HOCHSCHULE HEILBRONN

Statistikexamen

Zeit: *90 Minuten*

Name: ______________________

Matr. Nummer: ______________________

Hinweise:

1. Zugelassene Hilfsmittel: Open-Book: Aufschriebe, Formelsammlung, Skript, Taschenrechner (keine gespeicherten Formeln etc.!), Notizen.

2. Jede Antwort muss hinreichend begründet werden. Antworten ohne Begründung ergeben 0 Punkte.

3. Unleserliche Ergebnisse werden nicht gewertet. Nutzen Sie bei weiterem Platzbedarf bitte auch die Rückseiten der Klausurblätter!

4. Die geschätzte Bearbeitungszeit (in Minuten) für eine Aufgabe entspricht der Punktzahl. Somit sind die Aufgaben insgesamt 90 Punkte wert.

5. **Viel Glück!!!**

Aufgabe	**Punkte**	**Erreichte Punkte**
1	15	
2	20	
3	5	
4	10	
5	20	
6	20	
Gesamt	**90**	

Aufgabe 1: Deskriptive Statistik – Skalenniveau und Lagemaße (15 Punkte)

Die folgende Tabelle zeigt Daten von 5 Personen:

Daten					
Augenfarbe	Braun	Grün	Braun	Braun	Blau
Einkommens-klasse	2'000 - 5'000 €	>10'000 €	< 2'000 €	2'000 - 5'000 €	>10'000 €
Größe	170 cm	155 cm	155 cm	160 cm	180 cm

a) Geben Sie das Skalenniveau für jede Variable (Augenfarbe, Einkommensklasse, Größe) an.

b) Berechnen Sie für jede Variable diejenigen Lagemaße (Modalwert, Mittelwert, Median), welche gemäß Skalenniveau berechnet werden KÖNNEN.

Lösung

a) Augenfarbe: nominal, Einkommensklasse: ordinal, Größe: verhältnis

b)

- Augenfarbe: Nur Modalwert (Modalwert = Braun)
- Einkommensklasse: Modalwert und Median
  - Modalwerte = 2000 to 5000 and 10000
  - Median = 2000 to 5000
- Größe: Alle
  - Modalwert = 155
  - Median = 160
  - Mittelwert = 164

Aufgabe 2: Deskriptive Statistik – Lage- und Streuungsmaße (20 Punkte)

Die folgende Tabelle zeigt den Umsatz von 5 Firmen (in Millionen EUR):

Firma	1	2	3	4	5
Umsatz	7	6	9	7	3

Geben Sie bei Ihren Antworten Rechenschritte oder eine Erklärung an.

a) Berechnen Sie die Lagemaße (Modalwert, Median, Mittelwert)

b) Berechnen Sie die Streuungsmaße (Varianz, Standardabweichung)

c) Geben Sie die Einheit an in welcher die Varianz gemessen wird und die Einheit der Standardabweichung an. Interpretieren Sie die Standardabweichung. Was sagt diese aus?

d) Wie würde sich die Varianz ändern, wenn wir Umsatz in 100'000 EUR (statt Millionen EUR) messen würden?

Lösung

a)

$$Mittelwert = 6.4$$

$$Median \; = 7$$

$$Modalwert = 7$$

b)

$$Varianz = 3.84$$

$$Standardabweichung = 1.96$$

c) Varianz: (Millionen EUR)^2, Standardabweichung: Million EUR. Standardabweichung: Mittlerer Abstand der Daten vom Mittelwert.

d) Varianz würde auf 384 ansteigen.

Aufgabe 3: Deskriptive Statistik – Korrelation (5 Punkte)

Tragen Sie in die untere Tabelle Werte ein, sodass die Variablen X und Y perfekt negativ korreliert sind.

X	**Y**

Lösung (Beispiel):

X	**Y**
1	0
0	1

Aufgabe 4: Wahrscheinlichkeitstheorie - Normalverteilung (10 Punkte)

Die Nachfrage nach einem Produkt sei normalverteilt mit einem Erwartungswert von μ = 9300 Tausend € und einer Standardabweichung von σ = 2750 Tausend €.

a) Bestimmen Sie die Wahrscheinlichkeit, dass die Nachfrage zwischen 9300 und 12000 Tausend € liegt.

b) Bestimmen Sie die Wahrscheinlichkeit, dass die Nachfrage größer oder gleich 13000 Tausend € ist.

c) Bestimmen Sie die Wahrscheinlichkeit, dass die Nachfrage zwischen 12500 und 15000 Tausend € liegt.

d) Bestimmen Sie die Wahrscheinlichkeit, dass die Nachfrage zwischen 5000 und 11000 Tausend € liegt.

