Η ΠΡΩΤΗ MPLE ΤΩΝ GOOD ΚΑΙ GASKINS

ΜΑΝΩΛΙΑ ΚΟΚΟΛΟΓΙΑΝΝΑΚΗ
Μαθηματικό Τμήμα, Τ.Θ 2208, Πανεπιστήμιο Κρήτης
71003 Ηράκλειο-Κρήτη, ΕΛΛΑΣ.

16 Δεκεμβρίου, 1998

Περιεχόμενα

Περίληψη

Στην εργασία αυτή υπολογίζουμε την εκτιμήτρια της άγνωστης συνάρτησης πυκνότητας f, μεγιστοποιώντας το συναρτησοειδές του λογαρίθμου της πιθανοφάνειας μείον ένα όρο ποινής. Στην περίπτωση που ο όρος ποινής είναι η L_2 νόρμα της πρώτης παραγώγου της $f^{1/2}$ στον RKHS $H^1(-\infty, +\infty)$, αποδεικνύουμε ότι η MPL εκτιμήτρια που προκύπτει υπάρχει, είναι μοναδική και είναι συνεπής ως πρός την απόσταση Hellinger. Η DMPL εκτιμήτρια που προκύπτει από μια διακριτή προσέγγιση του παραπάνω όρου ποινής, βλέπουμε ότι υπάρχει μονοσήμαντα και είναι συνεπής στο χώρο των ιστογραμμάτων. Τέλος, παρουσιάζουμε μια σειρά από αριθμητικά αποτελέσματα που σχετίζονται με τον υπολογισμό της DMPLE στο χώρο των ιστοπολυγώνων με όρο ποινής τη διακριτή δεύτερη παράγωγο.

Κεφάλαιο 1

Εισαγωγή

Εστω $X_1, X_2, ..., X_n$ ανεξάρτητες και ισόνομες παρατηρήσεις, από μια κατανομή με πυκνότητα πιθανότητας f. Ενα από τα βασικά προβλήματα με τα οποία ασχολείται ο τομέας της Στατιστικής, είναι η εκτίμηση της άγνωστης συνάρτησης f. Αν με $\hat{f}$ συμβολίσουμε την εκτιμήτρια της συνάρτησης πυκνότητας f, τότε η εκτιμήτρια της συνάρτησης κατανομής F του τυχαίου δείγματος, θα είναι $\hat{F}(x) = \int_{-\infty}^{x} \hat{f}(z)dz$. Η ℓ δεν είναι φραγμένη στο σύνολο των πυκνοτήτων και μια απευθείας μεγιστοποίηση του συναρτησοειδούς της πιθανοφάνειας:

$$\ell(f) = \prod_{i=1}^{n} f(X_i),$$

στο σύνολο $\{f \ : \ f \geq 0, \ \int f = 1\}$, οδηγεί στην "λύση Dirac", $\frac{1}{n}\sum_{i=1}^{n} \delta(x - X_i)$.

Για να αποφύγουν την "λύση Dirac" οι Good και Gaskins το 1971 [1], ενσωμάτωσαν στο συναρτησοειδές της πιθανοφάνειας ένα όρο που περιγράφει το μη ομαλόν της εκτιμήτριας. Εστω λοιπόν $\Phi(f)$ ένα συναρτησοειδές ποινής, όπως ονομάστηκε, το οποίο ποσοτικοποιεί την μεταβλητότητα της εκτιμήτριας και απειρίζεται στην "λύση Dirac". Οι Good και Gaskins πρότειναν τα ακόλουθα συναρτησοειδή ποινής:

$$\Phi(f) = \int_{-\infty}^{+\infty} (f^{(k)})^2$$

για $k = 1$ ή 2. Στα πλαίσια της παραπάνω ιδέας τους οι Good και Gaskins εισήγαγαν την Maximum Penalized Likelihood Μέθοδο για την εκτίμηση της άγνωστης συνάρτησης πυκνότητας. Με βάση τη μεθόδο αυτή, μεγιστοποιούμε το ποινικοποιημένο συναρτησοειδές του λογαρίθμου της πιθανοφάνειας, που ορίζεται :

$$L(f) = \sum_{i=1}^{n} \log f(X_i) - \alpha\Phi(f), \tag{1.1}$$

όπου α είναι μια θετική παράμετρος, η οποία καθορίζει το πόσο ομαλή θα είναι η εκτιμήτρια μας. Θεωρούμε δηλαδή σαν εκτιμήτρια της άγνωστης συνάρτησης πυκνότητας f, το μέγιστο του συναρτησοειδούς του λογαρίθμου της πιθανοφάνειας μείον ένα όρο ποινής και όχι αυτού καθ'εαυτού του συναρτησοειδούς του λογαρίθμου της πιθανοφάνειας. Το συναρτησοειδές L ποσοτικοποιεί την αντιπαλότητα μεταξύ της ομαλότητας της εκτιμήτριας και της προσαρμογής της στα δεδομένα, μια που ο όρος $\sum_{i=1}^{n} \log f(X_i)$, ρυθμίζει το πόσο καλά προσεγγίζει η συνάρτηση f τα δεδομένα.

Πρίν από τους Good και Gaskins, ο τρόπος που προσεγγιζόταν η "λύση Dirac", ήταν με την προσέγγιση της κατανομής $\delta(x - X_i)$ από μια συνάρτηση της μορφής $\frac{1}{h}K\left(\frac{x-X_i}{h}\right)$, όπου K ήταν μια τυχαία συνάρτηση πυκνότητας στο $\mathbf{R}$ και h μια παράμετρος συνδεδεμένη με το n, καθώς αυτό τείνει στο άπειρο. Αυτή η προσέγγιση της "λύσης Dirac", δηλαδή η:

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{h}K\left(\frac{x-X_i}{h}\right) = \int_{-\infty}^{+\infty}\frac{1}{h}K\left(\frac{x-y}{h}\right)dF_n(y),$$

δίνει τις κλασικές εκτιμήτριες πυκνότητας, τις εκτιμήτριες των πυρήνων, οι οποίες βέβαια εμφανίστηκαν αρχικά με άλλο διαισθητικό υπόβαθρο.

Σαν Maximum Penalized Likelihood Εκτιμήτρια (MPLE) ορίζεται κάθε λύση του ακόλουθου προβλήματος βελτιστοποίησης:

$$\max_f\left\{\prod_{i=1}^{n} f(X_i)\exp(-\alpha\Phi(f))\right\} \Leftrightarrow \max_f\left\{n\int_{-\infty}^{+\infty}\log f(x)dF_n(x) - \alpha\Phi(f)\right\} \tag{1.2}$$

υπό τους περιορισμούς:

$$f \geq 0, \quad \int f = 1. \tag{1.3}$$

Είναι φανερό ότι η εκτιμήτρια που παίρνουμε με την MPL μέθοδο θα είναι εξ'ορισμού μια συνάρτηση πυκνότητας. Η παράμετρος α, ρυθμίζει το πόσο ομαλή θα είναι η εκτιμήτρια μας μια και καθορίζει, την υπεροχή της ομαλότητας της εκτιμήτριας ή της καλής προσέγγισης της f στα δεδομένα. Οσο μικρότερη είναι η τιμή της παραμέτρου, τόσο ανώμαλη θα είναι η αντίστοιχη MPLE. Εν συνεχεία, για λόγους απλότητας ενσωματώνουμε την παράμετρο α, στο συναρτησοειδές $\Phi(f)$.

Οι Scott, Tapia και Thompson το 1980 [2] (βλέπε επίσης [3]), πρότειναν μια διακριτοποίηση του ανωτέρω προβλήματος που κάνει την διάσταση του πεπερασμένη και εισάγει διακριτά συναρτησοειδή ποινής, ανάλογα των Good και Gaskins. Οπως θα δούμε στο κεφάλαιο 3, οι εκτιμήτριες που προκύπτουν είναι συνεπείς. Επιπλέον στο ίδιο κεφάλαιο παρουσιάζουμε μια σειρά από αριθμητικά αποτελέσματα, σχετικά με τον υπολογισμό της MPLE στην περίπτωση που σαν όρο ποινής θεωρήσουμε την διακριτή δεύτερη παράγωγο. Ο Silverman [4], [5], για να αποφύγει την μη γραμμικότητα του συναρτησοειδούς της πιθανοφάνειας, χρησιμοποίησε ως μεταβλητή τον $\log f$ και κατά συνέπεια οδηγήθηκε σε ένα συναρτησοειδές ποινής επί του $\log f$, όπως το ακόλουθο:

$$\Phi(g) = \int_{-\infty}^{+\infty}(g'')^2$$

όπου $g = \log f$. Το 1975 οι De Montricher, Tapia και Thompson [6] (βλέπε επίσης [2]), απέδειξαν τόσο την ύπαρξη όσο και την μοναδικότητα και των δυο MPLE που εισήγαγαν οι Good και Gaskins και γενίκευσαν τα συναρτησοειδή ποινής. Η θεωρία τους αναπτύχθηκε κυρίως στους Reproducing Kernel Hilbert Space $H^1(-\infty,+\infty)$ και $H^2(-\infty,+\infty)$. Ενας χώρος Hilbert $H(\Omega)$ συναρτήσεων ορισμένων στο Ω, ονομάζεται Reproducing Kernel Hilbert Space (RKHS) αν υπάρχει ένα συναρτησοειδές $K(\cdot,\cdot)$ ορισμένο στο $\Omega \times \Omega$ τέτοιο ώστε :

1. $K(\cdot,t) \in H(\Omega), t \in \Omega$ και

2. $f(t) = \langle f(\cdot), K(\cdot,t)\rangle_H$, για κάθε $f \in H(\Omega)$ και $t \in \Omega$.

