ΑΣΥΜΜΕΤΡΙΑ

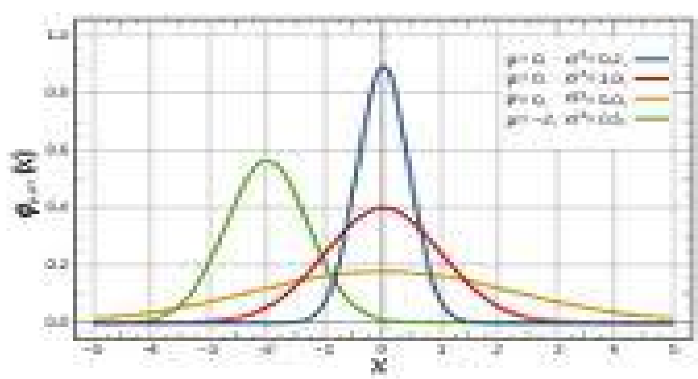
Ας υποθέσουμε, ότι κατά την μελέτη της κατανομής δύο μεταβλητών, καταλήγουμε στα παρακάτω ιστογράμματα.

Στα παραπάνω ιστογράμματα , παρατηρούμε , ότι αν και υπάρχει διαφορά στη διασπορά των τιμών (διαφορετική μεταβλητικότητα) κατά τα άλλα, υπάρχουν πολλές ομοιότητες:

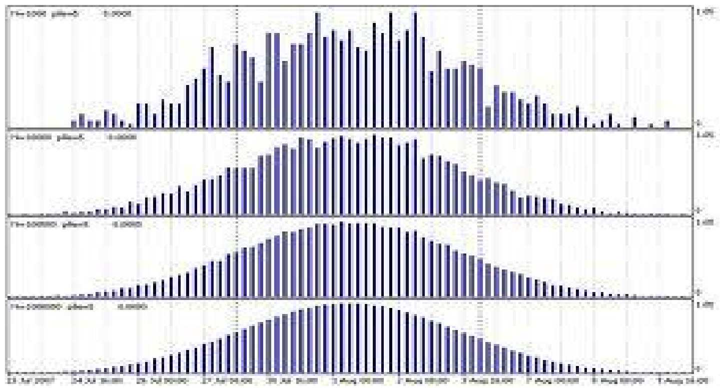
α) και τα δύο είναι συμμετρικά, με άξονα συμμετρίας κοντά στο 20

β) καταλήγουν ομαλά και προς τις δύο πλευρές, χωρίς να υπάρχουν απομονωμένες τιμές

Αν τώρα σχεδιάσουμε τα αντίστοιχα πολύγωνα συχνοτήτων, τότε θα παρατηρήσουμε ότι αυτά "τείνουν" να ταυτιστούν με μία καμπύλη, που περνάει από τις κορυφές των στηλών
Αυτές, οι καμπύλες, ονομάζονται **κανονικές καμπύλες** και έχουν σχήματα της μορφής των παρακάτω σχεδιαγραμμάτων.



Πρακτικά , η ταύτιση αυτή (πολυγώνου-καμπύλης) είναι εφικτή , για κατανομές , συνεχών μεταβλητών, με άπειρο πλήθος τιμών.
Αποδεικνύεται, ότι όλες οι κατανομές, όταν αναφέρονται σε πολύ μεγάλο αριθμό επαναλήψεων (όταν το μέγεθος του δείγματος είναι πολύ μεγάλο, δηλ. $\nu \to +\infty$)ταυτίζονται τελικά ,είτε απευθείας είτε με έναν απλό μετασχηματισμό, με την κανονική κατανομή, που έχει γραφική παράσταση κανονικής καμπύλης.



