Άσκηση 1: Έστω $X \sim B(a,\beta)$, με συνάρτηση πυκνότητας

$$f(x;a,\beta) = \frac{1}{B(a,\beta)} \cdot x^{a-1} \cdot (1-x)^{\beta-1} \quad , \quad a>0,\ \beta>0\ ,\ S_x=[0,1],$$

όπου $B(a,\beta) = \int_0^1 u^{a-1}(1-u)^{\beta-1}\, du$, η σταθερά ολοκλήρωσης.

Έστω $Y \sim \Gamma(a,\beta)$, με συνάρτηση πυκνότητας

$$f(y;a,\beta) = \frac{\beta^a}{\Gamma(a)} \cdot y^{a-1} \cdot e^{-\beta y}, \quad a>0,\ \beta>0,\quad S_y=[0,+\infty)$$

όπου $\Gamma(a) = \int_0^{+\infty} u^{a-1} \cdot e^{-u}\, du$, η σταθερά ολοκλήρωσης.

Να αποδειχθεί ότι οι σταθερές ολοκλήρωσης συνδέονται με τη σχέση:

$$B(a,\beta) = \frac{\Gamma(a)\cdot\Gamma(\beta)}{\Gamma(a+\beta)}.$$

Απόδειξη:

$$\Gamma(a)\cdot\Gamma(\beta) = \int_0^{+\infty} u^{a-1}\cdot e^{-u}\,du \cdot \int_0^{+\infty} v^{\beta-1}\cdot e^{-v}\,dv = \int_{v=0}^{+\infty}\int_{u=0}^{+\infty} u^{a-1}\cdot v^{\beta-1}\cdot e^{-(u+v)}\,du\,dv$$

Κάνω αλλαγή μεταβλητής: $\left\{\begin{matrix} x := u+v \\ y := \frac{u}{u+v} \end{matrix}\right\} \iff \left\{\begin{matrix} u = xy \\ v = x(1-y) \end{matrix}\right\}$

Ορίζω $\varphi: \overset{x}{(0,+\infty)} \times \overset{y}{(0,1)} \to \mathbb{R}^2$, $\varphi(x,y) = (\overset{u}{xy}, \overset{v}{x(1-y)})$,

$$\det(J_\varphi) = \det\begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix} = \det\begin{bmatrix} y & x \\ 1-y & -x \end{bmatrix} = -xy - x(1-y) = -x$$

$$\Gamma(a)\cdot\Gamma(\beta) \overset{F-T}{=} \iint_{(u,v)\in(0,+\infty)\times(0,+\infty)} u^{a-1} v^{\beta-1} e^{-(u+v)}\,du\,dv = \iint_{(x,y)\in(0,+\infty)\times(0,1)} (xy)^{a-1}\,[x(1-y)]^{\beta-1}\, e^{-x} \cdot \overset{|\det(J_\varphi)|}{|(-x)|}\,dx\,dy$$

$$\overset{F-T}{\Longrightarrow} \int_{x=0}^{+\infty} \int_{y=0}^{1} (xy)^{a-1} \, [x(1-y)]^{\beta-1} \, e^{-x} \cdot x \; dy \; dx =$$

$$= \int_{x=0}^{+\infty} \int_{y=0}^{1} x^{a+\beta-1} \, y^{a-1} \cdot (1-y)^{\beta-1} \cdot e^{-x} \; dy \; dx =$$

$$= \int_{x=0}^{+\infty} x^{a+\beta-1} \cdot e^{-x} \, dx \cdot \int_{y=0}^{1} y^{a-1} (1-y)^{\beta-1} \, dy = \Gamma(a+\beta) \cdot B(a,\beta).$$

Τελικά,

$$\Gamma(a)\Gamma(\beta) = \Gamma(a+\beta) \cdot B(a,\beta) \iff B(a,\beta) = \frac{\Gamma(a) \cdot \Gamma(\beta)}{\Gamma(a+\beta)}.$$

Άσκηση 2: Να αποδείξετε ότι για την συνάρτηση Γάμμα,

$$\Gamma(a) = \int_0^{+\infty} x^{a-1} e^{-x}\,dx, \quad a>0, \quad \text{ισχύει η σχέση:}$$

$$\Gamma(a+1) = a \cdot \Gamma(a), \qquad a>0.$$

Απόδειξη:

$$\Gamma(a+1) = \int_0^{+\infty} \underline{x^a e^{-x}}\,dx = \int_0^{+\infty} a x^{a-1} e^{-x}\,dx - \int_0^{+\infty} (x^a e^{-x})'\,dx =$$

$$\hookrightarrow (x^a e^{-x})' = a x^{a-1} e^{-x} - x^a e^{-x}$$

$$x^a e^{-x} = a x^{a-1} e^{-x} - (x^a e^{-x})'$$

$$= a \int_0^{+\infty} x^{a-1} e^{-x}\,dx - \left[x^a e^{-x} \right]_0^{+\infty} = a \cdot \Gamma(a) - 0 = a \cdot \Gamma(a)$$

