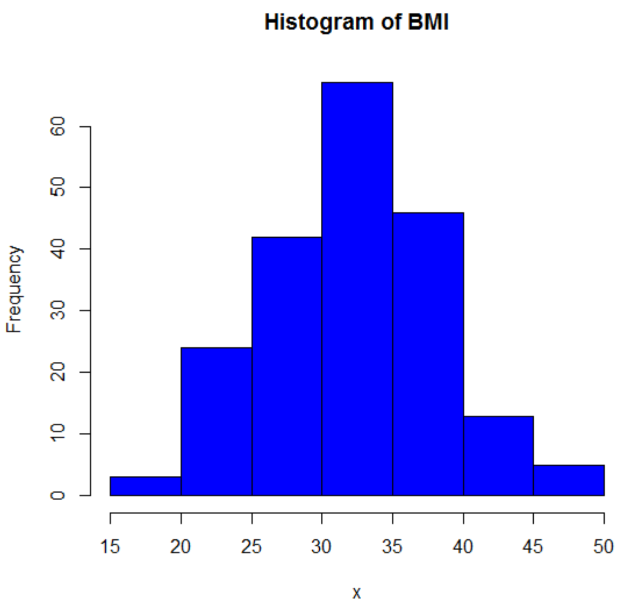
Ασκήσεις σε Περιγραφική Στατιστική

Άσκηση 2η: Υπολογίσαμε τον δείκτη μάζας σώματος (BMI) (βάρος σε Kgr /(ύψος σε m)^2) για 200 εγκύους στον τελευταίο μήνα της κυήσεως. Τα δεδομένα αυτά δίδονται από τον παρακάτω πίνακα:

BMI σε kgr/m^2	15-20	20-25	25-30	30-35	35-40	40-45	45-50
Πλήθος εγκύων	3	24	42	67	46	13	5



$L_i - U_i$	x_i	f_i	F_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
15-20	17.5	3	3	52.5	918.75
20-25	22.5	24	27	540.0	12150.00
25-30	27.5	42	69	1155.0	31762.50
30-35	32.5	67	136	2177.5	70768.75
35-40	37.5	46	182	1725.0	64687.50
40-45	42.5	13	195	552.5	23481.25
45-50	47.5	5	200	237.5	11281.25
Σύνολο		200		6440	215050

β. Να βρεθούν η δειγματική μέση τιμή, η διασπορά, η τυπική απόκλιση, το πρώτο τεταρτημόριο, η διάμεσος, το 90% ποσοστιαίο σημείο και η κορυφή του δείκτη μάζας σώματος των εγκύων.

Έτσι ο δειγματικός μέσος, , για τον δείκτη μάζας σώματος (BMI)

$$\overline{x} = \frac{\sum_{i=1}^{\kappa} x_i \cdot f_i}{\sum_{i=1}^{\kappa} f_i} \overset{\text{εδώ}}{\underset{\kappa=6}{=}} \frac{17.5 \times 3 + 22.5 \times 24 + \cdots + 47.5 \times 5}{3 + 24 + \cdots + 5}$$

$$= \frac{6440}{200} = 32.2 \text{ (βάρος σε Kgr /(ύψος σε m)}^2\text{)}$$

Εάν συμβολίσουμε με $n = \sum_{i=1}^{7} f_i$ το μέγεθος του δείγματος,

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{\kappa}(x_i - \overline{x})^2 \cdot f_i \overset{\text{ή}}{=} \frac{1}{n-1}\left\{\sum_{i=1}^{\kappa} x_i^2 \cdot f_i - \frac{\left(\sum_{i=1}^{\kappa} x_i \cdot f_i\right)^2}{n}\right\}$$

$$\overset{\text{εδώ}}{=} \frac{1}{199}\left\{215050 - \frac{6440^2}{200}\right\} = \frac{1}{199}\{215050 - 207368\} = \frac{7682}{119} = 38.60 = (6.21)^2$$

τυπική απόκλιση είναι: S=6.21 βάρος σε Kgr /(ύψος σε m)^2

Για το πρώτο τεταρτημόριο, επειδή $\frac{n}{4} = 50$ και $F_2 = 27 < 50 < 69 = F_3$, αυτό βρίσκεται στην

3η κλάση $\quad Q_1 = L_3 + \left(\frac{\frac{n}{4} - F_2}{f_3}\right) c \stackrel{\text{εδώ}}{=\!=} 25 + \left(\frac{50-27}{42}\right) 5 = 27.7381 \simeq 27.7$ (βάρος σε Kg

για την διάμεσο, παρατηρούμε ότι επειδή $\frac{n}{2} = 100$ και $F_3 = 69 < 100 < 136 = F_4$,

βρίσκεται στην 4η κλάση

$\delta($ ή $Q_2) = L_4 + \left(\frac{\frac{n}{2} - F_3}{f_4}\right) c \stackrel{\text{εδώ}}{=\!=} 30 + \left(\frac{100-69}{67}\right) 5 = 32.31343 \simeq 32.3$ (βάρος σε Kgr /(ύψ

για το 90-στο εκατοστημόριο έχουμε $\frac{90\,n}{100} = 180$ και $F_4 = 136 < 180 < 182 = F_5$,

