ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
Διατμηματικό Πρόγραμμα Μεταπτυχιακών Σπουδών
των τμημάτων Μαθηματικών και Εφαρμοσμένων
Μαθηματικών

Μεταπτυχιακή Εργασία

Μη παραμετρική εκτίμηση συναρτήσεων πυκνότητας πιθανότητας με μεθόδους ποινικοποιημένης πιθανοφάνειας σύμφωνα με το μοντέλο Koziol – Green

Ξιφανταράκης Ιωάννης

Επιβλέπων Καθηγητής: Κλωνιάς Βασίλειος

(Κατεύθυνση: Επιχειρησιακά Μαθηματικά)

Ηράκλειο, 2005

Περιεχόμενα

Η εργασία αυτή παρουσιάστηκε στις 29 Νοεμβρίου του 2005, ενώπιον της τριμελούς κριτικής επιτροπής, η οποία αποτελούνταν από την κα Σουζάνα Παπαδοπούλου, την κα Δούκισσα Κρητικού και τον κ. Βασίλειο Κλωνιά (επιβλέπων καθηγητής).

Πριν προχωρήσω στην ουσία αυτής της εργασίας, θα ήθελα να ευχαριστήσω τα μέλη της προαναφερθείσας τριμελούς κριτικής επιτροπής, καθώς και όλους τους καθηγητές του τμήματος Μαθηματικών του Πανεπιστημίου Κρήτης, διαλέξεις των οποίων είχα την τύχη να παρακολουθήσω κατά την διάρκεια των προπτυχιακών και μεταπτυχιακών σπουδών μου.

Επίσης, θα ήθελα να αφιερώσω αυτήν την εργασία στην οικογένεια μου και στους φίλους μου, χωρίς την συμπαράσταση και την υποστήριξη των οποίων δεν θα μπορούσα να ολοκληρώσω τις μεταπτυχιακές μου σπουδές.

1. ΕΙΣΑΓΩΓΗ

Έστω $X_1, X_2, ..., X_n$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές από μία άγνωστη κατανομή F πάνω στον $\mathbb{R}_+ = (0, +\infty)$, με μία συνάρτηση πυκνότητας f. Επίσης, έστω $Y_1, Y_2, ..., Y_n$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές κοπής δεδομένων, ανεξάρτητες των X_i, από μία άγνωστη κατανομή G πάνω στον $\mathbb{R}_+$. Στο μοντέλο της ανεξάρτητης τυχαίας κοπής δεδομένων, τα δεδομένα είναι οι παρατηρήσεις (Z_i, δ_i), όπου

$$Z_i := \min\{X_i, Y_i\},\ \delta_i := \mathbf{1}(X_i \le Y_i),\ i = 1, 2, ..., n.$$

Χωρίς καμία επιπλέον πληροφορία για την G, ο εκτιμητής μεγίστης πιθανοφάνειας για την F, ο οποίος εισήχθη απ' τους Kaplan και Meier (1958), δίνεται από την εξής σχέση

$$\hat{F}_n(t) := 1 - \prod_{i=1}^{n} \left(1 - \frac{\mathbf{1}(Z_{ni} \le t)}{n+1-i}\right)^{\delta_{ni}},\ t \in \mathbb{R}_+ \tag{1.1}$$

όπου $Z_{n1} \le \cdots \le Z_{nn}$ είναι το διατεταγμένο δείγμα των Z_i και $\delta_{n1}, \ldots, \delta_{nn}$ είναι τα αντίστοιχα δ_i.

Εδώ, μας ενδιαφέρει η εκτίμηση της συνάρτησης πυκνότητας πιθανότητας f, με τεχνικές ποινικοποιημένης μεγίστης πιθανοφάνειας, υπό το πιο συγκεκριμένο μοντέλο της αναλογικότητας των συναρτήσεων κινδύνου των F και G, δηλαδή

$$1 - G = (1 - F)^c, \tag{1.2}$$

για κάποιο ακαθόριστο $c > 0$. Οι εκτιμητές είναι ανάλογοι αυτών στο άρθρο των Κλωνιά και Weierman (1985), που παράγονται από την σχέση (1.1), για το μη περιορισμένο μοντέλο τυχαίας κοπής δεδομένων.

Οι Koziol και Green (1976), οι οποίοι μελέτησαν το μοντέλο τυχαίας κοπής δεδομένων υπό τη συνθήκη (1.2) της αναλογικότητας των κινδύνων, παρέχουν επαρκές κίνητρο για την κατανόησή του και δίνουν στοιχεία για την ευρεία εφαρμοσιμότητά του. Σύμφωνα με τους Csorgo και Horvath (1981), αυτό το μοντέλο συχνά αναφέρεται ως μοντέλο Koziol – Green. Κάτω απ' την συνθήκη (1.2), τα

$$Z_1, Z_2, ..., Z_n \text{ και τα } \delta_1, \delta_2, ..., \delta_n \text{ είναι ανεξάρτητα.} \tag{1.3}$$

Για να το δείξουμε αυτό ακολουθούμε την εξής διαδικασία: (ξεχνάμε προς το παρόν τους δείκτες για λόγους απλότητας)

$$P(\delta = 1) = P(X \le Y) = \int_0^{+\infty} (1 - G)\,dF = \int_0^{+\infty} (1 - F)^c\,dF$$
$$= \int_0^1 (1 - t)^c\,dt = \frac{1}{c+1},$$

επίσης

$$P(Z \ge z) = P(\min\{X, Y\} \ge z) = (1 - F(z))(1 - G(z))$$
$$= (1 - F(z))^{c+1},$$

χρησιμοποιώντας ότι οι X, Y είναι ανεξάρτητες. Οπότε παίρνοντας την περίπτωση $\delta = 1$ (όμοια για $\delta = 0$) έχουμε

$$
\begin{aligned}
P\left(Z \geq z,\ \delta=1\right) &= P\left(\min\{X,Y\} \geq z,\ X \leq Y\right) = P\left(z \leq X \leq Y\right) \\
&= \int_z^{+\infty}\left[F(y)-F(z)\right]dG(y) \\
&= \left(1-F(z)\right)\left(1-G(z)\right)-\int_z^{+\infty}\left(1-F\right)dG \\
&= \left(1-F(z)\right)^{c+1}+\int_z^{+\infty}\left(1-F\right)d\left(1-G\right) \\
&= \left(1-F(z)\right)^{c+1}+c\int_{1-F(z)}^{0} t d\left(t^c\right) \\
&= \frac{1}{c+1}\left(1-F(z)\right)^{c+1} = P\left(\delta=1\right)P\left(Z \geq z\right).
\end{aligned}
$$

Σ' αυτό το πλαίσιο, οι Abdushukurov (1984), Cheng και Lin (1984) και (1987) έδειξαν ότι ο εκτιμητής μεγίστης πιθανοφάνειας της F δίνεται από τη σχέση

$$\tilde{F}_n(t) := 1-\prod_{i=1}^{n}\left(1-\frac{\mathbf{1}(Z_{ni} \leq t)}{n+1-i}\right)^{\bar{\delta}_n} = 1-\{1-H_n(t)\}^{\hat{p}_n},\ t \in \mathbb{R}_+ \tag{1.4}$$

όπου $H_n(t) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(Z_i \leq t)$ είναι η συνήθης εμπειρική συνάρτηση κατανομής των Z_i και $\hat{p}_n = \overline{\delta}_n = \frac{1}{n}\sum_{i=1}^{n}\delta_i$ είναι ο εκτιμητής μεγίστης πιθανοφάνειας της $p := P(\delta_1 = 1) = \int_0^{+\infty} FdG = (1+c)^{-1}$. Επίσης δείχνουν ότι υπό τη συνθήκη (1.2), η $\tilde{F}_n$ είναι ασυμπτωτικά περισσότερο επαρκής από την $\hat{F}_n$. Ο εκτιμητής $\tilde{F}_n$ είναι γνωστός και ως ACL εκτιμητής, από τα αρχικά των ονομάτων των ανωτέρω συγγραφέων.

Οι Csorgo και Mielniczuk (1988) μελέτησαν τις ασυμπτωτικές ιδιότητες των εκτιμητών πυκνοτήτων πυρήνων και κοντινοτέρων περιοχών, στο περιορισμένο – από την (1.2) – μοντέλο τυχαίας κοπής δεδομένων, όπου τον ρόλο της εμπειρικής συνάρτησης κατανομής τον παίζει τώρα η (1.4). Επιπρόσθετα αποτελέσματα πάνω σ' αυτήν τη μέθοδο εκτίμησης παρουσιάζονται στο ερευνητικό άρθρο του Csorgo (1988), όπου μελετούνται πολλά νέα αποτελέσματα και το bootsrapping. Οι Ghorai και Rejto (1987) μελετούν την εκτίμηση της μέσης υπολειπόμενης ζωής σ' αυτό το πλαίσιο και επεκτείνοντας ένα αποτέλεσμα των Cheng και Lin (1984, 1987), βρήκαν τον ακριβή ρυθμό ομοιόμορφης σύγκλισης πάνω στο $\mathbb{R}_+$ της $\tilde{F}_n$ στην F – αποτέλεσμα ανάλογο αυτού που χρειαζόμαστε εδώ για την ασυμμετρική q-norm σύγκλιση της $\tilde{F}_n$ στην F (βλέπε λήμμα 3.1).

Η μέθοδος εκτίμησης πυκνοτήτων που θεωρούμε εδώ στηρίζεται σε τεχνικές ποινικοποιημένης μεγίστης πιθανοφάνειας, ανάλογες μ' εκείνες των Κλωνιά και Weierman (1985). Η μέθοδος εκτίμησης της ποινικοποιημένης μεγίστης πιθανοφάνειας (MPLE) συναρτήσεων πυκνότητας εισήχθη από τους Good και Gaskins (1971, 1980) στην περίπτωση των μη λογοκριμένων δεδομένων. Η MPLE f_n της f είναι εκείνη που μεγιστοποιεί την πιθανοφάνεια $\prod_{i=1}^{n} f(X_i)$ πάνω σ' ένα χώρο «ομαλών» συναρτήσεων – επομένως αποφεύγοντας την άφρακτη λύση του Dirac-δέλτα στην περίπτωση του μη περιορισμένου προβλήματος. Οπότε η f_n είναι η συνάρτηση που μεγιστοποιεί την παράσταση

$$n\int \log f dF_n - \Phi(f), \tag{1.5}$$

υπό τον περιορισμό $\int f = 1$ και $f \geq 0$, όπου $\Phi(f)$ είναι μία *ποινή "μη ομαλότητας"* και η F_n συμβολίζει την εμπειρική συνάρτηση κατανομής των X_i. Οι De Montricher, Tapia και Thompson (1975) απέδειξαν αυστηρά την ύπαρξη και τη μοναδικότητα των MPLE, χρησιμοποιώντας χώρους Sobolev. Ο Silverman (1982) μελέτησε την συνέπεια και την θεωρία ασυμπτωτικών κατανομών για μία ομάδα εκτιμητών, την οποία πρότεινε, με ποινές μη ομαλότητας πάνω στο $\log f$. Ο Κλωνιάς (1984) μελέτησε τις ιδιότητες συνέπειας των MPLE πάνω στο $\mathbb{R}$ για μία ευρεία κλάση συναρτησοειδών ποινής πάνω στο $f^{\frac{1}{2}}$. Οι MPLE που θεωρούμε εδώ αντιστοιχούν στην πρώτη MPLE των Good και Gaskins για την περίπτωση των μη λογοκριμένων δεδομένων, δηλαδή η $\Phi(f)$ είναι ανάλογη του συναρτησοειδούς της πληροφορίας κατά Fisher, οι ιδιότητες συνέπειας του οποίου μελετήθηκαν στο άρθρο του Κλωνιά (1982).

