Γραπτή Εξέταση Περιόδου Ιανουαρίου 2008
στο Μάθημα <u>Στατιστική</u>

27.2.2008

Α΄ ΣΕΙΡΑ ΘΕΜΑΤΩΝ

1. [20] Ο ανθρώπινος οργανισμός χρειάζεται καθημερινά από 2000 έως 6000 mg (milligrams) καλίου. Η ποσότητα καλίου που περιέχεται στα τρόφιμα ποικίλει. Έχει διαπιστωθεί πειραματικά ότι η κατανομή της ποσότητας καλίου που περιέχεται σε μια μπανάνα είναι κανονική με μέση τιμή $\mu = 630\,mg$ και τυπική απόκλιση $\sigma = 40\,mg$ ανά μπανάνα. **α)** Ποια είναι η πιθανότητα μια μπανάνα να περιέχει i) περισσότερο από 700 mg κάλιο ii) λιγότερο από 600 mg κάλιο.
β) Κάποιος καταναλώνει καθημερινά τρεις μπανάνες και έστω T η συνολική ποσότητα καλίου που παίρνει από τις τρεις μπανάνες. i) Βρείτε τη μέση τιμή και την τυπική απόκλιση της τυχαίας μεταβλητής T. ii) Ποια είναι η πιθανότητα η συνολική ποσότητα καλίου που παίρνει ο άνθρωπος αυτός καθημερινά από τις τρεις μπανάνες να είναι τουλάχιστον 2000 mg.

2. [15] Ένα σύστημα αποτελείται από 3 εξαρτήματα τύπου I και 5 εξαρτήματα τύπου II που λειτουργούν ανεξάρτητα το ένα από το άλλο. Η αξιοπιστία (πιθανότητα λειτουργίας) κάθε εξαρτήματος τύπου I είναι 0.95 ενώ η αξιοπιστία κάθε εξαρτήματος τύπου II είναι 0.90. **α)** Ποια είναι η πιθανότητα να λειτουργούν συγχρόνως τουλάχιστον 2 εξαρτήματα τύπου I. **β)** Ποια είναι η πιθανότητα να λειτουργούν συγχρόνως τουλάχιστον 4 εξαρτήματα τύπου II. **γ)** Το σύστημα λειτουργεί μόνο αν λειτουργούν συγχρόνως τουλάχιστον 2 εξαρτήματα τύπου I και τουλάχιστον 4 εξαρτήματα τύπου II. Να βρεθεί η αξιοπιστία R του συστήματος.

3. [15] α) Εξηγείστε αν δύο ξένα ενδεχόμενα A, B με $P(A) > 0$ και $P(B) > 0$ είναι ανεξάρτητα ή εξαρτημένα.
β) Ο αριθμός των επισκέψεων των γεωπόνων-ελεγκτών σε μια γεωργική μονάδα ανά μήνα είναι διακριτή τυχαία μεταβλητή X με συνάρτηση πιθανότητας:

x	0	1	2	3
$f(x)$	0.1	0.6	0.2	?

i) Βρείτε την πιθανότητα οι ελεγκτές σε ένα μήνα να επισκεφθούν τη μονάδα 3 φορές ακριβώς. ii) Βρείτε τον αναμενόμενο αριθμό επισκέψεων των ελεγκτών στην μονάδα ανά μήνα. iii) Υπολογίστε το διάστημα $\mu \pm 2\cdot\sigma$ και βρείτε την πιθανότητα η τυχαία μεταβλητή X να πάρει τιμή στο διάστημα αυτό.

4. [15] Δύο διαφορετικά κράματα μετάλλων I και II χρησιμοποιούνται για την κατασκευή εξαρτημάτων μιας οικιακής συσκευής. Εκατό δείγματα από το κάθε ένα υπόκεινται σε ένα τεστ αντοχής. Ελαττώματα παρατηρήθηκαν σε 18 από εκείνα που γίνονται με το κράμα I και σε 26 από εκείνα που γίνονται με το κράμα II.
α) Σε στάθμη σημαντικότητας 5% μπορούμε να συμπεράνουμε ότι το κράμα τύπου I είναι ανθεκτικότερο από το κράμα τύπου II;
β) Δώστε ένα 98% διάστημα εμπιστοσύνης για το ποσοστό των ελαττωμάτων που παρατηρούνται στα εξαρτήματα που κατασκευάζονται με το κράμα I.

5. [20] Το ύψος των δένδρων ενός δάσους μετριέται συνήθως από το έδαφος. Η μέθοδος αυτή είναι ακριβής αλλά δαπανηρή και για αυτό προτείνεται να υπολογίζεται το ύψος των δένδρων από αεροφωτογραφίες. Για να μελετήσουμε την ακρίβεια της μεθόδου με αεροφωτογραφίες, μετρήσαμε το ύψος 10 δένδρων και με τους δύο τρόπους και βρήκαμε τα εξής μεγέθη:

	Ύψος δένδρου (σε πόδια)									
Από το έδαφος	43	39	39	41	46	43	38	44	54	43
Από αεροφωτογραφία	37	35	34	42	39	37	35	40	46	35

α) Υπάρχει διαφορά μεταξύ των δύο μεθόδων σε επίπεδο σημαντικότητας 5%;
β) Ελέγξτε αν το μέσο ύψος των δέντρων, όταν υπολογίζεται από αεροφωτογραφίες είναι μεγαλύτερο από 37 πόδια (α=0.05).
γ) Δώστε ένα 98% διάστημα εμπιστοσύνης για το μέσο ύψος του πληθυσμού των δένδρων , όταν αυτό υπολογίζεται από αεροφωτογραφίες.

6. [15] Στον επόμενο πίνακα δίνονται οι συγκεντρώσεις κάποιου αμινοξέος (αλανίνη) (σε mg/100ml) στο αίμα τριών ειδών κάποιου οργανισμού και για τα δύο φύλα.

	Είδος B1	Είδος B2	Είδος B3
Αρσενικά A1	21.5	14.5	16.0
	19.6	17.4	20.3
	20.9	15.0	18.5
	22.8	17.8	19.3
Θηλυκά A2	14.8	12.1	14.4
	15.6	11.4	14.7
	13.5	12.7	13.8
	16.4	14.5	12.0

Αφού διατυπώσετε κατάλληλους ελέγχους υποθέσεων, ελέγξετε σε επίπεδο σημαντικότητας 5% αν υπάρχει διαφορά στη συγκέντρωση της αλανίνης που να οφείλεται στα διαφορετικά είδη, στα διαφορετικά φύλα και αν υπάρχει αλληλεπίδραση μεταξύ των παραγόντων του πειράματος.
(Δίνονται: SSA = 139, SSB = 55, SSAB = 7, SST = 239).