βρίσκεται στην 5η κλάση:

$p_{90} = L_5 + \left(\frac{\frac{90n}{100} - F_4}{f_5}\right) c = 35 + \left(\frac{180-136}{46}\right) 5 = 39.78261 \simeq 39.8$ (βάρος σε Kgr /(ύψος σ

Για την κορυφή, παρατηρούμε ότι $\max\limits_{j=1,\ldots,7} f_j = f_4 = f_i$ οπότε αυτή βρίσκεται στην $i = 4$η κλάσ

$M = L_4 + \frac{\Delta_1}{\Delta_1 + \Delta_2} c \stackrel{\text{εδώ}}{=\!=} 30 + \frac{67-42}{2 \cdot 67 - 42 - 46} \cdot 5 = 32.71739 \simeq 32.7$ (βάρος σε Kgr /(ύψος σ

όπου $\Delta_1 = f_i - f_{i-1}$ και $\Delta_2 = f_i - f_{i+1}$

γ. Εάν σε τρεις μήνες μετά την εγκυμοσύνη οι γυναίκες χάσουν το 10% του βάρους τους, πώς μεταβάλλονται η μέση τιμή, η τυπική απόκλιση και η κορυφή;

$$\text{BMI Αρχικό} = \frac{\text{Αρχικό Βάρος σε Kgr}}{\text{Ύψος σε m}^2}$$

$$\text{BMI Μετά} = \frac{(\text{Αρχικό Βάρος} - 0.1 \cdot \text{Αρχικό Βάρος})\text{σε Kgr}}{\text{Ύψος σε m}^2} = \frac{0.9 \cdot \text{Αρχικό Βάρος σε Kgr}}{\text{Ύψος σε m}^2}$$

$$= 0.9 \cdot \text{BMI Αρχικό}$$

μεταβάλλονται τα άκρα των κλάσεων, $L_i^* = 0.9\, L_i$ και $U_i^* = 0.9\, L_i$ και το εύρος τους, $c^* = 0.9\, c$

Παραμένουν ίδιες οι συχνότητες και οι αθροιστικές συχνότητες των κλάσεων

$$\overline{x^*} = \frac{\sum_{i=1}^{\kappa} x_i^* \cdot f_i}{n} = \frac{\sum_{i=1}^{\kappa} 0.9\, x_i \cdot f_i}{n} = 0.9\, \frac{\sum_{i=1}^{\kappa} x_i \cdot f_i}{n} \Rightarrow \overline{x^*} = 0.9\, \overline{x}$$

$$\overline{x^*} = 0.9 \cdot 32.2 = 28.98$$

$$S^* = \sqrt{\frac{1}{n-1}\sum_{i=1}^{\kappa}\left(x_i^* - \overline{x^*}\right)^2 \cdot f_i} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{\kappa}\left(0.9\,x_i - 0.9\,\overline{x}\right)^2 \cdot f_i}$$

$$= \sqrt{\frac{1}{n-1}\sum_{i=1}^{\kappa} 0.9^2\,\left(x_i - \overline{x}\right)^2 \cdot f_i} \quad = 0.9\sqrt{\frac{1}{n-1}\sum_{i=1}^{\kappa}\left(x_i - \overline{x}\right)^2 \cdot f_i} \Rightarrow S^* = 0.9\,S$$

$$S^* = 0.9 \cdot 6.21 = 5.59$$

Για οποιοδήποτε κ-στο εκατοστημόριο των νέων δεδομένων

$$p_\kappa^* = L_i^* + \left(\frac{\frac{\kappa n}{100} - F_{i-1}}{f_i}\right) c^* = 0.9\,L_i + \left(\frac{\frac{\kappa n}{100} - F_{i-1}}{f_i}\right) 0.9\,c$$

$$= 0.9\left\{L_i + \left(\frac{\frac{\kappa n}{100} - F_{i-1}}{f_i}\right) c\right\} \Rightarrow p_\kappa^* = 0.9\,p_\kappa$$

$$Q_1^* = 0.9\,Q_1 = 24.96, \quad Q_2^* = 0.9\,Q_2 = 29.08 \quad \text{και} \quad p_{90}^* = 0.9\,p_{90} = 0.9 \cdot 39.78 = 35.8$$

Ανάλογα για την κορυφή

$$M^* = L_4^* + \frac{\Delta_1}{\Delta_1+\Delta_2}c^* = 0.9L_4 + \frac{\Delta_1}{\Delta_1+\Delta_2}0.9c$$

$$= 0.9\left\{L_4 + \frac{\Delta_1}{\Delta_1+\Delta_2}c\right\} = 0.9 \cdot M = 0.9 \cdot 32.7 = 29.44$$

Παρατήρηση Για να γίνουν πιο εύκολα οι πράξεις

μετασχηματισμό $x_i^* = \frac{x_i - 32.5}{5} \Rightarrow x_i = 5x_i^* + 32.5$

$L_i^* - U_i^*$	x_i^*	f_i	F_i	$x_i^* \cdot f_i$	$x_i^{*2} \cdot f_i$
-3.5 - -2.5	-3	3	3	-9	27
-2.5 - -1.5	-2	24	27	-48	96
-1.5 - -0.5	-1	42	69	-42	42
-0.5 - 0.5	0	67	136	0	0
0.5 - 1.5	1	46	182	46	46
1.5 - 2.5	2	13	195	26	52
2.5 - 3.5	3	5	200	15	45
Σύνολο		200		-12	308

