

Επαναληπτικές Ασκήσεις

Από το Αρχείο με παλιά θέματα κλπ στο eclass

Παλιά θέματα, ερώτηση 3:

Για τα παρακάτω δεδομένα

1	1	1	3	2	2	1	2	2	0	1
1	2	0	2	2	3	3	4	5	5	0
1	2	2	1	1	1	1	0	0	0	4
3	3	4	5	5	1	2	2			

- α) να βρείτε τη διάμεσο, τη μέση τιμή και την επικρατούσα τιμή
- β) να σχεδιάσετε ένα απλό ραβδόγραμμα συχνοτήτων
- γ) να βρείτε τα τεταρτημόρια
- δ) τι μορφή παρουσιάζει η κατανομή;

Μετράμε πόσα 0, πόσα 1, Έχουμε και προκύπτει ο πίνακας συχνοτήτων

Τιμή	Συχνότητα
0	6
1	12
2	11
3	5
4	3
5	4
Σύνολο	41

Επικρατούσα Τιμή

Το 2 γιατί εμφανίζεται τις περισσότερες φορές (12 φορές)

Μέση Τιμή

Μέση τιμή = $(0+0+0+.....+5+5+5+5+5) / 41 =$

$81/41 =$ περίπου 2

(ακριβώς βγαίνει 1,98)

Διάμεσος (50% σημείο)

Τα γράφω στη σειρά

0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 4 4 4 5 5 5 5

Διαγράφω από πάνω και από κάτω και

~~0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 2 2~~ **2** ~~2 2 2 2 2 2 2 2 2 3 3 3 3 3 4 4 4 5 5 5 5~~

- Διάμεσος είναι το 2

Τεταρτημόρια Q1 και Q3

Το πρώτο τεταρτημόριο αφήνει 25% κάτω.

Οπότε αν πάρω τα πρώτα μισά νούμερα από πάνω που τα έβαλα στη σειρά, έχω

0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 2 2 2

Σβήνοντας ένα από πάνω κι ένα από κάτω έχουμε

~~0 0 0 0 0 1 1 1 1~~ **1** ~~1 1 1 1 1 1 1 1 2 2 2~~

Μένει στη μέση το 1, που είναι στη μέση του 50% δηλαδή είναι το 25% σημείο, δηλαδή το 1ο τεταρτημόριο (Q1)

Το τρίτο τεταρτημόριο αφήνει 75% κάτω.

Οπότε αν πάρω τα δεύτερα μισά νούμερα από πάνω που τα έβαλα στη σειρά, έχω

2 2 2 2 2 2 2 2 2 3 3 3 3 3 4 4 4 5 5 5 5

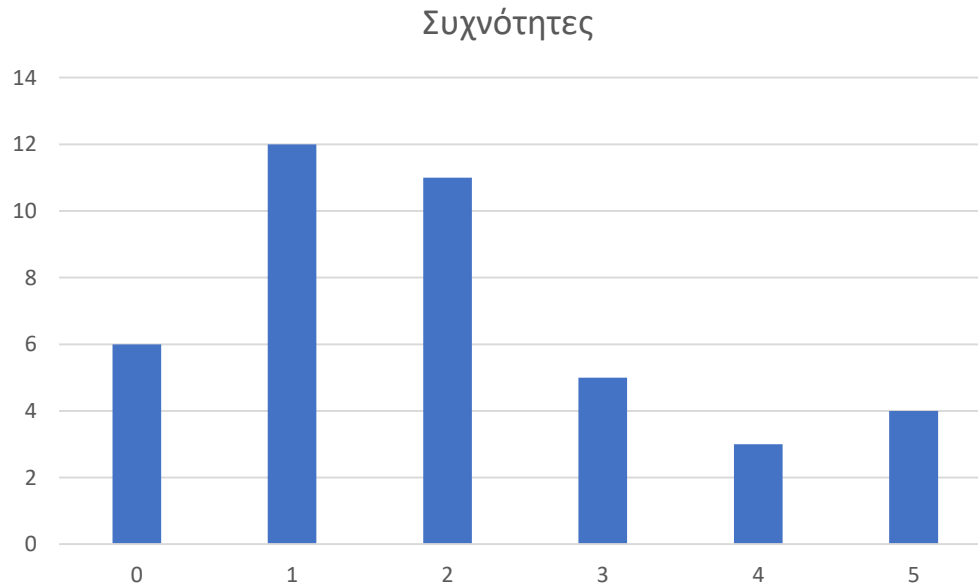
Σβήνοντας ένα από πάνω κι ένα από κάτω έχουμε

