Wiederholung

- Erwartungswert

$$\mathbb{E}[X] \;=\; \sum_{x \in W_X} x \cdot \Pr[X = x] \;=\; \sum_{\omega \in \Omega} X(\omega) \cdot \Pr[\omega].$$

- Varianz

$$\mathrm{Var}[X] := \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Indikatorvariable

Eine Zufallsvariable, die nur die Werte 0 und 1 annimmt, nennt man *Indikatorvariable* des Ereignisses A, wobei $A = \{\omega \in \Omega \mid X(\omega) = 1\}$.

Für eine Indikatorvariable I_A des Ereignisses A gilt:

$$\mathbb{E}[I_A] = 0 \cdot \Pr[X = 1] + 1 \cdot \Pr[X = 1] = \Pr[A]$$

Eigenschaften von Zufallsvariablen

- Linearität des Erwartungswertes

  Für Zufallsvariablen $X_1, \ldots, X_n$ und $X := X_1 + \ldots + X_n$ gilt

$$\mathbb{E}[X] = \mathbb{E}[X_1] + \ldots + \mathbb{E}[X_n].$$

- Linearität der Varianz

  Für **unabhängige** Zufallsvariablen $X_1, \ldots, X_n$ und $X := X_1 + \ldots + X_n$ gilt

$$\mathrm{Var}[X] = \mathrm{Var}[X_1] + \ldots + \mathrm{Var}[X_n].$$

- Multiplikativität des Erwartungswertes

  Für **unabhängige** Zufallsvariablen $X_1, \ldots, X_n$ und $X := X_1 \cdot \ldots \cdot X_n$ gilt

$$\mathbb{E}[X] = \mathbb{E}[X_1] \cdot \ldots \cdot \mathbb{E}[X_n].$$

BEWEISE: Nachrechnen durch Einsetzen der Definition und elementare Umformungen.

BSP. Ein Beispiel aus dem Alltag

Szenario:

In der Verwaltung einer großen Universität sollen die Immatrikulationsunterlagen (Semesterbögen) für n Studenten verschickt werden. Diese sind bereits gedruckt. Auch n Umschläge wurden schon mit den Anschriften der Studenten versehen. Beim „Eintüten“ passiert allerdings ein kleines Malheur. Der zuständige Mitarbeiter leidet unter einem Hitzschlag und nimmt die Zuordnung vor ohne auf die Namen zu achten. (D.h., jede der Zuordnungen von Semesterbögen zu Umschlägen sei gleich wahrscheinlich).

Frage:

Wieviele Studenten erhalten im Mittel ihren eigenen Bogen?

Lösung:

Wir bezeichnen mit X die Anzahl der Studenten, die ihren eigenen Bogen erhalten. Dann definieren wir geeignete 0-1 Zufallsvariable X_i so daß gilt $X = X_1 + \ldots + X_n$ und nutzen die Linearität des Erwartungswertes aus.

Genauer: Setze

$$X_i := \begin{cases} 1 & \text{falls Student } i \text{ seinen eigenen Bogen erhält} \\ 0 & \text{sonst} \end{cases}$$

Dann gilt

$$X := X_1 + \ldots + X_n$$

und

$$\Pr[X_i = 1] = \frac{1}{n}$$

(da alle Zuordnungen gleichwahrscheinlich sind)

Daraus folgt

$$\mathbb{E}[X_i] = 0 \cdot \Pr[X_i = 0] \; + \; 1 \cdot \Pr[X_i = 1] = \frac{1}{n}$$

und somit

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \sum_{i=1}^{n} \frac{1}{n} = 1$$

BEISPIEL: Average Case Analyse des QuickSort-Algorithmus

Randomisierten Quicksort: Wähle das Pivot-Element in jedem Schritt zufällig und gleichwahrscheinlich unter allen Elementen des Arrays.

Satz: Der randomisierte Quicksort-Algorithmus benötigt zum Sortieren von n Elementen im Mittel $\mathcal{O}(n \log n)$ Vergleiche.

BEWEIS:

$X :=$ „Anzahl der von Quicksort durchgeführten Vergleiche“

Für die folgende Analyse betrachten wir eine beliebige Eingabe S. Mit $S_{(i)}$ bezeichnen wir das i-te Element von S bezüglich der *sortierten* Reihenfolge.

Idee: Stelle X als Summe von Indikatorvariablen dar:

$$X := \sum_{i=1}^{n} \sum_{j=i+1}^{n} X_{ij}$$

$X_{ij} = 1$ gdw. $S_{(i)}$ mit $S_{(j)}$ verglichen wird.

Dann gilt:

$$\mathbb{E}[X] = \sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{E}[X_{ij}] = \sum_{i=1}^{n} \sum_{j=i+1}^{n} p_{ij}.$$

wobei $p_{ij} = \Pr[x_{ij} = 1]$.

