

Στατιστική
Μέρος Δεύτερο: Σημειακή Εκτίμηση

Διδάσκουσα: Λουκία Μελιγκοτσίδου
Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών
Τμήμα Μαθηματικών

November 8, 2017

1 Εισαγωγικά

Η ενότητα της *Εκτίμησης σε Σημείο* αποτελείται από δύο μέρη:

- (a) μεθοδοι εύρεσης εκτιμητών και
- (b) αξιολόγηση των εκτιμητών.

Η ουσία είναι απλή. Όταν διαθέτουμε ένα τ.δ. από πληθυσμο με σ.π. $f(x|\underline{\theta})$, τότε γνώση του $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_x)$ (παράμετρος) μας δίνει γνώση του πληθυσμού.

[Ορισμός]: Οποιαδήποτε σ.σ. $T(X_1, X_2, \dots, X_n)$ ενός τ.δ. $X_1, X_2, \dots, X_n$ που χρησιμοποιείται για την εκτίμηση του $g(\theta)$ ονομάζεται *σημειακός εκτιμητής* του $g(\theta)$.

- Ο εκτιμητής ως σ.σ. είναι τ.μ. αλλά και γνωστός αριθμός που στηρίζετε στο τ.δ. (παίρνει συγκεκριμένη τιμή σε συγκεκριμένο δείγμα).
- Η τ.μ. $T(X_1, X_2, \dots, X_n)$ καλείται *εκτιμητρια* ενώ η τιμή της στο συγκεκριμένο δείγμα, $T(x_1, x_2, \dots, x_n)$, καλείται *εκτίμηση*.

Αν και υπάρχουν φορές που οι εκτιμητές είναι προφανείς (όπως ο δειγματικός μέσος για τον αντίστοιχο του πληθυσμού), πολλές φορές δεν είναι.

1.1 Μέθοδος των ροπών

- Παλαιότερη μέθοδος και απλούστερη όλων.
- Σχεδόν πάντα δίνει εκτιμητή.
- Πολλές φορές δεν δίνει "ποιοτικούς" εκτιμητές.

[Περιγραφή]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. από πληθυσμό με $f(x|\theta)$ όπου $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_k)$. Οι εκτιμητές βρίσκονται εξισώνοντας τις πρώτες k -ροπές του πληθυσμού με τις αντίστοιχες k -ροπές του δείγματος και λύνοντας το σύστημα των k -εξισώσεων με k -αγνώστους. Πιο συγκεκριμένα:

Ροπές Πληθυσμού	Ροπές Δείγματος
$\mu_1 = E(X)$	$m_1 = \frac{1}{n} \sum_{i=1}^n X_i$
$\mu_2 = E(X^2)$	$m_2 = \frac{1}{n} \sum_{i=1}^n X_i^2$
$\vdots$	$\vdots$
$\mu_k = E(X^k)$	$m_k = \frac{1}{n} \sum_{i=1}^n X_i^k.$

Οι ροπές του πληθυσμού θα είναι συνάρτηση των $\theta_1, \theta_2, \dots, \theta_k$, οπότε οι εκτιμήτριες προκύπτουν από την λύση του συστήματος

$$\begin{aligned}\mu_1(\theta_1, \theta_2, \dots, \theta_k) &= m_1 \\ \mu_2(\theta_1, \theta_2, \dots, \theta_k) &= m_2 \\ &\vdots \\ \mu_k(\theta_1, \theta_2, \dots, \theta_k) &= m_k.\end{aligned}$$

[Παράδειγμα 1]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. από πληθυσμό με κατανομή $Bernoulli(p)$. Να βρεθεί εκτιμήτρια του p με τη μέθοδο των ροπών.

[Λύση]: $X \sim Bernoulli(p)$:

$$f(x; p) = p^x(1-p)^{1-x}, \quad x = 0, 1, \quad 0 < p < 1.$$

Έχουμε μια άγνωστη παράμετρο και συνεπώς θα εξισώσουμε τον μέσο του πληθυσμού $\mu = E(X)$ με το δειγματικό μέσο. Συνεπώς:

$$\mu = E(X) = p = \bar{x} \Rightarrow \hat{p} = \bar{x}.$$

[Παράδειγμα 2]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. από πληθυσμό με κατανομή $Poisson(\lambda)$. Να βρεθεί εκτιμήτρια του λ με τη μέθοδο των ροπών.

[Λύση]: $X \sim Poisson(\lambda)$:

$$f(x; \lambda) = e^{-\lambda} \frac{\lambda^x}{x!}, \quad x = 0, 1, \dots, \quad \lambda > 0.$$

Μία η άγνωστη παράμετρος, οπότε

$$\mu = E(X) = \lambda = \bar{x} \Rightarrow \hat{\lambda} = \bar{x}.$$

[Παρατήρηση]:

Αν έχω να εκτιμήσω 2 άγνωστες παραμέτρους, συχνά εξισώνω το $E(x)$ με το $\bar{x}$ και για δεύτερη εξίσωση $E[(x - \mu)^2] = V(x) = \frac{1}{n} \sum (x_i - \bar{x})^2$ (δηλ. δεύτερη ροπή περί το μέσο). Δηλ. μπορώ να έχω

$$\begin{aligned} E(X) = \bar{x} \quad \text{είτε} \quad E(X) = \bar{x} \\ V(X) = \frac{1}{n} \sum (x_i - \bar{x})^2 \quad E(X^2) = \frac{1}{n} \sum x_i^2. \end{aligned}$$

Γενικότερα μπορώ να πάρω ροπές περί το μέσο:

$$\mu_k = E[(x - \mu)^k] \quad \text{με τις} \quad M_k = \frac{1}{n} \sum (x_i - \bar{x})^k.$$

[Παράδειγμα 3]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. από πληθυσμό με κατανομή $N(\theta, \sigma^2)$. Ποιοί οι εκτιμητές των θ και σ^2 .

[Λύση]: Βάση συμβολισμού έχουμε $\theta_1 = \theta$ και $\theta_2 = \sigma^2$. Ακόμα έχουμε:

$$\mu_1 = E(X) = \theta \quad \& \quad m_1 = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

$$\mu_2 = E(X^2) = Var(X) + [E(X)]^2 \quad \& \quad m_2 = \frac{1}{n} \sum_{i=1}^n X_i^2.$$

Οπότε:

$$\left. \begin{matrix} \mu_1 = m_1 \\ \mu_2 = m_2 \end{matrix} \right\} \Rightarrow \left. \begin{matrix} \theta = \bar{X} \\ \sigma^2 + \theta^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 \end{matrix} \right\} \Rightarrow$$

$$\left. \begin{matrix} \hat{\theta} = \bar{X} \\ \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2 \end{matrix} \right\} \Rightarrow \left. \begin{matrix} \hat{\theta} = \bar{X} \\ \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \end{matrix} \right\}.$$

[Παράδειγμα 4]: Έστω $x_1, x_2, \dots, x_n$ τ.δ. από πληθυσμό με κατανομή $Bin(k, p)$. Με την μεθοδο των ροπών να βρεθεί ο εκτιμητής (α) του p (k -γνωστό) και (β) των p και k .

[Λύση]:

(i) Έχουμε:

$$\left. \begin{matrix} \mu_1 = E(X) = kp \\ m_1 = \bar{x} \end{matrix} \right\} \Rightarrow kp = \bar{x} \Rightarrow \hat{p} = \frac{\bar{x}}{k}$$

(ii) Επιπλέον:

$$\left. \begin{matrix} \mu_2 = E(X^2) = Var(X) + [E(X)]^2 = kp(1-p) + k^2p^2 \\ m_2 = \frac{1}{n} \sum_{i=1}^n x_i^2 \end{matrix} \right\}.$$

Εξισώνοντας $\mu_2 = m_2$ μαζί με την εξίσωση του (i) καταλήγουμε στο σύστημα:

$$\left. \begin{matrix} kp = \bar{x} \\ kp(1-p) + k^2p^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 \end{matrix} \right\} \Rightarrow$$

$$\left. \begin{aligned} \hat{p} &= \frac{\bar{x}}{k} \\ \bar{x}(1 - \frac{\bar{x}}{k}) + \bar{X}^2 &= \frac{1}{v} \sum_{i=1}^v x_i^2 \end{aligned} \right\} \Rightarrow$$

$$\left. \begin{aligned} \hat{p} &= \frac{\bar{x}}{k} \\ \bar{x}(1 - \frac{\bar{x}}{k}) &= \frac{1}{v} \sum_{i=1}^v (x_i - \bar{x})^2 \end{aligned} \right\} \Rightarrow$$

$$\left. \begin{aligned} \hat{p} &= \frac{\bar{x}}{k} \\ \hat{k} &= \bar{x}^2 / \left[\bar{x} - \frac{1}{v} \sum_{i=1}^v (x_i - \bar{x})^2 \right] \end{aligned} \right\}.$$

Αλλιώς η δεύτερη εξίσωση στο πιο πάνω σύστημα μπορεί να αντικατασταθεί με:

$$E[(x - \mu)^2] = kp(1 - p) = \frac{1}{v} \sum (x_i - \bar{x})^2 = M_2.$$

Διαιρώ τις δύο εξισώσεις και παίρνω

$$\frac{kp(1 - p)}{kp} = \frac{M_2}{\bar{x}} \Rightarrow \hat{p} = 1 - \frac{M_2}{\bar{x}},$$

και από τη πρώτη εξίσωση παίρνουμε

$$\hat{k} = \frac{\bar{x}}{\hat{p}} = \frac{\bar{x}}{1 - \frac{M_2}{\bar{x}}} = \frac{\bar{x}^2}{\bar{x} - M_2}.$$

[Παρατηρήσεις]:

- Το k πρέπει να είναι ακέραιος. Όταν είναι άγνωστο και το εκτιμώ, παίρνω σαν εκτιμήτρια τον πλησιέστερο ακέραιο.
- Οι παραπάνω εκτιμητές μάλλον δεν είναι οι καλύτεροι. Βλέπουμε ότι το k και p μπορούν θεωρητικά να λάβουν και αρνητικές τιμές, αν η μεταβλητότητα του δείγματος είναι μεγαλύτερη της δειγματικής μέσης τιμής ($M_2 > \bar{x}$). Αν ισχύει αυτό τότε μάλλον το μοντέλο μας δεν είναι το κατάλληλο για να περιγράψει τα δεδομένα
 - Αν $M_2 \simeq \bar{x}$, τότε Poisson (όπου $E(X) = V(X)$)
 - Αν $M_2 > \bar{x}$, τότε Αρνητική Διωνυμική (όπου $V(X) > E(X)$).

Μέθοδος Μεγίστης Πιθανοφάνειας

[O]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. από πληθυσμό με σ.π.π. $f(x; \theta)$ (ή σ.π. $p(x; \theta)$). Τότε η συνάρτηση

$$L(\theta|\underline{x}) = \prod_{i=1}^n f(x_i, \theta),$$

θεωρείται συνάρτηση της παραμέτρου θ και ονομάζεται *συνάρτηση πιθανοφάνειας*.

[O]: Έστω $L(\underline{\theta}|\underline{x})$ η συνάρτηση πιθανοφάνειας ενός τ.δ. $X_1, X_2, \dots, X_n$. Ο εκτιμητής του $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_x)$ λέγεται εκτιμητής μεγίστης πιθανοφάνειας (ΕΜΠ) του $\underline{\theta}$ αν

$$L(\hat{\underline{\theta}}|\underline{x}) = \max_{\underline{\theta} \in \Theta} L(\underline{\theta}|\underline{x}).$$

Συχνά συμβολίζουμε τη $L(\underline{\theta}|\underline{x})$ με $L(\underline{\theta})$.

[Παρατηρήσεις]

- Συνήθως για λόγους ευκολίας αντί να μεγιστοποιούμε την $L(\underline{\theta}|\underline{x})$ μεγιστοποιούμε την $\ln L(\underline{\theta}|\underline{x})$ που λαμβάνει μέγιστο στο ίδιο σημείο.
- Όταν μεγιστοποιούμε την $L(\underline{\theta}|\underline{x})$ μπορεί να έχουμε τις εξής περιπτώσεις
 - Δεν υπάρχει πεπερασμένο μέγιστο
 - Υπάρχει μόνον ένα
 - Υπάρχουν περισσότερα από ένα μέγιστα.

