Vorlesung 18 | 12.1.2021 | 14:15–16:00 via Zoom

Ein frohes neues Jahr!

6 Das Gesetz der großen Zahlen (Ende)

(Kapitel 6 in Bovier Skript)

6.6 Große Abweichungen von GGZ

Seien $X_1, X_2, \ldots$ i.i.d Z.V. integrierbar. Sei $S_n = \sum_{k=1}^n X_k$ mit $\mathbb{E}[X_i] = m$, und sei $\psi(t) = \mathbb{E}[e^{tX_1}]$. Dann wegen GGZ:

$$\frac{S_n}{n} \xrightarrow[f.s.]{} m.$$

Satz 15. *Für alle $n \geqslant 1$, $a \in \mathbb{R}$*

$$(a)\quad \mathbb{P}\left(\frac{S_n}{n} \geqslant a\right) \leqslant e^{-nI(a)}, \qquad \textit{für } a \geqslant m$$
$$(b)\quad \mathbb{P}\left(\frac{S_n}{n} \leqslant a\right) \leqslant e^{-nI(a)}, \qquad \textit{für } a \leqslant m$$

wobei die exponentielle Abfallrate $I(a)$ ist gegeben durch

$$I(a) = \sup_{t \in \mathbb{R}} [at - \log \psi(t)].$$

Bemerkung. Oft wir können sagen dass für „gute Menge“ $G \subseteq \mathbb{R}$:

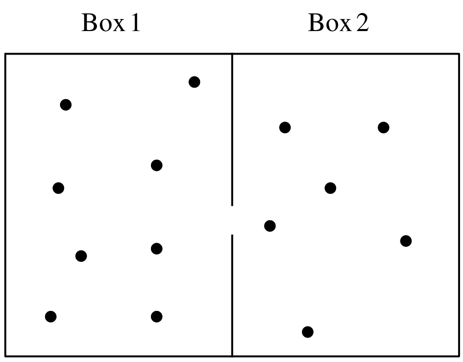
$$\lim_{n \to \infty} -\frac{1}{n} \log \mathbb{P}\left(\frac{S_n}{n} \in G\right) = \inf_{a \in G} I(a).$$

Das ist die Theorie von Große Abweichungen. Das bedeuted dass typischeweise wir haben

$$\mathbb{P}\left(\frac{S_n}{n} \geqslant a\right) = e^{-n(I(a) + o(1))}, \qquad \text{als } n \to \infty.$$

Beispiel. Modell von ein Gas. Jede Moleküle hat W-keit $1/2$ in Box 1 zu sein und sind unabhängig (ideal gas), d.h. $X_1, \ldots, X_n$ i.i.d. Ber$(1/2)$ Z.V. wobei

$$X_k = \begin{cases} 1 & \text{falls Moleküle } \#k \text{ in Box 1 ist} \\ 0 & \text{falls Moleküle } \#k \text{ in Box 2 ist} \end{cases}$$



$n = 10^{23}$ Moleküle (Teilchen)

$\mathrm{Vol}(\mathrm{Box}\,1) = \mathrm{Vol}(\mathrm{Box}\,2)$

Frage:

$$\mathbb{P}\left(\frac{\#\text{Teilchen in Box 1}}{n} \geqslant \frac{1}{2} + 10^{-10}\right) \leqslant p = ?$$

Tchebichev Abschätzung: ($\mathrm{Var}(X_k) = 1/4$)

$$p \leqslant \frac{\mathrm{Var}\left(\frac{S_n}{n}\right)}{(10^{-10})^2} = \frac{10^{20}}{4\,10^{23}} = \frac{1}{4000}.$$

Abschatzüng mit Theorem 15. $m = 1/2$ un ein explizit Rechnüng gibt

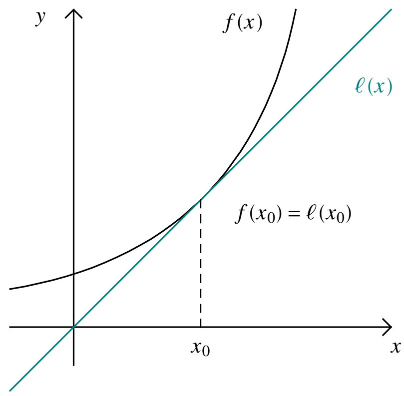
$$I(a) = a\log(2a) + (1-a)\log(2(1-a))$$

$$p \leqslant e^{-2n(10^{-10})^2} = e^{-2000}!!!$$

Proposition 16. *(Jensen's Ungleichung) Sei X eine reelle Z.V., integrierbar, und $f\colon \mathbb{R} \to \mathbb{R}$ eine konvexe Funktion. Dann*

$$f(\mathbb{E}[X]) \leqslant \mathbb{E}(f(X)).$$

Beweis.



