K 13 Wie gut ist der ZGS?

13.1 Wahrscheinlichkeitsmetrik
Um zwei Verteilungen zu vergleichen, benutzen wir

$$d(F,G) = \sup_x |F(x) - G(x)|$$

den maximalen Unterschied zwischen den Verteilungsfunktionen. (Es gibt natürlich andere Möglichkeiten auch.)

z.B. **U(0,1) und N(0,1)** $\quad X \sim U(0,1)$
Um mit einer standard Normalverteilung zu vergleichen, nehmen wir

$$Y = \frac{X - \frac{1}{2}}{\sqrt{\frac{1}{12}}}$$

weil

$$E[Y] = 0 \quad \text{und} \quad V[Y] = 1$$

$$f(y) = \frac{1}{2\sqrt{3}} \qquad -\sqrt{3} < y < \sqrt{3}$$

$$\begin{aligned} d(F(y),\phi) &= \max\left(\phi(-\sqrt{3}), \max_{(-\sqrt{3},0)}\left(\phi(y) - \frac{\sqrt{3}+y}{2\sqrt{3}}\right)\right) \\ &\approx 0.057 \end{aligned}$$

13.2 Berry-Esséen

Seien $X_1, X_2, \ldots$ u.i.v. Ist $0 < \sigma^2 = V(X_i) < \infty$ und $\gamma = E[|X-\mu|^3] < \infty$, so gilt

$$d(S_n^*, \phi) \leq \frac{0.7655\gamma}{\sigma^3\sqrt{n}}$$

$\left(S_n^* \text{ ist die Verteilungsfunktion von } \frac{\sum X_i - n\mu}{\sigma\sqrt{n}} = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)$

13.2.1 Berry-Esséen und die Gleichverteilung

$$X_i \sim U(0,1)$$

$$\mu = \frac{1}{2} \quad \text{und} \quad \sigma^2 = \frac{1}{12}$$

$$\gamma = E[|X-\frac{1}{2}|^3] = 2\int_{\frac{1}{2}}^{1}(x-\frac{1}{2})^3 \mathrm{d}x = \frac{1}{32}$$

$$d(S_n^*, \phi) \leq \frac{0.8 * \frac{1}{32}}{(\frac{1}{12})^{\frac{3}{2}}\sqrt{n}}$$

$$n = 12 \Rightarrow d(S_{12}^*, \phi) \leq 0.3$$

$$n = 300 \Rightarrow d(S_{300}^*, \phi) \leq 0.06$$

13.2.2 Berry-Esséen und die Exponentialverteilung

$$X \sim E(\lambda)$$

$$\mu = \lambda^{-1} \quad \text{und} \quad \sigma^2 = \lambda^{-2}$$

$$\gamma = E[|X - \mu|^3] = \frac{1}{\lambda^3}(12e^{-1} - 2)$$

$$d(S_n^*, \phi) \leq \frac{0.8 * \frac{1}{\lambda^3}(12e^{-1} - 2)}{(\frac{1}{\lambda})^3\sqrt{n}}$$

$$\approx \frac{1.9316}{\sqrt{n}}$$

$$n = 12 \Rightarrow d(S_{12}^*, \phi) \leq 0.5576$$

$$n = 300 \Rightarrow d(S_{300}^*, \phi) \leq 0.1115$$

$$n = 30000 \Rightarrow d(S_{30000}^*, \phi) \leq 0.01115$$

Sei $F(s)$ die genaue Verteilungsfunktion für $n = 300$, dann sagt das Resultat, dass

$$|F(s) - \phi(s)| \leq 0.1115$$

$$\phi(s) - 0.1115 < F(s) < \phi(s) + 0.1115$$

13.3 Das Gesetz vom iterierten Logarithmus

$$\{X_i\} \quad \text{u.i.v. } E[X] = 0 \text{ und } V[X] = 1$$

$$S_n = \sum_{i=1}^{n} X_i$$

Wir wissen aus dem SGGZ, dass

$$S_n/n \to 0 \quad \text{fast sicher}$$

und aus dem ZGS, dass

$$U_n = S_n/\sqrt{n} \to N(0,1) \quad \text{in Verteilung}$$

Wie gross dürfen die (sehr seltenen) Fluctuationen von U_n sein? Es kann gezeigt werden, dass

$$\lim_{n\to\infty} \sup \frac{S_n}{\sqrt{2n \log \log n}} = 1 \quad \text{fast sicher} \quad (\lim \inf = -1)$$

Der Satz gilt auch für $\{X_i\}$ u.i.v. im allgemeinen. Zum Beweis muß gezeigt werden, dass das Ereignis

$$A_n = \{S_n \geq c\sqrt{2n \log \log n}\}$$

unendlich oft passiert für $c < 1$ und nur endlich oft für $c > 1$ mit Wahrscheinlichkeit 1.

