5 Der zentrale Grenzwertsatz

Satz 3.18 (Der zentrale Grenzwertsatz) *Es seien $X_1, \ldots, X_n$ ($n \in \mathbb{N}$) unabhängige, identisch verteilte zufällige Variablen mit*

$$\mu := \mathbf{E}X_i; \qquad \sigma^2 := \operatorname{Var} X_i.$$

Wir definieren für alle $n \in \mathbb{N}$ Zufallsgrößen Z_n, $\overline{Z}_n$ und Y_n durch: $Z_n := \sum_{i=1}^{n} X_i$ bzw. $\overline{Z}_n := \frac{Z_n}{n}$ und

$$Y_n = \sqrt{n} \cdot \frac{\overline{Z}_n - \mu}{\sigma}$$

Dann gilt:

$$\begin{aligned}\lim_{n\to\infty} P\left(\frac{Z_n - n\cdot\mu}{\sqrt{n}\cdot\sigma} < x\right) &= \lim_{n\to\infty} P\left(Y_n < x\right) = \Phi(x)\\ &= \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt.\end{aligned}$$

Beweis: (Als Hilfsmittel werden charakteristische Funktionen verwendet, siehe unten, für den interessierten Leser) □

Bem.: Die Folge $\{Y_n\}_{n\in\mathbb{N}}$ konvergiert in Verteilung gegen eine Zufallsgröße Z, $Y_n \longrightarrow^D Z$, $Z \sim \mathcal{N}(0,1)$.

Anwendungen:

Simulation bei der Erzeugung einer normalverteilten Zufallsgröße aus Pseudozufallszahlen

Approximation von Wkt.-verteilungen (insbes. von Teststatistiken)

Genauigkeitsabschätzung:

Satz 3.19 (Berry-Esséen) *Es seien die Voraussetzungen des zentralen Grenzwertsatzes erfüllt und* $M := \mathbf{E}|X_i - \mu|^3 < \infty$. *Dann gilt:*

$$\left| P\left(\frac{Z_n - n\cdot\mu}{\sqrt{n}\cdot\sigma}\right) - \Phi(x)\right| < K,$$

wobei $K = \frac{0{,}8\cdot M}{\sigma^3\cdot\sqrt{n}}$ *ist.*

Bsp. 3.8 *Es seien* $X_i \sim R(0,1)$,

$$\begin{aligned} \mu &= \mathbf{E}X_i = \tfrac{1}{2} \\ \sigma^2 &= \mathbf{E}X_i^2 - \mu^2 = \tfrac{1}{12} \end{aligned}$$

Wir bestimmen die Zahl M:

$$\begin{aligned} M &= \mathbf{E}|X_i-\mu|^3 \\ &= \int\limits_{-\infty}^{+\infty} |x-\mu|^3 \cdot f(x)\,dx \\ &= \int\limits_0^1 |x-\tfrac{1}{2}|^3\,dx \\ &= 2\cdot \int\limits_{\frac{1}{2}}^{1} (x-\tfrac{1}{2})^3\,dx \\ &= \tfrac{1}{32} \end{aligned}$$

Daraus ergeben sich beispielsweise folgende BERRY–ESSÉEN–*Schranken* K:

n	K
12	0.3
100	0.104
1000	0.033

Bsp. 3.9 *Seien* $X_i \sim Poi(\lambda)$,

$$\mathbf{E}X_i = \operatorname{Var} X_i = \lambda$$

Wir schätzen die Zahl M *ab:*

$$\begin{aligned} M^{\frac{1}{3}} &= \left(\mathbf{E}|X_i - \lambda|^3\right)^{\frac{1}{3}} \\ &\leq \left(\mathbf{E}|X_i - \lambda|^4\right)^{\frac{1}{4}} \quad \textit{Lemma 3.9)} \\ &= \left(\mathbf{E}(X_i - \lambda)^4\right)^{\frac{1}{4}} \\ &= (\lambda + 3\lambda^2)^{\frac{1}{4}} \end{aligned}$$

