



ΓΡΑΦΗΜΑΤΑ ΚΑΤΑΝΟΜΩΝ

Καθηγητής Αντώνης Κ. Τραυλός (B.A., M.A., Ph.D.)
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΛΟΠΟΝΝΗΣΟΥ
 Σχολή Επιστημών Ανθρώπινης Κίνησης και Ποιότητας Ζωής
 Τμήμα Οργάνωσης και Διαχείρισης Αθλητισμού

1

1

ΓΡΑΦΗΜΑΤΑ ΚΑΤΑΝΟΜΩΝ

- ▶ Μετά την πινακοποίηση και την αριθμητική ανάπτυξη μιας κατανομής ακολουθεί συνήθως η γραφηματική της ανάλυση, η οποία περιλαμβάνει μια σειρά από γεωμετρικές απεικονίσεις της μορφής της.
 - ▶ Οι απεικονίσεις αυτές (γραφήματα) χρησιμεύουν όχι μόνο για τον ποιοτικό έλεγχο της συμμετρικότητας και της κυρτότητας των κατανομών, αλλά και για την επίτευξη ορισμένων αναλύσεων, μέσω των οποίων είναι δυνατός ο κατά προσέγγιση υπολογισμός των βασικών αριθμητικών ιδιοτήτων τους.
- ▶ Η γραφηματική ανάλυση έχει εφαρμογές σε πληθυσμιακές ή εκτεταμένες βάσεις δειγματικών δεδομένων, με στόχο την καθιέρωση νορμών για μεγάλης σημασίας μεταβλητές.
 - ▶ Μερικές φορές μπορούμε να συνδυάσουμε τα γραφήματα δύο, τριών ή και περισσότερων κατανομών, προκειμένου να εντοπίσουμε κάποιες αρχικές ποιοτικές διαφορές τους, πριν προβούμε σε καθαρά στατιστική σύγκριση τους.

2

2

Τρόποι Γραφικής Αναπαράστασης

- ▶ Οι τρόποι γραφικής αναπαράστασης των κατανομών έχουν σχέση με τα αποτελέσματα που παράγονται μετά την πλήρη πινακοποίησή τους.
- ▶ Παραδείγματα αυτών των πινακοποιήσεων είδαμε στο προηγούμενο μάθημα στον πίνακα 2.5 για απλές κατανομές και στον πίνακα 2.6 για ομαδοποιημένες κατανομές.

Πίνακας 2.5 - Απλή κατανομή του βάρους: πλήρης ανάλυση (N = 115)

Κλάση (k)	Βάρος (kg) X	Συχνότητα f	Αθροιστική Συχνότητα c	Σχετική Συχνότητα %f	Αθροιστική Σχετική Συχνότητα c%f
1	64	1	1	0.9	0.9
2	65	3	4	2.6	3.5
3	66	0	4	0.0	3.5
4	67	1	5	0.9	4.3
5	68	3	8	2.6	7.0
6	69	5	13	4.3	11.3
7	70	5	18	4.3	15.6
8	71	6	24	5.2	20.9
9	72	6	30	5.2	26.1
10	73	2	32	1.7	27.8
11	74	8	40	7.0	34.8
12	75	13	53	11.3	46.1
13	76	7	60	6.1	52.2
14	77	9	69	7.8	60.0
15	78	3	72	2.6	62.6
16	79	10	82	8.7	71.3
17	80	3	85	2.6	73.9
18	81	7	92	6.1	80.0
19	82	5	97	4.3	84.3
20	83	4	101	3.5	87.8
21	84	4	105	3.5	91.3
22	85	5	110	4.3	95.6
23	86	1	111	0.9	96.5
24	87	1	112	0.9	97.4
25	88	2	114	1.7	99.1
26	89	0	114	0.0	99.1
27	90	1	115	0.9	100.0
--	Σf =	115	Σ%f =	100	--

Οι τιμές στις στήλες %f και c%f στρογγυλοποιήθηκαν στο 1ο δεκαδικό ψηφίο.

3

Απλή Κατανομή

- ▶ Από τα αποτελέσματα των πινακοποιήσεων των απλών κατανομών μπορούν να παραχθούν τέσσερα (4) διαφορετικά γραφήματα.
- ▶ Αυτά είναι η γραφική παράσταση της κατανομής των απόλυτων συχνοτήτων (f), η γραφική παράσταση της κατανομής των αθροιστικών συχνοτήτων (cf), η γραφική παράσταση της κατανομής των εκατοστιαίων (σχετικών) συχνοτήτων (%f) και η γραφική παράσταση της κατανομής των αθροιστικών εκατοστιαίων συχνοτήτων (c%f).

