Multivariate Verteilungen

Gerhard Tutz
LMU München

Inhaltsverzeichnis

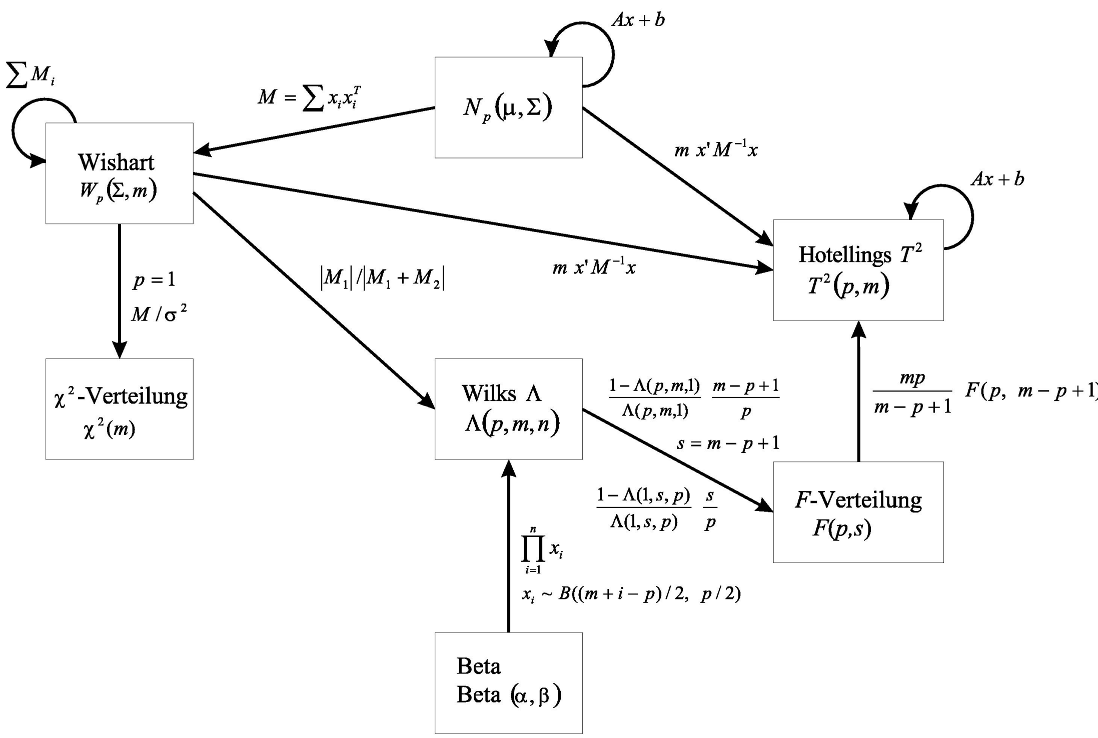

Abb. 1: Übersicht über multivariate Verteilungen

1 Multivariate Normalverteilung

Ein Zufallsvektor $\boldsymbol{x} = (X_1, \ldots, X_p)$ heißt (nicht entartet) multivariat normalverteilt, $\boldsymbol{x} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, wenn die Dichte gegeben ist durch

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

wobei $\boldsymbol{\mu}^T = (\mu_1, \ldots, \mu_p) = E(\boldsymbol{x}^T)$ den Erwartungswert und $\boldsymbol{\Sigma} = cov(\mathbf{x})$ die Kovarianzmatrix des Zufallsvektors $\boldsymbol{x}$ bezeichnet. Vorausgesetzt ist, dass $\boldsymbol{\Sigma}$ positiv definit ist, $\mid \boldsymbol{\Sigma} \mid$ bezeichnet die Determinante der Kovarianzmatrix.

Alternativ lässt sich die Normalverteilung charakterisieren durch die Projektionen. Ein Zufallsvektor ist normalverteilt, wenn $\boldsymbol{a}^T\boldsymbol{x}$ eindimensional normalverteilt ist für jedes $\boldsymbol{a} \in \mathbb{R}^p$ (Für $\boldsymbol{a} = \mathbf{0}$ wird die entartete Verteilung zugelassen). Faßt man $\boldsymbol{a}^T\boldsymbol{x}$ als Projektion auf den Vektor $\boldsymbol{a}$ auf, ergibt sich damit, dass jede beliebige Projektion einer multivariaten Normalverteilung eine eindimensionale Normalverteilung ergibt.

