Definition 35

Sei X eine Zufallsvariable und A ein Ereignis mit $\Pr[A] > 0$. Die bedingte Zufallsvariable $X|A$ besitzt die Dichte

$$f_{X|A}(x) := \Pr[X = x \mid A] = \frac{\Pr[„X = x" \cap A]}{\Pr[A]}\,.$$

Die Definition von $f_{X|A}$ ist zulässig, da

$$\sum_{x \in W_X} f_{X|A}(x) = \sum_{x \in W_X} \frac{\Pr[„X = x" \cap A]}{\Pr[A]} = \frac{\Pr[A]}{\Pr[A]} = 1\,.$$

Der Erwartungswert $\mathbb{E}[X|A]$ der Zufallsvariablen $X|A$ berechnet sich entsprechend:

$$\mathbb{E}[X|A] = \sum_{x \in W_X} x \cdot f_{X|A}(x)\,.$$

Satz 36

Sei X eine Zufallsvariable. Für paarweise disjunkte Ereignisse $A_1, \ldots, A_n$ mit $A_1 \cup \ldots \cup A_n = \Omega$ und $\Pr[A_1], \ldots, \Pr[A_n] > 0$ gilt

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X|A_i] \cdot \Pr[A_i]\,.$$

Für paarweise disjunkte Ereignisse $A_1, A_2, \ldots$ mit $\bigcup_{i=1}^{\infty} A_k = \Omega$ und $\Pr[A_1]$, $\Pr[A_2], \ldots > 0$ gilt analog

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} \mathbb{E}[X|A_i] \cdot \Pr[A_i],$$

sofern die Erwartungswerte auf der rechten Seite alle existieren und die Summe $\sum_{i=1}^{\infty} |\mathbb{E}[X|A_i]| \cdot \Pr[A_i]$ konvergiert.

Beweis:

$$\begin{aligned}\mathbb{E}[X] &= \sum_{x \in W_X} x \cdot \Pr[X = x] = \sum_{x \in W_X} x \cdot \sum_{i=1}^{n} \Pr[X = x | A_i] \cdot \Pr[A_i] \\ &= \sum_{i=1}^{n} \Pr[A_i] \sum_{x \in W_X} x \cdot \Pr[X = x | A_i] = \sum_{i=1}^{n} \Pr[A_i] \cdot \mathbb{E}[X | A_i].\end{aligned}$$

Der Beweis für den unendlichen Fall verläuft analog. □

Beispiel 37

Wir werfen eine Münze so lange, bis zum ersten Mal „Kopf“ erscheint. Dies geschehe in jedem Wurf unabhängig mit Wahrscheinlichkeit p. Wir definieren dazu die Zufallsvariable $X :=$ „Anzahl der Würfe“. Wir haben bereits gesehen, dass

$$\Pr[X = k] = p(1-p)^{k-1}$$

und damit

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} k \cdot p(1-p)^{k-1} = p \cdot \frac{1}{(1-(1-p))^2} = \frac{1}{p}\,.$$

Beispiel 37

Andere Berechnungsmethode: (gestützt auf Satz 36)
Definiere das Ereignis

$$K_1 := \text{„Im ersten Wurf fällt Kopf"}\,.$$

Offensichtlich gilt $\mathbb{E}[X|K_1] = 1$.
Nehmen wir nun an, dass im ersten Wurf *nicht* „Kopf" gefallen ist.
Wir starten das Experiment neu.

Beispiel 37

Sei X' die Anzahl der Würfe bis zum ersten Auftreten von „Kopf" im neu gestarteten Experiment. Wegen der Gleichheit der Experimente gilt $\mathbb{E}[X'] = \mathbb{E}[X]$. Damit schließen wir

$$\mathbb{E}[X|\bar{K}_1] = 1 + \mathbb{E}[X'] = 1 + \mathbb{E}[X]$$

und erhalten mit Satz 18:

$$\begin{aligned}\mathbb{E}[X] &= \mathbb{E}[X|K_1] \cdot \Pr[K_1] + \mathbb{E}[X|\bar{K}_1] \cdot \Pr[\bar{K}_1] \\ &= 1 \cdot p + (1 + \mathbb{E}[X]) \cdot (1 - p)\,.\end{aligned}$$

Daraus ergibt sich wiederum $\mathbb{E}[X] = 1/p$.

