Tentamen Toegepaste Stochastiek

28 Juni 2018, 8:30-11:30

- Er zijn 90 punten te verdienen. U krijgt er 10 gratis.
- Maak elke opgave op een apart vel, vergeet niet uw naam en studienummer op elk vel te vermelden.

Opgave 1 (8+8=16 punten)

Beschouw de volgende twee uitspraken (waarbij $n \to \infty$):

$$f(n) = (n + o(n))^{\frac{1}{2}} \qquad \text{en} \qquad f(n) = n^{\frac{1}{2}+o(1)}.$$

(a) Bewijs of weerleg dat de eerste uitspraak de tweede impliceert.

(b) Bewijs of weerleg dat de tweede uitspraak de eerste impliceert.

Uitwerking

(a) Neem aan dat de eerste uitspraak waar is. Dan bestaat er een functie $g(n)$ met

$$\lim_{n\to\infty} \frac{g(n)}{n} = 0 \qquad \text{en} \qquad f(n) = (n + g(n))^{\frac{1}{2}}.$$

Derhalve

$$f(n) = \left(1 + \frac{g(n)}{n}\right)^{\frac{1}{2}} n^{\frac{1}{2}}.$$

De eerste term gaat naar 1, dus er bestaat n_0 zodat voor alle $n > n_0$ geldt

$$n^{\frac{1}{2}-\frac{\log 2}{\log n}} = \frac{1}{2} n^{\frac{1}{2}} < f(n) < 2n^{\frac{1}{2}} = n^{\frac{1}{2}+\frac{\log 2}{\log n}}.$$

Dus inderdaad $f(n) = n^{\frac{1}{2}+o(1)}$, waarmee is aangetoond dat de eerste uitspraak de tweede impliceert.

(b) Neem

$$f(n) = 3n^{\frac{1}{2}} = n^{\frac{1}{2}+\frac{\log 3}{\log n}} = n^{\frac{1}{2}+o(1)}.$$

Dus deze functie voldoet aan de tweede uitspraak. Er geldt tevens

$$f(n) = (n + 8n)^{\frac{1}{2}}.$$

Het is duidelijk dat $8n \neq o(n)$, dus ook $f(n) \neq (n + o(n))^{\frac{1}{2}}$.

Opgave 2 (16 punten)

Zij $\mathcal{F}$ een eindige collectie binaire rijtjes met eindige lengte. Neem aan dat geen enkel rijtje in deze collectie een prefix is van een ander rijtje in de collectie. D.w.z. voor $u, v \in \mathcal{F}$ met $|u| < |v|$ is u ongelijk aan het beginstuk van v. Zij N_k het aantal rijtjes in $\mathcal{F}$ met lengte k. Bewijs dat

$$\sum_{k=1}^{\infty} \frac{N_k}{2^k} \leq 1.$$

Uitwerking

Zij $\omega = (\omega_i)_{i=1}^{\infty}$ een willekeurige oneindige binaire rij. Zij $\omega^k, k \geq 1$ de prefix van deze rij met lengte k. Definieer $X_k = \mathbb{1}_{\{\omega^k \in \mathcal{F}\}}$ en

$$X = \#\left\{k : \omega^k \in \mathcal{F}\right\} = \sum_{k=1}^{\infty} X_k.$$

Dan geldt

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} \mathbb{E}[X_k] = \sum_{k=1}^{\infty} \mathbb{P}(\omega^k \in \mathcal{F}) = \sum_{k=1}^{\infty} \frac{N_k}{2^k}.$$

Verder is duidelijk dat er ten hoogste één prefix van ω in $\mathcal{F}$ kan zitten, dus $X \leq 1$, en dus ook $\mathbb{E}[X] \leq 1$, waarmee het gevraagde bewezen is.

Opgave 3 (5+16 = 21 punten)

Voor een graaf $G = (V, E)$ met $n = 2k$ knopen definiëren we de graafeigenschap $\mathcal{P}$ als volgt:

Er bestaat $U \subseteq V$ met $|U| = k$ zodat $(U \times U^c) \cap E = \emptyset$.

