15/4/13

Μέση Τιμή - Ιδιότητες

1. Υπενθυμίσεις

$$E[X] = \begin{cases} \sum_x x P_X(x) &, X: \text{διακριτή} \\ \int_{-\infty}^{\infty} x f_X(x)\, dx &, X: \text{συνεχής} \end{cases}$$

- $E[aX+b] = aE[X]+b$

$$g(E[X]) \neq E[g(X)] = \begin{cases} \sum_x g(x) P_X(x) &, X: \text{διακριτή} \\ \int_{-\infty}^{\infty} g(x) f_X(x)\, dx &, X: \text{συνεχής} \end{cases}$$

2. Ιδιότητες

(i) $X \geq 0 \Rightarrow E[X] \geq 0$

ii) $a \leq X \leq b \Rightarrow a \leq E[X] \leq b$

Απόδειξη

$$a \leq X \leq b \Rightarrow \begin{matrix} X-a \geq 0 \\ b-X \geq 0 \end{matrix} \Rightarrow \begin{matrix} E[X-a] \geq 0 \\ E[b-X] \geq 0 \end{matrix} \Rightarrow \begin{matrix} E[X] \geq a \\ E[X] \leq b \end{matrix}$$

(iii) (X,Y) διδιάστατη τυχαία μεταβλητή και $g: \mathbb{R}^2 \to \mathbb{R}$
$$(x,y) \mapsto g(x,y)$$

Τότε $$E[g(X,Y)] = \begin{cases} \sum_x \sum_y g(x,y) P_{X,Y}(x,y) & (X,Y): \text{διακριτή} \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\,dx\,dy, & (X,Y): \text{συνεχής} \end{cases}$$

Απόδειξη (για (X,Y): διακριτή)

$Z = g(X,Y)$

$$P_Z(z) = P(Z=z) = P(g(X,Y)=z) = \sum_{(x,y):g(x,y)=z} P(X=x, Y=y)$$
$$= \sum_{(x,y):g(x,y)=z} P_{X,Y}(x,y)$$

$$E[g(X,Y)] = E[Z] = \sum_z z P_Z(z)$$
$$= \sum_z z \sum_{(x,y):g(x,y)=z} P_{X,Y}(x,y)$$
$$= \sum_z \sum_{(x,y):g(x,y)=z} z P_{X,Y}(x,y)$$
$$= \sum_z \sum_{(x,y):g(x,y)=z} g(x,y) P_{X,Y}(x,y)$$
$$= \sum_x \sum_y g(x,y) P_{X,Y}(x,y)$$

(iv) $E[X+Y] = E[X] + E[Y]$

Απόδειξη (για (X,Y): διακριτή)

$$\underbrace{E[X+Y]}_{g(X,Y)} = \sum_x\sum_y (x+y)P_{X,Y}(x,y) = \sum_x\sum_y x P_{X,Y}(x,y) + \sum_x\sum_y y P_{X,Y}(x,y)$$

$$= \sum_x x \underbrace{\boxed{\sum_y P_{X,Y}(x,y)}}_{P_X(x)} + \sum_y y \underbrace{\boxed{\sum_x P_{X,Y}(x,y)}}_{P_Y(y)}$$

$$= E[X] + E[Y]$$

Γενίκευση

$$E[X_1 + X_2 + \dots + X_n] = E[X_1] + E[X_2] + \dots + E[X_n]$$

$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$$

(v) $E[XY] \neq E[X]\,E[Y]$ γενικά (μπορεί να ισχύει αλλά μόνο σε ειδική περίπτωση)

π.χ. (X,Y) διακριτή με σ.π.

