

Mittel		ungew.	gewichtet
arithmeth.	\bar{x}	$\frac{1}{n} \sum_{i=1}^n x_i$	$\frac{\sum_{i=1}^n w_i \cdot x_i}{\sum_{i=1}^n w_i}$
geometr.	$\mathbf{G}(\mathbf{x})$	$\sqrt[n]{\prod_{i=1}^n x_i}$	$\sqrt[n]{\prod_{i=1}^n x_i^{w_i}}$
harmon.	$\mathbf{H}(\mathbf{x})$	$\frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$	$\frac{\sum_{i=1}^n \frac{w_i}{x_i}}{\sum_{i=1}^n \frac{w_i}{x_i}}$

Es gilt: $H(x) \leq G(x) \leq \bar{x}$

p%-Quantile: $u_{p\%} = \begin{cases} \frac{1}{2}(x_{[np]} + x_{[np+1]}) & , np \in \mathbb{Z} \\ x_{[[np]]} & , \text{sonst} \end{cases}$

Median: $\tilde{x} := \begin{cases} x_{[\frac{n+1}{2}]} & , n \bmod 2 = 1 \\ \frac{1}{2}\{x_{[\frac{n}{2}]} + x_{[\frac{n}{2}+1]}\} & , n \bmod 2 = 0 \end{cases}$

Standardabweich.: $\sigma_x^n = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$

effizienter mit: $\sigma_x^n = \sqrt{\frac{1}{n} \left(\sum_{i=1}^n x_i^2 - \frac{1}{n} \left(\sum_{i=1}^n x_i \right)^2 \right)}$

empir. Streuung: $s_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$

Varianz: σ_x^{n2} **empir. Varianz:** s_x^2

Datenpaare (x_i, y_i)

empir. Kovarianz: $s_{xy} = \frac{1}{n-1} \sum_{i=1}^n ((x_i - \bar{x})(y_i - \bar{y}))$

wobei $s_{xx} = s_x^2$



Boxplot

Wahrscheinlichkeitsmodelle

a) Ergebnismenge Ω

Menge aller möglichen Ausgänge des Experiments

b) Ereignis-System \mathcal{A} mit Ereignissen $A_i \subset \Omega$

Menge aller Ereignisse A_i heißt Ereignis-System \mathcal{A}

c) Wahrscheinlichkeit P

(Mengen)- σ -Algebra

$\Omega \in \mathcal{A}; \quad A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$

$A_1, A_2, \dots \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$

Wahrscheinlichkeiten

$0 \geq P(A) \leq 1; \quad P(\Omega) = 1, \quad P(\emptyset) = 0$

$P(A_1 + A_2 + \dots) = P(A_1) + P(A_2) + \dots; \quad P(A) = \frac{|A|}{|\Omega|}$

stoch. Unabhängigkeit: $P(AB) = P(A)P(B)$

rel. Häufigkeit: $H_N = \frac{A_N}{N}$

Bedingte Wahrscheinlichkeit: $P(A|B) = \frac{P(A \cap B)}{P(B)}$

totale W.: $P(A) = \sum_{i \in I} P(AB_i) = \sum_{i \in I} P(B_i)P(A|B_i)$

Verkett.regel: $P(ABC) = P(A) \cdot P(B|A) \cdot P(C|AB)$

Formel von Bayes: $P(B_k|A) = \frac{P(B_k)P(A|B_k)}{\sum_{i \in I} P(B_i)P(A|B_i)}$

Laplace-Experiment: $P(1) = \dots = P(N) = \frac{1}{N}$

Varianz

$\text{Var}(x) = E(X - EX)^2 = EX^2 - (EX)^2$

$\text{Var}(a + X) = \text{Var}(X); \quad \text{Var}(aX) = a^2 \text{Var}(X)$

$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2 \cdot \text{Kov}(X, Y)$

Kovarianz $\text{Kov}(X, Y) = EXY - EXEY$

X, Y st.u. $\rightarrow \text{Kov}(X, Y) = 0$

Korrelationskoeffizient $\text{korr}(X, Y) = \frac{\text{Kov}(X, Y)}{\text{Str}X \text{Str}Y}$

Riemann-Dichte: $P((a, b]) = \int_a^b f(x) dx$

Erwartungswert

$EX = \sum_{k \in \Omega'} k \cdot P(X = k); \quad EX = \int_{-\infty}^{\infty} x \cdot f^X(x) dx$

monoton: $a \leq X \leq Y \Rightarrow a \leq EX \leq EY$

linear: $E(aX + b) = a \cdot EX + b$

$E(X + Y) = EX + EY$

X, Y st.u. $\Leftrightarrow E(XY) = EX \cdot EY$

$Eg(X) = \int_{-\infty}^{\infty} g(x) f^X(x) dx$

$Eh(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) f^{(X, Y)}(x, y) dx dy$

