03.05.10 25° μάθημα

Μέση Τιμή - Διασπορά - Συνδιακύμανση

① Υπενθυμίσεις

$$E[X] = \begin{cases} \sum_x x p(x), & X \text{ διακριτή με σ.π. } p(x) \\ \int_{-\infty}^{\infty} x f(x) dx, & X \text{ συνεχής με σ.π.π. } f(x) \end{cases}$$

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

$$E[g(X)] = \begin{cases} \sum_x g(x) p(x) \\ \int_{-\infty}^{\infty} g(x) f(x) dx \end{cases}$$

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

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

$$Var[X] = E[X^2] - E^2[X]$$

$$SD[X] = \sqrt{Var[X]}$$

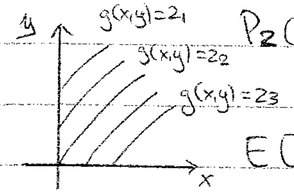
② Επιπλέον Ιδιότητες

1. $X \geq 0 \Rightarrow E[X] \geq 0$

2. $a \leq X \leq b \Rightarrow a \leq E[X] \leq b$

3. $Z = g(X,Y) \Rightarrow E[g(X,Y)] = \begin{cases} \sum_x \sum_y g(x,y) p(x,y), & (X,Y) \text{ διακριτή} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f(x,y) dx dy, & (X,Y) \text{ συνεχής} \end{cases}$

Απόδειξη : $Z = g(X,Y)$, (X,Y) διακριτή
Διακριτή



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

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

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

Απόδειξη : $E[X+Y] = \sum_x \sum_y (x+y) p_{X,Y}(x,y) = \sum_x \sum_y x\, p_{X,Y}(x,y) + \sum_x \sum_y y\, p_{X,Y}(x,y) =$
Διακριτή

$$= \sum_x x\, p_X(x) + \sum_y y\, p_Y(y) = E[X] + E[Y]$$

Γενικά $E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i]$ (n αριθμός, όχι τ.μ., n πεπερασμένο)

5. Γενικά $E[XY] \neq E[X]E[Y]$
Αλλά μπορεί και να ισχύει
π.χ. $P(X=-1) = P(X=0) = P(X=1) = 1/3$
$P(Y=0) = 1/2$ και $P(Y=1) = 1/2$
$P(XY=0) = 1/2 + 1/2 \cdot 1/3 = 2/3$ ← $P(XY=0|Y=0)\,P(Y=0) + P(XY=0|Y=1)\,P(Y=1)$
$P(XY=-1) = 1/6$
$P(XY=1) = 1/6$

$E[X] = 0$
$E[Y] = 1/2$
$E[XY] = 0$
Εδώ ισχύει $E[XY] = E[X]E[Y]$

X, Y ανεξ $\underset{\not\Leftarrow}{\Rightarrow}$ $E[XY] = E[X]E[Y]$

Απόδειξη: Διακριτές: $E[XY] = \sum_x \sum_y xy\, P_{XY}(x,y) \overset{\text{ανεξ}}{=} \sum_x \sum_y xy\, P_X(x)\, P_Y(y) =$

$$= \left(\sum_x x\, P_X(x)\right)\left(\sum_y y\, P_Y(y)\right) = E[X]E[Y]$$

Γενικά $X_1, X_2, \ldots, X_n$ ανεξ $\Rightarrow E\left[\prod_{i=1}^{n} X_i\right] = \prod_{i=1}^{n} E[X_i]$ (n πεπερασμένο, n αριθμός, όχι τμ)

③ Παράδειγμα

(X,Y) διακριτή με σ.π.