$x \backslash y$	0	1	$P_X(x)$
0	0	1/3	1/3
1	1/2	1/6	2/3
$P_Y(y)$	1/2	1/2	1

$$E[XY] = \sum_{x=0}^{1}\sum_{y=0}^{1} x \cdot y\, P_{X,Y}(x,y) = 1 \cdot 1/6 = 1/6$$

$$E[X] = \sum_{x=0}^{1} x P_X(x) = 1 \cdot 2/3 = 2/3$$

$$E[Y] = \sum_{y=0}^{1} y P_Y(y) = 1 \cdot 1/2 = 1/2$$

$$E[XY] = 1/6 \neq E[X] \cdot E[Y] = 2/3 \cdot \frac{1}{2} = 1/3$$

(vi) X, Y ανεξάρτητες $\Rightarrow E[X,Y] = E[X]E[Y]$

"Η συνεπαγωγή δεν πάει ανάποδα" $\not\Leftarrow$

Απόδειξη (για (X,Y): διακριτή)

$$E[X,Y] = \sum_x \sum_y xy P_{X,Y}(x,y) \overset{X,Y:\text{ ανεξάρτητες}}{=} \sum_x \sum_y xy P_X(x) P_Y(y)$$

$$= \boxed{\sum_x x P_X(x)} \boxed{\underbrace{\sum_y y P_Y(y)}_{E[Y]}} = E[X]E[Y]$$

(vii) Γενίκευση

$X_1, X_2, \ldots, X_n$: ανεξάρτητες $\Rightarrow E[X_1 X_2 \ldots X_n] = E[X_1]E[X_2]\ldots E[X_n]$

$$E\left[\prod_{i=1}^n X_i\right] = \prod_{i=1}^n E[X_i]$$

(viii) $E[X_1 X_2 \ldots X_n] = E[X_1]\ldots E[X_n] \not\Rightarrow X_1, X_2, \ldots X_n$: ανεξάρτητες

3. Υπενθύμιση

$A \in \mathcal{A}$, A^c: , 1_A: Δείκτρια του A

$$1_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$$

$E[1_A] = 1 \cdot P(A) + 0 \cdot P(A^c) =$
$= P(A)$

4. Παράδειγμα - Matching Problem

n άτομα με τα καπέλα τους.
Τα βάζουν στο κέντρο - διαλέγουν στην τύχη

X = # ατόμων που βρίσκουν το καπέλο τους
E[X] = ;

$A_i = \{$το άτομο i βρίσκει το καπέλο του, $i = 1, 2, \ldots, n\}$

$$X = \sum_{i=1}^{n} 1_{A_i} \Rightarrow E[X] = E\left[\sum_{i=1}^{n} 1_{A_i}\right] = \sum_{i=1}^{n} E[1_{A_i}] = \sum_{i=1}^{n} P(A_i) = \sum_{i=1}^{n} \frac{1}{n} = \cancel{n} \cdot \frac{1}{\cancel{n}} = 1$$

$P(A_i) = \frac{1}{n}$, $i = 1, 2, \ldots, n$

5. Παράδειγμα - Coupon Collector problem

Υπάρχουν n τύποι κουπονιών
Ένας πελάτης αγοράζει προϊόντα.
Κάθε προϊόν έχει 1 κουπόνι που είναι το i με πιθανότητα 1/n

X = # προϊόντων που αγοράζει μέχρι να μαζεύει τουλάχιστον 1 από κάθε είδος κουπονιού.

E[X] = ;

Έστω X_i = # προϊόντων που θα αγοράσει από τη στιγμή που βρήκε το i-1 διαφορετικό τύπο κουπονιού ως τη στιγμή που βρήκε τον i τύπο κουπονιού.

$X_1 = 1$

$X_2 = $ # δοκιμών Bernoulli μέχρι την πρώτη επιτυχία $\sim Geom\left(\frac{n-1}{n}\right)$

επιτυχία = νέος τύπος κουπονιού

$$P(X_2 = i) = \left(\frac{1}{n}\right)^{i-1}\left(\frac{n-1}{n}\right), \quad i \geq 1$$

$X_3 \sim Geom\left(\frac{n-2}{n}\right)$

Γενικά: $X_1 = 1 \Rightarrow E[X_1] = 1$

$X_i \sim Geom\left(\frac{n-i+1}{n}\right), \quad i \geq 2 \Rightarrow E[X_i] = \frac{n}{n-i+1}$