Τρόπος Εργασίας για την Εύρεση Ε.Μ.Π.

- Εάν $X_1, X_2, \dots, X_n$ τ.δ. από $f(x; \theta)$, τότε γράφουμε τη συνάρτηση πιθανοφάνειας

$$L(\theta) = f(x_1, x_2, \dots, x_n; \theta) = f(\underline{x}; \theta) = \prod_{i=1}^n f(x_i; \theta)$$

- Μεγιστοποίηση της $l(\theta) = \log L(\theta)$ ως προς θ , όπου $\hat{\theta}$ η τιμή για την οποία

$$\frac{\partial}{\partial \theta} l(\theta) = 0.$$

- Η $\frac{\partial}{\partial \theta} l(\theta) = 0$ δίνει ακρότατο! Για να είναι μέγιστο θα πρέπει

$$\frac{\partial^2}{\partial \theta^2} l(\theta) \Big|_{\theta=\hat{\theta}} < 0.$$

Παρατήρηση: Για τις γνωστές κατανομές αυτό ισχύει!

[Παραδείγματα]

[Παράδειγμα 1]: Έστω X ο αριθμός των επιτυχιών ενός διωνυμικού πειράματος με πιθανότητα επιτυχίας p . Ποιός ο ΕΜΠ του p ;

[Λύση 1]: Ο ΕΜΠ θα είναι η τιμή του p που μεγιστοποιεί την $L(p|x)$ ή $\ln L(p|x)$. Έχουμε:

$$L(p|x) = \binom{n}{x} p^x (1-p)^{n-x},$$

και συνεπώς

$$\ln L(p|x) = \ln \binom{n}{x} + x \ln p + (n-x) \ln(1-p)$$

και

$$\frac{\partial \ln L(p|x)}{\partial p} = \frac{x}{p} - \frac{n-x}{1-p}.$$

Από την λύση της εξίσωσης

$$\frac{\partial \ln L(p|x)}{\partial p} \Big|_{p=\hat{p}} = 0$$

ως προς p έπεται ότι

$$\hat{p} = \frac{x}{n}.$$

Ελέγχοντάς την δεύτερη παράγωγο βλέπουμε ότι

$$\frac{\partial^2 \ln L(p|x)}{\partial p^2} \Big|_{p=\hat{p}} = -\frac{x}{\hat{p}^2} - \frac{n-x}{(1-\hat{p})^2} < 0,$$

και συνεπώς το $\hat{p} = \frac{x}{n}$ είναι μέγιστο.

[Παράδειγμα 2]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. από πληθυσμό με κατανομή την $Exp(1/\theta)$.
Ποιός ο ΕΜΠ του θ ;

[Λύση 2]: Έχουμε:

$$L(\theta|\underline{x}) = \prod_{i=1}^n f(x_i; \theta) = \frac{1}{\theta^n} \exp\left\{-\frac{1}{\theta} \sum_{i=1}^n x_i\right\},$$

οπότε λαμβάνουμε:

$$\ln L(\theta|\underline{x}) = -n \ln \theta - \frac{1}{\theta} \sum_{i=1}^n x_i.$$

Επομένως:

$$\frac{\partial \ln L(\theta|\underline{x})}{\partial \theta} = -\frac{n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^n x_i,$$

και από την σχέση $\left. \frac{\partial \ln L(\theta|\underline{x})}{\partial \theta} \right|_{\theta=\hat{\theta}} = 0$ παίρνουμε

$$\hat{\theta} = \frac{\sum_{i=1}^n x_i}{n} = \bar{X}.$$

Ελέγχοντας την δεύτερη παράγωγο βλέπουμε ότι όντως έχουμε μέγιστο:

$$\left. \frac{\partial^2 \ln L(\theta|\underline{x})}{\partial \theta^2} \right|_{\theta=\hat{\theta}} = \frac{n}{\hat{\theta}^2} - 2 \frac{1}{\hat{\theta}^3} \sum_{i=1}^n x_i = -\frac{n^3}{(\sum_{i=1}^n x_i)^2} < 0.$$

[Παράδειγμα 3]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. από πληθυσμό με κατανομή την $Poisson(\theta)$. Ποιά η ΕΜΠ του θ ;

[Λύση 3]: Έχουμε:

$$L(\theta|\underline{x}) = \prod_{i=1}^n e^{-\theta} \frac{\theta^{x_i}}{x_i!} = \frac{e^{-n\theta} \theta^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}$$

και λαμβάνουμε:

$$\ln L(\theta|\underline{x}) = -n\theta + \sum_{i=1}^n x_i \ln \theta - \ln \prod_{i=1}^n x_i!.$$

Επομένως:

$$\frac{\partial \ln L(\theta|\underline{x})}{\partial \theta} \Big|_{\theta=\hat{\theta}} = -n + \frac{\sum_{i=1}^n x_i}{\hat{\theta}} = 0$$

και συνεπώς

$$\hat{\theta} = \frac{\sum_{i=1}^n x_i}{n} = \bar{X}.$$

Ελέγχοντας την δεύτερη παράγωγο βλέπουμε ότι όντως έχουμε μέγιστο:

$$\frac{\partial^2 \ln L(\theta|\underline{x})}{\partial \theta^2} \Big|_{\theta=\hat{\theta}} = -\frac{n^2}{\sum_{i=1}^n x_i} < 0.$$

Παρατήρηση: Για την εύρεση ΕΜΠ δεν χρησιμοποιούμε πάντα τη μέθοδο παραγωγίσης!

[Παράδειγμα 4]: Οι καραμέλες smarties βγαίνουν σε ω ισοπίθανα χρώματα. Έστω ότι δεν γνωρίζουμε το ω . Παίρνουμε ένα δείγμα από 3 καραμέλες και αυτές είναι με τη σειρά κόκκινη-πράσινη-κόκκινη.

Η πιθανοφάνεια για αυτά τα δεδομένα είναι

x : το χρώμα της δεύτερης καραμέλας διαφέρει από αυτό της πρώτης ενώ το χρώμα της τρίτης είναι το ίδιο με της πρώτης.

Άρα

$$\begin{aligned} L(\omega) &= P(x|\omega) \\ &= P(2\eta \text{ διαφέρει από } 1\eta)P(3\eta \text{ ταιριάζει με τη } 1\eta) \\ &= \left(\frac{\omega-1}{\omega}\right) \frac{1}{\omega} \\ &= \frac{\omega-1}{\omega^2}, \end{aligned}$$

η οποία παίρνει τιμές $1/4, 2/9, 3/16$ για $\omega=2,3,4$ και συνεχίζει να μειώνεται για $\omega > 4$. Συνεπώς, η εκτιμήτρια μεγίστης πιθανοφάνειας είναι $\hat{\omega} = 2$.

Αν μιά τέταρτη καραμέλα είχε χρώμα μπλε, τότε

$$\begin{aligned} L(\omega) &= \left(\frac{\omega-1}{\omega}\right) \frac{1}{\omega} \frac{\omega-2}{\omega} \\ &= \frac{(\omega-1)(\omega-2)}{\omega^3}, \end{aligned}$$

η οποία παίρνει τιμές $2/27, 3/32, 12/125, 5/54$ για $\omega=3,4,5,6$ και συνεχίζει να μειώνεται για $\omega > 6$. Συνεπώς, η εκτιμήτρια μεγίστης πιθανοφάνειας είναι $\hat{\omega} = 5$.

[Παράδειγμα 5]: Έστω τ.δ $X_1, X_2, \dots, X_n \sim \text{Geometric}(p)$, όπου p προς εκτίμηση.
Τότε

$$\begin{aligned} l(p) &= \log f(x_1, x_2, \dots, x_n | p) \\ &= \log \prod_{i=1}^n p(1-p)^{x_i-1} \\ &= \left(\sum_{i=1}^n x_i - n \right) \log(1-p) + n \log p \end{aligned}$$

η οποία παίρνει μέγιστη τιμή όταν

$$-\frac{\sum_i x_i - n}{1 - \hat{p}} + \frac{n}{\hat{p}} = 0 \Rightarrow \hat{p} = \frac{1}{\bar{X}}.$$

Όμως η ΕΜΠ είναι biased. Έστω $n=1$, τότε

$$E[\hat{p}] = E[1/X_1] = \sum_{i=1}^{\infty} \frac{1}{x} p(1-p)^{x-1} = \frac{p}{1-p} \log p > p.$$

Προσοχή: $E[1/X_1] \neq \frac{1}{E[X_1]}.$

Σημείωση: Εάν η συνάρτηση πιθανοφάνειας έχει s παραμέτρους, έτσι ώστε $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_s)$, τότε η ΕΜΠ του $\underline{\theta}$ είναι το διάνυσμα των τ.μ. $(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_s)$ για το οποίο μεγιστοποιείται η

$$L(\theta_1, \theta_2, \dots, \theta_s) = f(x_1, x_2, \dots, x_n; \theta_1, \theta_2, \dots, \theta_s).$$

Κάτω από μερικές συνθήκες ομαλότητας, το σημείο που μεγιστοποιείται η πιθανοφάνεια βρίσκεται από την λύση

$$\begin{aligned} \frac{\partial}{\partial \theta_1} l(\underline{\theta}) &= 0 \\ \vdots &= \vdots \\ \frac{\partial}{\partial \theta_s} l(\underline{\theta}) &= 0. \end{aligned}$$

Για να έχουμε μέγιστο, θα πρέπει:

$$\frac{\partial^2}{\partial \theta_i^2} l(\underline{\theta}) < 0, \quad i = 1, \dots, s.$$

[Παράδειγμα]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. από πληθυσμό με κατανομή την $N(\mu, \sigma^2)$.
Ποιές οι ΕΜΠ των μ και σ^2 ;

[Λύση]: Έχουμε:

$$L(\underline{\theta}|\underline{x}) = \prod_{i=1}^n f(x_i; \theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x_i - \mu)^2}{2\sigma^2}\right\}$$

ή

$$L(\underline{\theta}|\underline{x}) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right\},$$

και λαμβάνουμε:

$$\ln L(\underline{\theta}|\underline{x}) = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2.$$

Επομένως:

$$\frac{\partial \ln L(\underline{\theta}|\underline{x})}{\partial \mu} \Big|_{\theta=\hat{\theta}} = \frac{1}{\hat{\sigma}^2} \sum_{i=1}^n (x_i - \hat{\mu}) = 0$$

και

$$\frac{\partial \ln L(\underline{\theta}|\underline{x})}{\partial \sigma^2} \Big|_{\theta=\hat{\theta}} = -\frac{n}{2\hat{\sigma}^2} + \frac{1}{2\hat{\sigma}^4} \sum_{i=1}^n (x_i - \hat{\mu})^2 = 0.$$

Συνεπώς έχουμε δύο εξισώσεις με δύο αγνώστους, και μετά από λίγες πράξεις παίρνουμε:

$$\begin{cases} \hat{\mu} = \bar{X} \\ \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2 \end{cases}$$

Εδώ για να έχουμε μέγιστο χρειαζόμαστε να έχουμε $B^2 - A\Gamma < 0$ και $A < 0$ ($\Gamma < 0$), με $A = \frac{\partial^2 \ln L(\underline{\theta}|\underline{x})}{\partial \mu^2}$, $B = \frac{\partial^2 \ln L(\underline{\theta}|\underline{x})}{\partial \mu \partial \sigma^2}$ και $\Gamma = \frac{\partial^2 \ln L(\underline{\theta}|\underline{x})}{\partial \sigma^2}$, κατι που ισχύει.

Εκτιμητές Bayes

Θεώρημα του Bayes για ενδεχόμενα:

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{P(B|A)P(A)}{P(B)}.$$

- Στην κλασική στατιστική, η παράμετρος θ θεωρείται άγνωστη αλλά σταθερή ποσότητα που πρέπει να εκτιμηθεί.
- Στη Μπεϋζιανή στατιστική η παράμετρος θ θεωρείται μια τ.μ. που ακολουθεί συγκεκριμένη κατανομή.
- Η κατανομή αυτή συμβολίζει την αρχική μας θεώρηση για την συγκεκριμένη παράμετρο και ονομάζεται *εκ των προτέρων κατανομή* ή *a-priori distribution*.
- Η *εκ των προτέρων κατανομή* είναι μια υποκειμενική κατανομή (άποψη ειδικού) η οποία και θα αναπροσαρμοστεί με βάση την πληροφορία που θα πάρουμε από το δείγμα.
- Η αναπροσαρμοσμένη *εκ των προτέρων κατανομή* κατανομή ονομάζεται *εκ των υστέρων κατανομή* ή *a-posteriori distribution*.
- Η αναπροσαρμογή γίνεται με βάση τον κανόνα του Bayes.

