# Στοχαστικές Διεργασίες

## 5ο Εξάμηνο – Επιλογής

Διδάσκων: Επαμεινώνδας Διαμαντόπουλος (Ε.ΔΙ.Π)

# Διαδικασίες Διακλάδωσης

# Εισαγωγή

# Γεωμετρική κατανομή

Έστω **p η πιθανότητα επιτυχίας** σε μία δοκιμή ενός πειράματος. Ως "**γεωμετρική κατανομή**" περιγράφονται οι εξής δύο διακριτές κατανομές πιθανοτήτων:

- Η κατανομή πιθανότητας του πλήθους X ϵ N* των δοκιμών Bernoulli(p) που απαιτούνται για να επιτευχθεί μία επιτυχία. Η πιθανότητα ότι η n – οστή δοκιμή είναι η πρώτη επιτυχία είναι

$$\mathbf{P(X = n) = (1 – p)^{n-1} \cdot p,}\ (E(X) = 1/p,\ Var(X) = (1 – p)/p^2)$$

- Η κατανομή πιθανότητας του πλήθους Y = X − 1 (ϵ N) αποτυχιών πριν από την πρώτη επιτυχία. Στην περίπτωση αυτή η συνάρτηση πιθανότητας της Y είναι

$$\mathbf{P(Y = n) = (1 – p)^n \cdot p},\ (E(Y) = (1 – p)/p,\ Var(Y) = (1 – p)/p^2)$$

Και στις δύο περιπτώσεις γράφουμε **X ~ Geometric(p)** ή **Y ~ Geometric(p)** αντίστοιχα και θα διευκρινίζεται από το πλαίσιο αναφοράς το ακριβές είδος της κατανομής.

Σημείωση: Η προέλευση της ονομασίας είναι φανερή, καθώς και στις δύο περιπτώσεις η ακολουθία των πιθανοτήτων αποτελεί γεωμετρική πρόοδο.

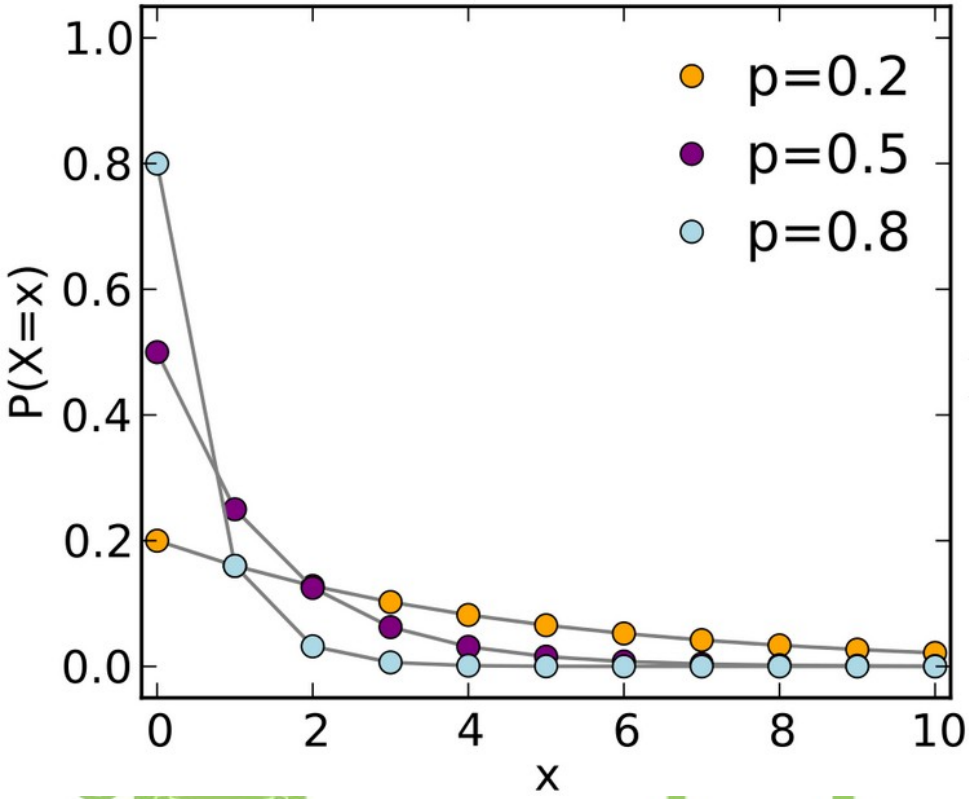
# Γεωμετρική κατανομή

Ενδεικτική συνάρτηση πιθανότητας μίας τ.μ.

