Εκτίμηση τυχαίας μεταβλητής. Ελάχιστο μέσο τετραγωνικό σφάλμα.

Εκτιμήτρια ελαχίστου μέσου τετραγωνικού σφάλματος

- Έστω ότι θέλουμε να εκτιμήσουμε την τυχαία μεταβλητή X με βάση τη γνώση που μπορούμε να έχουμε για μια άλλα τυχαία μεταβλητή την Y εφόσον όμως γνωρίζουμε την δεσμευμένη πυκνότητα (πιθανότητας) $f_{X|Y}(x|y)$.

- Τότε η εκτιμήτρια $\hat{x}_M = E[X|Y = y]$ ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα (Minimum Mean Squared Error, MMSE, estimate)

- Επίσης καλείται εκτιμήτρια ελαχίστων τετραγώνων (Least Mean Squares, LMS, estimate).

- Για να μπορέσουμε να βρούμε το $\hat{x}_M = E[X|Y = y]$ πρέπει είτε να γνωρίζουμε τη δεσμευμένη πυκνότητα (πιθανότητας) $f_{X|Y}(x|y)$ είτε να μας δίνεται η πυκνότητα (πιθανότητας) $f_X(x)$ και η δεσμευμένη πυκνότητα (πιθανότητας) $f_{Y|X}(y|x)$.

Εκτιμήτρια ελαχίστου μέσου τετραγωνικού σφάλματος

<u>Θεώρημα:</u> Έστω $\hat{X} = g(Y)$ μια εκτιμήτρια της τυχαίας μεταβλητής X, δεδομένου ότι έχουμε παρατηρήσει την τυχαία μεταβλητή Υ. Τότε το μέσο τετραγωνικό σφάλμα αυτής της εκτιμήτριας ορίζεται ως

$$E\left[\left(X - \hat{X}\right)^2\right] = E\left[\left(X - g(Y)\right)^2\right].$$

Η εκτιμήτρια $\hat{x}_M = E[X|Y = y]$ έχει το ελάχιστο μέσο τετραγωνικό σφάλμα από όλες τις δυνατές εκτιμήτριες.

Εκτιμήτρια ελαχίστου μέσου τετραγωνικού σφάλματος - Ιδιότητες

- $E[\hat{x}_M] = E[E[X|Y = y]] = E[X]$
- Η τυχαία μεταβλητή $\tilde{X}$ ορίζεται ως $\tilde{X} = X - \hat{x}_M$ με $E[\tilde{X}] = 0$.
- Τότε $E[\tilde{X}|Y] = 0$ και για κάθε συνάρτηση $g(Y)$ της τυχαίας μεταβλητής Y έχουμε ότι $E[\tilde{X}g(Y)] = E[E[\tilde{X}g(Y)|Y]] = 0$.
- $Cov(\tilde{X}, \hat{x}_M) = 0$
- $Var(X) = Var(\tilde{X} + \hat{x}_M) = Var(\tilde{X}) + Var(\hat{x}_M)$
- $E[X^2] = E[\tilde{X}^2] + E[{\hat{x}_M}^2]$

Γραμμική εκτίμηση ελαχίστου μέσου τετραγωνικού σφάλματος – ευθεία ελαχίστων τετραγώνων

Υπάρχουν περιπτώσεις όπου η εκτίμηση του $\hat{x}_M = E[X|Y=y]$ μέσω της δεσμευμένης $f_{X|Y}(x|y)$ (όπως είπαμε παραπάνω) να είναι δύσκολο ή αδύνατο να γίνει. Σε αυτή τη περίπτωση επιδιώκουμε να βρούμε μια γραμμική (Linear) εκτιμήτρια $\hat{X}_L = aY + b$ η οποία θα ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα (Mean Square Error, MSE):

$$MSE = E[(X - \hat{X}_L)^2] = E[(X - aY - b)^2]$$

Γραμμική εκτίμηση ελαχίστου μέσου τετραγωνικού σφάλματος – ευθεία ελαχίστων τετραγώνων

<u>Θεώρημα:</u> Η ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος $MSE = E[(X - \hat{X}_L)^2] = E[(X - aY - b)^2]$ οδηγεί στις σχέσεις:

$$a = A = \frac{Cov(X,Y)}{Var(Y)} \text{ και } b = B = EX - aEY = EX - \frac{Cov(X,Y)}{Var(Y)} EY$$

- Με ελάχιστο $MSE = E[(X - AY - B)^2] = (1 - \rho^2) Var(X),$ όπου $\rho = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$.
- $E[(X - AY - B)Y] = 0$

Γραμμική εκτίμηση ελαχίστου μέσου τετραγωνικού σφάλματος – ευθεία ελαχίστων τετραγώνων

Συνοπτικά έχουμε:

- $\hat{X}_L = \frac{Cov(X,Y)}{Var(Y)}(Y - EY) + EX = \frac{\rho\ \sigma_X}{\sigma_Y}(Y - EY) + EX$
- Για το σφάλμα (estimation error) $\tilde{X} = X - \hat{X}_L$έχουμε $E[\tilde{X}] = 0$ και $Cov(\tilde{X}, Y) = E[\tilde{X}Y] = 0$
- Το ελάχιστο μέσο τετραγωνικό σφάλμα είναι min$\{MSE\} = (1 - \rho^2)\sigma_X^2$

