Δίδονται $\Phi(0.2) = 0.579$, $\Phi(0.4) = 0.655$, $\Phi(0.5) = 0.691$, $\Phi(0.8) = 0.788$, $\Phi(1) = 0.841$, $\Phi(1.2) = 0.885$, $\Phi(1.4) = 0.919$, $\Phi(1.5) = 0.933$, $\Phi(1.645) = 0.95$, $\Phi(1.96) = 0.975$, $\Phi(2) = 0.977$, $\Phi(2.5) = 0.994$, $\mathcal{X}_6^2(0.1) = 10.64$, $\mathcal{X}_2^2(0.1) = 4.61$, $\mathcal{X}_6^2(0.05) = 12.59$, $\mathcal{X}_2^2(0.05) = 5.99$, $t_6(0.05) = 1.943$, $t_6(0.1) = 1.439$, $t_7(0.05) = 1.895$, $t_7(0.1) = 1.414$.

Θέμα 1ο: Για την μελέτη του πληθυσμού των αρουραίων της ερήμου (Gerbillinae) καταγράφηκε ο αριθμός των νεογνών που γεννήθηκαν ζωντανά σε 68 γέννες. Τα δεδομένα αυτά δίνονται στον πίνακα που ακολουθεί:

ζωντανά νεογνά	1	2	3	4	5	6	7
Πλήθος γεννών	3	17	8	10	14	9	7

α. Για τον αριθμό των νεογνών αυτού του είδους αρουραίου, να υπολογιστεί η (δειγματική) μέση τιμή, η τυπική απόκλιση, το πρώτο τεταρτημόριο και η διάμεσος των παρατηρήσεων.

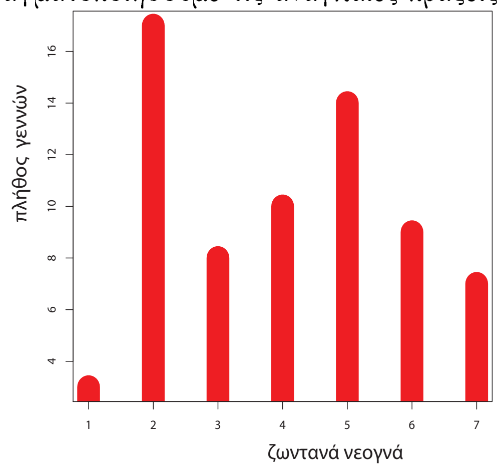
β. Αν μαζί με τα νεογνά προσμετρήσουμε και τη μητέρα, πώς μεταβάλλονται τα παραπάνω μέτρα;

γ. Υποθέτοντας ότι ο αριθμός των νεογνών αυτών ακολουθεί κατανομή με μέσο μ, να δοθεί διάστημα εμπιστοσύνης, συντελεστού εμπιστοσύνης 90%, για τον μέσο μ. Ποιας κατανομής τα ποσοστιαία σημεία χρησιμοποιήσατε για την κατασκευή του διαστήματος αυτού και γιατί;

δ. Με βάση το διάστημα εμπιστοσύνης που κατασκευάστηκε, μπορούμε να αποδεχτούμε την υπόθεση $H_o : \mu = 3.8$ έναντι της $H_1 : \mu \neq 3.8$; Αν ναι, σε ποιο επίπεδο σημαντικότητας;

Λύση: Τα δεδομένα είναι διακριτά και παρουσιάζονται στον πίνακα που μας δίδεται ομαδοποιημένα, κατά φυσικό τρόπο. Έτσι δημιουργούμε τον ακόλουθο πίνακα (για να πραγματοποιήσουμε τις αναγκαίες πράξεις):

x_i	f_i	F_i	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
1	3	3	3	3
2	17	20	34	68
3	8	28	24	72
4	10	38	40	160
5	14	52	70	350
6	9	61	54	324
7	7	68	49	343
Σύνολο	68		274	1320



