Κατανομή συνάρτησης τυχαίας μεταβλητής

Πρόβλημα

- ✓ **Δίνεται** η τυχαία μεταβλητή ***X*** με γνωστή κατανομή, δηλ. **γνωρίζουμε** την συνάρτηση πυκνότητας πιθανότητας $f_X(x)$

- ✓ **Ζητείται** η κατανομή της συνάρτησης $Y=g(X)$ της μεταβλητής αυτής, δηλ. **ψάχνουμε** την συνάρτηση πυκνότητας πιθανότητας $f_Y(y)$

Κατανομή συνάρτησης τυχαίας μεταβλητής

- Έστω τ.μ. ***X*** με **γνωστή κατανομή**. Δηλαδή **γνωρίζουμε** την συνάρτηση πυκνότητας πιθανότητας $f_X(x)$ και την αθροιστική συνάρτηση κατανομής $F_X(x)$.

Συνεχής *τυχαία μεταβλητή*

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)\,dt$$

$$f(x) = \frac{dF(x)}{dx}$$

Διακριτή *τυχαία μεταβλητή*

$$F(x) = P(X \le x) = \sum_{t \le x} f(t)$$

$$f(x) = F(x) - F(x-1)$$

$$f(x) = P(X = x)$$

Κατανομή συνάρτησης τυχαίας μεταβλητής

- Έστω τ.μ. ***X*** με **γνωστή κατανομή**. Δηλαδή **γνωρίζουμε** την συνάρτηση πυκνότητας πιθανότητας $f_X(x)$ και την αθροιστική συνάρτηση κατανομής $F_X(x)$.

- Δίνεται η **συνάρτηση της τυχαίας μεταβλητής** $Y=g(X)$

- Η ***Y*** είναι επίσης **τυχαία μεταβλητή** καθώς είναι συνάρτηση μιας τυχαίας μεταβλητής.

  (καθώς μεταβάλλεται η ***X*** μεταβάλλεται και η ***Y***)

Μελέτη συνάρτησης τυχαίας μεταβλητής *Y=g(X)*

- **Α. Περίπτωση διακριτής τ.μ. *X***

- Έστω τυχαία μεταβλητή ***X*** με **πεδίο τιμών**

$$\Omega_X = \{x_1, x_2, \ldots, x_n\}$$

και **συνάρτηση πυκνότητας πιθανότητας**

$$f_X(x_i) = P(X = x_i) \quad \forall \;\; x_i \in \Omega_X$$

- Τότε η τ.μ. ***Y=g(X)*** έχει πεδίο τιμών

$$\Omega_Y = \{y_1, y_2, \ldots, y_k\} \subseteq \{g(x_1), g(x_2), \ldots, g(x_n)\}$$

➢ Α. Περίπτωση **διακριτής** τ.μ. Χ (συν.)

- Η **συνάρτηση πυκνότητας πιθανότητας** ***$f_Y(y)$*** της τυχαίας συνάρτησης *Y=g(X)* υπολογίζεται από τις **πιθανότητες** των τιμών της μεταβλητής ***X*** που αντιστοιχούν στις τιμές της μεταβλητής ***Y***.

  Έτσι:

$$\Omega_Y = \{y_1, y_2, \ldots, y_k\} \subseteq \{g(x_1), g(x_2), \ldots, g(x_n)\}$$

$$f_Y(y_j) = P(Y = y_j) = P(\{x_i\} : g(x_i) = y_j) =$$
$$= \sum_{x_i \,:\, g(x_i)=y_j} P(X = x_i) = \sum_{x_i \,:\, g(x_i)=y_j} f_X(x_i)$$

όλες οι τιμές x_i της τ.μ. ***X*** που αντιστοιχούν στη τιμή y_j της τ.μ. ***Y***

