

Multivariate Verteilungen

Literatur: Fahrmeir, Hammerle und Tutz: Multivariate statistische Verfahren. De Gruyter, 1996., Kap 2, Anhang A11

Wir betrachten Zufalls-Vektoren \mathbf{X} mit der mehrdimensionale Dichte- und Verteilungsfunktion:

$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

$$F : \mathbb{R}^n \rightarrow \mathbb{R}$$

Die Kovarianzmatrix von \mathbf{X} wird mit $V(\mathbf{X}) = \mathbf{\Sigma}$ bezeichnet

Spektralzerlegung der Kovarianzmatrix Σ

$$\Sigma = PLP^T \quad \text{und} \quad \Sigma^{-1} = PL^{-1}P^T,$$

P : die orthogonale Matrix der orthonormalen Eigenvektoren ist,

L : $\text{Diag}(\lambda_1, \dots, \lambda_p)$ die Diagonalmatrix der Eigenwerte.

$$c^2 = (\mathbf{X} - \mu)^T \mathbf{P} \mathbf{L}^{-1/2} \mathbf{L}^{-1/2} \mathbf{P}^T (\mathbf{X} - \mu) = \mathbf{Y}^T \mathbf{L}^{-1} \mathbf{Y} = \sum_{i=1}^p \frac{y_i^2}{\lambda_i},$$

$$\mathbf{Y} = P^T (\mathbf{X} - \mu).$$

Die Gleichung $\sum_{i=1}^p \frac{y_i^2}{\lambda_i} = c^2$ ist eine Ellipsoidengleichung mit den Hauptachsenlängen $\sqrt{\lambda_i} c$

$$\mathbf{Y} = (v_1^T (\mathbf{X} - \mu), \dots, v_p^T (\mathbf{X} - \mu)),$$

sind Projektionen von $\mathbf{X} - \mu$ auf die Eigenvektoren v_1, \dots, v_p .

Sphärische und Elliptische Verteilungen

Verallgemeinerung von eindimensionalen Dichten durch Rotation um Zentrum x_0 :

f sei symmetrisch um 0

$$f(\mathbf{x}) = \text{const} \cdot g(\sqrt{\mathbf{x}^T \mathbf{x}})$$

Verallgemeinerung auf elliptische Verteilungen:

$$f(\mathbf{x}) = |\boldsymbol{\Sigma}|^{-\frac{1}{2}} (\mathbf{g}(\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} \mathbf{g}(\mathbf{x} - \boldsymbol{\mu})) \cdot \text{const}$$

Kurven gleicher Dichte sind Ellipsen.

Dichte der multivariaten Normalverteilung

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{p/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right\},$$

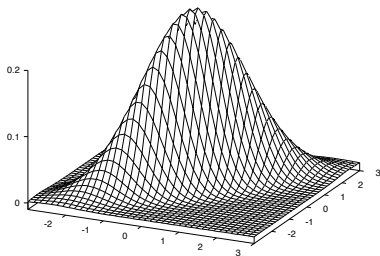
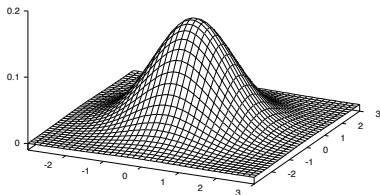
$\boldsymbol{\mu}^T = (\mu_1, \dots, \mu_p) = E(\mathbf{X}^T)$ Erwartungswert.

$\boldsymbol{\Sigma}$: Kovarianzmatrix

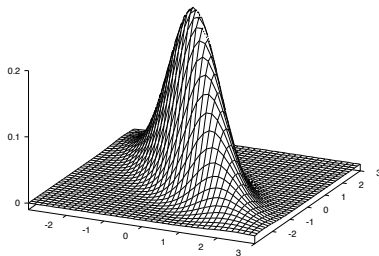
Höhenlinien der Normalverteilung:

$$(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) = \mathbf{c}^2,$$

Normalverteilungsdichte I



Normalverteilungsdichte II



Einige Eigenschaften

- Gilt $\mathbf{X} \sim \mathbf{N}_p(\mu, \Sigma)$, dann ist $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b}$ mit $(q \times p)$ -Matrix \mathbf{A} und $(q \times 1)$ -Vektor \mathbf{b} wiederum normalverteilt mit $\mathbf{Y} \sim N_q(\mathbf{A}\mu + \mathbf{b}, \mathbf{A}\Sigma\mathbf{A}^T)$.
- Gilt $\mathbf{X} \sim \mathbf{N}_p(\mu, \Sigma)$, dann ist $\mathbf{Y} = \Sigma^{-1/2}(\mathbf{X} - \mu)$ standardnormalverteilt, d.h. $\mathbf{Y} \sim N_p(\mathbf{0}, \mathbf{I})$. Die quadratische Form $(\mathbf{X} - \mu)^T \Sigma^{-1}(\mathbf{X} - \mu)$ ist damit χ^2 -verteilt, $(\mathbf{X} - \mu)^T \Sigma^{-1}(\mathbf{X} - \mu) \sim \chi^2(p)$.

