ΣΤΑΤΙΣΤΙΚΗ ΙΙΙ

Θεωρία 09

Δήμητρα Κυριακοπούλου

Τμήμα Οικονομικών Επιστημών

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

- Είδαμε ότι στην περίπτωση που θέλουμε να πραγματοποιήσουμε τον έλεγχο των υποθέσεων

$$H_0 : \theta \leq \theta_0$$
$$H_1 : \theta > \theta_0$$

μπορούμε πάλι να εφαρμόσουμε το λήμμα Neyman-Pearson, ώστε να βρούμε Ομοιόμορφο Ισχυρότατο Ελεγχο (ΟΙΕ) για τις απλές υποθέσεις

$$H_0 : \theta = \theta_0$$
$$H_1 : \theta = \theta_1 \quad \text{με } \theta_1 > \theta_0.$$

- Αν επιπλέον δείξουμε ότι η κρίσιμη περιοχή του ελέγχου δεν εξαρτάται από το θ_1 και ότι η συνάρτηση ισχύος $\pi(\theta)$ είναι αύξουσα ως προς θ, τότε ο έλεγχος είναι ΟΙΕ και για τον έλεγχο των υποθέσεων $H_0 : \theta \leq \theta_0$ έναντι $H_1 : \theta > \theta_0$.

- Αντίστοιχα για $H_0 : \theta \geq \theta_0$ έναντι $H_1 : \theta < \theta_0$: αν δείξουμε ότι η κρίσιμη περιοχή του ελέγχου δεν εξαρτάται από το θ_1 και ότι η συνάρτηση ισχύος $\pi(\theta)$ είναι φθίνουσα ως προς θ, τότε ο έλεγχος αυτών των υποθέσεων είναι ΟΙΕ.

- Γενικά στους ελέγχους υποθέσεων, δεν έχουμε κάποια συγκεκριμένη βάση για την επιλογή της τιμής της παραμέτρου θ στην εναλλακτική υπόθεση H_1.

- Χρειαζόμαστε για αυτό το λόγο μια διαδικασία να ελέγχουμε σύνθετες υποθέσεις, κατά προτίμηση με έναν ΟΙΕ.

- Ένας τέτοιος βέλτιστος έλεγχος έχει τη μεγαλύτερη ισχύ από όλους τους ελέγχους σε επίπεδο σημαντικότητας α.

- Η ιδιότητα του μονότονου λόγου πιθανοφανειών (ΜΛΠ) παίζει σημαντικό ρόλο σε αυτό το κομμάτι της θεωρίας (σύνθετες υποθέσεις μονόπλευρων ελέγχων).

- Οικογένειες κατανομών που φέρουν την ιδιότητα του ΜΛΠ με επαρκές στατιστικό $T(X) = \sum_{i=1}^{n} x_i$ είναι: η Κανονική (με σ γνωστό), η Εκθετική, η Διωνυμική, η Poisson.

Έστω τυχαίο δείγμα $(X_1, ..., X_n)$ από πληθυσμό που ακολουθεί κατανομή $f(x;\theta)$, η οποία έχει μονότονο λόγο πιθανοφανειών (ΜΛΠ) ως προς τη στατιστική συνάρτηση $T = T(X)$. Τότε έχουμε τα εξής:

1. Για το πρόβλημα ελέγχου:

$$H_0 : \theta \leq \theta_0$$
$$H_1 : \theta > \theta_0$$

υπάρχει ΟΙΕ που ορίζεται από την κρίσιμη περιοχή:

$$C : \{x : T(X) > \kappa\}.$$

2. Για το πρόβλημα ελέγχου:

$$H_0 : \theta \geq \theta_0$$
$$H_1 : \theta < \theta_0$$

υπάρχει ΟΙΕ που ορίζεται από την κρίσιμη περιοχή:

$$C : \{x : T(X) < \kappa\}.$$

- Και στις δύο περιπτώσεις η σταθερά κ καθορίζεται έτσι ώστε:

$$P\left(\text{Σφάλμα Τύπου Ι} \mid \theta = \theta_0\right) = \alpha.$$

- Τότε, οι έλεγχοι με κρίσιμες περιοχές που δόθηκαν στα προηγούμενα είναι οι ΟΙΕ μεγέθους α και ικανοποιούν τη σχέση:

$$P\left(\text{Απορρίπτουμε την} H_0 \mid H_0\right) \leq \alpha, \quad \forall \theta \in H_0.$$

