

Inhaltsverzeichnis

3.1 Problemstellung . . . . . 71

3.2 Univariate Zufallsvariablen . . . . . 71

3.3 Zufallsmatrizen und Zufallsvektoren . . . . . 76

3.4 Multivariate Normalverteilung . . . . . 88

3.1 Problemstellung

In Kap. 2 haben wir einfache Verfahren zur Darstellung hochdimensionaler Datensätze kennengelernt. Bei diesen Datensätzen handelt es sich in der Regel um Stichproben aus Populationen, wie sie bei Kauermann & Küchenhoff (2011) detailliert beschrieben werden. Populationen werden auch Grundgesamtheiten genannt. Um Schlüsse von einer Stichprobe über die zugrunde liegenden Populationen ziehen zu können, muss man Annahmen über die Merkmale machen. Hierzu benötigen wir das Konzept der *Zufallsvariablen*. Wir werden in diesem Kapitel zunächst *univariate* Zufallsvariablen betrachten. Anschließend werden wir die wesentlichen Eigenschaften von *mehrdimensionalen* Zufallsvariablen herleiten, die wir im weiteren Verlauf des Buches immer wieder benötigen werden.

3.2 Univariate Zufallsvariablen

**Beispiel 18** In Beispiel 2 wurden die 20 Studenten unter anderem danach befragt, ob sie den Leistungskurs Mathematik besucht haben. Kodiert man  $j$  mit 1 und  $n$  mit 0, so erhält man folgende Daten:

0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 . □

In Kap. 2 haben wir gesehen, wie wir diesen Datensatz beschreiben können. Hier wollen wir die Daten als Zufallsstichprobe aus einer Grundgesamtheit auffassen, über die man Aussagen treffen will. In der Regel will man Parameter schätzen oder Hypothesen testen. Um zu sinnvollen Schlussfolgerungen zu gelangen, muss man für das Merkmal eine Wahrscheinlichkeitsverteilung unterstellen. Das Merkmal wird dann zu einer Zufallsvariablen. Man unterscheidet *diskrete* und *stetige* Zufallsvariablen.

**Definition 3.1** Eine Zufallsvariable  $Y$ , die höchstens abzählbar viele Werte annehmen kann, heißt *diskret*. Dabei heißt  $P(Y = y)$  die *Wahrscheinlichkeitsfunktion* von  $Y$ .

Für die Wahrscheinlichkeitsfunktion einer diskreten Zufallsvariablen  $Y$  gilt

$$\sum_y P(Y = y) = 1$$

und

$$P(Y = y) \geq 0$$

für alle  $y \in \mathbb{R}$ .

*Beispiel 18 (Fortsetzung)* Es liegt nahe, die Zufallsvariable  $Y$  zu betrachten, die wir Leistungskurs nennen wollen. Sie kann die Werte 0 und 1 annehmen. Die Wahrscheinlichkeit des Wertes 1 sei  $p$ . Somit ist die Wahrscheinlichkeit des Wertes 0 gleich  $1 - p$ . Wir erhalten somit die Wahrscheinlichkeitsfunktion

$$P(Y = 0) = 1 - p,$$

$$P(Y = 1) = p.$$

□

**Definition 3.2** Die Zufallsvariable  $Y$  heißt *Bernoulli-verteilt* mit dem Parameter  $p$ , wenn gilt

$$P(Y = 0) = 1 - p,$$

$$P(Y = 1) = p.$$

Man kann die Wahrscheinlichkeitsverteilung der *Bernoulli-Verteilung* auch kompakter schreiben. Es gilt

$$P(Y = y) = p^y (1 - p)^{1-y} \quad \text{für } y = 0, 1.$$

Oft ist die Frage von Interesse, ob eine Zufallsvariable  $Y$  Werte annimmt, die kleiner oder gleich einem vorgegebenen Wert  $y$  sind.

**Definition 3.3** Sei  $Y$  eine Zufallsvariable. Dann heißt

$$F_Y(y) = P(Y \leq y)$$

die Verteilungsfunktion von  $Y$ .

Betrachten wir stetige Zufallsvariablen.

**Definition 3.4** Eine Zufallsvariable  $Y$  heißt stetig, wenn eine Funktion  $f_Y : \mathbb{R} \rightarrow \mathbb{R}$  existiert, sodass für die Verteilungsfunktion  $F_Y(y)$  von  $Y$  gilt

$$F_Y(y) = \int_{-\infty}^y f_Y(u) du.$$

Die Funktion  $f_Y(y)$  heißt *Dichtefunktion* der Zufallsvariablen  $Y$ .

Die Dichtefunktion  $f_Y(y)$  erfüllt folgende Bedingungen:

1.

$$f_Y(y) \geq 0 \quad \text{für alle } y \in \mathbb{R},$$

2.

$$\int_{-\infty}^{\infty} f_Y(y) dy = 1.$$

Man kann zeigen, dass jede Funktion, die diese Bedingungen erfüllt, als Dichtefunktion einer stetigen Zufallsvariablen aufgefasst werden kann.

Das wichtigste Verteilungsmodell für eine stetige Zufallsvariable ist die *Normalverteilung*.