Πίνακας 2.5 - Απλή κατανομή του βάρους: πλήρης ανάλυση (N = 115)

Κλάση (k)	Βάρος (kg) X	Συχνότητα f	Αθροιστική Συχνότητα c	Σχετική Συχνότητα %f	Αθροιστική Σχετική Συχνότητα c%f
1	64	1	1	0.9	0.9
2	65	3	4	2.6	3.5
3	66	0	4	0.0	3.5
4	67	1	5	0.9	4.3
5	68	3	8	2.6	7.0
6	69	5	13	4.3	11.3
7	70	5	18	4.3	15.6
8	71	6	24	5.2	20.9
9	72	6	30	5.2	26.1
10	73	2	32	1.7	27.8
11	74	8	40	7.0	34.8
12	75	13	53	11.3	46.1
13	76	7	60	6.1	52.2
14	77	9	69	7.8	60.0
15	78	3	72	2.6	62.6
16	79	10	82	8.7	71.3
17	80	3	85	2.6	73.9
18	81	7	92	6.1	80.0
19	82	5	97	4.3	84.3
20	83	4	101	3.5	87.8
21	84	4	105	3.5	91.3
22	85	5	110	4.3	95.6
23	86	1	111	0.9	96.5
24	87	1	112	0.9	97.4
25	88	2	114	1.7	99.1
26	89	0	114	0.0	99.1
27	90	1	115	0.9	100.0
--	Σf =	115	Σ%f =	100	--

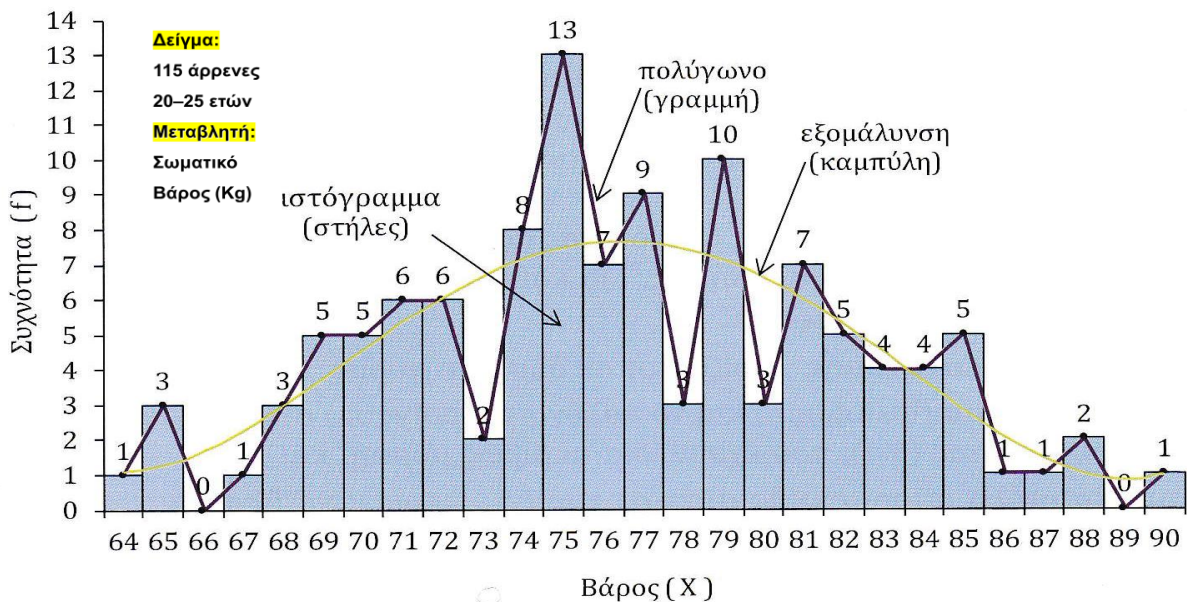
Οι τιμές στις στήλες %f και c%f στρογγυλοποιήθηκαν στο 1ο δεκαδικό ψηφίο.