Berry-Esseen Schranke:

$$K \leq \frac{0.8(\lambda + 3\lambda^2)^{\frac{3}{4}}}{\lambda^{\frac{3}{2}}\sqrt{n}} \rightarrow_{\lambda\to\infty} \frac{0.8 \cdot 3^{\frac{3}{4}}}{\sqrt{n}}$$

Bsp. 3.10 *Seien* $X_i \sim \mathrm{B}(1,p)$, $i = 1, \ldots, n$, *unabhängig,*

$$X_i : \quad \begin{pmatrix} 0 & 1 \\ \\ 1-p & p \end{pmatrix}$$

- $\mathbf{E}X_i = \mu = p$;
- $\mathrm{Var}\, X_i = \sigma^2 = p(1-p)$.

Wir definieren nun für alle $n \in \mathbb{N}$ *eine Zufallsgröße*

$$Z_n := \sum_{i=1}^{n} X_i.$$

Die Zufallsgrößen Z_n *(*$n \in \mathbb{N}$*) haben also folgende Gestalt:*

$$Z_n : \quad \begin{pmatrix} 0 & 1 & 2 & \ldots & n \\ \\ p_0 & p_1 & p_2 & \ldots & p_n \end{pmatrix}$$

Wir zeigen jetzt: Für alle $n \in \mathbb{N}$ *gilt:* $Z_n \sim \mathrm{B}(n,p)$,*d.h.* $p_i = \binom{n}{i} p^i (1-p)^{n-i}$. *Beweis mittels vollständiger Induktion.*

IA: Es sei $n = 2$. Dann gilt:

$$Z_2 = X_1 + X_2 : \quad \begin{pmatrix} 0 & 1 & 2 \\ \\ p_0 & p_1 & p_2 \end{pmatrix}$$

Wir ermitteln die Wktn. p_0, p_1 und p_2:

$$\begin{aligned}
p_0 &= P(Z_2 = 0) \\
&= P(X_1 = 0, X_2 = 0) \\
&= P(X_1 = 0) \cdot P(X_2 = 0) \quad \text{(Unabh. von } X_1 \text{ und } X_2\text{)} \\
&= (1-p) \cdot (1-p) = (1-p)^2 = \binom{2}{0} p^0 (1-p)^{2-0} \\
p_1 &= P(Z_2 = 1) \\
&= P(\{X_1 = 1, X_2 = 0\} \cup \{X_1 = 0, X_2 = 1\}) \\
&= P(X_1 = 1, X_2 = 0) + P(X_1 = 0, X_2 = 1) \\
&\qquad \text{(Unvereinbarkeit der Ereignisse)} \\
&= P(X_1 = 1) \cdot P(X_2 = 0) + P(X_1 = 0) \cdot P(X_2 = 1) \\
&= p \cdot (1-p) + (1-p) \cdot p = \binom{2}{1} p^1 (1-p)^{2-1} \\
p_2 &= P(Z_2 = 2) = P(X_1 = 1, X_2 = 1) \\
&= P(X_1 = 1) \cdot P(X_2 = 1) = p^2 = \binom{2}{2} p^2 (1-p)^{2-2}
\end{aligned}$$

IS: ÜA

Satz 3.20 (MOIVRE–LAPLACE**)** *Es seien $X_i \sim Bi(1,p)$, unabhängig. Dann gilt für $Z_n = \sum_{i=1}^n X_i$ ($\sim Bi(n,p)$):*

$$\lim Z_n \to^D Z \sim \mathcal{N}(np, np(1-p))$$

Bem.: Der Satz sagt aus, daß für ausreichend großes $n \in \mathbb{N}$ die Binomialverteilung durch die (einfachere) (Standard–)Normalverteilung ersetzt werden kann,

$$P(Z_n < y) \approx \Phi\left(\frac{y-n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}}\right).$$

