# Ex. 2 & 4

March 26, 2021

## EX. 2

Μας δίνεται η ακολουθία i.i.d. ζευγών τυχαίων μεταβλητών $(X_1, Y_1), ..., (X_n, Y_n), ...$ με αναμενόμενη τιμή $E(X, Y) = (\mu, \nu)$ και $\delta = Cov(X_i, Y_i)$ για κάθε $n \in \{2, 3, ...\}$

Ορίζουμε τον εκτιμητή της δ : $W_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)(Y_i - \nu)$

$$\sum_{i=1}^{n}(X_i - \mu)(Y_i - \nu) = \sum_{i=1}^{n} X_iY_i - \nu\sum_{i=1}^{n} X_i - \mu\sum_{i=1}^{n} Y_i + n\mu\nu$$

$$= \sum_{i=1}^{n} X_iY_i - 2n\mu\nu + n\mu\nu = \sum_{i=1}^{n} X_iY_i - n\mu\nu$$

Αρα για την αναμενόμενη τιμή έχουμε:

$$E(W_n) = \frac{1}{n}(\sum_{i=1}^{n} X_iY_i - n\mu\nu)$$

$$= \frac{1}{n}(\sum_{i=1}^{n} E(X_iY_i) - nE(X_i)E(Y_i))$$

$$= \frac{1}{n}(nE(X_iY_i) - nE(X_i)E(Y_i))$$

$$= Cov(X_i, Y_i) = \delta$$

Διότι $E(X_iY_i) = Cov(X_i, Y_i) + E(X_i)E(Y_i)$

-Συνεπώς ο $W_n$ είναι ***unbiased estimator*** της δ.

Επιπλέον ορίζουμε τον εκτιμητή της δ:

$$S_n = \frac{1}{n-1}\sum_{i=1}^{n}[X_i - (\frac{1}{n}\sum_{i=1}^{n} X_i)][Y_i - (\frac{1}{n}\sum_{i=1}^{n} Y_i)]$$

και θέτουμε $M(X) = \frac{1}{n}\sum_{i=1}^{n} X_i$ και $M(Y) = \frac{1}{n}\sum_{i=1}^{n} Y_i$
Παίρνουμε το άθροισμα:

$$\sum_{i=1}^{n}(X_i - M(X))(Y_i - M(Y)) = \sum_{i=1}^{n} X_i Y_i - M(Y)\sum_{i=1}^{n} X_i - M(X)\sum_{i=1}^{n} Y_i + nM(X)M(Y)$$

$$= \sum_{i=1}^{n} X_i Y_i - 2nM(X)M(Y) + nM(X)M(Y) = \sum_{i=1}^{n} X_i Y_i - nM(X)M(Y)$$

Η αναμενόμενη τιμή του $S_n$ είναι:

$$E(S_n) = \frac{1}{n-1}E(\sum_{i=1}^{n} X_i Y_i - nM(X)M(Y))$$

$$= \frac{1}{n-1}(\sum_{i=1}^{n} E(X_i Y_i) - nE(M(X)M(Y))) = \frac{n}{n-1}[E(X_i Y_i) - E(M(X)M(Y))]$$

Ισχύει οτι:

$$E(X_i Y_i) = Cov(X_i, Y_i) + E(X_i)E(Y_i) = \delta + \mu\nu$$
$$E(M(X)M(Y)) = Cov(M(X), M(Y)) + E(M(X))E(M(Y)) = \frac{\delta}{n} + \mu\nu$$

Επομένως:

$$E(S_n) = \frac{1}{n-1}(n\delta + n\mu\nu - n\frac{\delta}{n} + n\mu\nu) = \frac{n-1}{n-1}\delta = \delta$$

-Συνεπώς ο $S_n$ είναι ***unbiased estimator*** της δ.

**b)** Μελτάμε την συνέπεια των παραπάνω ετκτιμητών.
Για τον $W_n$ αρκεί να δείξουμε ότι:

$$Var(W_n) = \frac{1}{n}(\delta_2 - \delta^2) \text{ for every } n \in N_+$$

όπου:

$$Cov(X_i, Y_i) = \delta$$
$$E[(X-\mu)^2(Y-\nu)^2] = \delta_2$$

$$
\begin{aligned}
Var((X-\mu)(Y-\nu) &= E[(X-\mu)^2(Y-\nu)^2] - (E[(X-\mu)(Y-\nu)])^2 \\
&= \delta_2 - \delta^2
\end{aligned}
$$

Αρα για το $W_n$ έχουμε:
$Var(W_n) = \frac{1}{n^2}\sum_{i=1}^{n} Var((X-\mu)(Y-\nu)) = \frac{n}{n^2}(\delta_2-\delta^2) = \frac{1}{n}(\delta_2-\delta^2)$

-Δηλαδή η ακολουθία του $W_n$ είναι **συνεπής**.

