

Πανεπιστήμιο Κρήτης - Τμήμα Επιστήμης Υπολογιστών
ΗΥ-217 - Θεωρία Πιθανοτήτων
Διδάσκων: Π. Τσακαλίδης
Λύσεις Προόδου- 14 Νοεμβρίου 2015

Θέμα 1 - (10 μονάδες)

(α') Προφανώς, $P(R) = P(B) = P(G) = \frac{2}{4} = \frac{1}{2}$, καθώς 2 στις 4 πλευρές του τετράεδρου περιέχουν κόκκινο, μπλε ή πράσινο χρώμα.

(β') Σαφώς, $P(R \cap G) = P(G \cap B) = P(B \cap R) = \frac{1}{4}$, καθώς μόνο η τέταρτη πλευρά περιέχει ταυτόχρονα 2 χρώματα. Καθώς,

$$\begin{aligned}P(R \cap G) &= \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(R) \cdot P(G) \\P(G \cap B) &= \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(G) \cdot P(B) \\P(B \cap R) &= \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(B) \cdot P(R)\end{aligned}$$

τα γεγονότα R, B, G είναι ανεξάρτητα ανά δύο.

(γ') $P(R \cap B \cap G) = \frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = P(R) \cdot P(G) \cdot P(B)$.
Συνεπώς, τα γεγονότα R, B, G δεν είναι ανεξάρτητα.

Θέμα 2 - (20 μονάδες)

Ορίζουμε τα γεγονότα:

$C = \{\text{Μια γυναίκα ηλικίας 30-39 έχει καρκίνο του μαστού}\}$

$M = \{\text{Η εξέταση μαστογραφίας είναι θετική}\}$

Σύμφωνα με την εκφώνηση,

$$\begin{aligned}P(C) &= 0.008, & P(M/C) &= 0.9 \\P(M/C^c) &= 0.07\end{aligned}$$

(α') Ζητάμε την πιθανότητα του γεγονότος M . Εφαρμόζουμε το Θεώρημα Ολικής Πιθανότητας:

$$\begin{aligned}P(M) &= P(C) \cdot P(M/C) + P(C^c) \cdot P(M/C^c) \\&= 0.008 \times 0.9 + 0.992 \times 0.07 = 0.0766 \simeq 7.7\%\end{aligned}$$

(β') Εφαρμόζοντας τον κανόνα του Bayes,

$$P(C/M) = \frac{P(C) \cdot P(M/C)}{P(M)} = \frac{0.008 \times 0.9}{0.0766} = 0.094 = 9.4\%$$

(γ') (i) Από τον πληθυσμό των 1000 γυναικών περιμένουμε ότι οι $1000 \times P(C) = 1000 \times 0.008 = 8$ έχουν καρκίνο.

(ii) Η πιθανότητα σωστής ανίχνευσης είναι

$$P_d = P(C \cap M) = P(C) \cdot P(M/C) = 0.008 \times 0.9 = 0.0072$$

Επομένως, από το συνολικό πληθυσμό των 1000 γυναικών, οι $1000 \times P_d = 1000 \times 0.0072 = 7.2$ διαγιγνώσκονται σωστά ότι έχουν καρκίνο του μαστού.

(iii) Η πιθανότητα λάθος συναγερμού είναι:

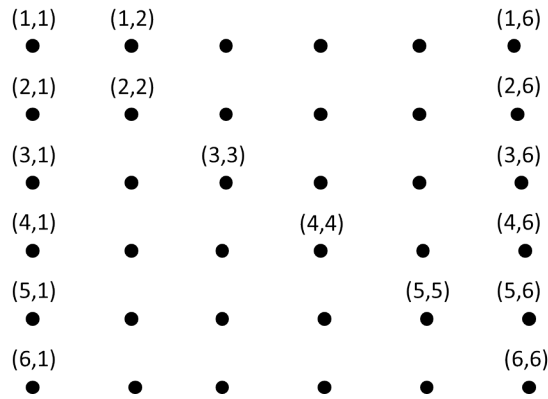
$$P_{FA} = P(C^c \cap M) = P(C^c) \cdot P(M/C^c) = 0.992 \times 0.07 = 0.0694$$

Συνεπώς, από το συνολικό πληθυσμό των 1000 γυναικών, οι $1000 \times P_{FA} = 1000 \times 0.0694 = 69.4$ λάθος διαγιγνώσκονται να έχουν καρκίνο ενώ δεν έχουν. Οι γυναίκες αυτές θα πρέπει να κάνουν επιπλέον εξετάσεις (με κόστος και κόπο) πριν διαπιστωθεί ότι τελικά δε έχουν καρκίνο.