Όμως $X = \sum_{i=1}^{n} X_i \Longrightarrow E[X] = \sum_{i=1}^{n} E[X_i]$

$$= \frac{n}{n} + \frac{n}{n-1} + \dots + \frac{n}{1} = n\left(1 + \frac{1}{2} + \dots + \frac{1}{n}\right) \simeq n \ln n$$

6. Παράδειγμα (Ροές)

00 11 0 1 000 1 00 1 000 11 0 1

m	n
00	11
00	1

m "0", n "1"

Πείραμα Τύχης: Εξαγωγή χωρίς επανάθεση

Αποτέλεσμα: Ακολουθία από m "0" και n "1"

π.χ. $m = 3$, $n = 8$: 0 111 00 1111 → 4 ροές

Ροή = Block από συνεχόμενα 0 ή συνεχόμενα 1

$R = \#$ ροών , $E[R] = ;$

$R(0) = \#$ ροών με συνεχόμενα "0"

$R(1) = \#$ ροών με συνεχόμενα "1"

$E[R] = E[R(0) + R(1)] = E[R(0)] + E[R(1)]$

$$\overbrace{0001100011110}^{m+n \text{ στοιχεία}}$$

$A_i = \{$στο σημείο i αρχίζει μια ροή από "0"$\}$

$$R(0) = \sum_{i=1}^{m+n} 1_{A_i} \Rightarrow E[R(0)] = \sum_{i=1}^{m+n} E[1_{A_i}] = \sum_{i=1}^{m+n} P(A_i)$$

$P(A_1) = \frac{m}{m+n}$ Πιθανότητα στην $1^{η}$ θέση να βγει "0"

$P(A_i) = \frac{n}{m+n} \cdot \frac{m}{m+n-1}$ Πιθανότητα να έχω "0" στην i-θέση και "1" στην $i-1$ θέση.

$$E[R(0)] = \sum_{i=1}^{m+n} P(A_i) = \frac{m}{m+n} + \sum_{i=2}^{m+n} \frac{n \cdot m}{(n+m)(n+m-1)} = \frac{m}{m+n} + \frac{n \cdot m}{n+m} = \frac{m(n+1)}{m+n}$$

Ομοίως $E[R(1)] = \frac{n(m+1)}{n+m}$

Τελικά $E[R] = \frac{n(m+1) + m(n+1)}{n+m}$

19/4/13

Υπολογισμοί Μέσων Τιμών - Διασπορών

Συνδιακύμανση

1. Μέση Τιμή - Διασπορά - Συνδιακύμανση

$$E[X] = \begin{cases} \sum_x x P_X(x), & X: \text{διακριτή} \\ \int_{-\infty}^{\infty} x f_X(x)\, dx, & X: \text{συνεχής} \end{cases} \longleftarrow \text{Μέτρο Θέσης της } X.$$

$$Var[X] = E[(X - E[X])^2] \longleftarrow \text{Μέτρο Μεταβλητότητας της } X.$$

↱ Covariance = Συνδιακύμανση

$$Cov[X,Y] = E[(X - E[X])(Y - E[Y])] \longleftarrow \text{Μέτρο συσχέτισης των } X, Y.$$

Διαισθητικά

- $Cov[X,Y] > 0 \Rightarrow$ Θετικά συσχετισμένες X, Y $(X\uparrow \Leftrightarrow Y\uparrow)$

- $Cov[X,Y] < 0 \Rightarrow$ Αρνητικά συσχετισμένες X, Y $(X\downarrow \Leftrightarrow Y\uparrow)$ $(X\uparrow \Leftrightarrow Y\downarrow)$

Ορισμός

X, Y ασυσχέτιστες $\Longleftrightarrow Cov[X,Y] = 0$

2. Ιδιότητες

(i) $Var[X] = E[X^2] - E[X]^2$

(ii) $Cov[X,Y] = E[XY] - E[X]E[Y]$

(iii) $Var[X] = 0 \Rightarrow X = E[X]$ με πιθανότητα 1.

