26-4-2023

Control Variate

$X \sim F$: $E(X) = \theta$

Y : $E(Y) = \mu_Y$: γνωστή

$Cov(X,Y) \neq 0$.

$\tilde{X} = X + c(Y - \mu_Y)$, $c = \ldots$

$E(\tilde{X}) = E(X) = \theta$.

$Var(\tilde{X})$ min για $c^* = -\dfrac{Cov(X,Y)}{Var(Y)}$

και για $c = c^*$ $\dfrac{Var(\tilde{X})}{Var(X)} = 1 - \rho^2(X,Y)$

Αλγόριθμος

n επαναλήψεις

$$\begin{array}{cc} X_1 & Y_1 \\ \vdots & \\ X_n & Y_n \end{array}$$

Εκτιμούμε $\widehat{Cov}(X,Y)$, $\widehat{Var}(Y)$:

$$\hat{c}^* = -\frac{\widehat{Cov}(X,Y)}{\widehat{Var}(Y)}$$

$$\tilde{X}: \boxed{\tilde{X}_j = X_j + \hat{c}^* \cdot Y_j}, \quad j = 1, \dots, n$$

$$\downarrow$$

$$\hat{\theta} = \overline{\tilde{X}}$$

Εναλλακτικά Μοντέλο παλινδρόμησης $X = b_0 + b_1 Y + \varepsilon$

Εκτιμούμε τα $\hat{b}_0$, $\hat{b}_1$

$$\Longrightarrow \dots \Longrightarrow \tilde{\theta} = \hat{b}_0 + \hat{b}_1 \cdot \mu_Y$$

Πρόβλημα αξιοπιστίας

$$S = \varphi(S_1, \ldots, S_n)$$

$$\hat{\theta} = E(S) = P(S=1)$$

$$Y = \sum_{j=1}^{n} S_j \quad , \quad E(Y) = \mu_Y = \sum_{j=1}^{n} p_j$$

3η μέθοδος : Δεσμευμένη

Έστω τμ. X, Y (π.χ. συνεχείς)

$$E(X|Y=y) = \int_x x\, f_{X|Y}(x|y)\, dx = m(y)$$

$E(X|Y) = m(Y)$ τυχαία μεταβλητή : δεσμευμένη μέση τιμή της X δοθ. Y.

$E(X)$

$$\boxed{E(X) = E_Y\left[E(X|Y)\right]} = \int_y E(X|Y=y)\, f_Y(y)\, dy$$

$$\theta = E_X(X) = \int_y m(y)\, f_Y(y)\, dy = E_Y\left[m(Y)\right]$$

1. Δημιουργούμε n παρατ. από $X \Rightarrow \hat{\theta}_1 = \overline{X}_n$

2. " n παρ. από Y $\quad y_1, \dots, y_n$

$$Z_j = m(y_j), \quad j = 1, \dots, n$$

$$\hat{\theta}_2 = \overline{Z}_n$$

Πρέπει $\begin{cases} \exists \text{ γεννήτρια της } Y \sim F_Y \\ m(Y) \text{ γνωστή συνάρτηση.} \end{cases}$

Σύγκριση διασπορών : $Var_X(X)$, $Var_Y(E(X|Y))$
$= Var_Y(m(Y))$

$$\boxed{Var(X) = \underbrace{E_Y(Var(X|Y))}_{\geq 0} + Var_Y[E(X|Y)]}$$ ανάλυση διασποράς

