Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Εαρινό Εξάμηνο 2023-2024

ΠΙΘΑΝΟΤΗΤΕΣ ΚΑΙ ΣΤΑΤΙΣΤΙΚΗ

Λύσεις Θεμάτων

Ιωάννης (Χουάν) Τσαντήλας

03120883

Github-Repo

Contents

Κ22

Ε22

Π22

Πολλαπλής Επιλογής (σκόρπια)

Ερώτημα 1

Έστω A, B, C τρία ενδεχόμενα και P ένα μέτρο πιθανότητας. Ποιο/α από τα παρακάτω είναι αληθή;

α. $P(A \cup B \cup C) = P(A) + P(B) + P(C)$
β. $P(A \cup B \cup C) = P(B) + P(B' \cap C) + P(B' \cap C' \cap A)$
γ. $P(A \cup B \cup C) = P(A) + P(A' \cap B) + P((A \cup B)' \cap C)$

Λύση

Το **(γ)**.

Ερώτημα 2

Έστω μία συνηθισμένη τράπουλα με 52 φύλλα. Ποιο/α από τα παρακάτω είναι αληθή;

α. Αν μοιράσουμε με τη σειρά 6 φύλλα υπάρχουν $52!/(6!\,46!)$ δυνατές εξάδες.
β. Αν μοιράσουμε 6 φύλλα υπάρχουν $52!/46!$ δυνατές εξάδες.
γ. Αν μοιράσουμε με τη σειρά 9 φύλλα υπάρχουν $52!/43!$ δυνατές εννιάδες.

Λύση

Το **(γ)**.

Ερώτημα 3

Έστω X, Y τυχαίες μεταβλητές που αναπαριστούν τα αποτελέσματα των ρίψεων δύο δίκαιων ζαριών. Ορίζουμε την τυχαία μεταβλητή $V = X + Y$. Ποιο/α από τα παρακάτω είναι αληθή;

α. Η πιθανότητα η V να πάρει τιμή 2 είναι ίση με 1/36.
β. Η κατανομή της V μπορεί να είναι ομοιόμορφη στο σύνολο {2,4,6,8,10,12}.
γ. Η κατανομή της V μπορεί να συγκεντρώνεται στο μονοσύνολο {7}.

Λύση

Τα **(α, γ)**.

Ερώτημα 4

Έστω X μία διακριτή τυχαία μεταβλητή με μέση τιμή μ και διασπορά σ^2. Ποιο/α από τα παρακάτω είναι αληθή;

α. $P(X \geq c) \leq \mu/c, \quad \forall c > 0$
β. $P(|X - \mu| \geq c) \leq \sigma^2/c^2, \forall c > 0$
γ. $P(|X - \mu| \geq c) \leq \sigma^2/d^2$, για κάποιο d>0.

Λύση

Τα **(α, β)**.

Ερώτημα 5

Οι παρακάτω συναρτήσεις είναι πυκνότητες πιθανότητας συνεχών τυχαίων μεταβλητών:

α. $f(x) = \begin{cases} \frac{3}{4}(1-x^2), -1 \le x \le 1 \\ 0, else \end{cases}$

β. $f(x) = \begin{cases} \frac{3}{4}(1-x), 0 \le x \le 1 \\ 0, else \end{cases}$

γ. $f(x) = \begin{cases} \left(\frac{1}{2} - x\right), -1 \le x \le 1 \\ 0, else \end{cases}$

Λύση

Το **(γ)**.

Ερώτημα 6

Έστω A, B, C τρία ενδεχόμενα και P ένα μέτρο πιθανότητας. Ποιο/α από τα παρακάτω είναι αληθή;

α. $P(A \cup B \cup C) = P(A) + P(B) + P(C)$

β. $P(A \cup B \cup C) = P(B) + P(B) + P(C) - P(A \cap B \cap C)$

γ. $P(A \cup B \cup C) = P(A) + +P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$

Λύση

Το **(γ)**.

Ερώτημα 7

Έστω X, Y τυχαίες μεταβλητές που αναπαριστούν τα αποτελέσματα των ρίψεων δύο δίκαιων ζαριών. Ορίζουμε την τυχαία μεταβλητή $V = X + Y$. Ποιο/α από τα παρακάτω είναι αληθή;

α. Η πιθανότητα η V να πάρει τιμή 3 είναι ίση με 2/36.

β. Η πιθανότητα η V να πάρει τιμή 7μπορεί να είναι ίση με 1.

γ. Η κατανομή της V μπορεί να είναι ομοιόμορφη στο σύνολο {2,4,6,8,10,12}.

Λύση

Το **(α, β)**.

Ερώτημα 8

Έστω X τ.μ. που ακολουθεί την Εκθετική κατανομή με παράμετρο 9. Ποιο/α από τα παρακάτω είναι αληθή;

α. $P(X \geq 7|X \geq 4) = e^{-1/3}$
β. $P(X \geq 7|X \geq 4) = e^{-4/9}$
γ. $P(X \geq 7|X \geq 4) \leq 3$

Λύση

Το **(γ)**.

Ερώτημα 9

Έστω $X_1, X_2, \dots, X_n$ ανεξάρτητες τ.μ. που ακολουθούν την Εκθετική κατανομή με άγνωστη παράμετρο θ.

α. Ο εκτιμητής μέγιστης πιθανοφάνειας για την θ δίνεται από $\widehat{\theta_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$.
β. Ο εκτιμητής μέγιστης πιθανοφάνειας για την θ δίνεται από $\widehat{\theta_n} = \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2$.
γ. Ο δειγματικός μέσος όρος είναι αμερόληπτος και συνεπής εκτιμητής για την μέση τιμή θ.

Λύση

Το **(α, γ)**.

Ερώτημα 10

Έστω 16 ομάδες που συμμετέχουν στο μουντιάλ. Οι 3 πρώτες ομάδες παίρνουν τα μετάλλια. Ποιο/α από τα παρακάτω είναι αληθή;

α. Αν δεν μας ενδιαφέρει η σειρά, υπάρχουν 16!/13! δυνατές τριάδες.
β. Αν δεν μας ενδιαφέρει η σειρά, υπάρχουν 16!/(3!13!) δυνατές τριάδες.
γ. Αν μας ενδιαφέρει η σειρά, υπάρχουν 16!/(3!13!) δυνατές τριάδες.

Λύση

Το **(β)**.

