Schema für die Lösung von Klausuraufgaben

Stand: 12/2024

Sachverhalt:

Welches ist das erhobene Merkmal? Skala?
oder gibt es mehrere Merkmale? Skalen? (eventuell: mehrdimensionale Verteilung, Kontingenztafel)

A) deskriptive (beschreibende) Statistik

a) für die Beschreibung von Merkmalen kommen folgende Verfahren in Betracht:

	Nominalskala	Ordinalskala	Kardinalskala
Mittelwerte	Modus	Modus Median	Modus (wenn diskret) Median arithm. Mittel geometrisches Mittel (in der Verhältnisskala)
Streuungsmaße			Spannweite Quartilsabstand Standardabweichung
Häufigkeiten	h p, f	h, H f, F p,P	h, H f, F p,P (i.d.R. werden stetige Merkmale in Klassen dargestellt)
Diagrammtyp	Säulendiagramm	Säulendiagramm	Histogramm

Summenhäufigkeit F
- bei kardinalskalierten, stetigen Merkmalen kann eine lineare Interpolation von Quantilen durchgeführt werden
- eine grafische Darstellung von F kann in Klausuraufgaben verlangt werden, einschl. grafischer Bestimmung von Quantilen

b) Untersuchung des Zusammenhangs zwischen <u>zwei Merkmalen</u>

1) wenn beide Merkmale metrisch skaliert sind

- lineare Regression durch Minimierung der Residuen und Minimierung der Summe der kleinsten Quadrate, Regressionsgerade $\hat{y} = a + b*x$
- Korrelationskoeffizient r nach Bravais-Pearson, Bestimmtheitsmaß $B = r^2$
  (jeweils Interpretation der Werte erforderlich)

2) wenn ein Merkmal oder beide Merkmale nicht metrisch ist / sind

- quadratische Kontingenz χ^2
- Kontingenzkoeffizient Cramers V zur Interpretation von χ^2

B) induktive (schließende) Statistik

In der induktiven Statistik geht es darum,
- unbekannte Parameter der Grundgesamtheit (GG) aufgrund einer Stichprobe zu schätzen,
- Hypothesen zu testen

Typische Formulierungen sind: „auf dem Signifikanzniveau / Konfidenzniveau 1- α“, „bei α=0,05“, „Testen Sie, ob…“, „es wird behauptet, dass in der Grundgesamtheit…“, „Schätzen Sie…“

Die folgenden Methoden a und b erfordern Kenntnisse der (Standard-)Normalverteilung, Tabellenwerte (Quantile) von ϕ müssen abgelesen / interpretiert werden.
Die folgende Methode c erfordert Kenntnis der Berechnung von χ^2 (siehe deskriptive Statistik).

a) metrische Skala
- Schätzen einer unbekannten Standardabweichung $\hat{\sigma}$ der GG als Punktschätzung
- Schätzen eines unbekannten Mittelwertes $\hat{\mu}$ der GG als Punktschätzung: $\hat{\mu} = \bar{x}$
- Schätzen eines unbekannten Mittelwertes $\hat{\mu}$ der GG als Intervallschätzung
  auf dem Signifikanzniveau 1- α,
  mit dem Standardfehler der Schätzung des unbekannten Mittelwertes $\sigma_{\bar{x}} = \frac{\hat{\sigma}}{\sqrt{n}}$,
  der unteren Grenze UG = $\bar{x} - z^*\sigma_{\bar{x}}$ und der oberen Grenze OG = $\bar{x} + z^*\sigma_{\bar{x}}$
  sowie dem z-Wert aus der Tabelle mit den symmetrischen Quantilen, so dass $\phi(z) = 1-\alpha$

  → $W(\bar{x} - z^*\sigma_{\bar{x}} < \hat{\mu} < \bar{x} + z^*\sigma_{\bar{x}}) = 1-\alpha$ (bei linksseitigen Quantilen: z-Wert zu $1 - \frac{\alpha}{2}$)

- Berechnen des notwendigen Stichprobenumfangs n für eine geforderte Genauigkeit der Schätzung

- Hypothesentest: z-Wert der Behauptung berechnen, wenn |z| > Quantil → ablehnen

b) dichotomes Merkmal, Binomialverteilt B(n;p) $\mu = n^*p$ $\sigma = \sqrt{n \cdot p \cdot (1-p)}$
Sei p der Anteilswert aus der Stichprobe
- Schätzen eines unbekannten Anteilswertes $\hat{p}$ der GG als Punktschätzung $\hat{p} = p$
- Schätzen eines unbekannten Anteilswertes $\hat{p}$ als Intervallschätzung
  Annäherung durch Normalverteilung:
  $W(p - z^*\sigma_p < \hat{p} < p + z^*\sigma_p) = 1-\alpha$ wobei σ_p der Standardfehler der Schätzung des Anteilswertes p der Grundgesamtheit ist $\sigma_p = \sqrt{\frac{p*(1-p)}{n-1}}$

- Berechnen des notwendigen Stichprobenumfangs n für eine geforderte Genauigkeit der Schätzung

- Hypothesentest: z-Wert der Behauptung berechnen, wenn |z| > Quantil → ablehnen

c) Zusammenhang zwischen zwei Merkmalen

Hypothese H: Die Merkmale sind voneinander unabhängig.

Entscheidungsregel: Werte für χ^2 ab einem bestimmten Wert (Quantil) führen dazu, dass die Hypothese abgelehnt wird

wenn $\chi^2 > \chi^2_{(1-\alpha,\text{Freiheitsgrad})}$ dann H verwerfen, anderenfalls beibehalten (nicht verwerfen)

$\chi^2_{(1-\alpha,\text{Freiheitsgrad})}$: Tabellenwert (Quantil) der χ^2-Verteilung (Konfidenzniveau und Freiheitsgrad beachten)