$$\Rightarrow \boxed{Var_Y(E(X|Y)) \leq Var(X)}$$

Γραμμικό Μοντέλο

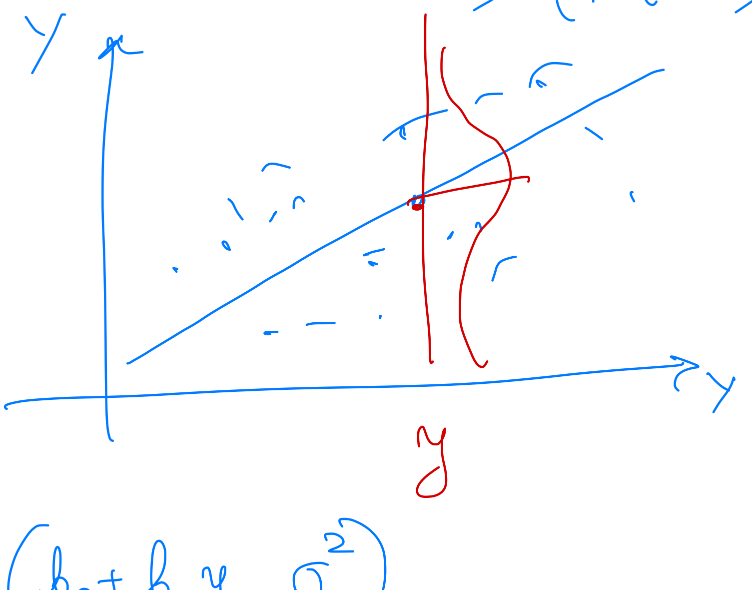
$SST = SSR + SSE$

$SSR \to Var(E(X|Y))$

$SSE \to E(Var(X|Y))$

$X = b_0 + b_1 Y + \varepsilon$

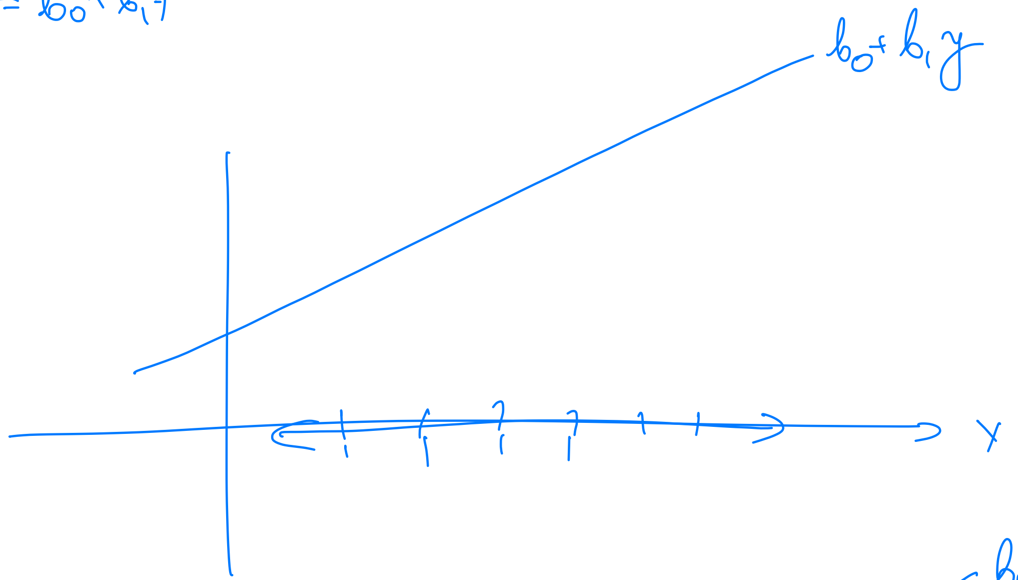
$\varepsilon \sim N(0, \sigma^2)$



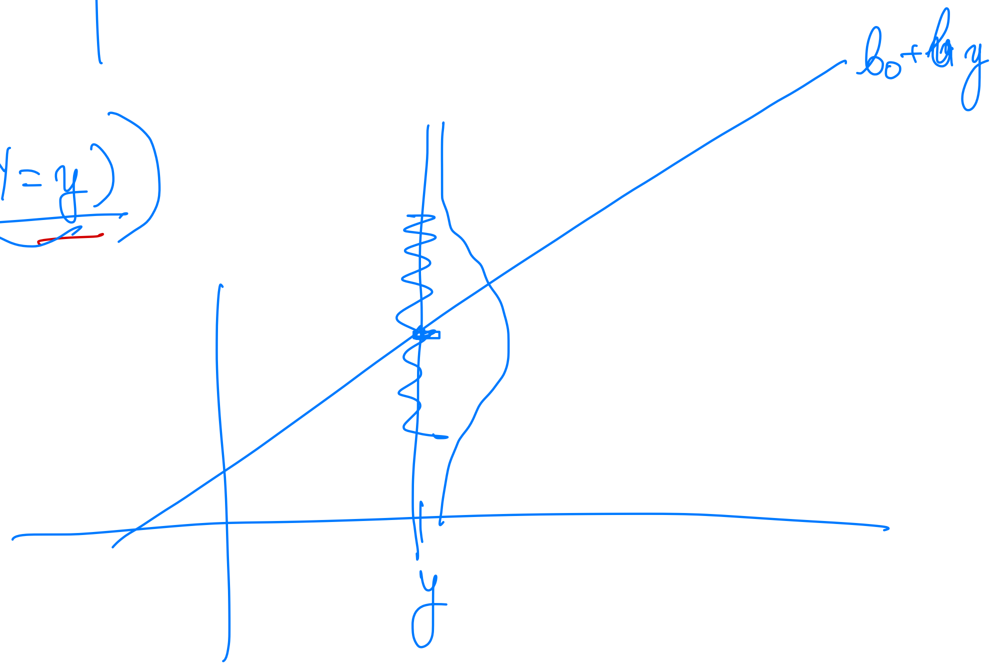
$\Rightarrow X|Y=y \sim N(b_0 + b_1 y, \sigma^2)$