Δίνονται:

- Τιμές της συνάρτησης κατανομής της τυποποιημένης κανονικής κατανομής:
Φ(0.75)=0.7734, Φ(0.78)=0.7823, Φ(0.8)=0.7881, Φ(1.24)=0.8925, Φ(1.33)=0.9082, Φ(1.75)=0.9599, Φ(1.59)=0.9441, Φ(2.33)=0.9901.

- Κριτικές τιμές Z_α της τυποποιημένης κανονικής κατανομής σε επίπεδο σημαντικότητας α

$Z_{0.005}=2.57$ $Z_{0.01}=2.33$ $Z_{0.02}=2.05$ $Z_{0.025}=1.96$ $Z_{0.05}=1.64$ $Z_{0.10}=1.28$

- Κριτικές τιμές X_κ^2 (α) της X^2 κατανομής με k βαθμούς ελευθερίας και για επίπεδο σημαντικότητας α.

$X_1^2(0.05)=3.8$ $X_2^2(0.05)=6.0$ $X_3^2(0.05)=7.8$ $X_4^2(0.05)=9.5$ $X_5^2(0.05)=11.1$

- Κριτικές τιμές $t_k(\alpha)$ της t κατανομής με k βαθμούς ελευθερίας και για επίπεδο σημαντικότητας α.

$t_9(0.01)=2.82$ $t_9(0.025)=2.26$ $t_9(0.05)=1.83$
$t_{18}(0.01)=2.55$ $t_{18}(0.025)=2.10$ $t_{18}(0.05)=1.73$

- Κριτικές τιμές $F_{\mu,\nu}(\alpha)$ της F κατανομής με μ και ν βαθμούς ελευθερίας για επίπεδο σημαντικότητας α.

$F_{1,18}(0.05)=4.41$ $F_{2,18}(0.05)=3.55$ $F_{1,23}(0.05)=4.28$ $F_{2,23}(0.05)=3.42$
$F_{2,27}(0.05)=3.35$ $F_{4,27}(0.05)=2.73$ $F_{2,35}(0.05)=3.28$ $F_{4,35}(0.05)=2.65$

<u>Ενδεικτικές απαντήσεις (Α΄ Σειρά)</u>

1^o Θέμα

Έστω X η ποσότητα καλίου που περιέχεται σε μια μπανάνα. Δίνεται ότι $X \sim N(630,\ 40^2)$.

α) i) $P(X > 700) = P\left(Z > \frac{700-630}{40}\right) = P(Z > 1.75) = 1 - \Phi(1.75) = 1 - 0.9599 = 0.04$

ii) $P(X < 600) = P\left(Z < \frac{600-630}{40}\right) = P(Z < -0.75) = \Phi(-0.75) = 1 - \Phi(0.75) = 1 - 0.7734 = 0.2266$

β) i) Η μέση τιμή της T είναι $3 \cdot 630 = 1890\,mg$ και η διασπορά της είναι $3 \cdot 40^2\,mg^2$ άρα η τυπική απόκλισή της είναι $\sqrt{3 \cdot 40^2} = 69.28$. ii) Είναι γνωστό ότι $T \sim N(1890,\ 69.28^2)$ άρα

$P(T > 2000) = P\left(Z > \frac{2000-1890}{69.28}\right) = P(Z > 1.59) = 1 - \Phi(1.59) = 1 - 0.9441 = 0.0559$.

2^o Θέμα

α) Από τα 3 εξαρτήματα τύπου Ι έστω ότι X_1 από αυτά λειτουργούν συγχρόνως. Προφανώς $X_1 \sim B(3,\ 0.95)$ και επομένως,

$P(X_1 \geq 2) = P(X_1 = 2) + P(X_1 = 3) = \binom{3}{2} \cdot 0.95^2 \cdot 0.05^1 + \binom{3}{3} \cdot 0.95^3 \cdot 0.05^0 = 0.9928$.

β) Από τα 5 εξαρτήματα τύπου ΙΙ έστω ότι X_2 από αυτά λειτουργούν συγχρόνως. Προφανώς $X_2 \sim B(5,\ 0.9)$ και επομένως,

$P(X_2 \geq 4) = P(X_2 = 4) + P(X_2 = 5) = \binom{5}{4} \cdot 0.9^4 \cdot 0.1^1 + \binom{5}{5} \cdot 0.9^5 \cdot 0.1^0 = 0.9185$.

γ) Έστω A_1 το ενδεχόμενο: λειτουργούν συγχρόνως τουλάχιστον δύο εξαρτήματα τύπου Ι και A_2 το ενδεχόμενο: λειτουργούν συγχρόνως τουλάχιστον τέσσερα εξαρτήματα τύπου ΙΙ. Προφανώς $P(A_1) = P(X_1 \geq 2)$ και $P(A_2) = P(X_2 \geq 4)$.

Από τη συνθήκη λειτουργίας του συστήματος έχουμε $R = P(A_1 A_2)$ και επειδή τα A_1, A_2 είναι ανεξάρτητα $R = P(A_1)P(A_2) = 0.9928 \cdot 0.9185 = 0.9119$.

3^o Θέμα

α) Είναι προφανώς εξαρτημένα διότι $P(A/B) = P(B/A) = 0$ που σημαίνει ότι η γνώση για την πραγματοποίηση του ενός επηρεάζει την πιθανότητα πραγματοποίησης του άλλου. Μάλιστα, όταν πραγματοποιείται το ένα είμαστε **βέβαιοι** ότι δεν μπορεί να πραγματοποιηθεί το άλλο.

β) i) Πρέπει $f(0) + f(1) + f(2) + f(3) = 1 \Rightarrow f(3) = 1 - 0.9 = 0.1$. Άρα $P(X = 3) = f(3) = 0.1$. ii) $\mu = E(X) = 0 \cdot 0.1 + 1 \cdot 0.6 + 2 \cdot 0.2 + 3 \cdot 0.1 = 1.3$. iii) $E(X^2) = 0 \cdot 0.1 + 1 \cdot 0.6 + 4 \cdot 0.2 + 9 \cdot 0.1 = 2.3$. Άρα $\sigma^2 = E(X^2) - [E(X)]^2 = 2.3 - 1.3^2 = 0.61 \Rightarrow \sigma = 0.78$ και επομένως το διάστημα $\mu \pm 2 \cdot \sigma$ είναι το διάστημα $(-0.26,\ 2.86)$. Ζητείται η πιθανότητα $P(-0.26 < X < 2.86) = P(X = 0) + P(X = 1) + P(X = 2) = 0.9$.

4^o Θέμα

α) Πρόκειται για έλεγχο δύο άγνωστων ποσοστών: του ποσοστού p_1 των ελαττωματικών κραμάτων στον πληθυσμό Ι και του ποσοστού p_2 των ελαττωματικών κραμάτων στον πληθυσμό ΙΙ. Συγκεκριμένα, για τα p_1 και p_2 θα κάνουμε τον έλεγχο της υπόθεσης $H_0 : p_1 = p_2$ έναντι της εναλλακτικής $H_1 : p_1 < p_2$.

