

POISSON eine gute Näherung, da $np = 1 \leq 10$ und $1500p = 15 \leq 100 = n$. Wir erhalten somit als Näherung

$$P(2 \leq X \leq 10) = \sum_{k=2}^{10} \pi_1(k) = 0,264241.$$

Exakte Rechnung ergibt

$$P(2 \leq X \leq 10) = \sum_{k=2}^{10} \binom{100}{k} 0,01^k 0,99^{100-k} = 0,264238.$$

1.4.2.4 Bemerkung zur hypergeometrischen Verteilung

Es lassen sich auch Grenzverteilungssätze für eine hypergeometrisch verteilte Zufallsgröße $X \sim \mathbf{H}(n, N, M)$ formulieren. Wenn z.B. $N \rightarrow \infty$, $M \rightarrow \infty$ und $\frac{M}{N} \rightarrow p$ für $n \rightarrow \infty$ gilt, so nutzt man die Näherung

$$P(X = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}} \approx \binom{n}{k} p^k (1-p)^{n-k},$$

d.h. man ersetzt die hypergeometrische Verteilung näherungsweise durch eine entsprechende Binomialverteilung. Es gibt jedoch keine handhabbaren Faustregeln für die Nutzung dieser Approximation. Um die so erhaltene Binomialverteilung auszuwerten, kann man dann den Grenzverteilungssatz von MOIVRE/LAPLACE oder den Grenzverteilungssatz von POISSON verwenden.

1.5 Mehrdimensionale Verteilungen

Bisher wurden ausschließlich reellwertige, d.h. eindimensionale, Zufallsgrößen betrachtet. Wir haben also stets nur ein Merkmal des zu beobachtenden Objekts bei unseren Untersuchungen berücksichtigt. Häufig sind aber bei praktischen Modellierungsproblemen mehrere Merkmale der Beobachtungsobjekte gleichermaßen von Interesse. Somit benötigen wir mehrdimensionale Zufallsgrößen.

1.5.1 Einführung

Definition 1.5.1. Sei $\underline{X} = (X_1, X_2, \dots, X_n)$ eine Zusammenfassung von n Zufallsgrößen X_1, X_2, \dots, X_n . Dann heißt das Objekt \underline{X} *n-dimensionale Zufallsgröße* oder *zufälliger Vektor*.

Definition 1.5.2. Die durch

$$F(x_1, \dots, x_n) := P(X_1 < x_1, \dots, X_n < x_n)$$

für alle Vektoren $(x_1, \dots, x_n) \in \mathbb{R}^n$ definierte reelle Funktion F heißt *Verteilungsfunktion des zufälligen Vektors* $\underline{X} = (X_1, \dots, X_n)$.

Wir werden im Folgenden nicht immer den allgemeinen Fall $\underline{X} = (X_1, \dots, X_n)$ betrachten, sondern uns auf $n = 2$, d.h. $\underline{X} = (X, Y)$, beschränken, wenn dies das Verständnis erleichtert. Ein Großteil der Aussagen gilt dann analog für allgemeines n .

Definition 1.5.3. Für einen zufälligen Vektor $\underline{X} = (X, Y)$ heißen die Funktionen

$$F_X(x) := \lim_{y \rightarrow \infty} F(x, y) \quad \text{und} \quad F_Y(y) := \lim_{x \rightarrow \infty} F(x, y)$$

Randverteilung von X bzw. Y . Die zu einer Randverteilung gehörende Wahrscheinlichkeits- bzw. Dichtefunktion bezeichnen wir ebenfalls als Randverteilung.

Eigenschaften der Verteilungsfunktion

- $0 \leq F(x, y) \leq 1 \quad \forall x, y \in \mathbb{R}$.
- $\lim_{x \rightarrow -\infty} F(x, y) = 0 \quad \forall y \in \mathbb{R}, \quad \lim_{y \rightarrow -\infty} F(x, y) = 0 \quad \forall x \in \mathbb{R}$.
- Die Randverteilungen F_X und F_Y sind die Verteilungsfunktionen der eindimensionalen Zufallsgrößen X und Y , d.h.

