Mit Hilfe von zwei gegebenen Funktionen $X_1: \Omega \to \Re$ und $X_2: \Omega \to \Re$ lässt sich eine zweidimensionale Funktion $X: \Omega \to \Re^2$ definieren, die jedem $\omega \in \Omega$ das geordnete Paar $(X_1(\omega), (X_1(\omega)) \in \Re^2$ zuordnet.

Diese Funktion *X* heißt 2-dimensionale Zufallsvariable, wenn das Urbild jeder 2-dimensionalen Borel-Menge (Rechtecke und andere einfache Teilmengen des $\Re^2$) ein Ereignis ist. In diesem Fall sind automatisch auch X_1 und X_2 Zufallsvariablen.

Gemeinsame Verteilungsfunktion von X_1 und X_2:

$$F(x_1, x_2) = P(X_1 \leq x_1 \wedge X_2 \leq x_2)$$

$$= P(X^{-1}(-\infty, x_1] \times (-\infty, x_2]))$$

Die Randverteilungsfunktionen

$$F_1(x_1) = P(X_1 \leq x_1) = P(X^{-1}((-\infty, x_1] \times \Re),$$

$$F_2(x_2) = P(X_2 \leq x_2) = P(X^{-1}(\Re \times (-\infty, x_2])$$

sind identisch mit den üblichen (1-dimensionalen) Verteilungsfunktionen von X_1 und X_2.

Die 2-dimensionale Zufallsvariable *X* heißt diskret, falls $X(\Omega)$ endlich oder abzählbar unendlich ist, und stetig, falls eine *2*-dimensionale Dichtefunktion *f* existiert, sodass

$$F(x_1, x_2) = \int_{-\infty}^{x_2} \left(\int_{-\infty}^{x_1} f(t_1, t_2) dt_1 \right) dt_2.$$

Im letzteren Fall sind die Randdichtefunktionen

$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, t_2)\, dt_2,$$

$$f_2(x_2) = \int_{-\infty}^{\infty} f(t_1, x_2)\, dt_1.$$

identisch mit den üblichen (1-dimensionalen) Dichtefunktionen von X_1 und X_2.

Beispiel:

Gemeinsame Dichtefunktion $f(x,y)=x+10y$ für $0<x<4$, $0<y<0.1$ und $f(x,y)=0$ sonst, $2F(1,7) = ?$

Lösung:

$$2F(1,7) = 2\,P(X_1 \leq 1 \wedge X_2 \leq 7)$$

$$= 2\int_{-\infty}^{7}\left(\int_{-\infty}^{1} f(x,y)dx\right)dy$$

$$= 2\int_{-\infty}^{7}\left(\int_{0}^{1} f(x,y)dx\right)dy \quad \text{\# weil } f(x,y) = 0 \text{ für } x < 0$$

$$= 2\int_{0}^{7}\left(\int_{0}^{1} f(x,y)dx\right)dy \quad \text{\# weil } f(x,y) = 0 \text{ für } y < 0)$$

$$= 2\int_{0}^{0.1}\left(\int_{0}^{1} f(x,y)dx\right)dy \quad \text{\# weil } f(x,y) = 0 \text{ für } y > 0.1$$

$$= 2\int_{0}^{0.1}\left(\int_{0}^{1}(x + 10y)dx\right)dy$$

weil der Integrationsbereich (0,1)×(0,0.1) dieses Integrals eine Teilmenge des Rechtecks (0,4)×(0,0.1) ist, auf dem die Dichtefunktion gegeben ist durch $f(x,y)x + 10y$

$$= 2\int_{0}^{0.1}\left(\left\{\frac{x^2}{2} + (10y)\cdot x\right\}\Big|_{0}^{1}\right)dy$$

$$= \int_{0}^{0.1}(\{x^2 + (20y)\cdot x\}|_{0}^{1})\,dy$$

$$= \int_{0}^{0.1}(1^2 + (20y)\cdot 1 - (0^2 + (20y)\cdot 0))\,dy$$

$$= \int_{0}^{0.1}(1 + 20y)\,dy$$

$$= \{y + 10y^2\}|_{0}^{0.1}$$

$$= 0.1 + 10\cdot 0.1^2 - (0 + 10\cdot 0^2)$$

$$= \mathbf{0.2}$$

Beispiel:

Gemeinsame Dichtefunktion $f(x,y)=0.2+sy$ für $-3<x<-2.5$, $1<y<3$ und $f(x,y)=0$ sonst, s = ?