$$X \sim \text{Geometric}(p),$$

για p = 0,2, 0,5, 0,8, όπου το X εκφράζει το πλήθος αποτυχιών πριν από την πρώτη επιτυχία για x = 0, 1, .., ,10 αποτυχίες (p: πιθανότητα επιτυχίας)

$$P(X = n) = (1 - p)^n \cdot p$$



# Πιθανογεννήτρια συνάρτηση

Έστω X μία διακριτή μεταβλητή που λαμβάνει ακέραιες τιμές. Η **πιθανογεννήτρια συνάρτηση (probability-generating function – PGF**) της τ.μ. X ορίζεται να είναι η συνάρτηση

$$G_X(z) = E(z^X) = \Sigma_{\kappa = 0, 1, \ldots} P(X = \kappa) \cdot z^\kappa$$

Καθώς $0 \leq P(X = \kappa) \leq 1$, $\kappa \in N$, η συνάρτηση $G_X$ συγκλίνει πάντα για κάθε $|z| < 1$. Είναι ωστόσο πιθανό να συγκλίνει σε χωρίο μεγαλύτερης ακτίνας, ακόμα και στο σύνολο του μιγαδικού επιπέδου C.

Σημειώσεις

Η πιθανογεννήτρια συνάρτηση $G_X(z)$ στη συνέχεια θα αξιοποιηθεί ως μία πραγματική συνάρτηση μίας πραγματικής μεταβλητής.

Η πιθανογεννήτρια $G_X(z)$ συνδέεται στενά με τη ροπογεννήτρια $M_X$ με τη σχέση $G_X(e^t) = M_X(t)$ ή ισοδύναμα $G_X(z) = M_X(\log z)$

# Πιθανογεννήτρια συνάρτηση

**Παράδειγμα**

Αν $X \sim B(n, p)$ τότε $P(X = \kappa) = (n \text{ ανά } \kappa)\, p^{\kappa} (1 - p)^{n - \kappa}$ και

$$G_X(z) = E(z^X) = \Sigma_{\kappa = 0, 1, \ldots, n} P(X = \kappa) \cdot z^{\kappa}$$

$$= \Sigma_{\kappa = 0, 1, \ldots, n} (n \text{ ανά } \kappa)\, p^{\kappa} (1 - p)^{n - \kappa} \cdot z^{\kappa}$$

$$= \Sigma_{\kappa = 0, 1, \ldots, n} (n \text{ ανά } \kappa)\, (zp)^{\kappa} (1 - p)^{n - \kappa}$$

$$= (zp + 1 - p)^n$$

$$= (1 + (z - 1)\, p)^n.$$

Άρα, $G_X(z) = (1 + (z - 1)\, p)^n$ για **κάθε z ϵ C**, ειδικότερα,

$G_X(x) = (1 + (x - 1)\, p)^n$, για **κάθε x ϵ R**.

# Πιθανογεννήτρια συνάρτηση

**Παράδειγμα**

Αν $X \sim \text{Geometric}(p)$ τότε $P(X = \kappa) = (1 - p)^{\kappa} \cdot p$, $\kappa = 0, 1, 2, \ldots$, και

$$G_X(z) = E(z^X) = \Sigma_{\kappa = 0, 1, \ldots} (1 - p)^{\kappa} \cdot p \cdot z^{\kappa}$$

$$= p \cdot \Sigma_{\kappa = 0, 1, \ldots} [z(1 - p)]^{\kappa}$$

$$= p / [1 - z(1 - p)]$$

Άρα, $G_X(z) = p / [1 - z(1 - p)]$ **για κάθε $|z(1 - p)| < 1$**, δηλαδή **$|z| < 1 / (1 - p)$**.

Σημείωση

Άθροισμα γεωμετρικής προόδου $\Sigma_{\kappa = 0, 1, \ldots} \alpha^{\kappa} = 1 + \alpha + \alpha^2 + \ldots = 1/(1 - \alpha)$ για $|\alpha| < 1$.

