3/7/17

## Konvergenz von $\Delta$-Schemata

**Def 7.2** Für $n \in \mathbb{N}$ sei $k_n \in \mathbb{N}$ und $(X_{n,\ell})_{n\in\mathbb{N}, 1\le \ell \le k_n}$ ZV. $(X_{n,\ell})_{n\in\mathbb{N}, 1\le\ell\le k_n}$ heißt $\Delta$-Schema. Ein $\Delta$-Schema ist

- unabhängig, falls $\forall n$ $(X_{n,\ell})_{\ell\le k_n}$ unabhängig
- zentriert, falls $\forall n \,\forall \ell$: $\mathbb{E}[X_{n,\ell}] = 0$.
- normiert, falls $\sum_{1\le\ell\le k_n} \mathrm{Var}[X_{n,\ell}] = 1 \quad \forall n \in \mathbb{N}$.
- standardisiert, falls zentriert + normiert.

$$\begin{array}{llll} X_{1,1} & \dots & X_{1,m_1} & \\ X_{2,1} & \dots & \dots & X_{2,m_2} \\ \vdots & & & \\ X_{n,1} & \dots & \dots & X_{n,k_n} \end{array}$$

Bsp.

$$\begin{array}{llll} X_1 & & & \\ X_1 & X_2 & & \\ X_1 & X_2 & X_3 & \\ \vdots & & & \\ X_1 & \dots & & X_n \end{array}$$

Sei $S_n = \sum_{1\le\ell\le k_n} X_{n,\ell}$.

zentr. $\Rightarrow \mathbb{E}[S_n] = 0$

norm $\Rightarrow \mathrm{Var}[S_n] = 1$

$\boxed{\text{keine Einschränkung}}$

**Def 7.3** Ein $\Delta$-Schema heißt

1) asymptotisch vernachlässigbar, falls
$$\forall \varepsilon > 0: \lim_{n\to\infty} \max_{\ell \le k_n} P(|X_{n,\ell}| > \varepsilon) = 0$$

2) mit $X_{n,\ell} \in L^2$ erfüllt es Lindenberg-Bedingung, falls
$$\forall \varepsilon > 0: \frac{1}{\underbrace{\mathrm{Var}[S_n]}_{=1}} \sum_{1\le\ell\le k_n} \mathbb{E}\left[X_{n,\ell}^2 \cdot \mathbb{1}\left[X_{n,\ell}^2 \ge \varepsilon \underbrace{\mathrm{Var}[S_n]}_{=1}\right]\right] \xrightarrow{n\to\infty} 0$$

**Satz 7.4.** (Lindenberg-Feller). Sei $X$ ein unabhängiges, stand. $\Delta$-Schema. Dann äquivalent:

1) $X$ erfüllt Lindenberg

2) $X$ asymp. vernachlässigbar und $S_n \xrightarrow{d} \mathcal{N}(0,1)$.

**Beweis.** Nur 1) $\Rightarrow$ 2). Für $\varepsilon > 0$ gilt:

$$\max_{\ell\le k_n} P(|X_{n,\ell}| > \varepsilon) \le \sum_{1\le\ell\le k_n} P(|X_{n,\ell}| > \varepsilon)$$
$$= \sum_{1\le\ell\le k_n} \mathbb{E}[\mathbb{1}[X_{n,\ell}^2 > \varepsilon^2]]$$
$$\le \frac{1}{\varepsilon^2} \sum_{1\le\ell\le k_n} \mathbb{E}[X_{n,\ell}^2 \cdot \mathbb{1}[X_{n,\ell}^2 \cdot \mathbb{1}[|X_{n,\ell}| > \varepsilon]] \to 0$$

Wir zeigen: $\varphi_{S_n}(t) \to e^{-t^2/2}$, wodurch 2) gezeigt ist.

Mit der Unabhängigkeit:
$$\varphi_{S_n}(t) = \prod_{1\le\ell\le k_n} \varphi_{X_{n\ell}}(t) =: \prod_{\ell=1}^{k_n} \varphi_{n,\ell}(t).$$