- *Combinations*: **<u>ΔΕΝ</u>** μας νοιάζει η σειρά: $C(n,k) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$
- *Permutations*: μας νοιάζει η σειρά: $P(n,k) = \frac{n!}{(n-k)!}$

Κανονική 23

Θέμα 1

Μια ασφαλιστική εταιρεία πιστεύει ότι οι άνθρωποι μπορούν να χωριστούν σε δύο κατηγορίες: εκείνους που είναι επιρρεπείς σε ατυχήματα και εκείνους που δεν είναι. Κατά τη διάρκεια μιας χρονιάς, ένας επιρρεπής σε ατυχήματα άνθρωπος έχει ατύχημα με πιθανότητα 0,4, ανεξάρτητα για διαφορετικές χρονιές. Η αντίστοιχη πιθανότητα για έναν άνθρωπο που δεν είναι επιρρεπής σε ατυχήματα είναι 0,2. Υποθέτουμε ότι το 30% των ασφαλισμένων είναι επιρρεπείς σε ατυχήματα.

α. Ποια είναι η πιθανότητα ένας πελάτης της εταιρείας να έχει ατύχημα κατά τον πρώτο χρόνο του συμβολαίου του;
β. Δεδομένου ότι ένας πελάτης της εταιρείας είχε ατύχημα κατά τον πρώτο χρόνο του συμβολαίου του, ποια είναι η πιθανότητα να είναι επιρρεπής σε ατυχήματα;
γ. Ποια είναι η πιθανότητα ένας πελάτης της εταιρείας να έχει ατύχημα τον δεύτερο χρόνο του συμβολαίου του, δεδομένου ότι είχε ατύχημα τον πρώτο χρόνο;
δ. Ανάμεσα σε 300 πελάτες της εταιρείας, ποια είναι προσεγγιστικά η πιθανότητα τουλάχιστον 88 από αυτούς να έχουν ατύχημα κατά τον πρώτο χρόνο του συμβολαίου τους;

Λύση

Ερώτημα (α)

Έστω Α το ενδεχόμενο ένας πελάτης να είναι επιρρεπής σε ένα ατύχημα και Α' να μην είναι. Έστω επίσης Β το ενδεχόμενο ένας πελάτης να έχει ατύχημα μέσα στη 1^η χρονιά και Β΄ να μην έχει. Ισχύει:

$$P(A) = 0.3, \qquad P(A') = 0.7$$

$$P(B|A) = 0.4, \qquad P(B|A') = 0.2$$

Ψάχνουμε την πιθανότητα ένας πελάτης να πάθει ατύχημα τον 1^ο χρόνο, δηλαδή το P(B).

$$P(B) = P(B|A) \cdot P(A) + P(B|A') \cdot P(A') = 0.4 \cdot 0.3 + 0.2 \cdot 0.7 = 0.26$$

Ερώτημα (β)

Ψάχνουμε την πιθανότητα ένας πελάτης να είναι επιρρεπής, δεδομένου ότι έπαθε ατύχημα το 1^ο χρόνο, δηλαδή P(A|B).

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} = \frac{0.4 \cdot 0.3}{0.26} = 0{,}4615$$

Ερώτημα (γ)

Έστω C το ενδεχόμενο να έχει ατύχημα μέσα στη 2^η χρονιά. Ψάχνουμε την πιθανότητα ένας πελάτης να πάθει ατύχημα τον 2^ο χρόνο, δεδομένου ότι έπαθε ατύχημα τον 1^ο χρόνο, δηλαδή το P(C|B). Ωστόσο, αφού οι χρονιές είναι μεταξύ τους ανεξάρτητες, θα ισούται απλά με 0,26.

Ερώτημα (δ)

Έστω X_i ο i-οστός πελάτης της εταιρείας, με $X_i = 0$ εάν δεν πάθει ατύχημα και $X_i = 1$ εάν πάθει ατύχημα, την 1η χρονιά. Το συνολικό πλήθος των πελατών που έπαθαν ατύχημα δίνεται από τον τύπο:

$$S = \sum_{i=1}^{300} X_i$$

Η X_i ακολουθεί κατανομή Bernoulli, με πιθανότητα $p = P(B) = 0.26$. Άρα έχουμε μέση τιμή $\mu = 0.26$ και διασπορά $\sigma^2 = p \cdot (1-p) = 0.1898$. Η S επίσης ακολουθεί κατανομή Bernoulli, ως γραμμικό άθροισμα των X. Ψάχνουμε την πιθανότητα $P(S \geq 88)$, ενώ έχουμε μέγεθος δείγματος N=300.

$$P(S \geq 88) = P(S - \mu \cdot N \geq 88 - \mu \cdot N) = P\left(\frac{S - \mu \cdot N}{\sqrt{\sigma^2 \cdot N}} \geq \frac{88 - \mu \cdot N}{\sqrt{\sigma^2 \cdot N}} = \frac{88 - 0.26 \cdot 300}{\sqrt{0.1898 \cdot 300}} = 1.3252\right)$$

Θέτω:

$$Y = \frac{S - \mu \cdot N}{\sqrt{\sigma^2 \cdot N}} \rightarrow P(Y \geq 1.3252) = 1 - P(Y \leq 1.3252)$$

Η Y ακολουθεί κανονική κατανομή N(0,1) και από τον πίνακα της κανονικής κατανομής έχουμε:

$$1 - P(Y \leq 1.3252) = 1 - \Phi(1.3252) = 1 - 0.9059 = 0.0941$$

Θέμα 2

Ένας πάροχος φυσικού αερίου πιστεύει ότι σε μία μέρα με τις σημερινές καιρικές συνθήκες, η κατανάλωση φυσικού αερίου ενός νοικοκυριού (σε κατάλληλη μονάδα μέτρησης) είναι τυχαία μεταβλητή (τ.μ.), X, με συνάρτηση πυκνότητας πιθανότητας (σ.π.π.):

$$f(x) = \begin{cases} c \cdot (1-x)^4, & 0 < x < 1 \\ 0, & otherwise \end{cases}$$

Ενώ οι καταναλώσεις διαφορετικών νοικοκυριών είναι ανεξάρτητες, ισόνομες τ.μ..

α. Να υπολογίσετε:
   i. Την τιμή της σταθεράς c.
   ii. Την συνάρτηση κατανομής πιθανότητας της τ.μ. X.

β. Να υπολογίσετε τη μέση τιμή και τη διασπορά της τ.μ. X.

