Στατιστική

Ορισμός

Στατιστική είναι το σύνολο των μεθόδων και θεωριών που εφαρμόζονται σε αριθμητικά δεδομένα προκειμένου να ληφθεί κάποια απόφαση σε συνθήκες αβεβαιότητας.

Βασικές έννοιες

- Η μελέτη ενός ***πληθυσμού*** είναι συνήθως δύσκολη ή αδύνατη.
- Τις περισσότερες φορές αντί του πληθυσμού εξετάζεται ένα ***δείγμα***.
- Η επιλογή δείγματος με μια ορισμένη μέθοδο ονομάζεται ***δειγματοληψία***.

- Για την εξέταση ενός δείγματος ή πληθυσμού χρησιμοποιούνται ***αριθμητικά δεδομένα*** τα οποία ονομάζονται και ***μεταβλητές.***

- Οι μεταβλητές αυτές μπορεί να είναι:

  ποσοτικές ή ***ποιοτικές***

  συνεχείς ή ***διακριτές***

  μετρήσιμες ή ***προσεγγίσεις***.

- Όλες οι θεωρίες και μέθοδοι που χρησιμοποιούνται για διερευνήσουν ένα δείγμα ή ένα πληθυσμό αποτελούν την ***Περιγραφική Στατιστική***

- Όλες οι θεωρίες και μέθοδοι που χρησιμοποιούνται για να εξάγουν συμπεράσματα για τον πληθυσμό με βάση τα στοιχεία ενός δείγματος αποτελούν την ***Επαγωγική Στατιστική***

Επαγωγική Στατιστική:

Εκτίμηση παραμέτρων πληθυσμού

Έλεγχος υποθέσεων

- Η εξαγωγή συμπερασμάτων για τον πληθυσμό μπορεί να αφορά ***χαρακτηριστικά μεταβλητών*** ή ***αιτιώδεις σχέσεις***

- Χρήση συμπερασμάτων

  Προβλέψεις

  Λήψη αποφάσεων

Εφαρμογές στα οικονομικά

Η Στατιστική έχει κυρίως εφαρμογή σε εκείνη την κατηγορία φαινομένων που ***δεν ελέγχονται πλήρως*** από τον ερευνητή. Εκεί δηλαδή που υπάρχει κάποιος βαθμός ***αβεβαιότητας***.

Τα οικονομικά φαινόμενα ανήκουν στην κατηγορία αυτή.

Συνήθως το φαινόμενο που μελετάται έχει συμβεί και είναι αδύνατο να επαναληφθεί.

Τα στατιστικά στοιχεία που χρησιμοποιούνται συνήθως έχουν την μορφή

Χρονολογικών σειρών

Διαστρωματικών δεδομένων

Δυναμικών διαστρωματικών δεδομένων

Παραδείγματα:

- Μέτρηση οικονομικής δραστηριότητας (δείκτες)
- Ποσοτικός προσδιορισμός σχέσεων οικονομικών μεταβλητών (συναρτήσεις ζήτησης, προσφοράς, κόστους κ.λ.π.)
- Ανάλυση διαχρονικών διακυμάνσεων
- Έρευνα αγοράς

Τυχαίες μεταβλητές και κατανομές πιθανότητας

Τυχαία μεταβλητή

Η μεταβλητή της οποίας οι τιμές καθορίζονται τυχαία.

Σε κάθε τιμή της τυχαίας μεταβλητής αντιστοιχεί και μια ***πιθανότητα***, η πιθανότητα η μεταβλητή να πάρει την συγκεκριμένη τιμή.

Η σχέση μεταξύ των τιμών μιας τυχαίας μεταβλητής (X_i) και των αντίστοιχων πιθανοτήτων $f(X_i)$ μας δίνει την ***κατανομή πιθανότητας*** της τυχαίας μεταβλητής.

$$f(X_i) = P(X = X_i)$$

Βασικές ιδιότητες

$$f(X_i) \geq 0$$

$$\sum_i f(X_i) = 1$$

Παράδειγμα 1 :

200 διαφορετικά εισοδήματα και συχνότητα εμφάνισης

X_i	f_i	$f(X_i)$	$F(X_i)$
180	10	0.050	0.050
190	15	0.075	0.125
200	20	0.100	0.225
210	25	0.125	0.350
220	25	0.125	0.475
230	30	0.150	0.625
240	35	0.175	0.800
250	25	0.125	0.925
260	10	0.050	0.975
270	5	0.025	1.00
	200	1.00	

Εμπειρική κατανομή

Πιθανότητα = Σχετική Συχνότητα

$$f(X_i) = \frac{f_i}{\sum f_i}$$

$$F(X_i) = P(X \le X_i)$$

Παράδειγμα 2 : Συνολικός αριθμός «Γραμμάτων» σε 2 ρίψεις ενός νομίσματος.