# Εισαγωγή

**Παράδειγμα**

Αν $X \sim \text{Poisson}(\lambda)$ τότε $P(X = \kappa) = \lambda^{\kappa} \cdot e^{-\lambda} / \kappa!$, $\kappa = 0, 1, 2, \ldots$, και

$$G_X(z) = E(z^X) = \sum_{\kappa = 0, 1, \ldots} P(X = \kappa) \cdot z^{\kappa}$$

$$= \sum_{\kappa = 0, 1, \ldots} \lambda^{\kappa} \cdot e^{-\lambda} / \kappa! \cdot z^{\kappa}$$

$$= e^{-\lambda} \cdot \sum_{\kappa = 0, 1, \ldots} (z\lambda)^{\kappa} / \kappa!$$

$$= e^{-\lambda} \cdot e^{z\lambda}$$

$$= e^{(z-1)\lambda}.$$

Άρα, $G_X(z) = e^{(z-1)\lambda}$ για **κάθε z ϵ C**, ειδικότερα, $G_X(x) = e^{(x-1)\lambda}$, για **κάθε x ϵ R**.

# Πιθανογεννήτρια συνάρτηση

Η συνάρτηση $G_X(z) = E(z^X) = \Sigma_{\kappa = 0, 1, \ldots} P(X = \kappa) \cdot z^\kappa$ στο χωρίο που ορίζεται καλώς, παραγωγίζεται ως εξής:

$$G^{(1)}_X(z) = \Sigma_{\kappa = 1, 2, \ldots} \kappa \cdot P(X = \kappa) \cdot z^{\kappa - 1}$$

$$= P(X = 1) + 2 \cdot P(X = 2) \cdot z + 3 \cdot P(X = 3) \cdot z^2 + \ldots$$

$$G^{(2)}_X(z) = \Sigma_{\kappa = 2, 3, \ldots} \kappa \cdot (\kappa - 1) \cdot P(X = \kappa) \cdot z^{\kappa - 2}$$

$$= 2 \cdot P(X = 2) + 6 \cdot P(X = 3) \cdot z + 12 \cdot P(X = 4) \cdot z^2 + \ldots$$

$$G^{(3)}_X(z) = \Sigma_{\kappa = 3, 4, \ldots} \kappa \cdot (\kappa - 1) \cdot (\kappa - 2) \cdot P(X = \kappa) \cdot z^{\kappa - 3}$$

$$= 6 \cdot P(X = 3) + 24 \cdot P(X = 4) \cdot z + 60 \cdot P(X = 5) \cdot z^2 + \ldots$$

...

$$G^{(n)}_X(z) = \Sigma_{\kappa = n, n+1, \ldots} \kappa! / (\kappa - n)! \cdot P(X = \kappa) \cdot z^{\kappa - n}$$

# Πιθανογεννήτρια συνάρτηση

**Η ονομασία "πιθανογεννήτρια συνάρτηση" για την $G_X(z)$ δικαιολογείται από τις εξής ιδιότητες:**

$G_X(0) = \Sigma_{\kappa = 0, 1, \ldots} P(X = 0) \cdot 0^{\kappa} = P(X = 0)$

$G^{(1)}_X(0) = \Sigma_{\kappa = 1, 2, \ldots} \kappa \cdot P(X = \kappa) \cdot 0^{\kappa-1} = P(X = 1)$

$G^{(n)}_X(0) / n! = 1 / n! \cdot \Sigma_{\kappa = n, n+1, \ldots} \kappa! / (\kappa - n)! \cdot P(X = \kappa) \cdot 0^{\kappa - n} = P(X = n)$

Συμπεραίνουμε ότι:

**Η πιθανογεννήτρια συνάρτηση $G_X(z)$ της διακριτής τυχαίας μεταβλητής X, προσδιορίζει πλήρως τη συνάρτηση πιθανότητάς $f(n) = P(X = n)$, $n = 0, 1, 2, \ldots$.**

Ιδιαίτερα, αν δύο τ.μ. έχουν την ίδια πιθανογεννήτρια τότε θα έχουν και την ίδια κατανομή πιθανότητας.

