

## Erwartungswert, Varianz, Kovarianz

In einem Spiel wie in Beispiel F.26 interessiert uns der zu erwartende Gewinn und allgemein der „mittlere Wert“ einer reellen Zufallsvariablen.

### Definition F.32 (Erwartungswert einer reellen Zufallsvariablen)

Sei  $X$  eine reelle Zufallsvariable auf dem Wahrscheinlichkeitsraum  $(\Omega, P)$ . Der **Erwartungswert von  $X$**  ist definiert als

$$EX := E(X) := \sum_{\omega \in \Omega} X(\omega) \cdot P(\omega) = \sum_{x \in \mathbb{R}} x \cdot P_X(x). \quad (10)$$

### Bemerkung: Erwartungswert einer Verteilung

In (10) ist  $P_X$  die Verteilung von  $X$  (s. Definition F.28). Lediglich solche Summanden sind ungleich 0, für die  $P_X(x) > 0$ . Dies sind aber nur endlich viele, da der Definitionsbereich und somit der Bildbereich von  $X$  endlich ist. In (10) wird der „steuernde Wahrscheinlichkeitsraum“  $\Omega$  nicht explizit erwähnt. Der Erwartungswert ist also eine Eigenschaft der Verteilung. Durch (10) ist der **Erwartungswert der Verteilung  $P_X$**  definiert, und analog definiert man allgemein den **Erwartungswert eines Wahrscheinlichkeitsmaßes auf endlichen Mengen reeller Zahlen**.

- 245 -

## Beispiel F.34 (für Erwartungswerte spezieller Verteilungen)

1. Wir berechnen den Erwartungswert der **Binomialverteilung** zu den Parametern  $n$  und  $p$  (s. (8)) auf zwei verschiedene Weisen.

1. *Methode:*

$$\begin{aligned} E(X) &= \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} \\ &= np \sum_{k=1}^n \frac{(n-1)!}{(k-1)!((n-1)-(k-1))!} p^{(k-1)} (1-p)^{(n-1)-(k-1)} \\ &= np \sum_{\tilde{k}=0}^{\tilde{n}} \binom{\tilde{n}}{\tilde{k}} p^{\tilde{k}} (1-p)^{\tilde{n}-\tilde{k}} \\ &= np (p + (1-p))^{\tilde{n}} \\ &= np. \end{aligned}$$

Dabei haben wir die Substitution  $n-1 = \tilde{n}$  und  $k-1 = \tilde{k}$  verwendet.

- 247 -

## Satz F.33 (Eigenschaften des Erwartungswertes)

- Der Erwartungswert ist *linear*, d.h. für reelle Zufallsvariablen  $X, Y$  und  $\lambda \in \mathbb{R}$  gilt

$$E(\lambda X + Y) = \lambda \cdot E(X) + E(Y). \quad (11)$$

- Sind  $X, Y$  unabhängig, so gilt

$$E(X \cdot Y) = E(X) \cdot E(Y).$$

Hierbei bezeichnet  $X \cdot Y$  das Produkt der beiden Zufallsvariablen. Diese durch  $(X \cdot Y)(\omega) = X(\omega) \cdot Y(\omega)$  definierte Produktfunktion ist wieder eine reelle Zufallsvariable auf demselben Wahrscheinlichkeitsraum.

- 246 -

- Methode:* Wir verwenden (11) (Linearität von  $E$ ). Es gilt

$$X = X_1 + \dots + X_n$$

mit  $X_i : \Omega \rightarrow \{0, 1\}$ ,  $P(\{X_i = 1\}) = p$ ,  $P(\{X_i = 0\}) = 1 - p$ , also  $E(X_i) = p$  und somit

$$E(X) = \sum_{i=1}^n E(X_i) = np.$$

2. Wir berechnen den Erwartungswert für die Augenzahl beim **Laplace-Würfel**, gegeben durch  $\Omega = \{1, \dots, 6\}$  und  $P(\omega) = \frac{1}{6}$  für  $\omega \in \Omega$ . Die Zufallsvariable  $X$  gibt die Augenzahl an. (S. Beispiel F.26) Wir erhalten

$$E(X) = \sum_{i=1}^6 i \cdot \frac{1}{6} = 3.5. \quad (12)$$

Insbesondere sehen wir, dass der Erwartungswert i.a. nicht als Wert von der Zufallsvariablen angenommen wird.