Άσκηση 3: Έστω $X \sim B(a,\theta)$, $a,\theta>0$. Να αποδείξετε ότι η ροπή κ-τάξης δίνεται από τον τύπο:

$$E[X^k] = \frac{(a+k-1)\cdot \ldots \cdot a}{(a+\theta+k-1)\cdot \ldots \cdot (a+\theta)}, \qquad k=1,2,\ldots$$

Απόδειξη:

$$E[X^k] = \int_0^1 x^k \cdot \frac{1}{B(a,\theta)} \cdot x^{a-1}(1-x)^{\theta-1}\,dx = \frac{1}{B(a,\theta)} \int_0^1 x^{(a+k)-1}(1-x)^{\theta-1}\,dx =$$

$$= \frac{B(a+k,\theta)}{B(a,\theta)} = \frac{\dfrac{\Gamma(a+k)\cdot\Gamma(\theta)}{\Gamma(a+\theta+k)}}{\dfrac{\Gamma(a)\cdot\Gamma(\theta)}{\Gamma(a+\theta)}} = \frac{\Gamma(a+k)}{\Gamma(a)} \cdot \frac{\Gamma(a+\theta)}{\Gamma(a+\theta+k)}$$

$$= \frac{(a+k-1)\cdot \ldots \cdot a\,\cancel{\Gamma(a)}}{\cancel{\Gamma(a)}} \cdot \frac{\cancel{\Gamma(a+\theta)}}{(a+\theta+k-1)\cdot \ldots \cdot (a+\theta)\,\cancel{\Gamma(a+\theta)}} = \frac{(a+k-1)\cdot \ldots \cdot a}{(a+\theta+k-1)\cdot \ldots \cdot (a+\theta)}.$$

Άσκηση 4: Έστω $X \sim \mathcal{B}(a,\beta)$, $a,\beta > 0$. Να υπολογιστούν οι ποσότητες $E(X)$, $E(X^2)$, $Var(X)$.

Λύση: Από τον τύπο (Άσκηση 3):

$$E[X^k] = \frac{(a+k-1)\cdot \ldots \cdot a}{(a+\beta+k-1)\cdot \ldots \cdot (a+\beta)},$$

προκύπτει άμεσα ότι:

$$E[X] = \frac{a}{a+\beta}, \qquad E[X^2] = \frac{(a+1)\cdot a}{(a+\beta+1)(a+\beta)},$$

$$Var(X) = E[X^2] - E[X]^2 = \frac{a(a+1)}{(a+\beta)(a+\beta+1)} - \frac{a^2}{(a+\beta)^2} = \ldots = \frac{a\beta}{(a+\beta)^2(a+\beta+1)}. \blacksquare$$

Άσκηση 5: Έστω $\underset{\sim}{X} = (X_1, X_2, \dots, X_n)$ τυχαίο δείγμα από την κατανομή $B(a,\theta)$, $a,\theta > 0$ άγνωστα. Να βρεθούν οι εκτιμήτριες ροπών των a και θ.

Λύση:

Θα λύσω το σύστημα των εξισώσεων:

$$\left\{ \begin{array}{l} \overline{X} = E[X] \\ \overline{X^2} = E[X^2] \end{array} \right\}, \text{ ή ισοδύναμα το } \left\{ \begin{array}{l} \overline{X} = E[X] \\ M_2 = Var(X) \end{array} \right\}.$$

$$\left\{ \begin{array}{l} \overline{X} = \dfrac{a}{a+\theta} \; (1) \implies (a+\theta)\overline{X} = a \; (2) \overset{??}{\implies} \theta = \left(\dfrac{1}{\overline{X}} - 1\right) a \implies \theta = \dfrac{1-\overline{X}}{\overline{X}} a \; (3) \\ \overline{X^2} = \dfrac{a(a+1)}{(a+\theta)(a+\theta+1)} \overset{(1)}{\implies} \overline{X^2} = \overline{X} \cdot \dfrac{a+1}{a+\theta+1} \implies \overline{X^2} = \overline{X}^2 \cdot \dfrac{a+1}{(a+\theta)\overline{X} + \overline{X}} \overset{(2)}{\implies} \end{array} \right.$$

$$\implies \overline{X^2} = \overline{X}^2 \frac{a+1}{a+\overline{X}} \implies (a+\overline{X})\overline{X^2} = (a+1)\overline{X}^2 \implies$$

$$\implies a\left(\overline{X^2} - \overline{X}^2\right) = \overline{X}^2 - \overline{X} \cdot \overline{X^2} \overset{??}{\implies} \boxed{\tilde{a} = \frac{\overline{X}\left(\overline{X} - \overline{X^2}\right)}{\overline{X^2} - \overline{X}^2}} \; (4).$$

Από (3),(4): $\boxed{\tilde{\theta} = \dfrac{(1-\overline{X})\left(\overline{X} - \overline{X^2}\right)}{\overline{X^2} - \overline{X}^2}}$.