Έστω $\pi(\theta)$ η *εκ των προτέρων κατανομή* της παραμέτρου και $f(\underline{x}|\theta)$ η δειγματική κατανομή. Τότε η *εκ των υστέρων κατανομή* δίνεται από:

$$\pi(\theta|\underline{x}) = \frac{f(\underline{x}|\theta)\pi(\theta)}{g(\underline{x})},$$

όπου $g(\underline{x})$ είναι η περιθώρια συνάρτηση των X , δηλαδή

$$g(\underline{x}) = \int f(\underline{x}, \theta) d\theta = \int f(\underline{x}|\theta)\pi(\theta) d\theta.$$

[Σημειώσεις]:

- Η *εκ των υστέρων κατανομή* της παραμέτρου θ είναι μια δεσμευμένη κατανομή.
- Η συνάρτηση $f(\underline{x}|\theta)$ είναι η γνωστή μας συνάρτηση πιθανοφάνειας.
- Για σημειακό εκτιμητή της θ μπορούμε να χρησιμοποιήσουμε χαρακτηριστικά της *εκ των υστέρων κατανομής*, όπως ο μέσος ή ο διάμεσος.

[Ο]: Έστω $\mathfrak{S}$ η τάξη όλων των σ.π. ή σ.π.π. $f(x|\theta)$. Μία τάξη Π εκ των προτέρων κατανομών ονομάζεται συζυγής οικογένεια της $\mathfrak{S}$ (conjugate family), αν η εκ των υστέρων κατανομή του θ ανήκει στην τάξη Π για κάθε $f \in \mathfrak{S}$, για όλες τις εκ των προτέρων κατανομές του Π και για όλα τα $x \in X$.

[Π.Χ. - Θεώρημα]: Έστω $\bar{X}$ είναι ο μέσος ενός τ.δ. μεγέθους n από κανονικό πληθυσμό με γνωστή διασπορά σ^2 και με εκ των προτέρων κατανομή για τον μέσο M την κανονική κατανομή με μέσο μ_0 και διασπορά σ_0^2 . Τότε, η εκ των υστέρων κατανομή του M δεδομένου του $\bar{X} = \bar{x}$ είναι κανονική με μέσο μ_1 και διασπορά σ_1^2 , όπου

$$\mu_1 = \frac{n\bar{x}\sigma_0^2 + \mu_0\sigma^2}{n\sigma_0^2 + \sigma^2} \quad \text{και} \quad \frac{1}{\sigma_1^2} = \frac{n}{\sigma^2} + \frac{1}{\sigma_0^2}.$$

[Απόδειξη]:

Για τον μέσο $M = \mu$ έχουμε

$$f(\bar{x}|\mu) = \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \right)^2 \right\}, \quad -\infty < \bar{x} < \infty,$$

και

$$h(\mu) = \frac{1}{\sigma_0\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{\mu - \mu_0}{\sigma_0} \right)^2 \right\}, \quad -\infty < \mu < \infty.$$

Οπότε έχουμε

$$\begin{aligned} \varphi(\mu|\bar{x}) &= \frac{f(\bar{x}|\mu)h(\mu)}{g(\bar{x})} \\ &= \frac{\sqrt{n}}{2\pi\sigma\sigma_0g(\bar{x})} \exp \left\{ -\frac{1}{2} \left(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \right)^2 - \frac{1}{2} \left(\frac{\mu - \mu_0}{\sigma_0} \right)^2 \right\}. \end{aligned}$$

Τώρα συγκεντρώνουμε τις δυνάμεις στην εκθετική συνάρτηση και παίρνουμε

$$-\frac{1}{2} \left(\frac{n}{\sigma^2} + \frac{1}{\sigma_0^2} \right) \mu^2 + \left(\frac{n\bar{x}}{\sigma^2} + \frac{\mu_0}{\sigma_0^2} \right) \mu - \frac{1}{2} \left(\frac{n\bar{x}^2}{\sigma^2} + \frac{\mu_0^2}{\sigma_0^2} \right)$$

και αν θέσουμε

$$\frac{1}{\sigma_1^2} = \frac{n}{\sigma^2} + \frac{1}{\sigma_0^2} \quad \text{και} \quad \mu_1 = \frac{n\bar{x}\sigma_0^2 + \mu_0\sigma^2}{n\sigma_0^2 + \sigma^2}$$

βγάλουμε το $-\frac{1}{2\sigma_1^2}$ κοινό παράγοντα και κάνουμε τις πράξεις παίρνουμε τον εκθέτη της εκθετικής συνάρτησης

$$-\frac{1}{2\sigma_1^2}(\mu - \mu_1)^2 + R,$$

όπου R δεν περιλαμβάνει το μ . Συνεπώς η εκ των υστέρων κατανομή γίνεται

$$\varphi(\mu|\bar{x}) = \frac{\sqrt{n}e^R}{2\pi\sigma\sigma_0g(\bar{x})}e^{-\frac{1}{2\sigma_1^2}(\mu-\mu_1)^2}$$

η οποία είναι μια κανονική κατανομή με μέσο μ_1 και διασπορά σ_1^2 , και μπορεί να γραφεί

$$\varphi(\mu|\bar{x}) = \frac{1}{\sigma_1\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\mu-\mu_1}{\sigma_1}\right)^2}.$$

Είναι φανερό ότι η $g(\bar{x})$ δεν ορίστηκε και απορροφήθηκε στον σταθερο όρο στο τέλος.

Μέθοδοι Αξιολόγησης Εκτιμητών

- Υπάρχει η δυνατότητα να έχουμε περισσότερους από έναν εκτιμητές, κανοντας χρήση διαφόρων τεχνικών.
- Συνεπώς υπάρχει η ανάγκη να εξετάζουμε τους εκτιμητές και να επιλέγουμε τον καλύτερο με βάση κάποια κριτήρια.

[Αμερόληπτη εκτίμηση]: Ο εκτιμητής $T(\underline{X})$ της συνάρτησης $g(\underline{\theta})$ λέγεται αμερόληπτος αν ισχύει

$$E[T(\underline{X})] = g(\underline{\theta}), \quad \forall \underline{\theta} \in \Theta.$$

- Οι εκτιμητές, γενικά, θα συμβολίζονται με $\hat{\theta}$.

[Παράδειγμα] Έχουμε ήδη δείξει ότι αν $X_1, X_2, \dots, X_n$ τ.δ. από πληθυσμό που ακολουθεί $B(1, p)$, τότε η στατιστική συνάρτηση

$$\hat{p} = \frac{\sum_{i=1}^n x_i}{n}$$

είναι εκτιμητής της παραμέτρου p . Επίσης έχουμε

$$E \left[\frac{\sum_{i=1}^n x_i}{n} \right] = \frac{E[\sum_{i=1}^n x_i]}{n} = \frac{1}{n} n E[x_i] = p,$$

οπότε είναι και αμερόληπτη εκτιμήτρια της παραμέτρου p .

[Παράδειγμα] Ο δειγματικός μέσος $\bar{X}$ είναι πάντα (για κάθε κατανομή) αμερόληπτος εκτιμητής του πληθυσμιακού μέσου. Έχουμε δείξει ότι

$$E(\bar{X}) = \mu.$$

Ελάχιστη Διασπορά

- Ως σφάλμα ενός εκτιμητή T της $g(\underline{\theta})$ μπορεί να θεωρηθεί η ποσότητα $|T - g(\underline{\theta})|$ (απόλυτο σφάλμα).
- Αντ' αυτής, το τετράγωνο $[T - g(\underline{\theta})]^2$ ονομάζεται *τετραγωνικό σφάλμα*.
- Το $E[T - g(\underline{\theta})]^2$ λέγεται *μέσο τετραγωνικό σφάλμα* (MTΣ) του εκτιμητή (Mean Squared Error - MSE).

[Σημείωση]: Η ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος δεν προσφέρεται ως κριτήριο εκτίμησης, μιας και αυτή η ποσότητα ελαχιστοποιείται όταν $\theta = \theta_0$, δηλαδή στην αληθινή τιμή της θ , η οποία και είναι άγνωστη.

Το μέσο τετραγωνικό σφάλμα γράφεται

$$\begin{aligned} E[T - g(\underline{\theta})]^2 &= E[T - E(T) + E(T) - g(\underline{\theta})]^2 \\ &= E\{[T - E(T)] - [g(\underline{\theta}) - E(T)]\}^2 \\ &= E[T - E(T)]^2 + E[g(\underline{\theta}) - E(T)]^2 \\ &\quad - 2E\{[T - E(T)][g(\underline{\theta}) - E(T)]\} \\ &= Var(T) + [g(\underline{\theta}) - E(T)]^2 \\ &= Var(T) + b^2(T), \end{aligned}$$

όπου η ποσότητα

$$b(T) = g(\underline{\theta}) - E(T)$$

ονομάζεται *μεροληψία*. Συνεπώς, αν έχουμε αμερόληπτους εκτιμητές ($b(T) = 0$), τότε το MTΣ ισούται με την διασπορά.

[Σημείωση]: Αν θέσουμε λοιπόν σε αμερόληπτους εκτιμητές την συνθήκη η διασπορά να είναι όσο το δυνατό μικρότερη, τότε βάση των πιο πάνω το MTΣ θα γίνει μικρό αφού

$$E[T - g(\underline{\theta})]^2 = Var(T).$$

[Ορισμός]: Ένας εκτιμητής $T(X_1, X_2, \dots, X_n)$ της $g(\underline{\theta})$ ονομάζεται *αμερόληπτος εκτιμητής ελαχίστης διασποράς* (α.ε.ε.δ.) εάν είναι αμερόληπτος και έχει ελάχιστη διασπορά μεταξύ όλων των αμερόληπτων εκτιμητών για κάθε $\theta \in \Theta$. Συνεπώς, αν T_1 ένας άλλος εκτιμητής της ίδιας ποσότητας, τότε

$$Var(T) \leq Var(T_1), \quad \forall \underline{\theta} \in \Theta.$$

[Παράδειγμα]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. από πληθυσμό με κατανομή με $E(X_i) = \mu$ και $Var(X_i) = \sigma^2$. Από την κλάση των γραμμικών συνδυασμών της μορφής $T = \sum \alpha_i X_i$, αμερόληπτων εκτιμητών της μ , ναδειχθεί ότι ο $\bar{X}$ έχει την ελάχιστη διασπορά.

[Λύση]: Ο εκτιμητής T είναι αμερόληπτος, συνεπώς

$$E(T) = \sum \alpha_i E(X_i) = \mu,$$

οπότε παίρνουμε ότι

$$\sum \alpha_i = 1.$$

Επίσης

$$Var(T) = \sum \alpha_i^2 Var(X_i) = \sigma^2 \sum \alpha_i^2.$$

Επομένως, μας ενδιαφέρει να ελαχιστοποιήσουμε την διασπορά. Η ελαχιστοποίηση γίνεται με πολλαπλασιαστή Lagrange, και

$$\Phi(\alpha_1, \alpha_2, \dots, \alpha_n, \lambda) = \sigma^2 \sum \alpha_i^2 + \lambda(\sum \alpha_i - 1).$$

Παίρνουμε

$$\begin{aligned} \frac{\partial \Phi}{\partial \alpha_1} &= 2\alpha_1 \sigma^2 + \lambda = 0 \\ \frac{\partial \Phi}{\partial \alpha_2} &= 2\alpha_2 \sigma^2 + \lambda = 0 \\ &\vdots = \vdots \\ \frac{\partial \Phi}{\partial \alpha_n} &= 2\alpha_n \sigma^2 + \lambda = 0 \\ \frac{\partial \Phi}{\partial \lambda} &= \sum \alpha_i = 1. \end{aligned}$$

Από τα παραπάνω λαμβάνουμε ότι $\alpha_1 = \alpha_2 = \dots = \alpha_n$ και από την $\sum \alpha_i = 1$ παίρνουμε ότι $\alpha_1 = \alpha_2 = \dots = \alpha_n = \frac{1}{n}$. Συνεπώς η εκτιμητρια

$$T = \frac{1}{n} \sum X_i = \bar{X}$$

έχει την ελάχιστη διασπορά.