Δειγματικός χώρος S={(Γ,Γ), (Κ,Γ), (Γ,Κ), (Κ,Κ)}

Αποτελέσματα	**Συνολικός αριθμός Γραμμάτων** X_i	**Πιθανότητα** $f(X_i)=P(X=X_i)$
(Γ,Γ)	2	1/4
(Κ,Γ)	1	1/4
(Γ,Κ)	1	1/4
(Κ,Κ)	0	1/4

f(0)=1/4

f(1)=1/4+1/4=1/2

f(2)=1/4

Συνάρτηση πιθανότητας κ συνάρτηση αθροιστικής κατανομής

X	$f(X_i)$	$F(X)$
0	**1/4**	**1/4**
1	**1/2**	**3/4**
2	1/4	4/4

- Η $f(X_i)$ ονομάζεται ***συνάρτηση πιθανότητας*** ή ***συνάρτηση πυκνότητας.***
- Η $F(X_i)$ ονομάζεται ***συνάρτηση αθροιστικής κατανομής***
- Στην περίπτωση ***συνεχούς*** μεταβλητής η πιθανότητα μιας συγκεκριμένης τιμής είναι ίση με το ***μηδέν***.

  Για τον λόγο αυτό χρησιμοποιείται η πιθανότητα διαστήματος

Χαρακτηριστικά κατανομών πιθανότητας

Μαθηματική Ελπίδα τυχαίας μεταβλητής

$$E(X) = \sum X_i f(X_i)$$

ή $E(X) = \int_{-\infty}^{+\infty} X_i f(X_i)\, dX$ όταν η μεταβλητή είναι συνεχής

Ονομάζεται και ***προσδοκώμενη τιμή***

$$E(X) = \mu$$

Η μαθηματική ελπίδα δεν είναι παρά ο σταθμικός μέσος όταν ως συντελεστές στάθμισης χρησιμοποιούνται οι αντίστοιχες πιθανότητες.

Παραδείγματα:

Έστω ότι η μεταβλητή X είναι ο αριθμός που εμφανίζεται στην όψη ενός ζαριού. Η μαθηματική ελπίδα της X είναι:

$$\mathrm{E}\left[\mathrm{X}\right]=1\times\frac{1}{6}+2\times\frac{1}{6}+3\times\frac{1}{6}+4\times\frac{1}{6}+5\times\frac{1}{6}+6\times\frac{1}{6}=3.5$$

Θεωρήματα για την μαθηματική ελπίδα:

➢<u>Θεώρημα 1</u>:

Αν c είναι μια σταθερά τότε:

α) $\mathrm{E}c = c$

b) $\mathrm{E}\left[cg(X)\right] = c\mathrm{E}\left[g(X)\right]$

➢<u>Θεώρημα 2</u>:

$$\mathrm{E}\left[\sum_{i=1}^{K} g_i(X)\right] = \sum_{i=1}^{K} \mathrm{E}g_i(X)$$

Δηλαδή, η προσδοκώμενη τιμή του αθροίσματος Κ συναρτήσεων είναι ίση με το άθροισμα των προσδοκώμενων τιμών των συναρτήσεων.

Διακύμανση τυχαίας μεταβλητής (συμβολίζεται και ως σ^2)

$Var(X) = E(X - E(X))^2$ ή αν μ είναι ο μέσος τότε:

$\sigma^2 = E(X^2) - \mu^2$

Απόδειξη:

$$
\begin{aligned}
\mathrm{V}(X) &= \mathrm{E}\left(X - \mathrm{E}(X)\right)^2 \\
&= \mathrm{E}(X - \mu)^2 \\
&= \mathrm{E}\left(X^2 + \mu^2 - 2\mu X\right) \\
&= \mathrm{E}\left(X^2\right) + \mu^2 - 2\mu \mathrm{E}(X) \\
&= \mathrm{E}\left(X^2\right) + \mu^2 - 2\mu\mu \\
&= \mathrm{E}\left(X^2\right) - \mu^2
\end{aligned}
$$

Τυπική Απόκλιση

$$\sigma_{\mathrm{X}} = \sqrt{Var(X)}$$

Συνέπειες των Θεωρημάτων 1 και 2

$$\mathrm{E}(aX+b)=\mathrm{E}(aX)+\mathrm{E}b=a\mathrm{E}(X)+b$$

$$\begin{aligned}
Var(aX+b) &= E\big((aX+b)-E(aX+b)\big)^2 \\
&= E\big(aX+b-aE(X)-b\big)^2 \\
&= E\big(aX-aE(X)\big)^2 = a^2E\big(X-E(X)\big)^2 \\
&= a^2Var(X)
\end{aligned}$$

