07.05.10 26^{o} μάθημα

Διασπορά - Συνδιακύμανση

① Ορισμός : $Var[X] = E[(X-E[X])^2]$ ← Μετράει μεταβλητότητα της X

$Cov[X,Y] = E[(X-E[X])(Y-E[Y])]$ ← Μετράει συσχέτιση X,Y.

↑ Συνδιακύμανση
(Covariance)

② Ιδιότητες

1. $Var[X] = E[X^2] - E^2[X]$
2. $Cov[X,Y] = E[XY] - E[X]E[Y]$

Απόδειξη:

$$Cov[X,Y] = E[XY - XE[Y] - YE[X] + E[X]E[Y]] =$$
$$= E[XY] - E[XE[Y]] - E[YE[X]] + E[E[X]E[Y]] =$$
$$= E[XY] - \cancel{E[Y]E[X]} - E[X]E[Y] + \cancel{E[X]E[Y]}$$

3. $Var[X]=0 \Rightarrow X = c$ με πιθ. 1. $(c=E[X])$

Απόδειξη:

Για διακριτές: $Var[X]=0 \Rightarrow E[(X-E[X])^2] = \sum_x (x-E[X])^2 P(x) = 0$ ← σ.π. της f

$\Rightarrow x - E[X] = 0, \ \forall x$ με $p(x)>0$

$\Rightarrow x = E[X] \quad \forall x$ με $P(x)>0$

4. $Cov[X,Y]=0 \overset{\text{ορισμός}}{\Leftrightarrow}$ X,Y ασυσχέτιστες
5. X,Y ανεξάρτητες $\overset{\not\Leftarrow}{\Rightarrow}$ X,Y ασυσχέτιστες

Απόδειξη:

$(\Rightarrow)$ X,Y ανεξάρτητες $\Rightarrow E[XY] = E[X]E[Y] \overset{②}{\Rightarrow} Cov[X,Y]=0$

$(\not\Leftarrow)$ X,Y όχι ανεξάρτητες αλλά ασυσχέτιστες

(X,Y) διακριτή

$P(X=-1) = P(X=0) = P(X=1) = 1/3$

$$Y = \begin{cases} 0, & X \neq 0 \\ 1, & X = 0 \end{cases}$$

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

X, Y όχι ανεξάρτητες

Όμως $\left.\begin{array}{l} E[X]=0 \Rightarrow E[X]E[Y]=0 \\ XY=0 \Rightarrow E[XY]=0 \end{array}\right\} \Rightarrow Cov[X,Y]=0$

6. $Cov[X,Y] = Cov[Y,X]$
7. $Var[X] = Cov[X,X]$
8. $Var[aX+b] = a^2 Var[X]$
9. $Cov[aX+b, cY+d] = ac\, Cov[X,Y]$
10. $Cov\left[\sum_{i=1}^{m} X_i, \sum_{j=1}^{m} Y_j\right] = \sum_{i=1}^{m}\sum_{j=1}^{m} Cov[X_i, Y_j]$

π.χ.

$$Cov[X_1+X_2, Y_1+Y_2] = Cov[X_1,Y_1] + Cov[X_1,Y_2] + Cov[X_2,Y_1] + Cov[X_2,Y_2]$$

απόδειξη:

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

$$= \sum_{i=1}^{m}\sum_{j=1}^{m} E[X_i Y_j] - \sum_{i=1}^{m}\sum_{j=1}^{m} E[X_i]E[Y_j]$$

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

γιατί $\sum_{i=1}^{m}\sum_{j=1}^{m} Cov[X_i, X_j] = Cov[X_1,X_1] + Cov[X_1,X_2] + Cov[X_1,X_3]$
$Cov[X_2,X_1] + Cov[X_2,X_2] + Cov[X_2,X_3]$
$Cov[X_3,X_1] + Cov[X_3,X_2] + Cov[X_3,X_3] + \dots$

και

$$(X_1+X_2+\dots+X_m)^2 = (X_1^2+X_2^2+\dots+X_m^2) + 2(X_1X_2 + X_1X_3 + \dots + X_1X_m)$$

12. $X_1, X_2, \dots, X_m$ ανεξ $\Rightarrow Var\left[\sum_{i=1}^{m} X_i\right] = \sum_{i=1}^{m} Var[X_i]$

③ Υπολογισμοί μέσων τιμών και διασπορών ειδικών διακριτών

1. $X \sim Bin(n,p)$

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

$E[X] = ;$

$Var[X] = ;$

$X = \sum_{i=1}^{n} I_i, \quad I_i = \begin{cases} 0 & \text{με πιθ } p \\ 1 & \text{με πιθ } 1-p \end{cases}, \quad I_i$ ανεξ.

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

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

$\Rightarrow E[X] = E[\sum_{i=1}^{n} I_i] = \sum_{i=1}^{n} E[I_i] = np$

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

2. $X \sim NegBin(n,p)$

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

$E[X] = ;$

$Var[X] = ;$

$X = \sum_{i=1}^{n} X_i, \quad X_i$ = # δοκιμών από την $i-1$ ως την i επιτυχία

$X_i \sim$ Γεωμ(p)

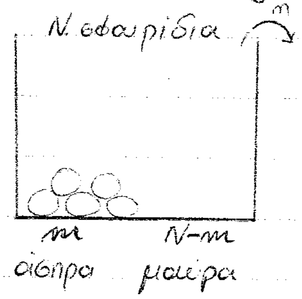
$E[X_i] = \frac{1}{p}$

$Var[X_i] = \frac{1-p}{p^2}$

Άρα $E[X] = n \cdot \frac{1}{p}$

$Var[X] = n \cdot \frac{1-p}{p^2}$

3. $X \sim \text{Hypergeom}(n, N, m)$



$\frac{m}{N}$ = ποσοστό άσπρων = πιθανότητα επιτυχίας

$X = \#$ άσπρων

$$P(X=x) = \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}, \quad 0 \le x \le m$$

$X = \sum_{i=1}^{n} I_i$, $\quad I_i = \begin{cases} 1, & \text{αν το } i \text{ σφαιρ. είναι άσπρο} \\ 0, & \text{διαφορετικά} \end{cases}$, $\quad I_i$ όχι ανεξ.

$E[I_i] = P(\text{το } i \text{ σφαιρίδιο να είναι άσπρο}) = \frac{m}{N} = p$

$Var[I_i] = E[I_i^2] - E^2[I_i] = \frac{m}{N} - \frac{m^2}{N^2} = \frac{m}{N}\left(1 - \frac{m}{N}\right) = p(1-p)$

$Cov[I_i, I_j] = E[I_i I_j] - E[I_i]E[I_j] = \frac{m}{N}\frac{m-1}{N-1} - \frac{m^2}{N^2} = \frac{m}{N}\left(\frac{m-1}{N-1} - \frac{m}{N}\right)$

γιατί

$I_i I_j = \begin{cases} 1, & i, j \text{ λευκά} \\ 0, & \text{διαφορετικά} \end{cases}$ άρα $E[I_i I_j] = P(i,j \text{ σφ λευκά}) = \frac{m}{N} \cdot \frac{m-1}{N-1}$

$$E[X] = E\left[\sum_{i=1}^{n} I_i\right] = \sum_{i=1}^{n} E[I_i] = np = \frac{n \cdot m}{N}$$

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

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

$$= np(1-p) + n(n-1)\,\frac{(m-1)N - m(N-1)}{N(N-1)} =$$

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