ΕΛΛΗΝΙΚΟ ΑΝΟΙΚΤΟ ΠΑΝΕΠΙΣΤΗΜΙΟ

ΑΚΑΔΗΜΑΪΚΟ ΕΤΟΣ 2013 -2014

ΘΕΜΑΤΙΚΗ ΕΝΟΤΗΤΑ

Βασικά Εργαλεία και Μέθοδοι για τον Έλεγχο της Ποιότητας [ΔΙΠ 50]

2η ΓΡΑΠΤΗ ΕΡΓΑΣΙΑ

Προσοχή: *Οι απαντήσεις των ασκήσεων πρέπει να φθάσουν στον Καθηγητή-Σύμβουλο ιδανικά μέχρι την Κυριακή 15/12/2013 (καταληκτική ημερομηνία) και σε καμιά περίπτωση αργότερα από την Τρίτη 17/12/2013 (τελική καταληκτική ημερομηνία (με την παράταση)).*

ΟΔΗΓΙΕΣ ΓΙΑ ΤΗΝ ΥΠΟΒΟΛΗ - ΕΚΠΟΝΗΣΗ ΕΡΓΑΣΙΩΝ ΣΤΗ ΔΙΠ 50

1. Μπορείτε να μην υποβάλλετε μόνο 1 από τις 5 εργασίες και πρέπει να συγκεντρώσετε 25 μονάδες στις 4 ή 5 εργασίες που θα υποβάλλετε για να έχετε δικαίωμα συμμετοχής στις εξετάσεις.
2. Πρέπει να τηρείτε τις προθεσμίες υποβολής των εργασιών και να ακολουθείτε πιστά τις οδηγίες που γράφονται στην αρχή κάθε άσκησης ή ερωτήματος.
3. Οι γραπτές εργασίες υποβάλλονται αποκλειστικά ηλεκτρονικά στο χώρο «ΗΛΕΚΤΡΟΝΙΚΟΙ ΧΩΡΟΙ ΕΚΠΑΙΔΕΥΤΙΚΗΣ ΔΙΑΔΙΚΑΣΙΑΣ» (http://study.eap.gr). Οι απαντήσεις των ερωτημάτων να γράφονται σε χαρτί μεγέθους Α4. Τα φύλλα της εργασίας θα πρέπει να είναι αριθμημένα και στην πρώτη σελίδα της εργασίας θα πρέπει να αναφέρεται το ονοματεπώνυμό σας. Η εργασία <u>συνιστάται</u> να είναι δακτυλογραφημένη. Αν είναι χειρόγραφη είναι <u>υποχρεωτικό</u> να είναι ευανάγνωστη. Μη γράφετε περισσότερα από αυτά που ζητούνται σε κάθε άσκηση, αφού τα επιπλέον, αν μεν είναι σωστά δεν λαμβάνονται υπ' όψιν, αν όμως είναι λάθος, επηρεάζουν αρνητικά τη βαθμολογία του θέματος.
4. Να βάζετε τις απαντήσεις σας στις ασκήσεις – υποερωτήματα με τη σειρά, όχι ανάκατα.
5. Η σειρά των θεμάτων και των ερωτημάτων στις γραπτές εργασίες δεν ακολουθεί απαραίτητα τη σειρά των περιεχομένων του εκπαιδευτικού υλικού.
6. Μην αντιγράφετε τις εκφωνήσεις! Να γράφετε μόνο τις αιτιολογημένες απαντήσεις σας βάζοντας στην αρχή την αντίστοιχη αρίθμηση της άσκησης και του ερωτήματος που απαντάτε.
7. Σε κάθε ερώτημα να δίνετε έναν (1) μόνο τρόπο λύσης.
8. Αν ένα ερώτημα ζητά σχήμα, να βάζετε το σχήμα στο τέλος της απάντησής σας <u>για το ερώτημα</u>, όχι στο τέλος της άσκησης ή της εργασίας.
9. Αν σε ένα ερώτημα χρησιμοποιείται το πακέτο MINITAB, θα πρέπει να περιέχεται στην απάντησή σας (α) πλήρης περιγραφή της διαδικασίας του MINITAB που ακολουθήσατε, (β) αντίγραφο της εκτύπωσης του session window του MINITAB και (γ) σχολιασμός ή ερμηνεία του αποτελέσματος του MINITAB. Τα (α) και (γ) πρέπει να μπαίνουν στην απάντηση του ερωτήματος, ενώ το (β) πρέπει να περιέχεται στο τέλος της άσκησης στην οποία ανήκει το ερώτημα (σε καμιά περίπτωση στο τέλος της εργασίας).
10. Δεν θα βαθμολογούνται απαντήσεις στις οποίες γίνεται χρήση άλλου στατιστικού πακέτου (εκτός του MINITAB). Επιτρέπεται η χρήση του Excel για τη διενέργεια <u>μόνον</u> απλών αριθμητικών υπολογισμών.
11. Στα ερωτήματα που χρειάζεται να χρησιμοποιήσετε μια σχέση (έναν τύπο), πρέπει να γράφετε πρώτα τη λογική ή το επιχείρημα για την επιλογή της σχέσης, ακολούθως να γράφετε τη γενική μορφή της σχέσης, μετά να κάνετε αντικατάσταση των τιμών και, τέλος, να βρίσκετε το αποτέλεσμα.
12. Στις προτάσεις Σωστού-Λάθους πρέπει απαραίτητα να γράφετε την επιλογή σας (Σωστή ή Λάθος) και στις προτάσεις που δίνετε απάντηση «Λάθος» να δίνετε σαφή αιτιολόγηση, η οποία να περιέχει <u>εντοπισμό του λάθους</u>, και όχι να παραπέμπετε σε ολόκληρες παραγράφους ή σελίδες του εκπαιδευτικού υλικού. Ομοίως, στα ερωτήματα πολλαπλής επιλογής, πρέπει απαραίτητα να επιλέξετε τη σωστή επιλογή, π.χ. «η (i) είναι η Σωστή» και να αιτιολογείτε με σαφήνεια την απάντησή σας. Αν δεν υπάρχει ξεκάθαρη επιλογή σε ένα ερώτημα Σωστού-Λάθους (Σωστή ή Λάθος) ή σε ένα ερώτημα πολλαπλής επιλογής (ποια επιλογή είναι η Σωστή) το ερώτημα θα μηδενίζεται.
13. Μπορείτε να ανταλλάσσετε απόψεις για τη σωστή απάντηση των ασκήσεων, αλλά <u>δεν επιτρέπονται σε καμιά περίπτωση αντιγραφές</u>. Εάν υποβληθούν από δύο ή περισσότερους φοιτητές πανομοιότυπες απαντήσεις σε μια ή περισσότερες ασκήσεις (έστω και με διακοσμητικές αλλαγές για να φαίνονται δήθεν διαφορετικές), <u>θα μηδενίζονται οι εργασίες όλων των εμπλεκόμενων φοιτητών</u> (απόφαση της Ομάδας Διδακτικού Προσωπικού της ΔΙΠ 50).

Βαθμολόγηση: *Οι μονάδες που αντιστοιχούν σε κάθε άσκηση και σε κάθε ερώτημα χωριστά δίνονται μέσα σε παρένθεση (σύνολο 100 μονάδες).*

ΑΣΚΗΣΕΙΣ 2^{ης} ΓΡΑΠΤΗΣ ΕΡΓΑΣΙΑΣ

Άσκηση 1 (12 μονάδες)

Δώστε την κατάλληλη απάντηση (ΣΩΣΤΗ ή ΛΑΘΟΣ) στις προτάσεις (γ), (δ), (ε), (στ), (ζ), και στις τελευταίες προτάσεις των (α), (β), (η), (θ), (ι), (κ), (λ). Αιτιολογήστε σύντομα μόνο τις απαντήσεις στις οποίες επιλέξατε ΛΑΘΟΣ.

(α-1) Έστω $X \sim B(n, p)$. Το σύνολο $\{1, 2, \ldots, n\}$ περιλαμβάνει όλες τις δυνατές τιμές που μπορεί να πάρει η τ.μ. *X*.