Ferner gilt:
$$e^{iy} = 1 + iy + \frac{\theta_1 y^2}{2} \qquad e^{iy} = 1 + iy - \frac{y^2}{2} + \frac{\theta_2 |y|^3}{6}$$
mit $|\theta_1|, |\theta_2| \le 1$. (hängen von $y$ ab.) Insbesondere:

$$\varphi_{n,\ell}(t) = \int e^{itx}\, dP_{n,\ell}(x)$$
$$= \int_{|x|\ge\varepsilon} \left(1 + itx + \frac{\theta_1 (tx)^2}{2}\right) dP_{n,\ell}(x) + \int_{|x|<\varepsilon} \left(1 + itx - \frac{t^2 x^2}{2} + \frac{\theta_2 |tx|^3}{6}\right) dP_{n,\ell}(x)$$
$$= 1 - \underbrace{\frac{t^2}{2} \int_{|x|<\varepsilon} x^2\, dP_{n,\ell}(x)}_{=:A_{n,\ell}} + \frac{t^2}{2} \int_{|x|\ge\varepsilon} \theta_1 x^2\, dP_{n,\ell}(x) + \frac{|t|^3}{6} \int_{|x|<\varepsilon} x^3 \theta_2\, dP_{n,\ell}(x)$$

Sei $B_{n,\ell} = \int_{|x|\ge\varepsilon} x^2\, dP_{n,\ell}(x) \Rightarrow A_{n,\ell} + B_{n,\ell} = \mathrm{Var}[X_{n,\ell}]$.

Zusätzlich gilt:
$$\left| \int_{|x|\ge\varepsilon} \theta_1 x^2\, dP_{n,\ell}(x) \right| \le \mathbb{E}[X_{n,\ell}^2\, \mathbb{1}[X_{n,\ell}^2 \ge \varepsilon]] = B_{n,\ell}$$

und $\left| \int_{|x|<\varepsilon} \theta_2 x^3 \, dP_{n,\ell}(x) \right| \le \varepsilon \cdot A_{n,\ell}$.

$\Rightarrow \varphi_{n,\ell}(t) = 1 - \frac{t^2}{2} A_{n,\ell} + \frac{t^2}{2} H_{n,\ell} + \frac{|t|^3}{6} \tilde{H}_{n,\ell}$ mit

$|H_{n,\ell}| \le B_{n,\ell}$, $|\tilde{H}_{n,\ell}| \le \varepsilon \cdot A_{n,\ell}$.

Für $z \in \mathbb{C}$ mit $|z| \le 1/2$ ist $\log(1+z) = z + \theta(z)|z|^2$ mit $|\theta(z)| \le 1$. Damit

$$\log \varphi_{S_n}(t) = \sum_{\ell=1}^{k_n} \underbrace{\left( -\frac{t^2}{2} A_{n,\ell} + \frac{t^2}{2} H_{n,\ell} + \frac{|t|^3}{6} \tilde{H}_{n,\ell} \right)}_{=: q_\ell} + \sum_{\ell=1}^{k_n} \theta_\ell \cdot \underbrace{\left( \frac{t^2}{2} A_{n,\ell} - \frac{t^2}{2} H_{n,\ell} + \frac{|t|^3}{6} \tilde{H}_{n,\ell} \right)^2}_{q_\ell^2}.$$

Da
- $\sum_\ell (A_{n,\ell} + B_{n,\ell}) = 1 \Rightarrow \sum_\ell A_{n,\ell} \xrightarrow{n\to\infty} 1$ [Lindenberg]
- $\left| \sum_\ell H_{n,\ell} \right| \le \sum_\ell |B_{n,\ell}| \to 0$ [Lindenberg]
- $\left| \sum_\ell \tilde{H}_{n,\ell} \right| \le \varepsilon \cdot \sum_\ell |A_{n,\ell}| \le \varepsilon \cdot C$ [$|A_{n,\ell}| \le \mathbb{E}[X_{n,\ell}^2]$ und $\mathrm{Var}[S_n] = 1$]

Ferner ist $A_{n,\ell} \le \varepsilon^2$ und somit

$$\left( \frac{t^2}{2} A_{n,\ell} - \frac{t^2}{2} H_{n,\ell} + \frac{|t|^3}{6} \tilde{H}_{n,\ell} \right)^2 \le \varepsilon (t^2 + |t|^3) \cdot q_\ell + t^2 \cdot |B_{n,\ell}| q_\ell$$

$$\Rightarrow \left| \sum_{\ell=1}^{k_n} \theta_\ell \cdot q_\ell^2 \right| \le \varepsilon (t^2 + |t|^3) \sum q_\ell + t^2 \sum_{1 \le \ell \le k_n} |B_{n,\ell}| q_\ell$$

Insgesamt: da $\varepsilon > 0$ beliebig

$\log \varphi_{S_n}(t) \xrightarrow{n\to\infty} -\frac{t^2}{2} \overset{\log \text{ stetig}}{\Rightarrow} \varphi_{S_n}(t) \to e^{-t^2/2}$.

