Tentamen Inleiding Kansrekening

06 juni 2023, 14.00–17.00 uur

Docent: Prof. dr. F. den Hollander

---

Bij dit tentamen is het gebruik van boek, aantekeningen en rekenmachine *niet* toegestaan. Wel is het toegestaan om een persoonlijke lijst van belangrijke notaties, definities en stellingen te consulteren.

Er zijn 8 vragen, waarvan sommige bestaan uit onderdelen. Elk onderdeel is een aantal punten waard, dat vet gedrukt is aangegeven. Het totaal aantal punten is 100. Er zijn vragen in drie categorieën: **R** = reproductie, **T** = toepassing, **I** = inzicht.

Schrijf je naam, studentnummer en studierichting op *elk* blad dat je inlevert. Motiveer steeds je antwoorden: een los antwoord zonder uitleg is *niet* voldoende. Elke berekening dient van een *toelichting* te worden voorzien. Het is niet nodig om formele bewijzen te geven.

---

(1) [**R**][**8**] Zij $\mathcal{F}$ een gebeurtenissenruimte. Laat zien dat voor elke $B \in \mathcal{F}$ de verzameling $\mathcal{G} = \{A \cap B\colon\ A \in \mathcal{F}\}$ opnieuw een gebeurtenissenruimte is.

(2) [**R**][**10**] Een zuivere dobbelsteen wordt 3 keer gegooid. Zij $A_{ij}$ de gebeurtenis dat worp $i$ en worp $j$ dezelfde uitkomst opleveren. Laat zien dat $A_{ij}$, $1 \le i < j \le 3$, paarsgewijs onafhankelijk zijn, maar niet tripelsgewijs onafhankelijk.

(3) [**T**] Op zaterdag en zondag doe ik boodschappen. Op zaterdag ga ik met kans $\frac{1}{3}$ naar Jumbo (J) en met kans $\frac{2}{3}$ naar Albert Heijn (AH). Gegeven dat ik op zaterdag naar J ga, kies ik zondag voor J met kans $\frac{1}{5}$ en voor AH met kans $\frac{4}{5}$. Gegeven echter dat ik zaterdag naar AH ga, kies ik op zondag voor AH met kans $\frac{1}{4}$ en voor J met kans $\frac{3}{4}$.

(a) [**8**] Bereken de kans dat ik op zondag naar J ga.

(b) [**8**] Bereken de kans dat ik op zaterdag naar AH ga, gegeven dat ik op zondag naar J ga.

(4) [**I**] Een machine gaat kapot na $T$ dagen, waarbij $T$ kansmassafunctie $f_T$ heeft. Gegeven dat de machine na $t$ dagen nog werkt, bereken het gemiddeld aantal dagen na dag $t$ dat de machine nog blijft werken wanneer $f_T$ gelijk is aan:

(a) [**8**] $f_T(s) = 1/N$, $s \in \{1, \ldots, N\}$, $N \in \mathbb{N}$.

(b) [**8**] $f_T(s) = (\frac{1}{2})^s$, $s \in \mathbb{N}$.

(5) **[I]** $X_1, \ldots, X_n$ zijn onafhankelijke stochasten, waarbij $X_k$ Bernoulli verdeeld is met parameter $p_k \in (0,1)$.

(a) **[6]** Bereken de verwachting en de variantie van de som $\sum_{k=1}^{n} X_k$.

(b) **[8]** Laat zien dat voor vaste verwachting de variantie maximaal is dan en slechts dan wanneer alle $p_k$ aan elkaar gelijk zijn.

(6) **[R]** $X$ en $Y$ hebben gezamenlijke kansdichtheidsfunctie $f_{X,Y}(x,y) = \lambda^2 e^{-\lambda y}$, $0 \leq x \leq y$.

(a) **[6]** Bereken $f_X(x)$, de marginale kansdichtheidsfunctie van $X$.

(b) **[6]** Bereken $f_{Y|X}(y\,|\,x) = f_{X,Y}(x,y)/f_X(x)$, de conditionele kansdichtheidsfunctie van $Y$ gegeven $X = x$.

(7) **[R][8]** Bereken de kansdichtheidsfunctie van $Z = X + Y$ wanneer $X$ en $Y$ de volgende gezamenlijke kansdichtheidsfunctie hebben:

$$f_{X,Y}(x,y) = \begin{cases} \frac{1}{2}(x+y)e^{-(x+y)}, & x, y \geq 0, \\ 0, & \text{elders.} \end{cases}$$

(8) **[T]** De gemeente Leiden voert een onderzoek uit naar de bierconsumptie van jongeren op 3 oktober. Het doel van het onderzoek is om tot een schatting te komen van het gemiddeld aantal glazen dat jongeren tussen 18 en 21 jaar op die feestdag drinken. Daartoe wordt een vrijwillige enquête gehouden onder 900 jongeren. De uitkomsten van die enquête zijn $N_1, \ldots, N_{900}$. Volgens de wet van de grote aantallen geldt dat $\frac{1}{900}\sum_{i=1}^{900} N_i \approx \mu$, waar $\mu$ het gezochte maar onbekende gemiddelde is.