Παρατήρηση

- Καθώς γνωρίζουμε ότι $\sum_{x_i \in \Omega_X} f_X(x_i) = 1$ τότε

  είμαστε βέβαιοι ότι ισχύει $\sum_{y_j \in \Omega_Y} f_Y(y_j) = 1$

  καθώς:

$$\sum_{y_j \in \Omega_Y} f_Y(y_j) = \sum_{y_j \in \Omega_Y} \left[\sum_{x_i :\ g(x_i) = y_j} f_X(x_i) \right] = \sum_{x_i \in \Omega_X} f_X(x_i) = 1$$

- Βιβλιοπωλείο αγοράζει 10 βιβλία προς 10 ευρώ το καθένα και τα πουλάει προς 15 ευρώ. Ωστόσο, δεν επιτρέπεται να επιστρέψει τα απούλητα βιβλία. Αν η καθημερινή ζήτηση είναι μια *Διωνυμική* τυχαία μεταβλητή *X* με *n=10* και *ρ=0.6*, ποιο είναι το καθημερινό **αναμενόμενο κέρδος** και η **τυπική απόκλισή** του?

Λύση

- Κέρδος ***K***: τυχαία μεταβλητή

$$K = 5X - 10(10 - X) = 15X - 100$$

- Επομένως:

$$E(K) = 15E(X) - 100 \qquad V(K) = 15^2 V(X) = 225V(X)$$

όπου ***X*** είναι ***Διωνυμική*** τ.μ. και έτσι:

$$E(X) = n\rho = 6 \qquad V(X) = n\rho(1-\rho) = 2.4$$

$$\mu_K = E(K) = -10 \qquad \sigma_K = \sqrt{V(K)} = 15\sqrt{V(X)} = 23.24$$

- Ένα βιβλιοπωλείο αγοράζει 10 αντίτυπα ενός βιβλίου προς 20 ευρώ το ένα και τα πουλάει προς 30 ευρώ. Μετά από ένα έτος υπάρχει η δυνατότητα επιστροφής των απούλητων βιβλίων προς 15 ευρώ το ένα. Αν η κατανομή του αριθμού των αντίτυπων *X* που πουλιούνται δίνεται από τον τύπο:

  f(x) = (2x+1)/120 , *x={1, 2, 3, …, 10}* ,

  να υπολογιστεί η μέση τιμή του κέρδους του βιβλιοπωλείου σε ένα έτος από τις πωλήσεις του συγκεκριμένου βιβλίου.

Λύση

- Κέρδος Κ: τυχαία μεταβλητή $K = 10X - 5(10 - X) = 15X - 50$
- Επομένως: $E(K) = 15E(X) - 50$

  εύρεση **μέσης τιμής** της τ.μ. Χ

$$E(X) = \sum_{x=1}^{10} x f(x) = \frac{1}{120}\left(2\sum_{x=1}^{10} x^2 + \sum_{x=1}^{10} x\right) =$$

$$= \frac{1}{120}\left(2\frac{10(10+1)(20+1)}{6} + 55\right) = 6.875$$

- Ένα βιβλιοπωλείο αγοράζει 10 αντίτυπα ενός βιβλίου προς 20 ευρώ το ένα και τα πουλάει προς 30 ευρώ. Μετά από ένα έτος υπάρχει η δυνατότητα επιστροφής των απούλητων βιβλίων προς 15 ευρώ το ένα. Αν η κατανομή του αριθμού των αντίτυπων *X* που πουλιούνται δίνεται από τον τύπο:

  f(x) = (2x+1)/120 , *x={1, 2, 3, …, 10}* ,

  να υπολογιστεί η μέση τιμή του κέρδους του βιβλιοπωλείου σε ένα έτος από τις πωλήσεις του συγκεκριμένου βιβλίου.

