Einige Konzepte aus der Wahrscheinlichkeitstheorie (Review)

Diskrete Zufallsvariablen (Random variables)

- Eine **Zufallsvariable** $X(c)$ ist eine Variable (genauer eine Funktion), deren Wert vom Ergebnis c eines Zufallsprozesses abhängt
- Eine diskrete Zufallsvariable X kann nur abzählbar viele Zustände $\{x_1, x_2, \ldots\}$ annehmen

Diskrete Zufallsvariablen (2)

- Eine **Wahrscheinlichkeitsverteilung** spezifiziert, mit welcher Wahrscheinlichkeit eine Zufallsvariable einen bestimmten Zustand einnimmt

- Eine Wahrscheinlichkeitsverteilung von X kann durch eine **Wahrscheinlichkeitsfunktion** $f(x)$ definiert werden:

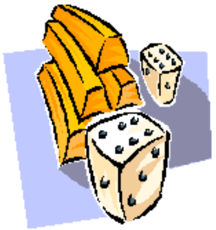
$$P(X = x) = P(\{c : X(c) = x\}) = f(x) \tag{1}$$

- $f(x)$ ist die Wahrscheinlichkeitsfunktion und x ist eine **Realisierung** von X. Typischerweise macht man sich keine Gedanken über den zugrundeliegenden Zufallsprozess und man definiert einfach eine Wahrscheinlichkeitsverteilung.

- Wir schreiben in der Regel

$$f(x) = P_X(x) = P(x) \tag{2}$$

Beispiel: 2 Würfe eines fairen Würfels: Elementarereignisse



6	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
1	1/36	1/36	1/36	1/36	1/36	1/36
2.Wurf 1. Wurf	1	2	3	4	5	6

- Elementarereignis: Ergebnis zweier Würfe
- $c = \{c_1, c_2\}$: erster Wurf ist c_1, der zweite ist c_2
- Die Wahrscheinlichkeit jedes Elementarereignisses ist $1/36$

Beispiel: Zufallsvariable X

6	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
1	1/36	1/36	1/36	1/36	1/36	1/36
2.Wurf 1. Wurf	1	2	3	4	5	6

$X = 0$ falls $c_1 \in \{1, 2, 3\}$ und $X = 1$ anderenfalls

$$P(X = 0) = P(\{c : X(c) = 0\}) =$$

$$1/36 \times 3 \times 6 = 1/2 = f(0)$$

$$P(X = 1) = P(\{c : X(c) = 1\}) =$$

$$1/36 \times 3 \times 6 = 1/2 = f(1)$$

Beispiel: Zufallsvariable Y

6	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
1	1/36	1/36	1/36	1/36	1/36	1/36
2.Wurf 1.Wurf	1	2	3	4	5	6

- $Y = 0$ falls $c_2 \in \{1,2,3,4\}$ und $Y = 1$ anderenfalls

  $$P(Y=0) = P(\{c : Y(c) = 0\}) =$$

  $$1/36 \times 4 \times 6 = 2/3 = f(0)$$

  $$P(Y=1) = P(\{c : Y(c) = 1\}) =$$

  $$1/36 \times 2 \times 6 = 1/3 = f(1)$$

Beispiel: Zufallsvariable Z

6	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
1	1/36	1/36	1/36	1/36	1/36	1/36
2.Wurf 1. Wurf	1	2	3	4	5	6

- $Z = 0$ falls $c_1 + c_2 \leq 5$ und $Z = 1$ anderenfalls

  $$P(Z = 0) = P(\{c : Z(c) = 0\}) =$$

  $$1/36 \times 10 = 10/36 = f(0)$$

  $$P(Z = 1) = P(\{c : Z(c) = 1\}) =$$

  $$1/36 \times 26 = 26/36 = f(1)$$

Multivariate Verteilungen

- Man definiere zwei Zufallsvariablen $X(c)$ und $Y(c)$. Eine multivariate Verteilung ist definiert durch:

$$P(x,y) = P(X=x, Y=y) = P(X=x \wedge Y=y) \tag{3}$$

$$= P(\{c : X(c) = x \wedge Y(c) = y\}) \tag{4}$$

Beispiel: Multivariate Verteilung X, Y

$Y = 1$

6	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
1	1/36	1/36	1/36	1/36	1/36	1/36
2.Wurf / 1.Wurf	1	2	3	4	5	6

$X = 1$

$$P(X = 1, Y = 1)$$

$$= P(X = 1 \wedge Y = 1)$$

$$= P(\{c : X(c) = 1 \wedge Y(c) = 1\})$$

$$1/36 \times 6 = 1/6$$

1	1/6	1/6
0	1/3	1/3
Y/X	0	1

Beispiel: Multivariate Verteilung X, Z

$Z = 1$

6	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
1	1/36	1/36	1/36	1/36	1/36	1/36
	1	2	3	4	5	6