Συχνά όμως, μία κατανομή δεν είναι συμμετρική. Δεν υπάρχει τιμή της μεταβλητής, ως προς την οποία η συχνότητα να είναι συμμετρική
Σε τέτοιες περιπτώσεις, η κατανομή παρουσιάζει συχνά μία "ουρά" προς τα δεξιά ή προς τα αριστερά
Τότε λέμε ότι έχουμε **κατανομή ασύμμετρη προς τα δεξιά** ή **ασύμμετρη κατανομή προς τα αριστερά** , αντίστοιχα

Ροπές κατανομής συχνοτήτων

Γνωρίζουμε από τα προηγούμενα , ότι ο αριθμητικός μέσος των τιμών μιας μεταβλητής Χ , $\overline{\chi} = \dfrac{\sum_{i=1}^{\nu} \chi_i}{\nu}$, και η διασπορά (διακύμανση) $V(\chi) = \dfrac{\sum_{i=1}^{\nu} (\chi_i - \overline{\chi})^2}{\nu}$ μας κατατοπίζουν, σε μεγάλο βαθμό, για τη θέση και τη διασπορά των τιμών μιας κατανομής
Υπάρχουν όμως κατανομές, όπως οι ασύμμετρες, στις οποίες χρειάζονται , περισσότερες χαρακτηριστικές σταθερές , για τον καθορισμό αυτών των θέσεων
Ορίζουμε ως **ροπή τ τάξεως της μεταβλητής Χ περί την αρχή** την ποσότητα $\mu'_\tau = \dfrac{\sum_{i=1}^{\nu} \chi_i^{\tau}}{\nu}$ όπου τ = 0,1,2,3,…..
Είναι προφανές ότι : $\mu'_0 = 1$ και $\mu'_1 = \overline{\chi}$

Με τον ίδιο τρόπο ορίζεται και η ροπή τ τάξεως της μεταβλητής X ως προς την μέση τιμή ή αλλιώς κεντρική ροπή τ τάξεως , ως η ποσότητα

$$\mu_\tau = \frac{\sum_{i=1}^{\nu}(\chi_i - \overline{\chi})^\tau}{\nu}$$ όπου τ = 0,1,2,3,…..

Είναι προφανές ότι : $\mu_0 = 1$, $\mu_1 = 0$ και $\mu_2 = \frac{\sum_{i=1}^{\nu}(\chi_i - \overline{\chi})^2}{\nu} = V(\chi)$

Επιστρέφοντας , τώρα στην ασυμμετρία μιας κατανομής, μπορούμε να καθορίσουμε την ποσότητα που μπορεί να χρησιμοποιηθεί σαν μέτρο ασυμμετρίας.
Ονομάζουμε **συντελεστή ασυμμετρίας** ή **<u>ασυμμετρία</u>** ή **συντελεστή λοξότητας** ή απλά **<u>λοξότητα</u>** την ποσότητα :

$$\alpha_3 = \frac{\sum_{i=1}^{\nu}(\chi_i - \overline{\chi})^2}{\nu} \div s^3 = \frac{\mu^3}{s^3}$$ όπου s η τυπική απόκλιση ($\sqrt{V(\chi)}$)

- Αν $\alpha_3 > 0$ τότε η κατανομή είναι ασύμμετρη προς τα δεξιά
- Αν $\alpha_3 < 0$ τότε η κατανομή είναι ασύμμετρη προς τα αριστερά
- Αν $\alpha_3 = 0$ τότε η κατανομή είναι συμμετρική

Κύρτωση κατανομής

Αν παρατηρήσουμε τα σχήματα 3 και 4 βλέπουμε ότι οι κατανομές μπορεί να είναι συγκεντρωμένες κοντά στη μέση, οπότε η καμπύλη που την παριστάνει, παρουσιάζει μία έντονη κύρτωση (καμπούριασμα) όπως στο σχήμα 4
Μπορεί , όμως, μία κατανομή να μην είναι συγκεντρωμένη κατά αυτό τον τρόπο, οπότε η καμπύλη, δεν παρουσιάζει σαφή κύρτωση, όπως στο σχήμα 3
Μία ποσότητα που μπορεί να χρησιμοποιηθεί σαν μέτρο της κύρτωσης, καλείται **συντελεστής κύρτωσης** ή απλά **<u>κύρτωση</u>**

Μία τέτοια ποσότητα, είναι η : $\alpha_4 = \frac{\frac{\sum_{i=1}^{\nu}(\chi_i - \overline{\chi})^4}{\nu}}{s^4} = \frac{\mu^4}{s^4}$

