ΣΤΑΤΙΣΤΙΚΗ ΙΙΙ

Θεωρία 10

Δήμητρα Κυριακοπούλου

Τμήμα Οικονομικών Επιστημών

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών

- Σε συνέχεια του Παραδείγματος [Θεωρία 09, σελ. 21] της κατανομής Poisson:

  Η κρίσιμη περιοχή του ελέγχου είναι:

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

όπου το κ προσδιορίζεται για δεδομένο α:

$$P\left(Y > \kappa \mid \lambda = \lambda_0\right) = \alpha,$$

όπου

$$Y = \sum_{i=1}^{n} x_i \sim P\left(n\lambda_0\right).$$

Για $\alpha = 4.87$ και $\lambda_0 = 1, n = 10$, έχουμε $\kappa = 15$ διότι:

$$P\left(Y > 15\right) = P\left(Y \geq 16\right) = 0.049.$$

Για $\alpha = 6.38$ και $\lambda_0 = 2, n = 4$, έχουμε $\kappa = 12$ διότι:

$$P\left(Y > 12\right) = P\left(Y \geq 13\right) = 0.064.$$

Για $\alpha = 1.4\%$ και $\lambda_0 = 1, n = 10$, έχουμε $\kappa = 17$.

(βλ. τους Πίνακες της Poisson κατανομής)

Για $\lambda_0 = 1, n = 5$:

$$
\begin{aligned}
P\left(\sum x_i > 9 \mid H_0\right) &= 1 - P\left(\sum x_i \leq 9\right) \\
&= 1 - 0.968 \\
&= 0.032.
\end{aligned}
$$

Η πιθανότητα να κάνουμε σφάλμα τύπου I δίνεται από:

$$
\begin{aligned}
\begin{matrix} P\left(\sum x_i > 9 \mid H_0\right) P\left(\sum x_i > 9\right) \\ +P\left(\sum x_i = 9 \mid H_0\right) P\left(\sum x_i = 9\right) \end{matrix} &= \alpha \Rightarrow \\
0.032 + 0.036\gamma &= 0.05 \Rightarrow \\
\gamma &= 0.5
\end{aligned}
$$

Οπότε έχουμε:

$$
\Psi = \begin{cases} 1, & \sum x_i > 9 \\ 0.5, & \sum x_i = 9 \\ 0, & \sum x_i < 9 \end{cases}
$$

Για $\lambda_0 = 3, n = 3$:

$$
\begin{aligned}
P\left(\sum x_i > 13 \mid H_0\right) &= 1 - P\left(\sum x_i \leq 13\right) \\
&= 1 - 0.959 \\
&= 0.041.
\end{aligned}
$$

Η πιθανότητα να κάνουμε σφάλμα τύπου I δίνεται από:

$$
\begin{aligned}
\begin{matrix} P\left(\sum x_i > 13 \mid H_0\right) P\left(\sum x_i > 13\right) \\ +P\left(\sum x_i = 13 \mid H_0\right) P\left(\sum x_i = 13\right) \end{matrix} &= \alpha \Rightarrow \\
0.041 + 0.023\gamma &= 0.05 \Rightarrow \\
\gamma &= 0.3913
\end{aligned}
$$

Οπότε έχουμε:

$$
\Psi = \begin{cases} 1, & \sum x_i > 13 \\ 0.3913, & \sum x_i = 13 \\ 0, & \sum x_i < 13 \end{cases}
$$

- <u>Παράδειγμα:</u>

Έστω τυχαίο δείγμα $(X_1, X_2, ..., X_n) \sim P(\lambda)$ με $n = 10$. Να βρεθεί ο ΟΙΕ μεγέθους $\alpha = 2.9\%$ για την υπόθεση:

$$H_0 : \lambda \geq 1$$
$$H_1 : \lambda < 1$$

Βάσει του Θεωρήματος Karlin-Rubin και επειδή η κατανομή Poisson έχει την ιδιότητα του ΜΛΠ ως προς $T = \sum x_i$, η κρίσιμη περιοχή του ΟΙΕ δίνεται από:

$$C : \sum_{i=1}^{10} x_i < \kappa.$$

Το μέγεθος του ελέγχου προσδιορίζεται από την ισότητα:

$$\alpha = P\left(\text{Απορρίπτουμε } H_0 \mid H_0\right) = P\left(\sum_{i=1}^{10} x_i < \kappa \mid \lambda = 1\right)$$
$$= 0.029.$$