Ευθεία ελαχίστων τετραγώνων $y = a + b \cdot x$ για εφαρμογές στη Φυσική (με ελαχιστοποίηση του χ^2)

$$y(x) = y(x; a, b) = a + bx$$

$$\chi^2(a,b) = \sum_{i=1}^{N} \left(\frac{y_i - a - bx_i}{\sigma_i} \right)^2$$

$$0 = \frac{\partial \chi^2}{\partial a} = -2 \sum_{i=1}^{N} \frac{y_i - a - bx_i}{\sigma_i^2}$$

$$0 = \frac{\partial \chi^2}{\partial b} = -2 \sum_{i=1}^{N} \frac{x_i(y_i - a - bx_i)}{\sigma_i^2}$$

$$S \equiv \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \quad S_x \equiv \sum_{i=1}^{N} \frac{x_i}{\sigma_i^2} \quad S_y \equiv \sum_{i=1}^{N} \frac{y_i}{\sigma_i^2}$$

$$S_{xx} \equiv \sum_{i=1}^{N} \frac{x_i^2}{\sigma_i^2} \quad S_{xy} \equiv \sum_{i=1}^{N} \frac{x_i y_i}{\sigma_i^2}$$

$$aS + bS_x = S_y$$

$$aS_x + bS_{xx} = S_{xy}$$

$$\Delta \equiv S S_{xx} - (S_x)^2$$

$$a = \frac{S_{xx} S_y - S_x S_{xy}}{\Delta}$$

$$b = \frac{S S_{xy} - S_x S_y}{\Delta}$$

$$\sigma_f^2 = \sum_{i=1}^{N} \sigma_i^2 \left(\frac{\partial f}{\partial y_i} \right)^2$$

$$\mathrm{Cov}(a,b) = -S_x / \Delta$$

$$\frac{\partial a}{\partial y_i} = \frac{S_{xx} - S_x x_i}{\sigma_i^2 \Delta}$$

$$\frac{\partial b}{\partial y_i} = \frac{S x_i - S_x}{\sigma_i^2 \Delta}$$

$$\sigma_a^2 = S_{xx} / \Delta$$

$$\sigma_b^2 = S / \Delta$$

$$r_{ab} = \frac{-S_x}{\sqrt{S S_{xx}}}$$

Ευθεία ελαχίστων τετραγώνων $y=a+b\cdot x$ για εφαρμογές στη Φυσική (με ελαχιστοποίηση του χ^2)

$$y(x) = y(x;a,b) = a + bx$$

$$\chi^2(a,b) = \sum_{i=1}^{N} \left(\frac{y_i - a - bx_i}{\sigma_i} \right)^2$$

$$0 = \frac{\partial \chi^2}{\partial a} = -2 \sum_{i=1}^{N} \frac{y_i - a - bx_i}{\sigma_i^2}$$

$$0 = \frac{\partial \chi^2}{\partial b} = -2 \sum_{i=1}^{N} \frac{x_i(y_i - a - bx_i)}{\sigma_i^2}$$

$$S \equiv \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \qquad S_x \equiv \sum_{i=1}^{N} \frac{x_i}{\sigma_i^2} \qquad S_y \equiv \sum_{i=1}^{N} \frac{y_i}{\sigma_i^2}$$

$$aS + bS_x = S_y$$

$$aS_x + bS_{xx} = S_{xy}$$

$$\Delta \equiv SS_{xx} - (S_x)^2$$

$$a = \frac{S_{xx}S_y - S_xS_{xy}}{\Delta}$$

$$b = \frac{SS_{xy} - S_xS_y}{\Delta}$$

$$\boxed{\sigma_f^2 = \sum_{i=1}^{N} \sigma_i^2 \left(\frac{\partial f}{\partial y_i} \right)^2}$$

$$\mathrm{Cov}(a,b) = -S_x/\Delta$$

$$\frac{\partial a}{\partial y_i} = \frac{S_{xx} - S_x x_i}{\sigma_i^2 \Delta}$$

$$\sigma_a^2 = S_{xx}/\Delta$$

$$-S_x$$

ΣΗΜΑΝΤΙΚΟ: Τύπος διάδοσης των σφαλμάτων

Βιβλιογραφία

- Kontogiannis, I., & Toumpis, S. (2015). ELEMENTS OF PROBABILITY [Undergraduate textbook]. Kallipos, Open Academic Editions. http://hdl.handle.net/11419/2810
- Δαμιανού Χ., Χαραλαμπίδης Χ., Παπαδάτος Ν. (2010). Εισαγωγή στις Πιθανότητες και τη Στατιστική, ΕΚΔΟΣΕΙΣ ΣΥΜΜΕΤΡΙΑ Μ. ΑΘΑΝΑΣΟΠΟΥΛΟΥ Σ. ΑΘΑΝΑΣΟΠΟΥΛΟΣ Ο.Ε
- Κάκουλλου, Θ. (1969). Μαθήματα Θεωρίας Πιθανοτήτων, Εκδόσεις Παπαζήση, Αθήνα
- Strait, P. T. (1989). A First Course in Probability and Statistics With Applications, Harcourt Brace Jovanovich, Inc Orlando, FL, USA
- H. Pishro-Nik, "Introduction to probability, statistics, and random processes", available at https://www.probabilitycourse.com, Kappa Research LLC, 2014.
- Press, W. H., Flannery, B. P., Teukolsky, S. A.,, Vetterling, W. T. (1992). Numerical Recipes in FORTRAN 77: The Art of Scientific Computing. Cambridge University Press. Cambridge, MA, USA