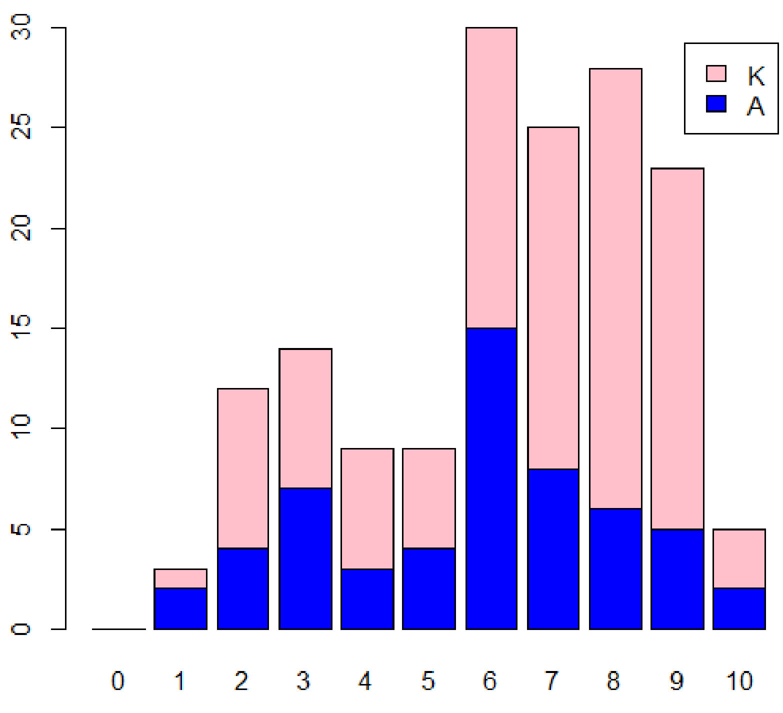
$$\overline{x^*} = -0.06, \quad S^{*2} = 1.544121, \quad S^* = 1.242626$$

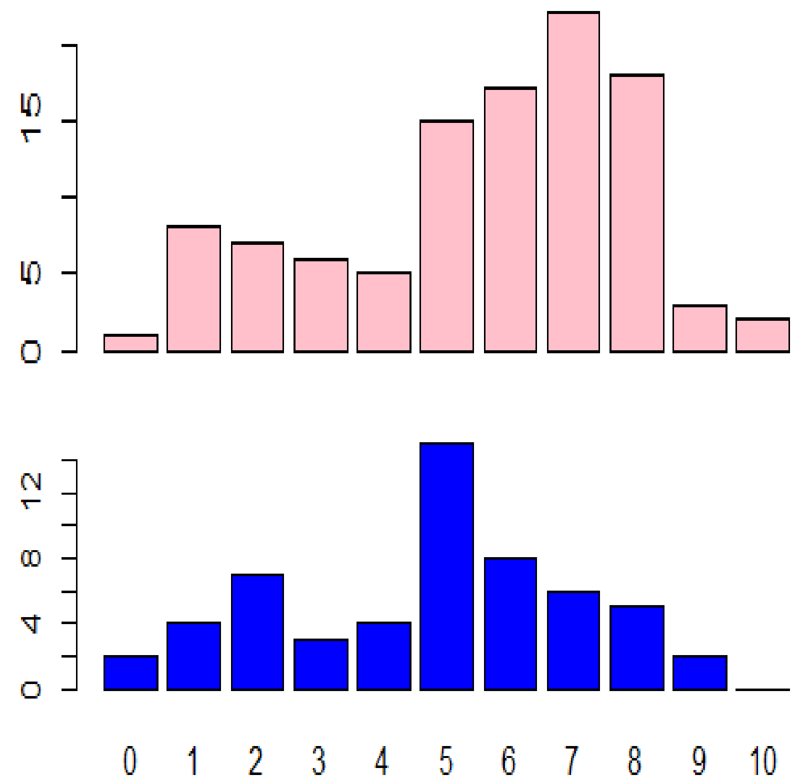
$\overline{x} = 5\overline{x^*} + 32.5 = 5(-0.06) + 32.5 = 32.3,$

$S^2 = 5^2 S^{*2} = 25 \cdot 1.544121 = 38.60, \quad S = 5S^* = 5 \cdot 1.242626 = 6.21$

Άσκηση 3η: Οι βαθμοί των φοιτητριών και των φοιτητών του Τμήματος Φαρμακευτικής που εξετάστηκαν στο μάθημα των Εφαρμοσμένων Μαθηματικών, κατά την εξεταστική περίοδο του Ιανουαρίου 2013, δίνονται στον παρακάτω πίνακα

Βαθμός	0	1	2	3	4	5	6	7	8	9	10
Πλήθος φοιτητριών	1	8	7	6	5	15	17	22	18	3	2
Πλήθος φοιτητών	2	4	7	3	4	15	8	6	5	2	0