Ο έλεγχος θα γίνει με βάση τα πειραματικά δεδομένα: στον πληθυσμό Ι σε $n_1 = 100$ δοκιμές διαπιστώθηκαν $x = 18$ ελαττωματικά και στον πληθυσμό ΙΙ σε $n_2 = 100$ δοκιμές διαπιστώθηκαν $y = 26$ ελαττωματικά. Επειδή $n_1, n_2 \geq 30$, η απορριπτική περιοχή για την H_0 ορίζεται από τη σχέση: $\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}} < -Z_\alpha$, όπου $\hat{p}_1 = \frac{x}{n_1} = \frac{18}{100} = 0.18$, $\hat{p}_2 = \frac{y}{n_2} = \frac{26}{100} = 0.26$, $\hat{p} = \frac{x+y}{n_1+n_2} = \frac{44}{200} = 0.22$ και $Z_\alpha = Z_{0.05} = 1.64$.

Επειδή, $\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}} = \frac{0.18 - 0.26}{\sqrt{0.22(1-0.22)(\frac{1}{100} + \frac{1}{100})}} = -1.37 > -Z_{0.05} = -1.64$ η H_0 σε επίπεδο σημαντικότητας 5% δεν απορρίπτεται. Επομένως, με βάση τα πειραματικά δεδομένα, δεν μπορούμε να ισχυρισθούμε σε επίπεδο σημαντικότητας 5% ότι το κράμα τύπου Ι είναι ανθεκτικότερο από το κράμα τύπου ΙΙ.

β) Ζητείται 98% διάστημα εμπιστοσύνης για το άγνωστο ποσοστό p_1 των ελαττωματικών κραμάτων στον πληθυσμό Ι. Επειδή για το παρατηρούμενο στο δείγμα ποσοστό $\hat{p}_1 = 0.18$ ισχύει ότι $n_1\hat{p}_1 = 100 \cdot 0.18 = 18 \geq 5$ και $n_1(1-\hat{p}_1) = 100 \cdot 0.82 = 82 \geq 5$ το ζητούμενο 98% διάστημα εμπιστοσύνης είναι το $\hat{p}_1 \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}}$ ή $0.18 \pm z_{0.01}\sqrt{\frac{0.18(1-0.18)}{100}}$ ή $0.18 \pm 2.33\sqrt{\frac{0.18(1-0.18)}{100}}$ ή 0.18 ± 0.09.

5$^{\text{ο}}$ Θέμα

α) Πρόκειται για σύγκριση των μέσων δύο πληθυσμών από ζευγαρωτές παρατηρήσεις διότι μετράμε το ίδιο δένδρο με δύο διαφορετικούς τρόπους οπότε οι δύο μετρήσεις ανά δένδρο δεν μπορεί να θεωρηθούν ανεξάρτητες μεταξύ τους. Έτσι, παίρνουμε το δείγμα των διαφορών $z_i = x_i - y_i$: 6, 4, 5, −1, 7, 6, 3, 4, 8, 8 (όπου x_i οι μετρήσεις των υψών από το έδαφος και y_i οι μετρήσεις των υψών από αεροφωτογραφία) και κάνουμε έλεγχο για έναν πληθυσμό με ένα δείγμα, το δείγμα των διαφορών. Αν ονομάσουμε μ_1 τον μέσο του πληθυσμού των υψών που μετράμε από το έδαφος και μ_2 τον μέσο του πληθυσμού των υψών που μετράμε από αεροφωτογραφίες, θα ελέγξουμε την υπόθεση $H_0: \mu_1 - \mu_2 = 0$ έναντι της εναλλακτικής $H_1: \mu_1 - \mu_2 \neq 0$. Ή αλλιώς, αν ονομάσουμε $\mu = \mu_1 - \mu_2$ τη διαφορά των δύο μέσων, θα ελέγξουμε την υπόθεση $H_0: \mu = 0$ έναντι της εναλλακτικής $H_1: \mu \neq 0$, με βάση το δείγμα των διαφορών $z_i = x_i - y_i$: 6, 4, 5, −1, 7, 6, 3, 4, 8, 8.

Ο μέσος αυτού του δείγματος είναι $\bar{z} = \frac{6+4+5-1+7+6+3+4+8+8}{10} = 5$ και η διασπορά του

$$s_z^2 = \frac{1}{n-1}\sum_{i=1}^{n}(z_i - \bar{z})^2 = \frac{1}{9}(1^2 + 1^2 + 0^2 + 6^2 + 2^2 + 1^2 + 2^2 + 1^2 + 3^2 + 3^2) = 7.33.$$

Επειδή $n = 10 < 30$ (μικρό δείγμα) και η διασπορά σ^2 του πληθυσμού των διαφορών είναι άγνωστη, η απορριπτική περιοχή της H_0 ορίζεται από τη σχέση: $\frac{|\bar{z}|}{s_z/\sqrt{n}} > t_{n-1,a/2}$

όπου $\bar{z} = 5$, $s_z = \sqrt{7.33} = 2.71$ και $a = 0.05$.

Έτσι, επειδή $\frac{|\bar{z}|}{s_z/\sqrt{n}} = \frac{5}{2.71/\sqrt{10}} = 5.83 > t_{9,0.025} = 2.26$ η μηδενική υπόθεση H_0, σε επίπεδο σημαντικότητας 5%, απορρίπτεται και συμπεραίνουμε ότι οι παρατηρήσεις, σε επίπεδο σημαντικότητας 5%, δείχνουν να υπάρχει διαφορά μεταξύ των δύο μεθόδων μέτρησης.

β) Πρόκειται για έλεγχο του μέσου ενός πληθυσμού. Πιο συγκεκριμένα πρόκειται για τον έλεγχο της υπόθεσης $H_0 : \mu_2 = 37$ έναντι της εναλλακτικής $H_1 : \mu_2 > 37$ με βάση το δείγμα y_i : 37, 35, 34, 42, 39, 37, 35, 40, 46, 35. Ο μέσος αυτού του δείγματος είναι $\bar{y} = \frac{37 + 35 + ... + 46 + 35}{10} = 38$ και η διασπορά του

$$s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \frac{1}{9}(1^2 + 3^2 + ... + 8^2 + 3^2) = 14.44.$$

Επειδή $n = 10 < 30$ (μικρό δείγμα) και η διασπορά σ^2 του πληθυσμού είναι άγνωστη, με την υπόθεση ο πληθυσμός (των μετρήσεων των υψών από αεροφωτογραφία) είναι κανονικός, η απορριπτική περιοχή της H_0 ορίζεται από τη σχέση: $\frac{\bar{y} - \mu_0}{s_y/\sqrt{n}} > t_{n-1,\alpha}$ όπου $\bar{y} = 38$, $s_y = \sqrt{14.44} = 3.8$, $\mu_0 = 37$ και $a = 0.05$.

