Ειδικές περιπτώσεις κατανομής συνάρτησης πολλών μεταβλητών

[I]. Κατανομή αθροίσματος δύο μεταβλητών $Z=X+Y$

- Έστω 2 τυχαίες μεταβλητές X, Y με από-κοινού συνάρτηση πυκνότητας $f_{X,Y}(x,y)$
- Εύρεση της κατανομής του αθροίσματος $Z=X+Y$
- Εφαρμόζουμε τα **4 βήματα της γενικής μεθοδολογίας**

Βήμα 0. Ορίζουμε μία **βοηθητική μεταβλητή**, ***W=X***.

Έτσι έχουμε πρόβλημα εύρεσης από-κοινού κατανομής δύο συναρτήσεις τυχαίων μεταβλητών:

$$\boxed{Z = X + Y \qquad , \qquad W = X}$$

Βήμα 1.

$$\left.\begin{aligned} Z &= X+Y \\ W &= X \end{aligned}\right| \Rightarrow \begin{aligned} X &= W = h_1(z,w) \\ Y &= Z-W = h_2(z,w) \end{aligned}$$

Βήμα 2.

$$\left|J(z,w)\right| = \begin{vmatrix} \dfrac{\partial}{\partial z} h_1(z,w) & \dfrac{\partial}{\partial w} h_1(z,w) \\ \dfrac{\partial}{\partial z} h_2(z,w) & \dfrac{\partial}{\partial w} h_2(z,w) \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix} = \left|-1\right| = 1$$

Βήμα 3.

$$f_{Z,W}(z,w) = \left|J(z,w)\right| f_{x,y}\left(x = h_1(z,w), y = h_2(z,w)\right) =$$
$$= f_{x,y}(w, z-w)$$

Βήμα 4.

$$\boxed{f_Z(z) = \int_{-\infty}^{+\infty} f_{Z,W}(z,w)\,dw = \int_{-\infty}^{+\infty} f_{X,Y}(w, z-w)\,dw}$$

Υπολογίζουμε την περιθώρια κατανομή του *Z*

- **Κατανομή αθροίσματος δύο τυχαίων μεταβλητών**

- **γενικός τύπος:**

$$f_Z(z) = \int_{-\infty}^{+\infty} f_{Z,W}(z, w)\,dw = \int_{-\infty}^{+\infty} f_{X,Y}(w, z-w)\,dw$$

[ΙΙ]. Κατανομή αθροίσματος δύο ανεξάρτητων μεταβλητών $Z=X+Y$

- **Κατανομή αθροίσματος:** $f_Z(z)=\int_{-\infty}^{+\infty} f_{Z,W}(z,w)dw=\int_{-\infty}^{+\infty} f_{X,Y}(w,z-w)dw$

- Αν επιπλέον οι δύο τ.μ. ***X, Y*** **είναι** **ανεξάρτητες** τότε:

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

που είναι ένα ολοκλήρωμα - συνάρτηση ως προς z :

$$H(z)=\int_{-\infty}^{+\infty} f(x)g(z-x)dx$$

- Το **ολοκλήρωμα** αυτής της μορφής:

$$H(z) = \int_{-\infty}^{+\infty} f(x) g(z-x) dx$$

αποτελεί ένα **γνωστό ολοκλήρωμα** το οποίο ονομάζεται **συνέλιξη** (***convolution***) και συμβολίζεται ως :

$$f * g = \int_{-\infty}^{+\infty} f(x) g(z-x) dx$$

Επομένως ισχύει η παρακάτω **πρόταση**:

- **Πρόταση:** Η κατανομή του αθροίσματος ***Z=X+Y*** δύο ανεξάρτητων τυχαίων μεταβλητών ***X, Y***

$$f_Z(z) = \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x) dx$$

προκύπτει από το **ολοκλήρωμα της συνέλιξης** (***convolution***) **των περιθώριων συναρτήσεων πυκνότητας πιθανότητας** των δύο ανεξάρτητων μεταβλητών.

(II). Κατανομή συνάρτησης $Z=aX+bY$

Γενική μορφή

- Εργαζόμενοι όπως και προηγουμένως βρίσκουμε ότι:

$$f_Z(z) = \int_{-\infty}^{+\infty} f_{Z,W}(z,w)dw = \frac{1}{|b|}\int_{-\infty}^{+\infty} f_{X,Y}\left(x, \frac{1}{b}z - \frac{a}{b}x\right)dx$$

Βήμα 1.

$$\left.\begin{array}{c} Z = aX + bY \\ W = X \end{array}\right| \Rightarrow \begin{array}{c} X = W = h_1(z,w) \\ Y = \dfrac{1}{b}Z - \dfrac{a}{b}W = h_2(z,w) \end{array}$$

