
Exkurs: Multivariate Verteilungen

Dr. Dirk Hasenclever

Mehrdimensionale / Multivariate Verteilungen

Zufallsvariable X mit Realisierungen $\begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathbb{R}^p$

Dichte: $f(x_1, x_2, \dots, x_p) dx_1 dx_2 \dots dx_p$

$$A \subset \mathbb{R}^p: \quad p(X \in A) = \int_A f(x_1, \dots, x_p) dx_1 \dots dx_p$$



Bei Koordinaten Transformationen Betrag der Determinante der
Jakobi matrix nicht vergessen

Normal Verteilung

X hat Realisierungen $x \in \mathbb{R}$

$$f(x) = \frac{1}{\sqrt{2\pi} \sigma_x} \exp\left(-\frac{1}{2} \frac{(x - \mu_x)^2}{\sigma_x^2}\right) dx$$

$\mu_x \in \mathbb{R}$ Mittelwert

$\sigma_x^2 \in \mathbb{R}_+$ Varianz

$$X \sim \mathcal{N}(\mu_x, \sigma_x^2)$$

Multinormal Verteilung

X hat Realisierungen $x \in \mathbb{R}^p$

$$f(x) = \frac{1}{(2\pi)^p \sqrt{\det(\sum_x)}} \exp\left(-\frac{1}{2} (x - \mu_x)^T \sum_x^{-1} (x - \mu_x)\right) dx_1 \dots dx_p$$

$\mu_x \in \mathbb{R}^p$ Mittelwert

$\sum_x = \begin{pmatrix} \sigma_{11} & \dots & \sigma_{1p} \\ \vdots & & \vdots \\ \sigma_{p1} & \dots & \sigma_{pp} \end{pmatrix}$ Kovarianzmatrix
(positiv definit)

$$X \sim \mathcal{N}(\mu_x, \sum_x)$$

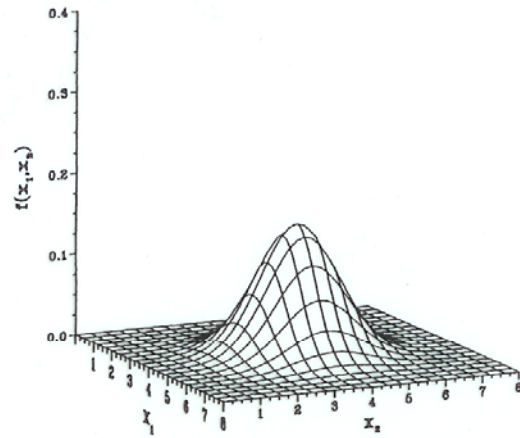


Abb.7: Dichte der bivariaten Normalverteilung $N\left(\begin{pmatrix} 4 \\ 4 \end{pmatrix}; \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$

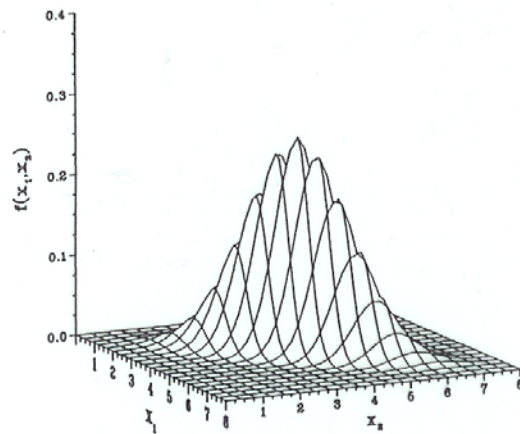


Abb.8: Dichte der bivariaten Normalverteilung $N\left(\begin{pmatrix} 4 \\ 4 \end{pmatrix}; \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}\right)$

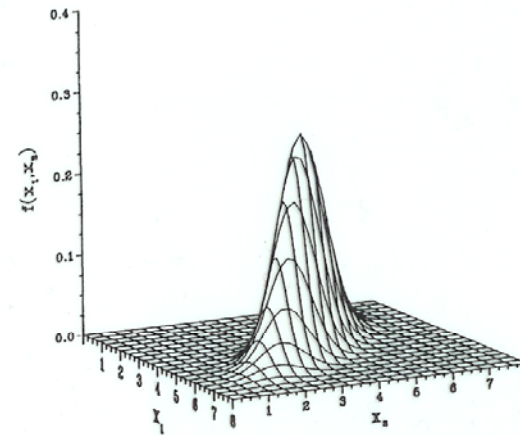


Abb.9: Dichte der bivariaten Normalverteilung $N\left(\begin{pmatrix} 4 \\ 4 \end{pmatrix}; \begin{pmatrix} 1 & -0.8 \\ -0.8 & 1 \end{pmatrix}\right)$

