Κεφάλαιο 7

Εκτίμησης παραμέτρων στην πράξη

7.1 Επιλογή της μεθόδου εκτίμησης

Στο Κεφάλαιο 6 συζητήσαμε μεθόδους εκτίμησης παραμέτρων. Εν γένει διαφορετικές μέθοδοι οδηγούν σε διαφορετικούς εκτιμητές. Προκειμένου να επιλέξουμε την πλέον ενδεδειγμένη μέθοδο, ιδιαίτερα εάν πρόκειται να αναλύσουμε μεγάλο όγκο πειραματικών δεδομένων, θα πρέπει να συνυπολογίσουμε πλήθος παραγόντων που αφορούν τις ιδιότητες των εκτιμητών αλλά και την δυνατότητα πρακτικής εφαρμογής της μεθόδου αλλά και των αποτελεσμάτων της εκτίμησης. Τα ακόλουθα πέντε κριτήρια επιλογής[1] μεθόδου εκτίμησης δεν είναι τα μοναδικά αλλά συμπυκνώνουν τις σημαντικότερες «αρετές» τις οποίες επιζητούμε συνήθως από έναν εκτιμητή.

1. **Συνέπεια και προκατάληψη:** Ο εκτιμητής θα πρέπει να είναι συνεπής ώστε η εκτίμηση να τείνει προς την αληθή τιμή των παραμέτρων καθώς το πλήθος των πειραματικών δεδομένων αυξάνει. Ο εκτιμητής θα πρέπει να είναι απαλλαγμένος από προκατάληψη. Στην περίπτωση όπου η εκτίμηση είναι προκατελημένη, η προκατάληψη θα πρέπει να εκτιμάται και η εκτίμηση να διορθώνεται, όπως θα δείξουμε στην επόμενη υποενότητα.
2. **Περιεχόμενο σε Πληροφορία:** Συνήθως με την μέθοδο της εκτίμησης παραμέτρων επιδιώκουμε την συμπύκνωση της πληροφορίας, η οποία εμπεριέχεται στα διαθέσιμα πειραματικά δεδομένα αναφορικά με την τιμή κάποιων φυσικών μεγεθών, σε ένα ή μερικά νούμερα. Συνεπώς το αποτέλεσμα της εκτίμησης θα πρέπει θα πρέπει να περιέχει όλη την δυνατή πληροφορία[2].
3. **Αποδοτικότητα:** Επιδιώκουμε την επιλογή εκτιμητών με την ελάχιστη δυνατή διασπορά. Όσο μικρότερη είναι η διασπορά της εκτίμησης τόσο ακριβέστερη είναι η προσέγγιση της αληθούς τιμής.
4. **Εξάρτηση από παραδοχές.** Στις περισσότερες των περιπτώσεων, είμαστε αναγκασμένοι να προβούμε σε υποθέσεις, π.χ. όσον αφορά την συναρτησιακή μορφή της πυκνότητας πιθανότητας ή το πρότυπο προσαρμογής. Επιδιώκουμε να επιλέγουμε τέτοιες συναρτησιακές μορφές ή να επιβάλουμε περιορισμούς (π.χ. πεδίο τιμών των μεταβλητών) με τρόπο που δεν θα «καταστρέψει» εντελώς την διαδικασία εκτίμησης[3].

---

[1] Τα κριτήρια αυτά προτείνοντα από τους Eadie et all, Section 8.1, page 143

[2] Επί παραδείγματι, ας υποθέσουμε ότι θέλουμε να εκτιμήσουμε το σφάλμα μέτρησης μίας πειραματικής συσκευής, η οποία μετράει τις τιμές του φυσικού μεγέθους X, διαθέτοντας N μετρήσεις x_i (i=1,2,3,…N) και τις αντίστοιχες αληθείς τιμές του φυσικού μεγέθους. Είναι προφανές ότι θα πρέπει να εξετάσουμε την περίπτωση όπου το μετρητικό σφάλμα της συσκευής εξαρτάται από την αληθή τιμή του φυσικού μεγέθους. Η πληροφορία αυτή εμπεριέχεται στα πειραματικά δεδομένα και η επιλεγείσα μέθοδος εκτίμησης θα πρέπει να έχει την δυνατότητα να προσδιορίσει το μετρητικό σφάλμα ως συνάρτηση της αληθούς τιμής.

[3] Ας υποθέσουμε ότι οι τιμές του φυσικού μεγέθους Y εξαρτάται από το μέγεθος X με σχέση εκθετικής μορφής: $e^{-a\cdot x}$. Εάν επιλέξουμε ως πρότυπο προσαρμογής με την μέθοδο των ελαχίστων τετραγώνων της μορφής: $\sin(b\cdot x + c)$ και επιδιώξουμε να εκτιμήσουμε τις παραμέτρους b και c, κατά πάσα πιθανότητα, θα αποτύχουμε. Αντίθετα, εάν επιλέξουμε ως πρότυπο προσαρμογής την πολυονυμική συνάρτηση : $\sum_{i=0}^{k} w_i \cdot x^i$ μπορούμε να εκτιμήσουμε τις παραμέτρους w_i και να περιγράψουμε με ικανοποιητική ακρίβεια την εξάρτηση των μεγεθών Y και X. Παρατηρήσετε ότι η ανάλυση σε σειρά Teylor του αληθούς προτύπου έχει πολυονυμική μορφή. Αρκεί το πλήθος των πειραματικών σημείων να είναι αρκετά μεγάλο ώστε να μπορέσουμε να χρησιμοποιήσουμε ικανοποιητικό αριθμό όρων,k.

Βεβαίως, η γενικότητα στην επιλογή συναρτήσεων και περιορισμών θα αποβεί εις βάρος της ακρίβειας (σφάλμα) της εκτίμησης.

5. **Ελάχιστο κόστος σε χρόνο και απλότητα:** Επιδιώκουμε την χρήση μεθόδων που η εφαρμογή τους δεν απαιτεί μεγάλο χρόνο επεξεργασίας ή ιδιαίτερη δυσκολία στην εκτέλεση της επεξεργασίας. Επίσης, το αποτέλεσμα της εκτίμησης θα πρέπει να συνοδεύεται από τις ελάχιστες παραδοχές (π.χ. για την ισχύ φυσικών θεωριών και μοντέλων) και πάντοτε θα πρέπει να συνοδεύεται από τις απαραίτητες επεξηγήσεις ώστε να είναι εύκολο από κάθε άλλο ενδιαφερόμενο να το χρησιμοποιήσει.

Επισημαίνεται, ότι δεν υπάρχουν γενικές οδηγίες για την επιλογή μεθόδου εκτίμησης σε κάθε πρόβλημα ανάλυσης πειραματικών δεδομένων. Εν τέλει, η επιλογή και η χρήση μεθόδων εκτίμησης επαφίονται στην εμπειρία και δημιουργικότητα του επιστήμονα.

7.2 Άρση της προκατάληψης.

Έστω ότι διαθέτουμε N μετρήσεις της ποσότητας X, x_i (i=1,2,3,…N), με συνάρτηση πυκνότητας πιθανότητας $f(x;\theta)$. Έστω επίσης η συνεπής εκτίμηση $\hat{\theta}$ της παραμέτρου θ, με αληθή τιμή θ_0, η οποία όμως πάσχει από προκατάληψη: $b = E\left[\hat{\theta}\right] - \theta_0$

Εάν η συνάρτηση πυκνότητας πιθανότητας είναι γνωστή, μπορούμε να υπολογίσουμε την ποσότητα b και στην συνέχεια να ορίσουμε τον μη προκαταλημένο εκτιμητή:

$$\hat{\theta}_1 = \hat{\theta} - b \qquad 7.2.1$$

Στην περίπτωση αυτή η διασπορά της εκτίμησης θα παραμένει η ίδια:

$$V\left[\hat{\theta}_1\right] = V\left[\hat{\theta}\right] - V\left[b\right] = V\left[\hat{\theta}\right]$$

αρκεί να γνωρίζουμε την αληθή τιμή της παραμέτρου ώστε να υπολογίσουμε την προκατάληψη χωρίς σφάλμα. Ωστόσο, σε κάθε πρακτική εφαρμογή, η αληθής τιμή θ_0 δεν είναι γνωστή. Το μόνο που μπορούμε να κάνουμε είναι να εκτιμήσουμε την προκατάληψη, οπότε το σφάλμα της εκτίμησης της προκατάληψης συμμετέχει στην αύξηση του σφάλματος της εκτίμησης της παραμέτρου.

Παράδειγμα 7.2.1:

Έχουμε δείξει ότι η σύγχρονη εκτίμηση, με την μέθοδο της μεγίστης πιθανοφάνειας, των παραμέτρων μ και σ της κανονικής συνάρτησης πιθανότητας των μετρήσεων $\{x_1, x_2, K\ , x_N\}$, οδηγεί σε εκτιμήσεις $\hat{\mu}$ και $\hat{\sigma}$ εκ των οποίων η δεύτερη:

$$\hat{\sigma}^2 = \frac{1}{N} \cdot \sum_{i=1}^{N} \left(< x > - x_i \right)^2 \quad , \quad E\left[\hat{\sigma}^2\right] = \sigma^2 - \frac{\sigma^2}{N}$$

πάσχει από προκατάληψη.

Προκειμένου να ορίσουμε νέο εκτιμητή χωρίς προκατάληψη εφαρμόζουμε την σχέση 7.2.1:

$$\hat{s}^2 = \frac{1}{N-1} \cdot \sum_{i=1}^{N} \left(< x > - x_i \right)^2 = \hat{\sigma}^2 + \frac{\hat{\sigma}^2}{N} \qquad 7.2.2$$

Στην περίπτωση της κανονικής συνάρτησης πυκνότητας πιθανότητας βρίσκεται ότι:

$$V\left[\hat{\sigma}^2\right] = 2 \cdot \sigma^4 \cdot \frac{N-1}{N^2} \qquad 7.2.3$$

Συνδυάζοντας τις σχέσεις 7.2.2 και 7.2.3 καταλήγουμε ότι:

$$V\left[\hat{s}^2\right] = 2\cdot\sigma^4\cdot\frac{1}{N-1} \qquad 7.2.4$$

Η σχέση 7.2.4 αντιστοιχεί σε αύξηση του σφάλματος στην εκτίμηση κατά:

$$\sqrt{V\left[\hat{s}^2\right]} - \sqrt{V\left[\hat{\sigma}^2\right]} = \frac{\sigma^2}{N}\cdot\sqrt{\frac{2}{N-1}} \qquad 7.2.5$$

Ερώτηση: Δείξετε ότι για N>3 η αύξηση 7.2.5 του σφάλματος είναι μικρότερη από την προκατάληψη.

Στην περίπτωση όπου η συνάρτηση πυκνότητας πιθανότητας των δεδομένων δεν είναι γνωστοί υπάρχουν διάφοροι τρόποι για την μείωση της προκατάληψης. Στο επόμενα αναφέρεται η πλέον εύχρηστη από αυτές τις μεθόδους.

Αναλύοντας[4] την αναμενόμενη τιμή της εκτίμησης σε σειρά, ως προς 1/N, εκφράζουμε ως:

$$E\left[\hat{\theta}_N\right] = \theta_0 + \frac{1}{N}\cdot\beta + d\,(\frac{1}{N^2}) \qquad 7.2.6$$

όπου χρησιμοποιήσαμε το σύμβολο $\hat{\theta}_N$ για να δηλώσουμε ότι η εκτίμηση επετεύχθη με N μετρήσεις. Παρατηρήστε ότι η κύρια συνεισφορά στην απόκλιση από την αληθή τιμή, προέρχεται από τον όρο που είναι ανάλογος το 1/N.

Ας χωρίσουμε τα δεδομένα σε δύο ομάδες και ας επαναλάβουμε την εκτίμηση της ιδίας παραμέτρου, με τα δύο ήμισυ.

$$E\left[\hat{\theta}^1_{N/2}\right] = E\left[\hat{\theta}^2_{N/2}\right] = \theta_0 + \frac{2}{N}\cdot\beta + d\,(\frac{1}{N^2}) \qquad 7.2.7$$

Συνεπώς:

$$E\left[2\cdot\hat{\theta}_N - \frac{1}{2}\cdot\left(\hat{\theta}^1_{N/2} + \hat{\theta}^2_{N/2}\right)\right] = \theta_0 + d\,(\frac{1}{N^2}) \quad 7.28$$

Δηλαδή ο εκτιμητής $2\cdot\hat{\theta}_N - \frac{1}{2}\cdot\left(\hat{\theta}^1_{N/2} + \hat{\theta}^2_{N/2}\right)$ έχει μικρότερη προκατάληψη (τάξης $\frac{1}{N^2}$) από τον εκτιμητή $\hat{\theta}_N$.

