23-5-2023

ΑΝΑΚΕΦΑΛΑΙΩΣΗ:

Μάθημα 23 - 23.05.2023
ΓΡΑΦΕΑΣ
Χρύσα Νούση

Έστω τ.μ $X \sim N(\mu_x, \sigma^2_x)$

$Y \sim N(\mu_y, \sigma^2_y)$

$$\Rightarrow f_X(x) = \frac{1}{\sigma_x\sqrt{2\pi}} e^{-\frac{(x-\mu_x)}{2\sigma^2_x}}$$

$$f_Y(y) = \frac{1}{\sigma_y\sqrt{2\pi}} e^{-\frac{(y-\mu_y)^2}{2\sigma^2_y}}$$

Αν $X \perp\!\!\!\perp Y \Rightarrow f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left[\frac{(x-\mu_x)^2}{2\sigma^2_x} + \frac{(y-\mu_y)^2}{2\sigma^2_y}\right]\right\}$

ΔΙΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ:

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma^2_y} - 2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right]\right\}$$

Να βρεθούν οι

α) Περιθώριες κατανομές $f_x(x)$, $f_y(y)$

β) Η δεσμευμένη σ.π.π $f_{x|y}(x|y)$

Λύση:

α) $f_x(x) = \int_{-\infty}^{\infty} f(x,y)dy$ $=$

$$= \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-p^2}} \underbrace{\exp\left\{-\frac{1}{2(1-p^2)}\frac{(x-\mu_x)^2}{\sigma_x^2}\right\}\int_{-\infty}^{\infty}\exp\left\{\frac{1}{2(1-p^2)}\left[\frac{(y-\mu_y)^2}{\sigma_y^2}-2p\frac{(x-\mu_x)(y-\mu)}{\sigma_x\sigma_y}\right]\right\}dy}$$

Θέλουμε να το φέρουμε στην μορφή $\exp -\frac{(y-?)^2}{?}$

Εστιάζω στον ευθέτη:

$$\frac{(y-\mu_y)^2}{\sigma_y^2} - 2p\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + p^2\frac{(x-\mu_x)^2}{\sigma_x^2} - p^2\frac{(x-\mu_x)^2}{\sigma_x^2} =$$

$$= \left(\frac{y-\mu_y}{\sigma_y} - p\frac{x-\mu_x}{\sigma_x}\right)^2 - p^2\frac{(x-\mu_x)^2}{\sigma_x^2} =$$

$$= \frac{\left[y-\mu_y-p(\sigma_x/\sigma_y)(x-\mu_x)\right]^2}{\sigma_y^2} - p^2\frac{(x-\mu_x)^2}{\sigma_x^2}$$

Άρα:

$$f_X(x) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2}\underbrace{\left(\frac{1}{1-\rho^2} - \frac{\rho^2}{1-\rho^2}\right)}_{\to 1}\frac{(x-\mu_x)^2}{\sigma_x^2}\right\}\exp\left\{\frac{1}{2(1-\rho^2)}\frac{[y-\mu_y-\rho(\sigma_y/\sigma_x)(x-\mu_x)]^2}{\sigma_y^2}\right\}dy$$

$$\underbrace{\frac{1}{\sigma_x\sqrt{2\pi}}\exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right\}}_{\checkmark}\underbrace{\boxed{\frac{1}{\sqrt{2\pi}\,\sigma_y\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{2}\frac{[y-\mu_y-\rho(\sigma_y/\sigma_x)(x-\mu_x)]^2}{(1-\rho^2)\sigma_y^2}\right\}dy}}_{\to 1}$$

$\hookrightarrow$ σ.π.π $Z \sim N\left(\mu_y + \rho\frac{\sigma_y}{\sigma_x}(x-\mu_x), (1-\rho^2)\sigma_y^2\right)$

$$\Rightarrow \frac{1}{\sqrt{2\pi}\,\sigma_y\sqrt{1-\rho^2}}\int \dots\, dy = 1$$

$$\Rightarrow f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma_x}\exp\left\{-\frac{(x-\mu_x)^2}{2\sigma_x^2}\right\}$$