- Επίσης, στην περίπτωση 1. (στην προηγούμενη διαφάνεια):

$$H_0 : \theta \leq \theta_0$$
$$H_1 : \theta > \theta_1$$

για κάποια τιμή $\theta_1 > \theta_0$, θα λέμε ότι, αν $L(\theta_1, x)$ ικανοποιεί την ιδιότητα του ΜΛΠ ως προς $L(\theta_0, x)$, όσο υψηλότερη η παραστηρούμενη τιμή του $T(X)$, τόσο πιο πιθανό το $T(X)$ να προέρχεται από κατανομή με πιθανοφάνεια $L(\theta_1, x)$ παρά $L(\theta_0, x)$.

- Παρατήρηση:

  Ουσιαστικά, ένας ΟΙΕ για τις υποθέσεις:

$$H_0 : \theta = \theta_0$$

$$H_1 : \theta > \theta_0$$

  μπορεί να επεκταθεί στην περίπτωση του πιο γενικού μονόπλευρου ελέγχου, όπου:

$$H_0 : \theta \leq \theta_0$$

$$H_1 : \theta > \theta_0$$

  για τον οποίο ισχύει ότι:

$$\alpha = \max P\left(\text{απορρίπτουμε } H_0 \mid \theta = \theta_0\right).$$

  Αυτό ωστόσο αποτυγχάνει στον δίπλευρο έλεγχο υποθέσεων, όπου σε μια τέτοια περίπτωση δεν υπάρχει ΟΙΕ.

Τυχαιοποιημένος Έλεγχος

- Σε ορισμένες περιπτώσεις που δεν μπορούμε να επιτύχουμε ακριβώς το επίπεδο σημαντικότητας που θέλουμε, για παράδειγμα όταν η συνάρτηση πιθανότητας προέρχεται από διακριτή κατανομή, μπορούμε να πραγματοποιήσουμε τυχαιοποιημένο έλεγχο (randomized test).

- Έστω

$$\Psi(x) = P\,(\text{Απορρίψουμε } H_0) = \begin{cases} 1, & \sum_{i=1}^{n} x_i > \kappa \\ \gamma, & \sum_{i=1}^{n} x_i = \kappa \\ 0, & \sum_{i=1}^{n} x_i < \kappa \end{cases}$$

δηλαδή, αν $\sum_{i=1}^{n} x_i > \kappa$ θα απορρίψουμε την H_0(με πιθανότητα 1), ενώ αν $\sum_{i=1}^{n} x_i = \kappa$ θα απορρίψουμε την H_0 με πιθανότητα γ.

- Η πιθανότητα να κάνουμε σφάλμα τύπου I, α, δίνεται από:

$$\begin{aligned}
&P\left(\text{Απορρίψουμε } H_0 \mid \textstyle\sum_{i=1}^{n} x_i > \kappa\right) P\left(\textstyle\sum_{i=1}^{n} x_i > \kappa\right) = \\
&+P\left(\text{Απορρίψουμε } H_0 \mid \textstyle\sum_{i=1}^{n} x_i = \kappa\right) P\left(\textstyle\sum_{i=1}^{n} x_i = \kappa\right) \\
&= 1 \times P\left(\sum_{i=1}^{n} x_i > \kappa \mid H_0\right) + \gamma \times P\left(\sum_{i=1}^{n} x_i = \kappa \mid H_0\right)
\end{aligned}$$

- Αν θέλουμε $\alpha = 0.05$, τότε:

$$1 \times P\left(\sum_{i=1}^{n} x_i > \kappa \mid H_0\right) + \gamma \times P\left(\sum_{i=1}^{n} x_i = \kappa \mid H_0\right) = 0.05$$

από την οποία σχέση θα προκύψει η πιθανότητα γ.