α. Ο δειγματικός μέσος, $\overline{X}$, για τον ο αριθμός των νεογνών που γεννήθηκαν ζωντανά ανά γέννα, δίνεται από τον τύπο:

$$\overline{x} = \frac{\sum_{i=1}^{\kappa} x_i \cdot f_i}{\sum_{i=1}^{\kappa} f_i} \underset{\kappa=7}{\overset{\varepsilon\delta\acute{\omega}}{=\!=}} \frac{1\times 3 + 2\times 17 + \cdots + 7\times 7}{3+17+\cdots+7} = \frac{274}{68} = \underline{\underline{4.03 \text{ νεογνά}}},$$

όπου κ το πλήθος των διαφορετικών τιμών των παρατηρήσεων, εδώ $\kappa = 7$. Εάν συμβολίσουμε με $n = \sum_{i=1}^{7} f_i$ το μέγεθος του δείγματος, η δειγματική διασπορά δίδεται από τον τύπο:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{\kappa}(x_i - \overline{x})^2 \cdot f_i \overset{\acute{\eta}}{=\!=} \frac{1}{n-1}\left\{\sum_{i=1}^{\kappa} x_i^2 \cdot f_i - \frac{\left(\sum_{i=1}^{\kappa} x_i \cdot f_i\right)^2}{n}\right\} \overset{\varepsilon\delta\acute{\omega}}{=\!=} \frac{1}{67}\left\{1320 - \frac{274^2}{68}\right\}$$

$$= \frac{1}{67}\left\{1320 - \frac{75076}{68}\right\} = \frac{1}{67}\{1320 - 1104.059\} = \frac{215.94}{67} = 3.22 = \underline{\underline{1.795}}^2.$$

Οπότε η ζητούμενη τυπική απόκλιση είναι $S = \underline{\underline{1.795}}$.

Για το πρώτο τεταρτημόριο, παρατηρούμε ότι επειδή $\frac{n}{4} = 17$ και $F_1 = 3 < 17 < 20 = F_2$, αυτή ταυτίζεται με την δεύτερη μεγαλύτερη τιμή, δηλαδή $Q_1 = x_{(17)} = 2$.
Για τη διάμεσο, παρατηρούμε ότι επειδή $\frac{n}{2} = 34$ και $F_3 = 28 < 34 < 38 = F_4$, αυτή ταυτίζεται με την τέταρτη μεγαλύτερη τιμή, δηλαδή $\delta = x_{(34)} = 4$.

β. Τα νέα δεδομένα, νεογνά και μητέρες, είναι: $y_i = x_i + 1$, $i = 1, ..., 68$ (άτομα)

οπότε: $\overline{Y} = \dfrac{\sum_{i=1}^{n}(X_i+1)}{n} = \dfrac{\sum_{i=1}^{n} X_i}{n} + \dfrac{n}{n} = \overline{X} + 1 \overset{\text{εδώ}}{=} \underline{\underline{5.03 \text{ άτομα}}}$

Ανάλογα,

$S_y^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}\left(X_i + 1 - (\overline{X}+1)\right)^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 = S_x^2$

$\Rightarrow S_y = S_x = \underline{\underline{1.795 \text{ άτομα}}},$

επίσης $Q_1^y = y_{(17)} = x_{(17)} + 1 = Q_1^x + 1 = \underline{\underline{3 \text{ άτομα}}}$ και $\quad \delta_y = y_{(34)} = x_{(34)} + 1 = \delta_x + 1 = \underline{\underline{5 \text{ άτομα}}}$