4.2.2 Varianz

Wir betrachten die beiden folgenden Zufallsexperimente:

1. Wir würfeln (mit einem fairen Würfel), bei gerader Augenzahl erhalten wir 1 Euro, bei ungerader Augenzahl müssen wir 1 Euro bezahlen.
2. Wir würfeln (mit einem fairen Würfel), bei 6 Augen erhalten wir 5 Euro, ansonsten müssen wir 1 Euro bezahlen.

Beobachtung:
In beiden Fällen ist der erwartete Gewinn $= 0$.

Dennoch sind die „Schwankungen“ im ersten Fall geringer als im zweiten.

Eine nahe liegende Lösung wäre,

$$\mathbb{E}[|X - \mu|]$$

zu berechnen, wobei $\mu = \mathbb{E}[X]$ sei. Dies scheitert jedoch meist an der „unhandlichen“ Betragsfunktion. Aus diesem Grund betrachtet man stattdessen $\mathbb{E}[(X - \mu)^2]$, also die quadratische Abweichung vom Erwartungswert.

Definition 38

Für eine Zufallsvariable X mit $\mu = \mathbb{E}[X]$ definieren wir die *Varianz* $\mathrm{Var}[X]$ durch

$$\mathrm{Var}[X] := \mathbb{E}[(X - \mu)^2] = \sum_{x \in W_X} (x - \mu)^2 \cdot \Pr[X = x]\,.$$

Die Größe $\sigma := \sqrt{\mathrm{Var}[X]}$ heißt *Standardabweichung* von X.

Satz 39

Für eine beliebige Zufallsvariable X gilt

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \, .$$

Beweis:

Sei $\mu := \mathbb{E}[X]$. Nach Definition gilt

$$\begin{aligned} \mathrm{Var}[X] &= \mathbb{E}[(X-\mu)^2] = \mathbb{E}[X^2 - 2\mu \cdot X + \mu^2] \\ &= \mathbb{E}[X^2] - 2\mu \cdot \mathbb{E}[X] + \mu^2 \\ &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \, . \end{aligned}$$

□

Beispiel 40

1. Wir würfeln (mit einem fairen Würfel), bei gerader Augenzahl erhalten wir 1 Euro, bei ungerader Augenzahl müssen wir 1 Euro bezahlen. Es ist

$$\mu = 0 \text{ und } \mathrm{Var}[X] = \frac{1}{2} \cdot 1^2 + \frac{1}{2} \cdot (-1)^2 = 1\,.$$

2. Wir würfeln (mit einem fairen Würfel), bei 6 Augen erhalten wir 5 Euro, ansonsten müssen wir 1 Euro bezahlen.
Es ist

$$\mu = 0 \text{ und } \mathrm{Var}[X] = \frac{1}{6} \cdot 5^2 + \frac{5}{6} \cdot (-1)^2 = 5\,.$$

Satz 41

Für eine beliebige Zufallsvariable X und $a, b \in \mathbb{R}$ gilt

$$\mathrm{Var}[a \cdot X + b] = a^2 \cdot \mathrm{Var}[X]\,.$$

Beweis:

Aus der in Satz 33 gezeigten Linearität des Erwartungswerts folgt $\mathbb{E}[X+b] = \mathbb{E}[X] + b$.
Zusammen mit der Definition der Varianz ergibt sich damit sofort

$$\mathrm{Var}[X+b] = \mathbb{E}[(X+b-\mathbb{E}[X+b])^2] = \mathbb{E}[(X-\mathbb{E}[X])^2] = \mathrm{Var}[X]\,.$$

Weiter folgt mit Satz 39:

$$\mathrm{Var}[a\cdot X] = \mathbb{E}[(aX)^2] - \mathbb{E}[aX]^2 = a^2\mathbb{E}[X^2] - (a\mathbb{E}[X])^2 = a^2\cdot\mathrm{Var}[X]\,,$$

und daraus zusammen die Behauptung. □

Der Erwartungswert und die Varianz gehören zu den so genannten Momenten einer Zufallsvariablen:

Definition 42

Für eine Zufallsvariable X nennen wir $\mathbb{E}[X^k]$ das k-te Moment und $\mathbb{E}[(X - \mathbb{E}[X])^k]$ das k-te zentrale Moment.