- Παράδειγμα Διωνυμικής κατανομής:

Έστω τυχαίο δείγμα $(X_1, X_2, ..., X_n) \sim B(p)$ με συνάρτηση κατανομής $f(x;p) = p^x (1-p)^{1-x}, \quad x = 0,1, \quad p \in [0,1]$.

Και έστω ότι θέλουμε να ελέγξουμε την υπόθεση:

$$H_0 : p = 0.2$$

$$H_1 : p = 0.4$$

Έχουμε ότι η συνάρτηση πιθανοφάνειας δίνεται από:

$$L_p(x) = \prod_{i=1}^{n} f(x_i;p) = \prod_{i=1}^{n} p^{x_i} (1-p)^{1-x_i}$$
$$= p^{\sum x_i} (1-p)^{n-\sum x_i}.$$

Οπότε ισχύει για το λόγο των πιθανοφανειών της H_1 προς την H_0:

$$\begin{aligned}
\lambda = \frac{L_1(x)}{L_0(x)} &= \frac{(0.4)^{\sum x_i}(0.6)^{n-\sum x_i}}{(0.2)^{\sum x_i}(0.8)^{n-\sum x_i}} \\
&= 2^{\sum x_i}\left(\frac{0.6}{0.8}\right)^{n}\left(\frac{0.6}{0.8}\right)^{-\sum x_i} \\
&= \left(\frac{2}{\frac{0.6}{0.8}}\right)^{\sum x_i}\left(\frac{0.6}{0.8}\right)^{n} \\
&= \left(\frac{8}{3}\right)^{\sum x_i}\left(\frac{3}{4}\right)^{n},
\end{aligned}$$

και άρα $\lambda(T)$ είναι αυστηρώς αύξουσα συνάρτηση ως προς $T = \sum x_i$ (οπότε ισχύει η ιδιότητα του ΜΛΠ).

Η περιοχή απόρριψης $\lambda > \kappa$ ισοδυναμεί με:

$$\begin{aligned}
&\sum x_i > \frac{\log\kappa - n\log\left(\frac{3}{4}\right)}{\log\left(\frac{8}{3}\right)} \iff \\
&\sum x_i > \kappa^* \iff \\
&\quad C : \left\{x_i \mid \sum x_i > \kappa^*\right\}.
\end{aligned}$$

Η τυχαία μεταβλητή:

$$Y = \sum x_i \sim B\left(n, p\right)$$

$$\begin{aligned} P\left(\text{Σφάλμα Τύπου I}\right) &= P\left(\text{Απορρίπτουμε την} H_0 \mid H_0 \text{ αληθής}\right) \\ &= P\left(\sum x_i > \kappa^* \mid p = 0.2\right) \\ &= P\left(Y > \kappa \mid p = 0.2\right). \end{aligned}$$

Η συνάρτηση πιθανότητας είναι:

$$f\left(y; p\right) = P\left(Y = y\right) = \binom{n}{y} p^y \left(1 - p\right)^{n-y}, \quad y = 0, 1, ..., n.$$

Άρα,

$$\begin{aligned} P\left(Y > \kappa \mid p = 0.2\right) &= \sum_{y=\kappa+1}^{n} f\left(y; p\right) \\ &= 1 - P\left(Y \leq \kappa \mid p = 0.2\right) \\ &= \Pi\left(\kappa, p\right). \end{aligned}$$

Επομένως αναζητούμε τον αριθμό κ τέτοιο ώστε, η πιθανότητα να κάνουμε σφάλμα τύπου I:

$$P\left(Y > \kappa \mid p = 0.2\right) = \alpha.$$

Ας υποθέσουμε ότι $\alpha = 0.05$ και $n = 10$. Τότε, από τους πίνακες της Διωνυμικής συνάρτησης κατανομής, οι αθροιστικές πιθανότητες:

- Για $n = 10, \quad \kappa = 4$:

$$\Pi\left(\kappa, p\right) = 1 - 0.9672 = 0.0328.$$

- Για $n = 10, \quad \kappa = 3$:

$$\Pi\left(\kappa, p\right) = 1 - 0.8791 = 0.1209.$$

Οπότε:

$$\begin{aligned} P\left(Y = 4\right) = f\left(4; 0.2\right) &= 0.9672 - 0.8791 \\ &= 0.9672 - 0.8791 \\ &= 0.0881. \end{aligned}$$

$$
\begin{aligned}
P\left(\text{Απορρίπτουμε την} H_0 \mid p=0.2\right)=0.05 \Rightarrow \\
P\left(Y>4 \mid p=0.2\right) \times P\left(\text{Απορρίπτουμε την} H_0 \mid Y>4\right) \\
+P\left(Y=4 \mid p=0.2\right) \times P\left(\text{Απορρίπτουμε την} H_0 \mid Y=4\right)=0.05 \Rightarrow \\
0.0328+0.0881\gamma=0.05 \Rightarrow \\
\gamma=0.1952.
\end{aligned}
$$

Το παραπάνω έχει προκύψει από τον τυχαιοποιημένο έλεγχο, όπου θα απορρίψουμε την H_0 αν $Y>4$ ή θα δεχθούμε την H_0 αν $Y<4$.

Αν $Y=4$, τότε θα απορρίψουμε τη μηδενική υπόθεση με πιθανότητα γ, ενώ θα τη δεχθούμε με πιθανότητα $1-\gamma$.

Κάτω από τη μηδενική υπόθεση:

$$Y = \sum x_i \sim B\left(n, p_0\right),$$

όπου

$$E\left(Y\right) = np_0,$$
$$Var\left(Y\right) = np_0q_0, \quad \left(q_0 = 1 - p_0\right)$$

Τότε:

$$z = \frac{\sum x_i - np_0}{\sqrt{np_0q_0}} = \frac{\bar{x} - p_0}{\sqrt{\frac{p_0q_0}{n}}},$$

όπου κάτω από τη μηδενική υπόθεση:

$$z \sim N\left(0, 1\right).$$

Το κριτήριο για την κρίσιμη περιοχή είναι:

$$\sum x_i > \kappa \Longleftrightarrow \bar{x} > \kappa^* \Longleftrightarrow$$
$$\frac{\bar{x} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} > \kappa^{**}.$$

Οπότε ισχύει:

$$P\left(z > \kappa^{**} \mid H_0\right) = \alpha \Longleftrightarrow \kappa^{**} = z_\alpha \Longleftrightarrow$$
$$\frac{\bar{x} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} > z_\alpha \Longleftrightarrow$$
$$\bar{x} > z_\alpha \sqrt{\frac{p_0 q_0}{n}} + p_0 \Longleftrightarrow$$
$$\begin{aligned} \sum x_i &> z_\alpha \sqrt{n p_0 q_0} + n p_0 \\ &= 1.645\sqrt{10 \cdot 0.2 \cdot 0.8} + 10 \cdot 0.2 \\ &= 4.08. \end{aligned}$$

Επίσης, από τον πίνακα της Διωνυμικής κατανομής για

$$n = 10, \quad p = 0.2$$

έχουμε τις αθροιστικές πιθανότητες:

$$P(4) = 1 - 9672 = 0.0328$$

$$P(3) = 1 - 8791 = 0.1209.$$

Άρα αν θέσουμε $\kappa = 4$, έχουμε την κρίσιμη περιοχή:

$$C : \sum_{i=1}^{n} x_i > 4,$$

που αντιστοιχεί σε επίπεδο σημαντικότητας $\alpha = 3.28\%$.

Αν θέσουμε $\kappa = 3$, έχουμε την κρίσιμη περιοχή:

$$C : \sum_{i=1}^{n} x_i > 3,$$

που δίνει επίπεδο σημαντικότητας $\alpha = 0.1209$ ή ισοδύναμα $\alpha = 12.09\%$.