γ. Θα κατασκευάσω διάστημα εμπιστοσύνης (δ.ε.), ίσων ουρών, για το μ, συντελεστού εμπιστοσύνης (σ.ε.) $1-\alpha = 0.90$, δηλαδή δ.ε. για τον μέσο άγνωστης κατανομής με άγνωστη διασπορά και μεγάλο μέγεθος δείγματος $n = 68$. Ως εκ τούτου, χρησιμοποιώ τον τύπο

$$\left[\overline{X} - z_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}},\ \ \overline{X} + z_{\frac{\alpha}{2}}\frac{S}{\sqrt{n}}\right] \overset{\text{εδώ}}{=} \left[\overline{X} - z_{0.05}\frac{S}{\sqrt{68}},\ \ \overline{X} + z_{0.05}\frac{S}{\sqrt{68}}\right] =$$
$$\left[4.029 - 1.645\frac{1.795}{\sqrt{68}},\ \ 4.029 + 1.645\frac{1.795}{\sqrt{68}}\right] = [4.029 - 1.645 \cdot 0.217,\ \ 4.029 + 1.645 \cdot 0.217] =$$
$$[4.029 - 0.358,\ \ 4.029 + 0.358] = \underline{\underline{[3.671,\ \ 4.387]}} \tag{1}$$

εφόσον, $1 - \alpha = 0.9 \Rightarrow \frac{\alpha}{2} = 0.05$ και $z_{0.05} = 1.645$ διότι, για $Z \sim N(0,1)$, $P(Z > z_{0.05}) = 0.05 \Rightarrow \Phi(z_{0.05}) = P(Z \leq z_{0.05}) = 1 - 0.05 \Rightarrow \Phi(z_{0.05}) = 0.95$, όμως $\Phi(1.645) = 0.95$. Για την κατασκευή του διαστήματος αυτού, επειδή το μέγεθος του δείγματος είναι μεγάλο ($n = 68$), μπορεί να εφαρμοστεί το κεντρικό οριακό θεώρημα και λόγω αυτού η ποσότητα $\dfrac{\overline{X} - \mu}{\sqrt{\frac{S^2}{n}}} \underset{n\to\infty}{\overset{\text{προσεγ.}}{\sim}} N(0,1)$.

δ. Επειδή έχω αμφίπλευρο έλεγχο, αποδέχομαι την μηδενική υπόθεση $H_o : \mu = \mu_0$ σε ε.σ. α αν το δ.ε. ίσων ουρών για το μ σ.ε. $1 - \alpha$ περιέχει το μ_0, διαφορετικά την απορρίπτω σε ε.σ. α. Εδώ, το δ.ε. (1) σ.ε. 90% περιέχει το 3.8, άρα η H_0 γίνεται δεκτή σε ε.σ. 10%.

Θέμα 2ο: **α.** Το ποσοστό μιας συγκεκριμένης ασθένειας σε ένα πληθυσμό είναι 8%. Το 85% από εκείνους που έχουν την ασθένεια εμφανίζουν ένα συγκεκριμένο εργαστηριακό εύρημα (θετικό τεστ), ενώ μόνο το 10% από τους μη-ασθενείς παρουσιάζουν το ίδιο εύρημα (ψευδή θετικά).

1. Να βρεθεί η πιθανότητα ένα άτομο του πληθυσμού να μην εμφανίσει αυτό το εργαστηριακό εύρημα.
2. Ποια είναι η πιθανότητα ένα άτομο που εμφανίζει το εργαστηριακό εύρημα να έχει πράγματι την ασθένεια;

β. Το ποσοστό του λίπους των ατόμων, σε έναν πληθυσμό ανδρών ηλικίας 25 έως 35 ετών, ακολουθεί την κανονική κατανομή με μέση τιμή $\mu = 20\%$ και τυπική απόκλιση $\sigma = 3\%$.

1. Να βρεθεί η πιθανότητα ένα άτομο να έχει ποσοστό λίπους από 18.5% έως 23%.
2. Παίρνουμε δείγμα από 8 άτομα. Ποια η πιθανότητα τουλάχιστον 2 άτομα από αυτά να έχουν ποσοστό λίπους από 18.5% έως 23% (να αιτιολογήσετε);

Λύση: **α.** Εάν συμβολίσουμε με

A: το ενδεχόμενο επιλογής ασθενούς ατόμου

Y: το ενδεχόμενο επιλογής μη-ασθενούς ατόμου, δηλαδή $\mathbf{Y}' = \mathbf{A}$

$\mathbf{T}^+$: το ενδεχόμενο το τεστ να είναι θετικό

$\mathbf{T}^-$: το ενδεχόμενο το τεστ να είναι αρνητικό, δηλαδή $\mathbf{T}^{-'} = \mathbf{T}^+$