Για $s = 1, 2, ...$ ο χώρος

$$H^s(-\infty,+\infty) = \left\{ f \quad : \quad f^{(j)} \in L^2(-\infty,+\infty), \quad j = 0, ..., s \right\}$$

με εσωτερικό γινόμενο

$$\langle f, g\rangle_{H^s} = \sum_{j=0}^{s} w_j \left\langle f^{(j)}, g^{(j)} \right\rangle_{L^2(-\infty,+\infty)}, \quad w_j \geq 0, \quad w_0, w_s > 0$$

είναι ένας RKHS (βλέπε [2]). Στη συνέχεια θα συμβολίζουμε την $L^2(-\infty,+\infty)$ νόρμα με $\|\cdot\|$ για λόγους απλότητας. Τα αποτελέσματα σχετικά με την πρώτη MPLE των Good και Gaskins παρουσιάζονται αναλυτικά στο Κεφαλαίο 2. Το 1982 ο Κλωνιάς [7], απέδειξε την ισχυρή συνέπεια της πρώτης MPLE των Good και Gaskins υπό διάφορες μετρικές, η δε συνέπεια της δεύτερης έπεται από το άρθρο [8] του ιδίου το 1984. Από την δουλειά αυτή παρουσιάζουμε στο κεφάλαιο 4 μόνο την συνέπεια ως πρός την απόσταση Hellinger της συνάρτησης πυκνότητας

Κεφάλαιο 2

Η πρώτη MPLE των Good και Gaskins

Δεδομένου του τυχαίου δείγματος $X_1, X_2, ..., X_n$, από μια συνάρτηση πυκνότητας f, με πεπερασμένη πληροφορία κατά Fisher, στόχος μας στο κεφάλαιο αυτό είναι να μελετήσουμε κυρίως την πρώτη εκτιμήτρια των Good και Gaskins [1] και να αναφέρουμε λίγα πράγματα για τη δεύτερη εκτιμήτρια. Οι εκτιμήτριες αυτές, λαμβάνονται σαν τη λύση του ακόλουθου προβλήματος βελτιστοποίησης:

$$\max_f \left\{ \sum_{i=1}^{n} \log f(X_i) - \Phi(\sqrt{f}) \right\}$$

υπό τους περιορισμούς:

$$f^{1/2} \in H(-\infty, +\infty), \quad \int_{-\infty}^{+\infty} f(t)dt = 1, \quad f(t) \geq 0, \quad \forall t \in \mathbf{R}$$

όπου $H(-\infty, +\infty)$ είναι ένας RKHS και το συναρτησοειδές Φ, δίνεται αντίστοιχα από τις σχέσεις:

$$\Phi(\sqrt{f}) = 4\alpha \left\| \left(\sqrt{f}\right)' \right\|_2^2, \quad \alpha > 0$$

και

$$\Phi(\sqrt{f}) = \alpha \left\| \left(\sqrt{f}\right)' \right\|_2^2 + \quad \beta \left\| \left(\sqrt{f}\right)'' \right\|_2^2, \quad \alpha \geq 0, \quad \beta > 0.$$

Παρατηρούμε ότι και στις δύο περιπτώσεις το συναρτησοειδές Φ του όρου ποινής είναι ένα συναρτησοειδές της τετραγωνικής ρίζας της συνάρτησης πυκνότητας.

Το παραπάνω πρόβλημα βελτιστοποίησης γράφεται ισοδύναμα:

$$\max_f \left\{ n \int_{-\infty}^{+\infty} \log f(x) dF_n(x) - \Phi(\sqrt{f}) \right\}$$

κάτω από τους ίδιους περιορισμούς.

Οι Good και Gaskins έκφρασαν το παραπάνω πρόβλημα σε όρους της τετραγωνικής ρίζας της συνάρτησης πυκνότητας κυρίως για μαθηματική ευκολία και από Bayesian διαίσθηση. Οι De Montricher, Tapia και Thompson [6], [2], ακολούθησαν την τεχνική τους κύρια για τεχνικούς λόγους και συγκεκριμένα, για να αποφύγουν τη μεγιστοποίηση του ποινικοποιημένου συναρτησοειδούς στο θετικό κώνο των συναρτήσεων πυκνότητας. Στην προσπάθεια τους να μελετήσουν τις MPLE, πρότειναν την ακόλουθη εναλλακτική διατύπωση του παραπάνω προβλήματος:

$$\max_u \left\{ n \int_{-\infty}^{+\infty} \log u^2(x) dF_n(x) - \Phi(u) \right\} \tag{2.1}$$

υπό τους περιορισμούς:

$$u \in H(-\infty, +\infty), \quad \int_{-\infty}^{+\infty} u^2(t)dt = 1, \; u(X_i) \geq 0 \; \forall \; i = 1, 2, ..., n,$$

όπου $f^{1/2} = u$.

Στη συνέχεια, βασιζόμενοι στη δουλειά των De Montricher, Tapia και Thompson [6], [2], θα παρουσιάσουμε μια σειρά από αποτελέσματα τα οποία εξασφαλίζουν τόσο την ύπαρξη όσο και τη μοναδικότητα της πρώτης MPLE. Επιπλέον, θα δούμε ότι είναι μια θετική εκθετική spline με κόμβους ακριβώς στα δειγματικά σημεία και ότι σε αυτή την περίπτωση η εναλλακτική διατύπωση των Good και Gaskins δίνει την MPLE. Στην περίπτωση όμως που το συναρτησοειδές Φ είναι γραμμικός συνδυασμός της L_2 νόρμας της πρώτης και δεύτερης παραγώγου του u, η MPLE υπάρχει και είναι μοναδική αλλά η εναλλακτική διατύπωση δίνει μια ψεύδο-MPLE. Να σημειώσουμε ότι εκθετική spline με κόμβους στα σημεία $X_1, X_2, ..., X_n$, ονομάζεται κάθε συνάρτηση της μορφής

$$E(x) = \sum_{i=1}^{n} \left(\alpha_i + \beta_i (x - X_i) + \gamma_i \cosh\left(\kappa_i \left| x - X_i \right|\right) + \delta_i \sinh\left(\kappa_i \left| x - X_i \right|\right) \right), \tag{2.2}$$

όπου $\alpha_i, \beta_i, \gamma_i, \delta_i, \kappa_i$ είναι σταθερές.

Μια τεχνική, για προβλήματα ανάλογα με το δικό μας όπου έχουμε τον περιορισμό η συνάρτηση πυκνότητας να είναι θετική, είναι να δουλεύουμε με ένα πρόβλημα εκφρασμένο σε όρους της τετραγωνικής ρίζας της άγνωστης συνάρτησης πυκνότητας. Με το μετασχηματισμό αυτό, ο περιορισμός η συνάρτηση πυκνότητας να είναι θετική, είναι περιττός. Στη συνέχεια, τα προβλήματα βελτιστοποίησης που αφορούν την πρώτη και δεύτερη MPLE των Good και Gaskins έχουν την ακόλουθη μορφή:

$$\max_f J(f) \tag{2.3}$$

υπό τους περιορισμούς:

$$f^{1/2} \in H(-\infty, +\infty), \int_{-\infty}^{+\infty} f(t)dt = 1, \;\; f(t) \geq 0, \forall t \in \mathbf{R},$$

όπου το συναρτησοειδές J είναι ορισμένο στο χώρο $H(-\infty,+\infty)$. Ακολουθώντας την παραπάνω τεχνική οι Good και Gaskins θεώρησαν και το ακόλουθο πρόβλημα βελτιστοποίησης:

$$\max_{u} J(u^2) \tag{2.4}$$

υπό τους περιορισμούς:

$$u \in H(-\infty,+\infty), \int_{-\infty}^{+\infty} u^2(t)dt = 1.$$

Μεταξύ των προβλημάτων (2.3) και(2.4) υπάρχει μια διαφορά, εξ' αιτίας της οποίας τα δύο αυτά προβλήματα δεν είναι πάντα ισοδύναμα. Συγκεκριμένα,

- **Λήμμα 2.1:** Αν v^* είναι λύση του προβλήματος (2.4), τότε $f^* = (v^*)^2$ είναι λύση του προβλήματος (2.3) αν και μόνο αν έχουμε επιπλέον ότι $|v^*| \in H(-\infty,+\infty)$.

Απόδειξη : Υποθέτουμε ότι v^* είναι λύση του προβλήματος (2.4.) Αρχικά έστω ότι το $f^* = (v^*)^2$ είναι λύση του προβλήματος (2.3). Τότε από τους περιορισμούς του προβλήμαος (2.3) θα έχουμε ότι $(f^*)^{1/2} \in H(-\infty,+\infty)$. Ομως $(f^*)^{1/2} = |v^*|$, άρα $|v^*| \in H(-\infty,+\infty)$. Αντίστροφα, έστω ότι $|v^*| \in H(-\infty,+\infty)$. Τότε από το γεγονός ότι το v^* είναι λύση του προβλήματος (2.4) θα έχουμε ότι το $f^* = (v^*)^2$ θα είναι λύση του προβλήματος (2.3).□

Οτι η επιπλέον αυτή συνθήκη είναι αναγκαία, φαίνεται από το γεγονός ότι στο πρόβλημα (2.3) θέλουμε $\sqrt{f} \in H(-\infty,+\infty)$. Παρατηρούμε ότι αν το v^* είναι θετική συνάρτηση, τότε η συνθήκη αυτή ικανοποιείται. Αν όμως το πρόβλημα (2.4) έχει λύση που παίρνει θετικές και αρνητικές τιμές, η συνθήκη αυτή δεν ικανοποιείται πάντα.

