

# Monster Energy Ultra Rosá Zero Sugar: (Ik ben verder door aan het draaien)

Zwansrekenen & Statisfreak II van Cécile Adam

16 juni 2025, 8:00. Getranscribeerd door François Slutsky.

Hierbij de vragen van het examen en oplossingen door een groep studenten. Deze zijn dus niet perfect omdat wij ook maar gewoon een hoop mensen zijn die dit examen eens gemaakt hebben. Wees dus kritisch! Het examen was volledig schriftelijk en verdeeld in een deel theorie (gesloten boek) op 50% en een deel oefeningen (open boek) op 50%. Het examen zelf duurde vier uur. Je mocht pas een deel afgeven als je het een uur lang had. Je had maximaal twee uur om het theorie deel af te geven. Veel succes!

## Theorie

### Vraag 1: 12 punten

(a) Geef de definitie van bijna zekere convergentie. Leg bij alle gebruikte notatie uit en wat het betekent.

(b) Bewijs het lemma van Slutsky (*het lemma zelf was gegeven*).

(c) Bewijs dat convergentie in kans ook convergentie in verdeling impliceert. Dat wil zeggen, bewijs dat

$$X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{D} X.$$

(d) Stel dat  $X_n \xrightarrow{D} X$  en  $Y_n \xrightarrow{P} c$  met  $c > 0$ . Bewijs dat  $X_n Y_n \xrightarrow{D} cX$ .

### Vraag 2: 3 punten

(a) Geef de definitie van een sufficiënte statistiek voor  $\theta$  in een statistisch model.

(b) Stel dat  $X_1, X_2 \sim \text{Poisson}(\theta)$ . Toon aan dat  $S = X_1 + X_2$  een sufficiënte statistiek is voor  $\theta$ . (*Tabel met verdelingen, verwachte waarden, ... was gegeven*)

### Vraag 3: 5 puntent

(a) Geef de algemene definitie van een stochastisch proces. Definieer hierin alle grootheden.

(b1) Geef de definitie van een Markovketen met discrete tijd.

(b2) Beschouw een Markovketen met toestanden  $(0, 1, 2)$  en matrix

$$P = \begin{pmatrix} 0.5 & 0.5 & 0 \\ 0.25 & 0.5 & 0.25 \\ 0 & 0.5 & 0.5 \end{pmatrix}.$$

Ga na of dit een geldige overgangsmatrix voor een Markovketen is. Bereken  $P(X_2 = 2 | X_0 = 0)$ .

## Theorie: Oplossingen/Waar in het boek?

### Vraag 1

- (a) Definitie 4.3, pagina 84.
- (b) Lemma 4.1, pagina 89.
- (c) Stelling 4.3 (i), pagina 90.
- (d) Stelling 4.9 (ii), pagina 101.

### Vraag 2

- (a) Definitie 5.9, pagina 141.
- (b) Voorbeeld 5.14, pagina 142. Alternatief kan je de factorisatiestelling van Fisher-Neyman gebruiken, of beargumenteren dat de Poissonverdeling een lid is van de exponentiële klasse.

### Vraag 3

- (a) Definitie 7.1, pagina 183.
- (b1) Definitie 7.2, pagina 184.
- (b2) Ga na dat alle elementen in de matrix groter of gelijk aan nul zijn, en dat de som over elke rij één is. Hier is aan voldaan. Verder is deze kans gelijk aan  $(P^2)_{02} = 0.125$ .

# Oefeningen

## Vraag 1: 7 punten

Beschouw de karakteristieke functie gegeven door

$$\varphi_X(t) = \frac{1}{1 + a^2 t^2}$$

met  $a > 0$ . We noteren  $X \sim \text{LP}(a)$ .

(a) Bereken  $E(X)$  en  $\text{Var}(X)$  (je mag aannemen dat deze momenten bestaan).

(b) Zij  $Y \sim \text{Expo}(\frac{1}{a})$  en  $Z \sim \text{Bernoulli}(\frac{1}{2})$ . Bewijs dat  $Y(2Z - 1) \sim \text{LP}(a)$ . (Hint: voor twee willekeurige s.v.  $U, V$  geldt  $E(U) = E_T(E(U|T))$ )

(c) Zij  $V_i, W_i \sim \Gamma(\alpha, \beta)$  en beschouw

$$S_n = \frac{\sum_{i=1}^n (V_i - W_i)}{\sqrt{2\alpha n\beta}}.$$

Wat is de verdeling van  $S_n$  in de limiet  $n \rightarrow \infty$ ?