Όμως η τυχαία μεταβλητή $Y = \sum_{i=1}^{10} x_i$ έχει κατανομή:

$$Y \sim P(n\lambda) = P(10 \times 1) = P(10).$$

Άρα από τους πίνακες της Poisson για $\lambda = 10$ αναζητούμε την τιμή κ για την οποία ισχύει:

$$P(Y < \kappa) = 0.029.$$

Από τους σχετικούς πίνακες έχουμε:

$$P(Y < 5) = P(Y \leq 4) = F(4) = 0.029,$$

άρα $\kappa = 5$.

Αμφίπλευροι Έλεγχοι - Η γενικότερη Περίπτωση

- Έχουμε ήδη δει από το παράδειγμα της Κανονικής κατανομής ότι <u>δεν υπάρχει ΟΙΕ για την αμφίπλευρη υπόθεση</u>:

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$

  διότι:

- για την μονόπλευρη εναλλακτική υπόθεση $H_1 : \mu > \mu_0$, η κρίσιμη περιοχή είναι

$$\bar{x} > \frac{\sigma}{\sqrt{n}} z_\alpha + \mu_0$$

- ενώ για την μονόπλευρη εναλλακτική υπόθεση $H_1 : \mu < \mu_0$, η κρίσιμη περιοχή είναι

$$\bar{x} < \mu_0 - \frac{\sigma}{\sqrt{n}} z_\alpha .$$

- Είδαμε στις προηγούμενες διαλέξεις ότι μέσω του λήμματος Neyman-Pearson μπορούμε να κατασκευάζουμε ισχυρότατους ελέγχους απλών υποθέσεων.

- Διαπιστώσαμε επίσης ότι σε συνδυασμό με το λήμμα αυτό, μπορεί να αναπτυχθεί κατάλληλη μεθοδολογία για τη δημιουργία ΟΙΕ στην Εκθετική Οικογένεια Κατανομών για τον έλεγχο μονόπλευρων υποθέσεων της μορφής:

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

και

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

- Ωστόσο, στη γενικότερη περίπτωση όπου για παράδειγμα μπορεί να έχουμε **αμφίπλευρες υποθέσεις** της μορφής:

$$H_0 : \theta = \theta_0, H_1 : \theta \neq \theta_0$$

  ή μπορεί στους ελέγχους να εμπλέκονται άγνωστες παράμετροι (π.χ. στον έλεγχο για το μ της κανονικής κατανομής με το σ^2 να είναι άγνωστο) δεν μπορούμε να εργαστούμε μέσω του λήμματος Neyman-Pearson για τη δημιουργία ΟΙΕ.

- Σε αυτή την περίπτωση θα δούμε παρακάτω ότι μπορεί να χρησιμοποιηθεί μια επέκταση του λήμματος Neyman-Pearson που, αν και δεν οδηγεί πάντα σε βέλτιστο έλεγχο, δίνει αρκετά ικανοποιητικά αποτελέσματα στις περισσότερες περιπτώσεις.

- Συνοψίζοντας αυτά τα συμπεράσματα:
  - ο μονότονος λόγος πιθανοφανειών (ΜΛΠ) που έχουμε δει στα προηγούμενα δεν είναι εφαρμόσιμος για την υπόθεση της μορφής: $H_0 : \theta = \theta_0$ έναντι $H_0 : \theta \neq \theta_0$.
  - εφαρμόζεται μόνο για μονόπλευρες υποθέσεις.
  - όταν μας ενδιαφέρει να πραγματοποιήσουμε δίπλευρες υποθέσεις, εφαρμόζουμε ένα άλλο κριτήριο, αυτό του γενικευμένου λόγου πιθανοφανειών, που θα δούμε στη συνέχεια.