- 248 -

3. Wir vergleichen das letzte Beispiel mit der Zufallsvariablen  $Y$ , definiert auf demselben  $(\Omega, P)$  durch

$$Y(\omega) = 3.5 \quad \text{für } \omega \in \{1, \dots, 6\}.$$

Diese Zufallsvariable hat den gleichen Erwartungswert wie der Laplace-Würfel:

$$E(Y) = 3.5.$$

Dennoch sind die beiden Zufallsvariablen nicht gleichverteilt. Wie durch die **Stabdiagramme** in der folgenden Abbildung veranschaulicht wird, ist die Verteilung  $P_Y$  deterministisch, wohingegen  $P_X$  um den Erwartungswert streut.

### Definition F.35

#### (Varianz, Streuung, Kovarianz, Korrelationskoeffizient)

Seien  $(\Omega, P)$  ein endlicher Wahrscheinlichkeitsraum und  $X, Y$  reelle Zufallsvariablen.

1. Die **Varianz** von  $X$  ist

$$\text{Var}(X) = E((X - E(X))^2).$$

2. Die **Streuung** (oder **Standardabweichung**) von  $X$  ist

$$\sigma = \sqrt{\text{Var}(X)}.$$

3. Die **Kovarianz** von  $X$  und  $Y$  ist

$$\text{Cov}(X, Y) = E((X - E(X)) \cdot (Y - E(Y))).$$

4. Der **Korrelationskoeffizient** von  $X$  und  $Y$  (mit  $\sigma_x, \sigma_y \neq 0$ ) ist

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sigma_x \sigma_y}. \quad (13)$$

5. Zufallsvariablen  $X, Y$  mit  $\text{Cov}(X, Y) = 0$  heißen **unkorreliert**.

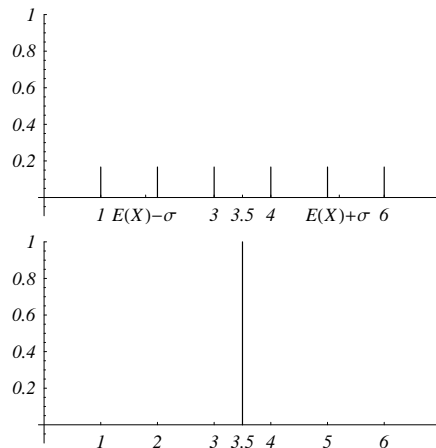


Abbildung: Stabdiagramme für den Laplace-Würfel und für eine deterministische Zufallsvariable

### Satz F.36 (Eigenschaften von Varianz und Kovarianz)

Seien  $X, Y, X_i$  (für  $1 \leq i \leq n$ ) reelle Zufallsvariablen und  $a, b, c, d \in \mathbb{R}$ .

Dann gilt:

1.

$$\text{Var}(X) = E(X^2) - (E(X))^2. \quad (14)$$

2.

$$\text{Var}(aX + b) = a^2 \cdot \text{Var}(X). \quad (15)$$

3.

$$\text{Cov}(X, Y) = E(XY) - E(X) \cdot E(Y). \quad (16)$$

4.

$$\text{Cov}(aX + b, cY + d) = a \cdot c \cdot \text{Cov}(X, Y), \quad (17)$$

5.

$$\text{Var}(X_1 + \dots + X_n) = \sum_{i=1}^n \text{Var}(X_i) + \sum_{\substack{(i,j) \\ i \neq j}} \text{Cov}(X_i, X_j), \quad (18)$$

wobei in der letzten Summe die Summanden  $\text{Cov}(X_1, X_2)$  und  $\text{Cov}(X_2, X_1)$  etc. auftreten.

