

Toets 4 Inleiding Kansrekening wi1614

Rice Chapter 1-5, Markov chains Section 1.1&1.2.

3 mei 2010, 15:45-16:45 uur

Bij deze toets is het gebruik van een (eventueel grafische) rekenmachine en van het formuleblad wi1614 toegestaan.

Een antwoord alleen is *niet* voldoende: er dient een berekening, toelichting en/of motivatie aanwezig te zijn. Dit alles goed leesbaar en in goed Nederlands (of goed Engels).

Dit jaar mag er gekozen worden: óf Opgave 4. maken (Stof: Ch.1-5 Rice + begin Markovketens), óf Opgave 5. maken (Stof: alleén Ch.1-5 Rice).

Normering: Vragen 1 en 2 en 3a, 3b, 3c, 4a, 4c 5b en 5c tellen voor 10 punten, 4b voor 20, 5a voor 5 en 5d voor 15 punten.

Je krijgt 10 punten cadeau. Cijfer: aantal punten gedeeld door 10.

1. Laten A en B gebeurtenissen zijn met $P(A) = \frac{1}{3}$, $P(B) = \frac{1}{5}$ en $P(A|B) = \frac{1}{4}$.

Wat is $P(B|A)$?

2. Gegeven zijn onafhankelijke stochasten X en Y die uniform verdeeld zijn over $[0, 1]$.

Bereken $E[|X - Y|]$.

Hint: Verdeel het eenheidsvierkant in de twee driehoeken waar $x > y$, resp. $x < y$.

3. Laat $(X_1), (X_2), \dots$ een rij onafhankelijke stochasten zijn allemaal met een uniforme verdeling over het interval $[0, 1]$.

- a. Wat zegt de Wet van de Grote Aantallen over het gedrag van

$$\frac{1}{n} \sum_{i=1}^n X_i \quad \text{als } n \rightarrow \infty?$$

- b. Wat zegt diezelfde Wet van de Grote Aantallen over het gedrag van

$$\frac{1}{n} \sum_{i=1}^n |X_{2i} - X_{2i-1}| \quad \text{als } n \rightarrow \infty?$$

- c. Als we willen weten hoe

$$\frac{1}{n} \sum_{i=1}^n |X_{i+1} - X_i|$$

zich gedraagt als $n \rightarrow \infty$, dan kunnen we niet zonder meer de Wet van de Grote Aantallen gebruiken. Waarom niet?

Kun je tóch een onderbouwde uitspraak doen over het gedrag van deze gewogen sommen $(\sum_{i=1}^n |X_{i+1} - X_i|)/n$?

Aanwijzing: Je mag zonder bewijs gebruiken dat als (Y_i) en (Z_i) in kans convergeren, dat dan ook $(Y_i + Z_i)$ in kans convergeert, naar de som van de twee limieten.

4. We bekijken een fysisch wat realistischer variant van het Ehrenfest model aangaande de N moleculen verdeeld over twee vaten A en B met een kleine opening er tussen.

Zij p een getal met $0 \leq p \leq 1$. Tussen tijdstip n en $n + 1$ gaat met kans p elk molecuul door de kleine opening naar het andere vat, onafhankelijk van de andere $N - 1$ moleculen (elk molecuul in A gaat dus met kans p naar B, en elk molecuul in B met kans p naar A).

- a. Wat zijn de kansverdelingen van de stochasten W_{AB}^j , die aangeven hoeveel moleculen er tussen tijdstip n en $n + 1$ van vat A naar vat B gaan als er op tijdstip n in vat A j moleculen waren? Hier is $0 \leq j \leq N$.

- b. Zij (X_n) de Markovketen met als toestanden het aantal deeltjes in kamer A. Geef de overgangsmatrix P van deze keten als $N = 2$.

Aanwijzing: Bijvoorbeeld $p_{10} = P(W_{AB}^1 = 1) P(W_{BA}^1 = 0) = p(1 - p)$, omdat het ene molecuul in vat A dan naar B moet, en onafhankelijk daarvan het ene molecuul in vat B dan in vat B moet blijven.

- c. Bewijs dat nog steeds (voor alle p) de $Bin(2, 1/2)$ verdeling de stationaire verdeling is van de Markovketen (X_n) (doe dit dus alleen als $N = 2$).

