

I. Το άθροισμα

$$\alpha_1 + \alpha_2 + \cdots + \alpha_\nu$$

συμβολίζεται ως :

$$\sum_{i=1}^{\nu} \alpha_i \quad \text{ή} \quad \sum_{j=1}^{\nu} \alpha_j \quad \text{ή} \quad \sum_{\kappa=1}^{\nu} \alpha_{\kappa} \quad \text{ή} \quad \text{πιο απλά} \quad \sum \alpha_i \quad \text{ή} \quad \sum \alpha_j \quad \text{ή} \quad \sum \alpha_{\kappa}$$

όταν τα όρια της άθροισης εννοούνται. Δηλαδή:

$$\alpha_1 + \alpha_2 + \cdots + \alpha_\nu = \sum_{i=1}^{\nu} \alpha_i$$

II. Παραδείγματα:

1) Να γράψετε συμβολικά τα παρακάτω αθροίσματα:

a) $\alpha_1 + \alpha_2 + \cdots + \alpha_{100}$

b) $\chi_{10} + \chi_{11} + \cdots + \chi_{20}$

c) $\chi_1^2 + \chi_2^2 + \cdots + \chi_{50}^2$

d) $(\alpha_1 - \mu)^2 + (\alpha_2 - \mu)^2 + \cdots + (\alpha_{100} - \mu)^2$

e) $\alpha_1\beta_2 + \alpha_2\beta_3 + \cdots + \alpha_{100}\beta_{101}$

2) Να αναλύσετε τα παρακάτω συμβολικά αθροίσματα:

$\alpha)$ $\sum_{i=1}^{10} \alpha_i$

$\beta)$ $\sum_{i=1}^{10} \chi_i^2$

$\gamma)$ $\sum_{i=1}^{10} \chi_i \gamma_i$

$$\delta) \sum_{i=1}^{10} (\chi_i + m)$$

III. Ιδιότητες του συμβόλου Σ

$$1) \sum_{i=1}^v (\alpha_i + \beta_i) = \sum_{i=1}^v \alpha_i + \sum_{i=1}^v \beta_i$$

$$2) \sum_{i=1}^v \lambda \cdot \alpha_i = \lambda \cdot \sum_{i=1}^v \alpha_i, \quad \text{όπου } \lambda \in \mathbb{R}$$

$$3) \sum_{i=1}^v (\alpha_i + \mu) = \sum_{i=1}^v \alpha_i + v\mu, \quad \text{όπου } \mu \in \mathbb{R}$$

$$4) \sum_{i=1}^v (\chi_i - \bar{\chi})(y_i - \bar{y}) = \sum_{i=1}^n \chi_i y_i - v \cdot \bar{\chi} \cdot \bar{y}$$

$$5) \sum_{i=1}^v (\chi_i - \bar{\chi})^2 = \sum_{i=1}^n \chi_i^2 - v\bar{\chi}^2$$

Τύποι από την Περιγραφική Στατιστική

Ας υποθέσουμε ότι έχουμε ένα δείγμα μεγέθους n το οποίο εξετάζεται ως προς μία μεταβλητή X και έστω $x_1, x_2, \dots, x_n$ οι τιμές του δείγματος ως προς τη μεταβλητή αυτή. Τότε ορίζουμε:

$$\bar{x} = \frac{\text{Μέση τιμή του δείγματος}}{\text{ως προς τη μεταβλητή } X} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

ή αλλιώς

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{\sum x_i}{n}$$

$$s^2 = \frac{\text{Διασπορά του δείγματος}}{\text{ως προς τη μεταβλητή } X} \\ = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n - 1}$$

ή αλλιώς

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

Αν το δείγμα εξετάζεται και ως προς μία δεύτερη μεταβλητή Y με $y_1, y_2, \dots, y_n$ τις αντίστοιχες τιμές της, τότε εκτός από τη μέση τιμή $\bar{y}$ και την αντίστοιχη δειγματική διασπορά s_y^2 ορίζεται και η δειγματική συνδιακύμανση (s_{xy}) των X, Y ως:

$$s_{xy} = \frac{\text{Συνδιακύμανση του δείγματος}}{\text{ως προς τις μεταβλητές } X, Y} \\ = \frac{(x_1 - \bar{x})(y_1 - \bar{y}) + (x_2 - \bar{x})(y_2 - \bar{y}) + \dots + (x_n - \bar{x})(y_n - \bar{y})}{n - 1}$$

ή αλλιώς

$$s_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n - 1} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n - 1}$$

Ο δειγματικός συντελεστής συσχέτισης Pearson ορίζεται ως:

$$r = \frac{s_{xy}}{s_x \cdot s_y}.$$

Να αποδείξετε ότι:

$$1) \quad r = \frac{\sum (\chi_i - \bar{\chi})(y_i - \bar{y})}{\sqrt{\sum (\chi_i - \bar{\chi})^2} \cdot \sqrt{\sum (y_i - \bar{y})^2}}$$

$$2) \quad r = \frac{\sum \chi_i y_i - v \cdot \bar{\chi} \cdot \bar{y}}{\sqrt{\sum \chi_i^2 - v \bar{\chi}^2} \cdot \sqrt{\sum y_i^2 - v \bar{y}^2}}$$

$$3) \quad r = \frac{\sum \chi_i y_i - \frac{(\sum \chi_i)(\sum y_i)}{v}}{\sqrt{\sum \chi_i^2 - \frac{(\sum \chi_i)^2}{v}} \cdot \sqrt{\sum y_i^2 - \frac{(\sum y_i)^2}{v}}}$$

$$4) \quad r = \frac{v \sum \chi_i y_i - (\sum \chi_i)(\sum y_i)}{\sqrt{v \sum \chi_i^2 - (\sum \chi_i)^2} \cdot \sqrt{v \sum y_i^2 - (\sum y_i)^2}}$$

Ένα άλλο μέτρο που χρησιμοποιείται για την ερμηνεία της συσχέτισης δύο μεταβλητών είναι ο συντελεστής προσδιορισμού r^2 . Ο συντελεστής αυτός εκφράζεται συνήθως σαν ποσοστό, δηλαδή $r^2 \cdot 100\%$. Ο συντελεστής προσδιορισμού r^2 δίνει το ποσοστό της μεταβλητότητας των τιμών της τ.μ. Υ που ερμηνεύεται από τη μεταβλητότητα των τιμών της τ.μ. Χ (και αντίστροφα). Για παράδειγμα, αν $r=0,969$, τότε $r^2=0,939$. Αυτό σημαίνει ότι η μεταβλητότητα της μίας τ.μ. μπορεί να εξηγηθεί από τη συσχέτισή της με την άλλη κατά ποσοστό 93,9%.