week3b: Kapitel 2.2: Der zentrale Grenzwertsatz

Wir betrachten eine Folge von unabhängigen und identisch verteilten Zufallsvariablen

$$X_1\,,\,X_2\,,\,\cdots\,,\,X_n$$

mit Mittelwert μ und Varianz σ^2,

$$\begin{array}{rcll} \mathsf{E}[X_i] & = & \mu & \forall i \\ \mathsf{V}[X_i] & = & \sigma^2 & \forall i \end{array}$$

Die X_i können diskrete Zufallszahlen sein, etwa mit Werten in den natürlichen Zahlen, oder auch kontinuierliche Zufallszahlen, mit Werten in $\mathbb{R}$. Machen wir kurz zwei Beispiele, auf die der zentralen Grenzwertsatz dann angewendet werden kann, bevor wir den Satz formulieren:

Beispiel 2.2.2: a) Diskrete Verteilung: Eine Zufallsvariable oder Zufallszahl X heisst Poisson verteilt mit Parameter λ, wenn X nur Werte in den natürlichen Zahlen (mit Null) annehmen kann, und das mit den folgenden Wahrscheinlichkeiten:

$$\mathsf{Prob}[\,X = k\,] \;=\; \tfrac{\lambda^k}{k!}\,e^{-\lambda} \;=:\; p_\lambda(k)\,, \qquad k = 0,1,2,\cdots$$

Der Erwartungswert von X ist dann gegeben durch

$$\mathsf{E}[X] \;=\; \sum_{k=1}^{\infty} k \times \mathsf{Prob}[X=k] \;=\; \sum_{k=1}^{\infty} k \times p_\lambda(k) \;\overset{\text{UeBlatt3}}{=}\; \lambda$$

Und für die Varianz findet man

$$\begin{aligned} \mathsf{V}[X] \;&=\; \mathsf{E}\big[\,(X-\mathsf{E}[X])^2\,\big] \;=\; \mathsf{E}[X^2] \,-\, \mathsf{E}[X]^2 \\ &=\; \sum_{k=1}^{\infty} k^2 \times \mathsf{Prob}[X=k] \,-\, \lambda^2 \;=\; \sum_{k=1}^{\infty} k^2 \times p_\lambda(k) \,-\, \lambda^2 \;\overset{\text{UeBlatt3}}{=}\; \lambda\,. \end{aligned}$$

b) Kontinuierliche Verteilung: Eine Zufallsvariable oder Zufallszahl X heisst auf dem Intervall $[a,b]$ gleichverteilt oder uniform verteilt, wenn X jede reelle Zahl zwischen a und b mit gleicher Wahrscheinlichkeit annehmen kann[1],

$$\mathsf{Prob}\big[\,X \in [x, x+dx)\,\big] \;=\; \tfrac{1}{b-a}\,\chi(a \le x \le b)\,dx \;=:\; p_{a,b}(x)\,dx$$

[1] χ ist die Indikator-Funktion gegeben durch $\chi(A) := \begin{cases} 1 & \text{falls } A = true \\ 0 & \text{falls } A = false \end{cases}$.

Der Erwartungswert von X ist dann gegeben durch

$$\mathsf{E}[X] \;=\; \int_{\mathbb{R}} x \times \mathsf{Prob}\big[\,X \in [x, x+dx)\,\big] \;=\; \int_{\mathbb{R}} dx\; x\; p_{a,b}(x) \;=\; \tfrac{a+b}{2}$$

Und für die Varianz findet man

$$\begin{aligned}
\mathsf{V}[X] \;&=\; \mathsf{E}\big[\,(X-\mathsf{E}[X])^2\,\big] \;=\; \mathsf{E}[X^2] \,-\, \mathsf{E}[X]^2 \\
&=\; \int_{\mathbb{R}} x^2 \times \mathsf{Prob}\big[\,X \in [x, x+dx)\,\big] \;-\; \tfrac{(a+b)^2}{4} \;=\; \int_{\mathbb{R}} dx\; x^2\; p_{a,b}(x) \;-\; \tfrac{(a+b)^2}{4} \\
&=\; \tfrac{1}{3}\,\tfrac{b^3-a^3}{b-a} \;-\; \tfrac{a^2+2ab+b^2}{4} \\
&=\; \tfrac{b^2+ab+a^2}{3} \;-\; \tfrac{a^2+2ab+b^2}{4} \;=\; \tfrac{4b^2+4ab+4a^2}{12} \;-\; \tfrac{3a^2+6ab+3b^2}{12} \\
&=\; \tfrac{b^2-2ab+a^2}{12} \;=\; \tfrac{(b-a)^2}{12} \;.
\end{aligned}$$