Da f konvex ist, $\forall x_0 \in \mathbb{R}\ \exists$ gerade $\ell(x)$ s.d.

$$\ell(x_0) = f(x_0), \quad \ell(x) \leqslant f(x), \quad x \in \mathbb{R}.$$

Wählen wir $x_0 = \mathbb{E}[X]$. Dann,

$$f(\mathbb{E}[X]) = f(x_0) = \ell(x_0) = \ell(\mathbb{E}[X]) \underset{\text{Linearität}}{=} \mathbb{E}[\ell(X)] \leqslant \mathbb{E}[f(X)].$$

□

Beweis. (von Thm. 15) Fall $a \geqslant m$. Den Fall $a \leqslant m$ ist analog. OBdA $m = 0$.

$$\mathbb{P}\left(\frac{S_n}{n} \geqslant a\right) = \mathbb{P}(S_n \geqslant na) \underset{\text{Markov}}{\leqslant} \inf_{\lambda \geqslant 0} \mathbb{E}\left[\frac{e^{\lambda S_n}}{e^{\lambda na}}\right] \underset{i.i.d.}{=} \inf_{\lambda \geqslant 0} e^{-\lambda na}(\mathbb{E}[e^{\lambda X_1}])^n = \inf_{\lambda \geqslant 0} \exp\left\{-n\,(\lambda\, a - \log \mathbb{E}[e^{\lambda X_1}])\right\}$$

$$= \exp\left(-n \sup_{\lambda \geqslant 0} (a\lambda - \log \psi(\lambda))\right)$$

Für $\lambda \leqslant 0$, $a \geqslant 0$,

$$a\lambda - \log \psi(\lambda) \leqslant -\log \psi(\lambda) = -\log \mathbb{E}[e^{\lambda X_1}] \underset{\text{Jensen}}{\leqslant} -\log e^{\lambda \mathbb{E}[X_1]} = 0.$$

$$\sup_{\lambda \geqslant 0} (a\lambda - \log \psi(\lambda)) = \sup_{\lambda \in \mathbb{R}} (a\lambda - \log \psi(\lambda)) = I(a).$$

$\square$

Eigenschaften von $I(a)$.

a) I ist convex.

b) $I(a) \geqslant 0$ für alle $a \in \mathbb{R}$.

c) Falls $\psi(\lambda) < \infty$ auf $(-\delta, \delta)$ $\delta > 0$, dann

$$f_a(\lambda) = a\lambda - \log\psi(\lambda) \in C^\infty((-\delta, \delta))$$

mit $f_a(0) = 0$, $f_a'(0) = a - m$

$$\Rightarrow I(a) > 0, \qquad a \neq m.$$

Im Fall (c): exponentieller Abfall der W-keit großer Abweichungen!

Beispiel.

- $X_k \sim \mathcal{N}(m, \sigma^2) \Rightarrow I(a) = \frac{(a-m)^2}{2\sigma^2}$
- $X_k \sim \mathrm{Exp}(\lambda) \Rightarrow I(a) = \begin{cases} \lambda a - 1 - \log(\lambda a), & a > 0 \\ +\infty & a < 0 \end{cases}$
- $X_k \sim \mathrm{Ber}(p) \Rightarrow I(a) = a\log(a/p) + (1-a)\log((1-a)/(1-p))$, $a \in (0,1)$.

7 Der zentrale Grenzwertsatz

(Kapitel 7 in Bovier Skript)

7.1 Einleitung

Das GGZ sagt und, dass für $X_1, X_2, \dots$ i.i.d. integrierbar Z.V. und $S_n = \sum_{k=1}^n X_k$

$$\frac{S_n}{n} \xrightarrow{f.s.} \mathbb{E}[X_1], \qquad \text{als } n \to \infty.$$

Die typische Abweichungen (Fluktuationen) von S_n bzlg. $n\mathbb{E}[X_1]$ sind $o(n)$ (d.h. $f(n) = o(n) \Leftrightarrow f(n)/n \to 0$).