γ. Σε μια πολυκατοικία, υπάρχουν 2 νοικοκυριά που χρησιμοποιούν φυσικό αέριο. Ποια είναι η πιθανότητα η σημερινή κατανάλωση της πολυκατοικίας να ξεπεράσει την τιμή 1 (στις μονάδες που μετράμε την κατανάλωση);

Λύση

Ερώτημα (α)

Για να βρούμε την σταθερά, θα πρέπει να κανονικοποιήσουμε την σ.π.π. ως ακολούθως:

$$1 = \int_{-\infty}^{\infty} f(x)\,dx = \int_0^1 c \cdot (1-x)^4\,dx = \left(c \cdot \frac{-(1-x)^5}{5}\right)_0^1 = \frac{c}{5} \to c = 5$$

Για να βρούμε την σ.κ.π.:

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_0^x 5 \cdot (1-t)^4\,dt = \left(5 \cdot \frac{-(1-t)^5}{5}\right)_0^x = 1-(1-x)^5$$

Δηλαδή:

$$F(x) = P(X \le x) = \begin{cases} 0, & x \le 0 \\ 1-(1-x)^5, & 0 \le x \le 1 \\ 1, & x \ge 1 \end{cases}$$

Ερώτημα (β)

Για να βρούμε την μέση τιμή της X:

$$\mu = E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\,dx = \int_0^1 x \cdot 5 \cdot (1-x)^4\,dx = \cdots = \frac{1}{6}$$

Ενώ για την διασπορά:

$$E[X^2] = \int_{-\infty}^{\infty} x^2 \cdot f(x)\,dx = \int_0^1 x^2 \cdot 5 \cdot (1-x)^4\,dx = \cdots = \frac{1}{21}$$

$$\sigma^2 = E[X^2] - \mu^2 = \frac{1}{21} - \left(\frac{1}{6}\right)^2 = \frac{5}{252}$$

Ερώτημα (γ)

Έστω πως η κατανάλωση των 2 νοικοκυριών είναι X_1, X_2 αντίστοιχα. Η συνολική κατανάλωση την πολυκατοικίας είναι ίση με $S = X_1 + X_2$. Ψάχνουμε την πιθανότητα η συνολική κατανάλωση τους να ξεπερνά το 1, δηλαδή $P(X_1 + X_2 \geq 1) = P(S \geq 1)$. Η S ακολουθεί την ίδια κατανομή με την X, αφού είναι γραμμικός συνδυασμός των X. Επιπλέον, έχουμε μέγεθος δείγματος N=2.

$$P(S \geq 1) = P(S - \mu \cdot N \geq 1 - \mu \cdot N) = P(\frac{S - \mu \cdot N}{\sqrt{\sigma^2 \cdot N}} \geq \frac{1 - \frac{1}{6} \cdot 2}{\sqrt{\frac{5}{252} \cdot 2}} = 3.3466)$$

Θέτω:

$$Y = \frac{S - \mu \cdot N}{\sqrt{\sigma^2 \cdot N}} \rightarrow P(Y \geq 3.3466) = 1 - P(Y \leq 3.3466)$$

Η Y ακολουθεί κανονική κατανομή N(0,1) και από τον πίνακα της κανονικής κατανομής έχουμε:

$$1 - P(Y \leq 3.3466) = 1 - \Phi(3.3466) = 1 - 0.9995 = 0.0005$$

Θέμα 3

Θέλουμε να εκτιμήσουμε την παράμετρο λ_0 μιας εκθετικής κατανομής από ένα τυχαίο δείγμα της, $X_1, \dots, X_n$. Υιοθετούμε την bayesian προσέγγιση και υποθέτουμε ότι η παράμετρος της κατανομής είναι μια τ.μ. Λ, της οποίας η εκ των προτέρων κατανομή είναι Εκθ(α) και δεδομένου ότι $\Lambda = \lambda$, οι τ.μ. $X_1, \dots, X_n$ είναι ανεξάρτητες, ισόνομες με κατανομή Εκθ(λ).

α. Ποια είναι η από κοινού σ.π.π. των $X_1, \dots, X_n$ και Λ, με βάση την προσέγγισή μας;
β. Ποια είναι η δεσμευμένη κατανομή της Λ δεδομένου ότι $X_1 = x_1, \dots, X_n = x_n$;
γ. Πώς περιμένετε να μοιάζει τυπικά η δεσμευμένη κατανομή του ερωτήματος (β), καθώς $n \to \infty$, και γιατί;

Λύση

Ερώτημα (α)

Η εκ των προτέρων κατανομή της Λ είναι Εκθ(α), δηλαδή:

$$f_\Lambda(\lambda) = ae^{-a\lambda}, \qquad \lambda \geq 0$$

Η υπό συνθήκη κατανομή της X_i δεδομένης της Λ=λ είναι:

$$f_{X_i|\Lambda}(x_i \mid \lambda) = \lambda e^{-\lambda x_i}, \qquad x_i \geq 0$$

Επειδή οι $X_1, \dots, X_n$ είναι ανεξάρτητες, η κοινή υπό συνθήκη σ.π.π. είναι:

$$f_{X_1,\dots,X_n|\Lambda}(x_1, \dots, x_n \mid \lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \hat{X}}, \qquad \hat{X} = \sum_{i=1}^{n} x_i$$

Επομένως, η κοινή σ.π.π. των $X_1, \dots, X_n$ και Λ είναι:

$$f_{X_1,\dots,X_n,\Lambda}(x_1, \dots, x_n, \lambda) = f_{X_1,\dots,X_n|\Lambda}(x_1, \dots, x_n \mid \lambda) \cdot f_\Lambda(\lambda) =$$

$$= \lambda^n e^{-\lambda \hat{X}} \cdot ae^{-a\lambda} = a\lambda^n e^{-\lambda(a+\hat{X})}$$

Ερώτημα (β)

Η εκ των υστέρων κατανομή $f_{\Lambda|X_1,\dots,X_n}(\lambda \mid x_1, \dots, x_n)$ είναι ανάλογη της κοινής κατανομής της Λ και των $X_1, \dots, X_n$. Χρησιμοποιώντας το θεώρημα του Bayes:

$$f_{\Lambda|X_1,\dots,X_n}(\lambda \mid x_1, \dots, x_n) \propto f_{X_1,\dots,X_n|\Lambda}(x_1, \dots, x_n \mid \lambda) f_\Lambda(\lambda)$$

Από το Ερώτημα (γ):

$$f_{\Lambda|X_1,\dots,X_n}(\lambda \mid x_1, \dots, x_n) \propto \lambda^n e^{-\lambda \hat{X}} \cdot ae^{-a\lambda} \to f_{\Lambda|X_1,\dots,X_n}(\lambda \mid x_1, \dots, x_n) \propto \lambda^n e^{-\lambda(a+\hat{X})}$$

Πρόκειται για τον πυρήνα μιας κατανομής Γάμμα, με παραμέτρους n+1 και $a + \hat{X}$:

$$\Lambda \mid X_1 = x_1, \dots, X_n = x_n \sim \text{Gamma}\left(n + 1, a + \sum_{i=1}^{n} x_i\right)$$

Ερώτημα (γ)

Καθώς το μέγεθος του δείγματος n αυξάνεται, η εκ των υστέρων κατανομή της Λ συνήθως γίνεται πιο συγκεντρωμένη γύρω από την πραγματική τιμή της παραμέτρου. Αυτό συμβαίνει επειδή:

- Με περισσότερα δεδομένα, η επιρροή της εκ των προτέρων μειώνεται και η εκ των υστέρων μεταβλητή κυριαρχείται από την πιθανότητα.
- Ο μέσος όρος της εκ των υστέρων κατανομής θα προσεγγίσει τον πραγματικό μέσο όρο της εκθετικής κατανομής, λ_0.