Beweis: Mit $\mathbf{E}X_i = np$ und $\operatorname{Var} X_i = np(1-p)$ folgt unter Anwendung des Zentralen Grenzwertsatzes:

$$\begin{aligned} P(Z_n < y) &= P\left(\frac{Z_n - n\cdot\mu}{\sqrt{n}\cdot\sigma} < \frac{y-n\cdot\mu}{\sqrt{n}\cdot\sigma}\right) \\ &= P\left(\frac{Z_n - n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}} < \frac{y-n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}}\right) \\ &\approx \Phi\left(\frac{y-n\cdot p}{\sqrt{n\cdot p\cdot(1-p)}}\right) \end{aligned}$$

$\square$

Bsp. 3.11 *Es seien* $n = 1000$ *und* $p = 0.4$. *Gesucht werde die Wahrscheinlichkeit* $P(Z_n < 300)$. *Es gilt:*

$$\begin{aligned}
P(Z_n < 300) &= \sum_{x<300} P(Z_n = x) \\
&= \sum_{i=0}^{299} \binom{1000}{i} 0.4^i (1-0.4)^{1000-i}
\end{aligned}$$

großer Rechenaufwand.

besser: Anwendung des Satzes von MOIVRE–LAPLACE.

Es gilt:

$$\begin{aligned}
P(Z_n < 300) &\approx \Phi\left(\frac{300 - 1000 \cdot 0{,}4}{\sqrt{1000 \cdot 0{,}4 \cdot (1-0{,}4)}}\right) \\
&= \Phi\left(\frac{-100}{\sqrt{240}}\right) \approx \Phi\left(\frac{-100}{15{,}49}\right) \\
&= \Phi(-6.45) = 1 - \underbrace{\Phi(6.45)}_{\approx 1} \approx 0
\end{aligned}$$

Bem.: Die Anwendung des Satzes von MOIVRE–LAPLACE setzt voraus, daß $n \in \mathbb{N}$ hinreichend groß ist.

Faustregel: $n \cdot p \geq 10$ und $n \cdot (1-p) \geq 10$.

Bsp. 3.12 *Wir betrachten* POISSON*–verteilte unabhängige Zufallsgrößen* $X_i \sim Poi(\lambda_i)$ *(*$i = 1, \ldots, n$*,*

$$X_i : \quad \begin{pmatrix} 0 & 1 & 2 & \ldots & k & \ldots \\ p_0 & p_1 & p_2 & \ldots & p_k & \ldots \end{pmatrix}$$

mit $p_j = \frac{\lambda_i^j}{j!} \cdot e^{-\lambda_i}$ *(*$i = 1, \ldots, n$*).*
$\mathbf{E}X_i = \mathrm{Var}\, X_i = \lambda_i$*.*

$$Z_n := \sum_{i=1}^{n} X_i$$

Für den Erwartungswert dieser Zufallsgrößen gilt:

$$\mathbf{E}Z_n = \mathbf{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbf{E}X_i = \sum_{i=1}^{n} \lambda_i$$

Wir nehmen nun an, $\lambda_i = \lambda$*, für alle* $i = 1, \ldots, n$*. Ohne diese Annahme haben die Zufallsgrößen* X_i *verschiedene Erwartungswerte und Varianzen, so daß der zentrale Grenzwertsatz (in der angegebenen Form) nicht anwendbar ist.*

Es gilt also unter dieser Annahme:
$\mathbf{E}X_i = \mu = \lambda$*;* $\quad \mathrm{Var}\, X_i = \sigma^2 = \lambda$*.*

Lemma 3.21 *Es seien* X_1 *und* X_2 *unabhängig,* $X_1, X_2 \sim Poi(\lambda_i)$, $i = 1, 2$*). Dann ist die Zufallsgröße* $Z_2 := X_1 + X_2$ *ebenfalls* POISSON*–verteilt und es gilt:* $Z_2 \sim Poi(\lambda_1 + \lambda_2)$.