✷

Η έννοια της ***Συνδιακύμανσης***

Η συνδιακύμανση σ_{XY} ή Cov(X,Y) είναι μέτρο του βαθμού συσχετίσεως δύο μεταβλητών X και Y.

$$\begin{aligned} Cov(XY) &= E\left[\left(X - E(X)\right)\left(Y - E(Y)\right)\right] \\ &= E\left[XY - YE(X) - XE(Y) + E(X)E(Y)\right] \\ &= E(XY) - E(X)E(Y) \end{aligned}$$

Αν X και Y ***ανεξάρτητες*** τυχαίες μεταβλητές

$$E(XY) = E(X)E(Y) \quad \Rightarrow \quad Cov(XY) = 0$$

Το αντίστροφο δεν ισχύει γιατί $Cov(XY)=0$ σημαίνει ότι δεν υπάρχει ***γραμμική σχέση*** ενώ ***ανεξαρτησία*** σημαίνει ότι δεν υπάρχει ***καμία σχέση***.

$$Cov(XY) = \sigma_{\mathrm{XY}}$$

Για την μέτρηση του βαθμού συσχετίσεως δύο μεταβλητών χρησιμοποιούμαι τον συντελεστή συσχέτισης ρ.

$$\rho = \frac{Cov(X,Y)}{\sqrt{V(X)}\sqrt{V(Y)}} = \frac{\sigma_{\mathrm{XY}}}{\sigma_{\mathrm{X}}\sigma_{\mathrm{Y}}}$$

Ο συντελεστής συσχέτισης είναι καθαρός αριθμός (ανεξάρτητος από μονάδες μέτρησης) και επομένως επιτρέπει συγκρίσεις του βαθμού συσχέτισης ανάμεσα σε διάφορες μεταβλητές.

Παίρνει τιμές: $-1 < \rho < 1$

ρ=0, οι μεταβλητές δε συσχετίζονται

ρ=1, τέλεια θετική συσχέτιση

ρ= -1, τέλεια αρνητική συσχέτιση

Στατιστική επαγωγή: Εκτίμηση

Ας θεωρήσουμε μια τυχαία μεταβλητή X, η οποία χαρακτηρίζεται από μια παράμετρο θ (π.χ. μέσος, διακύμανση κτλ). Για να εκτιμήσουμε την θ πρέπει να διαλέξουμε μια συνάρτηση των παρατηρήσεων του δείγματος.

Εκτιμητής Ο τύπος (συνάρτηση) με βάση τον οποίο υπολογίζεται η εκτίμηση της παραμέτρου του πληθυσμού:

$$\hat{\theta}=\hat{\theta}\left(x_1,x_2,\ldots,x_n\right)$$

Εκτίμηση Η αριθμητική τιμή $\hat{\theta}$ που παίρνει ο εκτιμητής για δεδομένο δείγμα

Οι εκτιμητές είναι τυχαίες μεταβλητές αφού είναι συναρτήσεις των παρατηρήσεων του δείγματος που είναι τυχαίες μεταβλητές.

Οι εκτιμήσεις διαφέρουν από δείγμα σε δείγμα.

ΑΛΓΕΒΡΑ ΜΗΤΡΩΝ

Βασικοί ορισμοί

Μήτρα διαστάσεων MxN είναι μια ορθογώνια κατάταξη αριθμών σε M γραμμές και N στήλες

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ a_{M1} & a_{M2} & \dots & a_{MN} \end{bmatrix}$$

Μήτρες διαστάσεων Mx1 ή 1xM ονομάζονται διάνυσμα στήλης και γραμμής αντίστοιχα.

Αν M=N η μήτρα ονομάζεται *τετραγωνική*. Μια τετραγωνική μήτρα με όλα τα στοιχεία ίσα με μηδέν εκτός από την κύρια διαγώνιο ονομάζεται *διαγώνια* μήτρα.

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ 0 & 0 & \dots & a_{NN} \end{bmatrix}$$

Βασικοί ορισμοί

Μια διαγώνια μήτρα με όλα τα στοιχεία στην κύρια διαγώνιο ίσα με την μονάδα ονομάζεται *μοναδιαία*

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

✷

Βασικοί ορισμοί

Γενικά $\mathbf{AB} \neq \mathbf{BA}$ αν όμως $\mathbf{AA} = \mathbf{A}$ η A λέγεται *ταυτοδύναμη*.

Ανάστροφη μήτρα είναι η μήτρα που προκύπτει όταν οι γραμμές πάρουν την θέση των στηλών και οι στήλες την θέση των γραμμών.

Αν η A έχει διαστάσεις MxN, τότε η A′ έχει διαστάσεις NxM.

Ισχύουν τα παρακάτω:

$$(\mathbf{A}')' = \mathbf{A}$$

$$(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$$

$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$$

Αν A′=A η μήτρα λέγεται *συμμετρική*.

