Kapitel 5

Mehrdimensionale Zufallsgrößen

Eine zweidimensionale Zufallsgröße besteht aus einem Paar (X, Y) von Zufallsgrößen. Ziel dieses Abschnitts ist es zu zeigen, wie die *gemeinsame Verteilung* der Größen X und Y modelliert werden kann. Allgemeiner können auch mehrdimensionale Größen behandelt werden, die durch einen Vektor $(X_1, \ldots, X_n)$ dargestellt werden.

5.1 Der diskrete Fall:

Die beiden diskreten Zufalls größen mögen die folgenden Werte annehmen:

Werte für X: $x_1, x_2, \ldots, x_k$;

Werte für Y: $y_1, y_2, \ldots, y_l$.

Analog zu den gemeinsamen relativen Häufigkeiten der beschreibenden Statistik definieren wir die gemeinsamen Wahrscheinlichkeiten

$$p_{ij} = P(X = x_i, Y = y_j).$$

Die **Randwahrscheinlichkeiten** sind gegeben durch

$$\begin{aligned} p_{i\cdot} &= \sum_{j=1}^{l} p_{ij} &= \sum_{j=1}^{l} P(X = x_i, Y = y_j) = P(X = x_i) \\ p_{\cdot j} &= \sum_{i=1}^{k} p_{ij} &= \sum_{i=1}^{k} P(X = x_i, Y = y_j) = P(Y = y_j) \end{aligned}$$

Die gesamte Information kann in einer Wahrscheinlichkeitstabelle zusammengefasst werden:

$X \setminus Y$		y_j			
		$\vdots$			
x_i	$\dots$	p_{ij}	$\dots$	$p_{i\cdot}$	$\left(\sum_j\right)$
		$\vdots$			
		$p_{\cdot j}$		1	
		$\left(\sum_i\right)$			

Wir können zunächst wieder die **Erwartungswerte** und **Varianzen** der beiden Zufallsgrößen einzeln berechnen:

$$\begin{aligned}
\mu_X &= \mathrm{E}(X) = \sum_{i=1}^{k}\sum_{j=1}^{l} x_i p_{ij} \\
\mu_Y &= \mathrm{E}(Y) = \sum_{i=1}^{k}\sum_{j=1}^{l} y_j p_{ij} \\
\sigma_X^2 &= \mathrm{V}(X) = \sum_{i=1}^{l}\sum_{j=1}^{k} (x_i-\mu_X)^2 p_{ij} \\
\sigma_Y^2 &= \mathrm{V}(Y) = \sum_{i=1}^{l}\sum_{j=1}^{k} (y_j-\mu_Y)^2 p_{ij}
\end{aligned}$$

Der wichtigste Parameter der gemeinsamen Verteilung ist die **Kovarianz**

$$\sigma_{XY} = \mathrm{COV}(X,Y) = \sum_{i=1}^{l}\sum_{j=1}^{k}(x_i-\mu_X)(y_j-\mu_Y)p_{ij}.$$

Die Bildung der **Kovarianzmatrix**

$$\mathrm{C}(X,Y) = \begin{bmatrix} \mathrm{V}(X) & \mathrm{COV}(X,Y) \\ \mathrm{COV}(X,Y) & \mathrm{V}(Y) \end{bmatrix} = \begin{bmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{bmatrix}.$$

wird sich als nützlich erweisen. Der *Korrelationskoeffizient* ist

$$\varrho(X,Y) = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} = \frac{\mathrm{COV}(X,Y)}{\sqrt{\mathrm{V}(X)\mathrm{V}(Y)}}.$$

Wenn die Kovarianz – oder der Korrelationskoeffizient – zweier Größen X, Y gleich Null ist, so heißen X, Y **unkorreliert**.

