

§31 Zufallsvariable, Erwartungswert, Varianz

Oft will man den Ergebnissen von Zufallsversuchen reelle Zahlen zuordnen (beispielsweise als Gewinn bei einem Glücksspiel). Hierzu benötigen wir den Begriff der Zufallsvariablen.

Definition 31.1

Sei Ω ein Stichprobenraum. Eine Funktion X , die jedem Ergebnis $\omega \in \Omega$ eine reelle Zahl $X(\omega)$ zuordnet, heißt **Zufallsvariable**.

Bemerkung:

Eine Zufallsvariable ist weder zufällig noch eine Variable, sondern eine Funktion. Man kann sie sich stets als Gewinn bei einem Glücksspiel vorstellen.

31.2 Beispiel

Eine faire Münze mit Seiten 0 und 1 werde 3 Mal geworfen. Die Anzahl der Einsen sei der Gewinn. Man kann $\Omega = \{000, 001, \dots, 111\}$ als Stichprobenraum und den Gewinn als Zufallsvariable $X(\omega)$ auffassen.

Ergebnis ω	000	001	010	011	100	101	110	111
Gewinn $X(\omega)$	0	1	1	2	1	2	2	3
Wahrscheinlichkeit $P(\omega)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

Verschiedene Ereignisse ω_1, ω_2 können zum selben Gewinn führen.

Definition 31.3

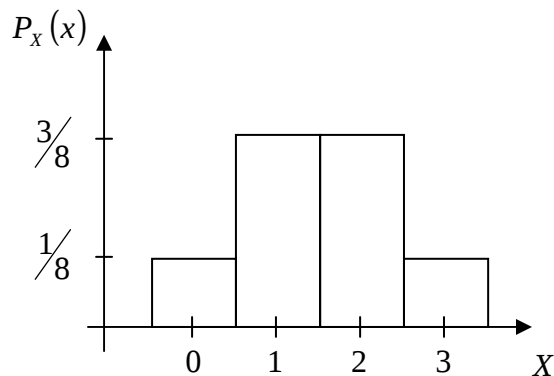
Eine **Verteilung** P_X einer Zufallsvariablen X ordnet jedem Wert $x \in X$ eine Wahrscheinlichkeit $P_X(x)$ zu.

31.4 Beispiel

Im Beispiel 31.2 können wir dem Gewinn folgende Wahrscheinlichkeiten zuordnen:

Gewinn X	0	1	2	3
Wahrscheinlichkeit $P_X(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Oft veranschaulicht man die Verteilung einer Zufallsvariablen durch ein Histogramm:



Bemerkung:

Da wir hier meist nur diskrete Zufallsvariablen betrachten, sind die Verteilungen ebenfalls diskret.

Interessiert man sich für den Durchschnittsgewinn je Versuchswiederholung, gelangt man zu dem Begriff des Erwartungswertes.

Definition 31.5

Unter dem **Erwartungswert** $E(X)$ einer (diskreten) Zufallsvariablen X versteht man das gewichtete Mittel der Funktion X über Ω , wobei jeder Wert mit seiner Wahrscheinlichkeit gewichtet wird.

$$E(X) = \sum_{\omega \in \Omega} X(\omega) \cdot P(\omega)$$

31.6 Beispiel

Für die Zufallsvariable X aus 31.2 erhält man den Erwartungswert

$$\begin{aligned} E(X) &= 0 \cdot \frac{1}{8} + 1 \cdot \frac{1}{8} + 1 \cdot \frac{1}{8} + 2 \cdot \frac{1}{8} + 1 \cdot \frac{1}{8} + 2 \cdot \frac{1}{8} + 2 \cdot \frac{1}{8} + 3 \cdot \frac{1}{8} \\ &= \frac{3}{2} \end{aligned}$$

Bequemer hätte man den Erwartungswert mit der Verteilung P_X berechnet (vgl. 31.4):

$$\begin{aligned} E(X) &= 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} \\ &= \frac{3}{2} \end{aligned}$$

Es gilt also auch:

$$E(X) = \sum_{x \in X(\Omega)} x \cdot P_X(x)$$

Bemerkung:

Bei kontinuierlichen Zufallsvariablen verwendet man Integrale statt Summen.