Συγκεκριμένα, για την εκ των υστέρων κατανομή:

$$\Lambda \mid X_1 = x_1, \dots, X_n = x_n \sim \mathrm{Gamma}\left(n+1, a + \sum_{i=1}^{n} x_i\right)$$

Ο μέσος όρος της κατανομής Γάμμα είναι:

$$\frac{n+1}{a + \sum_{i=1}^{n} x_i}$$

Καθώς το n τείνει στο άπειρο, το άθροισμα θα μεγαλώνει, ενώ το α θα γίνει αμελητέο. Έτσι, ο παραπάνω μέσος όρος θα συγκλίνει στο:

$$\frac{n}{n \cdot \lambda_0} = \frac{1}{\lambda_0}$$

Θέμα 4

Αν $\{U_n\}_{n\in N}$ είναι μια ακολουθία από ανεξάρτητες, ισόνομες τ.μ. με ομοιόμορφη κατανομή στο (0,1) και η τυχαία μεταβλητή N_ε , ακολουθεί κατανομή Γεω(ε) και είναι ανεξάρτητη από τις $U_1, U_2, \ldots$, να αποδείξετε ότι:

$$\varepsilon^{-1} \min_{1\le k\le N_\varepsilon} U_k \xrightarrow{d} X, \qquad \varepsilon \to 0$$

και να βρείτε τη σ.π.π. του ορίου X.

Λύση

Αν το $U_{(1)}$ είναι το ελάχιστο της ακολουθίας, τότε:

$$P\big(U_{(1)} > x\big) = (1-x)^n, \qquad x \in (0,1)$$

Άρα, η σ.κ.π. της $U_{(1)}$ δεδομένου ότι $N_\epsilon = n$ είναι:

$$F_{U_{(1)}|N_\epsilon=n}(x) = 1-(1-x)^n$$

Έστω $M_\epsilon = \min\limits_{1\le k\le N_\epsilon} U_k$. Η σ.κ.π. της $\epsilon^{-1}M_\epsilon$ προκύπτει από:

$$P(\epsilon^{-1}M_\epsilon \le x) = P(M_\epsilon \le \epsilon x)$$

Δεδομένου ότι η N_ϵ είναι μια γεωμετρική τ.μ. ανεξάρτητη από τις $U_1, U_2, \ldots,$, η σ.κ.π. της M_ϵ είναι:

$$P(M_\epsilon \le y) = \sum_{n=1}^{\infty} P(\, M_\epsilon \le y \mid N_\epsilon = n\,)P(N_\epsilon = n)$$

Δεδομένου ότι $N_\epsilon = n$:

$$P(\, M_\epsilon \le y \mid N_\epsilon = n\,) = 1-(1-y)^n$$

και η σ.μ.π. της N_ϵ είναι:

$$P(N_\epsilon = \mathrm{n}) = \epsilon(1-\epsilon)^{n-1}$$

Άρα:

$$P(M_\epsilon \le y) = \sum_{n=1}^{\infty} [1-(1-y)^n]\cdot\epsilon\cdot(1-\epsilon)^{n-1}$$

Όταν $\epsilon \to 0$ έχουμε: $(1-\epsilon)^{n-1} \approx e^{-\epsilon(n-1)}$. Άρα:

$$P(M_\epsilon \le y) \approx \epsilon \sum_{n=1}^{\infty} (1-(1-y)^n)\cdot e^{-\epsilon(n-1)}$$

Αυτό το άθροισμα κυριαρχείται από τις μικρές τιμές n, και χρησιμοποιώντας την εκθετική προσέγγιση, μπορεί να αποδειχθεί ότι:

$$\epsilon^{-1}M_\epsilon \xrightarrow{d} X$$

όπου η X έχει εκθετική κατανομή με ρυθμό 1. Η εκθετική κατανομή με ρυθμό 1 έχει σ.π.π.:

$$f_X(x) = e^{-x}, x \geq 0$$

Επαναληπτική 23

Θέμα 1

Η διάρκεια ζωής (σε έτη), Τ, ενός φωτοβολταϊκού στοιχείου (ΦΒΣ) είναι τυχαία μεταβλητή (τ.μ.) με συνάρτηση πυκνότητας πιθανότητας (σ.π.π.):

$$f(x) = \begin{cases} \frac{1}{6} \cdot (1 - k \cdot x^2), & 0 \le x \le 9 \\ 0, & otherwise \end{cases}$$

α. Να αποδείξετε ότι $k = 1/81$ και να βρείτε τη συνάρτηση κατανομής πιθανότητας (σ.κ.π.) της τ.μ. Τ.

β. Να υπολογίσετε τη μέση διάρκεια ζωής, μ, ενός ΦΒΣ αυτού του τύπου και την τυπική της απόκλιση, σ.

γ. Αν ένα ΦΒΣ τοποθετήθηκε πριν 3 χρόνια και λειτουργεί ακόμα, να υπολογίσετε την πιθανότητα το ΦΒΣ να είναι λειτουργικό για τουλάχιστον 3 χρόνια ακόμα.

δ. Σε ένα ηλιακό πάρκο τοποθετήθηκαν 728 τέτοια ΦΒΣ πριν από 3 χρόνια. Να υπολογίσετε προσεγγιστικά την πιθανότητα τουλάχιστον 360 από αυτά να είναι ακόμα λειτουργικά.

Λύση

Ερώτημα (α)

Για να βρούμε την σταθερά, θα πρέπει να κανονικοποιήσουμε την σ.π.π. ως ακολούθως:

$$1 = \int_{-\infty}^{\infty} f(x)\, dx = \int_0^9 \frac{1}{6} \cdot (1 - k \cdot x^2)\, dx = \left(\frac{x}{6} - \frac{k \cdot x^3}{18} \right)_0^9 = \cdots \to k = 1/81$$

Για να βρούμε την σ.κ.π.:

$$F(x) = \int_{-\infty}^{x} f(t)\, dt = \int_0^x \frac{1}{6} \cdot \left(1 - \frac{1}{81} \cdot t^2\right) dt = \cdots = \frac{x}{6} - \frac{x^3}{1458}$$

Δηλαδή:

$$F(x) = P(X \le x) = \begin{cases} 0, & x \le 0 \\ \frac{x}{6} - \frac{x^3}{1458}, & 0 < x < 9 \\ 1, & x \ge 9 \end{cases}$$

Ερώτημα (β)