(β-1) Ο αριθμός των ατυχημάτων που συμβαίνουν κάθε Κυριακή σε μία διασταύρωση έχει την κατανομή Poisson με μέση τιμή 0.2. Η πιθανότητα να μην συμβεί ατύχημα σε δύο διαδοχικές Κυριακές στη συγκεκριμένη διασταύρωση είναι ίση με $e^{-0.2}$.

(γ-1) Η τυχαία μεταβλητή *X* που μετράει το πλήθος των βολών που ρίχνει ένας καλαθοσφαιριστής μέχρις ότου βάλει 5 καλάθια έχει τη γεωμετρική κατανομή.

(δ-1) Στο θηκόγραμμα απεικονίζονται με οριζόντιες γραμμές από κάτω προς τα πάνω, το πρώτο τεταρτημόριο, η μέση τιμή και το τρίτο τεταρτημόριο των δεδομένων μας.

(ε-1) Το καταλληλότερο μέτρο προσδιορισμού της κεντρικής τάσης ενός συνόλου δεδομένων με ακραίες τιμές είναι η διάμεσος.

(στ-1) Η εκθετική κατανομή είναι η μόνη κατανομή που έχει την ιδιότητα της αμνησίας.

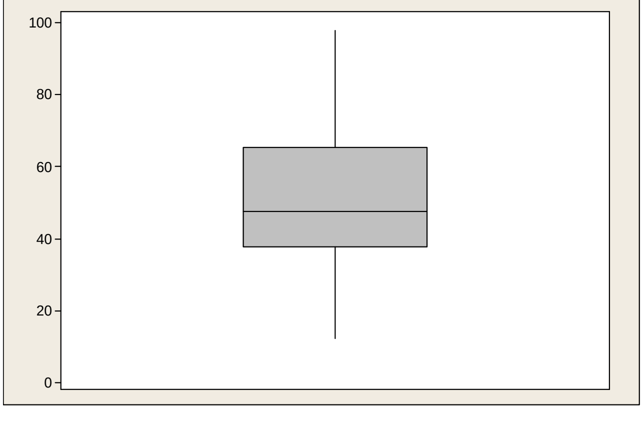
(ζ-1) Για το 0.80 ποσοστιαίο σημείο της κατανομής $N(0,2)$ ισχύει ότι $x_{0.80} = 0.84$.

(η-1) Έστω $\{N(t), t \geq 0\}$ μια διαδικασία Poisson με μέσο ρυθμό *ν*. Η πιθανότητα να συμβούν ακριβώς 2 συμβάντα στο χρονικό διάστημα $(2,4]$ είναι ίση με $\dfrac{2\nu^2}{e^{2\nu}}$.

(θ-1) Ένα τυχαίο ζεύγος (X, Y) έχει τη διδιάστατη κανονική κατανομή με παραμέτρους $\mu_X, \mu_Y, \sigma_X, \sigma_Y, \rho$. Αν οι τ.μ. *X*, *Y* είναι ανεξάρτητες, τότε η κοινή συνάρτηση πυκνότητας πιθανότητας του ζεύγους δίνεται από τον τύπο

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y}\exp\left(-\left[\frac{(x-\mu_X)^2}{2\sigma_X^2} + \frac{(y-\mu_Y)^2}{2\sigma_Y^2}\right]\right), \quad (x,y) \in R^2.$$

(ι-1) Το παρακάτω θηκόγραμμα δίνει μια απεικόνιση της κατανομής της ηλικίας ενός δείγματος 50 κατοίκων μιας περιοχής. Από το θηκόγραμμα συμπεραίνουμε ότι η κατανομή είναι λοξή προς τα αριστερά.



(κ-1) Για $i = 1, 2, \ldots, n$, έστω $(x_i - \bar{x})$ οι αποκλίσεις των τιμών ενός δείγματος μεγέθους *n* από τη δειγματική μέση τιμή τους. Τότε $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$.

(λ-1) Δίνεται το παρακάτω διάγραμμα στελέχους-φύλλου που αφορά τις ηλικίες ενός δείγματος 30 κατοίκων μιας περιοχής. Συμπεραίνουμε ότι στο δείγμα υπάρχουν 7 κάτοικοι με ηλικία μεγαλύτερη των 40 ετών και μικρότερη των 50 ετών.

Stem-and-Leaf Display: Ηλικία

```
Stem-and-leaf of Ηλικία N = 30
Leaf Unit = 1,0

 3   1  256
 7   2  3356
 12  3  23689
(7)  4  1225568
 11  5  0016
 7   6  0669
 3   7  5
 2   8  4
 1   9  0
```

Άσκηση 2 (13 μονάδες)

Ο δείκτης νοημοσύνης (IQ) των φοιτητών του ΕΑΠ έχει την κανονική κατανομή μέση τιμή 110 και τυπική απόκλιση 15.

(α-3) Αν επιλέξουμε τυχαία έναν φοιτητή, ποια είναι η πιθανότητα το IQ του να βρίσκεται στο διάστημα από 95 έως 125; Απαντήστε χωρίς MINITAB σε αυτό το ερώτημα.

(β-3) Ποια είναι η τιμή του IQ πάνω από την οποία κατανέμεται το 5% των φοιτητών; Απαντήστε χωρίς MINITAB σε αυτό το ερώτημα.

(γ-3) Επαληθεύστε με τη χρήση MINITAB τις απαντήσεις που δώσατε στα ερωτήματα (α) και (β).

(δ-4) Αν επιλέξουμε τυχαία 5 φοιτητές, ποια είναι η πιθανότητα τουλάχιστον 4 από αυτούς να έχουν IQ μεγαλύτερο από 125;

Άσκηση 3 (15 μονάδες)

Το 10% των συγκολλήσεων που γίνονται σε ένα μεγάλο συνεργείο είναι ελαττωματικές.

(α-3) Αν επιλέξουμε στην τύχη 18 συγκολλήσεις, ποια είναι η πιθανότητα το πολύ 4 συγκολλήσεις να είναι ελαττωματικές; Απαντήστε χωρίς MINITAB σε αυτό το ερώτημα.

(β-2) Ποια είναι η μέση τιμή και η διασπορά του αριθμού των ελαττωματικών συγκολλήσεων του ερωτήματος (α);

(γ-3) Εάν ελέγχουμε τη μία κατόπιν της άλλης όλες τις συγκολλήσεις που γίνονται στο συνεργείο, ποια είναι η πιθανότητα η δεύτερη ελαττωματική συγκόλληση να βρεθεί στο δωδέκατο κατά σειρά έλεγχο που θα κάνουμε; Απαντήστε χωρίς MINITAB σε αυτό το ερώτημα.

(δ-4) Θεωρείστε ότι το ποσοστό των ελαττωματικών συγκολλήσεων που γίνονται στο συνεργείο είναι 1%. Να υπολογιστεί προσεγγιστικά (χρησιμοποιώντας δηλαδή κατάλληλη προσεγγιστική κατανομή) η πιθανότητα να βρεθούν το πολύ 4 ελαττωματικές συγκολλήσεις σε ένα δείγμα 300 τυχαία επιλεγμένων συγκολλήσεων. Απαντήστε χωρίς MINITAB σε αυτό το ερώτημα.

(ε-3) Επιβεβαιώσετε με τη χρήση MINITAB τις απαντήσεις που δώσατε στα ερωτήματα (α), (γ) και (δ).

Άσκηση 4 (16 μονάδες)

Μια ηλεκτρική συσκευή, που θέλουμε να λειτουργεί συνεχώς (365 ημέρες τον χρόνο, 24 ώρες την μέρα), παρουσιάζει βλάβες που συμβαίνουν σύμφωνα με μια διαδικασία Poisson με μέσο ρυθμό 0.0002 βλάβες την ώρα.

(α-2) Ποια είναι η πιθανότητα να μην παρουσιάσει βλάβη η συσκευή κατά το πρώτο χρόνο λειτουργίας της; Απαντήστε χωρίς MINITAB σε αυτό το ερώτημα.

(β-2) Ποια είναι η πιθανότητα να παρουσιαστεί το πολύ μια βλάβη στη συσκευή κατά το πρώτο χρόνο λειτουργίας της; Απαντήστε χωρίς MINITAB σε αυτό το ερώτημα.

