Vorlesung 28.5

Peter Nejjar

Hier wird der Inhalt der Vorlesung vom 28.5 wiedergegeben. Insofern sich die Vorlesung an [1] orientierte, werden die Inhalte anhand der dortigen Bezeichnungen/Nummern nur kurz genannt.

Kapitel 3.3

Wir berechnen die Varianz einiger Zufallsvariablen bzw. geben diese an: Vorab stellen wir die Beobachtung an, dass wegen $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$ der Erwartungswert einer Zufallsvariable X existieren muss, damit die Varianz existiert. Insbesondere gilt:

$$\mathbb{V}(X) < \infty \Rightarrow \mathbb{E}(X) < \infty, \tag{1}$$

wobei $\Rightarrow$ die logische Implikation darstellt, und $\mathbb{V}(X) < \infty$ (bzw. $\mathbb{E}(X) < \infty$) eine Kurzschreibweise dafür ist, dass Varianz (bzw. Erwartungswert) existieren. Umgekehrt gilt das nicht, d.h. aus $\mathbb{E}(X) < \infty$ folgt nicht, dass $\mathbb{V}(X) < \infty$.

Beispiele zur Varianz:

1.) Varianz der Gleichverteilung auf $\Omega = \{1, \ldots, n\}$:

Es sei $\Omega = \{1, \ldots, n\}, \mathbb{P}(\{\omega\}) = 1/n$, und $X : \Omega \to \Omega, X(\omega) = \omega$.

Um die Varianz zu berechnen, nutzen wir aus, dass für eine Zufallsvariable $Y : \Omega \to C$, mit $\Omega = \{1, \ldots, n\}$, gilt (siehe Aufgabe 1 des 6. Zettels bzw. der Ausdruck (4) vom Skript vom 21.5)

$$\mathbb{E}(Y) = \sum_{c \in C} \mathbb{P}(Y = c) * c = \sum_{\omega \in \Omega} Y(\omega)\mathbb{P}(\{\omega\}) = \sum_{i=1}^{n} Y(i)\mathbb{P}(\{i\}).$$

Wir wenden diese Formel mit $Y(i) = (X(i) - \mathbb{E}(X))^2 = \left(i - \frac{n+1}{2}\right)^2$ an. Es ist

$$\begin{aligned}\sum_{i=1}^{n} Y(i)\mathbb{P}(\{i\}) = \sum_{i=1}^{n} \left(i - \frac{n+1}{2}\right)^2 \frac{1}{n} &= \sum_{i=1}^{n} \left(i^2 - (n+1) * i + \frac{(n+1)^2}{4}\right) \frac{1}{n} \\ &= \frac{1}{n} \sum_{i=1}^{n} \left(\frac{n(n+1)(2n+1)}{6} - \frac{n(n+1)^2}{2} + \frac{n(n+1)^2}{4}\right) \\ &= \frac{n^2 - 1}{12}.\end{aligned}$$

(In der Vorlesung wurde direkt mit $\sum_{c \in C} \mathbb{P}(Y = c) * c$ gerechnet).

2.) Varianz der Bernoulliverteilung auf $\Omega = \{0, 1\}$: Es sei $\mathbb{P}(\{1\}) = p \in [0, 1]$ und $X : \Omega \to \Omega, X(\omega) = \omega$. Es ist $X^2 = X, \mathbb{E}(X) = p$, daher ist $\mathbb{V}(X) = p - p^2$.

3.) Varianz der geometrischen Verteilung : Sei X eine Zufallsvariable, sodass $\mathbb{P}_X$ die geomtrische Verteilung mit Parameter q ist. Wir hatten in der Vorlesung bewiesen, dass $\mathbb{E}(X) = 1/(1 - q)$ gilt, und wir verwenden, dass $\mathbb{V}(X) = \frac{q}{(1-q)^2}$ ist.

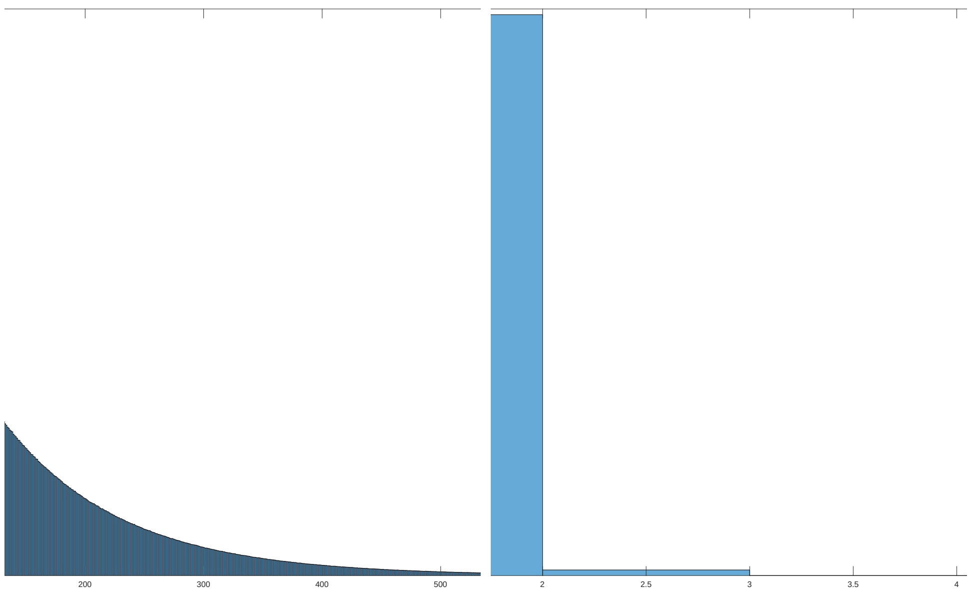
4.) Varianz der Normalverteilung : Sei X eine Zufallsvariable, sodass $\mathbb{P}_X$ auf $\mathbb{R}$ die Dichtefunktion