Ένας άλλος εκτιμητής, γενικής χρήσης, είναι ο $N\cdot\hat{\theta}_N - (N-1)\cdot<\hat{\theta}'_{N-1}>$, όπου $<\hat{\theta}'_{N-1}>$ είναι ο μέσος όρος των εκτιμήσεων που επιτυγχάνονται από όλους τους συνδυασμούς N-1 γεγονότων.

7.3 Εκτίμηση παραμέτρων με την μέθοδο των ορθογωνίων συναρτήσεων.

Μία ειδική περίπτωση της μεθόδου των ροπών προκύπτει όταν η συνάρτηση πυκνότητας πιθανότητας των παρατηρήσεων, f(x;θ), μπορεί να εκφρασθεί ως ανάπτυγμα ορθογωνίων συναρτήσεων g_k (x).

Σ΄ αυτή την περίπτωση η συνάρτηση πυκνότητας πιθανότητας θα έχει την μορφή:

[4] Βλέπε Eadie et al, Section 7.3.2 page 126

$$f(x;\underline{\theta}) = a + \sum_{k=1}^{r} \theta_k \cdot g_k(x)$$

$$\int g_j(x) \cdot g_i(x) dx = \delta_{ij} \qquad 7.3.1$$

$$\int g_i(x) dx = 0$$

Η αναμενόμενη τιμή των ορθογωνίων συναρτήσεων, $g_k(x)$, είναι:

$$E\left[g_k(x)\right] = \int g_k(x) \cdot \underbrace{\left[a + \sum_{\lambda=1}^{r} \theta_\lambda \cdot g_k(x)\right]}_{f(x;\underline{\theta})} dx$$

$$= a \cdot \int g_k(x) dx + \sum_{\lambda=1}^{r} \left(\theta_\lambda \cdot \int g_k(x) \cdot g_\lambda(x) dx\right) \qquad 7.3.2$$

$$= \theta_k$$

Οι αναμενόμενες τιμές, $E\left[g_k(x)\right]$, μπορούν να υπολογισθούν προσεγγιστικά από τις πειραματικές παρατηρήσεις και τον νόμο των μεγάλων αριθμών ως:

$$E\left[g_k(x)\right] = \theta_k \quad \rightarrow \quad \hat{\theta}_k = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} g_k(x_i) \qquad 7.3.3$$

Επειδή η εκτίμηση της παραμέτρου θ_k ανάγεται, σε άθροισμα N τυχαίων μεταβλητών, $g_k(x_i)$ i=1,2,..N, η εκτίμηση, $\hat{\theta}_k$,θα κατανέμεται σύμφωνα με κανονική συνάρτηση πιθανότητας, με μέση τιμή ίση με την αναμενόμενη τιμή της ορθογώνιας συνάρτησης:

$$E\left[\frac{1}{N} \sum_{i=1}^{N} g_k(x_i)\right] = E\left[g_k(x)\right] = \theta_k \quad 7.3.4$$

όπως υπαγορεύει το θεώρημα του κεντρικού ορίου.. Η σχέση 7.3.4 δηλώνει επίσης ότι η εκτίμηση, με την μέθοδο των ορθογωνίων συναρτήσεων, είναι συνεπής.

Η διασπορά των εκτιμήσεων (δηλαδή η ακρίβεια) θα εκφράζεται ως το ακόλουθο άθροισμα :

$$V(\hat{\theta}_k) = \frac{1}{N} \sum_{i} V\left[g_k(x_i)\right]$$

όπου

$$V\left[g_k(x_i)\right] = \frac{1}{N-1} \sum_{j=1}^{N} \left[g_k(x_j) - \hat{\theta}_k\right]^2$$

Τελικά η διασπορά της εκτίμησης θ_k θα είναι :

$$V\left[\hat{\theta}_k\right] = \frac{1}{N \cdot (N-1)} \cdot \sum_{i=1}^{N} \left[\hat{\theta}_k - g_k(x_i)\right]^2 \qquad 7.3.5$$

Παράδειγμα 7.3.1:

Η έκφραση f (x|θ) = α + Σ θ_k $g_k(x)$, με τις $g_k(x)$ να είναι ορθογώνιες συναρτήσεις, προκύπτει πολύ συχνά στην μελέτη γωνιακών κατανομών, λόγω διατήρησης της στροφορμής.

Επί παραδείγματι, κατά την διάσπαση του βαρυονίου Λ^0 , σε ένα πρωτόνιο (p) και ένα πιόνιο (π), το συνημίτονο της γωνίας (cosφ) που σχηματίζει η διεύθυνση κίνησης του πιονίου στο σύστημα κέντρου μάζας του Λ^0 με την διεύθυνση κίνησης του Λ^0 στο σύστημα εργαστηρίου, χαρακτηρίζεται από συνάρτηση πυκνότητας πιθανότητας της μορφής: $f(\cos\phi; P) = \frac{1}{2}\cdot\left[1 + a\cdot P\cdot\cos\phi\right]$

όπου P είναι η παράμετρος πόλωσης και a είναι η παράμετρος ασυμμετρίας, την οποία θα θεωρήσουμε ότι είναι γνωστή.

Ας μελετήσουμε την συνάρτηση g: $g(\cos\phi) = R\cdot\cos\phi$

αναζητώντας την τιμή της παραμέτρου ώστε να πληρούνται οι απαιτήσεις της σχέσεις 7.3.1.

$$\left.\begin{aligned} &\int_{-1}^{1}\left(R\cdot\cos\varphi\right)d\left(\cos\varphi\right) = 0 \\ &\int_{-1}^{1}\left(R\cdot\cos\varphi\right)\cdot\left(R\cdot\cos\varphi\right)d(\cos\varphi) = \frac{R^2}{3}2 = 1 \end{aligned}\right\} \Rightarrow g\left(\cos\varphi\right) = \frac{\sqrt{3}\cdot\cos\varphi}{\sqrt{2}}$$

Επιλέγοντας $R=(3/2)^{1/2}$, η συνάρτηση, g(cosφ) =R·cosφ, είναι ορθογώνια και η συνάρτηση πυκνότητας πιθανότητας γράφεται ως:

$$f(\cos\varphi;P) = \left[\frac{1}{2} + a\cdot P\cdot\cos\varphi\cdot\frac{\sqrt{3}}{\sqrt{3}}\cdot\frac{1}{\sqrt{2}\cdot\sqrt{2}}\right] = \left[\frac{1}{2} + \frac{a\cdot P}{\sqrt{3}\cdot\sqrt{2}}\cdot g\left(\cos\varphi\right)\right]$$

Κατά συνέπεια ο πολλαπλασιαστικός όρος $\frac{\alpha\cdot P}{\sqrt{3}\cdot\sqrt{2}}$ και κατ΄ επέκταση ο συντελεστής πόλωσης P, μπορεί να εκτιμηθεί ως:

$$\left(\frac{a\cdot P}{\sqrt{3}\cdot\sqrt{2}}\right) = E\left(\frac{\sqrt{3}}{\sqrt{2}}\cdot\cos\phi\right) \cong \sum_{i=1}^{N}\frac{\sqrt{3}}{\sqrt{2}}\frac{\cos\phi_i}{N} \qquad 7.3.6$$

$$\hat{P} = \frac{3}{a\cdot N}\sum\cos\phi_i$$

Για απλούστευση ας γράψουμε την συνάρτηση πυκνότητας πιθανότητας του προηγουμένου παραδείγματος ως f (x;θ) = ½ (1 + θ x), όπου -1≤ x ≤ 1 και x = cosφ. Η παράμετρος θ έχει απορροφήσει τους σταθερούς όρους και τις παραμέτρους πόλωσης και ασυμμετρίας. Εφαρμόζοντας τις σχέσεις 7.3.5 και 7.3.6, η διασπορά της εκτίμησης της

παραμέτρου θ, με την μέθοδο των ορθογωνίων συναρτήσεων, βρίσκεται ότι θα είναι ίση με $(3-\theta^2)/N$.

Ας επαναλάβουμε την εκτίμηση της ιδίας παραμέτρου με την μέθοδο της μέγιστης πιθανοφάνειας . Χρησιμοποιώντας τις ασυμπτωτικές ιδιότητες της μεθόδου βρίσκουμε ότι:

$$
\begin{aligned}
V\left[\hat{\theta}\right] &= -E\left[\frac{\partial^2 \ln L}{\partial \theta^2}\right]^{-1}_{\theta=\hat{\theta}} \\
&= \left(-E\left[\frac{d^2}{d\theta^2}\ln\left(\prod_i \frac{1}{2}(1+\theta\cdot x_i)\right)\right]\right)^{-1}_{\theta=\hat{\theta}} \\
&= \frac{1}{N}\left(-E\left[\frac{d^2}{d\theta^2}\ln(1+\theta\cdot x_i)\right]\right)^{-1}_{\theta=\hat{\theta}} \\
&\cong \frac{1}{N}\left[\frac{1}{\theta^2}\left[-1+\frac{1}{2\theta}\ln\frac{1+\theta}{1-\theta}\right]\right]^{-1} \\
&\cong \frac{1}{N}\left(3-\frac{9}{5}\hat{\theta}^2+K\right) < \frac{1}{N}(3-\hat{\theta}^2)
\end{aligned}
$$

Δηλαδή, η διασπορά της εκτίμησης με την μέθοδο της μεγίστης πιθανοφάνειας είναι μικρότερη της διασποράς που αντιστοιχεί στην μέθοδο των ορθογωνίων συναρτήσεων.

7.3 Εκτεταμένη μέθοδος μεγίστης πιθανοφάνειας

Η συνάρτηση της πιθανοφάνειας εξαρτάται αποκλειστικά από την συναρτησιακή μορφή της συνάρτησης πυκνότητας πιθανότητας των τυχαίων μεταβλητών που χαρακτηρίζουν το πρόβλημα. Δεν μεταφέρει καμία πληροφορία αναφορικά με τον αριθμό των γεγονότων που αναμένει να συλλέξει κανείς στο συγκεκριμένο πείραμα, έστω και εάν το αναμενόμενο πλήθος γεγονότων εξαρτάται από τις υπό εκτίμηση παραμέτρους

Παράδειγμα 7.3.1

Ας υποθέσουμε ότι σε ένα πείραμα ανιχνεύονται ακτίνες Χ και μετρείται η ενέργεια, Ε, κάθε ενός από τα φωτόνια που ανιχνεύονται στην ενεργειακή περιοχή $\left[E_1, E_1\right]$. Οι ακτίνες Χ παράγοντα από δύο διαφορετικούς φυσικούς μηχανισμούς, Α και Β. Η ενεργειακή κατανομή των φωτονίων τύπου Α χαρακτηρίζεται από την συνάρτηση πυκνότητας πιθανότητας $f_A(E)$ και τα φωτόνια από τον μηχανισμό Β έχουν ενέργειες σύμφωνα με την συνάρτηση πυκνότητας πιθανότητας $f_B(E)$. Προφανώς:

$$\int_{E_1}^{E_2} f_A\left(E\right)dE = 1$$

$$\int_{E_1}^{E_2} f_B\left(E\right)dE = 1 \qquad 7.3.1$$

Επιπλέον, ο αναμενόμενος αριθμός των φωτονίων τύπου Β που αντιστοιχεί στο χρόνο διάρκειας του πειράματος είναι μ_B ενώ ο αντίστοιχος αριθμός φωτονίων τύπου Α δίνεται από την ακόλουθη σχέση:

$$\mu_A = \theta \cdot R \qquad 7.3.2$$

όπου R είναι γνωστή σταθερά και η παράμετρος θ θα εκτιμηθεί από τα πειραματικά δεδομένα.

Τα πειραματικά δεδομένα συνίστανται από μετρήσεις της ενέργειας N φωτονίων: $\underline{E} = \left\{E_1, E_1, K\ , E_N\right\}$.