$X = 1$

$$P(X = 1, Z = 1)$$

$$= P(X = 1 \wedge Z = 1)$$

$$= P(\{c : X(c) = 1 \wedge Z(c) = 1\})$$

$$= 17/36$$

1	9/36	17/36
0	9/36	1/36
Z/X	0	1

Bedingte Verteilungen

- *Definition* einer **bedingten Verteilung**

$$P(Y = y|X = x) = \frac{P(X = x, Y = y)}{P(X = x)} \quad \text{mit} \quad P(X = x) > 0$$

Beispiel: Bedingte Verteilung $P(Y|X)$

$Y=1$

6	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
1	1/36	1/36	1/36	1/36	1/36	1/36
2.Wurf 1.Wurf	1	2	3	4	5	6

$X=1$

$$P(Y=1|X=1)=\frac{P(X=1,Y=1)}{P(X=1)}$$

$$=\frac{1/6}{1/2}=1/3$$

1	1/6	1/6
0	1/3	1/3
Y/X	0	1

Beispiel: Bedingte Verteilung $P(Z|X)$

$Z = 1$

6	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
1	1/36	1/36	1/36	1/36	1/36	1/36
	1	2	3	4	5	6

$X = 1$

$$P(Z = 1|X = 1) = \frac{P(X = 1, Z = 1)}{P(X = 1)}$$

$$= \frac{17/36}{1/2} = 34/36$$

1	9/36	17/36
0	9/36	1/36
Z/X	0	1

Produktzerlegungen und Kettenregel

- Es folgen: **Produktsatz**

$$P(x, y) = P(x|y)P(y) = P(y|x)P(x) \tag{5}$$

- und die **Kettenregel**

$$P(x_1, \ldots, x_M) = P(x_1)P(x_2|x_1)P(x_3|x_1, x_2) \ldots P(x_M|x_1, \ldots, x_{M-1}) \tag{6}$$

Bayes'sche Regel

- **Bayes'sche Regel**

$$P(y|x) = \frac{P(x,y)}{P(x)} = \frac{P(x|y)P(y)}{P(x)} \quad P(x) > 0 \tag{7}$$

Beispiel: Bayes'sche Regel

$Z=1$

6	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
1	1/36	1/36	1/36	1/36	1/36	1/36 $X=1$
	1	2	3	4	5	6

$$P(X=1|Z=1) = \frac{P(Z=1|X=1)P(X=1)}{P(Z=1)}$$

$$= \frac{34/36 \times 18/36}{26/36} = 17/26$$

1	9/36	17/36
0	9/36	1/36
Z/X	0	1

Marginale Verteilungen

- Da die Ereignisse sich ausschließen,

$$P(X = x) = P(X = x, [Y = y_1 \vee Y = y_2]) = \tag{8}$$

$$P(\{c : X(c) = x \wedge [Y(c) = y_1 \vee Y(c) = y_2]\}) = P(x, y_1) + P(x, y_2) \tag{9}$$

- Daraus folgt, dass sich die **marginale (Rand-) Verteilungen** berechnen läßt zu:

$$P(x) = \sum_y P(x, y) \tag{10}$$

Beispiel: Marginale Verteilung

$$P(X = 1) = P(Z = 0, X = 1) + P(Z = 1, X = 1)$$

$$= 1/36 + 17/36 = 1/2$$

$Z = 1$

6	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
1	1/36	1/36	1/36	1/36	1/36	1/36 $X = 1$
	1	2	3	4	5	6

P(Z=1) =(9+17)/36 =26/36	9/36	17/36
P(Z=0) =(9+1)/36= 10/36	9/36	1/36
Z/X	P(X=0) =(9+9)/36 =1/2	P(X=1) =(17+1)/36 =1/2

Unabängige Zufallsvariablen

- **Unabhängigkeit**: zwei Zufallsvariablen sind unabhängig, falls gilt,

$$P(x, y) = P(x)P(y|x) = P(x)P(y) \tag{11}$$

Beispiel: unabhängige Würfelwürfe.

Beispiel: 2 Würfe eines Würfels

$Y=1$

6	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
1	1/36	1/36	1/36	1/36	1/36	1/36
2.Wurf 1.Wurf	1	2	3	4	5	6

$X=1$

$$P(Y=1, X=1) = P(X=1)P(Y=1)$$

$$= 1/2 \times 2/6 = 1/6$$

1	1/6	1/6
0	1/3	1/3
Y/X	0	1

Erwartungswerte

- **Erwartungswert**

$$E(X) = E_{P(x)}(X) = \sum_i x_i P(X = x_i) \tag{12}$$

Beispiel: Erwartungswerte

6	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
1	1/36	1/36	1/36	1/36	1/36	1/36
2.Wurf 1. Wurf	1	2	3	4	5	6

$$E(X) = 0 \times 1/2 + 1 \times 1/2 = 1/2$$

$$E(Y) = 0 \times 2/3 + 1 \times 1/3 = 1/3$$

$$E(Z) = 0 \times 10/36 + 1 \times 26/36 = 26/36$$

Varianzen

- Definition der **Varianz** einer Zufallsvariable:

$$var(X) = \sum_i (x_i - E(X))^2 P(X = x_i) \qquad (13)$$

Beispiel: Varianzen

6	1/36	1/36	1/36	1/36	1/36	1/36
5	1/36	1/36	1/36	1/36	1/36	1/36
4	1/36	1/36	1/36	1/36	1/36	1/36
3	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	1/36	1/36	1/36	1/36	1/36
1	1/36	1/36	1/36	1/36	1/36	1/36
2.Wurf 1. Wurf	1	2	3	4	5	6

$$var(X) = (0 - 1/2)^2 \times 1/2 + (1 - 1/2)^2 \times 1/2 = 1/4$$

$$var(Y) = (0 - 1/3)^2 \times 2/3 + (1 - 1/3)^2 \times 1/3 = 6/27$$

$$var(Z) = (0 - 26/36)^2 \times 10/36 + (1 - 26/36)^2 \times 26/36 \approx 0.20$$

Kovarianzen

- **Kovarianz**:

$$cov(X,Y) = \sum_i \sum_j (x_i - E(X))(y_j - E(Y))P(X = x_i, Y = y_j) \quad (14)$$

- **Kovarianzmatrix**:

$$\Sigma_{[XY],[XY]} = \begin{pmatrix} var(X) & cov(X,Y) \\ cov(Y,X) & var(Y) \end{pmatrix} \quad (15)$$

- Falls $cov(X,Y) = 0$, sind X und Y unkorreliert (aber nicht notwendigerweise unabhängig). Aus Unabhängigkeit folgt Unkorreliertheit. Das Inverse gilt nicht!

- Beachte, das gilt:

$$cov(X,Y) = E(XY) - E(X)E(Y)$$

wobei man $E(XY)$ als Korrelation bezeichnet. Beachte jedoch, dass der Korrelationskoeffizient definiert ist als

$$r = \frac{cov(X,Y)}{\sqrt{var(X)}\sqrt{var(Y)}}$$

Beispiel: Kovarianzen

1	1/6	1/6
0	1/3	1/3
Y/X	0	1

1	9/36	17/36
0	9/36	1/36
Z/X	0	1

$$cov(X,Y) = (0-1/2)(0-1/3)\times 1/3 + (0-1/2)(1-1/3)\times 1/6$$

$$+(1-1/2)(0-1/3)\times 1/3 + (1-1/2)(1-1/3)\times 1/6 = 0$$

$$cov(X,Z) = (0-1/2)(0-26/36)\times 9/36 + (0-1/2)(1-26/36)\times 1/36$$

$$+(1-1/2)(0-26/36)\times 9/36 + (1-1/2)(1-26/36)\times 17/36 = 0.0617$$

Kontinuierliche (stetige) Zufallsvariablen

- Zufallsvariablen können auch kontinuierlich (stetig) sein. Hier definiert man die **Wahrscheinlichkeitsdichte** als

$$f(x) = \lim_{\Delta x \to 0} \frac{P(x \le X \le x + \Delta x)}{\Delta x}$$

- Man erhält

$$P(a < x < b) = \int_a^b f(x) dx$$

- Die Verteilungsfunktion ist

$$F(x) = \int_{-\infty}^{x} f(x) dx = P(X \le x)$$

Wir werden in der Notation keinen Unterschied machen zwischen einer Wahrscheinlichkeitsfunktion für diskrete Variablen und Wahrscheinlichkeitsdichten für stetige Variablen und schreiben ebenfalls

$$f(x) = P(x) \tag{16}$$

Erwartungswerte für stetige Variablen

- Erwartungswert

$$E(X) = E_{P(x)}(X) = \int xP(x)dx$$

- Varianz

$$var(X) = \int (x - E(x))^2 P(x)dx \tag{17}$$

- Kovarianz:

$$cov(X, Y) = \int (x - E(X))(y - E(Y))P(x, y)dxdy \tag{18}$$

Statistik

- Bisher hatten wir angenommen, dass z.B. $P(X = x)$ bekannt ist
- In der **frequentistischen Statistik** gibt es unbekannte Parameter, die geschätzt werden: $P(X = 1) = \theta$. Wenn θ der wahre unbekannte Parameter ist, so ist $\hat{\theta}$ die Schätzung: wenn ich genügend viele Würfelergebnisse beobachte, so ist z.B. $\hat{\theta} = N_{true}/N$ eine vernünftige Schätzung, wobei N die Anzahl der Würfe ist und N_{true} die Anzahl der Würfe, in der X den Wert 1 annimmt
- In der **Bayes'schen Statistik** erweitert man das Wahrscheinlichkeitsmodell um eine Verteilung über die Parameter $P(\theta)$ und wendet die Regeln der Wahrscheinlichkeitslehre an; die W.-Verteilung ist dann $P(\theta) \prod_{i=1}^{N} P(X = x_i|\theta)$ wobei $x_i \in \{0, 1\}$ der Wert ist, den X im i-ten Wurf annimmt