

Formelsammlung

Eindimensionale Zufallsvariablen

Nicht-zentrierte Momente von X (diskret)	$E(X^k) = \sum_{i=1}^{\infty} x_i^k \cdot P(X = x_i)$
Nicht-zentrierte Momente von X (stetig)	$E(X^k) = \int_{-\infty}^{\infty} x^k \cdot f_X(x) dx$
Varianz von X (Verschiebesatz)	$Var(X) = E(X^2) - E^2(X)$

Mehrdimensionale Zufallsvariablen

Randdichte von X	$f_X(x) = \int_{-\infty}^{\infty} f_{(X,Y)}(x,y) dy$
Erwartungswert des Produkts von X und Y (diskret)	$E(XY) = \sum_{i=1}^n \sum_{j=1}^m x_i y_j f_{(X,Y)}(x_i, y_j)$
Kovarianz von X und Y	$Cov(X, Y) = E(XY) - E(X)E(Y)$
Korrelationskoeffizient von X und Y	$\rho_{X,Y} = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$

Punktschätzung

	ML-Schätzung	Korrigierte Schätzung
$E(X)$	$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$	—
$Var(X)$	$S_X^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$	$\check{S}_X^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{n}{n-1} \cdot S_X^2$

Intervallschätzung

Erwartungswert	GG normalverteilt	GG beliebig verteilt ($n \geq 30$)
σ_X^2 bekannt	$\left[\bar{X} - z_{1-\alpha/2} \frac{\sigma_X}{\sqrt{n}}, \bar{X} + z_{1-\alpha/2} \frac{\sigma_X}{\sqrt{n}} \right]$	$\left[\bar{X} - z_{1-\alpha/2} \frac{\sigma_X}{\sqrt{n}}, \bar{X} + z_{1-\alpha/2} \frac{\sigma_X}{\sqrt{n}} \right]$
σ_X^2 unbekannt	$\left[\bar{X} - t_{1-\alpha/2}(n-1) \frac{\check{S}_X}{\sqrt{n}}, \bar{X} + t_{1-\alpha/2}(n-1) \frac{\check{S}_X}{\sqrt{n}} \right]$	$\left[\bar{X} - z_{1-\alpha/2} \frac{\check{S}_X}{\sqrt{n}}, \bar{X} + z_{1-\alpha/2} \frac{\check{S}_X}{\sqrt{n}} \right]$
Anteilswert	$\left[P - z_{1-\alpha/2} \sqrt{\frac{P(1-P)}{n}}, P + z_{1-\alpha/2} \sqrt{\frac{P(1-P)}{n}} \right]$	
Varianz	$\left[\frac{(n-1)\check{S}^2}{\chi_{1-\alpha/2}^2(n-1)}, \frac{(n-1)\check{S}^2}{\chi_{\alpha/2}^2(n-1)} \right]$	

Hypothesentests

Test	Voraussetzungen	Prüfgröße	Prüfverteilung
Approx. Binomialtest	$n\pi_0 \geq 5 ; n(1 - \pi_0) \geq 5$	$G = \frac{\bar{X} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$	$\mathcal{N}(0, 1)$
χ^2-Anpassungstest	$n\theta_k \geq 1$ für alle $k = 1, \dots, K$ $n\theta_k \geq 5$ für mind. 80% der k	$G = \sum_{k=1}^K \frac{(X_k - n\theta_k)^2}{n\theta_k}$	$\chi^2(K - 1 - L)$
χ^2-Unabhängigkeitstest		$G = \sum_{k=1}^K \sum_{l=1}^L \frac{(N_{k,l} - \tilde{N}_{k,l})^2}{\tilde{N}_{k,l}}$	$\chi^2((K - 1)(L - 1))$
Prüfung einer Varianz		$G = \frac{1}{\sigma_0^2} \sum_{i=1}^n (X_i - \bar{X})^2$	$\chi^2(n - 1)$
Vergleich zweier Varianzen		$G = \frac{\check{S}_X^2}{\check{S}_Y^2}$	$F(n - 1, m - 1)$
Vergl. von Anteilswerten	$n, m \geq 30$	$G = \frac{\bar{X} - \bar{Y} - \delta_0}{\sqrt{\frac{\bar{X}(1-\bar{X})}{n} + \frac{\bar{Y}(1-\bar{Y})}{m}}}$	approx. $\mathcal{N}(0, 1)$
Vergl. von Anteilswerten	$n, m \geq 30$	$G = \frac{\bar{X} - \bar{Y}}{\sqrt{\hat{P}(1-\hat{P})\left(\frac{1}{n} + \frac{1}{m}\right)}}$ $\hat{P} = \frac{n\bar{X} + m\bar{Y}}{n+m}$	approx. $\mathcal{N}(0, 1)$
$\delta_0 = 0$			

