n_{ij}	$(n_{ij} - E_{ij})^2$
$n_{ij} - E_{ij}$	E_{ij}

1 B. Statistkik

- Qualitative Merkmale:
  - Variieren nach Beschaffenheit
  - Bspw. Geschlecht
- Quantitative Merkmale:
  - Variieren nach Wert/Zahlen
  - Bspw. Alter, Einkommen
- Diskrete Merkmale:
  - abgestufte Werte
  - Bspw. Einkommensklasse
- Stetige Merkmale:
  - können im Intervall jeden reellen Wert annehmen
  - Bspw. Körpergröße

Skalenniveaus

- Nominal
  - nur Gleichheit oder Andersartigkeit feststellbar (keine Bewertung)
  - stets qualitativ (Religion, Beruf etc.)
- Ordinal
  - natürliche oder festzulegende Rangfolge
  - IQ, Schulnoten
- Kardinal
  - numerischer Art
  - Ausprägung und Unterschied sind messbar
  - verhältnisskaliert (Absoluter Nullpunkt vorhanden; Gewicht, Preis (Doppelt so viel.))
  - intervallskaliert (Kein Nullpunkt, nur Differenzen; Temperatur (10 Grad wärmer als gestern))

Werte

- Arithmetisches Mittel $\overline{x}$
  - $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} a_i = \frac{a_1+a_2+\cdots+a_n}{n}$
  - Summe aller Abweichungen vom Mittel $= 0$
  - Verschiebung um kostanten Wert a $a+\overline{x}$
  - Multiplikation mit konstantem Wert $a\cdot\overline{x}$
- Median $\widetilde{x}$
  - Mittleres Element der geordneten Liste
  - Bei gerader Anzahl, Durchschnitt der mittleren Elemente
- Quartile (FEHLT)
  - Unteres Quartil $\widetilde{x}_{0,25}$
  - Oberes Quartil $\widetilde{x}_{0,75}$
- Varianz σ^2
  - Populations Varianz $\sigma^2 = \frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}$
  - Sample Varianz $S^2_{n-1} = \frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{n-1}$
  - Altn. Formel $\sigma^2 = \overline{x^2} - \overline{x}^2$
  - Eigenschaften:
    - Immer >= 0
    - Addition mit a, Varianz unverändert
    - Multiplikation mit b, $Varianz * b^2$
- Standardabweichung σ
  - $\sigma = \sqrt{\sigma^2}$
  - StichprobenSTD $S = \sqrt{S^2_{n-1}}$
- Quartilsabstand (FEHLT)

Zweidimensionale Häuffigkeitstabellen

- Statistische Variablen X und Y mit versch.Auspräungen
- Spaltensummen sowie Zeilensummen = n
- Relative Häufigkeit $h_{ij} = \frac{n_{ij}}{n}$
- Randverteilung = Betrachtung einer einzigen Variable
- $Z = X + Y$; $\overline{z} = \overline{x} + \overline{y}$;

Kovarianz

- Arithmetisches Mittel des Produkts der Abweichung der einzelnen Beobachtungen von ihrem Mittel
- $C_{XY} := \frac{1}{n}\sum_{j=1}^{n}(x_j - \overline{x})(y_j - \overline{y})$
- $C_{XY} = \overline{xy} - \overline{x} * \overline{y}$
- $C_{XY} > 0$ "große X-Werte zu großen Y-Werten"
- $C_{XY} < 0$ "große Werte zu kleine Werten"
- Sind zwei Variablen statistisch unabhängig ist die Kovarianz = 0

Korrelation

- Normal (Pearson) $r_{XY} = \frac{C_{XY}}{\sigma_x * \sigma_y}$
  - normiertes Maß für Strenge des linearen statistischen Zusammenhangs
  - r_{XY} hat das gleiche Vorzeichen wie C_{XY}
  - Bleibt unverändert bei linearer Transformation
  - $r_{XY} = r_{YX}$
- Rangkorrelation (Spearman) $r^{Sp}_{XY} = r_{rg(X),rg(Y)}$
  - für ordinale Variablen
  - misst monotonen Anteil des stat. Zusammenhangs
  - Ränge müssen vorher berechnet werden
- Kovarianz und Korrelation bedeuten nicht zwangsweise eine kausale Beziehung!

Kontingenzkoeffizient

- beschreibt die Stärke des Zusammenhangs zweier Merkmale, nicht deren Richtung
- Chi-Quadrat $QK = \sum_{i=1}^{k}\sum_{j=1}^{l}\frac{(n_{ji}-E_{ij})^2}{E_{ij}}$
  - $E_{ij} = \frac{1}{n} * n_i * n_j = \frac{1}{n} n(x_i) * n(y_j)$
  - Siehe Erweiterte Kontingeztabelle
  - X und Y unabhängig: $QK = 0$
  - Sonst $QK > 0$
  - Für 2x2 Matrix: $QK = \frac{n(ad-bc)^2}{(a+b)(a+c)(b+d)(c+d)}$
  - a bis d sind Inhalte der Tabelle, Summen sind Randhäufigkeiten
- Kontingenzkoeffizient $K := \sqrt{\frac{QK}{QK+n}}$
  - normiertes Maß
  - X und Y unabhängig: $K = 0$
  - $0 <= K <= K_{max} = \sqrt{\frac{m-1}{m}} < 1$
  - m = Minimum von Zeilenzahl und Spaltenzahl
- Korrigierter K.-koeffizient $K^* := \frac{K}{K_{max}} = \sqrt{\frac{QK*m}{(QK+n)(m-1)}}$
- – $0 <= K^* <= 1$
  - Vergleichbar mit anderen K-Tabellen

2 Regression

- Lineare Regression $y(x) = a + bx$
- $b = \frac{c_{XY}}{s^2_X}$ und $a = \overline{y} - b\overline{x}$

3 S. Statistik

Hallo

4 Taschenrechner

Hallo