5.2 Der kontinuierliche Fall:

Die gemeinsame Verteilung wird durch eine **gemeinsame Dichtefunktion** $f(x,y)$ beschrieben. Für eine Menge $M \subset \mathbb{R}^2$ ist die Wahrscheinlichkeit, dass das Ereignis $(x,y) \in$

$P\big((X,Y)\in M\big) = \iint_M f(x,y)\,\mathrm{d}x\mathrm{d}y$

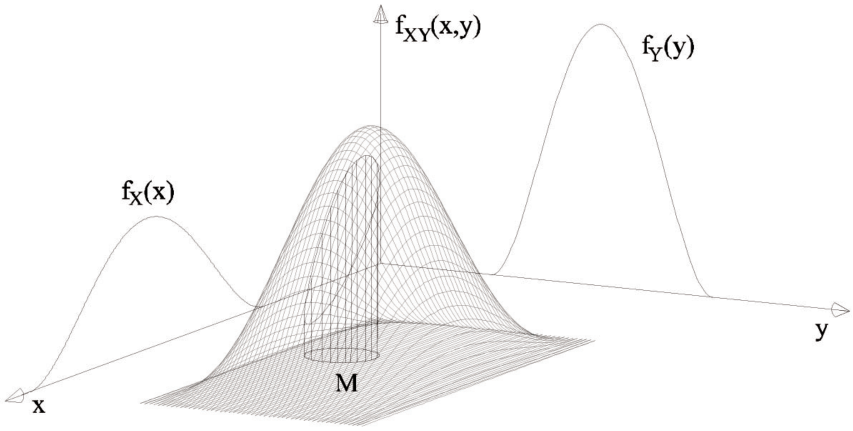

Abbildung 5.1: Dichte einer zweidimensionalen Zufallsgröße
mit Darstellung der Randdichten.

M eintritt, gerade das Volumen unter der Fläche $z = f(x,y)$, wie in Abbildung 5.1 dargestellt.

Insbesondere gilt für das Gesamtvolumen:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x\mathrm{d}y = 1.$$

Die **gemeinsame Verteilungsfunktion** ist

$$F(x,y) = P(X\leq x, Y\leq y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(\xi,\eta)\,\mathrm{d}\eta\mathrm{d}\xi.$$

Die **Randverteilungen** sind gegeben durch

$$F_X(x) = P(X\leq x) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f(\xi,\eta)\,\mathrm{d}\eta\mathrm{d}\xi,$$

$$F_Y(y) = P(Y\leq y) = \int_{-\infty}^{y}\int_{-\infty}^{\infty} f(\xi,\eta)\,\mathrm{d}\xi\mathrm{d}\eta$$

mit den **Randdichten**

$$f_X(x) = \int_{-\infty}^{\infty} f(x,\eta)\,\mathrm{d}\eta, \qquad f_Y(y) = \int_{-\infty}^{\infty} f(\xi,y)\,\mathrm{d}\xi.$$

Die **Erwartungswerte** der beiden Zufallsgrößen erhält man durch

$$\mu_X = \mathrm{E}(X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f(x,y)\,\mathrm{d}x\mathrm{d}y,$$

$$\mu_Y = \mathrm{E}(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f(x,y)\,\mathrm{d}x\mathrm{d}y,$$

die **Varianzen** durch

$$\sigma_X^2 = \mathrm{V}(X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_X)^2 f(x,y)\,\mathrm{d}x\mathrm{d}y,$$

$$\sigma_Y^2 = \mathrm{V}(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (y-\mu_Y)^2 f(x,y)\,\mathrm{d}x\mathrm{d}y$$

und die **Kovarianz** durch

$$\sigma_{XY} = \mathrm{COV}(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_x)(y-\mu_y) f(x,y)\,\mathrm{d}x\mathrm{d}y.$$

Die **Kovarianzmatrix** und der **Korrelationskoeffizient** sind wie im diskreten Fall definiert durch

$$\mathrm{C}(X,Y) = \begin{bmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{bmatrix}, \qquad \varrho(X,Y) = \frac{\sigma_{XY}}{\sigma_X\sigma_Y}.$$