Mit dem Erwartungswert lässt sich gut arbeiten, da er linear ist:

Satz 31.7 (Linearität des Erwartungswertes)

Seien X, Y Zufallsvariablen und $\alpha, \beta \in \mathbb{R}$. Dann gilt:

$$E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$$

Beweis:

$$\begin{aligned} E(\alpha X + \beta Y) &= \sum_{\omega \in \Omega} (\alpha X(\omega) + \beta Y(\omega)) \cdot P(\omega) \\ &= \alpha \sum_{\omega \in \Omega} X(\omega) P(\omega) + \beta \sum_{\omega \in \Omega} Y(\omega) P(\omega) \\ &= \alpha E(X) + \beta E(Y) \end{aligned} \quad \text{q.e.d.}$$

31.8 Unabhängigkeit von Zufallsvariablen

Wir wissen: $E(X + Y) = E(X) + E(Y)$.

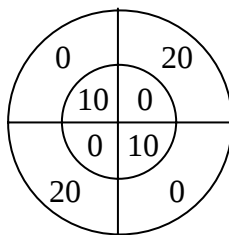
Gilt jedoch auch $E(XY) = E(X)E(Y)$?

Man kann zeigen, dass dies dann gilt, wenn X und Y **unabhängig** sind, d.h.

$$P((X = a) \cap (Y = b)) = P(X = a)P(Y = b) \quad \forall a, b$$

31.9 Beispiele

a)



Glücksrad mit 4 Ergebnissen, auf denen die Zufallsvariablen X (äußerer Ring) und Y (innerer Ring) definiert sind

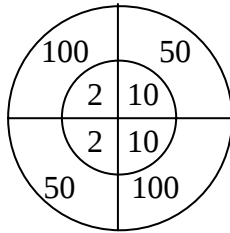
$$E(X) = \frac{1}{4} \cdot 20 + \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot 20 + \frac{1}{4} \cdot 0 = 10$$

$$E(Y) = \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot 10 + \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot 10 = 5$$

$$E(XY) = \frac{1}{4} \cdot 20 \cdot 0 + \frac{1}{4} \cdot 0 \cdot 10 + \frac{1}{4} \cdot 20 \cdot 0 + \frac{1}{4} \cdot 0 \cdot 10 = 0 \neq E(X)E(Y)$$

X und Y sind hier nicht unabhängig: Das Ereignis $Y = 0$ hat die Wahrscheinlichkeit $\frac{1}{2}$. Weiß man jedoch, dass $X = 20$ eingetreten ist, dann hat $Y = 0$ die Wahrscheinlichkeit 1.

b)



$$E(X) = \frac{1}{4} \cdot 50 + \frac{1}{4} \cdot 100 + \frac{1}{4} \cdot 50 + \frac{1}{4} \cdot 100 = 75$$

$$E(Y) = \frac{1}{4} \cdot 10 + \frac{1}{4} \cdot 10 + \frac{1}{4} \cdot 2 + \frac{1}{4} \cdot 2 = 6$$

$$\Rightarrow E(X)E(Y) = 450$$

$$E(XY) = \frac{1}{4} \cdot 50 \cdot 10 + \frac{1}{4} \cdot 100 \cdot 10 + \frac{1}{4} \cdot 50 \cdot 2 + \frac{1}{4} \cdot 100 \cdot 2 = 450 = E(X)E(Y)$$

X und Y sind unabhängig:

$Y = 2$ und $Y = 10$ sind gleich wahrscheinlich. Weiß man z.B., dass $X = 50$ eingetreten ist, so sind $Y = 2$ und $Y = 10$ noch stets gleich wahrscheinlich.

