Continue kansvariabelen

Statistiek en Probabiliteit
Academiejaar 2019-2020

Theoretische oefeningen

Oefening 1. Stel dat $X_1, ..., X_n$ onafhankelijke variabelen zijn met dezelfde verwachtingswaarde μ en dezelfde variantie σ^2. Wat zijn de verwachtingswaarde en de variantie van de som $S = X_1 + ... + X_n$ en het gemiddelde $\bar{X} = S/n$?

Oefening 2. Met welke verdeling zou je volgende voorbeelden modelleren?

a. De lengte van de Belgische bevolking.

b. Het tijdstip op de dag waarop iemand geboren wordt.

c. De tijd nodig om een bestand te downloaden van het internet.

d. De tijd om 3 gelijkaardige bestanden te downloaden van het internet.

e. De tijd om 100 gelijkaardige bestanden te downloaden van het internet.

Oefening 3. Voor een exponentieel verdeelde kansvariabele X geldt voor alle $x \in \mathbb{R}^+$ dat

$$f(x) = \lambda e^{-\lambda x}$$

voor een zeker parameter $\lambda > 0$. X is positief, m.a.w. is de kansdichtheid gelijk aan nul voor alle strikt negatieve waarden.

a. Toon aan dat dit een geldige kansdichtheid is.

b. Bereken de verdelingsfunctie F van X.

c. Bepaal de verwachtingswaarde van X.

Oefening 4. Een uniform verdeelde variabele X heeft als kansdichtheid

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{als } a \leq x \leq b \\ 0 & \text{anders} \end{cases}$$

voor zekere $a, b \in \mathbb{R}, a < b$.

a. Toon aan dat dit een geldige kansdichtheid is.

b. Bereken de verdelingsfunctie F van X.

c. Bepaal de verwachtingswaarde van X.

Oefening 5. De tijd X, in minuten, die nodig is om een bepaald systeem opnieuw op te starten is een continue variabele met dichtheid

$$f(x) = \begin{cases} C(10-x)^2 & \text{als } 0 < x < 10 \\ 0 & \text{anders} \end{cases}$$

a. Bepaal de constante C. (Maak gebruik van de eigenschappen van een dichtheidsfunctie)

b. Bepaal de cumulatieve distributiefunctie F

c. Bereken de verwachtingswaarde van X.

d. Bereken de kans dat het opstarten meer dan 5 minuten duurt.

Praktische oefeningen

Oefening 6. De beroemde geiser "Old Faithful" uit het Amerikaanse Yellowstone National Park blaast om het uur hete damp de lucht in. Ik passeer toevallig langs de geiser

a. Hoe lang zal ik gemiddeld moeten wachten om een uitbarsting van de geiser te zien?

b. Ik blijf 15 minuten wachten, wat is de kans dat ik een uitbarsting van de geiser zie?

Oefening 7. De IT-Helpdesk van een bepaald bedrijf krijgt gemiddeld 8 telefoons binnen per uur.

a. Welke verdeling zal je gebruiken om de tijd tussen twee telefoons te modelleren?

b. Wat is de kans dat de tijd tussen twee inkomende telefoons meer dan 15 minuten bedraagt?

c. Vergelijk dit antwoord met het antwoord van oefening 14c uit vorige les. Wat stel je vast?

Oefening 8. Voor een bepaalde elektronische component is de tijd totdat de component uitvalt Gamma verdeeld met parameters $\alpha = 2$ en $\lambda = 2(/jaar)$. Bereken de kans dat de component uitvalt in de eerste 6 maanden. (Hint: gebruik de functie *pgamma*)

Oefening 9. De tijd die een printer nodig heeft om een bepaald document af te drukken is exponentieel verdeeld met gemiddelde 12 seconden.

a. Wat is de kans dat het afdrukken van mijn document meer dan 20 seconden duurt?

b. Veronderstel dat mijn document pas het derde is in de wachtlijn. Welke verdeling volgt de tijd tot wanneer het document afgeprint is? Wat is de kans dat ik langer dan 1 minuut moet wachten?

c. Welke verdeling volgt het aantal afgewerkte documenten per minuut?

d. Bereken de kans uit b. aan de hand van de verdeling uit c.

Oefening 10. Als $Z \sim N(0, 1)$, bepaal dan

a. $P(Z \leq 2.71) = \Phi(2.71)$

b. $P(Z < 2.71)$

c. $P(Z < -2.71)$

d. $P(|Z| < 2.71)$

e. $P(Z > 0.41)$

f. $P(|Z| > 0.41)$

Oefening 11. Als $Z \sim N(0, 1)$, bepaal dan a zodat

a. $P(Z < a) = 0.95$ (met andere woorden: bepaal $z_{0.95}$)

b. $P(Z > a) = 0.10$

c. $P(|Z| < a) = 0.5$

d. $P(|Z| > a) = 0.3$

Oefening 12. Als $X \sim N(-1, 4)$, bepaal dan b zodat

a. $P(X < b) = 0.05$

b. $P(X > b) = 0.05$

c. $P(|X + 1| > b) = 0.05$

Oefening 13. De gemiddelde lengte van een professionele basketbalspeler is 2 meter, en de standaarddeviatie is 10 cm. Als we veronderstellen dat de lengte van de personen in deze groep normaal verdeeld is,

a. Welk percentage van de professionele basketbalspelers is groter dan 210 cm?

b. Als jouw favoriete speler bij de 20% grootste basketbalspelers is, hoe groot moet hij dan minstens zijn?

Oefening 14. De leeftijd van een elektronische component is een toevalsveranderlijke met verwachtingswaarde 5000 uren, een standaarddeviatie van 1000 uren en is bij benadering normaal verdeeld. Wat is de (benaderde) kans dat de gemiddelde leeftijd van 400 onafhankelijke componenten minder is dan 5012 uren?