$\underbrace{E(X|Y=y) = b_0 + b_1 y}_{m(y)}$ $\leftarrow$ συνάρτηση παλινδρόμησης

$$\mathrm{Var}\big(\underbrace{E(X|Y=y)}_{=b_0+b_1 y}\big) = \mathrm{Var}(b_0+b_1 Y)$$



$$\underbrace{E\big(\mathrm{Var}(X|Y=y)\big)}_{=\sigma^2}$$

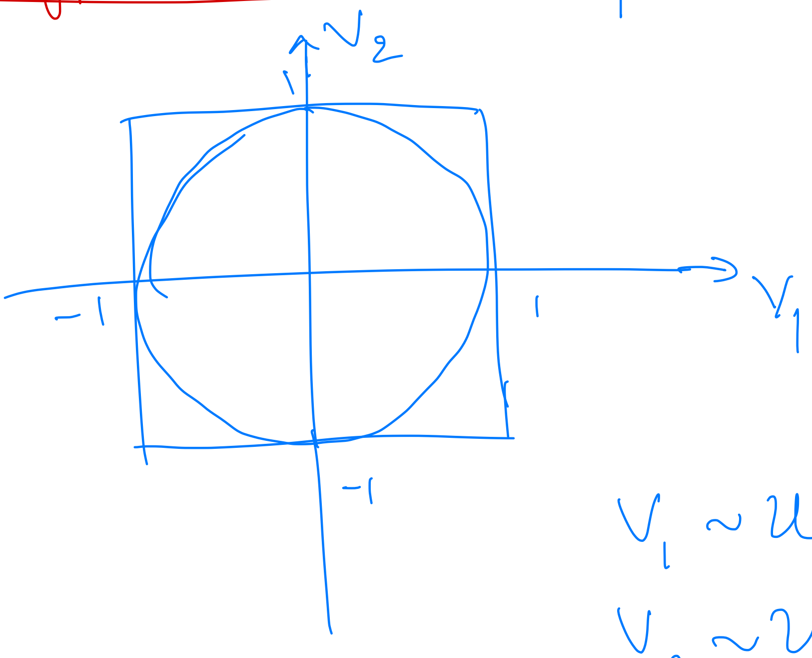


Στο γραμμικό μοντέλο $E(\mathrm{Var}(X|Y)) = \sigma^2$

$$\hat{\sigma}^2 = MSE$$

Παράδειγμα 1

Εκτίμηση π



$V_1 \sim \mathcal{U}(-1,1)$ ανεξ.

$V_2 \sim \mathcal{U}(-1,1)$

$$P(V_1^2+V_2^2 \le 1) = E[X], \quad X = \begin{cases} 1, & V_1^2+V_2^2 \le 1 \\ 0, & V_1^2+V_2^2 > 1 \end{cases}$$

$$E(X) = \frac{\pi}{4} := \theta$$

---

1 Δημιουργούμε η σειρά $(V_{1j}, V_{2j})_{j=1,\dots,n}$

$$\hat{\theta} = \frac{1}{n} \sum_{j=1}^{n} 1(V_{1j}^2 + V_{2j}^2 \le 1) = \frac{1}{n} \sum_{j=1}^{n} X_j$$

2 Έστω $Y = V_1$ $\qquad V_1 \in [-1,1]$

$$E(X \mid V_1 = v) = P(V_1^2 + V_2^2 \le 1 \mid V_1 = v)$$

$$= P(v^2 + V_2^2 \le 1 \mid V_1 = v)$$

$$= P\left(V_2^2 \le 1-v^2\right) =$$

$$= P\left(-\sqrt{1-v^2} \le V_2 \le \sqrt{1-v^2}\right) =$$

$$= \int_{-\sqrt{1-v^2}}^{\sqrt{1-v^2}} \frac{1}{2}\, du = \sqrt{1-v^2}$$

$$\Rightarrow \quad m(v) = E(X \mid Y = v) = \sqrt{1-v^2}$$

$$E\left(m(V_1)\right) = E\left(\sqrt{1-V_1^2}\right) = E(X) = \frac{\pi}{4}$$

$$Var\left(m(V_1)\right) < Var(X)$$

<u>Προσομοίωση</u>

Δημιουργούμε $\quad V_1 \sim \mathcal{U}(-1,1)$

$$\tilde{X} = \sqrt{1-V_1^2}$$

$$Var(\tilde{X}) = E(\tilde{X}^2) - \left(E(\tilde{X})\right)^2$$

$$= E\left(1-V_1^2\right) - \left(\frac{\pi}{4}\right)^2 = \frac{2}{3} - \left(\frac{\pi}{4}\right)^2$$