Για το γενικό μοντέλο τυχαίας κοπής δεδομένων, οι Lubecke και Padgett (1985) πρότειναν την εκτίμηση της πυκνότητας από τον μεγιστοποιητή της ποινικοποιημένης υπό συνθήκη πιθανοφάνειας, δεδομένου των παρατηρήσεων που έχουν λογοκριθεί. Τα ζητήματα τις συνέπειας και του αριθμητικού υπολογισμού του προκύπτοντα εκτιμητή δεν αντιμετωπίστηκαν εκεί. Επίσης, είναι πιθανόν ότι η ίδια προσέγγιση μπορεί να γίνει και στο πλαίσιο του μοντέλου τυχαίας κοπής δεδομένων των Koziol – Green. Όμως, η ανάλυση – στατιστική και αριθμητική – ενός τέτοιου εκτιμητή φαίνεται αρκετά πιο περίπλοκη από του εκτιμητή που προτείνουμε εδώ. Αυτό συμβαίνει κυρίως επειδή, στο δικό μας πλαίσιο, ένας τέτοιος εκτιμητής θα οριζόταν πεπλεγμένα, μέσω μίας εξίσωσης που θα εμπεριείχε και τους εκτιμητές της συνάρτησης πυκνότητας, αλλά και της συνάρτησης κατανομής ταυτόχρονα. Με τη δική μας προσέγγιση, που θα περιγράψουμε στην επόμενη ενότητα, λύνουμε ένα πρόβλημα βελτιστοποίησης ανάλογο του (1.5), με τον ACL εκτιμητή $\tilde{F}_n$ στη θέση της εμπειρικής συνάρτησης κατανομής F_n. Αυτή η αντικατάσταση μας οδηγεί στην μεγιστοποίηση όχι του συναρτησοειδούς της ποινικοποιημένης πιθανοφάνειας, αλλά ενός εμπειρικού συναρτησοειδούς πληροφορίας τύπου Kullback – Leibler. Η ποινή μη ομαλότητας λαμβάνεται να είναι ανάλογη της πληροφορίας κατά Fisher $\int \frac{f'}{f} dF = 4\int (v')^2$, όπου $v = f^{\frac{1}{2}}$. Τότε ο εκτιμητής της f είναι η λύση αυτού του προβλήματος και συμβολίζεται με $f_n = u_n^{\,2}$, όπου u_n είναι ο εκτιμητής της τετραγωνικής ρίζας της πυκνότητας v – βλέπε Πρόταση 2.1. Στην ενότητα 3 παίρνουμε την σ.β. συνέπεια της u_n με τις L_p νόρμες, $p = 2, \infty$ και τη Sobolev $W^{2,1}$ νόρμα. Ο υπολογισμός των εκτιμητών περιγράφεται στο τέλος της ενότητας 2 και βασίζεται στους αλγορίθμους των Κλωνιά και Nash (1987). Οι ρυθμοί σύγκλισης που παίρνουμε εξαρτώνται από το μέγεθος της κοπής των δεδομένων μέσω του p.

2. ΔΙΑΤΥΠΩΣΗ ΤΩΝ ΕΚΤΙΜΗΤΩΝ

Στο παρόν πλαίσιο του μοντέλου Koziol – Green, το πρόβλημα μεγιστοποίησης, ανάλογο του (1.5), είναι εκείνο της μεγιστοποίησης της παράστασης

$$n\int \log f d\tilde{F}_n - \alpha I\left(f\right) \tag{2.1$^{\alpha}$}$$

υπό τον περιορισμό

$$\int f = 1 \text{ και } f \geq 0, \tag{2.1β}$$

όπου $\alpha > 0$ και $I\left(f\right)$ είναι η πληροφορία κατά Fisher, η οποία δίνεται από τον εξής τύπο:

$$I\left(f\right) = \int \left[\left(f^{\frac{1}{2}}\right)'\right]^2 = \frac{1}{4}\int \frac{\left(f'\right)^2}{f}.$$

Παρατηρείστε ότι η παράσταση $n\int \log f d\tilde{F}_n$ δεν είναι η log-πιθανοφάνεια των δεδομένων $\left(Z_i, \delta_i\right)$, $i = 1, 2, ..., n,$ όπως ήταν στην (1.5) στην περίπτωση ελλιπών δεδομένων. Το κίνητρο μας εδώ έχει περισσότερο να κάνει με το γεγονός ότι το $\int \log f dF$ μεγιστοποιείται πάνω σε όλες τις πυκνότητες πιθανότητας f, για $f = F'$ σ.β. Συνεπώς, μέσω του προβλήματος μεγιστοποίησης (2.1), ψάχνουμε για ένα εκτιμητή πυκνότητας f που να βρίσκεται κοντά στην "παράγωγο" του ACL εκτιμητή $\tilde{F}_n$, ο οποίος είναι ο εκτιμητής μεγίστης πιθανοφάνειας της F στο πλαίσιο του μοντέλου Koziol – Green. Το συναρτησοειδές ποινής "μη ομαλότητας", το οποίο στην περίπτωσή μας είναι η πληροφορία κατά Fisher $I\left(f\right)$, μας εγγυάται την φραξιμότητα του συναρτησοειδούς στο πρόβλημα (2.1).

Ο ACL εκτιμητής $\tilde{F}_n$ της F μπορεί να παρασταθεί με την εξής μορφή:

$$\tilde{F}_n(t) = \frac{1}{n}\sum_{i=1}^{n} w_{ni} 1(Z_{ni} \leq t), \ t \in \mathbb{R}_+ \tag{2.2}$$

όπου Z_{ni}, $i = 1, 2, \ldots, n$ είναι το διατεταγμένο δείγμα των Z_i, $i = 1, 2, \ldots, n$ και τα αντίστοιχα βάρη δίνονται από τη σχέση

$$w_{ni} := n\left\{\left(1 - \frac{i-1}{n}\right)^{\hat{p}_n} - \left(1 - \frac{i}{n}\right)^{\hat{p}_n}\right\}, i = 1, 2, \ldots, n. \tag{2.3}$$

Τότε, χρησιμοποιώντας την (2.2), παίρνουμε ένα πρόβλημα βελτιστοποίησης ανάλογο του (2.1), που είναι εκείνο της μεγιστοποίησης της παράστασης

$$\sum_{i=1}^{n} w_{ni} \log u^2(Z_{ni}) - \alpha \int_A (u')^2, \;\; u \in H(A) \tag{2.4α}$$

υπό τον περιορισμό

$$\int_A u^2 = 1\,,\; u(Z_{ni}) \geq 0, i = 1,2,\ldots,n, \tag{2.4β}$$

όπου θεωρούμε $u := f^{\frac{1}{2}} \geq 0$ και συμβολίζουμε $H(A) := \{u \in L_2(A) : u' \in L_2(A)\}$, $A = \mathbb{R}_+$ ή $\mathbb{R}$, δηλαδή τον χώρο Hilbert των τετραγωνικά ολοκληρώσιμων συναρτήσεων με τετραγωνικά ολοκληρώσιμες παραγώγους πάνω στο *A*. Το εσωτερικό γινόμενο του $H(A)$ ορίζεται να είναι

$$\langle u, v \rangle := \int_A uv + \beta \int_A u'v', \;\; \beta > 0 \tag{2.5}$$

και η επαγόμενη νόρμα είναι

$$\|u\| := \langle u, u \rangle^{\frac{1}{2}} = \left\{ \|u\|_2^2 + \beta \|u'\|_2^2 \right\}^{\frac{1}{2}}, \;\; u \in H(A), \tag{2.6}$$

όπου

$$\|u\|_2^2 := \int_A u^2. \tag{2.7}$$

Ο χώρος $\left(H(A), \langle \cdot, \cdot \rangle\right)$ στην πραγματικότητα είναι ένας χώρος Hilbert αναπαραγόμενου πυρήνα (reproducing kernel Hilbert space), δηλαδή υπάρχει $k(t,s;\beta^{\frac{1}{2}}),\;\; t,s \in A,$ τέτοιο ώστε

$$k(\cdot, s; \beta^{\frac{1}{2}}) \in H(A), \;\; \forall s \in A$$

και

$$\left\langle u, k(\cdot, s; \beta^{\frac{1}{2}}) \right\rangle = u(s) \;\; \forall u \in H(A), s \in A.$$

Άρα, για $A = \mathbb{R},$ η παραπάνω εξίσωση (ως προς k) λύνεται ως εξής:

Για το πρώτο μέλος έχουμε

$$\left\langle k(\cdot,s;\beta^{\frac{1}{2}}),u(\cdot)\right\rangle = \int_{\mathbb{R}} k(t,s;\beta^{\frac{1}{2}})u(t)dt + \beta\int_{\mathbb{R}} \frac{\partial}{\partial t}k(t,s;\beta^{\frac{1}{2}})u'(t)dt$$

$$= \int_{\mathbb{R}} k(t,s;\beta^{\frac{1}{2}})u(t)dt - \beta\int_{\mathbb{R}} \frac{\partial^2}{\partial t^2}k(t,s;\beta^{\frac{1}{2}})u(t)dt + \underbrace{\left.\frac{\partial}{\partial t}k(t,s;\beta^{\frac{1}{2}})u(t)\right|_{-\infty}^{+\infty}}_{=0}$$

$$= \int_{\mathbb{R}} \left(k(t,s;\beta^{\frac{1}{2}}) - \beta\frac{\partial^2}{\partial t^2}k(t,s;\beta^{\frac{1}{2}}) \right)u(t)dt.$$

Ο όρος $\left.\frac{\partial}{\partial t}k(t,s;\beta^{\frac{1}{2}})u(t)\right|_{-\infty}^{+\infty}$ ισούται με μηδέν επειδή η u είναι πυκνότητα.

Για το δεύτερο μέλος της εξίσωσης έχουμε

$$u(s) = \int_{\mathbb{R}} \delta(t-s)u(t)dt,$$

όπου δ είναι το δέλτα του Dirac.

Εξισώνοντας τα δύο μέλη, παίρνουμε την εξής μερική διαφορική εξίσωση:

$$k(t,s;\beta^{\frac{1}{2}}) - \beta\frac{\partial^2}{\partial t^2}k(t,s;\beta^{\frac{1}{2}}) = \delta(t-s) \quad \text{σ.π.}$$

Περνώντας στους μετασχηματισμούς Fourier, έχουμε

$$\left.\begin{array}{r} \tilde{k}_s(\xi) - \beta\tilde{k}_s''(\xi) = \int_{\mathbb{R}} e^{-i\xi t}\delta(t-s)\,dt = e^{-i\xi s} \\ \tilde{k}_s''(\xi) = (-i\xi)^2\tilde{k}_s(\xi) = -\xi^2\tilde{k}_s(\xi) \end{array}\right\} \Rightarrow$$

$$\Rightarrow (1+\beta\xi^2)\tilde{k}_s(\xi) = e^{-i\xi s}$$

$$\Rightarrow \tilde{k}_s(\xi) = \frac{e^{-i\xi s}}{1+\beta\xi^2}.$$

Όμως, το δεύτερο μέλος της τελευταίας σχέσης είναι ο μετασχηματισμός Fourier της παράστασης $\frac{\beta^{\frac{-1}{2}}}{2}e^{-\beta^{\frac{-1}{2}}|t-s|}$.