Έτσι, επειδή $\frac{\bar{y} - \mu_0}{s_y/\sqrt{n}} = \frac{38 - 37}{3.8/\sqrt{10}} = 0.83 < t_{9,0.05} = 1.83$ η μηδενική υπόθεση H_0, σε επίπεδο σημαντικότητας 5%, δεν απορρίπτεται και συμπεραίνουμε ότι οι παρατηρήσεις, σε επίπεδο σημαντικότητας 5%, δεν δείχνουν ότι το μέσο ύψος των δένδρων όταν η μέτρηση γίνεται από αεροφωτογραφία είναι μεγαλύτερο από 37 πόδια.

γ) Ζητείται 98% διάστημα εμπιστοσύνης για τον άγνωστο μέσο μ_2 του πληθυσμού των μετρήσεων των υψών από αεροφωτογραφίες. Επειδή, $n = 10 < 30$ (μικρό δείγμα) και η διασπορά σ^2 του πληθυσμού είναι άγνωστη, με την υπόθεση ο πληθυσμός (των μετρήσεων των υψών από αεροφωτογραφία) είναι κανονικός, το ζητούμενο 98% διάστημα εμπιστοσύνης είναι το $\bar{y} \pm \frac{s_y}{\sqrt{n}} t_{n-1,a/2}$ ή $38 \pm \frac{3.8}{\sqrt{10}} t_{9,0.01}$ ή

$38 \pm \frac{3.8}{\sqrt{10}} 2.82$ ή 38 ± 3.39.

6^{o} Θέμα

Πρόκειται για πρόβλημα ανάλυσης διασποράς για δύο παράγοντες (Α: το φύλο και Β: το είδος) με αλληλεπίδραση. Υποθέτουμε ότι οι τέσσερις μετρήσεις αλανίνης σε κάθε «κελί», δηλαδή, για κάθε συνδυασμό φύλου-είδους, είναι ένα δείγμα μεγέθους 4 από έναν πληθυσμό που ακολουθεί κανονική κατανομή. Δηλαδή υποθέτουμε ότι για κάθε συνδυασμό φύλου-είδους ο αντίστοιχος πληθυσμός με αυτό το φύλο και είδος ακολουθεί κανονική κατανομή. Υποθέτουμε επίσης ότι όλοι αυτοί οι πληθυσμοί έχουν ίδια διασπορά και ότι τα δείγματα από κελί σε κελί είναι ανεξάρτητα. Με την προϋπόθεση ότι ικανοποιούνται οι παραπάνω υποθέσεις, θα κάνουμε τους ελέγχους:

i) $H_{0a}: \alpha_1 = \alpha_2$ (ο παράγοντας A - το φύλο - δεν επιδρά στη συγκέντρωση αλανίνης, δηλαδή, η αλανίνη είναι ίδια στα δύο φύλα)

$H_{1a}: \alpha_1 \neq \alpha_2$ (ο παράγοντας A επιδρά στη συγκέντρωση αλανίνης, δηλαδή, η αλανίνη είναι διαφορετική στα δύο φύλα)

ii) $H_{0\beta}: \beta_1 = \beta_2 = \beta_3$ (ο παράγοντας B -το είδος- δεν επιδρά στη συγκέντρωση αλανίνης, δηλαδή, η αλανίνη είναι ίδια στα τρία είδη)

$H_{1\beta}: \tau\alpha\ \beta_i, i = 1,2,3$ δεν είναι όλα ίσα (ο παράγοντας B επιδρά στη συγκέντρωση αλανίνης, δηλαδή, η αλανίνη δεν είναι ίδια στα τρία είδη)

iii) $H_{0\gamma}: \gamma_{ij} = 0$ για κάθε *i, j* (δεν υπάρχουν αλληλεπιδράσεις μεταξύ των παραγόντων A και B)

$H_{1\gamma}: \gamma_{ij} \neq 0$ για κάποια *i, j* (υπάρχουν αλληλεπιδράσεις μεταξύ των παραγόντων A και B).

Ο πίνακας ανάλυσης διασποράς είναι ο παρακάτω:

Πηγή Μεταβολής	Άθροισμα Τετραγώνων	Βαθμοί Ελευθερίας	Μέσα Τετράγωνα	F
Παράγοντας Α (Φύλο)	*SSA=139*	*k-1=1*	*MSA=139/1=139*	$F_A = \frac{MSA}{MSE} = 65.88$
Παράγοντας Β (Είδος)	*SSB=55*	*λ-1=2*	*MSB=55/2=27.5*	$F_B = \frac{MSB}{MSE} = 13.03$
Αλληλεπίδραση	*SSAB=7*	*(k-1)(λ-1)=2*	*MSAB=7/2=3.5*	$F_{AB} = \frac{MSAB}{MSE} = 1.66$
Τυχαία σφάλματα	*SSE=38*	*kλ(r-1)=18*	*MSE=38/18=2.11*	
Σύνολο	*SST=239*	*kλr-1=23*		

Το *SSE* υπολογίσθηκε από τη σχέση *SSA+SSB+SSAB+SSE=SST.*

Από τις τιμές που υπολογίσθηκαν στον πίνακα, έχουμε:

i) Η μηδενική υπόθεση H_{0a} σε επίπεδο σημαντικότητας 5% απορρίπτεται διότι $F_A = 65.88 > F_{1,18,0.05} = 4.41$ και επομένως με βάση τα πειραματικά δεδομένα, σε επίπεδο σημαντικότητας 5%, το φύλο επιδρά στη συγκέντρωση αλανίνης δηλαδή τα πειραματικά δεδομένα δείχνουν ότι η συγκέντρωση αλανίνης είναι διαφορετική στα δύο φύλα.

ii) Η μηδενική υπόθεση $H_{0\beta}$ σε επίπεδο σημαντικότητας 5% απορρίπτεται διότι $F_B = 13.03 > F_{2,18,0.05} = 3.55$ και επομένως με βάση τα πειραματικά δεδομένα, σε επίπεδο σημαντικότητας 5%, το είδος επιδρά στη συγκέντρωση αλανίνης δηλαδή τα πειραματικά δεδομένα δείχνουν ότι η συγκέντρωση αλανίνης δεν είναι ίδια στα τρία είδη.

iii) Η μηδενική υπόθεση $H_{0\gamma}$ σε επίπεδο σημαντικότητας 5% δεν απορρίπτεται διότι $F_{AB} = 1.66 < F_{2,18,0.05} = 3.55$ και επομένως με βάση τα πειραματικά δεδομένα, σε επίπεδο σημαντικότητας 5%, δεν υπάρχουν αλληλεπιδράσεις μεταξύ του φύλου και του είδους.

