

- ! Να ξεκαθαρίσετε τις έννοιες
- ανεξάρτητα ενδεχόμενα
 - ζεύγος ενδεχόμενα
 - ανεξάρτητες τυχαίες μεταβλητές.

Ορισμός: X, Y : τ.μ. : ανεξάρτητες αν γνωρίζοντας πότες τιμές έχει πάρει η μία δεν παρέχει καμία πληροφορία σχετικά με το πότες τιμές παίρνει η άλλη

$$X, Y: \text{ανεξ.} \Leftrightarrow F_{X,Y}(x,y) = F_X(x) \cdot F_Y(y) \Leftrightarrow f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y) \begin{matrix} \nearrow \text{6.π.π} \\ \searrow \text{6.μ.π.} \end{matrix}$$

Ισχύουν: Έστω X, Y : ανεξάρτητες τ.μ. με πεπερασμένες μέσες τιμές.
 τότε $E[XY] = E[X] \cdot E[Y]$ (προσοχή! \neq). (*)

• X, Y : τ.μ. (αυτοί και εξαρτημένες).
 τότε $E[aX + bY] = aE[X] + bE[Y]$

(θυμηθείτε: $E[X] = \int x f_X(x) dx$, $E[XY] = \iint (xy) f_{X,Y}(x,y) dx dy$)

• $E[g(X)] = \int g(x) f_X(x) dx$, $E[g(X,Y)] = \iint g(x,y) f_{X,Y}(x,y) dx dy$.

• X, Y : ανεξάρτητες τότε $V(X+Y) = V(X) + V(Y)$

$$V(X+Y) = E[(X+Y)^2] - (E[X+Y])^2 =$$

$$= E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2 =$$

$$= E[X^2] + 2E[XY] + E[Y^2] - (E[X])^2 - 2E[X]E[Y] - (E[Y])^2$$

$$= V(X) + V(Y) + 2(E[XY] - E[X]E[Y]) \stackrel{(*)}{=} V(X) + V(Y)$$

• Ερώτηση: Τι συμβαίνει εάν έχουμε $V(X-Y)$ $\stackrel{X,Y \text{ ανεξ.}}{=} \dots$???

Συνδιακύρμανση (Covariance) ή Συνδιασπορά

Θυμηθείτε: Για μια ζμ. $V(X) = \sigma^2(X) = E[(X - \mu)^2] =$
(≥ 0)
 $= E[X^2] - (E[X])^2$

Ορισ | X, Y ζμ. με πεπερασμένες μέσες τιμές $\left(\begin{matrix} E[X] = \mu_x < \infty \\ E[Y] = \mu_y < \infty \end{matrix} \right)$

$$Cov(X, Y) := E[(X - \mu_x)(Y - \mu_y)]$$

Πρόταση: $E[(X - \mu_x)^m (Y - \mu_y)^n]$, $m, n \in \mathbb{N}$: κεντρική στιγμή (m, n) τάξης

Παρατήρηση: Εάν X, Y : διακριτές ζμ με finite τιμές (x_1, x_2, \dots) ,
 (y_1, y_2, \dots) αντίστοιχα και από κοινά σημ $f_{X,Y}(x_i, y_j)$, $i, j \in \mathbb{N}$,

τότε:
$$Cov(X, Y) = \sum_i \sum_j (x_i - \mu_x)(y_j - \mu_y) f_{X,Y}(x_i, y_j)$$

Εάν X, Y : συνεχείς ζμ με finite τιμές $\omega \in \mathbb{R}$ και

από κοινά σημ $f_{X,Y}(x, y)$, τότε:

$$Cov(X, Y) = \int_{-\omega}^{+\omega} \int_{-\omega}^{+\omega} (x - \mu_x)(y - \mu_y) f_{X,Y}(x, y) dx dy$$

(και στις 2 περιπτώσεις η Covariance < ∞ μόνο όταν
οι βερεές - ομοσυνεπείς συνεπίκριναν άσπιδους)

Ιδιότητες

(18)

• Η συνδιαστροφή είναι ένα ποσοτικό μέτρο της σχέσης μεταξύ δύο τ.μ.