5. Gegeven is dat de continue stochast X een kansdichtheid f heeft, die gegeven wordt door:

$$f(x) = \begin{cases} 1 + x & \text{als } -1 \leq x \leq 0 \\ 1 - x & \text{als } 0 \leq x \leq 1 \\ 0 & \text{als } x < -1 \text{ of } x > 1. \end{cases}$$

- a. Schets de dichtheid f en bereken de verdelingsfunctie F van X .

- b. Bereken de variantie van X .

Zij $X_1, X_2, \dots, X_{100}$ een steekproef uit de verdeling met kansdichtheid f .

Als u onderdeel **b.** niet hebt kunnen doen, reken dan bij de onderdelen **c.** en **d.** verder met variantie $6/11$.

- c. Geef een schatting van de kans $P(|\bar{X}_{100} - E[\bar{X}_{100}]| > 0.6) = P(|\bar{X}_{100}| > 0.6)$ met behulp van de ongelijkheid van Chebyshev.

- d. Benader de kans $P(|\bar{X}_{100}| > 0.6)$ met behulp van de centrale limietstelling (*central limit theorem*).

Antwoorden:

1 $P(B|A) = P(A \cap B) / P(A) = P(A|B)P(B) / P(A) = \frac{1}{4} \cdot \frac{1}{5} / \frac{1}{3} = \frac{3}{20}.$

2 Direct rekenen:

$$\begin{aligned} E[|X - Y|] &= \int_0^1 \int_0^1 |x - y| \, dy \, dx = \int_0^1 \int_0^x (x - y) \, dy \, dx + \int_0^1 \int_x^1 (y - x) \, dy \, dx \\ &= 2 \int_0^1 \int_0^x (x - y) \, dy \, dx = 2 \int_0^1 \int_0^x (x^2 - \frac{1}{2}x^2) \, dx = 2 \frac{1}{6}x^3 \Big|_0^1 = \frac{1}{3}. \end{aligned}$$

3a De (X_i) zijn onafhankelijke stochasten, met

$$E[X_i] = \frac{1}{2}.$$

Dus het steekproefgemiddelde $\bar{X}_n$ convergeert in kans naar $E[X_1] = \frac{1}{2}$:

$$\forall \varepsilon > 0 \quad \lim_{n \rightarrow \infty} P\left(\left|\bar{X}_n - \frac{1}{2}\right| > \varepsilon\right) \rightarrow 0.$$

N.B. De manier waarop de Wet van de Grote Aantallen geformuleerd is vereist dat $\text{Var}(X_1) < \infty$, maar hier is zeker aan voldaan.

3b Zij $Y_i = |X_{2i} - X_{2i-1}|$. Dan zijn de (Y_i) zijn onafhankelijke stochasten, met (zie Opgave 2)

$$E[Y_i] = E[|X_{2i} - X_{2i-1}|] = \frac{1}{3}.$$

Dus het steekproefgemiddelde $\bar{Y}_n$ convergeert in kans naar $E[Y_1] = \frac{1}{3}$:

$$\forall \varepsilon > 0 \quad \lim_{n \rightarrow \infty} P\left(\left|\bar{Y}_n - \frac{1}{3}\right| > \varepsilon\right) \rightarrow 0.$$

(En weer is voldaan aan $\text{Var}(Y_1) < \infty$.)

3c De stochasten $|X_{i+1} - X_i|$ en $|X_{i+2} - X_{i+1}|$ zijn niet onafhankelijk!

Neem $Y_i = |X_{2i} - X_{2i-1}|$, en $Z_i = |X_{2i+1} - X_{2i}|$. Dan zijn de (Y_i) , en ook de (Z_i) onafhankelijke stochasten, met $E[Y_i] = E[Z_i] = \frac{1}{3}$. Dus de steekproefgemiddelden $\bar{Y}_n$, en $\bar{Z}_n$ convergeren in kans naar $E[Y_1] = E[Z_1] = \frac{1}{3}$. Schrijf dan voor even n :

$$\frac{1}{n} \sum_{i=1}^n |X_{i+1} - X_i| = \frac{1}{2} \frac{1}{n/2} \sum_{i=1}^{n/2} Y_i + \frac{1}{2} \frac{1}{n/2} \sum_{i=1}^{n/2} Z_i.$$

Volgens de aanwijzing convergeren die naar $\frac{1}{2} \frac{1}{3} + \frac{1}{2} \frac{1}{3} = \frac{1}{3}$. Bij oneven n doen we iets dergelijks. Conclusie: $(\sum_{i=1}^n |X_{i+1} - X_i|)/n$ convergeert ook naar $\frac{1}{3}$.

