Επαναληπτικό διαγώνισμα στη Στατιστική – Λύσεις

aftermathsgr aftermathsgr aftermathsgr

17 Οκτωβρίου 2020

Θέμα 1 (25 μονάδες).

1. Αν $x_1, x_2, \ldots, x_k$ είναι οι τιμές μίας μεταβλητής X που αφορά άτομα ενός δείγματος μεγέθους ν, $k \leq n$, να αποδείξετε ότι για τις αντίστοιχες σχετικές συχνότητες f_i ισχύει ότι (6 μονάδες):

$$f_1 + f_2 + \ldots + f_k = 1.$$

2. Να χαρακτηρίσετε τις ακόλουθες προτάσεις ως **Σωστές** ή **Λανθασμένες** (10 μονάδες):

   (α΄) Το ραβδόγραμμα (barchart) χρησιμοποιείται για τη γραφική παράσταση των τιμών μίας ποσοτικής μεταβλητής.

   (β΄) Η διακύμανση/διασπορά μετριέται στις ίδιες μονάδες που μετριούνται και οι παρατηρήσεις του δείγματος.

   (γ΄) Σε ένα δείγμα 5 παρατηρήσεων με μέση τιμή $\overline{x} = 6$ αν αυξήσουμε μία παρατήρηση κατά 2 μονάδες τότε θα αυξηθεί και η μέση τιμή κατά 2 μονάδες.

   (δ΄) Η διάμεσος ενός δείγματος είναι πάντοτε ίση με μία από τις παρατηρήσεις του δείγματος.

   (ε΄) Το εύρος (R) ενός δείγματος μας δείχνει την απόσταση της μεγαλύτερης από την μικρότερη παρατήρηση.

3. Ποιες μεταβλητές ονομάζουμε **ποσοτικές** και σε ποιες κατηγορίες τις διακρίνουμε; Να δώσετε από τρία παραδείγματα για κάθε κατηγορία που θα γράψετε. (6 μονάδες)

4. Να χαρακτηρίσετε την ακόλουθη πρόταση ως **Σωστή** ή **Λανθασμένη** και να **αιτιολογήσετε** την απάντησή σας (3 μονάδες):

   «Αν οι 400 παρατηρήσεις ενός δείγματος ακολουθούν περίπου κανονική κατανομή με μέση τιμή $\overline{x} = 46$ και τυπική απόκλιση $s = 4$ τότε περίπου 480 από αυτές είναι μικρότερες ή ίσες από 54».

Θέμα 2 (25 μονάδες). Οι θερμοκρασίες των τελευταίων 10 ημερών στην πόλη της Χαλκίδας ήταν οι εξής:

$$27, 24, 27, 27, 26, 25, 24, 25, 28, 27.$$

1. Να κατασκευάσετε έναν πίνακα συχνοτήτων για το παραπάνω δείγμα που να συμπεριλαμβάνει τις απόλυτες (ν_i) και σχετικές (f_i) συχνότητες καθώς και τις αντίστοιχες αθροιστικές συχνότητες. (8 μονάδες)

2. Να σχεδιάσετε τα διαγράμματα και πολύγωνα σχετικών και αθροιστικών σχετικών συχνοτήτων για το παραπάνω δείγμα. (5 μονάδες)

3. Να υπολογίσετε τη μέση και τη διάμεση θερμοκρασία καθώς και την τυπική απόκλιση για το παραπάνω δεκαήμερο. (9 μονάδες)

4. Να εξετάσετε αν το παραπάνω δείγμα είναι ομοιογενές. (3 μονάδες)

   Δίνεται: $\sqrt{1.8} = 1.34$.