$\mathbf{T}^+ \mid \mathbf{A}$: το ενδεχόμενο το τεστ να είναι θετικό, όταν όντος το άτομο αυτό έχει τη νόσο

$\mathbf{T}^+ \mid \mathbf{Y}$: το ενδεχόμενο το τεστ να είναι θετικό, όταν το άτομο αυτό δεν πάσχει από τη νόσο

Από τα δεδομένα του προβλήματος γνωρίζουμε ότι

$$P(\mathbf{A}) = 0.08 \Rightarrow P(\mathbf{Y}) = 0.92 \quad P(\mathbf{T}^+ \mid \mathbf{A}) = 0.85 \Rightarrow P(\mathbf{T}^- \mid \mathbf{A}) = 0.15 \quad P(\mathbf{T}^+ \mid \mathbf{Y}) = 0.1 \Rightarrow P(\mathbf{T}^- \mid \mathbf{Y}) = 0.9$$

1. Επειδή τα ενδεχόμενα **A, Y** αποτελούν μια διαμέριση του Ω, εφαρμόζοντας το θεώρημα ολικής πιθανότητας παίρνουμε ότι:

$$P(\mathbf{T}^-) = P(\mathbf{T}^- \mid \mathbf{A})P(\mathbf{A}) + P(\mathbf{T}^- \mid \mathbf{Y})P(\mathbf{Y}) = 0.15 \times 0.08 + 0.9 \times 0.92 = 0.012 + 0.828 = \underline{\underline{0.84}}$$

2. Χρησιμοποιώντας τον τύπο του Bayes παίρνουμε ότι:

$$P(\mathbf{A} \mid \mathbf{T}^+) = \frac{P(\mathbf{T}^+ \mid \mathbf{A})P(\mathbf{A})}{P(\mathbf{T}^+)} = \frac{0.85 \times 0.08}{1 - P(\mathbf{T}^-)} = \frac{0.068}{0.16} = \underline{\underline{0.425}}$$

β. **1.** Συμβολίζουμε με την (τυχαία μεταβλητή) τ.μ. X,

X: το ποσοστό του λίπους των ατόμων ενός πληθυσμού ανδρών ηλικίας 25 έως 35 ετών.

Τότε, σύμφωνα με τα δεδομένα του προβλήματος, $X \sim N(\mu = 20, \sigma^2 = 3^2)$ (επί τοις 100), οπότε η τυποποιημένη τ.μ. $Z = \frac{X-\mu}{\sigma} = \frac{X-20}{3} \sim N(0,1)$ και άρα την α.σ.κ. τη συμβολίζουμε με $P(Z \le z) = \Phi(z)$.

$$\begin{aligned} P(18.5 < X < 23) &= P\left(\frac{18.5-20}{3} < \frac{X-20}{3} < \frac{23-20}{3}\right) = P(-0.5 < Z < 1) \\ &= P(Z < 1) - P(Z \le -0.5) = \Phi(1) - \Phi(-0.5) \overset{\text{N(0,1) συμμ. ως προς 0}}{=\!=\!=} \Phi(1) - (1 - \Phi(0.5)) \\ &= 0.841 - 1 + 0.691 = \underline{\underline{0.532}} \end{aligned}$$

2. Αν συμβολίσουμε με την τ.μ.

W: το πλήθος, από τα 8 άτομα, που έχουν ποσοστό λίπους από 18.5% έως 23%. Τότε:

$$W \sim \text{Binomial}\,(n = 8, p = P(18.5 < X < 23) = 0.532)$$

με $P(W = w) = \binom{8}{w} 0.532^w 0.468^{8-w}, \;\; w = 0, 1, \ldots, 8.$

Αυτό ισχύει, επειδή η μέτρηση του ποσοστού λίπους ενός ατόμου είναι ανεξάρτητη από τη μέτρηση για ένα άλλο άτομο και βρίσκεται κάθε μέτρηση στα ζητούμενα όρια (εδώ από 18.5% έως 23%) με την ίδια πιθανότητα $p = P(18.5 < X < 23) = 0.532$. Έτσι, έχουμε ανεξάρτητες και ισόνομες δοκιμές Bernoulli, από τις οποίες προκύπτει η διωνυμική κατανομή.