Θέμα 3 - (25 μονάδες)

(α') $\Omega = \{(i, j) : i = 1, 2, \dots, 6 \text{ και } j = 1, 2, \dots, 6\}, \quad P(i, j) = \frac{1}{36}$

Γραφικά, στο Σχήμα 1 παρακάτω:



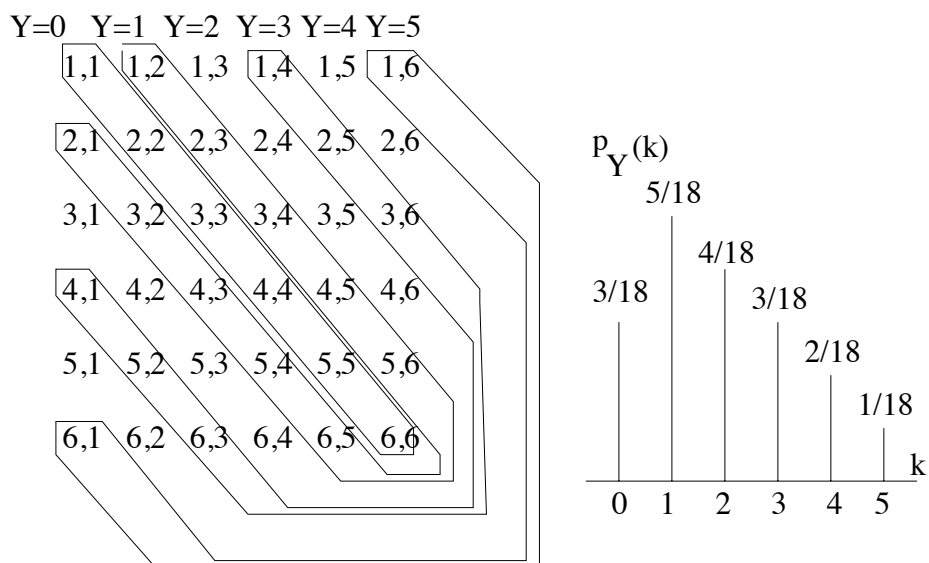
Σχήμα 1: Γραφική αναπαράσταση του δειγματικού χώρου Ω

- (β') • $\{Y = 0\} = \{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\}$
 $\therefore p_Y(0) = P(Y = 0) = \frac{6}{36} = \frac{1}{6}$
- $\{Y = 1\} = \{(1, 2), (2, 3), (3, 4), (4, 5), (5, 6), (6, 5), (5, 4), (4, 3), (3, 2), (2, 1)\}$
 $\therefore p_Y(1) = P(Y = 1) = \frac{10}{36} = \frac{5}{18}$
- $\{Y = 2\} = \{(1, 3), (2, 4), (3, 5), (4, 6), (6, 4), (5, 3), (4, 2), (3, 1)\}$
 $\therefore p_Y(2) = P(Y = 2) = \frac{8}{36} = \frac{2}{9}$
- $\{Y = 3\} = \{(1, 4), (2, 5), (3, 6), (6, 3), (5, 2), (4, 1)\}$
 $\therefore p_Y(3) = P(Y = 3) = \frac{6}{36} = \frac{1}{6}$
- $\{Y = 4\} = \{(1, 5), (2, 6), (6, 2), (5, 1)\}$
 $\therefore p_Y(4) = P(Y = 4) = \frac{4}{36} = \frac{1}{9}$
- $\{Y = 5\} = \{(1, 6), (6, 1)\}$
 $\therefore p_Y(5) = P(Y = 5) = \frac{2}{36} = \frac{1}{18}$