Βαθμός	0	1	2	3	4	5	6	7	8	9	10
Πλήθος φοιτητριών	1	8	7	6	5	15	17	22	18	3	2
Πλήθος φοιτητών	2	4	7	3	4	15	8	6	5	2	0

β. Να βρεθεί η δειγματική μέση τιμή του βαθμού για τις φοιτήτριες, για τους φοιτητές και για το σύνολο των εξεταζομένων. Για το σύνολο των παρατηρήσεων να βρεθούν η διασπορά, η τυπική απόκλιση, το πρώτο τεταρτημόριο, η διάμεσος, το τρίτο τεταρτημόριο και η κορυφή.

x_i	f_i^K	F_i^K	$x_i * f_i^K$	f_i^A	F_i^A	$x_i * f_i^A$	$f_i^\Sigma = f_i^K + f_i^A$	F_i^Σ	$x_i^2 * f_i^\Sigma$
0	1	1	0	2	2	0	3	3	0
1	8	9	8	4	6	4	12	15	12
2	7	16	14	7	13	14	14	29	56
3	6	22	18	3	16	9	9	38	81
4	5	27	20	4	20	16	9	47	144
5	15		75	15		75	30	77	750
6	17		102	8		48	25	102	900
7	22		154	6		42	28	130	1372
8	18		144	5		40	23	153	1472
9	3		27	2		18	5	158	405
10	2		20	0		0	2	160	200
Σύνολο	104		582	56		266	160		5392

Έτσι ο δειγματικός μέσος, $\overline{X_K}$, για τις φοιτήτριες δίνεται από τον τύπο:

$$\overline{x_K} = \frac{\sum_{i=1}^{11} x_i * f_i^K}{\sum_{i=1}^{11} f_i^K} \overset{\text{εδώ}}{=\!=} \frac{0 \times 1 + 1 \times 8 + \cdots + 10 \times 2}{1 + 8 + \cdots + 2} = \frac{582}{104} = 5.5962 \simeq 5.6$$

Αντιστοίχως ο δειγματικός μέσος, $\overline{X_A}$, για τους φοιτητές είναι:

$$\overline{x_A} = \frac{\sum_{i=1}^{11} x_i * f_i^A}{\sum_{i=1}^{11} f_i^A} \overset{\text{εδώ}}{=\!=} \frac{0 \times 2 + 1 \times 4 + \cdots + 10 \times 0}{2 + 4 \cdots + 0} = \frac{266}{56} = 4.75$$

Ο δειγματικός μέσος, $\overline{X_\Sigma}$, για το σύνολο των εξεταζομένων μπορεί να υπολογιστεί

$$\overline{x_\Sigma} = \frac{\sum_{i=1}^{11} x_i * f_i^K + \sum_{i=1}^{11} x_i * f_i^A}{\sum_{i=1}^{11} f_i^K + \sum_{i=1}^{11} f_i^A} \overset{\text{εδώ}}{=\!=} \frac{582 + 266}{104 + 56} = \frac{848}{160} = 5.3$$

$n = \sum_{i=1}^{11} f_i^\Sigma$ το μέγεθος του δείγματος

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{\kappa} (x_i - \overline{x})^2 \cdot f_i^\Sigma \overset{\text{ή}}{=\!=} \frac{1}{159} \left\{ \sum_{i=1}^{11} x_i^2 \cdot f_i^\Sigma - \frac{\left(\sum_{i=1}^{11} x_i \cdot f_i^\Sigma \right)^2}{160} \right\}$$

$$\overset{\text{εδώ}}{=\!=} \frac{1}{159} \left\{ 5392 - \frac{848^2}{160} \right\} = \frac{1}{159} \{5392 - 4494.4\} = \frac{897.6}{159} = 5.6452 = (2.376)^2$$

η τυπική απόκλιση είναι: S=2.376 μονάδες

Για τον υπολογισμό της διαμέσου, από την αθροιστική συχνότητα, F_i^{Σ},

$n = 160$, είναι άρτιος αριθμός έχουμε: $\delta = \dfrac{X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}}{2} \overset{\varepsilon\delta\acute{\omega}}{=\!=} \dfrac{x_{(80)} + x_{(81)}}{2} = 6$

για τις διατεταγμένες παρατηρήσεις ισχύει ότι: $x_{(78)} = \cdots = x_{(106)} = 6$.

Για το πρώτο τεταρτημόριο, επειδή $\frac{n}{4} = 40$ και $F_4 = 38 < 40 < 47 = F_5$

παρατηρούμε ότι $x_{(39)} = x_{(40)} = \cdots = x_{(47)} = 4$ άρα $Q_1 = x_{(40)} \left[+\frac{1}{4}(x_{(41)} - x_{(40)})\right] = 4$

Για το τρίτο τεταρτημόριο, επειδή $\frac{3n}{4} = 120$ και $F_7 = 102 < 120 < 130 = F_8$

παρατηρούμε ότι $x_{(103)} = x_{(104)} = \cdots = x_{(130)} = 7$ άρα $Q_3 = x_{(120)} \left[-\frac{1}{4}(x_{(120)} - x_{(119)})\right] = 7$