$E(X - a)^2 = \text{Var}(X) + (EX - a)^2$

Integration im \mathbb{R}^n

Substitutionsregel $\int \int_G f(x, y) dx dy$

$\frac{\partial(x, y)}{\partial(u, v)} = \begin{pmatrix} \frac{\partial x(u, v)}{\partial u} & \frac{\partial x(u, v)}{\partial v} \\ \frac{\partial y(u, v)}{\partial u} & \frac{\partial y(u, v)}{\partial v} \end{pmatrix}$

$\int \int_G f(x, y) d(x, y) =$

$\int \int_H f(x(u, v), y(u, v)) \left| \det \left(\frac{\partial(x, y)}{\partial(u, v)} \right) \right| d(u, v)$

Mehrdim. Normalverteilung $\mathcal{N}(\vec{a}, \mathbf{K})$

Lin. Transformation $Y = g(X) = \vec{a} + \mathbf{A}X$

$EY_i = a_i$

$EX_i^2 = 1, \quad EX_k X_l = 0 \rightarrow \text{Kov}(Y_i, Y_j) = \mathbf{A} \mathbf{A}^T$

Dichte

$f^Y(\vec{y}) = \frac{1}{\sqrt{|\det \mathbf{K}|}} \left(\frac{1}{\sqrt{2\pi}} \right)^n e^{-\frac{1}{2}(\vec{y} - \vec{a})^T (\mathbf{K}^{-1})(\vec{y} - \vec{a})}$

$Z = \vec{b} + \mathbf{B}Y$ wobei $Y \sim \mathcal{N}(\vec{a}, \mathbf{K})$

$\rightarrow Z \sim \mathcal{N}(\vec{b} + \mathbf{B}\vec{a}, \mathbf{B}\mathbf{K}\mathbf{B}^T)$

zufällige Summe $S = \sum_{i=1}^Y X_i$

$ES = EY \cdot EX_1$

$\text{Var}S = EY \cdot \text{Var}X_1 + \text{Var}Y \cdot (\text{Var}EX_1)^2$

Gesetze der großen Zahl

stochastisch $Y_n \xrightarrow{st} Y$

Es gilt: $\lim_{n \rightarrow \infty} P(|Y_n - Y| \geq \epsilon) = 0 \quad \forall \epsilon > 0$

im r-ten Mittel $Y_n \xrightarrow{(r)} Y$

Es gilt: $E|Y_n - Y|^r \rightarrow 0$

nach Verteilung $Y_n \xrightarrow{V} Y$

Es gilt: $F^{Y_n}(x) \rightarrow F^Y(x)$

Ungl. Chebychew-Markow: $P(|Y| \geq \epsilon) \leq \frac{1}{\epsilon^r} E|Y|^r$

für $r = 2, EY$ exist.: $P(|Y - EY| \geq \epsilon) \leq \frac{1}{\epsilon^2} \text{Var}Y$

Zentraler Grenzwertsatz

$\frac{S_n - ES_n}{\text{Str}S_n} = \frac{\sum_{i=1}^n X_i - n \cdot EX_1}{\sqrt{n \text{Str}X_1}} \xrightarrow{n \rightarrow \infty} \mathcal{N}(0, 1)$

i-te Randdichte

$f^{X_i}(\omega_i) = \sum_{\omega_1 \in \Omega_1} \dots \sum_{\omega_{i-1} \in \Omega_{i-1}} \sum_{\omega_{i+1} \in \Omega_{i+1}} \dots \sum_{\omega_n \in \Omega_n} f(\omega_1, \dots, \omega_n)$

$f^{X_i}(x_i) = \int_{\mathbb{R}} \dots \int_{\mathbb{R}} f(x_1, \dots, x_n) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n$

Gamma-Funktion

$\Gamma(\nu + 1) = \nu \cdot \Gamma(\nu), \quad \Gamma(1) = 1, \quad \Gamma(0.5) = \sqrt{\pi}$

$\Gamma(\nu) = (\nu - 1)! \quad \nu \in \mathbb{N}$

Schätzverfahren

rel. Wirksamkeit $\eta = \frac{\text{Var}(\hat{\Theta}_1)}{\text{Var}(\hat{\Theta}_2)}$