$$F_X(x) = P(X < x) \quad \forall x \in \mathbb{R} \quad \text{und} \quad F_Y(y) = P(Y < y) \quad \forall y \in \mathbb{R}.$$

- $\lim_{\substack{x \rightarrow \infty \\ y \rightarrow \infty}} F(x, y) = 1$.
- $F(x, y)$ ist in beiden Komponenten monoton wachsend, d.h.

$$\begin{aligned} x_1 < x_2 &\Rightarrow F(x_1, y) \leq F(x_2, y) \quad \forall y \in \mathbb{R}, \\ y_1 < y_2 &\Rightarrow F(x, y_1) \leq F(x, y_2) \quad \forall x \in \mathbb{R}. \end{aligned}$$

- $F(x, y)$ ist in beiden Komponenten linksseitig stetig, d.h.

$$\lim_{\substack{h \rightarrow 0 \\ h > 0}} F(x - h, y) = F(x, y) = \lim_{\substack{h \rightarrow 0 \\ h > 0}} F(x, y - h) \quad \forall x, y \in \mathbb{R}.$$

1.5.1.1 Kovarianz und Korrelationskoeffizient

Für einen zufälligen Vektor (X, Y) bezeichnen wir mit $\mu_X := \mathbf{E}X$ den Erwartungswert und mit $\sigma_X^2 := \mathbf{D}^2X$ die Varianz der Randverteilung von X . Analog $\mu_Y := \mathbf{E}Y$ und $\sigma_Y^2 := \mathbf{D}^2Y$ für die Randverteilung von Y .

Definition 1.5.4. Seien X und Y zwei Zufallsgrößen. Dann heißt die Größe

$$\sigma_{XY} := \mathbf{Cov}(X, Y) := \mathbf{E}(X - \mathbf{E}X)(Y - \mathbf{E}Y)$$

Kovarianz der Zufallsgrößen X und Y . Die Kovarianz ist ein Maß für das stochastische Verhalten der Zufallsgrößen zueinander. Die Zufallsgrößen X und Y heißen *unkorreliert*, wenn $\mathbf{Cov}(X, Y) = 0$ gilt.

Eigenschaften der Kovarianz

- $\mathbf{Cov}(X, Y) = \mathbf{E}(XY) - \mathbf{E}X\mathbf{E}Y = \mathbf{Cov}(Y, X)$,
- $\mathbf{Cov}(a + bX, c + dY) = bd\mathbf{Cov}(X, Y)$,

- $\text{Cov}(X, X) = \mathbf{D}^2 X$.

Satz 1.5.5. *Existieren die zweiten Momente der Randverteilungen von X und Y , so existiert auch die Kovarianz $\text{Cov}(X, Y)$.*

Definition 1.5.6. Besitzen die beiden Zufallsgrößen X und Y endliche Streuungen, so heißt die Größe

$$\rho_{XY} := \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

Korrelationskoeffizient von X und Y .

Satz 1.5.7. *Für den Korrelationskoeffizienten zweier Zufallsgrößen X und Y gilt*

$$-1 \leq \rho_{XY} = \rho_{YX} \leq 1.$$

Der Korrelationskoeffizient ρ_{XY} zweier Zufallsgrößen X und Y ist also ein normiertes Maß für das stochastische Verhalten beider Zufallsgrößen zueinander. X und Y sind offensichtlich genau dann unkorreliert, wenn $\rho_{XY} = 0$ gilt. $\rho_{XY} \approx -1$ drückt starke negative Korrelation und $\rho_{XY} \approx 1$ starke positive Korrelation aus.

1.5.1.2 Diskrete Verteilungen

Definition 1.5.8. Ein zufälliger Vektor $\underline{X} = (X_1, \dots, X_n)$ heißt *diskret verteilt*, wenn alle eindimensionalen Randverteilungen diskrete Verteilungen sind.