(γ-4) Ποια είναι η πιθανότητα ο χρόνος μεταξύ δύο διαδοχικών βλαβών της συσκευής να είναι μεγαλύτερος του ενός έτους και μικρότερος των δύο ετών; Απαντήστε χωρίς MINITAB σε αυτό το ερώτημα.

(δ-3) Επιβεβαιώσετε με τη χρήση MINITAB τις απαντήσεις που δώσατε στα ερωτήματα (α), (β) και (γ).

(ε-5) Υποθέστε ότι η ηλεκτρική συσκευή συνοδεύεται από εγγύηση η οποία προβλέπει ότι θα αποκαθίστανται δωρεάν μόνο οι πρώτες n ($n \geq 1$) βλάβες που συμβαίνουν εντός του πρώτου έτους λειτουργίας της συσκευής. Να βρεθεί ο ελάχιστος αριθμός των βλαβών που πρέπει να αποκαθίστανται δωρεάν κατά τον πρώτο χρόνο λειτουργίας της συσκευής, έτσι ώστε η πιθανότητα να χρειαστεί να πληρώσουμε έξοδα επισκευής να είναι το πολύ 5%.

Άσκηση 5 (16 μονάδες)

Θέλουμε να δημιουργήσουμε μια συμβουλευτική τεχνική για να μειώνουμε το άγχος των μαθητών που πρόκειται να λάβουν μέρος στις πανελλήνιες εξετάσεις. Επιλέγουμε τυχαία δύο δείγματα τέτοιων μαθητών από τα λύκεια μιας ευρείας περιοχής. Στην 1η ομάδα κάνουμε *Ομαδικές Συμβουλευτικές Συναντήσεις* και τους δίνουμε να διαβάσουν υλικό σχετικό με το άγχος. Στην 2η ομάδα κάνουμε *Ατομικές Συμβουλευτικές Συναντήσεις* και τους δίνουμε να διαβάσουν και αυτοί υλικό σχετικό με το άγχος. Στο τέλος της «θεραπευτικής» περιόδου, μετράμε το επίπεδο του άγχους στα μέλη των δύο ομάδων (0: μηδενικό άγχος – 50: απόλυτο άγχος). Οι μετρήσεις δίνονται στον παρακάτω πίνακα.

Ομάδα 1η							
23	17	29	29	26	24	27	33
25	14	21	26	20	27	26	32
20	32	17	23	20	30	26	12

Ομάδα 2η							
21	26	28	31	14	27	29	23
18	25	32	23	16	21	17	20
26	23	7	18	29	32	24	19

(α-4) Κατασκευάστε το διάγραμμα ποσοστιαίας για καθένα από τα δείγματα δεδομένων. Σχολιάστε τη μορφή των διαγραμμάτων.

(β-4) Κατασκευάστε τα θηκογράμματα των δύο δειγμάτων στο ίδιο διάγραμμα και συγκρίνετε τις αντίστοιχες κατανομές.

(γ-4) Συγκρίνετε τις κατανομές των δύο δειγμάτων με χρήση του διαγράμματος ποσοστιαίας-ποσοστιαίας.

(δ-4) Χρησιμοποιείστε κατάλληλο θεωρητικό διάγραμμα Q-Q για να εξετάσετε εάν η κανονική κατανομή προσεγγίζει ικανοποιητικά την κατανομή του δείγματος της 2ης Ομάδας.

Άσκηση 6 (15 μονάδες)

Επιλέξτε τη σωστή συμπλήρωση (i), (ii), (iii) ή (iv) στα ερωτήματα (α) – (ε) παρακάτω, αιτιολογώντας την απάντησή σας.

(α-2.5) Τρεις ίδιες μηχανές λειτουργούν ανεξάρτητα η μια από την άλλη. Η πιθανότητα να πάθει βλάβη μια μηχανή μέσα σε ένα έτος ισούται με p. Όταν χαλάσει μια μηχανή δεν επισκευάζεται. Η πιθανότητα να μην παρατηρηθεί πάνω από μια βλάβη συνολικά και στις τρεις μηχανές μέσα σε ένα έτος είναι ίση με

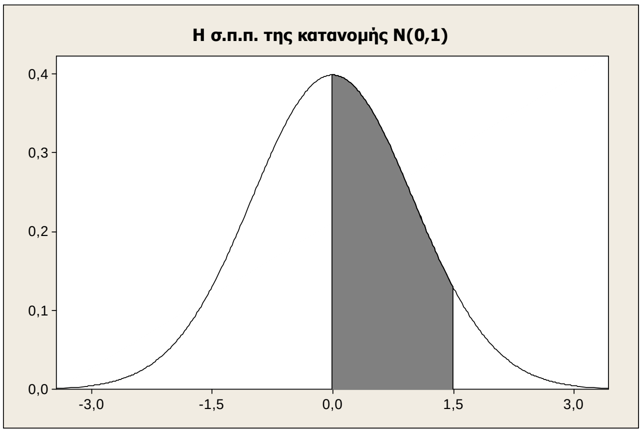
(i) $p(1-p)^2$.

(ii) $(1-p)^3 + 3p(1-p)^2$.

(iii) $(1-p)^3 + p(1-p)^2$.

(iv) $3(1-p) + p(1-p)^2$.

(β-2.5) Το παρακάτω γράφημα αναπαριστά τη σ.π.π. της τυπικής κανονικής κατανομής.



Το εμβαδό του σκιασμένου χωρίου είναι ίσο με

(i) $0.5 - P(Z < -1.5)$.

(ii) $0.5 - P(Z < 1.5)$.

(iii) $P(Z < -1.5)$.

(iv) $P(Z < 1.5)$.

(γ-2.5) Δίνεται το παρακάτω διάγραμμα στελέχους-φύλλου.

Stem-and-Leaf Display: C1

```
Stem-and-leaf of C1  N  = 50
Leaf Unit = 1,0

 1   -0  3
 7   -0  111000
 13  0   000011
 18  0   22223
(9)  0   444455555
 23  0   66666666677
 12  0   88999
 7   1   0011
 3   1   22
 1   1   4
```

Ποια από τις παρακάτω προτάσεις **<u>δεν ισχύει</u>**;

(i) Η διάμεσος των δεδομένων βρίσκεται στο 5^{ο} στέλεχος.

(ii) Τα δεδομένα έχουν χωριστεί σε 10 στελέχη.

(iii) Το 5^{ο} στέλεχος περιέχει 18 δεδομένα.

(iv) Τα τελευταία 3 στελέχη περιέχουν 7 δεδομένα.

(δ-2.5) Θεωρούμε ότι σε ένα εργοστάσιο παραγωγής λαμπτήρων η πιθανότητα να παραχθεί ελαττωματικός λαμπτήρας είναι 0.01. Έστω X η τυχαία μεταβλητή που δηλώνει το πλήθος των ελαττωματικών λαμπτήρων σε μια παρτίδα 100 λαμπτήρων. Η κατανομή της τ.μ. X προσεγγίζεται ικανοποιητικά από την

(i) αρνητική διωνυμική κατανομή.

(ii) κατανομή Poisson.

(iii) εκθετική κατανομή.

(iv) ομοιόμορφη κατανομή.

(ε-2.5) Μία μπάλα μπιλιάρδου που κυλάει πάνω σε μία γραμμή μήκους ενός μέτρου σταματά στο σημείο y $(0 < y < 1)$. Η τ.μ. Y που δηλώνει το σημείο που σταματά η μπάλα έχει την ομοιόμορφη κατανομή στο διάστημα $(0,1)$. Η ίδια μπάλα κυλάει n φορές κάτω από τις ίδιες ακριβώς συνθήκες όπως και την πρώτη φορά. Η τ.μ. X που δηλώνει το πλήθος των φορών που σταματά η μπάλα πριν από το σημείο y έχει την κατανομή

(i) $\mathcal{P}(ny)$.

(ii) $U(0, y)$.

(iii) $B(n, y)$.

(iv) $NB(n, y)$.