Neem nu voor G de Erdős-Renyi graaf $G_{n,p}$ met n even.

(a) Is de eigenschap $\mathcal{P}$ monotoon? Zo ja, dalend of stijgend?

(b) Laat zien dat voor

$$p \geq 1 - \left(\frac{1}{4}\right)^{1/k}$$

geldt dat $G_{n,p}$ asymptotisch bijna zeker (a.a.s.) *niet* de eigenschap $\mathcal{P}$ heeft.

U mag gebruik maken van Stirling's benadering: $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$.

Uitwerking

(a) Als G_1 eigenschap $\mathcal{P}$ niet heeft en $G_1 \subseteq G_2$, dan heeft G_2 eigenschap $\mathcal{P}$ ook niet. Anders gezegd, als we kanten toevoegen aan G_1, kan er nooit een graaf ontstaan die $\mathcal{P}$ wel heeft. Dus $\mathcal{P}$ is monotoon dalend.

Equivalent: Als G_1 eigenschap $\mathcal{P}$ heeft en $G_2 \subseteq G_1$, dan heeft G_2 eigenschap $\mathcal{P}$ ook. Anders gezegd, als we kanten verwijderen uit G_1, kan er nooit een graaf ontstaan die $\mathcal{P}$ niet heeft. Dus $\mathcal{P}$ is monotoon dalend.

(b) We moeten laten zien dat

$$\lim_{n\to\infty} \mathbb{P}(G_{n,p} \text{ heeft } \mathcal{P}) = 0,$$

dus we willen deze kans van boven begrenzen. Voor vaste U met $|U| = k$ geldt $|U \times U^c| = k^2$, dus er zijn k^2 potentiële kanten van U naar U^c. Derhalve

$$\mathbb{P}((U \times U^c) \cap E = \emptyset) = (1-p)^{k^2}.$$

Nu vinden we

$$\begin{aligned}
\mathbb{P}(G \text{ heeft } \mathcal{P}) &\leq \sum_{|U|=k} \mathbb{P}((U \times U^c) \cap E = \emptyset) \\
&= \binom{2k}{k}(1-p)^{k^2} = \frac{(2k)!}{k!k!}(1-p)^{k^2} \\
&\sim \frac{\sqrt{4\pi k}}{2\pi k}\left(\frac{2k}{e}\right)^{2k}\left(\frac{e}{k}\right)^{2k}(1-p)^{k^2} \\
&= \frac{1}{\sqrt{\pi k}} 4^k (1-p)^{k^2}.
\end{aligned}$$

Als $p \geq 1 - \left(\frac{1}{4}\right)^{1/k}$, dan is dit ten hoogste $1/\sqrt{\pi k}$, dus gaat naar 0.

Als u dit niet meteen ziet, neemt u de logaritme

$$\begin{aligned}
\log(\mathbb{P}(G \text{ heeft } \mathcal{P})) &\leq -\frac{1}{2}\log(\pi k) + k\log(4) + k^2\log(1-p) \\
&= -\frac{1}{2}\log(\pi k) + k(\log(4) + k\log(1-p))
\end{aligned}$$

Als $\log(4) + k\log(1-p) \leq 0$, dan

$$\lim_{n\to\infty} \log(\mathbb{P}(G \text{ heeft } \mathcal{P})) = -\infty, \qquad \text{en dus} \qquad \lim_{n\to\infty} \mathbb{P}(G \text{ heeft } \mathcal{P}) = 0.$$

Herschrijven van de voorwaarde geeft $4(1-p)^k \leq 1$, en dus

$$p \geq 1 - \left(\frac{1}{4}\right)^{1/k}.$$

Opgave 4 (4+4+4+10+15=37 punten)

Beschouw het Barabási-Albert model met $m = 1$. De graaf G_t heeft knopen $v_0, \ldots, v_t$. Uit deze graaf ontstaat G_{t+1} door de nieuwe knoop v_{t+1} met 1 kant te verbinden met een bestaande knoop, met kansen evenredig met de graden. De graaf G_0 bestaat uit één knoop v_0 en G_1 bestaat uit twee knopen, v_0 en v_1, verbonden door 1 kant. Zij $D_i(t)$ de graad van v_i in G_t voor $t \geq i$. Voor $t < i$ definiëren we $D_i(t) = 0$.