Επομένως,

$$k(t,s;\beta^{\frac{1}{2}}) = \beta^{\frac{-1}{2}}\kappa(\beta^{\frac{-1}{2}}(t-s)), \qquad (2.8)$$

όπου $\kappa(t) = \frac{1}{2}e^{-|t|}, t \in \mathbb{R}$

Με ανάλογο τρόπο, για $A=\mathbb{R}_+$,

$$k(t,s;\beta^{\frac{1}{2}})=(4\beta)^{\frac{-1}{2}}\left\{\kappa(\beta^{\frac{-1}{2}}(t-s))+\kappa(\beta^{\frac{-1}{2}}(t+s))\right\}. \qquad (2.9)$$

Η μόνη διαφορά ανάμεσα στα προβλήματα (2.1) και (2.4) είναι ότι στο τελευταίο απαιτούμε το u να είναι μη αρνητικό μόνο στα σημεία του δείγματός. Στην πραγματικότητα τα προβλήματα (2.1) και (2.4) είναι ισοδύναμα (βλέπε παρακάτω την Πρόταση 2.1).

Το παρακάτω λήμμα οφείλεται στους De Montricher, Tapia και Thompson (1975) και θα μας βοηθήσει στην απόδειξη της Πρότασης 2.1.

Λήμμα 2.1

Έστω Η ένα υποσύνολο του $L_2(\Omega)$ και J ένα συναρτησοειδές ορισμένο πάνω στο Η. Θεωρούμε το Πρόβλημα Ι, της μεγιστοποίησης του $J\left(\upsilon^{\frac{1}{2}}\right)$, υπό τους περιορισμούς $\upsilon\in \mathrm{H}$, $\int_\Omega \upsilon(t)dt=1$ και $\upsilon(t)\geq 0,\ \forall t\in\Omega$ και το Πρόβλημα ΙΙ της μεγιστοποίησης του $J(u)$, υπό τους περιορισμούς $u\in \mathrm{H}$ και $\int_\Omega u^2(t)dt=1$. Έστω u^* η λύση του Προβλήματος ΙΙ. Τότε η $\upsilon^*=\left(u^*\right)^2$ λύνει το Πρόβλημα Ι, αν και μόνο αν $\left|u^*\right|\in \mathrm{H}$ και $J\left(u^*\right)=J\left(\left|u^*\right|\right)$.

Απόδειξη:

Παρατηρούμε ότι $\left(\upsilon^*\right)^{\frac{1}{2}}=\left|u^*\right|$. Οπότε αν $\left|u^*\right|\notin \mathrm{H}$, τότε το $J\left(\left(\upsilon^*\right)^{\frac{1}{2}}\right)$ δεν ορίζεται και δεν έχει νόημα. Ενώ, αν $\left|u^*\right|\in \mathrm{H}$ και $J\left(u^*\right)=J\left(\left|u^*\right|\right)$, τότε για κάθε $\upsilon\geq 0$ τέτοιο ώστε $\upsilon^{\frac{1}{2}}\in \mathrm{H}$ έχουμε ότι

$$J\left(\upsilon^{\frac{1}{2}}\right)\leq J\left(u^*\right)=J\left(\left|u^*\right|\right)=J\left(\left(\upsilon^*\right)^{\frac{1}{2}}\right),$$

καθώς το u^* είναι ο μεγιστοποιητής. Επίσης παρατηρούμε ότι οι περιορισμοί ικανοποιούνται, άρα έχουμε το λήμμα. Ο.Ε.Δ.

■

Θεώρημα: [7, Παραρτήματος Ι, Tapia, Thompson (1978)]

Έστω S είναι ένα κλειστό και κυρτό υποσύνολο ενός χώρου Hilbert H. Θεωρούμε μία συνάρτηση $f : H \to \mathbb{R}$, η οποία να είναι συνεχής στο S και δύο φορές Gateaux διαφορίσιμη στο S, όπου το δεύτερο Gateaux διαφορικό της είναι ομοιόμορφα θετικά ορισμένο. Τότε το πρόβλημα της ελαχιστοποίησης της f πάνω στο S έχει μία μοναδική λύση.

Πρόταση 2.1

(α) Το πρόβλημα (2.4) έχει μοναδική λύση, η οποία είναι μία συνάρτηση spline που δίνεται πεπλεγμένα στην εξής μορφή:

$$\begin{aligned} u_n(t) &= \lambda_n^{-1} \sum_{i=1}^{n} w_{ni} [u_n(Z_{ni})]^{-1} k(t, Z_{ni}; h) \\ &= \frac{n}{\lambda_n} \int u_n(s)^{-1} k(t, s; h) d\tilde{F}_n(s), \quad t \in A, \end{aligned} \tag{2.10}$$

όπου $h := \left(\dfrac{\alpha}{\lambda_n} \right)^{\frac{1}{2}}$ και λ_n είναι ο πολλαπλασιαστής Lagrange που αφορά την περιοριστική ισότητα του προβλήματος (2.4).

(β) Το πρόβλημα (2.1) έχει μοναδική λύση, που δίνεται από τη σχέση

$$f_n = u_n^2. \tag{2.11}$$

Απόδειξη:

Το δεύτερο μέρος της πρότασης είναι άμεσο επακόλουθο του μέρους (α) και του Λήμματος 2.1, που αποδείξαμε προηγουμένως, αφού $u_n \geq 0$. Για το μέρος (α), παρατηρούμε ότι αν κάποιο από τα $u(Z_{ni})$ ήταν μηδέν, τότε το συναρτησοειδές του προβλήματος (2.4) θα ήταν $-\infty$ και οι πολλαπλασιαστές Lagrange, που αφορούν τις περιοριστικές ανισότητες του προβλήματος (2.4), θα ήταν όλοι ίσοι με μηδέν. Τότε, η Langragian του προβλήματος δίνεται από τη σχέση

$$L_\lambda(u) = \sum_{i=1}^{n} w_{ni} \log u^2(Z_{ni}) - \lambda \|u\|_2^2 - \alpha \|u'\|_2^2, \tag{2.12}$$

και αρκεί να λύσουμε το ακόλουθο πρόβλημα μεγιστοποίησης για κάθε λ>0:

$$\max \{ L_\lambda(u), u \in B \}, \tag{2.13}$$

όπου $B := \{u \in H(A) : u(Z_{ni}) \geq 0, i = 1, 2, \dots, n\}$, ένα κλειστό, κυρτό σύνολο. Η ύπαρξη μίας μοναδικής λύσης του (2.13) προκύπτει, σύμφωνα με το Θεώρημα 7 του παραρτήματος I των Tapia και Thompson (1978), από την συνέχεια της $L_\lambda(u)$ και από το γεγονός ότι η δεύτερη κατά Gateaux παράγωγος της $L_\lambda(u)$ είναι ομοιόμορφα αρνητικά ορισμένη. Για να δείξουμε το πρώτο, χρησιμοποιούμε την τριγωνική ανισότητα και βλέπουμε ότι, για $u_1, u_2 \in H(A)$,

$$\left| \|u_1\| - \|u_2\| \right| \leq \|u_1 - u_2\|$$

και από την ανισότητα Cauchy – Schwarz έχουμε:

$$|u_1(Z_{ni}) - u_2(Z_{ni})| = \left|\langle k(\cdot, Z_{ni}; h), u_1 - u_2 \rangle\right| \leq \|k(\cdot, Z_{ni}; h)\| \|u_1 - u_2\|.$$

Για το δεύτερο, κατ' αρχήν υπολογίζουμε το πρώτης τάξεως διαφορικό Gateaux της Langrangian, ως εξής:

$$\begin{aligned}
\nabla L_\lambda(u)(\eta) &= \frac{d}{d\varepsilon} L_\lambda(u + \varepsilon\eta)\bigg|_{\varepsilon=0} \\
&= \frac{d}{d\varepsilon}\left\{\sum_{i=1}^{n} w_{ni} \log\big(u(Z_{ni}) + \varepsilon\eta(Z_{ni})\big)^2 - \lambda\|u + \varepsilon\eta\|_2^2 - \alpha\|u' + \varepsilon\eta'\|_2^2\right\}\Bigg|_{\varepsilon=0} \\
&= \left\{2\sum_{i=1}^{n} w_{ni} \frac{\eta(Z_{ni})}{u(Z_{ni}) + \varepsilon\eta(Z_{ni})} - 2\lambda\int(u + \varepsilon\eta)\eta - 2\alpha\int(u' + \varepsilon\eta')\eta'\right\}\Bigg|_{\varepsilon=0} \\
&= 2\left\{\sum_{i=1}^{n} w_{ni} \frac{\eta(Z_{ni})}{u(Z_{ni})} - \left(\lambda\int u\eta + \alpha\int u'\eta'\right)\right\} \\
&= 2\left\{\sum_{i=1}^{n} w_{ni} \frac{\langle k(\cdot, Z_{ni}; h), \eta\rangle}{u(Z_{ni})} - \lambda\langle u, \eta\rangle\right\} \\
&= 2\left\langle \sum_{i=1}^{n} w_{ni} \frac{k(\cdot, Z_{ni}; h)}{u(Z_{ni})} - \lambda u, \eta \right\rangle.
\end{aligned}$$

Ύστερα, συνεχίζοντας από την τρίτη ισότητα των ανωτέρω υπολογισμών, υπολογίζουμε το δεύτερο διαφορικό Gateaux, ως ακολούθως:

$$\begin{aligned}
\nabla^2 L_\lambda(u)(\eta, \eta) &= \frac{d^2}{d\varepsilon^2} L_\lambda(u + \varepsilon\eta)\bigg|_{\varepsilon=0} \\
&= 2\left\{\sum_{i=1}^{n} w_{ni} \frac{\eta(Z_{ni})\big(-\eta(Z_{ni})\big)}{u^2(Z_{ni})} - \lambda\int\eta^2 - \alpha\int(\eta')^2\right\} \\
&= -2\left\{\sum_{i=1}^{n} w_{ni} \frac{\eta^2(Z_{ni})}{u^2(Z_{ni})} + \lambda\|\eta\|^2\right\},
\end{aligned}$$

η οποία παρατηρούμε ότι είναι αρνητική. Επομένως, η $L_\lambda(u)$ παρουσιάζει μέγιστο, το οποίο θα το βρούμε θέτοντας το πρώτο διαφορικό Gateaux ταυτοτικά ίσο με το μηδέν, δηλαδή

$$\nabla L_\lambda\left(u\right)\left(\eta\right)=2\left\langle \sum_{i=1}^{n} w_{ni}\frac{k\left(\cdot,Z_{ni};h\right)}{u\left(Z_{ni}\right)}-\lambda u,\eta\right\rangle=0,\ \text{για κάθε}\ \eta\in H\left(A\right),$$

το οποίο συνεπάγεται ότι

$$\left\|\lambda^{-1}\sum_{i=1}^{n} w_{ni}[u(Z_{ni})]^{-1}k(\cdot,Z_{ni};h)-u\right\|=0,$$

απ' όπου παίρνουμε

$$u_\lambda(t)=\lambda^{-1}\sum_{i=1}^{n} w_{ni}[u_\lambda(Z_{ni})]^{-1}k(t,Z_{ni};h),\ \ t\in A, \tag{2.14}$$

για κάθε $\lambda>0$. Η τιμή λ_n του πολλαπλασιαστή Lagrange λ στο βέλτιστο σημείο προσδιορίζεται από την εξίσωση της αντίστοιχης συνθήκης Kuhn – Tucker:

$$\left\|u_\lambda\right\|_2=1,\ \ \lambda>0, \tag{2.15}$$

η οποία δίδει $\lambda=n\left\|u_n\right\|_2^2$, όπου το u_n λαμβάνεται από την (2.14) για $\lambda=n$.

Η ύπαρξη και η μοναδικότητα της λύσης της εξίσωσης (2.15) προκύπτει από την Πρόταση 2.2, η οποία ακολουθεί παρακάτω. Χρησιμοποιούμε την προαναφερθείσα πρόταση, θεωρώντας

$$D=\left\{u_\lambda\in H\left(\Omega\right):u_\lambda\geq 0,\left\|u_\lambda\right\|_2=1\right\},$$

όπου το D είναι ένα κλειστό και κυρτό σύνολο.