Στην περίπτωση της πρώτης MPLE, ο χώρος H είναι ο

$$H^1(-\infty,+\infty) = \left\{ f \ : \ f, f' \in L^2(-\infty,+\infty) \right\},$$

ενώ στην περίπτωση της δεύτερης MPLE είναι ο

$$H^2(-\infty,+\infty) = \left\{ f : f, f', f'' \in L^2(-\infty,+\infty) \right\}.$$

Το εσωτερικό γινόμενο στους παραπάνω χώρους ορίζεται ως εξής:

$$\langle f, g \rangle_{H^s} = \sum_{j=0}^{s} w_j \left\langle f^{(j)}, g^{(j)} \right\rangle_{L^2(-\infty,+\infty)}, \quad w_j \geq 0, \ \ w_0, w_s > 0$$

με $s = 1, 2$ και η νόρμα

$$\|f\|_{H^s} = (\langle f, f \rangle_{H^s})^{1/2}.$$

Οι χώροι αυτοί είναι RKHS (βλέπε [2]). Στη συνέχεια θα υπολογίσουμε τον πυρήνα $K(\cdot, X_i)$ για τον χώρο $H^1(-\infty,+\infty)$, με νόρμα

$$\langle f, g \rangle_{\lambda} = 2\alpha \int_{-\infty}^{+\infty} f'(t) g'(t) dt + \lambda \int_{-\infty}^{+\infty} f(t) g(t) dt, \ \ \alpha > 0, \ \ \lambda > 0$$

που είναι ισοδύναμη με τη νόρμα που έχουμε ορίσει στον $H^1(-\infty,+\infty)$. Ο πυρήνας $K(\cdot,X_i) \equiv K_i(\cdot)$, από τις ιδιότητες του RKHS θα πρέπει να ικανοποιεί την

$$\langle f(x), K(x,X_i)\rangle_\lambda = f(X_i)$$

για κάθε $f \in H^1(-\infty,+\infty)$. Ισοδύναμα:

$$2\alpha\int_{-\infty}^{+\infty} f^{'}(t)K_i^{'}(t)dt + \lambda\int_{-\infty}^{+\infty} f(t)K_i(t)dt = f(X_i)$$

για κάθε $f \in H^1(-\infty,+\infty)$. Ομως με ολοκλήρωση κατά παράγοντες θα έχουμε ότι

$$\int_{-\infty}^{+\infty} f^{'}(t)K_i^{'}(t)dt = -\int_{-\infty}^{+\infty} f(t)K_i^{''}(t)dt,$$

άρα η προηγούμενη σχέση θα γίνει

$$\int_{-\infty}^{+\infty}\left\{-2\alpha K_i^{''}(t)+\lambda K_i(t)\right\}f(t)dt = f(X_i) = \int_{-\infty}^{+\infty}\delta_i(t)f(t)dt$$

για κάθε $f \in H^1(-\infty,+\infty)$. Αρα

$$-2\alpha K_i^{''} + \lambda K_i = \delta_i, \;\; i=1,2,...,n \tag{2.5}$$

όπου $\delta_i(t) = \delta_0(t-X_i)$ και δ_0 είναι η Dirac κατανομή, δηλαδή $\int\delta_0(t)\eta(t)dt = \eta(0)$. Τότε για $i=0$ η

$$K_0(t) = \frac{1}{2(2\alpha\lambda)^{1/2}}\exp\{-(\lambda/(2\alpha))^{1/2}\,|t|\}, \tag{2.6}$$

είναι λύση του (2.5). Αρα ο πυρήνας $K_i(\cdot)$ του χώρου $H^1(-\infty,+\infty)$, υπολογισμένος στο X_i θα δίνεται από τη σχέση:

$$K_i(t) = \frac{1}{2(2\alpha\lambda)^{1/2}}\exp\left\{-(\lambda/2\alpha)^{1/2}\,|t-X_i|\right\}, \tag{2.7}$$

μια που $K_i(t) = K_0(t-X_i), i=1,2...,n$.

Οταν H είναι ένας χώρος Hilbert και αναφερόμαστε στο ποινικοποιημένο συναρτησοειδές του λογαρίθμου της πιθανοφάνειας χωρίς να προσδιορίζουμε το συναρτησοειδές ποινής Φ, θα υποθέτουμε ότι Φ είναι το τετράγωνο της νόρμας στον χώρο H.

Πριν συνεχίσουμε με τη μελέτη της πρώτης εκτιμήτριας των Good και Gaskins, όπως αυτή έγινε από τους De Montricher, Tapia και Thompson, θα αποδείξουμε την Πρόταση 2.1 η οποία εξασφαλίζει τόσο την ύπαρξη όσο και τη μοναδικότητα των εκτιμητριών μας.

• **Πρόταση 2.1:** Εστω $H(\Omega)$, με Ω ένα διάστημα της πραγματικής ευθείας, ένας RKHS και D ένα κλειστό και κυρτό υποσύνολο του συνόλου

$$\{f \in H(\Omega) : f(X_i) \geq 0, i=1,...,n\}$$

με την ιδιότητα ότι περιέχει τουλάχιστον μια συνάρτηση, η οποία είναι θετική στα δειγματικά σημεία. Τότε το ποινικοποιημένο συναρτησοειδές της πιθανοφάνειας έχει μοναδικό μέγιστο στο D.

Απόδειξη : Αφού $H(\Omega)$ είναι ένας RKHS θα έχουμε $|f(X_i)| \leq \|K_i\|_H \|f\|_H , \quad f \in H(\Omega), \quad i = 1, 2, ...n$, (βλέπε [2]). Αρα αν

$$L(f) = \prod_{i=1}^{n} f(X_i) \exp(-\Phi(f))$$

είναι το ποινικοποιημένο συναρτησοειδές της πιθανοφάνειας, θα έχουμε

$$\begin{aligned} |L(f)| &\leq \prod_{i=1}^{n} |f(X_i)| \exp(-\Phi(f)) \\ \Rightarrow |L(f)| &\leq c_1 \|f\|_{\mathrm{H}}^n \exp(-\|f\|_{\mathrm{H}}^2). \end{aligned} \tag{2.8}$$

Παρατηρούμε ότι η συνάρτηση $\lambda^n \exp(-\lambda^2)$ είναι φραγμένη πάνω από το $(n/2)\exp(-n/2)$, άρα $|L(f)| \leq c_2$. Αν $M = \sup\{L(f), f \in D\}$, τότε υπάρχει ακολουθία $\{f_j\} \subset D$ τέτοια ώστε $L(f_j) \to M$, καθώς το $j \to +\infty$. Προφανώς $M > 0$ και $\|f_j\|_{\mathrm{H}} \leq c_3$, για κάθε j. Θεωρούμε την μπάλα

$$\{f \in H(\Omega) : \|f\|_{\mathrm{H}} \leq c_3\}, \tag{2.9}$$

η οποία εύκολα βλέπουμε ότι είναι κλειστή και φραγμένη, άρα ασθενώς συμπαγής. Αρα η $\{f_j\}$ έχει μια ασθενώς συγκλίνουσα υπακολουθία, έστω $\{f_j\}$ (για λόγους απλότητας). Δηλαδή $f_j \to f^*$ ασθενώς, άρα $f_j(\mathrm{X}_i) \to f^*(\mathrm{X}_i), j \to \infty, i = 1, ..., n$ Η νόρμα είναι weakly lower semicontinuous, άρα $\liminf_j \|f_j\|_{\mathrm{H}} \geq \|f^*\|_{\mathrm{H}}$. Συνεπώς,

$$\limsup_j \left(\prod_{i=1}^{n} f_j(X_i) \exp(-\|f_j\|^2) \right) \leq \prod_{i=1}^{n} f^*(X_i) \exp(-\|f^*\|^2) \tag{2.10}$$

Παρατηρούμε ότι το αριστερό μέρος της (2.10) είναι ίσο με M, ενώ το δεξιό ισούται με $L(f^*)$, άρα $M \leq L(f^*)$. Εχουμε όμως ισότητα μια που $f^* \in D$. Αρα το ποινικοποιημένο συναρτησοειδές της πιθανοφάνειας έχει μέγιστο στο D. Από την υπόθεση, υπάρχει τουλάχιστον ένα $f \in D$ τέτοιο ώστε $L(f) > 0$. Αρα αντί να μεγιστοποιήσουμε το L πάνω στο D, αρκεί να ελαχιστοποιήσουμε το $J = -\log L$ πάνω στο D. Με σχετικά απλές πράξεις βρήσκουμε ότι η δεύτερη Gateaux παράγωγος του J είναι

$$J''(f)(h, h) = \sum_{i=1}^{n} \frac{h(X_i)h(\mathrm{X}_i)}{f^2(\mathrm{X}_i)} + 2\langle h, h\rangle > 0, \tag{2.11}$$

άρα το J θα έχει το πολύ ένα ελάχιστο στο κυρτό σύνολο D. Συνεπώς το ποινικοποιημένο συναρτησοειδές της πιθανοφάνειας θα έχει μοναδικό μέγιστο στο κυρτό σύνολο D, με βάση το Θεώρημα 2 από το βιβλίο των Thompson και Tapia [2](σελ. 267). □

Στην επόμενη παράγραφο θα ασχοληθούμε με την ύπαρξη και τη μοναδικότητα της πρώτης εκτιμήτριας των Good και Gaskins, όπως και με συγκεκριμένες ιδιότητες που τη χαρακτηρίζουν. Για παράδειγμα θα δούμε ότι στο χώρο $H^1(-\infty, +\infty)$ είναι μια θετική εκθετική spline με κόμβους ακριβώς στα σημεία του δείγματος και ότι η εναλλακτική διατύπωση των Good και Gaskins δίνει την MPLE.

2.1 Η πρώτη εκτιμήτρια των Good και Gaskins

Στο χώρο $H^1(-\infty,+\infty)$, που όπως προαναφέραμε είναι ένας RKHS, ορίζουμε το συναρτησοειδές

$$\Phi_1(f) = \alpha \int_{-\infty}^{+\infty} \frac{(f')^2(t)}{f(t)} dt = 4\alpha \left\| \left(\sqrt{f}\right)' \right\|_2^2, \ \ \alpha > 0. \tag{2.12}$$

Χρησιμοποιώντας σαν όρο ποινής το $\Phi_1(f)$, δηλαδή ένα συναρτησοειδές της τετραγωνικής ρίζας της συνάρτησης πυκνότητας, παίρνουμε την πρώτη MPLE των Good και Gaskins [1], η οποία ορίζεται σαν η λύση του ακόλουθου προβλήματος βελτιστοποίησης,

$$\max_f \left\{ n \int_{-\infty}^{+\infty} \log f(x) dF_n(x) - \Phi_1(f) \right\} \tag{2.13}$$

υπό τους περιορισμούς

$$f^{1/2} \in H^1(-\infty,+\infty), \ \int_{-\infty}^{+\infty} f(t)dt = 1, \ f(t) \geq 0, \ \forall \ t \in \mathbf{R}.$$

Συνχά βοηθάει να βλέπουμε το παραπάνω πρόβλημα εκφρασμένο με την ακόλουθη ισοδύναμη μορφή του

$$\max_f \left\{ \prod_{i=1}^{n} f(X_i) \exp(-\Phi_1(f)) \right\} \Leftrightarrow \max_f \left\{ \sum_{i=1}^{n} \log f(X_i) - \Phi_1(f) \right\}, \tag{2.14}$$

με τους ίδιους περιορισμούς.