Λύση

- Κέρδος Κ: τυχαία μεταβλητή $K = 10X - 5(10 - X) = 15X - 50$

- Επομένως:

$$E(X) = 6.875$$

$$E(K) = 15E(X) - 50 = 53.125$$

➢ Β. Περίπτωση συνεχής τ.μ. *X*

Αναγκαίες συνθήκες:

- ✓ Η *g(x)* είναι 1-1 (**γνησίως μονότονη**) και **παραγωγίσιμη**
- ✓ Υπάρχει η **αντίστροφος συνάρτηση** $g^{-1}(y)$

Τότε η συνάρτηση πυκνότητας πιθανότητας της τ.μ. ***Y*** είναι:

$$f_Y(y) = f_X\left(g^{-1}(y)\right)\left|\frac{dg^{-1}(y)}{dy}\right|$$

Κατανομή συνάρτησης Y=g(X)

1η περίπτωση: Η *g(x)* είναι **γνησίως αύξουσα**

- Δηλαδή (**i**) ***θετική παράγωγο*** και (**ii**) $g(x) \le y \Rightarrow x \le g^{-1}(y)$
- Χρήση της **αθροιστικής συνάρτησης κατανομής *$F_Y(y)$***:

$$F_Y(y) = P(Y \le y) = P(g(X) \le y) = P\left(X \le g^{-1}(y)\right) = F_X\left(g^{-1}(y)\right)$$

- **Παραγωγίζοντας την *$F_Y(y)$*** παίρνουμε την **σ.π.π. *$f_Y(y)$***:

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{dF_X\left(g^{-1}(y)\right)}{dy} = \boxed{\frac{dF_X\left(g^{-1}(y)\right)}{dg^{-1}(y)}}\frac{dg^{-1}(y)}{dy} =$$

$$= f_X\left(g^{-1}(y)\right)\frac{dg^{-1}(y)}{dy} \quad \ge 0 \qquad (1)$$

2η περίπτωση: Η g(x) είναι **γνησίως φθίνουσα**

- Δηλαδή (**i**) ***αρνητική παράγωγο*** (**ii**) $g(x) \le y \Rightarrow x \ge g^{-1}(y)$

- Χρησιμοποιώντας την συνάρτηση κατανομής ***F_Y(y)***:

$$F_Y(y) = P(Y \le y) = P\big(g(X) \le y\big) = P\big(X \ge g^{-1}(y)\big) =$$
$$= 1 - P\big(X \le g^{-1}(y)\big) = 1 - F_X\big(g^{-1}(y)\big)$$

- **Παραγωγίζοντας την *F_Y(y)*** παίρνουμε την σ.π.π. ***f_Y(y)***:

$$f_Y(y) = \frac{dF_Y(y)}{dy} = -\frac{dF_X\big(g^{-1}(y)\big)}{dy} = -\boxed{\frac{dF_X\big(g^{-1}(y)\big)}{dg^{-1}(y)}}\frac{dg^{-1}(y)}{dy} =$$
$$= -f_X\big(g^{-1}(y)\big)\boxed{\frac{dg^{-1}(y)}{dy}} \quad \le 0 \qquad (2)$$

Συνδυάζοντας τις (1), (2) παίρνουμε τον **γενικό τύπο:**

Αν η g(x) είναι **γνησίως αύξουσα**

$$f_Y(y) = f_X\left(g^{-1}(y)\right)\frac{dg^{-1}(y)}{dy} \qquad (1)$$

Αν η g(x) είναι **γνησίως φθίνουσα**

$$f_Y(y) = -f_X\left(g^{-1}(y)\right)\frac{dg^{-1}(y)}{dy} \qquad (2)$$

Συνδυάζοντας τις (1), (2) παίρνουμε τον **γενικό τύπο:**

$$\boxed{f_Y(y) = f_X\left(g^{-1}(y)\right)\left|\frac{dg^{-1}(y)}{dy}\right|}$$

ή εναλλακτικά

$$\boxed{f_Y(y) = f_X(x)\left(\frac{dg(x)}{dx}\right)^{-1}_{x=g^{-1}(y)}}$$

$$\left(\boldsymbol{f}^{-\mathbf{1}}\right)'(y) = \frac{1}{\boldsymbol{f}'(x)} \quad \text{όπου } x = g^{-1}(y)$$