Lösung:

a)
- z = $\frac{12000-9300}{2750}$ = 0,98 → A= 0,3365
- P([9300,12000]) = 33,65%

b)
- z = $\frac{13000-9300}{2750}$ = 1,35 → A = 0,4115
- P([13000,∞))= 0,5 - 0,4115 = 8,85%

c)
- z1 = $\frac{12500-9300}{2750}$ = 1,16 → A1 = 0,3770
- z2 = $\frac{15000-9300}{2750}$ = 2,07 → A2 = 0,4808
- P([12500,15000])= 0,4808 – 0,377 = 10,34%

d)
- z1 = $\frac{11000-9300}{2750}$ = 0,62 → A1 = 0,2324
- z2 = $\frac{|5000-9300|}{2750}$ = 1,56 → A2 = 0,4406
- P([5000,11000])= 0,4406 + 0,2324 = 67,3%

Aufgabe 5: Inferenzstatistik – Schätzer (20 Punkte)

Es wird eine Zufallsstichprobe mit 81 Personen aus Deutschland erhoben. Die Personen geben im Schnitt 953 € (Stichprobenvarianz: 47'089 €^{2}) für ihren Sommerurlaub aus.

a) Bestimmen Sie den Punktschätzer für die durchschnittlichen Urlaubsausgaben in Deutschland.

b) Sind die Schätzer normalverteilt? Begründen Sie Ihre Antwort.

c) Bestimmen Sie den Punktschätzer für die Populationsstandardabweichung.

Tipp: Bestimmen Sie zuerst den Punktschätzer für die Populationsvarianz und dann daraus den Schätzer für die Populationsstandardabweichung.

d) Bestimmen Sie den 95%-Intervallschätzer (95%-Konfidenzintervall) für die durchschnittlichen Urlaubsausgaben in Deutschland.

Tipp: Verwenden Sie den Schätzer für die Populationsstandardabweichung aus c), um den Standardfehler zu bestimmen.

e) Würde das Konfidenzintervall größer oder kleiner werden, wenn wir die Stichprobengröße (n) erhöhen? Begründen Sie Ihre Antwort.

Lösung:

a) Punktschätzer = Stichprobenmittelwert = 953

b) Ja, da 81 > 30 (zentraler Grenzwertsatz)

c) $Schätzer\ Varianz = Stichprobenvarianz * \frac{n-1}{n} = 47089 * \frac{80}{81} \approx 46508$

$$\text{Schätzer Standardabweichung} = \sqrt{Schätzer\ Varianz} \approx 216$$

d)

$$95\% - KI = Punktschätzer \pm ErrorBound$$

- Punktschätzer = 953
- $ErrorBound = z * Standardfehler = z * \frac{\sigma}{\sqrt{n}}$
- $z = 1.96 (Tabelle), n = 81, \sigma = 216\ (siehe\ c)$:

$$\text{ErrorBound} = 47{,}04$$

Damit:

$$95\% - \text{KI} = 953 \pm 47 = [906, 1000]$$

e) Kleiner. Wenn n erhöht wird, dann sinkt der Standardfehler, dadurch sinkt der ErrorBound, dadurch wird das Intervall enger.

Aufgabe 6: Inferenzstatistik - Hypothesentest (20 Punkte)

Es soll die Hypothese untersucht werden, dass Männer im Schnitt mindestens 500 € mehr monatlich verdienen als Frauen. Hierzu wird eine Stichprobe erhoben. Die Stichprobe enthält 212 Frauen und 165 Männer. Das mittlere Monatsgehalt der Frauen beträgt 2201 € (Varianz: 605), das mittlere Monatsgehalt der Männer ist 2704 € (Varianz: 325).

a) Schreiben Sie die Null- und die Alternativhypothese in Bezug auf den Kontext auf.

b) Handelt es sich um einen Ein- oder Zweistichprobentest? Begründen Sie Ihre Antwort.

c) Handelt es sich um einen einseitigen oder zweiseitigen Test? Begründen Sie Ihre Antwort.

d) Berechnen Sie den gepoolten Standardfehler.

e) Berechnen Sie den p-Wert. Liegt Evidenz vor, dass Männer im Schnitt mindestens 500 € mehr verdienen? Bergünden Sie Ihre Antwort.

Lösung:

a)

- $H_0: \mu_M - \mu_F \leq 500$
- $H_1: \mu_M - \mu_F > 500$

b) Zweistichprobentest (zwei Gruppen)

c) Einseitig, "mehr verdienen" impliziert Richtung

d)

$$SE_{POOL} = \left(\frac{s_M^2}{n_M} + \frac{s_F^2}{n_F}\right)^{0.5} = \left(\frac{325}{165} + \frac{605}{212}\right)^{0.5} \approx 2.2$$

e)

- Teststatistik:

$$z = \frac{x_M - x_F - 500}{SE_{POOL}} \approx 1{,}36$$

- Fläche Teststatistik: A = 0,4131
- P-Wert = 0,5 – A = 8,69%

Der P-Wert ist deutlich größer als 5%. Daher können wir die Nullhypothese nicht ablehnen. Es liegt keine Evidenz für die Alternativhypothese vor.