Θεώρημα Karlin-Rubin

- Το Θεώρημα Karlin-Rubin (KR) μπορεί να θεωρηθεί ως μια επέκταση του λήμματος Neyman-Pearson για σύνθετες υποθέσεις.

- Σύμφωνα με το Θεώρημα: έστω ότι εξετάζουμε την υπόθεση:

$$
\begin{aligned}
H_0 &: \theta \leq \theta_0 \\
H_1 &: \theta > \theta_0
\end{aligned}
$$

  και έστω ότι $T(X)$ είναι ένα επαρκές στατιστικό για το θ και ότι η οικογένεια κατανομών με άγνωστη παράμετρο θ έχει ΜΛΠ ως προς $T(X)$. Τότε, για κάθε κ, ο έλεγχος με κρίσιμη περιοχή $T(X) > \kappa$ είναι ΟΙΕ μεγέθους α, όπου α η πιθανότητα να κάνουμε Σφάλμα Τύπου Ι.

- Το Θεώρημα αυτό είναι γνωστό για το γενικό πρόβλημα ελέγχων σύνθετων υποθέσεων που έχουν την παραπάνω μορφή (right-sided alternative hypotheses).

- Άσκηση:

  Έστω ότι $X_1, ..., X_n \sim N(\theta, 1)$. Να βρεθεί ΟΙΕ μεγέθους α

  1. για την υπόθεση:

  $$H_0 : \theta = \theta_0 \text{ vs } H_1 : \theta = \theta_1$$

     όπου $\theta_1 > \theta_0$

  2. για την υπόθεση:

  $$H_0 : \theta = \theta_0 \text{ vs } H_1 : \theta > \theta_0$$

  3. για την υπόθεση:

  $$H_0 : \theta \leq \theta_0 \text{ vs } H_1 : \theta > \theta_0$$

(Εφαρμόστε το Θεώρημα Karlin-Rubin σε κάθε μία από τις τρεις αυτές υποθέσεις για να αποδείξετε ότι όλες χρησιμοποιούν την ίδια στατιστική ελέγχου και την ίδια κρίσιμη περιοχή.)

- <u>Παράδειγμα:</u>

Έστω τυχαίο δείγμα $(X_1, ..., X_n) \sim \Gamma(p, \theta)$, με $p > 0$ γνωστό. Να βρεθεί ΟΙΕ για τις υποθέσεις:

$$H_0 : \theta \leq \theta_0$$
$$H_1 : \theta < \theta_0$$

Η Κατανομή Γάμμα ανήκει στην Εκθετική Οικογένεια Κατανομών και ισχύει:

$$f(x;\theta) = \exp\left\{-\theta \sum_{i=1}^{n} x_i + np \log \theta\right\} \frac{1}{[\Gamma(p)]^n} \prod_{i=1}^{n} x_i^{p-1},$$

όπου η συνάρτηση $c(\theta) = \theta$ είναι γνησίως αύξουσα και $d(x) = -\sum_{i=1}^{n} x_i$.

Από το Θεώρημα Karlin-Rubin, ισχύει για την κρίσιμη περιοχή:

$$T\left(X\right) < \kappa \Longleftrightarrow -\sum_{i=1}^{n} x_i < \kappa \Longleftrightarrow$$

$$\sum_{i=1}^{n} x_i > \kappa^*.$$

Για $\theta = \theta_0$:

$$Y = \sum_{i=1}^{n} X_i \sim \Gamma\left(np, \theta_0\right).$$

- <u>Παράδειγμα:</u>

  Έστω $X_1, X_2, ..., X_n$ τυχαίο δείγμα από την κατανομή Poisson(λ) και έστω ο έλεγχος:

$$H_0 : \lambda \leq \lambda_0$$

$$H_1 : \lambda > \lambda_0$$

  Να δειχθεί ότι υπάρχει ΜΛΠ για την οικογένεια κατανομών:

$$f(x; \lambda) = e^{-\lambda} \frac{\lambda^x}{x!}.$$

  Έχουμε ότι για $\lambda_1 < \lambda_2$:

$$\frac{L_2}{L_1} = \frac{e^{-n\lambda_2} \frac{\lambda_2^{\sum x_i}}{\prod x_i!}}{e^{-n\lambda_1} \frac{\lambda_1^{\sum x_i}}{\prod x_i!}} = e^{n(\lambda_1 - \lambda_2)} \left(\frac{\lambda_2}{\lambda_1}\right)^{\sum x_i},$$