Likelihood-Fkt $L(x_1, \dots, x_n; \Theta) = \prod_{i=1}^n P(X = x_i; \Theta)$

Zum Schätzen: $\frac{dL}{d\Theta} \stackrel{!}{=} 0$ oder $\frac{d \ln L}{d\Theta} \stackrel{!}{=} 0$

Erwartungstreue: $E(\hat{\Theta}) = \Theta$

asympt. Erw.treue: $\lim_{n \rightarrow \infty} E(\hat{\Theta}(X_1, \dots, X_n)) = \Theta$

Konsistent: $\lim_{n \rightarrow \infty} P(|\hat{\Theta}(X_1, \dots, X_n) - \Theta| < \epsilon) = 1$

Konsistenzkriterium: $\text{Var}(\hat{\Theta}(X_1, \dots, X_n)) \xrightarrow{n \rightarrow \infty} 0 \Rightarrow$

$P(|\hat{\Theta}(X_1, \dots, X_n) - \Theta| > \epsilon) \leq \frac{\text{Var}(\hat{\Theta})}{\epsilon^2} \xrightarrow{n \rightarrow \infty} 0$

Konfidenzschätzung \mathcal{N}

für μ, σ bekannt: $(\bar{x} - z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}; \bar{x} + z_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}})$

für μ, σ unbek.: $(\bar{x}_n - \frac{s_n}{\sqrt{n}} t_{1-\frac{\alpha}{2}, n-1}; \bar{x}_n + \frac{s_n}{\sqrt{n}} t_{1-\frac{\alpha}{2}, n-1})$

für σ^2 : $(\frac{(n-1) \cdot s^2}{\chi_{1-\frac{\alpha}{2}; n-1}^2}; \frac{(n-1) \cdot s^2}{\chi_{\frac{\alpha}{2}; n-1}^2})$

Konfidenzschätzung $B(p)$ $z := z_{1-\frac{\alpha}{2}}$

$p_{max/min} = \frac{n}{n-z} \left(\bar{X} + \frac{z^2}{2n} \pm z \sqrt{\frac{\bar{X}(1-\bar{X})}{n} + \left(\frac{z}{2n}\right)^2} \right)$

Hypothesentests

	nicht abgelehnt	abgelehnt
H_0 richtig	\rightarrow ok $p_1 = 1 - \alpha$	Fehler 1.Art $p_2 = \alpha$
H_0 falsch	Fehler 2.Art $p_3 = \beta$	\rightarrow ok $p_1 = 1 - \beta$

Signifikanztest: Verzicht auf Fehler 2.Art

Test für $\mathcal{N}(\mu, \sigma^2)$, $H_0 = \{\mathcal{N}(\mu, \sigma^2), \mu = \mu_0\}$, σ bek.

$u = \frac{\bar{X} - \mu_0}{\sigma} \sqrt{n} \sim \mathcal{N}(0, 1)$

zweiseit.: $P(|U| \geq z_{1-\frac{\alpha}{2}} | H_0) = \alpha; |u| \geq z_{1-\frac{\alpha}{2}} \rightarrow \neg H_0$

einseit.: $P(U \geq z_{1-\alpha} | H_0) = \alpha; u \geq z_{1-\alpha} \rightarrow \neg H_0$

oder $P(U \leq z_{1-\alpha} | H_0) = \alpha; u \leq z_{1-\alpha} \rightarrow \neg H_0$

Test für $\mathcal{N}(\mu, \sigma^2)$, $H_0 = \{\mathcal{N}(\mu, \sigma^2), \mu = \mu_0\}$, σ unbek.

$u = \frac{\bar{X} - \mu_0}{s} \sqrt{n} \sim t_m$ mit $m = n - 1$

zweis.: $P(|U| \geq t_{1-\frac{\alpha}{2}, m} | H_0) = \alpha; |u| \geq t_{1-\frac{\alpha}{2}, m} \rightarrow \neg H_0$

eins.: $P(U \geq t_{1-\alpha, m} | H_0) = \alpha; u \geq t_{1-\alpha, m} \rightarrow \neg H_0$

oder $P(U \leq t_{1-\alpha, m} | H_0) = \alpha; u \leq t_{1-\alpha, m} \rightarrow \neg H_0$

Gleichheitstest

$u = \frac{\bar{X} - \bar{Y}}{\sqrt{(n_1-1)s_x^2 + (n_2-1)s_y^2}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} \sim t_{m=n_1+n_2-2}$

zweis.: $P(|U| \geq t_{1-\frac{\alpha}{2}, m} | H_0) = \alpha; |u| \geq t_{1-\frac{\alpha}{2}, m} \rightarrow \neg H_0$