Η πιθανότητα ώστε ένα από τα φωτόνια που ανιχνεύθηκαν, ανεξαρτήτως της ενέργειάς του, να προέρχεται από τον μηχανισμό Α, είναι:

$$w_A = \frac{\mu_A}{\mu_A + \mu_B} = \frac{\theta \cdot R}{\theta \cdot R + \mu_B} \qquad 7.3.3$$

Η αντίστοιχη πιθανότητα για τα φωτόνια τύπου Β δίνεται από την σχέση:

$$w_B = \frac{\mu_B}{\mu_A + \mu_B} = \frac{\mu_B}{\theta \cdot R + \mu_B} \qquad 7.3.4$$

Οι πιθανότητες ώστε ένα από τα φωτόνια που ανιχνεύθηκαν, στην ενεργειακή περιοχή [E, E +dE], να προέρχεται από τον μηχανισμό Α ή Β, θα εκφράζεται ώς:

$$P_A(E) = w_A \cdot f_A(E)dE = \frac{\theta \cdot R}{\theta \cdot R + \mu_B} \cdot f_A(E)dE$$

$$P_B(E) = w_B \cdot f_B(E)dE = \frac{\mu_B}{\theta \cdot R + \mu_B} \cdot f_B(E)dE \qquad 7.3.5$$

Συνεπώς, η συνάρτηση πυκνότητας πιθανότητας, F(E), που περιγράφει τα πειραματικά δεδομένα είναι:

$$F(E) = \frac{\theta \cdot R}{\theta \cdot R + \mu_B} \cdot f_A(E) + \frac{\mu_B}{\theta \cdot R + \mu_B} \cdot f_B(E)$$

$$\int_{E_1}^{E_2} F(E)dE = 1 \qquad 7.3.6$$

Η συνάρτηση πιθανοφάνειας των πειραματικών δεδομένων εκφράζεται ως:

$$L(\underline{E};\theta) = \prod_{i=1}^{N} \left[\frac{\theta \cdot R}{\theta \cdot R + \mu_B} \cdot f_A(E_i) + \frac{\mu_B}{\theta \cdot R + \mu_B} \cdot f_B(E_i) \right] \qquad 7.3.7$$

Παρατηρήστε ότι η πιθανοφάνεια (7.3.7) εξαρτάται από την παράμετρο θ μόνο όταν οι δύο φυσικοί μηχανισμοί παράγουν ακτίνες X με διαφορετικές ενεργειακές κατανομές.
Επί παραδείγματι, εάν $f_A(E) = f_B(E)$ η σχέση 7.3.6 και η συνάρτηση πιθανοφάνειας (7.3.7) δεν μεταφέρουν καμία πληροφορία για την παράμετρο θ ($F(E) = f_A(E) = f_B(E)$).

Ωστόσο, μπορούμε να αντλήσουμε πληροφορία για την παράμετρο θ από το πλήθος των φωτονίων που ανιχνεύθηκαν. Η πιθανότητα ώστε τα N φωτόνια που ανιχνεύθηκαν να εμπεριέχουν φωτόνια από τον μηχανισμό A, με τιμή της παραμέτρου ίση με θ, θα δίνεται από την ακόλουθη συνάρτηση πιθανότητας:

$$P_N = \frac{e^{-(\theta \cdot R + \mu_B)} \cdot (\theta \cdot R + \mu_B)^N}{N!} \qquad 7.3.8$$

Όπου $\theta \cdot R + \mu_B$ είναι ο αναμενόμενος αριθμός φωτονίων στον ανιχνευτή και από τους δύο μηχανισμούς..

Η μεγιστοποίηση της πιθανότητας 7.3.8, κατά αντίστοιχο τρόπο της μεθόδου μεγίστης πιθανοφάνειας, καταλήγει στην εκτίμηση της αληθούς τιμής της παραμέτρου θ.

Προφανώς, η σχέση 7.3.8 εκφράζει την πιθανότητα ανίχνευσης N φωτονίων ανεξάρτητα της μορφής των ενεργειακών κατανομών, $f_A(E)$ και $f_B(E)$. Κατά συνέπεια, ο συνδυασμός των συναρτήσεων πιθανοφάνειας, 7.3.7 και 7.3.8, αναμένεται να ενισχύσει την ακρίβεια της εκτίμησης.

Προκειμένου να συμπεριλάβουμε στη εκτίμηση των παραμέτρων και την πληροφορία που εμπεριέχει ο συνολικός αριθμός των δεδομένων, πολλαπλασιάζουμε τη συνάρτηση πιθανοφάνειας με την πιθανότητα ανίχνευσης N γεγονότων, η οποία εκφράζεται από Poissonian συνάρτηση πιθανότητας,.

$$L(\underline{x};\theta)=\prod_{i=1}^{N} f(x_i,\theta) \rightarrow L_E(\underline{x};\theta)=\frac{\left(\nu(\theta)\right)^N\cdot e^{-\nu(\theta)}}{N!}\cdot\prod_{i=1}^{N} f(x_i,\theta)$$

$$\int L(\underline{x};\theta)dx_1dx_2\,\mathrm{L}\,dx_N=1 \quad\rightarrow\quad \sum_{N=0}^{\infty}\int L_E(\underline{x};\theta)dx_1dx_2\,\mathrm{L}\,dx_N=1 \qquad 7.3.9$$

όπου, $\underline{x}=\{x_1,x_2,\mathrm{K}\ ,x_N\}$ είναι οι N μετρήσεις και ν(θ) είναι η αναμενόμενη τιμή του πλήθους των μετρήσεων, η οποία εν γένει είναι συνάρτηση της παραμέτρου θ. Η στατιστική συνάρτηση $L_E(\underline{x};\theta)$, που ορίσθηκε στην σχέση 7.3.9, καλείται εκτεταμένη (extended) συνάρτηση πιθανοφάνειας.

Κατ΄ αντιστοιχία της σχέσης 6.6.6, η αναμενόμενη τιμή της ποσότητας $\frac{\partial \ln L_E(\underline{x};\theta)}{\partial\theta}$, ισούται με μηδέν.

$$E\left(\frac{\partial \ln L_E(\underline{x};\theta)}{\partial\theta}\right)=\sum_{i=1}^{N}\int\frac{\partial \ln L_E(\underline{x};\theta)}{\partial\theta}\cdot L_E(\underline{x};\theta)dx_1dx_2\mathrm{K}\ dx_N$$

$$=\frac{\partial}{\partial\theta}\sum_{i=1}^{N}\int L_E(\underline{x};\theta)dx_1dx_2\mathrm{K}\ dx_N=0 \qquad 7.3.10$$

Ως συνέπεια του θεωρήματος του κεντρικού ορίου, η ποσότητα $\frac{\partial \ln L_E(\underline{x};\theta)}{\partial\theta}$, ως άθροισμα τυχαίων μεταβλητών: $\frac{\partial \ln L_E(\underline{x};\theta)}{\partial\theta}=\sum_{i=1}^{N}\frac{\partial}{\partial\theta}\left[\ln\left(\nu(\theta)\cdot f(x_i;\theta)\right)\right]-\frac{\partial\nu(\theta)}{\partial\theta}$, θα κατανέμεται σύμφωνα με κανονική συνάρτηση πυκνότητας πιθανότητας καθώς το πλήθος των μετρήσεων θα τείνει προς το άπειρο. Η διασπορά της μεταβλητής $\frac{\partial \ln L_E(\underline{x};\theta)}{\partial\theta}$ ευρίσκεται κατ' αντιστοιχία της σχέσης 6.6.7 ως:

$$V\left(\frac{\partial \ln L_E(\underline{x};\theta)}{\partial\theta}\right)=E\left(\left(0-\frac{\partial \ln L_E(\underline{x};\theta)}{\partial\theta}\right)^2\right)=\sum_{N=0}^{\infty}\int\left(\frac{\partial \ln L_E(\underline{x};\theta)}{\partial\theta}\right)^2\cdot L_E(\underline{x};\theta)dx_1dx_2\mathrm{K}\ dx_N$$

$$=-\sum_{N=0}^{\infty}\int\frac{\partial^2 \ln L_E(\underline{x};\theta)}{\partial\theta^2}\cdot L_E(\underline{x};\theta)dx_1dx_2\mathrm{K}\ dx_N=-E\left(\frac{\partial^2 \ln L_E(\underline{x};\theta)}{\partial\theta^2}\right)=\frac{1}{I_{\theta,E}} \qquad 7.3.11$$

Αναπτύσσοντας την συνθήκη μεγιστοποίησης, $\left[\frac{\partial \ln L_E(\partial)}{\partial\theta}\right]_{\theta=\hat{\theta}_E}=0$, σε σειρά Taylor, σε μία περιοχή τιμών της παραμέτρου γύρω από την αληθή της τιμή, αγνοώντας όρους ανώτερης της πρώτης τάξεως ως προς θ, καταλήγουμε ότι:

$$\left[\frac{\partial \ln L_E(\underline{x};\theta)}{\partial\theta}\right]_{\theta=\hat{\theta}_E}=\left[\left(\frac{\partial \ln L_E(\underline{x};\theta)}{\partial\theta}\right)_{\theta=\theta_0}+(\hat{\theta}_E-\theta_0)\left(\frac{\partial^2 \ln L_E(\underline{x};\theta)}{\partial\theta^2}\right)_{\theta=\theta_0}+\underbrace{\mathrm{L}}_{\text{όροι μεγαλύτερης τάξης}}\right]=0$$

$$\hat{\theta}_E - \theta_0 = -\frac{\left(\frac{\partial \ln L_E\left(\underline{x};\theta\right)}{\partial\theta}\right)_{\theta=\theta_0}}{\left(\frac{\partial^2 \ln L_E\left(\underline{x};\theta\right)}{\partial\theta^2}\right)_{\theta=\theta_0}} \qquad 7.3.12$$

Εάν ο παρανομαστής στην σχέση 7.3.12 είναι σταθερά ή μεταβάλλεται πολύ λίγο για τιμές της παραμέτρου θ πλησίον της τιμής θ_0, η απόκλιση της εκτίμησης από την πραγματική τιμή, $(\hat{\theta} - \theta_0)$, θα είναι ανάλογη της πρώτης παραγώγου του λογαρίθμου της πιθανοφάνειας. Κατά συνέπεια, θα κατανέμεται όπως και η μεταβλητή $\frac{\partial \ln L_E\left(\underline{x};\theta\right)}{\partial\theta}$, δηλαδή σύμφωνα με κανονική συνάρτηση πιθανότητας με μέση τιμή ίση με το μηδέν και διασπορά ίση με:

$$V_E\left[\hat{\theta}_E\right] = \frac{V\left[\left(\frac{\partial^2 \ln L_E\left(\underline{x};\theta\right)}{\partial\theta^2}\right)_{\theta=\theta_0}\right]}{\left(\frac{\partial^2 \ln L_E\left(\underline{x};\theta\right)}{\partial\theta^2}\right)^2_{\theta=\theta_0}} = \frac{-E\left(\frac{\partial^2 \ln L_E\left(\underline{x};\theta\right)}{\partial\theta^2}\right)_{\theta=\theta_0}}{\left(\frac{\partial^2 \ln L_E\left(\underline{x};\theta\right)}{\partial\theta^2}\right)^2_{\theta=\theta_0}} \qquad 7.3.13$$

Επαναλαμβάνοντας τους συλλογισμούς της σχέσης 6.6.13 συμπεραίνουμε ότι για μεγάλο πλήθος μετρήσεων ισχύει ότι:

$$E\left(\frac{\partial^2 \ln L_E\left(\underline{x};\theta\right)}{\partial\theta^2}\right) = E\left(\sum_{i=1}^{N}\frac{\partial^2}{\partial\theta^2}\left[\ln\left(\nu(\theta)\cdot f\left(x_i;\theta\right)\right)\right] - \frac{\partial^2\nu(\theta)}{\partial\theta^2}\right) = N\cdot E\left(\frac{\partial^2}{\partial\theta^2}\ln\left(\nu(\theta)\cdot f\left(x_i;\theta\right)\right)\right) - \frac{\partial^2\nu(\theta)}{\partial\theta^2}$$

ώστε μπορούμε να γράψουμε την σχέση 7.3.13 ως:

$$= \lim_{\substack{N\to\infty \\ M\to\infty}} N\cdot\frac{1}{M}\sum_{j=1}^{M}\frac{\partial^2}{\partial\theta^2}\left[\ln\left(\nu(\theta)\cdot f\left(x_i;\theta\right)\right)\right] - \frac{\partial^2\nu(\theta)}{\partial\theta^2} = \frac{\partial^2 \ln L_E\left(\underline{x};\theta\right)}{\partial\theta^2}$$

$$V_E\left[\hat{\theta}_E\right] = \frac{-\left(\frac{\partial^2 \ln L_E\left(\underline{x};\theta\right)}{\partial\theta^2}\right)_{\theta=\theta_0}}{\left(\frac{\partial^2 \ln L_E\left(\underline{x};\theta\right)}{\partial\theta^2}\right)^2_{\theta=\theta_0}} = -\frac{1}{\left(\frac{\partial^2 \ln L_E\left(\underline{x};\theta\right)}{\partial\theta^2}\right)_{\theta=\theta_0}} \qquad 7.3.14$$

Παράδειγμα 7.3.2

Στο Παράδειγμα 7.3.1 παρουσιάστηκε ένα υποθετικό πείραμα ανίχνευσης φωτονίων, τα οποία παράγονται από δύο διαφορετικούς φυσικούς μηχανισμούς. Στόχος του πειράματος είναι η εκτίμηση της παραμέτρου θ, η οποία καθορίζει το ποσοστό φωτονίων τύπου Α.