Παραγώγιση με μήτρες

$$Y = b_1 x_1 + b_2 x_2 + \ldots + b_n x_n$$

$$\Rightarrow \frac{\partial Y}{\partial x_1} = b_1, \quad \frac{\partial Y}{\partial x_2} = b_2, \quad \ldots \quad \frac{\partial Y}{\partial x_n} = b_n$$

$$Y = \left[b_1, b_2, \ldots b_n \right] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{bx}$$

$$\Rightarrow \frac{\partial Y}{\partial \mathbf{x}} = \frac{\partial \left(\mathbf{bx} \right)}{\partial \mathbf{x}} = \mathbf{b} = \left[b_1, b_2, \ldots b_n \right]$$

Παραγώγιση με μήτρες

$$Y = b_{11}x_1^{\;2} + 2b_{12}x_1x_2 + b_{22}x_2^{\;2} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{12} & b_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{x}'\mathbf{B}\mathbf{x}$$

$$\frac{\partial Y}{\partial x_1} = 2b_{11}x_1 + 2b_{12}x_2 = 2\left(b_{11}x_1 + b_{12}x_2\right)$$

$$\frac{\partial Y}{\partial x_2} = 2b_{12}x_1 + 2b_{22}x_2 = 2\left(b_{12}x_1 + b_{22}x_2\right)$$

$$\begin{bmatrix} \dfrac{\partial Y}{\partial x_1} \\ \dfrac{\partial Y}{\partial x_2} \end{bmatrix} = 2\begin{bmatrix} b_{11} & b_{12} \\ b_{12} & b_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \;\Rightarrow\; \frac{\partial Y}{\partial \mathbf{x}} = \frac{\partial\left(\mathbf{x}'\mathbf{B}\mathbf{x}\right)}{\partial \mathbf{x}} = 2\mathbf{B}\mathbf{x}$$

Λύση συστημάτων γραμμικών εξισώσεων

Έστω το γραμμικό σύστημα:

$$\begin{matrix} a_{11}X_1 + a_{12}X_2 + \ldots + a_{1n}X_n = c_1 \\ a_{21}X_1 + a_{22}X_2 + \ldots + a_{2n}X_n = c_2 \\ \vdots \qquad \vdots \qquad \ldots \qquad \vdots \qquad \vdots \\ a_{n1}X_1 + a_{n2}X_2 + \ldots + a_{nn}X_n = c_n \end{matrix}$$

Το οποίο μπορεί να με την μορφή μητρών να ξαναγραφεί:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ a_{1n} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

ή **AX=C**

Λύση συστημάτων γραμμικών εξισώσεων

Μέθοδος Cramer:

Η λύση της άγνωστης X_j δίνεται από την σχέση: $X_j = \dfrac{\Delta_j}{\Delta}, \quad \Delta \neq 0$

Όπου Δ=|**A**|

Και Δ_j η ορίζουσα που προκύπτει αν αντικαταστήσουμε την *j* στήλη του πίνακα **A** με το διάνυσμα στήλη **C**.

Λύση συστημάτων γραμμικών εξισώσεων

$$\begin{aligned} 2x_1 - x_2 + 3x_3 &= -3 \\ -x_1 - x_2 + 3x_3 &= -6 \\ x_1 - 2x_2 - x_3 &= -2 \end{aligned} \Rightarrow \begin{bmatrix} 2 & -1 & 3 \\ -1 & -1 & 3 \\ 1 & -2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -3 \\ -6 \\ -2 \end{bmatrix}$$

$$x_1 = \frac{\Delta_1}{\Delta} = \frac{\begin{vmatrix} -3 & -1 & 3 \\ -6 & -1 & 3 \\ -2 & -2 & -1 \end{vmatrix}}{\begin{vmatrix} 2 & -1 & 3 \\ -1 & -1 & 3 \\ 1 & -2 & -1 \end{vmatrix}} = \frac{21}{21} = 1 \qquad x_2 = \frac{\Delta_2}{\Delta} = \frac{\begin{vmatrix} 2 & -3 & 3 \\ -1 & -6 & 3 \\ 1 & -2 & -1 \end{vmatrix}}{\begin{vmatrix} 2 & -1 & 3 \\ -1 & -1 & 3 \\ 1 & -2 & -1 \end{vmatrix}} = \frac{42}{21} = 2$$

$$x_3 = \frac{\Delta_3}{\Delta} = \frac{\begin{vmatrix} 2 & -1 & -3 \\ -1 & -1 & -6 \\ 1 & -2 & -2 \end{vmatrix}}{\begin{vmatrix} 2 & -1 & 3 \\ -1 & -1 & 3 \\ 1 & -2 & -1 \end{vmatrix}} = \frac{-21}{21} = -1$$