  που είναι αύξουσα συνάρτηση ως προς $T = \sum x_i$ και συνάρτηση του τυχαίου δείγματος μέσω του T.

- <u>Παράδειγμα:</u>

  Έστω η συνάρτηση πιθανοφάνειας της Κανονικής κατανομής για τυχαίο δείγμα $(X_1, X_2, ..., X_n)$:

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

  Και έστω ότι θέλουμε να κάνουμε τον έλεγχο ότι κάτω από τη μηδενική υπόθεση H_0:

$$\hat{\mu}_{MLE} = \bar{x}$$
$$\hat{\sigma^2}_{MLE} = \sigma_0^2$$

  ένω κάτω από την εναλλακτική H_1 ισχύει:

$$\hat{\mu}_{MLE} = \bar{x}$$
$$\hat{\sigma^2}_{MLE} = \hat{\sigma^2} = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2.$$

Τότε η συνάρτηση πιθανοφάνειας κάτω από την H_0 είναι:

$$L_0 = \left(\frac{1}{2\pi\sigma_0^2}\right)^{n/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{\sigma_0}\right)^2\right\}$$

και κάτω από την H_1:

$$\begin{aligned} L_1 &= \left(\frac{1}{2\pi\hat{\sigma}^2}\right)^{n/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{\hat{\sigma}}\right)^2\right\} \\ &= \left(\frac{1}{2\pi\hat{\sigma}^2}\right)^{n/2} \exp\left\{-\frac{n}{2}\right\}. \end{aligned}$$

Οπότε ο λόγος των πιθανοφανειών δίνεται από:

$$\frac{L_0}{L_1} = \left(\frac{\hat{\sigma}^2}{\sigma_0^2}\right)^{n/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{\sigma_0}\right)^2 + \frac{n}{2}\right\}.$$

Θέτουμε:

$$\begin{aligned} T &= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{\sigma_0}\right)^2 \\ &= \frac{\hat{\sigma^2}}{\sigma_0^2}. \end{aligned}$$

Τότε

$$\begin{aligned} \frac{L_0}{L_1} &= T^{n/2}\exp\left\{\frac{n}{2}-\frac{n}{2}T\right\} \\ &= \left(Te^{1-T}\right)^{n/2} < \kappa \Rightarrow \\ Te^{1-T} &< \kappa^* \Rightarrow \\ T &< t_1 \text{ ή } T > t_2. \end{aligned}$$

Από το οποίο προκύπτει ότι:

$$\hat{S_0^2} < \kappa_1 \text{ ή } \hat{S_0^2} > \kappa_2, \qquad \hat{S_0^2} = \sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{\sigma_0}\right)^2.$$

Η πιθανότητα να κάνουμε Σφάλμα Τύπου I:

$$P\left(\text{Απορρίψουμε } H_0 \mid H_0\right) = \alpha$$

$$\left\{\begin{array}{c} P\left(\hat{S_0^2} < \kappa_1 \mid \hat{\sigma^2} = \sigma_0^2\right) \\ +P\left(\hat{S_0^2} > \kappa_2 \mid \hat{\sigma^2} = \sigma_0^2\right) \end{array}\right\} = \alpha.$$

Όμως,

$$\hat{S_0^2} = \sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{\sigma_0}\right)^2 \sim \mathcal{X}_{n-1}^2,$$

οπότε η μία κρίσιμη περιοχή είναι:

$$C_1 = \mathcal{X}_{n-1,1-\frac{\alpha}{2}}^2,$$

και η δεύτερη:

$$C_2 = \mathcal{X}_{n-1,\frac{\alpha}{2}}^2.$$

- Παράδειγμα:

Έστω $(X_1, X_2, ..., X_n) \sim f(x;\theta) = \frac{1}{\theta} \exp\left\{-\frac{x}{\theta}\right\}$.

