**Θεωρία Πιθανοτήτων και Στατιστική**
**Θέματα Εξετάσεων Σεπτεμβρίου 2023 και ενδεικτικές λύσεις**

**Θέμα 1**
Γνωρίζουμε ότι 5% των ασθενών σε μία κλινική παρουσιάζουν λοίμωξη κάποιου είδους. Επίσης είναι γνωστό ότι το 40% των ασθενών που έχουν λοίμωξη έχουν μειωμένα λευκά αιμοσφαίρια ενώ το 1% των ασθενών που δεν έχουν λοίμωξη έχουν μειωμένα λευκά αιμοσφαίρια. Να υπολογισθούν οι πιθανότητες ένας ασθενής στην κλινική:
(α) να έχει μειωμένα λευκά αιμοσφαίρια.
(β) να έχει λοίμωξη δεδομένου ότι έχει μειωμένα λευκά αιμοσφαίρια.
(γ) να έχει λοίμωξη δεδομένου ότι δεν έχει μειωμένα λευκά αιμοσφαίρια.

**Βαθμολογία: (α) 0,8μ (β) 0,9μ (γ) 0,8μ**

**Ενδεικτική λύση**

A = {ο ασθενής έχει λοίμωξη}
B = {ο ασθενής έχει μειωμένα λευκά αιμοσφαίρια}

Από τα δεδομένα έχουμε: $P(A) = 0{,}05$, $P(B \mid A) = 0{,}4$, $P(B \mid A') = 0{,}01$.

(α) $P(B) = P(B \mid A)\, P(A) + P(B \mid A')\, P(A') = 0{,}4 \cdot 0{,}05 + 0{,}01 \cdot 0{,}95 = 0{,}0295 = 2{,}95\%$

(β) $P(A \mid B) = P(B \mid A) \cdot P(A) / P(B) = 0{,}4 \cdot 0{,}05 / 0{,}0295 = 0{,}678 = 67{,}8\%$.

(γ) $P(A \mid B') = P(B' \mid A) \cdot P(A) / P(B') = [1 – 0{,}4] \cdot 0{,}05 / [1 – 0{,}0295] = 0{,}031 = 3{,}1\%$.

**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 X 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**

Ένα νέο φάρμακο για την αντιμετώπιση της υπέρτασης δόθηκε πειραματικά σε 200 άτομα που πάσχουν από υπέρταση. Το αποτέλεσμα της φαρμακευτικής αγωγής για κάθε ασθενή ταξινομήθηκε σε μια από τέσσερις κατηγορίες:

Α: Βαθμιαία μείωση Β: Μέτρια μείωση Γ: Μικρή μείωση Δ: Μικρή αύξηση.

Οι συχνότητες των τεσσάρων κατηγοριών στα 200 άτομα φαίνονται στον πίνακα που ακολουθεί.

| | Α | Β | Γ | Δ |
|---|---|---|---|---|
| Συχνότητα | 120 | 50 | 20 | 10 |

Από σχετικές μελέτες είναι γνωστό ότι ένα αντίστοιχο φάρμακο που ήδη κυκλοφορεί και χρησιμοποιείται, έχει την εξής αποτελεσματικότητα: Α:50%, Β:30%, Γ:19% και Δ:1%.

Θέλουμε να ελέγξουμε αν το νέο φάρμακο διαφέρει ως προς την αποτελεσματικότητά του από το φάρμακο που ήδη κυκλοφορεί.

(α) Να αναφέρετε τη στατιστική δοκιμασία που θα εφαρμόσουμε για να ελέγξουμε την παραπάνω υπόθεση.
(β) Να καταγράψετε την μηδενική και την εναλλακτική υπόθεση.
(γ) Να βρείτε αν το νέο φάρμακο διαφέρει ως προς την αποτελεσματικότητά του από το φάρμακο που ήδη κυκλοφορεί. ($\alpha = 0{,}05$).

**Βαθμολογία: ($\alpha_1$) 0,2μ ($\alpha_2$) 0,5μ (β) 1,8μ**

**Ενδεικτική Λύση**

(α) Είναι η δοκιμασία ομοιογένειας (ή προσαρμογής) $\chi^2$.

