ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΙ ΜΕΤΑΒΛΗΤΩΝ

Όπως έχει αποδειχθεί *(βλέπε π.χ. Ε. Ξεκαλάκη και Ι. Πανάρετο, 1993)*, οι αναμενόμενες τιμές $E(X)$ και $E(X - m^2)$ παρέχουν σημαντικές πληροφορίες σχετικά με την κατανομή μιας πραγματικής τυχαίας μεταβλητής X. Επίσης, οι αναμενόμενες τιμές $E(s^X) = G_X(s)$, $E(e^{sX}) = M_X(s)$ και $E(e^{isX}) = \varphi(s)$ έχουν ένα πολύ σημαντικό ρόλο όσο αφορά τον μονοσήμαντο προσδιορισμό της κατανομής της τυχαίας μεταβλητής X. Κάθε μία από αυτές αποτελεί την αναμενόμενη τιμή μιας συνάρτησης, ή μετασχηματισμού της τυχαίας μεταβλητής X, είναι δηλαδή της μορφής $E(g(X))$. Οι αναμενόμενες τιμές, ως ολοκληρώματα, έχουν την μορφή μετασχηματισμών, οι οποίοι οδηγούν σε συναρτήσεις που περιέχουν μεγάλη ποσότητα πληροφόρησης για την κατανομή της τυχαίας μεταβλητής. Οι μετασχηματισμοί αυτοί βεβαίως σχετίζονται με κατανομές μάλλον παρά με τυχαίες μεταβλητές. Συχνά, όμως, ενδιαφερόμαστε για ένα άλλο είδος μετασχηματισμών οι οποίοι αναφέρονται στις τυχαίες μεταβλητές αυτές καθεαυτές. Ενδιαφερόμαστε δηλαδή για μετασχηματισμούς της μορφής $Y = g(X)$, οι οποίοι οδηγούν από την τυχαία μεταβλητή X στην τυχαία μεταβλητή Y. Η σημασία και η χρησιμότητα αυτών των μετασχηματισμών γίνεται φανερή στις περιπτώσεις όπου οι τυχαίες μεταβλητές που χρησιμοποιούμε σε σχέση με κάποιο φαινόμενο δεν εκπροσωπούν πάντα τις μετρήσεις που ελήφθησαν. Συχνά, χρησιμοποιούνται συναρτήσεις ή μετασχηματισμοί αυτών των μετρήσεων. Για παράδειγμα, κατά την μέτρηση μιας γωνίας, συχνά καταγράφεται η εφαπτομένη της γωνίας αυτής. Αντίστοιχα, μία μέτρηση της θερμοκρασίας στην κλίμακα Fahrenheit μπορεί να απαιτείται να μετασχηματισθεί σε μέτρηση της κλίμακας Κελσίου.

Στη συνέχεια, θα εξετάσουμε μεθόδους προσδιορισμού των κατανομών μετασχηματισμένων τυχαίων μεταβλητών $Y = g(X)$ με βάση την γνώση της κατανομής της αρχικής τυχαίας μεταβλητής X.

ΓΡΑΜΜΙΚΟΙ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΙ

Ας θεωρήσουμε την ειδική περίπτωση όπου η τυχαία μεταβλητή Y είναι ένας γραμμικός συνδυασμός $Y = aX + b$, $a \neq 0$, της τυχαίας μεταβλητής X.
Η συνάρτηση κατανομής της τυχαίας μεταβλητής Y μπορεί να προσδιορισθεί βάσει της συνάρτησης κατανομής της τυχαίας μεταβλητής X ως εξής:

Αν $a > 0$, τότε

$$F_Y(y) = P(Y \le y) = P(aX + b \le y) = P(X \le \frac{y-b}{a}) = F_X(\frac{y-b}{a}).$$

Αν $a < 0$, τότε

$$F_Y(y) = P(aX + b \le y) = P(X \ge \frac{y-b}{a}) = 1 - P(X < \frac{y-b}{a})$$

$$= 1 - P(X \le \frac{y-b}{a}) + P(X = \frac{y-b}{a}) = 1 - F_X(\frac{y-b}{a}) + P(X = \frac{y-b}{a}).$$