[Παράδειγμα]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. από πληθυσμό με κατανομή την $N(\mu, \sigma^2)$. Να δειχθεί ότι ο S^2 είναι αμερόληπτος εκτιμητής του σ^2 . Στη συνέχεια να βρεθεί το ΜΤΣ του S^2 όπως και της σ.σ. $T = cS^2$, όπου c σταθερός αριθμός. Να βρεθεί το c που ελαχιστοποιεί το ΜΤΣ της T και να γίνει σύγκριση των εκτιμητών.

[Λύση]: Είναι γνωστό ότι

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2,$$

με

$$E[\chi_{n-1}^2] = n-1 \quad \text{και} \quad Var(\chi_{n-1}^2) = 2(n-1).$$

Τότε

$$E\left[\frac{(n-1)S^2}{\sigma^2}\right] = n-1 \Rightarrow E[S^2] = \sigma^2.$$

Συνεπώς

$$E[S^2 - \sigma^2]^2 = Var(S^2).$$

Όμως

$$\begin{aligned} Var\left[\frac{(n-1)S^2}{\sigma^2}\right] &= 2(n-1) \Rightarrow \\ \frac{(n-1)^2 Var(S^2)}{\sigma^4} &= 2(n-1) \Rightarrow \\ Var(S^2) &= \frac{2\sigma^4}{n-1}, \end{aligned}$$

που είναι και το ΜΤΣ.

Ομοίως

$$E[cS^2] = c\sigma^2,$$

και συνεπώς το ΜΤΣ είναι

$$E[cS^2 - \sigma^2]^2 = E[c(S^2 - \sigma^2) - \sigma^2(1-c)]^2 = \left[\frac{2c^2}{n-1} + (1-c)^2\right] \sigma^4.$$

Η παραπάνω ποσότητα γίνεται ελάχιστη όταν

$$c^* = \frac{\nu - 1}{\nu + 1},$$

και συνεπώς ο εκτιμητής

$$T^* = c^* S^2 = \frac{\nu - 1}{\nu + 1} S^2 = \frac{\sum (X_i - \bar{X})^2}{\nu + 1}$$

έχει ΜΤΣ ίσο με $\frac{2\sigma^4}{\nu+1}$. Επίσης παρατηρούμε ότι

$$MT\Sigma(c^* S^2) < MT\Sigma(S^2).$$

Από τα παραπάνω δεν μπορούμε να αποφανθούμε ποιος εκτιμητής από τους δυο είναι προτιμότερος. Άν και οι δύο ήταν αμερόληπτοι, τότε σίγουρα ο T^* θα ήταν και ο καταλληλότερος.

Cramer-Rao κάτω φράγμα

Έστω $X_1, X_2, \dots, X_n$ τ.δ. από μονοδιάστατη τ.μ. X και θ μονοδιάστατο. Επίσης $f(x; \theta)$ η σ.π.π. του X .

[Ορισμός]: Η ποσότητα

$$I_X(\theta) = E \left[\frac{\partial \ln f(x; \theta)}{\partial \theta} \right]^2$$

λέγεται μέτρο πληροφορίας του Fisher που περιέχεται στο X για την παράμετρο θ .

Θεωρούμε τις παρακάτω συνθήκες ομαλότητας:

- Θ ανοικτό διάστημα.
- η $f(x; \theta)$ είναι θετική σε ένα σύνολο D των X , το οποίο είναι ανεξάρτητο του θ .
- για κάθε $\theta \in \Theta$ και για κάθε $x \in D$ υπάρχει η $\frac{\partial f(x, \theta)}{\partial \theta}$
- η ποσότητα

$$\int \dots_D \int f(x_1; \theta) \dots f(x_n; \theta) dx_1 \dots dx_n$$

μπορεί να παραγωγιστεί υπό το ολοκλήρωμα (αθροισμα)

- η ποσότητα

$$\int \dots_D \int T(x_1, x_2, \dots, x_n) f(x_1; \theta) \dots f(x_n; \theta) dx_1 \dots dx_n$$

μπορεί να παραγωγιστεί υπό το ολοκλήρωμα (αθροισμα)

- Οι Hodges και Lehmann έχουν δείξει ότι οι συνθήκες ομαλότητας ισχύουν για τις κατανομές $N(0, \theta)$, $N(\theta, 1)$, Διωνυμική και Poisson για κάθε αμερόληπτο εκτιμητή της παραμέτρου θ .

[Θεώρημα]: Ανισότητα Cramer-Rao

Εάν $E[T(X_1, X_2, \dots, X_n)] = \theta + b(\theta)$ και οι πιο πάνω συνθήκες ομαλότητας ισχύουν, τότε

$$\text{Var}(T) \geq \frac{[1 + b'(\theta)]^2}{n I_X(\theta)}.$$

[Απόδειξη]: [ΕΚΤΟΣ ΥΛΗΣ] Έστω X συνεχής τ.μ. με σ.π.π. $f(x, \theta)$. Τότε έχουμε

$$E(T) = \int \dots \int T(X_1, \dots, X_v) f(x_1, \theta) \dots f(x_v, \theta) dx_1 \dots dx_v = \theta + b(\theta).$$

Παραγωγίζοντας ως προς θ παίρνουμε

$$\begin{aligned} 1 + b'(\theta) &= \int \dots \int T(X_1, \dots, X_v) \frac{\partial f(x_1, \theta)}{\partial \theta} \dots f(x_v, \theta) d\mathbf{x} \\ &\quad + \int \dots \int T(X_1, \dots, X_v) f(x_1, \theta) \dots \frac{\partial f(x_v, \theta)}{\partial \theta} d\mathbf{x} \\ &= \int \dots \int T(X_1, \dots, X_v) \left[\frac{1}{f(x_1, \theta)} \frac{\partial f(x_1, \theta)}{\partial \theta} + \dots \right. \\ &\quad \left. \dots + \frac{1}{f(x_v, \theta)} \frac{\partial f(x_v, \theta)}{\partial \theta} \right] f(x_1, \theta) \dots f(x_v, \theta) d\mathbf{x} \\ &= \int \dots \int T(X_1, \dots, X_v) \sum_{i=1}^v \left[\frac{1}{f(x_i, \theta)} \frac{\partial f(x_i, \theta)}{\partial \theta} \right] \\ &\quad \prod_{i=1}^v f(x_i, \theta) d\mathbf{x} \\ &= \int \dots \int T(X_1, \dots, X_v) \sum_{i=1}^v \frac{\partial \ln f(x_i, \theta)}{\partial \theta} \prod_{i=1}^v f(x_i, \theta) d\mathbf{x} \\ &= E \left[T(X_1, \dots, X_v) \sum_{i=1}^v \frac{\partial \ln f(x_i, \theta)}{\partial \theta} \right]. \end{aligned}$$

Επίσης γνωρίζουμε ότι

$$\int \dots \int f(x_1, \theta) \dots f(x_v, \theta) d\mathbf{x} = 1,$$

και παραγωγίζοντας όπως και πριν παίρνουμε

$$E \left[\sum_{i=1}^v \frac{\partial \ln f(x_i, \theta)}{\partial \theta} \right] = 0.$$

Τότε αν

$$W = \sum_{i=1}^v \frac{\partial \ln f(x_i, \theta)}{\partial \theta}$$

έχουμε

$$1 + b'(\theta) = E[TW],$$

με

$$E[W] = 0.$$

Επειδή όμως

$$\begin{aligned} Cov(T, W) &= E(TW) - E(T)E(W) \\ &= E(TW), \end{aligned}$$

παίρνουμε ότι

$$Cov(T, W) = 1 + b'(\theta).$$

Όμως έχουμε

$$\begin{aligned} Var(W) &= \sum_{i=1}^{\nu} Var \left[\frac{\partial \ln f(x_i, \theta)}{\partial \theta} \right] \\ &= \sum_{i=1}^{\nu} E \left[\frac{\partial \ln f(x_i, \theta)}{\partial \theta} \right]^2 \\ &= \nu I_X(\theta). \end{aligned}$$

Από την ανισότητα Cauchy-Schwarz έχουμε

$$[Cov(T, W)]^2 \leq Var(T)Var(W)$$

οπότε

$$Var(T) \geq \frac{[Cov(T, W)]^2}{Var(W)}$$

και με αντικατάσταση παίρνουμε

$$Var(T) \geq \frac{[1 + b'(\theta)]^2}{\nu I_X(\theta)}.$$

[Παρατηρήσεις]:

- Αν $E[T(X_1, X_2, \dots, X_n)] = g(\theta)$ τότε η C-R ανισότητα παίρνει την μορφή

$$Var(T) \geq \frac{[g'(\theta)]^2}{nI_X(\theta)}.$$

Επίσης, αν $E[T(X_1, X_2, \dots, X_n)] = \theta$, τότε

$$Var(T) \geq \frac{1}{nI_X(\theta)}.$$

- Η ανισότητα Cramer-Rao δεν εγγυάται την ύπαρξη αμερόληπτου εκτιμητή ελάχιστης διασποράς (αεεδ). Απλά δίνει τρόπο για την αναζήτηση του ως εξής:
 1. Βρίσκουμε το κατώτερο φράγμα για τις διασπορές
 2. Αναζητούμε αμερόληπτο εκτιμητή του οποίου η διασπορά συμπίπτει με το κατώτερο φράγμα.

[Παράδειγμα]: Έστω τ.δ. $X_1, X_2, \dots, X_n$ από πληθυσμό με κατανομή $Poisson(x; \theta) = e^{-\theta} \frac{\theta^x}{x!}$. Να βρεθεί αεεδ της θ .

[Λύση]: Το κατώτερο φράγμα από την C-R ανισότητα βρίσκεται από

$$\begin{aligned} I_X(\theta) &= E \left[\frac{\partial \ln P(x, \theta)}{\partial \theta} \right]^2 \\ &= E \left[\frac{\partial}{\partial \theta} (-\theta + x \ln \theta - \ln x!) \right]^2 \\ &= E \left[-1 + \frac{x}{\theta} \right]^2 \\ &= \frac{1}{\theta^2} E [X - \theta]^2 \\ &= \frac{Var(X)}{\theta^2} \\ &= \frac{1}{\theta}. \end{aligned}$$

Οπότε

$$Var(T) \geq \frac{1}{n \frac{1}{\theta}} = \frac{\theta}{n}.$$

Ο αμερόληπτος εκτιμητής που έχουμε για το θ είναι το $\bar{X}$, οπότε:

$$Var(\bar{X}) = \frac{1}{v^2} \sum_{i=1}^v Var(X_i) = \frac{\theta}{v}.$$

Άρα, $\bar{X}$ είναι αεεδ εκτιμητής.

[Ορισμός]: Αποτελεσματικότητα (efficiency):

Έστω $U = U(\underline{X})$ ένας αμερόληπτος εκτιμητής της $g(\underline{\theta})$, ο οποίος επιτυγχάνει το C-R κάτω φράγμα της διασποράς. Τότε ο U ονομάζεται αποτελεσματικός (efficient) (με την έννοια της διασποράς). Έστω τώρα T ένας άλλος αμερόληπτος εκτιμητής με πεπερασμένη διασπορά. Τότε το πηλίκο

$$\frac{Var(U)}{Var(T)}$$

ονομάζεται σχετική αποτελεσματικότητα ή αποδοτικότητα του αμερόληπτου εκτιμητή T σε σχέση με τον (αποτελεσματικό) εκτιμητή U . Η σχετική αποτελεσματικότητα παίρνει τιμές στο $(0, 1]$.

[Παράδειγμα]: Έστω X_1, X_2, X_3 τυχαίο δείγμα από πληθυσμό με $Poisson(\theta)$. Να δειχθεί ότι η σ.σ. $U = \bar{X}$ είναι αποτελεσματικός εκτιμητής της θ και να βρεθεί η σχετική αποτελεσματικότητα του $T = \frac{X_1 + 2X_2 + 3X_3}{6}$ ως προς την U .