Wir wollen nun die Inverse der Kovarianzmatrix berechnen. Dazu schreiben wir μ_1, μ_2, σ_1, σ_2 und σ_{12} statt μ_X, μ_Y, σ_X, σ_Y und σ_{XY}. Wir fassen das Paar (x,y) zu einem Vektor $\mathbf{x} = [x\ y]^\mathsf{T}$ zusammen und ebenso die beiden Erwartungswerte: $\boldsymbol{\mu} = [\mu_1\ \mu_2]^\mathsf{T}$. Aus $\varrho = \sigma_{12}/\sigma_1\sigma_2$ folgt $\sigma_{12} = \varrho\,\sigma_1\sigma_2$ und wir erhalten

$$\mathrm{C} = \begin{bmatrix} \sigma_1^2 & \varrho\,\sigma_1\sigma_2 \\ \varrho\,\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}, \qquad \det(\mathrm{C}) = \sigma_1^2\sigma_2^2(1-\varrho^2)$$

und weiter

$$\mathrm{C}^{-1} = \frac{1}{\sigma_1^2\sigma_2^2(1-\varrho^2)} \begin{bmatrix} \sigma_2^2 & -\varrho\,\sigma_1\sigma_2 \\ -\varrho\,\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix} = \frac{1}{1-\varrho^2} \begin{bmatrix} \frac{1}{\sigma_1^2} & -\frac{\varrho}{\sigma_1\sigma_2} \\ -\frac{\varrho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2} \end{bmatrix}.$$

Beispiel 5.1 (Die zweidimensionale Normalverteilung)

Ihre Dichte lässt sich am einfachsten mit Hilfe der Kovarianzmatrix schreiben:

$$f(x,y) = \frac{1}{2\pi\sqrt{\det \mathrm{C}}} \exp\Big(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^\mathsf{T}\mathrm{C}^{-1}(\mathbf{x}-\boldsymbol{\mu})\Big).$$

Setzt man die eben hergeleiteten Formeln für die Inverse der Kovarianzmatrix ein, so erhält man für $f(x,y)$ weiter

$$= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\varrho^2}} \exp\left(\frac{-1}{2(1-\varrho^2)} \begin{bmatrix} x-\mu_1 & y-\mu_2 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{\sigma_1^2} & -\frac{\varrho}{\sigma_1\sigma_2} \\ -\frac{\varrho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2} \end{bmatrix} \cdot \begin{bmatrix} x-\mu_1 \\ y-\mu_2 \end{bmatrix}\right)$$

$$= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\varrho^2}} \exp\left(\frac{-1}{2(1-\varrho^2)}\left(\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\varrho}{\sigma_1\sigma_2}(x-\mu_1)(y-\mu_2) + \frac{(y-\mu_2)^2}{\sigma_2^2}\right)\right).$$

Integration nach y bzw. x zeigt, dass die *Randverteilungen* eindimensionale Normalverteilungen sind:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma_1}\,\mathrm{e}^{-(x-\mu_1)^2/2\sigma_1^2}\,, \qquad f_Y(y) = \frac{1}{\sqrt{2\pi}\,\sigma_2}\,\mathrm{e}^{-(y-\mu_2)^2 2\sigma_2^2}\,.$$

Im *unkorrelierten* Fall $\varrho = 0$ ergibt sich

$$\begin{aligned} f(x,y) &= \frac{1}{2\pi\sigma_1\sigma_2}\,\exp\Big(-\frac{(x-\mu_1)^2}{2\sigma_1^2} - \frac{(y-\mu_2)^2}{2\sigma_2^2}\Big) \\ &= \frac{1}{\sqrt{2\pi}\,\sigma_1}\,\mathrm{e}^{-(x-\mu_1)^2/\sigma_1^2} \cdot \frac{1}{\sqrt{2\pi}\,\sigma_2}\,\mathrm{e}^{-(y-\mu_2)^2/2\sigma_2^2}\,, \end{aligned}$$

also

$$f(x,y) = f_X(x) f_Y(y). \tag{5.1}$$

Zwei *normalverteilte* Zufallsgrößen X, Y sind genau dann unkorreliert, wenn die gemeinsame Dichte gleich dem Produkt der Randdichten ist.

5.3 Unabhängigkeit von Zufallsgrößen:

In Anlehnung an den entsprechenden Begriff aus der beschreibenden Statistik (Abschnitt **??**) führen wir nun den Begriff der Unabhängigkeit zweier Zufallsgrößen ein.

