week3a: Kapitel 2.2: Der zentrale Grenzwertsatz, Teil1

Wir wollen zunächst in dem folgenden Theorem beweisen, dass die Summe von n normalverteilten Zufallszahlen oder Zufallsvariablen wieder normalverteilt ist. Dazu brauchen wir das Integral aus dem Theorem 2.1.4, was wir letztes Mal noch nicht bewiesen hatten. Dann:

Theorem 2.2.1: Es seien $\phi_1\,,\,\phi_2\,,\,\cdots\,,\,\phi_n$ normalverteilte Zufallsvariablen mit Mittelwerten $\mu_1\,,\,\mu_2\,,\,\cdots\,,\,\mu_n$ und Standardabweichungen $\sigma_1\,,\,\sigma_2\,,\,\cdots\,,\,\sigma_n$ Dann gilt: Die Summe

$$\phi \;:=\; \phi_1 + \phi_2 + \cdots + \phi_n$$

ist normalverteilt mit Mittelwert

$$\mu \;=\; \mu_1 + \mu_2 + \cdots + \mu_n \tag{1}$$

und Varianz

$$\sigma^2 \;=\; \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2 \tag{2}$$

oder Standardabweichung

$$\sigma \;=\; \bigl\{\,\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2\,\bigr\}^{1/2}$$

Das heisst genauer, für eine beliebige Funktion $f:\mathbb{R}\to\mathbb{R}$ gilt

$$\int_{\mathbb{R}^n} f(\phi_1+\cdots+\phi_n)\,\prod_{k=1}^n e^{-\frac{(\phi_k-\mu_k)^2}{2\sigma_k^2}}\,\frac{d\phi_k}{\sqrt{2\pi\sigma_k^2}} \;=\; \int_{\mathbb{R}} f(\phi)\,e^{-\frac{(\phi-\mu)^2}{2\sigma^2}}\,\frac{d\phi}{\sqrt{2\pi\sigma^2}}\;. \tag{3}$$

mit μ und σ^2 gegeben durch (1) und (2).

Beweis: Wir substituieren

$$\begin{aligned} y_i \;&:=\; \phi_i - \mu_i \\ \Leftrightarrow \qquad \phi_i \;&=\; y_i + \mu_i \end{aligned}$$

und auf der rechten Seite von (3) $y := \phi - \mu$ oder $\phi = y + \mu$. Gleichung (3) ist dann äquivalent zu

$$\int_{\mathbb{R}^n} f(y_1+\cdots+y_n+\mu)\,\prod_{k=1}^n e^{-\frac{y_k^2}{2\sigma_k^2}}\,\frac{dy_k}{\sqrt{2\pi\sigma_k^2}} \;=\; \int_{\mathbb{R}} f(y+\mu)\,e^{-\frac{y^2}{2\sigma^2}}\,\frac{dy}{\sqrt{2\pi\sigma^2}}$$

oder, mit der Abkürzung $\tilde{f}(y) := f(y+\mu)$,

$$\int_{\mathbb{R}^n} \tilde{f}(y_1+\cdots+y_n)\ \prod_{k=1}^{n} e^{-\frac{y_k^2}{2\sigma_k^2}}\ \frac{dy_k}{\sqrt{2\pi\sigma_k^2}} \quad = \quad \int_{\mathbb{R}} \tilde{f}(y)\ e^{-\frac{y^2}{2\sigma^2}}\ \frac{dy}{\sqrt{2\pi\sigma^2}} \tag{4}$$

Die Gültigkeit von (4) können wir folgendermassen verifizieren: Wir substituieren

$$\begin{aligned}
x_1 &:= y_1 \\
x_2 &:= y_1+y_2 \quad = \quad x_1+y_2 \\
x_3 &:= y_1+y_2+y_3 \quad = \quad x_2+y_3 \\
&\vdots \\
x_n &:= y_1+y_2+\cdots+y_{n-1}+y_n \quad = \quad x_{n-1}+y_n
\end{aligned}$$

oder

$$\begin{aligned}
y_1 &= x_1 \\
y_2 &= x_2-x_1 \\
y_3 &= x_3-x_2 \\
&\vdots \\
y_n &= x_n-x_{n-1}
\end{aligned}$$

Wir haben dann

$$dy_1\cdots dy_n \quad = \quad \det\left[\tfrac{\partial y}{\partial x}\right]\, dx_1\cdots dx_n$$

mit der Funktionaldeterminante

$$\begin{aligned}
\det\left[\tfrac{\partial y}{\partial x}\right] &= \det\begin{pmatrix} - & \nabla_x y_1 & - \\ & \vdots & \\ - & \nabla_x y_n & - \end{pmatrix} \\
&= \det\begin{pmatrix} 1 & 0 & \cdots & 0 \\ -1 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} \\
&= 1
\end{aligned}$$

