Θεωρία Πιθανοτήτων και Στατιστική
Θέματα Εξετάσεων Ιουνίου 2023 και ενδεικτικές λύσεις

Θέμα 1
Ένας εργαζόμενος, για να πάει στη δουλειά του, χρησιμοποιεί το αυτοκίνητό του το 30% των ημερών του έτους, πάει με τα πόδια το 30% των ημερών και παίρνει το λεωφορείο το 40% του ημερών του έτους. Η πιθανότητα να καθυστερήσει είναι 3% όταν πηγαίνει με το αυτοκίνητο, 10% όταν πηγαίνει με τα πόδια και 7% όταν παίρνει το λεωφορείο.

(α) Ποια είναι η πιθανότητα ο εργαζόμενος να φτάσει αργοπορημένος στην εργασία του;

(β) Στις 365 ημέρες του χρόνου, πόσες ημέρες αναμένεται να καθυστερήσει ο εργαζόμενος;

(γ) Ποια είναι η πιθανότητα να πήρε το λεωφορείο αν έχει καθυστερήσει;

(δ) Ποια είναι η πιθανότητα να πήρε το λεωφορείο αν δεν έχει καθυστερήσει;

Βαθμολογία: (α) 0,8μ (β) 0,2μ (γ) 0,8μ (δ) 0,7μ

Ενδεικτική λύση

Α = {πήγε με το αυτοκίνητο}, Π = {πήγε με τα πόδια}, Λ = {πήγε με το λεωφορείο}
Κ = {καθυστέρησε στην άφιξή του}

Από τα δεδομένα έχουμε:

$P(A) = 0,3$, $P(Π) = 0,3$, $P(Λ) = 0,4$

$P(K \mid A) = 0,03$, $P(K \mid Π) = 0,1$, $P(K \mid Λ) = 0,07$.

(α) $P(K) = P(K \mid A)\, P(A) + P(K \mid Π)\, P(Π) + P(K \mid Λ)\, P(Λ)$
$= 0,03 \cdot 0,3 + 0,1 \cdot 0,3 + 0,07 \cdot 0,4$
$= 0,067 = 6,7\%$.

(β) Αναμενόμενο πλήθος: $365 \cdot 0,067 = 24,455$ ημέρες.

(γ) Κανόνας Bayes: $P(Λ \mid K) = P(K \mid Λ) \cdot P(Λ) / P(K)$
$= 0,07 \cdot 0,4 / 0,067$
$= 0,418$

(δ) Κανόνας Bayes: $P(Λ \mid K') = P(K' \mid Λ) \cdot P(Λ) / P(K')$
$= [1 - P(K \mid Λ)] \cdot P(Λ) / [1 – P(K)]$
$= (1 – 0,07) \cdot 0,4 / (1 – 0,067)$
$= 0,399$.

Θέμα 2
Τρεις λαμπτήρες λειτουργούν ανεξάρτητα ο ένας από τον άλλον έχοντας διάρκεια ζωής που ακολουθεί κανονική κατανομή με μέση διάρκεια 100 ώρες και τυπική απόκλιση 10 ώρες.
(α) Να δείξετε ότι η πιθανότητα ένας λαμπτήρας να συνεχίσει να λειτουργεί μετά από 110 ώρες λειτουργίας είναι 0,159.

(β) Να βρεθεί η πιθανότητα τουλάχιστον 2 από τους 3 λαμπτήρες να συνεχίζουν να λειτουργούν μετά από 110 ώρες λειτουργίας.

Βαθμολογία: (α) 1μ (β) 1μ

Ενδεικτική Λύση

Αν Χ ο χρόνος λειτουργίας ενός λαμπτήρα, τότε $X \sim N(100, 10^2)$.

(α)
$P(X > 110) = P(Z > (110 – 100)/10) = P(Z > 1) = 1 – \Phi(1) = 1 – 0,8413 = 0,1587 \approx 0,159$.

(β) Αν Y = {πλήθος λαμπτήρων από τους 3 που συνεχίζουν να λειτουργούν μετά από 110 ώρες λειτουργίας}, τότε $Y \sim B(3, 0.159)$ και

P({τουλάχιστον 2 από τους 3 να λειτουργούν μετά από 110 ώρες}) = $P(Y \geq 2)$
$= P(Y = 2) + P(Y = 3)$
= (3 ανά 2) $0.159^2(1 – 0.159)$ + (3 ανά 3) 0.159^3
$= 3\ 0.159^2\ 0,841 + 0.159^3$
= 0,068.

