Zufällige Graphen

Christoph Schumacher

04. Juni 2020

RGCN I, Kap. 5.4:
Grad-Verteilung Erdős-Rényi-Graph

Zusammenfassung

Grad-Verteilung im Erdős-Rényi-Graph

Fragen

1. —

Was bedeutet $n\varepsilon_n^2 \to \infty$? —

$$\varepsilon_n \sqrt{n} = \frac{\varepsilon_n}{1/\sqrt{n}} \to \infty \iff \varepsilon_n = \omega(1/\sqrt{n}) \tag{1}$$

Also: $(\varepsilon_n)_n$ darf gegen 0 konvergieren, aber nicht so schnell wie $(1/\sqrt{n})_n$. Insbesondere ist $\sum_n \varepsilon_n = \infty$, also folgt aus (5.4.3) keine Aussage über die Totalvariation $d_{\mathrm{TV}}\big((P_k^{(n)})_k, (p_k)_k\big) = \frac{1}{2}\sum_k |P_k^{(n)} - p_k|$.

Zwischenschritte zu (5.4.6):

$$\sum_{k=0}^{\infty} |\mathbb{P}(X_n = k) - \mathbb{P}(Y_n = k)| \tag{2}$$

$$= \sum_{k=0}^{\infty} |\mathbb{P}(X_n = k) - \mathbb{P}(X_n = k, I_n = 0) - \mathbb{P}(X_n = k-1, I_n = 1)| \tag{3}$$

$$= \sum_{k=0}^{\infty} |\mathbb{P}(X_n = k) - \mathbb{P}(X_n = k) \underbrace{\mathbb{P}(I_n = 0)}_{=1-\frac{\lambda}{n}} - \mathbb{P}(X_n = k-1) \underbrace{\mathbb{P}(I_n = 1)}_{=\frac{\lambda}{n}}| \tag{4}$$

$$= \frac{\lambda}{n} \sum_{k=0}^{\infty} |\mathbb{P}(X_n = k) - \mathbb{P}(X_n = k-1)| \tag{5}$$

$$\le \frac{\lambda}{n} \sum_{k=0}^{\infty} \big(\mathbb{P}(X_n = k) + \mathbb{P}(X_n = k-1)\big) \tag{6}$$

$$= \frac{2\lambda}{n} \tag{7}$$

Zur letzten Abschätzung in (5.4.7): Wir betrachten eine Kopplung $(\hat{X}, \hat{Y})$ von X^* und Y_n wie in Theorem 2.10, so dass $\mathbb{P}(\hat{X} \neq \hat{Y}) \le \frac{\lambda^2}{n}$. In Kapitel 2.2 über

Kopplungen haben wir den Totalvariationsabstand getroffen. Mit Theorem 2.9 können wir abschätzen:

$$\frac{1}{2}\sum_{k=0}^{\infty}|\mathbb{P}(X^*=k)-\mathbb{P}(Y_n=k)| = d_{\mathrm{TV}}\big(\mathrm{Poi}(\lambda),\mathrm{Bin}(n,\lambda/n)\big) \tag{8}$$

$$\leq \mathbb{P}(\hat{X}\neq\hat{Y}). \tag{9}$$

Da ist jetzt ein zusätzlicher Faktor 2 drin. Die Fußgängermethode ist etwas länger: Um den Betrag abzuschätzen, rechnen wir

$$\mathbb{P}(X^*=k)-\mathbb{P}(Y_n=k) = \mathbb{P}(\hat{X}=k)-\mathbb{P}(\hat{Y}=k) \tag{10}$$

$$\leq \mathbb{P}(\hat{X}=k)-\mathbb{P}(\hat{X}=k,\hat{Y}=k) \tag{11}$$

$$= \mathbb{P}(\hat{X}=k,\hat{Y}\neq k) \tag{12}$$

sowie

$$-\big(\mathbb{P}(X^*=k)-\mathbb{P}(Y_n=k)\big) = \mathbb{P}(\hat{Y}=k)-\mathbb{P}(\hat{X}=k) \tag{13}$$

$$\leq \mathbb{P}(\hat{Y}=k)-\mathbb{P}(\hat{X}=k,\hat{Y}=k) \tag{14}$$

$$= \mathbb{P}(\hat{X}\neq k,\hat{Y}\neq k). \tag{15}$$

Damit können wir zeigen, dass

$$\sum_{k=0}^{\infty}\big|\mathbb{P}(X^*=k)-\mathbb{P}(Y_n=k)\big| \tag{16}$$

$$\leq \sum_{k=0}^{\infty}\max\{\mathbb{P}(\hat{X}=k,\hat{Y}\neq k),\mathbb{P}(\hat{X}\neq k,\hat{Y}\neq k)\} \tag{17}$$

$$\leq \sum_{k=0}^{\infty}\mathbb{P}(\hat{X}=k,\hat{Y}\neq k)+\sum_{k=0}^{\infty}\mathbb{P}(\hat{X}\neq k,\hat{Y}\neq k) \tag{18}$$