$$f_{\mu,\sigma^2}(x) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{\sigma^2}\right) \tag{2}$$

hat. Das ist gerade die Normalverteilung mit Parametern $\mu \in \mathbb{R}$ und $\sigma^2 > 0$. Wir verwenden ohne Beweis, dass $\mathbb{E}(X) = \mu, \mathbb{V}(X) = \sigma^2$ gilt.

Anschließend wurden für $q = 0.01, q = 0.99$ Histogramme geplottet von 100 Millionen Pseudozufallszahlen mit (Pseudo) geometrischer Verteilung mit Parametern $q = 0.01, q = 0.99$. Da bei $q = 0.01$ die Varianz "klein" (genau: $0.01/(1-0.01)^2$) ist, zeigte sich, dass die Pseudozufallszahlen sehr nah am Erwartungswert 1/0.99 lagen (siehe den rechten Plot unten) , was daran liegt, dass wir die Wahrscheinlichkeit einer Abweichung vom Erwartungswert mit der Tchebycheff Ungleichung nach oben beschränken können.

Für $q = 0.99$ hingegen "streuen" die Zufallszahlen um den Erwartungswert 100, d.h. es traten durchaus Werte nahe der 1 und über 400 auf, siehe den linken Plot:



Wichtig: Natürlich kann man in diesem Fall, die Wahrscheinlichkeit, dass $\mathbb{P}(|X - \mathbb{E}(X)| > K)$ für beliebige K einfach exakt angeben, ohne jede Simulation. Die Simulation verdeutlicht aber das Prinzip, dass bei "kleiner "Varianz, die Datenpunkte nah beim Erwartungswert bleiben (die mathematisch präzise Aussage ist die Ungleichung von Tchebycheff), während bei großer Varianz dies nicht der Fall sein muss.

Anschließend wurde Dichtefunktion der Normalverteilung mit $\mu = 0$ und Varianzen 1 und 100 gezeigt. Es zeigt sich, dass die Normalverteilung für $\sigma^2 = 1$ eine deutlich größere Wahrscheinlichkeit hat, nah am Erwartungswert 0 zu bleiben als im Fall $\sigma^2 = 100$: zur Illustration wurde für $\sigma^2 = 1$ die Wahrscheinlichkeit $\int_{-1}^{1} \mathrm{d}x f_{0,1}(x) \approx 0.68$ und für $\sigma^2 = 100$ die Wahrscheinlichkeit $\int_{-1}^{1} \mathrm{d}x f_{0,100}(x) \approx 0.0079787$ numerisch berechnet (für die Definition von f_{μ,σ^2} siehe (2)).

Kapitel 4 - Bedingte Wahrscheinlichkeiten

Der Erwartungswert einer Zufallsvariablen X wurde mehrfach in der Vorlesung als 'Durchschnittswert' bezeichnet: Sind $X_1, \dots, X_n$ mit n "groß" und $\mathbb{P}_{X_i} = \mathbb{P}_X, i = 1, \dots, n$ könnte man meinen, dass

$$\frac{\sum_{i=1}^n X_i}{n} \approx \mathbb{E}(X) \tag{3}$$

gelten soll. Abgesehen davon, dass komplett unklar ist, was das $\approx$ hier überhaupt bedeuten soll, ist jedoch nicht zu erwarten, dass $\frac{\sum_{i=1}^n X_i}{n}$ in der Nähe von $\mathbb{E}(X)$ liegt, wenn wir bloß fordern, dass $\mathbb{P}_{X_i} = \mathbb{P}_X$ für alle i gilt: Wählen wir z.B. $X_i = X$ für alle i, ist $\frac{\sum_{i=1}^n X_i}{n} = X$, und wir wissen von Aufgabe 3 , Zettel 7, dass man für jedes noch so große K eine Zufallsvariable X finden kann, sodass $\mathbb{P}(|X - \mathbb{E}(X)| > K) = 1$ gilt. Damit ein Ergebnis der Art (3) gelten kann, sollten die Werte der X_i noch zusätzlich "unabhängig" sein: Werfen wir n Mal eine Münze, erwarten wir im Schnitt 50% Kopf , eben weil die Ergebnisse der verschiedenen Münzwürfe "unabhängig" sind.