Για να βρούμε την μέση τιμή της Χ:

$$\mu = E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\, dx = \int_0^9 \frac{x^2}{6} - \frac{x^4}{1458} dx = \cdots = \frac{27}{8}$$

Ενώ για την διασπορά:

$$E[X^2] = \int_{-\infty}^{\infty} x^2 \cdot f(x)\, dx = \int_0^9 \frac{x^3}{6} - \frac{x^5}{1458} dx = \cdots = \frac{81}{5}$$

$$\sigma^2 = E[X^2] - \mu^2 = \frac{81}{5} - \left(\frac{27}{8}\right)^2 = 4.8093$$

Ερώτημα (γ)

Ψάχνουμε την πιθανότητα να είναι λειτουργικό μετά από 6 χρόνια, δεδομένου ότι είναι λειτουργικό μετά από 3 χρόνια, δηλαδή $P(X \geq 6|X \geq 3)$.

$$P(X \geq 6|X \geq 3) = P(X \geq 3|X \geq 6) \cdot \frac{P(X \geq 6)}{P(X \geq 3)}$$

$$P(X \geq 6) = 1 - F(6) = 1 - \frac{6}{6} + \frac{6^3}{1458} = \frac{4}{27}$$

$$P(X \geq 3) = 1 - F(3) = 1 - \frac{3}{6} + \frac{3^3}{1458} = \frac{14}{27} = 0.\overline{518}$$

Η πιθανότητα $P(X \geq 3|X \geq 6)$ ισούται με 1, αφού εάν είναι λειτουργικό μετά από 6 χρόνια, τότε προφανώς θα ήταν και μετά από 3 χρόνια. Άρα:

$$P(X \geq 6|X \geq 3) = \frac{\frac{4}{27}}{\frac{14}{27}} = \frac{2}{7}$$

Ερώτημα (δ)

Έστω X_i το i-οστό ΦΒΣ, με $X_i = 0$ εάν δεν λειτουργεί μετά από 3 χρόνια και $X_i = 1$ εάν λειτουργεί. Το συνολικό πλήθος των ΦΒΣ που λειτουργούν μετά από 3 χρόνια δίνεται από τον τύπο:

$$S = \sum_{i=1}^{728} X_i$$

Η X_i ακολουθεί κατανομή Bernoulli, με πιθανότητα $p = P(X \geq 3) = 0.518$. Άρα έχουμε μέση τιμή $\mu = 0.518$ και διασπορά $\sigma^2 = p \cdot (1-p) = 0.2496$. Η S επίσης ακολουθεί κατανομή Bernoulli, ως γραμμικός συνδυασμός των X. Ψάχνουμε την πιθανότητα $P(S \geq 360)$, ενώ έχουμε μέγεθος δείγματος $N = 728$.

$$P(S \geq 360) = P(S - \mu \cdot N \geq 360 - \mu \cdot N) = P\left(\frac{S - \mu \cdot N}{\sqrt{\sigma^2 \cdot N}} \geq \frac{360 - \mu \cdot N}{\sqrt{\sigma^2 \cdot N}} = \frac{360 - 0.518 \cdot 728}{\sqrt{0.2496 \cdot 728}}\right)$$

Θέτω:

$$Y = \frac{S - \mu \cdot N}{\sqrt{\sigma^2 \cdot N}} \to P(Y \geq -1.2688) = 1 - P(Y \leq -1.2688)$$

Η Y ακολουθεί κανονική κατανομή N(0,1) και από τον πίνακα της κανονικής κατανομής έχουμε:

$$1 - P(Y \leq -1.2688) = 1 - \Phi(-1.2688) = 1 - \big(1 - \Phi(1.2688)\big) = \Phi(1.2688) = 0.9026$$

Θέμα 2

Σε ένα δάσος, ζουν άτομα από ένα απειλούμενο με εξαφάνιση είδος. Κάθε άτομο του είδους επιλέγει κάθε μέρα, ανεξάρτητα από τα άλλα άτομα και από προηγούμενες επιλογές που έχει κάνει, μία από τις 5 πηγές του δάσους και την επισκέπτεται για νερό. Έχετε τοποθετήσει κάμερες σε μία πηγή Π, και θα την παρακολουθείτε επί μία εβδομάδα.

α. Αν το πλήθος των ατόμων που ζουν στο δάσος είναι Ν, να αποδείξετε ότι το πλήθος των ατόμων που επισκέπτονται την πηγή Π κάθε μέρα είναι τ.μ. με κατανομή Διων(Ν,1/5).

β. Να υπολογίσετε τη συνάρτηση μάζας πιθανότητας (σ.μ.π.) του πλήθους, Ι, των διαφορετικών ατόμων που θα επισκεφτούν την πηγή Π κατά τη διάρκεια της εβδομάδας που ακολουθεί.

Θέλουμε να εκτιμήσουμε το άγνωστο πλήθος, Ν, με τη βοήθεια της τ.μ. Ι που μπορούμε να παρατηρήσουμε. Υιοθετούμε την Bayesian προσέγγιση και υποθέτουμε για τη Ν την εκ των προτέρων κατανομή $Po(\lambda = 50)$.

γ. Να υπολογίσετε τη σ.μ.π. της τ.μ. Ι με αυτή την προσέγγιση.

δ. Να υπολογίσετε την εκ των υστέρων κατανομή της Ν, δεδομένου ότι $I = 60$.

Λύση

Ερώτημα (α)

Έχουμε Ν άτομα, 5 πηγές και πιθανότητα επιλογής οποιασδήποτε πηγής 0,2 (συμπεριλαμβανομένου και της Π). Έχουμε δηλαδή Ν δοκιμές Bernoulli για την πηγή Π, με πιθανότητα 0.2 (δηλαδή ένα άτομο είτε να επιλέξει την Π είτε όχι). Έστω Χ το πλήθος των ατόμων που επέλεξαν την Π. Η Χ θα ακολουθεί διωνυμική κατανομή με παραμέτρους Ν και 0,2.

Ερώτημα (β)

Η πιθανότητα ένα άτομο να **<u>μην</u>** επιλέξει την Π μία οποιαδήποτε ημέρα είναι 0,8. Γνωρίζουμε πως η επιλογή σε σχέση με τα άλλα άτομα είναι ανεξάρτητη, όπως επίσης είναι ανεξάρτητη από τις προηγούμενες επιλογές του ατόμου. Επομένως, η πιθανότητα ένα άτομο να **<u>μην</u>** επιλέξει την Π για 7 ημέρες θα είναι απλά $0{,}8^7 = 0{,}2097152$.

Συμπληρωματικά, η πιθανότητα ένα άτομο να επιλέξει την Π **<u>τουλάχιστον</u>** μία μέρα μέσα στην εβδομάδα είναι $1 - 0{,}2097152 = 0{,}7902848$.