Μας ενδιαφέρει ο υπολογισμός της πιθανότητας

$$\begin{aligned} P(W \ge 2) &= \sum_{w=2}^{8} P(W = w) = 1 - P(W < 2) = 1 - P(W \le 1) = 1 - (P(W = 0) + P(W = 1)) \\ &= 1 - \binom{8}{0} 0.468^8 - \binom{8}{1} 0.532 \cdot 0.468^7 = 1 - (0.0023 + 0.0207) = 1 - 0.0230 = \underline{\underline{0.977}} \end{aligned}$$

Θέμα 3ο: **α.** Για την μελέτη της αποτελεσματικότητας ενός εμβολίου για την παρωτίτιδα, επιλέχθηκε ένας πληθυσμός 395 παιδιών εκ των οποίων τα 200 εμβολιάστηκαν με ένα εικονικό εμβόλιο (placebo) και τα 195 με το υπό μελέτη νέο εμβόλιο. Στον πίνακα που ακολουθεί έχουν καταγραφεί, για το διάστημα των 24 μηνών μετά τον εμβολιασμό, τα κρούσματα παρωτίτιδας στον πληθυσμό αυτό:

	δεν νόσησαν	νόσησαν ελαφρά	νόσησαν βαριά
placebo	100	71	29
νέο εμβόλιο	146	32	17

1. Να ελεγχθεί, σε επίπεδο σημαντικότητας α=10%, εάν το νέο εμβόλιο είναι αποτελεσματικό. Τι μπορείτε να πείτε για το p-value των δεδομένων για τον έλεγχο αυτόν;

2. Είναι ίδια η απόφασή σας σε επίπεδο σημαντικότητας α=5%; (να σχολιάσετε)
3. Εάν συμβολίσουμε με p_B το ποσοστό των παιδιών που εμβολιάστηκαν με το νέο εμβόλιο, τα οποία νόσησαν βαριά, να ελέγξετε εάν σε επίπεδο σημαντικότητας α=5% δεχόμαστε την υπόθεση $H_o : p_B \leq 0.08$ έναντι της $H_1 : p_B > 0.08$
4. Εάν συμβολίσουμε με p_{BP} το ποσοστό των παιδιών που εμβολιάστηκαν με εικονικό εμβόλιο (placebo), τα οποία νόσησαν βαριά, να ελέγξετε εάν σε επίπεδο σημαντικότητας α=5% δεχόμαστε την υπόθεση $H_o : p_B = p_{BP}$ έναντι της $H_1 : p_B \neq p_{BP}$

β. Δώστε ένα παράδειγμα στο οποίο μπορεί να χρησιμοποιηθεί η κατανομή Poisson. Εάν η τυχαία μεταβλητή X ακολουθεί κατανομή Poisson και ισχύει ότι $\Delta(X) = 1$, να υπολογιστούν: $P(X \leq 2)$, $E(5X + 1)$ και η $\Delta(-3X + 2)$.

Λύση: α. **1.** Ο έλεγχος που ζητείται είναι ένας έλεγχος ομοιογένειας:
$H_o : P$(κατηγορία νόσησης|εμβολ. με placebo) $= P$(κατηγορία νόσησης|εμβολ. με το νέο εμβόλιο),
H_1 : διαφορετικά,
με κατηγορία νόσησης $\in$ {δεν νόσησε, νόσησε ελαφρά, νόσησε βαριά}