- Έστω ότι έχουμε τον έλεγχο:

$$H_0 : \theta = \theta_0$$
$$H_1 : \theta \neq \theta_0$$

- Έστω θ^* η πραγματική τιμή του θ και $\hat{\theta}_{MLE}$ η εκτιμήτρια μεγίστης πιθανοφάνειας.

- Τότε:

$$\hat{\theta}_{MLE} \to \theta^*$$

και

$$\max_{\theta} L(\theta) \to L(\theta^*)$$
$$L\left(\hat{\theta}_{MLE}\right) \to L(\theta^*)$$

όπου $\max_{\theta} L(\theta) = L\left(\hat{\theta}_{MLE}\right)$.

- Άρα για μεγάλο n (μέγεθος του δείγματος), αν ισχύει η H_0 θα έχουμε:

$$\max_{\theta\in\Theta_0} L(\theta) \simeq \max_{\theta\in\Theta} L(\theta),$$

διότι:

$$\max_{\theta\in\Theta} L(\theta) = L\left(\hat{\theta}_{MLE}\right)$$
$$\to L(\theta^*).$$

Γενικευμένος Λόγος Πιθανοφανειών

- Στις περιπτώσεις όπου δεν μπορούν να εφαρμοστούν τα αποτελέσματα όσων έχουμε δει στα προηγούμενα [Λήμμα Neyman-Pearson, ΜΛΠ και Θεώρημα Karlin-Rubin] για την έυρεση ΟΙΕ, καταφεύγουμε στη χρήση του κριτηρίου του Γενικευμένου Λόγου Πιθανοφανειών, το οποίο γενικά δεν μας οδηγεί σε ομοιόμορφα ισχυρότατους ελέγχους.

- Ένα τέτοιο παράδειγμα είναι ο έλεγχος των υποθέσεων $H_0 : \theta = \theta_0$ έναντι $H_0 : \theta \neq \theta_0$.

- Ορισμός: Έστω τυχαίο δείγμα X από κατανομή με άγνωστη παράμετρο $\theta \in \Theta$. Θέλουμε να πραγματοποιήσουμε τον έλεγχο των υποθέσεων:

$$H_0 : \theta \in \Theta_0$$
$$H_1 : \theta \in \Theta_1$$

όπου $\Theta_0 \cup \Theta_1 = \Theta$ και $\Theta_0 \cap \Theta_1 = \emptyset$.

- Ορίζουμε τη στατιστική συνάρτηση:

$$\lambda = \frac{\max_{\theta \in \Theta_0} L(\theta)}{\max_{\theta \in \Theta} L(\theta)},$$

η οποία καλείται **γενικευμένος λόγος πιθανοφανειών** (ΓΛΠ). Ισχύει ότι:

1. $\lambda \in [0,1], \quad \forall x \epsilon \mathbb{R}^n$.

2. εάν η μηδενική υπόθεση είναι αληθής, τότε $\lambda \simeq 1$.

- Πρόταση:

  Εάν υπάρχουν η εκτιμήτρια MLE $\hat{\theta}$ του θ και η εκτιμήτρια MLE $\hat{\theta}_0$ του θ υπό την $H_0 : \theta \in \Theta_0$, δηλαδή $L\left(\hat{\theta}_0\right) = \max_{\theta \in \Theta_0} L(\theta)$, τότε:

$$\lambda = \frac{L\left(\hat{\theta}_0\right)}{L\left(\hat{\theta}\right)}.$$

- Διαισθητικά, ο αριθμητής του γενικευμένου λόγου πιθανοφανειών εκφράζει τη μέγιστη δυνατή πιθανοφάνεια υπό την ισχύ της μηδενικής υπόθεσης, ενώ ο παρανομαστής εκφράζει τη μέγιστη δυνατή πιθανοφάνεια συνολικά.