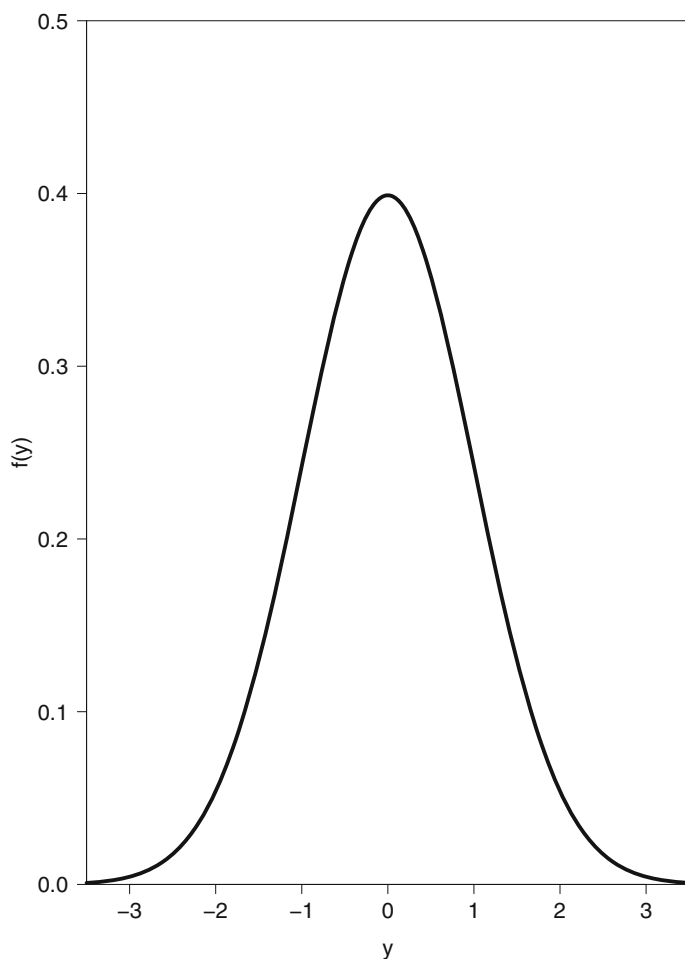
**Definition 3.5** Die Zufallsvariable  $Y$  heißt normalverteilt mit den Parametern  $\mu$  und  $\sigma^2$ , wenn ihre Dichtefunktion gegeben ist durch

$$f_Y(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(y-\mu)^2}{2\sigma^2}\right\} \quad \text{für } y \in \mathbb{R}. \quad (3.1)$$

Wir schreiben  $Y \sim N(\mu, \sigma^2)$ .

Abb. 3.1 zeigt die Dichtefunktion der Normalverteilung mit  $\mu = 0$  und  $\sigma = 1$ , die *Standardnormalverteilung* heißt.

Charakteristika von Verteilungen werden durch *Maßzahlen* beschrieben. Die beiden wichtigsten sind der *Erwartungswert* und die *Varianz*. Der Erwartungswert ist die wichtigste Maßzahl für die Lage einer Verteilung.



**Abb. 3.1** Dichtefunktion der Standardnormalverteilung

**Definition 3.6** Sei  $Y$  eine Zufallsvariable mit Wahrscheinlichkeitsfunktion  $P(Y = y)$  beziehungsweise Dichtefunktion  $f_Y(y)$ . Der Erwartungswert  $E(Y)$  von  $Y$  ist definiert durch

$$E(Y) = \sum_y y P(Y = y),$$

falls  $Y$  diskret ist, und durch

$$E(Y) = \int_{-\infty}^{\infty} y f_Y(y) dy,$$

falls  $Y$  stetig ist.

Schauen wir uns die Bernoulli-Verteilung und die Normalverteilung an. Für eine mit Parameter  $p$  Bernoulli-verteilte Zufallsvariable  $Y$  gilt

$$E(Y) = p . \quad (3.2)$$

Dies sieht man folgendermaßen:

$$E(Y) = 0 \cdot (1 - p) + 1 \cdot p = p .$$

Für eine mit den Parametern  $\mu$  und  $\sigma^2$  normalverteilte Zufallsvariable  $Y$  gilt

$$E(Y) = \mu . \quad (3.3)$$

Der Beweis ist bei Mood et al. (1974) gegeben.

Der Erwartungswert  $E(g(Y))$  einer Funktion  $g(Y)$  der Zufallsvariablen  $Y$  ist definiert durch

$$E(g(Y)) = \sum_y g(y) P(Y = y) ,$$

falls  $Y$  diskret ist, und durch

$$E(g(Y)) = \int_{-\infty}^{\infty} g(y) f_Y(y) dy ,$$

falls  $Y$  stetig ist.

Der Erwartungswert besitzt folgende wichtige Eigenschaft:

$$E(aY + b) = a E(Y) + b . \quad (3.4)$$

Wir zeigen dies für eine diskrete Zufallsvariable  $Y$  mit Wahrscheinlichkeitsfunktion  $P(Y = y)$ :

$$\begin{aligned} E(aY + b) &= \sum_y (ay + b) P(Y = y) = \sum_y (ay P(Y = y) + b P(Y = y)) \\ &= \sum_y ay P(Y = y) + \sum_y b P(Y = y) \\ &= a \sum_y y P(Y = y) + b \sum_y P(Y = y) = a E(Y) + b . \end{aligned}$$

**Definition 3.7** Sei  $Y$  eine Zufallsvariable. Dann ist die Varianz von  $Y$  definiert durch

$$\text{Var}(Y) = E\left((Y - E(Y))^2\right) . \quad (3.5)$$

Für eine mit Parameter  $p$  Bernoulli-verteilte Zufallsvariable  $Y$  gilt

$$\text{Var}(Y) = p \cdot (1 - p) .$$

Dies sieht man folgendermaßen:

$$\begin{aligned} \text{Var}(Y) &= (0 - p)^2 \cdot (1 - p) + (1 - p)^2 \cdot p \\ &= p \cdot (1 - p) \cdot (p + 1 - p) \\ &= p \cdot (1 - p) . \end{aligned}$$

Für eine mit den Parametern  $\mu$  und  $\sigma^2$  normalverteilte Zufallsvariable  $Y$  gilt

$$\text{Var}(Y) = \sigma^2 .$$

Der Beweis ist bei Mood et al. (1974) gegeben. Die Parameter  $\mu$  und  $\sigma^2$  sind also gerade der Erwartungswert und die Varianz einer normalverteilten Zufallsvariablen.

Die Varianz besitzt folgende Eigenschaft:

$$\text{Var}(a Y + b) = a^2 \text{Var}(Y) . \quad (3.6)$$

Dies sieht man mit (3.4) folgendermaßen:

$$\begin{aligned} \text{Var}(a Y + b) &= E \left[ (a Y + b - E(a Y + b))^2 \right] \\ &= E \left[ (a Y + b - a E(Y) - b)^2 \right] = E \left[ (a (Y - E(Y)))^2 \right] \\ &= a^2 E \left[ (Y - E(Y))^2 \right] = a^2 \text{Var}(Y) . \end{aligned}$$

---

### 3.3 Zufallsmatrizen und Zufallsvektoren

Wir wollen nun mehrere Zufallsvariablen gleichzeitig betrachten. Beginnen wir mit einem Theorem, dessen Aussage wir im Folgenden immer wieder benötigen.