~~2 2 2 2 2 2 2 2 2 3~~ **3** ~~3 3 3 4 4 4 5 5 5 5~~

Μένει στη μέση το 3, που είναι στη μέση του πάνω 50% κομματιού, δηλαδή είναι το 75% σημείο, δηλαδή το 3ο τεταρτημόριο (Q3)

Συμμετρία;

Επειδή διάμεσος και μέση τιμή είναι σχεδόν ίσες (2 και 1,98) έχουμε συμμετρική κατανομή.



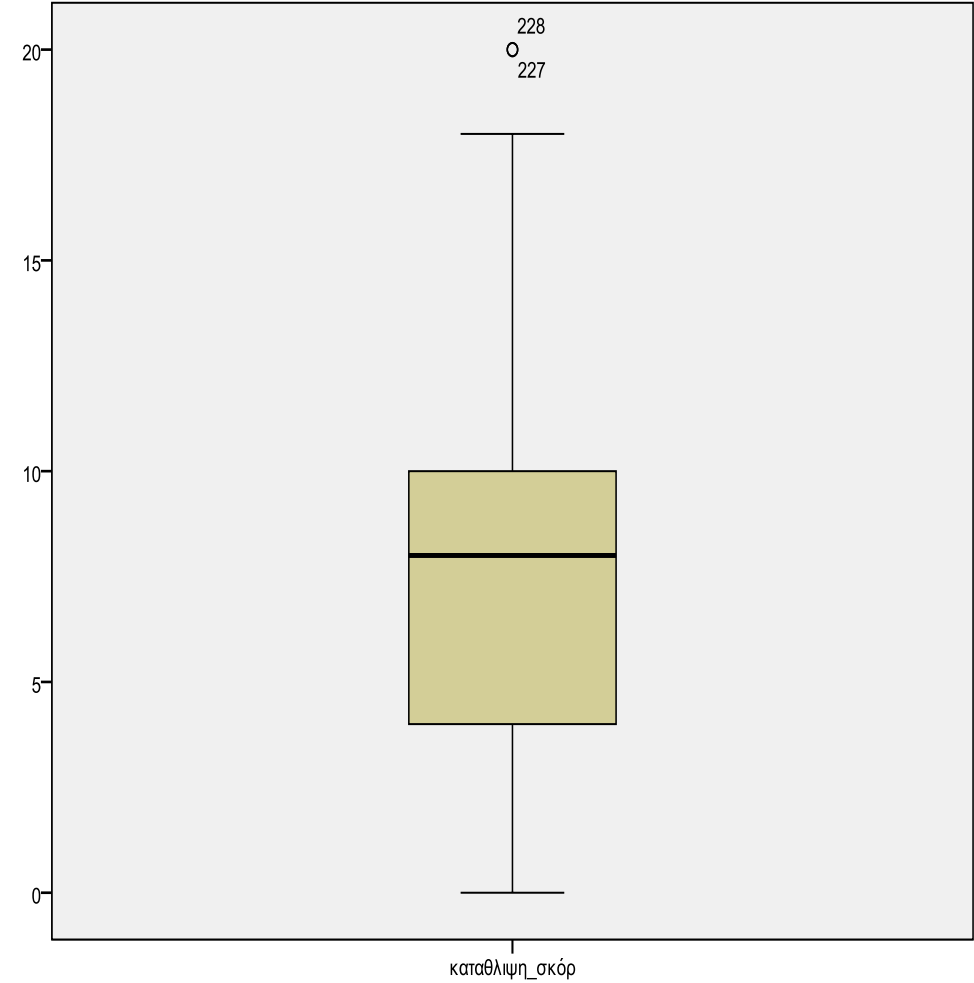
Ραβδόγραμμα συχνοτήτων

Παλιά θέματα, ερώτηση 8:

Σε μια μελέτη καταγράφηκαν 250 άτομα από περιοχές της Αθήνας. Οι ερευνητές χρησιμοποίησαν μια κλίμακα για την κατάθλιψη με σκόρ που είχε δυνατό εύρος από 0 μέχρι 20.