Παρατηρούμε ότι ο παρονομαστής είναι η M_2, αφού

$$M_2 := \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X})^2 = \frac{1}{n}\Big[\sum_{i=1}^n X_i^2 - 2\overline{X}\underbrace{\sum_{i=1}^n X_i}_{=n\overline{X}} + n\overline{X}^2\Big] = \underbrace{\frac{1}{n}\sum X_i^2}_{\overline{X^2}} - \frac{1}{n}2n\overline{X}^2 + \frac{n}{n}\overline{X}^2 = \overline{X^2} - \overline{X}^2.$$

?? Προσοχή: Αν $X_1 = X_2 = \dots = X_n = 0$, τότε $\overline{X} = \overline{X^2} = 0$, άρα $\tilde{a} = 0$ και δεν μπορούμε να εκτιμήσουμε την παράμετρο θ. Αν $X_1 = X_2 = \dots = X_n = 1$, τότε $\overline{X} = \overline{X^2} = 1$, άρα $\tilde{\theta} = 0$ και δεν μπορούμε να εκτιμήσουμε την a. Τα δύο αυτά ενδεχόμενα έχουν πιθανότητα 0.

Άσκηση 6: Έστω $X \sim Unif(a,b)$, $a<b$, $a,b \in \mathbb{R}$. Να βρεθεί η κατανομή της $Y := \frac{X-a}{b-a}$. Ως ειδική περίπτωση, να βρείτε κατάλληλες συναρτήσεις $g: S_X \to S_Y$ τέτοιες ώστε:

i) Αν $X \sim Unif(-\theta,\theta)$ τότε $Y := g(X) \sim Unif(0,1)$

ii) Αν $X \sim Unif(0,\theta)$ τότε $Y := g(X) \sim Unif(0,1)$

iii) Αν $X \sim Unif(\theta,\theta+1)$ τότε $Y := g(X) \sim Unif(0,1)$.

Παρατηρήστε ότι η g είναι $g(x;a,b)$, δηλαδή δεν είναι στατιστική συνάρτηση.

Λύση:

Αν $X \sim Unif(a,b)$ τότε $F_X(x) = P(X \le x) = \begin{cases} 0 & x<a \\ \frac{x-a}{b-a}, & a \le x < b \\ 1 & b \le x \end{cases}$

και $f_X(x) = \begin{cases} 1/_{b-a}, & x \in (a,b) \\ 0, & x \notin (a,b) \end{cases}$

Επομένως:

$$F_Y(x) = P(Y \le x) = P\left(\frac{X-a}{b-a} \le x\right) = P\left(X \le a+(b-a)x\right) = \begin{cases} 0 & a+(b-a)x \le a \\ x & a \le a+(b-a)x < b \\ 1 & b \le a+(b-a)x \end{cases}$$

ή ισοδύναμα $F_Y(x) = \begin{cases} 0 & x < 0 \\ x & 0 \le x < 1 \\ 1 & 1 \le x \end{cases}$, άρα $Y \sim Unif(0,1)$.

Ειδικές περιπτώσεις:

i) $g(x) = \frac{x+\theta}{2\theta}$, $Y := \frac{X+\theta}{2\theta} \sim Unif(0,1)$

ii) $g(x) = \frac{x}{\theta}$, $Y := X/\theta \sim Unif(0,1)$

iii) $g(x) = x-\theta$, $Y := X-\theta \sim Unif(0,1)$

Άσκηση 7: Έστω $X \sim Unif(-\theta,\theta)$, $\theta>0$. Να βρεθεί η κατανομή της $Y := |X|$.

Λύση:

Αν $X \sim Unif(-\theta,\theta)$ τότε $f_X(x) = \frac{1}{2\theta} 1_{(-\theta,\theta)}(x)$. Άρα

$$F_Y(x) = P(Y \le x) = P(|X| \le x)$$

- Αν $x<0$, $F_Y(x) = 0$.
- Αν $0 \le x < \theta$, τότε

$$F_Y(x) = P(-x \le X \le x) = \int_{-x}^{x} \frac{1}{2\theta} 1_{(-\theta,\theta)}(u)\, du = \int_{-x}^{x} \frac{1}{2\theta}\, du = \left[\frac{u}{2\theta}\right]_{-x}^{x} = \frac{x}{\theta}$$

- Αν $\theta \le x$, τότε

$$F_Y(x) = P(-x \le X \le x) = \int_{-x}^{x} \frac{1}{2\theta} 1_{(-\theta,\theta)}(u)\, du = \int_{-\theta}^{\theta} \frac{1}{2\theta}\, du = \left[\frac{u}{2\theta}\right]_{-\theta}^{\theta} = 1.$$

Άρα $F_Y(x) = \begin{cases} 0 & x<0 \\ x/\theta & 0 \le x < \theta \\ 1 & \theta \le x \end{cases}$, άρα $Y \sim Unif(0,\theta)$.

$\blacksquare$

Άσκηση 8: Έστω $\underset{\sim}{X} = (X_1, X_2, \dots, X_n)$ τυχαίο δείγμα από την κατανομή $Unif(-\theta, \theta)$. Να βρεθεί μία εκτιμήτρια ροπών του θ.

Λύση:

Α' τρόπος: Λύνω την εξίσωση $\overline{X_n} = E[X] \Rightarrow \overline{X_n} = 0$ και παρατηρώ ότι η $E[X]$ δεν είναι συνάρτηση της παραμέτρου. Προχωρώ στην

$$M_2 = Var(X) \Rightarrow M_2 = \frac{(b-a)^2}{12} \Rightarrow M_2 = \frac{(2\theta)^2}{12} \Rightarrow \boxed{\tilde{\theta} = \sqrt{3M_2}}.$$

Β' τρόπος: Κάνω τον μετασχηματισμό $Y_i = |X_i|$, $i = 1, 2, \dots, n$.