παρατηρούμε ότι: $\max\limits_{i=1,\ldots,11} f_i^{\Sigma} = f_7^{\Sigma} = 30$, η τιμή που εμφανίζεται

με αυτήν την συχνότητα είναι η κορυφή, δηλαδή $M = x_6 = 5$

γ. Επιλέγουμε τυχαία έναν εξεταζόμενο. Ποια είναι η πιθανότητα να έχει εξεταστεί επιτυχώς στο μάθημα; Εάν γνωρίζουμε ότι αυτός έχει εξεταστεί επιτυχώς στο μάθημα, τι είναι πιο πιθανό να είναι γυναίκα ή να είναι άνδρας;

με βάση τα χαρακτηριστικά: 1. επιτυχής ή μη εξέταση στο μάθημα

2. το φύλο των εξεταζομένων,

παίρνουμε τον ακόλουθο πίνακα συνάφειας:

	βαθμός< 5 (**ΜΕΕ**)	βαθμός≥ 5 (**ΕΕ**)	Σύνολο
φοιτήτριες (**Κ**)	27	77	104
φοιτητές (**Α**)	20	36	56
Σύνολο	47	113	160

Κ: το ενδεχόμενο να επιλέξουμε φοιτήτρια $P(\mathbf{K}) = \frac{104}{160} = 0.65$

Α: το ενδεχόμενο να επιλέξουμε φοιτητή, δηλαδή $\mathbf{K}'=\mathbf{A}$ $P(\mathbf{A}) = \frac{56}{160} = 0.35$

ΜΕΕ: το ενδεχόμενο ο εξεταζόμενος να μην εξετάστηκε επιτυχώς (βαθμός< 5)

ΕΕ: το ενδεχόμενο ο εξεταζόμενος να εξετάστηκε επιτυχώς (βαθμός≥ 5), δηλαδή $\mathbf{MEE}'=\mathbf{E}$

$$P(\mathbf{K} \mid \mathbf{EE}) = \frac{77}{113} = 0.6814 \qquad P(\mathbf{A} \mid \mathbf{EE}) = \frac{36}{113} = 0.3186$$

δ. Εάν επιλέξουμε στην τύχη μια φοιτήτρια και έναν φοιτητή από τους εξεταζόμενους, ποια είναι η πιθανότητα να έχουν και οι δύο εξεταστεί επιτυχώς στο μάθημα; Να αιτιολογήσετε την απάντησή σας.

$$\frac{\text{ευνοϊκά ενδεχόμενα}}{\text{πλήθος ισοπ. ενδεχομένων}} = \frac{77\times 36}{104\times 56} = 0.4760$$

$$\overset{\text{ή}}{=} \frac{77}{104} \times \frac{36}{56} = 0.74 \times 0.643 = P(\mathbf{EE}|\mathbf{K}) \times P(\mathbf{EE}|\mathbf{A})$$

επειδή οι δύο πληθυσμοί είναι ανεξάρτητοι.

Άσκηση 3η: Ένας συγκεκριμένος πληθυσμός ανθρώπων αποτελείται από 52% γυναίκες και 48% άνδρες. Επιλέγουμε τυχαία ένα από τα άτομα αυτού του πληθυσμού και βρίσκουμε ότι έχει αχρωματοψία. Εάν υποθέσουμε ότι το ποσοστό των ατόμων με αχρωματοψία είναι 25% για τις γυναίκες και 5% για τους άνδρες, ποια είναι η πιθανότητα το άτομο που επελέγει να είναι άνδρας;

Εάν συμβολίσουμε με

$\mathbf{\Gamma}$: το ενδεχόμενο επιλογής γυναίκας

$\mathbf{A}$: το ενδεχόμενο επιλογής άνδρα, δηλαδή $\mathbf{\Gamma}' = \mathbf{A}$

$\mathbf{AX}|\ \mathbf{\Gamma}$: το ενδεχόμενο επιλογής ατόμου με αχρωματοψία, όταν το άτομο αυτό είναι γυναίκα

$\mathbf{AX}|\ \mathbf{A}$: το ενδεχόμενο επιλογής ατόμου με αχρωματοψία, όταν το άτομο αυτό είναι άνδρας

Από τα δεδομένα του προβλήματος γνωρίζουμε ότι

$$P(\mathbf{\Gamma}) = 0.52 \qquad P(\mathbf{A}) = 0.48 \qquad P(\mathbf{AX}|\ \mathbf{\Gamma}) = 0.25 \qquad P(\mathbf{AX}|\ \mathbf{A}) = 0.05$$

Επειδή τα ενδεχόμενα **Γ, Α** αποτελούν μια διαμέριση του Ω,

εφαρμόζοντας το θεώρημα ολικής πιθανότητας

$$\begin{aligned} P(\mathbf{AX}) &= P(\mathbf{AX} \mid \mathbf{\Gamma})P(\mathbf{\Gamma}) + P(\mathbf{AX} \mid \mathbf{A})P(\mathbf{A}) \\ &= 0.25 \times 0.52 + 0.05 \times 0.48 = 0.13 + 0.024 = 0.154 \end{aligned}$$