Θέμα 3 (25 μονάδες). Δίνεται ο παρακάτω πίνακας συχνοτήτων για ένα δείγμα ν παρατηρήσεων:

$[\alpha_i, \beta_i)$	x_i	ν_i	f_i	$f_i\%$	N_i	F_i	$F_i\%$
[-4,0)		4					
[0,4)			0,25				
[4,8)							90
[8,12)		2					
Σύνολο	—				—	—	—

1. Να συμπληρώσετε τον πίνακα. (9 μονάδες)

2. Να βρείτε τη διάμεσο (δ) του δείγματος. (7 μονάδες)

3. Έστω $\overline{x}$ και s η μέση τιμή και η τυπική απόκλιση του παραπάνω δείγματος αντίστοιχα. Αν $\overline{x}'$ και s' είναι η μέση τιμή και η τυπική απόκλιση του δείγματος αφού πολλαπλασιάσουμε κάθε παρατήρησή του με -4, να λύσετε την εξίσωση (8 μονάδες):

$$x^2 - \frac{s' - s}{s}x - \frac{4\overline{x} - \overline{x}'}{\overline{x}'} = 0.$$

Θέμα 4 (25 μονάδες). Δίνεται το παρακάτω δείγμα:

$$2, 4, 3, a, 0,$$

όπου $a \in \mathbb{R}$.

1. Να βρείτε την τιμή του a για την οποία για τη μέση τιμή ($\overline{x}$) του δείγματος ισχύει $\overline{x} = 3$. (4 μονάδες)

2. Να βρείτε τις τιμές του a για τις οποίες για τη διάμεσο (δ) του δείγματος ισχύει $\delta = 3$. (5 μονάδες)

3. Να βρείτε τις τιμές του a για τις οποίες $\overline{x} = \delta$. (7 μονάδες)

4. Για $a = 6$:

   (α΄) Να δείξετε ότι το δείγμα είναι ανομοιογενές. (2 μονάδες)

   (β΄) Να βρείτε την ελάχιστη τιμή της θετικής σταθεράς c που πρέπει να προσθέσουμε σε κάθε παρατήρηση έτσι ώστε το δείγμα να γίνει ομοιογενές. (7 μονάδες)

Λύσεις

Λύση Θέματος 1.

1. Σ.Β., σελ. 65.

2. (α΄) Λάθος

   (β΄) Λάθος

   (γ΄) Λάθος

   (δ΄) Λάθος

   (ε΄) Σωστό

3. Σ.Β., σελ. 59.

4. Η πρόταση είναι **Λανθασμένη**. Πράγματι, αφού οι παρατηρήσεις ακολουθούν περίπου κανονική κατανομή με μέση τιμή $\bar{x} = 46$ και τυπική απόκλιση $s = 4$ τότε το $2,5\%$ αυτών είναι μεγαλύτερες από $\bar{x} + 2s = 46 + 8 = 54$, συνεπώς περίπου $400 \cdot 2.5\% = 400 \cdot \frac{2.5}{100} = 10$ παρατηρήσεις είναι μεγαλύτερες από 54 άρα οι υπόλοιπες (περίπου) 490 είναι μικρότερες ή ίσες από 54.

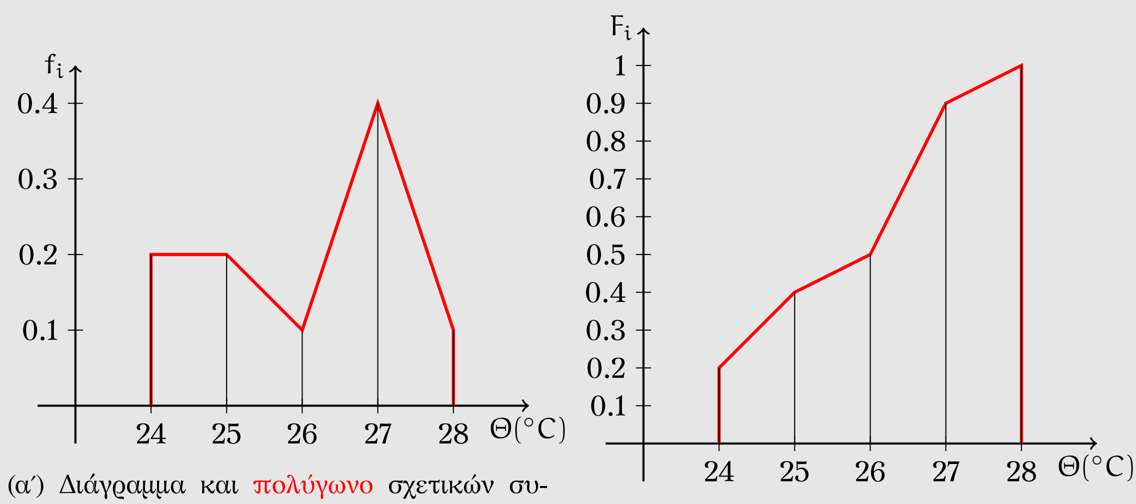
Λύση Θέματος 2.