$$= \mathbb{P}\Big(\bigcup\nolimits_{k=0}^{\infty}\{\hat{X}=k,\hat{Y}\neq k\}\Big)+\mathbb{P}\Big(\bigcup\nolimits_{k=0}^{\infty}\{\hat{X}\neq k,\hat{Y}=k\}\Big) \tag{19}$$

$$= 2\,\mathbb{P}(\hat{X}\neq\hat{Y}). \tag{20}$$

Warum ist der Faktor 2 da drin? Für ZVn X, Y mit $X=0$ und $Y=1$ fast sicher ist

$$\sum_{k=0}^{\infty}\big|\mathbb{P}(X=k)-\mathbb{P}(Y=k)\big| = \mathbb{P}(X=0)+\mathbb{P}(Y=1) = 2 = 2\,\mathbb{P}(X\neq Y), \tag{21}$$

also brauchen wir hier diesen Faktor 2. Die Rechnung mit dem Totalvariationsabstand legt auch nahe, dass der Faktor 2 sich nicht vermeiden lässt. Zum Glück macht diese zusätzliche 2 im Folgenden keine Probleme.

Warum gilt

$$\frac{2\lambda+2\lambda^2}{n} \leq \frac{\varepsilon_n}{2}? \tag{22}$$

Nach Voraussetzung gilt

$$n\Big(\frac{2\lambda+2\lambda^2}{n}\Big)^2 \xrightarrow{n\to\infty} 0 \quad \text{und} \quad n\Big(\frac{\varepsilon_n}{2}\Big)^2 \xrightarrow{n\to\infty} \infty, \tag{23}$$

also gibt es ein n_0, so dass (22) für alle $n \geq n_0$ stimmt.

Warum genügt (5.4.8)?

$$\begin{aligned}
&\mathbb{P}_\lambda\big(\max_{k\geq 0}|P_k^{(n)} - p_k| \geq \varepsilon_n\big) && (24)\\
&\leq \mathbb{P}_\lambda\big(\max_{k\geq 0}\big(\big|P_k^{(n)} - \mathbb{E}[P_k^{(n)}]\big| + \big|\mathbb{E}[P_k^{(n)}] - p_k\big|\big) \geq \varepsilon_n\big) && (25)\\
&\leq \mathbb{P}_\lambda\big(\max_{k\geq 0}\big|P_k^{(n)} - \mathbb{E}[P_k^{(n)}]\big| + \varepsilon_n/2 \geq \varepsilon_n\big) && (26)\\
&= \mathbb{P}_\lambda\big(\max_{k\geq 0}\big|P_k^{(n)} - \mathbb{E}[P_k^{(n)}]\big| \geq \varepsilon_n/2\big). && (27)
\end{aligned}$$

Zu (5.4.11):

$$\begin{aligned}
\mathrm{Var}_\lambda(P_k^{(n)}) &= \mathrm{Cov}_\lambda\Big(\frac{1}{n}\sum\nolimits_{i\in[n]} \mathbf{1}_{\{D_i=k\}}, \frac{1}{n}\sum\nolimits_{j\in[n]} \mathbf{1}_{\{D_j=k\}}\Big) && (28)\\
&= \frac{1}{n^2}\sum_{i,j\in[n]} \mathrm{Cov}_\lambda(\mathbf{1}_{\{D_i=k\}}, \mathbf{1}_{\{D_j=k\}}) && (29)\\
&= \frac{1}{n^2}\big(n\,\mathrm{Var}_\lambda(\mathbf{1}_{\{D_1=k\}}) + n(n-1)\,\mathrm{Cov}_\lambda(\mathbf{1}_{\{D_1=k\}}, \mathbf{1}_{\{D_2=k\}})\big) && (30)\\
&= \frac{1}{n}\big(\mathbb{P}_\lambda(D_1=k) - (\mathbb{P}_\lambda(D_1=k))^2\big)\\
&\quad + \frac{n-1}{n}\big(\mathbb{P}_\lambda(D_1=D_2=k) - \mathbb{P}_\lambda(D_1=k)\,\mathbb{P}_\lambda(D_2=k)\big) && (31)
\end{aligned}$$

Zu (5.4.17): Wir kürzen ab: $A := \{(X_1+I_1, X_2+I_1) = (k,k)\}$, $B := \{(X_1+I_1, X_2+I_2) = (k,k)\}$:

$$\begin{aligned}
\mathbb{P}_\lambda(D_1=D_2=k) - \big(\mathbb{P}_\lambda(D_2=k)\big)^2 &= \mathbb{P}(A) - \mathbb{P}(B) && (32)\\
&\leq \mathbb{P}(A) - \mathbb{P}(A\cap B) && (33)\\
&= \mathbb{P}(A\setminus B) && (34)
\end{aligned}$$