(**Bem:** Vergleichen Sie mit der Faltungsformel für stetige Zufallsvariablen)

Beweis: Es gilt für alle $k \in \mathbb{N}$:

$$
\begin{aligned}
P(Z_2 = k) &= \sum_{t=0}^{k} p_1(t) \cdot p_2(k-t) \\
&= \sum_{t=0}^{k} \left(\frac{\lambda_1^t}{t!} \cdot e^{-\lambda_1} \cdot \frac{\lambda_2^{k-t}}{(k-t)!} \cdot e^{-\lambda_2} \right) \\
&= \sum_{t=0}^{k} \left(\frac{\lambda_1^t \cdot \lambda_2^{k-t}}{t! \cdot (k-t)!} \cdot e^{-(\lambda_1+\lambda_2)} \right) \\
&= e^{-(\lambda_1+\lambda_2)} \cdot \frac{1}{k!} \cdot \sum_{t=0}^{k} \frac{\lambda_1^t \cdot \lambda_2^{k-t} \cdot k!}{t! \cdot (k-t)!} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{k!} \cdot (\lambda_1 + \lambda_2)^k \quad \text{(Binomischer Lehrsatz)}
\end{aligned}
$$

$\square$

Bem. 20 *Die Funktionen p_1 und p_2 heißen auch* *Faltungsdichten.*

Mittels dieses Lemmas können wir nun unter unserer Annahme $\lambda_i = \lambda$ ($i = 1, \ldots, n$) schlußfolgern:

$$Z_n = \sum_{i=1}^{n} X_i \sim Poi(n \cdot \lambda).$$

Wir wenden jetzt den Zentralen Grenzwertsatz an. Dann erhalten wir für hinreichend großes $\lambda' := n \cdot \lambda$:

$$P\left(\frac{Z_n - n\cdot\mu}{\sqrt{n}\cdot\sigma} < x\right) = P\left(\frac{Z_n - \lambda'}{\sqrt{\lambda'}} < x\right) \approx \Phi(x).$$

Also kann auch eine POISSON–Verteilung durch eine einfachere (Standard–)Normalverteilung ersetzt werden, falls die Parameter λ_i ($i = 1, \ldots, n$) alle gleich λ sind und der Faktor $n \cdot \lambda$ hinreichend groß ist (etwa $n \cdot \lambda \geq 10$).

Bem.: Sind die Parameter λ_i ($i = 1, \ldots, n$) nicht alle gleich, so gilt die Aussage trotzdem, falls ihre Summe hinreichend groß ist (≥ 10).

Bsp. 3.13 *Seien* X_i *unabhängig,* $X_i \sim \mathcal{N}(0,1)$, $i = 1,\ldots,n$.

$$Y = \sum_{i=1}^{n} X_i^2 \sim \chi_n^2,$$

d.h. Y *ist* χ^2 *verteilt mit* n *Freiheitsgraden.*

$$f_Y(y) = \begin{cases} \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n-2}{2}} e^{-\frac{x}{2}}, & \textit{falls} \quad x \geq 0 \\ 0 & \textit{sonst.} \end{cases}$$

$$\begin{aligned}
\mathbf{E}Y &= n \\
\operatorname{Var} Y &= \mathbf{E}(Y-n)^2 = \mathbf{E}\sum_{i=1}^{n}(X_i^2-1)^2 = n\mathbf{E}(X_1^2-1)^2 \\
&= n\mathbf{E}(X_1^4 - 2\mathbf{E}X_1^2 + 1) = n(3-2+1) = 2n. \\
&\Rightarrow \lim_{n\to\infty} P\Big(\frac{\sum_{i=1}^n X_i^2 - n}{\sqrt{2n}} < y\Big) = \Phi(y).
\end{aligned}$$

$$P(\sum_{i=1}^{n} X_i^2 < x) \approx \Phi\Big(\frac{x-n}{\sqrt{2n}}\Big)$$

z.B. $n = 30, x = 23.364$: $P(\sum_{i=1}^n X_i^2 < x) = 0.2$

$$\Phi\Big(\frac{x-n}{\sqrt{2n}}\Big) = \Phi(-0.8567) = 1 - 0.8042 = 0.1958.$$

bleibt z.z.: $\mathbf{E}X_i^4 = 3$.

