Stochastik und Statistik

Jil Zerndt, Lucien Perret
December 2024

Begriffe

Grundlegende Begriffe

- $\Omega =$ Grundgesamtheit
- $n =$ Anzahl Objekte
- $X =$ Stichprobenwerte
- $a =$ Ausprägungen
- $h =$ Absolute Häufigkeit
- $f =$ Relative Häufigkeit
- $H =$ Kumulative Absolute Häufigkeit
- $F =$ Kumulative Relative Häufigkeit

Bivariate Daten (Merkmale)

- 2x kategoriell $\rightarrow$ Kontingenztabelle + Mosaikplot
- 1x kategoriell + 1x metrisch $\rightarrow$ Boxplot oder Striptchart
- 2x metrisch $\rightarrow$ Streudiagramm

Beschreibende Statistik

Absolute Häufigkeiten

$$H = \sum_{1}^{n} h_i$$

Relative Häufigkeiten

$$F = \sum_{1}^{m} f_i, \quad F(x) = \frac{H(x)}{n}$$

Kennwerte (Lagemasse)

Quantil

$$i = \lceil n \cdot q \rceil, Q = x_i = x_{\lceil n \cdot q \rceil}$$

Interquartilsabstand

$$IQR = Q_3 - Q_1$$

Modus

$x_{\text{mod}} =$ Häufigste Wert

Arithmetisches Mittel und Median

Arithmetisches Mittel	Median
$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i = \sum_{i=1}^{m} a_i \cdot f_i$	$\begin{cases} x_{\left[\frac{n+1}{2}\right]} & n \text{ ungerade} \\ 0.5 \cdot \left(x_{\left[\frac{n}{2}\right]} + x_{\left[\frac{n}{2}+1\right]} \right) & n \text{ gerade} \end{cases}$

Stichprobenvarianz s^2 (Streumasse)

$$s^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 = \overline{x^2} - \bar{x}^2, \quad (s_{kor})^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$(s_{kor})^2 = \frac{n}{n-1} \cdot s^2$$

Standardabweichung s (Streumasse)

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2} = \sqrt{\overline{x^2} - \bar{x}^2}, \quad s_{kor} = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2}$$

PDF + CDF

Nicht klassierte Daten (PMF und CDF)

Die absolute Häufigkeit kann als Funktion $h : \mathbb{R} \rightarrow \mathbb{R}$ bezeichnet werden.

$$h_i$$

Die relative Häufigkeit kann als Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ bezeichnet werden.

$$f_i = \frac{h_i}{n}$$

Kombinatorik

Fakultät

$$n! = 1 \cdot 2 \cdot \ldots \cdot n = \prod_{k=1}^{n} k$$

Binomialkoeffizient

Wie viele Möglichkeiten gibt es k Objekte aus einer Gesamtheit von n Objekten auszuwählen.

$$\binom{n}{k} = \frac{n!}{(n-k)! \cdot k!}$$

Systematik

Grundbegriffe

- k Anzahl Stellen
- n Anzahl Optionen pro Stelle

Variation (mit Reihenfolge)		Kombination (ohne Reihenfolge)	
Mit Wiederholung	Ohne Wiederholung	Mit Wiederholung	Ohne Wiederholung
n^k	$\frac{n!}{(n-k)!}$	$\binom{n+k-1}{k}$	$\binom{n}{k}$
Zahlenschloss	Schwimmwettkampf	Zahnarzt	Lotto

Wahrscheinlichkeitsrechnung

Spezialfälle der Kombinatorik

Romme Beispiel Beim Rommé spielt man mit 110 Karten: sechs davon sind Joker. Zu Beginn eines Spiels erhält jeder Spieler genau 12 Karten.
Wahrscheinlichkeit für genau zwei Joker:

$$\frac{\binom{6}{2} \cdot \binom{104}{10}}{\binom{110}{12}}$$

Wahrscheinlichkeit für mindestens einen Joker:

$$1 - \frac{\binom{104}{12}}{\binom{110}{12}}$$

Glühbirnen Beispiel Von 100 Glühbirnen sind genau drei defekt. Es werden nun 6 Glühbirnen zufällig ausgewählt.
Anzahl Möglichkeiten mit mindestens einer defekten Glühbirne:

$$\binom{100}{6} - \binom{97}{6} = 203'880'032$$

Wahrscheinlichkeit für keine defekte Glühbirne:

$$\frac{\binom{97}{6}}{\binom{100}{6}}$$

Wahrscheinlichkeitstheorie

Ergebnisraum

Ergebnisraum Ω ist die Menge aller möglichen Ergebnisse des Zufallsexperiments. Zähldichte $\rho : \Omega \rightarrow [0, 1]$ ordnet jedem Ereignis seine Wahrscheinlichkeit zu.
Für ein Laplace-Raum (Ω, P) gilt:

$$P(M) = \frac{|M|}{|\Omega|}$$

Stochastische Unabhängigkeit

Zwei Ereignisse A und B heissen stochastisch unabhängig, falls:

$$P(A \cap B) = P(A) \cdot P(B)$$

Zwei Zufallsvariablen $X : \Omega \rightarrow \mathbb{R}$ und $Y : \Omega \rightarrow \mathbb{R}$ heissen stochastisch unabhängig, falls:

$$P(X = x, Y = y) = P(X = x) \cdot P(Y = y), \quad \text{für alle } x, y \in \mathbb{R}$$

Bedingte Wahrscheinlichkeit

Bedingte Wahrscheinlichkeit

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)}$$

Multiplikationssatz

$$P(A \cap B) = P(A) \cdot P(B \mid A) = P(B) \cdot P(A \mid B)$$

Satz von der Totalen Wahrscheinlichkeit

$$P(B) = P(A) \cdot P(B \mid A) + P(\bar{A}) \cdot P(B \mid \bar{A})$$

Satz von Bayes

$$P(A \mid B) = \frac{P(A) \cdot P(B \mid A)}{P(B)}$$

Spezielle Verteilungen

Verteilungen und Erwartungswerte Für diskrete Verteilungen:

$$E(X) = \sum_{x \in \mathbb{R}} f(x) \cdot x$$

$$V(X) = \sum_{x \in \mathbb{R}} f(x) \cdot (x - E(X))^2$$

Für stetige Verteilungen:

$$E(X) = \int_{-\infty}^{\infty} f(x) \cdot x dx$$

$$V(X) = \int_{-\infty}^{\infty} f(x) \cdot (x - E(X))^2 dx$$

Bernoulliverteilung Bernoulli-Experimente sind Zufallsexperimente mit nur zwei möglichen Ergebnissen (1 und 0):

$$P(X = 1) = p, \quad P(X = 0) = 1 - p = q$$

Es gilt:
1. $E(X) = E(X^2) = p$
2. $V(X) = p \cdot (1 - p)$

Normalverteilung

Gauss-Verteilung Die stetige Zufallsvariable X folgt der Normalverteilung mit den Parametern $\mu, \sigma \in \mathbb{R}, \sigma > 0$:

$$\varphi_{\mu,\sigma}(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$$

Standardnormalverteilung ($\mu = 0$ und $\sigma = 1$):

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}x^2}$$

Approximation durch die Normalverteilung
- Binomialverteilung: $\mu = np, \sigma^2 = npq$
- Poissonverteilung: $\mu = \lambda, \sigma^2 = \lambda$

$$P(a \leq X \leq b) = \sum_{x=a}^{b} P(X = x) \approx \phi_{\mu,\sigma}(b + \frac{1}{2}) - \phi_{\mu,\sigma}(a - \frac{1}{2})$$

Zentraler Grenzwertsatz Für eine Folge von Zufallsvariablen $X_1, X_2, \ldots, X_n$ mit gleichem Erwartungswert μ und gleicher Varianz σ^2 gilt:

$$E(S_n) = n \cdot \mu, \quad V(S_n) = n \cdot \sigma^2, \quad E(\bar{X}_n) = \mu, \quad V(\bar{X}_n) = \frac{\sigma^2}{n}$$

Die standardisierte Zufallsvariable:

$$U_n = \frac{((X_1 + X_2 + \cdots + X_n) - n\mu)}{\sqrt{n} \cdot \sigma} = \frac{(\bar{X} - \mu)}{\frac{\sigma}{\sqrt{n}}}$$

konvergiert in Verteilung gegen die Standardnormalverteilung.