## 7.2. Berry-Esseen Ungleichung

ZGS: $\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} N(0,1)$

**Satz 7.5.** $(X_n)_{n\in\mathbb{N}}$ unabhängig und identisch verteilt mit $\mathbb{E}[X_i] = 0$, $0 < \sigma^2 = \mathbb{E}[X_i^2] < \infty$, $\gamma = \mathbb{E}[|X_i|^3] < \infty$. Dann

$$\sup_{x\in\mathbb{R}} \left| P\left( \frac{S_n}{\sigma\sqrt{n}} \le x \right) - \Phi(x) \right| \le \frac{C \cdot \gamma}{\sigma^3 \sqrt{n}} \qquad C = 3.$$

Besseres $C$ möglich: Shevtsova '12: $C = 0.4748$.

**Bemerkung.** Die Abschätzung ist bestmöglich im Sinne dass $1/\sqrt{n}$ nicht durch eine schneller fallende Funktion ersetzt werden kann. Betrachte dazu

$P(X_i = -1) = 1/2 = P(X_i = 1)$

Dann $\mathbb{E}[X_i] = 0$, $\mathbb{E}[X_i^2] = 1$, $\mathbb{E}[|X_i|^3] = 1$ und

$1 = P(S_{2n} = 0) + P(S_{2n} > 0) + P(S_{2n} < 0)$
$= P(S_{2n} = 0) + 2P(S_{2n} < 0)$ [Symmetrie]

$\Rightarrow P(S_{2n} \le 0) = \frac{1}{2} + \frac{1}{2} P(S_{2n} = 0)$, und somit

$$\sup_x \left| P\left( \frac{S_{2n}}{\sqrt{2n}} \le x \right) - \Phi(x) \right| \ge \frac{1}{2} P(S_{2n} = 0) \underset{\uparrow}{=} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2n}} + o\left(\frac{1}{\sqrt{n}}\right)$$

Mit Stirling: $P(S_{2n} = 0) = \binom{2n}{n} \cdot 2^{-2n} = \frac{1}{\sqrt{\pi n}} + o\left(\frac{1}{\sqrt{n}}\right)$

Somit ist im Satz 7.5 $C = \frac{1}{\sqrt{2\pi}}$ bestmöglich.

**Def 7.6.** Sei $X$ ZV. $X$ heißt Polya-verteilt mit Parameter $L > 0$, falls die Lebesgue-Dichte von $X$

$$h_L(x) = \frac{1 - \cos(Lx)}{\pi L x^2}.$$

Die Char. Funktion ist $\varphi_L(t) = \left(1 - \frac{|t|}{L}\right)^+ \left( = \max\left\{0, 1 - \frac{|t|}{L}\right\} \right)$ (Übung!)

Lem 7.7. Seien $F, G$ Verteilungsfkt. mit $G'(x) \leq \lambda < \infty$.

$\Delta = \sup_x |F(x) - G(x)|$

$\Delta_L = \sup_x \left| \int F(x-u) h_L(u)\, du - \int G(x-u) h_L(u)\, du \right|$

Dann:

$$\Delta \leq 2 \cdot \Delta_L + \frac{24\lambda}{\pi L} \quad \text{für} \quad L > 0.$$

Beweis. Ohne Beschränkung: $\Delta > 0$. Da $F(x) - G(x) \xrightarrow{x \to \infty} 0$

gibt es $x_0$ mit

$F(x_0) - G(x_0) = \Delta$ oder $F(x_0^-) - G(x_0^-) = -\Delta$

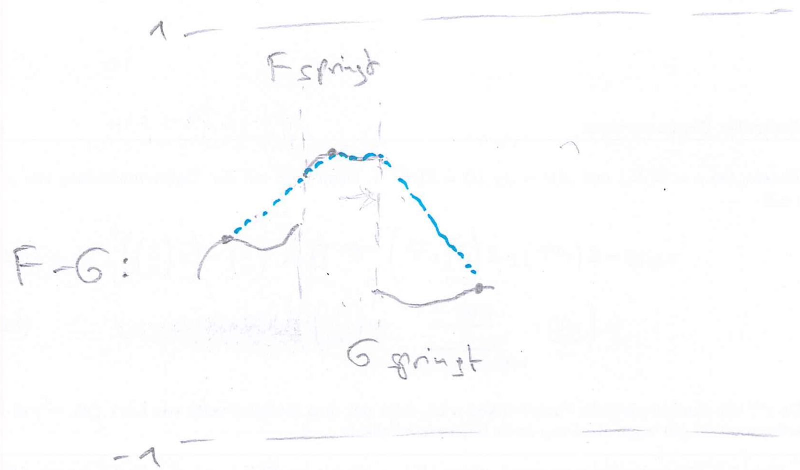
Da $F \nearrow$ und $G' \leq \lambda$

$$F(x_0 + s) - G(x_0 + s) \underbrace{\geq}_{[\geq F(x_0) - (G(x_0) + \lambda s)]} \Delta - \lambda s, \quad s \geq 0$$