Isodensiten

Die Isodensiten oder Höhenlinien sind diejenigen Kurven, die dieselbe Dichte besitzen. Für die Normalverteilung sind die Isodensiten Ellipsoide, die bestimmt sind durch

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

wobei c eine beliebige Konstante ist.

Die Spektralzerlegung von $\boldsymbol{\Sigma}$ liefert

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

wobei $\boldsymbol{P}$ die orthogonale Matrix der orthonormalen Eigenvektoren ist, und $\boldsymbol{L} = \text{ Diag }(\lambda_1, \ldots, \lambda_p)$ die Diagonalmatrix der Eigenwerte. Damit ergibt sich

$$c^2 = (\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{PL}^{-1/2}\boldsymbol{L}^{-1/2}\boldsymbol{P}^T(\boldsymbol{x}-\boldsymbol{\mu}) = \boldsymbol{y}^T\boldsymbol{L}^{-1}\boldsymbol{y} = \sum_{i=1}^{p} \frac{y_i^2}{\lambda_i},$$

wobei $\boldsymbol{y} = \boldsymbol{P}^T(\boldsymbol{x}-\boldsymbol{\mu})$. Die Gleichung $\sum_{i=1}^{p} \frac{y_i^2}{\lambda_i} = c^2$ ist eine Ellipsoidengleichung mit den Hauptachsenlängen $\sqrt{\lambda_i}c$.

Seien die Eigenvektoren bestimmt durch $\boldsymbol{P} = (\mathbf{v}_1, \ldots, \mathbf{v}_p)$, so sind die Komponenten von $\boldsymbol{y}$ bestimmt durch

$$\boldsymbol{y} = (\mathbf{v}_1^T(\boldsymbol{x}-\boldsymbol{\mu}), \ldots, \mathbf{v}_p^T(\boldsymbol{x}-\boldsymbol{\mu})),$$

sind also Projektionen von $\boldsymbol{x}-\boldsymbol{\mu}$ auf die Eigenwerte $\mathbf{v}_1, \ldots, \mathbf{v}_p$.

Einige Eigenschaften:

1. Gilt $\boldsymbol{x} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, dann ist $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}$ mit $(q \times p)$–Matrix $\boldsymbol{A}$ und $(q \times 1)$–Vektor $\mathbf{b}$ wiederum normalverteilt mit $\boldsymbol{y} \sim N_q(\boldsymbol{A}\boldsymbol{\mu} + \boldsymbol{b}, \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}^T)$.

2. Gilt $\boldsymbol{x} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, dann ist $\boldsymbol{y} = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{x}-\boldsymbol{\mu})$ standardnormalverteilt, d.h. $\boldsymbol{y} \sim N_p(\mathbf{0}, \boldsymbol{I})$. Die quadratische Form $(\boldsymbol{x}-\boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})$ ist damit χ^2–verteilt, $(\boldsymbol{x}-\boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) \sim \chi^2(p)$.

3. Sei $\boldsymbol{x} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ partitioniert in $\boldsymbol{x}^T = (\boldsymbol{x}_1^T, \boldsymbol{x}_2^T)$ mit zugehörigen Partitionen

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

Dann gilt für die bedingte Verteilung

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

wobei

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

Für den Spezialfall $\boldsymbol{x}_2 = y$ (eindimensional), $\boldsymbol{x}_1^T = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{p-1})$ gilt $\boldsymbol{\Sigma}_{22} = \sigma_y^2$, $\boldsymbol{\Sigma}_{21} = (\text{cov}(y, x_1), \ldots, \text{cov}(y, x_p)) = \boldsymbol{\Sigma}_{12}^T$ und daher mit $\sigma_{yx_i} = \text{cov}(y, x_i)$

$$
\begin{aligned}
\mu_{2.1} &= E(y) + \sum_{i=1}^{p} \sigma_{yx_i}(x_i - \mu_i)/\sigma_y^2 \\
&= E(y) - \sum_{i=1}^{p} \sigma_{yx_i}\mu_i/\sigma_y^2 + \sum_{i=1}^{p} \frac{\sigma_{yx_i}}{\sigma_y^2} x_i. \\
\mathbf{\Sigma}_{2.1} &= \sigma_y^2 - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}.
\end{aligned}
$$