Λύνω την εξίσωση $\overline{Y_n} = E[Y] \Rightarrow \overline{Y_n} = \frac{\theta}{2} \Rightarrow \boxed{\theta^* = 2\overline{Y_n}}.$

Άσκηση 9: Έστω $X_1, \dots, X_n$ τυχαίες μεταβλητές. Να αποδείξετε ότι:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} (X_i - X_j)^2 = 2n \sum_{i=1}^{n} (X_i - \overline{X_n})^2 = 2n^2\left(\overline{X_n^2} - \overline{X_n}^2\right).$$

Απόδειξη:

$$\sum_{i=1}^{n}\sum_{j=1}^{n} (X_i - X_j)^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} (X_i^2 - 2X_iX_j + X_j^2) = n^2\overline{X_n^2} - 2n^2\overline{X_n}^2 + n^2\overline{X_n^2} = 2n^2\left(\overline{X_n^2} - \overline{X_n}^2\right).$$

$$2n\sum_{i=1}^{n} (X_i - \overline{X_n})^2 = 2n\sum_{i=1}^{n} (X_i^2 - 2X_i\overline{X_n} + \overline{X_n}^2) = 2n\left(n\overline{X_n^2} - 2n\overline{X_n}^2 + \overline{X_n}^2\right) = 2n^2\left(\overline{X_n^2} - \overline{X_n}^2\right).$$

$\blacksquare$

Άσκηση 10: Έστω $\underline{X} = (X_1, X_2, \dots, X_n)$ τυχαίο δείγμα από μία κατανομή.
Να υπολογιστεί η μέση τιμή και η διασπορά των M_2 και S^2.
Δίνεται ότι $E[X^4] < +\infty$.

Λύση:

$$E[M_2] = \frac{n-1}{n}\sigma^2, \quad E[S^2] = \sigma^2$$

Για την διασπορά, ισχύει ότι $S^2 := \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X_n})^2 = \frac{1}{2n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}(X_i - X_j)^2$,

άρα:

$$Var(S^2) = E\left[\left(S^2 - E(S^2)\right)^2\right] = E\left[\left(\frac{1}{2n(n-1)}\sum_{i=1}\sum_{j=1}(X_i - X_j)^2 - \sigma^2\right)^2\right] =$$

$$= E\left[\left(\frac{1}{2n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[(X_i - X_j)^2 - 2\sigma^2\right]\right)^2\right] = \frac{1}{4n^2(n-1)^2}E\left[\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\left[(X_i - X_j)^2 - 2\sigma^2\right]\right)^2\right] =$$

$$= \frac{1}{4n^2(n-1)^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} E\left[\left[(X_i - X_j)^2 - 2\sigma^2\right]\left[(X_k - X_l)^2 - 2\sigma^2\right]\right].$$

Οι n^4 όροι του παραπάνω αθροίσματος ανήκουν σε ακριβώς μία από τις παρακάτω περιπτώσεις:

1) $i=j$ ή $k=l$. Υπάρχουν $2n^3-n^2$ τέτοιοι όροι, οι οποίοι δίνουν $E[\dots]=0$. Τα παρακάτω αφορούν $i\neq j$ <u>και</u> $k\neq l$.

2) $|\{i,j\}\cap\{k,l\}|=0$. Τότε i,j,k,l ξένα ανά δύο. Υπάρχουν $n(n-1)(n-2)(n-3)$ τέτοιοι όροι και δίνουν:

$$E\Big[\big[(X_i-X_j)^2-2\quad\sigma^2\big]\big[(X_k-X_l)^2-2\quad\sigma^2\big]\Big]=E\big[(X_i-X_j)^2-2\quad\sigma^2\big]^2=$$

$$=\Big\{E[(X_i-X_j)^2]-2\quad\sigma^2\Big\}^2=\big(2\sigma^2-2\quad\sigma^2\big)^2=0.$$

3) $|\{i,j\}\cap\{k,l\}|=1$. Υπάρχουν $4\cdot n(n-1)(n-2)$ τέτοιοι όροι, και χωρίς βλάβη της γενικότητας υποθέτουμε $i=k$, $j\neq l$. Τότε:

$$E\Big[\big[(X_i-X_j)^2-\quad 2\sigma^2\big]\big[(X_i-X_l)^2-\quad 2\sigma^2\big]\Big]=$$

$$=E\big[(X_i-X_j)^2(X_i-X_l)^2\big]-\quad 4\sigma^2E\big[(X_i-X_j)^2\big]+\quad 4\sigma^4=$$

$$\overset{(*)}{=}\mu_4+3\sigma^4-8\sigma^4+4\sigma^4=\mu_4-\sigma^4.$$

(*) Δείξτε ότι $E\big[(X_i-X_j)^2(X_i-X_l)^2\big]=\mu_4+3\sigma^4$, όπου $\mu_4:=E[(X-E(X))^4]$.