**Theorem 3.1** Seien  $Y_1, \dots, Y_p$  univariate Zufallsvariablen. Dann gilt

$$E \left( \sum_{i=1}^p Y_i \right) = \sum_{i=1}^p E(Y_i) . \quad (3.7)$$

Der Beweis des Theorems ist bei Rice (2006) zu finden.

Man kann mehrere Zufallsvariablen zu einer *Zufallsmatrix* oder einem *Zufallsvektor* zusammenfassen.

**Definition 3.8** Seien  $W_{11}, \dots, W_{1n}, \dots, W_{m1}, \dots, W_{mn}$  univariate Zufallsvariablen. Dann heißt

$$\mathbf{W} = \begin{pmatrix} W_{11} & \dots & W_{1n} \\ W_{21} & \dots & W_{2n} \\ \vdots & \ddots & \vdots \\ W_{m1} & \dots & W_{mn} \end{pmatrix}$$

Zufallsmatrix.

Wie auch bei univariaten Zufallsvariablen ist bei Zufallsmatrizen der Erwartungswert von Interesse:

**Definition 3.9** Sei  $\mathbf{W}$  eine Zufallsmatrix. Dann heißt

$$E(\mathbf{W}) = \begin{pmatrix} E(W_{11}) & \dots & E(W_{1n}) \\ E(W_{21}) & \dots & E(W_{2n}) \\ \vdots & \ddots & \vdots \\ E(W_{m1}) & \dots & E(W_{mn}) \end{pmatrix}$$

der Erwartungswert von  $\mathbf{W}$ .

Der Erwartungswert einer Zufallsmatrix besitzt eine wichtige Eigenschaft:

**Theorem 3.2** Seien  $\mathbf{W}$  eine  $(m, n)$ -Zufallsmatrix,  $\mathbf{A}$  eine  $(l, m)$ -Matrix und  $\mathbf{B}$  eine  $(n, p)$ -Matrix, dann gilt

$$E(\mathbf{AWB}) = \mathbf{A}E(\mathbf{W})\mathbf{B}. \quad (3.8)$$

**Beweis:**

Das Element in der  $i$ -ten Zeile und  $j$ -ten Spalte von  $\mathbf{AWB}$  erhält man, indem man das innere Produkt der  $i$ -ten Zeile von  $\mathbf{AW}$  und der  $j$ -ten Spalte der Matrix  $\mathbf{B}$  bildet. Die  $i$ -te Zeile der Matrix  $\mathbf{AW}$  ist

$$\left( \sum_{r=1}^m a_{ir} W_{r1}, \dots, \sum_{r=1}^m a_{ir} W_{rn} \right)$$

und die  $j$ -te Spalte von  $\mathbf{B}$  ist

$$\begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix}.$$

Bildet man das innere Produkt dieser beiden Vektoren, so erhält man

$$\begin{aligned} b_{1j} \sum_{r=1}^m a_{ir} W_{r1} + \dots + b_{nj} \sum_{r=1}^m a_{ir} W_{rn} &= \sum_{r=1}^m a_{ir} W_{r1} b_{1j} + \dots + \sum_{r=1}^m a_{ir} W_{rn} b_{nj} \\ &= \sum_{s=1}^n \sum_{r=1}^m a_{ir} W_{rs} b_{sj}. \end{aligned}$$

Nun gilt

$$E \left( \sum_{s=1}^n \sum_{r=1}^m a_{ir} W_{rs} b_{sj} \right) = \sum_{s=1}^n \sum_{r=1}^m a_{ir} E(W_{rs}) b_{sj}.$$

Dies ist aber gerade das Element in der  $i$ -ten Zeile und  $j$ -ten Spalte der Matrix  $\mathbf{A}E(\mathbf{W})\mathbf{B}$ .

Diese Eigenschaft von Zufallsmatrizen werden wir gleich verwenden. Wir werden uns im Folgenden aber nicht mit Zufallsmatrizen, sondern mit Zufallsvektoren beschäftigen, die wir auch als *mehrdimensionale Zufallsvariablen* bezeichnen.

**Definition 3.10** Seien  $Y_1, \dots, Y_p$  univariate Zufallsvariablen. Dann heißt

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_p \end{pmatrix}$$

$p$ -dimensionale Zufallsvariable.

In Analogie zum Erwartungswert einer Zufallsmatrix definieren wir den Erwartungswert einer  $p$ -dimensionalen Zufallsvariablen.

**Definition 3.11** Sei  $\mathbf{Y}$  eine  $p$ -dimensionale Zufallsvariable. Dann heißt

$$E(\mathbf{Y}) = \begin{pmatrix} E(Y_1) \\ \vdots \\ E(Y_p) \end{pmatrix}$$

Erwartungswert von  $\mathbf{Y}$ .

Wir haben bei univariaten Zufallsvariablen  $Y$  gesehen, dass der Erwartungswert einer Lineartransformation von  $Y$  gleich der Lineartransformation des Erwartungswertes ist. Eine entsprechende Eigenschaft gilt für mehrdimensionale Zufallsvariablen.



**Theorem 3.3** Sei  $\mathbf{Y}$  eine  $p$ -dimensionale Zufallsvariable,  $\mathbf{A}$  eine  $(m, p)$ -Matrix,  $\mathbf{b}$  ein  $m$ -dimensionaler Vektor. Dann gilt

$$E(\mathbf{A}\mathbf{Y} + \mathbf{b}) = \mathbf{A}E(\mathbf{Y}) + \mathbf{b}. \quad (3.9)$$

**Beweis:**

Die  $i$ -te Komponente des Vektors  $\mathbf{A}\mathbf{Y} + \mathbf{b}$  ist

$$\sum_{k=1}^p a_{ik} Y_k + b_i.$$

Nun gilt

$$E\left(\sum_{k=1}^p a_{ik} Y_k + b_i\right) = \sum_{k=1}^p a_{ik} E(Y_k) + b_i.$$

Dies ist aber gerade die  $i$ -te Komponente von  $\mathbf{A}E(\mathbf{Y}) + \mathbf{b}$ .