Der Erwartungswert ist also identisch zum ersten Moment, während die Varianz dem zweiten zentralen Moment entspricht.

4.3 Mehrere Zufallsvariablen

Beispiel 43

Aus einem Skatblatt mit 32 Karten ziehen wir zufällig eine Hand von zehn Karten sowie einen Skat von zwei Karten. Unter den Karten gibt es vier Buben. Die Zufallsvariable X zählt die Anzahl der Buben in der Hand, während Y die Anzahl der Buben im Skat angibt. Die Werte von X und Y hängen offensichtlich stark voneinander ab. Beispielsweise muss $Y = 0$ sein, wenn $X = 4$ gilt.

Wie kann man mit mehreren Zufallsvariablen über demselben Wahrscheinlichkeitsraum rechnen, auch wenn sie, wie im obigen Beispiel, sehr voneinander abhängig sind?
Wir untersuchen Wahrscheinlichkeiten der Art

$$\Pr[X = x, Y = y] = \Pr[\{\omega;\ X(\omega) = x, Y(\omega) = y\}]\,.$$

Beispiel 44

Wenn wir nur die Zufallsvariable X betrachten, so gilt für $0 \leq x \leq 4$

$$\Pr[X = x] = \frac{\binom{4}{x}\binom{28}{10-x}}{\binom{32}{10}}\,.$$

Allgemein nennt man Zufallsvariablen mit der Dichte

$$\Pr[X = x] = \frac{\binom{b}{x}\binom{a}{r-x}}{\binom{a+b}{r}}$$

hypergeometrisch verteilt. Durch diese Dichte wird ein Experiment modelliert, bei dem r Elemente ohne Zurücklegen aus einer Grundmenge der Mächtigkeit $a+b$ mit b besonders ausgezeichneten Elementen gezogen werden.

Beispiel 44 (Forts.)

Die Zufallsvariable Y ist für sich gesehen ebenfalls hypergeometrisch verteilt mit $b = 4$, $a = 28$ und $r = 2$.
Für X und Y zusammen gilt jedoch z.B.

$$\Pr[X = 4, Y = 1] = 0,$$

und allgemein

$$\Pr[X = x, Y = y] = \frac{\binom{4}{x}\binom{28}{10-x}\binom{4-x}{y}\binom{28-(10-x)}{2-y}}{\binom{32}{10}\binom{22}{2}}\,.$$

Bemerkung: Die Schreibweise $\Pr[X = x, Y = y]$ stellt eine Abkürzung von $\Pr[„X = x \wedge Y = y“]$ dar. Ein anderes Beispiel ist

$$\Pr[X \leq x, Y \leq y_1, \sqrt{Y} = y_2]\,.$$

Die Funktion

$$f_{X,Y}(x,y) := \Pr[X = x, Y = y]$$

heißt gemeinsame Dichte der Zufallsvariablen X und Y.
Aus der gemeinsamen Dichte $f_{X,Y}$ kann man ableiten

$$f_X(x) = \sum_{y \in W_Y} f_{X,Y}(x,y) \quad \text{bzw.} \quad f_Y(y) = \sum_{x \in W_X} f_{X,Y}(x,y)\,.$$

Die Funktionen f_X und f_Y nennt man Randdichten.

Die Ereignisse „$Y = y$“ bilden eine Partitionierung des Wahrscheinlichkeitsraumes, und es gilt daher

$$\Pr[X = x] = \sum_{y \in W_Y} \Pr[X = x, Y = y] = f_X(x)\,.$$

Die Dichten der einzelnen Zufallsvariablen entsprechen also genau den Randdichten.