4

Γράφημα κατανομής των απόλυτων συχνοτήτων

- ▶ Το γράφημα της κατανομής των **απόλυτων συχνοτήτων** (f) είτε ως **ιστόγραμμα** (εικ. 3.1) είτε ως **συχνοτικό πολύγωνο** (εικ. 3.2).
- ▶ Η διαφορά μεταξύ των δύο αυτών τύπων γραφικής απεικόνισης της μορφής της κατανομής των συχνοτήτων έγκειται μόνο στην εμφάνιση.
- ▶ Το ιστόγραμμα απεικονίζει τις διαφορετικές τιμές της κατανομής με μεγαλύτερη σαφήνεια από την καμπύλη του συχνοτικού πολυγώνου, η οποία λόγω της συνέχειάς της είναι κατάλληλη για εξομάλυνση και κατά συνέπεια για ρεαλιστικότερη παρουσίαση της μορφής της κατανομής σε ένα δεύτερο στάδιο ανάλυσης.
- ▶ Η περίπτωση αυτή αφορά κυρίως κατανομές με μεγάλο πλήθος, οι οποίες χρησιμοποιούνται ως πρότυπα (π.χ. νόρμες), για την εκτίμηση ποσοστού συμμετοχής επιλεγμένων τιμών της υπό εξέταση μεταβλητής (X) σε μελλοντικά δείγματα.

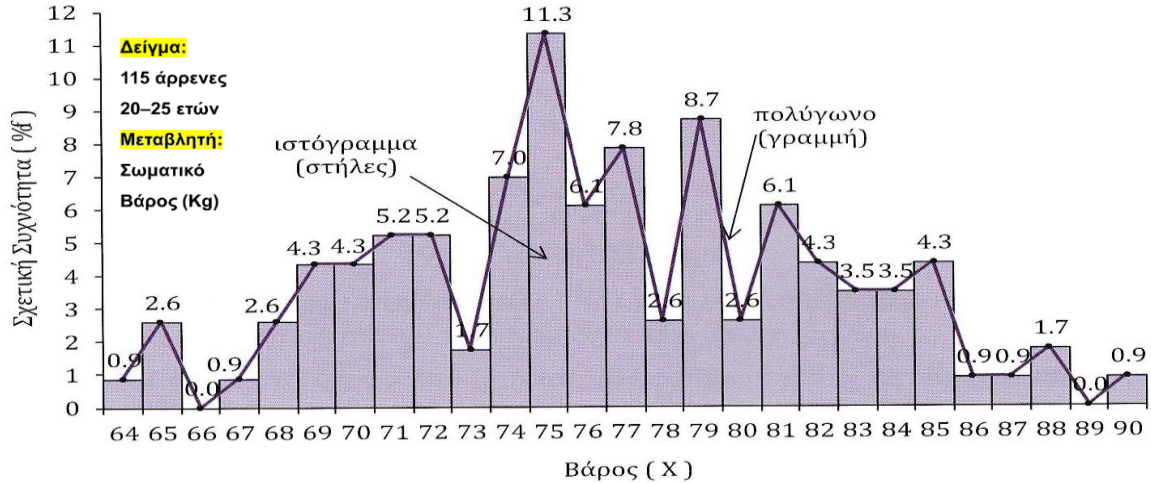
5



Σχήμα 3.1 - Ιστόγραμμα και πολύγωνο **συχνοτήτας** (f) της **απλής κατανομής** του Βάρους ($N = 115$)

6

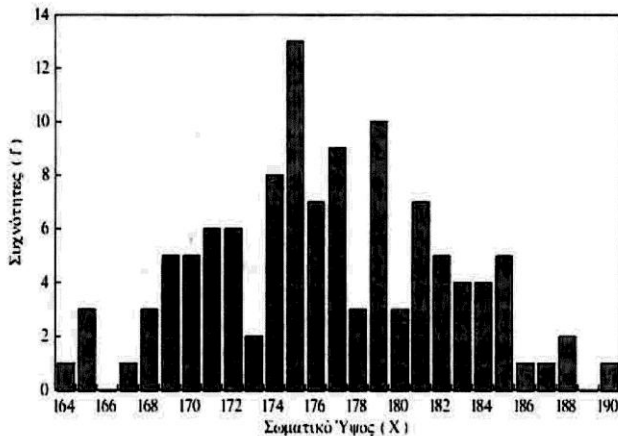
Ιστόγραμμα και πολύγωνο συχνότητας (f) της απλής κατανομής του βάρους (N=115)