Η ποσότητα αυτή, συνήθως συγκρίνεται με τον αριθμό 3, που είναι ο αντίστοιχος συντελεστής κύρτωσης για την κανονική κατανομή

Συνεχείς κατανομές

Στην περίπτωση των συνεχών κατανομών, δεν υπάρχει μεγάλη διαφοροποίηση των τύπων, οι οποίοι μετατρέπονται ως εξής :

$$\mu'_\tau = \int_{-\infty}^{+\infty} \chi^\tau f(\chi) dx$$

$$\mu_\tau = \int_{-\infty}^{+\infty} (\chi - \overline{\chi})^\tau f(\chi) dx$$

$$\alpha_3 = \frac{\int_{=\infty}^{+\infty} (\chi - \overline{\chi})^3 f(\chi) dx}{s^3}$$

$$\alpha_4 = \frac{\int_{=\infty}^{+\infty} (\chi - \overline{\chi})^4 f(\chi) dx}{s^4}$$

όπου $f(\chi)$ η συνάρτηση που μας δίνει την τιμή της μεταβλητής (σχετική συχνότητα)

<u>Εφαρμογές</u>

1) Αν $f(\chi) = \begin{cases} 2\chi & \text{αν } 0<\chi<1 \\ 0 & \text{αλλού} \end{cases}$

Να υπολογιστούν : η διακύμανση , η κεντρική ροπή 3ης τάξης , η κεντρική ροπή 4ης τάξης , η ασυμμετρία και η κύρτωση

<u>Απάντηση</u>

- $\overline{\chi} = \mu'_1 = \int_{-\infty}^{+\infty} \chi f(\chi) dx = \int_0^1 \chi \cdot 2\chi dx = \int_0^1 2\chi^2 dx = \left[\frac{2\chi^3}{3}\right]_0^1 = \frac{2 \cdot 1^3}{3} - \frac{2 \cdot 0^3}{3} = \frac{2}{3}$
- $V(\chi) = \mu_2 = \int_{-\infty}^{+\infty} (\chi - \overline{\chi})^2 dx = \int_0^1 \left(\chi - \frac{2}{3}\right)^2 2\chi dx = \int_0^1 \left(2\chi^3 - \frac{8}{3}\chi^2 + \frac{8}{9}\chi\right) dx = \left[\frac{\chi^4}{2} - \frac{8\chi^3}{9} + \frac{4\chi^2}{9}\right]_0^1 =$
  $= \left(\frac{1^4}{2} - \frac{8 \cdot 1^3}{9} + \frac{4 \cdot 1^2}{9}\right) - \left(\frac{0^4}{2} - \frac{8 \cdot 0^3}{9} + \frac{4 \cdot 0^2}{9}\right) = \frac{1}{2} - \frac{8}{9} + \frac{4}{9} = \frac{1}{18}$
- $\mu_3 = \int_{-\infty}^{+\infty} (\chi - \overline{\chi})^3 f(\chi) dx = \int_0^1 \left(\chi - \frac{2}{3}\right)^3 2\chi dx = \int_0^1 \left(\chi^3 - 2\chi^2 + \frac{4}{3}\chi - \frac{8}{27}\right) 2\chi dx =$

$$= \int_0^1 \left(2\chi^4 - 4\chi^3 + \frac{8}{3}\chi^2 - \frac{16}{27}\chi \right) dx = \left[\frac{2\chi}{5} - \chi^4 + \frac{8\chi^3}{9} - \frac{8\chi^2}{27} \right]_0^1 = -\frac{1}{135}$$

- $\mu_4 = \int_{-\infty}^{+\infty} (\chi - \overline{\chi})^4 f(\chi) dx = \int_0^1 \left(\chi - \frac{2}{3} \right)^4 2\chi dx = \int_0^1 \left(\chi^4 - \frac{8}{3}\chi^3 + \frac{8}{3}\chi^2 - \frac{32}{27}\chi + \frac{16}{81} \right) 2\chi dx =$

$$= \int_0^1 \left(2\chi^5 - \frac{16}{3}\chi^4 + \frac{16}{3}\chi^3 - \frac{64}{27}\chi^2 + \frac{32}{81}\chi \right) dx = \left[\frac{\chi^6}{3} - \frac{16\chi^5}{15} + \frac{4\chi^4}{3} - \frac{64\chi^3}{81} + \frac{16\chi^2}{81} \right]_0^1 = \frac{13}{105}$$