$$\Rightarrow F(t-x) - G(t-x) \geq \begin{cases} \Delta/2 + \lambda x, & |x| \leq \delta \\ -\Delta, & \text{sonst} \end{cases}$$

mit $\delta = \Delta / 2\lambda$, $t = x_0 + \delta$.

Einsetzen in $\Delta_L$ ergibt die gewünschte Schranke. ∎



6/7/17

**Lem 7.8.** $F_1, F_2$ Verteilungsfkt. mit integrierbaren char. Funktionen. Dann

$$F_1(x) - F_2(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \frac{\varphi_1(t) - \varphi_2(t)}{it} \, dt, \quad x \in \mathbb{R}$$

Bew. Mit Kor. 6.7. ∎

Bew. (Satz 7.5) Mit Hölder:

$$\gamma = \mathbb{E}[|X_1|^3] \geq \mathbb{E}[X_1^2]^{3/2} = \sigma^3$$

Damit ist nichts zu zeigen für $\frac{3}{\sqrt{n}} \geq 1 \Rightarrow n \leq 9$.
Sei ohne Beschränkung $\sigma = 1$. Setze

$$F_L(x) = (F * h_L)(x), \quad G_L(x) = (G * L_L(x))$$

Mit Lem 7.8:

$$|F_L(x) - G_L(x)| \leq \frac{1}{2\pi} \int |\varphi_F(t)\varphi_L(t) - \varphi_G(t)\varphi_L(t)| \frac{1}{|t|} \, dt$$

$$\leq \frac{1}{2\pi} \int_{-L}^{L} |\varphi_F(t) - \varphi_G(t)| \frac{1}{|t|} \, dt$$

Mit Lem 7.7:

$$\sup_x |F(x) - G(x)| \leq \frac{1}{\pi} \int_{-L}^{L} |\varphi_F(t) - \varphi_G(t)| \frac{1}{|t|} \, dt + \frac{24\lambda}{\pi L}.$$

Setze $F_n(x) = P\left(\frac{S_n}{\sqrt{n}} \leq x\right)$, $G(x) = \Phi(x)$, $G' = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \leq \frac{1}{\sqrt{2\pi}}$

$$\Rightarrow |F_n(x) - G(x)| \leq \frac{1}{\pi} \int_{-L}^{L} \left| \varphi\left(\frac{t}{\sqrt{n}}\right)^n - e^{-t^2/2} \right| \frac{1}{|t|} \, dt + \frac{24}{\sqrt{2}\pi^{3/2} L}$$

($\varphi$ die char. Fkt von $X_1$)

Mit $|\varphi(t) - 1 + t^2/2| \leq \gamma \, |t|^3/6$ erhalten wir

$$\left|\varphi\left(\frac{t}{\sqrt{n}}\right)\right| \leq 1 - \frac{t^2}{2n} + \frac{\gamma \cdot |t|^3}{6 n^{3/2}} \leq 1 - \frac{5t^2}{18n} \leq e^{-5t^2/18n}.$$

und für $|t| \leq L =: \frac{4\sqrt{n}}{3\gamma}$

Für $\alpha, \beta \in \mathbb{C}$ gilt

$$\alpha^n - \beta^n = (\alpha - \beta) \sum_{m=0}^{n-1} \alpha^{n-m-1} \cdot \beta^m \qquad \left(\text{aus } \sum_{i=0}^{n-1} x^i = \frac{1-x^n}{1-x}\right)$$

also $|\alpha^n - \beta^n| \leq n \cdot |\alpha - \beta| \cdot \max\{|\alpha|, |\beta|\}^{n-1}$.