Δηλαδή $X \sim N(\mu_x, \sigma_x^2)$

Εντελώς ανάλογα προκύπτει ότι $Y \sim N(\mu_y, \sigma_y^2)$

"Τι κρατάμε"

Η διακανονική

1] Μας περιγράφει την από κοινού κατανομή 2 κανονικών τ.μ που απλώς έχουν κάποια εξάρτηση. Στην διακανονική, η περιθώριες είναι κανονικές

2] Συμπλήρωση τετραγώνου όταν έχουμε κανονικές, διακανονικές κατανομές!!

β) Υπολογισμός της δεσμευμένης $f_{X|Y}(x|y)$

$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$$

$$= \frac{\dfrac{1}{\cancel{\sqrt{2\pi}}\sqrt{2\pi}\,\sigma_x\cancel{\sigma_y}\sqrt{1-\rho^2}}\exp\left\{-\dfrac{1}{2(1-\rho^2)}\left[\dfrac{(x-\mu_x)^2}{\sigma_x^2} + \dfrac{(y-\mu_y)^2}{\sigma_y^2} - 2\rho\dfrac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right]\right\}}{\dfrac{1}{\cancel{\sqrt{2\pi}\,\sigma_y}}\exp\left\{-\dfrac{1-\rho}{2(1-\rho^2)}\dfrac{(y-\mu_y)^2}{\sigma_y^2}\right\}} =$$

* Αν $\rho=0$ τότε $f_{X|Y} = f_X$

$$= \frac{1}{\sqrt{2\pi}\sigma_x\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2(1-\rho^2)}\left[\frac{(x-\mu_x)^2}{\sigma_x^2} - 2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \rho^2\frac{(y-\mu_y)^2}{\sigma_y^2}\right]\right\}$$

$$= \frac{1}{\sqrt{2\pi}\sigma_x\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x-\mu_x}{\sigma_x} - \rho\frac{y-\mu_y}{\sigma_y}\right)^2\right\}$$

$$= \frac{1}{\sqrt{2\pi}\sigma_x\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)\sigma_x^2}\left[(x-\mu_x-\rho(\sigma_x/\sigma_y)(y-\mu_y)\right]^2\right\}$$

$$\Rightarrow \boxed{x|y \sim N \quad \mu_x + \rho\frac{\sigma_x}{\sigma_y}(y-\mu_y), \quad \sigma_x^2(1-\rho^2)}$$

*

Πολυκανονική Κατανομή

$$f(x_1, \dots, x_n) = \frac{1}{(2\pi)^{n/2}\sqrt{|\Sigma|}} \exp\left\{-\frac{1}{2}(\bar{x}-\bar{\mu})^T \Sigma^{-1}(\bar{x}-\bar{\mu})\right\}$$

$$\Sigma = \begin{matrix} \sigma_x^2 & \rho\sigma_x\rho\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{matrix}$$

ΠΟΣΟΤΙΚΟΠΟΙΗΣΗ ΕΛΛΕΙΨΗΣ ΑΝΕΞΑΡΤΗΣΙΑΣ + ΣΥΝΔΙΑΚΥΜΑΝΣΗ

- Διασπορά = Μέτρο τυχαιότητας μιας τυχαίας μεταβλητής

$Var(x) = \mathbb{E}[x^2] - \mathbb{E}[x] = \mathbb{E}\left[[x-\mathbb{E}[x]]^2\right]$

- Συνδιακύμανση = Μέτρο της εξάρτησης δυο τ.μ. ή μέτρο της εναπομείνουσας τυχαιότητας όταν η μια τ.μ. δεσμευθεί ως προς την άλλη

Ορισμός:
Έστω X, Y τ.μ. Ως <u>συνδιακύμανση</u> των X, Y ορίζεται η ποσότητα

$$\boxed{\underset{\text{covariance}}{Cov}(X,Y) = \mathbb{E}\left[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])\right]}$$

Βασική Ιδιότητα:

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

Απόδειξη:

$$\begin{aligned}
Cov(X,Y) &= \mathbb{E}\left[[X-\mathbb{E}[X]]\cdot[Y-\mathbb{E}[Y]]\right] \\
&= \mathbb{E}\left[XY-\mathbb{E}[X]Y - X\mathbb{E}[Y]+\mathbb{E}[X]\mathbb{E}[Y]\right] \\
&= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] - \mathbb{E}[Y]\mathbb{E}[X] + \mathbb{E}[X]\cdot\mathbb{E}[Y] \\
&= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \quad *
\end{aligned}$$

Πόρισμα:

Αν $X \perp\!\!\!\perp Y$, τότε $Cov(X,Y) = 0$

ΙΔΙΟΤΗΤΕΣ ΣΥΝΔΙΑΚΥΜΑΝΣΗΣ:

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

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

3. $Cov(X,X) = Var(X)$

* Για $x=y$ έχουμε $\mathbb{E}[x^2] - \mathbb{E}[x]^2$

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

5) $Cov(aX, bY) = ab\, Cov(X, Y)$.

Για την 4η ιδιότητα:

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

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

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

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

$$= \sum_{i=1}^{n}\sum_{j=1}^{m} Cov(X_i, Y_j).$$

ΙΔΙΟΤΗΤΕΣ ΔΙΑΚΥΜΑΝΣΗΣ

1) $Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

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

3) $Var(aX) = a^2 Var(X)$

$Var\left(\sum_{i=1}^{n} X_i\right) = Cov\left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i\right) =$

Για την 2η ιδιότητα:

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} Cov(X_i, X_j)$$

$$= \sum_{i=1}^{n} Cov(X_i, X_i) + \sum_{i \neq j} Cov(X_i, X_j)$$

$$= \sum_{i=1}^{n} Var(X_i) + 2 \sum_{i<j} Cov(X_i, X_j)$$ (συμβολισμός στον Ross)

Πορισματα στην περιπτωση ανεξάρτητων μεταβλητων

1) $Cov(X,Y) = 0$ αν $X \perp\!\!\!\perp Y$

2) $Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i)$ αν $(X_1, \dots, X_n) \perp\!\!\!\perp$

ΠΡΟΣΟΧΗ !!!

$X \perp\!\!\!\perp Y \Rightarrow Cov(X,Y) = 0$. Το αντίστροφο δεν ισχυει !!

▸ Ασκηση: Εστω οτι X, Y ακολουθούν διμανονική κατανομή με παραμέτρους $\mu_x, \mu_y, \sigma_x, \sigma_y, \rho$. Να υπολογισθεί η συνδιακύμανση $Cov(X,Y)$.

ΑΘΡΟΙΣΜΑΤΑ ΤΥΧΑΙΩΝ ΜΕΤΑΒΛΗΤΩΝ

Γιατί μας ενδιαφέρουν;

→ Πολλα προβλήματα έχουν να κάνουν με αθροισματα μεταβλητων (ζάρια)

→ Επίσης πολλά προβλήματα απαρίθμησης ανάγονται σε αθροισμάτα τ. μ.
(# επιτυχιών $= \sum_{i=1}^{n} 1\{\text{επιτυχία}\}$)

ΑΝΕΞΑΡΤΗΤΕΣ ΤΥΧΑΙΕΣ ΜΕΤΑΒΛΗΤΕΣ

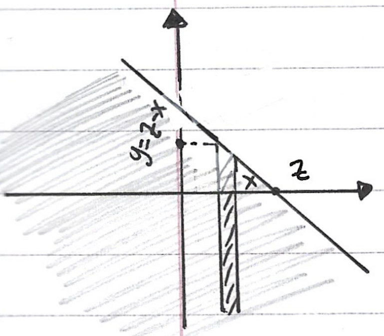
Αν $X \perp\!\!\!\perp Y$ ποια η κατανομή της $Z = X+Y$?