Für zwei Zufallsvariablen definiert man die gemeinsame Verteilung

$$\begin{aligned} F_{X,Y}(x,y) &= \Pr[X \le x, Y \le y] = \Pr[\{\omega;\, X(\omega) \le x, Y(\omega) \le y\}] \\ &= \sum_{x' \le x} \sum_{y' \le y} f_{X,Y}(x', y')\,. \end{aligned}$$

Die Randverteilung ergibt sich gemäß

$$F_X(x) = \sum_{x' \leq x} f_X(x') = \sum_{x' \leq x} \sum_{y \in W_Y} f_{X,Y}(x', y)$$

sowie

$$F_Y(y) = \sum_{y' \leq y} f_Y(y') = \sum_{y' \leq y} \sum_{x \in W_X} f_{X,Y}(x, y') \,.$$

4.3.1 Unabhängigkeit von Zufallsvariablen

Definition 45

Die Zufallsvariablen $X_1, \ldots, X_n$ heißen unabhängig, wenn für alle $(x_1, \ldots, x_n) \in W_{X_1} \times \ldots \times W_{X_n}$ gilt

$$\Pr[X_1 = x_1, \ldots, X_n = x_n] = \Pr[X_1 = x_1] \cdot \ldots \cdot \Pr[X_n = x_n] .$$

Alternativ:

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdot \ldots \cdot f_{X_n}(x_n) .$$

Bei unabhängigen Zufallsvariablen ist also die gemeinsame Dichte gleich dem Produkt der Randdichten. Ebenso gilt

$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = F_{X_1}(x_1) \cdot \ldots \cdot F_{X_n}(x_n) .$$

Satz 46

Seien $X_1, \ldots, X_n$ unabhängige Zufallsvariablen und $S_1, \ldots, S_n$ beliebige Mengen mit $S_i \subseteq W_{X_i}$. Dann sind die Ereignisse „$X_1 \in S_1$“, ..., „$X_n \in S_n$“ unabhängig.

Beweis:

$$
\begin{aligned}
&\Pr[X_1 \in S_1, \ldots, X_n \in S_n] \\
&= \sum_{x_1 \in S_1} \ldots \sum_{x_n \in S_n} \Pr[X_1 = x_1, \ldots, X_n = x_n] \\
&\stackrel{\text{Unabh.}}{=} \sum_{x_1 \in S_1} \ldots \sum_{x_n \in S_n} \Pr[X_1 = x_1] \cdot \ldots \cdot \Pr[X_n = x_n] \\
&= \left(\sum_{x_1 \in S_1} \Pr[X_1 = x_1] \right) \cdot \ldots \cdot \left(\sum_{x_n \in S_n} \Pr[X_n = x_n] \right) \\
&= \Pr[X_1 \in S_1] \cdot \ldots \cdot \Pr[X_n \in S_n] \,.
\end{aligned}
$$

Satz 47

$f_1, \ldots, f_n$ seien reellwertige Funktionen ($f_i : \mathbb{R} \to \mathbb{R}$ für $i = 1, \ldots, n$). Wenn die Zufallsvariablen $X_1, \ldots, X_n$ unabhängig sind, dann gilt dies auch für $f_1(X_1), \ldots, f_n(X_n)$.

Beweis:

Sei $z_i \in W_{f(X_i)}$ für $i = 1, \ldots, n$ und $S_i = \{x;\ f(x) = z_i\}$.

$$\begin{aligned}
&\Pr[f_1(X_1) = z_1, \ldots, f_n(X_n) = z_n] \\
&= \Pr[X_1 \in S_1, \ldots, X_n \in S_n] \\
&\overset{\text{Unabh.}}{=} \Pr[X_1 \in S_1] \cdot \ldots \cdot \Pr[X_n \in S_n] \\
&= \Pr[f_1(X_1) = z_1] \cdot \ldots \cdot \Pr[f_n(X_n) = z_n]\,.
\end{aligned}$$

□