(στ-2.5) Το πιθανότερο πρότυπο κατανομής που περιγράφει τα δεδομένα των οποίων ορισμένα περιγραφικά στατιστικά μέτρα δίνονται στον παρακάτω πίνακα

Descriptive Statistics: C1

```
Variable    N  N*    Mean  SE Mean   StDev  Minimum      Q1  Median      Q3
C1        500   0  1,0079   0,0458  1,0248   0,0001  0,2917  0,6604  1,4018

Variable  Maximum  Skewness  Kurtosis
C1         5,9548      1,67      2,84
```

είναι η

(i) κανονική κατανομή.

(ii) ομοιόμορφη κατανομή.

(iii) κατανομή $WEI(5,1)$.

(iv) εκθετική κατανομή.

<u>Άσκηση 7 (13 μονάδες)</u>

Για να μελετηθούν τα εβδομαδιαία έξοδα μετακίνησης (σε ευρώ) των μελών Δ.Ε.Π. ενός μεγάλου ελληνικού πανεπιστημίου επιλέχθηκαν τυχαία 52 μέλη του. Τα εβδομαδιαία έξοδα μετακίνησής τους δίνονται στον ακόλουθο πίνακα.

34	39	38	31
35	35	37	37
34	40	32	40
43	35	36	21
23	24	31	33
27	31	30	33
38	47	35	32
30	28	46	32
53	29	33	28

40	34	37	29
43	26	37	22
44	32	33	39
26	42	36	25

(α-4) Υπολογίστε τη δειγματική μέση τιμή, τη δειγματική διάμεσο, το πρώτο τεταρτημόριο, το τρίτο τεταρτημόριο, τη δειγματική τυπική απόκλιση, τη δειγματική έκταση και το δειγματικό συντελεστή λοξότητας. Να σχολιαστεί ο δειγματικός συντελεστής λοξότητας.

(β-2) Κατασκευάστε ένα ιστόγραμμα σχετικών συχνοτήτων χρησιμοποιώντας 7 κλάσεις δεδομένων. Σχολιάστε τη λοξότητα της κατανομής των δεδομένων και αναφέρετε αν υπάρχει συμφωνία με τα ευρήματα του ερωτήματος (α).

(γ-2) Κατασκευάστε το διάγραμμα στελέχους-φύλλου των δεδομένων και σχολιάστε τη μορφή του.

(δ-3) Κατασκευάστε το θηκόγραμμα των δεδομένων και σχολιάστε τη μορφή του.

(ε-2) Διερευνήστε την καταλληλότητα της κανονικής κατανομής για την περιγραφή των δεδομένων με χρήση ιδιοτήτων της κανονικής κατανομής.

Κ Α Λ Η   Ε Π Ι Τ Υ Χ Ι Α !

ΑΠΑΝΤΗΣΕΙΣ 2ης ΕΡΓΑΣΙΑΣ
ΕΑΠ ΔΙΠ 50 2013-2014

Άσκηση 1

(α) Λ το σύνολο είναι $\{0,1,2,\dots,n\}$

(β) Λ $\lambda=0{,}2$ $P(0)=e^{-0{,}2}\frac{0{,}2^0}{0!}=e^{-0{,}2}$, $P(0)\cdot P(0)=e^{-0{,}2}\cdot e^{-0{,}2}=e^{-0{,}4}$

(γ) Λ, είναι η αρνητική Διωνυμική

(δ) Λ, όχι την μέση τιμή αλλά διάμεσο

(ε) Σ

(στ) Λ, αφού η Weibull για $a=1$ συμπίπτει με την εκθετική $E(1/B)$

(ζ) Λ, Για να βρούμε το σημείο $X_{0.80}$ της κατανομής $N(0,2)$ βρίσκουμε πρώτα το 0,80 ποσοστιαίο σημείο της κατανομής $N(0,1)$. Η πλησιέστερη τιμή είναι 0,7995. Αυτή αντιστοιχεί σε $Z=0{,}84$. Άρα $X_{0.80}=\mu+\sigma Z_{0.80}$

$\Leftrightarrow X_{0.80}=0+2\cdot 0{,}84 \Leftrightarrow X_{0.80}=1{,}68$

(η) Σ, Έχω $(2,4]$ άρα $s=2$ και $t=2$ και $\lambda=2v$

άρα $P[N(4)-N(2)=2]=\dfrac{e^{-2v}(2v)^2}{2!}=\dfrac{e^{-2v}\cdot 4\cdot v^2}{2}=\dfrac{2v^2}{e^{2v}}$

(θ) Σ, $f(x,y)=\dfrac{1}{2\pi(1-\rho^2)^{1/2}\sigma_X\sigma_Y}\exp\left\{-\dfrac{1}{2(1-\rho^2)}\left[\left(\dfrac{x-\mu_x}{\sigma_x}\right)^2-2\rho\left(\dfrac{x-\mu_x}{\sigma_x}\right)\left(\dfrac{y-\mu_y}{\sigma_y}\right)+\left(\dfrac{y-\mu_y}{\sigma_y}\right)^2\right]\right\}$

Αφού οι X,Y είναι ανεξάρτητες $\rho=0$

$$f(x,y)=\frac{1}{2\pi\sigma_X\sigma_Y}\exp\left\{-\left[\frac{(x-\mu_x)^2}{2\sigma_X^2}+\frac{(y-\mu_y)^2}{2\sigma_Y^2}\right]\right\}$$

(ι) Λ, Η κατανομή είναι λοξή προς τα δεξιά

(κ) Σ, $\sum_{i=1}^{n}(x_i-\bar{x}) = \sum_{i=1}^{n}x_i - \sum_{i=1}^{n}\bar{x} = n\bar{x} - n\bar{x} = 0$

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n}x_i \Leftrightarrow \sum_{i=1}^{n}x_i = n\bar{x}$

(λ) Σ

Άσκηση 2

(α) $X \sim N(110, 15^2)$

$$P(95 \le X \le 125) = P\left(\frac{95-110}{15} \le Z \le \frac{125-110}{15}\right)$$

$$= P(-1 \le Z \le 1) = \Phi(1) - \Phi(-1) = \Phi(1) - P(Z>1)$$

$$= \Phi(1) - [1-\Phi(1)] = \Phi(1) - 1 + \Phi(1) = 2\Phi(1) - 1$$

$$= 2 \cdot 0{,}8413 - 1 = 0{,}6826$$

(β) $P(X>c) = 0{,}05 \Leftrightarrow 1 - P(X \le c) = 0{,}05$

$$\Leftrightarrow P(X \le c) = 0{,}95 \Leftrightarrow P\left(\frac{X-\mu}{\sigma} \le \frac{c-\mu}{\sigma}\right) = 0{,}95$$

$$\Leftrightarrow P\left(Z \le \frac{c-110}{15}\right) = 0.95 \Leftrightarrow \Phi\left(\frac{c-110}{15}\right) = 0.95$$

$$\Leftrightarrow \frac{c-110}{15} = 1{,}65 \Leftrightarrow c - 110 = 24{,}75 \Leftrightarrow c = 134{,}75$$

(γ) Minitab

(δ) $p = P(X>125) = P\left(\frac{X-\mu}{\sigma} > \frac{125-110}{15}\right) = P(Z>1)$

$$= 1 - \Phi(1) = 1 - 0{,}8413 \Leftrightarrow p = 0.1587$$

Άρα $B(15, 0.1587)$

Θέλω $P(X \geq 4) = 1 - P(X \leq 3)$

$$P(X \leq 3) = P(0) + P(1) + P(2) + P(3)$$

$$P(0) = \binom{5}{0} 0.1587^0 \cdot 0.8413^5 = 0.421458$$

$$P(1) = \binom{5}{1} 0.1587^1 \cdot 0.8413^4 = 0.397512$$

$$P(2) = \binom{5}{2} 0.1587^2 \cdot 0.8413 = 0.149971$$

$$P(3) = \binom{5}{3} 0.1587^3 \cdot 0.8413^2 = 0.02829$$

$P(4) = 0{,}002668$, $P(5) = 0{,}000101$

Άρα $P(X \geq 4) = 1 - (P(0) + P(1) + P(2) + P(3))$

$= 1 - 0.997231 = 0{,}002769$

<u>$2^{\text{ος}}$ τρόπος</u>

$$P(X \geq 4) = P(4) + P(5) = 0.002769$$

$\boxed{\text{Άσκηση 3}}$

(a) Έχουμε $B(18, 0.1)$

$$P(X \leq 4) = P(0) + P(1) + P(2) + P(3) + P(4)$$

$$= \binom{18}{0} 0.1^0 \cdot 0.9^{18} + \binom{18}{1} 0.1^1 \cdot 0.9^{17} + \binom{18}{2} 0.1^2 \cdot 0.9^{16}$$

$$+ \binom{18}{3} 0.1^3 \cdot 0.9^{15} + \binom{18}{4} 0.1^4 \cdot 0.9^{14}$$

$$= 0.150095 + 0.300189 + 0.283512 + 0.168067$$

$$+ 0.0700029 = 0.971806$$

(β) $E(X) = np = 18 \cdot 0,1 = 1.8$

$Var(X) = npq = 18 \cdot 0.1 \cdot 0,9 = 1,62$

(γ) Αρνητική Διωνυμική $X \sim NB(r,p)$ $x=12$ και $r=2$

και $p = 0.1$

Άρα $P(X=12) = \binom{12-1}{2-1} \cdot 0.1^2 \cdot 0.9^{10} = \binom{11}{1} 0,01 \cdot 0,3487$