Αντιστοιχα για το $S_n$ αρκεί να δείξουμε οτι :
$Var(S_n) = \frac{1}{n}(\delta_2 + \frac{\sigma^2\tau^2}{n-1} - \frac{n-2}{n-1}\delta^2)$ for every $n \in N_+$

Το $S_n$ μπορεί να γραφεί και ως:

$$S_n = \frac{1}{2n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}(X_i - X_j)(Y_i - Y_j)$$

Άρα:

$$Var(S_n) = \frac{1}{4n^2(n-1)^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n}(X_i - X_j)(Y_i - Y_j)(X_k - X_l)(Y_k - Y_l)$$

Για τις διαφορες τιμές των $i, j, k, l$ έχουμε:
1)$Cov((X_i - X_j)(Y_i - Y_j)(X_k - X_l)(Y_k - Y_l)) = 0$ για $i = j$και $k = l$.
Υπάρχουν $n^2(2n-1)$ τέτοιες περιπτώσεις.

2)$Cov((X_i - X_j)(Y_i - Y_j)(X_k - X_l)(Y_k - Y_l)) = 0$ για $i, j, k, l$ διαφορετικά μεταξύ τους.
Υπάρχουν $n(n-1)(n-2)(n-3)$ τέτοιες περιπτώσεις.

3)$Cov((X_i - X_j)(Y_i - Y_j)(X_k - X_l)(Y_k - Y_l)) = 2\delta_2 + 2\sigma^2\tau^2$ για $i \neq j$ και $\{k, l\} = \{i, j\}$
Υπάρχουν $2n(n-1)$ τέτοιες περιπτώσεις.

4)$Cov((X_i - X_j)(Y_i - Y_j)(X_k - X_l)(Y_k - Y_l)) = \delta_2 - \delta^2$ για $i \neq j, k \neq l$ και $|\{k, l\} \cap \{i, j\}| = 1$
Υπάρχουν $4n(n-1)(n-2)$ τέτοιες περιπτώσεις.

Άρα αντικαθιστώντας καταλήγουμε ότι:

$$
\begin{aligned}
Var(S_n) &= \frac{1}{4n^2(n-1)^2}(2n(n-1)(2\delta_2 + 2\sigma^2\tau^2) + 4n(n-1)(n-2)(2\delta_2 - \delta^2)) \\
&= \frac{1}{n(n-1)}(\delta_2 + \sigma^2\tau^2) + \frac{1}{n}\frac{n-2}{n-1}(\delta_2 - \delta^2) \\
&= \frac{1}{n}(\delta_2 + \frac{\sigma^2\tau^2}{n-1} - \frac{n-2}{n-1}\delta^2)
\end{aligned}
$$

**c)** Μελετάμε την αποδοτικότητα των $W_n$ και $S_n$ παίρνωντας την διαφορά τους:

$$
\begin{aligned}
Var(S_n) - Var(W_n) &= \frac{1}{n}(\delta_2 + \frac{\sigma^2\tau^2}{n-1} - \frac{n-2}{n-1}\delta^2 - \delta_2 + \delta^2) \\
&= \frac{1}{n(n-1)}(\sigma^2\tau^2 + \delta^2) > 0
\end{aligned}
$$

Για κάθε n ισχύει $Var(S_n) > Var(W_n)$ και συνεπώς ο $W_n$ είναι αποδοτικότερος εκτιμητής της δ.

**d)** Για την ασυμπτωτική σχετική απόδοση των $W_n, S_n$ έχουμε:

$$
\lim_{n\to\infty} Eff(W_n, S_n) = \lim_{n\to\infty} \frac{Var(S_n)}{Var(W_n)} = \lim_{n\to\infty}(\frac{\delta_2 + \frac{\sigma^2\tau^2}{n-1} - \frac{n-2}{n-1}\delta^2}{\delta_2 - \delta^2}) = \frac{\delta_2 - \delta^2}{\delta_2 - \delta^2} = 1
$$

## EX. 4

Έστω ακολουθία i.i.d r.v.s $X_1, ...X_n, ...$ με ομοιόμορφη συνεχή κατανομή στο $(0, \theta)$
Προτείνονται οι εξής εκτιμητές για την παράμετρο ϑ :

$$\theta_n = \frac{2}{n}\sum_{i=1}^{n} X_i \quad and \quad \zeta_n = (n+1)min\{X_1, ...X_n\}$$

Για τις τυχαίες μεταβλητές έχουμε την συνάρτηση πυκνότητας πιθανότητας :

$$f_x(x; \theta) = \frac{1}{\theta} 1(x \in (0, \theta))$$

και

$$E(X_i) = \int_{-\infty}^{+\infty} x f_x(x; \theta) dx = \int_0^{\theta} \frac{x}{\theta} dx = \frac{\theta^2}{2\theta} = \frac{\theta}{2}$$

Αρχικά ϑα μελετήσουμε τον εκτιμητή $\theta_n$ :

$$E(\theta_n) = E(\frac{2}{n}\sum_{i=1}^{n} X_i) = \frac{2}{n}\sum_{i=1}^{n} E(X_i) = \frac{2}{n} n \frac{\theta}{2} = \theta$$

Αρα ο $\theta_n$ είναι **unbiased** εκτιμητής της ϑ Μελετάμε και την συνέπεια του $\theta_n$

$$E((\theta_n - \theta)^2) = Var(\theta_n) = Var(\frac{2}{n}\sum_{i=1}^{n} X_i) = \frac{4}{n^2}\sum_{i=1}^{n} Var(X_i)$$
$$= \frac{4}{n^2} n \frac{1}{12}\theta^2 = \frac{\theta^2}{3n}$$

και

$$\lim_{n\to\infty}(E((\theta_n - \theta)^2)) = \lim_{n\to\infty} \frac{\theta^2}{3n} = 0$$

Συνεπώς η ακολουθία του $\theta_n$ είναι μεσοτετραγωνικά **συνεπής**.