$$X = 1\left(V_1^2 + V_2^2 \le 1\right) \qquad E(X) = \frac{\pi}{4}$$

$$X \sim \mathrm{Bern}\left(\frac{\pi}{4}\right)$$

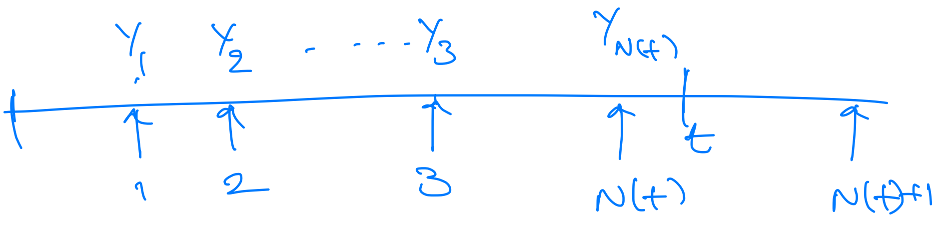
$$\mathrm{Var}(X) = \frac{\pi}{4}\left(1 - \frac{\pi}{4}\right) = \frac{\pi}{4} - \left(\frac{\pi}{4}\right)^2$$

$$\frac{2}{3} < \frac{\pi}{4} \Rightarrow \boxed{\mathrm{Var}(\tilde{X}) < \mathrm{Var}(X)}$$

Παράδειγμα 2 : Έστω $\{N(t), t \ge 0\}$ ανανεωτική διαδικασία.

ενδ. χρόνοι $T_1, T_2, \ldots, \sim F$.

$N(t) = \max\{n : X_1 + \cdots + X_n \le t\}$ = αρ. γεγονότων στο $[0, t]$.

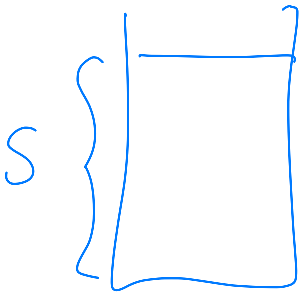


$Y_1, Y_2, \ldots, Y_n$: ανεξ. ισόνομες $\sim F_Y$
ανεξάρτητες της $\{N(t), t \ge 0\}$

π.χ. ① $\{N(t)\}$: ατυχήματα

Y_j = μέγεθος ζημιάς στο ατύχημα j

② Αποθέματα



απόθεμα

Ζήτηση: $\{N(t), t \geq 0\}$: αφίξεις ζήτησης

Y_j := μέγεθος j-οστής ζήτησης

Έστω $C(t) = \sum_{j=1}^{N(t)} Y_j$

$$\theta = E(X), \quad X = h(C(t))$$

π.χ. $h(x) = \mathbb{1}(x > m) \Rightarrow \theta = P(C(t) > m)$

---

① Monte Carlo: (t: fixed)

Προσομοιώνουμε $T_1, T_2, \ldots \Longrightarrow N(t) = n$

$j = 1, \ldots, n \Rightarrow$ προσομοιώνουμε $Y_1, \ldots, Y_n \Rightarrow \ldots$

$\Rightarrow C = Y_1 + \cdots + Y_n \Longrightarrow h(C) = X$

π.χ. (A) Αν $h(C) = C$

$$\theta = E(C) = \underbrace{E(Y)}_{\downarrow \; m_Y} \cdot E(N(t))$$

$$\underline{E(C \mid N(t) = n)} = n \, m_Y$$ γνωστή συνάρτηση.

(2) Προσομοιώνουμε $N(t) = n$

$$X = n \cdot m_Y$$

(B) Αν $h(C) = 1(C > m) = X$

$$E(X \mid N(t) = n) = P(C > m \mid N(t) = n)$$

$$= P(Y_1 + \dots + Y_n > m)$$? αναλυτική έκφραση.

π.χ. Αν $Y \sim Exp(v) \Rightarrow Y_1 + \dots + Y_n \sim Erlang(n, v)$

$$P(Y_1 + \dots + Y_n > m) = \int \dots$$

$$= 1 - \text{pgamma}$$