Wie das folgende Theorem zeigt, gilt die Aussage von Theorem 3.1 auch für  $p$ -dimensionale Zufallsvariablen.

**Theorem 3.4** Seien  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$   $p$ -dimensionale Zufallsvariablen mit

$$\mathbf{Y}_i = \begin{pmatrix} Y_{i1} \\ \vdots \\ Y_{ip} \end{pmatrix}.$$

Dann gilt

$$E\left(\sum_{i=1}^n \mathbf{Y}_i\right) = \sum_{i=1}^n E(\mathbf{Y}_i). \quad (3.10)$$

**Beweis:**

$$\begin{aligned} E\left(\sum_{i=1}^n \mathbf{Y}_i\right) &= E\left[\begin{pmatrix} \sum_{i=1}^n Y_{i1} \\ \vdots \\ \sum_{i=1}^n Y_{ip} \end{pmatrix}\right] = \begin{pmatrix} E(\sum_{i=1}^n Y_{i1}) \\ \vdots \\ E(\sum_{i=1}^n Y_{ip}) \end{pmatrix} \\ &= \begin{pmatrix} \sum_{i=1}^n E(Y_{i1}) \\ \vdots \\ \sum_{i=1}^n E(Y_{ip}) \end{pmatrix} = \sum_{i=1}^n \begin{pmatrix} E(Y_{i1}) \\ \vdots \\ E(Y_{ip}) \end{pmatrix} = \sum_{i=1}^n E(\mathbf{Y}_i). \end{aligned}$$

In Kap. 2 haben wir die empirische Kovarianz als Maß für den linearen Zusammenhang zwischen zwei Merkmalen kennengelernt. Sie ist definiert durch

$$s_{ij} = \frac{1}{n-1} \sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j).$$

Diesen Ausdruck können wir direkt auf zwei Zufallsvariablen übertragen.

**Definition 3.12** Seien  $Y$  und  $Z$  univariate Zufallsvariablen. Dann ist die Kovarianz  $\text{Cov}(Y, Z)$  zwischen  $Y$  und  $Z$  definiert durch

$$\text{Cov}(Y, Z) = E[(Y - E(Y))(Z - E(Z))]. \quad (3.11)$$

Offensichtlich gilt

$$\text{Cov}(Y, Z) = \text{Cov}(Z, Y). \quad (3.12)$$

Setzen wir in (3.11)  $Z$  gleich  $Y$ , so ergibt sich die Varianz von  $Y$ :

$$\text{Var}(Y) = \text{Cov}(Y, Y). \quad (3.13)$$

Das folgende Theorem gibt wichtige Eigenschaften der Kovarianz an.

**Theorem 3.5** Seien  $U, V, Y$  und  $Z$  univariate Zufallsvariablen und  $a, b, c$  und  $d$  reelle Zahlen. Dann gilt

$$\begin{aligned} \text{Cov}(U + V, Y + Z) &= \text{Cov}(U, Y) + \text{Cov}(U, Z) \\ &\quad + \text{Cov}(V, Y) + \text{Cov}(V, Z) \end{aligned} \quad (3.14)$$

und

$$\text{Cov}(aV + b, cY + d) = ac \text{Cov}(V, Y). \quad (3.15)$$

**Beweis:**

Wir beweisen zunächst (3.14):

$$\begin{aligned} \text{Cov}(U + V, Y + Z) &= E[(U + V - E(U + V))(Y + Z - E(Y + Z))] \\ &= E[(U + V - E(U) - E(V))(Y + Z - E(Y) - E(Z))] \\ &= E[(U - E(U) + V - E(V))(Y - E(Y) + Z - E(Z))] \\ &= E[(U - E(U))(Y - E(Y)) + (U - E(U))(Z - E(Z)) \\ &\quad + (V - E(V))(Y - E(Y)) + (V - E(V))(Z - E(Z))] \end{aligned}$$

$$\begin{aligned}
&= E[(U - E(U))(Y - E(Y))] + E[(U - E(U))(Z - E(Z))] \\
&\quad + E[(V - E(V))(Y - E(Y))] + E[(V - E(V))(Z - E(Z))] \\
&= \text{Cov}(U, Y) + \text{Cov}(U, Z) + \text{Cov}(V, Y) + \text{Cov}(V, Z) .
\end{aligned}$$

(3.15) gilt wegen

$$\begin{aligned}
\text{Cov}(aV + b, cY + d) &= E[(aV + b - E(aV + b))(cY + d - E(cY + d))] \\
&= E[(aV + b - aE(V) - b)(cY + d - cE(Y) - d)] \\
&= ac E[(V - E(V))(Y - E(Y))] \\
&= ac \text{Cov}(V, Y) .
\end{aligned}$$

Setzen wir in (3.14)  $U$  gleich  $Y$  und  $V$  gleich  $Z$ , so gilt

$$\begin{aligned}
\text{Cov}(Y + Z, Y + Z) &= \text{Cov}(Y, Y) + \text{Cov}(Y, Z) + \text{Cov}(Z, Y) + \text{Cov}(Z, Z) \\
&= \text{Var}(Y) + 2\text{Cov}(Y, Z) + \text{Var}(Z) .
\end{aligned}$$

Es gilt also

$$\text{Var}(Y + Z) = \text{Var}(Y) + \text{Var}(Z) + 2\text{Cov}(Y, Z) . \quad (3.16)$$

Entsprechend kann man zeigen:

$$\text{Var}(Y - Z) = \text{Var}(Y) + \text{Var}(Z) - 2\text{Cov}(Y, Z) . \quad (3.17)$$