Lösung:

$$1 = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(x,y) dx \right) dy$$

das Integral von f über dem ganzen $\Re^2$ muss gleich 1 sein

$$= \int_{1}^{3} \left(\int_{-3}^{-2.5} f(x,y) dx \right) dy$$

$f(x,y) = 0$ außerhalb des Rechtecks $(-3, -2.5)\times(1,3)$

$$= \int_{1}^{3} \left(\int_{-3}^{-2.5} (0.2 + sy) dx \right) dy$$

$f(x,y) = 0.2 + sy$ innerhalb dieses Rechtecks

$$= \int_1^3 (0.2 + sy) \cdot (-2.5 - (-3))\, dy$$

inneres Integral ist Fläche eines Rechtecks

$$= 0.5 \int_1^3 (0.2 + sy)\, dy$$

$$= 0.5 \left\{ 0.2y + s\frac{y^2}{2} \right\} \Big|_1^3$$

$$= 0.5 \left\{ 0.2 \cdot 3 + s\frac{3^2}{2} - \left(0.2 \cdot 1 + s\frac{1^2}{2} \right) \right\}$$

$$= 0.5 \left\{ 0.4 + s\frac{8}{2} \right\} = 0.2 + 2s$$

Aus $1 = 0.2 + 2s$ folgt, dass $s = \mathbf{0.4}$.

Beispiel:

Gemeinsame Dichtefunktion $f(x,y)=0.8+0.3x$ für $1<x<2$, $c<y<d$ und $f(x,y)=0$ sonst, $d-c=?$

Lösung:

$$1 = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(x,y)dx\right)dy$$

das Integral von f über dem ganzen $\Re^2$ muss gleich 1 sein

$$= \int_{c}^{d}\left(\int_{1}^{2} f(x,y)dx\right)dy$$

$f(x,y) = 0$ außerhalb des Rechtecks $(1, 2)\times(c,d)$

$$= \int_{c}^{d}\left(\int_{1}^{2} (0.8 + 0.3x)dx\right)dy$$

$f(x,y) = 0.8 + 0.3x$ innerhalb dieses Rechtecks

$$= \int_1^2\left(\int_c^d (0.8 + 0.3x)dy\right)dx$$

einfachheitshalber Integrationsreihenfolge vertauscht

$$= \int_1^2 (0.8 + 0.3x)\cdot(d-c)\,dx$$

inneres Integral ist Fläche eines Rechtecks

$$= (d-c)\left\{0.8x + 0.3\frac{x^2}{2}\right\}\Big|_1^2$$

$$= (d-c)\left\{0.8\cdot 2 + 0.3\frac{2^2}{2} - \left(0.8\cdot 1 + 0.3\frac{1^2}{2}\right)\right\}$$

$$= 1.25\,(d-c)$$

Aus $1 = \frac{5}{4}(d-c)$ folgt, dass $d-c = \frac{4}{5} = \mathbf{0.8}$.

Die stetigen Zufallsvariablen X und Y sind genau dann unabhängig, wenn die gemeinsame Dichtefunktion f gleich dem Produkt der beiden Randdichtefunktionen f_X und f_Y ist:

$$X \text{ und } Y \text{ unabhängig} \Leftrightarrow f(x,y) = f_X(x) f_Y(y)$$

Sind X und Y unabhängig, dann gilt:

$$\begin{aligned}
E(XY) &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} xy\, f(x,y)\, dx \right) dy \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} xy\, f_X(x)\, f_Y(y)\, dx \right) dy \\
&= \int_{-\infty}^{\infty} y\, f_Y(y) \left(\int_{-\infty}^{\infty} x\, f_X(x)\, dx \right) dy \\
&= \int_{-\infty}^{\infty} x\, f_X(x)\, dx \int_{-\infty}^{\infty} y\, f_Y(y)\, dy \\
&= E(X)E(Y)
\end{aligned}$$

Das letzte Resultat gilt auch allgemein, also auch im nicht-stetigen Fall. Weiters gilt allgemein:

$$X \text{ und } Y \text{ unabhängig} \Leftrightarrow F(x,y) = F_X(x) F_Y(y)$$

Die Kovarianz

$$\sigma_{XY} = \operatorname{cov}(X,Y) = E(X-\mu_X)(Y-\mu_Y)\,.$$

ist ein Maß für die Beziehung zwischen zwei Zufallsvariablen X und Y mit Erwartungswerten μ_X und μ_Y.