Στη συνέχεια, θα δούμε ότι η λύση αυτή υπάρχει, είναι μοναδική και είναι μια θετική εκθετική spline με κόμβους ακριβώς στα σημεία του δείγματος.

Για τη μελέτη του προβλήματος (2.13) οι De Montricher, Tapia και Thompson διατύπωσαν το ακόλουθο βοηθητικό πρόβλημα εκφρασμένο σε όρους του $u = f^{1/2}$

$$\max_u \left\{ n \int_{-\infty}^{+\infty} \log u(x) dF_n(x) - \Phi(u) \right\} \tag{2.15}$$

υπό τους περιορισμούς:

$$u \in H^1(-\infty,+\infty), \ \int_{-\infty}^{+\infty} u^2(t)dt = 1, \ u(t) \geq 0, \ \forall \ t \in \mathbf{R},$$

όπου $\Phi(u) = 2\alpha \int_{-\infty}^{+\infty} \left(u'\right)^2 (t)dt$. Παρατηρούμε ότι το πρόβλημα (2.15) προκύπτει από το (2.14) αν θέσουμε $u = f^{1/2}$ και πάρουμε την τετραγωνική ρίζα της ποσότητας που μεγιστοποιούμε.

• **Λήμμα 2.1.1:** Αν u^* είναι λύση του (2.15) τότε $f^* = (u^*)^2$ είναι λύση του (2.13).

Απόδειξη: Το πρόβλημα (2.13) για $u = f^{1/2}$, γίνεται:

$$\max_u \left\{ n \int_{-\infty}^{+\infty} \log u^2(x) dF_n(x) - 4\alpha \int_{-\infty}^{+\infty} \left(u'\right)^2 (t)dt \right\} \tag{2.16}$$

υπό τους περιορισμούς:

$$u \in H^1(-\infty, +\infty), \int_{-\infty}^{+\infty} u^2(t)dt = 1.$$

Παρατηρούμε ότι αν u^* λύση του προβλήματος (2.15) τότε u^* θα είναι λύση και του προβλήματος (2.16). Αρα $f^* = (u^*)^2$ θα είναι λύση του (2.13), με βάση το Λημμα 2.1. □

Για να αποφύγουν τη μεγιστοποίηση του ποινικοποιημένου συναρτησοειδούς στο θετικό κώνο, θεώρησαν μια εναλλακτική διατύπωση του προβλήματος (2.15). Συγκεκριμμένα διατύπωσαν το ακόλουθο πρόβλημα βελτιστοποίησης:

$$\max_u \left\{ n \int_{-\infty}^{+\infty} \log u(x) dF_n(x) - \Phi(u) \right\} \tag{2.17}$$

υπό τους περιορισμούς:

$$u \in H^1(-\infty, +\infty), \int_{-\infty}^{+\infty} u^2(t)dt = 1, \; u(X_i) \geq 0, \; \forall \; i = 1, 2, ..., n.$$

Οπως θα δούμε στη συνέχεια η εναλλακτική αυτή διατύπωση δίνει την πρώτη MPLE των Good και Gaskins.

Θεώρησαν ακόμα το ακόλουθο βοηθητικό πρόβλημα:

$$\max_u \left\{ L_\lambda(u) = n \int_{-\infty}^{+\infty} \log u(x) dF_n(x) - 2\alpha \int_{-\infty}^{+\infty} \left(u'\right)^2 (t)dt - \lambda \int_{-\infty}^{+\infty} u^2(t)dt \right\} \tag{2.18}$$

υπό τους περιορισμούς:

$$u \in H^1(-\infty, +\infty), \; u(\mathrm{X}_i) \geq 0, \; \forall \; i = 1, 2, ..., n,$$

για κάθε $\lambda > 0$.

Στην Πρόταση 2.1.1 αποδεικνύουμε ότι το πρόβλημα (2.18) έχει μοναδική λύση u_λ για την οποία

$$\|u_\lambda\|_2^2 \geq \frac{n}{4\lambda}$$

για κάθε $\lambda > 0$.

Στην Πρόταση 2.1.2 αποδεικνύουμε, χρησιμοποιώντας την Πρόταση 2.1.1, ότι το πρόβλημα (2.17) έχει μοναδική λύση. Σύμφωνα με την Πρόταση 2.1.3 η λύση αυτή θα είναι μια θετική εκθετική spline με κόμβους στα σημεία του δείγματος και άρα θα είναι λύση και του προβλήματος (2.15). Συνεπώς το τετράγωνο της θα είναι λύση του αρχικού μας προβλήματος (2.13) με βάση το Λήμμα 2.1.1.

• **Πρόταση 2.1.1:** Το πρόβλημα (2.18) έχει μοναδική λύση. Επιπλέον αν u_λ είναι αυτή η λύση τότε:

1. u_λ είναι μια εκθετική spline με κόμβους στα σημεία του δείγματος,
2. $u_\lambda(t) > 0$,για κάθε $t \in \mathbf{R}$
3. $\|u_\lambda\|_2 \geq (\frac{n}{4\lambda})^{1/2}$.

Απόδειξη: Ο $H^1(-\infty, +\infty)$ είναι ένας RKHS και αν θεωρήσουμε τη νόρμα

$$\|u\|_\lambda^2 = 2\alpha \int_{-\infty}^{+\infty} \left(u^{'}\right)^2 (t)dt + \lambda \int_{-\infty}^{+\infty} u^2(t)dt$$

η οποία είναι ισοδύναμη με την νόρμα του $H^1(-\infty, +\infty)$, το πρόβλημα μας θα έχει μοναδική λύση, έστω u_λ. Πράγματι, αυτό είναι άμεσο από την Πρόταση 2.1, αν πάρουμε σαν

$$D = \left\{u \in H^1(-\infty, +\infty) : u(X_i) \geq 0, i = 1, 2..., n\right\}.$$

Εστω $K_i(x)$ ο πυρήνας του $H^1(-\infty, +\infty)$ υπολογισμένος στο X_i. Αφού το u_λ είναι η λύση του προβλήματος βελτιστοποίησης (2.18), θα έχουμε ότι $u_\lambda(X_i) > 0$ για κάθε $i = 1, ..., n$ μια που αν ήταν ίσο με το μηδέν η log πιθανοφάνεια θα γινόταν μείον άπειρο και άρα δεν θα ήταν σημείο μεγίστου. Ακόμα

$$\nabla L_\lambda(u_\lambda) = 0, \tag{2.19}$$

όπου

$$L_\lambda(u_\lambda) = \sum_{i=1}^{n} \log u_\lambda(X_i) - 2\alpha \int_{-\infty}^{+\infty} \left(u_\lambda^{'}\right)^2 (t)dt - \lambda \int_{-\infty}^{+\infty} u_\lambda^2(t)dt.$$

Δεν μένει παρά να υπολογίσουμε το $\nabla L_\lambda(u_\lambda)$. Εχουμε

$$\begin{aligned} L_\lambda(u_\lambda + \varepsilon h) &= \sum_{i=1}^{n} \log(u_\lambda(X_i) + \varepsilon h(X_i)) \\ &- 2\alpha \int_{-\infty}^{+\infty} (u_\lambda'(t) + \varepsilon h'(t))^2 dt - \lambda \int_{-\infty}^{+\infty} (u_\lambda(t) + \varepsilon h(t))^2 dt. \end{aligned}$$

Αρα (βλέπε [2])

$$\begin{aligned} \nabla L_\lambda(u_\lambda)(h) &= \lim_{\varepsilon \to 0} \frac{d}{d\varepsilon} L_\lambda(u_\lambda + \varepsilon h) \\ &= \sum_{i=1}^{n} \frac{h(X_i)}{u_\lambda(X_i)} - 4\alpha \int_{-\infty}^{+\infty} u_\lambda'(t)h'(t)dt - 2\lambda \int_{-\infty}^{+\infty} u_\lambda(t)h(t)dt \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} \frac{1}{u_\lambda(X_i)} \langle h, K_i \rangle_\lambda - 2 \langle u_\lambda, h \rangle_\lambda \\
&= \left\langle \sum_{i=1}^{n} \frac{K_i}{u_\lambda(X_i)} - 2u_\lambda, h \right\rangle_\lambda , \qquad (2.20)
\end{aligned}
$$

για κάθε $h \in H^1(-\infty, +\infty)$. Αρα

$$\nabla L_\lambda(u_\lambda) = 2u_\lambda - \sum_{i=1}^{n} \frac{K_i}{u_\lambda(X_i)},$$

οπότε από τη (2.19) θα έχουμε

$$2u_\lambda - \sum_{i=1}^{n} \frac{K_i}{u_\lambda(X_i)} = 0 \Rightarrow u_\lambda = \frac{1}{2} \sum_{i=1}^{n} \frac{K_i}{u_\lambda(X_i)} \qquad (2.21)$$

Συνεπώς, από τη σχέση (2.7), βλέπουμε ότι το u_λ είναι μια θετική εκθετική spline με κόμβους ακριβώς στα σημεία του δείγματος, όπου οι σταθερές της σχέσης (2.2) είναι $\alpha_i = 0$, $\beta_i = 0$, $\gamma_i = \frac{1}{4\sqrt{2\alpha\lambda}} \frac{1}{u_\lambda(X_i)}$, $\delta_i = -\gamma_i$ και $\kappa_i = \sqrt{\frac{\lambda}{2\alpha}}$.