(a) **[8]** Gebruik de centrale limietstelling om een schatting te geven van de kans

$$\mathbb{P}\left(\left|\frac{1}{900}\sum_{i=1}^{900} N_i - \mu\right| > 0.1\right)$$

wanneer gegeven is dat $N_1, \ldots, N_{900}$ onafhankelijk en identiek verdeeld zijn met standaarddeviatie 1 glas. Gebruik dat $\Phi(3) \approx 0.99865$, waar $\Phi$ de cumulatieve kansverdelingsfunctie van de standaard normale verdeling is.

(b) **[8]** De uitkomst van de enquête is dat er 4800 glazen bier zijn gedronken. Geef aan hoe we hiermee het gemiddelde $\mu$ kunnen schatten, en wat de kans is dat deze schatting correct is.

SOLUTIONS

(1) Volgens Definitie 1.1 van het boek moeten we laten zien dat $\mathcal{G}$: (1) niet leeg is; (2) gesloten is onder het nemen van complementen; (3) gesloten is onder het nemen van aftelbare verenigingen. Eigenschap (1) geldt omdat $\mathcal{F}$ niet leeg is. Eigenschap (2) geldt omdat $\Omega\backslash(A\cap B) = (\Omega\backslash A)\cup(\Omega\backslash B)$ en $\mathcal{F}$ aan eigenschappen (2) en (3) voldoet. Eigenschap (3) geldt omdat $\cup_{i=1}^{\infty}(A_i \cap B) = (\cup_{i=1}^{\infty} A_i)\cap B$ en $\mathcal{F}$ aan eigenschap (3) voldoet.

(2) Zij $X_m$ de uitkomst van worp $m$. Dan hebben we $A_{ij} = \{X_i = X_j\}$. Voor $1 \le i < j \le 3$ geldt dat

$$\mathbb{P}(A_{ij}) = \mathbb{P}(X_i = X_j) = \sum_{u=1}^{6} \mathbb{P}(X_i = X_j = u) = \sum_{u=1}^{6} (\tfrac{1}{6})^2 = \tfrac{1}{6}.$$

Voor $1 \le i < j \le 3$ en $1 \le k < l \le 3$ met $(i,j) \neq (k,l)$ geldt dat

$$\begin{aligned}\mathbb{P}(A_{ij}\cap A_{kl}) &= \mathbb{P}(X_i = X_j, X_k = X_l)\\ &= \mathbb{P}(X_1 = X_2 = X_3) = 6\,(\tfrac{1}{6})^3 = (\tfrac{1}{6})^2,\end{aligned}$$

dus is er paarsgewijze onafhankelijkheid (zie Definitie 1.38 van het boek). Anderzijds geldt ook dat

$$\mathbb{P}(A_{12}\cap A_{23}\cap A_{13}) = \mathbb{P}(X_1 = X_2 = X_3) = (\tfrac{1}{6})^2,$$

dus is er geen tripelgewijze onafhankelijkheid.

(3) Zijn $X_1, X_2$ de supermarkten die ik op zaterdag en zondag bezoek. Dan geldt dus

$$\begin{aligned}&\mathbb{P}(X_1 = J) = \tfrac{1}{3}, \quad \mathbb{P}(X_1 = AH) = \tfrac{2}{3},\\ &\mathbb{P}(X_2 = J \mid X_1 = J) = \tfrac{1}{5}, \quad \mathbb{P}(X_2 = AH \mid X_1 = J) = \tfrac{4}{5},\\ &\mathbb{P}(X_2 = AH \mid X_1 = AH) = \tfrac{1}{4}, \quad \mathbb{P}(X_2 = J \mid X_1 = AH) = \tfrac{3}{4}.\end{aligned}$$

(a) Bereken

$$\begin{aligned}\mathbb{P}(X_2 = J) &= \mathbb{P}(X_2 = J \mid X_1 = J)\,\mathbb{P}(X_1 = J)\\ &\quad + \mathbb{P}(X_2 = J \mid X_1 = AH)\,\mathbb{P}(X_1 = AH)\\ &= \tfrac{1}{5}\times\tfrac{1}{3} + \tfrac{3}{4}\times\tfrac{2}{3} = \tfrac{17}{30}.\end{aligned}$$

(b) Bereken, met behulp van de regel van Bayes,

$$\begin{aligned}
&\mathbb{P}(X_1 = AH \mid X_2 = J) = \mathbb{P}(X_1 = AH, X_2 = J)\,\frac{1}{\mathbb{P}(X_2 = J)} \\
&= \mathbb{P}(X_2 = J \mid X_1 = AH)\,\frac{\mathbb{P}(X_1 = AH)}{\mathbb{P}(X_2 = J)} \\
&= \tfrac{3}{4} \times \frac{\frac{2}{3}}{\frac{17}{30}} = \tfrac{15}{17}.
\end{aligned}$$