Es gilt:

(i) $\sigma_{XY} = E(X-\mu_X)(Y-\mu_Y)$

$$\begin{aligned}
&= E(XY - X\mu_Y - \mu_X Y + \mu_X\mu_Y)\\
&= E(XY) - E(X\mu_Y) - E(\mu_X Y) + E(\mu_X\mu_Y)\\
&= E(XY) - \mu_Y E(X) - \mu_X E(Y) + \mu_X\mu_Y\\
&= E(XY) - \mu_Y\mu_X - \mu_X\mu_Y + \mu_X\mu_Y\\
&= E(XY) - \mu_X\mu_Y
\end{aligned}$$

(ii) $\sigma_{XY} = E(X-\mu_X)Y = EX(Y-\mu_Y)$

(iii) $\forall a,b,c,d \in \Re:\ \operatorname{cov}(a+bX, c+dY) = bd\ \operatorname{cov}(X,Y)$

(iv) $\operatorname{cov}(X+Y,Z) = \operatorname{cov}(X,Z) + \operatorname{cov}(Y,Z)$

(v) $\operatorname{var}(X+Y) = \operatorname{var}(X) + 2\operatorname{cov}(X,Y) + \operatorname{var}(Y)$

(vi) X, Y ua. $\Rightarrow$ $\operatorname{cov}(X,Y) = 0$

(vii) X, Y ua. $\Rightarrow$ $\operatorname{var}(X+Y) = \operatorname{var}(X) + \operatorname{var}(Y)$

Standardisiert man die beiden Zufallsvariablen X und Y, bevor man die Kovarianz berechnet, dann erhält man den Korrelationskoeffizienten

$$\begin{aligned}
\rho_{XY} &= \operatorname{cov}\left(\frac{X-\mu_X}{\sigma_X}, \frac{Y-\mu_Y}{\sigma_Y}\right)\\
&= \operatorname{cov}\left(\frac{X}{\sigma_X} - \frac{\mu_X}{\sigma_X}, \frac{Y}{\sigma_Y} - \frac{\mu_Y}{\sigma_Y}\right)\\
&= \frac{1}{\sigma_X}\frac{1}{\sigma_Y}\operatorname{cov}(X,Y)\\
&= \frac{\sigma_{XY}}{\sigma_X\sigma_Y}\,.
\end{aligned}$$

Im Unterschied zur Kovarianz hängt der Korrelationskoeffizient nicht von der Skalierung der Zufallsvariablen ab. Er liegt immer zwischen -1 und 1.

Beispiel:

EX = 13, EY = 15, var(X) = 9, var(Y) = 1, ρ = −0.2,
var(8−2X−3Y) = ?

Lösung:

Gegeben sind die Erwartungswerte $\mu_X = 13$, $\mu_Y = 15$ und die Varianzen $\sigma_X^2 = 9$, $\sigma_Y^2 = 1$ der beiden Zufallsvariablen X, Y sowie der Korrelationskoeffizient

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = -0.2.$$

Für die Kovarianz erhält man

$$\sigma_{XY} = \rho_{XY}\sigma_X\sigma_Y = \rho_{XY}\sqrt{\sigma_X^2}\sqrt{\sigma_Y^2} = -0.2\sqrt{9}\sqrt{1} = -0.6$$

und weiters

$$\begin{aligned} &\mathrm{var}(8 - 2X - 3Y) = \mathrm{var}(-2X - 3Y) \\ &= \mathrm{var}(-2X) + 2\mathrm{cov}(-2X, -3Y) + \mathrm{var}(-3Y) \\ &= (-2)^2\mathrm{var}(X) + 2(-2)(-3)\mathrm{cov}(X, Y) + (-3)^2\mathrm{var}(Y) \\ &= 4 \cdot 9 + 12 \cdot (-0.6) + 9 \cdot 1 = \mathbf{37.8} \end{aligned}$$

Die Erwartungswerte wurden nicht benötigt.

Eine Zufallsstichprobe vom Umfang n besteht aus n i.i.d. Zufallsvariablen $X_1,\ldots,X_n$. Eine Zufallsvariable der Form

$$T = g(X_1,\ldots,X_n)$$

heißt Statistik. Statistiken können zum Schätzen von unbekannten Parametern einer Verteilung verwendet werden.

Beispielsweise ist das arithmetische Mittel

$$\bar{X} = \frac{X_1 + \ldots + X_n}{n}$$

eine Statistik, mit der der Erwartungswert μ_X geschätzt werden kann.

