Beiblätter zur Vorlesung Mathematik für Biologen

Erwartungswert und Varianz:

Erwartungswert

Sei X eine reelle Zufallsvariable. Der Erwartungswert von X ist

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

falls Ω abzählbar ist.

Ist X eine diskret verteilte ZV (d.h. es existiert eine Folge $(x_i)_{i\in\mathbb{N}}$ von Punkten in $\mathbb{R}$ mit $P(X = x_i) = p_i$ für $i \in \mathbb{N}$, wobei $0 \le p_i \le 1$ und $\sum\limits_{i\in\mathbb{N}} p_i = 1$ ist), so ist der Erwartungswert gegeben durch

$$E(X) = \sum_{i\in\mathbb{N}} x_i\, P(X = x_i) = \sum_{i\in\mathbb{N}} x_i\, p_i.$$

Für $\varphi : \mathbb{R} \to \mathbb{R}$ ist dann $E(\varphi \circ X) = \sum\limits_{i\in\mathbb{N}} \varphi(x_i)\, p_i$.

Ist X absolutstetig verteilt mit Dichtefunktion f (d.h. es ist $P(X \le a) = \int_{-\infty}^{a} f(x)\, dx$ für alle $a \in \mathbb{R}$), so ist

$$E(X) = \int x f(x)\, dx.$$

Für $\varphi : \mathbb{R} \to \mathbb{R}$ ist dann $E(\varphi \circ X) = \int \varphi(x) f(x)\, dx$.

Rechenregeln für den Erwartungswert

Für $c \in \mathbb{R}$ ist $E(cX) = cE(X)$.
Für Zufallsvariablen $X_1, \dots, X_n$ ist

$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i).$$

Für das arithmetische Mittel gilt somit

$$E(\overline{X}_n) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i).$$

Für unabhängige Zufallsvariable X und Y ist $E(XY) = E(X)\, E(Y)$.

Varianz und Kovarianz

Die Varianz $Var(X)$ einer reellen Zufallsvariable X ist definiert durch

$$Var(X) = E\left((X - E(X))^2\right) = E(X^2) - E(X)^2.$$

Für zwei Zufallsvariablen X und Y ist die Kovarianz $Cov(X,Y)$ definiert durch

$$Cov(X,Y) = E\left((X - E(X))(Y - E(Y))\right) = E(XY) - E(X)E(Y).$$

Der Quotient $\rho_{X,Y} = \dfrac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$ heißt der Korrelationskoeffizient der Zufallsvariablen X und Y.
Gilt $Cov(X,Y) = 0$, also auch $\rho_{X,Y} = 0$, so heißen die beiden Zufallsvariablen unkorreliert.
Sind X und Y unabhängig, dann sind sie unkorreliert.

Rechenregeln für die Varianz

Für $a \in \mathbb{R}$ ist

$$Var(X + a) = Var(X).$$

Für $c \in \mathbb{R}$ ist

$$Var(cX) = c^2 Var(X).$$

Für die Varianz einer Summe von zwei Zufallsvariablen X, Y gilt:

$$Var(X + Y) = Var(X) + Var(Y) + 2\,Cov(X,Y).$$

Sind X und Y unkorrelierte Zufallsvariable, d.h. ist $Cov(X,Y) = 0$, so folgt

$$Var(X + Y) = Var(X) + Var(Y).$$

Sind $X_1, X_2, \ldots, X_n$ paarweise unkorrelierte Zufallsvariable (d.h. gilt $Cov(X_i, X_j) = 0$ für $i \neq j$), so ist

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) \quad \text{und} \quad Var(\overline{X}_n) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i).$$