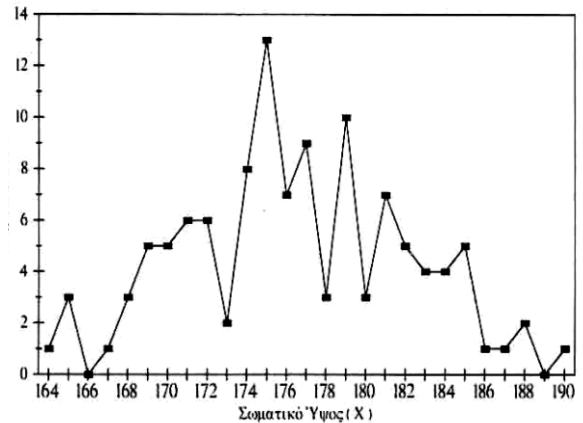
Σχήμα 3.2 - Ιστόγραμμα και πολύγωνο σχετικής συχνότητας (%f) της απλής κατανομής του Βάρους (N = 115)

7

Παραδείγματα γραφημάτων κατανομής των απόλυτων συχνοτήτων



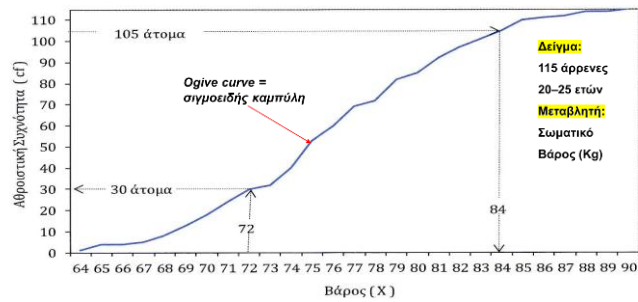
Εικ. 3.1 - Ιστόγραμμα συχνοτήτων κατανομής (N=115)



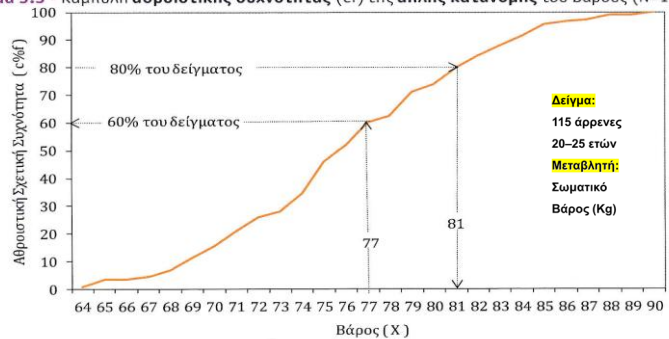
Εικ. 3.2 - Πολύγωνο συχνοτήτων κατανομής

8

- ▶ Στην επιστημονική έρευνα, πριν αποφασίσουμε ποια στατιστική ανάλυση είναι καταλληλότερη για τον έλεγχο μιας ερευνητικής υπόθεσης, ελέγχουμε αν η κατανομή της υπό διερεύνηση μεταβλητής ιδιότητας (X) πληροί κάποιες προϋποθέσεις (στατιστικές παραδοχές).
- ▶ Μία απ' αυτές είναι η **κανονικότητα της κατανομής**, δηλαδή κατά πόσο η κατανομή είναι συμμετρική και μεσόκυρτη. Τον έλεγχο αυτό τον κάνουμε αρχικά μέσω του ιστογράμματος της κατανομής.



Σχήμα 3.3 - Καμπύλη αθροιστικής συχνότητας (cf) της απλής κατανομής του Βάρους (N=115)



Σχήμα 3.4 - Καμπύλη αθροιστικής σχετικής συχνότητας (c%f) της απλής κατανομής του Βάρους (N = 115)

9

Πίνακας 2.6 - Ομαδοποιημένη κατανομή του Βάρους: ανάλυση (N = 115)