(iv) X,Y ανεξάρτητες $\Rightarrow$ X,Y ασυσχέτιστες

(v) X,Y ανεξάρτητες $\not\Leftarrow$ X,Y ασυσχέτιστες

(vi) $Cov[X,Y] = Cov[Y,X]$

(vii) $Var[aX+b] = a^2 Var[X]$

(viii) $Cov[aX+b, cY+d] = ac\, Cov[X,Y]$

(ix) $Cov\left[\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right] = \sum_{i=1}^{n}\sum_{j=1}^{m} Cov[X_i, Y_j]$

(x) $Var\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} Var[X_i] + 2\sum_{1\le i<j\le n} Cov[X_i, X_j]$

π.χ. $Var[X+Y] = Var[X] + Var[Y] + 2\,Cov[X,Y]$

(xi) X_j ανεξάρτητες, $j = 1, 2, \ldots, n \Rightarrow Var\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} Var[X_i]$

Αποδείξεις

(ii) $Cov[X,Y] = E[(X-E[X])(Y-E[Y])]$

$= E[XY - XE[Y] - E[X]Y + E[X]E[Y]]$

$= E[XY] - \cancel{E[X]E[Y]} - E[X]E[Y] + \cancel{E[X]E[Y]}$

$= E[XY] - E[X]E[Y]$

(iii) X: διακριτή

$Var[X] = 0 \Rightarrow \sum (x - E[X])^2 P_X(x) = 0$

$\Rightarrow (x - E[X])^2 P_X(x) = 0 \quad \forall x$

$\Rightarrow$ Αν x τέτοιο ώστε $P_X(x) > 0 \Rightarrow P(X=x) > 0$

τότε αναγκαστικά $x = E[X]$

(iv) X,Y ανεξάρτητες $\Rightarrow E[X,Y] = E[X]E[Y]$

$$Cov[X,Y] = E[XY] - E[X]E[Y] = 0$$

$\Rightarrow X,Y$ ασυσχέτιστες

(v) X,Y με $P(X=-1) = P(X=0) = P(X=1) = 1/3$

και $Y = \begin{cases} 0, & x \neq 0 \\ 1, & x=0 \end{cases}$ Τότε X,Y ασυσχέτιστες αλλά όχι ανεξάρτητες

$x \backslash y$	0	1	$P_X(x)$
-1	1/3	0	1/3
0	0	1/3	1/3
1	1/3	0	1/3
$P_Y(y)$	2/3	1/3	1

$P_{X,Y}(0,0) = 0 \neq P_X(0) \cdot P_Y(0) = 1/3 \cdot 2/3$

$\Rightarrow X,Y$: όχι ανεξάρτητες

Όμως $E[X]=0$

$E[XY]=0$

$\Rightarrow Cov[X,Y] = 0 \Rightarrow X,Y$: ασυσχέτιστες

(viii) $Cov[aX+b, cY+d] = E[(aX+b)(cY+d)]$

$E[ax+b]E[cY+d]$

$= E[acXY + adX + bcY + bd] - (aE[X]+b)(cE[Y]+d)$

$= acE[XY] + \cancel{adE[X]} + \cancel{bcE[Y]} + \cancel{bd} - acE[X]E[Y] - \cancel{adE[X]} - \cancel{bcE[Y]} - \cancel{bd}$

$= ac(E[XY] - E[X]E[Y]) = ac\,Cov[XY]$

(ix) $Cov(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j) = E[(\sum_{i=1}^{n} X_i)(\sum_{j=1}^{m} Y_j)] - E[\sum_{i=1}^{n} X_i]E[\sum_{j=1}^{m} Y_j]$

$$= \sum_{i=1}^{n}\sum_{j=1}^{m} (\underbrace{E[X_iY_j] - E[X_i]E[Y_j]}_{Cov[X_i,Y_j]})$$

(x) $Var[\sum_{i=1}^{n} X_i] = Cov[\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j] = \sum_{i=1}^{n}\sum_{j=1}^{m} Cov[X_i, X_j]$