# Πιθανογεννήτρια συνάρτηση

**Δραστηριότητα**

Έστω X διακριτή τ.μ. που λαμβάνει τιμές φυσικούς αριθμούς με πιθανογεννήτρια συνάρτηση $G_X(z) = z(2 + 3z^2) / 5$. Να βρεθεί η συνάρτηση πιθανότητας $P(X = κ)$ της X, κ = 0, 1, 2, ….

Υπόδειξη

Υπολογίστε τις τιμές $G^{(n)}_X(0) / n!$ για διάφορές τιμές του n.

# Πιθανογεννήτρια συνάρτηση

**Η ονομασία πιθανογεννήτρια αιτιολογείται <u>και</u> από τις παρακάτω ιδιότητες**

Αν υπάρχει το όριο $\lim_{z \to 1} G^{(1)}_X(z)$ και $\lim_{z \to 1} G^{(1)}_X(z) = G^{(1)}_X(1)$, τότε για την πιθανογεννήτρια $G_X(z)$ ισχύουν τα εξής:

- $E(X) = \Sigma_{\kappa = 1, 2, \ldots} \kappa \cdot P(X = \kappa) = \lim_{z \to 1} G^{(1)}_X(z) = G^{(1)}_X(1)$
- $E(X \cdot (X - 1) \cdot \ldots \cdot (X - n + 1)) = \Sigma_{\kappa = n, n+1, \ldots} \kappa! / (\kappa - n)! \cdot P(X = \kappa) = \lim_{z \to 1} G^{(n)}_X(z) = G^{(n)}_X(1)$
- $\mathrm{Var}(X) = E(X - E(X))^2$

  $= E(X^2) - (EX)^2$

  $= E(X(X - 1)) + E(X) - (E(X))^2$

  $= \lim_{z \to 1} G^{(2)}_X(z) + \lim_{z \to 1} G^{(1)}_X(z) - [\lim_{z \to 1} G^{(1)}_X(z)]^2$

  $= G^{(2)}_X(1) + G^{(1)}_X(1) - [G^{(1)}_X(1)]^2$

# Η πιθανογεννήτρια ως πραγματική συνάρτηση μίας πραγματικής μεταβλητής

Δύο χρήσιμες παρατηρήσεις:

**Θεώρημα**

Η πιθανογεννήτρια ως πραγματική συνάρτηση μίας πραγματικής μεταβλητής:

(i) είναι αύξουσα στο διάστημα [0, 1].

(ii) Αν η τ.μ. X παίρνει έστω και μία τιμή μεγαλύτερη του 1, τότε η $G_X$ είναι κυρτή στο διάστημα [0, 1].

**Απόδειξη**

Πράγματι, είναι $G_X(s) = E(s^X) = \Sigma_{\kappa = 0, 1, \ldots} P(X = \kappa) \cdot s^\kappa$ , άρα

$$G^{(1)}_X(s) = \Sigma_{\kappa = 1, 2 \ldots} P(X = \kappa) \cdot \kappa \cdot s^{\kappa-1} \geq 0, \text{ και}$$

$$G^{(2)}_X(s) = \Sigma_{\kappa = 2, 3, \ldots} P(X = \kappa) \cdot \kappa \cdot (\kappa - 1) \cdot s^{\kappa-2} \geq 0,$$

για $0 \leq s \leq 1$.

# Πιθανογεννήτρια συνάρτηση αθροίσματος τ.μ.

Αν $X_1, X_2, \ldots, X_n$ είναι ανεξάρτητες τυχαίες μεταβλητές και $Y = X_1 + X_2 + \ldots + X_n$, τότε

$$G_Y(z) = G_{X1}(z) \cdot G_{X2}(z) \cdot \ldots G_{Xn}(z)$$

Πράγματι, $G_Y(z) = E(z^Y) = E(z^{X1 + X2 + \ldots + Xn}) = E(z^{X1} \cdot z^{X2} \cdot \ldots \cdot z^{Xn})$

$$= E(z^{X1}) \cdot E(z^{X2}) \cdot \ldots \cdot E(z^{Xn}) = G_{X1}(z) \cdot G_{X2}(z) \cdot \ldots G_{χn}(z),$$

γιατί $X_1, X_2, \ldots, X_n$ είναι ανεξάρτητες τ.μ. , άρα και οι $z^{X1}, z^{X2} \cdot \ldots \cdot z^{Xn}$ είναι ανεξάρτητες.