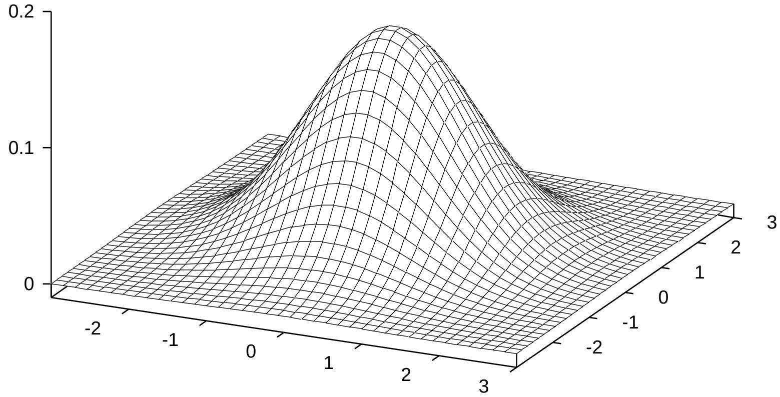

Abb. 2: Zweidimensionale Normalverteilungsdichte für unkorrelierte Merkmale, $\rho = 0$, mit $\mu_1 = \mu_2 = 0, \sigma_1 = \sigma_2 = 1.0$

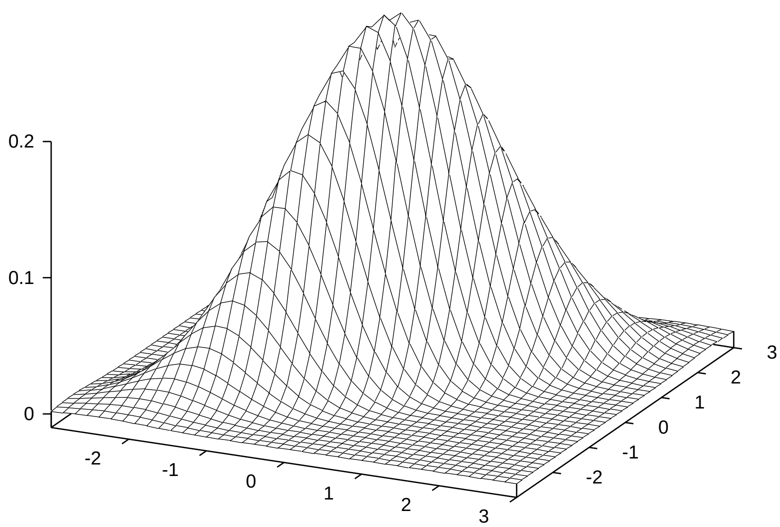

Abb. 3: Zweidimensionale Normalverteilungsdichte, $\rho = 0.8$, $\mu_1 = \mu_2 = 0, \sigma_1 = \sigma_2 = 1.0$

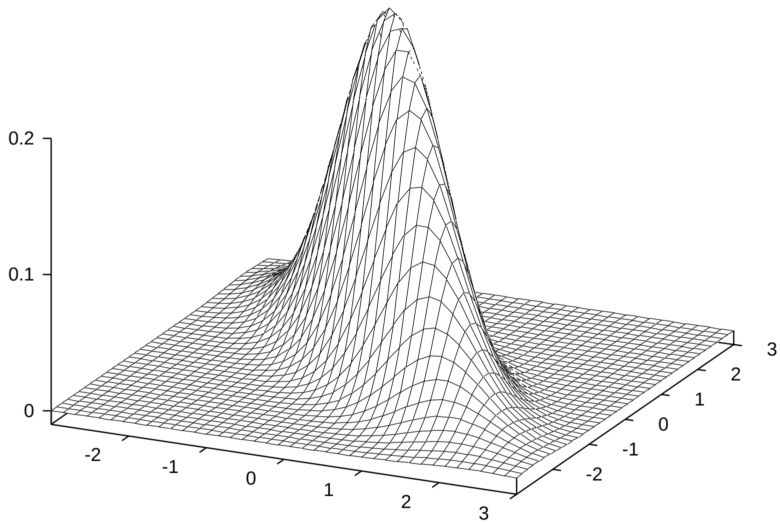

Abb. 4: Zweidimensionale Normalverteilungsdichte,
$\rho = -0.8, \mu_1 = \mu_2 = 0, \sigma_1 = \sigma_2 = 1.0$

2 Wishart–Verteilung

Die Wishart–Verteilung ist für Zufallsmatrizen definiert, d.h. für Matrizen, deren sämtliche Komponenten Zufallsvariablen sind.