(3.15) zeigt, dass die Kovarianz nicht skaleninvariant ist. Das haben wir auch schon in Kap. 2 gesehen. Misst man die Körpergröße in Zentimetern und das Körpergewicht in Gramm und bestimmt die Kovarianz zwischen diesen beiden Zufallsvariablen, so ist die Kovarianz 100 000-mal so groß, als wenn man die Körpergröße in Metern und das Körpergewicht in Kilogramm bestimmt. Eine skaleninvariante Maßzahl für den Zusammenhang zwischen zwei Zufallsvariablen erhält man, indem man die beiden Zufallsvariablen standardisiert. Wir bilden

$$Y^* = \frac{Y - E(Y)}{\sqrt{\text{Var}(Y)}} \quad (3.18)$$

und

$$Z^* = \frac{Z - E(Z)}{\sqrt{\text{Var}(Z)}} . \quad (3.19)$$

Wie man mit (3.4) und (3.6) leicht zeigen kann, gilt

$$E(Y^*) = E(Z^*) = 0 \quad (3.20)$$

und

$$\text{Var}(Y^*) = \text{Var}(Z^*) = 1. \quad (3.21)$$

Für die Kovarianz zwischen  $Y^*$  und  $Z^*$  gilt

$$\text{Cov}(Y^*, Z^*) = \frac{\text{Cov}(Y, Z)}{\sqrt{\text{Var}(Y)} \sqrt{\text{Var}(Z)}}. \quad (3.22)$$

Dies sieht man mit (3.15) folgendermaßen:

$$\begin{aligned} \text{Cov}(Y^*, Z^*) &= \text{Cov}\left(\frac{Y - E(Y)}{\sqrt{\text{Var}(Y)}}, \frac{Z - E(Z)}{\sqrt{\text{Var}(Z)}}\right) \\ &= \frac{1}{\sqrt{\text{Var}(Y)} \sqrt{\text{Var}(Z)}} \text{Cov}(Y - E(Y), Z - E(Z)) \\ &= \frac{\text{Cov}(Y, Z)}{\sqrt{\text{Var}(Y)} \sqrt{\text{Var}(Z)}}. \end{aligned}$$

**Definition 3.13** Seien  $Y$  und  $Z$  Zufallsvariablen. Der Korrelationskoeffizient  $\rho_{Y,Z}$  zwischen  $Y$  und  $Z$  ist definiert durch

$$\rho_{Y,Z} = \frac{\text{Cov}(Y, Z)}{\sqrt{\text{Var}(Y)} \sqrt{\text{Var}(Z)}}. \quad (3.23)$$

Der Korrelationskoeffizient ist skaleninvariant. Sind  $a$  und  $c$  positive reelle Zahlen, so gilt

$$\rho_{aY,cZ} = \rho_{Y,Z}.$$

Mit (3.6) und (3.15) sieht man dies folgendermaßen:

$$\begin{aligned} \rho_{aY,cZ} &= \frac{\text{Cov}(aY, cZ)}{\sqrt{\text{Var}(aY)} \sqrt{\text{Var}(cZ)}} = \frac{a c \text{Cov}(Y, Z)}{\sqrt{a^2 \text{Var}(Y)} \sqrt{c^2 \text{Var}(Z)}} \\ &= \frac{a c \text{Cov}(Y, Z)}{a c \sqrt{\text{Var}(Y)} \sqrt{\text{Var}(Z)}} = \frac{\text{Cov}(Y, Z)}{\sqrt{\text{Var}(Y)} \sqrt{\text{Var}(Z)}} = \rho_{Y,Z}. \end{aligned}$$

Das folgende Theorem gibt eine wichtige Eigenschaft des Korrelationskoeffizienten an.

**Theorem 3.6** Für den Korrelationskoeffizienten  $\rho_{Y,Z}$  zwischen den Zufallsvariablen  $Y$  und  $Z$  gilt

$$-1 \leq \rho_{Y,Z} \leq 1 . \quad (3.24)$$

Dabei ist  $|\rho_{Y,Z}| = 1$  genau dann, wenn Konstanten  $a$  und  $b \neq 0$  existieren, sodass gilt

$$P(Z = a \pm b Y) = 1 .$$

**Beweis:**

Seien  $Y^*$  und  $Z^*$  die standardisierten Variablen. Dann gilt wegen (3.16)

$$\text{Var}(Y^* + Z^*) = \text{Var}(Y^*) + \text{Var}(Z^*) + 2 \text{Cov}(Y^*, Z^*) = 2 + 2 \rho_{Y,Z} .$$

Da die Varianz nichtnegativ ist, gilt

$$2 + 2 \rho_{Y,Z} \geq 0$$

und somit

$$\rho_{Y,Z} \geq -1 .$$

Außerdem gilt wegen (3.17)

$$\text{Var}(Y^* - Z^*) = \text{Var}(Y^*) + \text{Var}(Z^*) - 2 \text{Cov}(Y^*, Z^*) = 2 - 2 \rho_{Y,Z} .$$

Hieraus folgt

$$\rho_{Y,Z} \leq 1 .$$

Also gilt

$$-1 \leq \rho_{Y,Z} \leq 1 .$$

Ist

$$\rho_{Y,Z} = 1 ,$$

so gilt

$$\text{Var}(Y^* - Z^*) = 0 .$$

Somit gilt

$$P(Y^* - Z^* = 0) = 1 ,$$

siehe dazu [Rice \(2006\)](#).

Also gilt

$$P(Y = a + b Z) = 1$$

mit

$$a = E(Y) - \frac{\sqrt{\text{Var}(Y)}}{\sqrt{\text{Var}(Z)}} E(Z)$$

und

$$b = \frac{\sqrt{\text{Var}(Y)}}{\sqrt{\text{Var}(Z)}} .$$

Eine analoge Beziehung erhält man für  $\rho_{Y,Z} = -1$ .

Das Konzept der Kovarianz kann auch auf mehrdimensionale Zufallsvariablen übertragen werden.

**Definition 3.14** Sei  $\mathbf{Y}$  eine  $p$ -dimensionale Zufallsvariable und  $\mathbf{Z}$  eine  $q$ -dimensionale Zufallsvariable, dann heißt

$$\text{Cov}(\mathbf{Y}, \mathbf{Z}) = \begin{pmatrix} \text{Cov}(Y_1, Z_1) & \dots & \text{Cov}(Y_1, Z_q) \\ \vdots & \ddots & \vdots \\ \text{Cov}(Y_p, Z_1) & \dots & \text{Cov}(Y_p, Z_q) \end{pmatrix} \quad (3.25)$$

die Kovarianzmatrix von  $\mathbf{Y}$  und  $\mathbf{Z}$ .