$= 11 \cdot 0,01 \cdot 0,3487 = 0,038357$

(δ) $p = 0.01$ $n = 300$ αφού $np = 3 < 5$

Η κατανομή που θα το προσεγγίσουμε είναι η Poisson

με $\lambda = 3$.

$P(X \leq 4) = P(0) + P(1) + P(2) + P(3) + P(4)$

$$= e^{-3}\frac{3^0}{0!} + e^{-3}\frac{3^1}{1!} + e^{-3}\frac{3^2}{2!} + e^{-3}\frac{3^3}{3!} + e^{-3}\frac{3^4}{4!}$$

$= 0.049787 + 0.149361 + 0.224042 \cdot 2 + 0.168031$

$= 0.815263$

(ε) Minitab

Άσκηση 4

(α) $\lambda = 365 \cdot 24 \cdot 0{,}0002 = 1{,}752$

$$P(X=0) = e^{-1.752}\frac{1.752^0}{0!} = 0.1734$$

(β) $$P(X\le 1) = P(0)+P(1) = e^{-1.752}\frac{1.752^0}{0!} + e^{-1.752}\frac{1.752^1}{1!}$$

$$= 0.1734 + 0.30695 = 0.48035$$

(γ) Έστω T_1 ο χρόνος (σε χρόνια) από σήμερα μέχρι την εμφάνιση της πρώτης βλάβης και T_2 ο χρόνος (σε χρόνια) που μεσολαβεί μεταξύ της $1^{ης}$ και της $2^{ης}$ από σήμερα βλάβης. Από το Θεώρημα 4.3 γνωρίζουμε ότι οι τ.μ T_1 και T_2 είναι ανεξάρτητες και έχουν την κατανομή $P(v)$, όπου $v = 1{,}752$

$$P(T_1 \ge 1, T_2 < 2) = P(T_1 \ge 1)\, P(T_2 < 2)$$

$$= e^{-1{,}752}\cdot\left(1-e^{-3{,}504}\right) = 0{,}1734\cdot(1-0{,}030)$$

$$= 0{,}1734 \cdot 0{,}97 = 0{,}1682$$ ($2^{ος}$ τρόπος) (Σελ 16)

(δ) Minitab

(ε) Έστω $X(t)$ ο αριθμός των ζημιών που συμβαίνουν στην ηλεκτρική συσκευή σε ένα χρόνο

Η $\{X(t), t \ge 0\}$ είναι μια διαδικασία Poisson με ρυθμό $\lambda_p = 1{,}752$

Άρα θέλω μέχρι n βλάβες σε 1 έτος έχω:

$$P(X(1)=n) = e^{-1,752\cdot 1}\,\frac{(1,752\cdot 1)^n}{n!}$$

Και ισχύει $P(X(1)=n) \le 0,05$

$$\Leftrightarrow e^{-1,752}\cdot\frac{(1,752)^n}{n!} \le 0,05$$

$$\Leftrightarrow \frac{(1,752)^n}{n!} \le 0,05\cdot e^{1,752} \Leftrightarrow \frac{(1,752)^n}{n!} \le 0,2883$$

- Για $n=1$ $\quad 1,752 \not\le 0,2883$
- Για $n=2$ $\quad 1,5348 \not\le 0,2883$
- Για $n=3$ $\quad 0,8963 \not\le 0,2883$
- Για $n=4$ $\quad 0,3926 \not\le 0,2883$
- Για $n=5$ $\quad 0,1376 \le 0,2883$

Άρα η εταιρία αποζημιώνει μέχρι το πολύ 5 βλάβες

2ος τρόπος

Θέλω $P(X \ge n) = R(n) = \exp(-\lambda\cdot n)$, $\frac{1}{\lambda} = 1,752$

$\Leftrightarrow \lambda = 0,5708$

$R(1) = e^{-0,5708\cdot 1} = 0,565087 > 0,05$

$R(2) = e^{-0,5708\cdot 2} = 0,3193 > 0,05$

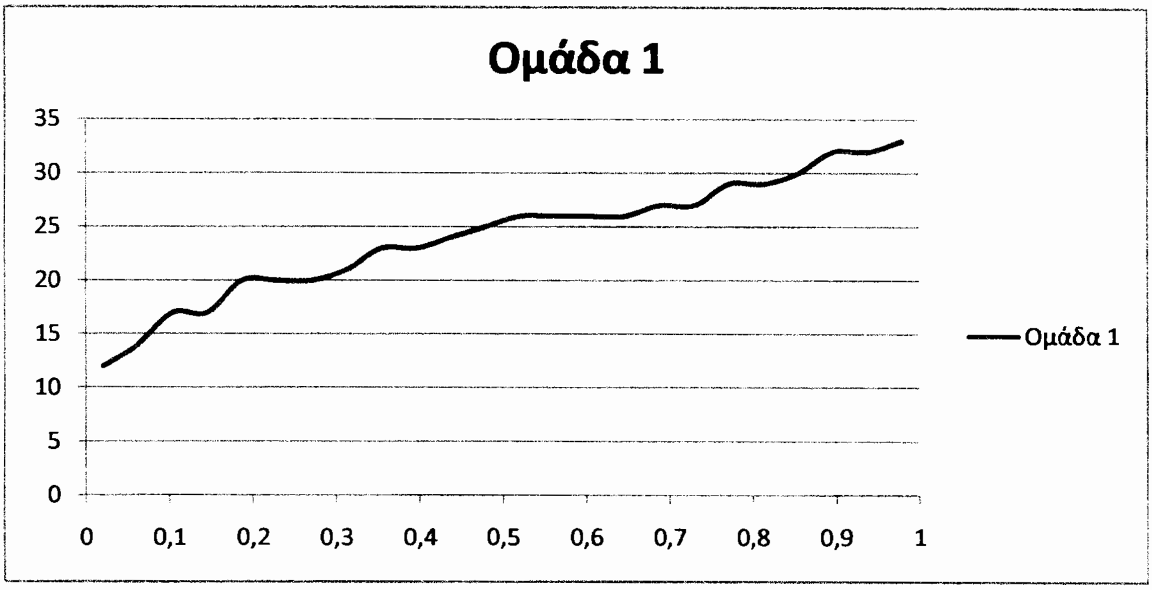
$R(3) = 0,1804 > 0,05$, $R(4) = 0,1020 > 0,05$