Γραπτή Εξέταση Περιόδου Ιανουαρίου 2008
στο Μάθημα <u>Στατιστική</u>

27.2.2008

Β΄ ΣΕΙΡΑ ΘΕΜΑΤΩΝ

1. [20] Ο ανθρώπινος οργανισμός χρειάζεται καθημερινά από 2000 έως 6000 mg καλίου. Η ποσότητα καλίου που περιέχεται στα τρόφιμα ποικίλει. Έχει διαπιστωθεί πειραματικά ότι η κατανομή της ποσότητας καλίου που περιέχεται σε ένα ποτήρι χυμό πορτοκαλιού είναι κανονική με μέση τιμή $\mu = 440\,mg$ και τυπική απόκλιση $\sigma = 30\,mg$ ανά ποτήρι. **α)** Ποια είναι η πιθανότητα σε ένα ποτήρι χυμό πορτοκαλιού να περιέχεται i) περισσότερο από 510 mg κάλιο ii) λιγότερο από 400 mg κάλιο.
β) Κάποιος καταναλώνει καθημερινά τρία ποτήρια χυμό πορτοκαλιού και έστω T η συνολική ποσότητα καλίου που παίρνει από αυτά τα τρία ποτήρια χυμό. i) Βρείτε τη μέση τιμή και την τυπική απόκλιση της τυχαίας μεταβλητής T. ii) Ποια είναι η πιθανότητα η συνολική ποσότητα καλίου που παίρνει ο άνθρωπος αυτός καθημερινά από τα τρία ποτήρια χυμό πορτοκαλιού να είναι τουλάχιστον 2000 mg.

2. [15] Ένα σύστημα αποτελείται από 8 εξαρτήματα που λειτουργούν ανεξάρτητα το ένα από το άλλο. Η αξιοπιστία (πιθανότητα λειτουργίας) κάθε εξαρτήματος είναι 0.9 και το σύστημα λειτουργεί μόνο αν τουλάχιστον 6 εξαρτήματά του λειτουργούν συγχρόνως. **α)** Βρείτε την αξιοπιστία R του συστήματος. **β)** Βρείτε την πιθανότητα να υποστούν συγχρόνως βλάβη τουλάχιστον δύο εξαρτήματα δεδομένου ότι έχει υποστεί βλάβη τουλάχιστον ένα.

3. [15] α) Δώστε το νόημα της πλήρους ή τέλειας ανεξαρτησίας τριών ενδεχομένων σε αντιδιαστολή με το νόημα της ανεξαρτησίας κατά ζεύγη (δεν ζητούνται τύποι-συνθήκες αλλά το νόημά τους).
β) Ο αριθμός των επισκέψεων των γεωπόνων-ελεγκτών σε μια γεωργική μονάδα ανά μήνα είναι διακριτή τυχαία μεταβλητή X με συνάρτηση πιθανότητας:

x	0	1	2	3
$f(x)$	0.1	0.7	?	0.1

i) Βρείτε την πιθανότητα οι ελεγκτές σε ένα μήνα να επισκεφθούν τη μονάδα 2 φορές ακριβώς. ii) Βρείτε τον αναμενόμενο αριθμό επισκέψεων των ελεγκτών στην μονάδα ανά μήνα. iii) Υπολογίστε το διάστημα $\mu \pm 2 \cdot \sigma$ και βρείτε την πιθανότητα η τυχαία μεταβλητή X να πάρει τιμή στο διάστημα αυτό.

4. [15] Δύο διαφορετικά κράματα μετάλλων Ι και ΙΙ χρησιμοποιούνται για την κατασκευή εξαρτημάτων μιας οικιακής συσκευής. Εκατό δείγματα από το κάθε ένα υπόκεινται σε ένα τεστ αντοχής. Ελαττώματα παρατηρήθηκαν σε 18 από εκείνα που γίνονται με το κράμα Ι και σε 26 από εκείνα που γίνονται με το κράμα ΙΙ.
α) Σε στάθμη σημαντικότητας 5% μπορούμε να συμπεράνουμε ότι το κράμα τύπου Ι είναι ανθεκτικότερο από το κράμα τύπου ΙΙ;
β) Δώστε ένα 98% διάστημα εμπιστοσύνης για το ποσοστό των ελαττωμάτων που παρατηρούνται στα εξαρτήματα που κατασκευάζονται με το κράμα Ι.

5. [15] Στον επόμενο πίνακα δίνεται ο ρυθμός αναπνοής τριών ειδών καβουριών, σε τρεις διαφορετικές θερμοκρασίες. Αφού διατυπώσετε κατάλληλους ελέγχους υποθέσεων, ελέγξετε σε επίπεδο σημαντικότητας 5% αν υπάρχει διαφορά στο ρυθμό αναπνοής που να οφείλεται στα διαφορετικά είδη, στις διαφορετικές θερμοκρασίες και αν υπάρχει αλληλεπίδραση μεταξύ των παραγόντων του πειράματος.

	Είδη (Α)					
Θερμοκρασία(Β)	Α1		Α2		Α3	
Χαμηλή	1.9	1.6	2.1	1.8	1.1	1.0
Β1	1.8	1.4	2.0	2.2	1.2	1.4
Μεσαία	2.3	2.0	2.4	2.7	2.0	1.9
Β2	2.1	2.6	2.6	2.3	2.1	2.2
Υψηλή	2.9	3.4	3.6	3.4	2.9	3.0
Β3	2.8	3.2	3.1	3.2	2.8	3.1

(Δίνονται: SSA=1.9, SSB=13.4, SSAB=0.3, SST=16.7).

6. [20] Για να ελέγξουμε εάν το μήκος των αριστερών μπροστινών ποδιών είναι ίσο με το μήκος των αριστερών πίσω ποδιών των ελαφιών, μετρήσαμε τα αριστερά μπροστινά και πίσω πόδια 10 ελαφιών μιας περιοχής.

Ελάφι	1	2	3	4	5	6	7	8	9	10
Μήκος μπροστινών ποδιών	138	136	147	139	143	141	141	145	136	146
Μήκος πίσω ποδιών	142	140	144	144	142	146	142	150	142	148

(α) Ελέγξτε σε επίπεδο σημαντικότητας 5% αν τα αριστερά πίσω πόδια των ελαφιών της περιοχής είναι μεγαλύτερα από τα μπροστινά.
(β) Ελέγξτε αν το μήκος των αριστερών πίσω ποδιών είναι μεγαλύτερο από 142εκ (α =0.05).
(γ) Δώστε 98% διάστημα εμπιστοσύνης για το μέσο μήκος των αριστερών πίσω ποδιών των ελαφιών.