Man kann die Kovarianzmatrix von  $\mathbf{Y}$  und  $\mathbf{Z}$  auch mithilfe des äußeren Produkts darstellen:

$$\text{Cov}(\mathbf{Y}, \mathbf{Z}) = E \left[ (\mathbf{Y} - E(\mathbf{Y})) (\mathbf{Z} - E(\mathbf{Z}))' \right] . \quad (3.26)$$

Dies sieht man folgendermaßen:

$$\begin{aligned}
 \text{Cov}(\mathbf{Y}, \mathbf{Z}) &= \begin{pmatrix} \text{Cov}(Y_1, Z_1) & \dots & \text{Cov}(Y_1, Z_q) \\ \vdots & \ddots & \vdots \\ \text{Cov}(Y_p, Z_1) & \dots & \text{Cov}(Y_p, Z_q) \end{pmatrix} \\
 &= \begin{pmatrix} E[(Y_1 - E(Y_1))(Z_1 - E(Z_1))] & \dots & E[(Y_1 - E(Y_1))(Z_q - E(Z_q))] \\ \vdots & \ddots & \vdots \\ E[(Y_p - E(Y_p))(Z_1 - E(Z_1))] & \dots & E[(Y_p - E(Y_p))(Z_q - E(Z_q))] \end{pmatrix} \\
 &= E \begin{pmatrix} (Y_1 - E(Y_1))(Z_1 - E(Z_1)) & \dots & (Y_1 - E(Y_1))(Z_q - E(Z_q)) \\ \vdots & \ddots & \vdots \\ (Y_p - E(Y_p))(Z_1 - E(Z_1)) & \dots & (Y_p - E(Y_p))(Z_q - E(Z_q)) \end{pmatrix} \\
 &= E \left[ \begin{pmatrix} Y_1 - E(Y_1) \\ \vdots \\ Y_p - E(Y_p) \end{pmatrix} (Z_1 - E(Z_1)) \dots (Z_q - E(Z_q)) \right] \\
 &= E [(\mathbf{Y} - E(\mathbf{Y}))(\mathbf{Z} - E(\mathbf{Z}))'] .
 \end{aligned}$$

In Theorem 3.5 wurden Eigenschaften von  $\text{Cov}(Y, Z)$  bewiesen.  $\text{Cov}(\mathbf{Y}, \mathbf{Z})$  besitzt analoge Eigenschaften.

**Theorem 3.7** Seien  $\mathbf{U}$  und  $\mathbf{V}$   $p$ -dimensionale Zufallsvariablen,  $\mathbf{Y}$  und  $\mathbf{Z}$   $q$ -dimensionale Zufallsvariablen,  $\mathbf{A}$  eine  $(m, p)$ -Matrix,  $\mathbf{C}$  eine  $(n, q)$ -Matrix,  $\mathbf{b}$  ein  $m$ -dimensionaler Vektor und  $\mathbf{d}$  ein  $n$ -dimensionaler Vektor. Dann gilt

$$\text{Cov}(\mathbf{U} + \mathbf{V}, \mathbf{Y} + \mathbf{Z}) = \text{Cov}(\mathbf{U}, \mathbf{Y}) + \text{Cov}(\mathbf{V}, \mathbf{Y}) + \text{Cov}(\mathbf{U}, \mathbf{Z}) + \text{Cov}(\mathbf{V}, \mathbf{Z}) \quad (3.27)$$

und

$$\text{Cov}(\mathbf{A}\mathbf{V} + \mathbf{b}, \mathbf{C}\mathbf{Y} + \mathbf{d}) = \mathbf{A} \text{Cov}(\mathbf{V}, \mathbf{Y}) \mathbf{C}' . \quad (3.28)$$

**Beweis:**

Wir zeigen zunächst (3.27):

$$\begin{aligned}
 \text{Cov}(\mathbf{U} + \mathbf{V}, \mathbf{Y} + \mathbf{Z}) &= E[(\mathbf{U} + \mathbf{V} - E(\mathbf{U} + \mathbf{V}))(\mathbf{Y} + \mathbf{Z} - E(\mathbf{Y} + \mathbf{Z}))'] \\
 &= E[(\mathbf{U} + \mathbf{V} - E(\mathbf{U}) - E(\mathbf{V}))(\mathbf{Y} + \mathbf{Z} - E(\mathbf{Y}) - E(\mathbf{Z}))'] \\
 &= E[(\mathbf{U} - E(\mathbf{U}) + \mathbf{V} - E(\mathbf{V}))(\mathbf{Y} - E(\mathbf{Y}) + \mathbf{Z} - E(\mathbf{Z}))'] \\
 &= E[(\mathbf{U} - E(\mathbf{U}))(\mathbf{Y} - E(\mathbf{Y}))' + (\mathbf{U} - E(\mathbf{U}))(\mathbf{Z} - E(\mathbf{Z}))' \\
 &\quad + (\mathbf{V} - E(\mathbf{V}))(\mathbf{Y} - E(\mathbf{Y}))' + (\mathbf{V} - E(\mathbf{V}))(\mathbf{Z} - E(\mathbf{Z}))']
 \end{aligned}$$

$$\begin{aligned}
&= E[(\mathbf{U} - E(\mathbf{U}))(\mathbf{Y} - E(\mathbf{Y}))'] + E[(\mathbf{U} - E(\mathbf{U}))(\mathbf{Z} - E(\mathbf{Z}))'] \\
&\quad + E[(\mathbf{V} - E(\mathbf{V}))(\mathbf{Y} - E(\mathbf{Y}))'] + E[(\mathbf{V} - E(\mathbf{V}))(\mathbf{Z} - E(\mathbf{Z}))'] \\
&= \text{Cov}(\mathbf{U}, \mathbf{Y}) + \text{Cov}(\mathbf{U}, \mathbf{Z}) + \text{Cov}(\mathbf{V}, \mathbf{Y}) + \text{Cov}(\mathbf{V}, \mathbf{Z}).
\end{aligned}$$