Die Komponenten eines diskret verteilten zufälligen Vektors $\underline{X} = (X_1, \dots, X_n)$ nehmen die Werte $X_1 = x_{1i_1}, \dots, X_n = x_{ni_n}$ an, wobei für endliches X_j $i_j \in \{1, \dots, l_j\} =: I_{X_j}$ mit $l_j \in \mathbb{N}$ und für abzählbar unendliches X_j $i_j \in \mathbb{N} =: I_{X_j}$ gilt. Die Wahrscheinlichkeitsfunktion von \underline{X} lässt sich als n -dimensionales Feld schreiben:

$$p_{i_1 \dots i_n} = P(X_1 = x_{1i_1}, \dots, X_n = x_{ni_n}).$$

Im Fall $n = 2$ bezeichnen wir die Werte der beiden Komponenten X und Y des zufälligen Vektors \underline{X} mit x_j und y_k , wobei $j \in I_X$ und $k \in I_Y$. Die Wahrscheinlichkeitsfunktion von \underline{X} lässt sich dann als endliche oder unendliche Matrix mit Einträgen der Form

$$p_{jk} = P(X = x_j, Y = y_k)$$

schreiben. Für die Randverteilungen von X und Y erhalten wir

$$p_{j\cdot} = \sum_{k \in I_Y} p_{jk} \quad \text{und} \quad p_{\cdot k} = \sum_{j \in I_X} p_{jk}.$$

Für Erwartungswert, Varianz und Kovarianz von X und Y gilt:

- $\mathbf{E}X = \mu_X = \sum_{j \in I_X} x_j p_{j\cdot}, \quad \mathbf{E}Y = \mu_Y = \sum_{k \in I_Y} y_k p_{\cdot k},$
- $\mathbf{D}^2 X = \sigma_X^2 = \sum_{j \in I_X} (x_j - \mu_X)^2 p_{j\cdot}, \quad \mathbf{D}^2 Y = \sigma_Y^2 = \sum_{k \in I_Y} (y_k - \mu_Y)^2 p_{\cdot k},$
- $\text{Cov}(X, Y) = \sigma_{XY} = \sum_{j \in I_X} \sum_{k \in I_Y} (x_j - \mu_X)(y_k - \mu_Y) p_{jk}.$

Beispiel. In einer Urne befinden sich 12 Lose: 2 Geldgewinne, 4 Freilose und 6 Nieten. Es werden ohne Zurücklegen 2 Lose gezogen. Wir bezeichnen mit X die Anzahl der gezogenen Geldgewinne und mit Y die Anzahl der gezogenen Freilose. Gesucht werden die Wahrscheinlichkeitsfunktion $p_{jk} = P(X = x_j, Y = y_k)$ des zufälligen Vektors (X, Y) und die zugehörigen Randverteilungen $p_{j\cdot}$ und $p_{\cdot k}$ seiner Komponenten X und Y . Aus kombinatorischen Überlegungen ergibt sich:

$$p_{jk} = \begin{cases} \frac{\binom{2}{x_j} \binom{4}{y_k} \binom{6}{2-x_j-y_k}}{\binom{12}{2}}, & 0 \leq x_j + y_k \leq 2 \\ 0, & x_j + y_k > 2 \end{cases}.$$

Die Wahrscheinlichkeitsfunktion lässt sich zusammen mit den Randverteilungen wie folgt als Tabelle darstellen:

		x_k			$p_{\cdot k}$
		0	1	2	\downarrow
y_j	0	$\frac{5}{22}$	$\frac{2}{11}$	$\frac{1}{66}$	$\frac{14}{33}$
	1	$\frac{4}{11}$	$\frac{4}{33}$	0	$\frac{16}{33}$
	2	$\frac{1}{11}$	0	0	$\frac{1}{11}$
$p_{j\cdot} \rightarrow$		$\frac{15}{22}$	$\frac{10}{33}$	$\frac{1}{66}$	$\Sigma = 1$

Die Randverteilung $p_{j\cdot}$ bzw. $p_{\cdot k}$ von X bzw. Y erhält man dabei aus den Spalten- bzw. Zeilensummen der 3×3 -Matrix. Da die Randverteilungen eindimensionale Wahrscheinlichkeitsfunktionen sind, sind die Summen der Werte der Randverteilungen stets 1.

