Wahrscheinlichkeitstheorie

X endliche Menge, $p : X \to \mathbb{R}^{\geq 0}$, $\sum_{x \in X} p(x) = 1$.

1. p heißt Wahrscheinlichkeitsverteilung.
2. (X, p) heißt Wahrscheinlichkeitsraum, kann als zufälliges Experiment aufgefaßt werden.
3. $A \subseteq X$ heißt Ereignis. Die Wahrscheinlichkeit von A ist $\Pr(A) = \sum_{x \in A} p(x)$.
4. Komplementärereignis von A: $\overline{A} = X \setminus A$.
5. $0 \leq \Pr(A) \leq 1$, $\Pr(\{\}) = 0$, $\Pr(X) = 1$, $\Pr(\overline{A}) = 1 - \Pr(A)$.
6. $A, B \subseteq X$: $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$.
7. $A \subseteq B \Rightarrow \Pr(A) \leq \Pr(B)$.
8. Gleichverteilung: $p(x) = 1/\#X$, $\Pr(A) = \#A/\#X$.

Wahrscheinlichkeitstheorie

(X, p) Wahrscheinlichkeitsraum, $A, B \subseteq X$, $\Pr(B) > 0$.

1. Bedingte Wahrscheinlichkeit: $\Pr(A \mid B) := \Pr(A \cap B) / \Pr(B)$.
2. A und B heißen unabhängig, wenn $\Pr(A \cap B) = \Pr(A)\Pr(B)$.
3. Satz von Bayes: $\Pr(A \mid B) = \Pr(A)\Pr(B \mid A) / \Pr(B)$.

Y Menge, $f : X \to Y$.
Dann heißt f heißt Y-wertige Zufallsvariable auf X.

1. Ereignis $f = y$ ist $f^{-1}(\{y\}) = \{x \in X \mid f(x) = y\}$ für $y \in Y$.
   Ereignis $f \in A$ ist $f^{-1}(A) = \{x \in X \mid f(x) \in A\}$ für $A \subseteq Y$.
2. $\Pr(f = y) = \Pr(f^{-1}(\{y\})$, $\Pr(f \in A) = \Pr(f^{-1}(A))$.
3. Erwartungswert („Durchschnittswert“) von f für $Y \subseteq \mathbb{R}$:
   $E(f) := \sum_{x \in X} f(x)p(x) = \sum_{y \in Y} y \Pr(f = y)$.

Wahrscheinlichkeitstheorie

f, g zwei $\mathbb{R}$-wertige ZV auf X.

1. $f + g$ durch $(f + g)(x) := f(x) + g(x)$ definiert, etc.
2. $f \leq g$ genau dann, wenn $f(x) \leq g(x)$ für alle x, etc.

Prop 1:
1. $E(f + g) = E(f) + E(g)$, $E(af) = aE(f)$ für $a \in \mathbb{R}$.
2. $f \leq g$ impliziert $E(f) \leq E(g)$.

Wahrscheinlichkeitstheorie

Konstruktionen mit W-Räumen und Z-Variablen.

(X, p), $f_i : X \to Y_i$, $1 \leq i \leq n$.

1. Ereignis $f_1 = y_1, \ldots, f_n = y_n$ ist $\{x \in X \mid f_i(x) = y_i \text{ für alle } i\}$.
   Dann $\Pr(f_1 = y_1, \ldots, f_n = y_n) = \Pr(\{x \in X \mid f_i(x) = y_i \text{ für alle } i\})$.
2. f_i heißen unabhängig, wenn $\Pr(f_1 = y_1, \ldots, f_n = y_n) = \prod_i \Pr(f_i = y_i)$.

Prop 2: Für $Y_i \subseteq \mathbb{R}$ und unabhängige f_i gilt $E(\prod_{i=1}^n f_i) = \prod_{i=1}^n E(f_i)$.

Bew:

$$E(\prod_i f_i) = \sum_{x \in X} f_1(x) \cdots f_n(x) p(x) = \sum_{y_1 \in Y_1, \ldots, y_n \in Y_n} y_1 \cdots y_n \Pr(f_1 = y_1, \ldots, f_n = y_n)$$
$$= \prod_{i=1}^n \sum_{y_i \in Y_i} y_i \Pr(f_i = y_i) = \prod_{i=1}^n E(f_i).$$

Wahrscheinlichkeitstheorie

(X_i, p_i), $f_i : X_i \to Y_i$, $1 \leq i \leq n$.