Η εκτίμηση της παραμέτρου με την μέθοδο μεγιστοποίησης της πιθανοφάνειας καταλήγει ότι:

$$\left(\frac{\partial L(\underline{E};\theta)}{\partial\theta}\right)_{\theta=\hat{\theta}} = \sum_{i=1}^{N}\frac{\partial}{\partial\theta}\ln\left(\frac{\theta\cdot R\cdot f_A(E_i)+\mu_B\cdot f_B(E_i)}{\theta\cdot R+\mu_B}\right)_{\theta=\hat{\theta}}$$

$$= \sum_{i=1}^{N}\left(\frac{R\cdot f_A(E_i)}{\hat{\theta}\cdot R\cdot f_A(E_i)+\mu_B\cdot f_B(E_i)}\right)-\frac{N\cdot R}{\hat{\theta}\cdot R+\mu_B}=0 \qquad 7.3.15$$

Έναν συνήθη τρόπο για να επιλυθούν εξισώσεις όπως η 7.3.5 προσφέρει η μέθοδος των διαδοχικών προσεγγίσεων[1].

Παρατηρήσετε ότι στην περίπτωση όπου $f_A(E)$; $f_B(E)$ η εξίσωση 7.3.15 καταλήγει σε ταυτότητα και δεν προσφέρει καμία πληροφορία για την παράμετρο θ.
Αντίθετα, εάν οι ενεργειακές κατανομές των φωτονίων από τους δύο μηχανισμούς διαφέρουν μεταξύ τους, μεταφέρεται πληροφορία για την αληθή τιμή της παραμέτρου.
Παραδείγματος χάριν, εάν οι ενεργειακές κατατομές είναι της μορφής:

$$f_A(E)\begin{cases}=\phi_A(E); & E\in[\varepsilon_1,\varepsilon_2]\\ =0; & E\notin[\varepsilon_1,\varepsilon_2]\end{cases}$$

$$f_B(E)\begin{cases}=\phi_A(E); & E\in[\varepsilon_3,\varepsilon_4]\\ =0; & E\notin[\varepsilon_3,\varepsilon_4]\end{cases} \qquad 7.3.16$$

όπου $[\varepsilon_1,\varepsilon_2]$ και $[\varepsilon_3,\varepsilon_4]$ είναι διαφορετικές και μη αλληλοκαλυπτόμενες υπο-περιοχές του ενεργειακού διαστήματος $[E_1,E_2]$, είναι εύκολο να πιστοποιήσουμε τον μηχανισμό προέλευσης κάθε φωτονίου. Έστω ότι ανιχνεύθηκαν n φωτόνια στην περιοχή $[\varepsilon_1,\varepsilon_2]$ (δηλαδή φωτόνια που προέρχονται από τον μηχανισμό A) και N-n φωτόνια στην περιοχή $[\varepsilon_3,\varepsilon_4]$. Στην περίπτωση αυτή η σχέση 7.3.15 γράφεται:

---

[1] Αναλύσετε το δεξιό μέλος της εξίσωσης 7.3.15 σε σειρά του $\hat{\theta}$, γύρω από την τιμή $\theta=h_0$. Επιλέξετε την τιμή h_0 πλησίον της αληθούς τιμής της παραμέτρου, επιστρατεύοντας όλη την πρότερη γνώση που διαθέτετε .

$$\left[\sum_{i=1}^{N}\frac{R\cdot f_A(E_i)}{h_0\cdot R\cdot f_A(E_i)+\mu_B\cdot f_B(E_i)}+\sum_{i=1}^{N}\frac{R}{h_0\cdot R+\mu_B}\right]$$
$$-\left[\sum_{i=1}^{N}\frac{\left(R\cdot f_A(E_i)\right)^2}{\left(h_0\cdot R\cdot f_A(E_i)+\mu_B\cdot f_B(E_i)\right)^2}+\sum_{i=1}^{N}\frac{R^2}{\left(h_0\cdot R+\mu_B\right)^2}\right]\cdot\left(\hat{\theta}-h_0\right)$$
$$+d\left(\left(\hat{\theta}-h_0\right)^{k>1}\right)=0$$

Κρατήστε όρους μέχρι πρώτης τάξης και επιλύσετε την γραμμική εξίσωση που προκύπτει ως προς $\hat{\theta}$. Θέσετε $h_1=\hat{\theta}$, αναλύσετε σε σειρά στην γειτονία του h_1 και επιλύσετε την γραμμική εξίσωση. Επαναλάβατε την διαδικασία μέχρι η διαφορά $\delta=h_{j+1}-h_j$ να είναι αρκετά μικρότερη από το στατιστικό σφάλμα που αναμένετε.

$$\sum_{i=1}^{n}\left(\frac{R\cdot\phi_A\left(E_i\right)}{\hat{\theta}\cdot R\cdot\phi_A\left(E_i\right)}\right)-\frac{R\cdot N}{\hat{\theta}\cdot R+\mu_B}=0$$

$$\frac{n}{\hat{\theta}}-\frac{R\cdot N}{\hat{\theta}\cdot R+\mu_B}=0$$

και

$$\hat{\theta}=\frac{n\cdot\mu_B}{R\cdot\left(N-n\right)} \qquad 7.3.17.$$

Επειδή, για μεγάλο πλήθος παρατηρήσεων, N, ισχύει με καλή προσέγγιση ότι μ_B ; $\left(N-n\right)$,η σχέση 7.3.17 καταλήγει στην αναμενόμενη εκτίμηση: $\hat{\theta}=\frac{n}{R}$.

Για σχετικά μεγάλο N, όπως υπαγορεύουν οι ασυμπτωτικές ιδιότητες της μεθόδου μεγίστης πιθανοφάνειας, η εκτίμηση $\hat{\theta}$ κατανέμεται σύμφωνα με κανονική συνάρτηση πυκνότητας πιθανότητας, με αναμενόμενη τιμή ίση με την αληθή τιμή και διασπορά ίση με:

$$V\left[\hat{\theta}\right]=-\frac{1}{\left(\frac{\partial^2\ln L\left(\underline{x};\theta\right)}{\partial\theta^2}\right)_{\theta=\theta_0}}$$

$$=\left[\sum_{i=1}^{N}\frac{\left(R\cdot f_A\left(E_i\right)\right)^2}{\left(\theta_0\cdot R\cdot f_A\left(E_i\right)+\mu_B\cdot f_B\left(E_i\right)\right)^2}-\frac{N\cdot R^2}{\left(\theta_0\cdot R+\mu_B\right)^2}\right]^{-1} \qquad 7.3.18$$

Βεβαίως, επειδή η αληθής τιμή δεν είναι γνωστή, χρησιμοποιούμε αντ΄ αυτής την εκτίμηση $\hat{\theta}$:

$$\hat{V}\left[\hat{\theta}\right]=\left[\sum_{i=1}^{N}\frac{\left(R\cdot f_A\left(E_i\right)\right)^2}{\left(\hat{\theta}\cdot R\cdot f_A\left(E_i\right)+\mu_B\cdot f_B\left(E_i\right)\right)^2}-\frac{N\cdot R^2}{\left(\hat{\theta}\cdot R+\mu_B\right)^2}\right]^{-1} \qquad 7.3.19$$

Ας αντιμετωπίσουμε το ίδιο πρόβλημα χρησιμοποιώντας την μέθοδο της εκτεταμένης συνάρτησης πιθανοφάνειας:

$$L_E(\underline{E};\theta)=\frac{\left(\theta\cdot R+\mu_B\right)^N\cdot e^{-\left(\theta\cdot R+\mu_B\right)}}{N!}\cdot\prod_{i=1}^{N}\frac{\theta\cdot R\cdot f_A\left(E_i\right)+\mu_B\cdot f_B\left(E_i\right)}{\theta\cdot R+\mu_B} \qquad 7.3.20$$

Η συνθήκη μεγιστοποίησης και εκτίμησης της αληθούς τιμής της παραμέτρου, γράφεται ως:

$$\left(\frac{\partial L_E(\underline{E};\theta)}{\partial\theta}\right)_{\theta=\hat{\theta}} = \sum_{i=1}^{N}\frac{\partial}{\partial\theta}\ln\left(\theta\cdot R\cdot f_A\left(E_i\right)+\mu_B\cdot f_B\left(E_i\right)\right)_{\theta=\hat{\theta}} - \left(\theta\cdot R+\mu_B\right)_{\theta=\hat{\theta}}$$

$$= \sum_{i=1}^{N}\left(\frac{R\cdot f_A\left(E_i\right)}{\hat{\theta}\cdot R\cdot f_A\left(E_i\right)+\mu_B\cdot f_B\left(E_i\right)}\right) - R = 0 \qquad 7.3.21$$

Η εξίσωση 7.3.21 λύνεται με την μέθοδο των διαδοχικών προσεγγίσεων, αλλά εύκολα παρατηρεί κανείς ότι, στην περίπτωση όπου $f_A(E)$; $f_B(E)$, η εκτίμηση της αληθούς τιμής είναι εφικτή και βρίσκεται να είναι: $\hat{\theta} = \frac{N-\mu_B}{R}$.

Παρατηρήστε ακόμα ότι, για μεγάλο πλήθος παρατηρήσεων, όταν με καλή προσέγγιση ισχύει ότι $\hat{\theta}\cdot R+\mu_B$; N, η σχέση 7.3.15 είναι ισοδύναμη της σχέσης 7.3.21. Πράγματι η μέθοδος της εκτεταμένης συνάρτησης πιθανοφάνειας προσφέρει σημαντική επιπλέον πληροφορία όταν το πλήθος των παρατηρήσεων είναι μικρό.

Εάν οι ενεργειακές κατανομές των φωτονίων εκφράζονται από τις συναρτήσεις πυκνότητας πιθανότητας που ορίσθηκαν με την σχέση 7.3.16 και κατά την διάρκεια εκτέλεσης του πειράματος ανιχνεύθηκαν n φωτόνια στην περιοχή , $\left[\varepsilon_1,\varepsilon_2\right]$, τότε η σχέση 7.3.21 καταλήγει ότι:

$$\sum_{i=1}^{n}\left(\frac{1}{\hat{\theta}}\right) - R = 0 \Rightarrow \hat{\theta} = \frac{n}{R} \qquad 7.3.22$$

και ευρίσκεται σε πλήρη συμφωνία με τα αποτελέσματα της κλασικής μεθόδου μεγίστης πιθανοφάνειας.

Επισημαίνεται όμως ότι, ενώ η σχέση 7.3.22 ισχύει για κάθε N, καταλήξαμε στα αποτελέσματα της κλασικής μεθόδου 7.3.17 υποθέτοντας ότι το πλήθος των μετρήσεων είναι αρκετά μεγάλο ώστε να ισχύει με καλή προσέγγιση: μ_B ; $\left(N-n\right)$

Όπως εδείχθη σ΄ αυτή την υποενότητα, οι εκτιμήσεις με την μέθοδο της εκτεταμένης συνάρτησης πιθανοφάνειας έχουν ανάλογες ασυμπτωτικές ιδιότητες με τις εκτιμήσεις της κλασσικής μεθόδου μέγιστης πιθανοφάνειας. Δηλαδή, οι εκτιμήσεις δεν πάσχουν από προκατάληψη και κατανέμονται σύμφωνα με κανονική συνάρτηση πιθανότητας. Η διασπορά της εκτίμησης, σ΄ αυτό το παράδειγμα, βρίσκεται να είναι:

$$V\left[\hat{\theta}_E\right] = -\frac{1}{\left(\frac{\partial^2 \ln L_E\left(\underline{x};\theta\right)}{\partial\theta^2}\right)_{\theta=\theta_0}} \qquad 7.3.23$$

$$= \left[\sum_{i=1}^{N}\frac{\left(R\cdot f_A\left(E_i\right)\right)^2}{\left(\theta_0\cdot R\cdot f_A\left(E_i\right)+\mu_B\cdot f_B\left(E_i\right)\right)^2}\right]^{-1}$$

και για κάθε πρακτική χρήση, προσεγγίζεται ως:

$$\hat{V}\left[\hat{\theta}_E\right] = \left[\sum_{i=1}^{N} \frac{\left(R \cdot f_A\left(E_i\right)\right)^2}{\left(\hat{\theta}_E \cdot R \cdot f_A\left(E_i\right) + \mu_B \cdot f_B\left(E_i\right)\right)^2}\right]^{-1} \qquad 7.3.24$$

Απλή σύγκριση των σχέσεων 7.3.24 και 7.3.19 αρκεί να πείσει τον αναγνώστη ότι στο συγκεκριμένο παράδειγμα η μέθοδος της εκτεταμένης συνάρτησης πιθανοφάνειας καταλήγει σε ακριβέστερη εκτίμηση.

Ερώτηση: Υπολογίσετε την ακρίβεια της εκτίμησης κάθε μεθόδου, όταν οι ενεργειακές κατανομές είναι της μορφής 7.3.16.

Υπόδειξη: Θέσετε $\hat{\theta}_E = \frac{n}{R}$ και $\hat{\theta} \approx \frac{n}{R}$ στις σχέσεις 7.3.24 και 7.3.19 για να καταλήξετε ότι:

$$\hat{V}\left[\hat{\theta}_E\right] = \frac{n}{R^2}, \hat{V}\left[\hat{\theta}\right] = \frac{n}{R^2} \cdot \frac{N}{N-n}$$

Ερώτηση: Από τους N μαθητές μίας τάξης οι E προήχθησαν ενώ οι A (=N-E) απέτυχαν στις εξετάσεις. Εκτιμήσετε πλήθος των επιτυχόντων, των πλήθος αποτυχόντων και τα στοιχεία του πίνακα συνδιασποράς των εκτιμήσεων.