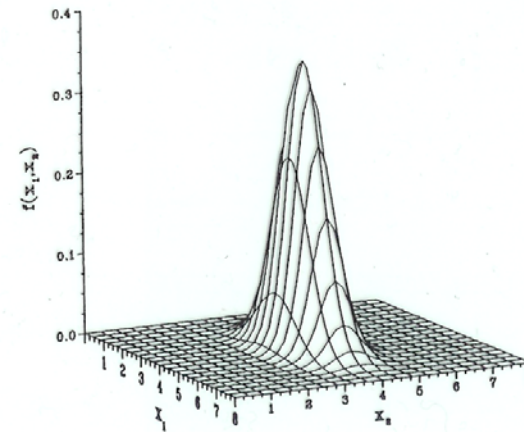


Abb.10: Dichte der bivariaten Normalverteilung $N\left(\begin{pmatrix} 4 \\ 4 \end{pmatrix}; \begin{pmatrix} 1 & 0 \\ 0 & 0.2 \end{pmatrix}\right)$

Aus Hartung: Multivariate Statistik

Marginale / Randverteilungen

$$(X_1, X_2) \sim f(x_1, x_2) dx_1 dx_2$$

$$f_{X_1}(x_1) dx_1 := \left(\int_{-\infty}^{+\infty} f(x_1, x_2) dx_2 \right) dx_1$$

$$f_{X_2}(x_2) dx_2 := \left(\int_{-\infty}^{+\infty} f(x_1, x_2) dx_1 \right) dx_2$$

Projektion auf jeweils eine Komponente

Konditionale Verteilungen:

$$(X_1, X_2) \sim f(x_1, x_2) dx_1 dx_2$$

Wie ist die Verteilung von X_1 nach der Information $X_2 = x_2^*$?

$$f_{X_1|X_2=x_2^*}(x_1) dx_1 = \frac{f(x_1, x_2) dx_1 dx_2}{\int_{-\infty}^{+\infty} f(x_1, x_2) dx_1} dx_2 = \frac{f(x_1, x_2^*)}{\int_{-\infty}^{+\infty} f(x_1, x_2^*) dx_1} dx_1$$

↑
Marginaldichte von x_2

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

Normal Verteilung

(1) $a, b \in \mathbb{R}$

$$aX + b \sim \mathcal{N}(a\mu_X + b, a^2\sigma_X^2)$$

(2) $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$

$$(X + Y) \sim \mathcal{N}(\mu_X + \mu_Y, (\sigma_X^2 + \sigma_Y^2))$$

(3) —

Multi-normal Verteilung

A ($q \times p$) Matrix $b \in \mathbb{R}^q$

$$AX + b \sim \mathcal{N}(A\mu_X + b, A \Sigma_X A^T)$$

Insbesondere sind Projektionen und Randverteilungen wieder normal.

$$Y \sim \mathcal{N}(\mu_Y, \Sigma_Y)$$

$$(X + Y) \sim \mathcal{N}(\mu_X + \mu_Y, (\Sigma_X + \Sigma_Y))$$

konditionale Verteilungen
sind wieder normal.

Binormal Verteilung

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix}$$

mit dem Korrelationskoeffizienten: $\rho = \frac{\sigma_{12}}{\sigma_1 \cdot \sigma_2} \in [-1, 1]$

$$\Sigma^{-1} = \begin{pmatrix} \frac{1}{\sigma_1^2(1-\rho^2)} & \frac{-\rho}{\sigma_1 \sigma_2(1-\rho^2)} \\ \frac{-\rho}{\sigma_1 \sigma_2(1-\rho^2)} & \frac{1}{\sigma_2^2(1-\rho^2)} \end{pmatrix}$$

$$f(x_1, x_2) = \frac{1}{2\pi \sigma_1 \sigma_2 (1-\rho^2)} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left(\frac{(x_1 - \mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1 \sigma_2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} \right) \right\}$$

Binomiale Verteilung II

Die konditionelle Verteilung von X_1 gegeben $X_2 = x_2$
ist univariat normal

$$(X_1 | X_2 = x_2) \sim N\left(\mu_1 + (x_2 - \mu_2)\rho \frac{\sigma_1}{\sigma_2}, \sigma_1^2(1 - \rho^2)\right)$$

Beweis: Umschreiben des Exponenten der Dichte:

$$\frac{1}{2(1-\rho^2)} \left\{ \left(\frac{x_1 - \mu_1}{\sigma_1} - \rho \frac{x_2 - \mu_2}{\sigma_2} \right)^2 + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} (1 - \rho^2) \right\}$$

und Vergleich mit der funktionalen Form der univariaten Normalverteilung.