Θέμα 3

(α) Ένας σάκος περιέχει μπάλες από 4 διαφορετικά χρώματα σύμφωνα με την κατανομή

x	Κόκκινο	Κίτρινο	Πράσινο	Μπλε
$f_X(x)$	2θ/3	θ/3	2(1 – θ)/3	(1 – θ)/3

Για να εκτιμήσουμε την άγνωστη παράμετρο θ, υλοποιούμε δειγματοληψία 10 ανεξάρτητων τιμών και παίρνουμε το δείγμα

(Μπλε, Κόκκινο, Πράσινο, Κίτρινο, Μπλε, Πράσινο, Κίτρινο, Κόκκινο, Πράσινο, Κίτρινο).

($α_1$) Να δείξετε ότι η συνάρτηση πιθανοφάνειας είναι η $L(\theta) = \left(\frac{2}{9}\right)^5 \theta^5 (1-\theta)^5$.

($α_2$) Να δείξετε ότι ο εκτιμητής μέγιστης πιθανοφάνειας για την παράμετρο θ έχει τιμή 0,5.

(β) (Για θ = 0,5): Σε μία αποθήκη υπάρχουν 100 σακιά όπως αυτά του ερωτήματος (α). Επιλέγουμε μία μπάλα από κάθε ένα. Να βρεθεί η πιθανότητα να επιλέξουμε περισσότερες από 30 κόκκινες.

Βαθμολογία: ($α_1$) 0,8μ ($α_2$) 0,7μ (β) 1μ

Ενδεικτική Λύση

($α_1$) Η ανεξαρτησία των παρατηρήσεων μας επιτρέπει να γράψουμε

$$L(\theta) = f(3)f(0)f(2)f(1)f(3)f(2)f(1)f(0)f(2)f(1) = f(0)^2 f(1)^3 f(2)^3 f(3)^2 = \left(\frac{2}{9}\right)^5 \theta^5 (1-\theta)^5.$$

($α_2$) Αν $l(\theta) = \ln L(\theta) = 5\ln\left(\frac{2}{9}\right) + 5\ln\theta + 5\ln(1-\theta)$, υπολογίζουμε $l'(\theta) = \frac{5}{\theta} + \frac{5}{1-\theta}$.

Η l'(θ) έχει ρίζα το ½ και μελετώντας το πρόσημο της παραγώγου βρίσκουμε ότι το ½ είναι σημείο μεγίστου για την l(θ). Άρα, ο εκτιμητής μέγιστης πιθανοφάνειας για την παράμετρο θ είναι 0,5.

(β) Για θ = 0,5, η κατανομή δίνεται από τον πίνακα:

x	Κόκκινο	Κίτρινο	Πράσινο	Μπλε
$f_X(x)$	1/3	1/6	1/3	1/6

Αν X_i = {επιλέγω κόκκινη μπάλα στο i σακί}, τότε X_i ~ Bernoulli(1/3) με $E(X_i)$ = 1/3, $Var(X_i)$ = 1/3·2/3 = 2/9. Το πλήθος των κόκκινων μπαλών στις 100 δίνεται από τη τ.μ.

$Y = X_1 + X_2 + \ldots X_{100}$ και από το ΚΟΘ είναι Y ~ N(100/3, 200/9). Υπολογίζουμε

$$P(Y > 30) = P\left(Z > \frac{30 - \frac{100}{3}}{\sqrt{\frac{200}{9}}}\right) = P(Z > -0{,}707) = \Phi(0{,}707) = \frac{0{,}758 + 0{,}7611}{2} = 0{,}760 = 76\%.$$

Εναλλακτικά για το (β)
Ορίζουμε X = {πλήθος κόκκινων μπαλών στις 100 επιλογές}. Τότε $X \sim B(100, 1/3)$.

Είναι $n = 100$, $p = 1/3$, $q = 2/3$, $np = 33,3 > 5$, $npq = 22,2 > 5$, άρα $B(n, p) \approx N(np, npq)$.

Συμπεραίνουμε ότι $B(100, 1/3) \approx N(100/3, 200/9)$ και αναζητούμε την πιθανότητα
$P(X > 30)$. Οι υπόλοιποι υπολογισμοί ακολουθούν με τον ίδιο τρόπο.