Χρησιμοποιώντας τον τύπο του Bayes παίρνουμε ότι:

$$P(\mathbf{A} \mid \mathbf{AX}) = \frac{P(\mathbf{AX} \mid \mathbf{A})P(\mathbf{A})}{P(\mathbf{AX})} = \frac{0.05 \times 0.48}{0.154} = \frac{0.024}{0.154} = \underline{\underline{0.1558}}$$

Άσκηση 2η: Σε τρία καλάθια ενός νοικοκυριού βρίσκονται 15 φρούτα, αχλάδια και μήλα, τοποθετημένα ως εξής:

	Καλάθι 1ο	Καλάθι 2ο	Καλάθι 3ο
Αχλάδια	2	2	3
Μήλα	4	2	2

Η νοικοκυρά επιλέγει στην τύχη ένα καλάθι και στη συνέχεια επιλέγει από αυτό ένα φρούτο.

Εάν συμβολίσουμε με

$\mathbf{M}$: το ενδεχόμενο επιλογής μήλου

$\mathbf{A}$: το ενδεχόμενο επιλογής αχλαδιού, δηλαδή $\mathbf{M}' = \mathbf{A}$

$\mathbf{K}_i$: το ενδεχόμενο επιλογής του i καλαθιού, για $i = 1, 2, 3$

και ισχύει ότι: $K_1 \bigcup K_2 \bigcup K_3 = \Omega$ με $K_i \bigcap K_j = \varnothing$ για $i \neq j$

$\mathbf{A}|\mathbf{K}_i$: το ενδεχόμενο επιλογής αχλαδιού, όταν η επιλογή γίνεται από το i καλάθι, για $i = 1, 2, 3$.

τα δεδομένα του προβλήματος γνωρίζουμε ότι, εφόσον επιλέγουμε στην τύχη ένα καλάθι,

$$P(\mathbf{K}_1) = P(\mathbf{K}_2) = P(\mathbf{K}_3) \overset{(1)}{\Longrightarrow} P(\mathbf{K}_1) + P(\mathbf{K}_2) + P(\mathbf{K}_3) = P(\Omega) = 1 \Rightarrow P(\mathbf{K}_i) = \frac{1}{3}, \quad i = 1, 2, 3,$$

$$P(\mathbf{A} \mid \mathbf{K}_1) = \frac{2}{2+4} = \frac{2}{6} = 0.3333,\ P(\mathbf{A} \mid \mathbf{K}_2) = \frac{2}{2+2} = \frac{1}{2} = 0.5,\ P(\mathbf{A} \mid \mathbf{K}_3) = \frac{3}{3+2} = \frac{3}{5} = 0.6.$$

α. Ποια είναι η πιθανότητα το φρούτο αυτό να είναι αχλάδι;

Επειδή λόγω της (1) τα K_1, K_2, K_3 αποτελούν μια διαμέριση του Ω, εφαρμόζοντας το θεώρημα ολικής πιθανότητας παίρνουμε ότι:

$$\begin{aligned} P(\mathbf{A}) &= P(\mathbf{A} \mid \mathbf{K}_1)P(\mathbf{K}_1) + P(\mathbf{A} \mid \mathbf{K}_2)P(\mathbf{K}_2) + P(\mathbf{A} \mid \mathbf{K}_3)P(\mathbf{K}_3) \\ &= \frac{1}{3} \times \frac{1}{3} + \frac{1}{2} \times \frac{1}{3} + \frac{3}{5} \times \frac{1}{3} = \frac{10+15+18}{30} \times \frac{1}{3} = \frac{43}{90} = 0.4778 \end{aligned}$$

β. Εάν τελικά επιλέξει ένα μήλο, τι είναι πιο πιθανό: να το έχει επιλέξει από το 2ο ή από το 3ο καλάθι;

$$P(\mathbf{M}) = 1 - P(\mathbf{A}) = \frac{47}{90} = 0.5222$$

Χρησιμοποιώντας τον τύπο του Bayes παίρνουμε ότι:

$$P(\mathbf{K}_2 \mid \mathbf{M}) = \frac{P(\mathbf{M} \mid \mathbf{K}_2)P(\mathbf{K}_2)}{P(\mathbf{M})} = \frac{\frac{1}{2} \times \frac{1}{3}}{\frac{47}{30} \times \frac{1}{3}} = \frac{30}{94} = 0.3191$$

$$P(\mathbf{K}_3 \mid \mathbf{M}) = \frac{P(\mathbf{M} \mid \mathbf{K}_3)P(\mathbf{K}_3)}{P(\mathbf{M})} = \frac{\frac{2}{5} \times \frac{1}{3}}{\frac{47}{30} \times \frac{1}{3}} = \frac{60}{235} = 0.2553$$

Οπότε είναι πιο πιθανό να έχει επιλεγεί από το δεύτερο καλάθι (στο συγκεκριμένο πρόβλημα, $P(\mathbf{K}_2) = P(\mathbf{K}_3)$, αρκούσε η σύγκριση $P(\mathbf{M} \mid \mathbf{K}_2) = \frac{1}{2} > \frac{2}{5} = P(\mathbf{M} \mid \mathbf{K}_3)$)