- Αν ο αριθμητής είναι πολύ μικρότερος από τον παρανομαστή, δηλαδή ο λόγος λ παίρνει τιμές κοντά στο μηδέν, τότε είναι λιγότερο πιθανό να έχουμε παρατηρήσει τα δεδομένα που παρατηρήσαμε κάτω από τις τιμές της παραμέτρου θ που ανήκουν στο σύνολο Θ_0 σε σύγκριση με τιμές της παραμέτρου που ανήκουν γενικά στον παραμετρικό χώρο Θ, οπότε και απορρίπτουμε την H_0.

- Από το κριτήριο του ΓΛΠ δεν προκύπτει απαραίτητα βέλτιστος έλεγχος (δηλαδή με τη μεγαλύτερη ισχύ από όλους τους ελέγχους σε επίπεδο σημαντικότητας α, για κάθε $\theta \in \Theta_1$).

- Κάτι τέτοιο εξάλλου δεν θα ήταν δυνατό αφού αποδεικνύεται ότι υπάρχουν περιπτώσεις (π.χ. έλεγχοι αμφίπλευρων υποθέσεων) όπου δεν μπορεί να υπάρξει τέτοιος έλεγχος.

- Παρόλα αυτά το κριτήριο του ΓΛΠ είναι ένα εύχρηστο εργαλείο που στις περισσότερες περιπτώσεις δίνει αρκετά ικανοποιητικά αποτελέσματα.

- Πρόταση:

  Έστω ότι θέλουμε να πραγματοποιήσουμε τον έλεγχο των υποθέσεων:

$$H_0 : \theta = \theta_0$$
$$H_1 : \theta \neq \theta_0$$

  Τότε το κριτήριο του γενικευμένου λόγου πιθανοφανειών μας οδηγεί σε έναν έλεγχο με κρίσιμη περιοχή της μορφής:

$$C = \{x : T(X) < \kappa_1 \quad \text{ή} \quad T(X) > \kappa_2\}$$

  για κάποια στατιστική συνάρτηση $T(X)$.

- Στη συνέχεια σκοπός είναι να πραγματοποιήσουμε σύνθετους ελέγχους υποθέσεων για τις παραμέτρους ενός κανονικού πληθυσμού, κάνοντας χρήση του γενικευμένου λόγου πιθανοφανειών.

- Με γνώμονα αυτό, υποθέτουμε ότι διαθέτουμε ένα τυχαίο δείγμα $X_1, ..., X_n$ από την Κανονική κατανομή με μέση τιμή μ και διακύμανση σ^2.

- Παράδειγμα 1:

  Έστω ότι η διακύμανση σ^2 είναι γνωστή και ότι θέλουμε να πραγματοποιήσουμε τον έλεγχο υποθέσεων:

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$

  Οπότε ισχύει ότι:

$$H_0 : \mu \in \Theta_0 = \{\mu_0\}$$
$$H_1 : \mu \in \Theta_1 = (-\infty, \mu_0) \cup (\mu_0, \infty)$$

Σύμφωνα με το ΓΛΠ, η περιοχή απόρριψης θα είναι της μορφής:

$$C = \{x : T(X) < \kappa_1 \quad \text{ή} \quad T(X) > \kappa_2\},$$

με τον ΓΛΠ να δίνεται από:

$$\lambda = \frac{\max_{\mu \in \Theta_0} L(\mu)}{\max_{\mu \in \Theta_0 \cup \Theta_1} L(\mu)} = \frac{L(\mu_0)}{L(\hat{\mu})}$$

Έχουμε επίσης ότι:

$$\max_{H_0} L(\mu) = \max_{\mu = \mu_0} L(\mu) = L(\mu_0)$$

και

$$\max L(\mu) = \max_{\mu \in \mathbb{R}^n} L(\mu) = L(\bar{x}),$$

αφού γνωρίζουμε ότι: $\hat{\mu} = \bar{x}$ και επειδή το μ που μεγιστοποιεί την πιθανοφάνεια $L(\mu)$ (ή ισοδύναμα την λογαριθμική μορφή της) σε όλο τον παραμετρικό χώρο ($\mu \in \mathbb{R}^n$) είναι ίσο με την εκτιμήτρια μεγίστης πιθανοφάνειας του μ.