4a De W_{AB}^j zijn sommen van j onafhankelijke Bernoulli's allemaal met parameter p , dus de W_{AB}^j zijn $\text{Bin}(j, p)$ verdeeld.

4b

$$P = \begin{pmatrix} (1-p)^2 & 2p(1-p) & p^2 \\ p(1-p) & (1-p)^2 + p^2 & p(1-p) \\ p^2 & 2p(1-p) & (1-p)^2 \end{pmatrix}.$$

4c We moeten laten zien dat $[\frac{1}{4}, \frac{1}{2}, \frac{1}{4}]$ een linkereigenvector van P is bij de eigenwaarde 1.
Dit volgt uit

$$\frac{1}{4} \cdot (1-p)^2 + \frac{1}{2} \cdot 2p(1-p) + \frac{1}{4} \cdot p^2 = \frac{1}{4} \cdot (1-p+p)^2 = \frac{1}{4},$$

en

$$\frac{1}{4} \cdot 2p(1-p) + \frac{1}{2} \cdot ((1-p)^2 + p^2) + \frac{1}{4} \cdot 2p(1-p) = \frac{1}{2} \cdot (1-p+p)^2 = \frac{1}{2}.$$

5a f is een driehoeksverdeling tussen -1 en $+1$. Omdat alle massa op $[-1, 1]$ ligt, is $F(x) = 0$ als $x < -1$, en $F(x) = 1$ als $x > 1$. Verder volgt door integreren dat $F(x) = \frac{1}{2}x^2 + x + \frac{1}{2}$ als $-1 \leq x \leq 0$ en $F(x) = -\frac{1}{2}x^2 + x + \frac{1}{2}$ als $0 \leq x \leq 1$.

5b Dit kan heel snel door te bedenken dat de driehoeksverdeling op $(0,2)$ de som is van twee uniform $(0,1)$ verdelingen, dus de variantie is twee keer zo groot (verschuiven naar $(-1,1)$ heeft geen invloed op de variantie): $2 \cdot 1/12 = 1/6$.

Direct uitrekenen: $E[X] = 0$, dus

$$\text{Var}(X) = E[X^2] = \int_{-1}^0 x^2(1+x) dx + \int_0^1 x^2(1-x) dx = \frac{1}{6}.$$

5c Invullen:

$$P(|\bar{X}_{100} - E[\bar{X}_{100}]| > 0.6) \leq \frac{\text{Var}(\bar{X}_{100})}{[0.6]^2} = \frac{1}{216} = 0.004629...,$$

want $\text{Var}(\bar{X}_{100}) = \text{Var}(X_1)/100 = 1/600$.

5d We zagen al dat $E[\bar{X}_{100}] = 0$, dus als Z standaard normaal verdeeld is, dan is

$$\begin{aligned} P(|\bar{X}_{100}| > 0.6) &= P\left(\frac{|\bar{X}_{100}|}{\sqrt{\text{Var}(\bar{X}_{100})}} > \frac{0.6}{\sqrt{\text{Var}(\bar{X}_{100})}}\right) \\ &= P\left(\frac{|\bar{X}_{100}|}{\sqrt{1/600}} > \frac{0.6}{\sqrt{1/600}}\right) \\ &\approx P\left(|Z| > \frac{0.6}{\sqrt{1/600}}\right) \\ &= P(Z > 6\sqrt{6}) - P(Z < -6\sqrt{6}) \\ &= P(Z > 14.69...) - P(Z < -14.69...) \\ &= 0.000000000000.... \end{aligned}$$

(N.B. Deze kans staat niet meer in de tabel, maar is natuurlijk heel klein; je weet so wie so dat hij veel kleiner zal zijn dan het antwoord bij c.)