1.5.1.3 Stetige Verteilungen

Definition 1.5.9. Ein zufälliger Vektor $\underline{X} = (X_1, \dots, X_n)$ heißt *stetig verteilt* mit der Verteilungsfunktion F , wenn es eine integrierbare Dichtefunktion f gibt mit

$$F(x_1, \dots, x_n) = P(X_1 < x_1, \dots, X_n < x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(t_1, \dots, t_n) dt_1 \cdots dt_n.$$

Definition 1.5.10. Sei g eine reelle Funktion von n reellen Veränderlichen und sei $\underline{X} = (X_1, \dots, X_n)$ eine stetig verteilte n -dimensionale Zufallsgröße. Dann ist der Erwartungswert $\mathbf{E}g(\underline{X})$ gegeben durch

$$\mathbf{E}g(\underline{X}) := \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \dots, x_n) f(x_1, \dots, x_n) dx_1 \cdots dx_n.$$

Definition 1.5.11. Sei (X, Y) ein stetig verteilter zufälliger Vektor und f seine Dichtefunktion. Dann heißen die Funktionen

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad \text{und} \quad f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

Randdichten der Komponenten X und Y .

Die Randdichten sind eindimensionale Dichtefunktionen. Für Erwartungswert, Varianz und Kovarianz von X und Y erhalten wir:

- $\mathbf{E}X = \mu_X = \int_{-\infty}^{\infty} x f_X(x) dx, \quad \mathbf{E}Y = \mu_Y = \int_{-\infty}^{\infty} y f_Y(y) dy,$
- $\mathbf{D}^2 X = \sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx, \quad \mathbf{D}^2 Y = \sigma_Y^2 = \int_{-\infty}^{\infty} (y - \mu_Y)^2 f_Y(y) dy,$
- $\mathbf{Cov}(X, Y) = \sigma_{XY} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f(x, y) dx dy.$

Beispiel. Als Beispiel für eine mehrdimensionale Verteilung betrachten wir die *zweidimensionale Normalverteilung*. Ein zufälliger Vektor (X, Y) ist normalverteilt, wenn er die Dichtefunktion

$$f(x, y) = \frac{\exp\left(\frac{-1}{2(1-\rho^2)} \left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho \left(\frac{x-\mu_X}{\sigma_X}\right) \left(\frac{y-\mu_Y}{\sigma_Y}\right) + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2 \right]\right)}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}$$

mit den Parametern $\sigma_X > 0$, $\sigma_Y > 0$ und $-1 < \rho < 1$ besitzt. Als Randdichten ergeben sich die Funktionen

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left(-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right) \quad \text{und} \quad f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma_Y} \exp\left(-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}\right),$$

d.h. die Randverteilungen sind eindimensionale Normalverteilungen. Der Parameter $\rho = \rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y}$ ist der Korrelationskoeffizient und somit ist $\mathbf{Cov}(X, Y) = \sigma_{XY} = \rho\sigma_X\sigma_Y$. Ist $\rho = 0$, so sind X und Y also unkorreliert und für die Dichtefunktion f gilt die Produktdarstellung

$$f(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y} \exp\left(-\frac{(x-\mu_X)^2}{2\sigma_X^2} - \frac{(y-\mu_Y)^2}{2\sigma_Y^2}\right) = f_X(x)f_Y(y).$$

1.5.1.4 Stochastische Unabhängigkeit

Definition 1.5.12. Zwei Zufallsgrößen X und Y heißen *stochastisch unabhängig*, wenn für den zufälligen Vektor (X, Y) gilt:

$$F(x, y) = F_X(x)F_Y(y) \quad \forall x, y \in \mathbb{R}.$$

Bemerkung. Handelt es sich bei den stochastisch unabhängigen Zufallsgrößen X und Y um diskrete Verteilungen, so gilt

$$p_{jk} = p_{j \cdot} p_{\cdot k} \quad \forall x, y \in \mathbb{R}.$$

Im stetigen Fall gilt analog

$$f(x, y) = f_X(x)f_Y(y) \quad \forall x, y \in \mathbb{R}.$$

Satz 1.5.13. Seien X und Y stochastisch unabhängige Zufallsgrößen mit $\mathbf{D}^2 X < \infty$ und $\mathbf{D}^2 Y < \infty$. Dann gilt $\mathbf{Cov}(X, Y) = 0$, d.h. aus der stochastischen Unabhängigkeit folgt stets die Unkorreliertheit.