Βήμα 2.

$$|J(z,w)| = \begin{vmatrix} \dfrac{\partial}{\partial z}h_1(z,w) & \dfrac{\partial}{\partial w}h_1(z,w) \\ \dfrac{\partial}{\partial z}h_2(z,w) & \dfrac{\partial}{\partial w}h_2(z,w) \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ \dfrac{1}{b} & -\dfrac{a}{b} \end{vmatrix} = \left|-\frac{1}{b}\right| = \left|\frac{1}{b}\right|$$

Βήμα 3.

$$f_{Z,W}(z,w) = |J(z,w)| f_{x,y}\big(x = h_1(z,w), y = h_2(z,w)\big) =$$
$$= \left|\frac{1}{b}\right| f_{x,y}\left(w, \frac{1}{b}z - \frac{a}{b}w\right)$$

Βήμα 4.

$$\boxed{f_Z(z) = \int_{-\infty}^{+\infty} f_{Z,W}(z,w)dw = \left|\frac{1}{b}\right| \int_{-\infty}^{+\infty} f_{x,y}\left(w, \frac{1}{b}z - \frac{a}{b}w\right) dw}$$

Η περιθώρια κατανομή της τυχαίας μεταβλητής *Z*

[III]. Κατανομή Z = X² + Y² αν X, Y ανεξάρτητες μεταβλητές που ακολουθούν την τυπική κανονική N(0,1)

- Είναι **γνωστό** ότι αν *X~N(0,1)*, τότε η *Y=X²* ακολουθεί την Χι-τετράγωνο κατανομή με ν=*1* βαθμό ελευθερίας, ***X²(1)***.

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

Σημείωση: Η κατανομή Χι-τετράγωνο έχει συνάρτηση πυκνότητας:

$$f(x)=\frac{1}{2^{\nu/2}\Gamma(\nu/2)}x^{\nu/2-1}e^{-x/2}$$

(υπενθύμιση)

Περίπτωση της **συνάρτησης** **$Y=X^2$**

1ο βήμα: Βρίσκουμε την αντίστροφη συνάρτηση (**2 λύσεις**)

$$y = x^2 \Rightarrow x = g^{-1}(y) = \pm\sqrt{y}$$

2ο βήμα: Παίρνουμε την παράγωγο

$$\frac{dg^{-1}(y)}{dy} = \pm\frac{1}{2\sqrt{y}}$$

3ο βήμα: Εφαρμόζουμε τον γνωστό τύπο

$$f_Y(y) = \sum_{i=1}^{2} f_X\left(g_i^{-1}(y)\right)\left|\frac{dg_i^{-1}(y)}{dy}\right| = \frac{1}{2\sqrt{y}}\left(f_X\left(\sqrt{y}\right) + f_X\left(-\sqrt{y}\right)\right)$$

(υπενθύμιση)

Στην ειδική περίπτωση όπου *X* ακολουθεί την τυπική κανονική,

δηλ. ***X~N(0,1)*** και άρα $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

Αντικαθιστώντας στον τύπο: $f_Y(y) = \frac{1}{2\sqrt{y}}\left(f_X(\sqrt{y}) + f_X(-\sqrt{y})\right)$

και επειδή $f_X(\sqrt{y}) = f_X(-\sqrt{y}) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}}$

παίρνουμε:

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

που είναι η σ.π.π. της **κατανομής** Χι-τετράγωνο **με 1 βαθμό ελευθερίας (** ***X²(1)*** **)**

$$f(x) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\nu/2-1} e^{-x/2}$$

(III). Κατανομή $Z = X^2 + Y^2$ *αν X, Y ανεξάρτητες μεταβλητές που ακολουθούν την τυπική κανονική N(0,1)*

Πρόταση

- Αν οι ***X, Y ~***τότε είναι δύο **ανεξάρτητες** τυχαίες μεταβλητές που ακολουθούν την **τυπική κανονική** κατανομή, ***N(0,1)***, τότε η συνάρτηση η $Z=X^2+Y^2$ ακολουθεί την **Χι-τετράγωνο** κατανομή με ***2*** βαθμούς ελευθερίας, $X^2(2)$.

$$f_Z(z) = \frac{1}{2} e^{-\frac{z}{2}}$$

Κατανομή $Z = X^2 + Y^2$ αν X, Y ανεξάρτητες μεταβλητές που ακολουθούν την τυπική κανονική N(0,1)

- Προηγουμένως βρήκαμε ότι η $Y=X^2$ είναι κατανομή *χι-τετράγωνο* με 1 βαθμό ελευθερίας (**β.ε.**).