(d) We herhalen de vorige vraag, echter is nu  $\alpha = 1/n$ . Wat is nu de verdeling van  $S_n$  als  $n \rightarrow \infty$ ? (Hint: gebruik de karakteristieke functie)

(e) Er was een stuk R-code, die ik niet heb onthouden, gegeven die steekproeven genereert van de situatie uit de vorige deelvragen. Hiervoor werd de empirische cumulatieve verdelingsfunctie en de empirische kansdichtheid geplot in twee verschillende figuren, samen met een theoretische afleiding van de limietverdeling. Je moest bespreken wat deze code deed en wat de figuren toonden.

## Vraag 2: 6.5 punten

Beschouw het gebied  $A = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1, x - 1 \leq y \leq 1 - x\}$  met een uniforme kansmaat op  $A$ .

(a) Verantwoord waarom  $f_{X,Y}(x, y) = 1$  als  $(x, y) \in A$  en 0 elders.

(b1) Toon aan dat voor de ordestatistieken van  $X_1, \dots, X_5$  geldt dat

$$E(X_{(5)} \mid X_{(4)}) = \frac{2X_{(4)} + 1}{3}.$$

(b2) Zij  $Z_i = 2X_i - X_i^2$ . Vind de verwachtingswaarde van de range op  $Z_1, \dots, Z_5$ .

(c) Vind de gezamenlijke<sup>1</sup> verdeling van de stochastische vector  $(U, V) = \left(X + Y, \frac{X}{X+Y}\right)$ .

(d) Zonder iets te berekenen, leg uit hoe je de verdeling van de s.v.  $\frac{X}{X+Y}$  zou vinden.

---

<sup>1</sup>Merk op dat  $V$  mogelijks niet goed gedefinieerd is, bv voor  $(x, y) = (1/2, -1/2)$ , maar dit negeren we :).

### Vraag 3: 6.5 punten

Beschouw

$$f(x; \theta) = \sqrt{\frac{\theta}{2\pi x^3}} \exp\left(-\frac{\theta(x-1)^2}{2x}\right)$$

met  $x > 0$  en  $\theta > 0$ .

(a) Toon aan dat de maximum likelihood estimator van  $\theta$  gegeven wordt door

$$\hat{\theta}_{\text{MLE}} = \frac{n}{\sum_{i=1}^n \frac{(X_i-1)^2}{X_i}}.$$

(b) Vind de asymptotische verdeling van  $\hat{\theta}_{\text{MLE}}$  en ook die van  $\log(1 + \hat{\theta}_{\text{MLE}})$ .  
Gegeven is dat

$$\frac{\theta(X-1)^2}{X} \sim \chi^2(1).$$

(c1) Zoek op basis van de MLE-schatter een onvertekende schatter voor  $\theta$ .

(c2) Geef de efficiëntie van deze onvertekende schatter.

(c3) Is deze onvertekende schatter zwak consistent? Verklaar.

## Oefeningen: Oplossingen

### Vraag 1

(a) Gebruik Stelling 1.5 om  $E(X)$  en  $E(X^2)$  te berekenen. Je moet uitkomen  $E(X) = 0$  en  $\text{Var}(X) = E(X^2) = 2a^2$ .

(b) We berekenen de karakteristieke functie van  $K = Y(2Z - 1)$ .

$$\begin{aligned}\varphi_K(t) &= E(\exp(itY(2Z - 1))) \\ &= E(E(\exp(itY(2Z - 1)) \mid Z))\end{aligned}$$

Merk nu op dat

$$E(\exp(itY(2Z - 1)) \mid Z = 0) = E(\exp(-itY)) = \varphi_Y(-t).$$

Verder is

$$E(\exp(itY(2Z - 1)) \mid Z = 1) = E(\exp(itY)) = \varphi_Y(t).$$

Dus is

$$\varphi_K(t) = \frac{1}{2}(\varphi_Y(t) + \varphi_Y(-t)) = \frac{1}{1 + a^2 t^2}.$$

(c) Gebruik de Centrale Limietstelling en vind dat  $S_n \xrightarrow{D} Z$  met  $Z \sim N(0, 1)$ .

(d) Met de Stelling van Lévy-Cramér vind je dat  $S_n \xrightarrow{D} \text{LP}\left(\frac{1}{\sqrt{2}}\right)$ . Sterker nog, je vindt dat  $S_n \sim \text{LP}\left(\frac{1}{\sqrt{2}}\right)$  voor alle  $n \in \mathbb{N}$ .

### Vraag 2

(a) Dit is omdat

$$\iint_A f_{X,Y}(x, y) dx dy = \int_0^1 \int_{x-1}^{1-x} c dx dy = c.$$

(b1) Breinrot rekenvraag. Gebruik Stelling 3.7 en Stelling 3.8 om  $f_{Y_5|Y_4}(y_5 \mid y_4)$  te berekenen. Dit is

$$f_{Y_5|Y_4}(y_5 \mid y_4) = \frac{f_X(y_5)}{1 - F_X(y_4)}.$$

Er resteert nog  $f_X$  en  $F_X$  te vinden.