Beweis (für stetige Zufallsgrößen). Da X und Y stochastisch unabhängig sind, gilt:

$$\begin{aligned}
\mathbf{Cov}(X, Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f(x, y) \, dx \, dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) f_X(x) f_Y(y) \, dx \, dy \\
&= \int_{-\infty}^{\infty} (x - \mu_X) f_X(x) \, dx \int_{-\infty}^{\infty} (y - \mu_Y) f_Y(y) \, dy \\
&= (\mathbf{E}(X - \mu_X))(\mathbf{E}(Y - \mu_Y)) = (\mu_X - \mu_X)(\mu_Y - \mu_Y) = 0
\end{aligned}$$

□

Wie wir am vorhergehenden Beispiel gesehen haben, folgt im Falle einer zweidimensionalen Normalverteilung aus der Unkorreliertheit der beiden Komponenten deren stochastische Unabhängigkeit. Zusammen mit dem vorhergehenden Satz erhalten wir also die nachfolgende Aussage.

Satz 1.5.14. *Sei (X, Y) ein (zweidimensional) normalverteilter zufälliger Vektor. Dann sind die Komponenten X und Y genau dann stochastisch unabhängig, wenn sie unkorreliert sind.*

1.5.2 Bedingte Verteilungen

In diesem Abschnitt betrachten wir nur den Fall $n = 2$, d.h. zufällige Vektoren der Form $\underline{X} = (X, Y)$. X und Y seien dabei stetig verteilte Zufallsgrößen.

Bei bedingten Verteilungen wird nur eine der beiden Komponenten eines zufälligen Vektors betrachtet, d.h. die andere Komponente bleibt konstant. Wir erhalten also z.B. Aussagen über die Verteilung der Zufallsgröße X , wenn Y einen festen Wert hat.

Definition 1.5.15. Seien X und Y stetig verteilte Zufallsgrößen mit den Randdichten f_X und f_Y , wobei $f_X(x) > 0$ und $f_Y(y) > 0$ für alle $x, y \in \mathbb{R}$. Dann heißen die Funktionen

$$f_{X|Y=y}(x) = \frac{f(x, y)}{f_Y(y)} \quad \text{und} \quad f_{Y|X=x}(y) = \frac{f(x, y)}{f_X(x)}$$

bedingte Dichten.

Bemerkung. $f_{X|Y=y}(x)$ ist eine eindimensionale Dichtefunktion, da $f_{X|Y=y}(x) \geq 0$ und

$$\int_{-\infty}^{\infty} f_{X|Y=y}(x) \, dx = \int_{-\infty}^{\infty} \frac{f(x, y)}{f_Y(y)} \, dx = \frac{1}{f_Y(y)} \int_{-\infty}^{\infty} f(x, y) \, dx = \frac{f_Y(y)}{f_Y(y)} = 1.$$

$f_{X|Y=y}$ ist also die Dichtefunktion der Zufallsgröße X unter der Bedingung $Y = y$. Analoges gilt für $f_{Y|X=x}(y)$

Definition 1.5.16. Die Größe

$$\mathbf{E}(X|Y = y) = \int_{-\infty}^{\infty} x f_{X|Y=y}(x) dx$$

heißt *bedingter Erwartungswert* der Zufallsgröße X unter der Bedingung, dass Y den Wert y annimmt.

Beispiel. Wir betrachten nochmals einen normalverteilten Zufallsvektor (X, Y) . Als bedingte Dichte haben wir

$$f_{X|Y=y}(x) = \frac{f(x, y)}{f_Y(y)} = \frac{1}{\sqrt{2\pi}\sigma_X\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2}\left(\frac{x - \left(\mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y - \mu_Y)\right)}{\sigma_X\sqrt{1-\rho^2}}\right)^2\right).$$

Daraus ergibt sich der bedingte Erwartungswert

$$\mathbf{E}(X|Y = y) = \mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y - \mu_Y).$$

Die Gerade $x = \mathbf{E}(X|Y = y)$ heißt *Regressionsgerade* im (x, y) -Koordinatensystem und gibt für jedes y den Wert (x, y) des Zufallsvektors an, für den X den Erwartungswert unter der Bedingung $Y = y$ annimmt. Ist $\rho = 0$, d.h. X und Y sind unkorreliert und damit stochastisch unabhängig, so ist die Regressionsgerade $x = \mu_x$ parallel zur y -Achse und es gilt $f_{X|Y=y}(x) = f_X(x)$.

