Lösungen Probeklausur I

Statistische Verfahren in der Geographie

Till Straube <straube@geo.uni-frankfurt.de>

Institut für Humangeographie
Goethe-Universität Frankfurt

Aufgabe 1

Bestimmen Sie das Skalenniveau der folgenden Variablen. *(5 Punkte)*

Kürzen Sie ab: N = Nominalskala; O = Ordinalskala; I = Intervallskala; V = Verhältnisskala

	Variable	Skalenniveau
a)	Verkaufsfläche (gemessen in Quadratmeter)	V
b)	Zufriedenheit mit der eigenen Wohnsituation (erfasst entlang einer mehrstufigen Skala mit den Polen „sehr unzufrieden“ = 0 und „sehr zufrieden“ = 10)	I
c)	Familienstand (erfasst mit den Merkmalsausprägungen „ledig“, „verheiratet“, „geschieden“, „verwitwet“)	N
d)	Lagerbestände an Rohöl (gemessen in Barrel; 1 Barrel = 159 Liter)	V
e)	Subjektive Schichteinstufung (erfasst mit den Schichteinstufungen „Unterschicht“, „untere Mittelschicht“, „Mittelschicht“, „obere Mittelschicht“, „Oberschicht“)	O

Sind die folgenden Aussagen richtig oder falsch? Kreuzen Sie das entsprechende Feld an. *(5 Punkte)*

	Aussage	richtig	falsch
f)	Ein Boxplot zeigt in der Regel an, wo das arithmetische Mittel einer Verteilung von Werten liegt.		×
g)	Wer nachweisen möchte, dass die Mittelwerte zweier Stichproben sich nicht signifikant unterscheiden, möchte die Nullhypothese beibehalten.	×	
h)	Es ist möglich, dass bei einer Verteilung von Werten Median und Modus aufeinander fallen.	×	
i)	Je größer der Stichprobenumfang, desto enger wird das Konfidenzintervall des Mittelwerts ausfallen.	×	
j)	Ähnliche arithmetische Mittel bei zwei Variablen bedeuten immer, dass auch die Streuungen der Merkmalsausprägungen der Variablen ähnlich sind.		×

Geben Sie an, welcher Signifikanztest zur Beantwortung der unten stehenden Fragestellungen bzw. Untersuchungsabsichten angemessen ist. *(5 Punkte)*

Verwenden Sie dafür folgende Zahlen: 1 = z-Test bzw. 1-Stichproben-t-Test; 2 = 2-Stichproben-t-Test; 3 = F-Test; 4 = χ^2-Test

	Fragestellung	Testverfahren
k)	Unterscheidet sich die durchschnittliche jährliche Niederschlagsmenge an der Klimastation Zugspitze signifikant von der durchschnittlichen jährlichen Niederschlagsmenge in Deutschland?	1
l)	Sie untersuchen, ob es einen signifikanten Unterschied in der Beurteilung des Einzelhandelsangebots in der Frankfurter Innenstadt (gut, mittel, schlecht) in Bezug auf das Alter der Befragten (bis 25 Jahre, 26–65 Jahre, über 65 Jahre) gibt.	4
m)	Sie möchten in Erfahrung bringen, ob die Kaufpreise für Eigentumswohnungen signifikant stärker im Stadtzentrum als im Umland der Stadt variieren.	3
n)	Sie fragen sich, ob das durchschnittliche Einkommen von Auszubildenden in Frankfurt am Main sich von dem durchschnittlichen Einkommen von Auszubildenden in Mainz signifikant unterscheidet.	2
o)	Sie untersuchen entlang einer Umfrage, ob ein signifikanter Zusammenhang zwischen der subjektiven Schichteinstufung und der Form des Schulabschlusses auszumachen ist.	4

Aufgabe 2

Die Stadt Flussmünde verfügt über zwei öffentliche Bibliotheken: Die Bibliothek Hafenviertel und die Bibliothek Waldsiedlung. Gegeben ist die Anzahl der Ausleihvorgänge in den beiden Bibliotheken, die in sechs zufällig ausgewählten Kalenderwochen im Jahr 2018 verzeichnet wurden:

Kalenderwoche	Bibliothek Hafenviertel	Bibliothek Waldsiedlung
KW 18	152	169
KW 24	157	161
KW 26	165	164
KW 32	307	168
KW 37	144	176
KW 46	155	173

a) Welche der beiden Bibliotheken weist über die zufällig ausgewählten Kalenderwochen hinweg im Durchschnitt mehr Ausleihvorgänge auf? *(5 Punkte)*

Hafenviertel:

$$\begin{aligned}\bar{x} &= \frac{\sum\limits_{i=1}^{n} x_i}{n} \\ &= \frac{1080}{6} \\ &= 180\end{aligned}$$

Waldsiedlung:

$$\begin{aligned}\bar{x} &= \frac{\sum\limits_{i=1}^{n} x_i}{n} \\ &= \frac{1011}{6} \\ &= 168{,}5\end{aligned}$$

Die Bibliothek Hafenviertel weist im Schnitt mehr Ausleihvorgänge auf.

---

b) In welcher der beiden Bibliotheken schwanken über die zufällig ausgewählten Kalenderwochen hinweg die Ausleihvorgänge stärker? Berechnen Sie Varianzen und Standardabweichungen. *(10 Punkte)*

Hafenviertel:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{19588}{5} = 3917{,}6$$

$$s = \sqrt{s^2} = \sqrt{3917{,}6} \approx 62{,}59$$

Waldsiedlung:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{153{,}5}{5} = 30{,}7$$

$$s = \sqrt{s^2} = \sqrt{30{,}7} \approx 5{,}54$$

In der Bibliothek Hafenviertel schwanken die Ausleihvorgänge stärker.

Aufgabe 3

Die „Vereinigung der Gewerbetreibenden der Stadt Flussmünde" kommt zu dem Ergebnis: Für einen Einkauf auf dem städtischen Wochenmarkt benötigen die Kund*innen einen durchschnittlichen Anfahrtsweg von 8,9 Kilometern mit einer Standardabweichung von 2,8 Kilometern. Dieser Wert wurde auf Basis einer Zufallsstichprobe von 300 Kund*innen berechnet.

Darüber hinaus sei bekannt: Die Standardabweichung von 2,8 Kilometer entspricht der Standardabweichung in der Population, und die Anfahrtswege sind annähernd normal verteilt.

a) Mit welcher Wahrscheinlichkeit kommen die in der Stichprobe erfassten Kund*innen aus der Stadt Flussmünde oder den umliegenden Orten, d. h. hier: beträgt ihr Anfahrtsweg 12 Kilometer oder weniger? *(5 Punkte)*

---

z-Transformation:

$$\begin{aligned} z_p &= \frac{x_p - \bar{x}}{s} \\ &= \frac{12 - 8{,}9}{2{,}8} \\ &\approx 1{,}11 \end{aligned}$$

Unterschreitungswahrscheinlichkeit:

$$\begin{aligned} P(x < x_p) &= P(z < z_p) \\ &\approx P(z < 1{,}11) \\ &\approx 0{,}8665 \end{aligned}$$

Die Wahrscheinlichkeit, dass der Anfahrtsweg weniger als 12 km beträgt, ist ca. 86,65%.

---

b) Mit welcher Wahrscheinlichkeit wird von den in der Stichprobe erfassten Kund*innen ein Anfahrtsweg von 16 Kilometern überschritten? *(5 Punkte)*

---

z-Transformation:

$$\begin{aligned} z_p &= \frac{x_p - \bar{x}}{s} \\ &= \frac{16 - 8{,}9}{2{,}8} \\ &\approx 2{,}54 \end{aligned}$$

Überschreitungswahrscheinlichkeit:

$$\begin{aligned} P(x > x_p) &= 1 - P(x < x_p) \\ &= 1 - P(z < z_p) \\ &\approx 1 - P(z < 2{,}54) \\ &\approx 1 - 0{,}9945 \\ &= 0{,}0055 \end{aligned}$$

Die Wahrscheinlichkeit, dass der Anfahrtsweg mehr als 16 km beträgt, ist ca. 0,55%.