1. Στον πίνακα 1 βλέπουμε τον ζητούμενο πίνακα συχχοτήτων.

x_i	ν_i	f_i	N_i	F_i
24	2	0.2	2	0.2
25	2	0.2	4	0.4
26	1	0.1	5	0.5
27	4	0.4	9	0.9
28	1	0.1	10	1
Σύνολο	10	1	–	–

Πίνακας 1: Ο πίνακας συχντοτήτων του ερωτήματος 2.

2. Τα ζητούμενα διαγράμματα και πολύγωνα φαίνονται στο σχήμα 1



(α΄) Διάγραμμα και πολύγωνο σχετικών συχνοτήτων.

(β΄) Διάγραμμα και πολύγωνο σχετικών συχνοτήτων.

Σχήμα 1: Διαγράμματα και πολύγωνα συχνοτήτων.

3. Για τη διάμεσο, παρατηρούμε ότι οι δύο μεσαίες παρατηρήσεις είναι οι 26 και 27, επομένως:
$$\delta = \frac{26+27}{2} = 26.5.$$
Για τη μέση τιμή έχουμε:
$$\bar{x} = \frac{2\cdot 24 + 2\cdot 25 + 1\cdot 26 + 4\cdot 27 + 1\cdot 28}{10} = \frac{260}{10} = 26.$$
Για τη διακύμανση έχουμε:
$$\begin{aligned} s^2 &= \frac{2(24-26)^2 + 2(25-26)^2 + 1(26-26)^2 + 4(27-26)^2 + 1(28-26)^2}{10} = \\ &= \frac{2\cdot 4 + 2\cdot 1 + 0 + 4\cdot 1 + 1\cdot 4}{10} = \frac{18}{10} = 1.8. \end{aligned}$$
Συνεπώς:
$$s = \sqrt{s^2} = \sqrt{1.8} = 1.34.$$

4. Για την ομοιογένεια του δείγματος έχουμε:
$$CV = \frac{s}{\bar{x}} = \frac{1.34}{26} < \frac{2.6}{26} = 0.1,$$
άρα το δείγμα είναι ομοιογενές.