Damit bekommen wir für die linke Seite von (4)

$$\int_{\mathbb{R}^n} \tilde{f}(y_1+\cdots+y_n)\ \prod_{k=1}^{n} e^{-\frac{y_k^2}{2\sigma_k^2}}\ \frac{dy_k}{\sqrt{2\pi\sigma_k^2}} \quad = \quad \int_{\mathbb{R}^n} \tilde{f}(x_n)\ \prod_{k=1}^{n} e^{-\frac{(x_k-x_{k-1})^2}{2\sigma_k^2}}\ \frac{dx_k}{\sqrt{2\pi\sigma_k^2}} \tag{5}$$

wobei wir wegen $y_1 = x_1$ die Definition

$$x_0 \quad := \quad 0$$

machen, dann gilt die Gleichung $y_k = x_k - x_{k-1}$ auch für $k = 1$. Wir erinnern uns an die Notation aus dem Theorem 2.1.4 vom letzten Mal,

$$p_{\sigma_k^2}(x_k, x_{k-1}) \quad := \quad \frac{1}{\sqrt{2\pi\sigma_k^2}}\; e^{-\frac{(x_k - x_{k-1})^2}{2\sigma_k^2}}$$

und erhalten aus (5)

$$\begin{aligned}\int_{\mathbb{R}^n} \tilde{f}(y_1 + \cdots + y_n) \prod_{k=1}^n e^{-\frac{y_k^2}{2\sigma_k^2}} \frac{dy_k}{\sqrt{2\pi\sigma_k^2}} \quad &= \quad \int_{\mathbb{R}^n} \tilde{f}(x_n) \prod_{k=1}^n p_{\sigma_k^2}(x_k, x_{k-1})\, dx_k \\ = \quad \int_{\mathbb{R}} dx_n\, \tilde{f}(x_n) \int_{\mathbb{R}} dx_{n-1}\, p_{\sigma_n^2}(x_n, x_{n-1}) \times \cdots \times &\int_{\mathbb{R}} dx_1\, p_{\sigma_2^2}(x_2, x_1)\, p_{\sigma_1^2}(x_1, x_0) \qquad (6)\end{aligned}$$

Die Integrationsvariablen $x_1, x_2, \cdots, x_{n-1}$ auf der rechten Seite von (6) können wir jetzt mit Hilfe des Theorems 2.1.4 vom letzten Mal ausintegrieren. Wir haben etwa

$$\int_{\mathbb{R}} dx_1\, p_{\sigma_2^2}(x_2, x_1)\, p_{\sigma_1^2}(x_1, x_0) \quad = \quad p_{\sigma_2^2+\sigma_1^2}(x_2, x_0)$$

und bekommen dann

$$\begin{aligned}&\int_{\mathbb{R}^n} \tilde{f}(y_1 + \cdots + y_n) \prod_{k=1}^n e^{-\frac{y_k^2}{2\sigma_k^2}} \frac{dy_k}{\sqrt{2\pi\sigma_k^2}} \\ &\quad = \int_{\mathbb{R}} dx_n\, \tilde{f}(x_n) \int_{\mathbb{R}} dx_{n-1}\, p_{\sigma_n^2}(x_n, x_{n-1}) \times \cdots \times \int_{\mathbb{R}} dx_1\, p_{\sigma_2^2}(x_2, x_1)\, p_{\sigma_1^2}(x_1, x_0) \\ &\quad = \int_{\mathbb{R}} dx_n\, \tilde{f}(x_n) \int_{\mathbb{R}} dx_{n-1}\, p_{\sigma_n^2}(x_n, x_{n-1}) \times \cdots \times \int_{\mathbb{R}} dx_2\, p_{\sigma_3^2}(x_3, x_2)\, p_{\sigma_2^2+\sigma_1^2}(x_2, x_0) \\ &\quad = \int_{\mathbb{R}} dx_n\, \tilde{f}(x_n) \int_{\mathbb{R}} dx_{n-1}\, p_{\sigma_n^2}(x_n, x_{n-1}) \times \cdots \times \int_{\mathbb{R}} dx_3\, p_{\sigma_4^2}(x_4, x_3)\, p_{\sigma_3^2+\sigma_2^2+\sigma_1^2}(x_3, x_0) \\ &\quad \vdots \\ &\quad = \int_{\mathbb{R}} dx_n\, \tilde{f}(x_n) \int_{\mathbb{R}} dx_{n-1}\, p_{\sigma_n^2}(x_n, x_{n-1})\, p_{\sigma_{n-1}^2+\cdots+\sigma_2^2+\sigma_1^2}(x_{n-1}, x_0) \\ &\quad = \int_{\mathbb{R}} dx_n\, \tilde{f}(x_n)\, p_{\sigma_n^2+\sigma_{n-1}^2+\cdots+\sigma_2^2+\sigma_1^2}(x_n, x_0) \qquad (7)\end{aligned}$$

Das x_n können wir auch zu einem y umbenennen und wir erinnern uns an die Definition $x_0 = 0$ und die Definition von σ^2 als die Summe der σ_i^2. Dann haben wir also

$$\begin{aligned}\int_{\mathbb{R}^n} \tilde{f}(y_1 + \cdots + y_n) \prod_{k=1}^n e^{-\frac{y_k^2}{2\sigma_k^2}} \frac{dy_k}{\sqrt{2\pi\sigma_k^2}} \quad &= \quad \int_{\mathbb{R}} dx_n\, \tilde{f}(x_n)\, p_{\sigma_n^2+\sigma_{n-1}^2+\cdots+\sigma_2^2+\sigma_1^2}(x_n, x_0) \\ &= \quad \int_{\mathbb{R}} dy\, \tilde{f}(y)\, p_{\sigma^2}(y, 0) \\ &= \quad \int_{\mathbb{R}} \tilde{f}(y)\, e^{-\frac{y^2}{2\sigma^2}} \frac{dy}{\sqrt{2\pi\sigma^2}} \qquad (8)\end{aligned}$$

und das ist identisch mit der rechten Seite von (4). Damit ist das Theorem bewiesen. ■

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,

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

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 dann in der nächsten Veranstaltung 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 annehmen kann, und das mit den folgenden Wahrscheinlichkeiten:

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

Der Erwartungswert von X ist dann gegeben durch

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

Und für die Varianz findet man

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

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],

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

Der Erwartungswert von X ist dann gegeben durch

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

---

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

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

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