$= \sum_{i=1}^{n} \underbrace{Cov[X_i, X_i]}_{Var[X_i]} + 2 \sum_{1 \le i < j \le n} Cov[X_i, X_j]$

$Var[X+Y] = Cov[X+Y, X+Y] = Cov[X,X] + Cov[X,Y] + Cov[Y,X] + Cov[Y,Y]$

(xi) X_j: ανεξάρτητες $\Rightarrow Cov[X_i, X_j] = 0 \quad \forall\, i \neq j$

3. Υπολογισμός E, Var στην Bin(n,p)

$X \sim Bin(n,p)$

↳ # επιτυχιών σε n ανεξάρτητες δοκιμές Bernoulli με πιθανότητα επιτυχίας p

$P(X=i) = \binom{n}{i} p^i (1-p)^{n-i}, \quad 0 \le i \le n$

$X = \sum_{i=1}^{n} I_i \qquad I_i = \begin{cases} 1, & \text{αν έχω επιτυχία στην δοκιμή } i \\ 0, & \text{διαφορετικά} \end{cases}$

↑ ανεξάρτητες

$E[X] = \sum_{i=1}^{n} E[I_i]$ (Πάντα) $= n \cdot p$

$Var[X] = \sum_{i=1}^{n} Var[I_i]$ (Ανεξάρτητες) $= np(1-p)$

$E[I_i] = 1 P(I_i = 1) + 0 \cdot P(I_i = 0) = p$

$Var[I_i] = E[I_i^2] - E[I_i]^2 = p - p^2 = p(1-p)$

4. Υπολογισμός E, Var στην NegBin(n,p)

$X \sim NegBin(n,p)$

↳ # ανεξάρτητων δοκιμών Bernoulli με πιθανότητα επιτυχίας p μέχρι να παρατηρήσω n επιτυχίες.

$$P(X=i) = \binom{i-1}{n-1} p^n (1-p)^{i-n}, \quad i \geq n$$

$$X = \sum_{i=1}^{n} X_i, \quad \underbrace{X_i}_{\sim\, Geom(p)} = \#\text{ δοκιμών από την } i-1 \text{ ως την } i\text{-οστή επιτυχία}$$

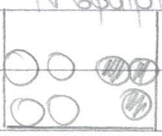
$$E[X_i] = \frac{1}{p}, \quad Var[X_i] = \frac{1-p}{p^2}, \quad X_i: \text{ανεξάρτητες}$$

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = \frac{n}{p}$$

$$Var[X] = Var\left[\sum_{i=1}^{n} X_i\right] \underset{\text{ανεξάρτητες}}{=} \sum_{i=1}^{n} Var[X_i] = n\frac{1-p}{p^2}$$

5. Υπολογισμός E, Var στην HyperGeom(n, N, m)

N σφαιρίδια



→ εξαγωγη n σφαιριδιων

m Λευκά (N−m) Μαύρα

$$X = \#\text{λευκών} \begin{cases} \xrightarrow{\text{με επανάθεση}} X \sim Bin\left(n, \frac{m}{N}\right) \\ \xrightarrow{\text{χ. επανάθεση}} X \sim HyperGeom(n, N, m) \end{cases}$$

$$P(X=i) = \frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}} \qquad 0 \leq i \leq n$$

$$X = \sum_{i=1}^{n} I_i \qquad I_i = \begin{cases} 1, \text{ αν το } i \text{ σφαιρίδιο είναι λευκό} \\ 0, \text{ διαφορετικά} \end{cases} \qquad I_i: \text{όχι ανεξάρτητες}$$

$$E[I_i] = P[I_i = 1] = P(i \text{ σφαιρίδιο λευκό}) = \frac{m}{N}$$

$$Var[I_i] = E[I_i^2] - E[I_i]^2 = \frac{m}{N} - \frac{m^2}{N^2}$$

$$\underset{i \neq j}{Cov[I_i, I_j]} = \underbrace{E[I_i I_j]}_{P(i,j \text{ σφαιρίδια είναι λευκά})} - E[I_i]E[I_j] = \frac{m}{N} \cdot \frac{m-1}{N-1} \cdot \frac{m^2}{N^2}$$