Bemerkung 2.2.3: Für den zentralen Grenzwertsatz ist es unerheblich, ob wir diskrete oder kontinuierliche Zufallsvariablen haben, wir werden nur die folgenden allgemeinen Eigenschaften von Erwartungswert und Varianz benutzen:

$$\mathsf{E}\big[\,X_1 + \cdots + X_n\,\big] \;=\; \mathsf{E}[X_1] \,+\, \cdots \,+\, \mathsf{E}[X_n] \;=\; n\,\mu$$

und

$$\begin{aligned}
\mathsf{V}\big[\,X_1 + \cdots + X_n\,\big] \;&=\; \mathsf{V}\big[\,\textstyle\sum_{k=1}^n X_k\,\big] \\
&=\; \mathsf{Cov}\big[\,\textstyle\sum_{k=1}^n X_k\,,\, \sum_{\ell=1}^n X_\ell\,\big] \\
&=\; \textstyle\sum_{k,\ell=1}^n \mathsf{Cov}\big[\,X_k\,,\, X_\ell\,\big] \\
&\overset{X_i \text{ unabhängig}}{=}\; \textstyle\sum_{k=1}^n \mathsf{Cov}\big[\,X_k\,,\, X_k\,\big] \\
&=\; \textstyle\sum_{k=1}^n \mathsf{V}\big[\,X_k\,\big] \\
&=\; n\,\sigma^2
\end{aligned}$$

Dabei erinnern wir uns noch einmal kurz an die Definition der Kovarianz von zwei Zufallsvariablen X und Y:

$$\mathsf{Cov}[X,Y] \;:=\; \mathsf{E}\big[\,(X-\mathsf{E}[X])(Y-\mathsf{E}[Y])\,\big]$$

Für unabhängige Zufallsvariablen X und Y gilt dann

$$\begin{aligned}
\mathsf{Cov}[X,Y] \;&=\; \mathsf{E}\big[\,(X-\mathsf{E}[X])(Y-\mathsf{E}[Y])\,\big] \\
&\overset{X \text{ und } Y \text{ unabhängig}}{=}\; \mathsf{E}\big[\,(X-\mathsf{E}[X])\,\big] \,\times\, \mathsf{E}\big[\,(Y-\mathsf{E}[Y])\,\big] \\
&=\; \big(\,\mathsf{E}[X]-\mathsf{E}[X]\,\big)\big(\,\mathsf{E}[Y]-\mathsf{E}[Y]\,\big) \;=\; 0 \;.
\end{aligned}$$

Für eine Folge von unabhängigen und identisch verteilten Zufallsvariablen

$$X_1\,,\,X_2\,,\,\cdots\,,\,X_n$$

mit Mittelwert μ und Varianz σ^2,

$$\begin{aligned}\mathsf{E}[X_i] &= \mu \quad \forall i\\ \mathsf{V}[X_i] &= \sigma^2 \quad \forall i\end{aligned}$$

definieren wir die Summe S_n und den Mittelwert mit M_n durch

$$\begin{aligned}S_n &:= \textstyle\sum_{k=1}^n X_k\\ M_n &:= \tfrac{1}{n}\textstyle\sum_{k=1}^n X_k\end{aligned}$$

Wir haben dann also

$$\begin{aligned}\mathsf{E}[S_n] &= n\,\mu\\ \mathsf{V}[S_n] &= n\,\sigma^2\end{aligned}$$

und

$$\begin{aligned}\mathsf{E}[M_n] &= \mathsf{E}\Big[\tfrac{1}{n}S_n\Big] = \tfrac{1}{n}\mathsf{E}\Big[S_n\Big] = \tfrac{n\,\mu}{n} = \mu\\ \mathsf{V}[M_n] &= \mathsf{V}\Big[\tfrac{1}{n}S_n\Big] = \tfrac{1}{n^2}\mathsf{V}\Big[S_n\Big] = \tfrac{n\,\sigma^2}{n^2} = \tfrac{\sigma^2}{n}\end{aligned}$$

