WuS - Lecture Notes Week 6

Ruben Schenk, ruben.schenk@inf.ethz.ch

June 22, 2022

0.1 Ungleichungen

0.1.1 Monotonie

Satz: Seien X, Y zwei Z.V., sodass

$$X \leq Y \; f.s.$$

gilt. Falls beide Erwartungswerte wohldefiniert sind foglt dann

$$\mathbb{E}[X] \leq \mathbb{E}[Y]. \; f.s.$$

0.1.2 Markov Ungleichung

Theorem (Markov-Ungleichung): Sei X eine nicht-negative Z.V. Für jedes $a > 0$ gilt dann

$$\mathbb{P}[X \geq a] \leq \frac{\mathbb{E}[X]}{a}.$$

0.1.3 Jensen Ungleichung

Theorem (Jensen Ungleichung): Sei X eine Z.V. Sei $\phi : \mathbb{R} \to \mathbb{R}$ eine konvexe Funktion. Falls $\mathbb{E}[\phi(x)]$ und $\mathbb{E}[X]$ wohldefiniert sind, gilt

$$\phi(\mathbb{E}[X]) \leq \mathbb{E}[\phi(X)].$$

Daraus folgt mit $\phi(x) = |x|$, dass $|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$ (Dreiecksungleichung) und mit $\phi(x) = x^2$, dass $\mathbb{E}[|X|] \leq \sqrt{\mathbb{E}[X^2]}$.

0.2 Varianz

Def: Sei X eine Zufallsvariable, sodass $\mathbb{E}[X]^2 < \infty$. Wir definieren die **Varianz von** X durch

$$\sigma_X^2 = \mathbb{E}[(X - m)^2], \text{ wobei } m = \mathbb{E}[X].$$

Die Wurzel aus σ_X^2 nennen wir gerade die **Standardabweichung von** X.

Die Standardabweichung ist ein Indikator für die Fluktuation von X um den *Mittelwert* $m = \mathbb{E}[X]$ herum. Allgemein ist eine Zufallsvariable mit geringer Varianz konzentriert um ihren Erwartungswert $m = \mathbb{E}[X]$. Die Tschebyscheffsche Ungleichung formalisiert diese Beobachtung.

Satz: Sei X eine Z.V. mit $\mathbb{E}[X^2] < \infty$. Dann gilt für jedes $a \geq 0$

$$\mathbb{P}[|X - m| \geq a] \leq \frac{\sigma_X^2}{a^2}, \text{ wobei } m = \mathbb{E}[X].$$

Satz (Grundlegende Eigenschaften der Varianz):

1. Sei X eine Z.V. mit $\mathbb{E}[X^2] < \infty$. Dann gilt
$$\sigma_X^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$
2. Sei X eine Z.V. mit $\mathbb{E}[X^2] < \infty$ und sei $\lambda \in \mathbb{R}$. Dann gilt
$$\sigma_{\lambda X}^2 = \lambda^2 \cdot \sigma_X^2.$$
3. Seien $X_1, ..., X_n$ n-viele paarweise unabhängige Z.V. und $S = X_1 + \cdots + X_n$. Dann gilt
$$\sigma_S^2 = \sigma_{X_1}^2 + \cdots + \sigma_{X_n}^2.$$

Anwendung: Sei S eine binomialverteilte Z.V. mit Parametern n und p. Was ist die Varianz von S? Wir benutzen, dass S die selbe Verteilung wie $S_n = X_1 + \cdots X_n$, mit $X_1, ..., X_n$ u.i.v. Bernoulli Z.V. mit Parameter p hat. Dann erhalten wir:

$$\sigma_S^2 = \sigma_{S_n}^2 = \sigma_{X_1}^2 + \cdots + \sigma_{X_n}^2 = n \cdot \sigma_{X_1}^2.$$

Zudem gilt $\sigma_{X_1}^2 = \mathbb{E}[X_i^2] - p^2 = p - 2 = p(1-p)$. Durch Einsetzen erhalten wir:

$$\sigma_S^2 = n \cdot p(1-p).$$

Im Allgemeinen erhalten wir für Summen von u.i.v. Z.V. stets $\mathbb{E}[S] = n \cdot p$ und $\sigma_S = \sqrt{n} \cdot \sqrt{p(1-p)}$.

0.3 Kovarianz

Def: Seien X, Y zwei Z.V. mit endlichen zweiten Momenten $\mathbb{E}[X^2] < \infty$ und $\mathbb{E}[Y^2] < \infty$. Wir definieren die **Kovarianz zwischen X und Y** durch

$$\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

Die Kovarianz verschwindet wenn X und Y unabhängig sind, somit gilt:

$$X, Y \text{ unabhängig} \implies \mathrm{Cov}(X, Y) = 0.$$

Achtung: Die umgekehrte Implikation ist falsch!

1 Gemeinsame Verteilungen

1.1 Gemeinsame diskrete Verteilungen

1.1.1 Definition

Def: Seien $X_1, ..., X_n$ n diskrete Zufallsvariablen, sei $W_i \subset \mathbb{R}$ endlich oder abzählbar, wobei $X_i \in W_i$ fast sicher gilt. Die gemeinsame Verteilung von $(X_1, ..., X_n)$ ist eine Familie $p = (p(x_1, ..., x_n))_{x_1 \in W_1, ..., x_n \in W_n}$, wobei jedes Mitglied definiert ist durch:

$$p(x_1, ..., x_n) = \mathbb{P}[X_1 = x_1, ..., X_n = x_n].$$

Beispiel: Seien X, Y zwei unabhängige Bernoulli Z.V. mit Parameter $1/2$. Die gemeinsame Verteilung von $(X,\, Y)$ ist gegeben durch

$$\forall x, y \in \{0,\, 1\} \quad p(x,\, y) = \frac{1}{4}.$$

Die gemeinsame Verteilung von $(X,\, X)$ ist gegeben durch

$$\forall x, y \in \{0,\, 1\} \quad p(x,\, y) = \begin{cases} \frac{1}{2}, & x = y \\ 0, & x \neq y. \end{cases}$$

Satz: Eine gemeinsame Verteilung von Z.V. $X_1, \ldots, X_n$ erfüllt

$$\sum_{x_1 \in W_1, \ldots, x_n \in W_n} p(x_1, \ldots, x_n) = 1.$$

Bemerkung: Man kann auch **Gewichtsfunktion** anstatt Verteilung sagen.