Θέμα 4
Για τα δεδομένα του πίνακα:

x	3	2	1
y	3	5	10

(α) Να βρεθεί η συνδιακύμανση των x, y.
(β) Να βρεθεί ο συντελεστής συσχέτισης των x, y.
(γ) Να βρεθεί γραμμικό μοντέλο πρόβλεψης της y από την x.
(δ) Να γίνει διάγραμμα διασποράς των τιμών στο οποίο να εμφανιστεί και η ευθεία γραμμικής παλινδρόμησης.
(ε) Να βρεθούν τα αθροίσματα τετραγώνων TSS, ESS, RSS.
(ζ) Να βρεθεί ο συντελεστής προσδιορισμού R^2.

Βαθμολογία: (α) 0,5μ (β) 0,5μ (γ) 0,5μ (δ) 0,5μ (ε) 0,8μ (ζ) 0,2μ

Ενδεικτική λύση

(α) Είναι $\bar{x} = 2,\ \bar{y} = 6,$ και $s_{XY}^2 = \dfrac{1}{3-1}\sum_{i=1}^{3}(x_i - 2)(y_i - 6) = -3,5.$

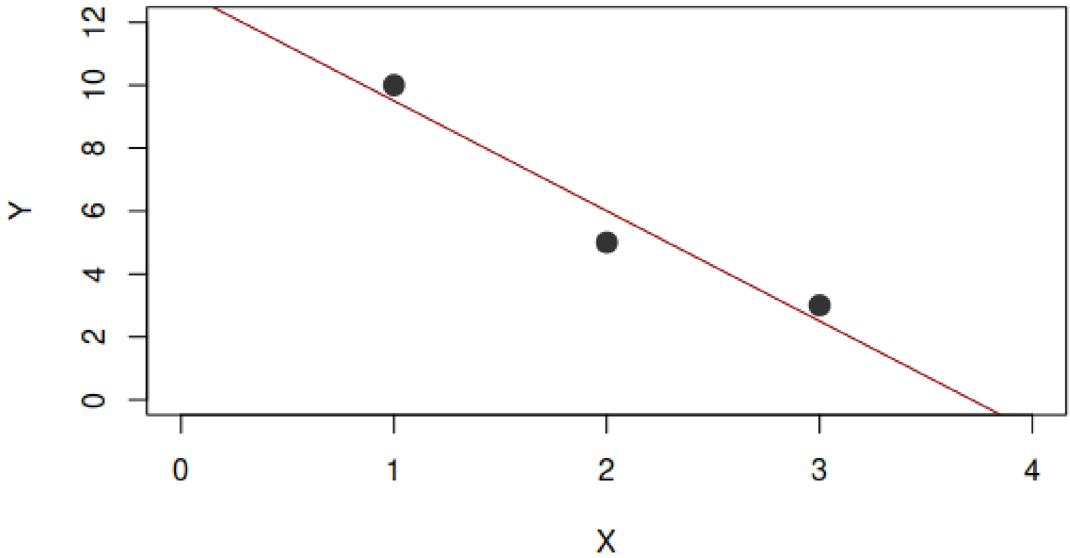
(β) $s_X^2 = \dfrac{1}{3-1}\sum_{i=1}^{3}(x_i - 2)^2 = 1,\ s_Y^2 = \dfrac{1}{3-1}\sum_{i=1}^{3}(y_i - 6)^2 = 13$ και

$r_{XY} = \dfrac{s_{XY}^2}{s_X \cdot s_Y} = -0,971.$

(γ) $\hat{b} = \dfrac{s_{XY}^2}{s_X^2} = r_{XY}\dfrac{s_Y}{s_X} = -3,5$ και $\hat{\alpha} = \bar{y} - \hat{b}\,\bar{x} = 13.$ Η εξίσωση της ευθείας γραμμικής παλινδρόμησης είναι η

$$Y = -3,5 \cdot X + 13$$

(δ) Διάγραμμα διασποράς



(ε)

X	3	2	1
Y	3	5	10
$\hat{Y}$	2,5	6	9,5
$Y-\hat{Y}$	0,5	-1	0,5

$$TSS = \sum_{i=1}^{n}(y_i - \bar{y})^2 = 26,\quad ESS = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 = 24{,}5,\quad RSS = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = 1{,}5$$

(ζ) $R^2 = \frac{ESS}{TSS} = \frac{24{,}5}{26} = 0{,}942.$