$$E[X] = E[\sum_{i=1}^{n} I_i] = \sum_{i=1}^{n} E[I_i] = n \overset{=p}{\frac{m}{N}} = n \cdot p$$

$$Var[X] = Var[\sum_{i=1}^{n} I_i] = \sum_{i=1}^{n} Var[I_i] + 2 \sum_{1 \le i < j \le n} Cov[I_i, I_j]$$

$$= n \overset{=p}{\frac{m}{N}} \left(1 - \overset{=p}{\frac{m}{N}}\right) + \cancel{2} \frac{n(n-1)}{\cancel{2}} \frac{m}{N} \left(\frac{m-1}{N-1} - \frac{m}{N}\right)$$

$$= n \frac{m}{N} \left[\left(1 - \frac{m}{N}\right) + (n-1) \frac{(m-1)N - m(N-1)}{N(N-1)}\right]$$

$$= n \frac{m}{N} \left[\frac{N(N-1) - m(N-1) + (n-1)(m-1)N - (n-1)m(N-1)}{N(N-1)}\right]$$

$$= \dots$$

$$= n \overset{=p}{\frac{m}{N}} \left(1 - \frac{m}{N}\right)\left(1 - \frac{n-1}{N-1}\right) = np(1-p)\left(1 - \frac{n-1}{N-1}\right)$$

$n=1 \longrightarrow$ ίδιο με Διωνυμική

$n=N \longrightarrow Var[X] = 0 \Rightarrow X = m$ με πιθανότητα 1.

25ο Μάθημα

22/4/13

Συντελεστής (γραμμικής) συσχέτισης και Συνδιακύμανση

① Παράδειγμα

Πείραμα Τύχης: Επιλογή ατόμου από πληθυσμό

X = Βάρος σε kgr

Y = ύψος σε m

$Cov[X, Y]$

Ένας άλλος μελετητής:

X' = Βάρος σε lb

Y' = Ύψος σε ft

$Cov[X', Y']$

<u>Γενικά</u> $Cov[X, Y] \neq Cov[X', Y']$

διότι $X' = aX$ a: σταθερά μετατροπής kgr → lb

$Y' = cY$ c: σταθερά μετατροπής m → ft

$$Cov[X', Y'] = Cov[aX, cY] = ac\, Cov[X, Y]$$

Ελάττωμα! της Cov: Εξάρτηση από τις μονάδες μετατροπής

② Ορισμός

Συντελεστής (γραμμικής) συσχέτισης των τ.μ. X, Y ορίζεται:

$$\rho(X, Y) = \frac{\overbrace{Cov[X, Y]}^{\sigma_{X,Y}}}{\underbrace{\sqrt{Var[X]}}_{\sigma_X}\ \underbrace{\sqrt{Var[Y]}}_{\sigma_Y}}$$

③ Ιδιότητες

1) $\rho(aX,bY)=\begin{cases}\rho(X,Y) & \text{, αν } ab>0\\ -\rho(X,Y) & \text{, αν } ab<0\end{cases}$

2) $\rho(X,Y)\in[-1,1]$ $\left(\begin{matrix}-1\\1\end{matrix}\right\rangle$ ακραίες συσχετίσεις$\Big)$

3) X,Y ασυσχέτιστες $\Rightarrow \rho(X,Y)=0$

4) $\rho(X,Y)=1 \Leftrightarrow Y=aX+b$, $a>0$

5) $\rho(X,Y)=-1 \Leftrightarrow Y=aX+b$, $a<0$

6) Προσοχή!