Γνωρίζουμε ότι η λογαριθμική συνάρτηση πιθανοφάνειας δίνεται από τη σχέση:

$$\mathcal{L}(\mu) = -\frac{n}{2}\log\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 .$$

Υπολογίζουμε:

$$\begin{aligned}
\log\lambda = \mathcal{L}(\mu_0) - \mathcal{L}(\hat{\mu}) &= -\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\mu_0)^2 - \sum_{i=1}^{n}(x_i-\bar{x})^2\right]\\
&= -\frac{1}{2\sigma^2}\left[n\mu_0^2 - 2\mu_0\sum_{i=1}^{n}x_i - n\bar{x}^2 + 2\bar{x}\sum_{i=1}^{n}x_i\right]\\
&= -\frac{1}{2\sigma^2}\left[n\mu_0^2 - 2n\mu_0\bar{x} - n\bar{x}^2 + 2n\bar{x}^2\right]\\
&= -\frac{n}{2\sigma^2}(\bar{x}-\mu_0)^2 \Rightarrow
\end{aligned}$$

$$\begin{aligned}
\lambda < \kappa &\iff \frac{n(\bar{x}-\mu_0)^2}{\sigma^2} > \kappa^* = -2\log\kappa\\
&\iff \frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} > z_{\alpha/2},
\end{aligned}$$

επειδή $Z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$ και:

$$
\begin{aligned}
P\left(\text{Απορρίψουμε } H_0 \mid H_0\right) = \alpha &\Rightarrow \\
P\left(\frac{|\bar{x}-\mu_0|}{\sigma/\sqrt{n}} > \kappa^*\right) = \alpha &\Rightarrow \\
P\left(Z > \kappa^*\right) + P\left(Z < -\kappa^*\right) = \alpha &\Rightarrow \\
1 - \Phi\left(\kappa^*\right) + \Phi\left(-\kappa^*\right) = \alpha &\Rightarrow \\
1 - \Phi\left(\kappa^*\right) + 1 - \Phi\left(\kappa^*\right) = \alpha &\Rightarrow \\
2\left[1 - \Phi\left(\kappa^*\right)\right] = \alpha &\Rightarrow \\
\Phi\left(\kappa^*\right) = 1 - \frac{\alpha}{2} &\Rightarrow \\
\Phi^{-1}\left(1 - \frac{\alpha}{2}\right) = z_{\frac{\alpha}{2}}&.
\end{aligned}
$$

- Συνοψίζοντας, τα βήματα που ακολουθήσαμε είναι:

1. ο προσδιορισμός του γενικευμένου λόγου πιθανοφανειών
2. η επίλυση της ανισότητας $\lambda < \kappa$ για τον προσδιορισμό της κρίσιμης περιοχής του ελέγχου.

- <u>Παράδειγμα 2:</u>

  Η υπόθεση που κάναμε στο προηγούμενο παράδειγμα ότι η διακύμανση είναι γνωστή δεν είναι μια ρεαλιστική υπόθεση. Στην πράξη, η διακύμανση του πληθυσμού είναι άγνωστη.

  Σε αυτό το παράδειγμα υποθέτουμε λοιπόν ότι η διακύμανση είναι άγνωστη και θέλουμε να πραγματοποιήσουμε τον έλεγχο:

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu \neq \mu_0$$

  δηλαδή να βρεθεί η κρίσιμη περιοχή του ελέγχου με βάση τον γενικευμένο λόγο πιθανοφανειών.