Προφανώς, εάν η τυχαία μεταβλητή X διακριτή, η συνάρτηση πιθανότητας της τυχαίας μεταβλητής Y μπορεί να προσδιορισθεί μέσω της συνάρτησης πιθανότητας της τυχαίας μεταβλητής X είτε με τη βοήθεια των παραπάνω τύπων είτε απευθείας ως αποτέλεσμα του γεγονότος ότι τα ενδεχόμενα {Y= y} και {X=(y-b)/a} ταυτίζονται, έτσι ώστε

$$P_Y(y) = P(Y = y) = P(X = \frac{y-b}{a}) = P_X(\frac{y-b}{a}).$$

Με όμοιο τρόπο, αν η τυχαία μεταβλητή X είναι συνεχής, η συνάρτηση πυκνότητας πιθανότητας $f_Y(y)$ της τυχαίας μεταβλητής Y μπορεί να προσδιορισθεί με βάση την συνάρτηση πυκνότητας πιθανότητας $f_X(x)$ της τυχαίας μεταβλητής X παραγωγίζοντας την συνάρτηση κατανομής της τυχαίας μεταβλητής Y ως προς y.
Έτσι, αν a>0,

$$f_Y(y) = \frac{d}{dy}F_Y(y) = F_X'(\frac{y-b}{a})\frac{d}{dy}(\frac{y-b}{a}) = \frac{1}{a}f_X(\frac{y-b}{a}),$$

ενώ, αν a<0,

$$f_Y(y) = -\frac{1}{a} f_X\left(\frac{y-b}{a}\right).$$

Δηλάδή, τελικά

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right).$$

Παράδειγμα *(διακριτή περίπτωση)*: Έστω ότι η τυχαία μεταβλητή X έχει συνάρτηση πιθανότητας

$$P_X(x) = \begin{cases} \frac{64}{85}\left(\frac{1}{4}\right)^x, & \text{αν} \quad x = 0,1,2,3 \\ 0, & \text{διαφορετικά} \end{cases}$$

Έστω $Y = 2X + 3$. Τότε,

$$P_Y(y) = P_X\left(\frac{y-3}{2}\right) = \begin{cases} \frac{64}{85}\left(\frac{1}{4}\right)^{(y-3)/2}, & \text{αν} \quad \frac{y-3}{2} = 0,1,2,3 \\ 0, & \text{διαφορετικά} \end{cases}$$

$$= \begin{cases} \frac{512}{85}\left(\frac{1}{2}\right)^y, & \text{αν} \quad y = 3,5,7,9 \\ 0, & \text{διαφορετικά} \end{cases}$$

Παράδειγμα *(συνεχής περίπτωση)*: Έστω ότι η τυχαία μεταβλητή X έχει συνάρτηση πυκνότητας πιθανότητας

$$f_X(x) = \begin{cases} e^{-x}, & \text{αν} \quad x > 0 \\ 0, & \text{διαφορετικά} \end{cases}$$

Αν $Y = \frac{5}{9}X - \frac{160}{9}$, τότε

$$f_Y(y)=\begin{cases}\frac{9}{5}e^{-(9y+160)/5}, & \text{αν} \quad y>-\frac{160}{9}\\ 0, & \text{διαφορετικά}\end{cases}$$

ΓΕΝΙΚΟΙ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΙ

ΔΙΑΚΡΙΤΗ ΠΕΡΙΠΤΩΣΗ

Όταν η τυχαία μεταβλητή X είναι διακριτή, η πιο άμεση μέθοδος καθορισμού της κατανομής της τυχαίας μεταβλητής Y= g(X), συνιστάται στον προσδιορισμό, για κάθε τιμή y, της πιθανότητας του ενδεχομένου {X: g(X)=y}. Προσδιορίζουμε, δηλαδή, για κάθε τιμή y όλες τις δυνατές τιμές της τυχαίας μεταβλητής X, οι οποίες οδηγούν σε αυτήν την τιμή του y και μετά προσθέτουμε τις πιθανότητες των τιμών αυτών. Η προκύπτουσα ποσότητα είναι η τιμή της πιθανότητας $P_Y(y)$. Έτσι,

$$P_Y(y)=\sum_{\{x:g(x)=y\}}P_X(x).$$

Παράδειγμα: Έστω ότι η τυχαία μεταβλητή X έχει συνάρτηση πιθανότητας

$$P_X(x)=\begin{cases}\frac{4}{31}(\frac{1}{2})^x, & \text{αν } x=-2,-1,0,1,2\\ 0, & \text{διαφορετικά}\end{cases}$$