γ. Είναι τα ενδεχόμενα επιλογή αχλαδιού και επιλογή του 3ου καλαθιού ανεξάρτητα;

Για να είναι ανεξάρτητα τα ενδεχόμενα $\mathbf{A}$ και $\mathbf{K}_3$ θα πρέπει να ισχύει (ισοδύναμα)

$$P(\mathbf{A} \mid \mathbf{K}_3) = P(\mathbf{A}) \quad ή \quad P(\mathbf{K}_3 \mid \mathbf{A}) = P(\mathbf{K}_3) \quad ή \quad P(\mathbf{A} \cap \mathbf{K}_3) = P(\mathbf{A})P(\mathbf{K}_3)$$

Όμως $P(\mathbf{K}_3 \mid \mathbf{A}) = \frac{\frac{3}{5}\times\frac{1}{3}}{\frac{43}{30}\times\frac{1}{3}} = \frac{90}{215} = 0.4186$ και $P(\mathbf{A} \cap \mathbf{K}_3) = \frac{3}{5} \times \frac{1}{3} = 0.2$

$$0.6 \neq 0.4778 \;\; ή \;\; 0.4186 \neq 0.3333 \;\; ή \;\; 0.2 \neq 0.1593 = \frac{43}{90} \times \frac{1}{3}$$

Άρα δεν είναι ανεξάρτητα.

Ασκήσεις σε Βασικές Κατανομές

Άσκηση 1η:

β. Δώστε ένα παράδειγμα στο οποίο μπορεί να χρησιμοποιηθεί η διωνυμική κατανομή. Εάν η τυχαία μεταβλητή X ακολουθεί διωνυμική κατανομή και ισχύει ότι $E(X) = 4$ και $\Delta(X) = 2.4$, να υπολογιστούν: $P(X > 2)$, $E(10 - X)$ και η $\Delta(10 - X)$.

Μας ενδιαφέρουν n ανεξάρτητες δοκιμές Bernoulli, με την ίδια πιθανότητα επιτυχίας p η κάθε μία.

Πχ. : πλήθος κοριτσιών από n επιλεγμένους στην τύχη φοιτητές

(θεωρούμε ότι ο πληθυσμός των φοιτητών είναι τόσο μεγάλος, ώστε η αφαίρεση μερικών από αυτούς δεν μεταβάλλει, από επιλογή σε επιλογή, το ποσοστό των κοριτσιών στον πληθυσμό, το οποίο δεν ξέρουμε εάν είναι $\frac{1}{2}$)

Εάν $X \sim \text{Binomial}(n, p)$, τότε $\mathrm{E}(X) = np$ και $\mathrm{Var}(X) = np(1-p)$.

Άρα $\dfrac{\mathrm{Var}(X)}{\mathrm{E}(X)} = 1 - p \overset{\text{εδώ}}{=} \dfrac{2.4}{4} = 0.6 \Rightarrow p = 0.4$ και $n = \dfrac{\mathrm{E}(X)}{p} \overset{\text{εδώ}}{=} \dfrac{4}{0.4} = 10.$

Οπότε $X \sim$Binomial$(10, 0.4)$ και άρα $P(X=x) = \binom{10}{x} 0.4^x 0.6^{10-x}, \;\; x = 0, 1, \ldots, 10$

$$\begin{aligned} P(X > 2) &= 1 - P(X \leq 2) = 1 - (P(X=0) + P(X=1) + P(X=2)) \\ &= 1 - (\tbinom{10}{0} 0.6^{10} + \tbinom{10}{1} 0.4 \cdot 0.6^9 + \tbinom{10}{2} 0.4^2 \cdot 0.6^8) \\ &= 1 - (0.6^{10} + 10 \cdot 0.4 \cdot 0.6^9 + 45 \cdot 0.4^2 \cdot 0.6^8) \\ &= 1 - (0.0060 + 0.0403 + 0.1209) = 1 - 0.1673 = \underline{\underline{0.8327}} \end{aligned}$$

από την γραμμικότητα της μέσης τιμής έχουμε

$$E(10 - X) \; = \; 10 - E(X) \; = \; 10 - 4 \; = \; \underline{\underline{6}}$$

$$\Delta(10 - X) = (\mathrm{Var}(10 - X) =)(-1)^2 \Delta(X) = \underline{\underline{2.4}}$$

γ. Εάν το αναμενόμενο πλήθος ακυρώσεων σε προγραμματισμένα χειρουργεία για ένα μήνα στο νοσοκομείο του Ρίου είναι 2.5, τι είναι πιο πιθανό να συμβεί: να έχουμε για τους επόμενους τρεις μήνες μία ακύρωση κάθε μήνα ή να έχουμε συνολικά στους τρεις μήνες τρεις ακυρώσεις; Ποιες κατανομές θα χρησιμοποιήσετε για τους υπολογισμούς σας και γιατί.