(4) (a) De gevraagde conditionele verwachting is gelijk aan (zie Definitie 6.68 van het boek)

$$\begin{aligned}
\mathbb{E}(T - t \mid T > t) &= \frac{\mathbb{E}((T-t)^+)}{\mathbb{P}(T > t)} = \frac{\sum_{s=t+1}^{\infty}(s-t)f_T(s)}{\sum_{s=t+1}^{\infty} f_T(s)} \\
&= \frac{\sum_{s=t+1}^{N}(s-t)}{\sum_{s=t+1}^{N} 1} = \frac{\frac{1}{2}(N-t)[1+(N-t)]}{N-t} \\
&= \tfrac{1}{2}(N-t+1), \qquad 0 < t < N.
\end{aligned}$$

(b) Nu geldt

$$\mathbb{E}(T - t \mid T \geq t) = \frac{\sum_{s=t+1}^{\infty}(s-t)(\frac{1}{2})^s}{\sum_{s=t+1}^{\infty}(\frac{1}{2})^s} = \frac{\sum_{u=1}^{\infty} u(\frac{1}{2})^u}{\sum_{u=1}^{\infty}(\frac{1}{2})^u} = 2.$$

Dit is de verwachting van de geometrische verdeling met parameter $\frac{1}{2}$ (zie Voorbeeld 2.36 van het boek).

(5) (a) Noteer $S_n = \sum_{k=1}^{n} X_k$. Dan geldt (zie p. 115 van het boek)

$$\begin{aligned}
\mathbb{E}(S_n) &= \sum_{k=1}^{n} \mathbb{E}(X_k) = \sum_{k=1}^{n} p_k, \\
\mathbb{V}\mathrm{ar}(S_n) &= \sum_{k=1}^{n} \mathbb{V}\mathrm{ar}(X_k) = \sum_{k=1}^{n} p_k(1-p_k).
\end{aligned}$$

(b) Zij $K$ de stochast die met kans $\frac{1}{n}$ de waarde $p_k$, $1 \leq k \leq n$, aanneemt. Met behulp van Jensen's ongelijkheid (Stelling 7.67 van het boek) volgt dan dat

$$\frac{1}{n}\sum_{k=1}^{n} p_k^2 = \mathbb{E}(K^2) \geq \mathbb{E}(K)^2 = \left(\frac{1}{n}\sum_{k=1}^{n} p_k\right)^2,$$

waarbij het gelijkteken geldt dan en slechts dan wanneer $K$ constant is, d.w.z. $p_k = p$ voor alle $1 \leq k \leq n$ en zekere $p \in (0,1)$. De laatste ongelijkheid kan gelezen worden als

$$\frac{1}{n}\,\mathbb{V}\mathrm{ar}(S_n) \leq \frac{1}{n}\,\mathbb{E}(S_n) - \left(\frac{1}{n}\,\mathbb{E}(S_n)\right)^2,$$

en dus is voor vaste verwachting de variantie maximaal dan en slechts dan wanneer alle $p_k$ aan elkaar gelijk zijn.

(6) (a) Bereken (zie p. 88 van het boek)

$$f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x,y)\,dy = \int_x^\infty \lambda^2 e^{-\lambda y}\,dy = \lambda e^{-\lambda x}, \qquad x \geq 0,$$

en $f_X(x) = 0$ elders.

(b) Er volgt dat

$$f_{Y|X}(y\,|\,x) = \frac{\lambda^2 e^{-\lambda y}}{\lambda e^{-\lambda x}} = \lambda e^{-\lambda(y-x)}, \qquad 0 \leq x \leq y,$$

en $f_{Y|X}(y\,|\,x) = 0$ elders.

(7) Volgens de formule voorafgaand aan Stelling 6.38 van het boek wordt de kansdichtheidsfunctie van $Z = X + Y$ gegeven door ($X, Y$ zijn niet onafhankelijk)

$$f_Z(z) = \int_{\mathbb{R}} dx\, f_{X,Y}(x, z-x) = \int_0^z dx\, \tfrac{1}{2} z e^{-z} = \tfrac{1}{2} z^2 e^{-z}, \qquad z \geq 0,$$

en 0 elders.

(8) (a) Definieer

$$Z_{900} = \frac{\sum_{i=1}^{900} N_i - 900\mu}{30}.$$

Volgens de centrale limietstelling is $Z_{900}$ bij benadering standaard normaal verdeeld. Ervan uitgaande dat deze benadering goed genoeg is (zonder extra informatie kunnen we dit niet nader kwantificeren), schatten we

$$\begin{aligned}
\mathbb{P}\left(\left|\tfrac{1}{900}\sum_{i=1}^{900} N_i - \mu\right| > 0.1\right) &= \mathbb{P}(|Z_n| > 3) \approx \mathbb{P}(|Z| > 3) \\
&= \int_{\mathbb{R}\setminus[-3,3]} \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2} = 2[1 - \Phi(3)] \approx 0.0027.
\end{aligned}$$

(b) De enquête levert $\frac{1}{900}\sum_{i=1}^{900} N_i = \frac{4800}{900} = 5\frac{1}{3}$. De kans is derhalve $\approx 0.0027$ dat $|5\frac{1}{3} - \mu| > 0.1$. M.a.w. $\mu \in [5\frac{7}{30}, 5\frac{13}{30}]$ met kans $\approx 0.9973$.