4) $|\{i,j\}\cap\{k,l\}|=2$. Υπάρχουν $2n(n-1)$ τέτοιοι όροι, και χωρίς βλάβη της γενικότητας υποθέτουμε $i=k$, $j=l$. Τότε:

$$E\Big[\big[(X_i-X_j)^2-2\sigma^2\big]\big[(X_i-X_j)^2-2\sigma^2\big]\Big]=E\big[(X_i-X_j)^2(X_i-X_j)^2\big]-8\sigma^4+4\sigma^4\overset{(**)}{=}2\mu_4+2\sigma^4$$

'Άσκηση 11: Έστω $\underline{X}=(X_1, X_2, \dots, X_n)$ τυχαίο δείγμα από κατανομή με συνάρτηση κατανομής F_X και συνάρτηση πυκνότητας f_X (δηλ X_i συνεχείς). Συμβολίζουμε με $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ το διατεταγμένο δείγμα, όπου $X_{(1)} := \min_{i=1,\dots,n}\{X_i\}$ και $X_{(n)} = \max_{i=1,\dots,n}\{X_i\}$.

i) Να βρεθούν οι συναρτήσεις κατανομής των $X_{(k)}$, $1 \le k \le n$.

ii) Να βρεθούν οι συναρτήσεις πυκνότητας των $X_{(k)}$, $1 \le k \le n$.

iii) Αν επιπλέον $X_i \sim Unif(0,1)$, να δείξετε ότι $X_{(k)} \sim B(k, n-k+1)$

Λύση:

i) Έστω $k \in T_n := \{1,2,\dots,n\}$. Για την $X_{(k)}$ ισχύει:

$$F_{X_{(k)}}(x) = P(X_{(k)} \le x) = P(\text{"τουλ. } k \text{ από τις } n \text{ τ.μ.} \le x\text{"}) =$$

$$= P\Bigg(\underbrace{\bigcup_{\substack{I:\, I \subseteq T_n,\\ k \le |I| \le n}}}_{\text{ξένα ενδεχόμενα}} \Big[\underbrace{\Big(\bigcap_{i \in I}\{X_i \le x\}\Big) \cap \Big(\bigcap_{j \notin I}\{X_j > x\}\Big)}_{\text{ανεξάρτητα ενδεχόμενα}}\Big]\Bigg)$$

$$= \sum_{\substack{I:\, I \subseteq T_n\\ k \le |I| \le n}} \underbrace{P\Big(\bigcap_{i \in I}\{X_i \le x\}\Big)}_{\text{ανεξάρτητα}} \cdot \underbrace{P\Big(\bigcap_{j \notin I}\{X_j > x\}\Big)}_{\text{ανεξάρτητα}} \quad m := |I|$$

$$= \sum_{m=k}^{n} \binom{n}{m} \cdot P(X_i \le x)^m \cdot P(X_j > x)^{n-m}$$

$$= \sum_{m=k}^{n} \binom{n}{m} F_X^m(x) \cdot \big[1 - F_X(x)\big]^{n-m}$$

Παρατηρώντας ότι $\sum_{m=0}^{n} \binom{n}{m} F_X^m(x)\big[1-F_X(x)\big]^{n-m} = \big[F_X(x) + 1 - F_X(x)\big]^n = 1$, έχουμε εναλλακτικό τύπο:

$$F_{X_{(k)}}(x) = 1 - \sum_{m=0}^{k-1} \binom{n}{m} F_X^m(x) \cdot \big[1 - F_X(x)\big]^{n-m}$$

(**) Δείξτε ότι $E\left[(X_i-X_j)^2(X_i-X_j)^2\right] = 2\mu_4+6\sigma^4$.

Τελικά, έχουμε:

$$4n^2(n-1)^2 Var(S^2) = 4n(n-1)(n-2)\cdot(\mu_4-\sigma^4) + 2n(n-1)(2\mu_4+2\sigma^4)$$

$$= \left[4n(n-1)(n-2)+4n(n-1)\right]\mu_4 + \left[4n(n-1) - 4n(n-1)(n-2)\right]\sigma^4$$

$$= 4n(n-1)^2\mu_4 \quad 4n(n-1)(n-3)\sigma^4$$

$$= 4n(n-1)\left[(n-1)\mu_4-(n-3)\sigma^4\right]$$

Δηλαδή:

$$Var(S^2) = \frac{\mu_4}{n} - \frac{n-3}{n(n-1)}\sigma^4$$

Από τη σχέση $M_2 = \frac{n-1}{n}S^2$ προκύπτει άμεσα ότι: $Var(M_2) = \left(\frac{n-1}{n}\right)^2 Var(S^2)$,