Η Ι ακολουθεί διωνυμική κατανομή με παραμέτρους Ν και $0{,}7902848$. Η συνάρτηση μάζας πιθανότητας είναι:

$$P[I = i] = \binom{N}{i} (0{,}79)^i (0{,}209)^{N-i}$$

Ερώτημα (γ)

Γνωρίζουμε από το Ερώτημα (β) την κατανομή της Ι. Επομένως, δεδομένου ότι το πλήθος Ν είναι n, η νέα συνάρτηση μάζας πιθανότητας της Ι είναι:

$$P(I = i \mid N = n) = \binom{n}{i} (0{,}79)^i (0{,}209)^{n-i}$$

Ερώτημα (δ)

Ψάχνουμε την πιθανότητα $P(N = n \mid I = 60)$. Από τον κανόνα του Bayes:

$$P(\,N = n \mid I = 60\,) = \frac{P(\,I = 60 \mid N = n\,) \cdot P(N = n)}{P(I = 60)}$$

Γνωρίζουμε πως η N ακολουθεί κατανομή Poisson με παράμετρο 50, άρα:

$$P(N = n) = \frac{e^{-50} 50^n}{n!}$$

Επιπλέον, από το Ερώτημα (γ) προκύπτει:

$$P(\,I = 60 \mid N = n\,) = \binom{n}{60} (0{,}79)^{60} (0{,}209)^{n-60}$$

Από LTP (Law of Total Probability) έχουμε:

$$P(I = 60) = \sum_{n=60}^{\infty} P(\,I = 60 \mid N = n\,) \cdot P(N = n)$$

Συνδυάζοντας όλους τους τύπους:

$$P(\,N = n \mid I = 60\,) = \frac{\binom{n}{60} p^{60} (1-p)^{n-60} \cdot \frac{e^{-50} 50^n}{n!}}{\sum_{m=60}^{\infty} \binom{m}{60} p^{60} (1-p)^{m-60} \cdot \frac{e^{-50} 50^m}{m!}}$$

<u>Σημείωση (όχι απαραίτητα σωστή)</u>

Στην πράξη, μπορούμε να το προσεγγίσουμε. Για μικρές τιμές του p, εδώ βέβαια είναι 0,79, έχουμε:

$$P(\,N = n \mid I = 60\,) \approx \text{Poisson}(N; \lambda = 50 \cdot p)$$

$$\lambda \approx 50 \cdot 0{,}79 \approx 39{,}5$$

Δηλαδή:

$$P(\,N = n \mid I = 60\,) \approx \frac{e^{-41.6} 41.6^n}{n!}$$

Θέμα 3

Οι τ.μ. X και Y είναι ανεξάρτητες και ακολουθούν την τυπική κανονική κατανομή N(0,1).

α. Να υπολογίσετε την από κοινού σ.π.π. των τ.μ. $U = X$ και $V = \frac{Y}{X}$.
β. Να βρείτε τη σ.κ.π. της τ.μ. $W = \max\{X, Y\}$.

Δύο πανομοιότυπα σώματα, A και B, κινούνται με σταθερή ταχύτητα κατά μήκος του άξονα x'x. Αρχικά βρίσκονται στις θέσεις $x_A = 0$ και $x_B = -1$ (σε m) και έχουν ταχύτητες u_A και u_B (σε m/sec) οι οποίες είναι ανεξάρτητες τυπικές κανονικές τ.μ.. Αν τα σώματα συγκρουστούν, η σύγκρουσή τους θα είναι ελαστική και, όπως γνωρίζουμε από τη Φυσική, θα ανταλλάξουν ταχύτητες.

γ. Ποια πιστεύετε ότι θα είναι η σ.π.π. της τελικής ταχύτητας του σώματος A και γιατί;
δ. Να βρείτε τη σ.κ.π. της θέσης του σώματος A μετά από 1 sec.

Λύση

Ερώτημα (α)

Έχουμε πως:

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}, \qquad V = Y/X \to Y = V \cdot X = V \cdot U$$

$$\hat{J} = \begin{pmatrix} \frac{\partial X}{\partial U} & \frac{\partial X}{\partial V} \\ \frac{\partial Y}{\partial U} & \frac{\partial X}{\partial V} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ V & U \end{pmatrix} \to det(\hat{J}) = U$$

Αφού οι X, Y είναι ανεξάρτητες, η από κοινού σ.π.π. τους θα είναι το γινόμενο των σ.π.π. τους:

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2+y^2}{2}}$$

Άρα η από κοινού σ.π.π. των U, V:

$$f_{U,V}(u,v) = f_{X,Y}(x,y) \cdot \left|det(\hat{J})\right| = f_{X,Y}(u, u \cdot v) \cdot \left|det(\hat{J})\right| = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2+(u \cdot v)^2}{2}} \cdot |u|$$

Ερώτημα (β)

Έχουμε πως:

$$F_W(w) = P[W \leq w] = P[max(X,Y) \leq w] = \begin{cases} P[X \leq w], & if X > Y \\ P[Y \leq w], & else \end{cases}$$

Το τελευταίο ισοδυναμεί με $P[X \leq w \cap Y \leq w]$ και αφού οι X, Y είναι ανεξάρτητες μεταξύ τους, η πιθανότητα ισούται με το γινόμενο των επιμέρους πιθανοτήτων:

$$F_W(w) = P[X \leq w] \cdot P[Y \leq w] = \Phi(w) \cdot \Phi(w) = \Phi^2(w)$$

Ερώτημα (γ)

Η θέση των δύο σωμάτων την στιγμή t (πριν την ενδεχόμενη κρούση) δίνεται από τις εξισώσεις:

$$x_A(t) = x_A(0) + u_A \cdot t = u_A \cdot t$$

$$x_B(t) = x_B(0) + u_B \cdot t = -1 + u_B \cdot t$$

Για να συγκρουστούν θα πρέπει κάποια στιγμή να έχουν την ίδια θέση, την ίδια χρονική στιγμή, δηλαδή:

$$x_A(t_{\kappa\rho}) = x_B(t_{\kappa\rho}) \rightarrow u_A \cdot t_{\kappa\rho} = -1 + u_B \cdot t_{\kappa\rho} \rightarrow t_{\kappa\rho} = \frac{1}{u_B - u_A}$$

Προφανώς, η χρονική στιγμή θα πρέπει να είναι θετική, δηλαδή:

$$t_{\kappa\rho} \geq 0 \rightarrow \frac{1}{u_B - u_A} \geq 0 \rightarrow u_B - u_A \geq 0$$

Άρα, η πιθανότητα να συγκρουστούν θα είναι $P[u_B - u_A \geq 0]$. Θέτω $D = u_B - u_A$, η οποία ακολουθεί την ίδια κατανομή, ως γραμμικός συνδυασμός τ.μ. Άρα:

$$P[u_B - u_A \geq 0] = P[u_A - u_B \leq 0] = \Phi(0) = 0.5$$

Δηλαδή έχουμε 50% πιθανότητα να συγκρουστούν, με τα εξής αποτελέσματα:

$$\mathrm{u}'_A = \begin{cases} u_A, & if\ u_A > u_B \\ u_B, & else \end{cases}$$

Σε κάθε περίπτωση, η τελική ταχύτητα θα είναι ίση με μία από τις αρχικές, οι οποίες ακολουθούν κανονική κατανομή. Άρα και η $\mathrm{u}'_A \sim N(0,1)$.

