheinlichkeit $\frac{4}{5}$ schlechtes Wetter korrekt als schlecht voraussagt.

Eine Person glaubt mit Wahrscheinlichkeit $\frac{2}{3}$, dass das Wetter Morgen schön wird. Dann hört sie den Wetterbericht sagen, dass es schlecht wird. Wie sollte die Person dadurch ihre subjektive Wahrscheinlichkeit für schönes Wetter ändern?

Wir bezeichnen mit A das Ereignis, dass es Morgen schön sein wird und mit B das Ereignis, dass der Wetterbericht schlechtes Wetter voraussagt.

Die Person erhält fünf Vorschläge für diese neue Wahrscheinlichkeit:

- Vorschlag (a): $P(A|B) = \frac{\frac{2}{3} \cdot \frac{1}{4}}{\frac{2}{3} \cdot \frac{1}{4} + \frac{1}{3} \cdot \frac{4}{5}}$ Vorschlag (b): $P(A|B) = \frac{2}{3} \cdot \frac{1}{4}$ Vorschlag (c): $P(A|B) = \frac{\frac{2}{3} \cdot \frac{1}{4}}{\frac{2}{3} \cdot \frac{1}{4} + \frac{3}{4} \cdot \frac{1}{5}}$

- Vorschlag (d): $P(A|B) = \frac{2}{3} \cdot \frac{4}{5}$ Vorschlag (e): $P(A|B) = \frac{\frac{2}{3}}{\frac{1}{4} + \frac{4}{5}}$

Welche dieser Rechnungen ist für die neue Wahrscheinlichkeit korrekt?

- Vorschlag (d)
- Vorschlag (c) 🔀

Bestmögliche Lösung:

Aus einer Zusammenstellung des Erfolgs der Wetterprognosen ist bekannt, dass der Wetterbericht

- mit Wahrscheinlichkeit $\frac{1}{4}$ schönes Wetter fälschlicherweise als schlecht voraussagt und
- mit Wahrscheinlichkeit $\frac{4}{5}$ schlechtes Wetter korrekt als schlecht voraussagt.

Eine Person glaubt mit Wahrscheinlichkeit $\frac{2}{3}$, dass das Wetter Morgen schön wird. Dann hört sie den Wetterbericht sagen, dass es schlecht wird. Wie sollte die Person dadurch ihre subjektive Wahrscheinlichkeit für schönes Wetter ändern?

Wir bezeichnen mit A das Ereignis, dass es Morgen schön sein wird und mit B das Ereignis, dass der Wetterbericht schlechtes Wetter voraussagt.

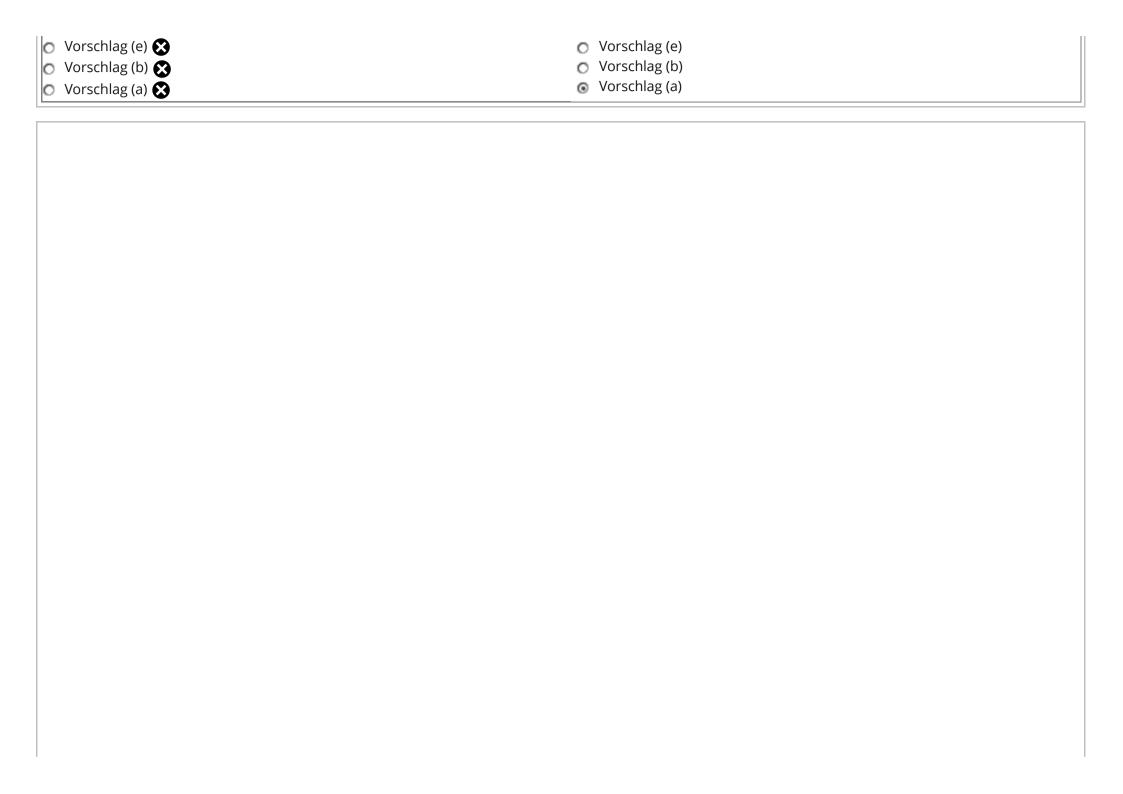
Die Person erhält fünf Vorschläge für diese neue Wahrscheinlichkeit:

- Vorschlag (a): $P(A|B) = \frac{\frac{2}{3} \cdot \frac{1}{4}}{\frac{2}{3} \cdot \frac{1}{4} + \frac{1}{3} \cdot \frac{4}{5}}$ Vorschlag (b): $P(A|B) = \frac{\frac{2}{3} \cdot \frac{1}{4}}{\frac{2}{3} \cdot \frac{1}{4} + \frac{3}{4} \cdot \frac{1}{5}}$ Vorschlag (c): $P(A|B) = \frac{\frac{2}{3} \cdot \frac{1}{4}}{\frac{2}{3} \cdot \frac{1}{4} + \frac{3}{4} \cdot \frac{1}{5}}$

- Vorschlag (d): $P(A|B) = \frac{2}{3} \cdot \frac{4}{5}$ Vorschlag (e): $P(A|B) = \frac{\frac{2}{3}}{\frac{1}{4} + \frac{4}{5}}$

Welche dieser Rechnungen ist für die neue Wahrscheinlichkeit korrekt?

- Vorschlag (d)
- Vorschlag (c)



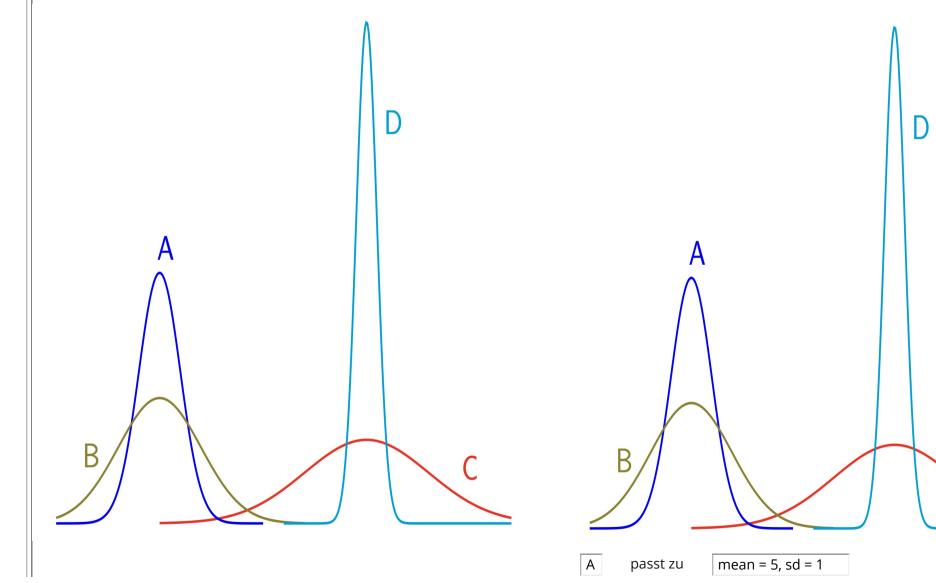
12. Normalverteilung [ID: 800287]

Ihre Antwort:

Ordnen Sie jedem Parameterpaar die richtige Normalverteilungskurve zu. Dazu ziehen Sie mit der Maus das passende Parameterpaar auf den Buchstaben auf der linken Seite.

Bestmögliche Lösung:

Ordnen Sie jedem Parameterpaar die richtige Normalverteilungskurve zu. Dazu ziehen Sie mit der Maus das passende Parameterpaar auf den Buchstaben auf der linken Seite.



В	passt zu	mean = 5, sd = 2	
С	passt zu	mean = 15, sd = 3	
D	passt zu	mean = 15, sd = 0.5	