Συμβολίζοντας την λύση της εξίσωσης (2.15) με λ_n και την αντίστοιχη u_{λ_n} με u_n, παίρνουμε τη λύση (2.10) για το πρόβλημα (2.4). Ο.Ε.Δ.

■

Πρόταση 2.2

Αν $H\left(A\right)$ είναι ένας χώρος Hilbert αναπαραγόμενου πυρήνα, όπου A είναι ένα διάστημα της πραγματικής ευθείας, και D είναι ένα κλειστό και κυρτό υποσύνολο του συνόλου

$$\left\{u\in H\left(\Omega\right):u\left(Z_i\right)\geq 0,\ i=1,2,...,n\right\},$$

το οποίο περιέχει τουλάχιστον μία u η οποία είναι θετική στα δειγματικά σημεία, τότε το ποινικοποιημένο συναρτησοειδές της πιθανοφάνειας έχει μοναδικό μέγιστο μέσα στο D.

<u>Απόδειξη:</u>

Από το γεγονός ότι ο $H\left(A\right)$ είναι χώρος Hilbert αναπαραγόμενου πυρήνα, έχουμε ότι

$$\left|u\left(Z_{ni}\right)\right| \leq \left\|k(\cdot, Z_{ni}, h)\right\| \left\|u\right\|, \text{ όπου } u \in H\left(\Omega\right),\ i = 1, 2, \ldots, n.$$

Αν θεωρήσουμε το συναρτησοειδές $L_\lambda\left(u\right)$, που ορίζεται στην (2.14), τότε θα έχουμε

$$\left|L_\lambda\left(u\right)\right| = \left|\sum_{i=1}^{n} w_{ni} \log u^2\left(Z_{ni}\right) - \lambda\left\|u\right\|\right| \leq 2n\log\left\|u\right\| - c_1\left\|u\right\|.$$

Παρατηρούμε ότι η συνάρτηση $g(x) = n\log x - x^2$ παρουσιάζει μέγιστο, οπότε είναι άνω φραγμένη. Οπότε $\left|L_\lambda\left(u\right)\right| \leq c_2$.

Έστω $M = \sup\left\{L_\lambda\left(u\right) : u \in D\right\}$. Τότε θα υπάρχει μία ακολουθία $\left\{u_j\right\}_{j\in\mathbb{N}} \subset D$, τέτοια ώστε $\lim_{j\to\infty} L_\lambda\left(u_j\right) = M$. Έχουμε ότι $M > 0$ και $\left\|u_j\right\| \leq c_3$, για κάθε $j \in \mathbb{N}$, αφού το D είναι φραγμένο. Αν θεωρήσουμε το σύνολο $\left\{u \in H\left(\Omega\right) : \left\|u\right\| \leq c_3\right\}$, τότε αυτό είναι κλειστό και φραγμένο. Επομένως υπάρχει υπακολουθία της $\left\{u_j\right\}_{j\in\mathbb{N}}$ - την οποία χάριν απλότητος θα την συμβολίζουμε και αυτήν με f_j - η οποία να συγκλίνει ασθενώς σε μία u^*. Η νόρμα του H είναι ασθενώς κάτω ημισυνεχής (υπενθυμίζουμε ότι σε ένα χώρο με νόρμα, V, μία συνάρτηση $f : K \to \mathbb{R}$, $K \subset V$ λέγεται *ασθενώς κάτω ημισυνεχής*, αν για κάθε ακολουθία $\left\{u_n\right\} \subset K$ με $u_n \to u \in K$, ισχύει $f(u) \leq \liminf_{n\to\infty} f\left(u_n\right)$), δηλαδή $\liminf_{j\to\infty}\left\|u_j\right\| \geq \left\|u^*\right\|$. Συνεπώς,

$$\limsup_{j\to\infty}\left(\sum_{j=1}^{n} w_{nj} \log u_j^2 - \lambda\left\|u_j\right\|\right) \leq \sum_{j=1}^{n} w_{nj} \log\left(u^*\right)^2 - \lambda\left\|u^*\right\|.$$

Όμως, το αριστερό μέλος της παραπάνω ανισότητας ισούται με M, ενώ το δεξί μέλος ισούται με $L_\lambda\left(u^*\right)$. Άρα $M \leq L_\lambda\left(u^*\right)$. Όμως, επειδή $u^* \in D$ και το M είναι το supremum, θα ισχύει η ισότητα, δηλαδή $M = L_\lambda\left(u^*\right)$. Άρα, το ποινικοποιημένο συναρτησοειδές της πιθανοφάνειας λαμβάνει μέγιστο μέσα στο D. Εξ υποθέσεως έχουμε ότι υπάρχει τουλάχιστον μία $u \in D$, δηλαδή μία u θετική στα δειγματικά σημεία, οπότε $L_\lambda\left(u\right) > 0$. Ξέρουμε ότι το πρόσημο του δεύτερου Gateaux διαφορικού είναι ταυτοτικά αρνητικό. Επομένως το L_λ λαμβάνει μέγιστο στο D. Αυτό το μέγιστο θα είναι μοναδικό επειδή το D είναι κυρτό και η συνάρτηση L_λ είναι γνήσια κοίλη. Ο.Ε.Δ.

■

Στην Πρόταση 2.1, λύνουμε το πρόβλημα (2.4) πάνω στον $H(\mathbb{R}_+)$ και στον $H(\mathbb{R})$ ταυτόχρονα. Οι αντίστοιχοι εκτιμητές για το $f^{\frac{1}{2}}$ συμβολίζονται με $u_n^{(1)}$ και $u_n^{(2)}$ αντίστοιχα και δίνονται και οι δύο από τη σχέση (2.10), αλλά με διαφορετικούς πυρήνες. Ο πυρήνας (2.9) δίνει τον $u_n^{(1)}$ και ο πυρήνας (2.8) δίνει τον $u_n^{(2)}$. Οι αντίστοιχοι εκτιμητές της f συμβολίζονται με $f_n^{(1)}$ και $f_n^{(2)}$ αντίστοιχα και υπολογίζονται από τη σχέση (2.11). Οι αποδείξεις για την συνέπεια, που βρίσκονται στην ενότητα 3, εφαρμόζονται και στις δύο περιπτώσεις ταυτόχρονα.

Στο μοντέλο τυχαίας κοπής δεδομένων, τα X_i και Y_i αντιπροσωπεύουν χρόνο, οπότε ο φυσιολογικός χώρος πάνω στον οποίο πρέπει να λυθεί το πρόβλημα (2.4) είναι ο $H(\mathbb{R}_+)$ και όχι ο $H(\mathbb{R})$. Όμως, όταν η $f^{\frac{1}{2}} \in H(\mathbb{R}) \bigcap H(\mathbb{R}_+)$ (όπως για παράδειγμα στην περίπτωση της Weibull με παράμετρο σχήματος $\gamma > 1$ – το $\gamma = 1$ αντιστοιχεί στην εκθετική κατανομή), αμφότεροι οι $f_n^{(1)}$ και $f_n^{(2)}$ είναι συνεπείς εκτιμητές της, όπως αποδεικνύεται στην ενότητα 3.

Για τον υπολογισμό του εκτιμητή, επιλέγουμε μια διαφορετική παραμετρικοποίηση για το πρόβλημα, δηλαδή αντί των παραμέτρων λ και α, επιλέγουμε τις λ και $h = \sqrt{\frac{\alpha}{\lambda}}$. Για τον εκτιμητή έχουμε την παρακάτω σχέση, η οποία βρίσκεται σε πεπλεγμένη μορφή:

$$u_\lambda(t) = \frac{1}{\lambda} \sum_{i=1}^{n} w_{ni} \frac{1}{u_\lambda(Z_{ni})} k(t, Z_{ni}; h).$$

Ισοδύναμα,

$$\sqrt{\lambda} u_\lambda(t) = \sum_{i=1}^{n} w_{ni} \frac{1}{\sqrt{\lambda} u_\lambda(Z_{ni})} k(t, Z_{ni}; h).$$

Οπότε για $t = Z_{ni}$, για κάποιο $i \in \{1, 2, ..., n\}$, παίρνουμε:

$$\sqrt{\lambda} u_\lambda(Z_{ni}) = \sum_{j=1}^{n} w_{nj} \frac{1}{\sqrt{\lambda} u_\lambda(Z_{nj})} k(Z_{ni}, Z_{nj}; h). \tag{2.16}$$

Τώρα, ορίζουμε

$$q_i := \frac{\sqrt{w_{ni}}}{\sqrt{\lambda} u_\lambda \left(Z_{ni}\right)}, \; i = 1, 2, ..., n.$$

Άρα, αντικαθιστώντας στη σχέση (2.16), έχουμε

$$\frac{1}{q_i} = \sum_{j=1}^{n} q_j \sqrt{\frac{w_{nj}}{w_{ni}}} k\left(Z_{ni}, Z_{nj}; h\right), \; i = 1, 2, ..., n.$$

Εδώ, το διάνυσμα $q = \left(q_1, q_2, ..., q_n\right)^T$ είναι άγνωστο και θέλουμε να το υπολογίσουμε.

Ορίζουμε τον πίνακα $C := \left[k\left(Z_{ni}, Z_{nj}; h\right)\right]_{\substack{i=1,2,...,n \\ j=1,2,...,n}}$, ο οποίος είναι ένας γνωστός, συμμετρικός και θετικά ορισμένος πίνακας.

Θεωρούμε την ποσότητα:

$$Q\left(q\right) := -2\sum_{i=1}^{n} \log q_i + q^T C q.$$

Για αυτήν την ποσότητα έχουμε

$$\nabla Q\left(q\right) = -2q^{-1} + 2Cq,$$

όπου $q^{-1} = \left(\frac{1}{q_1}, \frac{1}{q_2}, ..., \frac{1}{q_n}\right)^T$, ενώ ο πίνακας των δευτέρων μερικών παραγώγων είναι θετικά ορισμένος.

Συνεπώς, μπορούμε να ελαχιστοποιήσουμε την ποσότητα $Q\left(q\right)$, χρησιμοποιώντας αριθμητικές μεθόδους και μ' αυτόν τον τρόπο να υπολογίσουμε τα q_i, ουσιαστικά ως λύση του συστήματος $Cq = q^{-1}$. Κατόπιν, από το q μπορούμε να υπολογίσουμε την ποσότητα:

$$\sqrt{\lambda} u_\lambda \left(Z_{ni}\right) = \sum_{j=1}^{n} \sqrt{w_{nj}} q_j k\left(Z_{ni}, Z_{nj}; h\right).$$

Άρα, για τον υπολογισμό της u_λ αρκεί να υπολογίσουμε το λ:

Αν ολοκληρώσουμε και τα δύο μέλη της σχέσης

$$\sqrt{\lambda} u_\lambda \left(t\right) = \sum_{j=1}^{n} \sqrt{w_{nj}} q_j k\left(t, Z_{nj}; h\right)$$

αφού πρώτα τα υψώσουμε στο τετράγωνο, παίρνουμε

$$\int\left\{\sqrt{\lambda}u_\lambda(t)\right\}^2 dt=\int\left\{\sum_{j=1}^{n}\sqrt{w_{nj}}\,q_j k\left(t,Z_{nj};h\right)\right\}^2 dt=$$
$$=\sum_{i=1}^{n}\sum_{j=1}^{n}\sqrt{w_{ni}w_{nj}}\,q_i q_j\int k\left(t,Z_{ni};h\right)k\left(t,Z_{nj};h\right)dt.$$

Όμως, το πρώτο μέλος μπορεί να πάρει την εξής μορφή:

$$\int\left\{\sqrt{\lambda}u_\lambda(t)\right\}^2 dt=\lambda\int u_\lambda^2(t)\,dt=\lambda\left\|u_\lambda\right\|_2^2=\lambda.$$