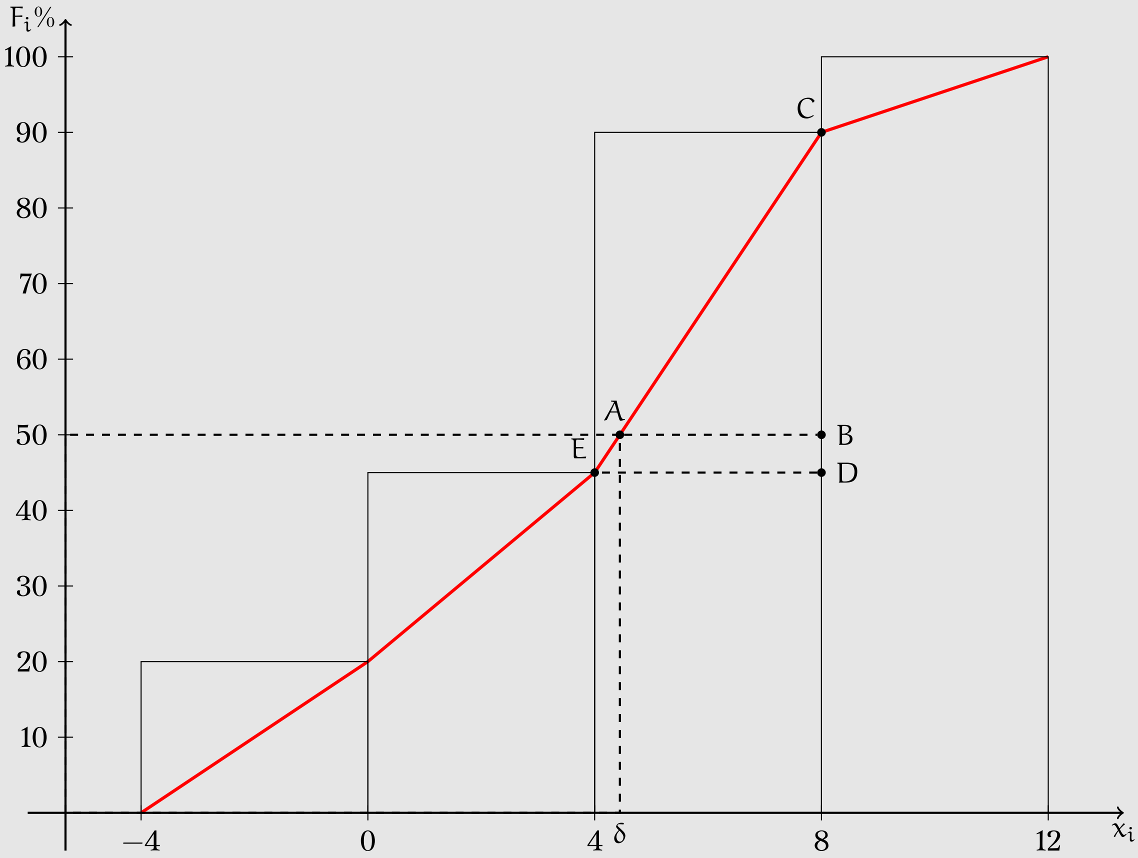
Λύση Θέματος 3. 1. Ο ζητούμενος πίνακας συχνοτήτων είναι πίνακας 2.

$[\alpha_i, \beta_i)$	x_i	ν_i	f_i	$f_i\%$	N_i	F_i	$F_i\%$
[-4,0)	-2	4	0.20	20	4	0.20	20
[0,4)	2	5	0.25	25	9	0.45	45
[4,8)	6	9	0.45	45	18	0.90	90
[8,12)	10	2	0.10	10	20	1	100
Σύνολο	—	20	1	100	—	—	—

Πίνακας 2: Ο ζητούμενος πίνακας συχνοτήτων.

2. Για να βρούμε τη διάμεσο θα σχεδιάσουμε πρώτα το πολύγωνο αθροιστικών σχετικών συχνοτήτων επί τοις εκατό ($F_i\%$) όπως φαίνεται στο σχήμα 2. Στη συνέχεια, αφού η διάμεσος αντιστοιχεί στην τιμή $F_i = 50\%$ φέρουμε ευθεία από το 50 του άξονα των $F_i\%$ και κατασκευάζουμε τα σημεία A, B όπως φαίνεται στο σχήμα 2. Από την ομοιότητα των τριγώνων $\triangle ABC$ και $\triangle CDE$ έχουμε:
$$\frac{AB}{ED} = \frac{BC}{CD} \Leftrightarrow \frac{8-\delta}{8-4} = \frac{90-50}{90-45} \Leftrightarrow \frac{8-\delta}{4} = \frac{40}{45} \Leftrightarrow \frac{8-\delta}{4} = \frac{8}{9} \Leftrightarrow 72 - 9\delta = 32 \Leftrightarrow \boxed{\delta = \frac{40}{9}.}$$

3. Αφού πολλαπλασιάσουμε κάθε τιμή του δείγματος με -4 η μέση τιμή και η τυπική απόκλιση μεταβάλλονται ως εξής:
$$\bar{x}'4\bar{x},\ \ s' = |-4|s = 4s.$$
Έτσι, έχουμε:
$$\frac{s'-s}{s} = \frac{4s-s}{s} = \frac{3s}{s} = 3 \tag{1}$$
$$\frac{4\bar{x}-\bar{x}'}{\bar{x}'} = \frac{4\bar{x}+4\bar{x}}{-4\bar{x}} = \frac{8\bar{x}}{-4\bar{x}} = -2 \tag{2}$$
Αντικαθιστώντας τις (1) και (2) στη δοσμένη εξίσωση έχουμε:
$$x^2 - 3x - (-2) = 0 \Leftrightarrow x^2 - 3x + 2 = 0 \xLeftrightarrow{\Delta=1>0} x = 1 \text{ ή } x = 2.$$



Σχήμα 2: Ιστόγραμμα και πολύγωνο $F_i\%$.

