Πολυδιάστατες τυχαίες μεταβλητές
Περίληψη

Αντώνης Οικονόμου
aeconom@math.uoa.gr

24 Απριλίου 2010

Η **από κοινού συνάρτηση κατανομής** ενός ζεύγους τυχαίων μεταβλητών X και Y ορίζεται ως

$$F(x,y) = P(X \leq x, Y \leq y), \quad -\infty < x, y < \infty.$$

Όλες οι πιθανότητες που αφορούν ενδεχόμενα σχετιζόμενα με τις τιμές αυτών των δυο μεταβλητών μπορούν να εκφραστούν μέσω της $F(x,y)$. Υπό αυτή την έννοια, η γνώση της από κοινού συνάρτησης κατανομής επιτρέπει τον υπολογισμό οποιουδήποτε ενδεχομένου σχετίζεται με τις τιμές αυτών των δυο μεταβλητών. Για να πάρουμε τις συναρτήσεις κατανομής των X και Y αντίστοιχα, που ονομάζονται **περιθώριες συναρτήσεις κατανομών** χρησιμοποιούμε τις σχέσεις

$$F_X(x) = P(X \leq x) = \lim_{y \to \infty} F(x,y), \quad F_Y(y) = P(Y \leq y) = \lim_{x \to \infty} F(x,y).$$

Αν οι X και Y είναι και οι δυο διακριτές τυχαίες μεταβλητές, τότε το ζεύγος (X,Y) αναφέρεται ως **διακριτή διδιάστατη τυχαία μεταβλητή** και οι υπολογισμοί πιθανοτήτων που αναφέρονται στις X, Y μπορούν να γίνουν χρησιμοποιώντας την **από κοινού συνάρτηση πιθανότητάς** τους που ορίζεται από τη σχέση

$$p(x,y) = P(X = x, Y = y).$$

Οι συναρτήσεις πιθανότητας των X και Y που ονομάζονται τώρα **περιθώριες συναρτήσεις πιθανότητας** υπολογίζονται από τις σχέσεις

$$p_X(x) = P(X = x) = \sum_y p(x,y), \quad p_Y(y) = P(Y = y) = \sum_x p(x,y).$$

Οι τυχαίες μεταβλητές X, Y λέμε ότι είναι **από κοινού συνεχείς** ή ότι το ζεύγος (X,Y) είναι μια **συνεχής διδιάστατη τυχαία μεταβλητή**, αν υπάρχει μια μη αρνητική συνάρτηση $f(x,y)$, τέτοια ώστε για κάθε υποσύνολο C,

$$P((X,Y) \in C) = \iint_C f(x,y)\,dx\,dy.$$

Η συνάρτηση $f(x,y)$ ονομάζεται τότε **από κοινού συνάρτηση πυκνότητας πιθανότητάς** των X και Y. Από τον προηγούμενο ορισμό είναι φανερό ότι για $f(x,y)$ συνεχή έχουμε τη σχέση-ερμηνεία

$$P(x < X \leq x + dx, y < Y \leq y + dy) \simeq f(x,y)\,dx\,dy, \quad dx\,dy \to 0.$$

Οι συναρτήσεις πυκνότητας πιθανότητας των X και Y που ονομάζονται τώρα **περιθώριες συναρτήσεις πυκνότητας πιθανότητας** υπολογίζονται από τις σχέσεις

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx.$$

Δυο τυχαίες μεταβλητές X και Y λέγονται **ανεξάρτητες** αν για οποιαδήποτε σύνολα A και B ισχύει

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B).$$

Αναγκαία και ικανή συνθήκη για να είναι δυο τυχαίες μεταβλητές X και Y ανεξάρτητες είναι η από κοινού συνάρτηση κατανομής τους, $F(x, y)$ (αντίστοιχα η από κοινού συνάρτηση πιθανότητάς τους για διακριτές ή η από κοινού συνάρτηση πυκνότητας πιθανότητάς τους για συνεχείς), να παραγοντοποιείται σε δυο μέρη που να εξαρτώνται το ένα μόνο από τη μεταβλητή x και το άλλο μόνο από τη μεταβλητή y. Γενικότερα, οι μεταβλητές $X_1, X_2, \ldots, X_n$ είναι ανεξάρτητες αν για οποιαδήποτε σύνολα $A_1, A_2, \ldots, A_n$, ισχύει

$$P(X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n) = P(X_1 \in A_1)P(X_2 \in A_2) \cdots P(X_n \in A_n).$$

Αν X και Y είναι διακριτές τυχαίες μεταβλητές με από κοινού συνάρτηση πιθανότητας $p(x, y)$, τότε η συνάρτηση πιθανότητας του αθροίσματός τους δίνεται από τη σχέση

$$p_{X+Y}(z) = \sum_x p(x, z - x) = \sum_y p(z - y, y).$$

Ανάλογα, αν οι X και Y είναι συνεχείς τυχαίες μεταβλητές με από κοινού συνάρτηση πυκνότητας πιθανότητας $f(x, y)$, τότε η συνάρτηση πυκνότητας πιθανότητας του αθροίσματός τους δίνεται από τη σχέση

$$f_{X+Y}(z) = \int_{-\infty}^{\infty} f(x, z - x)dx = \int_{-\infty}^{\infty} f(z - y, y)dy.$$

Αν οι X και Y είναι διακριτές τυχαίες μεταβλητές με από κοινού συνάρτηση πιθανότητας $p(x, y)$, τότε η δεσμευμένη συνάρτηση πιθανότητας της X, δοθέντος ότι $Y = y$, ορίζεται μέσω της σχέσης

$$p_{X|Y}(x|y) = P(X = x|Y = y) = \frac{p(x, y)}{p_Y(y)} = \frac{P(X = x, Y = y)}{P(Y = y)}.$$

Ανάλογα αν οι X και Y είναι συνεχείς τυχαίες μεταβλητές με από κοινού συνάρτηση πυκνότητας πιθανότητας $f(x, y)$, τότε η δεσμευμένη συνάρτηση πυκνότητας πιθανότητας της X, δοθέντος ότι $Y = y$, ορίζεται μέσω της σχέσης

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

Αν έχουμε τις ανεξάρτητες τυχαίες μεταβλητές $X_1, X_2, \ldots, X_n$ με αντίστοιχες συναρτήσεις κατανομής $F_{X_1}(x), F_{X_2}(x), \ldots, F_{X_n}(x)$, τότε η συνάρτηση κατανομής της μέγιστης, $X_{\max} = \max(X_1, X_2, \ldots, X_n)$, δίνεται από τη σχέση

$$F_{X_{\max}}(x) = F_{X_1}(x)F_{X_2}(x) \cdots F_{X_n}(x),$$

ενώ η συνάρτηση κατανομής της ελάχιστης, $X_{\min} = \min(X_1, X_2, \ldots, X_n)$, δίνεται από τη σχέση

$$F_{X_{\min}}(x) = 1 - (1 - F_{X_1}(x))(1 - F_{X_2}(x)) \cdots (1 - F_{X_n}(x)).$$