

Στατιστική 2^η - διαλέξη 1

Περιγραφική Στατιστική

Έστω $\{X_1, X_2, \dots, X_n\}$ τυχαίο δείγμα n στοιχείων πληθυσμού X .

Τα X_i δεν είναι κατ' ανάγκη διαφορετικά μεταξύ τους.

Συχνότητα τιμής $X_i \sim f_i$: Αριθμός φορές που η τιμή X_i εμφανίζεται στο δείγμα μας.

Σχετική συχνότητα τιμής $X_i \sim f_i/n$: Αριθμός φορές εμφάνισης της τιμής X_i , προς αριθμό στοιχείων του δείγματος.

$$\text{Ισχύουν } \sum_{i=1}^k f_i = n, \quad \sum_{i=1}^k f_i/n = 1.$$

Αθροιστική συχνότητα τιμής $X_i \sim F_i$: Αριθμός φορές (αθροισμα) εμφάνισης τιμής μικρότερης ή ίσης της X_i .

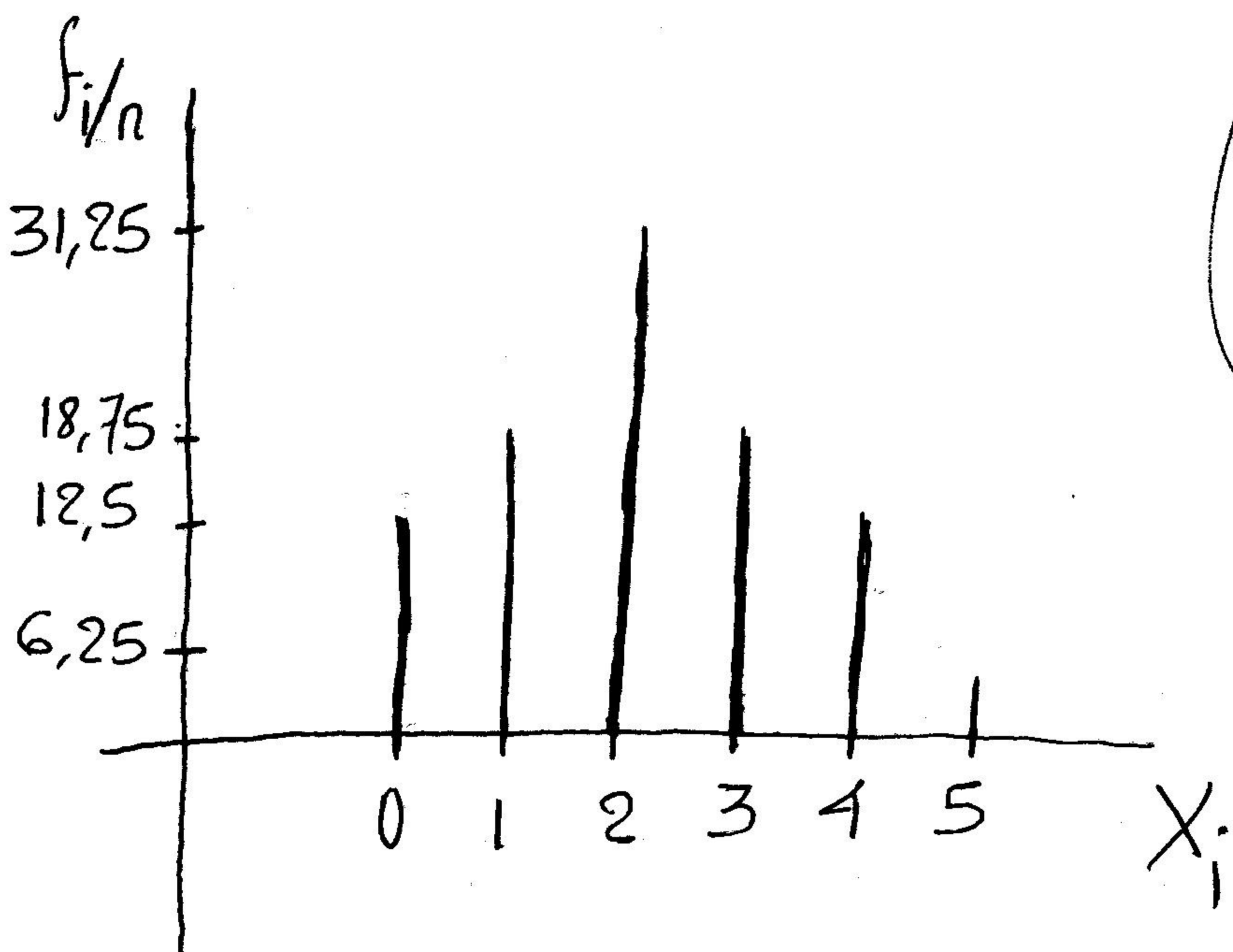
Αθροιστική σχετική συχνότητα τιμής $X_i \sim F_i/n$: Αριθμός φορές εμφάνισης (αθροισμα) εμφάνισης τιμής μικρότερης ή ίσης της X_i , προς αριθμό στοιχείων του δείγματος.