13.4 Das Arcussinus Gesetz

$$\text{Sei} \quad P(X_i = -1) = P(X_i = 1) = 0.5$$

und

$$S_n = \sum_{i=1}^{n} X_i$$

Sei $M_n(x_1, x_2, \ldots, x_n)$ die Anzahl jener Partiellsummen S_k, die positiv sind, dann gilt

$$P\left(a \leq \frac{M_n(x_1, \ldots, x_n)}{n} \leq b\right) \to \int_a^b \frac{1}{\pi\sqrt{x(1-x)}} \mathrm{d}x$$

die Arcussinusverteilung über (0,1).

Oder in anderer Form.
Sei $L_{2N} = \max\{2n \leq 2N : S_{2n} = 0\}$ der Zeitpunkt des letzten Nullpunkts. Für alle $0 < a < b < 1$ gilt

$$\lim_{N\to\infty} P\left(a \leq \frac{L_{2N}}{2N} \leq b\right) = \int_a^b \frac{1}{\pi\sqrt{x(1-x)}} \mathrm{d}x$$

Beweis vom Arcussinus Gesetz

(1) Sei $G_n = (S_{2n} = 0, S_{2k} \neq 0$ für $1 \leq k \leq n)$, die erste Rückkehr zu 0 nach $2n$ Schritten und sei

$$u_n = 2^{-2n}\binom{2n}{n}$$

dann ist

$$P(G_n) = u_{n-1} - u_n$$

Man veranschaulicht die Pfade von

$$S_1 = 1 \text{ bis } S_{2n-1} = 1$$

und von

$$S_1 = -1 \text{ bis } S_{2n-1} = 1$$

Mit Hilfe des Reflexionsprinzip fällt das Resultat aus.

(2) Sei $G_{>n} = (S_{2n} \neq 0$ für $1 \leq k \leq n)$, keine Rückkehr während der ersten $2n$ Schritte.

$$P(G_{>n}) = u_n$$

weil $P(G_{>n}) = \sum_{i=n+1}^{\infty} P(G_n)$

(3) Sei $P(L_{2N} = 2n)$ die Wahrscheinlichkeit dass die letzte Rückkehr nach $2n$ Schritten passiert und dass es keine weitere Rückkehr bis nach $2N$ gibt.

$$= u_n * u_{N-n}$$

$$= 2^{-2N} \binom{2n}{n} \binom{2(N-n)}{(N-n)}$$

Mit Hilfe der Stirling Formel

$$n! \sim \sqrt{2\pi n} n^n e^{-n}$$

folgt das Resultat.

Wahrscheinlichkeitstheorie bis Dez 2008

1. Einführung
   Ereignisse, Interpretationen, Axiome

2. Kombinatorik
   Permutationen und Kombinationen
   Urnen, Zellen, Runs

3. Diskrete Zufallsvariablen
   Binomial, Poisson, Geometrisch...

4. Bedingte Wahrscheinlichkeiten
   Unabhängigkeit, Bayes, Genetische Anwendungen

5. Wahrscheinlichkeitserzeugende Funktionen
   Summen von ZV, Verzweigungsprozesse

6. Einschluß-Ausschluß Formel

7. Wahrscheinlichkeitsmaße mit Dichten
   Wahrscheinlichkeitsräume, Verteilungsfunktionen und Dichten, Meßbare Funktionen
   Erwartungswerte

8. Modellierung
   Parameter, Korrelation
   Funktionen von ZV, Kombination von ZV

9. Bedingte Wahrscheinlichkeiten für stetige ZV

10. Momente und Charakteristische Funktionen

11. Ungleichungen
    Markov, Tschebychew

12. Grenzwertsätze
    GGZ(schwach), ZGS, Konvergenz, SGGZ

13. Wie gut ist der ZGS?
    Berry-Esséen, iteriertes Logarithmus, Arcussinus