  Οπότε ισχύει ότι:

$$H_0 : \mu \in \Theta_0 = \left\{\mu_0, \sigma_0^2\right\} = \{\mu_0\} \times \mathbb{R}_+$$
$$H_1 : \mu \in \Theta_1 = (-\infty, \mu_0) \cup (\mu_0, \infty) \times \mathbb{R}_+$$

Γνωρίζουμε ότι

$$\hat{\mu} = \bar{x},$$

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

Υπό την H_0, δηλαδή όταν $\mu = \mu_0$ γνωστό:

$$\hat{\sigma}_0^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \mu_0\right)^2.$$

Για να το δείξουμε αυτό, ξαναγράφουμε τη λογαριθμική συνάρτηση πιθανοφάνειας (κάτω από την H_0):

$$\mathcal{L}\left(\mu, \sigma^2\right) = -\frac{n}{2}\log\left(2\pi\sigma_0^2\right) - \frac{1}{2\sigma_0^2}\sum_{i=1}^{n}\left(x_i - \mu_0\right)^2,$$

και έχουμε:

$$\frac{\partial \mathcal{L}\left(\mu, \sigma^2\right)}{\partial \sigma^2} = 0 \Rightarrow -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (x_i - \mu_0)^2 = 0 \Rightarrow$$

$$\hat{\sigma}_0^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu_0)^2 .$$

Υπολογίζουμε ότι:

$$\begin{aligned}
\log \lambda &= \mathcal{L}\left(\mu_0, \hat{\sigma}_0^2\right) - \mathcal{L}\left(\hat{\mu}, \hat{\sigma}^2\right) \\
&= -\frac{n}{2} \log \frac{\hat{\sigma}_0^2}{\hat{\sigma}^2} - \frac{1}{2\hat{\sigma}_0^2} \sum_{i=1}^{n} (x_i - \mu_0)^2 + \frac{1}{2\hat{\sigma}^2} \sum_{i=1}^{n} (x_i - \bar{x})^2 \\
&= -\frac{n}{2} \log \frac{\hat{\sigma}_0^2}{\hat{\sigma}^2} - \frac{n}{2} + \frac{n}{2} \\
&= -\frac{n}{2} \log \frac{\hat{\sigma}_0^2}{\hat{\sigma}^2},
\end{aligned}$$

όπου:

$$\begin{aligned}
\hat{\sigma}_0^2 &= \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x} + \bar{x} - \mu_0\right)^2 \\
&= \frac{1}{n}\left[\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 + n\left(\bar{x} - \mu_0\right)^2 + 2\left(\bar{x} - \mu_0\right)\sum_{i=1}^{n}\left(x_i - \bar{x}\right)\right] \\
&= \hat{\sigma}^2 + \left(\bar{x} - \mu_0\right)^2 + \frac{2\left(\bar{x} - \mu_0\right)}{n}\left(\sum_{i=1}^{n} x_i - n\bar{x}\right) \\
&= \hat{\sigma}^2 + \left(\bar{x} - \mu_0\right)^2.
\end{aligned}$$

Άρα:

$$\log \lambda = -\frac{n}{2} \log \left[1 + \frac{(\bar{x} - \mu_0)^2}{\hat{\sigma}^2}\right] = -\frac{n}{2} \log \left[1 + \frac{n\,(\bar{x} - \mu_0)^2}{(n-1)\,S^2}\right] \Rightarrow$$

$$\lambda < \kappa \Longleftrightarrow \log \lambda < \kappa^* = \log \kappa \Longleftrightarrow$$

$$1 + \frac{n\,(\bar{x} - \mu_0)^2}{(n-1)\,S^2} > \kappa^{**} = e^{-\frac{2\kappa^*}{n}} \Longleftrightarrow$$

$$\frac{n\,(\bar{x} - \mu_0)^2}{S^2} > \kappa^{***} = (n-1)\,(\kappa^{**} - 1) \Longleftrightarrow$$

$$\frac{|\bar{x} - \mu_0|}{S/\sqrt{n}} > t_{\alpha/2}.$$

Υπό την H_0, γνωρίζουμε ότι $T = \frac{\bar{x}-\mu_0}{S/\sqrt{n}} \sim t_{n-1}$.