	κατηγορία νόσησης			
είδος εμβολίου	δεν νόσησαν	νόσησαν ελαφρά	νόσησαν βαριά	
placebo	100 (124.6)	71 (52.2)	29 (23.3)	$\pi_{1\bullet} = 200$
νέο εμβόλιο	146 (121.4)	32 (50.8)	17 (22.7)	$\pi_{2\bullet} = 195$
	$246 = \pi_{\bullet 1}$	$103 = \pi_{\bullet 2}$	$46 = \pi_{\bullet 3}$	$n = 395$

Για τον έλεγχο, χρησιμοποιείται η στατιστική συνάρτηση (σ.σ.)

$$X^2 = \sum_{i=1}^{\kappa_1} \sum_{j=1}^{\kappa_2} \frac{(\pi_{ij} - \theta_{ij})^2}{\theta_{ij}} \left(= \sum_{i=1}^{\kappa_1} \sum_{j=1}^{\kappa_2} \frac{\pi_{ij}^2}{\theta_{ij}} - n \right),$$

όπου π_{ij} οι παρατηρούμενες, $\theta_{ij} = \frac{\pi_{i\bullet} \times \pi_{\bullet j}}{n}$ οι αντίστοιχες θεωρητικές τιμές υπό την H_o, και κ_1, κ_2 το πλήθος των γραμμών και των στηλών αντιστοίχως του πίνακα συνάφειας. Στην περίπτωση που μελετούμε $\kappa_1 = 2, \kappa_2 = 3$

Όμως $X^2 \overset{\text{ασυμπτ.}}{\sim} \mathcal{X}^2_{(\kappa_1-1)\times(\kappa_2-1)}$, όταν ισχύει η H_o. Έτσι απορρίπτουμε την H_o, σε ε.σ. α, όταν $X^2 \geq \mathcal{X}^2_{(\kappa_1-1)\times(\kappa_2-1)}(\alpha)$. Εδώ,
$X^2 = \frac{(100-124.6)^2}{124.6} + \frac{(71-52.2)^2}{52.2} + \cdots + \frac{(32-50.8)^2}{50.8} + \frac{(17-22.7)^2}{22.7} = 4.8415 + 6.811 + \cdots + 6.9865 + 1.4352 = 26.44$
οπότε $X^2 = \underline{\underline{26.44 > 4.61 = \mathcal{X}^2_2(0.1)}}$, άρα απορρίπτουμε την H_o, δηλ. στα παιδιά η κατηγορία νόσησης διαφέρει ανάλογα με το εμβόλιο που έχουν κάνει (επιθυμητό!).

Άρα για το p-value των δεδομένων για τον έλεγχο αυτόν ισχύει: 10% > p-value, μιας και αυτό είναι το μικρότερο ε.σ. στο οποίο απορρίπτουμε την H_o, ή ισοδύναμα (εδώ όπου το κριτήριο ακολουθεί, υπό την H_o, συνεχή κατανομή) το μεγαλύτερο ε.σ. στο οποίο αποδεχόμαστε την H_o.

(p-value: $P(\mathcal{X}^2_2 > 26.44) = 0(= 1.81 \times 10^{-6})$, δηλ. δεν αποδέχομαι την H_o για κανένα ε.σ..)

2. Επειδή η υπόθεση H_o απορρίπτεται σε επίπεδο σημαντικότητας α=10% δεν γνωρίζουμε εάν απορρίπτεται επίσης και σε ε.σ. μικρότερο του 10%. Εδώ όντως, επειδή $X^2 = 26.44 > 5.99 = \mathcal{X}^2_2(0.05)$ και σε ε.σ. α=5% απορρίπτουμε την H_o, δηλ. στα παιδιά αυτά η κατηγορία νόσησης διαφέρει ανάλογα με το εμβόλιο που έχουν κάνει.

Αυτό σημαίνει ότι για το p-value των δεδομένων για τον έλεγχο αυτόν ισχύει: p-value < 5% < 10%.

3. Ο έλεγχος $H_o : p_B \leq 0.08$ έναντι της $H_1 : p_B > 0.08$ είναι μονόπλευρος.
Εδώ $\widehat{p_B} = \frac{17}{146+32+17} = \frac{17}{195} = \underline{\underline{0.0872}}$.