δηλαδή

$$Var(M_2) = \frac{(n-1)^2}{n^3}\mu_4 - \frac{(n-1)(n-3)}{n^3}\sigma^4$$

Ως ειδική περίπτωση, παίρνουμε:

$$F_{X_{(n)}}(x) = F_X(x)^n \quad , \quad F_{X_{(1)}}(x) = 1 - \left[1-F_X(x)\right]^n .$$

ii) Παραγωγίζουμε την πρώτη έκφραση της $F_{X_{(k)}}(x)$:

$$f_{X_{(k)}}(x) = F'_{X_{(k)}}(x) = \left(\sum_{m=k}^{n-1} \binom{n}{m} F_X^m(x) \left[1-F_X(x)\right]^{n-m} + F_X^n(x) \right)'$$

$$= \sum_{m=k}^{n-1} \binom{n}{m} \cdot m \cdot f_X(x) \cdot F_X^{m-1}(x) \cdot \left[1-F_X(x)\right]^{n-m} - \sum_{m=k}^{n-1} \binom{n}{m} \cdot f_X(x)\, F_X^m(x) \left[1-F_X(x)\right]^{n-m-1} (n-m) + n\, f_X(x) \cdot F_X^{n-1}(x) \quad (*)$$

Το πρώτο άθροισμα, χρησιμοποιώντας τη σχέση $\binom{n}{m} m = \binom{n-1}{m-1} n$, γράφεται

$$\sum_{m=k}^{n-1} \binom{n-1}{m-1} n \cdot f_X(x) \cdot F_X^{m-1}(x) \left[1-F_X(x)\right]^{n-m} = \sum_{m=k+1}^{n-1} \binom{n-1}{m-1} n \cdot f_X(x) \cdot F_X^{m-1}(x) \left[1-F_X(x)\right]^{n-m} + \binom{n-1}{k-1} n f_X(x) \cdot F_X^{k-1}(x) \left[1-F_X(x)\right]^{n-k}$$

Με αλλαγή μεταβλητής $\lambda = m-1$ παίρνουμε:

$$= \sum_{\lambda=k}^{n-2} \binom{n-1}{\lambda} n \cdot f_X(x)\, F_X^{\lambda}(x) \left[1-F_X(x)\right]^{n-\lambda-1} + \binom{n-1}{k-1} n\, f_X(x) \cdot F_X^{k-1}(x) \left[1-F_X(x)\right]^{n-k} \quad (1)$$

Το δεύτερο άθροισμα, χρησιμοποιώντας τη σχέση $\binom{n}{m}(n-m) = \binom{n-1}{m} \cdot n$, γράφεται

$$\sum_{m=k}^{n-2} \binom{n-1}{m} n\, f_X(x)\, F_X^m(x) \left[1-F_X(x)\right]^{n-m-1} + \cancelto{1}{\binom{n-1}{n-1}} n\, f_X(x)\, F_X^{n-1}(x) \cancelto{1}{\left[1-F_X(x)\right]^0} \quad (2)$$

Συνδυάζοντας την αρχική σχέση $(*)$ με τις (1), (2), έχουμε:

$$f_{X_{(k)}}(x) = \binom{n-1}{k-1} n\, f_X(x)\, F_X^{k-1}(x) \left[1-F_X(x)\right]^{n-k} - \cancel{n f_X(x)\, F_X^{n-1}(x)} + \cancel{n f_X(x) \cdot F_X^{n-1}(x)}$$

ή αλλιώς $f_{X_{(k)}}(x) = \dfrac{n!}{(k-1)!\,(n-k)!} f_X(x) \cdot F_X^{k-1}(x) \left[1-F_X(x)\right]^{n-k}$, $\quad x \in S_X$.

iii) Έστω $X_i \sim Unif(0,1)$. Τότε $f_X(x) = 1 \cdot 1_{(0,1)}(x)$ και

$$F_X(x) = \begin{cases} 0 & x<0 \\ x & 0 \le x < 1 \\ 1 & 1 \le x \end{cases}$$

. Χρησιμοποιώντας τον τύπο της $f_{X_{(k)}}(x)$

για κάποιο $x \in [0,1]$, έχουμε:
↳ S_x

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!\,(n-k)!}\, x^{k-1} \cdot (1-x)^{\overbrace{n-k}^{(n-k+1)-1}}$$

Παρατηρούμε ότι αν $Y \sim B(k, n-k+1)$, τότε

$$f_Y(x) = \frac{1}{B(k, n-k+1)}\, x^{k-1} (1-x)^{n-k} = \frac{\Gamma(n+1)}{\Gamma(k)\,\Gamma(n-k+1)}\, x^{k-1} (1-x)^{n-k} =$$

$$= \frac{n!}{(k-1)!\,(n-k)!}\, x^{k-1} (1-x)^{n-k} = f_{X_{(k)}}(x).$$

Άρα $X_{(k)} \sim B(k, n-k+1)$. ∎

<u>Άσκηση 12:</u> Έστω $\underset{\sim}{X} = (X_1, X_2, \dots, X_n)$ τυχαίο δείγμα από την $Unif(0,\theta)$ και $n = 2k+1$, $k \in \mathbb{N}_0$ (n περιττός). Να βρείτε τρεις εκτιμήτριες για τη διάμεσο $m(X)$ βασιζόμενοι:

i) στη μέθοδο ροπών

ii) στη δειγματική διάμεσο

iii) στη μέγιστη παρατήρηση.

Για κάθε εκτιμήτρια, να υπολογιστούν η μέση τιμή, η διασπορά, το αναμενόμενο μέσο σφάλμα (MSE) και το αναμενόμενο απόλυτο σφάλμα (MAE).