Παρατήρηση

- Σε περίπτωση όπου **η αντίστροφος συνάρτηση *g***

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

Ειδικές περιπτώσεις συναρτήσεων

(A). Περίπτωση **γραμμικής συνάρτησης** ***Y=aX+b***

1ο βήμα: Βρίσκουμε την **αντίστροφη συνάρτηση**

$$y = ax + b \Rightarrow x = g^{-1}(y) = \frac{y-b}{a}$$

2ο βήμα: Παίρνουμε την **παράγωγό της**

$$\frac{dg^{-1}(y)}{dy} = \frac{1}{a}$$

3ο βήμα: **Εφαρμόζουμε τον γνωστό τύπο**

$$f_Y(y) = f_X\left(g^{-1}(y)\right)\left|\frac{dg^{-1}(y)}{dy}\right| = \left|\frac{1}{a}\right| f_X\left(\frac{y-b}{a}\right)$$

Εφαρμογή στη περίπτωση της Κανονικής κατανομής

- Έστω τ.μ. X ~ N(μ,σ²) $f_X(x) = \boxed{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}}$
- Βρίσκουμε την κατανομή της ***Y=αX+b***
- Εφαρμόζοντας τον τύπο $f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$ παίρνουμε:

$$f_Y(y) = \frac{1}{|a|}\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\left(\frac{y-b}{a}-\mu\right)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}|a|\sigma} e^{-\frac{(y-b-a\mu)^2}{2a^2\sigma^2}}$$

$$\mu_Y = a\mu_X + b$$
$$\sigma_Y^2 = (a\sigma_X)^2$$

$$= \frac{1}{\sqrt{2\pi}(|a|\sigma)} e^{-\frac{(y-(b+a\mu))^2}{2(a\sigma)^2}} = \boxed{\frac{1}{\sqrt{2\pi}\sigma_y} e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}}}$$

N(y ; μ$_y$, σ^2_y)

Συμπέρασμα: Γραμμικότητα της Κανονικής

«Οποιοσδήποτε γραμμικός συνδυασμός μιας Κανονικής μεταβλητής είναι και αυτός Κανονική μεταβλητή»

$$\text{αν} \quad X \sim N\left(\mu_X, \sigma_X^2\right) \Rightarrow$$

$$Y = aX + b \sim N\left(\mu_Y = a\mu_X + b, \sigma_Y^2 = \left(a\sigma_X\right)^2\right)$$

Έτσι η αν $X \sim N\left(\mu, \sigma^2\right)$

Τότε: $$Y = \frac{(X-\mu)}{\sigma} = \overset{(a)}{\frac{1}{\sigma}} X + \overset{(b)}{\left(-\frac{\mu}{\sigma}\right)} \sim N(0,1)$$

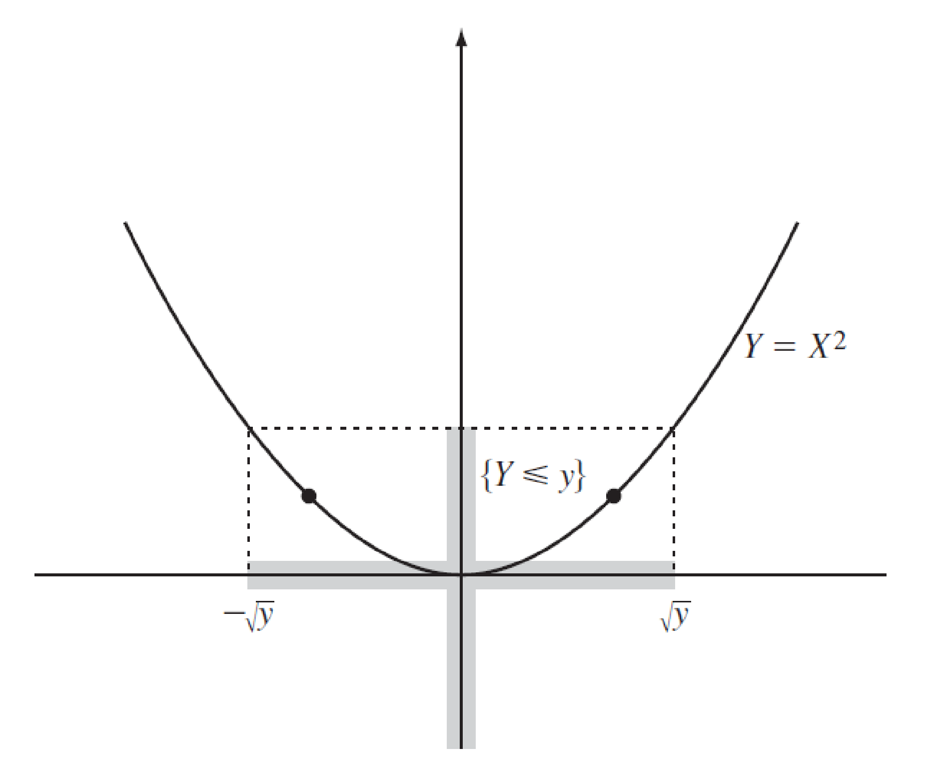
(***τυπική κανονική***)