Tests zu Lagealternativen

Voraussetzungen	Prüfgröße	Prüfverteilung
Normalverteilung Varianz σ_X^2 bekannt	$G = \frac{\bar{X} - \mu_0}{\frac{\sigma_X}{\sqrt{n}}}$	Standardnormalverteilung
Normalverteilung Varianz σ_X^2 unbekannt	$G = \frac{\bar{X} - \mu_0}{\frac{\check{S}_X}{\sqrt{n}}}$	t-Verteilung mit $(n - 1)$ -Freiheitsgraden
keine Verteilungsvoraussetzung Varianz σ_X^2 bekannt; $n \geq 30$	$G = \frac{\bar{X} - \mu_0}{\frac{\sigma_X}{\sqrt{n}}}$	Standardnormalverteilung
keine Verteilungsvoraussetzung approximativ Varianz σ_X^2 unbekannt; $n \geq 30$	$G = \frac{\bar{X} - \mu_0}{\frac{\check{S}_X}{\sqrt{n}}}$	Standardnormalverteilung

Tests zum Vergleich von Erwartungswerten

Voraussetzungen	Prüfgröße	Prüfverteilung
Normalverteilungen; Varianzen σ_X^2 und σ_Y^2 bekannt	$G = \frac{\bar{X} - \bar{Y} - \delta_0}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$	Standardnormalverteilung
keine Verteilungsvoraussetzung; Varianzen σ_X^2 und σ_Y^2 bekannt; $n, m \geq 30$; approximativ	$G = \frac{\bar{X} - \bar{Y} - \delta_0}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}}$	Standardnormalverteilung
keine Verteilungsvoraussetzung; Varianzen σ_X^2 und σ_Y^2 unbekannt; $n, m \geq 30$; approximativ	$G = \frac{\bar{X} - \bar{Y} - \delta_0}{\sqrt{\frac{\check{S}_X^2}{n} + \frac{\check{S}_Y^2}{m}}}$	Standardnormalverteilung
Normalverteilungen; Varianzen σ_X^2 und σ_Y^2 unbekannt, aber gleich	$G = \frac{\bar{X} - \bar{Y} - \delta_0}{\check{S} \sqrt{\frac{1}{n} + \frac{1}{m}}}$ $\check{S}^2 = \frac{(n-1)\check{S}_X^2 + (m-1)\check{S}_Y^2}{n+m-2}$	t-Verteilung mit $(n + m - 2)$ Freiheitsgraden

Lineare Einfachregression

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad s_x^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2 \quad s_{x,y} = \frac{1}{n} \sum_{i=1}^n x_i y_i - \bar{x} \cdot \bar{y}$$

$$\hat{y} = \hat{\alpha} + \hat{\beta} \cdot x \quad \hat{\beta} = \frac{s_{x,y}}{s_x^2} \quad \hat{\alpha} = \bar{y} - \hat{\beta} \cdot \bar{x} \quad R^2 = \hat{\beta}^2 \cdot \frac{s_x^2}{s_y^2}$$

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^n \hat{\epsilon}_i^2 \quad \hat{\sigma}_{\hat{\alpha}}^2 = \hat{\sigma}^2 \cdot \frac{\sum_{i=1}^n x_i^2}{n \left(\sum_{i=1}^n x_i^2 - n \cdot \bar{x}^2 \right)} \quad \hat{\sigma}_{\hat{\beta}}^2 = \frac{\hat{\sigma}^2}{\sum_{i=1}^n x_i^2 - n \cdot \bar{x}^2}$$

Hypothesentests

$$G_{\alpha} = \frac{\hat{\alpha} - \alpha_0}{\hat{\sigma}_{\hat{\alpha}}} \quad G_{\beta} = \frac{\hat{\beta} - \beta_0}{\hat{\sigma}_{\hat{\beta}}} \quad G_{\sigma^2} = (n-2) \cdot \frac{\hat{\sigma}^2}{\sigma_0^2}$$

$$G_{\alpha} \sim t(n-2) \quad G_{\beta} \sim t(n-2) \quad G_{\sigma^2} \sim \chi^2(n-2)$$

Konfidenzintervalle

$$\left[\hat{\alpha} - \hat{\sigma}_{\hat{\alpha}} t_{1-\alpha_0/2}(n-2), \hat{\alpha} + \hat{\sigma}_{\hat{\alpha}} t_{1-\alpha_0/2}(n-2) \right] \quad \left[\hat{\beta} - \hat{\sigma}_{\hat{\beta}} t_{1-\alpha_0/2}(n-2), \hat{\beta} + \hat{\sigma}_{\hat{\beta}} t_{1-\alpha_0/2}(n-2) \right]$$

Multiple lineare Regression

Testverfahren

$$t\text{-Test: } T_j = \frac{\hat{\beta}_j - \beta_{0,j}}{\hat{\sigma}_j} \sim t(n-k-1) \quad F\text{-Test: } F = \frac{R^2}{1-R^2} \cdot \frac{n-k-1}{k} \sim F(k, n-k-1)$$