<u>Λύση:</u>

i) Από τη μέθοδο ροπών, λύνουμε την εξίσωση $E[X] = \overline{X} \Rightarrow \frac{\theta}{2} = \overline{X} \Rightarrow \tilde{\theta}_n = 2\overline{X}_n$.

Αν $X \sim Unif(0,\theta)$, τότε $m(X) = \frac{\theta}{2}$ $(= E(X))$, επομένως αν αντικαταστήσουμε την ΕΡ $\tilde{\theta}_n$ στη θέση του θ παίρνουμε $\tilde{m}_n = \frac{2\overline{X}_n}{2}$ $\boxed{\tilde{m}_n = \overline{X}_n}$ (plug-in εκτιμήτρια).

ii) Επειδή το n είναι περιττός, η δειγματική διάμεσος είναι η μεσαία παρατήρηση του διατεταγμένου δείγματος, δηλαδή η $X_{(k+1)}$. Αν λοιπόν μιμηθούμε τη μέθοδο των ροπών που ταυτίζει τις δειγματικές ροπές με αυτές του πληθυσμού, παίρνουμε $\boxed{\hat{m}_n = X_{(k+1)}}$.

iii) Αν δεχθούμε την $X_{(n)}$ ως εκτιμήτρια του θ, τότε έχουμε την plug-in εκτιμήτρια $\boxed{m^*_n = \frac{X_{(n)}}{2}}$.

Για κάθε μία εκτιμήτρια θα υπολογίσουμε μέση τιμή, μεροληψία, διασπορά και αναμενόμενο τετραγωνικό σφάλμα.

$$\widetilde{m}_n = \overline{X}_n$$

$$E[\widetilde{m}_n] = E[\overline{X}_n] = E[X_1] = \frac{\theta}{2} = m$$

$X \sim Unif(a,b)$

$Var(X) = \frac{(b-a)^2}{12}$

$$bias(\widetilde{m}_n, m) = E[\widetilde{m}_n] - m = 0$$

$$Var(\widetilde{m}_n) = Var(\overline{X}_n) = \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right) = \frac{n \cdot Var(X_1)}{n^2} = \frac{1}{n} Var(X_1) = \frac{\theta^2}{12n}.$$

$$MSE(\widetilde{m}_n, m) = bias^2(\widetilde{m}_n, m) + Var(\widetilde{m}_n) = \frac{\theta^2}{12n}$$

$$\hat{m}_n = X_{(k+1)}$$

Θέτουμε $U_i = \frac{X_i}{\theta} \sim Unif(0,1)$, άρα $U_{(k+1)} \sim B(k+1, n-k)$ και $\hat{m}_n = X_{(k+1)} = \theta \cdot U_{(k+1)}$, οπότε

$$E[\hat{m}_n] = \theta E[U_{(k+1)}] = \theta \frac{k+1}{k+1+n-k} = \theta \frac{k+1}{n+1} = \theta \frac{k+1}{2k+2} = \frac{\theta}{2} = m$$

$$bias(\hat{m}_n, m) = E[\hat{m}_n] - m = 0$$

$$Var(\hat{m}_n) = Var(\theta \cdot U_{(k+1)}) = \theta^2 \cdot Var(U_{(k+1)}) = \theta^2 \cdot \frac{(k+1)(n-k)}{(k+1+n-k)^2(k+1+n-k+1)} =$$

$$= \theta^2 \cdot \frac{(k+1)(k+1)}{(n+1)^2(n+2)} = \theta^2 \frac{(k+1)^2}{4(k+1)^2(n+2)} = \frac{\theta^2}{4(n+2)}$$

$$MSE(\hat{m}_n, m) = bias^2(\hat{m}_n, m) + Var(\hat{m}_n) = \frac{\theta^2}{4(n+2)}$$

$m_n^* = X_{(n)}/2$

Θέτουμε $U_i = X_i/\theta \sim Unif(0,1)$, άρα $U_{(n)} \sim B(n,1)$ και $m_n^* = \frac{\theta U_{(n)}}{2}$, οπότε

$$E[m_n^*] = \frac{\theta}{2} E[U_{(n)}] = \frac{\theta}{2} \cdot \frac{n}{n+1} \quad (\neq m)$$

$$bias(m_n^*, m) = E[m_n^*] - m = \frac{\theta}{2} \cdot \frac{n}{n+1} - \frac{\theta}{2} = -\frac{1}{n+1} \cdot \frac{\theta}{2}$$

$$Var(m_n^*) = \frac{\theta^2}{4} Var(U_{(n)}) = \frac{\theta^2}{4} \cdot \frac{n \cdot 1}{(n+1)^2(n+2)} = \frac{n\theta^2}{4(n+1)^2(n+2)}$$

$$MSE(m_n^*, m) = bias^2(m_n^*, m) + Var(m_n^*) = \frac{\theta^2}{4(n+1)^2} + \frac{n\theta^2}{4(n+1)^2(n+2)} = \ldots = \frac{\theta^2}{2(n+1)(n+2)}$$