[Λύση]: Έχουμε ήδη δείξει ότι $\bar{X}$ είναι αμερόληπτος εκτιμητής της θ και αποτελεσματικός εκτιμητής υπό την έννοια της διασποράς. Επίσης

$$E(T) = E\left[\frac{X_1 + 2X_2 + 3X_3}{6}\right] = \theta,$$

άρα αμερόληπτος εκτιμητής, με

$$Var(T) = \frac{1}{36} [Var(X_1) + 4Var(X_2) + 9Var(X_3)] = \frac{7\theta}{18}.$$

Επίσης

$$Var(\bar{X}) = \frac{\theta}{3}.$$

Έτσι η σχετική αποτελεσματικότητα είναι

$$\frac{Var(U)}{Var(T)} = \frac{\theta/3}{7\theta/18} = \frac{6}{7}.$$

[Παρατηρήσεις]:

- Η ποσότητα $\frac{\partial \ln f(x; \theta)}{\partial \theta}$ ονομάζεται συνάρτηση σκορ (*score function*) και έχει μέση τιμή μηδέν, αφού

$$\begin{aligned} \int f(x; \theta) dx &= 1 \Rightarrow \\ \int \frac{\partial}{\partial \theta} f(x; \theta) dx &= 0 \Rightarrow \\ \int \frac{\frac{\partial}{\partial \theta} f(x; \theta)}{f(x; \theta)} f(x; \theta) dx &= 0 \Rightarrow \\ \int \frac{\partial}{\partial \theta} \ln f(x; \theta) f(x; \theta) dx &= 0 \Rightarrow \\ E \left[\frac{\partial}{\partial \theta} \ln f(x; \theta) \right] &= 0. \end{aligned}$$

Παίρνοντας μια παράγωγο ακόμα έχουμε

$$\begin{aligned} \int \left[\frac{\partial^2}{\partial \theta^2} \ln f(x; \theta) f(x; \theta) + \left(\frac{\partial}{\partial \theta} \ln f(x; \theta) \right) \left(\frac{\partial}{\partial \theta} f(x; \theta) \right) \right] dx &= 0 \Rightarrow \\ E \left[\frac{\partial^2}{\partial \theta^2} \ln f(x; \theta) \right] + E \left[\left(\frac{\partial}{\partial \theta} \ln f(x; \theta) \right)^2 \right] &= 0 \Rightarrow \\ -E \left[\frac{\partial^2}{\partial \theta^2} \ln f(x; \theta) \right] &= E \left[\left(\frac{\partial}{\partial \theta} \ln f(x; \theta) \right)^2 \right], \end{aligned}$$

συνεπώς η πληροφορία του *Fisher* γράφεται

$$I_X(\theta) = -E \left[\left(\frac{\partial}{\partial \theta} \ln f(x; \theta) \right) \right].$$

[Θεώρημα]: Στην ανισότητα C-R το ίσον επιτυγχάνεται εάν

$$\frac{\partial}{\partial \theta} \ln \prod_{i=1}^{\nu} f(x_i; \theta) = k(\theta, \nu)[U - g(\theta)].$$

Τότε, ο εκτιμητής U είναι αποτελεσματικός εκτιμητής του $g(\theta)$.

- Συνέπεια του πιο πάνω θεωρήματος είναι ότι εάν $x_1, x_2, \dots, x_n$ τ.δ. από κατανομή $f(x; \theta)$, τότε εάν $\hat{\theta}$ είναι ΕΜΠ του θ και $\delta = \delta(\underline{x})$ μία αποτελεσματική εκτιμήτρια του θ , τότε

$$\hat{\theta} = \delta.$$

- Η ανισότητα γενικεύεται όταν $\underline{\theta} \in \Theta \subset \mathcal{R}^k$, $k > 1$.

[Παράδειγμα] Έστω $X_1, X_2, \dots, X_n$ τ.δ. από $N(\mu, \theta)$, όπου μ γνωστό. Ποιός ο αεεδ εκτιμητής του θ ;

[Λύση]: Οι συνθήκες ομαλότητας ισχύουν (από προηγούμενη παρατήρηση) για κάθε αμερόληπτο εκτιμητή της θ . Υπολογίζουμε το C-R κ.φ. Έχουμε

$$f(x; \theta) = \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{1}{2\theta}(x-\mu)^2},$$

και

$$\ln f(x; \theta) = -\frac{1}{2} \ln 2\pi - \frac{1}{2} \ln \theta - \frac{1}{2\theta}(x-\mu)^2.$$

Συνεπώς

$$\frac{\partial}{\partial \theta} \ln f(x; \theta) = -\frac{1}{2\theta} + \frac{1}{2\theta^2}(x-\mu)^2,$$

και

$$\begin{aligned} \left[\frac{\partial}{\partial \theta} \ln f(x; \theta) \right]^2 &= \frac{1}{4\theta^2} + \frac{1}{4\theta^4}(x-\mu)^4 - \frac{1}{2\theta^3}(x-\mu)^2 \\ &= \frac{1}{4\theta^2} + \frac{1}{4\theta^2} \left(\frac{x-\mu}{\sqrt{\theta}} \right)^4 - \frac{1}{2\theta^2} \left(\frac{x-\mu}{\sqrt{\theta}} \right)^2. \end{aligned}$$

Οπότε

$$E \left[\frac{\partial}{\partial \theta} \ln f(x; \theta) \right]^2 = \frac{1}{4\theta^2} + \frac{1}{4\theta^2} E \left(\frac{x-\mu}{\sqrt{\theta}} \right)^4 - \frac{1}{2\theta^2} E \left(\frac{x-\mu}{\sqrt{\theta}} \right)^2.$$

Όμως $\frac{x-\mu}{\sqrt{\theta}} \sim N(0, 1)$ και $\left(\frac{x-\mu}{\sqrt{\theta}} \right)^2 \sim \chi_1^2$. Έτσι $E \left(\frac{x-\mu}{\sqrt{\theta}} \right)^2 = 1$ και $Var \left(\frac{x-\mu}{\sqrt{\theta}} \right)^2 = 2$, και βάσει της σχέσης $Var(X) = E(X^2) - E(X)^2$ παίρνουμε ότι

$$E \left(\frac{x-\mu}{\sqrt{\theta}} \right)^4 = 3.$$

Έτσι

$$\begin{aligned} I_X(\theta) &= E \left[\frac{\partial}{\partial \theta} \ln f(x; \theta) \right]^2 \\ &= \frac{1}{4\theta^2} + \frac{3}{4\theta^2} - \frac{1}{2\theta^2} \\ &= \frac{1}{2\theta^2}. \end{aligned}$$

Άρα το C-R κ.φ. είναι

$$C - R = \frac{1}{\nu I_X(\theta)} = \frac{2\theta^2}{\nu}.$$

Ομοίως (πιο εύκολα) έχουμε:

$$\begin{aligned} I_X(\theta) &= -E \left[\frac{\partial^2}{\partial \theta^2} \ln f(x; \theta) \right] \\ &= -E \left[\frac{1}{2\theta^2} - \frac{1}{\theta^3} (x - \mu)^2 \right] \\ &= -\frac{1}{2\theta^2} + \frac{1}{\theta^3} E[(x - \mu)^2] \\ &= -\frac{1}{2\theta^2} + \frac{1}{\theta^3} V[x] \\ &= -\frac{1}{2\theta^2} + \frac{1}{\theta^3} \theta \\ &= \frac{1}{2\theta^2}. \end{aligned}$$

Παίρνουμε τώρα την σ.σ. $T = \frac{1}{\nu} \sum (x_i - \mu)^2$ και έχουμε

$$E(T) = \frac{1}{\nu} \sum E(x_i - \mu)^2 = \frac{1}{\nu} \nu \theta = \theta,$$

άρα η σ.σ. T είναι αμερόληπτος εκτιμητής της θ . Επίσης

$$Var(T) = \frac{1}{\nu^2} \sum Var(x_i - \mu)^2.$$

Επειδή όμως $\left(\frac{x - \mu}{\sqrt{\theta}} \right)^2 \sim \chi_1^2$ έπεται ότι

$$\begin{aligned} Var \left(\frac{x_i - \mu}{\sqrt{\theta}} \right)^2 &= 2 \Rightarrow \\ \frac{1}{\theta^2} Var(x_i - \mu)^2 &= 2 \Rightarrow \\ Var(x_i - \mu)^2 &= 2\theta^2. \end{aligned}$$

Επομένως

$$Var(T) = \frac{1}{\nu^2} \nu 2\theta^2 = \frac{2\theta^2}{\nu},$$

το οποίο και παρατηρούμε ότι ισούται με το C-R κ.φ. και συνεπώς ο T είναι αποτελεσματικός εκτιμητής της παραμέτρου θ .

Αν τώρα παίρναμε το στατιστικό $S^2 = \frac{1}{\nu-1} \sum (x_i - \bar{X})^2$, γνωρίζουμε ότι $E(S^2) = \sigma^2 = \theta$, άρα α.ε. της θ με

$$Var(S^2) = \frac{2\theta^2}{\nu-1},$$

και συνεπώς: $Var(T) < Var(S^2)$.

Θα μπορούσαμε επίσης να εργαστούμε με βάση το προηγούμενο θεώρημα. Έτσι έχουμε

$$\begin{aligned} \frac{\partial}{\partial \theta} \ln \prod_{i=1}^{\nu} f(x_i; \theta) &= \frac{\partial}{\partial \theta} \left[-\frac{\nu}{2} \ln 2\pi\theta - \frac{1}{2\theta} \sum (x_i - \mu)^2 \right] \\ &= -\frac{\nu}{2\theta} + \frac{1}{2\theta^2} \sum (x_i - \mu)^2 \\ &= \frac{\nu}{2\theta^2} \left[\frac{1}{\nu} \sum (x_i - \mu)^2 - \theta \right] \\ &= k(\theta - \nu)[U - \theta]. \end{aligned}$$

Άρα $\frac{1}{\nu} \sum (x_i - \mu)^2$ αποτελεσματικός εκτιμητής της θ .

Στατιστική Επάρκεια (Sufficient Statistics)

- Η γνώση των τιμών ενός τ.δ. δεν μας δίνει τίποτε περισσότερο σχετικά με την εκτίμηση της παραμέτρου θ από ότι μας δίνει μια επαρκής στατιστική συνάρτηση.
- Επίσης, σημαντική είναι και η συμπτύξη της πληροφορίας (χαριν ευκολίας), χωρίς όμως να χάνουμε πληροφορία για τον άγνωστο πληθυσμό που περιέχετε στο αρχικό δείγμα.
- Συνεπώς, απορρίπτουμε το μέρος των δεδομένων που δεν περιέχουν πληροφορίες για την θ .

[Ορισμός] Έστω η τ.μ. $\underline{X} = (X_1, X_2, \dots, X_n)$ με σ.π.π. $f(x; \theta)$ και $\underline{T} = (T_1(\underline{X}), T_2(\underline{X}), \dots, T_n(\underline{X}))$ μια σ.σ. της $\underline{X}$. Τότε λέμε ότι η $\underline{T}$ είναι στατιστικά επαρκής ως προς $\underline{\theta} \in \Theta$ εάν η δεσμευμένη κατανομή της $\underline{X}$ όταν $\underline{T} = \underline{t}$ είναι ανεξάρτητη του $\underline{\theta}$ για όλες τις τιμές του $\underline{t}$ για τις οποίες ορίζεται η δεσμευμένη κατανομή.

- Η $\underline{T}$ παίρνει την ίδια τιμή για δύο διαφορετικά σημεία στο δειγματικό χώρο μόνο όταν αυτά τα δύο σημεία έχουν ανάλογες πιθανοφάνειες.

$$T(x) = T(z) \Rightarrow L(\theta, x) \propto L(\theta, z) \quad \text{για κάθε } \theta.$$

- Η στατιστική συνάρτηση $\underline{X} = (X_1, X_2, \dots, X_n)$ είναι πάντοτε επάρκης για την παράμετρο θ .

◇ *Διαισθητικά*: η $T(\cdot)$ περιέχει όλη την πληροφορία για το θ που υπάρχει στο δείγμα. Μέ άλλα λόγια, το να ξέρουμε πίο πολλά για το x από το ότι $T(X) = t$, δεν βοηθά στην εκτίμηση του θ .

[Ορισμός]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. από πληθυσμό με σ.π.π. $f(x; \theta)$. Μια σ.σ. $\underline{T} = T(\underline{X})$ είναι επαρκής για την θ εάν και μόνον εάν για κάθε άλλη σ.σ. $\underline{U}$ η κατανομή $\underline{U}|\underline{T}$ είναι ανεξάρτητη της $\underline{\theta}$. Ο συγκεκριμένος ορισμός είναι χρήσιμος για να δείξουμε ότι μια σ.σ. δεν είναι επαρκής.