Υπόδειξη: Η πιθανοφάνεια εκφράζεται από την δυονυμική συνάρτηση πιθανότητας:

$L\left(E, A; p\right) = \frac{N!}{E!A!} \cdot p^E \cdot \left(1-p\right)^A$, όπου p συμβολίζει την πιθανότητα επιτυχίας. Η πιθανότητα p εκτιμάται ως: $\left(\frac{\partial \ln L\left(E, A; p\right)}{\partial p}\right)_{p=\hat{p}} = 0 \Rightarrow \hat{p} = \frac{E}{A+E} = \frac{E}{N}$

Η διασπορά της εκτίμησης της πιθανότητας είναι:

$V\left[\hat{p}\right] = -\left[\frac{\partial \ln L(E, A; p)}{\partial p^2}\right]^{-1}_{p=p_0} = -\left[\frac{E}{p^2} + \frac{A}{\left(1-p\right)^2}\right]^{-1}$, όπου p_0 είναι η αληθής τιμή.

Η διασπορά εκτιμάται ως: $\hat{V}\left[\hat{p}\right] = -\left[\frac{E}{\hat{p}^2} + \frac{A}{\left(1-\hat{p}\right)^2}\right]^{-1} = \frac{\hat{p} \cdot \left(1-\hat{p}\right)}{N}$

Οι εκτιμήσεις του αριθμού επιτυχόντων και αποτυχόντων είναι: $\hat{E} = N \cdot \hat{p}$, $\hat{A} = N \cdot \left(1-\hat{p}\right)$ και τα στοιχεία του πίνακα συνδιασποράς είναι: $\hat{V}\left[\hat{E}\right] = N^2 \cdot \hat{V}\left[\hat{p}\right] = \frac{E \cdot A}{N}$ και $\hat{V}\left[\hat{A}\right] = N^2 \cdot \hat{V}\left[1-\hat{p}\right] = \frac{E \cdot A}{N}$

$\operatorname{cov}\left(\hat{E}, \hat{A}\right) = -\sqrt{V\left[\hat{E}\right] \cdot V\left[\hat{A}\right]}$.

Ερώτηση: Στην προηγούμενη ερώτηση χρησιμοποιήσετε την εκτεταμένη συνάρτηση πιθανοφάνειας, θεωρώντας ότι ο συνολικός αριθμός των μαθητών μίας τάξης δεν είναι σταθερός αλλά παίρνει τιμές σύμφωνα με Poissonian συνάρτηση πιθανότητας.
Υπόδειξη: Η εκτεταμένη συνάρτηση πιθανοφάνειας εκφράζεται ως:

$$L(E,A;p) = \frac{m^N \cdot e^{-m}}{N!} \cdot \frac{N!}{E!A!} \cdot p^E \cdot (1-p)^A$$
$$= \frac{m_E^N \cdot e^{-m_E}}{E!} \cdot \frac{m_A^N \cdot e^{-m_A}}{A!}$$

όπου m είναι η αναμενόμενη τιμή του αριθμού των μαθητών της τάξης και $m_E = p \cdot m$, $m_A = (1-p) \cdot m$ είναι οι αναμενόμενοι αριθμοί επιτυχόντων και αποτυχόντων αντίστοιχα. Από την συνθήκη μεγιστοποίησης κανείς καταλήγει ότι:

$$\left.\begin{aligned} \left(\frac{\partial \ln L(m_E, m_A; p)}{\partial m_E}\right)_{\substack{m_E=\hat{m}_E \\ m_A=\hat{m}_A}} &= 0 \\ \left(\frac{\partial \ln L(m_E, m_A; p)}{\partial m_A}\right)_{\substack{m_E=\hat{m}_E \\ m_A=\hat{m}_A}} &= 0 \end{aligned}\right\} \Rightarrow \begin{aligned} \hat{m}_E &= E \\ \hat{m}_A &= A \end{aligned}$$

Τα στοιχεία του πίνακα συνδιασποράς εύκολα ευρίσκονται ίσα με: $\hat{V}\left[\hat{E}\right] = E$, $\hat{V}\left[\hat{A}\right] = \hat{A}$, $\mathrm{cov}\left[\hat{E}, \hat{A}\right] = 0$

Στο παράδειγμα αυτό, η διασπορά των εκτιμήσεων βρίσκεται μεγαλύτερη από την διασπορά των αντίστοιχων εκτιμήσεων με την κλασική μέθοδο μεγίστης πιθανοφάνειας, διότι ουσιαστικά εκτιμούνται δύο μεταβλητές έναντι μίας που εκτιμάται με την κλασική πιθανοφάνεια. Επισημαίνεται όμως, ότι οι εκτιμήσεις των παραμέτρων E και A είναι στατιστικά ανεξάρτητες όταν χρησιμοποιείται η εκτεταμένη πιθανοφάνεια ενώ στην άλλη περίπτωση οι εκτιμήσεις είχαν 100% αντι-συσχέτιση (συντελεστής συσχέτισης ίσος με -1).

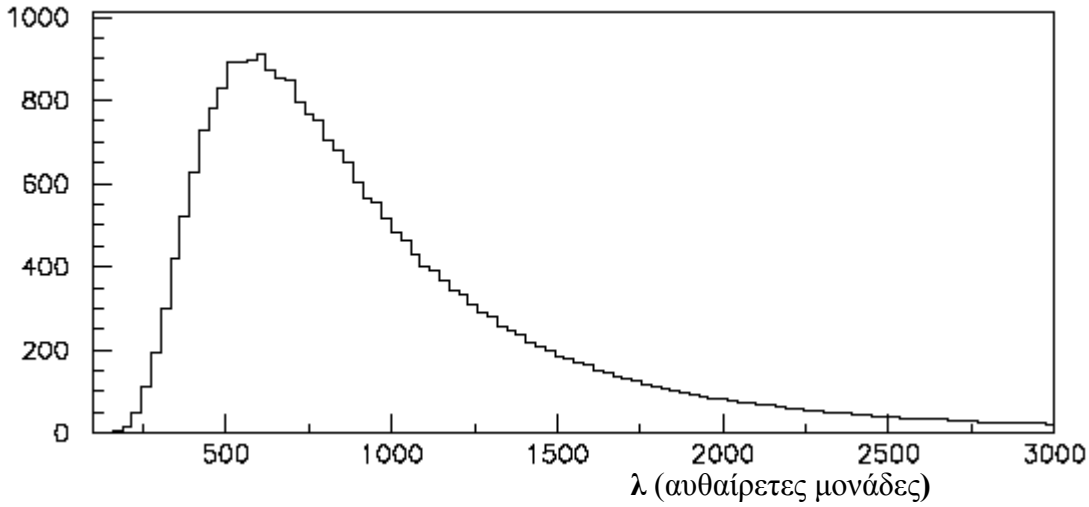
7.4 Εκτίμηση παραμέτρων από ιστογράμματα.

Έστω ότι σ΄ ένα πείραμα θέλουμε να προσδιορίσουμε την φυσική παράμετρο θ, η οποία καθορίζει το ενεργειακό φάσμα εκπομπής φωτονίων από ένα φυσικό σύστημα[2]. Έστω ότι η συνάρτηση πυκνότητας πιθανότητας του μήκους κύματος περιγράφεται από την συνάρτηση: f(λ,θ) και τα πειραματικά δεδομένα συνίστανται στις μετρήσεις του μήκους κύματος N φωτονίων: $\underline{\lambda} = \{\lambda_1, \lambda_2, \lambda_3, \ldots \lambda_N\}$, στην ενεργό περιοχή του οργάνου μέτρησης.

Είναι προφανές ότι η ανάλυση δεδομένων αυτού του πειράματος ανάγεται σε κλασική εκτίμηση παραμέτρων με την μέθοδο της μέγιστης πιθανοφάνειας. Δυστυχώς, οι πραγματικές πειραματικές συνθήκες δυσκολεύουν την εφαρμογή αυτής της μεθόδου και σε πολλές περιπτώσεις την καθιστούν αδύνατη. Επί παραδείγματι, η ανιχνευτική διάταξη πάσχει από

[2] Επί παραδείγματι, να προσδιορίσουμε την θερμοκρασία ενός ουρανίου σώματος από το φάσμα εκπομπής του, θεωρώντας ότι η εκπομπή περιγράφεται από το θεωρητικό πρότυπο του μέλανος σώματος.

μετρητικά σφάλματα και περιορισμένη απόδοση η οποία είναι συνάρτηση του φυσικού μεγέθους που μετράμε. Θα αντιμετωπίσουμε το γενικό πρόβλημα στην υποπαράγραφο 7.7.
Εν προκειμένω, ας υποθέσουμε ότι η ανιχνευτική συσκευή μπορεί μόνο να κατατάσσει το υπό μέτρηση μήκος κύματος σε μία από τις k περιοχές, εύρους Δλ, στις οποίες υποδιαιρείται η ενεργός περιοχή του οργάνου μέτρησης. Τα πειραματικά δεδομένα συνίστανται από το πληθος των φωτονίων $\underline{n} = \{n_1, n_2, n_3, K\ n_N\}$ που συλλέχθηκαν στις αντίστοιχες υποπεριοχές (ιστούς) με όρια $(\lambda_i, \lambda_{i+1}]$, i=0,1,2,3,…k και παρίστανται γραφικά στο Σχήμα 7.4.1.



Σχήμα 7.4.1: Η ανιχνευτική διάταξη κατατάσσει κάθε φωτόνιο σε μία από τις k=100 υποπεριοχές μήκους κύματος, ιδίου εύρους Δλ=29, που καλύπτουν την ενεργό περιοχή του οργάνου μέτρησης [100,3000]. Τα πειραματικά δεδομένα παρίστανται ως ιστόγραμμα.

7.4.1 Εκτίμηση με την μέθοδο μέγιστης πιθανοφάνειας

Η συνολική πληροφορία εμπεριέχεται στην συνάρτηση πιθανοφάνειας, που στην περίπτωση αυτής της ομαδοποίησης των δεδομένων ανάγεται στην πολυωνυμική συνάρτηση πυκνότητας πιθανότητας:

$$L(\underline{n};\theta) = \frac{N!}{n_1!n_2!L\ n_k!}\prod_{i=1}^{k}\left(p_i(\theta)\right)^{n_i}$$

$$p_i(\theta) = \int_{\lambda_{i-1}}^{\lambda_i} f(\lambda;\theta)d\lambda \qquad 7.4.1$$

Ερώτηση: Δώσετε την γενική έκφραση της διασποράς της εκτίμησης της παραμέτρου, με την μέθοδο της μεγίστης πιθανοφάνειας 7.4.1, για μεγάλο αριθμό ιστών,k.
Υπόδειξη: Η εκτίμηση της παραμέτρου προκύπτει από την μεγιστοποίηση της πιθανοφάνειας:

$$\left(\frac{\partial \ln L(\underline{n};\theta)}{\partial\theta}\right)_{\theta=\hat{\theta}} = \sum_{i=1}^{k}\frac{n_i}{p_i(\hat{\theta})}\cdot\left(\frac{\partial p_i(\theta)}{\partial\theta}\right)_{\theta=\hat{\theta}} = 0 \qquad 7.4.2$$

Για αρκετά μεγάλο αριθμό ιστών, ώστε να ισχύουν οι ασυμπτωτικές ιδιότητες της μεθόδου μέγιστης πιθανοφάνειας, η διασπορά της εκτίμησης της τιμής της παραμέτρου θα είναι:

$$V\left[\hat{\theta}\right] = \frac{-1}{\left(\frac{\partial \ln L\left(\underline{n};\theta\right)}{\partial\theta}\right)_{\theta=\theta_0}} = \left[\sum_{i=1}^{k}\left(\frac{n_i}{\left(p_i\left(\theta_0\right)\right)^2}\cdot\left(\frac{\partial p_i\left(\theta\right)}{\partial\theta}\right)^2_{\theta=\theta_0} - \frac{n_i}{p_i\left(\theta_0\right)}\cdot\left(\frac{\partial^2 p_i\left(\theta\right)}{\partial\theta^2}\right)_{\theta=\theta_0}\right)\right]^{-1} \quad 7.4.3.\alpha$$

όπου θ_0 είναι η πραγματική τιμή της παραμέτρου. Η σχέση 7.4.3.α προσεγγίζεται με την ακόλουθη έκφραση:

$$\hat{V}\left[\hat{\theta}\right] = \frac{-1}{\left(\frac{\partial \ln L\left(\underline{n};\theta\right)}{\partial\theta}\right)_{\theta=\hat{\theta}}} = \left[\sum_{i=1}^{k}\left(\frac{n_i}{\left(p_i\left(\hat{\theta}\right)\right)^2}\cdot\left(\frac{\partial p_i\left(\theta\right)}{\partial\theta}\right)^2_{\theta=\hat{\theta}} - \frac{n_i}{p_i\left(\hat{\theta}\right)}\cdot\left(\frac{\partial^2 p_i\left(\theta\right)}{\partial\theta^2}\right)_{\theta=\hat{\theta}}\right)\right]^{-1} \quad 7.4.3.\beta$$