Seien $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m \sim N_p(\mathbf{0}, \boldsymbol{\Sigma})$ und unabhängig. Man betrachtet die $(p \times p)$–Matrix $\mathrm{M} = \sum\limits_{i=1}^{m} \boldsymbol{x}_i \boldsymbol{x}_i^T = \boldsymbol{X}^T \boldsymbol{X}$, wobei

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{x}_1^T \\ \vdots \\ \boldsymbol{x}_m^T. \end{pmatrix}$$

Die $(p \times p)$–Matrix $\boldsymbol{M} = \sum\limits_{i=1}^{m} \boldsymbol{x}_i \boldsymbol{x}_i^T$ besitzt eine Wishart–Verteilung mit den Parametern $\boldsymbol{S}$ und m, $\boldsymbol{M} \sim W_p(\boldsymbol{\Sigma}, m)$.

Die Standardform der Verteilung liegt vor, wenn $\boldsymbol{\Sigma} = \boldsymbol{I}$ gilt. Gilt $\boldsymbol{M} = \sum\limits_{i=1}^{m} \boldsymbol{x}_i \boldsymbol{x}_i^T \sim W(\boldsymbol{\Sigma}, m)$ erhält man für $\boldsymbol{\Sigma}^{-1/2} \boldsymbol{M} \boldsymbol{\Sigma}^{-1/2} = \sum\limits_{i=1}^{m} (\boldsymbol{\Sigma}^{-1/2} \boldsymbol{x}_i)(\boldsymbol{\Sigma}^{-1/2} \boldsymbol{x}_i)^T \sim W_p(\boldsymbol{I}, m)$.

Anwendung: SSP–Matrix

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

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

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

Man erhält

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

mit

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

Empirische Varianz

Sind $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m$ unabhängige Wiederholungen von $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, dann ist $\bar{\boldsymbol{z}} = (\boldsymbol{z}_1 + \ldots + \boldsymbol{z}_m)/m$ normalverteilt mit $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}/n)$ und $\boldsymbol{z}_1 - \bar{\boldsymbol{z}}, \ldots, \boldsymbol{z}_m - \bar{\boldsymbol{z}}$ sind verteilt wie $N\left(\boldsymbol{0}, \frac{m-1}{m}\boldsymbol{\Sigma}\right)$, allerdings nicht mehr unabhängig. Die empirische Kovarianzmatrix

$$\boldsymbol{S} = \frac{1}{m} \sum_{i=1}^{m} (\boldsymbol{z}_i - \bar{\boldsymbol{z}})(\boldsymbol{z}_i - \bar{\boldsymbol{z}})^T$$

besitzt für $\boldsymbol{x}_i = \boldsymbol{z}_i - \bar{\boldsymbol{x}}$ die Struktur der Matrix $\boldsymbol{M}$. Man erhält als Verteilung

$$m\boldsymbol{S} \sim W_p(\boldsymbol{\Sigma}, m-1),$$

d.h. in $m\boldsymbol{\Sigma}$ sind nur $m-1$ unabhängige Komponenten enthalten.

Eigenschaften:

1. 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.

2. Gelte $\boldsymbol{M} = \boldsymbol{X}^T\boldsymbol{X} \sim W_p(\boldsymbol{\Sigma}, m)$ und $\boldsymbol{B}$ sei $(p \times q)$–Matrix, dann gilt mit $\boldsymbol{Y} = \boldsymbol{XB}$

$$\boldsymbol{B}^T\boldsymbol{MB} = \boldsymbol{B}^T\boldsymbol{X}^T\boldsymbol{XB} = \boldsymbol{Y}^T\boldsymbol{Y} \sim W_q(\boldsymbol{B}^T\hat{\boldsymbol{\Sigma}}\boldsymbol{B}, m) \; .$$

3. Als Spezialfall von (2) ergibt sich die Verteilung von quadratischen Formen. Gilt $\boldsymbol{M} \sim W_p(\hat{\boldsymbol{\Sigma}}, m)$ und $\boldsymbol{a}$ ist fester Vektor mit $\boldsymbol{a}^T\hat{\boldsymbol{\Sigma}}\boldsymbol{a} \neq 0$, dann gilt

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

4. Gilt $\boldsymbol{M}_1 \sim W_p(\hat{\boldsymbol{\Sigma}}, m_1)$, $\boldsymbol{M}_2 \sim W_p(\hat{\boldsymbol{\Sigma}}, m_2)$ und $\boldsymbol{M}_1$ und $\boldsymbol{M}_2$ sind unabhängig, dann gilt

$$\boldsymbol{M}_1 + \boldsymbol{M}_2 \sim W_p(\hat{\boldsymbol{\Sigma}}, m_1 + m_2) \; .$$

5. Sei $\boldsymbol{X}$ aus n unabhängigen Ziehungen aus $N_p(\boldsymbol{0}, \boldsymbol{\Sigma})$ konstruiert und $\boldsymbol{C}$ sei symmetrische $(n \times n)$–Matrix.