Faustregeln für Approximationen
- Die Approximation (Binomialverteilung) kann verwendet werden, wenn $npq > 9$
- Für grosses $n (n \geq 50)$ und kleines $p (p \leq 0.1)$ kann die Binomial- durch die Poisson-Verteilung approximiert werden:

$$B(n, p) \approx \text{Poi}(n \cdot p)$$

- Eine Hypergeometrische Verteilung kann durch eine Binomialverteilung angenähert werden, wenn $n \leq \frac{N}{20}$:

$$H(N, M, n) \approx B(n, \frac{M}{N})$$

Methode der kleinsten Quadrate

Lineare Regression Gegeben sind Datenpunkte $(x_i; y_i)$ mit $1 \leq i \leq n$. Die Residuen / Fehler $\epsilon_i = g(x_i) - y_i$ dieser Datenpunkte sind Abstände in y-Richtung zwischen y_i und der Geraden g. Die Ausgleichs- oder Regressiongerade ist diejenige Gerade, für die die Summe der quadrierten Residuen $\sum_{i=1}^{n} \epsilon_i^2$ am kleinsten ist.

Regressionsgerade Die Regressiongerade $g(x) = mx + d$ mit den Parametern m und d ist die Gerade, für welche die Residualvarianz s_ϵ^2 minimal ist.

Steigung: $m = \frac{s_{xy}}{s_x^2}$, y-Achsenabschnitt: $d = \bar{y} - m\bar{x}$, $s_\epsilon^2 = s_y^2 - \frac{s_{xy}^2}{s_x^2}$

Bestimmtheitsmass

Varianzaufspaltung Die Totale Varianz setzt sich zusammen aus der Residualvarianz und der Varianz der prognostizierten Werte:
- s_y^2 Totale Varianz
- $s_{\hat{y}}^2$ prognostizierte (erklärte) Varianz
- s_ϵ^2 Residualvarianz

$$s_y^2 = s_\epsilon^2 + s_{\hat{y}}^2$$

Bestimmtheitsmass Das Bestimmtheitsmass R^2 beurteilt die globale Anpassungsgüte einer Regression über den Anteil der prognostizierten Varianz $s_{\hat{y}}^2$ an der totalen Varianz s_y^2:

$$R^2 = \frac{s_{\hat{y}}^2}{s_y^2}$$

Das Bestimmtheitsmass R^2 entspricht dem Quadrat des Korrelationskoeffizienten:

$$R^2 = \frac{s_{xy}^2}{s_x^2 \cdot s_y^2} = (r_{xy})^2$$

Linearisierungsfunktionen

Transformationen

Ausgangsfunktion	Transformation
$y = q \cdot x^m$	$\log(y) = \log(q) + m \cdot \log(x)$
$y = q \cdot m^x$	$\log(y) = \log(q) + \log(m) \cdot x$
$y = q \cdot e^{m \cdot x}$	$\ln(y) = \ln(q) + m \cdot x$
$y = \frac{1}{q+m \cdot x}$	$V = q + m \cdot x; V = \frac{1}{y}$
$y = q + m \cdot \ln(x)$	$y = q + m \cdot U; u = \ln(x)$
$y = \frac{1}{q \cdot m^x}$	$\log(\frac{1}{y}) = \log(q) + \log(m) \cdot x$

Schliessende Statistik

Erwartungstreue Schätzfunktion Eine Schätzfunktion Θ eines Parameters θ heisst erwartungstreu, wenn:

$$E(\Theta) = \theta$$

Effizienz einer Schätzfunktion Gegeben sind zwei erwartungstreue Schätzfunktionen Θ_1 und Θ_2 desselben Parameters θ. Man nennt Θ_1 effizienter als Θ_2, falls:

$$V(\Theta_1) < V(\Theta_2)$$

Konsistenz einer Schätzfunktion Eine Schätzfunktion Θ heisst konsistent, wenn:

$$E(\Theta) \to \theta \text{ und } V(\Theta) \to 0 \text{ für } n \to \infty$$

Vertrauensintervalle

Vertrauensintervall Wir legen eine grosse Wahrscheinlichkeit γ fest (z.B. $\gamma = 95\%$). γ heisst statistische Sicherheit oder Vertrauensniveau. $\alpha = 1 - \gamma$ ist die Irrtumswahrscheinlichkeit.
Dann bestimmen wir zwei Zufallsvariablen Θ_u und Θ_o so, dass sie den wahren Parameterwert Θ mit der Wahrscheinlichkeit γ einschliessen:

$$P(\Theta_u \leq \Theta \leq \Theta_o) = \gamma$$

Intervallschätzung Verteilungstypen und zugehörige Quantile:

Verteilung	Parameter	Quantile
Normalverteilung (σ^2 bekannt)	μ	$c = u_p, p = \frac{1+\gamma}{2}$
t-Verteilung (σ^2 unbekannt)	μ	$c = t_{(p;f=n-1)}, p = \frac{1+\gamma}{2}$
Chi-Quadrat-Verteilung	σ^2	$c_1 = \chi^2_{(\frac{1-\gamma}{2};n-1)}, c_2 = \chi^2_{(\frac{1+\gamma}{2};n-1)}$

Berechnung eines Vertrauensintervalls Geben Sie das Vertrauensintervall für μ an (σ^2 unbekannt). Gegeben sind:

$$n = 10, \quad \bar{x} = 102, \quad s^2 = 16, \quad \gamma = 0.99$$

1. Verteilungstyp mit Param μ und σ^2 unbekannt $\rightarrow$ T-Verteilung
2. $f = n - 1 = 9$, $p = \frac{1+\gamma}{2} = 0.995$, $c = t_{(p;f)} = t_{(0.995;9)} = 3.25$
3. $e = c \cdot \frac{S}{\sqrt{n}} = 4.111$, $\Theta_u = \bar{X} - e = 97.89$, $\Theta_o = \bar{X} + e = 106.11$

Likelyhood-Funktion

Likelyhood-Funktion Wir betrachten eine Zufallsvariable X und ihre Dichte (PDF) $f_x(x|\theta)$, welche von x und einem oder mehreren Parametern θ abhängig sind.
Für eine Stichprobe vom Umfang n mit $x_1, \ldots, x_n$ nennen wir die vom Parameter θ abhängige Funktion die Likelyhood-Funktion der Stichprobe:

$$L(\theta) = f_x(x_1|\theta) \cdot f_x(x_2|\theta) \cdot \ldots \cdot f_x(x_n|\theta)$$

Vorgehen bei Maximum-Likelihood-Schätzung

1. Likelyhood-Funktion bestimmen
2. Maximalstelle der Funktion bestimmen:
   - (Partielle) Ableitung $L'(\theta) = 0$

Beispiele

Erwartungstreue Schätzfunktion Grundgesamtheit mit Erwartungswert μ, Varianz σ^2 und Zufallsstichprobe X_1, X_2, X_3. Die folgende Schätzfunktion ist gegeben:

$$\Theta_1 = \frac{1}{3} \cdot (2X_1 + X_2)$$

Ist diese Schätzfunktion erwartungstreu (Parameter: μ)?

$$E(\Theta_1) = E(\frac{1}{3} \cdot (2X_1 + X_2)) = \frac{1}{3} \cdot (2E(X_1) + E(X_2))$$

$$E(\Theta_1) = \frac{1}{3} \cdot (2\mu + \mu) = \frac{3\mu}{3} = \mu$$

Da $E(\Theta_1) = \mu$ ist die Funktion erwartungstreu.

Intervallschätzung für die Varianz Für die Varianz σ^2 einer Normalverteilung mit Stichprobenumfang $n = 10$ und Stichprobenvarianz $s^2 = 16$ soll ein 99%-Vertrauensintervall berechnet werden.

1. Verteilungstyp: Chi-Quadrat-Verteilung
2. Freiheitsgrade: $f = n - 1 = 9$
3. Quantile: $c_1 = \chi^2_{(0.005;9)} = 1.735$, $c_2 = \chi^2_{(0.995;9)} = 23.589$
4. Vertrauensintervall:

$$\frac{(n-1)s^2}{c_2} \leq \sigma^2 \leq \frac{(n-1)s^2}{c_1}$$

$$\frac{9 \cdot 16}{23.589} \leq \sigma^2 \leq \frac{9 \cdot 16}{1.735}$$

$$6.10 \leq \sigma^2 \leq 82.99$$

Bernoulli-Anteilsschätzung Ein Vertrauensintervall für den Parameter p einer Bernoulli-Verteilung soll aus einer Stichprobe mit $n = 100$ und $\bar{x} = 0.42$ bei einem Vertrauensniveau von 95% berechnet werden.

1. Prüfen der Voraussetzung: $n\hat{p}(1 - \hat{p}) = 100 \cdot 0.42 \cdot 0.58 = 24.36 > 9$
2. Quantil: $c = u_{0.975} = 1.96$
3. Standardfehler: $\sqrt{\frac{\bar{x}(1-\bar{x})}{n}} = \sqrt{\frac{0.42 \cdot 0.58}{100}} = 0.0494$
4. Vertrauensintervall:

$$0.42 \pm 1.96 \cdot 0.0494 = [0.323; 0.517]$$