Εάν επιπλέον του φάσματος και το πλήθος των φωτονίων που ανιχνεύονται εξαρτάται από την παράμετρο θ, τότε ενδείκνυται να χρησιμοποιήσουμε την εκετεταμένη συνάρτηση πιθανοφάνειας:

$$\begin{aligned} L_E\left(\underline{n};\theta\right) &= \frac{\left(M\left(\theta\right)\right)^N\cdot e^{-M(\theta)}}{N!}\cdot\frac{N!}{n_1!n_2!\ldots n_k!}\prod_{i=1}^{k}\left(p_i\left(\theta\right)\right)^{n_i} \\ &= \frac{e^{-\left(\sum_{i=1}^{k}\mu_i(\theta)\right)}}{n_1!n_2!\ldots n_k!}\prod_{i=1}^{k}\left(\underbrace{M\left(\theta\right)\cdot p_i\left(\theta\right)}_{\mu_i(\theta)}\right)^{n_i} \qquad 7.4.3 \\ &= \prod_{i=1}^{k}\frac{\left(\mu_i\left(\theta\right)\right)^{n_i}\cdot e^{-\mu_i(\theta)}}{n_i!} \end{aligned}$$

όπου $\mu_i\left(\theta\right) = M\left(\theta\right)\cdot p_i\left(\theta\right) = M\left(\theta\right)\cdot\int_{\lambda_{i-1}}^{\lambda_i} f\left(\lambda;\theta\right)d\lambda$ είναι το αναμενόμενο πλήθος παρατηρήσεων στην υποπεριοχή $\left(\lambda_i,\lambda_{i+1}\right]$.

Ερώτηση: Δείξετε ότι, στην περίπτωση όπου το πλήθος των γεγονότων σε κάθε ιστό είναι μεγάλο, η συνάρτηση εκτεταμένης πιθανοφάνειας μπορεί να γραφεί ώς γινόμενο κανονικών συναρτήσεων.

Υπόδειξη: Σύμφωνα με το θεώρημα του κεντρικού ορίου, για $n \to \infty$, η Poisonian συνάρτηση πιθανότητας : $\frac{\mu^{n_i}\cdot e^{-\mu}}{n!}$ τείνει προς κανονική συνάρτηση, με την ίδια αναμενόμενη τιμή και ίδια διασπορά, $\frac{1}{\sqrt{2\pi}\cdot\sqrt{\mu}}\cdot e^{\frac{-(\mu-n)^2}{2\cdot\mu}}$

Συνεπώς, η συνάρτηση εκτεταμένης πιθανοφάνειας 7.4.3 λαμβάνει την μορφή:

$$\lim_{\substack{n_1\to\infty \\ n_2\to\infty \\ \vdots \\ n_k\to\infty}} \left[L_E\left(\underline{n};\theta\right)\right] = \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi}\cdot\sqrt{\mu_i(\theta)}}\cdot e^{\frac{-\left(\mu_i(\theta)-n_i\right)^2}{2\cdot\mu_i(\theta)}} \qquad 7.4.4$$

Ερώτηση: Δώσετε την γενική έκφραση της διασποράς της εκτίμησης της παραμέτρου, που προκείπτει από την εκτεταμένη συνάρτηση πιθανοφάνειας 7.4.3, για μεγάλο αριθμό ιστών, k.

Υπόδειξη: Η εκτίμηση της παραμέτρου προκύπτει από την μεγιστοποίηση της εκτεταμένης πιθανοφάνειας:

$$\left(\frac{\partial}{\partial\theta}\ln L_E\left(\underline{n};\theta\right)\right)_{\theta=\hat{\theta}_E} = \left(\frac{\partial}{\partial\theta}\sum_{i=1}^{k}\left[n_i\cdot\ln\mu_i(\theta)-\mu_i(\theta)\right]\right)_{\theta=\hat{\theta}_E}$$

$$= \left(\sum_{i=1}^{k}\left[\frac{n_i}{\mu_i(\theta)}\cdot\frac{\partial\mu_i(\theta)}{\partial\theta}-\frac{\partial\mu_i(\theta)}{\partial\theta}\right]\right)_{\theta=\hat{\theta}_E} \qquad 7.4.5$$

$$= \left(\sum_{i=1}^{k}\frac{n_i}{\mu_i(\theta)}\cdot\frac{\partial\mu_i(\theta)}{\partial\theta}\right)_{\theta=\hat{\theta}_E} - \left(\frac{\partial M(\theta)}{\partial\theta}\right)_{\theta=\hat{\theta}_E} = 0$$

Η διασπορά της εκτίμησης της παραμέτρου υπολογίζεται, όταν $k\to\infty$, ως:

$$\hat{V}\left[\hat{\theta}_E\right] = \frac{-1}{\left(\frac{\partial\ln L_E\left(\underline{n};\theta\right)}{\partial\theta}\right)_{\theta=\hat{\theta}_E}}$$

$$= \left[\sum_{i=1}^{k}\left(\frac{n_i}{\left(\mu_i\left(\hat{\theta}_E\right)\right)^2}\cdot\left(\frac{\partial\mu_i(\theta)}{\partial\theta}\right)^2_{\theta=\hat{\theta}_E} - \frac{n_i}{\mu_i\left(\hat{\theta}_E\right)}\cdot\left(\frac{\partial^2\mu_i(\theta)}{\partial\theta^2}\right)_{\theta=\hat{\theta}_E}\right) + \left(\frac{\partial^2 M(\theta)}{\partial\theta^2}\right)_{\theta=\hat{\theta}_E}\right]^{-1} \qquad 7.4.6$$

Εύκολα μπορείτε να δείξετε ότι η χρήση της εκτεταμένης συνάρτησης πιθανοφάνειας καταλήγει σε εκτίμηση με μικρότερη διασπορά από εκείνη της κλασικής μεθόδου μεγίστης πιθανοφάνειας. Οι δύο μέθοδοι είναι εντελώς ισοδύναμες όταν το αναμενόμενο πλήθος συνολικών γεγονότων, M(θ), δεν εξαρτάται από την παράμετρο θ.

Ερώτηση: Δείξετε ότι η εκτίμηση με την μέθοδο της εκτεταμένης συνάρτησης πιθανοφάνειας καταλήγει σε εκτίμηση με μικρότερη διασπορά από εκείνη της κλασικής μεθόδου μεγίστης πιθανοφάνειας.

Υπόδειξη: Αντικαθιστώντας στην σχέση 7.4.6: $\mu_i(\theta) = M(\theta)\cdot p_i(\theta)$ και αντί του $\hat{\theta}$ την πραγματική τιμή της παραμέτρου, μετά από πράξεις, καταλήγουμε στην ακόλουθη σχέση:

$$V\left[\hat{\theta}_E\right]=\left[\sum_{i=1}^{k}\left(\frac{n_i}{\left(p_i\left(\theta_0\right)\right)^2}\cdot\left(\frac{\partial p_i\left(\theta\right)}{\partial\theta}\right)^2_{\theta=\theta_0}-\frac{n_i}{p_i\left(\theta_0\right)}\cdot\left(\frac{\partial^2 p_i\left(\theta\right)}{\partial\theta^2}\right)_{\theta=\theta_0}\right)\right.$$

$$\left.+\frac{\sum_{i=1}^{k}n_i}{\left(M\left(\theta_0\right)\right)^2}\cdot\left(\frac{\partial M\left(\theta\right)}{\partial\theta}\right)^2_{\theta=\theta_0}+\left(1-\frac{\sum_{i=1}^{k}n_i}{M\left(\theta_0\right)}\right)\left(\frac{\partial^2 M\left(\theta\right)}{\partial\theta^2}\right)_{\theta=\theta_0}\right]^{-1} \qquad 7.4.7$$

Ο όρος $\sum_{i=1}^{k}n_i$ στην σχέση 7.4.7 μπορεί να γραφεί ως:

$$\sum_{i=1}^{k}n_i=\underbrace{\sum_{i=1}^{k}\left(n_i-\mu_i\left(\theta_0\right)\right)}_{x_i}+\underbrace{\sum_{i=1}^{k}\mu_i\left(\theta_0\right)}_{M(\theta_0)} \qquad 7.4.8$$

$$=\sum_{i=1}^{k}x_i\left(\theta_0\right)+M\left(\theta_0\right)$$

Κάθε τυχαία μεταβλητή $x_i\left(\theta_0\right)$ έχει αναμενόμενη τιμή: $E\left[x_i\left(\theta_0\right)\right]=E\left[n_i-\mu_i\left(\theta_0\right)\right]=0$ ενώ η διασπορά της τυχαίας μεταβλητής $x_i\left(\theta_0\right)$ είναι: $V\left[x_i\left(\theta_0\right)\right]=V\left[n_i-\mu_i\left(\theta_0\right)\right]=V\left[n_i\right]=\mu_i\left(\theta_0\right)$. Εύκολα διαπιστώνεται ότι ισχύουν οι προϋποθέσεις για την εφαρμογή του νόμου των μεγάλων αριθμών για να γραφεί η σχέση 7.4.8 ως:

$$\sum_{i=1}^{k}n_i\approx k\cdot E\left[x_i\left(\theta_0\right)\right]+M\left(\theta_0\right)=M\left(\theta_0\right) \qquad 7.4.9$$

Συνεπώς , για μεγάλο αριθμό ιστών, η σχέση 7.4.7 παίρνει την μορφή.

$$V\left[\hat{\theta}_E\right]=\left[\sum_{i=1}^{k}\left(\frac{n_i}{\left(p_i\left(\theta_0\right)\right)^2}\cdot\left(\frac{\partial p_i\left(\theta\right)}{\partial\theta}\right)^2_{\theta=\theta_0}-\frac{n_i}{p_i\left(\theta_0\right)}\cdot\left(\frac{\partial^2 p_i\left(\theta\right)}{\partial\theta^2}\right)_{\theta=\theta_0}\right)\right.$$

$$\left.+\frac{1}{M\left(\theta_0\right)}\cdot\left(\frac{\partial M\left(\theta\right)}{\partial\theta}\right)^2_{\theta=\theta_0}\right]^{-1} \qquad 7.4.9$$

Συγκρίνοντας με την σχέση 7.4.3.α εύκολα καταλήγει κανείς ότι η εκτεταμένη συνάρτηση πιθανοφάνειας μεταφέρει περισσότερη πληροφορία αναφορικά με την παράμετρο από ότι η κλασική συνάρτηση πιθανοφάνειας και συνεπώς η διασπορά 7.4.9 είναι μικρότερη από την αντίστοιχη διασπορά της κλασικής μεθόδου.

Ερώτηση: Δείξετε ότι η κλασική μέθοδος μεγίστης πιθανοφάνειας και η μέθοδος της εκτεταμένης συνάρτησης πιθανοφάνειας καταλήγουν στα ίδια αποτελέσματα, όταν το αναμενόμενο πλήθος συνολικών γεγονότων δεν εξαρτάται από την παράμετρο,θ.

Υπόδειξη: Θέσετε στις σχέσεις 7.4.5 και 7.4.6 $\mu_i\left(\theta\right)=M\cdot p_i\left(\theta\right)$ και προβείτε σε απλοποιήσεις.

7.4.1 Εκτίμηση με την μέθοδο ελαχίστων τετραγώνων

Ας υποθέσουμε ότι το πλήθος των ιστών, k, είναι αρκετά μεγάλο ώστε οι πιθανότητες $p_i(\theta_0)$, να ανιχνευθεί ένα γεγονός στην υποπεριοχή i, είναι πολύ μικρές. Η υπόθεση αυτή εγγυάται ότι η συσχέτιση $(-N \cdot p_i(\theta_0) \cdot p_j(\theta_0))$ μεταξύ κάθε ζεύγους τυχαίων μεταβλητών, n_i και n_j, είναι αμελητέα. Επιπλέον, εάν το πλήθος των δεδομένων, N, είναι αρκετά μεγάλο ώστε κάθε τυχαία μεταβλητή n_i να κατανέμεται σύμφωνα με κανονική συνάρτηση πιθανότητας με μέση τιμή $m_i(\theta_0) = N \cdot p_i(\theta_0)$ και διασπορά ισούται με $V[n_i] = N \cdot p_i(\theta_0)$, ορίζουμε[3] την στατιστική συνάρτηση ελαχίστων τετραγώνων ως:

$$Q_1^2 = \sum_{i=1}^{k} \frac{\left(\mu_i(\theta) - n_i\right)^2}{\mu_i(\theta)} \qquad 7.4.10$$

Η ποσότητα Q_1^2 δεν θα πρέπει να έχει κατ΄ ανάγκη τις ιδιότητες του χ^2, διότι η διασπορά στον παρανομαστή εξαρτάται από την παράμετρο θ. Εύκολα μπορείτε επίσης να διαπιστώσετε ότι δεν είναι ισοδύναμη της συνάρτησης πιθανοφάνειας. Παρ΄ όλα αυτά, αποδεικνύεται ότι η εκτίμηση παραμέτρων με την ελαχιστοποίηση της 7.4.10, ασυμτωτικά, έχει ιδανικές ιδιότητες[4]: είναι συνεπής, η εκτίμηση συμπεριφέρεται σύμφωνα με κανονική συνάρτηση πυκνότητας πιθανότητας και είναι αποδοτική έχοντας την ελάχιστη διασπορά.