Definieer de *mediaanknoop* m_t van G_t als volgt:

$$m_t := \min \left\{ s : \sum_{i=0}^{s} D_i(t) \geq t \right\}.$$

(a) Toon aan dat $m_t \leq m_{t+1}$ voor alle $t \geq 0$.

(b) Zij $k \in \mathbb{N}$ vast. Definieer $X_t = \sum_{i=0}^{k}(D_i(t+1) - D_i(t))$.

Bepaal X_t voor $t = 0, \ldots, k-1$. (Komt u er niet uit, neem dan in het vervolg $X_t = k+3$.)

(c) Laat zien dat X_t voor $t \geq k$ gelijk is aan 0 of 1, waarbij $\mathbb{P}(X_t = 1) \leq \frac{1}{2} + \frac{k}{2t}$.

(d) Toon aan dat $\sum_{i=0}^{k} D_i(t) = \sum_{s=0}^{t-1} X_s$. Gebruik dit om te laten zien dat

$$\mathbb{E}\left[\sum_{i=0}^{k} D_i(t)\right] \leq \frac{1}{2}t + o(t).$$

(e) Definieer voor $t \geq k$ onafhankelijke stochasten $Y_t \sim \text{Ber}(\frac{1}{2} + \frac{k}{2t})$. Toon aan dat asymptotisch bijna zeker (a.a.s.) geldt $\lim_{t\to\infty} m_t = \infty$.

U mag gebruiken dat $\mathbb{P}\left(\sum_{s=k}^{t-1} X_s \geq a\right) \leq \mathbb{P}\left(\sum_{s=k}^{t-1} Y_s \geq a\right)$ voor elke $a \in \mathbb{R}$ en $t > k$.

Uitwerking

(a) Een informeel argument: de mediaanknoop is de oudste knoop waarvoor geldt dat minstens de helft van de halfkanten verbonden is met die knoop of een oudere. Van elke nieuwe kant komt in ieder geval één van de twee uiteinden aan een knoop die nieuwer is dan m_t, waaruit volgt dat m_t niet kan dalen. Een formeel argument:

$$\sum_{i=0}^{m_t-1} D_i(t+1) \leq 1 + \sum_{i=0}^{m_t-1} D_i(t) < 1 + t.$$

Daaruit volgt dat $m_t - 1 < m_{t+1}$, en dus $m_t \leq m_{t+1}$.

(b) X_t is de toename van de som van graden van de knopen $v_0, \ldots, v_k$ als we de knoop v_{t+1} en een nieuwe kant toevoegen aan G_t. Als $t < k$, dan is $v_{t+1} \in \{v_0, \ldots, v_k\}$. Met andere woorden, de nieuwe kant verbindt twee van de knopen $v_0, \ldots, v_k$ met elkaar, en dus is de toename van de som gelijk aan 2. Dus $X_t = 2$ voor $t = 0, \ldots, k-1$.

(c) Als $t \geq k$, dan $v_{t+1} \notin \{v_0, \ldots, v_k\}$. Eén uiteinde van de nieuwe kant is dus in ieder geval *niet* verbonden met een van de knopen $v_0, \ldots, v_k$. Het andere uiteinde misschien wel. De som van de graden neemt dus met 0 of 1 toe. De som van alle graden in G_t is $2t$ en elke knoop heeft minstens graad 1. De som van de graden van $v_0, \ldots, v_k$ in G_t is dus ten hoogste $2t - (t-k)$. De kans dat v_{t+1} verbonden wordt met één van deze knopen is derhalve begrensd door

$$\mathbb{P}(X_t = 1) = \frac{\sum_{i=0}^k D_i(t)}{2t} \leq \frac{2t - (t-k)}{2t} = \frac{1}{2} + \frac{k}{2t}.$$