(3.28) ist erfüllt wegen

$$\begin{aligned}
\text{Cov}(\mathbf{A}\mathbf{V} + \mathbf{b}, \mathbf{C}\mathbf{Y} + \mathbf{d}) &= E[(\mathbf{A}\mathbf{V} + \mathbf{b} - E(\mathbf{A}\mathbf{V} + \mathbf{b}))(\mathbf{C}\mathbf{Y} + \mathbf{d} - E(\mathbf{C}\mathbf{Y} + \mathbf{d}))'] \\
&= E[(\mathbf{A}\mathbf{V} - \mathbf{A}E(\mathbf{V}))(\mathbf{C}\mathbf{Y} - \mathbf{C}E(\mathbf{Y}))'] \\
&= E[(\mathbf{A}(\mathbf{V} - E(\mathbf{V})))(\mathbf{C}(\mathbf{Y} - E(\mathbf{Y})))'] \\
&= E[\mathbf{A}(\mathbf{V} - E(\mathbf{V}))(\mathbf{Y} - E(\mathbf{Y}))'\mathbf{C}'] \tag{3.29}
\end{aligned}$$

$$\begin{aligned}
&= \mathbf{A}E[(\mathbf{V} - E(\mathbf{V}))(\mathbf{Y} - E(\mathbf{Y}))'\mathbf{C}'] \tag{3.30} \\
&= \mathbf{A} \text{Cov}(\mathbf{V}, \mathbf{Y}) \mathbf{C}'.
\end{aligned}$$

Der Übergang von (3.29) zu (3.30) gilt aufgrund von (3.8).

**Definition 3.15** Sei  $\mathbf{Y}$  eine  $p$ -dimensionale Zufallsvariable. Dann nennt man  $\text{Cov}(\mathbf{Y}, \mathbf{Y})$  die Varianz-Kovarianz-Matrix  $\text{Var}(\mathbf{Y})$  von  $\mathbf{Y}$ .

Für die Varianz-Kovarianz-Matrix schreiben wir auch  $\mathbf{\Sigma}$ . Sie sieht folgendermaßen aus:

$$\mathbf{\Sigma} = \begin{pmatrix} \text{Var}(Y_1) & \text{Cov}(Y_1, Y_2) & \dots & \text{Cov}(Y_1, Y_p) \\ \text{Cov}(Y_2, Y_1) & \text{Var}(Y_2) & \dots & \text{Cov}(Y_2, Y_p) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(Y_p, Y_1) & \text{Cov}(Y_p, Y_2) & \dots & \text{Var}(Y_p) \end{pmatrix}.$$

Wegen  $\text{Cov}(Y_i, Y_j) = \text{Cov}(Y_j, Y_i)$  ist die Varianz-Kovarianz-Matrix symmetrisch. Diese Eigenschaft wird im Folgenden von zentraler Bedeutung sein.

**Theorem 3.8** Sei  $\mathbf{Y}$  eine  $p$ -dimensionale Zufallsvariable,  $\mathbf{A}$  eine  $(m, p)$ -Matrix und  $\mathbf{b}$  ein  $m$ -dimensionaler Vektor. Dann gilt

$$\text{Var}(\mathbf{A}\mathbf{Y} + \mathbf{b}) = \mathbf{A} \text{Var}(\mathbf{Y}) \mathbf{A}'. \tag{3.31}$$

**Beweis:**

Wegen (3.28) gilt:

$$\begin{aligned}
\text{Var}(\mathbf{A}\mathbf{Y} + \mathbf{b}) &= \text{Cov}(\mathbf{A}\mathbf{Y} + \mathbf{b}, \mathbf{A}\mathbf{Y} + \mathbf{b}) \\
&= \mathbf{A} \text{Cov}(\mathbf{Y}, \mathbf{Y}) \mathbf{A}' = \mathbf{A} \text{Var}(\mathbf{Y}) \mathbf{A}'.
\end{aligned}$$



Ist in (3.31)  $\mathbf{A}$  eine  $(1, p)$ -Matrix, also ein  $p$ -dimensionaler Zeilenvektor  $\mathbf{a}'$ , und  $\mathbf{b} = \mathbf{0}$ , so gilt

$$\text{Var}(\mathbf{a}'\mathbf{Y}) = \mathbf{a}'\text{Var}(\mathbf{Y})\mathbf{a} = \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}. \quad (3.32)$$

Da die Varianz nichtnegativ ist, gilt für jeden  $p$ -dimensionalen Vektor  $\mathbf{a}$

$$\mathbf{a}'\text{Var}(\mathbf{Y})\mathbf{a} \geq 0.$$

Also ist eine Varianz-Kovarianz-Matrix immer nichtnegativ definit.

Setzen wir in (3.32) für  $\mathbf{a}$  den Einservektor  $\mathbf{1}$  ein, erhalten wir

$$\text{Var}\left(\sum_{i=1}^p Y_i\right) = \text{Var}(\mathbf{1}'\mathbf{Y}) = \mathbf{1}'\text{Var}(\mathbf{Y})\mathbf{1} = \sum_{i=1}^p \text{Var}(Y_i) + \sum_{i \neq j} \text{Cov}(Y_i, Y_j).$$

Sind die Zufallsvariablen  $Y_1, \dots, Y_p$  also unkorreliert, so gilt

$$\text{Var}\left(\sum_{i=1}^p Y_i\right) = \sum_{i=1}^p \text{Var}(Y_i).$$

Die Varianz-Kovarianz-Matrix der standardisierten Zufallsvariablen nennt man auch *Korrelationsmatrix*  $\mathbf{P}$ :

$$\mathbf{P} = \begin{pmatrix} 1 & \rho_{12} & \dots & \rho_{1p} \\ \rho_{21} & 1 & \dots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \dots & 1 \end{pmatrix}. \quad (3.33)$$

Dabei gilt

$$\rho_{ij} = \frac{\text{Cov}(Y_i, Y_j)}{\sqrt{\text{Var}(Y_i)} \sqrt{\text{Var}(Y_j)}}.$$

Wir bezeichnen  $E(Y_i)$  mit  $\mu_i$  und  $\sqrt{\text{Var}(Y_i)}$  mit  $\sigma_i$ . Sei  $\tilde{\mathbf{Y}}$  die zentrierte  $p$ -dimensionale Zufallsvariable  $\mathbf{Y}$ . Es gilt also

$$\tilde{\mathbf{Y}} = \begin{pmatrix} Y_1 - \mu_1 \\ \vdots \\ Y_p - \mu_p \end{pmatrix}. \quad (3.34)$$

Die standardisierte  $p$ -dimensionale Zufallsvariable  $\mathbf{Y}$  bezeichnen wir mit  $\mathbf{Y}^*$ . Es gilt

$$\mathbf{Y}^* = \begin{pmatrix} \frac{Y_1 - \mu_1}{\sigma_1} \\ \vdots \\ \frac{Y_p - \mu_p}{\sigma_p} \end{pmatrix}. \quad (3.35)$$

Bilden wir die Diagonalmatrix

$$\mathbf{D} = \begin{pmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_p \end{pmatrix},$$

so gilt

$$\mathbf{Y}^* = \mathbf{D}^{-1} \tilde{\mathbf{Y}}.$$

Da Korrelationen gerade die Kovarianzen zwischen den standardisierten Zufallsvariablen sind, folgt

$$\mathbf{P} = \text{Var}(\mathbf{D}^{-1} \tilde{\mathbf{Y}}). \quad (3.36)$$

Diese Beziehung werden wir in Kap. 9 benötigen.