Einige Eigenschaften

- Sei $\mathbf{X} \sim \mathbf{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ partitioniert in $\mathbf{X}^T = (\mathbf{X}_1^T, \mathbf{X}_2^T)$ mit zugehörigen Partitionen

$$\begin{aligned}\boldsymbol{\mu}^T &= (\boldsymbol{\mu}_1^T, \boldsymbol{\mu}_2^T), \\ \boldsymbol{\Sigma} &= \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}.\end{aligned}$$

Dann gilt für die bedingte Verteilung

$$\mathbf{X}_2 \mid \mathbf{X}_1 \sim \mathbf{N}(\boldsymbol{\mu}_{2.1}, \boldsymbol{\Sigma}_{2.1}),$$

wobei

$$\begin{aligned}\boldsymbol{\mu}_{2.1} &= \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1), \\ \boldsymbol{\Sigma}_{2.1} &= \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21} \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12}.\end{aligned}$$

Wishart–Verteilung

Seien $x_1, \dots, x_m \sim N_p(\mathbf{0}, \Sigma)$ und unabhängig. Man betrachtet die $(p \times p)$ -Matrix $M = \sum_{i=1}^m x_i x_i^T = \mathbf{X}^T \mathbf{X}$, wobei

$$\mathbf{X} = \begin{pmatrix} x_1^T \\ \vdots \\ x_m^T \end{pmatrix}$$

Die $(p \times p)$ -Matrix $M = \sum_{i=1}^m x_i x_i^T$ besitzt eine Wishart–Verteilung mit den Parametern Σ und m , $M \sim W_p(\Sigma, m)$.

Die Standardform der Verteilung liegt vor, wenn $\Sigma = I$ gilt. Gilt

$M = \sum_{i=1}^m x_i x_i^T \sim W(\Sigma, m)$, erhält man für

$$\Sigma^{-1/2} M \Sigma^{-1/2} = \sum_{i=1}^m (\Sigma^{-1/2} x_i) (\Sigma^{-1/2} x_i)^T \sim W_p(I, m).$$

X besitzt die Form einer Datenmatrix mit m unabhängigen Wiederholungen einer p –dimensionalen normalverteilten Größe. M ist daher die SSP–Matrix (Matrix of sums of squares and products). Sei $\mathbf{x}_i^T = (x_{i1}, \dots, x_{im})$, dann gilt

$$\mathbf{X} = \begin{pmatrix} \mathbf{x}_1^T \\ \vdots \\ \mathbf{x}_m^T \end{pmatrix} = [\mathbf{x}_{(1)}, \dots, \mathbf{x}_{(m)}],$$

wobei $\mathbf{x}_{(j)}^T = (x_{1j}, \dots, x_{mj})$ zur Komponente j gehört.

SSP–Matrix II

Man erhält

$$\begin{aligned} M = (m_{ij}) &= \mathbf{X}^T \mathbf{X} = \begin{pmatrix} x_{(1)}^T \\ \vdots \\ x_{(p)}^T \end{pmatrix} [\mathbf{x}_{(1)}, \dots, \mathbf{x}_{(p)}] \\ &= \begin{pmatrix} x_{(1)}^T x_{(1)} & x_{(1)}^T x_{(2)} & \dots & x_{(1)}^T x_{(p)} \\ x_{(2)}^T x_{(1)} & & & \\ \vdots & & & \\ x_{(p)}^T x_{(1)} & & & x_{(p)}^T x_{(p)} \end{pmatrix} \end{aligned}$$

mit

$$\begin{aligned} x_{(j)}^T x_{(j)} &= \sum_{i=1}^m x_{ij}^2. \\ x_{(j)}^T x_{(s)} &= \sum_{i=1}^m x_{ij} x_{is}. \end{aligned}$$

Eigenschaften

1. Sei

$$M \sim W(\Sigma, m) \implies E(M) = m\Sigma$$

2. Für $p = 1$ gilt $M = \sum_{i=1}^m x_i^2$, wobei $x_i \sim N(0, \sigma^2)$, so dass $M/\sigma^2 = \sum_{i=1}^m (x_i/\sigma)^2$ eine $\chi^2(m)$ -Verteilung besitzt. Die Wishart-Verteilung $W_1(\sigma^2, m)$ ist somit äquivalent zur $\sigma^2\chi^2(m)$ -Verteilung.