Συνήθως, αντί της σχέσης 7.4.10 ελαχιστοποιείται η ποσότητα:

$$Q_1^2 = \sum_{i=1}^{k} \frac{\left(\mu_i(\theta) - n_i\right)^2}{n_i} \qquad 7.4.11$$

όπου η άγνωστη ποσότητα $\mu_i(\theta)$ προσεγγίζεται με την «μέτρηση» n_i. Για μεγάλο πλήθος ιστών, οι εκτιμητές 7.4.10 και 7.4.11 συμπίπτουν με την μέθοδο της μεγίστης πιθανοφάνειας.

Επισημαίνεται, ότι η μέθοδος της μεγίστης πιθανοφάνειας έχει ευρύτερο φάσμα εφαρμογών, με πολύ λιγότερες παραδοχές όσον αφορά στον αριθμό των υποπεριοχών και τον πληθυσμό γεγονότων σε κάθε υποπεριοχή.

---

[3] Όταν το πλήθος N τείνει στο άπειρο, η πολυονυμική συνάρτηση πιθανότητας τείνει προς κανονική κατανομή k διαστάσεων. Επιπλέον, λόγω της αμοιβαίας ανεξαρτησίας των n_i, η συνάρτηση πυκνότητας πιθανότητας ανάγεται σε γινόμενο κανονικών συναρτήσεων: $\prod_{i=1}^{k} \frac{1}{\sqrt{2\pi} \cdot \sqrt{m_i(\theta)}} \cdot e^{\frac{-\left(m_i(\theta) - n_i\right)^2}{2 \cdot m_i(\theta)}}$.

Εν άλλοις λόγοις, κάθε μία από τις αμοιβαία ανεξάρτητες τυχαίες μεταβλητές $\underline{n} = \{n_1, n_2, K\ n_k\}$ χαρακτηρίζεται από κανονική συνάρτηση πιθανότητας, με μέση τιμή $m_i(\theta_0) = N \cdot p_i(\theta_0)$ και διασπορά $V[n_i] = N \cdot p_i(\theta_0)$.

[4] Kendal Vol. II

7.5 Μέθοδοι εκτίμησης παραμέτρων υπό συνθήκες δέσμευσης.

Συνήθως, οι παράμετροι αντιστοιχούν σε φυσικές ποσότητες που ικανοποιούν φυσικούς νόμους. Παραδείγματος χάριν, στην περίπτωση που αποσκοπούμε στην σύγχρονη εκτίμηση της αληθούς τιμής του μέτρου της ορμής (P_0) και της ενέργειας (E_0) ενός σωματίου, από μετρήσεις της ενέργειας (E) και της ορμής (P), θα πρέπει να επιβάλουμε την συνθήκη ότι η αληθής τιμή της ενέργειας και η αληθής τιμή της ορμής συνδέονται με την σχέση: $E_0^2 = P_0^2 \cdot c^2 + m^2 \cdot c^4$, όπου m είναι η μάζα του σωματίου και c είναι η ταχύτητα του φωτός.

Σ΄ αυτή την υποενότητα θα συζητηθούν μέθοδοι εκτίμησης k παραμέτρων, $\underline{\theta} = \{\theta_1, \theta_2, K\ \theta_k\}$, οι οποίοι υπόκεινται σε m συνθήκες δέσμευσης, της μορφής $\underline{g}(\underline{\theta}) = 0$:

$$\begin{aligned} g_1(\underline{\theta}) &= g_1(\theta_1, \theta_2, K\ \theta_k) = 0 \\ g_2(\underline{\theta}) &= g_2(\theta_1, \theta_2, K\ \theta_k) = 0 \\ &K\ K\ K\ K \\ g_m(\underline{\theta}) &= g_m(\theta_1, \theta_2, K\ \theta_k) = 0 \end{aligned} \qquad 7.5.1$$

Η πλέον εύκολη και αποδοτική μέθοδος να συμπεριληφθούν οι συνθήκες δέσμεύσης 7.5.1 στην μέθοδο εκτίμησης συνίσταται σε κατάλληλη επιλογή νέων παραμέτρων.

Επί παραδείγματι, εάν οι παράμετροι θ_1 και θ_2 υφίστανται τον περιορισμό:

$$g(\theta_1, \theta_2) = \theta_1 + \theta_2 - 1 = 0 \qquad 7.5.2$$

θα μπορούσαμε να αντικαταστήσουμε, στην συνάρτηση πυκνότητας πιθανότητας ή στην έκφραση των ελαχίστων τετραγώνων, την τιμή της παραμέτρου θ_2 με $(1-\theta_1)$ και στη συνέχεια να εφαρμόσουμε την μέθοδο εκτίμησης μόνο για την εναπομένουσα παράμετρο, θ_1.

Μία άλλη περίπτωση αφορά περιορισμούς στο πεδίο τιμών των παραμέτρων. Συχνά αντιμετωπίζονται προβλήματα όπου η παράμετρος θ θα πρέπει να παίρνει τιμές μόνο στο διάστημα [a,b]. Στις περιπτώσεις αυτές ενδείκνυται η έκφραση της τιμής της παραμέτρου συναρτήσει μίας άλλης μεταβλητής, ψ, ως:

$$\theta = a + \frac{1}{2} \cdot (\sin\psi + 1) \cdot (b - a) \qquad 7.5.3$$

και η εφαρμογή της μεθόδου εκτίμησης για τον προσδιορισμό της τιμής της παραμέτρου ψ, στο διάστημα $(-\infty, \infty)$.

Άλλες φορές, το πρόβλημα συνίσταται στην εκτίμηση των ποσοστών συμμετοχής, β_i, επί μέρους κατηγοριών στο συνολικό αποτέλεσμα. Σ΄ αυτές τις περιπτώσεις, οι παράμετροι υπόκεινται στους ακόλουθους περιορισμούς:

$$0 \le \beta_i \le 1 \ (i = 1, 2, 3.....k) \quad \text{και} \quad \sum_{i=1}^{k} \beta_i = 1 \qquad 7.5.4$$

Ενδείκνυται η ακόλουθη σειρά μετασχηματισμού των παραμέτρων:

$$\begin{aligned}
\beta_1 &= r_1 \\
\beta_2 &= (1-r_1)\cdot r_2 \\
\beta_3 &= (1-r_1)\cdot(1-r_2)\cdot r_3 \\
M &= M\ M\ M\ M\ M\ M\ M \\
\beta_{k-1} &= (1-r_1)\cdot(1-r_2)L\ (1-r_{k-2})r_{k-1} \\
\beta_k &= (1-r_1)\cdot(1-r_2)L\ (1-r_{k-2})(1-r_{k-1})
\end{aligned} \qquad 7.4.6$$

όπου: $r_i = \frac{1}{2}\cdot(\sin\psi_i + 1)$ 7.4.7

και εκτιμώνται οι τιμές των παραμέτρων $\underline{\psi} = \{\psi_1, \psi_2, K\ , \psi_N\}$[5].

Ερώτηση: Δείξετε ότι οι παράμετροι β, όπως ορίζονται από τις σχέσεις 7.4.6 και 7.4.7, ικανοποιούν τις συνθήκες 7.5.4.

Στις περισσότερες περιπτώσεις εκτίμησης παραμέτρων υπό συνθήκες δέσμευσης, δεν είναι εύκολο να εφαρμοσθούν οι απλές μέθοδοι ορισμού νέων μεταβλητών. Μία γενική μέθοδος για να συμπεριλάβει κανείς τις συνθήκες δέσμευσης στους εκτιμητές αναφέρεται στην χρήση πολλαπλασιαστών Lagrange.

7.5.1 Μέθοδος της μεγίστης πιθανοφάνειας με Lagrangian πολλαπλασιαστές.

Έστω $x_i, (i = 1,2,3,K\ N)$ τα πειραματικά δεδομένα και $L(\underline{x};\underline{\theta}) = \prod_{=1}^{N} f(\underline{x};\underline{\theta})$ η συνάρτηση πιθανοφάνειας που εξαρτάται από k παραμέτρους $\theta_i, (i = 1,2,3,K\ k)$ οι οποίες πρέπει ικανοποιούν τις συνθήκες 7.5.1.

Ορίζουμε την στατιστική συνάρτηση:

$$F(\underline{x};\underline{\theta},\underline{\alpha}) = \ln L(\underline{x};\underline{\theta}) + \sum_{i=1}^{m} \alpha_i \cdot g_i(\underline{\theta}) \quad 7.4.8$$

όπου οι παράγοντες α_i (i=1,2,3,…,m) ονομάζονται πολλαπλασιαστές Lagrange και αναζητούμε τις τιμές των παραμέτρων $\theta_i\ (i = 1,2,3,K\ k)$ και $\alpha_j\ (j = 1,2,3,K\ m)$ για τις οποίες η έκφραση 7.4.8 έχει ακρότατο. Δηλαδή:

---

[5] Εύκολα μπορεί να δει κανείς ότι στην περίπτωση όπου η συνάρτηση πυκνότητας πιθανότητας είναι συμμετρική ως προς τις παραμέτρους β, δεν θα είναι κατ΄ ανάγκη συμμετρική και ως προς τις παραμέτρους ψ.
Εξετάσετε ως παράδειγμα την περίπτωση όπου η συνάρτηση πυκνότητας πιθανότητας είναι γραμμική συνάρτηση των παραμέτρων β, ή την περίπτωση όπου το πρότυπο προσαρμογής με την μέθοδο των ελαχίστων τετραγώνων είναι γραμμική συνάρτηση των παραμέτρων β.
Με τους μετασχηματισμούς 7.4.6 και 7.4.7 επιβάλλονται οι δεσμεύσεις στις παραμέτρους αλλά χάνεται η συμμετρία του προβλήματος.

$$\left(\frac{\partial F(\underline{x};\underline{\theta},\underline{\alpha})}{\partial\theta_i}\right)_{\substack{\underline{\theta}=\hat{\underline{\theta}}\\ \underline{\alpha}=\hat{\underline{\alpha}}}} = \left(\frac{\partial \ln L(\underline{x};\underline{\theta})}{\partial\theta_i}\right)_{\underline{\theta}=\hat{\underline{\theta}}} + \left(\sum_{j=1}^{m}\hat{\alpha}_j\cdot\frac{\partial g_j(\underline{\theta})}{\partial\theta_i}\right)_{\underline{\theta}=\hat{\underline{\theta}}} = 0 \ , \quad (i=1,2,3,\ldots,k)$$

$$\left(\frac{\partial F(\underline{x};\underline{\theta},\underline{\alpha})}{\partial\alpha_n}\right)_{\substack{\underline{\theta}=\hat{\underline{\theta}}\\ \underline{\alpha}=\hat{\underline{\alpha}}}} = g_n(\hat{\underline{\theta}}) = 0 \qquad (n=1,2,3,\ldots,m) \qquad 7.4.9$$

Οι εκτιμήσεις των παραμέτρων θ, που βρίσκονται ως λύσεις του συστήματος των εξισώσεων 7.4.9, ικανοποιούν τις συνθήκες δέσμευσης και έχουν τις συνήθεις ασυμπτωτικές ιδιότητες των εκτιμήσεων της μεθόδου μεγίστης πιθανοφάνειας. Επισημαίνεται ότι οι πολλαπλασιαστές Lagrange αποτελούν και αυτοί παραμέτρους με τιμές που εκτιμώνται από την 7.4.9.