Σημειώσεις

1. Αν $X_1, X_2, \ldots, X_n$ ανεξάρτητες τ.μ. τότε $f_1(X_1), f_2(X_2), \ldots, f_n(X_n)$ ομοίως ανεξάρτητες για οποιεσδήποτε πραγματικές συναρτήσεις $f_i$, i = 1, …, n. (Ισχύει και στη συνεχή περίπτωση τ.μ. , αρκεί οι $f_i$, i = 1, …, n, να είναι ολοκληρώσιμες στο R).

2. Επιπλέον, ισχύει ότι αν $G_{X+Y}(z) = G_X(z) \cdot G_Y(z)$, τότε οι X, Y είναι ανεξάρτητες τ.μ.

# Χρήσιμες Προτάσεις

Αν X, Y ανεξάρτητες τ.μ. τότε $E(X \cdot Y) = E(X) \cdot E(Y)$.

## Απόδειξη

Αν X, Y ανεξάρτητες, τότε $f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$ και

$$
\begin{aligned}
E(XY) &= \int_{-\infty}^{\infty} f_{XY}(x,y)xydxdy \\
&= \int_{-\infty}^{\infty} f_X(x)f_Y(y)xydxdy \\
&= \int_{-\infty}^{\infty} f_X(x)xdx \int_{-\infty}^{\infty} f_Y(y)ydy \\
&= E(X)E(Y)
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{E}[XY] &= \sum_{x,y} xy\Pr(X=x, Y=y) \\
&= \sum_{x}\sum_{y} xy\Pr(X=x, Y=y) \\
&= \sum_{x}\sum_{y} xy\Pr(X=x)\Pr(Y=y) \\
&= \sum_{x} x\Pr(X=x)\sum_{y} y\Pr(Y=y) \\
&= \sum_{x} x\Pr(X=x)\,\mathbf{E}[Y] \\
&= \mathbf{E}[Y]\sum_{x} x\Pr(X=x) \\
&= \mathbf{E}[Y]\,\mathbf{E}[X]
\end{aligned}
$$

Σημείωση: Η αντίστοιχη ιδιότητα της αναμενόμενης τιμής για το άθροισμα τ.μ. ισχύει χωρίς κάποια προυπόθεση για τις τ.μ. X, Y. Δηλαδή για κάθε τ.μ. X, Y ισχύει $E(X + Y) = E(X) + E(Y)$.

# Χρήσιμες Προτάσεις

Αν X, Y ανεξάρτητες τ.μ. τότε $E(X \mid Y = y) = E(X)$

Απόδειξη (ενδεικτικά για διακριτές τ.μ.)

Πράγματι, $E(X \mid Y) = \Sigma_x P(X = x \mid Y = y) = \Sigma_x P(X = x) = E(X)$.

Σημείωση
Δύο μεταβλητές είναι ανεξάρτητες αν η τιμή της μίας δεν επηρεάζει το δειγματοχώρο μέσα στον οποίο υπολογίζεται η πιθανότητα της άλλης.

# Χρήσιμες Προτάσεις

Για κάθε δύο τ.μ. X, Y ισχύει $E(X) = E_Y(E_X(X|Y))$.

Απόδειξη (α) Συνεχείς τ.μ. και (β) Διακριτές τ.μ.

$$
\begin{aligned}
\mathrm{E}(\mathrm{E}(X \mid Y)) &= \int_{-\infty}^{\infty} \mathrm{E}(X \mid Y = y) f_Y(y)\,dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx\, f_Y(y)\,dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{X|Y}(x|y) f_Y(y)\,dx\,dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{X,Y}(x,y)\,dx\,dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{X,Y}(x,y)\,dy\,dx \\
&= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy\,dx \\
&= \int_{-\infty}^{\infty} x f_X(x)\,dx \\
&= \mathrm{E}(X)
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}(\mathrm{E}(X \mid Y)) &= \mathrm{E}\left[\sum_x x \cdot \mathrm{P}(X = x \mid Y)\right] \\
&= \sum_y \left[\sum_x x \cdot \mathrm{P}(X = x \mid Y = y)\right] \cdot \mathrm{P}(Y = y) \\
&= \sum_y \sum_x x \cdot \mathrm{P}(X = x, Y = y).
\end{aligned}
$$

