19/4/13

Υπολογισμοί Μέσων Τιμών - Διασπορών

Συνδιακύμανση

1. Μέση Τιμή - Διασπορά - Συνδιακύμανση

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

$Var[X] = E[(X - E[X])^2]$ ← Μέτρο Μεταβλητότητας της X.

Covariance = Συνδιακύμανση

$Cov[X,Y] = E[(X - E[X])(Y - E[Y])]$ ← 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 $(X\searrow \Leftrightarrow Y\uparrow)$ $(X\uparrow \Leftrightarrow Y\searrow)$

Ορισμός

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

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

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

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

(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

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

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

$E[X] = \sum_{i=1}^{n} E[I_i]$ (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.

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

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

$\sim Geom(p)$

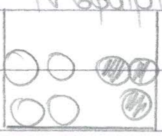
$E[X_i] = \frac{1}{p}$, $Var[X_i] = \frac{1-p}{p^2}$, X_i: ανεξάρτητες

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

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

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

N σφαιρίδια



→ εξαγωγή n σφαιριδίων

m Λευκά (N-m) Μαύρα

X = # λευκών $\xrightarrow{\text{με επανάθεση}} X \sim Bin(n, \frac{m}{N})$

X = # λευκών $\xrightarrow{\text{χ. επανάθεση}} X \sim HyperGeom(n, N, m)$

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

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

I_i: όχι ανεξάρτητες

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

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

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

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

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

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

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

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

$$= \ldots$$

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

$n = 1 \longrightarrow$ ίδιο με Διωνυμική

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