(d) We tonen eerst de gelijkheid van de twee sommen aan:

$$\begin{aligned}\sum_{s=0}^{t-1} X_s &= \sum_{s=0}^{t-1}\sum_{i=0}^{k}(D_i(s+1) - D_i(s)) = \sum_{i=0}^{k}\sum_{s=0}^{t-1}(D_i(s+1) - D_i(s)) \\ &= \sum_{i=0}^{k}(D_i(t) - D_i(0)) = \sum_{i=0}^{k} D_i(t).\end{aligned}$$

We gebruiken vervolgens (b) en (c). Omdat $t \to \infty$ mogen we aannemen dat $t > k$.

$$\begin{aligned}\mathbb{E}\left[\sum_{s=0}^{t-1} X_s\right] &= \mathbb{E}\left[\sum_{s=0}^{k-1} X_s + \sum_{s=k}^{t-1} X_s\right] \leq 2k + \sum_{s=k}^{t-1}\left(\frac{1}{2} + \frac{k}{2s}\right) \\ &\leq 2k + \frac{1}{2}(t-k) + \frac{k}{2}\sum_{s=k}^{t-1}\frac{1}{s} \leq \frac{1}{2}t + \frac{3}{2}k + \frac{k}{2}\int_{s=k-1}^{t-1}\frac{1}{s}ds \\ &= \frac{1}{2}t + \frac{3}{2}k + \frac{k}{2}(\ln(t-1) - \ln(k-1)) = \frac{1}{2}t + o(t).\end{aligned}$$

(e) Merk eerst op dat

$$\mathbb{P}(m_t \leq k) = \mathbb{P}\left(\sum_{i=0}^{k} D_i(t) \geq t\right) = \mathbb{P}\left(\sum_{s=0}^{t-1} X_s \geq t\right).$$

Als deze kans naar 0 gaat voor elke k, dan $m_t \to \infty$ bijna zeker. Met behulp van de gegeven stochasten Y_t vinden we voor $t > 2k$

$$\begin{aligned}
\mathbb{P}\left(\sum_{s=0}^{t-1} X_s \geq t\right) &= \mathbb{P}\left(\sum_{s=k}^{t-1} X_s \geq t - 2k\right) \\
&\leq \mathbb{P}\left(e^{\sum_{s=k}^{t-1} Y_s} \geq e^{t-2k}\right) \\
&\leq \frac{\mathbb{E}[e^{\sum_{s=k}^{t-1} Y_s}]}{e^{t-2k}} \\
&= e^{2k-t} \prod_{s=k}^{t-1} \mathbb{E}[e^{Y_s}].
\end{aligned}$$

Hier hebben we Markov's ongelijkheid gebruikt en het feit dat de stochasten Y_t onafhankelijk zijn. Verder geldt

$$\mathbb{E}[e^{Y_s}] = e \cdot \mathbb{P}(Y_s = 1) + 1 \cdot \mathbb{P}(Y_s = 0),$$

dus dit is altijd kleiner dan e. Voor $s \to \infty$ gaan de kansen naar $\frac{1}{2}$, dus de verwachting naar $\frac{e+1}{2}$. In het bijzonder geldt

$$\mathbb{E}[e^{Y_s}] = \left(\frac{1}{2} + \frac{k}{2s}\right) e + \left(\frac{1}{2} - \frac{k}{2s}\right) \leq \begin{cases} e, & k \leq s < 2k \\ \frac{3}{4}e + \frac{1}{4}, & s \geq 2k \end{cases}$$

Hiermee volgt

$$\begin{aligned}
\mathbb{P}\left(\sum_{s=0}^{t-1} X_s \geq t\right) &\leq e^{2k-t} \prod_{s=k}^{2k-1} \mathbb{E}[e^{Y_s}] \prod_{s=2k}^{t-1} \mathbb{E}[e^{Y_s}] \\
&\leq \frac{e^k \cdot \left(\frac{3}{4}e + \frac{1}{4}\right)^{t-2k}}{e^{t-2k}} = e^k \left(\frac{3e+1}{4e}\right)^{t-2k} \xrightarrow{t\to\infty} 0,
\end{aligned}$$

waarmee het gevraagde volgt.

EINDE