$$f_{Y=X^2}(y) = \frac{1}{\sqrt{2\pi}} y^{-\frac{1}{2}} e^{-\frac{y}{2}}$$

- Επομένως, η $Z = X^2 + Y^2$ είναι **άθροισμα 2 ανεξάρτητων μεταβλητών X_Γ, Y_Γ** :

$$Z = X_\Gamma + Y_\Gamma$$

που κάθε μία ακολουθεί την ***χι-τετράγωνο*** **με 1 β.ε.**

Κατανομή $Z = X^2 + Y^2 = X_\Gamma + Y_\Gamma$

- Η κατανομή της Z προκύπτει από την **συνέλιξη των δύο συναρτήσεων πυκνότητας** της κατανομής **χι-τετράγωνο** :

$$f_{X_\Gamma}(x) = \frac{1}{\sqrt{2\pi}} x^{-\frac{1}{2}} e^{-\frac{x}{2}} \qquad f_{Y_\Gamma}(y) = \frac{1}{\sqrt{2\pi}} y^{-\frac{1}{2}} e^{-\frac{y}{2}}$$

$$f_Z(z) = \int_{-\infty}^{+\infty} f_{X_\Gamma}(x) f_{Y_\Gamma}(z-x)dx = \frac{1}{2\pi}\int_0^{\infty} x^{-\frac{1}{2}} e^{-\frac{x}{2}} (z-x)^{-\frac{1}{2}} e^{-\frac{(z-x)}{2}} dx =$$

$$= \frac{1}{2\pi} e^{-\frac{z}{2}} \int_0^{\infty} x^{-\frac{1}{2}} (z-x)^{-\frac{1}{2}} dx = \frac{1}{2\pi} e^{-\frac{z}{2}} \int_0^{1} s^{-\frac{1}{2}} (1-s)^{-\frac{1}{2}} ds \quad \text{όπου } s = \frac{x}{z}$$

- Έτσι καταλήγουμε στο ολοκλήρωμα:

$$f_Z(z) = \frac{1}{2\pi} e^{-\frac{z}{2}} \int_0^1 s^{-\frac{1}{2}} (1-s)^{-\frac{1}{2}} ds$$

1

- Αλλά ισχύει ότι (***συνάρτηση Βήτα***)

$$\int_0^1 x^{a-1}(1-x)^{b-1} dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} = \mathrm{B}(a,b)$$

2

- Έτσι για $a = b = \frac{1}{2}$

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

- Έτσι καταλήγουμε στο ολοκλήρωμα:

$$f_Z(z)=\frac{1}{2\pi}e^{-\frac{z}{2}}\int_0^1 s^{-\frac{1}{2}}(1-s)^{-\frac{1}{2}}ds$$

1

- Αλλά ισχύει ότι (***συνάρτηση Βήτα***)

$$\int_0^1 x^{a-1}(1-x)^{b-1}dx=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}=\mathrm{B}(a,b)$$

2

- Έτσι για $a=b=\frac{1}{2}$

$$\mathrm{B}\left(\frac{1}{2},\frac{1}{2}\right)=\pi$$

3

- Συνδυάζοντας τις 1 και 2 παίρνουμε:

$$\boxed{f_Z(z)=\frac{1}{2\cancel{\pi}}e^{-\frac{z}{2}}\cancel{B\left(\frac{1}{2},\frac{1}{2}\right)}=\frac{1}{2}e^{-\frac{z}{2}}}$$

που είναι η συνάρτηση πυκνότητας πιθανότητας της **χι-τετράγωνο** **με** ***ν*****=2** **βαθμούς ελευθερίας,** ***Χ²(2)*** **.**

$$f(x)=\frac{1}{2^{\nu/2}\Gamma(\nu/2)}x^{\nu/2-1}e^{-x/2}$$

Γενίκευση για n ≥ 2 μεταβλητές

- Αν X_1, X_2, …, X_n ανεξάρτητες **τυπικές κανονικές** τυχαίες μεταβλητές, δηλ. $\boldsymbol{X_i \sim N(0, 1)}$, τότε η

$$Y = X_1^2 + X_2^2 + \ldots + X_n^2 = \sum_{i=1}^{n} X_i^2$$

ακολουθεί την ***χι-τετράγωνο με n β.ε.*** , $X^2(n)$:

$$\boxed{f_Y(y) = \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{n}{2}\right)} y^{\frac{n}{2}-1} e^{-\frac{n}{2}}}$$

Σημαντικό αποτέλεσμα στη Στατιστική

Ειδικές κατανομές πολυδιάστατων μεταβλητών

- Η **πολυωνυμική κατανομή** για *διακριτές μεταβλητές*

- Η **πολυδιάστατη κανονική κατανομή** για *συνεχείς μεταβλητές*