Oefening 15. De installatie van een software pakket vereist het downloaden van (achtereenvolgens) 82 bestanden. Het duurt gemiddeld 15 seconden om een bestand te downloaden met een variantie van 16 sec^2. Benader de kans dat de software geïnstalleerd wordt binnen de 20 minuten?

Oefening 16. Veronderstel X en Y onafhankelijk, $X \sim \chi^2_2$ en $Y \sim \chi^2_3$. Bepaal

a. $\mathbb{P}(X < 0.5754)$

b. $\mathbb{P}(Y < 2.3660)$

c. $\mathbb{P}(X + Y < 15.0863)$

d. a zodat $\mathbb{P}(X \leq a) = 0.95$ (met andere woorden: bepaal $\chi^2_{2,0.95}$)

e. b zodat $\mathbb{P}(Y \geq b) = 0.05$

f. c zodat $\mathbb{P}(X + Y < c) = 0.10$

g. d zodat $\mathbb{P}(X \leq dY) = 0.05$

Oefening 17. Veronderstel $X \sim t_4$ en $Y \sim t_{40}$. Bepaal dan

a. $\mathbb{P}(X > 8.610)$

b. $\mathbb{P}(X < -8.610)$

c. $\mathbb{P}(Y > 1.303)$

d. a zodat $\mathbb{P}(X < a) = 0.75$

e. a zodat $\mathbb{P}(X > a) = 0.05$

f. a zodat $\mathbb{P}(Y < a) = 0.01$

Oefening 18.

a. Zij $T \sim T_{23}$, wat is de kans dat T groter is dan 2.5

b. Zij $T \sim T_n$ en $P(T < d) = 0.95$. Bepaal de waarde van d voor $n \in \{3, 13, 30\}$.

c. Als X en Y onafhankelijk zijn en $X \sim N(0, 4)$ en $Y \sim \chi^2_9$, bepaal dan de waarde d zodat $P(X < d\sqrt{Y}) = 0.975$.

Oefening 19. Veronderstel dat $X \sim F_{1,2}$, $Y \sim F_{5,10}$ en $Z \sim F_{10,5}$. Bepaal

a. a zodat $\mathbb{P}(X < a) = 0.95$

b. a zodat $\mathbb{P}(Y > a) = 0.05$

c. a zodat $\mathbb{P}(Z > a) = 0.95$

d. a zodat $\mathbb{P}(Y < a) = 0.05$

e. $\mathbb{P}(0.005 < X < 18.5)$

Oefening 20.

a. Veronderstel dat Z_1 en Z_2 onafhankelijk en standaard normaal verdeeld zijn, bereken dan:
$$\mathbb{P}(Z_1^2 + Z_2^2 \leq 1.3863)$$

b. Zij $X \sim \chi^2_{13}$, bepaal $\mathbb{P}(|X| < 9.2991)$

c. Zij $X \sim N(0,3)$ en $Y \sim N(0,3)$ twee onafhankelijke variabelen, bepaal a zodat
$$\mathbb{P}(X^2 + Y^2 < a) = 0.975$$

d. Bepaal a zodat $\mathbb{P}(F < a) = 0.95$, als $F \sim F_{m,n}$ en voor
$$(m,n) \in \{(7,5); (5,7); (5,20)\}\,.$$

e. Bepaal a zodat $\mathbb{P}(F < a) = 0.05$, als $F \sim F_{17,30}$.

f. Als X_1, X_2, X_3, X_4, X_5 onderling onafhankelijke standaard normaal verdeelde veranderlijken zijn, bepaal dan a zodat
$$\mathbb{P}\left(\frac{X_1^2 + X_2^2}{X_3^2 + X_4^2 + X_5^2} < a\right) = 0.95.$$

Extra oefeningen

Oefening 21. Veronderstel dat $Z_1 \sim N(0,1)$ en $Z_2 \sim N(0,1)$, bereken dan $\mathbb{P}(Z_1^2 \leq 1 - Z_2^2)$ als je weet dat Z_1 en Z_2 onafhankelijk zijn.

Oefening 22. De installatietijd, in uren, van een bepaalde software heeft een dichtheidsfunctie $f(x) = k(1 - x^3)$ voor $0 < x < 1$. Bepaal k en bereken de kans dat het minder dat 30 minuten duurt om de software te installeren.

Oefening 23. Een computer voert taken uit in de volgorde dat hij ze ontvangt. De tijd nodig voor elke taak heeft een exponentiele verdeling met een gemiddelde van 2 minuten. Bereken de kans dat één taak minder dan 1 minuut duurt. Bereken ook de kans dat 5 taken minder dan 8 minuten duren.

Oefening 24. Als $X \sim N(1,4)$, bepaal dan $P(1 \leq X^2 \leq 9)$.

Oefening 25. We onderzoeken of de eieren van de Belgische pluimveehouders waar de insecticide fipronil is gevonden veilig zijn voor consumptie. We weten dat het gemiddelde fipronilgehalte 0.017 mg/kg is met standaarddeviatie 0.07 mg/kg. De concentraties fipronil zijn echter niet normaal verdeeld. We weten wel dat als we het natuurlijk logaritme van de concentraties nemen, de gegevens wel normaal verdeeld zijn, met gemiddelde -6.1 en standaarddeviatie 1.7.

a. Bereken het percentage van de eieren waarvan het fipronilgehalte hoger ligt dan de veiligheidsgrens 0.05 mg/kg en die dus niet bruikbaar zijn voor consumptie.

b. Veronderstel dat je een doos koopt met 36 eieren. Wat is dan de kans dat er minstens 2 eieren een te hoog fipronilgehalte (> 0.05 mg/kg) hebben?