1) $Cov(X, \zeta) = 0$, αν ζ : σταθερή τ.μ.

2) $Cov(X, Y) = Cov(Y, X)$ (συμμετρική)

3) $Cov(X+Y, Z) = Cov(X, Z) + Cov(Y, Z)$

4) $Cov(aX+b, cY+d) = ac Cov(X, Y)$

5) $Cov(X, X) = Var(X)$

• Η $Cov(X, Y)$ σημαίνει το βαθμό γραμμικής συσχέτισης μεταξύ των X, Y .

(i) $Cov(X, Y) > 0$

θετική συσχέτιση

(ii) $Cov(X, Y) < 0$

αρνητική συσχέτιση

(iii) $Cov(X, Y) = 0$

αδυσχέτιστες

Δείκτης

Έστω X_1, \dots, X_n τέμ με $E[X_i^2] < \infty$.

(19)

Πότε η διασπορά της τέμ $X_1 + \dots + X_n$ είναι ανεξάρτητη

$$V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i) + 2 \sum_{i < j} \text{Cov}(X_i, X_j)$$

και ισχύει
Εξίσωση

Ειδικά για $n=2$

$$V\left(\sum_{i=1}^2 X_i\right) = V(X_1 + X_2) = V(X_1) + V(X_2) + 2 \text{Cov}(X_1, X_2)$$

Πρόταση

$$V\left(\sum_{i=1}^n a_i X_i\right) = \sum_{i=1}^n a_i^2 V(X_i) + 2 \sum_{i < j} a_i a_j \text{Cov}(X_i, X_j)$$

$$\hookrightarrow \text{Ορίζουμε: } E\left[\sum_{i=1}^n a_i X_i\right] = \sum_{i=1}^n a_i E[X_i]$$

Πρόταση

Έστω X, Y τέμ με $E[X] < \infty, E[Y] < \infty$. Τότε

$$\text{Cov}(X, Y) = E[XY] - E[X]E[Y]$$

$$\text{Cov}(X, Y) \stackrel{\text{opg}}{=} E[(X - \mu_x)(Y - \mu_y)] = E[XY - \mu_y X - \mu_x Y + \mu_x \mu_y]$$

$$\frac{\mu_x \in \mathbb{R}}{\mu_y \in \mathbb{R}}$$

$$E[XY] - \mu_y E[X] - \mu_x E[Y] + \mu_x \mu_y \quad \frac{\mu_x = E[X]}{\mu_y = E[Y]}$$

$$= E[XY] - E[X]E[Y] - E[X]E[Y] + E[X]E[Y] \quad \square$$

Παρατήρηση

$$X, Y: \omega \in \Omega \Rightarrow E[XY] = E[X]E[Y] \Rightarrow$$

$$\Rightarrow \text{Cov}(X, Y) = 0, \text{ δηλ } X, Y: \text{ ανεξάρτητες}$$

$$\hookleftarrow \text{Το αντίστροφο}$$

Παράδειγμα $X \sim U(-1, 1) \leadsto f_X(\omega) = \begin{cases} \frac{1}{2}, & -1 \leq \omega \leq 1 \\ 0, & \text{αλλιώς} \end{cases}$ (20)

και $Y = X^2$.

• Τότε προφανώς $X, Y \in \{\text{αρτυμένες}\}$ (?!?) ^{γιατί?}

• $\text{Cov}(X, Y) = \text{Cov}(X, X^2) = \mathbb{E}[X \cdot X^2] - \mathbb{E}[X] \mathbb{E}[X^2]$
 $= \mathbb{E}[X^3] - (\mathbb{E}[X])^2 = 0$, αααα

$\mathbb{E}[X^3] = \int_{-1}^1 x^3 \frac{1}{2} dx = 0$ και $\mathbb{E}[X] = \int_{-1}^1 x \cdot \frac{1}{2} dx = 0$

• Επειδή $\text{Cov}(X, Y) = 0$ ενώ $X, Y \in \{\text{αρτυμένες}\}$. Έχουμε λοιπόν:

X, Y : ανεξάρτητες \Rightarrow αευσχεύσιμες

X, Y : αευσχεύσιμες \nRightarrow ανεξάρτητες

Συμπίπτει: X, Y : ανεξάρτητες $\Leftrightarrow \mathbb{E}[XY] = \mathbb{E}[X] \mathbb{E}[Y]$

Πρόταση Αν X_1, X_2, \dots, X_n : από 2 αευσχεύσιμες τότε

$$V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i)$$

Απόδ. $V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i) + 2 \sum_{i < j} \text{Cov}(X_i, X_j)$

• Επειδή (X_i, X_j) : αευσχ. $\forall i \neq j \Rightarrow \text{Cov}(X_i, X_j) = 0$
 $\forall i \neq j$

• Η πρόταση ισχύει και εάν είναι από δύο ανεξάρτητες, διότι τότε θα είναι και αευσχεύσιμες.

Ορισμός Αν X, Y : τ.μ με πεπερασμένες μέσες τιμές, τότε (21)

ορίζεται ως συντελεστής συσχέτισης αυτών:

$$\rho := \rho_{X,Y} = \frac{\text{Cov}(X,Y)}{\sigma_X \cdot \sigma_Y} = \frac{\text{Cov}(X,Y)}{\sqrt{V(X)} \cdot \sqrt{V(Y)}} = \rho(X,Y)$$

$\neq 0$ $\neq 0$ (γιατί?)

• $-1 \leq \rho_{X,Y} \leq 1$. (καθαρός αριθμός!)

• Εάν X, Y ανεξάρτητες $\leadsto \rho_{X,Y} = 0$ (το ίδιο ισχύει και εάν X, Y : ανεξάρτητες!)

• Εάν X, Y θετικά συσχετισμένες $\Leftrightarrow \rho > 0$
(αρνητικά) ($\rho < 0$)

• $\rho_{X,X} = \rho(X,X) = 1$. (προφανές)

• $\rho(aX+b, cY+d) = \rho(X,Y) \cdot \varepsilon$, όπου $\varepsilon = \begin{cases} +1, & \text{αν } ac > 0 \\ -1, & \text{αν } ac < 0 \end{cases}$

Απόδειξη: Άσκηση.

Κατανομή Αρροπών

Έστω X_1, X_2, \dots, X_n : ανεξ. & ισόνομες τυχ. με συν F_X . (27)

i) $Y := \min \{X_1, \dots, X_n\}$. Είναι:

$$\begin{aligned} F_Y(y) &= \mathbb{P}[Y \leq y] = 1 - \mathbb{P}[Y > y] = 1 - \mathbb{P}\left[\min_{i=1, \dots, n} \{X_i\} > y\right] \\ &= 1 - \mathbb{P}[X_1 > y, X_2 > y, \dots, X_n > y] \stackrel{\text{ανεξ.}}{=} \\ &= 1 - \prod_{i=1}^n \mathbb{P}(X_i > y) \stackrel{\text{ισόνομες}}{=} 1 - (1 - F_X(y))^n. \end{aligned}$$

Άρα $f_Y(y) = \frac{d}{dy} F_Y(y) = n (1 - F_X(y))^{n-1} \cdot f_X(y)$

ii) $Z := \max \{X_1, \dots, X_n\}$. Είναι:

$$\begin{aligned} F_Z(z) &= \mathbb{P}[Z \leq z] = \mathbb{P}[\max \{X_1, \dots, X_n\} \leq z] \\ &= \mathbb{P}[X_1 \leq z, \dots, X_n \leq z] \stackrel{\text{ανεξ.}}{=} \mathbb{P}[X_1 \leq z] \dots \mathbb{P}[X_n \leq z] \\ &= \prod_{i=1}^n \mathbb{P}[X_i \leq z] \stackrel{\text{ισόνομες}}{=} (F_X(z))^n. \end{aligned}$$

Άρα $f_Z(z) = n F_X(z) \cdot f_X(z)$

Παρατήρηση. Τα αποτελέσματα $f_Y(y)$ και $f_Z(z)$ ισχύουν
εάν οι X_i έχουν πυκνότητα.

• Τα αποτελέσματα $F_Y(y)$, $F_Z(z)$ ισχύουν
για διακριτές X_i και για συνεχείς X_i ($i=1, 2, \dots, n$).