Den Begriff der Unabhängigkeit zu präzisieren und zu untersuchen ist der Inhalt des folgenden Kapitels. Die vage Aussage (3) wird uns mathematisch präzise als *Gesetz der großen Zahlen* begegnen.

Kapitel 4.1

Als Beispiel betrachten wir $\Omega = \{1, \dots, 6\}^2$ mit der Gleichverteilung, $A = \{(i,j) : i + j \geq 7\}$ (Augensumme bei zwei Mal Würfeln größer gleich 7) und $B = \{(6,j) : j \in \{1, \dots, 6\}\}$ (im ersten Wurf kam eine 6). Was sollte die Wahrscheinlichkeit von A bedingt auf B sein ? Nun, wir wissen, dass wenn B eintritt, A eintreten muss, daher sollte diese Wahrscheinlichkeit 1 sein. Betrachten wir hingegen $C = \{(5,j) : j \in \{1, \dots, 6\}\}$, so sollte die Wahrscheinlichkeit von A bedingt auf B gleich 5/6 sein, denn sobald beim zweiten Wurf keine 1 kommt, wissen wir, dass A eintritt. Es ist

$$\frac{5}{6} = \frac{\mathbb{P}(A \cap C)}{\mathbb{P}(C)} = \frac{5/6 * 1/6}{1/6}. \tag{4}$$

Wir definieren daher die bedingte Wahrscheinlichkeit wie in Definition 4.1.1.

Wir definieren dann zwei Ereignisse A, B als unabhängig wie in Definition 4.1.3. Wir definieren zwei Zufallsvariablen X, Y als unabhängig wenn die Bedingung von Lemma 4.4.1 (i) erfüllt ist. Das ist die Definition 4.4.2, nur dass wir die Äquivalenz der Bedingungen Lemma 4.4.1 (i),(ii) nicht besprochen haben.

Nun wollen wir (s.o.) beliebige Folgen unabhängiger Zufallsvariablen haben, nicht nur zwei. Für $\Omega = \{1, \dots, 6\}^n$ mit der Gleichverteilung kann man nachweisen, dass mit $X_i : \Omega \to \{1, \dots, 6\}, X_i(\omega) = \omega_i$ (Projektion auf die i-te Koordinate) gilt, dass X_i, X_j unabhängig sind und dass die $X_i, i = 1, \dots, n$ als endliche Familie von Zufallsvariablen unabhängig sind (wobei wir noch nicht definiert haben, was es heißen soll, dass mehr als zwei Zufallsvariablen unabhängig sind). Allerdings können wir nicht einfach $n = \infty$ setzen, um eine unendliche Folge unabhängiger Zufallsvariablen zu haben: es gibt keine Gleichverteilung auf unendlichen Mengen wie etwa $\Omega = \{1, \dots, 6\}^\infty$. Ein tiefliegender Satz der Maßtheorie sagt aber aus, dass es eine $\sigma-$Algebra $\mathcal{E}$ und ein Wahrscheinlichkeitsmaß $\mathbb{P}$ auf $\mathcal{E}$ gibt, sodass für jedes $r \in \mathbb{N}$ und beliebige $A_i \subseteq \{1, \dots, 6\}$ gilt, dass **erstens**

$$\{\omega \in \{1, \dots, 6\}^\infty : \bigcap_{i=1}^r \{\omega_i \in A_i\}\} \in \mathcal{E}.$$

Indem wir manche der A_i als $A_i = \{1, \ldots, 6\}$ wählen, folgt dass für beliebige $j_1, \ldots, j_k \in \mathbb{N}$ und $A_{j_i} \subseteq \{1, \ldots, 6\}, i = 1, \ldots, k$ gilt:

$$\{\omega \in \{1, \ldots, 6\}^\infty : \bigcap_{i=1}^{k} \{\omega_{j_i} \in A_{j_i}\}\} \in \mathcal{E}.$$

Mengen der Form $\{\omega \in \{1, \ldots, 6\}^\infty : \cap_{i=1}^{k} \{\omega_{j_i} \in A_{j_i}\}\}$ heißen Zylindermengen. **Zweitens** hat das Wahrscheinlichkeitsmaß $\mathbb{P}$ die Eigenschaft, dass für Zylindermengen gilt

$$\mathbb{P}(\{\omega \in \{1, \ldots, 6\}^\infty : \cap_{i=1}^{r} \{\omega_i \in A_i\}\}) = \prod_{i=1}^{r} \frac{|A_i|}{6}. \quad (5)$$

Man bedenke, dass wenn $\mathbb{P}_n$ die Gleichverteilung auf $\{1, \ldots, 6\}^n$ ist, gilt (es sei $r \leq n$):

$$\mathbb{P}_n(\{\omega \in \{1, \ldots, 6\}^n : \cap_{i=1}^{r} \{\omega_i \in A_i\}\}) = \prod_{i=1}^{r} \frac{|A_i|}{6}. \quad (6)$$

In diesem Sinne ist $\mathbb{P}$ das Analogon von $\mathbb{P}_n$ für $n = \infty$.

Literatur

[1] E. Behrends. Elementare Stochastik. *Vieweg+Teubner Verlag, 2013,* `https://link.springer.com/book/10.1007/978-3-8348-2331-1`.