Παρατηρούμε ότι για $n \geq 3$ τα MSE των τριών εκτιμητριών ακολουθούν την παρακάτω διάταξη:

$$MSE(m_n^*, m) < MSE(\tilde{m}_n, m) < MSE(\hat{m}_n, m)$$

Παρ'όλα αυτά, η εκτιμήτρια m_n^* παρουσιάζει μεροληψία. Δημιουργούμε λοιπόν μία διορθωμένη έκδοση της m_n^*, την $\underbrace{m_{cn}^*}_{\text{corrected}} := \frac{n+1}{n} \cdot \frac{X_{(n)}}{2}$, για την οποία ισχύει

$$E[m_{cn}^*] = \frac{n+1}{n} \cdot \frac{\theta}{2} E[U_{(n)}] = \frac{\cancel{n+1}}{\cancel{n}} \cdot \frac{\theta}{2} \frac{\cancel{n}}{\cancel{n+1}} = \frac{\theta}{2} = m$$

$$bias(m_{cn}^*, m) = 0$$

$$Var(m_{cn}^*) = \left(\frac{n+1}{n}\right)^2 \cdot Var(m_n^*) = \frac{\cancel{(n+1)^2}}{n^2} \cdot \frac{\cancel{n}\theta^2}{4\cancel{(n+1)^2}(n+2)} = \frac{\theta^2}{4n(n+2)}$$

$$MSE(m_{cn}^*, m) = bias^2(m_{cn}^*, m) + Var(m_{cn}^*) = \frac{\theta^2}{4n(n+2)}$$

Παρατηρούμε ότι για $n \geq 2$ η m_{cn}^* παρουσιάζει το ελάχιστο MSE από τις τέσσερεις εκτιμήτριες του m.

Ο υπολογισμός του αναμενόμενου απόλυτου σφάλματος είναι γενικά πιο δύσκολος από αυτόν του αντίστοιχου τετραγωνικού.

$\widetilde{m}_n = \overline{X}_n$

Ο υπολογισμός της κατανομής του $\overline{Y}_n$ στην περίπτωση που $Y_i \overset{iid}{\sim} Unif(0,1)$ απαιτεί τον υπολογισμό της κατανομής του $\sum_{i=1}^{n} Y_i$ μέσω συνελίξεων. Εδώ θα μελετήσουμε μόνο την ειδική περίπτωση $n=2$, δηλαδή $\widetilde{m}_n = \frac{X_1+X_2}{2}$.

Η κατανομή της $Z := Y_1+Y_2$ ονομάζεται τριγωνική (triangular) και έχει σ.π.π. $f(z) = \begin{cases} 0 & z<0 \\ z & 0\le z<1 \\ 2-z & 1\le z<2 \\ 0 & 2\le z \end{cases}$.

$$\widetilde{m}_2 = \frac{X_1+X_2}{2} = \frac{\theta Y_1 + \theta Y_2}{2} = \frac{\theta}{2}\cdot Z, \text{ όπου } X_i \sim Unif(0,\theta),\ Y_i := \frac{X_i}{\theta} \sim Unif(0,1)$$

$$MAE(\widetilde{m}_2, m) = E\left[|\widetilde{m}_2 - m|\right] = E\left[\left|\frac{\theta}{2}Z - \frac{\theta}{2}\right|\right] = \frac{\theta}{2}\cdot E\left[|Z-1|\right].$$

Ο υπολογισμός της παραπάνω ποσότητας επιτυγχάνεται ως εξής:

$$E\left[|Z-1|\right] = \int_0^2 |z-1|\cdot f(z)dz = \int_0^1 |z-1|\cdot f(z)dz + \int_1^2 |z-1| f(z) dz =$$

$$= \int_0^1 (1-z)\cdot z\, dz + \int_1^2 (z-1)(2-z)\, dz = \int_0^1 z - z^2 dz - \int_1^2 z^2-3z+2\ dz =$$

$$= \left[\frac{z^2}{2} - \frac{z^3}{3}\right]_0^1 - \left[\frac{z^3}{3} - \frac{3z^2}{2} + 2z\right]_1^2 = \left(\frac{1}{2}-\frac{1}{3}\right) - \left(\frac{8}{3} - 6 + 4 - \frac{1}{3} + \frac{3}{2} - 2\right) = \frac{1}{3}$$

Τελικά, $MAE(\widetilde{m}_2, m) = \frac{\theta}{6}$.

$m_n^* = X_{(n)}/2$

Σε αυτήν την περίπτωση έχουμε, όπως πριν, $m_n^* = \frac{\theta}{2}\cdot U_{(n)}$, $U_{(n)} \sim Beta(n,1)$.

$$MAE(m_n^*, m) = E\left[|m_n^* - m|\right] = E\left[\left|\frac{\theta}{2}U_{(n)} - \frac{\theta}{2}\right|\right] = \frac{\theta}{2} E\left[|\underbrace{U_{(n)}}_{\le 1} - 1|\right] = \frac{\theta}{2} E\left[1 - U_{(n)}\right] =$$

$$= \frac{\theta}{2}\cdot\left[1 - E[U_{(n)}]\right] = \frac{\theta}{2}\left(1 - \frac{n}{n+1}\right) = \frac{\theta}{2n}.$$

∎