$R(5) = 0,05762 > 0,05$

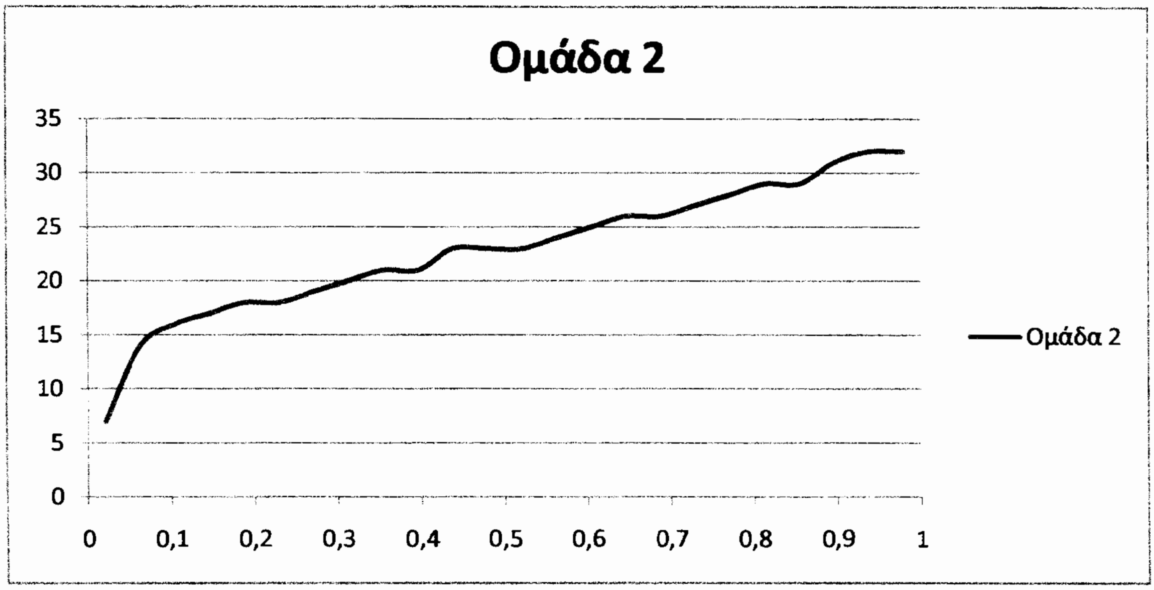
$R(6) = 0,03256 < 0,05$ άρα $n \le 6$

Άσκηση 5

		Ομάδα 1	Ομάδα 2	N(0,1)
i	pi	Q(pi)=x(i)	Q(pi)=y(i)	z(i)
1	0,020833333	12	7	-2,0368341
2	0,0625	14	14	-1,5341205
3	0,104166667	17	16	-1,2581616
4	0,145833333	17	17	-1,0544725
5	0,1875	20	18	-0,8871466
6	0,229166667	20	18	-0,741594
7	0,270833333	20	19	-0,6102946
8	0,3125	21	20	-0,4887764
9	0,354166667	23	21	-0,3740954
10	0,395833333	23	21	-0,264147
11	0,4375	24	23	-0,1573107
12	0,479166667	25	23	-0,0522452
13	0,520833333	26	23	0,0522452
14	0,5625	26	24	0,1573107
15	0,604166667	26	25	0,264147
16	0,645833333	26	26	0,3740954
17	0,6875	27	26	0,4887764
18	0,729166667	27	27	0,6102946
19	0,770833333	29	28	0,741594
20	0,8125	29	29	0,8871466
21	0,854166667	30	29	1,0544725
22	0,895833333	32	31	1,2581616
23	0,9375	32	32	1,5341205
24	0,979166667	33	32	2,0368341



(ΣΧ1)

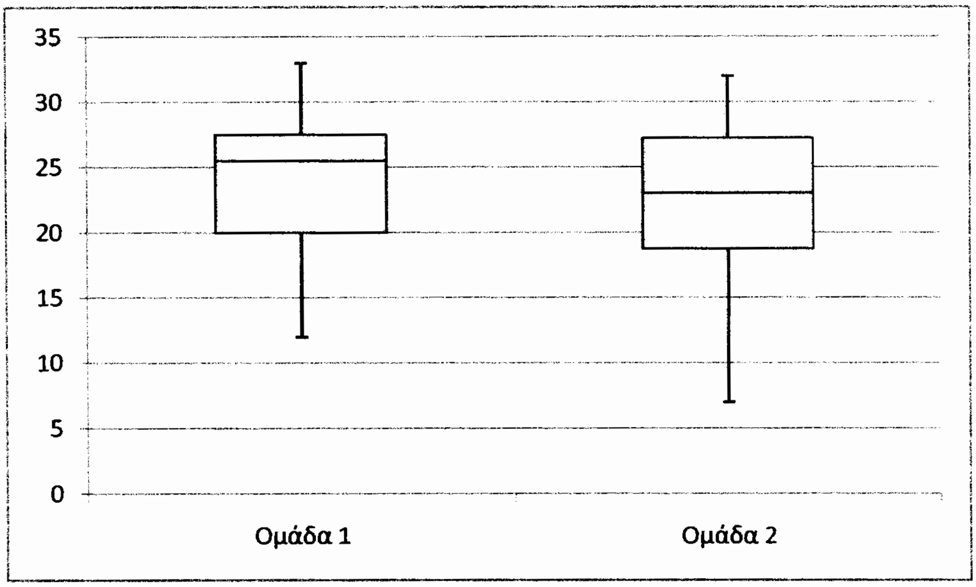


(ΣΧ2)

(α) Από το διάγραμμα ποσοστιαίας μπορούμε να συμπεράνουμε ότι για την Ομάδα 1 (ΣΧ1) φαίνεται να υπάρχει μια (κατά προσέγγιση) συμμετρική κατανομή σχετικά με το άνχος των ατόμων που εξετάστηκαν, ενώ για την Ομάδα 2 φαίνεται να υπάρχει μια ασυμμετρία προς τα αριστερά, αφού το κάτω μισό του διαγράμματος αυξάνεται περισσότερο απότομα.

(β) Παρατηρούμε ότι στο (ΣΧ3) η διάμεσος της Ομάδας 1 είναι αισθητά μεγαλύτερη από τη διάμεσο της Ομάδας 2. Αυτό μπορεί να σημαίνει ότι οι μέθοδοι που χρησιμοποιήθηκαν στην $2^η$ Ομάδα είχαν καλύτερο αποτέλεσμα, αφού τα άτομα παρουσιάζουν λιγότερο άνχος. Παρατηρούμε ότι το θηκόγραμμα για την Ομάδα 1 έχει τις περισσότερες τιμές του κάτω από την διάμεσο, κάτι που υποδεικνύει ότι η κατανομή των δεδομένων είναι λοξή προς τα αριστερά.

Σε αντίθεση με την Ομάδα 2 που φαίνεται να υπάρχει κάποια μορφής συμμετρία. Επιπρόσθετα υπάρχει μια μικρή διαφορά στις αριστερές ουρές, ενώ δε φαίνεται να υπάρχει σημαντική διαφορά στην μεταβλητότητα(κλίμακα) τους.

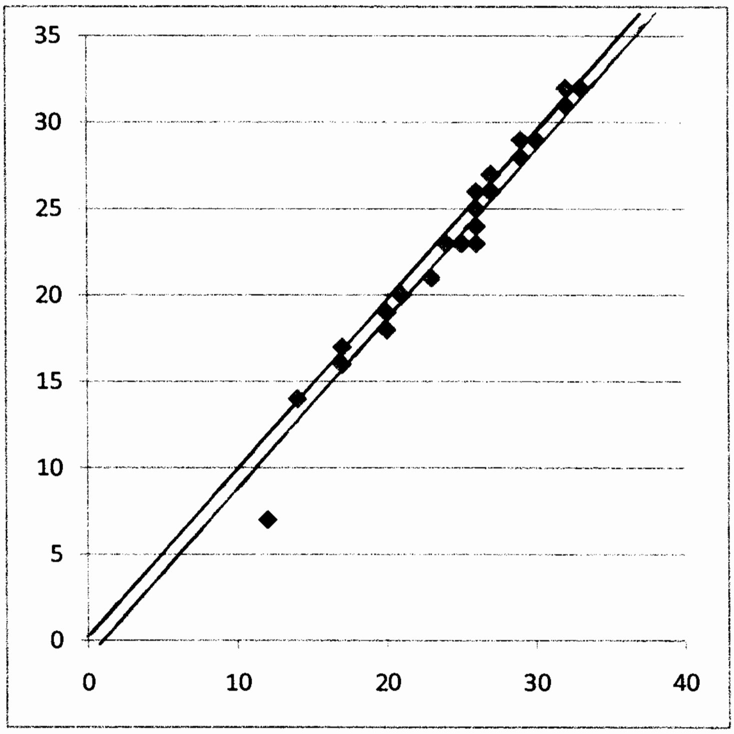


(ΣΧ3)