Για τον εκτιμητή $\zeta_n = (n+1)min\{X_1, ...X_n\}$

Έστω η $Y = min\{X_1, ...X_n\}$
Παίρνουμε απο την κατανομή του Y την συνάρτηση πυκνότητας πιθανότητας.
$F(y) = P(Y \leq y) = 1 - P(Y > y) = 1 - P(min\{X_1, ...X_n\} > y)$

Όμως έχουμε $min\{X_1, ...X_n\} > y$ μόνο όταν $X_i > y \ \forall i$
και εφόσον έχουμε i.i.d. rvs:

$$F(y) = 1 - P(X_1 > y)P(X_2 > y)...P(X_n > y) = 1 - P(X_i > y)^n$$

Τα $X_i$ έχουν uniform distribution στο $(0, \theta)$ και συνεπώς:

$$F(y) = \begin{cases} 1 - \left(\dfrac{\theta - y}{\theta}\right)^n & y \in (0, \theta) \\ 0 & y < 0 \\ 1 & y > \theta \end{cases}$$

και παραγωγίζοντας:

$$f(y) = \begin{cases} \dfrac{n}{\theta}\left(\dfrac{\theta - y}{\theta}\right)^{n-1} & y \in (0, \theta) \\ 0 & otherwise \end{cases}$$

Τώρα μπορούμε να παρουμε την αναμενόμενη τιμη της Y

$$E(Y) = \int_{-\infty}^{\infty} yf(y)dy = \int_0^{\theta} \frac{ny}{\theta}\left(\frac{\theta - y}{\theta}\right)^{n-1} dy = \frac{\theta}{n+1}$$

Αρα ο εκτιμητής μας είναι **unbiased** αφού έχει:

$$E(\zeta_n) = E((n+1)Y) = (n+1)E(Y) = (n+1)\frac{\theta}{n+1} = \theta$$

Μελετάμε και την συνέπεια του $\zeta_n$ . Έτσι παίρνουμε:

$$E((\zeta_n - \theta)^2) = Var(\zeta_n) = Var((n+1)Y) = (n+1)^2 Var(Y)$$

Υπολογίζω την Var(Y) απο την CDF της και έχω:

$$Var(Y) = 2\int_0^{\infty} y(1 - F(y))dy - \left(\int_0^{\infty}(1 - F(y))dy\right)^2$$

$$I_1 = \int_0^{\infty} y(1 - F(y))dy = \int_0^{\theta} y\left(\frac{\theta - y}{\theta}\right)^n dy = \frac{\theta^2}{n^2 + 3n + 2}$$

$$I_2 = \int_0^{\infty}(1 - F(y))dy = \int_0^{\theta}\left(\frac{\theta - y}{\theta}\right)^n dy = 0 + \frac{\theta}{n+1}$$

Άρα έχουμε:

$$\begin{aligned} Var(\zeta_n) &= (n+1)^2 Var(Y) = (n+1)^2(2I_1 - I_2^2) \\ &= (n+1)^2 \left( \frac{2\theta^2}{n^2+3n+2} - \frac{\theta^2}{(n+1)^2} \right) \\ &= \frac{n+1}{n+2} 2\theta^2 - \theta^2 \\ &= \frac{n}{n+1} \theta^2 \end{aligned}$$

$$\lim_{n\to\infty} Var(\zeta_n) = \lim_{n\to\infty} \frac{n}{n+1}\theta^2 = \theta^2$$

Συνεπώς η ακολουθία του $\zeta_n$ δεν είναι **συνεπής**.
Τέλος οσο αναφρά την αποδοτικότητα των 2 εκτιμητών έχουμε:

$$Var(\zeta_n) - Var(\theta_n) = \theta^2 \left( \frac{n}{n+1} - \frac{1}{3n} \right) = \theta^2 \left( \frac{3n^2-n-1}{3n(n+1)} \right) > 0 \;\; for \; n > 1/6(1+\sqrt{13})$$

$$O\mu\omega\varsigma \; n \in N_+ \; A\rho\alpha \; \iota\sigma\chi\upsilon\epsilon\iota \; \gamma\iota\alpha \; \kappa\theta\varepsilon \; n > 0$$

Συνεπώς ο $\theta_n$ είναι **αποδοτικότερος** εκτιμητής της παραμέτρου $\theta$ απο την $\zeta_n$