

Statistiek I

K. Van Lancker & G. Baklicharov

Academiejaar 2024-2025

U krijgt 4 uur voor het examen. Het examen is relatief lang omdat, ondanks de open boek vorm, een actieve kennis van de leerstof wordt verwacht. Geen van de vragen vergt echter veel rekenwerk. Alle vragen kunnen onafhankelijk van elkaar worden opgelost. **U mag verder telkens steunen op de resultaten bekomen in eerdere vragen.**

Gelieve uw naam zowel op deze examenkopij te schrijven, als op uw oplossingenblad. **Het is verplicht om alles in te dienen: ook deze examenkopij, uw kladpapier en alle notities die u tijdens het examen maakt.**

Alvast veel succes!

Gesloten boek: u kunt hier maximaal 1u aan werken.

1. (2 punten) Zij X en Y twee continue toevalsveranderlijken, die niet noodzakelijk onafhankelijk zijn. Neem aan dat de onderstaande uitdrukkingen bestaan. Vul het meest geschikte symbool $\leq$, $\geq$, of $=$ in op de lege plek. Toon aan.

(a) $P(|X + Y| > 2) \dots \frac{1}{10} E\{(X + Y)^4\}$

(b) $E[\{E(X | Y)\}^2] \dots \{E(X)\}^2$.

2. (2 punten) Zij $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ onafhankelijke en identisch verdeelde toevalsvectoren, waarbij bovendien $X_1, \dots, X_n, Y_1, \dots, Y_n$ allemaal onderling onafhankelijk zijn.

- (a) Veronderstel voor deze deelvraag dat $X_i \sim \text{Poisson}(\lambda)$, voor $i = 1, \dots, n$. Definieer Z_n als volgt:

$$Z_n = \sum_{i=1}^n \frac{X_i - \lambda}{\sqrt{n}}.$$

Onderzoek naar welke functie de karakteristieke functie van Z_n convergeert als $n \rightarrow \infty$. Geef een redenering voor je antwoord, zonder een formeel bewijs te geven.

- (b) Stel $V_i = (X_i - \bar{X}_n)Y_i$. Gaat u akkoord met onderstaande bewering? Verklaar waarom wel/niet.

Het volgt onmiddellijk uit de centrale limietstelling voor V_i dat $\frac{1}{\sqrt{n}} \sum_{i=1}^n V_i \xrightarrow{d} N(0, \sigma^2)$, met σ^2 een zekere variantie.

3. (1 punt) Toon formeel aan: als $f(A, B, C, D, E) = f(A)f(B|A)f(C|A)f(D|B, C)f(E|D)$, dan is $B \perp\!\!\!\perp E | A, D$.

Open boek:

1. (8.5 punten) Zij $X_1, X_2, \dots, X_n$ onderling onafhankelijke en identisch verdeelde toevalsveranderlijken, afkomstig van een Rayleigh verdeling met parameter $\tau > 0$. De dichtheid van de Rayleigh-verdeling is gegeven door:

$$f(x) = \tau x \exp\left(-\frac{1}{2}\tau x^2\right), \quad x \geq 0.$$

- (a) Toon aan dat de maximum-waarschijnlijkheidsschatter $\hat{\tau}$ voor τ gelijk is aan

$$\frac{2n}{\sum_{i=1}^n X_i^2}.$$

- (b) Toon aan dat X^2 exponentieel verdeeld is met parameter $\tau/2$. U mag dit resultaat gebruiken in de verdere vragen.
(c) Toon aan dat $1/\hat{\tau}$ een onvertekende schatter is voor $1/\tau$.
(d) Toon aan dat $\hat{\tau}$ een consistente schatter is voor τ .
(e) Toon aan dat

$$\sqrt{n}(\hat{\tau} - \tau) \xrightarrow{d} \mathcal{N}(0, \tau^2).$$

- (f) Gebruik voorgaand resultaat om een benaderend 95% betrouwbaarheidsinterval voor τ op te stellen.
(g) (Moeilijk) Bepaal tevens een exact 95% betrouwbaarheidsinterval voor τ . U mag er daarbij vanuit gaan dat u software (bv. R) ter beschikking heeft om de dichtheidsfunctie, cumulatieve distributiefunctie of percentielen van een verdeling naar keuze op te vragen.
Hint: denk eerst na wat de exacte verdeling van $2/\hat{\tau}$ is en hoe u op basis hiervan een exact 95% betrouwbaarheidsinterval voor $2/\tau$ kunt bekomen.

2. (4 punten) Zij X en Y continue toevalsveranderlijken en A een binaire toevalsveranderlijke die waarden aanneemt in $\{0, 1\}$. Stel $\psi = E\{E(Y|A=1, X)\}$, $\pi(X) = P(A=1|X)$ en

$$\phi(Y, A, X) = \frac{A}{\pi(X)} \{Y - E(Y|A=1, X)\} + E(Y|A=1, X) - \psi.$$

- (a) Bereken $E\{\phi(Y, A, X)\}$.
(b) Stel $\text{Var}(Y|A, X) = \sigma^2$ en $\mu_1(X) = E(Y|A=1, X)$, toon aan dat $\text{Var}\{\phi(Y, A, X)\} = \text{Var}\{\mu_1(X)\} + \sigma^2 E\left\{\frac{1}{\pi(X)}\right\}$.

3. (2.5 punten) Albuterol is een frequent gebruikt geneesmiddel voor kinderen met ademhalingsproblemen. De standaardmethode van toediening bij niet-astma patiënten is een verstuiver. Werken met een inhaleertoestel is echter gemakkelijker en goedkoper. Er werd een studie opgezet om beide toedieningsmethoden met elkaar te vergelijken. 32 patiënten werden op een niet-willekeurige wijze toegewezen aan de behandeling met verstuiver en 32 aan de behandeling met inhaleertoestel.

De primaire uitkomst is voor elke patiënt het verschil in PEF (Peak Expiratory Flow rate) na de studie in vergelijking met vóór toedoen van albuterol. Bij de groep met verstuiver vond men een gemiddeld verschil van 23.95 l/min met een standaardafwijking van 5 l/min en in de groep met inhaleertoestel een gemiddeld verschil van 24.46 l/min met een standaardafwijking van 6 l/min.

- (a) (1 punt) Stel een 95% betrouwbaarheidsinterval op voor de gemiddelde PEF-verandering in de groep waar toediening gebeurt via een inhaleertoestel.
- (b) (1 punt) Stel een 95% betrouwbaarheidsinterval op voor het gemiddeld verschil in de PEF-verandering tussen beide groepen. U mag er daarbij van uitgaan dat de PEF-verandering even variabel is in beide behandelingsgroepen.
- (c) (0.5 punt) Welk van de twee betrouwbaarheidsintervallen zou u verkiezen om iets te leren over de werking van het inhaleertoestel, en of dit al dan niet op de markt moet komen? Leg uit waarom.