Ο πίνακας συνδιασποράς των εκτιμήσεων των παραμέτρων:

$$\left(b_1\ b_2\ \ldots,b_k\ b_{k+1}\ b_{k+2}\ \ldots\ b_{k+m}\right)^T \equiv \left(\theta_1\ \theta_2\ \ldots,\theta_k\ \alpha_1\ \alpha_2\ \ldots\ \alpha_m\right)^T \qquad 7.4.10$$

δίνονται από τον αντίστροφο του πίνακα πληροφορίας ως:

$$\hat{\underset{\%}{V}} = \left(\underset{\%}{I}_{\theta,\alpha}\right)^{-1} = \left(-E\left[\begin{pmatrix} \underset{\%}{A} & \underset{\%}{B} \\ \underset{\%}{B}^T & \underset{\%}{0} \end{pmatrix}\right]\right)^{-1}_{\substack{\underline{\theta}=\hat{\underline{\theta}}\\ \underline{\alpha}=\hat{\underline{\alpha}}}}$$

$$\underset{\%}{A}_{\lambda\mu} = \left(\frac{\partial^2 F(\underline{x};\underline{\theta},\underline{\alpha})}{\partial\theta_\lambda\partial\theta_\mu}\right)_{\substack{\underline{\theta}=\hat{\underline{\theta}}\\ \underline{\alpha}=\hat{\underline{\alpha}}}} = \left(\frac{\partial^2 L(\underline{x};\underline{\theta})}{\partial\theta_\lambda\partial\theta_\mu}\right)_{\underline{\theta}=\hat{\underline{\theta}}}, \quad (\lambda=1,2,\ldots,k)\ \text{και}\ (\mu=1,2,\ldots,k)$$

$$\underset{\%}{B}_{\lambda\mu} = \left(\frac{\partial^2 F(\underline{x};\underline{\theta},\underline{\alpha})}{\partial\theta_\lambda\partial\alpha_\mu}\right)_{\substack{\underline{\theta}=\hat{\underline{\theta}}\\ \underline{\alpha}=\hat{\underline{\alpha}}}} = \left(\frac{\partial^2 g_m(\underline{\theta})}{\partial\theta_\lambda}\right)_{\underline{\theta}=\hat{\underline{\theta}}}, \quad (\lambda=1,2,\ldots,k)\ \text{και}\ (\mu=1,2,\ldots,m) \qquad 7.4.11$$

$$\underset{\%}{O}_{\lambda\mu} = \left(\frac{\partial^2 F(\underline{x};\underline{\theta},\underline{\alpha})}{\partial\alpha_\lambda\partial\alpha_\mu}\right)_{\substack{\underline{\theta}=\hat{\underline{\theta}}\\ \underline{\alpha}=\hat{\underline{\alpha}}}} = \underset{\%}{0} \qquad (\lambda=1,2,\ldots,m)\ \text{και}\ (\mu=1,2,\ldots,m)$$

Όταν ο πίνακας $\left(-E\left[\begin{pmatrix} \underset{\%}{A} & \underset{\%}{B} \\ \underset{\%}{B}^T & \underset{\%}{0} \end{pmatrix}\right]\right)^{-1}_{\substack{\underline{\theta}=\hat{\underline{\theta}}\\ \underline{\alpha}=\hat{\underline{\alpha}}}}$ αντιστρέφεται μπορεί κανείς, με επίπονες αλλά απλές πράξεις, να καταλήξει[6] στις ακόλουθες απλές εκφράσεις:

$$\hat{\underset{\%}{V}}\left[\hat{\theta}\right] = \underset{\%}{A}^{-1} - \underset{\%}{A}^{-1}\cdot\underset{\%}{B}\cdot\hat{\underset{\%}{V}}\left[\hat{\alpha}\right]\cdot\underset{\%}{B}^T\cdot\underset{\%}{A}^{-1}$$

$$\hat{\underset{\%}{V}}\left[\hat{\alpha}\right] = \left(\underset{\%}{B}^T\cdot\underset{\%}{A}^{-1}\cdot\underset{\%}{B}\right)^{-1} \qquad 7.4.12$$

καθώς και στη διαπίστωση ότι οι παράμετροι θ έχουν μηδενική συσχέτιση με τους πολλαπλασιαστές α.

Παρατηρήστε ότι ο πρώτος όρος στην έκφραση των στοιχείων του πίνακα συνδιασποράς των παραμέτρων θ αντιστοιχεί στα στοιχεία του αντίστοιχου πίνακα στην περίπτωση που δεν

---

[6] Αναπτύσσοντας τις μερικές παραγώγους, $\frac{\partial F(x;\underline{\theta},\underline{\alpha})}{\partial\theta_i}$ και $\frac{\partial F(x;\underline{\theta},\underline{\alpha})}{\partial\alpha_i}$ σε σειρά Taylor, γύρω από τις πραγματικές τιμές των παραμέτρων. Βλέπε Eadie, Sect 8.3.4

επιβάλλεται καμία συνθήκη δέσμευσης. Ο αρνητικός δεύτερος όρος εκφράζει την μείωση της διασποράς που επιφέρει η επιπλέον πληροφορία την οποία μεταφέρουν οι συνθήκες δέσμευσης.

Στην περίπτωση όπου ο πίνακας $\left(-E\left[\begin{pmatrix} \underset{\%}{A} & \underset{\%}{B} \\ \underset{\%}{B^T} & \underset{\%}{0} \end{pmatrix}\right]\right)^{-1}_{\substack{\underline{\theta}=\hat{\underline{\theta}} \\ \underline{\alpha}=\hat{\underline{\alpha}}}}$ δεν είναι αντιστρέψιμος ενδείκνυται να τροποποιηθεί η σχέση 7.4.8 ως:

$$F'(\underline{x};\underline{\theta},\underline{\alpha}) = \ln L(\underline{x};\underline{\theta}) + \sum_{i=1}^{m} \alpha_i \cdot g_i(\underline{\theta}) - \sum_{i=1}^{m} g_i^2(\underline{\theta}) \qquad 7.4.13$$

και να επαναληφθεί η μέθοδος που περιγράφηκε προηγουμένως.

7.5.2 Μέθοδος των ελαχίστων τετραγώνων με Lagrangian πολλαπλασιαστές.

Ας εξετάσουμε την περίπτωση όπου τα δεδομένα συνίστανται από ζεύγη μετρήσεων, $\{y_i, x_i\}, (i = 1,2,3K\ ,N)$ όπου οι μετρήσεις x_i δεν πάσχουν από μετρητικά σφάλματα. Επιδιώκουμε την προσαρμογή του γραμμικού προτύπου:

$$y(x) = \underset{\%}{H} \cdot \underset{\%}{\theta} = \sum_{\lambda=1}^{k} h_\lambda(x) \cdot \theta_\lambda \qquad 7.4.14$$

όπου ο μονοδιάστατος πίνακας (διάνυσμα) $\underset{\%}{\theta} = (\theta_1\ \theta_2\ K\ \theta_k)$ παριστά τις k παραμέτρους που θα εκτιμηθούν. Οι τυχαίες μεταβλητές, ε_i, που παριστούν τις αποκλίσεις των μετρήσεων από τις προβλέψεις του προτύπου προσαρμογής:

$$y_i = \sum_{\lambda=1}^{k} h_\lambda(x_i) \cdot \theta_\lambda + \varepsilon_\iota \qquad 7.4.15$$

έχουν συνδιασπορές που εκφράζονται από τα στοιχεία του πίνακα V.

Επιπλέον, θα υποθέσουμε ότι οι παράμετροι υπόκεινται στις ακόλουθες, m συνθήκες δέσμευσης:

$$\sum_{j=1}^{k} l_{ij} \cdot \theta_j = R_i\ , \quad (i = 1,2,3,...,m) \qquad 7.4.6$$

$$\underset{\%}{L} \cdot \underset{\%}{\theta} = \underset{\%}{R}$$

όπου ο πίνακας $\underset{\%}{L}$ έχει ως στοιχεία τους συντελεστές l_{ij}.

Θα χρησιμοποιήσουμε m Langrangian πολλαπλασιαστές, τους οποίους παριστούν στα στοιχεία του μονοδιάστατου πίνακα $\underset{\%}{\alpha}^T = (\alpha_1\ \alpha_1 K\ \alpha_m)^T$ και θα ορίσουμε την ακόλουθη στατιστική συνάρτηση :

$$\begin{aligned} Q_L^2 &= Q^2 + 2 \cdot \underset{\%}{\alpha}^T \cdot (\underset{\%}{L} \cdot \underset{\%}{\theta} - \underset{\%}{R}) \\ &= (\underset{\%}{y} - \underset{\%}{H} \cdot \underset{\%}{\theta})^T \cdot \underset{\%}{V}^{-1} \cdot (\underset{\%}{y} - \underset{\%}{H} \cdot \underset{\%}{\theta}) + 2 \cdot \underset{\%}{\alpha}^T \cdot (\underset{\%}{L} \cdot \underset{\%}{\theta} - \underset{\%}{R}) \end{aligned} \qquad 7.4.7$$

όπου το διάνυσμα $\underset{\sim}{y}$ παριστά τις μετρήσεις y_i $(i = 1, 2, 3K\ , N)$.

Οι συνθήκες ακρότατου της ποσότητας 7.4.7:

$\frac{\partial Q_L^2}{\partial \theta_i} = 0$ και $\frac{\partial Q_L^2}{\partial \alpha_\lambda} = 0$ (για $i = 1, 2, ...k$ και $\lambda = 1, 2, ..., m$) , καταλήγουν στις ακόλουθες κανονικές εξισώσεις:

$$\underset{\sim}{H}^T \cdot \underset{\sim}{V}^{-1} \cdot \underset{\sim}{H} \cdot \underset{\sim}{\hat{\theta}} + \underset{\sim}{L}^T \cdot \underset{\sim}{\hat{\alpha}} = \underset{\sim}{H}^T \cdot \underset{\sim}{V}^{-1} \cdot \underset{\sim}{y} \qquad 7.4.8$$
$$\underset{\sim}{L} \cdot \underset{\sim}{\hat{\theta}} = \underset{\sim}{R}$$

Οι εξισώσεις 7.4.8, μετά την αντικατάσταση:

$$\underset{\sim}{C} = \underset{\sim}{H}^T \cdot \underset{\sim}{V}^{-1} \cdot \underset{\sim}{H} \qquad 7.4.9$$
$$\underset{\sim}{S} = \underset{\sim}{H}^T \cdot \underset{\sim}{V}^{-1} \cdot \underset{\sim}{y}$$

γράφονται ως:

$$\begin{pmatrix} \underset{\sim}{C} & \underset{\sim}{L}^T \\ \underset{\sim}{L} & \underset{\sim}{0} \end{pmatrix} \cdot \begin{pmatrix} \underset{\sim}{\hat{\theta}} \\ \underset{\sim}{\hat{\alpha}} \end{pmatrix} = \begin{pmatrix} \underset{\sim}{S} \\ \underset{\sim}{R} \end{pmatrix} \qquad 7.4.10$$

Εάν οι πίνακες $\underset{\sim}{C}$ και $\underset{\sim}{L} \cdot \underset{\sim}{C}^{-1} \cdot \underset{\sim}{L}^T$ μπορούν να αντιστραφούν, η λύση της 7.4.10 δίνεται[7] ως:

$$\begin{pmatrix} \underset{\sim}{\hat{\theta}} \\ \underset{\sim}{\hat{\alpha}} \end{pmatrix} = \begin{pmatrix} \underset{\sim}{F} & \underset{\sim}{G}^T \\ \underset{\sim}{G} & \underset{\sim}{E} \end{pmatrix} \cdot \begin{pmatrix} \underset{\sim}{S} \\ \underset{\sim}{R} \end{pmatrix} \qquad 7.4.11$$

$$\underset{\sim}{W} = \left(\underset{\sim}{L} \cdot \underset{\sim}{C}^{-1} \cdot \underset{\sim}{L}^T\right)^{-1}$$

όπου $$\underset{\sim}{F} = \left(\underset{\sim}{1} - \underset{\sim}{C}^{-1} \cdot \underset{\sim}{L}^T \cdot \underset{\sim}{W} \cdot \underset{\sim}{L}\right) \cdot \underset{\sim}{C}^{-1} \qquad 7.4.12$$
$$\underset{\sim}{G} = \underset{\sim}{W} \cdot \underset{\sim}{L} \cdot \underset{\sim}{C}^{-1}$$
$$\underset{\sim}{E} = -\underset{\sim}{W}$$

Οι συνδιασπορές μεταξύ των παραμέτρων εκφράζονται[8] από τα στοιχεία των ακολούθων πινάκων:

$$\underset{\sim}{\hat{V}}\left[\underline{\hat{\theta}}\right] = \underset{\sim}{F}$$
$$\underset{\sim}{\hat{V}}\left[\underline{\hat{\alpha}}\right] = \underset{\sim}{W} \qquad 7.4.13$$
$$\mathrm{cov}\left(\underline{\hat{\theta}}, \underline{\hat{\alpha}}\right) = 0$$

Αποδεικνύεται[9], ότι οι εκτιμήσεις των παραμέτρων θ δεν πάσχουν από προκατάληψη, όπως ακριβώς και οι εκτιμήσεις με την μέθοδο ελαχίστων τετραγώνων με γραμμικό πρότυπο

---

[7] Βλέπε Eadie et al, Sect. 8.4.2
[8] Βλέπε Eadie et al, Sect. 8.4.2
[9] Βλέπε Eadie et al, Sect. 8.4.2

προσαρμογής, και ότι η αναμενόμενη τιμή των Lagrangian πολλαπλασιαστών είναι ίση με μηδέν.