X: το πλήθος των ακυρώσεων σε προγραμματισμένα χειρουργεία για ένα μήνα

δίδεται ότι $E(X) = 2.5$, άρα $X \sim$Poisson(2.5)

Οπότε $P(X = x) = e^{-2.5}\ \frac{2.5^x}{x!},\ x = 0, 1, \ldots,.$

P(να έχω μία ακριβώς ακύρωση κατά τη διάρκεια ενός μήνα)$=P(X = 1) = 2.5e^{-2.5} = 0.2052$

Y: το πλήθος των ακυρώσεων σε προγραμματισμένα χειρουργεία για τρεις μήνες

από τις ιδιότητες της κατανομής Poisson έχουμε ότι $Y \sim$Poisson$(3 \times 2.5 = 7.5)$

Αυτό που μας ζητούν είναι να συγκρίνουμε τις πιθανότητες:

$$P(X=1)^3 = 0.2052^3 = 2.5^3 e^{-3\times 2.5} = 0.0086$$

$$P(Y=3) = e^{-3\times 2.5}\ \frac{(3\times 2.5)^3}{3!} = \frac{3^3}{3!}\,0.0086 \quad = 0.0388$$

Οι κατανομές που χρησιμοποιήθηκαν είναι:

η Poisson γιατί έχω συμβάντα στο χρόνο

και ανεξάρτητες και ισόνομες δοκιμές Bernoulli

Άσκηση 3η: Οι μετρήσεις του ουρικού οξέως (σε mg/100ml) για άνδρες ηλικίας 35-50 ετών ακολουθούν κανονική κατανομή με μέση τιμή 5.4 και τυπική απόκλιση 1.

α. Να βρεθεί η πιθανότητα για έναν άνδρα αυτής της ηλικίας η μέτρηση του ουρικού οξέως να είναι μεγαλύτερη από 5, όταν γνωρίζουμε ότι αυτή είναι μικρότερη του μέσου.

X: τη μέτρηση του ουρικού οξέως (σε mg/100ml) για έναν άνδρα ηλικίας 35-50 ετών.

$X \sim N(\mu = 5.4, \sigma^2 = 1^2)$, οπότε $E(X) = 5.4$

τυποποιημένη τ.μ. $Z = \frac{X-\mu}{\sigma} = \frac{X-5.4}{1} \sim N(0,1)$,

άρα την α.σ.κ. τη συμβολίζουμε με $P(Z \leq z) = \Phi(z)$

$$P(X > 5 | X < E(X)) = \frac{P(X > 5 \cap X < E(X))}{P(X < E(X))} \overset{\mathrm{E(X)=5.4}}{=\!=\!=} \frac{P(5 < X < 5.4)}{P(X < 5.4)}$$

$$P(X<5.4) = P\left(\frac{X-5.4}{1} < 5.4-5.4\right) = P(Z<0) = \Phi(0) \overset{\text{N(0,1) συμμ. ως προς 0}}{\underset{\Phi(0)=1-\Phi(0)}{=\!=\!=\!=}} 0.5$$

$$P(5<X<5.4) = P\left(5-5.4 < \frac{X-5.4}{1} < 5.4-5.4\right) = P(-0.4<Z<0)$$

$$= P(Z \le 0) - P(Z \le -0.4)$$

$$= \Phi(0) - \Phi(-0.4) \overset{\text{N(0,1) συμμ. ως προς 0}}{=\!=\!=\!=} \Phi(0) - (1-\Phi(0.4))$$

$$= 0.5 - 1 + 0.655 = 0.155$$

Αυτά ισχύουν, επειδή η σ.π. της $N(0,1)$ είναι συμμετρική ως προς το 0, δηλαδή,

$$P(Z<0) = P(Z>0) \Rightarrow P(Z<0) = 1-P(Z<0) \Rightarrow \Phi(0) = 1-\Phi(0) \Rightarrow \Phi(0) = 0.5.$$

$$P(Z<-0.4) = P(Z>0.4) \Rightarrow P(Z<-0.4) = 1-P(Z<0.4) \Rightarrow \Phi(-0.4) = 1-\Phi(0.4)$$

$$P(X>5|X<E(X)) = \frac{P(5<X<5.4)}{P(X<5.4)} = \frac{0.155}{0.5} = 0.31$$

β. Παίρνουμε δείγμα 25 μετρήσεων. Ποια είναι η πιθανότητα ο μέσος όρος των μετρήσεων να είναι μεταξύ 5.3 και 5.6;

εάν $X_1, \ldots, X_{25}$ ανεξάρτητες παρατηρήσεις από $N(\mu, \sigma^2)$

τότε $\overline{X} = \frac{X_1 + \cdots + X_{25}}{25} \sim N\left(\mu, \frac{\sigma^2}{25}\right)$ δηλαδή $N(5.4, \frac{1}{25}) \equiv N(5.4, \frac{1}{5^2})$

και τυποποιούμε

$$P(5.3 < \overline{X} < 5.6) = P\left(\frac{5.3 - 5.4}{\frac{1}{5}} < \frac{\overline{X} - 5.4}{\frac{1}{5}} < \frac{5.6 - 5.4}{\frac{1}{5}}\right)$$

$$= P(-0.5 < Z < 1) = \Phi(1) - \Phi(-0.5)$$

$$\overset{\text{N(0,1) συμμ. ως προς 0}}{=\!=\!=\!=\!=} \Phi(1) - (1 - \Phi(0.5))$$

$$= 0.841 - 1 + 0.691 = 0.532$$