[Παρατήρηση]: Κάθε αμφιμονοσήματος μετασχηματισμός μιας επαρκούς σ.σ. είναι επίσης επαρκής σ.σ.

[Παράδειγμα] Έστω n -δοκιμές Bernoulli, δηλαδή $X_1, X_2, \dots, X_n$ με $X_i \sim B(1, \theta)$. Επίσης

$$T = X_1 + X_2 + \dots + X_n,$$

όπου $T \sim B(n, \theta)$. Το πείραμα εκτελείται και λαμβάνουμε

$$x_1 + x_2 + \dots + x_n = t.$$

Θέλουμε να εκτιμήσουμε την θ βάση των x_i .

[?] Πότε γνωρίζουμε περισσότερα, όταν ξέρουμε το t (αριθμό επιτυχιών) ή όταν γνωρίζουμε τις επιτυχίες και τότε αυτές έγιναν; Άρα, είναι τα $x_1, x_2, \dots, x_n$ και το t ισοδύναμα από πλευράς πληροφορίας για την θ ;

Σκεφτόμαστε ως εξής:

- όταν γνωρίζουμε το t , ας υπολογίσουμε την πιθανότητα καθενός από τους $\binom{n}{t}$ τρόπους με τους οποίους οι t επιτυχίες μπορούν να συμβούν.
- αν θεωρήσουμε ότι υπάρχουν τιμές της θ που δίνουν μεγαλύτερη πιθανότητα σε κάποιους συνδυασμούς επιτυχιών από άλλους, τότε δεχόμαστε ότι γνώση των $x_1, x_2, \dots, x_n$ είναι προτιμότερη από απλή γνώση του t .
- εάν όμως με γνώση του t όλα τα αποτελέσματα έχουν την ίδια πιθανότητα, τότε προφανώς η γνώση των $x_1, x_2, \dots, x_n$ δεν είναι απαραίτητη.

Έχουμε

$$\begin{aligned} P(X_1 = x_1, \dots, X_n = x_n | T = t) &= \frac{P(\underline{X} = \underline{x}, T = t)}{P(T = t)} \\ &= \frac{P(\underline{X} = \underline{x})}{P(T = t)} \\ &= \frac{\prod_{i=1}^n P(X_i = x_i)}{P(T = t)} \\ &= \frac{\theta^{\sum_{i=1}^n x_i} (1 - \theta)^{n - \sum_{i=1}^n x_i}}{\binom{n}{t} \theta^t (1 - \theta)^{n-t}} \\ &= \frac{1}{\binom{n}{t}}. \end{aligned}$$

Άρα, όλα τα αποτελέσματα έχουν την ίδια πιθανότητα όταν το t είναι γνωστό. Συνεπώς η σ.σ. $T = \sum_{i=1}^n X_i$ περιέχει όλες τις πληροφορίες για την θ που περιέχονται στο αρχικό δείγμα.

[Παρατήρηση]:

Οι σ.σ. $T_1 = (X_1, \sum_{i=2}^n X_i)$, $T_2 = (X_1, X_2, \sum_{i=3}^n X_i), \dots$ κλπ, είναι επίσης επαρκείς σ.σ.

[?]: Ποιά όμως προτιμάμε?

$\Rightarrow$ Την $T = \sum x_i$, γιατί "συμπυκνώνει" τα δεδομένα όσο το δυνατόν περισσότερο (Ελάχιστη Επαρκής).

Το παραγοντικό κριτήριο - Neyman's factorization criterion

Αν η $T(\cdot)$ είναι επαρκής, τότε η $L(\theta)$ εξαρτάται από το X μόνο μέσω της $T(X)$. Άρα, υπάρχει συνάρτηση $g(\cdot)$ τέτοια ώστε

$$L(\theta) \propto g(T(x), \theta).$$

Αφού όμως $L(\theta) = f(X|\theta)$, τότε συνεπάγεται ότι $f(x|\theta)/g(T(x), \theta)$ δεν εξαρτάται από το θ . Αν λοιπόν συμβολίσουμε με $h(x) = f(x|\theta)/g(T(x), \theta)$, τότε

$$f(x|\theta) = h(x)g(T(x), \theta), \quad (1)$$

για κάποιες συναρτήσεις $h(\cdot)$ και $g(\cdot)$. Σε αυτή την περίπτωση η $L(\theta)$ είναι συνάρτηση των δεδομένων X μόνο μέσω της $T(X)$, οπότε $T(X)$ είναι επαρκής. ■

[Θεώρημα]: Η συνάρτηση $T(\cdot)$ είναι επαρκής αν και μονον αν η $f(x|\theta)$ μπορεί να γραφεί στην μορφή (1).

[Απόδειξη]: (Για διακριτό δειγματικό χώρο)

($\Leftarrow$) Έστω $f(x|\theta) = P_\theta(X = x)$ μπορεί να γραφεί στην μορφή (1) και $T(x) = t$. Τότε

$$P_\theta(X = x) = P_\theta(T(x) = t)P_\theta(X = x|T(x) = t),$$

όπου

$$\begin{aligned} P_\theta(X = x|T(x) = t) &= \frac{P_\theta(X = x, T(x) = t)}{P_\theta(T(x) = t)} \\ &= \frac{P_\theta(X = x)}{P_\theta(T(x) = t)} \\ &= \frac{g(T(x), \theta)h(x)}{\sum_{x:T(x)=t} g(T(x), \theta)h(x)} \\ &= \frac{g(t, \theta)h(x)}{\sum_{x:T(x)=t} g(t, \theta)h(x)} \\ &= \frac{h(x)}{\sum_{x:T(x)=t} h(x)} \end{aligned}$$

το οποίο και δέν εξαρτάται από το θ .

($\Rightarrow$) Αντιστοίχως, αν η $T(\cdot)$ είναι επαρκής και $T(x) = t$, τότε

$$P_{\theta}(X = x) = P_{\theta}(T(x) = t)P_{\theta}(X = x|T(x) = t),$$

όπου λόγω επάρκειας ο δεύτερος όρος δεν εξαρτάται από το θ . Οπότε, συμβολίζουμε τον πρώτο όρο με $g(T(x), \theta)$ και τον δεύτερο με $h(x)$ και παίρνουμε την (1). ■

[Παράδειγμα]: Έστω $X_1, X_2, \dots, X_n \sim \text{Poisson}(\lambda)$, όπου το λ προς εκτίμηση. Έχουμε

$$L(\lambda) = f(x|\lambda) = \prod_{i=1}^n e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} = \lambda^{\sum_i x_i} \frac{e^{-n\lambda}}{\prod_{i=1}^n x_i!},$$

όπου $g(T(x), \theta) = \lambda^{\sum_i x_i} e^{-n\lambda}$ και $h(x) = \prod_{i=1}^n x_i!$. Άρα, $t = \sum_i x_i$ είναι επαρκής στατιστική συνάρτηση.

[Παρατήρηση]: Μια επαρκής στατιστική συνάρτηση δεν είναι μοναδική. Αν $T(x)$ είναι επαρκής σ.σ. τότε και οι συναρτήσεις $T(x)/n$ και $\log T(x)$ είναι επίσης επαρκής σ.σ. Η ΕΜΠ βρίσκεται από

$$\left. \frac{d}{d\lambda} \log L(\lambda) \right|_{\lambda=\hat{\lambda}} = \frac{\sum_i x_i}{\hat{\lambda}} - n = 0 \Rightarrow \hat{\lambda} = \bar{X}.$$

[Πόρισμα]: Έστω $\psi : \mathbb{R} \rightarrow \mathbb{R}$ και $\omega : \mathbb{R} \rightarrow \mathbb{R}$ δύο συναρτήσεις "1-1" (μονοσήμαντα ορισμένες) ώστε να υπάρχουν οι αντίστροφες συναρτήσεις ψ^{-1}, ω^{-1} . Αν η σ.σ. $\underline{T}$ είναι επαρκής για το $\underline{\theta}$, τότε

- η $\underline{T}_1 = \psi(\underline{T})$ είναι επαρκής για το $\underline{\theta}$
- η $\underline{T}$ είναι επαρκής για το $\underline{\theta}_1 = \omega(\underline{\theta})$

[Απόδειξη]:

- Επειδή η $\underline{T}$ επαρκής για το $\underline{\theta}$

$$\begin{aligned} f(\underline{x}; \theta) &= g(\underline{T}(\underline{x}); \underline{\theta})h(\underline{x}) \\ &= g(\underline{T}(\underline{x}); \omega^{-1}(\underline{\theta}_1))h(\underline{x}) \\ &= g_2(\underline{T}(\underline{x}); \underline{\theta}_1)h(\underline{x}) \end{aligned}$$

Άρα η $\underline{T}$ επαρκής για το $\underline{\theta}_1$.

- Επίσης

$$\begin{aligned} f(\underline{x}; \theta) &= g(\underline{T}(\underline{x}); \underline{\theta})h(\underline{x}) \\ &= g(\psi^{-1}\underline{T}_1(\underline{x}); \underline{\theta})h(\underline{x}) \\ &= g_1(\underline{T}_1(\underline{x}); \underline{\theta})h(\underline{x}) \end{aligned}$$

Άρα η $\underline{T}_1$ επαρκής για το $\underline{\theta}$.

[Παράδειγμα]: Εάν T επαρκής για το θ , τότε

- η σ.σ. $\bar{X} = \frac{T}{v}$ είναι επαρκής για το θ . Τα $\bar{x}^2, |\bar{x}|$ δεν γνωρίζω αν είναι επαρκής για το θ .
- η T είναι επαρκής για τα $\theta^3, 2\theta, k\theta, \log \theta, e^\theta$. Όχι όμως για το θ^2 .

[Παρατήρηση]: Η ΕΜΠ είναι πάντα συνάρτηση της $T(x)$. ■

[Παράδειγμα]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. από $N(\theta, \sigma^2)$, όπου σ^2 γνωστό. Να βρεθεί επαρκής σ.σ. για την θ .

[Λύση]: Έχουμε

$$f(\underline{x}; \theta) = \prod_{i=1}^n f(x_i; \theta) = \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \theta)^2\right\}.$$

Αλλά όμως

$$\sum_{i=1}^n (x_i - \theta)^2 = \sum_{i=1}^n x_i^2 + n\theta^2 - 2\theta \sum_{i=1}^n x_i,$$

και επομένως

$$\begin{aligned} f(\underline{x}; \theta) &= \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp\left\{-\frac{n\theta^2}{2\sigma^2} + \frac{\theta}{\sigma^2} \sum_{i=1}^n x_i\right\} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n x_i^2\right\} \\ &= g\left(\sum_{i=1}^n x_i, \theta\right) h(\underline{x}), \end{aligned}$$

όπου

$$g = \exp\left\{-\frac{n\theta^2}{2\sigma^2} + \frac{\theta}{\sigma^2} \sum_{i=1}^n x_i\right\}$$

και

$$h = \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^n \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n x_i^2\right\}.$$

Συνεπώς $\sum_{i=1}^n x_i$ επαρκής σ.σ. για την θ . Επίσης ο $\bar{X}$ είναι επαρκής σ.σ. για την θ .

[Θεώρημα **Rao-Blackwell**]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. από πληθυσμό με σ.π.π. $f(x; \theta)$ και $\underline{T} = (T_1(\underline{X}), \dots, T_x(\underline{X}))'$ μια επαρκής σ.σ. για την θ . Έστω ακόμα $U = U(\underline{X})$ μια αμερόληπτη εκτιμήτρια του θ . Τότε η σ.σ. $U^* = E(U|T)$ είναι αμερόληπτη εκτιμήτρια του θ και

$$Var(U^*) \leq Var(U),$$

για κάθε $\theta \in \Theta$.

