Lösungen Probeklausur II

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)	Schulbildung (erfasst entlang des höchsten erworbenen Schulabschlusses)	O
b)	Pendeldistanz zum Arbeitsplatz (in km Luftlinie)	V
c)	Tätigkeitsfeld laut Handwerksordnung (z.B. „Zimmerer“, „Dachdecker“, „Straßenbauer“, … [sic])	N
d)	Jahr der Fertigstellung eines Gebäudes (z.B. 1871)	I
e)	Wohnort (als Postleitzahl)	N

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

	Aussage	**richtig**	**falsch**
f)	Mit einem t-Test kann geprüft werden, ob die Mittelwerte von Stichproben sich signifikant voneinander unterscheiden.	×	
g)	Für eine Verteilung mit Extremwerten und geringen Fallzahlen ist das arithmetische Mittel das bestgeeignete Lagemaß.		×
h)	Die Standard-Normalverteilung hat immer ein arithmetisches Mittel von 0 und eine Standardabweichung von 1.	×	
i)	Es ist unmöglich, dass bei Verteilungen Median und arithmetisches Mittel identisch ausfallen.		×
j)	Bei einem χ^2-Test werden die Differenzen zwischen beobachteten und erwarteten Werten berechnet.	×	

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)	Sie untersuchen, ob die Erwartungshaltung gegenüber der geplanten Umgestaltung der innerstädtischen Fußgängerzone (positive oder negative Erwartungshaltung) signifikant von der Konsumbereitschaft der Befragten (hohe oder niedrige Konsumbereitschaft) abhängig ist.	4
l)	Schwankt die Pendelzeit der Frankfurter Studierenden zwischen Universität und Wohnort signifikant stärker als die der Heidelberger Studierenden?	3
m)	Sie fragen sich, ob die durchschnittliche Passagieranzahl pro Tag am Flughafen Frankfurt sich signifikant von der am Flughafen München unterscheidet.	2
n)	Besteht ein signifikanter Zusammenhang zwischen einem hohen bzw. niedrigen Gewerbesteueraufkommen in hessischen Kommunen und dem Umstand, das einige Kommunen ein kostenloses Schulessen ausgeben, andere nicht?	4
o)	Unterscheidet sich die durchschnittliche Umzugshäufigkeit von Studierenden von der durchschnittlichen Umzugshäufigkeit der Gesamtbevölkerung?	1

Aufgabe 2

Im Rahmen der Imagekampagne „Frankfurt: faszinierend“ subventioniert die Stadtverwaltung von Frankfurt am Main Kulturereignisse in den Stadtteilen Ostend und Bahnhofsviertel.

Im Zuge einer Untersuchung über die Rolle und Bedeutung städtischer Image- und Kulturpolitik haben Sie einen Überblick über die Höhe der Pro-Kopf-Zuwendungen in Euro bekommen können, die in den letzten Jahren zur Image- und Kulturförderung von der Stadtverwaltung an die Stadtteile Ostend und Bahnhofsviertel geflossen sind.

Die folgenden Werte seien erhoben worden:

Jahr	Ostend	Bahnhofsviertel
2012	96	101
2013	109	103
2014	110	109
2015	102	99
2016	123	110
2017	114	108

a) Welcher Stadtteil hat in den letzten Jahren im Durchschnitt mehr Zuwendungen von der Stadtregierung empfangen? *(5 Punkte)*

Ostend:

$$\begin{aligned}\bar{x} &= \frac{\sum\limits_{i=1}^{n} x_i}{n} \\ &= \frac{654}{6} \\ &= 109\end{aligned}$$

Bahnhofsviertel:

$$\begin{aligned}\bar{x} &= \frac{\sum\limits_{i=1}^{n} x_i}{n} \\ &= \frac{630}{6} \\ &= 105\end{aligned}$$

Das Ostend hat im Schnitt mehr Zuwendungen bekommen.

---

b) In welchem Stadtteil variierten die jährlichen Zuwendungen stärker? Berechnen Sie die die Varianzen. *(5 Punkte)*

Ostend:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{440}{5} = 88$$

Bahnhofsviertel:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{106}{5} = 21{,}2$$

Die Zuwendungen im Ostend variieren stärker.

c) Für den Stadtteil Gallus betragen die jährlichen Zuwendungen im genannten Zeitraum im Durchschnitt 100 Euro. Die Varianz beträgt 36 Euro. Für welchen der drei Stadtteile — Ostend, Bahnhofsviertel, Gallus — lässt sich die größte Streuung in Relation zu den durchschnittlichen Zuwendungen festhalten? Berechnen Sie die jeweiligen Variationskoeffizienten. *(5 Punkte)*

Ostend:

$$v = \frac{s}{|\bar{x}|} \cdot 100\% = \frac{\sqrt{88}}{109} \cdot 100\% \approx 8{,}61\%$$

Bahnhofsviertel:

$$v = \frac{s}{|\bar{x}|} \cdot 100\% = \frac{\sqrt{21{,}2}}{105} \cdot 100\% \approx 4{,}39\%$$

Gallus:

$$
\begin{aligned}
v &= \frac{s}{|\bar{x}|} \cdot 100\% \\
&= \frac{\sqrt{36}}{100} \cdot 100\% \\
&= 6{,}00\%
\end{aligned}
$$

Für das Ostend ist der Variationskoeffizient am größten.