∘ Θα κοιτάξουμε την περίπτωση οπου X, Y συνεχείς με σ.π.π. $f_X(x)$, $f_Y(y)$ αντιστοίχως

Τρόπος 1:

Θα υπολογίσουμε την αθροιστική F_Z.

$$F_Z(z) = \mathbb{P}(Z \le z) = \mathbb{P}(X+Y \le z) = \iint_{x+y\le z} f(x,y)\,dx\,dy \overset{X \perp Y}{=} \iint_{x+y\le z} f_X(x)\, f_Y(y)\,dx\,dy \overset{*}{=}$$



$$\mathcal{D} = \{(x,y) : x+y \le z\}$$
$$= \{(x,y) : x \in \mathbb{R},\ y \le z-x\}$$

$$\overset{(*)}{=} \int_{-\infty}^{\infty}\int_{-\infty}^{z-x} f_X(x)\, f_Y(y)\,dy\,dx = \int_{-\infty}^{\infty} f_X(x) \int_{-\infty}^{z-x} f_Y(y)\,dy\,dx = \int_{-\infty}^{\infty} f_X(x)\, F_Y(z-x)\,dx$$

$$\boxed{F_Z(z) = \int_{-\infty}^{\infty} f_X(x)\, F_Y(z-x)\,dx}$$

· Η σ.π.π της $Z = X+Y$ θα είναι $f_Z(z) = \dfrac{d}{dz}\mathbb{P}(Z \le z)$

$$= \frac{d}{dz}\int_{-\infty}^{\infty} f_X(x)\, F_Y(z-x)\,dx$$

$$= \int_{-\infty}^{\infty} f_X(x)\, \frac{d}{dz} F_Y(z-x)\,dx$$

$$= \int_{-\infty}^{\infty} f_X(x)\, F_Y'(z-x)\,dx$$

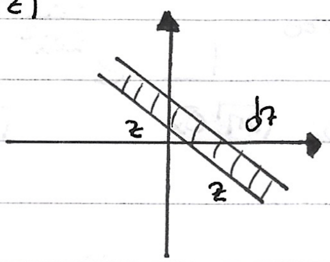
$$\boxed{f_Z(z) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(z-x)\,dx} \quad *$$

* Αυτό ονομάζεται συνέλιξη (convolution) των f_x, f_y.
* Συνέλιξη δυο συναρτήσεων f, g καλείται η συνάρτηση
$$(f * g)(x) = \int_{-\infty}^{\infty} f(u) \cdot g(x-u)\, du$$

Μετασχηματισμοί Laplace/Fourier
↳ Συνέλιξη παίζει το ρολο του πολλ/σμού στις μετ/νες συν/σεις

Τρόπος 2:

$\mathbb{P}(z \leq Z \leq z+dz)$



Δεν προλαβαίνουμε !!

▸ Εφαρμογή:

Εστω ανεξάρτητες τ.μ. $X_i \sim N(\mu_i, \sigma_i^2)$ $i = 1, 2, \ldots, n$ N.S.O
η $S_n = \sum_{i=1}^{n} X_i$ ακολουθεί κανονική κατανομή με μέσο $\mu_1 + \mu_2 + \ldots + \mu_n$.

και διασπορά $\sigma_1^2 + \ldots + \sigma_n^2$

$$S_n \sim N\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right)$$

Σημείωση:

∘ Ακομα και αν $X_1, \ldots, X_n$ οχι ανεξάρτητες

$$\boxed{\mathbb{E}[S_n] = \mathbb{E}[X_1 + \ldots + X_n] = \sum \mathbb{E}[X_i] = \sum_{i=1}^{n} \mu_i}$$

∘ Αν $(X_1, \ldots, X_n)$ ⫫ αλλά πιθανώς οχι κανονικές και πάλι

$$\boxed{Var(S_n) = \sum_{i=1}^{n} Var(X_i) = \sum_{i=1}^{n} \sigma_i^2}$$

Απλοποιήσεις

① Το πρόβλημα Χ.Β.Τ.Γ ανάγεται στην περίπτωση $\mu_1=\mu_2=\ldots=\mu_n=0$
(# Γιατί?)