Das arithmetische Mittel $\bar{X}$ ist ein **unverzerrter Schätzer** für den Erwartungswert μ_X, weil

$$E(\bar{X}) = \mu_X.$$

Die Stichprobenvarianz

$$s_X^2 = \frac{1}{n}\sum_{j=1}^{n}(X_j - \bar{X})^2$$

ist zwar kein unverzerrter Schätzer für die Varianz σ_X^2, weil

$$E(s_X^2) = \frac{n-1}{n}\sigma_X^2 \neq \sigma_X^2,$$

aber immerhin ein asymptotisch unverzerrter Schätzer, weil

$$E(s_X^2) = \frac{n-1}{n}\sigma_X^2 \to \sigma_X^2.$$

Ein unverzerrter Schätzer für die Varianz σ_X^2 ist gegeben durch

$$S_X^2 = \frac{1}{n-1}\sum_{j=1}^{n}(X_j - \bar{X})^2.$$

Die Rechenregeln für das arithmetische Mittel $\bar{X}$, die Stichprobenvarianz s_X^2 und die Stichprobenkovarianz

$$s_{XY} = \frac{1}{n}\sum_{j=1}^{n}(X_j - \bar{X})(Y_j - \bar{Y})$$

sind analog zu denen für den Erwartungswert, die Varianz und die Kovarianz.

Beispielsweise gilt:

$$s_{XY} = \frac{1}{n}\sum_{j=1}^{n}(X_j - \bar{X})\left(Y_j - \bar{Y}\right) = \frac{1}{n}\sum_{j=1}^{n} X_j Y_j - \bar{X}\bar{Y}$$

Die **Verzerrung** und der **mittlere quadratische Fehler** eines Schätzers $\hat{\theta}$ für θ sind gegeben durch

$$bias(\hat{\theta}) = E(\hat{\theta} - \theta)$$

und

$$MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2 .$$

Es gilt:

$$MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2 = E((\hat{\theta} - E(\hat{\theta})) + \underbrace{(E(\hat{\theta}) - \theta)}_{\in\Re})^2$$

$$= E(\hat{\theta} - E(\hat{\theta}))^2 + 2E(\hat{\theta} - E(\hat{\theta}))(E(\hat{\theta}) - \theta) + (E(\hat{\theta}) - \theta))^2$$

$$= \mathrm{var}(\hat{\theta}) + 2(E(\hat{\theta}) - \theta)\underbrace{E(\hat{\theta} - E(\hat{\theta}))}_{=E(\hat{\theta}) - E(\hat{\theta}) = 0} + bias^2(\hat{\theta}) .$$

$$= \mathrm{var}(\hat{\theta}) + bias^2(\hat{\theta})$$

Beispiel:

X, Y, Z i.i.d. $N(10,10^2)$, $\hat{\mu}=0.7X+0.5Y-0.1Z$, $MSE(\hat{\mu})=?$

Lösung:

Gegeben ist eine Zufallsstichprobe bestehend aus drei unabhängigen und identisch normalverteilten Zufallsvariablen X, Y, Z mit Erwartungswert $\mu = 10$ und Varianz $\sigma^2 = 10^2 = 100$.

Der Standardschätzer für den Erwartungswert μ ist das arithmetische Mittel $\bar{X}$. Hier betrachten wir allerdings einen anderen Schätzer, nämlich

$$\hat{\mu} = 0.7X + 0.5Y - 0.1Z.$$

Für die Verzerrung und die Varianz dieses Schätzers erhalten wir

$$\begin{aligned} bias(\hat{\mu}) &= E(\hat{\mu} - \mu) = E(0.7X + 0.5Y - 0.1Z - \mu) \\ &= 0.7E(X) + 0.5E(Y) - 0.1E(Z) - \mu \\ &= 0.7\mu + 0.5\mu - 0.1\mu - \mu \\ &= (0.7 + 0.5 - 0.1 - 1)\mu = 0.1 \cdot 10 = 1 \end{aligned}$$

und

$$\begin{aligned} \mathrm{var}(\hat{\mu}) &= \mathrm{var}(0.7X + 0.5Y - 0.1Z) \\ &= \mathrm{var}(0.7X) + \mathrm{var}(0.5Y) + \mathrm{var}(0.1Z) \\ &= 0.7^2\mathrm{var}(X) + 0.5^2\mathrm{var}(Y) + 0.1^2\mathrm{var}(Z) \\ &= 0.49\sigma^2 + 0.25\sigma^2 + 0.01\sigma^2 \\ &= (0.49 + 0.25 + 0.01)\sigma^2 \\ &= 0.75 \cdot 100 = 75. \end{aligned}$$

$$\Rightarrow \quad \begin{aligned} \mathrm{MSE}(\hat{\mu}) &= \mathrm{var}(\hat{\mu}) + bias^2(\hat{\mu}) \\ &= 75 + 1^2 = \mathbf{76} \end{aligned}$$