Wirtschaftsstatistik-Klausur am 03.07.2018

Aufgabe 1
Statistik-Dozent K.R. lehrt an einer privaten FH in Köln, wohnt aber in Frankfurt am Main. Er hat - wegen möglicher saisonaler Schwankungen - zwei Semester lang ausprobiert, ob es für ihn - nur aus dem Blickwinkel der arbeitstäglichen Fahrzeit (berechnet vom Verlassen seiner Wohnung bis zum Betreten des FH-Gebäudes in Köln) - günstiger ist, mit dem privaten PKW oder mit dem Intercity-Express (ICE) der Deutschen Bahn AG zu fahren.

60 Fahrten mit dem PKW und 40 Fahrten mit dem ICE ergaben folgende Fahrzeiten:

Benötigte Fahrzeit	Anzahl der Fahrten mit dem ...	
von ... bis unter ... Minuten	... eigenen PKW	... ICE
110 - 120	5	-
120 - 130	15	1
130 - 150	20	9
150 - 180	10	20
180 - 240	10	10
$\sum$	60	40

Für welches der beiden Verkehrsmittel wird er sich auf Grund der von ihm gesammelten Daten entscheiden, wenn er

a) nach dem Kriterium <u>kürzere Fahrzeit</u> entscheidet?

b) nach dem Kriterium <u>gleichmäßigere Fahrzeit</u> entscheidet?

Begründen Sie Ihre Antwort durch Berechnung und Interpretation geeigneter statistischer Maßzahlen!

Aufgabe 2
Bei einem Produktionsprozess will man anhand von 0,95-Konfidenzintervallen Aufschlüsse über die unbekannte wahre Ausschussquote von Maschine A und von Maschine B erhalten. Dazu werden aus der Produktion von Maschine A und aus der Produktion von Maschine B Stichproben erhoben und anschließend die Ausschussanteile in den Stichproben berechnet:

Maschine	Umfang der Stichprobe	Ausschussanteil in der Stichprobe
A	120	10%
B	100	12%

a) Berechnen Sie für die wahre Ausschussquote von Maschine A ein 0,95-Konfidenzintervall und interpretieren Sie anschließend das erhaltene Intervall.

b) Berechnen Sie für die wahre Ausschussquote von Maschine B ein 0,95-Konfidenzintervall und interpretieren Sie anschließend das erhaltene Intervall.

Lösung zu Aufgabe 1:
X = benötigte Fahrzeit mit PKW (in Minuten)
Y = benötigte Fahrzeit mit ICE (in Minuten)

a) Kriterium kürzere Fahrzeit

Benötigte Fahrzeit von ... bis unter ... Minuten	Klassen-mitte	Anzahl der Fahrten mit dem ... eigenen PKW	Dichte	Anzahl der Fahrten mit dem ... ICE	Dichte
110 - 120	115	5	0,0083	-	-
120 - 130	125	15	0,0250	1	0,0025
130 - 150	140	20	0,0167	9	0,0113
150 - 180	165	10	0,0056	20	0,0167
180 - 240	210	10	0,0028	10	0,0042
$\sum$		60		40	

arithmetisches Mittel:

$$\overline{x} \approx 115 \cdot \frac{5}{60} + 125 \cdot \frac{15}{60} + 140 \cdot \frac{20}{60} + 165 \cdot \frac{10}{60} + 210 \cdot \frac{10}{60} = \frac{9\,000}{60} = 150$$

$$\overline{y} \approx 125 \cdot \frac{1}{40} + 140 \cdot \frac{9}{40} + 165 \cdot \frac{20}{40} + 210 \cdot \frac{10}{40} = \frac{6\,785}{40} = 169{,}625 \approx 170$$

d.h. PKW-Fahrten sind im Durchschnitt kürzer.

50%-Punkte:

$x_{0,50} \approx 140$

$y_{0,50} \approx 165$

d.h. in 50% aller PKW-Fahrten lag die Fahrzeit unter 140 Minuten, während in 50% aller ICE-Fahrten die Fahrzeit weniger als 165 Minuten betrug; also sind die PKW-Fahrzeiten kürzer.

Modus:

$x_{\text{Modus}} \approx 125$

$y_{\text{Modus}} \approx 165$

d.h. PKW-Fahrzeiten sind kürzer.

b) Kriterium gleichmäßigere Fahrzeit

Da die Durchschnittswerte der PKW-Daten und der ICE-Daten „weit“ auseinander liegen, sollten hier der Variationskoeffizient oder die relative Quartilsabstand berechnet werden.