$$\begin{aligned}
\sqrt{2\pi}\mathbf{E}X_i^4 &= \int_{-\infty}^{\infty} x^4 e^{-\frac{x^2}{2}}\,dx \\
&= 2\int_0^{\infty} x^4 e^{-\frac{x^2}{2}}\,dx, \quad t = x^2,\ dx = \frac{1}{2}t^{-\frac{1}{2}}\,dt \\
&= \int_0^{\infty} t^{\frac{3}{2}} e^{-\frac{t}{2}}\,dt \\
&= \int_0^{\infty} t^{\frac{5}{2}-1} e^{-\frac{t}{2}}\,dt \\
&= \Gamma\Big(\frac{5}{2}\Big)2^{\frac{5}{2}} \\
&= \Gamma\Big(2+\frac{1}{2}\Big)2^{\frac{5}{2}} \\
&= 1\cdot 3\cdot \frac{\sqrt{\pi}}{4}\cdot 2^{\frac{5}{2}} \\
&= 3\cdot\sqrt{2\pi} \\
\mathbf{E}X_i^4 &= 3.
\end{aligned}$$

Dabei haben wir verwendet:

$$\int_0^{\infty} t^{\lambda-1} e^{-\alpha t}\,dt = \frac{\Gamma(\lambda)}{\alpha^\lambda}$$

$$\begin{aligned}
\Gamma(n+1) &= n\Gamma(n) = n! \\
\Gamma(n+\frac{1}{2}) &= 1\cdot 3\cdot 5\cdots(2n-1)\frac{\sqrt{\pi}}{2^n}
\end{aligned}$$

*Beweis des Zentralen Grenzwertsatzes

Sei $\phi_{X-\mu}$ die charakteristische Funktion von $X_i - \mu$. Da die ersten beiden Momente (μ, σ^2) existieren, folgt aus der Taylorreihendarstellung

$$\phi_{X-\mu}(t) = 1 - \frac{1}{2}\sigma^2 t^2 + o(t^2)$$

Die ZV

$$\frac{X_i - \mu}{\sqrt{n}\sigma}$$

haben die charakteristische Funktion

$$\phi_{X-\mu}\Big(\frac{t}{\sqrt{n}\sigma}\Big),$$

Die ZV $Y_n = \sum_{i=1}^n \frac{X_i-\mu}{\sqrt{n}\sigma}$ hat also die charakteristische Funktion

$$\phi_{Y_n}(t) = \left(\phi_{X-\mu}\Big(\frac{t}{\sqrt{n}\sigma}\Big)\right)^n = \Big(1 - \frac{t^2}{2n} + o\big(\frac{t^2}{n}\big)\Big)^n.$$

Es gilt:

$$\ln\Big(1 - \frac{t^2}{2n} + o\big(\frac{t^2}{n}\big)\Big)^n = n\ln\Big(1 - \frac{t^2}{2n} + o\big(\frac{t^2}{n}\big)\Big) \rightarrow -\frac{t^2}{2}.$$

(vgl. Taylorreihenentwicklung des Logarithmus)

$$\ln \phi_{Y_n}(t) \to -\frac{t^2}{2}$$

$$\phi_{Y_n}(t) \to e^{-\frac{t^2}{2}}.$$

D.h. die ch.Fkt. von Y_n konvergiert gegen die ch.Fkt. der Standard-Normalverteilung (sogar gleichmäßig).

Aus dem Konvergenzsatz folgt: $Y_n \to Z \sim \mathcal{N}(0,1)$.