3. Gelte $M = \mathbf{X}^T \mathbf{X} \sim \mathbf{W}_p(\Sigma, m)$ und B sei $(p \times q)$ -Matrix, dann gilt mit $\mathbf{Y} = \mathbf{X}B$

$$B^T M B = B^T \mathbf{X}^T \mathbf{X} B = \mathbf{Y}^T \mathbf{Y} \sim \mathbf{W}_q(B^T \Sigma B, m).$$

4. Als Spezialfall von (3) ergibt sich die Verteilung von quadratischen Formen. Gilt $M \sim W_p(\Sigma, m)$ und a ist fester Vektor mit $a^T \Sigma a \neq 0$, dann gilt

$$\frac{a^T M a}{a^T \Sigma a} \sim \chi^2(m) = W_1(a^T \Sigma a, m).$$

5. Gilt $M_1 \sim W_p(\Sigma, m_1)$, $M_2 \sim W_p(\Sigma, m_2)$ und M_1 und M_2 sind unabhängig, dann gilt

$$M_1 + M_2 \sim W_p(\Sigma, m_1 + m_2).$$

Hotellings T^2 -Verteilung

Bestimmung:

Sei $d \sim N_p(0, I)$ und $M \sim W_p(I, m)$ und d und M seien unabhängig, dann folgt

$$md^T M^{-1}d \sim T^2(p, m)$$

Hotellings T^2 -Verteilung, wobei p die Dimension des Vektors unabhängiger Normalverteilungen ist und m die Anzahl der Komponenten der Wishart-Verteilung.

Allgemeiner seien x und M unabhängig mit $x \sim N_p(\mu, \Sigma)$, $M \sim W_p(\Sigma, m)$, dann erhält man

$$m(x - \mu)^T M^{-1}(x - \mu) \sim T^2(p, m) .$$

Weitere Eigenschaften und Anwendungen

1. T^2 und t-Verteilung

Die $T^2(1, m)$ -Verteilung entspricht dem Quadrat der $t(m)$ -Verteilung und damit der $F(1, m)$ -Verteilung.

2. T^2 und F -Verteilung

Genereller gilt eine Äquivalenz zwischen der T^2 -Verteilung und der F -Verteilung in der Form

$$T^2(p, m) \triangleq \{(mp)/(m - p + 1)\} F(p, m - p + 1) .$$

bzw.

$$F(p, s) = \frac{s}{s + p - 1} T^2(p, s + p - 1)$$

3. Verteilung der empirischen Mahalanobis-Distanz zwischen \bar{x} und μ . Aus (2) folgt wegen

$$\begin{aligned} (n - 1)(\bar{x} - \mu)^T S^{-1}(\bar{x} - \mu) &\sim T^2(p, n - 1) \\ \frac{n - p}{p}(\bar{x} - \mu)^T S^{-1}(\bar{x} - \mu) &\sim F(p, n - p) . \end{aligned}$$

Wilks Λ -Verteilung erhält man aus zwei unabhängigen Wishart-verteilten Größen $A \sim W_p(I, m)$, $B \sim W_p(I, n)$, $m \geq p$. Die Größe

$$\Lambda = \frac{|A|}{|A + B|} \sim \Lambda(p, m, n)$$

folgt der Wilks-Verteilung $\Lambda(p, m, n)$ mit Parametern p, m, n .

Durch Erweitern mit $|A|^{-1}$ folgt wegen $|AB| = |A| |B|$ für Λ die Darstellung $\Lambda = |I + A^{-1}B|^{-1}$.

1. Für den eindimensionalen Spezialfall $A \sim \chi^2(1)$, $B \sim \chi^2(1)$ ergibt sich die Beta-Verteilung $\Lambda(1, 1, 1) \hat{=} B(0.5, 0.5)$.
2. Für $p = 1$ besitzt A eine $\chi^2(m)$ -Verteilung und B eine $\chi^2(n)$ -Verteilung, so dass

$$\Lambda(1, m, n) = B(m/2, n/2).$$

3. Die Verteilungen $\Lambda(p, m, n)$ und $\Lambda(n, m + n - p, p)$ sind identisch.
4. Die Verteilung von $\Lambda(p, m, n)$ ist identisch mit der Verteilung des Produkts

$$u_1 \cdot \dots \cdot u_n,$$

wobei $u_i \sim B((m + i - p)/2, p/2)$, $i = 1, \dots, n$.

5. Die Verteilung von $\Lambda(p, m, 1)$ ist äquivalent zu $B((m + 1 - p)/2, p/2)$.

- **Multinomialverteilung**
- **Multivariate Hypergeometrische Verteilung**

Diese Verteilung entspricht der Verallgemeinerung des SZiehen ohne Zurücklegen". Gezogen werden n aus N Objekten, die in K Klassen zerfallen. Es sind jeweils N_1, \dots, N_K Objekte vorhanden. Die Wahrscheinlichkeitsfunktion ist

$$P(X_1 = n_1, \dots, X_K = n_K) = \frac{\prod_{k=1}^K \binom{N_k}{n_k}}{\binom{N}{n}} \quad \text{für } \sum n_k = n$$