Bernoulli-Verteilung

Eine Zufallsvariable X mit $W_X = \{0, 1\}$ und der Dichte

$$f_X(x) = \begin{cases} p & \text{für } x = 1, \\ 1 - p & \text{für } x = 0. \end{cases}$$

heißt *Bernoulli-verteilt*.

Den Parameter p nennt man *Erfolgswahrscheinlichkeit*.

Die Komplementwahrscheinlichkeit $1 - p$ kürzt man meist mit $q := 1 - p$ ab.

Für eine Bernoulli-verteilte Zufallsvariable gilt:

$$\mathbb{E}[X] = p$$

und

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

Binomialverteilung

Ist eine Zufallsvariable $X := X_1 + \ldots + X_n$ als Summe von n unabhängigen, Bernoulli-verteilten Zufallsvariablen mit gleicher Erfolgswahrscheinlichkeit p definiert, dann heißt X *binomialverteilt* mit den Parametern n und p.

Anschaulich: Eine binomialverteilte Zufallsvariable mit Parametern n und p zählt die Anzahl „Erfolge" bei der Durchführung von n unabhängigen Bernoulli-Experimenten mit Erfolgswahrscheinlichkeit p.

Es gilt:

$$
\begin{aligned}
W_X &= \{0, \ldots, n\} \\
f_X(x) &= \binom{n}{x} p^x q^{n-x}, \qquad \text{wobei } q := 1 - p \\
\mathbb{E}[X] &= np \\
\mathrm{Var}[X] &= npq
\end{aligned}
$$

Geometrische Verteilung

Wiederholt man ein Bernoulli-Experiment mit Erfolgswahrscheinlichkeit p solange, bis Erfolg eintritt, so ist die Anzahl der insgesamt durchgeführten Experimente *geometrisch verteilt* mit Parameter p.

Es gilt:

$$\begin{aligned} W_X &= \mathbb{N} \\ f_X(i) &= pq^{i-1} \quad \text{für } i \in \mathbb{N} \\ \mathbb{E}[X] &= \frac{1}{p} \\ \mathrm{Var}[X] &= \frac{q}{p^2} \end{aligned}$$

Wichtige Eigenschaft: Die geometrissche Verteilung ist *gedächtnislos*, d.h. es gilt

$$\Pr[X > y + x \mid X > x] = \Pr[X > y] \qquad \text{für alle } x, y \in \mathbb{N}$$

BEISPIEL: **Warten auf den n-ten Erfolg**

Wir wieederholen ein Bernoulli-Experiment mit Erfolgswahrscheinlichkeit p solange, bis der n-te Erfolg eintritt. Wie ist die Anzahl Z der insgesamt durchgeführten Experimente verteilt?

$$f_Z(z) = \binom{z-1}{n-1} \cdot p^n (1-p)^{z-n}.$$

Die Zufallsvariable Z nennt man *negativ binomialverteilt* mit Ordnung n.

BEISPIEL: **Das Coupon-Collector-Problem**

Wenn es insgesamt n verschiedene Abziehbilder gibt, von denen jeweils eines einem Produkt beiliegt, wie viele Produkte muss man im Mittel erwerben, bis man eine vollständige Sammlung besitzt? Hierbei nehmen wir an, dass bei jedem Kauf die Abziehbilder unabhängig voneinander mit gleicher Wahrscheinlichkeit auftreten.

$X :=$ Anzahl der Käufe bis zur Komplettierung der Sammlung.

Weiter teilen wir die Zeit in Phasen ein:
Phase i bezeichne die Schritte vom Erwerb des $(i-1)$-ten Abziehbildes (ausschließlich) bis zum Erwerb des i-ten Abziehbildes (einschließlich).

$X_i :=$ Anzahl der Käufe in Phase i.

Dann gilt:

$$X = \sum_{i=1}^{n} X_i \qquad \text{und} \qquad \mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i]$$

Poisson-Verteilung

Eine Poisson-Verteilung mit dem Parameter λ hat den Wertebereich $W_X = \mathbb{N}_0$ und besitzt die Dichte

$$f_X(i) = \frac{e^{-\lambda}\lambda^i}{i!} \quad \text{für } i \in \mathbb{N}_0 .$$

Es gilt:

$$\begin{aligned} \mathbb{E}[X] &= \lambda \\ \mathrm{Var}[X] &= \lambda \end{aligned}$$

Wichtige Eigenschaft: Die Verteilung einer binomialverteilten Zufallsvariablen mit Parametern n und $p_n = \lambda/n$ nähert sich für $n \to \infty$ der Poisson-Verteilung mit Parameter λ an.