Frage: Wie gross sind die Fluktuationen? Falls $\mathrm{Var}(X_1) < \infty$

$$\mathrm{Var}(S_n) = n\,\mathrm{Var}(X_1) = O(n)$$

Dann

$$\frac{S_n - n\,\mathbb{E}[X_1]}{\sqrt{n\,\mathrm{Var}(X_1)}}$$

hat Varianz 1 und Mittelwert 0.

Frage: Im Allgemein,

a) Wann $\exists \gamma > 0$ s.d.

$$\frac{S_n - n\,\mathbb{E}[X_1]}{n^\gamma}$$

konvergiert (z.B. in Verteilung) gegen eine nicht triviale Z.V. (nicht 0 oder ∞).

b) Was sind die möglichen Limes Z.V.

Bevor wir versuchen, diese Frage zu beantworten, stellen wir ein nützliches Werkzeug zur Untersuchung des Gesetzes der Zufallsvariablen vor.

7.2 Charakteristiche Funktionen

Wir kennen schon die Momentenerzeugende Funktion $\psi(z) = \mathbb{E}[e^{zX}]$, $z \in \mathbb{R}$. Aber $\psi(z)$ ist manchmal $+\infty$ bis auf $z = 0$ (siehe die Cauchy Verteilung), weil $z \mapsto e^{zx}$ sehr schnell wächst (entweder für $x \to \infty$ oder $x \to -\infty$).

Dagegen ist

$$x \mapsto e^{itx} = \cos(tx) + i\sin(tx) = \sum_{n \geqslant 1} \frac{(itx)^n}{n!}$$

für $t \in \mathbb{R}$ beschränkt. Hier $i^2 = -1$.

Definition 17. *Sei X eine reelle Z.V.. Dann die charakteristische Funktion von X definiert durch*

$$\phi(t) = \phi_X(t) := \mathbb{E}[e^{itX}] = \mathbb{E}[\cos(tX)] + i\,\mathbb{E}[\sin(tX)] = \int_{\mathbb{R}} e^{itx}\mathbb{P}_X(\mathrm{d}x)$$

mit $\mathbb{P}_X$ die Verteilung von X.

Einige Eigenschaften

Lemma 18. $\phi_X\colon \mathbb{R} \to \mathbb{C}$ *ist (immer) wohldefiniert mit*

a) $\phi_X(0) = 1$

b) $|\phi_X(t)| \leqslant 1$

c)

$$\mathrm{Im}\,(\phi_X(t)) = \frac{\phi(t) - \phi(-t)}{2i}, \qquad \mathrm{Re}\,(\phi_X(t)) = \frac{\phi(t) + \phi(-t)}{2}$$

d) Falls X hat eine symmetrische Verteilung dann $\phi_X(t) \in \mathbb{R}$. (d.h. $X \sim -X$, z.B. $\mathcal{N}(0,1)$)

Beweis. Die Wohldefiniertheit folgt aus $\sin(tx)$ und $\cos(tx)$ beschränkt und messbar (⇒integrierbar bzgl jedes W-maß). Die Eigenschaften a) – d) sind ziemlich trivial. □

Bemerkung. Sei $f\colon \mathbb{R} \to \mathbb{C}$ dann

$$\mathbb{E}[f(X)] = \mathbb{E}[\mathrm{Re}\, f(X)] + i\,\mathbb{E}[\mathrm{Im}\, f(X)] \in \mathbb{C}$$

und

$$|\mathbb{E}[f(X)]|^2 = \big(\mathbb{E}[\mathrm{Re}\, f(X)]\big)^2 + \big(\mathbb{E}[\mathrm{Im}\, f(X)]\big)^2 \underset{\text{Jensen}}{\leqslant} \mathbb{E}[(\mathrm{Re}\, f(X))^2] + \mathbb{E}[(\mathrm{Im}\, f(X))^2] = \mathbb{E}[|f(X)|^2]. \tag{1}$$

Und auch für $\alpha \in \mathbb{R}$

$$|e^{i\alpha}| = 1.$$

Lemma 19. *Jede charakteristische Funktion ϕ eines W-maß μ ist* <u>*gleichmässig stetig*</u> *auf $\mathbb{R}$.*

Beweis. Es gilt

$$|e^{itx} - 1|^2 = \overline{(e^{itx} - 1)}(e^{itx} - 1) = (e^{-itx} - 1)(e^{itx} - 1) = 1 - e^{-itx} - e^{itx} + 1 = 2(1 - \mathrm{Re}\,(e^{itx})).$$