$$
\sum_x \sum_y x \cdot \mathrm{P}(X = x, Y = y) = \sum_x x \sum_y \mathrm{P}(X = x, Y = y)
$$

$$
= \sum_x x \cdot \mathrm{P}(X = x)
$$

$$
= \mathrm{E}(X).
$$

Για το νόημα του όρου $E_Y(E_X(X|Y))$, ο αναγνώστης παραπέμπεται εδώ:
https://stats.stackexchange.com/questions/118578/what-is-the-difference-between-exy-and-exy-y

# Νόμος ολικής διακύμανσης

**Νόμος ολικής διακύμανσης (Law of total variance ή Eve's Law)**

Για κάθε δύο τ.μ. X, Y είναι:

$$\mathbf{Var(Y) = E_X(Var_Y(Y \mid X)) + Var_X(E_Y(Y \mid X))}$$

**Απόδειξη**

$$\operatorname{Var}[Y] = \operatorname{E}\left[Y^2\right] - \operatorname{E}[Y]^2$$

$$\operatorname{E}\left[Y^2\right] = \operatorname{E}\left[\operatorname{E}[Y^2 \mid X]\right] = \operatorname{E}\left[\operatorname{Var}[Y \mid X] + [\operatorname{E}[Y \mid X]]^2\right]$$

$$\begin{aligned}\operatorname{E}\left[Y^2\right] - \operatorname{E}[Y]^2 &= \operatorname{E}\left[\operatorname{Var}[Y \mid X] + [\operatorname{E}[Y \mid X]]^2\right] - [\operatorname{E}[\operatorname{E}[Y \mid X]]]^2 \\ &= (\operatorname{E}[\operatorname{Var}[Y \mid X]]) + \left(\operatorname{E}\left[\operatorname{E}[Y \mid X]^2\right] - [\operatorname{E}[\operatorname{E}[Y \mid X]]]^2\right) \\ &= \operatorname{E}[\operatorname{Var}[Y \mid X]] + \operatorname{Var}[\operatorname{E}[Y \mid X]]\end{aligned}$$

Πηγή: https://en.wikipedia.org/wiki/Law_of_total_variance

# Νόμος ολικής διακύμανσης

**Eve's Law** όπως **E(xpected value) V(ariance)** Law: $\mathbf{Var(Y) = E_X(Var_Y(Y \mid X)) + Var_X(E_Y(Y \mid X))}$

**Ερμηνεία του διαμερισμού**

Αν οι X, Y είναι ανεξάρτητες τότε $Var_X(E_Y(Y \mid X)) = 0$, καθώς $E_Y(Y \mid X) = E_Y(Y)$ και $Var_X(E(Y \mid X)) = Var_X(E(Y)) = 0$ (δεν υπάρχει μεταβλητότητα ως προς X).

Αν ωστόσο οι X, Y δεν είναι ανεξάρτητες τότε η $E_Y(Y \mid X = x))$ διαφοροποιείται καθώς το x διατρέχει όλες τις πιθανές τιμές της τ.μ. X και το μέγεθος της διαφοροποίησης σχετίζεται με το μέγεθος της εξάρτησης των δύο μεταβλητών. Συνεπώς, η **$Var_X(E_Y(Y \mid X))$ που ποσοτικοποιεί τη μεταβλητότητα αυτών των τιμών, μπορεί να ερμηνευτεί ως το μέρος της μεταβλητότητας της τ.μ. Y <u>που εξηγείται από την τ.μ. X</u>.**

Καθώς **Var(Y) είναι το σύνολο της μεταβλητότητας** της Y και $E_X(Var_Y(Y \mid X)) + Var_X(E_Y(Y \mid X)) = Var(Y)$, η εναπομένουσα ποσότητα **$E_X(Var_Y(Y \mid X))$ προσδιορίζεται ως το μέρος της μεταβλητότητας της τ.μ. Y <u>που δεν εξηγείται από την τ.μ. X</u>.**