Επειδή $s = \sqrt{V(\chi)} = \sqrt{\frac{1}{18}} = \frac{\sqrt{2}}{6} = 0,2357$ άρα είναι

$$\alpha_3 = \frac{\mu_3}{s^3} = \frac{-\frac{1}{135}}{\left(\frac{\sqrt{2}}{6} \right)^3} = -\frac{4\sqrt{2}}{5} = -1,1314 \quad \text{και}$$

$$\alpha_4 = \frac{\mu_4}{s^4} = \frac{\frac{13}{105}}{\left(\frac{\sqrt{2}}{6} \right)^4} = \frac{1404}{35} = 40,1143$$

2) Μία μεταβλητή Χ, παίρνει τις τιμές 1,2,3 με αντίστοιχες σχετικές συχνότητες $\frac{1}{6}, \frac{1}{3}, \frac{1}{2}$. Να βρεθούν η μέση τιμή , η διακύμανση , η κεντρική ροπή 3ης τάξης και η ασυμμετρία

Απάντηση

Αν f_1, f_2, f_3 οι σχετικές συχνότητες τότε , ως γνωστόν , είναι :

$$f_1 + f_2 + f_3 = 1 \Leftrightarrow \frac{\nu_1}{\nu} + \frac{\nu_2}{\nu} + \frac{\nu_3}{\nu} = 1$$

Αλλά και $f_1 = \frac{\nu_1}{\nu} \Rightarrow \frac{\nu_1}{\nu} = \frac{1}{6} \Rightarrow \nu_1 = \frac{\nu}{6}$

$$f_2 = \frac{\nu_2}{\nu} \Rightarrow \frac{\nu_2}{\nu} = \frac{1}{3} \Rightarrow \nu_2 = \frac{\nu}{3}$$

$$f_3 = \frac{\nu_3}{\nu} \Rightarrow \frac{\nu_3}{\nu} = \frac{1}{2} \Rightarrow \nu_3 = \frac{\nu}{2}$$

Οπότε $\overline{\chi} = \dfrac{\sum_{i=1}^{3} \nu_i \chi_i}{\nu} = \dfrac{\nu_1\chi_1 + \nu_2\chi_2 + \nu_3\chi_3}{\nu} = \dfrac{\frac{\nu}{6}\cdot 1 + \frac{\nu}{3}\cdot 2 + \frac{\nu}{2}\cdot 3}{\nu} = \dfrac{1}{6} + \dfrac{2}{3} + \dfrac{3}{2} = \dfrac{7}{3}$

$$V(\chi) = \frac{\sum_{i=1}^{3} \nu_i (\chi_i - \overline{\chi})^2}{\nu} = \frac{\nu_1(\chi_1 - \overline{\chi})^2 + \nu_2(\chi_2 - \overline{\chi})^2 + \nu_3(\chi_3 - \overline{\chi})^2}{\nu} =$$

$$= \frac{\frac{\nu}{6}\left(1 - \frac{7}{3}\right)^2 + \frac{\nu}{3}\left(2 - \frac{7}{3}\right)^2 + \frac{\nu}{2}\left(3 - \frac{7}{3}\right)^2}{\nu} = \frac{\frac{16\nu}{54} + \frac{\nu}{27} + \frac{4\nu}{18}}{\nu} = \frac{30}{54} = \frac{5}{9}$$

$$\mu_3 = \frac{\sum_{i=1}^{3} \nu_i (\chi_i - \overline{\chi})^3}{\nu} = \frac{\nu_1(\chi_1 - \overline{\chi})^3 + \nu_2(\chi_2 - \overline{\chi})^3 + \nu_3(\chi_3 - \overline{\chi})^3}{\nu} =$$

$$= \frac{\frac{\nu}{6}\left(1 - \frac{7}{3}\right)^3 + \frac{\nu}{3}\left(2 - \frac{7}{3}\right)^3 + \frac{\nu}{2}\left(3 - \frac{7}{3}\right)^3}{\nu} = \frac{-\frac{64\nu}{162} - \frac{\nu}{81} + \frac{8\nu}{54}}{\nu} = -\frac{42}{162} = -\frac{7}{27}$$