6. Sind  $X, Y$  unabhängig, so sind sie auch unkorreliert.

7. (**Formel von Bienaymé**) Wenn  $X_1, \dots, X_n$  unabhängig sind, dann gilt

$$\text{Var}(X_1 + \dots + X_n) = \sum_{i=1}^n \text{Var}(X_i). \quad (19)$$

### Bemerkung

(Aus Unkorreliertheit folgt nicht Unabhängigkeit)

Aus der Unkorreliertheit von Zufallsvariablen folgt im Allgemeinen *nicht* deren Unabhängigkeit, wie wir in Beispiel F.41 sehen werden.

- 253 -

### Beispiel F.38 (Varianz der Binomialverteilung)

Mit Hilfe der Formel von Bienaymé (19) berechnen wir analog zur 2. Methode in Beispiel F.34 die Varianz der Binomialverteilung zu den Parametern  $n$  und  $p$ . Die Varianz von  $X_i$  ist

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

Aus der Unabhängigkeit der  $X_i$  folgt also

$$\text{Var}(X) = \text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i) = np(1 - p).$$

- 255 -

### Beispiel F.37 (Varianz bei der Augenzahl des Laplace-Würfels)

Es gilt für das **zweite Moment** der Augenzahl  $X$  des Laplace-Würfels:

$$E(X^2) = \sum_{i=1}^6 i^2 \cdot \frac{1}{6} = \frac{91}{6}.$$

Daraus erhalten wir nach (14) und unter Verwendung von (12)

$$\begin{aligned} \text{Var}(X) &= E(X^2) - (E(X))^2 \\ &= \frac{91}{6} - 3.5^2 = \frac{35}{12}. \end{aligned} \quad (20)$$

Die Streuung ist also  $\sigma_X \approx 1.71$ .

- 254 -

Zur Veranschaulichung von Korrelation führen wir noch den wichtigen Begriff der *gemeinsamen Verteilung* ein und beschränken uns dabei hier auf den Fall zweier reellwertiger Zufallsvariablen. Zur naheliegenden Verallgemeinerung auf den Fall von endlich vielen Zufallsvariablen mit Werten in beliebigen Mengen s. z.B. [Krengel]

### Definition F.39

(**Gemeinsame Verteilung zweier reeller Zufallsvariablen**)

Seien  $X, Y : \Omega \mapsto \mathbb{R}$  zwei auf derselben Ergebnismenge  $\Omega$  definierten reellwertigen Zufallsvariablen. Die Verteilung  $P_{X \times Y}$  (vgl. Definition F.28) der Produktfunktion

$$X \times Y : \Omega \mapsto \mathbb{R}^2$$

heißt **gemeinsame Verteilung** von  $X$  und  $Y$ . Die Funktion  $X \times Y$  nimmt genau die Werte  $(x, y) \in \mathbb{R}^2$  mit positiver Wahrscheinlichkeit an, für die  $P_X(x) > 0$  und  $P_Y(y) > 0$  gilt und gemäß Satz F.27 erhalten wir

$$P_{X \times Y}(x, y) = P(\omega \in \Omega : X(\omega) = x \text{ und } Y(\omega) = y).$$

- 256 -

### Beispiel F.40 (Korrelation bei Merkmalsverteilung)

Seien  $X_1$  und  $X_2$  Zufallsvariablen mit Werten in  $\{0, 1\}$ . Die Produktzufallsvariable  $X_1 \times X_2$  nehme die Werte  $(0, 0)$ ,  $(1, 0)$ ,  $(0, 1)$  und  $(1, 1)$  mit den Wahrscheinlichkeiten  $\frac{1}{10}$ ,  $\frac{1}{5}$ ,  $\frac{3}{10}$ ,  $\frac{2}{5}$ , respektive, an. Wir schreiben abkürzend  $P_{X_1 \times X_2}(1, 1)$  statt  $P_{X_1 \times X_2}(\{(1, 1)\})$  etc. Wir stellen die gemeinsame Verteilung sowie die Verteilungen von  $X_1$  und  $X_2$  tabellarisch dar:

	$X_2 = 0$	$X_2 = 1$	Verteilung von $X_2$ :
$X_1 = 0$	$1/10$	$3/10$	$2/5$
$X_1 = 1$	$1/5$	$2/5$	$3/5$
Verteilung von $X_1$ :	$3/10$	$7/10$	

Die Verteilung von  $X_1$  und  $X_2$  steht offensichtlich im oberen linken Teil der Tabelle. Die Verteilung von  $X_1$  steht in der unteren Zeile. Die Werte wurden als Summe der Zahlen der jeweiligen Spalten berechnet. Ebenso steht die Verteilung von  $X_2$  in der rechten Spalte. Diese Werte sind jeweils die Zeilensummen (aus dem Tabellenteil der gemeinsamen Verteilung). Eine Kontrollrechnung zeigt, dass die Summe der Werte der unteren Zeile (der rechten Spalte) jeweils 1 ergeben.

$$E(X_1 \cdot X_2) = \frac{2}{5},$$

$$\text{Cov}(X_1, X_2) = E(X_1 \cdot X_2) - E(X_1) \cdot E(X_2)$$

$$= \frac{2}{5} - \frac{3}{5} \cdot \frac{7}{10} = -\frac{1}{50},$$

$$\rho_{X_1, X_2} = \frac{-\frac{1}{50}}{\sqrt{\frac{6}{25} \cdot \frac{21}{100}}} \approx -0.089.$$

Die Zufallsvariablen  $X_1$  und  $X_2$  sind nicht voneinander unabhängig, da Ihre Kovarianz ungleich 0 ist. (Es gilt nämlich: „Unabhängigkeit  $\Rightarrow$  Kovarianz gleich 0“..) Der Betrag ihres Korrelationskoeffizienten ist allerdings auch nicht besonders groß, d.h. nahe bei 0.

Wir berechnen nun die Kenngrößen der Verteilungen.

$$E(X_1) = 0 \cdot \frac{2}{5} + 1 \cdot \frac{3}{5} = \frac{3}{5},$$

$$E(X_1^2) = \frac{3}{5},$$

$$\text{Var}(X_1) = \frac{3}{5} - \left(\frac{3}{5}\right)^2 = \frac{6}{25},$$

$$\sigma_{X_1} = \sqrt{\frac{6}{25}} \approx 0.49.$$

$$E(X_2) = \frac{7}{10}, \quad E(X_2^2) = \frac{7}{10},$$

$$\text{Var}(X_2) = \frac{7}{10} - \left(\frac{7}{10}\right)^2 = \frac{21}{100},$$

$$\sigma_{X_2} = \sqrt{\frac{21}{100}} \approx 0.46.$$

### Bemerkung: Interpretation von Korrelation

#### 1. (geometrische Sichtweise)

Wir können die Kovarianz als Skalarprodukt in  $\mathbb{R}^n$  mit  $n = |\Omega|$  auffassen. Hierzu nehmen wir an, dass alle Elementarereignisse eine positive Wahrscheinlichkeit haben. Dann gilt die Cauchy-Schwarz-Ungleichung

$$\text{Cov}(X, Y) \leq \sigma_x \sigma_y$$

und somit für  $\sigma_x, \sigma_y \neq 0$ :

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

Den Korrelationskoeffizienten können wir dann als „Kosinus des nicht-orientierten Winkels zwischen  $X$  und  $Y$ “ auffassen.

#### 2. (Korrelation als linearer Zusammenhang)

Für zwei Zufallsvariablen  $X$  und  $Y$  deutet ein Korrelationskoeffizient  $\rho_{X, Y}$  nahe bei 1 auf eine „Tendenz“ der Variablen  $X - E(X)$  und  $Y - E(Y)$  hin, gemeinsam große bzw. kleine bzw. stark negative Werte anzunehmen, also auf einen „linearen Zusammenhang“. Analoges gilt für  $\rho_{X, Y}$  nahe bei  $-1$ . Wir veranschaulichen dies in Beispiel F.41.