Diskreter Fall:

Die Größen X, Y heißen **unabhängig**, wenn stets gilt:

$$p_{ij} = p_{i\cdot} p_{\cdot j}$$

das heißt, die gemeinsamen Wahrscheinlichkeiten ergeben sich als Produkt der Randwahrscheinlichkeiten:

$$P(X = x_i, Y = y_j) = P(X = x_i)P(Y = y_j).$$

Kontinuierlicher Fall:

Die Größen X, Y sind **unabhängig**, wenn gilt:

$$f(x,y) = f_X(x) f_Y(y).$$

Die gemeinsame Dichte ist also das Produkt der Randdichten.

Wir wollen nun die Beziehung zwischen Unabhängigkeit und Unkorreliertheit klären. Wir erinnern, dass zwei Zufallsgrößen (X, Y) *unkorreliert* sind, wenn $\mathrm{COV}(X, Y) = 0$ ist.

Charakterisierung der Unkorreliertheit:

- *Zwei Zufallsgrößen (X, Y) sind genau dann unkorreliert, wenn gilt:*

$$\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y).$$

Dies zeigt die folgende Rechnung:

$$\begin{aligned} \mathrm{COV}(X,Y) &= \mathrm{E}\big((X-\mu_X)(Y-\mu_Y)\big) = \mathrm{E}(XY) - \mu_X \mathrm{E}(Y) - \mu_Y \mathrm{E}(X) + \mu_X\mu_Y \\ &= \mathrm{E}(XY) - \mu_X\mu_Y = \mathrm{E(XY)} - \mathrm{E}(X)\mathrm{E}(Y). \end{aligned}$$

Somit ist $\mathrm{COV}(X,Y) = 0$ gleichbedeutend mit $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$.

- *Unabhängige Zufallsgrößen sind auch unkorreliert.*

Dies folgt, jedenfalls im kontinuierlichen Fall, aus der Überlegung

$$\begin{aligned} \mathrm{E}(XY) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf_X(x)f_Y(y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{-\infty}^{\infty}\Big(\int_{-\infty}^{\infty} yf_Y(y)\,\mathrm{d}y\Big)\,xf_X(x)\,\mathrm{d}x \\ &= \int_{-\infty}^{\infty} \mathrm{E}(Y)\,xf_X\,\mathrm{d}x = \mathrm{E}(X)\mathrm{E}(Y). \end{aligned}$$

Eine ähnliche Rechnung ergibt dasselbe Ergebnis im diskreten Fall.

Aber Achtung! (Umkehrung gilt nicht!)

Das folgende Beispiel zeigt nämlich, dass unkorrelierte Größen im Allgemeinen *nicht* unabhängig sind.

Beispiel 5.2 Die Wahrscheinlichkeitstabelle zweier Zufallsgrößen X mit Werten $-1, 0, 1$ und Y mit Werten $0, 1$ sei

$X \setminus Y$	0	1	
-1	0.1	0.2	0.3
0	0.3	0.1	0.4
1	0.1	0.2	0.3
	0.5	0.5	

Die möglichen Werte für das Produkt XY sind dann $-1, 0, 1$ mit den Wahrscheinlichkeiten

XY	-1	0	1
	0.2	0.6	0.2

Wir erhalten daraus

$$\begin{aligned} \mathrm{E}(XY) &= (-1)\cdot 0.2 + 0\cdot 0.6 + 1\cdot 0.2 &&= 0 \\ \mathrm{E}(X) &= (-1)\cdot 0.3 + 0\cdot 0.4 + 1\cdot 0.3 &&= 0 \\ \mathrm{E}(Y) &= 0\cdot 0.5 + 1\cdot 0.5 &&= 0.5 \end{aligned}$$

Somit gilt $\mathrm{E}(XY) = 0 = \mathrm{E}(X)\mathrm{E}(Y)$, die Größen sind also unkorreliert. Sie sind jedoch nicht unabhängig, da zum Beispiel $p_{11} \neq p_{1\cdot} \cdot p_{\cdot 1}$ ist.

Im Falle normalverteilter Größen ist die Lage allerdings wieder anderst:

- *Zwei normalverteilte Zufallsgrößen sind genau dann unabhängig, wenn sie unkorreliert sind.*

  Wir haben nämlich in Gleichung (5.1) gesehen, dass normalverteilte Größen genau dann unkorreliert sind, wenn ihre gemeinsame Dichte das Produkt der Randdichten ist, und dies bedeutet gerade die Unabhängigkeit.