Όσο μεγαλύτερο ήταν το σκόρ τόσο περισσότερο καταθλιπτικός μπορούσε να χαρακτηριστεί κάποιος. Στο ερωτηματολόγιο υπήρχαν επίσης δημογραφικά στοιχεία (π.χ. ηλικία, φύλο). Τα αποτελέσματα απο τον υπολογιστή φαίνονται στους παρακάτω πίνακες (ή γραφικές παραστάσεις)

Descriptives			Statistic	Std. Error
καταθλιψη_σκόρ	Mean		7,50	,297
	95% Confidence Interval for Mean	Lower Bound	6,92	
		Upper Bound	8,08	
	5% Trimmed Mean		7,37	
	Median		8,00	
	Variance		22,002	
	Std. Deviation		4,691	
	Minimum		0	
	Maximum		20	
	Range		20	
	Interquartile Range		6	
	Skewness		,050	,154
	Kurtosis		-,423	,307



Απαντήστε σύντομα στις ερωτήσεις:

α) Πόση είναι η διάμεσος και τι πληροφορία δίνει;

Η διάμεσος (median) είναι 8 – μπορεί να φανεί από τον πίνακα ή ακόμη και από το θηκόγραμμα

Η πληροφορία που έχουμε είναι ότι τα μισά άτομα είχαν σκόρ κατάθλιψης κάτω από 8 και τα άλλα μισά πάνω από 8.

β) Πόση είναι η τυπική απόκλιση και τι πληροφορία δίνει;

Είναι 4,691 (SD-standard deviation)

Δίνει μια πληροφορία για το πόσο κοντά στο μέσο όρο κατανέμονται οι τιμές της κατάθλιψης για όλους τους συμμετέχοντες

Επιπλέον: μπορούμε πχ. να πούμε πως ανάμεσα στα σκόρ $7,5 \pm 4,691$ βρίσκεται το 68% των συμμετεχόντων στην έρευνα (ή ακόμα ανάμεσα στο $7,5 \pm 2 * 4,691$ βρίσκεται το 95% των συμμετεχόντων)

γ) Τι πληροφορία δίνει το 95% Διάστημα Εμπιστοσύνης για τη μέση τιμή (confidence interval for mean);

Ότι είμαστε σίγουροι 95% πως η πραγματική μέση τιμή της κατάθλιψης σε ολόκληρο τον πληθυσμό θα είναι μεταξύ 6,92 και 8,08

δ) πόσο είναι το $Q1$ και το $Q3$;

Αφού δεν αναγράφονται στον πίνακα με τα αποτελέσματα, μόνο από το θηκόγραμμα μπορούμε να τα βρούμε

$Q1=4$ (κάτω γραμμή του κουτιού) και

$Q3=10$ (πάνω γραμμή του κουτιού)

ε) Ποια είναι η μορφή της κατανομής της κατάθλιψης; Δικαιολογήστε την απάντησή σας.

Μέση τιμή = 7,5 και Διάμεσος = 8

Άρα φαίνεται να είναι αρνητικά ασύμμετρη

στ) τι είναι οι τιμές 227 και 228 στο παραπάνω γράφημα;

Είναι παράτυπες τιμές (outliers), υπερβολικά μεγάλες τιμές που δεν συμφωνούν με τους υπόλοιπους συμμετέχοντες στο δείγμα.

ζ) τι σημαίνουν τα ευρήματα του παρακάτω πίνακα;

Correlations			
		καταθλιψη_σκόρ	ηλικία
καταθλιψη_σκόρ	Pearson Correlation	1	,348**
	Sig. (2-tailed)		,000
	N	250	250
ηλικία	Pearson Correlation	,348**	1
	Sig. (2-tailed)	,000	
	N	250	250

**. Correlation is significant at the 0.01 level (2-tailed).

Σημαίνουν ότι ο ερευνητής υπολόγισε τον συντελεστή συσχέτισης του Pearson ανάμεσα στην ηλικία και την κατάθλιψη

Έβγαλε το r ίσο με 0,348 και το αντίστοιχο $p < 0,001$

Άρα η ηλικία έχει μια θετική σχέση με την κατάθλιψη, όσο πιο μεγάλη η ηλικία τόσο πιο μεγάλο το σκόρ της κατάθλιψης.

η) ο ερευνητής επίσης ανέφερε ότι συγκρίνοντας άνδρες-γυναίκες βρήκε $t=0.84$ και $p=0.402$. Τι ανάλυση έκανε και τι συμπέρασμα έβγαλε;

Έκανε t-test για να συγκρίνει την κατάθλιψη ανάμεσα σε δύο ομάδες (άνδρες – γυναίκες). Βρήκε το $p=0,402$ μεγαλύτερο από το 0,05.