(β) Μηδενική $H_0$: Α:50%, Β:30%, Γ:19% και Δ:1% έναντι της Εναλλακτικής $H_1$: όχι η $H_0$.

(γ) Συμπληρώνουμε τον παρακάτω πίνακα:

| | Α | Β | Γ | Δ | Σύνολο |
|---|---|---|---|---|---|
| Παρατηρούμενη συχνότητα ($O_i$) | 120 | 50 | 20 | 10 | 200 |
| Θεωρητικό ποσοστό σύμφωνα με την $H_0$ | 50% | 30% | 19% | 1% | 100% |
| Αναμενόμενη συχνότητα $E_i$ σύμφωνα με την $H_0$ | 100 | 60 | 38 | 2 | 200 |
| $O_i - E_i$ | 20 | -10 | -18 | 8 | 0 |
| $(O_i – E_i)^2 / E_i$ | 4 | 1,67 | 8,52 | 32 | 46,2 |

Είναι $\chi^2_0 = \Sigma_{i=1,2,3,4}(O_i – E_i)^2 / E_i = 46{,}2$ και $\chi^2 \sim \chi^2(4 – 1) = \chi^2(3)$.

Για $\chi^2 \sim \chi^2(3)$, είναι $p = P(\chi^2 > \chi^2_0) = P(\chi^2 > 46{,}2) < 0{,}001$, άρα η υπόθεση $H_0$ απορρίπτεται, δηλαδή η αποτελεσματικότητα του φαρμάκου είναι σημαντικά διαφορετική από αυτή που δηλώνει η $H_0$.

**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 $s_{XY}^2 = \dfrac{1}{3-1}\sum_{i=1}^{3}(x_i - 2)(y_i - 6) = -3{,}5.$

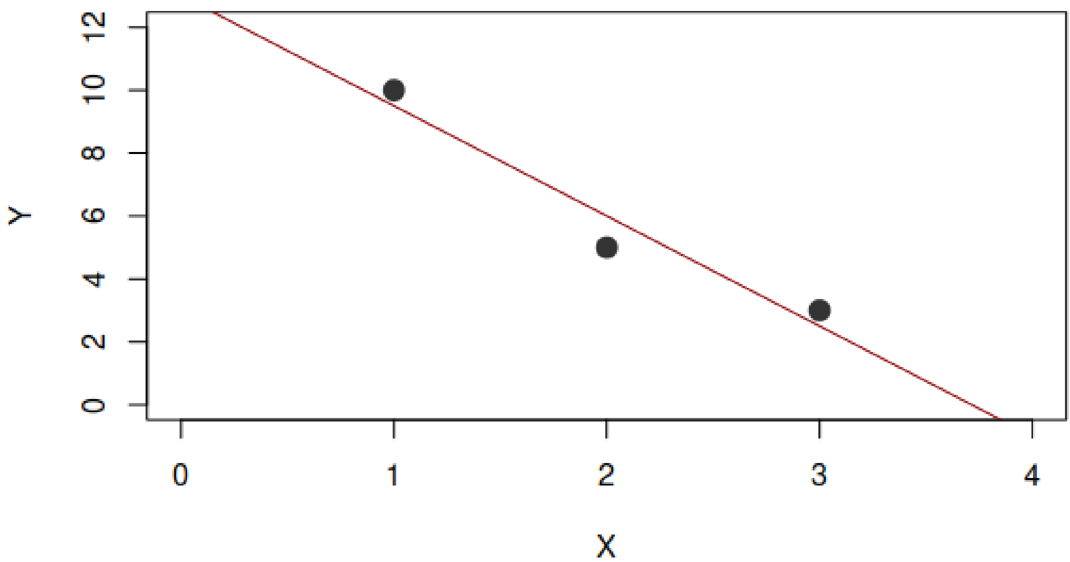
(β) $s_X^2 = \dfrac{1}{3-1}\sum_{i=1}^{3}(x_i - 2)^2 = 1,\ s_Y^2 = \dfrac{1}{3-1}\sum_{i=1}^{3}(y_i - 6)^2 = 13$ και

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

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