χ^2 - Anpassungstest

$u = \sum_{m=1}^k \frac{(h_m - np_m)^2}{np_m} \sim \chi_{k-r-1}^2$

k Klassen, r unbek. Variablen

$P(u \geq \chi_{1-\alpha; k-r-1}^2 | H_0) = \alpha$

$\rightarrow u \geq \chi_{1-\alpha; k-r-1}^2 \rightarrow H_0$ ablehnen

Markow-Ketten

Matrix: Zeilensumme = 1

irreduzibel: Jeder Zustand von jedem Zustand aus mit positiver Wahrsch. erreichbar.

$\exists n : \forall P_{ij}^n > 0 \rightarrow$ irreduzibel + aperiodisch

Gleichgewichtsverteilung

$\pi \cdot P = \pi; \pi_i \geq 0; \sum_{j \in I} \pi_j = 1$

Berechnen von π : $(P^T - 1 \cdot I) \cdot \pi \stackrel{!}{=} \vec{0}$

Summen unabh. ZVs & Faltung

$f^{X+Y}(z) = (f^X * f^Y)(z) = \sum_{x=0}^z f^X(x) f^Y(z-x)$

$f^{X+Y}(z) = (f^X * f^Y)(z) = \int_0^z f^X(x) f^Y(z-x) dx$

Binom.vert. $B(n+m, p) = B(n, p) * B(m, p)$

Poissonvert. $\pi(\lambda_1) * \pi(\lambda_2) = \pi(\lambda_1 + \lambda_2)$

$Nb^+(1, p) = Geo^+(p)$

$Nb^+(r_1 + r_2, p) = Nb^+(r_1, p) * Nb^+(r_2, p)$

$\mathcal{N}(a, \sigma^2) * \mathcal{N}(b, \tau^2) = \mathcal{N}(a+b, \sigma^2 + \tau^2)$

Zusammenhänge & Approximationen

$\mathcal{N}(a, \sigma^2) \rightarrow \mathcal{N}(0, 1) : \Phi_{(a, \sigma)}(X) \rightarrow \Phi\left(\frac{X-a}{\sigma}\right)$

$B(n, p) \xrightarrow{n \rightarrow \infty} \mathcal{N}(np, np(1-p))$

$B(n, p) \xrightarrow{p \rightarrow 0} \pi(n \cdot p)$

$Nb^+(1, p) = Geo^+(p)$

Regressionsanalyse \mathcal{N}

Regressionsgerade

$r(x) = ax + b$ mit $a = \frac{s_{xy}}{s_x^2}, b = \bar{y} - a\bar{x}$

emp. Restvarianz: $s_R^2 = \frac{1}{n-2} \sum_{i=1}^n (y_i - y(x_i))^2$

Prüfen $\beta_2: H_0 : \beta_2 = \beta_{20}$

$u = \frac{b_2 - \beta_{20}}{s_{\beta_2}} \sim t_m$ mit $s_{\beta_2}^2 = \frac{s_R^2}{(n-1)s_x^2}$

$|u| \geq t_{1-\frac{\alpha}{2}; n-2} \rightarrow H_0$ ablehnen

Prüfen $\beta_1: H_0 : \beta_1 = \beta_{10}$

$u = \frac{b_1 - \beta_{10}}{s_{\beta_1}} \sim t_m$ mit $s_{\beta_1}^2 = s_R^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{(n-1)s_x^2} \right)$

$|u| \geq t_{1-\frac{\alpha}{2}; n-2} \rightarrow H_0$ ablehnen

Konfidenzintervall β_2, β_1 analog

$P\left(\left| \frac{b_2 - \beta_2}{s_{\beta_2}} \right| < t_{1-\frac{\alpha}{2}; n-2}\right) = 1 - \alpha$

$\Rightarrow b_2 - t_{1-\frac{\alpha}{2}; n-2} s_{\beta_2} < \beta_2 < b_2 + t_{1-\frac{\alpha}{2}; n-2} s_{\beta_2}$

Konfidenzintervall η

$y(x) - t_{1-\frac{\alpha}{2}; n-2} s_\eta < \eta(x) < y(x) + t_{1-\frac{\alpha}{2}; n-2} s_\eta$

mit $s_\eta(x) = \sqrt{s_R^2 \left(\frac{1}{n} + \frac{(x-\bar{x})^2}{(n-1)s_x^2} \right)}$

Korrelationsanalyse \mathcal{N}

Ziel: Entscheiden ob X, Y st. u.