Αν $Y = X^2$, οι δυνατές τιμές της τυχαίας μεταβλητής Y είναι 0, 1, 4. Τότε,

$$P_Y(0)=\sum_{\{x:x^2=0\}}P_X(x)=P_X(0)=\frac{4}{31}$$

$$P_Y(1)=\sum_{\{x:x^2=1\}}P_X(x)=P_X(-1)+P_X(1)=\frac{8}{31}+\frac{2}{31}=\frac{10}{31}$$

$$P_Y(4)=\sum_{\{x:x^2=4\}}P_X(x)=P_X(-2)+P_X(2)=\frac{16}{31}+\frac{1}{31}=\frac{17}{31},$$

και $P_Y(y) = 0$, για όλες τις άλλες τιμές του y.

ΣΥΝΕΧΗΣ ΠΕΡΙΠΤΩΣΗ

Στην περίπτωση που η τυχαία μεταβλητή X είναι συνεχής, για τον προσδιορισμό της τιμής της τυχαίας μεταβλητής Y= g(X) απαιτείται η πρόσθετη υπόθεση ότι η συνάρτηση g(y) είναι γνησίως μονότονη και διαφορίσιμη. Η υπόθεση αυτή γίνεται για να εξασφαλισθεί η ύπαρξη της αντίστροφης συνάρτησης $y = g^{-1}(x) = h(x)$.
Στην περίπτωση αυτή έχουμε

$$F_Y(y) = P(g(X) \le y) = P(Y \le y)$$
$$= \begin{cases} P(X \le h(y)) = F_X(h(y)), & \text{αν η g είναι γνησίως αύξουσα} \\ P(X \ge h(y)) = 1 - F_X(h(y)), & \text{αν η g είναι γνησίως φθίνουσα} \end{cases}$$

Τότε,

$$f_Y(y) = \frac{dF_Y(y)}{dy}$$
$$= \begin{cases} f_X(h(y))\frac{dh(y)}{dy}, & \text{αν η g είναι γνησίως αύξουσα} \\ -f_X(h(y))\frac{dh(y)}{dy}, & \text{αν η g είναι γνησίως φθίνουσα} \end{cases}$$

Κατά συνέπεια,

$$f_Y(y) = \left|\frac{dh(y)}{dy}\right| f_X(h(y)).$$

<u>Παράδειγμα:</u> Έστω ότι μία συνεχής τυχαία μεταβλητή X έχει συνάρτηση πυκνότητας πιθανότητας

$$f_X(x) = \begin{cases} \theta e^{-\theta x}, & \text{αν } x > 0 \\ 0, & \text{διαφορετικά} \end{cases}$$

όπου θ>0. Έστω $Y = e^X$. Έχουμε δηλαδή εδώ ένα μετασχηματισμό της μορφής Y= g(X), όπου η συνάρτηση $g(x) = e^x$ είναι γνησίως αύξουσα και διαφορίσιμη. Επομένως, $y = e^x$ τότε και μόνο τότε εάν

$$x = g^{-1}(y) = h(y) = \log y .$$

Άρα, σύμφωνα με τα προηγούμενα, η συνάρτηση πυκνότητας πιθανότητας της τυχαίας μεταβλητής Y είναι

$$f_Y(y) = \left| \frac{d}{dy} \log y \right| f_x(\log y)$$

$$= \begin{cases} \frac{1}{\theta}\theta e^{-\theta \log y y}, & \text{αν } 0 \le \log y < \infty \\ 0, & \text{διαφορετικά} \end{cases}$$

$$= \begin{cases} \theta y^{-\theta-1}, & \text{αν } 1 \le y < \infty \\ 0, & \text{διαφορετικά} \end{cases}$$