Λύση Θέματος 4. 1. Έχουμε:

$$\overline{x} = \frac{2+4+3+a+0}{5} = \frac{9+a}{5},$$

επομένως:

$$\overline{x} = 3 \Leftrightarrow \frac{9+a}{5} = 3 \Leftrightarrow 9+a = 15 \Leftrightarrow a = 6.$$

2. Διατάσσουμε τις γνωστές παρατηρήσεις σε αύξουσα σειρά:

$$0, 2, 3, 4.$$

Δεδομένου ότι οι ππαρατηρήσεις μας είναι 5 στο πλήθος, η διάμεσος θα είναι μία από αυτές. Έτσι, για να ισχύει ότι η διάμεσος είναι ίση με 3 πρέπει είτε $a > 3$, οπότε η μεσαία παρατήρηση είναι το 3 είτε $a = 3$ οπότε και πάλι η μεσαία παρατήτηση είναι το 3. Ειδάλλως, αν δηλαδή $a < 3$ τότε η μεσαία παρατήρηση είναι το a, άρα $\delta = a < 3$. Συνεπώς, οι ζητούμενες τιμές είναι $a \geq 3$.

3. Όπως είδαμε παραπάνω:

$$\delta = \begin{cases} 3 & a \geq 3 \\ a & a < 3 \end{cases}$$

Διακρίνουμε, λοιπόν, δύο περιπτώσεις:

- Αν $a \geq 3$ τότε:

$$\overline{x} = \delta \Leftrightarrow \frac{9+a}{5} = 3 \Leftrightarrow 9+a = 15 \Leftrightarrow a = 6,$$

που είναι δεκτό.

- Αν $\alpha < 3$ τότε:

$$\bar{x} = \delta \Leftrightarrow \frac{9+\alpha}{5} = \alpha \Leftrightarrow 9+\alpha = 5\alpha \Leftrightarrow 4\alpha = 9 \Leftrightarrow \alpha = \frac{9}{4},$$

  που είναι δεκτό γιατί $\frac{9}{4} < 3$.

Συνεπώς, οι ζητούμενες τιμές είναι $\alpha = 6$ και $\alpha = \frac{9}{4}$.

4. Για $\alpha = 6$:

(α΄) Έχουμε $\bar{x} = 3$ όπως είδαμε παραπάνω, ενώ για τη διακύμανση έχουμε:

$$\begin{aligned} s^2 &= \frac{(0-3)^2+(2-3)^2+(3-3)^2+(4-3)^2+(6-3)^2}{5} = \\ &= \frac{9+1+0+1+9}{5} = \\ &= 4. \end{aligned}$$

Συνεπώς $s = \sqrt{s^2} = \sqrt{4} = 2$ άρα:

$$CV = \frac{s}{\bar{x}} = \frac{2}{3} > 0.1,$$

άρα το δείγμα είναι ανομοιογενές.

(β΄) Αν προσθέσουμε σε κάθε παρατήρηση μία θετική σταθερά c τότε η νέα μέση τιμή $\bar{x}'$ θα είναι $\bar{x}' = \bar{x} + c = 3 + c$ ενώ η τυπική απόκλιση θα παραμείνει αμετάβλητη. Συνεπώς:

$$CV' = \frac{s'}{\bar{x}'} = \frac{2}{3+c}.$$

Για να είναι το δείγμα μας ομοιογενές πρέπει:

$$CV' \leq 0.1 \Leftrightarrow \frac{2}{3+c} \leq \frac{1}{10} \overset{c>0}{\Longleftrightarrow} 20 \leq 3 + c \Leftrightarrow c \geq 7.$$

Έτσι, η ελάχιστη τιμή για την οποία συμβαίνει αυτό είναι $c = 7$.