Dann

$$|\phi(t) - \phi(s)|^2 = |\mathbb{E}[e^{itX} - e^{isX}]|^2 = |\mathbb{E}[e^{itX}(-e^{i(s-t)X})]|^2 \leqslant \mathbb{E}[|e^{itX}(1 - e^{i(s-t)X})|^2]$$

wegen (1), und dann

$$\mathbb{E}[|e^{itX}(1-e^{i(s-t)X})|^2] = \mathbb{E}\Big[\underbrace{|e^{itX}|^2}_{=1}|(1-e^{i(s-t)X})|^2\Big] = \mathbb{E}[|(1-e^{i(s-t)X})|^2] = \mathbb{E}[2(1-\mathrm{Re}\,(e^{i(s-t)X}))]$$

Dann

$$|\phi(t)-\phi(s)|^2 = 2(1-\mathrm{Re}\,\phi(s-t)).$$

Sei nun $N<\infty$. Dann,

$$|1-\mathrm{Re}\,\phi(u)| = |\mathbb{E}[1-e^{iuX}]| \leqslant \mathbb{E}[|1-e^{iuX}|] = \mathbb{E}[|1-e^{iuX}|\,\mathbb{1}_{|X|\leqslant N}] + \mathbb{E}\Big[\underbrace{|1-e^{iuX}|}_{\leqslant|1|+|e^{iuX}|=2}\,\mathbb{1}_{|X|>N}\Big]$$

$$\leqslant \sup_{|x|\leqslant N}|1-e^{iux}| + 2\,\mathbb{P}(|X|>N).$$

Für $\forall\varepsilon>0$ wählen wir $N=N(\varepsilon)$ und $\delta_0=\delta_0(N,\varepsilon)>0$ s.d.

$$\mathbb{P}(|X|>N)\leqslant\frac{\varepsilon^2}{4},$$

$$\forall u, |u|<\delta_0 \qquad \sup_{|x|\leqslant N}|1-e^{iux}|\leqslant\frac{\varepsilon^2}{2}.$$

Es folgt, dass für alle $\varepsilon>0$, existiert $\delta_0>0$ s.d. für $|t-s|\leqslant\delta_0$ wir haben

$$|\phi(t)-\phi(s)|^2 = 2(1-\mathrm{Re}\,\phi(s-t)) \leqslant \sup_{|x|\leqslant N}|1-e^{i(t-s)x}| + 2\,\mathbb{P}(|X|>N) \leqslant \frac{\varepsilon^2}{2}+\frac{\varepsilon^2}{2}\leqslant\varepsilon^2.$$

□

Wie bei der Momentenerzeugende Funktion, kann man die Momenten durch Ableitung von ϕ herleiten.

Proposition 20. *Sei ϕ die char. Fkt. einer Z.V. X mit $\mathbb{E}[|X|^n]<\infty$. Dann $\phi\in C^n(\mathbb{R};\mathbb{C})$ mit*

$$\phi(0)=1, \qquad \phi^{(n)}(0)=\frac{\mathrm{d}^n}{\mathrm{d}t^n}\phi(t)\Big|_{t=0} = i^n\mathbb{E}[X^n].$$

Beweis. $\phi(0)=\mathbb{E}[e^{i0X}]=1$. Setze $e(t,x)=e^{itx}$ und

$$e^{(n)}(t,x) := \left(\frac{\partial}{\partial t}\right)^n e(t,x) = i^n x^n e(t,x).$$

Dann

$$e(t,X) = e(0,X) + \int_0^t e^{(1)}(s,X)\mathrm{d}s$$

und wir können die $\mathbb{E}$ nehmen so

$$\phi(t) = \phi(0) + \mathbb{E}\int_0^t e^{(1)}(s,X)\mathrm{d}s$$

aber $|e^{(1)}(s,X)| = |iXe^{isX}|\leqslant|X|$ und $\mathbb{E}(|X|)<\infty$, dann $|e^{(1)}(s,X)|$ integrierbar ist bzgl. $\mathbb{P}\otimes\mathscr{U}_{[0,1]}$. Fubini–Lebesgue gibt uns

$$\phi(t) = \phi(0) + \int_0^t \mathbb{E}\big[e^{(1)}(s,X)\big]\mathrm{d}s$$

Aus dominierte konvergenz folgt dass $s\mapsto\mathbb{E}\big[e^{(1)}(s,X)\big]$ stetig ist (Übung!), dann wir haben

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(t) = \mathbb{E}\big[e^{(1)}(t,X)\big] = \mathbb{E}[iXe^{itX}]$$

und insbesonderes

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(t)\bigg|_{t=0} = \mathbb{E}[iX].$$