---

c) Geben Sie den Bereich an, in dem der durchschnittliche Anfahrtsweg *aller* Kund*innen mit einer Sicherheitswahrscheinlichkeit von 95 Prozent liegt. *(5 Punkte)*

Gefragt ist das Konfidenzintervall für eine Intervallschätzung vom Mittelwert μ. Standardfehler:

$$\begin{aligned}\sigma_{\bar{x}} &= \frac{\sigma}{\sqrt{n}} \\ &= \frac{2{,}8}{\sqrt{300}} \\ &\approx 0{,}162\end{aligned}$$

Halbe Konfidenzintervallbreite:

$$\begin{aligned}\frac{KIB}{2} &= z_{(1-\alpha/2)} \cdot \sigma_{\bar{x}} \\ &\approx z_{97{,}5\%} \cdot 0{,}162 \\ &\approx 1{,}96 \cdot 0{,}162 \\ &\approx 0{,}32\end{aligned}$$

Der Mittelwert liegt mit Konfidenzniveau von 95% in den Grenzen $\bar{x} \pm 0{,}32\mathrm{km}$, also zwischen 8,58 und 9,22 km.

Aufgabe 4

Die Bürgermeisterin der Stadt Flussmünde vermutet: Es besteht ein signifikanter Unterschied zwischen den Nutzungsgruppen der beiden öffentlichen Bibliotheken, und zwar in Hinblick auf das durchschnittliche Alter der beiden Gruppen.

Zur Überprüfung dieser Hypothese kann sie auf die Anmeldeinformationen von je 51 zufällig ausgewählten, bei den Bibliotheken registrierten Personen zurückgreifen. Demnach liegt das Durchschnittsalter der in der Bibliothek Hafenviertel angemeldeten 51 Personen bei 41,2 Jahren. Die Standardabweichung beträgt 4,8 Jahre. Die in der Bibliothek Waldsiedlung angemeldeten 51 Personen sind im Durchschnitt 38,0 Jahre alt. Die Standardabweichung beträgt 4,1 Jahre.

Prüfen Sie auf Basis dieser Angaben die Hypothese der Bürgermeisterin mit einem Signifikanzniveau von 5 Prozent. *(15 Punkte)*

---

1. Test wählen und Voraussetzungen prüfen

   Es handelt sich um den Vergleich von Mittelwerten anhand von zwei Stichproben. Es muss ein 2-Stichproben-t-Test durchgeführt werden.

2. Hypothesen formulieren

$$\begin{aligned}H_0 &: \mu_1 = \mu_2 \\ H_1 &: \mu_1 \neq \mu_2\end{aligned}$$

3. Signifikanzniveau entscheiden

$$\alpha = 0{,}05$$

4. Ablehnungsbereich bestimmen

   Freiheitsgrade:

$$\begin{aligned} df &= 2 \cdot n - 2 \\ &= 2 \cdot 51 - 2 \\ &= 100 \end{aligned}$$

   Ablehnungsbereich:

$$\begin{aligned} t \leq t_{df;\alpha/2} &\quad \text{und} \quad t \geq t_{df;(1-\alpha/2)} \\ t \leq t_{100;2,5\%} &\quad \text{und} \quad t \geq t_{100;97,5\%} \\ t \leq -1{,}984 &\quad \text{und} \quad t \geq 1{,}984 \end{aligned}$$

5. Prüfgröße berechnen

$$\begin{aligned} t &= \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2 + s_2^2}{n}}} \\ &= \frac{41{,}2 - 38{,}0}{\sqrt{\frac{4{,}8^2 + 4{,}1^2}{51}}} \\ &\approx \frac{3{,}2}{\sqrt{0{,}781}} \\ &\approx 3{,}621 \end{aligned}$$

6. Ergebnis interpretieren

   Der kritische Wert von 1,984 wurde deutlich überschritten, die Nullhypothese kann somit abgelehnt werden. Die Hypothese der Bürgermeisterin konnte statistisch belegt werden ($\alpha = 0{,}05$).