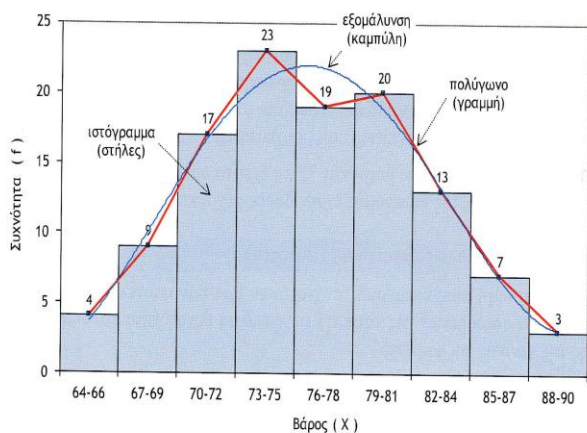
Κλάση	Διάστημα Τιμών	Μεσαία Τιμή	Συχνότητα	Αθροιστική Συχνότητα	Σχετική Συχνότητα	Αθροιστική Σχετική Συχνότητα
k	$X_L - X_U$	X	f	cf	%f	c%f
1	64-66	65	4	4	3.5	3.5
2	67-69	68	9	13	7.8	11.3
3	70-72	71	17	30	14.8	26.1
4	73-75	74	23	53	20.0	46.1
5	76-78	77	19	72	16.5	62.6
6	79-81	80	20	92	17.4	80.0
7	82-84	83	13	105	11.3	91.3
8	85-87	86	7	112	6.1	97.4
9	88-90	89	3	115	2.6	100.0
Σf =			115	Σ%f =	100	--

Ομαδοποιημένη Κατανομή Συχνότητας

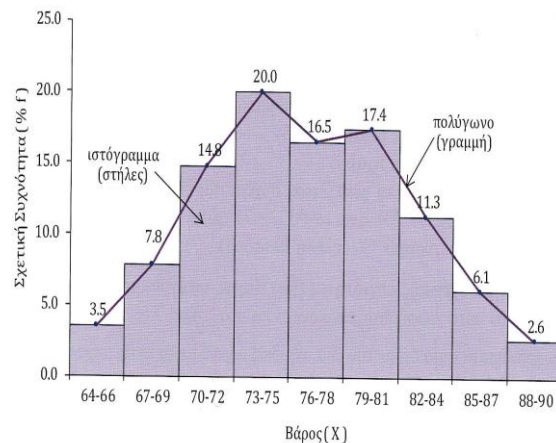
10

Ιστόγραμμα και πολύγωνο συχνότητας (f) της ομαδοποιημένης κατανομής του βάρους (N=115)

Δείγμα:
115 άρρενες
20–25 ετών
Μεταβλητή:
Σωματικό
Βάρος (Kg)



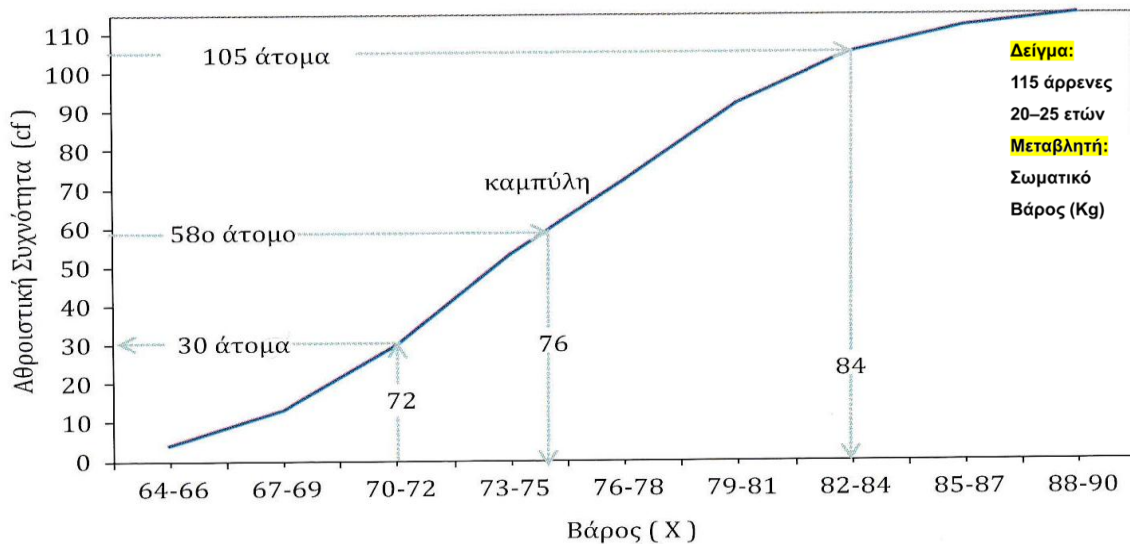
Σχήμα 3.5 - Ιστόγραμμα και πολύγωνο συχνότητας (f) της ομαδοποιημένης κατανομής του βάρους (N = 115)