Άρα,

$$\lambda=\sum_{i=1}^{n}\sum_{j=1}^{n}\sqrt{w_{ni}w_{nj}}\,q_i q_j\int k\left(t,Z_{ni};h\right)k\left(t,Z_{nj};h\right)dt.$$

3. ΙΣΧΥΡΗ ΣΥΝΕΠΕΙΑ ΤΩΝ ΕΚΤΙΜΗΤΩΝ

Ο σκοπός μας σ' αυτήν την ενότητα είναι να δείξουμε την ισχυρή συνέπεια του f_n, που δίνονται από τις (2.10) και (2.11), με την μετρική Hellinger [Θεώρημα 3.2(α)], τις L_p- νόρμες, $p=1,2,+\infty$ [Θεώρημα 3.3] και της Sobolev νόρμας με τάξη ένα [Θεώρημα 3.4(β)]. Θα δείξουμε την συνέπεια του f_n από εκείνη του εκτιμητή u_n της τετραγωνικής ρίζας της πυκνότητας, μέσω των ανισοτήτων (3.8), (3.9), (3.10) και (3.12) (βλ. Θεωρήματα 3.3 και 3.4). Απ' αυτά τα αποτελέσματα κάποιος θα μπορούσε να πάρει, μ' έναν ευθύ τρόπο, την συνέπεια των αντίστοιχων εκτιμητών

$$\tilde{F}_n^*(t) := \int_0^t f_n(s)ds, \ t \in \mathbb{R}_+, \tag{3.1}$$

και

$$\tilde{r}_n(t) := \frac{f_n(t)}{1-\tilde{F}_n^*}, \ t \in \mathbb{R}_+, \tag{3.2}$$

της συνάρτησης κατανομής $F(\cdot)$ και του ρυθμού κινδύνου $r(\cdot)$, αντίστοιχα.

Το αποτέλεσμα – κλειδί για τις αποδείξεις της συνέπειας είναι το Θεώρημα 3.1, για το οποίο πρέπει να δείξουμε ότι

$$\int \log\left(\frac{u_n^2}{f}\right) d(\tilde{F}_n - F) \to 0 \ \text{σ.β.}$$

καθώς το $n \to \infty$. Αυτό αποδεικνύεται στις Προτάσεις 3.1 και 3.2 και η δική τους απόδειξη στηρίζεται στο Λήμμα 3.1, το οποίο έχει ανεξάρτητο ενδιαφέρον και το οποίο παραθέτουμε παρακάτω χωρίς απόδειξη – για την απόδειξη βλέπε το άρθρο των Gijbels και Κλωνιά (1991).

Σ' αυτό το σημείο θα θέλαμε να διατυπώσουμε τις υποθέσεις μας:

A1. $p := P(X \le Y) \ne 0.$

A2. $E|X|^\tau < +\infty$ για $\tau > \gamma^{-1}$, όπου το $\gamma > 0$ ικανοποιεί τη σχέση $0 < (1-\gamma)p < \dfrac{1}{2}.$

A3. $\displaystyle\int \left(\frac{f'}{f}\right)^2 dF < +\infty.$

A4. $\displaystyle\int \left|\frac{f'}{f}\right| F^\beta (1-F)^\gamma < +\infty$ για κάποιο $\beta > 0$ και γ όπως στην A2.

Η πρώτη απ' αυτές τις υποθέσεις – η οποία χρειάζεται στην απόδειξη του Λήμματος 3.1 – φαίνεται απαραίτητη για συμπερασματολογία σύμφωνα με το μοντέλο τυχαίας κοπής δεδομένων. Η τρίτη μας υπόθεση συνεπάγεται ότι $\left(f^{\frac{1}{2}}\right)' \in L_2$ και φαίνεται φυσιολογική για το MPLE πρόβλημα. Όμως, η δεύτερη και η τέταρτη υπόθεση μας φαίνονται τεχνικές και είναι χρήσιμες μόνο για τις Προτάσεις 3.1 και 3.2. Στην πραγματικότητα, η Α4 σχετίζεται με την Α3, και οι δύο τους είναι ισοδύναμες, λαμβάνοντας υπόψη την Α2, στην ειδική περίπτωση όπου $\beta = \gamma = \frac{1}{2}$. Όμως, τυπικά πρέπει το β να είναι κοντά στο μηδέν και $(1-\gamma)p < \frac{1}{2}$.

Για τα παρακάτω, θεωρούμε ότι ο συντελεστής α του συναρτησοειδούς ποινής εξαρτάται από το n και θέτουμε $\frac{\alpha_n}{n} \to 0$ καθώς το $n \to \infty$, ή ισοδύναμα $h_n := \left(\frac{\alpha_n}{\lambda_n}\right)^{\frac{1}{2}} \to 0$ καθώς το $n \to \infty$. Θα χρησιμοποιήσουμε αποκλειστικά την παραμετρικοποίηση h_n, λ_n παρά την α_n, λ_n και θα συμβολίσουμε με ξ το ρυθμό σύγκλισης του h_n στο μηδέν, δηλαδή παίρνουμε

$$h_n = O(n^{-\xi}), \;\; \xi > 0.$$

Μετά αποδεικνύουμε την ισχυρή συνέπεια του $\tilde{F}_n$ στην F με την q-μετρική που ορίζεται παρακάτω.

<u>Λήμμα 3.1</u>

Υπό την υπόθεση Α1, έχουμε ότι

$$\left\| \frac{\tilde{F}_n - F}{F^{\beta}(1-F)^{\gamma}} \right\|_{\infty} = o(n^{-d}) \text{ σ.β.}$$

καθώς το $n \to \infty$, όπου $d < \left(\frac{1}{2} - \beta\right) \wedge (1-\gamma)p$, $\gamma > 0$ και $\beta, (1-\gamma)p \in \left(0, \frac{1}{2}\right)$.

Τώρα, θα χρησιμοποιήσουμε το Λήμμα 3.1 για να δείξουμε τη σύγκλιση των συναρτησοειδών που συνδέονται με τα συναρτησοειδή πληροφορίας Kullback – Leibler, τα οποία εμφανίζονται στο άνω φράγμα για την απόσταση του u_n από το ν στην (3.6).

<u>Πρόταση 3.1</u>

Υπό τις υποθέσεις Α1 και Α2, έχουμε ότι

$$\int_0^{\infty} \log u_n^2 d(\tilde{F}_n - F) = o\left(\left(n^d h_n\right)^{-1}\right) \text{ σ.β.}$$

καθώς το $n \to \infty$, όπου το d είναι όπως το Λήμμα 3.1 και συνεπώς $\xi \in (0, d)$.

<u>Απόδειξη:</u>

Πρώτα παρατηρούμε ότι $u_n(t) = \langle k(\cdot,t;h_n), u_n \rangle$, οπότε

$$\|u_n\|_\infty \le \|k(\cdot,t;h_n)\| \|u_n\| \le (2h_n)^{\frac{-1}{2}} \left(\frac{n}{\lambda_n} \right)^{\frac{1}{2}} \le h_n^{\frac{-1}{2}}, \tag{3.3}$$

όπου η πρώτη ανισότητα έπεται από την ανισότητα Cauchy – Schwarz, για την δεύτερη παρατηρούμε ότι:

$$\begin{aligned}
\|k(\cdot,t;h_n)\|^2 &= \langle k(\cdot,t;h_n), k(\cdot,t;h_n) \rangle = k(t,t;h_n) \\
&= \frac{1}{2h_n} e^{-\frac{1}{h_n}|t-t|} = \frac{1}{2h_n} \\
&\Rightarrow \|k(\cdot,t;h_n)\| = (2h_n)^{-\frac{1}{2}},
\end{aligned}$$

$$\begin{aligned}
\|u_n\|^2 = \langle u_n, u_n \rangle &= \left\langle \frac{1}{\lambda_n} \sum_{i=1}^{n} w_{ni} [u_n(Z_{ni})]^{-1} k(\cdot, Z_{ni}; h_n), u_n \right\rangle \\
&= \frac{1}{\lambda_n} \sum_{i=1}^{n} w_{ni} [u_n(Z_{ni})]^{-1} \langle k(\cdot, Z_{ni}; h_n), u_n \rangle \\
&= \frac{1}{\lambda_n} \sum_{i=1}^{n} w_{ni} [u_n(Z_{ni})]^{-1} u_n(Z_{ni}) = \frac{1}{\lambda_n} \sum_{i=1}^{n} w_{ni} = \frac{n}{\lambda_n} \\
&\Rightarrow \|u_n\| \le \left(\frac{n}{\lambda_n} \right)^{\frac{1}{2}}.
\end{aligned}$$

Για να δείξουμε ότι $\sum_{i=1}^{n} w_{ni} = n$ αρκεί να δείξουμε ότι

$$\sum_{i=1}^{n} \left\{ \left(1 - \frac{i-1}{n}\right)^{\hat{p}_n} - \left(1 - \frac{i}{n}\right)^{\hat{p}_n} \right\} = 1$$

και αυτό αποδεικνύεται ως εξής:

$$\begin{aligned}
\sum_{i=1}^{n} \left\{ \left(1 - \frac{i-1}{n}\right)^{\hat{p}_n} - \left(1 - \frac{i}{n}\right)^{\hat{p}_n} \right\} &= \sum_{i=1}^{n} \left(1 - \frac{i-1}{n}\right)^{\hat{p}_n} - \sum_{i=1}^{n} \left(1 - \frac{i}{n}\right)^{\hat{p}_n} \\
&= \sum_{i=1}^{n} \left(1 - \frac{i-1}{n}\right)^{\hat{p}_n} - \sum_{j=2}^{n+1} \left(1 - \frac{j-1}{n}\right)^{\hat{p}_n} \\
&= \left(1 - \frac{1-1}{n}\right)^{\hat{p}_n} - \left(1 - \frac{n+1-1}{n}\right)^{\hat{p}_n} = 1.
\end{aligned}$$

Για να πάρουμε την τελευταία ανισότητα στην (3.3), παρατηρούμε ότι παραγωγίζοντας την (2.10) παίρνουμε

$$\left| u_n'(t) \right| = \left| \frac{1}{\lambda_n} \sum_{i=1}^{n} \frac{w_{ni}}{u_n(Z_{ni})} \frac{1}{h_n} \operatorname{sgn}(Z_{ni} - t) k(t, Z_{ni}; h_n) \right| \le \frac{1}{h_n} |u_n(t)| \quad \text{σ.β.} \tag{3.4}$$

και επομένως,

$$\frac{n}{\lambda_n} = \|u_n\|^2 = \|u_n\|_2^2 + h_n^2 \left\| u_n' \right\|_2^2 = 1 + h_n^2 \left\| u_n' \right\|_2^2 \le 2,$$

που ισχύει, επειδή

$$\left\|u_n'\right\|_2^2=\int\left(u_n'\right)^2 \underset{(3.15)}{\leq} \frac{1}{h_n^2}\int u_n^2=\frac{1}{h_n^2}\left\|u_n\right\|_2^2=\frac{1}{h_n^2}.$$

Επίσης, παρατηρούμε, από ανωτέρω, ότι $\frac{n}{\lambda_n}\geq 1,$ και άρα

$$n\geq \lambda_n\geq \frac{n}{2}. \tag{3.4'}$$

Τώρα, θα δείξουμε ότι, $u_n(t)\geq \frac{1}{2}h_n^{\frac{-1}{2}}e^{\frac{-2t}{h_n}}$ για $t\geq T:=\max\left\{\left|Z_{ni}\right|,\ i=1,2,...,n\right\}$ και συνεπώς, λογαριθμώντας, θα έχουμε

$$-\log(2h_n^{\frac{1}{2}})-\frac{2t}{h_n}\leq \log u_n(t)\leq -\log h_n^{\frac{1}{2}},\ t\geq T. \tag{3.5}$$

Για την απόδειξη της (3.5), παρατηρούμε κατ' αρχάς ότι από την (3.3), $u_n\left(t\right)\leq\left\|u_n\right\|_\infty\leq h_n^{\frac{-1}{2}}$, $\forall \mathrm{t}\in \mathrm{A}$. Επίσης, για $t\geq T$, έχουμε ότι

$$u_n\left(t\right)=\frac{1}{\lambda_n}\sum_{i=1}^{n}\frac{w_{ni}}{u_n\left(Z_{ni}\right)}\frac{1}{2h_n}e^{\frac{-(t-Z_{ni})}{h_n}}\geq \frac{1}{2\sqrt{h_n}}\frac{1}{\lambda_n}e^{\frac{-2t}{h_n}}\sum_{i=1}^{n}w_{ni}e^{\frac{t+Z_{ni}}{h_n}}\geq$$

$$\geq \frac{1}{2\sqrt{h_n}}e^{\frac{-2t}{h_n}}\frac{1}{\lambda_n}\sum_{i=1}^{n}w_{ni}\cdot 1=\frac{1}{2\sqrt{h_n}}e^{\frac{-2t}{h_n}}\frac{n}{\lambda_n}\geq \frac{1}{2\sqrt{h_n}}e^{\frac{-2t}{h_n}},$$

που δίδουν την (3.5).