Ερώτημα (δ)

Παρατηρούμε πως έχουμε 2 περιπτώσεις:

<u>Περίπτωση 1: Δεν γίνεται κρούση ή γίνεται, αλλά μετά το 1 δευτερόλεπτο</u>

Η θέση του Α μετά από 1 δευτερόλεπτο είναι:

$$x_A(1) = u_A \cdot 1 = u_A \sim N(0,1)$$

Επομένως, η σ.κ.π. θα είναι:

$$\mathrm{F}_{x_A(1)}(x) = \Phi(x)$$

Αυτή η περίπτωση μπορεί να ισχύει αν η κρούση γίνει ακριβώς στο 1 δευτερόλεπτο, αφού τη θεωρούμε στιγμιαία.

<u>Περίπτωση 2: Γίνεται κρούση πριν το 1 δευτερόλεπτο</u>

Εάν έχουμε κρούση, έστω την χρονική στιγμή $t_{\kappa\rho}$, τότε η θέση του Α μετά την κρούση θα δίνεται από την εξίσωση:

$$x'_A(t) = u_A \cdot t_{\kappa\rho} + u_B \cdot (t - t_{\kappa\rho}) \rightarrow x'_A(1) = u_A \cdot t_{\kappa\rho} + u_B \cdot (1 - t_{\kappa\rho}) \xrightarrow[(\alpha)]{}$$

$$x'_A(1) = u_A \cdot \frac{1}{u_B - u_A} + u_B \cdot \left(1 - \frac{1}{u_B - u_A}\right) = \frac{u_A}{u_B - u_A} + u_B - \frac{u_B}{u_B - u_A} = u_B + \frac{u_A - u_B}{u_B - u_A} = u_B - 1$$

Άρα, η $x'_A(1)$ ακολουθεί επίσης κανονική κατανομή, απλά μετατοπισμένη, δηλαδή $x'_A(1) \sim N(-1,1)$. Επομένως:

$$\mathrm{F}_{x'_A(1)}(x) = \Phi(x + 1)$$

Επί Πτυχίω 23

Θέμα 1

Ένας φορητός υπολογιστής λειτουργεί ασταμάτητα. Η φόρτιση της μπαταρίας του (από πλήρως άδεια σε πλήρως φορτισμένη) διαρκεί 1 ώρα. Κάθε φορά που ο υπολογιστής φορτίζεται πλήρως σταματάμε την παροχή ρεύματος μέχρι η μπαταρία του να εξαντληθεί ξανά. Υποθέτουμε ότι ο χρόνος λειτουργίας (σε ώρες) ανάμεσα σε διαδοχικές πλήρεις φορτίσεις της μπαταρίας περιγράφεται από την τυχαία μεταβλητή (τ.μ.)

$$T = 4 - X^3$$

όπου η τ.μ. X ακολουθεί ομοιόμορφη κατανομή στο διάστημα [-3,1].

α. Ποια είναι η μέση τιμή και ποια είναι η διασπορά του χρόνου T;
β. Ποια είναι η πιθανότητα ο εν λόγω υπολογιστής να λειτουργήσει πάνω από 12 ώρες ανάμεσα σε διαδοχικές φορτίσεις;
γ. Ξεκινάτε να χρησιμοποιείτε τον υπολογιστή μετά από μία πλήρη φόρτιση και τον λειτουργείτε για 4 ώρες χωρίς να χρειαστεί φόρτιση. Ποια είναι η πιθανότητα ο υπολογιστής να λειτουργήσει για ακόμα 4 ώρες χωρίς να χρειαστεί φόρτιση;
δ. Αν αρχικά ο υπολογιστής ξεκινά με άδεια μπαταρία, να υπολογίσετε προσεγγιστικά την πιθανότητα να αποφορτιστεί τουλάχιστον 240 φορές τις πρώτες 100 ημέρες λειτουργίας του.
ε. Ποια είναι η συνάρτηση κατανομής πιθανότητας (σ.κ.π.) του χρόνου T;

Λύση

Ερώτημα (α)

Για μια ομοιόμορφη τυχαία μεταβλητή $X \sim U(a, b)$, ο μέσος όρος της X^k είναι:

$$E[X^k] = \frac{1}{b-a}\int_a^b x^k \, dx$$

Στην περίπτωσή μας, $a = -3$, $b = 1$, και $k = 3$:

$$E[X^3] = \frac{1}{1-(-3)}\int_{-3}^{1} x^3 \, dx = \frac{1}{4}\left[\frac{x^4}{4}\right]_{-3}^{1} = \frac{1}{4}\left(\frac{1^4}{4} - \frac{(-3)^4}{4}\right) = \cdots = -5$$

Έτσι, η μέση τιμή του T είναι:

$$E[T] = E[4 - X^3] = 4 - E[X^3] = 4 - (-5) = 4 + 5 = 9$$

Η διασπορά της X^3:

$$\mathrm{Var}(X^3) = E[X^6] - (E[X^3])^2 = E[X^6] - 125$$

Χρησιμοποιώντας τον ίδιο τύπο όπως και προηγουμένως:

$$E[X^6] = \frac{1}{4}\int_{-3}^{1} x^6 \, dx = \frac{1}{4}\left[\frac{x^7}{7}\right]_{-3}^{1} = \frac{1}{4}\left(\frac{1^7}{7} - \frac{(-3)^7}{7}\right) = \frac{547}{7}$$

Άρα $\mathrm{Var}(X^3) = 53.142857$.