Και έστω ότι ο λόγος των πιθανοφανειών δίνεται από:

$$\lambda = \frac{L(x;\theta_0)}{L\left(x;\hat{\theta}\right)}.$$

Αν θεωρήσουμε ότι:

$$\hat{\theta} := \text{η εκτιμήτρια MLE της παραμέτρου } \theta.$$

Η συνάρτηση πιθανοφάνειας είναι η εξής:

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

και η λογαριθμική πιθανοφάνεια:

$$\mathcal{L}(x;\theta) = -n\log\theta - \frac{1}{\theta}\sum x_i.$$

Υπλογίζοντας την πρώτη παράγωγο και εξισώνοντας με το μηδέν:

$$\frac{\partial \mathcal{L}}{\partial \theta} = 0 \Rightarrow$$
$$-\frac{n}{\theta} + \frac{1}{\theta^2} \sum x_i = 0 \Rightarrow$$
$$\hat{\theta} = \frac{\sum x_i}{n} = \bar{x}.$$

Οπότε τώρα ο λόγος των πιθανοφανειών δίνεται από:

$$\begin{aligned}
\lambda = \frac{\frac{1}{\theta_0^n} \exp\left\{-\frac{1}{\theta_0}\sum x_i\right\}}{\frac{1}{\bar{x}^n} \exp\left\{-\frac{1}{\bar{x}}\sum x_i\right\}} &= \left(\frac{\bar{x}}{\theta_0}\right)^n \frac{\exp\left\{-\frac{1}{\theta_0}\sum x_i\right\}}{\exp\{-n\}} \\
&= \left(\frac{\bar{x}}{\theta_0}\right)^n \exp\left\{n - \frac{n\bar{x}}{\theta_0}\right\} \\
&= \left(\frac{\bar{x}}{\theta_0}\right)^n \exp\left\{n\left(1 - \frac{\bar{x}}{\theta_0}\right)\right\} \\
&= \left(\frac{1}{n\theta_0}\sum x_i\right)^n \exp\left\{n\left(1 - \frac{\sum x_i}{n\theta_0}\right)\right\}.
\end{aligned}$$

$$
\begin{aligned}
\lambda < \Lambda \Longleftrightarrow \frac{1}{\kappa} x < \Lambda, \qquad & x : \sum x_i, \;\; \kappa : n\theta_0 \\
\Longleftrightarrow x < x_1 \text{ ή } x > x_2.
\end{aligned}
$$

Άρα η κρίσιμη περιοχή έχει την μορφή:

$$C : \sum x_i < \kappa_1 \text{ ή } \sum x_i > \kappa_2.$$

Προσδιορισμός των τιμών κ_1, κ_2:

$$
\begin{aligned}
P\left(\text{Απορρίψουμε} H_0 \mid H_0 \text{ισχύει}\right) &= P\left(\sum x_i < \kappa_1 \text{ ή } \sum x_i > \kappa_2 \mid \theta = \theta_0\right) \\
&= 1 - P\left(\kappa_1 \leq \sum x_i \leq \kappa_2 \mid \theta = \theta_0\right).
\end{aligned}
$$

Όμως αν $X_1, ...X_n \sim \mathrm{Exp}(\theta_0)$, τότε:

$$\sum_{i=1}^{n} X_i \sim \Gamma\left(n, \theta_0\right) : \qquad f\left(x; \alpha, \beta\right) = \frac{1}{\beta^{\alpha} \Gamma\left(\alpha\right)} x^{\alpha - 1} e^{-x/\beta}.$$