Zu (5.4.15):

$$\begin{aligned}
\mathbb{P}(A\setminus B) &= \mathbb{P}(I_1=1)\,\mathbb{P}(A\setminus B \mid I_1=1) + \mathbb{P}(I_1=0)\,\mathbb{P}(A\setminus B \mid I_1=0) && (35)\\
&= \frac{\lambda}{n}\,\mathbb{P}(X_1=X_2=k-1, I_2=0 \mid I_1=1)\\
&\quad + \Big(1-\frac{\lambda}{n}\Big)\,\mathbb{P}(X_1=X_2=k, I_2=1 \mid I_1=0) && (36)\\
&= \frac{\lambda}{n}\,\underbrace{\mathbb{P}(X_1=k-1)}_{\leq 1}\,\mathbb{P}(X_2=k-1)\,\underbrace{\mathbb{P}(I_2=0)}_{\leq 1}\\
&\quad + \underbrace{\Big(1-\frac{\lambda}{n}\Big)}_{\leq 1}\,\mathbb{P}(X_1=k)\,\underbrace{\mathbb{P}(X_2=k)}_{\leq 1}\,\underbrace{\mathbb{P}(I_2=1)}_{=\lambda/n} && (37)\\
&\leq \frac{\lambda}{n}\big(\mathbb{P}(X_2=k-1) + \mathbb{P}(X_1=k)\big) && (38)
\end{aligned}$$

Aufgabe 5.16: Es sei X^* eine Poi(λ)-verteilte ZV. Zeigen Sie: Für alle $A>0$ gibt es ein $C = C(A)$ so dass für alle $x\in\mathbb{N}_0$ gilt, dass

$$\mathbb{P}(X^*\geq x) \leq C\mathrm{e}^{-Ax}. \qquad (39)$$

Möglichkeit 1 (Large Deviations): Für alle $x \geq \lambda$ gilt

$$\begin{aligned}
\mathbb{P}(X^* \geq x) &\leq \inf_{s\geq 0} \mathrm{e}^{-sx}\, \mathbb{E}[\mathrm{e}^{sX^*}] && (40)\\
&= \inf_{s\geq 0} \mathrm{e}^{-sx} \exp((\mathrm{e}^s - 1)\lambda) && (41)\\
&= \inf_{s\geq 0} \exp(-sx + \lambda \mathrm{e}^s - \lambda) && (42)\\
&\overset{\mathrm{e}^s = \frac{x}{\lambda} \geq 1}{=} \exp\Big(-x \log \frac{x}{\lambda} + x - \lambda\Big) && (43)\\
&= \mathrm{e}^{-I_\lambda(x)} && (44)
\end{aligned}$$

mit $I_\lambda(x) = \lambda - x + x\log(x/\lambda)$, der Ratenfunktion der Poisson-Verteilung. Es gilt

$$\frac{I_\lambda(x)}{x} = \frac{\lambda}{x} - 1 + \log(x/\lambda) \xrightarrow{x\to\infty} \infty, \tag{45}$$

also finden wir ein $x_0 \geq \lambda$ mit $I_\lambda(x) \geq Ax$ für alle $x \geq x_0$. Wir setzen

$$C(A) := 1 \vee \max_{x\in\{0,1,\dots,x_0\}} \mathrm{e}^{Ax}\, \mathbb{P}(X^* \geq x) \tag{46}$$

und erkennen

$$\mathbb{P}(X^* \geq x) \leq C(A)\mathrm{e}^{-Ax}. \tag{47}$$

Möglichkeit 2 (Exponentialreihe und Sterling-Formel):

$$\begin{aligned}
\mathbb{P}(X^* \geq x) &= \mathrm{e}^{-\lambda} \sum_{k=x}^{\infty} \frac{\lambda^k}{k!} && (48)\\
&\overset{j:=k-x}{=} \mathrm{e}^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^{j+x}}{(j+x)!} && (49)\\
&\leq \mathrm{e}^{-\lambda} \frac{\lambda^x}{x!} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = \frac{\lambda^x}{x!} && (50)\\
&\leq \frac{\lambda^x}{(x/\mathrm{e})^x} = \Big(\frac{\lambda \mathrm{e}}{x}\Big)^x && (51)\\
&= \exp\Big(-x \log \frac{x}{\lambda \mathrm{e}}\Big) && (52)\\
&\leq \mathrm{e}^{-Ax} && (53)
\end{aligned}$$

für alle x mit $\log \frac{x}{\lambda \mathrm{e}} \geq A$, also für alle $x \geq \lambda \mathrm{e}^{1+A}$. Jetzt wählt man wieder

$$C := 1 \vee \max_{x\in\mathbb{N}_0\cap[0,\lambda\mathrm{e}^{1+A})} \mathrm{e}^{Ax}\, \mathbb{P}(X^* \geq x) \tag{54}$$

und erhält die gewünschte Ungleichung.