Άρα δεν διαπίστωσε καμία διαφορά στην κατάθλιψη ανάμεσα σε άνδρες και γυναίκες, ή με άλλα λόγια η κατάθλιψη δεν φαίνεται να έχει σχέση με το φύλο.

Άσκηση

Ο παρακάτω πίνακας παρουσιάζει τον αριθμό των κύριων δωματίων, που είχαν στην κατοικία τους 53 φοιτητές που έμεναν σε κάποια πόλη.

Αριθμός Δωματίων	Συχνότητα	Αθροιστική συχνότητα
1		25
2		45
3		50
Πάνω από 4		53
Σύνολο		

- α. Συμπληρώστε τον πίνακα. β. Σχεδιάστε ένα απλό ραβδόγραμμα με τις Συχνότητες
γ. Τι σημαίνει ο αριθμός 50 που βρίσκουμε στην τελευταία στήλη;

Αριθμός Δωματίων	Συχνότητα	Αθροιστική συχνότητα
1	25	25
2	20	45
3	5	50
Πάνω από 4	3	53
Σύνολο	53	

- Η πρώτη συχνότητα είναι πάντα ίδια με την Αθροιστική – άρα 25
- Η δεύτερη είναι η αφαίρεση των αθροιστικών, της κάτω αθροιστικής απο την πάνω, δηλαδή $45 - 25 = 20$
- Η επόμενη είναι $50 - 45 = 5$
- Και η τελευταία είναι $53 - 50 = 3$

Το 50 σημαίνει πως είχαμε 50 φοιτητές που είχαν στο σπίτι τους από 3 δωμάτια και κάτω

Άσκηση για εξάσκηση 1:

Μια μελέτη με ερωτηματολόγια, διερεύνησε την επαγγελματική ικανοποίηση σε απλό τυχαίο δείγμα 300 κοινωνικών λειτουργών που εργάζονται στην Αθήνα. Η επαγγελματική ικανοποίηση καθορίστηκε από το σκορ των ατόμων (δυνατό εύρος από 0 μέχρι 50). Οι δειγματικές στατιστικές για το σκορ επαγγελματικής ικανοποίησης υπολογίστηκαν ως εξής

μέση τιμή=16

τυπική απόκλιση=4

Ενδοτεταρτομοριακό εύρος: από 12,5 έως 20,5, Διάμεσος = 14

Απαντήστε σύντομα στις ερωτήσεις:

α) Τι πληροφορία δίνει η τιμή 12,5;

β) Είναι η κατανομή του σκορ επαγγελματικής ικανοποίησης συμμετρική; Δικαιολογήστε την απάντησή σας.

γ) Ανάμεσα σε ποιες τιμές σκορ, βρίσκεται το 68% των κοινωνικών λειτουργών που μετείχαν στην έρευνα;

δ) Να βρείτε την διακύμανση

ε) να υπολογίσετε τον συντελεστή μεταβλητότητας (CV)

στ) Συγκρίνοντας το σκορ επαγγελματικής ικανοποίησης ανάμεσα σε άντρες και γυναίκες, βρήκαμε $p=0,028$. Ποια κατάλληλη στατιστική διαδικασία χρησιμοποιήσαμε και τι συμπέρασμα βγάλαμε;

Άσκηση για εξάσκηση 2:

Μια μελέτη που έγινε σε ηλικιωμένους υπολογίστηκε ως ποσοτική μεταβλητή, το Σκόρ Ποιότητας Ζωής (στην κλίμακα 0-56) και βρέθηκαν:

μέση τιμή=24, διακύμανση=49, Διάμεσος=27, $Q1=22$, $Q3=36$

- α. Η κατανομή του Σκόρ είναι συμμετρική, αρνητικά ασύμμετρη ή θετικά ασύμμετρη;
- β. Μπορείτε να κατασκευάσετε ένα θηκόγραμμα ή λείπουν κάποια στοιχεία;
- γ. Πόση είναι η τυπική απόκλιση;
- δ. Ανάμεσα σε ποια Σκόρ Ποιότητας Ζωής, βρίσκεται το 95% των ηλικιωμένων της έρευνας;

αιτιολογήστε όλα τα παραπάνω (α,β,γ,δ)