1.5.3 Erwartungswertevektor, Kovarianzmatrix, Normalverteilung

Da wir in diesem Abschnitt mit Matrizen rechnen, schreiben wir Vektoren immer als Spaltenvektoren. Wir betrachten den n -dimensionalen Fall, d.h.

$$\underline{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}.$$

Definition 1.5.17. Den Vektor

$$\underline{\mu} = \mathbf{E}\underline{X} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = \begin{pmatrix} \mathbf{E}X_1 \\ \vdots \\ \mathbf{E}X_n \end{pmatrix}$$

der Erwartungswerte der Komponenten X_1, \dots, X_n eines zufälligen Vektors \underline{X} nennen wir *Erwartungswertevektor* von \underline{X} .

Definition 1.5.18. Sei \underline{X} ein zufälliger Vektor mit den Komponenten X_1, \dots, X_n . Dann heißt die Matrix

$$\underline{\Sigma} = \mathbf{Cov}\underline{X} = \left(\mathbf{Cov}(X_i, X_j)\right)_{i,j=1\dots n} = \mathbf{E}(\underline{X} - \mathbf{E}\underline{X})(\underline{X} - \mathbf{E}\underline{X})^T$$

Kovarianzmatrix von \underline{X} .

Satz 1.5.19. Existiert die Kovarianzmatrix $\underline{\Sigma}$ des Zufallsvektors \underline{X} , so ist sie symmetrisch, d.h. $\underline{\Sigma} = \underline{\Sigma}^T$, und positiv semidefinit, d.h. für alle $\underline{v} \in \mathbb{R}^n$ gilt $(\underline{\Sigma} \underline{v}, \underline{v}) = \underline{v}^T \underline{\Sigma} \underline{v} \geq 0$.

Beweis. Die Symmetrie der Kovarianzmatrix folgt direkt aus $\mathbf{Cov}(X_i, X_j) = \mathbf{Cov}(X_j, X_i)$ und es gilt

$$\underline{v}^T \underline{\Sigma} \underline{v} = \underline{v}^T (\mathbf{E}(\underline{X} - \mathbf{E}\underline{X})(\underline{X} - \mathbf{E}\underline{X})^T) \underline{v} = \mathbf{E} (\underline{v}^T (\underline{X} - \mathbf{E}\underline{X}))^2 \geq 0. \quad \square$$

Definition 1.5.20. Ein Zufallsvektor \underline{X} mit den Komponenten X_1, \dots, X_n heißt *nichtsingulär n-dimensional normalverteilt* mit den Parametern $\underline{\mu} = \mathbf{E}\underline{X}$ und $\underline{\Sigma} = \mathbf{Cov}(\underline{X})$, geschrieben $\underline{X} \sim \mathbf{N}(\underline{\mu}, \underline{\Sigma})$, wenn $\underline{\Sigma}$ symmetrisch und positiv definit ist und die Dichtefunktion f die Form

$$f(\underline{x}) = \frac{\sqrt{|\underline{\Sigma}^{-1}|}}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2}(\underline{x} - \underline{\mu})^T \underline{\Sigma}^{-1}(\underline{x} - \underline{\mu})\right)$$

besitzt, wobei $|\underline{\Sigma}^{-1}|$ die Determinante der Inversen der Kovarianzmatrix bezeichnet.

Falls für $\underline{X} \sim \mathbf{N}(\underline{\mu}, \underline{\Sigma})$ die Komponenten $X_i \sim \mathbf{N}(\mu_i, \sigma_i^2)$ paarweise disjunkt sind, d.h.

$$\underline{\Sigma} = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix},$$

so gilt $\sqrt{|\underline{\Sigma}^{-1}|} = \frac{1}{\sigma_1 \cdots \sigma_n}$ und

$$f(\underline{x}) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(x_i - \mu_i)^2}{2\sigma_i^2}\right).$$

Die Dichte f des Zufallsvektors \underline{X} ist also in diesem Fall gleich dem Produkt der Randdichten seiner Komponenten X_i .