Έτσι:

$$\alpha = 1 - \int_{\kappa_1}^{\kappa_2} \frac{1}{\theta_0^n \Gamma(n)} x^{n-1} e^{-x/\theta_0} dx \Longleftrightarrow$$

$$\int_{\kappa_1}^{\kappa_2} \frac{1}{\theta_0^n \Gamma(n)} x^{n-1} e^{-x/\theta_0} dx = 1 - \alpha.$$

Αν έχουμε ότι: $\alpha = 0.10, \;\; \theta_0 = 1, \;\; n = 10 \Rightarrow$

$$\kappa_1 = 5.45, \quad \kappa_2 = 15.7.$$

- <u>Παράδειγμα:</u>

Έστω τυχαία μεταβλητή $X \sim N(\mu, 1)$, $\quad f(x;\mu) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$ και έστω ότι έχουμε να εξετάσουμε τον παρακάτω έλεγχο:

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu = \mu_1$$

με $\mu_1 > \mu_0$.

$$\begin{aligned}
\frac{L_1(x)}{L_0(x)} &= \exp\left\{\frac{1}{2\sigma^2}\left[(x-\mu_0)^2 - (x-\mu_1)^2\right]\right\} \\
&= \exp\left\{\frac{1}{2\sigma^2}\left[\left(\mu_0^2 - \mu_1^2\right) + 2x(\mu_1 - \mu_0)\right]\right\} \\
&= \exp\left\{\frac{1}{2\sigma^2}\left(\mu_0^2 - \mu_1^2\right)\right\} \exp\left\{\frac{x}{\sigma^2}(\mu_1 - \mu_0)\right\} > \kappa \\
\iff \exp\left\{\frac{x}{\sigma^2}(\mu_1 - \mu_0)\right\} &> \kappa^* = \kappa \exp\left\{\frac{1}{2\sigma^2}\left(\mu_0^2 - \mu_1^2\right)\right\} \\
\iff x &> \kappa^{**} = \sigma^2 \frac{\log \kappa^*}{\mu_1 - \mu_0} = k.
\end{aligned}$$

Οπότε η κρίσιμη περιοχή ορίζεται ως:

$$C : x > k.$$

Για την εύρεση του k:

$$\begin{aligned}
P\left(\text{Απορρίψουμε}H_0 \mid H_0\text{ισχύει}\right) &= P\left(x > k \mid \mu = \mu_0\right) \\
&= P\left(\frac{x-\mu_0}{\sigma} > \frac{k-\mu_0}{\sigma}\right) \\
&= P\left(z > \frac{k-\mu_0}{\sigma}\right) = \alpha \Rightarrow \\
\frac{k-\mu_0}{\sigma} &= z_\alpha \Rightarrow \\
k &= \sigma z_\alpha + \mu_0.
\end{aligned}$$

Για την εύρεση ισχύος:

$$
\begin{aligned}
\beta &= P\left(\text{Αποδεχθούμε} H_0 \mid H_1 \text{ισχύει}\right) \\
&= P\left(x \leq k \mid H_1 \text{ισχύει}\right) \\
&= P\left(x \leq \sigma z_\alpha + \mu_0 \mid H_1 \text{ισχύει}\right) \\
&= P\left(\frac{x-\mu_1}{\sigma} \leq \frac{\sigma z_\alpha + \mu_0 - \mu_1}{\sigma}\right) \\
&= P\left(z \leq z_\alpha - \frac{\mu_1-\mu_0}{\sigma}\right)
\end{aligned}
$$

Και

$$
\begin{aligned}
\pi = 1-\beta &= P\left(z > z_\alpha - \frac{\mu_1-\mu_0}{\sigma}\right) \\
&= P\left(z > z_\alpha - \Delta\right), \quad \Delta = \frac{\mu_1-\mu_0}{\sigma}.
\end{aligned}
$$

Στο παράδειγμά μας, $\sigma = 1$. Αν θέσουμε, επίσης, $\mu_1 = 1, \mu_0 = 0, \alpha = 0.05$:

$$\Delta = 1, \quad z_\alpha = 1.645,$$

οπότε

$$\pi = P\left(z > 0.65\right) \simeq 0.26.$$