1.  $f_X(x) = \int_{x-1}^{1-x} dy = 2 - 2x$  als  $0 \leq x \leq 1$ , nul elders.
2.  $F_X(x) = \int_0^x f_X(u) du = 2x - x^2$  als  $0 \leq x \leq 1$ , nul als  $x < 0$  en één als  $x > 1$ .

We zijn nu klaar om de breinrot samen te brengen.

$$\begin{aligned}E(Y_5 \mid Y_4 = y_4) &= \int_{y_4}^1 y_5 \frac{f_X(y_5)}{1 - F_X(y_4)} dy_5, \\ &= \int_{y_4}^1 y_5 \frac{2 - 2y_5}{(1 - 2y_4 + y_4^2)} dy_5, \\ &= \frac{2y_4 + 1}{3} \cdot \frac{(1 - y_4)^2}{(1 - y_4)^2}, \\ &= \frac{2y_4 + 1}{3}.\end{aligned}$$

Als stochastische veranderlijke geldt dus inderdaad dat

$$E(X_{(5)} \mid X_{(4)}) = \frac{2X_{(4)} + 1}{3}.$$

**(b2)** Deze vraag ziet er veel enger uit dan hij eigenlijk is. Er is namelijk niets te berekenen dat je nog niet hebt gevonden in het eerste deel van deze vraag als je volgend trucje inziet. De truc is om in te zien dat  $Z_i = F_X(X_i)$ , en omdat  $X_i$  dezelfde verdeling volgt als  $X$ , is dit met Stelling 3.2 een kansintegraaltransformatie zoals gezien in Kans & Stat I. Alle  $Z_i$  volgen dus gewoon een uniforme verdeling op  $[0, 1]$ . De verwachtingswaarde van de range van de ordestatistieken wordt uitgerekend op pagina 82. Je vindt  $E(R) = 2/3$ .

**(c)** Breinrot rekenvraag nummer twee. Gebruik Stelling 3.3 en reken de inverse transformaties en Jacobiaan uit. Je vindt

$$f_{U,V}(u, v) = f_{X,Y}(uv, u(1-v))|u| = |u|.$$

Voor het definitie gebied hebben we alvast  $-1 \leq u \leq 1$ , gezien  $2x - 1 \leq x + y \leq 1$  zodat zeker  $u \leq 1$ . Verder vinden we dat  $(x, y) = (0, -1)$  de laagste waarde van  $u$  geeft in het definitie gebied, zodat  $-1 \leq u \leq 1$ . Verder is  $uv = x \leq 1$ , zodat  $v \leq \frac{1}{u}$  als  $u > 0$ ,  $v \geq \frac{1}{u}$  als  $u < 0$ . Verder moet in het eerste geval  $v \leq 0$ , in het tweede geval moet  $v \geq 0$ . Dus moet  $0 \leq |v| \leq \frac{1}{|u|}$  en moet het teken van  $v$  gelijk zijn aan het teken van  $u$ .

**(d)** Je moet de grenzen van  $U$  in functie van  $V$  zetten en de veranderlijke  $u$  uitintegreren uit de expressie van  $f_{U,V}$ .

### Vraag 3

**(a)** Merk op dat  $\mathcal{L}(x_1, \dots, x_n, \theta) = \frac{n}{2} \ln(\theta) - \sum_{i=1}^n \frac{1}{2} \ln(\pi x_i^3) - \frac{\theta}{2} \sum_{i=1}^n \frac{(x_i-1)^2}{x}$ . We<sup>2</sup> vinden enkel een maximum in een kritiek punt, dus we willen  $\frac{\partial \mathcal{L}}{\partial \theta} = 0$ , wat

$$\hat{\theta}_{MLE} = \frac{n}{\sum_{k=1}^n \frac{(X_k-1)^2}{X}}$$

geeft. Nu is ook  $\frac{\partial^2 \mathcal{L}}{\partial \theta^2} = -\frac{2n}{\theta^2} < 0$  zodat we inderdaad een maximum hebben.