Τότε, κάτω απ' την υπόθεση Α2, μετά από μία ολοκλήρωση κατά μέρη, χρησιμοποιώντας την (3.4) και την (3.5) για τον χειρισμό των ορίων στην ολοκλήρωση κατά μέρη, παίρνουμε ότι

$$\left|\int_0^\infty \log u_n d\left(\tilde{F}_n-F\right)\right|=\left|\int_0^\infty\frac{u_n'}{u_n}\left(\tilde{F}_n-F\right)\right|$$

$$\leq h_n^{-1}\int_0^\infty\left|\tilde{F}_n-F\right|\leq h_n^{-1}\left\|\frac{\tilde{F}_n-F}{F^\beta(1-F)^\gamma}\right\|_\infty\int_0^\infty F^\beta(1-F)^\gamma.$$

Τώρα, το συμπέρασμα προκύπτει από το Λήμμα 3.1 και από την παρατήρηση ότι η συνθήκη $\mathrm{E}\left|\mathrm{X}\right|^\tau<+\infty$, για κάποιο $\tau>\rho^{-1},\ \rho\in\left(\frac{1}{2},1\right),$ είναι ικανή ώστε να ισχύει $\int\left\{F\left(1-F\right)\right\}^\rho<+\infty.$ Αυτό φαίνεται αν παρατηρήσουμε ότι η συνθήκη της πεπερασμένης ροπής συνεπάγεται ότι $\left|x\right|^\tau F\left(x\right)\left\{1-F\left(x\right)\right\}\to 0,$ καθώς $\left|x\right|\to\infty$ και συνεπώς, υπάρχουν θετικές σταθερές c, m τέτοιες ώστε $\left|x\right|^\tau F\left(x\right)\left\{1-F\left(x\right)\right\}<c,$ για κάθε $x\in M:=\left\{x\in\mathbb{R}:\left|x\right|>m\right\},$ από το οποίο έπεται ότι

$$\int\left\{F\left(1-F\right)\right\}^\rho\leq\int_{M^c}\left\{F\left(1-F\right)\right\}^\rho+c^\rho\int_M\left|x\right|^{-\tau\rho}dx<+\infty.$$

Η συνεπαγωγή $E\left|X\right|^\tau<+\infty\Rightarrow\left|x\right|^\tau F\left(x\right)\left(1-F\left(x\right)\right)\xrightarrow[\left|x\right|\to\infty]{}0,$ μπορεί να αποδειχθεί ως εξής:

Έστω $X>0$ σ.β. (όμοια στην περίπτωση που είναι αρνητικό)

$$E\left|X\right|^\tau=\int_0^\infty P\left(X^\tau>x\right)dx=\int_0^\infty P\left(X>x^{\frac{1}{\tau}}\right)dx=\int_0^\infty\left[1-F\left(x^{\frac{1}{\tau}}\right)\right]dx=I$$

Κάνουμε την αλλαγή μεταβλητής $y = x^{\frac{1}{\tau}}$ και έχουμε

$$I = \tau \int_0^\infty y^{\tau-1}\left(1 - F\left(y\right)\right)dy \geq \tau \int_0^m y^{\tau-1}\left(1 - F\left(y\right)\right)dy$$

$$= m^\tau\left(1 - F\left(m\right)\right) + \int_0^m y^\tau dF\left(y\right),$$

κάνοντας ολοκλήρωση κατά μέρη και θεωρώντας $m > 0$. Άρα έχουμε ότι

$$E\left|X\right|^\tau \geq m^\tau\left(1 - F\left(m\right)\right) + \int_0^m y^\tau dF\left(y\right)$$

$$\Rightarrow m^\tau\left(1 - F\left(m\right)\right) \leq E\left|X\right|^\tau - \int_0^m y^\tau dF\left(y\right).$$

Οπότε περνώντας στα όρια για $m \to \infty$, παίρνουμε

$$\lim_{m\to\infty} m^\tau\left(1 - F\left(m\right)\right) \leq 0 \Rightarrow \lim_{m\to\infty} m^\tau\left(1 - F\left(m\right)\right) = 0$$

$$\Rightarrow \lim_{m\to\infty} m^\tau F\left(m\right)\left(1 - F\left(m\right)\right) = 0.$$

Ο.Ε.Δ.

■

Το ακόλουθο αποτέλεσμα είναι ανάλογο ενός ισχυρού νόμου, με συντελεστή ρυθμού, για το $\tilde{F}_n$.

Πρόταση 3.2

Υπό τις υποθέσεις A1, A2 και A4, έχουμε ότι

$$\int \log f d\left(\tilde{F}_n - F\right) = o\left(n^{-d}\right) \text{ σ.β.}$$

καθώς το $n \to \infty$, όπου ο ρυθμός d είναι όπως στο Λήμμα 3.1.

Απόδειξη:

Κάτω απ' την υπόθεση A2, $\int f \log f < +\infty$ - βλ. Πρόταση 4.6 των Κλωνιά και Weierman (1985) – και συνεπώς μία ολοκλήρωση κατά μέρη μας δίνει

$$\left|\int_0^\infty \log f d\left(\tilde{F}_n - F\right)\right| = \left|\int_0^\infty \frac{f'}{f}\left(\tilde{F}_n - F\right)\right|$$

$$\leq \left\|\frac{\tilde{F}_n - F}{F^\beta(1-F)^\gamma}\right\|_\infty \int_0^\infty \left|\frac{f'}{f}\right| \left\{F^\beta(1-F)^\gamma\right\}$$

και το συμπέρασμα ακολουθεί λόγω της υποθέσεως A4 και του Λήμματος 3.3. Ο.Ε.Δ.

■

Τώρα, είμαστε σε θέση να πάρουμε το πρώτο αποτέλεσμα συνέπειας για τον u_n, το οποίο μας εξυπηρετεί σαν λήμμα για τα υπόλοιπα αποτελέσματα συνέπειας που θα επακολουθήσουν. Παρατηρούμε ότι η νόρμα που χρησιμοποιούμε, δηλαδή η $\left\|g\right\| = \left(\left\|g\right\|_2^2 + h_n^2\left\|g'\right\|_2^2\right)^{\frac{1}{2}}$, εξαρτάται από το n μέσω του h_n. Σε ό,τι θα επακολουθήσει, θεωρούμε $\nu = f^{\frac{1}{2}}$ - η τετραγωνική ρίζα της συνάρτησης πυκνότητας f.

Θεώρημα 3.1

Υπό τις υποθέσεις A1 έως A4, έχουμε ότι

$$\|u_n - \nu\| = o\left(\left(n^d h_n\right)^{\frac{-1}{2}}\right) + O\left(h_n\right) \text{ σ.β.}$$

καθώς το $n \to \infty,$ με d όπως στο Λήμμα 3.1 και συνεπώς $\xi \in (0, d)$.

Απόδειξη:

Χρησιμοποιώντας τις ιδιότητες ενός αναπαραγόμενου πυρήνα, έχουμε ότι

$$\begin{aligned}
\|u_n - \nu\|^2 &= \|u_n\|^2 + \|\nu\|^2 - 2\langle u_n, \nu\rangle \\
&= n\lambda_n^{-1} + 1 + h_n^2 \|\nu'\|_2^2 - 2n\lambda_n^{-1} \int \nu u_n^{-1} d\tilde{F}_n \\
&\le 1 - n\lambda_n^{-1} + h_n^2 \|\nu'\|_2^2 - 2n\lambda_n^{-1} \int \left(\nu u_n^{-1} - 1\right) d\tilde{F}_n \\
&\le h_n^2 \|\nu'\|_2^2 + n\lambda_n^{-1}\left(\int \log u_n^2 d\tilde{F}_n - \int \log \nu^2 d\tilde{F}_n\right) \qquad (3.6) \\
&= h_n^2 \|\nu'\|_2^2 + n\lambda_n^{-1} \int \log\left(\frac{u_n}{\nu}\right)^2 d\left(\tilde{F}_n - F\right) + n\lambda_n^{-1}\left(\int \log u_n^2 dF - \int \log \nu^2 dF\right) \\
&\le h_n^2 \|\nu'\|_2^2 + n\lambda_n^{-1}\left(\int \log u_n^2 d\left(\tilde{F}_n - F\right) - \int \log f d\left(\tilde{F}_n - F\right)\right).
\end{aligned}$$

Για να περάσουμε από την τρίτη στην τέταρτη σειρά των υπολογισμών χρησιμοποιούμε την ανισότητα $\log(1+x) \le x,$ για $x > -1$ ως εξής:

$$\int \log \frac{\nu^2}{u_n^2} d\tilde{F}_n = 2\int \log\left(1 + \left(\frac{\nu}{u_n} - 1\right)\right) d\tilde{F}_n \le 2\int \left(\frac{\nu}{u_n} - 1\right) d\tilde{F}_n.$$

Η τελευταία ανισότητα είναι συνέπεια του γεγονότος ότι η παράσταση $\int \log g dF$ μεγιστοποιείται για $g = f = F'.$ Τώρα, το αποτέλεσμα ακολουθεί από τις Προτάσεις 3.1 και 3.2 και από την παρακάτω ανισότητα:

$$n\lambda_n^{-1} = \|u_n\|^2 = 1 + h_n^2 \left\|u_n'\right\|_2^2 \le 2, \qquad (3.7)$$

η οποία είναι συνέπεια της ανισότητας (3.4), διότι

$$\begin{aligned}
\|u_n - \nu\|^2 &= \mathrm{O}\left(h_n^2\right) + o\left(h_n^2\right)\left(o\left(\left(n^d h_n\right)^{-1}\right) + o\left(n^{-d}\right)\right) \\
&= \mathrm{O}\left(h_n^2\right) + o\left(n^{-d} h_n^{-1}\right) \text{ σ.β.}
\end{aligned}$$

Ο.Ε.Δ.

■

Θεώρημα 3.2

Υπό τις υποθέσεις του Θεωρήματος 3.1, έχουμε ότι

(α) $\|u_n - \nu\|_2 = o\left(\left(n^d h_n\right)^{\frac{-1}{2}}\right) + O\left(h_n\right)$ σ.β.,

(β) $\|u_n - \nu\|_\infty = o\left(\left(n^d h_n^2\right)^{\frac{-1}{2}}\right) + O\left(h_n^{\frac{1}{2}}\right)$ σ.β.

καθώς το $n \to \infty$, με d όπως στο Λήμμα 3.1. Τότε, $\xi \in (0, d)$ για τη σύγκλιση του μέρους (α) και $\xi \in \left(0, \frac{d}{2}\right)$ για εκείνη του μέρους (β).