1. $X_1 \times \cdots \times X_n$ wird durch $p((x_1, \ldots, x_n)) = \prod_i p_i(x_i)$ zum Wahrscheinlichkeitsraum.
2. $\Pr(f_i = y_i) = \Pr(f_i \circ \pi_i = y_i)$, wobei π_i die Projektion $X_1 \times \cdots \times X_n \to X_i$ ist.

$X^n = X \times \cdots \times X$ entspricht n-facher, unabhängiger Ausführung des Experiments X.

Wahrscheinlichkeitstheorie

Konstruktion und Notation aufeinanderfolgender (parametrisierter) Experimente.

(X, p), $(W_x, p_x)_{x \in X}$ W-Räume.

Dann $XW := \cup_{x \in X} \{x\} \times W_x$, $p_{XW}(x, w) := p(x) p_x(w)$.

$f : X \to Y$, $g : XW \to Y$ Zufallsvariablen.

1. $\Pr(f(x) = y : x \leftarrow X) := \Pr(f = y)$.
2. $\Pr(g(x, w) = y : x \leftarrow X, w \leftarrow W_x) := \Pr(g = y)$.

Wahrscheinlichkeitstheorie

Prop 3: $A, A_i \subseteq X$, $\Pr(A_i) > 0$, so daß X disjunkte Vereinigung der A_i ist. Dann $\Pr(A) = \sum_i \Pr(A_i) \Pr(A \mid A_i)$.

Bew: $\Pr(A) = \sum_i \Pr(A \cap A_i) = \sum_i \Pr(A_i) \Pr(A \mid A_i)$. $\square$

Einfaches Anwendungsbeispiel:

Wir wählen zufällig und gleichverteilt eine Zahl a aus $\{1 \ldots 6\}$. Falls wir $a = 1$ oder $a = 2$ erhalten, wählen wir zufällig und gleichverteilt eine Zahl b aus $\{1, 2\}$. Falls wir aber $a > 2$ erhalten, wählen wir zufällig und gleichverteilt b aus $\{1, \ldots, 15\}$.

Mit welcher Wahrscheinlichkeit erhalten wir $b = 1$?

Die Wahrscheinlichkeit ist $2/6 \cdot 1/2 + 4/6 \cdot 1/15 = 19/90$.

Wahrscheinlichkeitstheorie

Prop 4: $A \subseteq X$, $\Pr(A) > 0$ und $r \in \mathbb{Z}^{\geq 0}$. Sei f_r die Zufallsvariable, welche angibt, wie häufig X ausgeführt werden muß, um A genau r mal zu erhalten. Dann gilt $E(f_r) = r / \Pr(A)$.

Prop 5 (Markov Ungleichung): Sei f eine Zufallsvariable und $r > 0$. Dann $\Pr(f \geq r) \leq E(|f|)/r$ bzw. $\Pr(f \geq rE(|f|)) \leq 1/r$.

Einfaches Anwendungsbeispiel: Bei 11mal Würfeln erhalten wir mindestens eine 6 mit Wahrscheinlichkeit $\geq 1/2$.

Wahrscheinlichkeitstheorie

Bew Prop 4: Es gilt $f_r = w$, wenn die letzte und davor beliebige $r-1$ von insgesamt w Ausführungen von X das Ereignis A ergeben. Dies tritt folglich mit einer Wahrscheinlichkeit von

$$\Pr(f_r = w) = \binom{w-1}{r-1} \Pr(A)^r \,(1-\Pr(A))^{w-r}$$

ein. Für den Erwartungswert von f_r ergibt sich damit

$$\begin{aligned}
E(f_r) &= \sum_{w=0}^{\infty} w \Pr(f_r = w) = \sum_{w=r}^{\infty} w \binom{w-1}{r-1} \Pr(A)^r \,(1-\Pr(A))^{w-r} \\
&= \frac{\Pr(A)^r}{(r-1)!} \sum_{w=r}^{\infty} w(w-1)\cdots(w-r+1)\,(1-\Pr(A))^{w-r} \\
&= \frac{\Pr(A)^r}{(r-1)!} \cdot \frac{d^r}{dx^r}(1-x)^{-1}\big|_{x=(1-\Pr(A))} = \frac{\Pr(A)^r}{(r-1)!} \cdot \frac{r!}{(1-(1-\Pr(A)))^{r+1}} \\
&= r/\Pr(A). \quad \square
\end{aligned}$$

Wahrscheinlichkeitstheorie

Bew Prop 5: Sei g die durch

$$g = \begin{cases} 1 & \text{für } f \geq r \\ 0 & \text{für } f < r. \end{cases}$$

definierte Zufallsvariable. Dann gilt $|f| \geq rg$ und

$$E(|f|) \geq E(rg) = rE(g) = r\Pr(f \geq r). \quad \square$$