και $\alpha_3 = \dfrac{\mu_3}{s^3} = \dfrac{\mu_3}{\left(\sqrt{V(\chi)}\right)^3} = \dfrac{-\frac{7}{27}}{\left(\sqrt{\frac{5}{9}}\right)^3} = -\dfrac{7\sqrt{5}}{25}$

3) Κατά την μελέτη του ύψους , ενός δείγματος φυτών, ενός θερμοκηπίου , βρέθηκαν οι τιμές που αναφέρονται στον παρακάτω πίνακα

Ύψος φυτών σε cm	Αριθμός
0-10	4
10-20	3
20-30	2
30-40	1

Να υπολογιστούν : Η διακύμανση , η κεντρική ροπή 3ης τάξης, η κεντρική ροπή 4ης τάξης , η ασυμμετρία και η κύρτωση της κατανομής

<u>Απάντηση</u>

Ύψος	πλήθος ν_i	Κεντρ. τιμή χ_i	$\nu_i\chi_i$	$\chi_i - \overline{\chi}$	$(\chi_i - \overline{\chi})^2$	$\nu_i(\chi_i - \overline{\chi})^2$	$(\chi_i - \overline{\chi})^3$	$\nu_i(\chi_i - \overline{\chi})^3$	$(\chi_i - \overline{\chi})^4$	$\nu_i(\chi_i - \overline{\chi})^4$
0-10	4	5	20	-10	100	400	-1000	-4000	10000	40000
10-20	3	15	45	0	0	0	0	0	0	0
20-30	2	25	50	10	100	200	1000	2000	10000	20000
30-40	1	35	35	20	400	400	8000	8000	160000	160000
Σύνολο	10		150			1000		6000		220000

Ο πίνακας κατανομής των συχνοτήτων της κατανομής είναι ο παρακάτω :

Η μέση τιμή της κατανομής είναι : $\overline{\chi} = \dfrac{\sum_{i=1}^{4} \nu_i \chi_i}{\nu} = \dfrac{150}{10} = 15$

Η διακύμανση της κατανομής είναι : $V(\chi) = \dfrac{\sum_{i=1}^{4} \nu_i \left(\chi_i - \overline{\chi}\right)^2}{\nu} = \dfrac{1000}{10} = 100$

Η τυπική απόκλιση είναι : $s = \sqrt{V(\chi)} = \sqrt{100} = 10$

Η κεντρική ροπή 3ης τάξης είναι : $\mu_3 = \dfrac{\sum_{i=1}^{4} \nu_i \left(\chi_i - \overline{\chi}\right)^3}{\nu} = \dfrac{6000}{10} = 600$

Η κεντρική ροπή 4ης τάξης είναι : $\mu_4 = \dfrac{\sum_{i=1}^{4} \nu_i \left(\chi_i - \overline{\chi}\right)^4}{\nu} = \dfrac{220000}{10} = 22000$

Ο συντελεστής ασυμμετρίας (λοξότητα) είναι : $\alpha_3 = \dfrac{\mu_3}{s^3} = \dfrac{600}{10^3} = 0{,}6$

Ο συντελεστής κύρτωσης είναι : $\alpha_4 = \dfrac{\mu_4}{s^4} = \dfrac{22000}{10^4} = 2{,}2$

Επειδή $\alpha_3 = 0{,}6 > 0$ η κατανομή είναι ελαφρώς ασύμμετρη προς τα δεξιά

Επίσης , αφού $\alpha_4 = 2{,}2 < 3$ η καμπύλη έχει κορυφή ελαφρώς πιο οξεία της κανονικής κατανομής