Συνεπώς,

$$p_Y(k) = \begin{cases} \frac{1}{6}, & k = 0 \\ \frac{5}{18}, & k = 1 \\ \frac{2}{9}, & k = 2 \\ \frac{1}{6}, & k = 3 \\ \frac{1}{9}, & k = 4 \\ \frac{1}{18}, & k = 5 \end{cases}$$

Και η σχετική γραφική παράσταση στο Σχήμα 2 παρακάτω:



Σχήμα 2: Η ανάλυση της τ.μ. Y και η γραφική παράσταση της Σ.Π. της.

(γ') Έχουμε:

$$\begin{aligned}
 E[Y] &= \sum_{k=0}^5 k p_Y(k) = 0 \times \frac{3}{18} + 1 \times \frac{5}{18} + 2 \times \frac{4}{18} + 3 \times \frac{3}{18} + 4 \times \frac{2}{18} + 5 \times \frac{1}{18} = \frac{35}{18} \\
 E[Y^2] &= \sum_{k=0}^5 k^2 p_Y(k) = 0^2 \times \frac{3}{18} + 1^2 \times \frac{5}{18} + 2^2 \times \frac{4}{18} + 3^2 \times \frac{3}{18} + 4^2 \times \frac{2}{18} + 5^2 \times \frac{1}{18} \\
 &= \frac{105}{18} = 5.833 \\
 \therefore \quad \text{var}(Y) &= E[Y^2] - (E[Y])^2 = \frac{105}{18} - \left(\frac{35}{18}\right)^2 = 2.052
 \end{aligned}$$

(δ') (i) Για $Y = \{0, 1, 2, 3, 4, 5\}$ έχουμε τις εξής τιμές της $Z = \{-20, -10, 0, 10, 20, 30\}$.

Άρα:

$$p_Z(k) = \begin{cases} \frac{3}{18}, & k = -20 \\ \frac{5}{18}, & k = -10 \\ \frac{4}{18}, & k = 0 \\ \frac{3}{18}, & k = 10 \\ \frac{2}{18}, & k = 20 \\ \frac{1}{18}, & k = 30 \end{cases}$$

(ii) $P(Z > 0) = p_Z(10) + p_Z(20) + p_Z(30) = \frac{3}{18} + \frac{2}{18} + \frac{1}{18} = \frac{6}{18} = 0.333$

(iii) Ισχύει:

$$\begin{aligned}
 E[Z] &= E[10Y - 20] = 10E[Y] - 20 = 10 \times \frac{35}{18} - 20 = -0.556 \\
 \text{var}(Z) &= \text{var}(10Y - 20) = 100 \text{var}(Y) = 100 \times 2.052 = 205.2
 \end{aligned}$$

Φαίνεται ότι μάλλον δε μας συμφέρει και πολύ αυτό το παιχνίδι.

Θέμα 4 - (20 μονάδες)

(α') (i) Προφανώς,

$$p_X(k) = \begin{cases} \frac{1}{50}, & k = 10000 \\ \frac{49}{50}, & k = 0 \end{cases}$$

(ii) Έχουμε:

$$\begin{aligned} E[X] &= 0 \times \frac{49}{50} + 10000 \times \frac{1}{50} = 200\$ \\ E[X^2] &= 0^2 \times \frac{49}{50} + 10000^2 \times \frac{1}{50} = 2 \times 10^6 \\ \sigma_X &= \sqrt{\text{var}(X)} = \sqrt{E[X^2] - \mu_X^2} = \sqrt{2 \times 10^6 - 200^2} = 1400\$ \end{aligned}$$