Insbesondere,

$$\lim_{n\to\infty}\mathsf{V}[M_n] = 0$$

Mit der Tschebyscheff'schen Ungleichung

$$\mathsf{Prob}\Big[\,|\,M_n - \mathsf{E}[M_n]\,| > \varepsilon\,\Big] \le \frac{\mathsf{V}[M_n]}{\varepsilon^2} \overset{n\to\infty}{\to} 0$$

folgt dann das Gesetz der grossen Zahlen, das sagt gerade, dass die Mittelwerte M_n gegen das μ konvergieren müssen,

$$M_n \overset{n\to\infty}{\to} \mathsf{E}[M_n] = \mu\ .$$

Wir betrachten die normierten oder standardisierten Zufallsvariablen

$$\begin{aligned}Z_n &:= \frac{S_n - \mathsf{E}[S_n]}{\sqrt{\mathsf{V}[S_n]}}\\ \tilde{Z}_n &:= \frac{M_n - \mathsf{E}[M_n]}{\sqrt{\mathsf{V}[M_n]}}\end{aligned}$$

Wir haben dann

$$\begin{aligned}Z_n &= \frac{S_n - \mathsf{E}[S_n]}{\sqrt{\mathsf{V}[S_n]}}\\ &= \frac{S_n - n\mu}{\sqrt{n\sigma^2}}\\ &= \frac{\frac{1}{n}S_n - \mu}{\sqrt{\sigma^2/n}}\\ &= \frac{M_n - \mathsf{E}[M_n]}{\sqrt{\mathsf{V}[M_n]}} = \tilde{Z}_n\ ,\end{aligned}$$

die beiden Sachen sind also identisch, wir können beides mit Z_n bezeichnen und brauchen kein $\tilde{Z}_n$. Wir können das Z_n dann auch folgendermassen schreiben:

$$\begin{aligned} Z_n &= \frac{S_n - n\mu}{\sqrt{n\sigma^2}} && (1) \\ &= \frac{1}{\sqrt{n}} \frac{X_1 + \cdots + X_n \, - \, n\mu}{\sigma} \\ &= \frac{1}{\sqrt{n}} \left\{ \frac{X_1 - \mu}{\sigma} + \cdots + \frac{X_n - \mu}{\sigma} \right\} \\ &= \frac{Y_1 + \cdots + Y_n}{\sqrt{n}} && (2) \end{aligned}$$

mit der Abkürzung

$$Y_i := \frac{X_i - \mu}{\sigma} = \frac{X_i - \mathsf{E}[X_i]}{\sqrt{\mathsf{V}[X_i]}} \tag{3}$$

Die Y_i sind dann unabhängige, identisch verteilte Zufallsvariablen mit Erwartungswert 0 und Varianz 1,

$$\begin{aligned} \mathsf{E}[Y_i] &= 0 \quad \forall i \\ \mathsf{V}[Y_i] &= 1 \quad \forall i \end{aligned}$$

Wären die Y_i jetzt normalverteilt, dann wäre nach Theorem 2.2.1 vom letzten Mal das Z_n für jedes feste n ebenfalls eine standard-normalverteilte Zufallsvariable. Für beliebige Verteilungen ist das natürlich nicht mehr der Fall. Wenn man aber den Limes für grosse n betrachtet, bekommt man näherungsweise standard-normalverteilte Zufallsgrössen, das ist jetzt die Aussage des zentralen Grenzwertsatzes:

Theorem 2.2.4 (Zentraler Grenzwertsatz): Es seien $X_1, \cdots, X_n$ eine Folge von unabhängigen, identisch verteilten Zufallsvariablen mit Erwartungswert μ und Standardabweichung σ und die normierten oder standardisierten Zufallsgrössen Z_n seien gegeben durch die Gleichungen (1) oder (2,3) von oben. Dann gilt: Im Limes $n \to \infty$ sind die Z_n standard-normalverteilt. Das heisst genauer, für ein beliebiges $f : \mathbb{R} \to \mathbb{R}$ gilt:

$$\lim_{n\to\infty} \mathsf{E}\big[\, f(Z_n) \big] = \int_{\mathbb{R}} f(z)\, e^{-\frac{z^2}{2}} \, \tfrac{dz}{\sqrt{2\pi}} \tag{4}$$

Beweis: ..machen wir nächste Woche.

Mit Hilfe des zentralen Grenzwertsatzes lässt sich zum Beispiel die Stirlingsche Formel beweisen, das ist die folgende Aussage:

$$\lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n}\,\left(\frac{n}{e}\right)^n} \;=\; 1$$

Für grosse n kann man also Fakultäten approximieren durch

$$n! \;\approx\; \sqrt{2\pi n}\,\left(\frac{n}{e}\right)^n$$

Das schauen wir uns dann auf dem neuen Übungsblatt an.