Mit $\alpha = \varphi(t/\sqrt{n})$, $\beta = e^{-t^2/2n}$:

$$\left|\varphi\left(\frac{t}{\sqrt{n}}\right)^n - e^{-t^2/2}\right| \leq n \cdot \left|\varphi\left(\frac{t}{\sqrt{n}}\right) - e^{-t^2/2n}\right| \cdot \max\left\{\left|\varphi\left(\frac{t}{\sqrt{n}}\right)\right|, e^{-t^2/n}\right\}^n$$

Dann:

1) $$n\left|\varphi\left(\frac{t}{\sqrt{n}}\right) - e^{-t^2/2n}\right| \leq n \cdot \left|\varphi\left(\frac{t}{\sqrt{n}}\right) - 1 + t^2/2n\right| + n \cdot \left|e^{-t^2/2n} - 1 - t^2/2n\right|$$

$$\leq \gamma \cdot \frac{|t|^3}{6\sqrt{n}} + \frac{t^4}{8n} \qquad \left(|e^x - 1 - x| \leq x^2/2\right)$$

2) $$\max\{..\} \leq e^{-\frac{5t^2}{18n}(n-1)} \leq e^{-t^2/4}, \text{ für } n \geq 10.$$

Aus 1), 2):

$$\sup_x |F_n(x) - G(x)| \leq \frac{1}{\pi} \int_{-L}^{L} e^{-t^2/4} \left(\frac{\gamma |t|^3}{6\sqrt{n}} + \frac{t^4}{8n}\right) \frac{dt}{|t|} + \frac{24}{\sqrt{2}\pi^{3/2} L}$$

$$\leq \frac{\gamma}{\pi 6 \sqrt{n}} \underbrace{\int_{-\infty}^{\infty} t^2 e^{-t^2/4} \, dt}_{4\sqrt{\pi}} + \frac{1}{8\pi n} \underbrace{\int e^{-t^2/4} |t|^3 \, dt}_{16} + \frac{24}{\sqrt{2}\pi^{3/2} L}.$$

∎

## Poisson-Approximation

Seien $X_n$ unabh. mit $P(X_n = 1) = p$, $P(X_n = 0) = 1 - p$.

Mit Berry-Essen:

$$\sup_x \left| P\left( \frac{S_n - np}{\sqrt{np(1-p)}} \le x \right) - \Phi(x) \right| \le C \cdot \frac{p(1-p)(p^2+(1-p)^2)}{(p(1-p))^{3/2}\sqrt{n}}$$

Wenn $np \approx 1$, dann ist die Abschätzung unbrauchbar.

Seien $(X_{n,l})_{n\in\mathbb{N},\, 1\le l\le k_n}$ unabhängiges $\Delta$-Schema mit

$$P(X_{n,l}=1) = p_{n,l}, \quad P(X_{n,l}=0) = 1-p_{n,l}$$

mit $\mu_n = \sum_{1\le l\le k_n} p_{n,l}$

<u>Satz 7.4</u> Sei $X$ Poisson-Verteilt mit Par. $\mu_n$. Dann:

$$|P(S_n \le k) - P(X\le k)| \le 2 \cdot \sum_{1\le l\le k_n} p_{n,l}^2, \quad k\in\mathbb{N}_0.$$

<u>Bsp.</u> $p_{n,l} = 1/n \Rightarrow |P(S_n\le k) - P(X\le k)| \le 1/n$.

<u>Beweis.</u> Seien $(Z_{n,l})_{1\le l\le k_n}$ unabh. ZV mit:

$$P(Z_{n,l}=k) = \begin{cases} P(X_l = k), & k\ge 1 \\ 1-p_{n,l}, & k=0 \\ e^{-p_{n,l}} - (1-p_{n,l}), & k=-1 \end{cases}, \quad X_l \sim Po(p_{n,l})$$

Dann $X_{n,l} = \mathbb{1}[Z_{n,l}\ne 0]$, $X_l = \max\{Z_{n,l}, 0\}$

mit $X = \sum_{1\le l\le k_n} X_l$. Sei

$$D = \{\text{es gibt } l \text{ mit } X_{n,l} \ne X_l\} = \bigcup_{1\le l\le k_n} \{Z_{n,l} \notin \{0,1\}\}$$

Dann $P(D) \le \sum P(Z_{n,l}\notin\{0,1\}) = \sum 1-(1-p_{n,l}) - e^{-p_{n,l}} \cdot p_{n,l}$

$$= \sum_l p_{n,l} \cdot (1-e^{-p_{n,l}}) \le \sum_l p_{n,l}^2.$$

$$\Rightarrow |P(S_n\le k) - P(X_n \le k)|$$

$$= |P(\{S_n\le k\}\cap D) - P(\{X_n\le k\}\cap D)|$$

$$\le 2\, P(D)$$