[Απόδειξη]: Με βάση τη σχέση $E(Y) = E[E(Y|X)]$ έχουμε

$$E[U^*] = E[E(U|T)] = E[U] = \theta.$$

Άρα $U^* = E[U|T]$ αμερόληπτη εκτιμήτρια της θ . Η U^* δεν εξαρτάται από τη θ λόγω επάρκειας της T . Επίσης

$$Var(U^*) = Var E[U|T],$$

και

$$Var(U) = Var E[U|T] + E Var[U|T].$$

Όμως $E Var[U|T] \geq 0$ και έπεται ότι

$$Var(U^*) \leq Var(U).$$

[Παράδειγμα]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. από $Bin(1, \theta)$. Παρατηρούμε ότι $U = X_1$ είναι αμερόληπτη σ.σ. της θ και η $T = \sum_{i=1}^n x_i$ είναι επαρκής σ.σ. για το θ . Άς εφαρμόσουμε το θεώρημα R-B για την εύρεση άλλου αμερόληπτου εκτιμητή με διακύμανση μικρότερη του U .

Υπολογίζουμε

$$E(X_1|T=t) = 0 \times P(X_1=0|T=t) + 1 \times P(X_1=1|T=t).$$

Όμως

$$\begin{aligned}
 P(X_1 = 0|T = t) &= \frac{P(X_1 = 0, \sum_{i=1}^{\nu} x_i = t)}{P_T(t)} \\
 &= \frac{P(X_1 = 0, X_2 + \dots + X_{\nu} = t)}{P_T(t)} \\
 &= \frac{P(X_1 = 0)P(X_2 + \dots + X_{\nu} = t)}{P_T(t)} \\
 &= \frac{(1 - \theta) \binom{\nu-1}{t} \theta^t (1 - \theta)^{\nu-1-t}}{\binom{\nu}{t} \theta^t (1 - \theta)^{\nu-t}} \\
 &= \frac{\nu - t}{\nu}.
 \end{aligned}$$

Όμοίως $P(X_1 = 1|T = t) = \frac{t}{\nu}$. Άρα

$$E(X_1|T = t) = \frac{t}{\nu}.$$

Επομένως επιλέγουμε

$$U^* = \frac{T}{n} = \frac{\sum_{i=1}^{\nu} X_i}{\nu} = \bar{X}$$

ο οποίος είναι αμερόληπτος εκτιμητής της θ και έχει

$$Var(U^*) \leq Var(U),$$

αφού $Var(U) = \theta(1 - \theta)$ και $Var(U^*) = \frac{\theta(1-\theta)}{\nu}$.

Ελάχιστη Επαρκής - Minimal Sufficient Statistic

- Είναι αλήθεια ότι για κάθε πρόβλημα υπάρχουν περισσότερες από μια επαρκείς σ.σ.
- Η στατιστική συνάρτηση $\underline{X} = (X_1, X_2, \dots, X_n)$ είναι πάντοτε επαρκής για την παράμετρο θ κι αυτό γιατί

$$f(\underline{x}|\theta) = f(T(\underline{x})|\theta)h(x),$$

όπου θέτουμε $h(x) = 1$ και $T(\underline{x}) = \underline{x}$.

Λόγω του πλήθους των επαρκών σ.σ. πάντα θα υπάρχει το ερώτημα αν υπάρχει κάποια που είναι καλύτερη από όλες τις υπόλοιπες.

[Ορισμός]: Μια επαρκής σ.σ. $T(\underline{X})$ θα ονομάζεται ελάχιστη επαρκής (minimal sufficient statistic) αν και μόνον αν για κάθε άλλη επαρκή σ.σ. $T'(\underline{X})$, η $T(\underline{X})$ γράφεται ως συνάρτηση της $T'(\underline{X})$.

Πληρότητα (Completeness)

[Ορισμός]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. από πληθυσμό με σ.π.π. $f(x; \theta)$. Μια σ.σ. $\underline{T} = T(\underline{X})$ ή ακριβέστερα οικογένεια κατανομών της $\underline{T}$ είναι πλήρης εάν και μόνον εάν για κάθε συνάρτηση $h(\underline{T})$ του $\underline{T}$ μόνον, η σχέση

$$E[h(\underline{T})] = 0$$

για κάθε $\theta \in \Theta$, συνεπάγεται

$$h(\underline{t}) = 0$$

για κάθε $\underline{t}$, εκτός πιθανόν από σημεία για τα οποία η πιθανότητα είναι μηδέν για κάθε $\theta \in \Theta$.

[Παρατήρηση]: Με άλλα λόγια δεν υπάρχουν μη μηδενικοί αμερόληπτοι εκτιμητές του μηδενός ή η $\underline{T}$ είναι πλήρης εάν και μόνον εάν ο μόνος αμερόληπτος εκτιμητής του μηδενός, ο οποίος είναι συνάρτηση του $\underline{T}$ είναι ο εκτιμητής 0 με πιθανότητα 1.

[Παράδειγμα]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. από πληθυσμό με κατανομή $Poisson(\theta)$. Να δειχθεί ότι $T = \sum_{i=1}^n X_i$ είναι μια πλήρης σ.σ.

[Λύση]: Αρκεί να δείξουμε ότι

$$E[h(T)] = 0 \dots \forall \theta \Rightarrow h(T) = 0 \dots \forall t.$$

Είναι γνωστό ότι η σ.π. της T είναι $p(t) = e^{-n\theta} \frac{(n\theta)^t}{t!}$. Επομένως έχουμε

$$\begin{aligned} E[h(T)] &= \sum_{t=0}^{\infty} h(t) e^{-n\theta} \frac{(n\theta)^t}{t!} \\ &= e^{-n\theta} \sum_{t=0}^{\infty} h(t) \frac{(n\theta)^t}{t!} \\ &= h(0) + h(1) \frac{n\theta}{1!} + h(2) \frac{(n\theta)^2}{2!} + \dots \\ &= 0 \quad \forall \theta. \end{aligned}$$

Βάση γνωστού θεωρήματος που λέει ότι όταν η απειροσειρά

$$\alpha_0 + \alpha_1 z + \alpha_2 z^2 + \alpha_3 z^3 + \dots$$

συγκλίνει στο μηδέν για όλες τις τιμές του z σε ένα δοσμένο διάστημα, τότε

$$\alpha_0 = \alpha_1 = \alpha_2 = \alpha_3 = \dots = 0.$$

Οπότε, έχουμε

$$h(0) = h(1) = h(2) = \dots = 0$$

ή $h(t) = 0$ για κάθε t . Άρα $T = \sum X_i$ είναι πλήρης σ.σ.

[Παράδειγμα]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. από πληθυσμό με κατανομή $f(x; \theta)$ και $T = T(\underline{X})$ τέτοιο ώστε $T \sim U(0, \theta)$. Ναδειχθεί ότι η T είναι πλήρης.

[Λύση]: Αρκεί να δείξουμε ότι $E[h(T)] = 0$ για κάθε θ συνεπάγεται $h(t) = 0$ για κάθε t . Έχουμε

$$E[h(T)] = \int_0^\theta h(t) \frac{1}{\theta} dt = 0$$

ή

$$\int_0^\theta h(t) dt = 0 \quad \forall \theta \in \Theta.$$

Αν παραγωγίσουμε ως προς θ και τα δύο μέλη έχουμε

$$h(\theta) = 0$$

για κάθε θ . Άρα πλήρης.

[Θεώρημα **Lehmann-Scheffé**]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. από πληθυσμό με σ.π.π. $f(x; \theta)$. Έστω επίσης $\underline{T} = T(\underline{X})$ μια επαρκής και πλήρης σ.σ. και $U = U(T)$ μια αμερόληπτη εκτιμήτρια του $g(\underline{\theta})$, η οποία είναι συνάρτηση του $\underline{T}$. Τότε U είναι αεεδ εκτιμητής του $g(\underline{\theta})$.

[Απόδειξη]: Έστω $U^* = U^*(\underline{T})$ ένας άλλος αμερόληπτος εκτιμητής της $g(\underline{\theta})$. Τότε $E[U^*] = g(\underline{\theta})$ και συνεπώς

$$E[U^* - U] = E[U^*(\underline{T}) - U(\underline{T})] = 0,$$

και λόγω της πληρότητας παίρνουμε

$$U^* - U = 0 \Rightarrow U^* = U.$$

Άρα υπάρχει ένας μοναδικός αμερόληπτος εκτιμητής της $g(\underline{\theta})$ που να είναι συνάρτηση της $\underline{T}$.

Έστω τώρα U^{**} ένας άλλος οποιοσδήποτε αμερόληπτος εκτιμητής της $g(\underline{\theta})$. Τότε βάση του Θεωρ. R-B

$$Var E(U^{**}|\underline{T}) \leq Var(U^{**})$$

αλλά όμως

$$E(U^{**}|\underline{T}) = U(\underline{T}),$$

διότι υπάρχει μόνον ένας αμερόληπτος εκτιμητής που να είναι συνάρτηση του $\underline{T}$. Άρα $U(\underline{T})$ αεεδ εκτιμητής της $g(\underline{\theta})$. ■

[Παράδειγμα]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. από πληθυσμό με σ.π. $Bin(1, \theta)$. Να βρεθεί
(α) ο αεεδ εκτιμητής της θ και (β) ο αεεδ εκτιμητής της $\theta(1 - \theta)$.

[Λύση]: (α) Αρχικά βρίσκουμε μια επαρκή σ.σ. για την θ . Έχουμε

$$P(\underline{x}, \theta) = \theta^{\sum x_i} (1 - \theta)^{n - \sum x_i}.$$

Επομένως $T = \sum x_i$ επαρκής σ.σ. μιας και μπορούμε να γράψουμε

$$P(\underline{x}, \theta) = g(T, \theta)h(\underline{x}),$$

όπου $g(T, \theta) = \theta^T (1 - \theta)^{n - T}$ και $h(\underline{x}) = 1$. Θα δείξουμε επίσης ότι η σ.σ. T είναι πλήρης. Αρχεί να δείξουμε

$$E[h(T)] = 0 \quad \forall \theta \quad \Rightarrow \quad h(T) = 0 \quad \forall t.$$

Έχουμε

$$E[h(T)] = \sum_{t=0}^n h(t)P(T=t),$$

και επειδή $T \sim Bin(n, \theta)$ έχουμε

$$E[h(T)] = \sum_{t=0}^n h(t) \binom{n}{t} \theta^t (1 - \theta)^{n-t} = (1 - \theta)^n \sum_{t=0}^n h(t) \binom{n}{t} \left(\frac{\theta}{1 - \theta} \right)^t$$

ή

$$E[h(T)] = \sum_{t=0}^n h(t) \binom{n}{t} \alpha^t,$$

όπου $\alpha = \frac{\theta}{1 - \theta}$. Συνεπώς

$$h(0) \binom{n}{0} \alpha^0 + h(1) \binom{n}{1} \alpha^1 + \dots + h(n) \binom{n}{n} \alpha^n = 0 \quad \forall \alpha.$$

Επειδή αυτό είναι πολυώνυμο εκ ταυτότητας ίσο με μηδέν, οι συντελεστές α^t θα είναι καθένας ίσος με μηδέν. Δηλαδή

$$h(0) = h(1) = \dots = h(n) = 0$$

ή

$$h(t) = 0 \quad \forall t.$$

Άρα T είναι πλήρης. Τώρα πρέπει να βρούμε αμερόληπτο εκτιμητή ο οποίος να είναι συνάρτηση του T . Ός γνωστό $\bar{X}$ αμερόληπτος της θ , και

$$U = \bar{X} = \frac{T}{\nu}.$$

Άρα U αμερόληπτος και συνάρτηση του T , επομένως αεεδ της θ .

(β) Γνωρίζουμε ότι η διασπορά σ^2 είναι $\theta(1-\theta)$ και ως γνωστο το S^2 είναι αμερόληπρος εκτιμητής του σ^2 . Αρκεί λοιπόν να δείξουμε ότι S^2 είναι συνάρτηση του T , πλήρης και επαρκής σ.σ. Όμως

$$\begin{aligned} S^2 &= \frac{1}{\nu-1} \sum (X_i - \bar{X})^2 \\ &= \frac{1}{\nu-1} \left(\sum X_i^2 - \frac{(\sum X_i)^2}{\nu} \right) \\ &= \frac{1}{\nu-1} \left(\sum X_i - \frac{(\sum X_i)^2}{\nu} \right) \\ &= \frac{1}{\nu-1} \left(T - \frac{T^2}{\nu} \right), \end{aligned}$$

αφού $\sum X_i^2 = \sum X_i$ μιάς και $X = 0, 1$. Άρα το S^2 είναι συνάρτηση του T .