Μένει λοιπόν ότι:

$$\|u_\lambda\|_2 \geq (\frac{n}{4\lambda})^{1/2}.$$

Εχουμε

$$
\begin{aligned}
\|u_\lambda\|_\lambda^2 &= \langle u_\lambda, u_\lambda \rangle_\lambda = \left\langle \frac{1}{2} \sum_{i=1}^{n} \frac{K_i}{u_\lambda(X_i)}, u_\lambda \right\rangle_\lambda \\
&= \frac{1}{2} \sum_{i=1}^{n} \frac{1}{u_\lambda(X_i)} \langle K_i, u_\lambda \rangle_\lambda \qquad (2.22) \\
&= \frac{1}{2} \sum_{i=1}^{n} \frac{1}{u_\lambda(X_i)} u_\lambda(X_i) = \frac{n}{2}
\end{aligned}
$$

Από τη σχέση (2.7),

$$K_i'(t) = (\frac{\lambda}{2\alpha})^{1/2} K_i(t) \qquad (2.23)$$

για $t - X_i < 0$ και

$$K_i'(t) = -(\frac{\lambda}{2\alpha})^{1/2} K_i(t) \qquad (2.24)$$

για $t - X_i > 0$. Αρα

$$K_i'(t) K_j'(t) \leq \frac{\lambda}{2\alpha} K_i(t) K_j(t), i, j = 1, 2..., n.$$

Οπότε τελικά θα έχουμε, μέσω της (2.21), ότι

$$\left(u_\lambda'\right)^2 (t) = \frac{1}{4} \left\{ \sum_i \left(\frac{K_i'(t)}{u_\lambda(X_i)} \right)^2 + \sum_{i,j} \frac{K_i'(t) K_j'(t)}{u_\lambda(X_i) u_\lambda(\mathrm{X}_j)} \right\} \leq \frac{\lambda}{2\alpha} u_\lambda^2(t). \qquad (2.25)$$

Αν ολοκληρώσουμε την παραπάνω σχέση ως πρός t παίρνουμε

$$
\begin{aligned}
2\alpha \int_{-\infty}^{+\infty} \left(u_\lambda'\right)^2 (t) dt &\leq \lambda \int_{-\infty}^{+\infty} u_\lambda^2(t) dt \\
\Rightarrow 2\alpha \left\| u_\lambda' \right\|_2^2 &\leq \lambda \left\| u_\lambda \right\|_2^2 \qquad (2.26)
\end{aligned}
$$

Εξ' ορισμού της νόρμας $\|\cdot\|_\lambda$ και χρησιμοποιώντας τις σχέσεις (2.26) και (2.22) βλέπουμε ότι

$$\|u_\lambda\|_\lambda^2 \le 2\lambda \|u_\lambda\|_2^2 \Rightarrow \|u_\lambda\|_2 \ge (\tfrac{n}{4\lambda})^{1/2}. \tag{2.27}$$

Εχουμε λοιπόν το ζητούμενο αποτέλεσμα.□

• **Πρόταση 2.1.2:** Το πρόβλημα (2.17) έχει μοναδική λύση.

Απόδειξη: Εστω

$$\Theta = \left\{ u \in H^1(-\infty, +\infty) : \int_{-\infty}^{+\infty} u^2(t)dt \le 1, u(X_i) \ge 0, i = 1, 2..., n \right\}.$$

Είναι εύκολο να δούμε ότι το Θ είναι κλειστό και κυρτό. Με βάση την Πρόταση 2.1, το συναρτησοειδές L_λ έχει μοναδικό μέγιστο στο Θ, έστω μ_λ. Ενώ από την Πρόταση 2.1.1, αν u_λ είναι η λύση του (2.18), μπορώ να επιλέξω $0 < \lambda < \frac{n}{4}$ ώστε $\|u_\lambda\|_2 > 1$. Θα δείξουμε ότι:

$$\|\mu_\lambda\|_2 = 1$$

Εστω ότι $\|\mu_\lambda\|_2 < 1$. Θεωρούμε

$$u_\vartheta = \vartheta u_\lambda + (1-\vartheta)\mu_\lambda, \quad \vartheta \in (0,1).$$

Το συναρτησοειδές $(-L_\lambda)$ είναι γνήσια κυρτό. Αρα

$$L_\lambda(u_\vartheta) \ge \vartheta L_\lambda(u_\lambda) + (1-\vartheta) L_\lambda(\mu_\lambda). \tag{2.28}$$

Το μ_λ ικανοποιεί τους περιορισμούς του προβλήματος (2.18) άρα

$$L_\lambda(\mu_\lambda) \le L_\lambda(u_\lambda).$$

Συνεπώς η (2.28) δίνει

$$L_\lambda(u_\vartheta) \ge L_\lambda(\mu_\lambda), \tag{2.29}$$

για κάθε $\vartheta \in (0,1)$. Εστω

$$g(\vartheta) = \|u_\vartheta\|_2 \tag{2.30}$$

τότε

$$g(0) = \|\mu_\lambda\|_2 < 1$$

και

$$g(1) = \|u_\lambda\|_2 > 1$$

άρα υπάρχει $\vartheta_0 \in (0,1)$ τέτοιο ώστε

$$g(\vartheta_0) = \|u_{\vartheta_0}\|_2 = 1 \tag{2.31}$$

και

$$L_\lambda(u_{\vartheta_0}) \ge L_\lambda(\mu_\lambda). \tag{2.32}$$

Ατοπο, διότι $u_{\vartheta_0} \in \Theta$ και στο Θ το μ_λ είναι το μοναδικό μέγιστο του L_λ. Αρα $\|\mu_\lambda\|_2 = 1$. Συνεπώς το μ_λ είναι η μοναδική λύση του προβλήματος

$$\max_u L_\lambda(u) \tag{2.33}$$

υπό τους περιορισμούς:

$$u \in H^1(-\infty, +\infty), \int_{-\infty}^{+\infty} u^2(t)dt = 1, u(X_i) \geq 0, \forall i = 1, ..., n,$$

για $0 < \lambda < \frac{n}{4}$. Παρατηρούμε ότι ο όρος $\lambda \int_{-\infty}^{+\infty} u^2(t)dt$, στον οποίο τα προβλήματα (2.33) και (2.17) διαφέρουν, είναι σταθερός κάτω από τους περιορισμούς τους. Αρα τα προβλήματα (2.33) και (2.17) έχουν την ίδια λύση για κάθε $\lambda > 0$. Αρα το πρόβλημα (2.17) έχει μοναδική λύση.□

• **Πρόταση 2.1.3:** Το πρόβλημα (2.15) έχει μοναδική λύση η οποία είναι θετική εκθετική spline με κόμβους ακριβώς στα σημεία του δείγματος.

Απόδειξη: Αν δείξουμε ότι η μοναδική λύση v του προβλήματος (2.17)(γνωρίζουμε ότι υπάρχει μέσω της Πρότασης 2.1.2) είναι θετική εκθετική spline με κόμβους στα σημεία του δείγματος, τότε θα είναι λύση και του προβλήματος (2.15). Εστω λοιπόν v η λύση του (2.17). Εστω

$$g(u) = \int_{-\infty}^{+\infty} u^2(t)dt.$$

και

$$L(u) = n \int_{-\infty}^{+\infty} \log u(x) dF_n(x) - \Phi(u)$$

όπου $\Phi(u) = 2\alpha \int_{-\infty}^{+\infty} \left(u'\right)^2 (t)dt$, για $u \in H^1(-\infty, +\infty)$. Προφανώς $v(X_i) > 0$ για $i = 1, ..., n$. Αρα από τη μέθοδο των πολλαπλασιαστών Lagrange, υπάρχει λ ώστε το v να ικανοποιεί τις εξισώσεις:

$$L'(u) - \lambda g'(u) = 0 \tag{2.34}$$

και

$$g(u) = 1. \tag{2.35}$$

Οπου

$$\begin{aligned} \nabla L(u_\lambda)(h) &= \sum_{i=1}^n \frac{h(X_i)}{u(X_i)} - 4\alpha \int_{-\infty}^{+\infty} u'(t)h'(t)dt \\ &= \sum_{i=1}^n \frac{1}{u(X_i)} \int_{-\infty}^{+\infty} \delta_i(t)h(t)dt + 4\alpha \int_{-\infty}^{+\infty} u''(t)h(t)dt, \end{aligned} \tag{2.36}$$

με $\delta_i(t) = \delta_0(t - X_i)$ και

$$g'(u)(h) = 2\int_{-\infty}^{+\infty} u(t)h(t)dt \tag{2.37}$$

Αντικαθιστώντας τις (2.36) και (2.37) στην (2.34) θα έχουμε:

$$\sum_{i=1}^n \frac{1}{u(X_i)} \int_{-\infty}^{+\infty} \delta_i(t)h(t)dt + 4\alpha \int_{-\infty}^{+\infty} u''(t)h(t)dt - 2\lambda \int_{-\infty}^{+\infty} u(t)h(t)dt = 0$$

και αυτό για κάθε h. Αρα

$$\sum_{i=1}^{n} \frac{1}{u(X_i)}\delta_i + 4\alpha u'' - 2\lambda u = 0,$$

δηλαδή

$$-4\alpha u'' + 2\lambda u = \sum_{i=1}^{n} \frac{1}{u(X_i)}\delta_i \tag{2.38}$$