$H_0 : \rho_{XY} = \text{korr}(X, Y) = 0$

$u = \frac{r_{XY} \sqrt{n-2}}{\sqrt{1-r_{XY}^2}} \sim t_{n-2}$

$|u| \geq t_{1-\frac{\alpha}{2}; n-2} \rightarrow H_0$ ablehnen

oder: $W = \frac{1}{2} \ln \left(\frac{1+r_{XY}}{1-r_{XY}} \right)$

$\mu_W = \frac{1}{2} \ln \left(\frac{1+r}{1-r} \right) + \frac{r}{2(n-1)}$

$\sigma_W^2 = \frac{1}{n-3}$

$u = \frac{w - \mu_w}{\sigma_w} \sim \mathcal{N}(0, 1)$

$|u| \geq z_{\alpha-\frac{\alpha}{2}} \rightarrow H_0$ ablehnen

Name	Dichte	Verteilungsfunktion	E	Var
diskret				
Bernoulli	$f^{X_1}(1) = p$		p	$p(1-p)$
Binomial	$b(n, p; k) = \binom{n}{k} p^k (1-p)^{n-k}$		np	$np(1-p)$
Einpunkt	$f_a^X(a) = 1$		a	0
Geo.	$f^{W_1}(k) = p(1-p)^{k-1}$		$\frac{1}{p}$	$\frac{1-p}{p^2}$
	$f^{W_1'}(k) = p(1-p)^{k-1}$		$\frac{1-p}{p}$	$\frac{1-p}{p^2}$
	$f^{Z_n}(k) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}$		$\frac{n}{N}$	$\frac{n}{N} \frac{N-K}{N} \frac{N-n}{N-1}$
Hypergeo.	$f^X(k) = \frac{1}{N}$		$\frac{N+1}{2}$	$\frac{N^2-1}{12}$
Laplace	$f^{W_r}(k) = \binom{k-1}{r-1} p^r (1-p)^k$		$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$
Neg. Binomial	$f^{W_r'}(k) = \binom{k+r-1}{r-1} p^r (1-p)^k$		$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$
Poisson	$\pi(\lambda; k) = e^{-\lambda} \frac{\lambda^k}{k!}$	$F_\lambda(X) = e^{-\lambda} \sum_{k=0}^n \frac{\lambda^k}{k!}$	λ	λ
stetig				
Beta	$be_{\mu, \nu}(x) = \frac{\Gamma(\mu+\nu)}{\Gamma(\mu)\Gamma(\nu)} x^{\mu-1} (1-x)^{\nu-1} 1_{(0,1)}(x)$		$\frac{\mu}{\mu+\nu}$	$\frac{\mu\nu}{(\mu+\nu+1)(\mu+\nu)^2}$
Cauchy	$f^X(x) = \frac{\pi}{\pi(\alpha^2+x^2)}$		Ex. nicht	
Chi ²	$f_n(x) = \frac{x^{\frac{n}{2}-1} e^{-\frac{x}{2}}}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \cdot 1_{(0,\infty)}(x)$		n	$2n$
Exponential	$f^X(x) = \alpha e^{-\alpha x} 1_{(0,\infty)}(x)$	$F(x) = (1 - e^{-\alpha x}) \cdot 1_{(0,\infty)}(x)$	$\frac{1}{\alpha}$	$\frac{1}{\alpha^2}$
Fischer-F	$f^{Z_n}(t) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi} \Gamma(\frac{n}{2})} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}$			
Gamma	$\gamma(x) = \frac{\alpha^\nu}{\Gamma(\nu)} x^{\nu-1} e^{-\alpha x} 1_{(0,\infty)}(x)$		$\frac{\nu}{\alpha}$	$\frac{\nu}{\alpha^2}$
Normal	$\phi_{\alpha, \sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\alpha)^2}{2\sigma^2}}$	$\Phi(x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\alpha)^2}{2\sigma^2}} dt$	a	σ^2
Std.-Normal	$\phi_{0,1}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$	$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$	0	1
Std.-Normal ^r	$f^X(x_1, \dots, x_n) = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{x_1^2 + \dots + x_n^2}{2}}$			
Rechteck	$f^X(x) = \frac{1}{b-a} 1_{(a,b)}(x)$		$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Student-t	$f^{Z_n}(t) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi} \Gamma(\frac{n}{2})} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}$		0	$\frac{n}{n-2}$