$$
\begin{aligned}
P\left(\text{Απορρίψουμε } H_0 \mid H_0\right) = \alpha \Rightarrow \\
P\left(\frac{|\bar{x}-\mu_0|}{S/\sqrt{n}} > \kappa^{***}\right) = \alpha \Rightarrow \\
P\left(T > \kappa^{***}\right) + P\left(T < -\kappa^{***}\right) = \alpha \Rightarrow \\
1 - F\left(\kappa^{***}\right) + F\left(-\kappa^{***}\right) = \alpha \Rightarrow \\
2\left[1 - F\left(\kappa^{***}\right)\right] = \alpha \Rightarrow \\
F\left(\kappa^{***}\right) = 1 - \frac{\alpha}{2} \Rightarrow \\
F^{-1}\left(1 - \frac{\alpha}{2}\right) = t_{n-1,\frac{\alpha}{2}}.
\end{aligned}
$$

- <u>Παράδειγμα 3:</u>

Σε αυτό το παράδειγμα ο μέσος του πληθυσμού $\mu = \mu_0$ είναι γνωστός και θέλουμε να πραγματοποιήσουμε τον έλεγχο:

$$\begin{aligned} H_0 &: \sigma^2 = \sigma_0^2 \\ H_1 &: \sigma^2 \neq \sigma_0^2 \end{aligned}$$

Οπότε ισχύει ότι:

$$\begin{aligned} H_0 &: \sigma^2 \in \Theta_0 = \left\{\sigma_0^2\right\} \\ H_1 &: \sigma^2 \in \Theta_1 = \mathbb{R}_+ \setminus \left\{\sigma_0^2\right\} \end{aligned}$$

Σύμφωνα με το ΓΛΠ, η περιοχή απόρριψης θα είναι της μορφής:

$$C = \{x : T(X) < \kappa_1 \quad \acute{\eta} \quad T(X) > \kappa_2\},$$

με τον ΓΛΠ να δίνεται από:

$$\lambda = \frac{\max_{\sigma^2 \in \Theta_0} L\left(\sigma^2\right)}{\max_{\sigma^2 \in \Theta_0 \cup \Theta_1} L\left(\sigma^2\right)} = \frac{L\left(\sigma_0^2\right)}{L\left(\hat{\sigma}_0^2\right)}.$$

Έχουμε επίσης ότι:

$$\max_{H_0} L\left(\sigma^2\right) = \max_{\sigma^2=\sigma_0^2} L\left(\sigma^2\right) = L\left(\sigma_0^2\right)$$

και

$$\max L\left(\sigma^2\right) = \max_{\sigma^2>0} L\left(\sigma^2\right) = L\left(\hat{\sigma}_0^2\right),$$

όπου $\hat{\sigma}_0^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \mu_0\right)^2$.

Γνωρίζουμε ότι η συνάρτηση πιθανοφάνειας δίνεται από τη σχέση:

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

Υπολογίζουμε:

$$\begin{aligned}
\lambda = \frac{L\left(\sigma_0^2\right)}{L\left(\hat{\sigma}_0^2\right)} &= \left(\frac{\hat{\sigma}_0^2}{\sigma_0^2}\right)^{n/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu_0}{\sigma_0}\right)^2 + \frac{n}{2}\right\} \\
&= T^{n/2}\exp\left\{\frac{n}{2}\left(1-T\right)\right\} \\
&= \left(Te^{1-T}\right)^{n/2}
\end{aligned}$$

όπου

$$T = \frac{\hat{\sigma}_0^2}{\sigma_0^2} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\mu_0}{\sigma_0}\right)^2 \Rightarrow$$

$$\sum_{i=1}^{n}\left(\frac{x_i-\mu_0}{\sigma_0}\right)^2 = nT.$$

Συνεπώς:

$$\lambda < \kappa \Longleftrightarrow T < \kappa_1 \quad \text{ή} \quad T > \kappa_2,$$

Υπό την H_0, γνωρίζουμε ότι $T = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\mu_0}{\sigma_0}\right)^2 \sim \mathcal{X}_n^2$. Οπότε:

$$\kappa_1 = \mathcal{X}^2_{1-\alpha/2,n},$$
$$\kappa_2 = \mathcal{X}^2_{\alpha/2,n}.$$

(βλ. και Θεωρία 09, σελ. 24-25)