Σχήμα 3.6 - Ιστόγραμμα και πολύγωνο σχετικής συχνότητας (%f) της ομαδοποιημένης κατανομής του βάρους (N = 115)

11

Υπολογισμός τιμής X από την αθροιστική συχνότητα (cf)

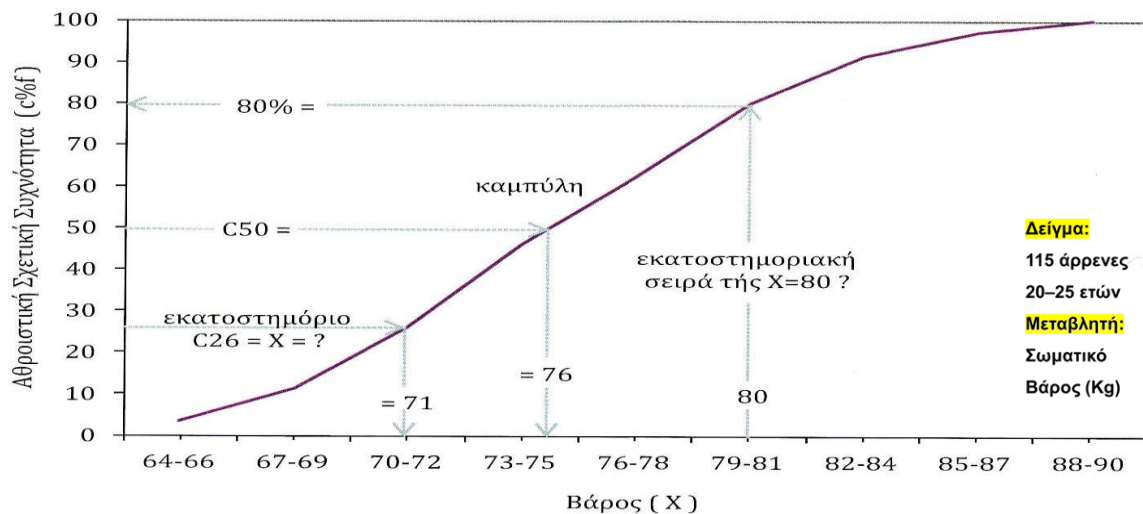


Δείγμα:
115 άρρενες
20–25 ετών
Μεταβλητή:
Σωματικό
Βάρος (Kg)

Σχήμα 3.7 - Καμπύλη αθροιστικής συχνότητας (cf) της ομαδοποιημένης κατανομής του βάρους (N = 115)

12

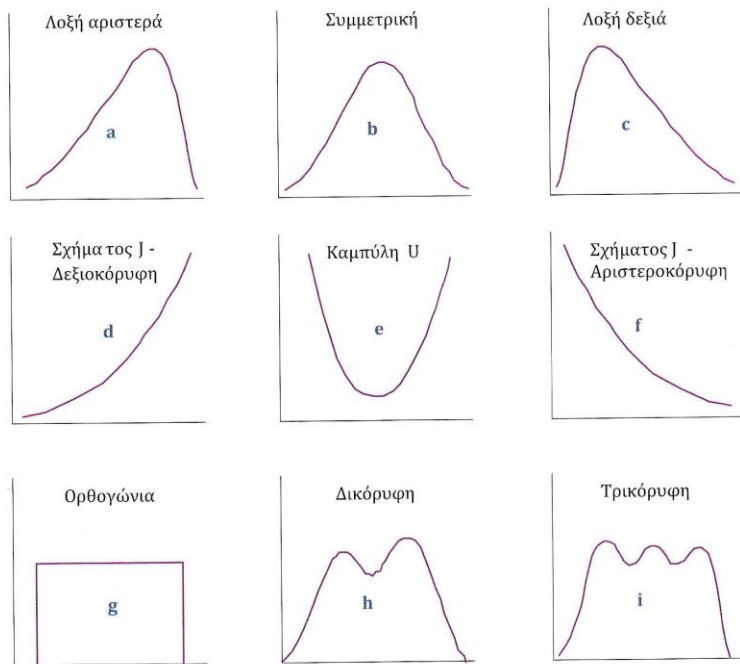
Υπολογισμός τιμής X από την αθροιστική σχετική



Σχήμα 3.8 - Καμπύλη αθροιστικής σχετικής συχνότητας (ogive, c%f) της ομαδοποιημένης κατανομής του Βάρους ($N = 115$): υπολογισμός τιμής X από εκατοστιαία σειρά (%) και αντίστροφα