Min	12	7
Q1	20	18,75
Median	25,5	23
Q3	27,5	27,25
Max	33	32

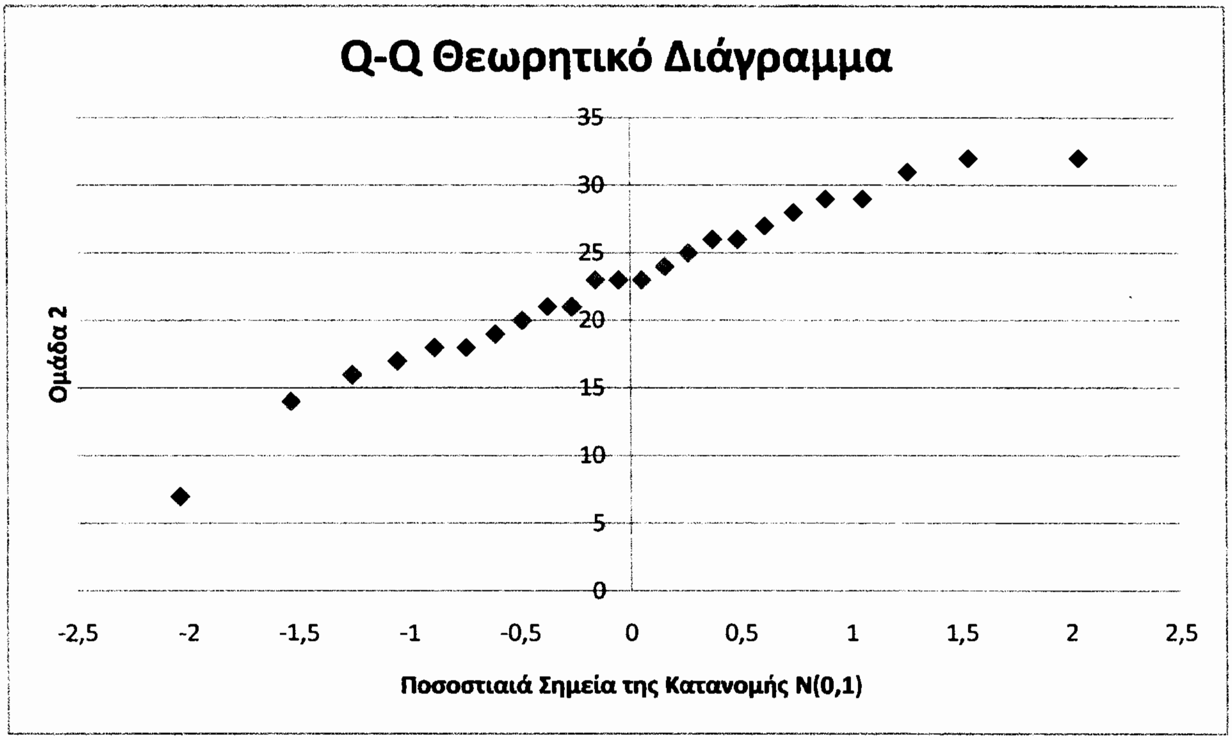
(γ) Στο διάγραμμα Q-Q (ΣΧ4), παρατηρούμε ότι τα σημεία βρίσκονται πολύ κοντά σε μια ευθεία παράλληλη της ευθείας y=x που κόβει τον κατακόρυφο άξονα κοντά στη μισή μονάδα. Επομένως οι κατανομές των δύο ομάδων διαφέρουν ουσιαστικά κατά c περίπου 0,5 που σημαίνει ότι οι κατανομές των αρχικών παρατηρήσεων διαφέρουν ουσιαστικά κατά την πολλαπλασιαστική σταθερή exp(c)=1.65. Δηλαδή $Y_{(i)}=1.65X_{(i)}$.

(10)



(ΣΧ4)

(δ) Στο Θεωρητικό διάγραμμα Q-Q στο (ΣΧ5), παρατηρούμε ότι η κατανομή του άνχους των ατόμων των δύο ομάδων διαφέρει ελαφρά από την κανονική, γιατί τα σημεία δεν πέφτουν πολύ κοντά σε μια ευθεία γραμμή. Παρατηρούμε επίσης ότι τα 4 πάνω δεξιά σημεία του διαγράμματος, πέφτουν πάνω από την νοητή ευθεία που περνά κοντά στα υπόλοιπα σημεία, πράγμα που σημαίνει ότι η κατανομή έχει μικρή λοξότητα προς τα δεξιά.



(ΣΧ5)

Άσκηση 6

(α) $X \sim B(3,p)$

Θέλω $P(X \leq 1) = P(0) + P(1) = \binom{3}{0} p^0 (1-p)^3 + \binom{3}{1} p^1 (1-p)^2$

$= (1-p)^3 + 3(1-p)^2$ άρα το (ii)

(β) $P(Z < -1{,}5) = P(Z > 1{,}5)$

$0{,}5 - P(Z > 1{,}5) = 0{,}5 - P(Z < -1{,}5)$ άρα το (i)

(γ) (iii)

(δ) $p = 0{,}01$ και $n = 100$, $np = 1 < 5$ άρα $\lambda = 1$

Η κατανομή Poisson (ii)

(ε) (iii) Εχουμε n ανεξάρτητες δοκιμές Bernoulli με σταθερή πιθανότητα επιτυχίας y σε καθε δοκιμή. Συνεπως $X \sim B(n, y)$

(στ) το (i) Δεν είναι αφου $a_3 \neq 0$

(ii) $E(X) = \frac{a+\beta}{2} = 1 \Leftrightarrow \boxed{a+\beta = 2}$ (1) και $Var(X) = \frac{(b-a)^2}{12} = 1$

$\Leftrightarrow (b-a)^2 = 12 \Leftrightarrow \boxed{b-a = \pm\sqrt{12}}$ (2)

Απο (1), (2)
- $a = 1-\sqrt{3}$ και $\beta = 1+\sqrt{3}$
- $a = 1+\sqrt{3}$ και $\beta = 1-\sqrt{3}$ απορρ αφού $a > \beta$

άρα $f(x) = \begin{cases} \frac{1}{\beta - a}, & a < x < \beta \\ 0, & \text{αλλιώς} \end{cases} = \begin{cases} \frac{\sqrt{3}}{6}, & 1-\sqrt{3} < x < 1+\sqrt{3} \\ 0, & \text{αλλιώς} \end{cases}$

Απορρίπτεται αφου δεν μπορεί να πάρει αρνητικές τιμές η κατανομή μας

(iii) Από τα δεδομένα έχω $a_3 > 0$ και $a_4 > 0$ άρα είναι λοξή προς τα δεξιά η $WEI(5,1)$ αφού $a=5$ είναι λοξή προς τα αριστερά

(iv) $E(X) = \frac{1}{\lambda} = 1{,}0079 \Leftrightarrow \lambda = 0{,}992162$

άρα $Var(X) = \frac{1}{\lambda^2} = \frac{1}{(0{,}992162)^2} = \frac{1}{0{,}984385277}$

$\Leftrightarrow Var(X) = 1{,}0159 \simeq 1{,}0248$

άρα Σωστή η (iv)

$\boxed{\text{Άσκηση 7}}$

(α) $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{52}(21+22+\dots+53) = 34{,}1923$

$\delta = \frac{X_{\nu/2} + X_{\frac{\nu+2}{2}}}{2} = \frac{X_{26}+X_{27}}{2} = \frac{34+34}{2} = 34 \Leftrightarrow \boxed{\delta = 34}$

$Q_1 = X_{(\frac{1}{4}(n+1))} = X_{\frac{53}{4}} = X_{13\frac{1}{4}}$. Άρα βρίσκεται μεταξύ της

$X_{(13)} = 30$ και της $X_{(14)} = 30$

$Q_1 = X_{(13)} + \frac{1}{4}\left(X_{(14)} - X_{(13)}\right) = 30 \Leftrightarrow \boxed{Q_1 = 30}$

$Q_3 = X_{(\frac{3}{4}(n+1))} = X_{\frac{159}{4}} = X_{39\frac{3}{4}} = X_{39} + \frac{3}{4}\left(X_{40} - X_{39}\right)$