Ενώ $Cov[X+Y,Z]=Cov[X,Z]+Cov[Y,Z]$

$\rho(X+Y,Z)\neq\rho(X,Z)+\rho(Y,Z)$

Αποδείξεις:

1) $$\rho(aX,bY)=\frac{Cov[aX,bY]}{\sqrt{Var[aX]}\sqrt{Var[bY]}}=\frac{ab}{|ab|}\underbrace{\frac{Cov[X,Y]}{\sqrt{Var[X]}\sqrt{Var[Y]}}}_{=\rho(X,Y)}$$

$$=\begin{cases}\rho(X,Y) & \text{, αν } ab>0\\ -\rho(X,Y) & \text{, αν } ab<0\end{cases}$$

2) $$Var\left(\frac{X}{\sigma_X}+\frac{Y}{\sigma_Y}\right)\geq 0\Rightarrow Var\left(\frac{X}{\sigma_X}\right)+Var\left(\frac{Y}{\sigma_Y}\right)+2Cov\left[\frac{X}{\sigma_X},\frac{Y}{\sigma_Y}\right]\geq 0$$

$$\Rightarrow \frac{1}{\sigma_X^2}Var[X]+\frac{1}{\sigma_Y^2}Var[Y]+2\frac{Cov[X,Y]}{\sigma_X\sigma_Y}\geq 0\Rightarrow$$

$$\Rightarrow 2+2\rho(X,Y)\geq 0\Rightarrow\rho(X,Y)\geq -1$$

και

$$Var\left(\frac{X}{\sigma_x}-\frac{Y}{\sigma_y}\right)\geq 0 \Rightarrow Var\left(\frac{X}{\sigma_x}\right)+Var\left(\frac{Y}{\sigma_y}\right)-2Cov\left[\frac{X}{\sigma_x},\frac{Y}{\sigma_y}\right]\geq 0$$

$$\Rightarrow \frac{1}{\sigma_x^2}Var[X]+\frac{1}{\sigma_y^2}Var[Y]-2\frac{Cov[X,Y]}{\sigma_x\sigma_y}\geq 0 \Rightarrow$$

$$\Rightarrow 2-2\rho(X,Y)\geq 0 \Rightarrow \rho(X,Y)\leq 1$$

C-S

$$\left(E[X-E[X]](Y-E[Y])\right)^2 \leq E[(X-E[X])^2]E[(Y-E[Y])^2]$$

3) X, Y ασυσχέτιστα $\Leftrightarrow Cov[X,Y]=0 \Leftrightarrow \rho(X,Y)=0$

$\Rightarrow)$

5) $\rho(X,Y)=-1 \Rightarrow Var\left[\frac{X}{\sigma_x}+\frac{Y}{\sigma_y}\right]=0 \Rightarrow \frac{X}{\sigma_x}+\frac{Y}{\sigma_y}=c$, σταθ

$$\Rightarrow Y=-\frac{\sigma_y}{\sigma_x}X+c\sigma_y \Rightarrow Y=aX+b,\ a<0$$

$\Leftarrow)$

$$Y=aX+b,\ a<0 \Rightarrow \rho(X,Y)=\frac{Cov[X,Y]}{\sqrt{Var[X]}\sqrt{Var[Y]}}$$

$$=\frac{Cov[X,aX+b]}{\sqrt{Var[X]}\sqrt{Var[aX+b]}}$$

$$=\frac{a\,Cov[X,X]}{|a|\sqrt{Var[X]}\sqrt{Var[X]}}=-1$$

4) ΟΜΟΙΑ με το 5)

$\rho(X,Y)=1 \Rightarrow Var\left[\frac{X}{\sigma_x}-\frac{Y}{\sigma_y}\right]=0 \Rightarrow \dots \Rightarrow Y=aX+b,\ a>0$

Ξέρουμε:

$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$$

$$Var\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} Var[X_i] + 2\sum_{1\le i\le j\le n} Cov[X_i, X_j]$$

$$Cov\left[\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right] = \sum_{i=1}^{n}\sum_{j=1}^{m} Cov[X_i, Y_j]$$

4 Εφαρμογή: Εκτίμηση $\mu = E[X]$, $\sigma_x^2 = Var[X]$

$X_1, X_2, \dots, X_n$ n παρατηρήσεις, ανεξ. ισόνομες

$\mu = ;$ $\sigma_x^2 = ;$

Για να εκτιμήσουμε το μ χρησιμοποιούμε τον $\underbrace{\text{μέσο όρο των παρατηρήσεων}}_{\text{Δειγματικός μέσος}}$.