Variationskoeffizient:

$$s_x^2 \approx (115-150)^2 \cdot \frac{5}{60} + (125-150)^2 \cdot \frac{15}{60} + (140-150)^2 \cdot \frac{20}{60} + (165-150)^2 \cdot \frac{10}{60} + (210-150)^2 \cdot \frac{10}{60} = 929{,}16$$

$$s_x = \sqrt{929{,}16} = 30{,}48$$

$$v_x = \frac{s_x}{\overline{x}} = \frac{30{,}48}{150} = 0{,}2032$$

$$s_y^2 \approx (125-170)^2 \cdot \frac{1}{40} + (140-170)^2 \cdot \frac{9}{40} + (165-170)^2 \cdot \frac{20}{40} + (210-170)^2 \cdot \frac{10}{40} = 665{,}48$$

$s_y = \sqrt{665{,}48} = 25{,}80$
$v_y = \frac{s_y}{\overline{y}} = \frac{25{,}80}{170} = 0{,}1518$
d.h. die Fahrzeiten mit dem ICE sind gleichmäßiger.

Relativer Quartilsabstand:
$\frac{x_{0,75} - x_{0,25}}{x_{0,50}} \approx 0{,}27$

$\frac{y_{0,75} - y_{0,25}}{y_{0,50}} \approx 0{,}18$
d.h. die Fahrzeiten mit dem ICE sind gleichmäßiger.

Lösung zu Aufgabe 2:

a) $\left[0{,}1 - 1{,}96 \cdot \sqrt{\frac{0{,}1 \cdot 0{,}9}{120}}; 0{,}1 + 1{,}96 \cdot \sqrt{\frac{0{,}1 \cdot 0{,}9}{120}}\right] =$
$[0{,}1 - 1{,}96 \cdot 0{,}0274; 0{,}1 + 1{,}96 \cdot 0{,}0274] =$
$[0{,}1 - 0{,}0537; 0{,}1 + 0{,}0537] =$
$[0{,}0463; 0{,}1537]$
$[4{,}63\%; 15{,}37\%]$ ist ein geschätzter Bereich für das Intervall, in dem die wahre Ausschussquote von Maschine A mit Wahrscheinlichkeit 0,95 liegt.

b) $\left[0{,}12 - 1{,}96 \cdot \sqrt{\frac{0{,}12 \cdot 0{,}88}{100}}; 0{,}12 + 1{,}96 \cdot \sqrt{\frac{0{,}12 \cdot 0{,}88}{100}}\right] =$
$[0{,}12 - 1{,}96 \cdot 0{,}0325; 0{,}12 + 1{,}96 \cdot 0{,}0325] =$
$[0{,}12 - 0{,}0637; 0{,}12 + 0{,}0637] =$
$[0{,}0563; 0{,}1637]$
$[5{,}63\%; 16{,}37\%]$ ist ein geschätzter Bereich für das Intervall, in dem die wahre Ausschussquote von Maschine B mit Wahrscheinlichkeit 0,95 liegt.

QM III-Klausur am 03.07.2018

Aufgabe 1

Ein Lebensversicherungsunternehmen verwendet für seine Kalkulationen unter anderem die folgende Kalkulationsgrundlage:

- 85 Prozent der 50-jährigen Männer werden 65 Jahre alt.

Diese Angabe lässt sich wie folgt verstehen: ein 50-jähriger Mann erreicht das Alter 65 (Erfolg) mit einer Wahrscheinlichkeit von 0,85. Im Gegenzug erreicht er mit einer Wahrscheinlichkeit von 0,15 dieses Alter nicht (Mißerfolg).

a) Wie groß ist die Wahrscheinlichkeit, dass von fünf 50-jährigen Versicherungsnehmern

1. alle fünf,
2. keiner,
3. mindestens vier

65 Jahre alt werden? Nennen Sie die Wahrscheinlichkeitsverteilung und ihre Parameter, wenn davon ausgegangen werden kann, dass die Lebensdauern der Versicherungsnehmer (VN) unabhängig voneinander sind!

b) Wie groß ist die Wahrscheinlichkeit dafür, dass von 1 000 Versicherungsnehmern im Alter von 50 Jahren mindestens 855 das Alter 65 erreichen? Welche Wahrscheinlichkeitsverteilung können Sie zur Approximation verwenden? Nennen Sie diese und geben Sie auch deren Paramenter an!