Bsp. 3.14 *Münzwurf: 1000 mal.*

Frage: Wie groß ist die Wkt., dass weniger als 475 mal Zahl fällt?

$$X_i = \begin{cases} 1 & \textit{falls Zahl} \\ 0 & \textit{sonst} \end{cases}$$

$$P(\sum_{i=1}^{1000} X_i \leq 475) \approx P(\underbrace{\sqrt{1000}\frac{\frac{1}{1000}\sum X_i - \frac{1}{2}}{\sqrt{\frac{1}{4}}}}_{\sim\mathcal{N}(0,1)} \leq \sqrt{1000}\frac{\frac{475}{1000} - \frac{1}{2}}{\sqrt{\frac{1}{4}}}$$

$$\approx \Phi(\sqrt{1000}\frac{0.475 - 0.5}{\frac{1}{2}})$$

$$= \Phi(-1.58) \approx 0.057.$$

Bedeutung des ZGWS in der Statistik

- beim Schätzen

Gesetz der Großen Zahlen: $\overline{X} \to \mu$.

Frage: Wie groß ist der Stichprobenumfang zu wählen,
um eine bestimmte Genauigkeit zu erreichen?

ε, δ vorgegeben, klein ($\varepsilon, \delta < 0.5$).

n ist so zu wählen, dass

$$P(|\overline{X} - \mu| \le \varepsilon) \ge 1 - \delta$$

$$\begin{aligned}
1 - \delta &\le P(|\overline{X} - \mu| \le \varepsilon) \\
&= P(\sqrt{n}\frac{(|\overline{X} - \mu|}{\sqrt{VarX}} \le \sqrt{n}\frac{\varepsilon}{\sqrt{VarX}}) \\
&= P(\sqrt{n}\frac{(|\overline{X} - \mu|}{\sigma} \le \sqrt{n}\frac{\varepsilon}{\sigma}) \\
&= \Phi(\sqrt{n}\frac{\varepsilon}{\sigma})
\end{aligned}$$

gdw.

$$\begin{aligned}
\Phi^{-1}(1 - \delta) &\le \sqrt{n}\frac{\varepsilon}{\sigma} \\
n &\ge \left(\frac{\sigma\Phi^{-1}(1 - \delta)}{\varepsilon}\right)^2
\end{aligned}$$

Bedeutung des ZGWS in der Statistik

- beim Testen

$\mu := \mathbf{E}X$. Wir testen z.B.

$$H_0 : \mu \leq \mu_0 \quad \text{gegen} \quad H_1 : \mu > \mu_0$$

Teststatistik:

$$T_n = \sqrt{n}\frac{\overline{X} - \mu_0}{\sigma}$$

T_n klein spricht für H_0, T_n groß gegen H_0.

Fehler 1. Art: H_0 ablehnen, obwohl richtig

möchte man begrenzen ($\leq \alpha$)

Fehler 2. Art: H_0 annehmen, obwohl falsch

sollte auch klein sein ($\leq \beta$)

$$P_{\mu_0}(T_n \geq u_{1-\alpha}) \to \alpha \quad \text{nach ZGWS}$$

denn

$$P_{\mu_0}(T_n < u_{1-\alpha}) \to \Phi(u_{1-\alpha}) = 1 - \alpha$$

(wenn $\mu < \mu_0$ so $P_\mu(T_n < u_{1-\alpha}) > P_{\mu_0}(T_n < u_{1-\alpha})$)

Wahrscheinlichkeitsverteilungen (Zus.)

Binomial $X \sim B(n,p)$

X : Anzahl von “Erfolgen” bei

n Versuchen und Erfolgswkt. p.

Poisson $X \sim Poi(\lambda)$

X : Anzahl von “Erfolgen” bei

n Versuchen und Erfolgswkt. p,

n groß und p klein, $n \cdot p = \lambda$.

X : $\#$ Ankünfte in einem Zeitintervall

Geometrische Verteilung $X \sim Geo(\lambda)$

X : Anzahl der Versuche bis zum ersten “Erfolg”

Gleichverteilung $X \sim R(a,b)$

Zufallszahlen

Exponential $X \sim Exp(\lambda)$

“gedächtnislose” stetige Verteilung.

Normal $X \sim N(\mu, \sigma^2)$

Zentraler Grenzwertsatz

Fehlergesetz (viele kleine unabh. Fehler)