Aufgabe 3

In einer Zufallserhebung wurden jeweils 16 Cafés in den Stadtteilen Ostend und Bahnhofsviertel der Stadt Frankfurt am Main untersucht. Sie weisen die folgende durchschnittliche Anzahl an Gästen pro Tag mit folgenden Standardabweichungen auf:

- Bahnhofsviertel: $\bar{x}_1 = 63{,}2; \quad s_1 = 13{,}1$
- Ostend: $\bar{x}_2 = 58{,}2; \quad s_2 = 14{,}5$

Unterscheidet sich die durchschnittliche Anzahl an Gästen in Cafés im Bahnhofsviertel signifikant von der im Ostend? Gehen Sie bei Überprüfung der Frage von einem Signifikanzniveau von 5 Prozent aus. *(15 Punkte)*

---

1. Test auswählen und Voraussetzungen prüfen

   Ziel ist die Überprüfung von zwei Stichproben auf einen signifikanten Unterschied der Mittelwerte, deshalb 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:

$$
\begin{aligned}
df &= 2 \cdot n - 2 \\
&= 2 \cdot 16 - 2 \\
&= 30
\end{aligned}
$$

   Ablehnungsbereich:

$$\begin{aligned} t &\leq t_{df;\alpha/2} \quad \text{und} \quad t \geq t_{df;(1-\alpha/2)} \\ t &\leq t_{30;2{,}5\%} \quad \text{und} \quad t \geq t_{30;97{,}5\%} \\ t &\leq -2{,}042 \quad \text{und} \quad t \geq 2{,}042 \end{aligned}$$

5. Prüfgröße berechnen

$$\begin{aligned} t &= \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2+s_2^2}{n}}} \\ &= \frac{63{,}2 - 58{,}2}{\sqrt{\frac{13{,}1^2+14{,}5^2}{16}}} \\ &\approx \frac{5}{\sqrt{24{,}03}} \\ &\approx 1{,}01 \end{aligned}$$

6. Ergebnis interpretieren

Der Ablehnungsbereich wurde nicht erreicht, die Nullhypothese muss beibehalten werden. Es konnte kein signifikanter Unterschied der durchschnittlichen Gästeanzahl festgestellt werden ($\alpha = 0{,}05$).

Aufgabe 4

Die Zentralbibliothek der Universität Frankfurt ist ständig überfüllt. Die Bibliotheksleitung plant die Einführung einer Pausenbeschränkung, die dazu zwingt, einen schon eingenommenen Platz zu räumen, wenn eine längere Pause eingelegt wird.

Dafür soll zunächst einmal mit einer Umfrage unter 20 zufällig ausgewählten Besucher*innen der Bibliothek das Pausenverhalten einschätzbar gemacht werden. Die Auswertung der Umfrage macht deutlich: Eine Pause dauert im Durchschnitt 55 Minuten an.

Es sei außerdem bekannt: Die Pausendauer ist annähernd normal verteilt mit einer Standardabweichung von 17 Minuten.

a) Vorausgesetzt, die durchschnittliche Pausendauer beträgt tatsächlich 55 Minuten: Mit welcher Wahrscheinlichkeit ist dann mit kurzen Pausen zu rechnen, d.h. hier: beträgt die Pause eine halbe Stunde oder weniger? *(5 Punkte)*

---

z-Transformation:

$$\begin{aligned} z_p &= \frac{x_p - \mu}{\sigma} \\ &= \frac{30 - 55}{17} \\ &\approx -1{,}47 \end{aligned}$$

Unterschreitungswahrscheinlichkeit:

$$\begin{aligned} P(x < x_p) &= P(z < z_p) \\ &= P(z < -1{,}47) \\ &\approx 1 - P(z < 1{,}47) \\ &\approx 1 - 0{,}9292 = 0{,}0708 \end{aligned}$$

Es ist mit einer Wahrscheinlichkeit von ca. 7,08% mit Pausen von weniger als 30 Minuten zu rechnen.

---

b) Wie groß müsste eine Stichprobe sein, um die durchschnittliche Pausendauer mit einer Wahrscheinlichkeit von 99% auf $\pm$ 1 Minute schätzen zu können? *(10 Punkte)*

Nötiger Standardfehler:

$$\begin{aligned} \frac{KIB}{2} &= z_{(1-\alpha/2)} \cdot \sigma_{\bar{x}} \\ \sigma_{\bar{x}} &= \frac{KIB}{2 \cdot z_{(1-\alpha/2)}} \\ &= \frac{2}{2 \cdot z_{(99{,}5\%)}} \\ &\approx \frac{1}{2{,}58} \approx 0{,}39 \end{aligned}$$

Nötige Stichprobengröße:

$$\begin{aligned} \sigma_{\bar{x}} &= \frac{\sigma}{\sqrt{n}} \\ n &= \left(\frac{\sigma}{\sigma_{\bar{x}}}\right)^2 \\ &= \left(\frac{17}{0{,}39}\right)^2 = 1900{,}07 \end{aligned}$$

Es müssten 1901 Besucher*innen befragt werden, um die durchschnittliche Pausenlänge mit einem Konfidenzniveau von 99% auf $\pm$ 1 Minute genau schätzen zu können.