Δίνονται:

- Τιμές της συνάρτησης κατανομής της τυποποιημένης κανονικής κατανομής:
Φ(0.75)=0.7734, Φ(0.78)=0.7823, Φ(0.8)=0.7881, Φ(1.24)=0.8925, Φ(1.33)=0.9082, Φ(1.75)=0.9599, Φ(1.59)=0.9441, Φ(2.33)=0.9901.

- Κριτικές τιμές $Z_α$ της τυποποιημένης κανονικής κατανομής σε επίπεδο σημαντικότητας α
$Z_{0.005}=2.57$ $Z_{0.01}=2.33$ $Z_{0.02}=2.05$ $Z_{0.025}=1.96$ $Z_{0.05}=1.64$ $Z_{0.10}=1.28$

- Κριτικές τιμές $X^2_κ$ (α) της X^2 κατανομής με k βαθμούς ελευθερίας και για επίπεδο σημαντικότητας α.
$X_1^2(0.05)=3.8$ $X_2^2(0.05)=6.0$ $X_3^2(0.05)=7.8$ $X_4^2(0.05)=9.5$ $X_5^2(0.05)=11.1$

- Κριτικές τιμές t_k(α) της t κατανομής με k βαθμούς ελευθερίας και για επίπεδο σημαντικότητας α.
$t_9(0.01)=2.82$ $t_9(0.025)=2.26$ $t_9(0.05)=1.83$
$t_{18}(0.01)=2.55$ $t_{18}(0.025)=2.10$ $t_{18}(0.05)=1.73$

- Κριτικές τιμές $F_{μ,ν}$(α) της F κατανομής με μ και ν βαθμούς ελευθερίας για επίπεδο σημαντικότητας α.
$F_{1,18}(0.05)=4.41$ $F_{2,18}(0.05)=3.55$ $F_{1,23}(0.05)=4.28$ $F_{2,23}(0.05)=3.42$
$F_{2,27}(0.05)=3.35$ $F_{4,27}(0.05)=2.73$ $F_{2,35}(0.05)=3.28$ $F_{4,35}(0.05)=2.65$

<u>Ενδεικτικές απαντήσεις (Β΄ Σειρά)</u>

1ο Θέμα

Έστω X η ποσότητα καλίου που περιέχεται σε ένα ποτήρι χυμό πορτοκαλιού. Δίνεται ότι $X \sim N(440,\ 30^2)$.

α) i) $P(X > 510) = P\left(Z > \dfrac{510-440}{30}\right) = P(Z > 2.33) = 1 - \Phi(2.33) = 1 - 0.9901 = 0.0099$

ii) $P(X < 400) = P\left(Z < \dfrac{400-440}{30}\right) = P(Z < -1.33) = \Phi(-1.33) = 1 - \Phi(1.33) = 1 - 0.9082 = 0.0918$

β) i) Η μέση τιμή της T είναι $3 \cdot 440 = 1320\,mg$ και η διασπορά της είναι $3 \cdot 30^2\,mg^2$ άρα η τυπική απόκλισή της είναι $\sqrt{3 \cdot 30^2} = 51.96$. ii) Είναι γνωστό ότι $T \sim N(1320,\ 51.96^2)$ άρα

$P(T > 2000) = P\left(Z > \dfrac{2000-1320}{51.96}\right) = P(Z > 13.09) = 1 - \Phi(13.09) \approx 0 \cdot$

2ο Θέμα

α) Έστω X ο αριθμός των εξαρτημάτων από τα 8 που λειτουργούν συγχρόνως. Προφανώς $X \sim B(8,\ 0.9)$ και επομένως,

$$R = P(X \geq 6) = P(X=6) + P(X=7) + P(X=8) = \binom{8}{6} \cdot 0.9^6 \cdot 0.1^2 + \binom{8}{7} \cdot 0.9^7 \cdot 0.1^1 + \binom{8}{8} \cdot 0.9^8 \cdot 0.1^0$$

β) Έστω Y ο αριθμός των εξαρτημάτων από τα 8 που έχουν πάθει βλάβη συγχρόνως. Προφανώς $Y \sim B(8,\ 0.1)$. Ζητείται η πιθανότητα $P(Y \geq 2 / Y \geq 1)$. Έχουμε:

$$P(Y \geq 2 / Y \geq 1) = \frac{P(Y \geq 2,\ Y \geq 1)}{P(Y \geq 1)} = \frac{P(Y \geq 2)}{P(Y \geq 1)} = \frac{1 - P(Y=0) - P(Y=1)}{1 - P(Y=0)} = \frac{1 - \binom{8}{0} 0.1^0 0.9^8 - \binom{8}{1} 0.1^1 0.9^7}{1 - \binom{8}{0} 0.1^0 0.9^8}$$

3ο Θέμα

α) Αν τρία ενδεχόμενα είναι ανεξάρτητα κατά ζεύγη τότε η γνώση για την ταυτόχρονη πραγματοποίηση δυο εξ αυτών μπορεί να επηρεάσει την πιθανότητα πραγματοποίησης του τρίτου κάτι το οποίο δε συμβαίνει αν τα τρία ενδεχόμενα είναι πλήρως ανεξάρτητα.

β) i) Πρέπει $f(0) + f(1) + f(2) + f(3) = 1 \Rightarrow f(2) = 1 - 0.9 = 0.1$. Άρα $P(X = 2) = f(2) = 0.1$. ii) $\mu = E(X) = 0 \cdot 0.1 + 1 \cdot 0.7 + 2 \cdot 0.1 + 3 \cdot 0.1 = 1.2$. iii) $E(X^2) = 0 \cdot 0.1 + 1 \cdot 0.7 + 4 \cdot 0.1 + 9 \cdot 0.1 = 2.$. Άρα $\sigma^2 = E(X^2) - [E(X)]^2 = 2 - 1.2^2 = 0.56 \Rightarrow \sigma = 0.75$ και επομένως το διάστημα $\mu \pm 2 \cdot \sigma$ είναι το διάστημα $(-0.3,\ 2.7)$. Ζητείται η πιθανότητα $P(-0.3 < X < 2.7) = P(X=0) + P(X=1) + P(X=2) = 0.9$.