Verallgemeinerung:

$$\begin{aligned}
e(t,X) &= e(0,X) + \int_0^t e^{(1)}(t_1,X)\mathrm{d}t_1 \\
&= e(0,X) + \int_0^t \left[e^{(1)}(0,X) + \int_0^{t_1} e^{(2)}(t_2,X)\mathrm{d}t_2\right]\mathrm{d}t_1 \\
&= e(0,X) + te^{(1)}(0,X) + \int_0^t \int_0^{t_1} e^{(2)}(t_2,X)\mathrm{d}t_2\mathrm{d}t_1 \\
&= e(0,X) + te^{(1)}(0,X) + \frac{t^2}{2}e^{(2)}(0,X) + \int_0^t \int_0^{t_1} \int_0^{t_2} e^{(3)}(t_3,X)\mathrm{d}t_3\mathrm{d}t_2\mathrm{d}t_1 \\
&\vdots \\
&= \sum_{k=0}^{n-1} \frac{t^k}{k!}\underbrace{e^{(k)}(0,X)}_{=(iX)^k} + \int_{0\leqslant t_n\cdots\leqslant t_1\leqslant t} e^{(n)}(t_n,X)\mathrm{d}t_n\cdots\mathrm{d}t_1
\end{aligned}$$

mit $|e^{(n)}(t_n,X)| \leqslant |X|^n$. Falls $\mathbb{E}[|X|^n] < \infty$ Fubini–Lebesgue gibt

$$\phi(t) = \sum_{k=0}^{n-1} \frac{t^k}{k!} i^k \mathbb{E}[X^k] + \int_{0\leqslant t_n\cdots\leqslant t_1\leqslant t} \mathbb{E}[e^{(n)}(t_n,X)]\mathrm{d}t_n\cdots\mathrm{d}t_1$$

wo $t_n \mapsto \mathbb{E}[e^{(n)}(t_n,X)]$ stetig ist (und dann gleichmässig stetig). Ableiten wir n-mal

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n}\phi(t)\bigg|_{t=0} = i^n \mathbb{E}[X^n e^{itX}]\Big|_{t=0} = i^n \mathbb{E}[X^n].$$

□

Für den Fall von summe unabhängiger Z.V. es gilt

Lemma 21.

a) Seien $X_1, X_2, \ldots$ unabhängige Z.V. mit char. Fkt. ϕ_{X_k}. Seien $S_n := \sum_{k=1}^n X_k$. Dann

$$\phi_{S_n}(t) = \prod_{k=1}^n \phi_{X_k}(t).$$

b)

$$\phi_{aX+b}(t) = e^{itb}\phi_X(at).$$

c) Insbesonderes, falls $\mathbb{E}[X_k] = \mu$ für $k \geqslant 1$, dann

$$\phi_{\frac{S_n-\mu n}{n^\gamma}}(t) = e^{-it\mu n^{1-\gamma}} \prod_{k=1}^n \phi_{X_k}(n^{-\gamma}t).$$

Beweis. (a) Folgt aus unabhängigkeit (da dann $e^{itX_1}, \ldots, e^{itX_k}$ auch unabhängig sind)

(b) Trivial.

(c)

$$\phi_{\frac{S_k-\mu n}{n^\gamma}}(t) = \mathbb{E}\left[e^{it\frac{S_k-\mu n}{n^\gamma}}\right] = e^{-it\mu n^{1-\gamma}}\mathbb{E}[e^{itn^{-\gamma}(X_1+\cdots+X_n)}] \underset{\text{unab.}}{=} e^{-it\mu n^{1-\gamma}} \prod_{k=1}^n \phi_{X_k}(n^{-\gamma}t).$$

□

Charakteristische Funktionen einiger Verteilungen

$\mathbb{P}_X$	$\phi^X(t)$
$\mathcal{N}(\mu, \sigma^2)$	$e^{it\mu - \frac{1}{2}\sigma^2 t^2}$
$\mathrm{Ber}(p)$	$(1 - p + pe^{it})$
$\mathrm{Bin}(n, p)$	$(1 - p + pe^{it})^n$
$\mathrm{Poi}(\lambda)$	$e^{-\lambda(e^{it} - 1)}$
$\mathrm{Exp}(\lambda)$	$(1 - it/\lambda)^{-1}$
$\mathrm{Geo}(q)$	$\frac{1-q}{1-qe^{it}}$
$\mathrm{Cauchy}(a)$	$e^{-a\|t\|}$

nicht Ableitbar in 0!