Häufig ist man nicht nur am Erwartungswert interessiert. Man möchte auch wissen, wie stark eine Verteilung um den Erwartungswert streut. Hierzu dienen die Begriffe Varianz und Standardabweichung.

Definition 31.10

Sei X eine Zufallsvariable mit Erwartungswert $\mu = E(X)$. Dann versteht man unter der **Varianz** $V(X) = \sigma^2$ den Erwartungswert von $(X - \mu)^2$:

$$\sigma^2 = V(X) = E((X - \mu)^2)$$

$\sigma = \sqrt{V(X)}$ nennt man die **Standardabweichung** (Streuung) von X .

31.11 Berechnung der Varianz

Wegen der Linearität des Erwartungswertes und wegen $E(\text{const.}) = \text{const.}$ gilt:

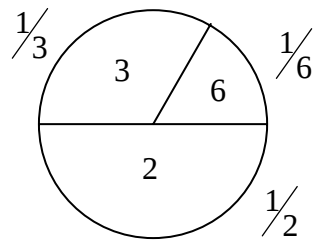
$$\begin{aligned} E((X - \mu)^2) &= E(X^2 - 2\mu X + \mu^2) \\ &= E(X^2) - 2\mu \underbrace{E(X)}_{=\mu} + \underbrace{E(\mu^2)}_{=\mu^2} = E(X^2) - \mu^2. \end{aligned}$$

Wir erhalten damit eine wichtige Formel zum Berechnen der Varianz:

$$\boxed{\sigma^2 = E(X^2) - \mu^2}$$

31.12 Beispiel

Sei X der Gewinn auf dem Glücksrad



mittlerer Gewinn:

$$\mu = E(X) = \frac{1}{2} \cdot 2 + \frac{1}{3} \cdot 3 + \frac{1}{6} \cdot 6 = 3$$

$$E(X^2) = \frac{1}{2} \cdot 4 + \frac{1}{3} \cdot 9 + \frac{1}{6} \cdot 36 = 11$$

Varianz:

$$\sigma^2 = E(X^2) - \mu^2 = 11 - 9 = 2$$

Standardabweichung:

$$\sigma = \sqrt{2}$$

Satz 31.13 (Eigenschaften der Varianz)

Sei X eine Zufallsvariable und seien $\alpha, \beta \in \mathbb{R}$. Dann gilt:

- i) $V(\alpha X) = \alpha^2 V(X)$
- ii) $V(X + \beta) = V(X)$

Beweis:

- i) $V(\alpha X) = E((\alpha X - \alpha \mu)^2) = E(\alpha^2 (X - \mu)^2) = \alpha^2 V(X)$
- ii) $V(X + \beta) = E((X + \beta - (\mu + \beta))^2) = V(X)$ q.e.d.

Letzte Vorlesung: Standardabweichung σ ist Maß dafür, wie stark X um seinen Erwartungswert streut. Lässt sich dies präzisieren?

Satz 31.14 (Tschebyschew-Ungleichung)

Die Wahrscheinlichkeit, dass eine Zufallsvariable X von ihrem Erwartungswert μ um mindestens $t\sigma$ abweicht ($t \in \mathbb{R}, t > 0$), ist kleiner als $\frac{1}{t^2}$:

$$P(|X - \mu| \geq t\sigma) \leq \frac{1}{t^2}$$

Beweis:

Sei Y eine Zufallsvariable mit $Y(\omega) \geq 0 \quad \forall \omega \in \Omega$ und $\varepsilon > 0$.

Wir setzen $A := \{\omega \in \Omega \mid Y(\omega) \geq \varepsilon\}$ und betrachten die Indikatorfunktion

$$I_A(x) := \begin{cases} 1 & (x \in A) \\ 0 & (x \notin A) \end{cases}$$

Dann gilt: $Y \geq \varepsilon I_A$; Quadrieren liefert $Y^2 \geq \varepsilon^2 I_A$.