Απόδειξη:

Επειδή $\|\cdot\|_2 \le \|\cdot\|$, το μέρος (α) είναι άμεση συνέπεια του Θεωρήματος 3.1. Για το μέρος (β), παρατηρούμε ότι από την ανισότητα Cauchy – Schwarz, παίρνουμε

$$|u_n(t) - v(t)| = |\langle k(\cdot, t; h_n), u_n - v\rangle| \le$$

$$\le \|k(\cdot, t; h_n)\| \|u_n - v\| \le (2h_n)^{-\frac{1}{2}} \|u_n - v\|,$$

όπως δείξαμε και στην απόδειξη της Πρότασης 3.1.
Δηλαδή

$$\|u_n - v\|_\infty \le (2h_n)^{\frac{-1}{2}} \|u_n - v\|,$$

το οποίο, σύμφωνα με το Θεώρημα 3.1, μας δίνει το συμπέρασμα. Ο.Ε.Δ.

■

Μετά, παρατηρούμε ότι η συνέπεια του $f_n = u_n^2$ μπορεί να εξαχθεί από εκείνη του u_n μέσω των ακόλουθων ανισοτήτων,

$$\|f_n - f\|_1 \le 2\|u_n - v\|_2, \tag{3.8}$$

$$\|f_n - f\|_2 \le 2\|v'\|_2^{\frac{1}{2}} \|u_n - v\|_2 + \|u_n - v\|_\infty \|u_n - v\|_2, \tag{3.9}$$

$$\|f_n - f\|_\infty \le 2\|v'\|_2^{\frac{1}{2}} \|u_n - v\|_\infty + \|u_n - v\|_\infty^2, \tag{3.10}$$

τις οποίες αποδεικνύουμε αμέσως τώρα.
Για την (3.8) παρατηρούμε ότι

$$\|f_n - f\|_1 = \int (u_n + v)(u_n - v) \le \|u_n + v\|_2 \|u_n - v\|_2$$

$$\le \left(\|u_n\|_2^2 + \|v\|_2^2 + 2\int u_n v\right)^{\frac{1}{2}} \|u_n - v\|_2 \le 2\|u_n - v\|_2.$$

Για την (3.9):

$$\begin{aligned}
\|f_n - f\|_2^2 &= \int (u_n + v)^2 (u_n - v)^2 \le \|u_n + v\|_\infty^2 \|u_n - v\|_2^2 \\
&= \|(u_n - v) + 2v\|_\infty^2 \|u_n - v\|_2^2 \\
&\le \|u_n - v\|_\infty^2 \|u_n - v\|_2^2 + 4\|v\|_\infty^2 \|u_n - v\|_2^2 \\
&\le \|u_n - v\|_\infty^2 \|u_n - v\|_2^2 + 4\|v'\|_2 \|u_n - v\|_2^2 \\
&\le \left\{\|u_n - v\|_\infty \|u_n - v\|_2 + 2\|v'\|_2^{\frac{1}{2}} \|u_n - v\|_2\right\}^2
\end{aligned}$$

$$\Rightarrow \|f_n - f\|_2 \le \|u_n - v\|_\infty \|u_n - v\|_2 + 2\|v'\|_2^{\frac{1}{2}} \|u_n - v\|_2.$$

Για την (3.10):

$$
\begin{aligned}
\left|f_n - f\right| &= \left|u_n - v\right|\left|u_n + v\right| = \left|u_n - v\right|\left|\left(u_n - v\right) + 2v\right| \\
&\le 2\left|u_n - v\right|\left|v\right| + \left|u_n - v\right|^2 \\
&\le 2\left\|u_n - v\right\|_\infty \left\|v'\right\|_2^{\frac{1}{2}} + \left\|u_n - v\right\|_\infty^2 \\
\Rightarrow \left\|f_n - f\right\|_\infty &\le 2\left\|u_n - v\right\|_\infty \left\|v'\right\|_2^{\frac{1}{2}} + \left\|u_n - v\right\|_\infty^2 .
\end{aligned}
$$

Συγκεντρωτικά, τα αποτελέσματα συνέπειας για τον f_n βρίσκονται στο παρακάτω θεώρημα.

<u>Θεώρημα 3.3</u>

Υπό τις υποθέσεις Α1 έως Α4, έχουμε ότι

(α) $\left\|f_n - f\right\|_1 = o\left(\left(n^d h_n\right)^{\frac{-1}{2}}\right) + O\left(h_n\right)$ σ.β.,

(β) $\left\|f_n - f\right\|_2 = o\left(\left(n^d h_n\right)^{\frac{-1}{2}}\right) + o\left(\left(n^{2d} h_n^3\right)^{\frac{-1}{2}}\right) + O\left(h_n\right)$ σ.β.,

(γ) $\left\|f_n - f\right\|_\infty = o\left(\left(n^d h_n^2\right)^{\frac{-1}{2}}\right) + O\left(h_n^{\frac{1}{2}}\right)$ σ.β.

καθώς το $n \to \infty$, με d όπως στο Λήμμα 3.1. Τότε, $\xi \in (0, d)$ για τη σύγκλιση του μέρους (α), $\xi \in \left(0, \frac{2d}{3}\right)$ για τη σύγκλισή του μέρους (β) και $\xi \in \left(0, \frac{d}{2}\right)$ για εκείνη του μέρους (γ).

<u>Απόδειξη:</u>
<u>(α):</u>
Είναι άμεσο επακόλουθο της ανισότητας (3.8) και του Θεωρήματος 3.2(α).
<u>(β):</u>
Χρησιμοποιώντας την ανισότητα (3.9) και το Θεώρημα 3.2 παίρνουμε

$$
\begin{aligned}
\left\|f_n - f\right\|_2 &= o\left(\left(n^d h_n\right)^{\frac{-1}{2}}\right) + O\left(h_n\right) + \left[o\left(\left(n^d h_n^2\right)^{\frac{-1}{2}}\right) + O\left(h_n^{\frac{1}{2}}\right)\right]\left[o\left(\left(n^d h_n\right)^{\frac{-1}{2}}\right) + O\left(h_n\right)\right] \\
&= o\left(\left(n^d h_n\right)^{\frac{-1}{2}}\right) + O\left(h_n\right) + o\left(\left(n^{2d} h_n^3\right)^{\frac{-1}{2}}\right) \text{ σ.β.}
\end{aligned}
$$

<u>(γ):</u>
Χρησιμοποιώντας την ανισότητα (3.10) και το Θεώρημα 3.2(β) παίρνουμε

$$
\begin{aligned}
\left\|f_n - f\right\|_\infty &= o\left(\left(n^d h_n^2\right)^{\frac{-1}{2}}\right) + O\left(h_n^{\frac{1}{2}}\right) + o\left(\left(n^d h_n^2\right)^{-1}\right) + O\left(h_n\right) \\
&= o\left(\left(n^d h_n^2\right)^{\frac{-1}{2}}\right) + O\left(h_n^{\frac{1}{2}}\right) \text{ σ.β.}
\end{aligned}
$$

Ο.Ε.Δ.

■

Κάτω από την επιπρόσθετη υπόθεση $\nu'' \in L_2$, θα δείξουμε την συνέπεια του f_n με την Sobolev νόρμα τάξης ένα.

Θεώρημα 3.4

Υπό τις υποθέσεις Α1 έως Α4 και υπό την επιπλέον υπόθεση $\|\nu''\|_2 < +\infty$, έχουμε ότι

(α) $\left\|u_n' - \nu'\right\|_2 = o\left(\left(n^d h_n^3\right)^{\frac{-1}{2}}\right) + o\left(\left(n^d h_n\right)^{\frac{-1}{4}}\right) + O\left(h_n^{\frac{1}{2}}\right)$ σ.β.,

(β) $\left\|f_n' - f'\right\|_2 = o\left(\left(n^d h_n^3\right)^{\frac{-1}{2}}\right) + o\left(\left(n^d h_n\right)^{\frac{-1}{4}}\right) + O\left(h_n^{\frac{1}{2}}\right)$ σ.β.

καθώς το $n \to \infty$, με d όπως στο Λήμμα 3.1 και $\xi \in \left(0, \frac{d}{3}\right)$.

Απόδειξη:

(α):

Παρατηρούμε ότι, αφού ο u_n είναι λύση του προβλήματος βελτιστοποίησης (2.12), έχουμε ότι

$$n\int \log u_n^2 d\tilde{F}_n - \lambda_n \|u_n\|^2 \geq n\int \log \nu d\tilde{F}_n - \lambda_n \|\nu\|^2$$

και συνεπώς,

$$\begin{aligned} \left\|u_n'\right\|_2^2 - \|\nu'\|_2^2 &\leq n\lambda_n^{-1} h_n^{-2} \int \log\left(\frac{u_n}{\nu}\right)^2 d\tilde{F}_n \\ &\leq 2h_n^{-2} \int \log\left(\frac{u_n}{\nu}\right)^2 d\left(\tilde{F}_n - F\right), \end{aligned} \tag{3.11}$$

όπου η τελευταία ανισότητα είναι συνέπεια της (3.7) και της ανισότητας $\int \log u_n^2 dF \leq \int \log \nu^2 dF$. Τότε,

$$\begin{aligned} \left\|u_n' - \nu'\right\|_2^2 &= \left\|u_n'\right\|_2^2 + \|\nu'\|_2^2 - 2\int u_n'\nu' \\ &= \left\|u_n'\right\|_2^2 - \|\nu'\|_2^2 - 2\int \nu'\left(u_n' - \nu'\right) \\ &= \left\|u_n'\right\|_2^2 - \|\nu'\|_2^2 + 2\int \nu''\left(u_n - \nu\right) \\ &\leq 2h_n^{-2} \int \log\left(\frac{u_n}{\nu}\right)^2 d\left(\tilde{F}_n - F\right) + 2\|\nu''\|_2 \|u_n - \nu\|_2, \end{aligned}$$

όπου στο προτελευταίο βήμα κάναμε ολοκλήρωση κατά μέρη ενώ το τελευταίο βήμα οφείλεται στη σχέση (3.11) και την ανισότητα Cauchy – Schwarz. Τώρα, το συμπέρασμα προκύπτει από τις Προτάσεις 3.1, 3.2 και το Θεώρημα 3.2(α), καθώς έχουμε

$$
\begin{aligned}
\left\| u_n{}' - v' \right\|_2 &= \left(2h_n^{-2}\right)^{\frac{1}{2}} \left[o\left(\left(n^d h_n\right)^{\frac{-1}{2}} \right) + o\left(n^{\frac{-d}{2}} \right) \right] + o\left(\left(n^d h_n\right)^{\frac{-1}{4}} \right) + O\left(h_n^{\frac{1}{2}} \right) \\
&= \left[o\left(\left(n^d h_n^3\right)^{\frac{-1}{2}} \right) + o\left(n^{\frac{-d}{2}} \right) \right] + o\left(\left(n^d h_n\right)^{\frac{-1}{4}} \right) + O\left(h_n^{\frac{1}{2}} \right) \qquad \text{σ.β.} \\
&= o\left(\left(n^d h_n^3\right)^{\frac{-1}{2}} \right) + o\left(\left(n^d h_n\right)^{\frac{-1}{4}} \right) + O\left(h_n^{\frac{1}{2}} \right)
\end{aligned}
$$