Aufgabe 2

a) Welche Testentscheidung trifft der t-Test zum Niveau 0,05, wenn das arithmetische Mittel in der Stichprobe mit μ_0 der Nullhypothese übereinstimmt?

b) Überprüfen Sie anhand der nachfolgenden Stichprobe vom Umfang 40 mit einem geeigneten Test zum Niveau 0,05, ob die Dauer X (in Sekunden) eines Downloads im Mittel sieben Sekunden beträgt. Stichprobe: $\sum_{i=1}^{40} x_i = 313{,}7933$ und $\sum_{i=1}^{40} x_i^2 = 2\,626{,}259$. Gehen Sie wie folgt vor:

1. Wie heißt der Test?
2. Wie lautet die Nullhypothese des Tests?
3. Überprüfen Sie, ob die Faustregel des Tests erfüllt ist.
4. Berechnen Sie den p-Wert des Tests.
5. Wie lautet die Testentscheidung aufgrund der obigen Stichprobe? (Begründung!) Interpretieren Sie in knappen Worten das Ergebnis.
6. Interpretieren Sie den Fehler 1. Art.

7. Führen Sie ggf. einen einseitigen Test zum Niveau 0,05 durch und interpretieren Sie die Testentscheidung.

Lösung zu Aufgabe 1:

a) Die gesuchten Wahrscheinlichkeiten betragen mit X als „Anzahl der 50-jährigen Versicherungsnehmer, die mindestens 65 Jahre alt werden“ und damit einer Binomialverteilung mit den Parametern $n = 5$ und $p = 0{,}85$.

   1. $P(\text{fünf 65-jährige VN}) = P(X = 5) = \binom{5}{5} \cdot 0{,}85^5 \cdot 0{,}15^0 = 0{,}4437$
      d.h. die gesuchte Wahrscheinlichkeit beträgt 0,4437.
   2. $P(\text{kein 65-jähriger VN}) = P(X = 0) = \binom{5}{0} \cdot 0{,}85^0 \cdot 0{,}15^5 = 0{,}000076$
      d.h. die gesuchte Wahrscheinlichkeit beträgt 0,000076.
   3. $P(X \geq 4) = P(X = 4) + P(X = 5) = \binom{5}{4} \cdot 0{,}85^4 \cdot 0{,}15^4 + 0{,}4437 = 0{,}3915 + 0{,}4437 = 0{,}8352$
      d.h. die gesuchte Wahrscheinlichkeit beträgt 0,8352.

b) Die Anzahl X der 50-jährigen Versicherungsnehmer, die mindestens 65 Jahre alt werden, ist auch hier binomialverteilt, jedoch mit den Parametern $n = 1\,000$ und $p = 0{,}85$. Diese Verteilung darf aufgrund des hohen Berechnungsaufwandes mit der Normalverteilung approximiert werden. Diese sollte denselben Erwartungswert und dieselbe Varianz wie die Binomialverteilung haben, also $\mu = n \cdot p = 850$ und $\sigma^2 = np \cdot (1-p) = 127{,}5$. Die Faustregel $n \cdot p = 850 \geq 10$ und $n \cdot (1-p) = 150 \geq 10$ ist erfüllt. Mittels der Normalverteilungsapproximation und der Tabelle der Standardnormalverteilung ergibt sich dann:
   $P(X \geq 855) = 1 - P(X < 855) = 1 - P(X \leq 854) \approx 1 - F_U\left(\frac{854 - +0{,}5 - 850}{\sqrt{127{,}5}}\right) = 1 - F_U(0{,}39085) = 1 - 0{,}655 = 0{,}345$
   d.h. die Wahrscheinlichkeit beträgt näherungsweise 0,345.

Lösung zu Aufgabe 2:

a) H_0 wird nicht abgelehnt, da gilt: p-Wert$= 2 \cdot 0{,}5 = 1$.

b) X=Dauer (in Sekunden) eines Downloads
   $n = 40$
   $\overline{x} = \frac{313{,}7933}{40} = 7{,}844834$
   $s^2 = \frac{2\,626{,}259}{40} - 7{,}844834^2 = 4{,}115063$, $s = \sqrt{4{,}115063} = 2{,}028562$

   1. Der Test heiß t-Test.
   2. $H_0 : E[X] = 7$
   3. Die Faustregel $n = 40 \geq 30$ ist erfüllt.

4. p-Wert $= 2 \cdot F_U \left(- \left| \frac{7{,}844834 - 7}{\frac{2{,}028562}{\sqrt{40}}} \right| \right) = 2 \cdot F_U(-2{,}6340) = 2 \cdot 0{,}004 = 0{,}008$
5. $0{,}008 \leq 0{,}05$, d. h. H_0 wird abgelehnt; d. h. die mittlere Dauer eines Downloads weicht signifikant von sieben Sekunden ab.
6. Der Test behauptet fälschlicherweise, die mittlere Dauer eines Downloads würde nicht sieben Sekunden betragen.
7. p-Wert $= 0{,}5 \cdot 0{,}008 = 0{,}004 \leq 0{,}05$; d. h. die mittlere Dauer eines Downloads ist signifikant länger als sieben Sekunden.