   Die Matrix $\boldsymbol{\Sigma}^T\boldsymbol{CX}$ besitzt eine Wishart–Verteilung, genau dann, wenn $\boldsymbol{C}$ idempotent ist. Es gilt dann

$$\mathbf{X}^T\boldsymbol{CX} \sim W_p(\boldsymbol{\Sigma}, r) \;\text{ mit }\; r = rg(\boldsymbol{C}) = tr(\boldsymbol{C}) \; .$$

6. Als Spezialfall von (5) ergibt sich für die empirische Kovarianzmatrix

$$\boldsymbol{S} = \frac{1}{n}\boldsymbol{X}^T\boldsymbol{HX} \;\text{ mit }\; \boldsymbol{H} = \boldsymbol{I} - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T$$

   die Aussage $n\boldsymbol{S} \sim W(\boldsymbol{\Sigma}, n-1)$, da die Zentrierungs–Matrix $\boldsymbol{H}$ den Rang $n-1$ besitzt.

Funktionen von Wishart–Verteilungen: Seien $\boldsymbol{A} \sim W_p(\boldsymbol{I}, m)$, $\boldsymbol{B} \sim W_p(\boldsymbol{I}, n)$, $m \geq p$ dann gilt

1. $\Lambda = \dfrac{|\, \boldsymbol{A} \,|}{|\, \boldsymbol{A} + \boldsymbol{B} \,|} \sim \Lambda(p, m, n)$
2. Der größte Eigenwert λ von $(\boldsymbol{A} + \boldsymbol{B})^{-1}\boldsymbol{B}$ besitzt eine Verteilung

$$\theta(p, m, n)$$

mit den Eigenschaften

$$\begin{aligned} \frac{\theta(1, m, n)}{1 - \theta(1, m, n)} &= \frac{1 - \Lambda(1, m, n)}{\Lambda(1, m, n)} \sim \frac{n}{m} F(n, m) \\ \frac{\theta(p, m, 1)}{1 - \theta(p, m, 1)} &= \frac{1 - \Lambda(p, m, 1)}{\Lambda(p, m, 1)} \sim \frac{p}{m - p + 1} F(p, m - p + 1). \end{aligned}$$

Ein anderer Zugang ergibt sich für den größten Eigenwert $\tilde{\lambda}$ von $\boldsymbol{A}^{-1}\boldsymbol{B}$, da $\lambda = \tilde{\lambda}/(1 + \tilde{\lambda})$ ein Eigenwert von $(\boldsymbol{A} + \boldsymbol{B})^{-1}B$ ist. Damit erhält man

$$\theta = \frac{\tilde{\lambda}}{1 + \tilde{\lambda}} \text{ mit } \tilde{\lambda} \text{ größter Eigenwert von } \boldsymbol{A}^{-1}\boldsymbol{B} \,.$$

3 Hotellings T^2–Verteilung

Bestimmung:

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

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

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

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

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

Die Aussage ergibt sich aus $\tilde{\boldsymbol{d}} = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{x} - \boldsymbol{\mu}) \sim N(0, \boldsymbol{I})$ und $\tilde{\boldsymbol{M}} = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{M}\boldsymbol{\Sigma}^{-1/2} \sim W_p(\boldsymbol{I}, m)$ aus $m\tilde{\boldsymbol{d}}^T\tilde{\boldsymbol{M}}^{-1}\tilde{\boldsymbol{d}} = m(\boldsymbol{x} - \boldsymbol{\mu})^T\boldsymbol{M}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})$.

