Statistiek I, Bachelor Wiskunde (Prof. C. Ley)
Examen Augustus 2020

Vraag 1 (10 punten)

Beantwoord volgende vragen met JA/NEE en een heel korte motivatie (max 2 lijnen per vraag). De juiste antwoord alleen geeft 0 punten, juiste antwoord met approximatieve verklaring 1 punt en juiste antwoord met juiste verklaring de volledige 2 punten per vraag.

1. Een test met power 40% kan nooit een p-waarde van 0.001 leveren.
2. Zij f and g dichtheidsfuncties. Dan is $|f+g|/2$ een geldige dichtheidsfunctie.
3. $\mathrm{E}[X+Y+Z] = \mathrm{E}[X] + \mathrm{E}[Y+Z]$ geldt enkel als X onafhankelijk is van de som $Y+Z$.
4. Bij een betrouwbaarheidsinterval voor de centrale ligging op het niveau 90% zit $\bar{X}$ in 90 van 100 gevallen in dat interval.
5. Zij f and g dichtheidsfuncties. Dan is $|f * g|/2$ (simpele vermenigvuldiging) een geldige dichtheidsfunctie.

Vraag 2

Zij $X_1, X_2, \ldots, X_n$ onafhankelijke en identiek verdeelde, continue toevalsveranderlijken die een symmetrische verdeling hebben ($X_i \stackrel{D}{=} -X_i$). Laat $Y_i = \mathrm{sgn}(X_i)$ voor $i = 1, ..., n$, waar $\mathrm{sgn}(z) = -1$ als $z < 0$, $\mathrm{sgn}(z) = 0$ als $z = 0$ en $\mathrm{sgn}(z) = 1$ als $z > 0$.

(a) Bepaal de verdeling van $S = \sum_{i=1}^n Y_i$.

(b) Toon dat $\overline{Y}_n (:= \frac{1}{n}\sum_{i=1}^n Y_i) \xrightarrow{P} 0$ als $n \to \infty$.

(c) Stel dat men S gebruikt als teststatistiek om de nulhypothese dat de verdeling van observaties X symmetrisch is al dan niet te verwerpen. Ga ervan uit dat deze testprocedure wordt uitgevoerd op het 5% significantieniveau. Bepaal het aanvaardingsgebied van deze teststatistiek indien $n = 12$.

(d) Wat is het percentage type I-fouten in dit geval?

(e) Bepaal het percentage type II-fouten indien X_i in werkelijkheid $N(1, 4)$-verdeeld is.

Vraag 3

Zij $(X_1, \ldots, X_n)$ onafhankelijke en identiek verdeelde observaties met dichtheidsfunctie

$$f_\theta(x) = \frac{1}{2\theta^3} x^2 \exp\left(-\frac{x}{\theta}\right), \quad x \in [0, \infty),$$

waar $\theta > 0$ de onbekende parameter is.

(a) Wat is de verdeling van $\sqrt{X}$?

(b) Bepaal $E[X]$ en $Var(X)$.

(c) Bepaal de maximum likelihood schatter van θ.

(d) Is $\hat{\theta}_{MLE}$ een consistente schatter voor θ?

(e) Bepaal de asymptotische verdeling van $\hat{\theta}_{MLE}$.

(f) Bepaal de asymptotische verdeling van $\log\left(\hat{\theta}_{MLE}\right)$.

(g) Bepaal een asymptotische $(1-\alpha)100\%$ betrouwbaarheidsinterval voor θ.

U kan bij deze vragen gebruik maken van de identiteit

$$\int_0^{+\infty} x^k e^{-x} = k!$$

Vraag 4

Zij $X_1, X_2, ..., X_n$ onafhankelijk en uniform getrokken uit het interval $[-\theta, \theta]$ met $\theta \in \mathbb{R}_0^+$. Dan geldt

$$n\left(\theta - \max_{i=1}^{n} |X_i|\right) \xrightarrow{d} Y$$

als $n \to \infty$ voor een zekere stochastische veranderlijke Y.

(a) Bepaal de verdeling van Y.

(b) Uit een steekproef van 200 observaties bekomt men als kleinste observatie -5 en als grootste 4.5, bepaal een asymptotisch tweezijdig 95% betrouwbaarheidsinterval voor θ.

Vraag 5

Beschouw het volgende statistische model

$$\frac{E(Y|A, X)}{E(Y|A = 0, X)} = \exp(\psi A),$$

waarbij $\psi \in \mathbb{R}$ ongekend is. Onderstel verder dat A binair is met $P \equiv P(A = 1|X)$ gekend en $1 - P \equiv P(A = 0|X)$. Onderstel dat we n i.i.d. metingen $(Y_i, A_i, X_i), i = 1, \ldots, n$ beschikbaar hebben, en noteer $P_i \equiv P(A_i = 1|X_i)$.

(a) Toon aan dat

$$\frac{1}{n}\sum_{i=1}^{n}(A_i - P_i)Y_i \exp(-\psi A_i)$$

gemiddeld 0 is.

(b) Gebruik de uitdrukking uit de vorige vraag als een teststatistiek voor de nulhypothese dat $\psi = 0$. Bepaal de asymptotische verdeling. Het is niet nodig om varianties en/of covarianties daarbij expliciet uit te rekenen van deze teststatistiek (bvb. termen van de vorm $\mathrm{Var}(A_i P_i Y_i)$ hoeven niet verder uitgewerkt te worden). Hint : standaardiseer de teststatistiek indien nodig onder de nulhypothese en gebruik dit om een criterium op te stellen wanneer de nulhypothese mag verworpen worden op het 5% significantieniveau.

(c) Omwille van het resultaat uit vraag (a), stellen we voor om ψ te schatten als de oplossing van

$$0 = \sum_{i=1}^{n}(A_i - P_i)Y_i \exp(-\hat{\psi} A_i).$$

Los hier een gesloten uitdrukking voor $\exp(\hat{\psi})$ uit op, wetende dat A_i binair is. Gebruik het resultaat om aan te tonen dat $\hat{\psi}$ een consistente schatter is voor ψ.