και

$$g(u) = \int_{-\infty}^{+\infty} u^2(t)dt = 1, \tag{2.39}$$

όπου για το δ_i έχουμε ότι $\int_{-\infty}^{+\infty} u(t)\delta_i(t)dt = u(X_i)$ για κάθε $i = 1, ..., n$. Αφού λοιπόν έχουμε αποδείξει στην Πρόταση 2.1.2 ότι το πρόβλημα (2.17) έχει μοναδική λύση, θα πρέπει οι εξισώσεις (2.38) και(2.39) να έχουν μοναδική λύση στο χώρο $H^1(-\infty, +\infty)$. Εστω ότι $\lambda \leq 0 \cdot$ θα καταλήξουμε σε άτοπο. Αν υποθέσουμε ότι $x < X_i, \ \forall i = 1, ..., n$, τότε στο διάστημα $(-\infty, x)$ η v θα είναι λύση της εξίσωσης $-4\alpha u'' + 2\lambda u = 0$ και άρα αυτή θα είναι άθροισμα τριγωνομετρικών συναρτήσεων, μια που δεν μπορεί να είναι η μηδενική συνάρτηση. Δηλαδή η v θα είναι μια περιοδική συνάρτηση και άρα δεν μπορεί να ικανοποιεί την (2.39). Συνεπώς πρέπει $\lambda > 0$. Είναι εύκολο να παρατηρήσουμε ότι

$$\begin{aligned} L(u) - \lambda g(u) &= \sum_{i=1}^{n} \log u(X_i) - 2\alpha \int_{-\infty}^{+\infty} \left(u'\right)^2 (t)dt - \lambda \int_{-\infty}^{+\infty} u^2(t)dt \\ &= L_\lambda(u), \end{aligned} \tag{2.40}$$

άρα το v θα είναι λύση και του (2.18) για αυτό το λ. Δηλαδή το v θα είναι λύση και του (2.18), οπότε με βάση την Πρόταση 2.1.1 το v θα είναι μια θετική εκθετική spline με κόμβους στα δειγματικά σημεία και αυτό ολοκληρώνει την απόδειξη μας.□

ωτης

Κεφάλαιο 3

Ενας διακριτός όρος ποινής

Οι Scott,Tapia και Thompson το 1980 [2] (βλέπε επίσης [3]), πρότειναν μια διακριτοποίηση του αρχικού προβλήματος βελτιστοποίησης, που κάνει τη διάσταση του χώρου πεπερασμένη και δίνει ως εκτιμήτρια μια εξομάλυνση του ιστογράμματος. Συγκεκριμένα, θεώρησαν τυχαίο δείγμα $X_1, X_2..., X_n$, από μια συνάρτηση κατανομής F, με συνάρτηση πυκνότητας f, για την οποία υπέθεσαν ότι συγκεντρώνεται στο φραγμένο διάστημα $[a, b]$, με $f(a) = f(b) = 0$. Για την εκτίμηση της f, μελέτησαν το πρόβλημα μεγιστοποίησης του συναρτησοειδούς του λογαρίθμου της πιθανοφάνειας ως πρός μια διακριτή πυκνότητα μείον ένα διακριτό όρο ποινής. Στο κεφάλαιο αυτό, θα παρουσιάσουμε το παραπάνω πρόβλημα βελτιστοποίησης και θα δούμε ότι έχει μοναδική λύση τόσο στο σύνολο των ιστογραμμάτων, όσο και στο σύνολο των ιστοπολυγώνων. Θα δούμε ακόμα τη συνέπεια των εκτιμητριών αυτών στο σύνολο των ιστογραμμάτων, καθώς και αριθμητικά αποτελέσματα σχετικά με τον υπολογισμό των παραπάνω εκτιμητριών στο χώρο των ιστοπολυγώνων.

Για να εφαρμόσουμε την παραπάνω ιδέα, διαμερίζουμε το διάστημα $[a, b]$ σε m ίσα τμήματα ,

$$a = t_0 < t_1 < ... < t_m = b,$$

όπου το κάθε ένα από αυτά έχει μήκος $h = t_j - t_{j-1} = \frac{b-a}{m}, j = 1, 2...m$. Συμβολίζουμε με T_i το διάστημα $[t_{i-1}, t_i)$, για $i = 1, ...m$ και με $I(T_i)$ την χαρακτηριστική συνάρτηση του διαστήματος T_i, δηλαδή $I(T_i)(x) = 0$ για $x \notin T_i$ και $I(T_i)(x) = 1$ για $x \in T_i$. Εστω $S_0(t_0, t_1, ...t_m)$, το σύνολο των συναρτήσεων με φορέα το διάστημα $[t_0, t_m) = [a, b)$ οι οποίες είναι σταθερές στα διάστηματα T_i. Τέλος, έστω $S_1(t_0, t_1, ...t_m)$ το σύνολο των συνεχών συναρτήσεων με φορέα το διάστημα $(t_0, t_m) = (a, b)$, οι οποίες είναι γραμμικές στα διαστήματα T_i. Παρατηρούμε ότι, κάθε συνάρτηση $p \in S_0(t_0, ...t_m)$ έχει τη μορφή

$$p(x) = \sum_{i=1}^{m} p_{i-1} I(T_i)(x),$$

στη συνέχεια συμβολίζουμε με p_i την τιμή της συνάρτησης στη θέση t_i, δηλαδή

$$p(t_i) = p_i$$

για καθε $i = 0, ...m-1$ και υποθέτουμε ότι $p_m = 0$. Αρα ο χώρος $S_0(t_0, ...t_m)$ είναι ισόμορφος με τον R^m.

Από τον τρόπο που ορίσαμε το χώρο $S_1(t_0, ...t_m)$, κάθε συνάρτηση $s \in S_1(t_0, ...t_m)$ θα έχει τη μορφή

$$s(x) = \sum_{i=1}^{m} \{s_{i-1} + h^{-1}(x - t_{i-1})(s_i - s_{i-1})\} I(T_i)(x)$$

όπου και πάλι έχουμε

$$s(t_i) = s_i$$

για $i = 1, ...m-1$ και υποθέτουμε ότι $s_0 = s(t_0) = 0$, $s_m = s(t_m) = 0$. Αρα ο χώρος $S_1(t_0, ...t_m)$ είναι ισόμορφος με τον R^{m-1}. Σημειώνουμε ότι τα μεν στοιχεία του $S_0(t_0, ...t_m)$ είναι γενικευμένα ιστογράμματα, τα δε στοιχεία του $S_1(t_0, ...t_m)$ είναι γενικευμένα ιστοπολύγωνα.

Δεδομένου του τυχαίου δείγματος $X_1, X_2..., X_n$, στο διάστημα $[a, b]$ και μιας θετικής σταθεράς α, θεωρούμε τα ακόλουθα προβλήματα βελτιστοποίησης (ορίζουμε $p_{-1} = 0$)

$$\max L(p_0, p_1, ..., p_{m-1}) = \max \left\{ \sum_{k=1}^{n} \log p(X_k) - \alpha h \sum_{i=0}^{m} \left(\frac{p_i - p_{i-1}}{h} \right)^2 \right\} \tag{3.1}$$

υπό τους περιορισμούς:

$$h \sum_{i=0}^{m-1} p_i = 1, \ p_i \geq 0, \ \forall \ i = 0, ..., m-1$$

και

$$\max L(s_1, ..., s_{m-1}) = \max \left\{ \sum_{i=1}^{n} \log s(X_i) - \frac{\alpha}{h^3} \sum_{i=1}^{m-1} \{s_{i+1} - 2s_i + s_{i-1}\}^2 \right\} \tag{3.2}$$

υπό τους περιορισμούς:

$$h \sum_{i=1}^{m-1} s_i = 1, \ s_i \geq 0, \ \forall \ i = 1, ..., m-1.$$

Ο όρος ποινής στο πρόβλημα (3.1) είναι μια διακριτή προσέγγιση του όρου $\int (f')^2$ όπου f είναι η συνάρτηση πυκνότητας, στο πρόβλημα (3.2) σαν όρο ποινής έχουμε χρησιμοποιήσει την διακριτή δεύτερη παράγωγο.

Τόσο η κατά τμήματα σταθερή spline (στοιχείο του $S_0(t_0, ...t_m)$) που αντιστοιχεί στην λύση του προβλήματος (3.1), όσο και η κατά τμήματα γραμμική spline (στοιχείο του $S_1(t_0, ...t_m)$) που αντιστοιχεί στην λύση του (3.2) και είναι γενικευμένο ιστοπολύγωνο ονομάζονται, διακριτές MPLE και στη συνέχεια θα αναφερόμαστε σε αυτές ως DMPLE, πρώτη και δεύτερη αντίστοιχα.

Στη συνέχεια θα δούμε το Θεώρημα 3.1, το οποίο εξασφαλίζει τόσο την ύπαρξη όσο και τη μοναδικότητα των εκτιμητριών μας.

• **Θεώρημα 3.1 :** Εστω $H(a,b)$ ένας RKHS και ας υποθέοουμε ότι το ολοκλήρωμα είναι ένας συνεχής τελεστής. Αν υπάρχει τουλάχιστον μια $f \in H(a,b)$ τέτοια ώστε

$$\int_a^b f(t)dt = 1, \;\; f(t) \geq 0 \;\; \forall t \in (a,b), \;\; f(\mathrm{X}_i) > 0,$$

τότε η MPLE που αντιστοιχεί στον $H(a,b)$ υπάρχει και είναι μοναδική.

Αν πάρουμε για $D = \left\{ f \in H(a,b) : \int_a^b f(t)dt = 1, f(t) \geq 0 \right\}$ το οποίο είναι κλειστό και κυρτό, το παραπάνω Θεώρημα έπεται από την Πρόταση 2.1.

• **Θεώρημα 3.2 :** Οι DMPLE που ορίζονται μέσω των προβλημάτων (3.1) και (3.2) υπάρχουν και είναι μοναδικές.

Απόδειξη: Η απόδειξη είναι άμεση από το Θεώρημα 3.1, αρκεί να σημειώσουμε ότι κάθε χώρος πεπερασμένης διάστασης με εσωτερικό γινόμενο είναι ένας RKHS (βλέπε [2]) και ότι όλα τα γραμμικά συναρτησοειδή σε χώρο πεπερασμένης διάστασης είναι συνεχή (βλέπε [2]).□

Στην επόμενη παράγραφο θα περιοριστούμε στο σύνολο $S_0(t_0, ...t_m)$ και θα δούμε τη συνέπεια της πρώτης DMPLE όπως αυτή παρουσιάστηκε από τους Scott-Tapia-Thompson [2],[3].