Anwendung:

Sei $\bar{\boldsymbol{x}} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\boldsymbol{x}_i \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}/n)$ der Mittelwert aus n unabhängigen normalverteilten Zufallsvariablen $\boldsymbol{x}_i \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ und $\boldsymbol{S} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(\boldsymbol{x}_i - \boldsymbol{x})(\boldsymbol{x}_i - \boldsymbol{x})^T$ die empirische Kovarianzmatrix. Dann gilt

$$(n-1)(\bar{\boldsymbol{x}} - \boldsymbol{\mu})^T\boldsymbol{S}^{-1}(\bar{\boldsymbol{x}} - \boldsymbol{\mu}) \sim T^2(p, n-1) \; .$$

Die Aussage ergibt sich aus $n\boldsymbol{S} \sim W_p(\boldsymbol{\Sigma}, n-1)$ und $n^{1/2}(\bar{\boldsymbol{x}} - \boldsymbol{\mu}) \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ aus $(n-1)n^{1/2}(\bar{\boldsymbol{x}} - \boldsymbol{\mu})^T(n\boldsymbol{S})^{-1}n^{1/2}(\bar{\boldsymbol{x}} - \boldsymbol{\mu}) = (n-1)(\bar{\boldsymbol{x}} - \boldsymbol{\mu})^T\boldsymbol{S}^{-1}(\bar{\boldsymbol{x}} - \boldsymbol{\mu})$.

Für den Spezialfall $p = 1$ ergibt sich

$$\begin{aligned}
&(n-1)(\bar{\boldsymbol{x}}-\boldsymbol{\mu})^T\boldsymbol{S}^{-1}(\bar{\boldsymbol{x}}-\boldsymbol{\mu}) \\
&= \frac{(\bar{x}-\mu)^2}{n\dfrac{1}{n-1}\sum_i (x_i-\bar{x})^2} \\
&= \left(\frac{\bar{x}-\mu}{s/\sqrt{n}}\right)^2 \sim T^2(1,n-1)\ .
\end{aligned}$$

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

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

   bzw.

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

3. Verteilung der empirischen Mahalanobis–Distanz zwischen $\bar{\boldsymbol{x}}$ und $\boldsymbol{\mu}$.
   Aus (2) folgt wegen

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

4. Verteilung der empirischen Mahalanobis–Distanz für zwei Stichproben.
   Aus unabhängigen Stichproben $\boldsymbol{x}_1^{(i)},\ldots,\boldsymbol{x}_{n_i}^{(i)} \sim N(\boldsymbol{\mu}_i,\boldsymbol{\Sigma}_i)\ ,\ i=1,2$ erhält man die Mittelwerte und empirischen Kovarianzen

   $$\bar{\boldsymbol{x}}_i = \frac{1}{n_i}\sum_{j=1}^{n_i}\boldsymbol{x}_j^{(i)}\ \ ,\ \ \boldsymbol{S}_i = \frac{1}{n_i}\sum_{j=1}^{n_i}(\boldsymbol{x}_j^{(i)}-\bar{\boldsymbol{x}}_i)(\boldsymbol{x}_j^{(i)}-\bar{\boldsymbol{x}}_i)^T$$

und mit der gepoolten Matrix $\boldsymbol{S} = (n_1\boldsymbol{S}_1 + n_2\boldsymbol{S}_2)(n-2)$ die quadrierte Mahalanobis–Distanz

$$D^2 = (\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2)^T \boldsymbol{S}^{-1} (\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2).$$

Gilt $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$, $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2$, so folgt

$$\frac{n_1 n_2}{n_1 + n_2} D^2 \sim T^2(p, n-2)$$

bzw.

$$\frac{n_1 n_2 (n-p-1)}{n(n-2)p} D^2 \sim F(p, n-p-1) \; .$$

5. T^2 ist invariant gegenüber Transformationen
$t(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} + b$, wobei $\boldsymbol{A}$ eine nichtreguläre Matrix ist.

4 Wilks Λ

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

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

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

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

Eigenschaften:

1. Für den eindimensionalen Spezialfall $\boldsymbol{A} \sim \chi^2(1)\ ,\ \boldsymbol{B} \sim \chi^2(1)$ ergibt sich die Beta–Verteilung $\Lambda(1, 1, 1) \mathrel{\widehat{=}} B(0.5, 0.5)$.

2. Für $p = 1$ besitzt $\boldsymbol{A}$ eine $\chi^2(m)$–Verteilung und $\boldsymbol{B}$ 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 der Verteilung des Produkts

$$u_1 \cdot \ldots \cdot u_n,$$

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

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

6. Weiter gilt

$$\begin{aligned} \frac{1 - \Lambda(p, m, 1)}{\Lambda(p, m, 1)} &\sim \frac{p}{m - p + 1}\, F(p, m - p + 1)\ . \\ \frac{1 - \Lambda(1, m, n)}{\Lambda(1, m, n)} &\sim \frac{n}{m}\, F(n, m)\ . \end{aligned}$$