

Deskriptive Statistik

Seminaraufgaben

- (1) Bei der Untersuchung der Auslastung einer Telefonzelle wird an fünf aufeinanderfolgenden Tagen jeweils für den Zeitraum einer Stunde zwischen 8 Uhr und 16 Uhr die Anzahl der Telefonkunden registriert:

1, 1, 2, 0, 6, 4, 3, 1, 0, 0, 3, 4, 3, 2, 2, 1, 1, 2, 3, 3,
0, 0, 0, 4, 3, 2, 1, 2, 6, 3, 3, 3, 0, 4, 1, 2, 2, 6, 2, 2

- (a) Bestimmen Sie die absoluten und relativen Häufigkeiten sowie die relativen Summenhäufigkeiten.
(b) Geben Sie die empirische Verteilungsfunktion an.
(c) Stellen Sie die absolute Häufigkeit und die empirische Verteilungsfunktion grafisch dar.

Bestimmen sie außerdem folgende Kennwerte:

- (d) arithmetisches Mittel
(e) empirischer Median
(f) empirischer Modalwert
(g) empirische Streuung
(h) empirischer Variationskoeffizient
(i) Das weitere Betreiben der Telefonzelle soll nur dann erfolgen, wenn zwischen 8 Uhr und 16 Uhr vier oder mehr Kunden mit einer „Wahrscheinlichkeit“ von wenigstens 0.6 kommen. Ist mit einer Schließung zu rechnen?

- (2) Bei einer Versuchsserie zur Prüfung von Bremsen wurden die Momentangeschwindigkeit X [km/h] und der Bremsweg Y [m] gemessen:

X	49.2	51	52.4	48.2	51.6	48.5	49.8	51.3	48.9	49.5
Y	30.8	33.9	35.3	29.9	34.6	30.6	31.4	33.8	31.2	32.1
X	50.9	51.4	51.1	48.6	49.4	52.8	52.1	50.7	50.3	50.4
Y	32.8	34.1	33.3	30.4	31.4	35.7	34.6	33.1	32.3	32.9

Bestimmen Sie

- (a) den empirischen Korrelationskoeffizienten.
(b) die Regressionsgerade in Abhängigkeit von der Geschwindigkeit.
(c) die durch die Regressionsgerade nicht erklärte Streuung (Reststreuung).
(d) den Anteil an der Gesamtstreuung des Bremsweges, der auf die Geschwindigkeit zurückzuführen ist.
(e) Welcher Bremsweg ist bei einer Geschwindigkeit von 55 km/h zu erwarten?

- (3) Ein Getränkekonzern will die Produktion an Mineralwasser der Nachfrage anpassen. Es hat sich durch umfangreiche statistische Untersuchungen gezeigt, dass es gewisse Abhängigkeiten zwischen der Tageshöchsttemperatur und der Nachfrage nach dem Mineralwasser der Firma gibt. Die Daten führten zu folgenden Größen: Y – Nachfrage in Mio. Liter, X – Tageshöchsttemperatur in $^{\circ}C$
 $\bar{y} = 1.8$, $s_y = 0.01$, $\bar{x} = 15.1$, $s_x = 3.1$, $s_{xy} = 0.0248$
- Berechnen Sie den empirischen Korrelationskoeffizienten.
 - Geben Sie die Gleichung der empirischen Regressionsgerade an.
 - Wie groß wird die Nachfrage danach bei einer Tageshöchsttemperatur von $25^{\circ}C$ sein?
 - Wie viel Prozent der Gesamtvarianz sind nicht durch die Regressionsgerade bestimmt?

Aufgaben zur Nachbereitung

- (1) Eine Textilfirma analysiert vor Beginn eines neuen Produktionsvorhabens die Größenverteilung eines gewissen Kundenkreises. Dabei wurden unter anderem auch die Körpergrößen (in cm) einer Gruppe von 30 Studenten ermittelt
- 180, 176, 167, 180, 177, 166, 160, 176, 176, 164,
 182, 165, 175, 172, 172, 173, 179, 166, 162, 168
 170, 177, 183, 179, 172, 166, 163, 176, 177, 181
- Bestimmen Sie für die Klasseneinteilung $[160, 165)$, $[165, 170)$, $[170, 175)$, $[175, 180)$, $[180, 185)$ die absoluten und relativen Klassenhäufigkeiten sowie die relativen Summenhäufigkeiten.
 - Geben Sie die empirische Verteilungsfunktion für die Klasseneinteilung an.
 - Stellen Sie die absolute Klassenhäufigkeit und die empirische Verteilungsfunktion grafisch dar.
- Bestimmen sie außerdem folgende Kennwerte bezüglich der Klasseneinteilung:
- arithmetisches Mittel
 - empirischer Median
 - empirischer Modalwert
 - empirische Streuung
 - empirische Standardabweichung
 - empirischer Variationskoeffizient
- (2) Für 6 verschiedene Monate im Jahr 1998 liegen Daten über den Hypothekenzinssatz X [%] sowie über den saisonbereinigten Auftragseingang Y [Mio Euro]

im Bauhauptgewerbe, welcher auf den privaten Wohnungsbau entfällt vor:

X	6	5	7	7	8	9
Y	3000	3200	2500	2300	2000	2000

Bestimmen Sie:

- (a) den empirischen Korrelationskoeffizienten.
- (b) die empirische Regressionsgerade
- (c) das Bestimmtheitsmaß
- (d) die Prognosewerte für den Auftragseingang, der bei einem Hypothekenzins von 4% und 7.5%.
- (e) die geschätzten Werte mittels der empirischen Regressionsgerade für die gegebenen Werte für X .

Lösungen der Nachbereitung

- (1a) absoluten Häufigkeiten 4, 6, 5, 10, 5
 relativen Häufigkeiten 0.1333, 0.2000, 0.1667, 0.3333, 0.1667
 relativen Summenhäufigkeiten 0.1333, 0.3333, 0.5, 0.8333, 1

$$(1b) \quad \hat{F}_X(x) = \begin{cases} 0, & x < 160 \\ 0.1333, & 160 \leq x < 165 \\ 0.3333, & 165 \leq x < 170 \\ 0.5, & 170 \leq x < 175 \\ 0.8333, & 175 \leq x < 180 \\ 1, & x \geq 180 \end{cases}$$

- (1c) siehe Konsultation

(1d) 173, 5

(1e) 172, 5

(1f) 177, 5

(1g) 43, 79

(1h) 6, 6176

(1i) 0, 0381

(2a) -0.9503

(2b) $-340x + 4880$

(2c) 0.9031

(2d) 3520 bei 4% 2330 bei 7.5%

(2e) [2840, 3180, 2500, 2500, 2160, 1820]