<u>(β):</u>

Παρατηρούμε ότι, από την τριγωνική ανισότητα, έχουμε ότι

$$
\begin{aligned}
\frac{1}{2}\left\| f_n{}' - f' \right\|_2 &= \left\| u_n{}' u_n - v'v \right\|_2 = \left\| u_n{}' \left(u_n - v\right) + \left(u_n{}' - v'\right) v \right\|_2 \le \\
&\le \left\| u_n{}' \left(u_n - v\right) \right\|_2 + \left\| v \left(u_n{}' - v'\right) \right\|_2 \le \\
&\le \left\| u_n - v \right\|_\infty \left\| u_n{}' \right\|_2 + \left\| v \right\|_\infty \left\| u_n{}' - v' \right\|_2 \le \\
&\le \left\| u_n - v \right\|_\infty \left\| u_n{}' \right\|_2 + \left\| v' \right\|_2^{\frac{1}{2}} \left\| u_n{}' - v' \right\|_2 ,
\end{aligned}
\tag{3.12}
$$

όπου η τελευταία ανισότητα έπεται από το ότι $\left\| v \right\|_\infty^2 \le \left\| v \right\|_2 \left\| v' \right\|_2$ - βλέπε (4.1) του Κλωνιά (1982). Μετά παρατηρούμε ότι , από την (3.11) και τις Προτάσεις 3.1 και 3.2, έχουμε ότι

$$
\left\| u_n{}' \right\|_2^2 \le \left\| v' \right\|_2^2 + o\left(\left(n^d h_n^3\right)^{-1} \right) + o\left(\left(n^d h_n^2\right)^{-1} \right) \text{ σ.β.} \tag{3.13}
$$

καθώς το $n \to \infty$. Οπότε,

$$
\begin{aligned}
\left\| f_n{}' - f' \right\|_2 &= o\left(\left(n^d h_n^2\right)^{\frac{-1}{2}} \right) + O\left(h_n^{\frac{1}{2}} \right) + o\left(\left(n^d h_n^3\right)^{\frac{-1}{2}} \right) + o\left(\left(n^d h_n\right)^{\frac{-1}{4}} \right) + O\left(h_n^{\frac{1}{2}} \right) = \\
&= o\left(\left(n^d h_n^2\right)^{\frac{-1}{2}} \right) + O\left(h_n^{\frac{1}{2}} \right) \text{ σ.β.}
\end{aligned}
$$

Ο.Ε.Δ.

■

4. ΑΛΓΟΡΙΘΜΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ

[**ΠΡΟΣΟΧΗ:** Οι αλγόριθμοι που ακολουθούν είναι πιθανόν να έχουν κάποια προγραμματιστικά λάθη, καθώς τα αποτελέσματα που δίνουν δεν είναι ικανοποιητικά.]

Σ' αυτήν την ενότητα παραθέτουμε δύο αλγορίθμους για την ελαχιστοποίηση του συναρτησοειδούς

$$Q(q) = -2\sum_{i=1}^{n}\log q_i + q^T Cq,$$

και την σχεδίαση του γραφήματος του εκτιμητή f_n.

Κατ' αρχήν υποθέτουμε ότι τα $X_1, X_2, ..., X_n$ ακολουθούν την εκθετική κατανομή με παράμετρο 1 και ότι το c στη σχέση (1.2) ισούται με 2, επομένως τα $Y_1, Y_2, ..., Y_n$ ακολουθούν εκθετική κατανομή με παράμετρο 2. Η γλώσσα προγραμματισμού που χρησιμοποιήθηκε είναι η MATLAB (version 7.0.1).

- 1η Περίπτωση: $A = \mathbb{R}$

  Στην σχέση (2.8), θεωρούμε $\kappa(t) = \varphi(t)$, όπου φ είναι η πυκνότητα πιθανότητας της τυπικής κανονικής κατανομής. Επίσης, παίρνουμε $h = \sqrt{\beta} = \dfrac{1}{80}$.

  Τότε, ο αλγόριθμος που παίρνουμε είναι ο εξής:

```
global C
global lambda
global w
global Zn
%Parakatw ypologizontai ta deigmata X,Y,Z kai d:
x1=rand(100,1);
y1=rand(100,1);
for i=1:100
    X(i)=icdf('exp',x1(i),1);
    Y(i)=icdf('exp',y1(i),2);
    Z(i)=min(X(i),Y(i));
    if X(i)<=Y(i)
        d(i)=1;
    else
        d(i)=0;
    end
end
%Taksinomoume to deigma Z sto Zn kai meta ypologizoume to p
%to opoio einai o mesos twn d:
Zn=sort(Z);
p=mean(d);
%Orizw to Wn:
for i=1:100
    w(i)=100*(((1-(i-1)/100)^p)-((1-i/100)^p));
end
%Orizoume ton pinaka C xrhsimopoiwntas ton pyrhna:
```

```
for i=1:100
    for j=1:100
        C(i,j)=80*normpdf(80*(Zn(i)-Zn(j)));
    end
end
%Orizoume th synarthsh pou 8eloume na elaxistopoihsoume:
Q = @(q)(-2)*sum(log(q))+q'*(C*q);
% To elaxisto ths parastashs Q to pairnoume sthn parakatw 8esh qmin:
qmin=fminsearch(Q,ones(100,1))
%Ypologizoume to lambda:
t=-10:0.1:10;
for i=1:100
    for j=1:100
        olokl=trapz(t,80.*normpdf(80*(t-Zn(i))).*80.*normpdf(80*(t-
Zn(j))));
        A(i,j)=sqrt(w(i)*w(j))*qmin(i)*qmin(j)*olokl;
    end
end
lambda=sum(sum(A))
%Pairnoume ta grafhmata tou ektimhth fn kai ths exponential(1)
parallhla:
fplot('fN2',[-10,10],'green')
grid on
hold on
fplot('exppdf',[-10,10],'red')
```

Ακόμα, για τον ορισμό της συνάρτησης του εκτιμητή κατασκευάσαμε την παρακάτω συνάρτηση στην MATLAB:

```
function Est=fN2(t)
global lambda
global w
global qmin
global Zn
for i=1:100
    B(i)=(sqrt(w(i))*80*normpdf(80*(t-Zn(i))))^2;
end
Est=sum(B);
Est=Est/lambda;
```

- <u>2η Περίπτωση:</u> $A=\mathbb{R}_+$

Όμοια, σ' αυτήν την περίπτωση, στη σχέση (2.9) παίρνουμε το ίδιο $\kappa(t)$ και για h παίρνουμε το 1/20. Τότε, ο αλγόριθμος γίνεται:

```
global C
global lambda
global w
global qmin
global Zn
x1=rand(100,1);
y1=rand(100,1);
for i=1:100
    X(i)=icdf('exp',x1(i),1);
```

```
    Y(i)=icdf('exp',y1(i),2);
    Z(i)=min(X(i),Y(i));
    if X(i)<=Y(i)
        d(i)=1;
    else
        d(i)=0;
    end
end
Zn=sort(Z);
p=mean(d);
for i=1:100
    w(i)=100*(((1-(i-1)/100)^p)-((1-i/100)^p));
end
for i=1:100
    for j=1:100
        C(i,j)=(1/2)*20*normpdf(20*(Zn(i)-
Zn(j)))+(1/2)*20*normpdf(20*(Zn(i)+Zn(j)));
    end
end
Q=@(q)(-2)*sum(log(q))+q'*(C*q);
% To elaxisto ths parastashs Q to pairnoume sthn parakatw 8esh qmin:
qmin=fminsearch(Q,ones(100,1))
t=0:0.1:10;
for i=1:100
    for j=1:100
        s1=Zn(i);
        s2=Zn(j);
        olokl=trapz(t,(20*20/4)*(normpdf(20*(t-
s1))+normpdf(20*(t+s1))).*(normpdf(20*(t-s2))+normpdf(20*(t+s2))));
        A(i,j)=sqrt(w(i)*w(j))*qmin(i)*qmin(j)*olokl;
    end
end
lambda=sum(sum(A))
fplot('fN3',[0,10],'green')
grid on
hold on
fplot('exppdf',[-10,10],'red')
```

και η επιπρόσθετη συνάρτηση για τον ορισμό του εκτιμητή είναι η εξής:

```
function Est=fN3(t)
global lambda
global w
global qmin
global Zn
for i=1:100
    B(i)=(sqrt(w(i))*qmin(i)*(1/2)*20*(normpdf(20*(t-
Zn(i)))+20*normpdf(20*(t+Zn(i)))))^2;
end
Est=sum(B)/lambda;
```

Βιβλιογραφία

1. Abdushukurov, A.A., (1984). On some estimates of the distribution function under random censorship (in Russian). Conference of Young Scientists, VINITY No. 8756- V. Math. Inst. Acad. Sci. Uzbek SSR, Tashkent.
2. Cheng, P.E. and Lin, G.D. (1984). Maximum likelihood estimation of a survival function under the Koziol – Green proportional hazards model. Technical Report B-84-5, Institute of Statistics, Academia Sinica, Taipei, Taiwan.
3. Cheng, P.E. and Lin, G.D. (1987). Maximum likelihood estimation of a survival function under the Koziol – Green proportional hazards model. Statist. Probab. Lett., 5, 75-80.
4. Csorgo, S. (1988). Estimation in the proportional hazards model of random censorship. Statistics, 19, 437-463.
5. Csorgo, S. and Horvath, L. (1981). On the Koziol – Green model of random censorship. Biometrika, 68, 391-401.
6. Csorgo, S. and Mielniczuk, J. (1988). Density estimation in the simple proportional hazards model. Statist. Probab. Lett., 6, 419-426.
7. De Montricher, G.F., Tapia, R.A. and Thompson, J.R. (1975). Nonparametric maximum likelihood estimation of probability densities by penalty function methods. Ann. Statist., 3, 1329-1348.
8. Ghorai, J.K. and Rejto, L. (1987). Estimation of mean residual life with censored data under proportional hazards model. Comm. Statist. A- Theory Methods, 16, 2097-2014.
9. Good, I.J. and Gaskins, R.A. (1971). Nonparametric roughness penalties for probability densities. Biometrika, 58, 255-277.
10. Good, I.J. and Gaskins, R.A. (1980). Density estimation and bump-hunting by the penalized likelihood method exemplified by scattering and meteorite data (invited paper). J. Amer. Statist. Assoc., 75, 42-73.
11. Kaplan, E.L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. J. Amer. Statist. Assoc., 53, 457-481.
12. Klonias, V.K. (1982). Consistency of two nonparametric maximum penalized likelihood estimators of the probability density function. Ann. Statist., 10, 811-824.
13. Klonias, V.K. (1984). On a class of nonparametric density and regression estimators. Ann. Statist.,12, 1263-1284.
14. Klonias, V.K. and Gijbels, I. (1991). Density estimation under the Koziol – Green model of censoring by penalized likelihood methods. The Canadian Journal of Statistics Vol.19, No.1, pages 23-28.
15. Klonias, V.K. and Nash, S.G. (1987). Some numerical techniques in nonparametric estimation. J. Statist. Comput. Simulation, 28, 97-126.
16. Klonias, V.K. and Wierman, J.C. (1985). On density estimation from censored data by penalized likelihood methods. John Hopkins University, Department of Mathematical Sciences. Technical Report No. 418.
17. Koziol, J.A. and Green, S.B. (1976). A Cramer – Von Mises statistic for randomly censored data. Biometrika, 63, 465-474.

18. Lubecke, A.M. and Padgett, W.J. (1985a). Nonparametric maximum penalized likelihood estimation of a density from arbitrarily right censored observations. Comm. Statist. A- Theory Methods, 14, 257-271.
19. Lubecke, A.M. and Padgett, W.J. (1985b). Correction to "Nonparametric maximum penalized likelihood estimation of a density from arbitrarily right censored observations." Comm. Statist. A- Theory Methods, 14, 2007.
20. Silverman, B.W., (1982). On the estimation of probability density function by the maximum penalized likelihood method. Ann. Statist., 10, 795-810.
21. Tapia, R.A. and Thompson, J.R.,(1978). Nonparametric Probability Density Estimation. John Hopkins University Press, Baltimore.