Δειγματικός μέσος ⟵ μέσος του δείγματος

$$\overline{X} = \overline{X}_n = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n}$$

$\mu = E[X_i]$ ⟵ μέσος του πληθυσμού

Είναι σωστό να χρησιμοποιούμε το $\overline{X}$ για να εκτιμήσουμε το μ;

Πρέπει $E[\overline{X}] = \mu$, $Var[X] \searrow$ ως προς n (φθίνουσα) και μάλιστα $\to 0$

Πράγματι,

$$E[\overline{X}] = E\left[\frac{\sum_{i=1}^{n} X_i}{n}\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \mu$$

$$Var[X] = Var\left[\frac{\sum_{i=1}^{n} X_i}{n}\right] \overset{\text{Ανεξ.}}{=} \frac{1}{n^2}\sum_{i=1}^{n} Var[X_i] = \frac{n\sigma_x^2}{n^2} = \frac{\sigma_x^2}{n} \searrow 0$$

Ερώτημα! Πώς να εκτιμήσουμε το $\sigma_x = E[(X-\mu)^2]$;

Κατά τη λογική του δειγματικού μέσου μία εκτιμήτρια (τυχαία μεταβλητή) για το σ_x^2 μέσω του δείγματος είναι η

$$T = \frac{(X_1-\bar{X})^2 + (X_2-\bar{X})^2 + \dots + (X_n-\bar{X})^2}{n}$$

Είναι η T καλή εκτιμήτρια του σ_x^2 ;

Για να είναι η T καλή εκτιμήτρια του σ_x^2 πρέπει:

$E[T] = \sigma_x^2$

$Var[T] \searrow 0$ για $n \to \infty$

$$E[T] = E\left[\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n}\right] = \frac{1}{n} E\left[\sum_{i=1}^{n}(X_i-\bar{X})^2\right] =$$

$$= \frac{1}{n}\sum_{i=1}^{n} E[(X_i-\bar{X})^2] = \frac{1}{n}\sum_{i=1}^{n} E[(X_i-\mu+\mu-\bar{X})^2] =$$

$$= \frac{1}{n}\sum_{i=1}^{n}\Big(E[(X_i-\mu)^2] + E[2(X_i-\mu)(\mu-\bar{X})] + E[(\mu-\bar{X})^2]\Big)$$

$$= \frac{1}{n}\left(\sum_{i=1}^{n} E[(X_i-\mu)^2]\right) + 2n\sum_{i=1}^{n}\frac{(X_i-\mu)(\mu-\bar{X})}{n} + n E[(\mu-\bar{X})^2]$$

$$= \frac{1}{n}\left(n\sigma_x^2 + 2n E\left[\frac{\sum_{i=1}^{n} X_i - n\mu}{n}(\mu-\bar{X})\right] + n E[(\mu-\bar{X})^2]\right)$$

$$= \frac{1}{n}\left(n\sigma_x^2 - 2n Var[\bar{X}] + n Var[\bar{X}]\right)$$

$$= \frac{1}{n}\left(n\sigma_x^2 - \cancel{n}\frac{\sigma_x^2}{\cancel{n}}\right) = \frac{(n-1)\sigma_x^2}{n}$$

Η T δεν είναι «καλή» διότι $E[T] = \frac{(n-1)\sigma_x^2}{n} \neq \sigma_x^2$

Αν θέσω $S = \frac{n}{n-1}\cdot T = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ τότε

$$E[S] = \frac{n}{n-1}E[T] = \sigma_x^2 .$$

Αποδεικνύεται ότι $Var[S] \longrightarrow$ καθώς $n \longrightarrow \infty$
Άρα για την εκτίμηση

$\mu = E[X_i]$ χρησιμοποιούμε $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ Δειγμ. μέσος

$\sigma_X^2 = Var[X_i]$ — " — $S = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}$ Αμερ. δειγμ. διασπορά