καθώς:

$$\mu_Y = a\mu_X + b = \frac{1}{\sigma}\mu + \left(-\frac{\mu}{\sigma}\right) = 0$$

$$\sigma_Y^2 = (a\sigma_X)^2 = \left(\frac{1}{\sigma}\sigma\right)^2 = 1$$

(B). Περίπτωση της **συνάρτησης** ***Y=X²***



(B). 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 (**2 λύσεις**)

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

- 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:

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

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

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

- Τότε:

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

που είναι η σ.π.π. της ***X²(ν=1)*** **(χι τετράγωνο) κατανομής με ν=1 βαθμό ελευθερίας**

- **Υπενθύμιση:** η *X²(ν)* έχει σ.π.π.

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

- **<u>Πόρισμα:</u>** Το τετράγωνο της τυπικής κανονικής κατανομής είναι ***Χι-τετράγωνο*** με **1** βαθμό ελευθερίας (β.ε.)

δηλ. $\boxed{\textbf{αν } \boldsymbol{X \sim N(0,1) \Rightarrow Y = X^2 \sim \mathrm{X}^2(v=1)}}$

- ***<u>Γενίκευση:</u>*** Το άθροισμα των τετραγώνων ***n*** τυπικών κανονικών μεταβλητών ακολουθεί την ***Χι-τετράγωνο*** κατανομή με ***n*** βαθμούς ελευθερίας (β.ε.)

$$\alpha\nu \quad X_i \sim N(0,1) \quad \Rightarrow \sum_{i=1}^{n} X_i^2 \sim \mathbf{X}^2(\boldsymbol{\nu} = \boldsymbol{n})$$

Παραδείγματα κατανομής συνάρτησης

- Έστω τ.μ. Χ ομοιόμορφα κατανεμημένη στο [0, 1]. Να βρεθεί η κατανομή της $Y = -\ln X$ και στη συνέχεια η μέση τιμή και η διακύμανσή της.

- Έστω τ.μ. X με σ.π.π. $f_X(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}$, *x>0*.

  Να βρεθεί η κατανομή του ***Y=X²*** και στη συνέχεια η μέση τιμή και η διακύμανσή του.

- Έστω τ.μ. Χ με σ.π.π. $f_X(x)=2x,\ 0<x<1$. Να βρεθεί η κατανομή της τυχαίας μεταβλητής $Y=e^{-X}$ και στη συνέχεια η μέση τιμή και η διακύμανσή της.

Το μέγεθος αποζημίωσης εκφράζεται ως μια τυχαία μεταβλητή X με συνάρτηση πυκνότητας: $f(x)=\begin{cases} k(1-x)^4 & 0<x<1 \\ 0 & \alpha\lambda\lambda o\upsilon \end{cases}$, όπου k μία σταθερά.

Αν *Y=100000*X* εκφράζει το ετήσιο ύψος αποζημίωσης, να βρεθεί η πιθανότητα το *Y* να ξεπεράσει το ποσό των *40000* εάν είναι γνωστό ότι έχει ξεπεράσει το ποσό των *10000*.