Formelsammlung

Satz von Bayes

Seien $A_1, A_2, \dots, A_n$ und B Ereignisse auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ mit $A_1 \uplus A_2 \uplus \dots \uplus A_n = \Omega$ und $\mathbb{P}(A_1), \mathbb{P}(A_2), \dots, \mathbb{P}(A_n), \mathbb{P}(B) > 0$. Dann gilt für $i = 1, 2, \dots, n$:

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)}{\sum_{j=1}^{n} \mathbb{P}(B|A_j) \cdot \mathbb{P}(A_j)}.$$

Erwartungswert

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $X\colon \Omega \to \{a_1, a_2, \dots, a_n\}$ eine diskrete Zufallsvariable. Dann gilt:

$$\mathbb{E}[X] = a_1 \cdot \mathbb{P}(X = a_1) + a_2 \cdot \mathbb{P}(X = a_2) + \dots + a_n \cdot \mathbb{P}(X = a_n),$$

$$\mathrm{Var}(X) = (a_1 - \mathbb{E}[X])^2 \cdot \mathbb{P}(X = a_1) + (a_2 - \mathbb{E}[X])^2 \cdot \mathbb{P}(X = a_2) + \dots + (a_n - \mathbb{E}[X])^2 \cdot \mathbb{P}(X = a_n).$$

Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und X eine stetige Zufallsvariable mit Dichte f_X. Dann gilt:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} t \cdot f_X(t) \mathrm{d}t,$$

$$\mathbb{E}[X^2] = \int_{-\infty}^{\infty} t^2 \cdot f_X(t) \mathrm{d}t.$$

Allgemein gilt:

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

Transformationen

erzeugende Funktion	$G_X(z) = \mathbb{E}[z^X]$.
momenterzeugende Funktion	$M_X(t) = \mathbb{E}[\mathrm{e}^{tX}]$.
charakteristische Funktion	$\chi_X(t) = \mathbb{E}[\mathrm{e}^{\mathrm{i}tX}]$. (Die Umkehrformell: X stetig $\Rightarrow f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \chi_X(t) \mathrm{e}^{-\mathrm{i}tX} \mathrm{d}t$.)

Ungleichungen

Tschebyscheff-Ungleichung: $\mathbb{P}\{|X - \mathbb{E}[X]| \geq \epsilon\} \leq \mathrm{Var}[X]/\epsilon^2$.

Verteilungen

diskrete Verteilungen:

Bernoulli-Verteilung mit Parameter p:	$\mathbb{P}(X = 1) = p$, $\mathbb{P}(X = 0) = 1 - p$,	$G_X(z) = 1 - p + zp$.
Binomialverteilung mit Parametern n, p:	$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, \dots, n$,	$G_X(z) = (1 - p + zp)^n$.
Geometrisch mit Parameter p:	$\mathbb{P}(X = k) = p(1-p)^k$, $k \in \mathbb{Z}_+$,	$G_X(z) = \frac{p}{1-(1-p)z}$.
Poisson mit Parameter λ:	$\mathbb{P}(X = k) = \lambda^k \exp(-\lambda)/k!$, $k \in \mathbb{Z}_+$,	$G_X(z) = \mathrm{e}^{\lambda(z-1)}$.

stetige Verteilungen:

Normalverteilung mit Parametern μ,σ^2:	$f_X(x) = \frac{1}{\sqrt{2\sigma^2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$,	$M_X(t) = \exp(\mu t + \sigma^2 t^2/2)$.
Gleichverteilung auf $[a, b]$:	$f_X(x) = \begin{cases} \frac{1}{b-a}, & x \in [a,b] \\ 0, & \text{sonst.} \end{cases}$,	$M_X(t) = \frac{\mathrm{e}^{bt} - \mathrm{e}^{at}}{t(b-a)}$.
Exponentialverteilung mit Parameter λ:	$f_X(x) = \begin{cases} \lambda \cdot e^{-\lambda x}, & x \geq 0 \\ 0, & \text{sonst.} \end{cases}$,	$M_X(t) = \frac{\lambda}{\lambda - t}$, $\quad t < \lambda$.

Lineare Regression (Methode der kleinsten Quadraten)

Das Optimierungsproblem:

$$\underset{\theta_0, \theta_1 \in \mathbb{R}}{\operatorname{argmin}} \sum_{i=1}^{n} (y_i - \theta_0 - \theta_1 x_i)^2 = ?$$

Lösung vom Optimierungsproblem:

$$\hat{\theta}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$$\hat{\theta}_0 = \bar{y} - \bar{\theta}_1 \bar{x}.$$