$x \backslash y$	0	1	$P_X(x)$
0	$1/6$	$2/6$	$1/2$
1	$1/4$	$1/4$	$1/2$
$P_Y(y)$	$5/12$	$7/12$	1

$E[X] = \sum_x x\, P_X(x) = 1/2$

$E[Y] = \sum_y y\, P_Y(y) = 7/12$

$E[XY] = 1/4$

$E[X]E[Y] = 1/2 \cdot 7/12 = 7/24 \neq 1/4 = E[XY]$

④ Εφαρμογή (Συλλογή κουπονιών)

Κάποιος αγοράζει προϊόντα

Κάθε προϊόν έχει 1 κουπόνι

Υπάρχουν n διαφορετικοί τύποι κουπονιών

(Όλα τα κουπόνια ισοπίθανα)

X = # προϊόντων ώστε να μαζέψει τουλ. 1 κουπόνι από κάθε είδος

$E[X] =$;

(i) $X \geq n$

(ii) $P(X=n) = 1 \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} \cdots \frac{1}{n} = \frac{n!}{n^n} = \frac{(n-1)!}{n^{n-1}}$ $\quad \left(1 = \frac{n}{n}\right)$

(iii) $X = X_0 + X_1 + X_2 + \dots + X_{n-1}$

X_i = # προϊόντων μέχρι να βρεθεί το επόμενο κουπόνι, όταν έχουν βρεθεί ήδη i κουπόνια

$X_0 = 1$

$X_1 \sim Geom\left(\frac{n-1}{n}\right)$

$X_2 \sim Geom\left(\frac{n-2}{n}\right)$

Γενικά $X_i \sim Geom\left(\frac{n-i}{n}\right)$, $i = 1, 2, 3, \dots$

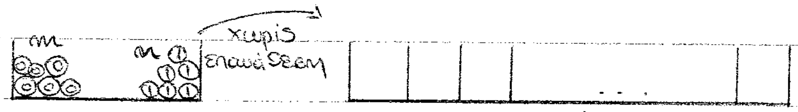
$E[X] = E[X_0 + X_1 + \dots + X_{n-1}] = E[X_0] + E[X_1] + \dots + E[X_{n-1}] =$

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

(5) Άσκηση (Μέσο πλήθος ροών (runs))

n μονάδες "1"

m μηδενικά "0"



$n+m$ θέσεις

Διαδοχικά "1" ~ ροή "1" $\qquad$ π.χ. 1110100110001

Διαδοχικά "0" ~ ροή "0" $\qquad$ $n=7$, $m=6$

4 ροές "1"

3 ροές "0"

R = # ροών

$R(1)$ = # ροών "1"

$R(0)$ = # ροών "0"

$R = R(0) + R(1)$

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

$R(1) = \sum_{i=1}^{n+m} I_i$, $\quad I_i = \begin{cases} 1, & \text{αν στη } i \text{ θέση αρχίζει ροή από "1"} \\ 0, & \text{διαφορετικά} \end{cases}$ όχι ακριβώς

$E[R(1)] = E\left[\sum_{i=1}^{n+m} I_i\right] = \sum_{i=1}^{n+m} E[I_i]$

$E[I_i] = P(I_i = 1) = P\left(\begin{smallmatrix}\text{στη } i \text{ θέση} \\ \text{αρχίζει ροή "1"}\end{smallmatrix}\right)$

$i=1: \; E[I_1] = P(I_1 = 1) = \dfrac{n}{n+m}$

$i \geq 2: \; E[I_i] = P(I_i = 1) = P\left(\begin{smallmatrix}\text{στη θέση } i-1 \text{ "0"} \\ \text{στη θέση } i \text{ "1"}\end{smallmatrix}\right) = \dfrac{m}{n+m} \cdot \dfrac{n}{n+m-1}$

$$E[R(1)] = \sum_{i=1}^{n+m} E[I_i] = \frac{n}{n+m} + \cancel{(n+m-1)} \frac{n \cdot m}{(n+m)\cancel{(n+m-1)}} = \frac{n}{n+m} + \frac{n \cdot m}{n+m}$$

Όμοια $E[R(0)] = \dfrac{m}{n+m} + \dfrac{nm}{n+m}$

Άρα $E[R] = E[R(0)] + E[R(1)] = 1 + \dfrac{2mn}{m+n}$

π.χ. αν έχουμε 1000 "1" και 1000 "0" και τα βάλω τυχαία στη σειρά

$$E[R] = 1 + \frac{2 \cdot 1000 \cdot 1000}{1000 + 1000} = 1001$$

$P(950 \leq R \leq 1050) = 99\%$