$= X_{39} = 38 \Leftrightarrow \boxed{Q_3 = 38}$

$R = X_{52} - X_1 = 53 - 21 = 32$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 = 44,5897$$

άρα $s = \sqrt{s^2} = 6,6776$

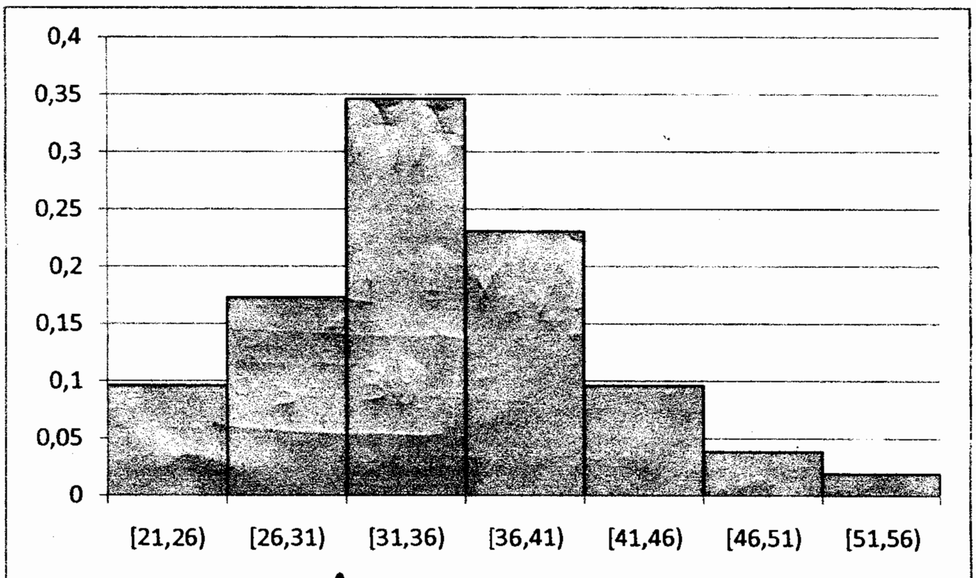
$$a_3 = \frac{1}{n s^3} \sum_{i=1}^{K} (X_i - \bar{x})^3 = 0,311933$$

Αφού $a_3 > 0$ η κατανομή είναι θετικά ασύμμετρη αλλά το $a_3 \simeq 0$ άρα θα είναι κοντά στην κανονική κατανομή

$\bar{x} - \delta = 0,192308 > 0$

(B) $C = \frac{R}{K} = \frac{32}{7} \simeq 5$

	Κλάσεις	Συχνότητα	Σχετική Συχνότητα
1	[21,26)	5	0,096153846
2	[26,31)	9	0,173076923
3	[31,36)	18	0,346153846
4	[36,41)	12	0,230769231
5	[41,46)	5	0,096153846
6	[46,51)	2	0,038461538
7	[51,56)	1	0,019230769
	Σύνολο	**52**	1



(ΣΧ1)

Παρατηρούμε από το ιστόγραμμα (ΣΧ1) ότι υπάρχει μια μικρή θετική ασυμμετρία κάτι που το είχαμε παρατηρήσει και στα προηγούμενα ερωτήματα. Αφού μόνο οι δύο τελευταίες κλάσεις είναι αυτές που χαλούν την συμμετρία

(γ) Επειδή οι αριθμοί που έχουμε είναι μεταξύ (20, 54) για την καλύτερη παρουσίαση του διαγράμματος μίσχου-φύλλου, θα θεωρήσουμε ως μίσχους $2^- \rightsquigarrow$ (20 έως 24), $2^+ \rightarrow$ (25 έως 29)

$3^- \rightsquigarrow$ (30 έως 34), $3^+ \rightarrow$ (35 έως 39)

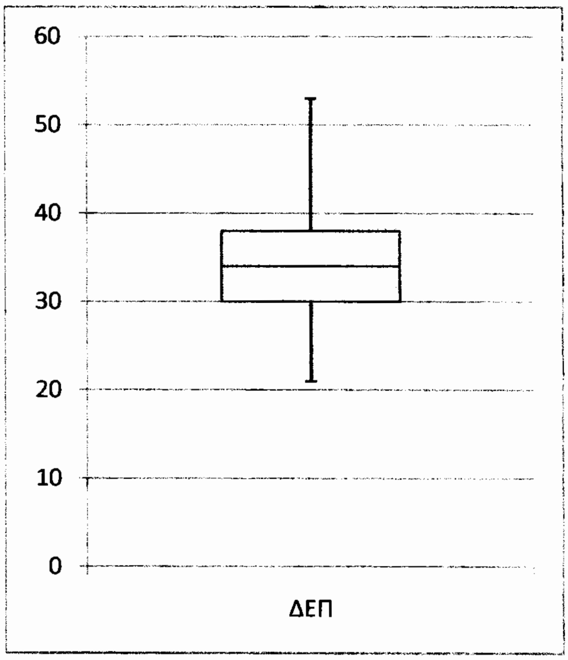
$4^- \rightsquigarrow$ (40 έως 44), $4^+ \rightarrow$ (45 έως 49)

2^- 1 2 3 4
2^+ 5 6 6 7 8 8 9 9
$(3)^-$ 0 0 1 1 1 2 2 2 2 3 3 3 3 4 4 4
3^+ 5 5 5 5 6 6 7 7 7 7 8 8 9 9
4^- 0 0 2 3 3 3 4
4^+ 6 7
5 3

Παρατηρούμε ότι υπάρχει θετική ασυμμετρία

(δ)

	ΔΕΠ
Min	21
Q1	30
Median	34
Q3	38
Max	53



Από το θηκόγραμμα παρατηρούμε ότι υπάρχει μια συμμετρία των δεδομένων.

(ε) Από μια γρήγορη παρατήρηση των παραπάνω διαγραμμάτων είναι προφανές ότι τα δεδομένα είναι πολύ κοντά στην κανονική κατανομή. Έτσι με την χρήση των ιδιοτήτων της κανονικής κατανομής έχουμε:

Χρήση των διαστημάτων $\bar{x} \pm ks$, $k = 1, 2, 3$

Για $k=1$: $(\bar{x} - s, \bar{x} + s) = (27.51, 40.87)$

Για $k=2$: $(\bar{x} - 2s, \bar{x} + 2s) = (20.84, 47.55)$

Για $k=3$: $(\bar{x} - 3s, \bar{x} + 3s) = (14.16, 54.22)$

Το ποσοστό των δεδομένων που ανήκουν στα παραπάνω διαστήματα είναι:

1ο Διάστημα: $\frac{36}{52} = 0,6923$ ή $69,23\% \simeq 68,26\%$

2ο Διάστημα: $\frac{51}{52} = 0,9808$ ή $98,08\%$ $(95,44\%)$

3ο Διάστημα: $\frac{52}{52} = 1$ ή 100% (99,74 %)

Παρατηρούμε ότι τα ποσοστά των τριών διαστημάτων των δεδομένων είναι πολύ κοντά στα ιδανικά της κανονικής κατανομής. Και που εξηγείται πλήρως από τον συντελεστή λοξότητας άρα $a_3 \simeq 0$ και από τα διαγράμματα. Τελικά μπορούμε να συμπεράνουμε την ελάχιστη θετική ασυμμετρία από τα ποσοστά που είναι λίγο μεγαλύτερα από αυτά της κανονικής

Άσκηση 4 (2ος τρόπος)

(γ) Έστω η τ.μ. T: ο ενδιάμεσος χρόνος μεταξύ δύο διαδοχικών βλαβών της συσκευής. Τότε η τ.μ. $T \sim \mathcal{E}(\lambda)$, όπου $\lambda = 1,752$

άρα $P(1 \leq T \leq 2) = \int_1^2 f_T(t)dt = \int_1^2 1.752\, e^{-1.752 \cdot t} dt$

$= \left[-e^{-1,752 \cdot t}\right]_1^2 = -e^{-1,752 \cdot 2} + e^{-1.752 \cdot 1} = -0,030 + 0,1734$

$= 0,1434$