(Τρόπος $1^{ος}$) Θα κατασκευάσω δ.ε., ίσων ουρών, για το ποσοστό p_B, σ.ε. $1-2\alpha = 0.90$, λαμβάνοντας υπόψη την H_o. Απορρίπτουμε την H_o όταν το 0.08 είναι μικρότερο από το κάτω φράγμα αυτού του διαστήματος (επειδή η H_1 είναι της μορφής $0.08 < p_B$).

Επειδή $n_E = 195$ (αρκετά μεγάλο μέγεθος δείγματος) το διάστημα θα δίδεται από τον τύπο:

$$\left[\widehat{p_B} - z_\alpha\sqrt{\frac{0.08(1-0.08)}{n_E}}\,,\widehat{p_B} + z_\alpha\sqrt{\frac{0.08(1-0.08)}{n_E}}\right] \tag{2}$$

όπου $z_\alpha \overset{\text{εδώ}}{=} z_{0.05} = 1.645$, εφόσον για $Z \sim N(0,1)$ $P(Z > z_{0.05}) = 0.05 \Rightarrow \Phi(z_{0.05}) = P(Z \leq z_{0.05}) = 1 - P(Z > z_{0.05}) = 1 - 0.05 \Rightarrow \Phi(z_{0.05}) = 0.95$, όμως $\Phi(1.645) = 0.95 \Rightarrow z_{0.05} = 1.645$.

Συνεπώς το δ.ε. (4) γίνεται:

$$\left[0.0872 - 1.645\sqrt{\frac{0.0736}{195}},\ \ 0.0872 + 1.645\sqrt{\frac{0.0736}{195}}\right] = [0.0872 - 1.645 \cdot 0.01942,\ \ 0.0872 + 1.645 \cdot 0.01942] =$$
$$[0.0872 - 0.0320,\ \ 0.0872 + 0.0320] = \underline{\underline{[0.0552,\ \ 0.1192]}} \tag{3}$$

Επειδή για το κάτω φράγμα του δ.ε. (5) ισχύει, $0.0552 < 0.08$, αποδεχόμαστε την H_0 σε ε.σ. 5%.

(Τρόπος $2^{ος}$) Απορρίπτουμε την H_o όταν ισχύει

$$T = \frac{\widehat{p_B} - 0.08}{\sqrt{\frac{0.08(1-0.08)}{n_E}}} > z_\alpha.$$

Εδώ $T = \frac{0.0872-0.08}{\sqrt{\frac{0.0736}{195}}} = \frac{0.0072}{0.01942} = \underline{\underline{0.3707}} < 1.645$ (το p-value είναι 0.3554).

Οπότε, αποδεχόμαστε την H_0 σε ε.σ. 5%.

4. Ο έλεγχος $H_o : p_B = p_{BP}$ έναντι της $H_1 : p_B \neq p_{BP}$ είναι αμφίπλευρος (και σύνθετος) και ισοδύναμα γίνεται $H_o : p_B - p_{BP} = 0$ έναντι της $H_1 : p_B - p_{BP} \neq 0$.

Εδώ $\widehat{p_{BP}} = \frac{29}{100+71+29} = 0.145$.

(Τρόπος $1^{ος}$) Θα κατασκευάσω δ.ε., ίσων ουρών, για τη διαφορά $p_B - p_{BP}$, σ.ε. $1-\alpha = 0.95$.