13

Καμπύλες των πιο αντιπροσωπευτικών μορφών κατανομών συχνότητας

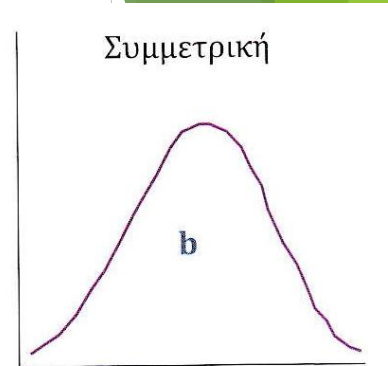


Σχήμα 3.9 - Διαφορετικοί τύποι κατανομών (πολύγωνα συχνότητας)

14

Τύποι & Ιδιότητες Κατανομών

- ▶ Παρατηρώντας τις πινακοποιήσεις και ιδιαίτερα τις γραφικές παραστάσεις ενός μεγάλου αριθμού κατανομών συχνότητας, διαπιστώνουμε ευκολα ότι, ενώ η γενική τάση είναι οι περισσότερες κατανομές να έχουν ομαλή μορφή (συμμετρική καμπύλη), μεταξύ τους παρουσιάζουν πολλές φορές αξιοσημείωτες διαφορές.
 - ▶ Όπως θα δούμε στο κεφάλαιο της σύγκρισης των μέσων τιμών, οι διαφορές αυτές μπορούν να ποσοτικοποιηθούν και να ελεγχθούν αν είναι στατιστικά σημαντικές.
- ▶ Η καμπύλη b αποτελεί παράδειγμα συμμετρικής (κωδωνοειδούς) κατανομής της οποίας ένα σπουδαίο παράδειγμα αποτελεί η κανονική καμπύλη.
 - ▶ Οι καμπύλες αυτές εκφράζουν τη συντριπτική πλειοψηφία των μεταβλητών που συναντώνται σε στατιστικές αναλύσεις, ιδιαίτερα τις μεταβλητές των οποίων οι τιμές μετρούνται με κλίμακες λόγου, έτσι που από τις μετρήσεις αυτές να προκύπτουν λίγες σχετικά τιμές με μικρό και μεγάλο μέγεθος (στα άκρα της κατανομής) και πολλές τιμές με μέτριο μέγεθος (στο κέντρο της κατανομής).



15

Γενικές ιδιότητες των κατανομών

- ▶ Ο τρόπος που οι συχνοτικές κατανομές διαφέρουν μεταξύ τους στηρίζεται στις τέσσερις γενικές ιδιότητες των κατανομών.
- ▶ Αυτές είναι:
 1. Η ιδιότητα της **κεντρικής τάσης**, που αφορά στις κεντρικές τιμές (μέσους όρους) της εξεταζόμενης μεταβλητής.
 2. Η ιδιότητα της **διασποράς**, που καθορίζει πόσο διάσπαρτες είναι οι τιμές της μεταβλητής γύρω από τις κεντρικές τιμές.
 3. Η ιδιότητα της **λοξότητας (ασυμμετρίας)**, που δείχνει αν η καμπύλη της κατανομής είναι ασύμμετρη και σε ποιο άκρο της κατανομής συγκεντρώνονται λιγότερες αρχικές τιμές (ουρά της καμπύλης):
 - ▶ στο άκρο των μικρών τιμών (αρχή της κλίμακας της κατανομής), οπότε υπάρχει **αρνητική λοξότητα** ή λοξότητα αριστερά, ή
 - ▶ στο άκρο των μεγάλων τιμών (τέλος της κλίμακας της κατανομής), οπότε υπάρχει **θετική λοξότητα** ή λοξότητα δεξιά.
 4. Η ιδιότητα της **κυρτότητας** της καμπύλης της κατανομής, η οποία αναφέρεται στο αν η κατανομή παρουσιάζει μορφή
 - ▶ **λεπτόκυρτη**, με υψηλές συχνότητες σε λίγες μόνο τιμές,
 - ▶ **μεσόκυρτη**, με μέτριες συχνότητες στις περισσότερες τιμές, ή
 - ▶ **πλατόκυρτη**, με μικρές συχνότητες σε πολλές τιμές.