Ερώτημα (β)

Ψάχνουμε την πιθανότητα $\mathrm{P}(T > 12) = P(4 - X^3 > 12) = P(X^3 < -8) = P(X < -2)$. Δεδομένου ότι $X \sim U(-3, 1)$:

$$P(X < -2) = \frac{-2 - (-3)}{1 - (-3)} = \frac{1}{4} = 0.25$$

Ερώτημα (γ)

Δεδομένου ότι ο υπολογιστής έτρεξε για 4 ώρες, θέλουμε την πιθανότητα να τρέξει άλλες 4 ώρες:

$$P(\,T > 8 \mid T > 4\,) = P(\,X^3 < -4 \mid X^3 < 0\,) = \frac{P(X^3 < -4 \cap X^3 < 0)}{P(X^3 < 0)} = \frac{P(X^3 < -4)}{P(X^3 < 0)} = \frac{P\big(X < -\sqrt[3]{4}\big)}{P(X < 0)}$$

Οι επιμέρους πιθανότητες:

$$P\big(X < -\sqrt[3]{4}\big) = \frac{-\sqrt[3]{4} - (-3)}{1 - (-3)} = \frac{3 - \sqrt[3]{4}}{4} = 0.3531$$

$$P(X < 0) = \frac{0 - (-3)}{1 - (-3)} = \frac{3}{4} = 0.75$$

Συνολικά, $P(\,T > 8 \mid T > 4\,) = 0{,}4708$.

Ερώτημα (δ)

Έστω πως συμβολίσουμε με Α το πλήθος των αποφορτίσεων του υπολογιστή. Δεδομένου ότι ο υπολογιστής λειτουργεί ασταμάτητα και ο μέσος χρόνος μεταξύ των φορτίσεων είναι 9 ώρες, ο αριθμός των αποφορτίσεων σε 100 ημέρες (2400 ώρες) ακολουθεί κατανομή Poisson με παράμετρο $\lambda = \frac{2400}{9} \approx 266{,}66$. Πρόχειρα, εάν το λ είναι πολύ μεγάλο, η κατανομή αυτή *προσεγγίζει* την κανονική, με μέση τιμή και διασπορά ίση με λ. Άρα, προκύπτει:

$$P(N > 240) = 1 - P(N \le 240) = 1 - P\left(\frac{N - 266{,}66}{\sqrt{266{,}66}} \le \frac{240 - 266{,}66}{\sqrt{266{,}66}} = -1{,}6326\right)$$

Θέτοντας $Z = \frac{N - 266{,}66}{\sqrt{266{,}66}}$, έχουμε:

$$P(N > 240) = 1 - P(Z \le -1{,}6326) = 1 - \Phi(-1{,}6326) = 1 - 0{,}0462 = 0{,}9538$$

Ερώτημα (ε)

Η συνάρτηση κατανομής πιθανότητας της Χ:

$$F_X(x) = \frac{x - (-3)}{1 - (-3)} = \frac{x - 3}{4}, \qquad -3 \le x \le 1$$

Η συνάρτηση πυκνότητας πιθανότητας της Χ:

$$f_X(x) = \frac{1}{4}, \qquad -3 \le x \le 1$$

Η συνάρτηση κατανομής πιθανότητας της Τ:

$$F_T(t) = \mathrm{P}(T \le t) = P\big(X \le (4 - t)^{1/3}\big) = \frac{(4 - t)^{1/3} + 3}{4}, \qquad -23 \le t \le 3$$

Η συνάρτηση πυκνότητας πιθανότητας της T:

$$f_T(t) = \frac{d}{dt}F_T(t) = \frac{-(4-t)^{2/3}}{12}, \qquad -23 \leq x \leq 3$$

Θέμα 2

Οι τ.μ. Χ και Υ είναι ανεξάρτητες, ισόνομες με συνάρτηση πυκνότητας πιθανότητας:

$$f(x) = \lambda e^{-\lambda x}, \qquad x > 0$$

για κάποιο λ>0.

α. Να βρείτε την από κοινού σ.π.π. των (X,Y).

β. Να αποδείξετε ότι $P[X > Y + 1] = \frac{1}{2} P[X > 1]$.

γ. Να υπολογίσετε τη σ.π.π. της τ.μ. $X + 2Y$.

δ. Κάνουμε $N = 2000$ ανεξάρτητες παρατηρήσεις $x_1, x_2, \ldots, x_N$ της μεταβλητής Χ. Αν για 400 ακριβώς $k \in 1,2, \ldots ,1000$ έχουμε $x_{2k-1} \leq x_{2k} + 1$, να βρείτε την εκτίμηση της παραμέτρου λ με τη μέθοδο της μέγιστης πιθανότητας.

Λύση

Ερώτημα (α)

Δεδομένου ότι οι Χ, Υ είναι ανεξάρτητες, η από κοινού σ.π.π. είναι το γινόμενο των επιμέρους σ.π.π. τους:

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y) = \lambda e^{-\lambda x} \cdot \lambda e^{-\lambda y} = \lambda^2 e^{-\lambda(x+y)}, \qquad x > 0, y > 0]$$

Ερώτημα (β)

$$P[X > Y + 1] = \int_0^{\infty} \int_{y+1}^{\infty} f_{X,Y}(x,y)\; dx\, dy = \int_0^{\infty} \int_{y+1}^{\infty} \lambda^2 e^{-\lambda(x+y)}\; dx\, dy = \cdots = \frac{e^{-\lambda}}{2}$$

$$P[X > 1] = \int_1^{\infty} \lambda e^{-\lambda x}\; dx = \left[-e^{-\lambda x}\right]_1^{\infty} = e^{-\lambda}$$

Άρα:

$$P[X > Y + 1] = \frac{1}{2} P[X > 1]$$

Ερώτημα (γ)

Έστω $Z = X + 2Y$. Για να βρούμε τη σ.π.π. της Ζ, χρησιμοποιούμε τη συνέλιξη των σ.π.π. των Χ, 2Υ:

$$f_X(x) = \lambda e^{-\lambda x}, \qquad x > 0$$

$$f_{2Y}(t) = \frac{\lambda}{2} e^{-\frac{\lambda}{2}t}, \qquad t > 0$$

Άρα:

$$f_Z(z) = \int_0^z f_X(z-t) f_{2Y}(t)\; dt = \int_0^z \lambda e^{-\lambda(z-t)} \cdot \frac{\lambda}{2} e^{-\frac{\lambda}{2}t}\, dt = \frac{\lambda^2}{2} e^{-\lambda z} \int_0^z e^{\frac{\lambda}{2}t}\; dt =$$

$$= \frac{\lambda^2}{2} e^{-\lambda z} \frac{2}{\lambda} \left[e^{\frac{\lambda}{2}t}\right]_0^z \rightarrow$$

$$f_Z(z) = \lambda e^{-\lambda z}\left(e^{\frac{\lambda}{2}z} - 1\right)$$

Ερώτημα (δ)

Από το Ερώτημα (β):

$$P(X \leq Y + 1) = 1 - P(X > Y + 1) = 1 - \frac{e^{-\lambda}}{2}$$

Δεδομένου ότι αυτό συνέβη 400 φορές σε 1000 δοκιμές:

$$\hat{p} = \frac{400}{1000} = 0{,}4$$

Το θέτουμε ίσο με τη θεωρητική πιθανότητα:

$$0{,}4 = 1 - \frac{e^{-\lambda}}{2} \to \cdots \to \lambda = -0{,}1823$$