Επειδή $n_E = 195$, $n_P = 200$ (αρκετά μεγάλα μεγέθη δειγμάτων) το διάστημα θα δίδεται από τον τύπο:

$$\left[\widehat{p_B} - \widehat{p_{BP}} - z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n_E} + \frac{\widehat{p}(1-\widehat{p})}{n_P}}\,,\widehat{p_B} - \widehat{p_{BP}} + z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n_E} + \frac{\widehat{p}(1-\widehat{p})}{n_P}}\right] \tag{4}$$

όπου p είναι (υπό την υπόθεση H_0) η κοινή πιθανότητα $p = p_B = p_{BP}$ με $\widehat{p} = \frac{17+29}{195+200} = 0.1165$,

και $z_{\frac{\alpha}{2}} \overset{\text{εδώ}}{=} z_{0.025} = 1.96$, διότι για $Z \sim N(0,1)$, $P(Z > z_{0.025}) = 0.025 \Rightarrow \Phi(z_{0.025}) = P(Z \leq z_{0.025}) = 1 - 0.025 \Rightarrow \Phi(z_{0.025}) = 0.975$, όμως $\Phi(1.96) = 0.975$.

Συνεπώς το δ.ε. (4) γίνεται:

$$\left(0.0872 - 0.145 - 1.96\sqrt{0.00053 + 0.00051},\ \ 0.0872 - 0.145 + 1.96\sqrt{0.00053 + 0.00051}\right) =$$
$$(-0.0578 - 1.96 \cdot 0.0323,\ \ -0.0578 + 1.96 \cdot 0.0323) = (-0.0578 - 0.0633,\ \ -0.0578 + 0.0633) =$$
$$\underline{\underline{(-0.1211,\ \ 0.0054)}} \tag{5}$$

Επειδή το δ.ε. (5) περιέχει το 0, <u>η H_0 γίνεται δεκτή σε ε.σ. 5%</u>.

(Τρόπος $2^{ος}$) Απορρίπτουμε την H_o όταν ισχύει

$$T = \left|\frac{\widehat{p_B} - \widehat{p_{BP}}}{\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n_E} + \frac{\widehat{p}(1-\widehat{p})}{n_P}}}\right| \geq z_{\frac{\alpha}{2}}.$$

$$\text{Εδώ} \quad T = \left|\frac{0.0872 - 0.145}{\sqrt{0.00053 + 0.00051}}\right| = \left|\frac{-0.0578}{0.0323}\right| = |-1.791| = \underline{\underline{1.791}} < 1.96 = z_{0.025},$$

όπου p είναι (υπό την υπόθεση H_0) η κοινή πιθανότητα $p = p_B = p_{BP}$ με $\widehat{p} = \frac{17+29}{195+200} = 0.1165$,

και $z_{\frac{\alpha}{2}} \overset{\text{εδώ}}{=} z_{0.025} = 1.96$, διότι για $Z \sim N(0,1)$, $P(Z > z_{0.025}) = 0.025 \Rightarrow \Phi(z_{0.025}) = P(Z \leq z_{0.025}) = 1 - 0.025 \Rightarrow$

$\Phi(z_{0.025}) = 0.975$, όμως $\Phi(1.96) = 0.975$.

Οπότε, $\underline{\underline{\text{αποδεχόμαστε την } H_0 \text{ σε ε.σ. } 5\%}}$. (το p-value είναι 0.0733)

β. Την Poisson κατανομή ακολουθούν γεγονότα που συμβαίνουν τυχαία στο χρόνο ή στο χώρο. Έτσι ένα παράδειγμα θα ήταν εάν θεωρήσουμε καφάσια (ίδιας χωρητικότητας) με ντομάτες από κάποιον προμηθευτή. Τότε, την Poisson κατανομή ακολουθεί το πλήθος από τις τομάτες, μέσα στο ίδιο καφάσι, οι οποίες είναι χαλασμένες.

Για την τ.μ. X μας δίδεται ότι $\Delta(X) = 1$, άρα $X \sim$Poisson(1) και κατά συνέπεια $E(X) = 1$. Οπότε

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

Κατά συνέπεια, $P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = e^{-1} + e^{-1} + \dfrac{e^{-1}}{2} = 2.5e^{-1} = \underline{\underline{0.9197}}$.

Επίσης, $E(5X + 1) = 5E(X) + 1 = 5 + 1 = \underline{\underline{6}}$ από την ιδιότητα της γραμμικότητας της μέσης τιμής,

και $\Delta(-3X + 2) = (-3)^2 \Delta(X) = \underline{\underline{9}}$.