4ο Θέμα

α) Πρόκειται για έλεγχο δύο άγνωστων ποσοστών: του ποσοστού p_1 των ελαττωματικών κραμάτων στον πληθυσμό Ι και του ποσοστού p_2 των ελαττωματικών κραμάτων στον πληθυσμό ΙΙ. Συγκεκριμένα, για τα p_1 και p_2 θα κάνουμε τον έλεγχο της υπόθεσης $H_0 : p_1 = p_2$ έναντι της εναλλακτικής $H_1 : p_1 < p_2$. Ο έλεγχος θα γίνει με βάση τα πειραματικά δεδομένα: στον πληθυσμό Ι σε $n_1 = 100$ δοκιμές διαπιστώθηκαν $x = 18$ ελαττωματικά και στον πληθυσμό ΙΙ σε $n_2 = 100$ δοκιμές διαπιστώθηκαν $y = 26$ ελαττωματικά. Επειδή $n_1,\ n_2 \geq 30$, η απορριπτική περιοχή για την H_0 ορίζεται από τη σχέση: $\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1}+\frac{1}{n_2})}} < -Z_\alpha$, όπου $\hat{p}_1 = \frac{x}{n_1} = \frac{18}{100} = 0.18$, $\hat{p}_2 = \frac{y}{n_2} = \frac{26}{100} = 0.26$, $\hat{p} = \frac{x+y}{n_1+n_2} = \frac{44}{200} = 0.22$ και $Z_\alpha = Z_{0.05} = 1.64$.

Επειδή, $\frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1}+\frac{1}{n_2})}} = \frac{0.18-0.26}{\sqrt{0.22(1-0.22)(\frac{1}{100}+\frac{1}{100})}} = -1.37 > -Z_{0.05} = -1.64$ η H_0 σε επίπεδο σημαντικότητας 5% δεν απορρίπτεται. Επομένως, με βάση τα πειραματικά δεδομένα, δεν μπορούμε να ισχυρισθούμε σε επίπεδο σημαντικότητας 5% ότι το κράμα τύπου I είναι ανθεκτικότερο από το κράμα τύπου II.

β) Ζητείται 98% διάστημα εμπιστοσύνης για το άγνωστο ποσοστό p_1 των ελαττωματικών κραμάτων στον πληθυσμό I. Επειδή για το παρατηρούμενο στο δείγμα ποσοστό $\hat{p}_1 = 0.18$ ισχύει ότι $n_1\hat{p}_1 = 100 \cdot 0.18 = 18 \geq 5$ και $n_1(1-\hat{p}_1) = 100 \cdot 0.82 = 82 \geq 5$ το ζητούμενο 98% διάστημα εμπιστοσύνης είναι το $\hat{p}_1 \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}}$ ή $0.18 \pm z_{0.01}\sqrt{\frac{0.18(1-0.18)}{100}}$ ή $0.18 \pm 2.33\sqrt{\frac{0.18(1-0.18)}{100}}$ ή 0.18 ± 0.09.

5ο Θέμα

Πρόκειται για πρόβλημα ανάλυσης διασποράς για δύο παράγοντες (Α: το είδος και Β: η θερμοκρασία) με αλληλεπίδραση. Υποθέτουμε ότι οι τέσσερις μετρήσεις του ρυθμού αναπνοής σε κάθε «κελί», δηλαδή, για κάθε συνδυασμό είδους-θερμοκρασίας, είναι ένα δείγμα μεγέθους 4 από έναν πληθυσμό που ακολουθεί κανονική κατανομή. Δηλαδή υποθέτουμε ότι για κάθε συνδυασμό είδους-θερμοκρασίας ο αντίστοιχος πληθυσμός με αυτό το είδος και αυτή τη θερμοκρασία ακολουθεί κανονική κατανομή. Υποθέτουμε επίσης ότι όλοι αυτοί οι πληθυσμοί έχουν ίδια διασπορά και ότι τα δείγματα από κελί σε κελί είναι ανεξάρτητα. Με την προϋπόθεση ότι ικανοποιούνται οι παραπάνω υποθέσεις, θα κάνουμε τους ελέγχους:

i) $H_{0\alpha} : \alpha_1 = \alpha_2 = \alpha_3$ (ο παράγοντας Α -το είδος- δεν επιδρά στο ρυθμό αναπνοής, δηλαδή, ο ρυθμός αναπνοής είναι ίδιος και στα τρία είδη)

$H_{1\alpha} : \tau\alpha\ a_i, i = 1,2,3$ δεν είναι όλα ίσα (ο παράγοντας Α επιδρά στο ρυθμό αναπνοής, δηλαδή, ο ρυθμός αναπνοής δεν είναι ίδιος στα τρία είδη)

ii) $H_{0\beta} : \beta_1 = \beta_2 = \beta_3$ (ο παράγοντας Β -η θερμοκρασία- δεν επιδρά στο ρυθμό αναπνοής, δηλαδή, ο ρυθμός αναπνοής είναι ίδιος στις τρεις θερμοκρασίες)

$H_{1\beta} : \tau\alpha\ \beta_i, i = 1,2,3$ δεν είναι όλα ίσα (ο παράγοντας Β επιδρά στο ρυθμό αναπνοής, δηλαδή, ο ρυθμός αναπνοής δεν είναι ίδιος στις τρεις θερμοκρασίες)

iii) $H_{0\gamma} : \gamma_{ij} = 0$ για κάθε *i, j* (δεν υπάρχουν αλληλεπιδράσεις μεταξύ των παραγόντων Α και Β)

$H_{1\gamma} : \gamma_{ij} \neq 0$ για κάποια *i, j* (υπάρχουν αλληλεπιδράσεις μεταξύ των παραγόντων Α και Β).

Ο πίνακας ανάλυσης διασποράς είναι ο παρακάτω:

Πηγή Μεταβολής	Άθροισμα Τετραγώνων	Βαθμοί Ελευθερίας	Μέσα Τετράγωνα	F
Παράγοντας Α (Είδος)	*SSA=1.9*	*k-1=2*	*MSA=1.9/2=0.95*	$F_A = \frac{MSA}{MSE} = 23.75$
Παράγοντας Β (Θερμοκρασία)	*SSB=13.4*	*λ-1=2*	*MSB=13.4/2=6.7*	$F_B = \frac{MSB}{MSE} = 167.5$
Αλληλεπίδραση	*SSAB=0.3*	*(k-1)(λ-1)=4*	*MSAB=0.3/4=0.075*	$F_{AB} = \frac{MSAB}{MSE} = 1.875$
Τυχαία σφάλματα	*SSE=1.1*	*kλ(r-1)=27*	*MSE=1.1/27=0.04*	
Σύνολο	*SST=16.7*	*kλr-1=35*		

Το *SSE* υπολογίσθηκε από τη σχέση *SSA+SSB+SSAB+SSE=SST*.
Από τις τιμές που υπολογίσθηκαν στον πίνακα, έχουμε:

i) Η μηδενική υπόθεση $H_{0\alpha}$ σε επίπεδο σημαντικότητας 5% απορρίπτεται διότι $F_A = 23.75 > F_{2,27,0.05} = 3.35$ και επομένως με βάση τα πειραματικά δεδομένα, σε επίπεδο σημαντικότητας 5%, το είδος επιδρά στο ρυθμό αναπνοής, δηλαδή τα πειραματικά δεδομένα δείχνουν ότι ο ρυθμός αναπνοής δεν είναι ίδιος στα τρία είδη.