(β') (i) Προφανώς, $Y = 100(X_1 + X_2 + \dots + X_{100}) = 100Z$ όπου $X_i \sim \text{Bernoulli}(p = \frac{1}{50})$. Καθώς οι X_i είναι ανεξάρτητες τ.μ., έχουμε ότι $Z \sim \Delta(n = 100, p = \frac{1}{50})$. Επομένως, το κέρδος $Y = 100Z$ παίρνει τιμές $0, 100, 200, \dots, 9.900, 10.000$ με πιθανότητες που αντιστοιχούν σε αυτές της διωνυμικής κατανομής με παραμέτρους $n = 100$ και $p = 1/50$:

$$p_Y(k) = \binom{100}{k/100} \left(\frac{1}{50}\right)^{k/100} \left(\frac{49}{50}\right)^{100-k/100}; \quad k = 0, 100, 200, 300, \dots, 10.000$$

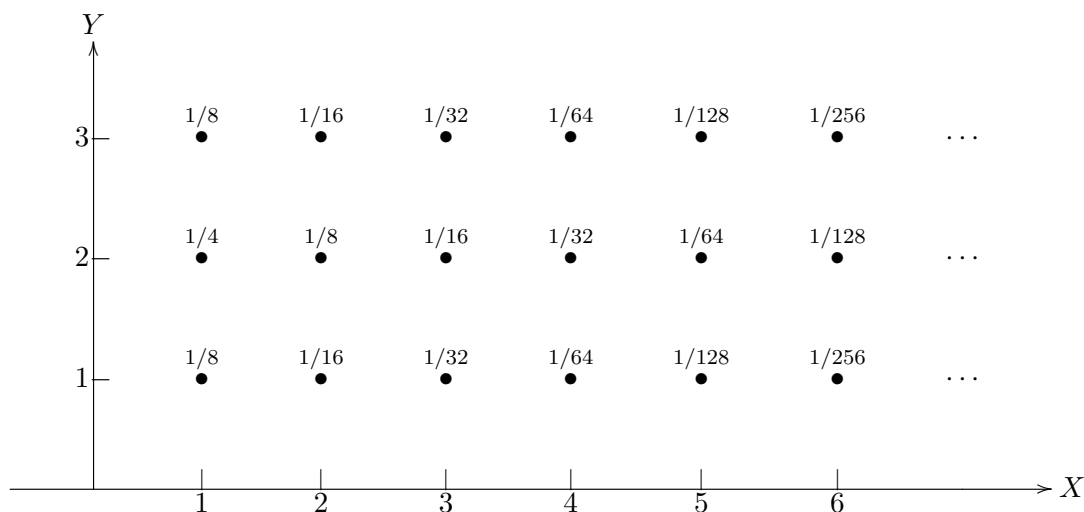
(ii) Ισχύει:

$$\begin{aligned} E[Y] &= E[100Z] = 100E[Z] = 100 \cdot n \cdot p = 100 \times 100 \times \frac{1}{50} = 200\$ \\ \sigma_Y &= \sqrt{\text{var}(100Z)} = \sqrt{100^2 \cdot \text{var}(Z)} = \sqrt{100^2 \cdot n \cdot p \cdot (1-p)} \\ &= \sqrt{100^2 \times 100 \times \frac{1}{50} \times \frac{49}{50}} = 140 \end{aligned}$$

Βλέπουμε ότι με τη δεύτερη στρατηγική, το μέσο κέρδος παραμένει το ίδιο ($= 200\$$), αλλά η τυπική απόκλιση είναι 10 φορές πιο μικρή!

Θέμα 5 - (25 μονάδες)

(α')



Είναι:

- $0 < p_{X,Y}(x, y) < 1, \quad \forall (x, y)$
- και

$$\begin{aligned} \sum_{x=1}^{+\infty} \sum_{y=1}^3 p_{X,Y}(x, y) &= \sum_{x=1}^{+\infty} \frac{1}{4} \left(\frac{1}{2}\right)^x + \sum_{x=1}^{+\infty} \frac{1}{2} \left(\frac{1}{2}\right)^x + \sum_{x=1}^{+\infty} \frac{1}{4} \left(\frac{1}{2}\right)^x \\ &= \frac{1}{4} \frac{1/2}{1-1/2} + \frac{1}{2} \frac{1/2}{1-1/2} + \frac{1}{4} \frac{1/2}{1-1/2} \\ &= \frac{1}{4} \cdot 1 + \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 1 = 1 \end{aligned}$$

Συνεπώς, η $p_{X,Y}(x, y)$ είναι μια έγκυρη συνάρτηση πιθανότητας.