---

### 3.4 Multivariate Normalverteilung

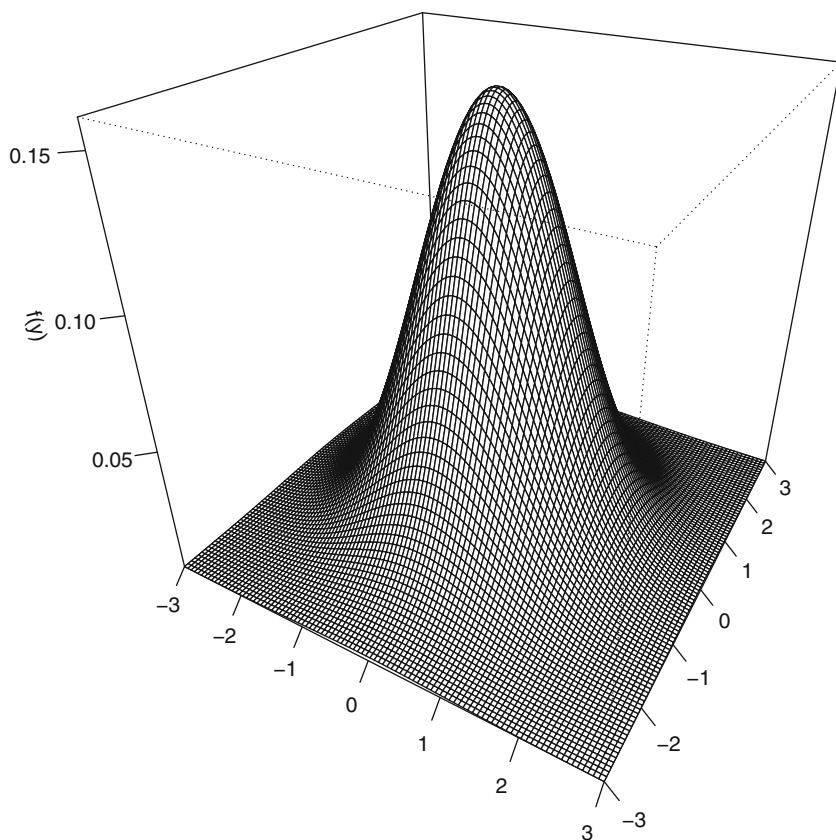
Die wichtigste stetige Verteilung ist die Normalverteilung. Die Dichtefunktion einer mit den Parametern  $\mu$  und  $\sigma^2$  normalverteilten Zufallsvariablen ist gegeben durch

$$f_Y(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(y - \mu)^2}{2\sigma^2}\right\} \quad \text{für } y \in \mathbb{R}.$$

Wir können diesen Ausdruck direkt übertragen auf eine  $p$ -dimensionale Zufallsvariable  $\mathbf{Y}$ .

**Definition 3.16** Die  $p$ -dimensionale Zufallsvariable  $\mathbf{Y}$  heißt  $p$ -variater normalverteilt mit den Parametern  $\mu$  und  $\Sigma$ , falls die Dichtefunktion von  $\mathbf{Y}$  gegeben ist durch

$$f_{\mathbf{Y}}(\mathbf{y}) = (2\pi)^{-p/2} |\Sigma|^{-0.5} \exp\{-0.5 (\mathbf{y} - \mu)' \Sigma^{-1} (\mathbf{y} - \mu)\}. \quad (3.37)$$



**Abb. 3.2** Dichtefunktion der bivariaten Standardnormalverteilung

Abb. 3.2 zeigt die Dichte einer bivariaten Standardnormalverteilung. Es gilt also

$$\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

und

$$\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Das folgende Theorem gibt den Erwartungswert und die Varianz-Kovarianz-Matrix einer  $p$ -dimensionalen Normalverteilung an.

**Theorem 3.9** *Sei  $\mathbf{Y}$  eine mit den Parametern  $\boldsymbol{\mu}$  und  $\boldsymbol{\Sigma}$   $p$ -variat normalverteilte Zufallsvariable. Dann gilt*

$$E(\mathbf{Y}) = \boldsymbol{\mu}$$

*und*

$$\text{Var}(\mathbf{Y}) = \boldsymbol{\Sigma}.$$

Ein Beweis dieses Theorems ist bei Seber (1977) gegen. Das folgende Theorem benötigen wir in Kap. 9.

**Theorem 3.10** *Sei  $\mathbf{Y}$  eine mit den Parametern  $\boldsymbol{\mu}$  und  $\boldsymbol{\Sigma}$   $p$ -variat normalverteilte Zufallsvariable,  $\mathbf{A}$  eine  $(m, p)$ -Matrix vom Rang  $m$  und  $\mathbf{b}$  ein  $m$ -dimensionaler Vektor. Dann ist  $\mathbf{A}\mathbf{Y} + \mathbf{b}$   $m$ -variat normalverteilt mit den Parametern  $\mathbf{A}\boldsymbol{\mu} + \mathbf{b}$  und  $\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}'$ .*

Ein Beweis dieses Theorems ist ebenfalls bei Seber (1977) enthalten.

Multivariate Analysemethoden

Theorie und Praxis mit R

Handl, A.; Kuhlenkasper, T.

2017, XVIII, 536 S. 99 Abb., Softcover

ISBN: 978-3-662-54753-3