[Παράδειγμα]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. από πληθυσμό με σ.π. $Poisson(\theta)$. Να βρεθεί (α) ο αεεδ εκτιμητής της θ , (β) ο αεεδ εκτιμητής της $g(\theta) = \theta^2$ και (γ) ο αεεδ εκτιμητής της $g(\theta) = e^{-\theta}$.

[Λύση]: (α) Βρίσκουμε πρώτα μια επαρκή σ.σ. Έχουμε

$$P(\underline{x}, \theta) = e^{-n\theta} \frac{\theta^{\sum x_i}}{\prod x_i!} = e^{-n\theta} \theta^{\sum x_i} \frac{1}{\prod x_i!},$$

όπου $g(\underline{x}, \theta) = e^{-n\theta} \theta^{\sum x_i}$ και $h(\underline{x}) = \frac{1}{\prod x_i!}$. Άρα $T = \sum x_i$ επαρκής σ.σ. για την θ . Επίσης T πλήρης (από προηγούμενο παράδειγμα). Επίσης $\bar{X}$ αμερόληπτος της θ και συνάρτηση του T

$$\bar{X} = \frac{T}{n},$$

και συνεπώς $\bar{X}$ αεεδ της θ .

(β) Αρκεί να βρούμε αμερόληπτο εκτιμητή της θ^2 , που να είναι συνάρτηση της σ.σ. T . Γνωρίζουμε $E(X_i) = \theta$ και $Var(X_i) = \theta$ και $T = \sum X_i \sim P(n\theta)$, όπου $E(T) = n\theta$ και $Var(T) = n\theta$. Όμως

$$Var(T) = E(T^2) - E(T)^2 = n\theta \Rightarrow E(T^2) = n\theta + (n\theta)^2.$$

Από αυτή τη σχέση παίρνουμε

$$E(T^2) - E(T) = n\theta + (n\theta)^2 - n\theta \Rightarrow E\left(\frac{T^2 - T}{n^2}\right) = \theta^2.$$

Άρα $U = E\left(\frac{T^2 - T}{n^2}\right)$ αμερόληπτος εκτιμητής της θ^2 και συνάρτηση της T . Επομένως αεεδ εκτιμητής της θ^2 .

Ας υπολογίσουμε την διασπορά του U . Έχουμε

$$\begin{aligned} Var(U) &= Var\left[\frac{T^2 - T}{n^2}\right] \\ &= \frac{1}{n^4} Var[T^2 - T] \\ &= \frac{1}{n^4} \left\{ E(T^2 - T)^2 - E[(T^2 - T)]^2 \right\} \\ &= \frac{1}{n^4} \left\{ E(T^4) - 2E(T^3) + E(T^2) \right. \\ &\quad \left. - [E(T^2)]^2 + 2E(T^2)E(T) - [E(T)]^2 \right\}. \end{aligned}$$

Γνωρίζουμε ότι $E(T) = \nu\theta$ και $E(T^2) = \nu\theta + (\nu\theta)^2$. Επίσης οι ροπές μπορούν να υπολογιστούν είτε βάση του ορισμού η μέσω των παραγώγων της ροπογεννήτριας

$$m_T(t) = e^{\nu\theta(e^t-1)}.$$

Μετά από λίγες πράξεις παίρνουμε

$$Var(U) = \frac{4\theta^3}{\nu} + \frac{2\theta^3}{\nu^2}.$$

Τώρα το C-R κ.φ. είναι

$$\frac{[g'(\theta)]^2}{\nu I_X(\theta)} = \frac{4\theta^3}{\nu},$$

διότι $[g'(\theta)]^2 = 4\theta^2$ και $I_X(\theta) = \frac{1}{\theta}$. Συνεπώς βλέπουμε ότι

$$Var(U) > \text{C-R κ.φ.}$$

παρόλο που είναι αεεδ εκτιμητής.

(γ) Ψάχνουμε αμερόληπτο εκτιμητή που να είναι συνάρτηση του T . Παρατηρούμε ότι $P(X = 0) = e^{-\theta}$ οπότε ορίζουμε την U ως εξής

$$u = \begin{cases} 1, & \text{άν } x_1 = 0; \\ 0, & \text{άν } x_1 > 0. \end{cases}$$

συνεπώς

$$E(U) = 1 \times P(X_1 = 0) + 0 \times P(X_1 = 1) = e^{-\theta}.$$

Όμως το U δεν είναι συνάρτηση του T που είναι επαρκής και πλήρης. Βάση όμως του Θεωρήματος R-B έχουμε

$$E(U|T = t) = 0 \times P(X_1 = 0|T = t) + 1 \times P(X_1 = 1|T = t),$$

όπου

$$\begin{aligned}
P(U = 1|T = t) &= P(X_1 = 0 | \sum X_i = t) \\
&= \frac{P(X_1 = 0, \sum X_i = t)}{P(\sum X_i = t)} \\
&= \frac{P(X_1 = 0, \sum_2^\nu X_i = t)}{P(\sum X_i = t)} \\
&= \frac{P(X_1 = 0)P(\sum_2^\nu X_i = t)}{P(\sum X_i = t)} \\
&= \frac{e^{-\theta}e^{-(\nu-1)\theta}[(\nu-1)\theta]^t/t!}{e^{-\nu\theta}(\nu\theta)^t/t!} \\
&= \left(\frac{\nu-1}{\nu}\right)^t.
\end{aligned}$$

Οπότε

$$E(U|T = t) = 1 \times \left(\frac{\nu-1}{\nu}\right)^t = \left(\frac{\nu-1}{\nu}\right)^t = U^*.$$

Άρα U^* αμερόληπτος της $e^{-\theta}$ και συνάρτηση του T , επομένως αεεδ εκτιμητής.

Παρατήρηση: Μπορεί ναδειχθεί ότι για την Κανονική, $N(\mu, \sigma^2 - \gamma \text{γνωστ})$, $Bernoulli(p)$, $Poisson(\lambda)$, $Exp(\theta)$, ο μέσος $\bar{x}$ είναι αεεδ της άγνωστης παραμέτρου.

Συνέπεια (Consistency)

Ας υποθέσουμε ότι έχουμε $\{T_n, n \geq 1\}$ μια ακολουθία εκτιμητών για την παράμετρο $\theta \in \Theta$.

[Ορισμός]: Η ακολουθία των εκτιμητών $\{T_n\}$ (ή ο εκτιμητής T_n) μιας παραμέτρου θ είναι (ασθενώς) συνεπής εάν για κάθε $\varepsilon > 0$ ισχύει

$$\begin{aligned}\lim_{n \rightarrow \infty} P(|T_n - \theta| < \varepsilon) &= 1, & \forall \theta \in \Theta, \varepsilon > 0 \\ \lim_{n \rightarrow \infty} P(-\varepsilon < T_n - \theta < \varepsilon) &= 1, & \forall \theta \in \Theta, \varepsilon > 0 \\ \lim_{n \rightarrow \infty} P(\theta - \varepsilon < T_n < \theta + \varepsilon) &= 1, & \forall \theta \in \Theta, \varepsilon > 0\end{aligned}$$

Η συνέπεια είναι μια ασυμπτωτική ιδιότητα, που λέει ότι ο T_n είναι συνεπής εάν η πιθανότητα ότι η T_n θα λάβει τιμή έξω από διάστημα μήκους ε γύρω από τη θ τείνει στο μηδέν όταν $n \rightarrow \infty$.

[Θεώρημα]: Έστω T_n μια εκτιμήτρια σ.σ. της θ . Εάν

i) η T_n είναι αμερόληπτη εκτιμήτρια της θ και

ii) $Var(T_n) \rightarrow 0$ όταν $n \rightarrow \infty$,

τότε η T_n είναι συνεπής εκτιμήτρια της θ .

[Παράδειγμα]: Έστω $X_1, X_2, \dots, X_\nu$ από την $N(\mu, \sigma^2)$. Ναδειχθεί ότι ο S^2 είναι συνεπής εκτιμητής του σ^2 .

[Λύση]: Γνωρίζουμε ήδη ότι

$$E\left[\frac{(\nu-1)S^2}{\sigma^2}\right] = \nu-1 \Rightarrow E[S^2] = \sigma^2,$$

και συνεπώς το S^2 είναι αμερόληπτος εκτιμητής του σ^2 . Επίσης

$$\text{Var}\left(\frac{(\nu-1)S^2}{\sigma^2}\right) = 2(\nu-1) \Rightarrow \text{Var}(S^2) = \frac{2\sigma^4}{\nu-1}.$$

Όμως

$$\text{Var}(S^2) = \frac{2\sigma^4}{\nu-1} \rightarrow 0 \quad \text{όταν } \nu \rightarrow \infty.$$

Άρα είναι συνεπής εκτιμητής του σ^2 .

[Θεώρημα (Γενίκευση)]: Έστω T_n μια εκτιμήτρια σ.σ. της $g(\theta)$. Εάν

i) η T_n είναι ασυμπτωτικά αμερόληπτη εκτιμήτρια της $g(\theta)$

$$\lim_{n \rightarrow \infty} E(T_n) = g(\theta)$$

ii) και

$$\lim_{n \rightarrow \infty} V(T_n) = 0,$$

τότε η T_n είναι συνεπής εκτιμήτρια της $g(\theta)$.

[Παράδειγμα]: Έστω $X_1, X_2, \dots, X_n$ από $f(x; \theta)$. Ο δειγματικός μέσος είναι πάντα συνεπής εκτιμήτρια της μέσης τιμής της κατανομής του X .

[Λύση]: Γνωρίζουμε ότι είναι αμερόληπτος:

$$E(\bar{X}) = E(X) = \mu.$$

Επίσης

$$V(\bar{X}) = \frac{1}{n} V(X)$$

άρα

$$\lim_{n \rightarrow \infty} V(\bar{X}) = \lim_{n \rightarrow \infty} \frac{1}{n} V(X) = 0.$$

Άρα, από το πιο πάνω θεώρημα, η $\bar{X}$ συνεπής εκτιμήτρια του μ .

Παρατήρηση: Το κριτήριο της συνέπειας δεν είναι κατάλληλο για την εκλογή καλών εκτιμητριών.

[Παράδειγμα]: Αν η T_n είναι συνεπής, τότε και η

$$T_n^* = T_n + Q(n), \quad \text{όπου: } Q(n) \rightarrow 0 \text{ όταν } n \rightarrow \infty$$

είναι επίσης συνεπής.

Άρα, αφού $\bar{X}$ συνεπής, τότε και $\bar{X} + \left(\frac{\sqrt{\bar{X}^2}}{n}\right)^6$ συνεπής!!!

Ιδιότητες Εκτιμητών Μεγίστης Πιθανοφάνειας

[Θεώρημα]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. με σ.π.π. $f(x; \underline{\theta})$ και έστω $\underline{T} = (T_1, T_2, \dots, T_r)$, όπου $T_j = T_j(\underline{X})$ με $j = 1, 2, \dots, r$, μια επαρκής σ.σ. για το $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_r)$. Εάν $\hat{\underline{\theta}} = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_r)$ είναι ο μονοσήμαντα ορισμένος Ε.Μ.Π. της $\underline{\theta}$, τότε ο $\hat{\underline{\theta}}$ είναι συνάρτηση μόνον του $\underline{T}$.

[Θεώρημα]: Έστω $X_1, X_2, \dots, X_n$ τ.δ. με σ.π.π. $f(x; \underline{\theta})$ και $\hat{\underline{\theta}} = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_r)$ ο Ε.Μ.Π. της $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_r)$. Εάν $\varphi(\underline{\theta}) = (\varphi_1(\underline{\theta}), \dots, \varphi_r(\underline{\theta}))$ ένας μετασχηματισμός του παραμετρικού χώρου Θ , τότε ο Ε.Μ.Π. της $\varphi(\underline{\theta})$ είναι $\varphi(\hat{\underline{\theta}}) = (\varphi_1(\hat{\underline{\theta}}), \dots, \varphi_r(\hat{\underline{\theta}}))$.

Παρατήρηση: Αυτό το Θεώρημα μας λέει ότι για να βρούμε τον Ε.Μ.Π. μιας συνάρτησης της άγνωστης παραμέτρου, αρκεί να βρούμε τον εκτιμητή της παραμέτρου και να αντικαταστήσουμε την τιμή στη συνάρτηση της οποίας ζητάμε τον Ε.Μ.Π.