(β') Ισχύουν τα εξής:

$$p_X(x) = \sum_{y=1}^3 p_{X,Y}(x, y) = \frac{1}{4} \left(\frac{1}{2}\right)^x + \frac{1}{2} \left(\frac{1}{2}\right)^x + \frac{1}{4} \left(\frac{1}{2}\right)^x = \left(\frac{1}{2}\right)^x, \quad x = 1, 2, 3, \dots$$

$$p_Y(y) = \sum_{x=1}^{+\infty} p_{X,Y}(x, y) = \begin{cases} \sum_{x=1}^{+\infty} \frac{1}{4} \left(\frac{1}{2}\right)^x, & y = 1, 3 \\ \sum_{x=1}^{+\infty} \frac{1}{2} \left(\frac{1}{2}\right)^x, & y = 2 \end{cases}$$

$$\Rightarrow p_Y(y) = \begin{cases} \frac{1}{4} \frac{1/2}{1-1/2}, & y = 1, 3 \\ \frac{1}{2} \frac{1/2}{1-1/2}, & y = 2 \end{cases} = \begin{cases} 1/4, & y = 1, 3 \\ 1/2, & y = 2 \end{cases}$$

Βλέπουμε ότι: $p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y)$ και συνεπώς οι τ.μ. X, Y είναι ανεξάρτητες.

(γ') Η γεωμετρική κατανομή έχει συνάρτηση πιθανότητας :

$$p_X(k) = (1-p)^{k-1} \cdot p, \quad k = 1, 2, 3, \dots$$

Άρα, είναι προφανές ότι η τ.μ. X του προβλήματος είναι γεωμετρική με παράμετρο $p = 1/2$.

Επομένως, $E[X] = \frac{1}{p} = 2$, $var(X) = \frac{1-p}{p^2} = \frac{1-1/2}{1/4} = 2$.

(δ') Χρειάζεται πρώτα να υπολογίσουμε τις $E[Y]$, $E[Y^2]$, $E[Y^4]$. Έχουμε:

$$\begin{aligned} E[Y] &= 1 \cdot \frac{1}{4} + 3 \cdot \frac{1}{4} + 2 \cdot \frac{1}{2} = 2 \\ E[Y^2] &= 1^2 \cdot \frac{1}{4} + 3^2 \cdot \frac{1}{4} + 2^2 \cdot \frac{1}{2} = \frac{9}{2} \\ E[Y^4] &= 1^4 \cdot \frac{1}{4} + 3^4 \cdot \frac{1}{4} + 2^4 \cdot \frac{1}{2} = \frac{57}{2} \end{aligned}$$

Εν τέλει θα έχουμε:

$$E[Z] = E[4X - Y^2] = 4E[X] - E[Y^2] = 4 \times 2 - \frac{9}{2} = \frac{7}{2}$$

και

$$\begin{aligned} \text{var}(Z) &= \text{var}(4X - Y^2) = \text{var}(4X) + \text{var}(Y^2) \quad (\text{X,Y ανεξάρτητες}) \\ &= 16 \cdot \text{var}(X) + \text{var}(Y^2) \\ &= 16 \cdot \text{var}(X) + E[Y^4] - E^2[Y^2] \\ &= 16 \cdot 2 + \frac{57}{2} - \frac{81}{4} = \frac{161}{4} \end{aligned}$$

(ε') Έχουμε:

$$\begin{aligned} p_{Z|X}(z|X=1) &= P(\{Z=z\}|\{X=1\}) = P(\{4X - Y^2 = z\}|\{X=1\}) \\ &= P(\{4 - Y^2 = z\}) = \begin{cases} 1/4, & z = 3, -5 \\ 1/2, & z = 0 \end{cases} \end{aligned}$$

αφού η τ.μ. Y παίρνει τις τιμές 1, 3, 2 με πιθανότητες $\frac{1}{4}, \frac{1}{4}, \frac{1}{2}$ αντίστοιχα.