**(b)** Merk op dat aan alle voorwaarden voor stelling 6.4 voldaan is (ik heb dit nagevraagd tijdens het examen, we moesten effectief alle voorwaarden nagaan, wat extreem cringe is). We vinden dus dat  $\sqrt{n}(\hat{\theta}_{MLE} - \theta) \xrightarrow{D} N\left(0, \frac{1}{I(\theta)}\right)$ . Merk nu op dat de regulariteits voorwaarden voor stelling 6.4 de voorwaarde voor lemma 6.1 bevatten, zodat  $I(\theta) = -E_\theta\left(\frac{\partial^2 \mathcal{L}}{\partial \theta^2}\right) = \frac{2}{\theta^2}$ . Dus zal  $\sqrt{n}(\hat{\theta}_{MLE} - \theta) \xrightarrow{D} N(0, 2\theta^2)$ . De delta methode geeft nu dat  $\sqrt{n}(\log(1 + \hat{\theta}_{MLE}) - \log(1 + \theta)) \xrightarrow{D} N(0, \frac{2\theta^2}{(1+\theta)^2})$  (je moet de voorwaarden eigenlijk opnieuw nagaan, maar dat kan je zelf wel).

**(c1)** Merk op dat

$$E_\theta(\hat{\theta}_{MLE}) = n\theta E_\theta\left(\frac{1}{\sum_{k=1}^n \frac{\theta(X_k-1)^2}{X_k}}\right) = n\theta E\left(\frac{1}{Y}\right)$$

---

<sup>2</sup>Hier gebruiken we  $\mathcal{L}$  voor het logaritme van de likelihood functie  $L$ .

met  $Y \sim \chi_n^2$  gezien alle  $\frac{\theta(X_i-1)^2}{X_i}$  onafhankelijk en  $\chi_1^2$  verdeeld zijn. Nu berekenen we even  $E(\frac{1}{Y})$  en doen we ook meteen  $E(\frac{1}{Y^2})$  gezien we dit zometeen nog nodig zullen hebben.

$$\begin{aligned} E\left(\frac{1}{Y}\right) &= \int_0^\infty \frac{2^{-\frac{n}{2}}}{\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} \frac{1}{x} dx \\ &= \frac{2^{-\frac{n}{2}}}{\Gamma(\frac{n}{2})} \int_0^\infty 2^{\frac{n}{2}-1} u^{\frac{n}{2}-2} e^{-u} du \\ &= \frac{1}{2\Gamma(\frac{n}{2})} \Gamma\left(\frac{n}{2} - 1\right) = \frac{1}{n-2} \end{aligned}$$

waarbij we  $u = \frac{x}{2}$  gesubstitueerd hebben en de definitie van de gamma functie gebruikt hebben (deze staat normaal in je formularium). Verder is volledig analoog

$$\begin{aligned} E\left(\frac{1}{Y^2}\right) &= \int_0^\infty \frac{2^{-\frac{n}{2}}}{\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} \frac{1}{x^2} dx \\ &= \frac{2^{-\frac{n}{2}}}{\Gamma(\frac{n}{2})} \int_0^\infty 2^{\frac{n}{2}-2} u^{\frac{n}{2}-3} e^{-u} du \\ &= \frac{1}{4\Gamma(\frac{n}{2})} \Gamma\left(\frac{n}{2} - 2\right) = \frac{1}{(n-2)(n-4)}. \end{aligned}$$

We hebben nu dus  $E_\theta(\hat{\theta}_{MLE}) = \frac{n\theta}{n-2}$ , zodat  $\hat{\theta}_* := \frac{n-2}{n}\hat{\theta}_{MLE}$  een onvertekende schatter is voor  $\theta$ .

**(c2)** We kennen de Fisher informatie al van dit statistisch model, namelijk  $I(\theta) = \frac{2}{\theta^2}$ . Nu is  $\text{Var}(\hat{\theta}_*) = \frac{(n-2)^2}{n^2} \text{Var}(\hat{\theta}_{MLE})$ . Verder is  $E(\hat{\theta}_{MLE}^2) = n^2 \theta^2 E\left(\frac{1}{Y^2}\right) = \frac{n^2 \theta^2}{(n-2)(n-4)}$  met opnieuw  $Y \sim \chi_n^2$ . Nu is dus  $\text{Var}(\hat{\theta}_{MLE}) = \frac{n^2 \theta^2}{(n-2)(n-4)} - \frac{n^2 \theta^2}{(n-2)^2} = \frac{n^2 \theta^2}{n-2} \left(\frac{1}{n-4} - \frac{1}{n-2}\right) = \frac{2n^2 \theta^2}{(n-2)^2(n-4)}$ . Dus is  $\text{Var}(\hat{\theta}_*) = \frac{2\theta^2}{n-4}$ , zodat de efficiëntie van  $\hat{\theta}_*$  gegeven is door  $\frac{n-4}{n}$ .

**(c3)** We vinden dat  $\hat{\theta}_*$  onvertekend is en dat de variantie van  $\hat{\theta}_*$  naar 0 gaat voor  $n \rightarrow \infty$ , zodat remark 5.7 zegt dat  $\hat{\theta}_*$  zwak consistent is.