3.1 Επίλυση και συνέπεια της πρώτης DMPLE

Εστω v_k ο αριθμός των παρατηρήσεων στο διάστημα $[t_k, t_{k+1})\,,\;\; k = 0, ..., m-1$. Τότε

$$\sum_{k=0}^{m-1} v_k = n$$

και το πρόβλημα (3.1) μετασχηματίζεται στην ακόλουθη μορφή:

$$\max L(p_0, p_1, ..., p_{m-1}) = \max \left\{ \sum_{i=0}^{m-1} v_i \log(p_i) - \frac{\alpha}{h} \sum_{i=0}^{m} (p_i - p_{i-1})^2 \right\}, \tag{3.3}$$

υπό τους περιορισμούς:

$$h \sum_{i=0}^{m-1} p_i = 1, \;\;\; p_i \geq 0, \; \forall i = 0, ..., m-1.$$

Εστω $[x:n]$ το διάστημα μήκους h που περιέχει εκείνο το x (για μια συγκεκριμένη τιμή του n) για το οποίο στο Θεώρημα 3.1.1 θα δείξουμε συνέπεια, και $v[x:n]$ ο αριθμός των δειγμάτων στο διάστημα αυτό. Στη συνέχεια αφήνουμε το h να εξαρτάται από το n και συμβολίζοντας το με h_n, μελετούμε τη συμπεριφορά του ιστογράμματος

$$\frac{v[x:n]}{nh_n}$$

καθώς το n τείνει στο άπειρο.

• **Λήμμα 3.1.1 :** Εστω $X_1, X_2..., X_n$ ένα τυχαίο δείγμα από μια συνεχή συνάρτηση πυκνότητας f με φραγμένο φορέα (a,b). Θεωρούμε μια διαμέριση του (a,b) όπως την προηγούμενη, με $m = [cn^q]$, $c > 0, 0 < q < \frac{1}{3}$, όπου $[d]$ είναι το ακέραιο μέρος του d. Τότε για $x \in (a,b)$,

$$\lim_{n\to+\infty} \frac{v[x:n]}{nh_n} = f(x) \quad a.s..$$

Απόδειξη : Ορίζουμε τις τυχαίες μεταβλητές

$$Y_{ni} = \frac{1[x:n](X_i) - p[x:n]}{h_n}$$

όπου

$$p[x:n] = \int_{[x:n]} f(t)dt.$$

Για κάθε n, οι $\{Y_{ni}, i = 1, ...n\}$ είναι ανεξάρτητες τυχαίες μεταβλητές. Παρατηρούμε ότι η $1[x:n](X_i)$ είναι μια Bernoulli τυχαία μεταβλητή με μέση τιμή ίση με $p[x:n]$. Αρα $EY_{ni} = 0$ και $Var(Y_{ni}) = \frac{1}{h_n^2} p[x:n](1 - p[x:n])$. Εστω

$$S_n = \sum_{i=1}^{n} Y_{ni},$$

ισχυριζόμαστε ότι

$$\frac{S_n}{n} \to 0 \quad a.s.$$

καθώς το n τείνει στο άπειρο. Για να το δούμε αυτό, αρκεί να δούμε ότι για κάθε $\epsilon > 0$

$$\sum_{n=1}^{\infty} P\{|S_n| > n\epsilon\} < +\infty.$$

Πράγματι, αυτό είναι άμμεσο από το πρώτο Λήμμα Borel Cantelli [12], σύμφωνα με το οποίο αν $\sum_n P(A_n) < +\infty$, τότε $P(\limsup_n A_n) = 0$.

Ισχυρίζομαστε ότι, αν μια τυχαία μεταβλητή $X \sim B(n,p)$, τότε $E(X - EX)^4 \leq 3pn^2$ για μεγάλα n. (Βλέπε την Σημείωση 3.1.1 που ακολουθεί). Στην περίπτωση μας, αφού $1[x:n](X_i)$ ακολουθεί μια $B(1, p[x:n])$ το $\sum_{i=1}^{n} 1[x:n](X_i)$ θα ακολουθεί μια $B(n, p[x:n])$ και άρα

$$E\left(\left\{\sum_{i=1}^{n} 1[x:n](X_i) - np\,[x:n]\right\}^4\right) \leq 3n^2 p\,[x:n]. \tag{3.4}$$

Η ανισότητα Markov μας δίνει ότι, για κάθε $\epsilon > 0$

$$P\{|S_n| > n\epsilon\} < \frac{E(S_n^4)}{(n\epsilon)^4}, \tag{3.5}$$

όπου

$$E(S_n^4) = \frac{1}{h_n^4} E\left(\left\{\sum_{i=1}^{n} 1[x:n](X_i) - np[x:n]\right\}^4\right).$$

Αρα η (3.5) μέσω της (3.4) θα γίνει

$$P\{|S_n| > n\epsilon\} < \frac{3p[x:n]}{n^2 \epsilon^4 h_n^4}. \tag{3.6}$$

Από το θεμελιώδες θεώρημα του ολοκληρωτικού λογισμού, υπάρχει $\xi \in [x:n]$ τέτοιο ώστε

$$p[x:n] = \int_{[x:n]} f(t)dt = h_n f(\xi).$$

Αρα η (3.6) θα γίνει

$$P\{|S_n| > n\epsilon\} < \frac{3f(\xi)}{n^2\epsilon^4 h_n^3}.$$

Αθροίζοντας πάνω σε όλα τα n θα έχουμε ότι

$$\begin{aligned}\sum_{n=1}^{\infty} P\{|S_n| > n\epsilon\} &< \sum_{n=1}^{\infty} \frac{3f(\xi)}{n^2\epsilon^4 h_n^3} \\ &= \sum_{n=1}^{\infty} O\left(\frac{1}{n^{2-3q}}\right),\end{aligned}$$

όπου για την τελευταία ισότητα αρκεί να παρατηρήσουμε ότι, $h_n = O(n^{-q})$ με $q > 0$, μια που $h = \frac{b-a}{m}$ και $m = [cn^q]$. Ομως η τελευταία σειρά συγκλίνει μια που $0 < q < 1/3$, άρα

$$\sum_{n=1}^{\infty} P\{|S_n| > n\epsilon\} < \infty.$$

Συνεπώς από το Λήμμα Borel Cantelli

$$\frac{S_n}{n} \to 0 \ \ a.s.$$

καθώς το n τείνει στο άπειρο. Ομως

$$\frac{S_n}{n} \equiv \frac{v[x:n]}{nh_n} - \frac{p[x:n]}{h_n}.$$

Λόγο της συνέχειας της συνάρτησης πυκνότητας f,

$$\lim_{n\to\infty} \frac{\int_{[x:n]} f(t)dt}{h_n} = f(x).$$

Αρα

$$\frac{v[x:n]}{nh_n} \to f(x) \ \ a.s.$$

καθώς το $n \to \infty$, και έχουμε ολοκληρώσει την απόδειξη του Λήμματος.□

• **Σημείωση 3.1.1 :** Αν η τυχαία μεταβλητή X ακολουθεί μια $B(n,p)$, τότε $E(X-EX)^4 \leq 3pn^2$.

Απόδειξη: Είναι εύκολο να δούμε ότι

$$E\left(X(X-1)(X-2)...(X-k)\right) = n(n-1)(n-2)...(n-k)p^{k+1} \tag{3.7}$$

Αν αναπτύξουμε τη σχέση $E(X-EX)^4$, θα έχουμε

$$E(X-EX)^4 = EX^4 - 4EX^3EX + 6EX^2(EX)^2 - 4EX(EX)^3 + (EX)^4 \tag{3.8}$$

Από τη σχέση (3.7) μπορούμε εύκολα να υπολογίσουμε τις EX^4, EX^3, EX^2. Πράγματι,

$$EX^2 = np(1-p+np),$$

$$EX^3 = n(n-1)(n-2)p^3 + 3EX^2 - 2np$$

και

$$\begin{aligned} EX^4 &= (n-1)(n-2)(n-3)p^4 + 6EX^3 - 11EX^2 + 6EX \\ &= n^4p^4 - 6n^3p^4 + 11n^2p^4 - 6np^4 + 6n^3p^3 - 18n^2p^3 \\ &+ 12np^3 + np - 7np^2 + 7n^2p^2. \end{aligned}$$

Αν αντικαταστήσουμε στη σχέση (3.8) τις παραπάνω μέσες τιμές, θα έχουμε μετά από πράξεις ότι

$$E(X-EX)^4 = n^2(3p^4 - 6p^3 + 3p^2) + n(-6p^4 + 12p^3 - 7p^2 + p).$$

Αρα θέλουμε

$$n^2(3p^4 - 6p^3 + 3p^2) + n(-6p^4 + 12p^3 - 7p^2 + p) - 3pn^2 \leq 0.$$

Αν τη σχέση αυτή τη διαιρέσουμε με n^2 θα έχουμε το ζητούμενο αποτέλεσμα για μεγάλα n, μια που ο όρος $3p^4 - 6p^3 + 3p^2 - 3p$ μπορεί να γίνει αρνητικός, για παράδειγμα όταν $p = 0.8$. □

Στη συνέχεια θα συμβολίζουμε τη σταθερή spline DMPLE που αντιστοιχεί στη λύση του παραπάνω προβλήματος με $p_{h_n}(x)$, για μια συγκεκριμένη τιμή του h_n.