ii) Η μηδενική υπόθεση $H_{0\beta}$ σε επίπεδο σημαντικότητας 5% απορρίπτεται διότι $F_B = 167.5 > F_{2,27,0.05} = 3.35$ και επομένως με βάση τα πειραματικά δεδομένα, σε επίπεδο σημαντικότητας 5%, η θερμοκρασία επιδρά στο ρυθμό αναπνοής, δηλαδή τα πειραματικά δεδομένα δείχνουν ότι ο ρυθμός αναπνοής δεν είναι ίδιος στις τρεις θερμοκρασίες.

iii) Η μηδενική υπόθεση $H_{0\gamma}$ σε επίπεδο σημαντικότητας 5% δεν απορρίπτεται διότι $F_{AB} = 1.875 < F_{4,27,0.05} = 2.73$ και επομένως με βάση τα πειραματικά δεδομένα, σε επίπεδο σημαντικότητας 5%, δεν υπάρχουν αλληλεπιδράσεις μεταξύ του είδους και της θερμοκρασίας.

6ο Θέμα

α) Πρόκειται για σύγκριση των μέσων δύο πληθυσμών από ζευγαρωτές παρατηρήσεις διότι οι δύο μετρήσεις ανά ελάφι δεν μπορεί να θεωρηθούν ανεξάρτητες μεταξύ τους αφού αφορούν στο ίδιο ελάφι. Έτσι, παίρνουμε το δείγμα των διαφορών $z_i = x_i - y_i : 4, 4, -3, 5, -1, 5, 1, 5, 6, 2$ (όπου x_i οι μετρήσεις των πίσω αριστερών ποδιών και y_i οι μετρήσεις των εμπρός αριστερών ποδιών) και κάνουμε έλεγχο για έναν πληθυσμό με ένα δείγμα, το δείγμα των διαφορών. Αν ονομάσουμε μ_1 τον μέσο του πληθυσμού των μηκών των πίσω αριστερών ποδιών και μ_2 τον μέσο του πληθυσμού των μηκών των εμπρός αριστερών ποδιών, θα ελέγξουμε την υπόθεση $H_0 : \mu_1 - \mu_2 = 0$ έναντι της εναλλακτικής $H_1 : \mu_1 - \mu_2 > 0$. Ή αλλιώς, αν ονομάσουμε $\mu = \mu_1 - \mu_2$ τη διαφορά των δύο μέσων, θα ελέγξουμε την υπόθεση $H_0 : \mu = 0$ έναντι της εναλλακτικής $H_1 : \mu > 0$, με βάση το δείγμα των διαφορών $z_i = x_i - y_i : 4, 4, -3, 5, -1, 5, 1, 5, 6, 2$.

Ο μέσος αυτού του δείγματος είναι $\bar{z} = \frac{4+4+...+6+2}{10} = 2.8$ και η διασπορά του

$$s_z^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} z_i^2 - n\bar{z}^2\right) = \frac{158 - 10\cdot 2.8^2}{9} = 8.84\,.$$

Επειδή $n = 10 < 30$ (μικρό δείγμα) και η διασπορά σ^2 του πληθυσμού των διαφορών είναι άγνωστη, η απορριπτική περιοχή της H_0 ορίζεται από τη σχέση: $\dfrac{\bar{z}}{s_z/\sqrt{n}} > t_{n-1,\alpha}$ όπου $\bar{z} = 2.8$, $s_z = \sqrt{8.84} = 2.97$ και $a = 0.05$.

Έτσι, επειδή $\dfrac{\bar{z}}{s_z/\sqrt{n}} = \dfrac{2.8}{2.97/\sqrt{10}} = 2.98 > t_{9,0.05} = 1.83$ η μηδενική υπόθεση H_0, σε επίπεδο σημαντικότητας 5%, απορρίπτεται και συμπεραίνουμε ότι οι παρατηρήσεις, σε επίπεδο σημαντικότητας 5%, δείχνουν ότι τα πίσω αριστερά πόδια των ελαφιών είναι μεγαλύτερα από τα εμπρός αριστερά.

β) Πρόκειται για έλεγχο του μέσου ενός πληθυσμού. Πιο συγκεκριμένα πρόκειται για τον έλεγχο της υπόθεσης $H_0 : \mu_1 = 142$ έναντι της εναλλακτικής $H_1 : \mu_1 > 142$ με βάση το δείγμα $x_i : 142, 140, 144, ..., 142, 148$. Ο μέσος αυτού του δείγματος είναι $\bar{x} = \dfrac{142 + 140 + ... + 142 + 148}{10} = 144$ και η διασπορά του

$$s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{1}{9}(2^2 + 4^2 + ... + 2^2 + 4^2) = 9.78\,.$$

Επειδή $n = 10 < 30$ (μικρό δείγμα) και η διασπορά σ^2 του πληθυσμού είναι άγνωστη, με την υπόθεση ο πληθυσμός (των μηκών των πίσω αριστερών ποδιών) είναι κανονικός, η απορριπτική περιοχή της H_0 ορίζεται από τη σχέση: $\dfrac{\bar{x} - \mu_0}{s_x/\sqrt{n}} > t_{n-1,\alpha}$ όπου $\bar{x} = 144$, $s_x = \sqrt{9.78} = 3.13$, $\mu_0 = 142$ και $a = 0.05$.

Έτσι, επειδή $\dfrac{\bar{x} - \mu_0}{s_x/\sqrt{n}} = \dfrac{144 - 142}{3.13/\sqrt{10}} = 2.023 > t_{9,0.05} = 1.83$ η μηδενική υπόθεση H_0, σε επίπεδο σημαντικότητας 5%, απορρίπτεται και συμπεραίνουμε ότι οι παρατηρήσεις, σε επίπεδο σημαντικότητας 5%, δείχνουν ότι το μέσο ύψος των αριστερών πίσω ποδιών των ελαφιών είναι μεγαλύτερο από 142 cm.

γ) Ζητείται 98% διάστημα εμπιστοσύνης για τον άγνωστο μέσο μ_1 του πληθυσμού των μηκών των αριστερών πίσω ποδιών των ελαφιών. Επειδή, $n = 10 < 30$ (μικρό δείγμα) και η διασπορά σ^2 του πληθυσμού είναι άγνωστη, με την υπόθεση ο πληθυσμός (των μηκών των αριστερών πίσω ποδιών) είναι κανονικός, το ζητούμενο 98% διάστημα εμπιστοσύνης είναι το $\bar{x} \pm \dfrac{s_x}{\sqrt{n}} t_{n-1,\alpha/2}$ ή $144 \pm \dfrac{3.12}{\sqrt{10}} t_{9,0.01}$ ή $144 \pm \dfrac{3.13}{\sqrt{10}} 2.82$ ή 144 ± 2.79.