Mit $E(I_A) = 0 \cdot p(\bar{A}) + 1 \cdot p(A) = p(A)$ folgt

$$\begin{aligned} E(Y^2) &\geq \varepsilon^2 E(I_A) \\ &\geq \varepsilon^2 p(A) \end{aligned}$$

und damit

$$P(Y \geq \varepsilon) = p(A) \leq \frac{1}{\varepsilon^2} E(Y^2)$$

Mit $Y := |X - \mu|$, $\mu := E(X)$ und $\sigma^2 := V(X)$ folgt

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

Mit $\varepsilon = t\sigma$ folgt die Behauptung.

q.e.d.

31.15 Bemerkungen

- a) Die Tschebyschew-Ungleichung ist meist recht grob. Weiß man Genaueres über die Verteilung, so lassen sich oft schärfere Abschätzungen gewinnen.
- b) Mit Hilfe der Tschebyschew-Ungleichung kann man das **schwache Gesetz der großen Zahlen** beweisen:
Wird ein Versuch mit Erfolgswahrscheinlichkeit p n -mal wiederholt, dann strebt die Wahrscheinlichkeit, dass die relative Erfolgshäufigkeit um weniger als ε von t abweicht, gegen 1 für $n \rightarrow \infty$.
Dabei ist ε eine beliebig kleine positive Zahl.

Definition 31.16

Seien X, Y Zufallsvariablen über demselben Stichprobenraum Ω mit Erwartungswerten μ_X, μ_Y und Varianzen $\sigma_X^2, \sigma_Y^2 > 0$.

Dann nennt man

$$\text{Cov}(X, Y) := \sigma_{XY}^2 := E((X - \mu_X)(Y - \mu_Y))$$

die **Kovarianz** von X und Y und

$$\rho_{XY} := \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

den **Korrelationskoeffizienten** von X und Y .

Ist $\text{Cov}(X, Y) = 0$, so heißen X und Y **unkorreliert**.

31.17 Bemerkungen

- a) $V(X) = \text{Cov}(X, X)$
- b) Fasst man $X - \mu_X$ und $Y - \mu_Y$ als Funktion über dem endlichen Stichprobenraum $\Omega = \{\omega_1, \dots, \omega_n\}$ auf, so kann sie auch als Vektoren aus dem $\mathbb{R}^n$ mit der i -ten Komponente $X(\omega_i) - \mu_X$ bzw. $Y(\omega_i) - \mu_Y$ interpretieren.

Dann bezeichnen (vgl. §7, §9)

- die Standardabweichungen die euklidischen Normen dieser Vektoren
- die Varianzen die quadrierten euklidischen Normen
- die Kovarianz das euklidische Skalarprodukt
- der Korrelationskoeffizient den Cosinus des Winkels zwischen beiden Vektoren (insbesondere ist $-1 \leq \rho_{XY} \leq 1$) bezüglich des gewichteten euklidischen

Skalarprodukts $\langle u, v \rangle = \sum_{i=1}^n p_i u_i v_i$ wobei p_i die Wahrscheinlichkeit für das Auftreten von ω_i ist.

- Unkorreliertheit entspricht Orthogonalität der Vektoren (bzgl. dieses Skalarproduktes)

Man kann folgende Resultate zeigen:

Satz 31.18 (Rechenregeln für die Kovarianz)

Seien X, Y Zufallsvariablen über Ω mit Erwartungswerten μ_X, μ_Y . Dann gilt:

- a) $\text{Cov}(X, Y) = E(XY) - \mu_X \mu_Y$
- b) $\text{Cov}(\alpha X + \beta, \gamma Y + \delta) = \alpha \gamma \text{Cov}(X, Y) \quad \forall \alpha, \beta, \gamma, \delta \in \mathbb{R}$
- c) $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
- d) Für m Zufallsvariablen $X_1, \dots, X_m$ gilt:

$$V(X_1 + \dots + X_m) = \sum_{i=1}^m V(X_i) + \sum_{i \neq j} \text{Cov}(X_i, X_j)$$

- e) Sind X, Y unabhängig, so sind sie auch unkorreliert.