② Θα δείξουμε ότι αν $X \perp\!\!\!\perp Y$, $X \sim N(0, \sigma_x^2)$ $Y(0, \sigma_y^2)$ τότε
$X+Y \sim N(0, \sigma_x^2+\sigma_y^2)$

Αφού $X \perp\!\!\!\perp Y$, η $Z=X+Y$ θα έχει σ.π.π:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(z-x)\, dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma_x} e^{\frac{-x^2}{2\sigma_x^2}} \frac{1}{\sqrt{2\pi}\,\sigma_y} e^{-\frac{(z-x)^2}{\sigma_y^2}} dx$$

$$= \frac{1}{2\pi \sigma_x \sigma_y} \int_{-\infty}^{\infty} \exp\left\{ -\frac{1}{2} \left[\frac{x^2}{\sigma_x^2} + \frac{(z-x)^2}{\sigma_y^2} \right] \right\} dx.$$

∘ Εστιάζουμε στον εκθέτη:

$$\frac{x^2}{\sigma_x^2} + \frac{(z-x)^2}{\sigma_y^2} = \frac{\sigma_y^2 x^2 + \sigma_x^2 z^2 + \sigma_x^2 x^2 - 2\sigma_x^2 zx}{\sigma_x^2 \sigma_y^2}$$

$$= \frac{\left[\sigma_x^2 + \sigma_y^2\right] x^2 - 2\sigma_x^2 zx + \sigma_x^2 z^2}{\sigma_x^2 \sigma_y^2}$$

$$= \frac{x^2 - 2\frac{\sigma_x^2}{\sigma_x^2+\sigma_y^2}\, zx + \frac{\sigma_x^2}{\sigma_x^2+\sigma_y^2} \frac{(\sigma_x^2+\sigma_y^2)}{\sigma_x^2+\sigma_y^2} z^2}{\frac{\sigma_x^2 \sigma_y^2}{\sigma_x^2 \sigma_y^2}}$$

$$\frac{\left(x - \frac{\sigma_x^2}{\sigma_x^2+\sigma_y^2} \cdot z\right)^2 + \frac{\sigma_y^2\sigma_x^2 z^2}{(\sigma_x^2+\sigma_y^2)^2}}{\frac{\sigma_x^2\sigma_y^2}{\sigma_x^2+\sigma_y^2}}$$

$$\frac{x - \frac{\sigma_x^2}{\sigma_x^2+\sigma_y^2}}{\frac{\sigma_x^2\sigma_y^2}{\sigma_x^2+\sigma_y^2}} + \frac{z^2}{\sigma_x^2+\sigma_y^2}$$

$$\Rightarrow f_z(z) = \frac{1}{2\pi\sigma_x\sigma_y}\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}\left[\frac{\left(x - \frac{\sigma_x^2}{\sigma_x^2+\sigma_y^2}\right)^2}{\sigma_x^2\sigma_y^2/(\sigma_x^2+\sigma_y^2)}\right]\right\}dx \, \exp\left(-\frac{1}{2}\frac{z^2}{\sigma_x^2+\sigma_y^2}\right)$$

$$f_z(z) = \frac{1}{\sqrt{2\pi}\cdot\sqrt{2\pi}\,\frac{\sigma_x\sigma_y}{\sqrt{\sigma_x^2+\sigma_y^2}}\sqrt{\sigma_x^2+\sigma_y^2}}\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}\left[\frac{x - \frac{\sigma_x^2}{\sigma_x^2+\sigma_y^2}}{\sigma_x^2\sigma_y^2/(\sigma_x^2+\sigma_y^2)}\right]\right\}dx \, \exp -\frac{1}{2}\frac{z^2}{\sigma_x^2+\sigma_y^2}$$

$$f_z(z) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_x^2+\sigma_y^2}}\exp\left\{-\frac{1}{2}\frac{z^2}{\sigma_x^2+\sigma_y^2}\right\}$$

$$\Rightarrow \boxed{Z \sim N(0, \sigma_x^2+\sigma_y^2)}$$