16

16

Παράγοντες που επηρεάζουν την πυκνότητα & την κανονικότητα των καμπυλών

- ▶ Στο σημείο αυτό ας δούμε ποιοι παράγοντες επηρεάζουν την **πυκνότητα** και την **κανονικότητα** των καμπυλών αυτών.
- ▶ Οι κυριότεροι παράγοντες είναι
 - ▶ το πλήθος N της κατανομής και
 - ▶ ο αριθμός των ομάδων k ή αντίστροφα το (κοινό) εύρος E_k των ομάδων όπως προκύπτει από την ομαδοποίηση της.
- ▶ Ένας δευτερεύων παράγοντας είναι η **ανάλυση** (πυκνότητα) της κλίμακας μέτρησης της κατανομής.
- ▶ Σχετικά με τον παράγοντα **πλήθος κατανομής**, γενικά, όσο μικρότερο είναι το πλήθος μιας κατανομής, τόσο περισσότερο ακανόνιστη εμφανίζεται η μορφή της, ενώ αντίθετα, όσο το πλήθος της κατανομής γίνεται μεγαλύτερο τόσο η μορφή της παρουσιάζεται όλο και περισσότερο εξομαλυσμένη.

17

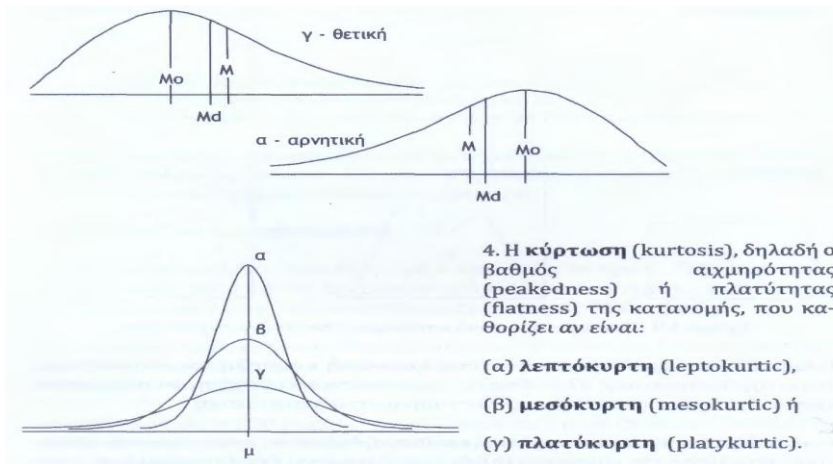
17

Οι κατανομές συχνότητας έχουν 4 θεμελιώδη γνωρίσματα.

1. Η **κεντρική τάση (central location)**, που αφορά στις κεντρικές τιμές (μέσους όρους) της μεταβλητής: μέσος (M), διάμεσος (M_d), κορυφή ή επικρατούσα τιμή (M_o).
2. Η **διασπορά (dispersion)**, που δείχνει πόσο διάσπαρτες είναι οι τιμές της μεταβλητής γύρω από τις κεντρικές τιμές: εύρος (E), διασπορά (s^2), τυπική απόκλιση (SD), μεταβλητότητα (CV), τιμές z .
3. Η **λοξότητα ή ασυμμετρία (skewness, asymmetry)**, που δείχνει αν η κατανομή είναι: (α) λοξή αριστερά, δηλαδή αρνητικά ασύμμετρα (left-, negatively skewed), (β) συμμετρική (symmetrical, σχ. 3.9b), ή (γ) λοξή δεξιά, δηλαδή θετικά ασύμμετρα (right-, positively skewed).

18

18



- ▶ Η **λοξότητα** και η **κύρτωση** αξιολογούνται **ποιοτικά** με διαγράμματα (π.χ. ιστόγραμμα), **ποσοτικά** με δείκτες (S_k , K_u) και τα τυπικά τους σφάλματα (standard errors) και **πιθανολογικά** με ειδικούς ελέγχους (π.χ. Shapiro-Wilk). Στο κεφ. 6 θα δούμε πως υπολογίζονται και στο κεφάλαιο 7 πως παράγονται με το στατιστικό πρόγραμμα SPSS στο πλαίσιο της πλήρους περιγραφικής στατιστικής ανάλυσης.

19

19

Με το αρχείο που βρίσκεται στην e-class (ΔΕΔΟΜΕΝΑ ΒΑΡΟΥΣ N_140) ...

.... μπορείτε να αρχίσετε του πειραματισμούς σας στο SPSS.



**Σας ευχαριστώ για την προσοχή
σας**

AKT

20