$$\text{Ισχύουν } F(n)_X = n, \quad F(n)/n_X = 1$$

Παράδειγμα 1].

{ 1, 0, 2, 1, 0, 2, 3, 2, 4, 3, 1, 2, 5, 3, 2, 2, 1, 2, 3, 2, 4, 1, 3, 0, 2, 5, 0, 3, 4, 1, 4, 2 } , 7α παιδιά 32 οικογενειών. = n = εὔρος δείγματος

Τιμή X_i (Αριθμός παιδιών)	Συχνότητα f_i	Σχ. Συχν. f_i/n	Αύρ. Συχν. F_i	Αύρ. Σχ. Συχν. F_i/n
0	4	$4/32 = 12,5\%$	4	$4/32 = 12,5\%$
1	6	$6/32 = 18,75\%$	10	$10/32 = 31,25\%$
2	10	$10/32 = 31,25\%$	20	$20/32 = 62,5\%$
3	6	$6/32 = 18,75\%$	26	$26/32 = 81,25\%$
4	4	$4/32 = 12,5\%$	30	$30/32 = 93,75\%$
5	2	$2/32 = 6,25\%$	<u>$32 = n$</u>	<u>$32/32 = 100\% = 1$</u>
	<u>$\sum f_i = 32 = n$</u>	<u>$\sum f_i/n = 100\% = 1$</u>		



(Η μπορώ να το αναπαράσω και με "πίτα"

Ραβδόγραμμα ($X_i \rightarrow f_i/n$)

Σχόλιο

Αν το εύρος n του δείγματος είναι ικανοποιητικά μεγάλο αριθμός (π.χ. $n > 100$), η προηγούμενη μέθοδος δε διευκολύνει την καταγραφή συχνοτήτων και την περαιτέρω μελέτη τους.

Σε τέτοιες περιπτώσεις ($n > 100$) δε μεριώ συχνοτήτες εμφάνισης τιμής X_i του δείγματος, αλλά μεριώ συχνοτήτα εμφάνισης ομάδας στην οποία περιέχεται η τιμή X_i (που είναι και το η/δόν λογικό)

Π.χ. Έστω ότι το δείγμα μου περιέχει τιμές X_i με $100 \leq X_i \leq 200$

Τότε χωρίσω το δείγμα μου σε επιμέρους ομάδες, διαφορετικών X_i η καθεμιά. Π.χ. $100 - 110$, $110 - 120$, ..., $190 - 200$ * και μεριώ η/δόν πόσες φορές συνάπτω αντιπρόσωπο των παραπάνω ομάδων στο αρχικό δείγμα. Π.χ. 3 φορές συνάπτησα στοιχείο X_i με $100 \leq X_i \leq 110$

* Η επιλογή των ομάδων που θα χωρίσω τις τιμές του δείγματός μου δεν είναι αυθαίρετη και προκαθορισμένη και εξαρτάται πάντα από τα δεδομένα μου. Οπότε απαιτείται πάντα να έχω ισογυκείς ομάδες.

Για αναλυτικότερη περιγραφή, βλέπε στο παραδειγμα Λούκας (18.02.16) Την έδωσα Συγγραφή σε φυλλάδιο.

Η γραφική αναπαράσταση θα δοθεί η/δόν μέσω ιστογράμματος.