- f) Für unabhängige $X_1, \dots, X_m$ gilt $V(X_1 + \dots + X_m) = \sum_{i=1}^m V(X_i)$

31.19 Bemerkungen

- a) Die Umkehrung zu 31.18.e) gilt **nicht!** Beispiel:

Ergebnis ω	1	2	3	4
Zufallsvariable X	1	-1	2	-2
Zufallsvariable Y	-1	1	2	-2
Wahrscheinlichkeit	$\frac{2}{5}$	$\frac{2}{5}$	$\frac{1}{10}$	$\frac{1}{10}$

X, Y sind nicht unabhängig, denn $X(\omega)$ bestimmt ω und damit $Y(\omega)$ eindeutig.

Aber: $E(X) = E(Y) = 0$ und $\text{Cov}(X, Y) = \frac{2}{5}(-1) + \frac{2}{5}(-1) + \frac{1}{10} \cdot 4 + \frac{1}{10} \cdot 4 = 0$

- b) In der Informatik tauchen Erwartungswerte und Varianzen z.B. bei der Zuverlässigkeitsanalyse von Systemen auf oder bei der Abschätzung von Wartezeiten bei Internetanfragen.

Kovarianzen sind unter anderem wichtig im Bereich des maschinellen Lernens.

31.20 Zuverlässigkeitsanalyse von Systemen

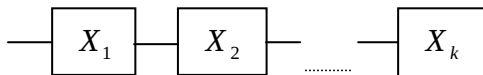
Wir wollen die Zuverlässigkeit eines Systems berechnen, das aus k Komponenten besteht, die zwei Zustände besitzen (arbeiten (1) oder arbeiten nicht (0)).

Der Zustand der Einzelkomponenten soll voneinander unabhängig sein.

Unter der Zuverlässigkeit eines Systems verstehen wir die Wahrscheinlichkeit, dass es arbeitet.

Sei p_k die Wahrscheinlichkeit, dass die k -te Komponente arbeitet, und $q_k = 1 - p_k$ die Wahrscheinlichkeit, dass sie ausfällt.

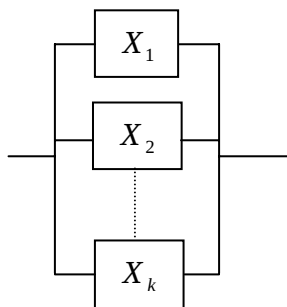
a) Reihensysteme



Ein Reihensystem arbeitet, wenn alle seine Komponenten arbeiten:

$$p = p_1 \cdot p_2 \cdot \dots \cdot p_k$$

b) Parallelsystem



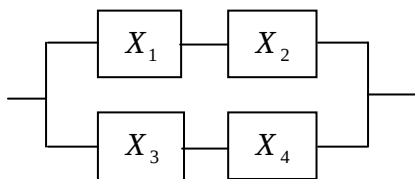
Ein Parallelsystem fällt aus, wenn all seine Komponenten ausfallen.

$$q = q_1 \cdot q_2 \cdot \dots \cdot q_k$$

$$\Rightarrow p = 1 - q = 1 - (1 - p_1)(1 - p_2) \cdot \dots \cdot (1 - p_k)$$

c) Gemischte Systeme

werden hierarchisch in Reihen- und Parallelsysteme zerlegt



oberes Reihensystem $p_1 p_2$

unteres Reihensystem $p_3 p_4$

äußeres Parallelsystem: $1 - (1 - p_1 p_2)(1 - p_3 p_4)$