• **Θεώρημα 3.1.1 :** Κάτω από τις υποθέσεις του Λήμματος 3.1.1 και με την επιπλέον υπόθεση ότι η f είναι γνήσια θετική στο (a,b), θα έχουμε:

$$\lim_{n\to\infty} p_{h_n}(x) = f(x) \quad a.s..$$

Απόδειξη: Αφού το $m = [cn^q], c > 0, 0 < q < \frac{1}{3}$, για το h_n θα έχουμε ότι $h_n = O(n^{-q})$. Με βάση την θεωρία των πολλαπλασιαστών Lagrange υπάρχουν $\lambda, \mu_0, ..., \mu_{m-1}$ τέτοια ώστε:

$$\frac{\partial L(p_{h_n})}{\partial p_i} + \lambda \frac{\partial}{\partial p_i}\left[\sum_{i=0}^{m-1} p_i - \frac{1}{h_n}\right] + \mu_i = 0,$$

με $\lambda\left(\sum_{i=0}^{m-1} p_i - \frac{1}{h_n}\right) = 0$ και $\mu_i p_i = 0$. Αρα, για $i = 0, ..., m-1$

$$\mu_i + \frac{v_i}{p_i} + \frac{2\alpha}{h_n}(p_{i-1} - 2p_i + p_{i+1}) + \lambda = 0, \tag{3.9}$$

όπου v_i είναι ο αριθμός των δειγμάτων στο διάστημα $[t_i, t_{i+1})$. Πολλαπλασιάζοντας με p_i και αθροίζοντας πάνω σε όλα τα i, θα έχουμε:

$$\sum_{i=0}^{m-1} v_i + \frac{2\alpha}{h_n}\sum_{i=0}^{m-1} p_i(p_{i-1} - 2p_i + p_{i+1}) + \lambda \sum_{i=0}^{m-1} p_i = 0,$$

μια που $\mu_i p_i = 0$ και άρα τελικά

$$n + \frac{2\alpha}{h_n} \sum_{i=0}^{m-1} p_i(p_{i-1} - 2p_i + p_{i+1}) + \frac{\lambda}{h_n} = 0. \tag{3.10}$$

Οπότε ο πολλαπλασιαστής Lagrange λ θα ισούται με

$$\lambda = -nh_n - 2\alpha \sum_{i=0}^{m-1} p_i(p_{i-1} - 2p_i + p_{i+1}).$$

Αν διαιρέσουμε την (3.9) με n και αντικαταστήσουμε με την τιμή που βρήκαμε για το λ, θα έχουμε:

$$\frac{\mu_i}{n} + \frac{v_i}{p_i n} + \frac{2\alpha}{nh_n}(p_{i-1} - 2p_i + p_{i+1}) - h_n - \frac{2\alpha}{n} \sum_{i=0}^{m-1} p_i(p_{i-1} - 2p_i + p_{i+1}) = 0. \tag{3.11}$$

Από τους περιορισμούς του προβλήματος, πρέπει $h_n \sum_{i=0}^{m-1} p_i = 1$ και $p_i \geq 0$, άρα $p_i \leq \frac{1}{h_n}$ οπότε:

$$p_{i-1} - 2p_i + p_{i+1} = O(\frac{1}{h_n}) \tag{3.12}$$

και

$$\sum_{i=0}^{m-1} p_i(p_{i-1} - 2p_i + p_{i+1}) = O(\frac{1}{h_n^3}), \tag{3.13}$$

μια που $m = \frac{b-a}{h_n}$. Αν διαιρέσουμε την (3.11) με h_n και χρησιμοποιήσουμε τις σχέσεις (3.12) και (3.13) θα έχουμε

$$\frac{\mu_i}{nh_n} + \frac{v_i}{p_i nh_n} - 1 + O(\frac{1}{nh_n^4}) = 0. \tag{3.14}$$

Η συνάρτηση f είναι γνήσια θετική, άρα από το Λήμμα 3.1.1 θα υπάρχει n_0 τέτοιο ώστε $v_i > 0$, $\forall n \geq n_0$. Αυτό σημαίνει ότι το αντίστοιχο p_i, θα είναι γνήσια θετικό, μια που αν αυτό το p_i ήταν ίσο με το μηδέν η log πιθανοφάνεια θα γινόταν μείον άπειρο και άρα δεν θα ήταν σημείο μεγίστου. Αρα $p_i > 0$, τότε όμως $\mu_i = 0$ και άρα

$$\frac{v_i}{p_i nh_n} - 1 + O(\frac{1}{nh_n^4}) = 0$$

για κάθε $n \geq n_0$. Πολλαπλασιάζοντας την παραπάνω σχέση με p_i, αφού $p_i > 0$, καταλήγουμε στην ακόλουθη σχέση

$$\frac{v_i}{nh_n} - p_i = O(\frac{1}{nh_n^5})$$

για κάθε $n \geq n_0$, μια που $p_i = O(\frac{1}{h_n})$

Στην παραπάνω σχέση βάζοντας όπου v_i το $v[x:n]$ και όπου p_i το $p_{h_n}(x)$, θα έχουμε ότι

$$\lim_{n\to\infty} p_{h_n}(x) = f(x) \quad a.s,$$

μια που από το Λήμμα 3.1.1

$$\frac{v[x:n]}{nh_n} \to f(x) \quad a.s.$$

και $nh_n^5 \to \infty$, καθώς το n τείνει στο άπειρο και έχουμε ολοκληρώσει την απόδειξη για την συνέπεια των DMPLE.□

Στην επόμενη παράγραφο, στόχος μας είναι να παρουσιάσουμε ορισμένα αριθμητικά αποτελέσματα σχετικά με τον υπολογισμό της γραμμικής spline DMPLE, όπου έχουμε χρησιμοποιήσει σαν όρο ποινής τη διακριτή δεύτερη παράγωγο.

3.2 Αριθμητικά αποτελέσματα

Σε αυτή την παράγραφο θα παρουσιάσουμε μια σειρά από αριθμητικά αποτελέσματα, τα οποία σχετίζονται με τον υπολογισμό της δεύτερης DMPLE η οποία προκύπτει από τη μεγιστοποίηση του ποινικοποιημένου συναρτησοειδούς του λογαρίθμου της πιθανοφάνεια, χρησιμοποιώντας τη διακριτή δεύτερη παράγωγο για όρο ποινής. Δεδομένου του τυχαίου δείγματος $X_1, X_2, ...X_n$, με συνάρτηση πυκνότητας f, στο διάστημα (a, b), μιας θετικής ποσότητας α και ενός θετικού ακεραίου m, ορίζουμε:

$$h = \frac{b-a}{m},$$

$$t_i = a + ih, \quad i = 0, 1, ..., m$$

και

$$s(t_0) = 0, \quad s(t_m) = 0.$$

Τότε όπως είδαμε και παραπάνω η γραμμική spline DMPLE είναι η λύση του ακόλουθου (m-1)-διάστατου προβλήματος βελτιστοποίησης

$$\max L(s_1, ..., s_{m-1}) = \max \left\{ \sum_{i=1}^{n} \log s(X_i) - \frac{\alpha}{h^3} \sum_{i=1}^{m-1} \{s_{i+1} - 2s_i + s_{i-1}\}^2 \right\} \tag{3.15}$$

υπό τους περιορισμούς:

$$h \sum_{i=1}^{m-1} s_i = 1, \;\; s_i \geq 0 \;\; \forall i = 1, ..., m-1.$$

Στην βιβλιοθήκη της IMSL [13], υπάρχει μια ρουτίνα η οποία λύνει το παραπάνω πρόβλημα βελτιστοποίησης. Τα αριθμητικά μας αποτελέσματα αναφέρονται στην Κανονική, τη Γάμμα και τη δίκορφη κανονική κατανομή και για να παράγουμε τυχαίους αριθμούς, χρησιμοποιούμε ρουτίνες από το βιβλίο *Numerical recipes in fortran* , [17]. Στο παράρτημα , παραθέτουμε τον κώδικα του προγράμματος που υπολογίζει την εκτιμήτρια της άγνωστης συνάρτησης πυκνότητας f, με τη μέθοδο της μεγιστοποίησης του ποινικοποιημένου συναρτησοειδούς του λογαρίθμου της πιθανοφάνειας, στην περίπτωση που σαν όρο ποινής θεωρήσουμε τη διακριτή δεύτερη παράγωγο

Στα σχήματα 3.1 έως 3.3 παρουσιάζουμε τα αποτελέσματα που αφορούν την τυπική κανονική κατανομη $N(0, 1)$, για 100,200 και 300 παρατηρήσεις με την τιμή της παραμέτρου α να είναι 3,3 και 7 αντίστοιχα. Τα αποτελέσματα για την κατανομή γάμμα $\Gamma(2, 1)$ τα παρουσιάζουμε στα σχήματα 3.4 έως 3.6 όπου έχουμε πάρει 100,200 και 300 παρατηρήσεις ενώ η τιμή της παραμέτρου α, είναι 7,7 και 13 αντίστοιχα. Στα σχήματα 3.7 έως 3.9 παρουσιάζουμε τα αποτελέσματα σχετικά με τη δίκορφη κανονική κατανομή $\frac{1}{2}N(0, 1) + \frac{1}{2}N(3, 1/4)$ για 100,200 και 300 παρατηρήσεις. Οι τιμές της παραμέτρου α, στην περίπτωση αυτή είναι 6,15 και 7 αντοίστοιχα.

Σχήμα 3.1. Η διακεκομμένη γραμμή παριστάνει μια τυπική Κανονική κατανομή, N(0,1) για 100 σημεία . Η συνεχής γραμμή παριστάνει την DMPLE με $\alpha = 3$.

Σχήμα 3.2. Η διακεκομμένη γραμμή παριστάνει μια τυπική Κανονική κατανομή, N(0,1) για 200 σημεία . Η συνεχής γραμμή παριστάνει την DMPLE με $\alpha = 3$.

Σχήμα 3.3. Η διακεκομμένη γραμμή παριστάνει μια τυπική Κανονική κατανομή, N(0,1) για 300 σημεία . Η συνεχής γραμμή παριστάνει την DMPLE με $\alpha = 5$.

Σχήμα 3.4. Η διακεκομμένη γραμμή παριστάνει μια τυπική $\Gamma(2,1)$ για 100 σημεία . Η συνεχής γραμμή παριστάνει την DMPLE με $\alpha = 7$.

Σχήμα 3.5. Η διακεκομμένη γραμμή παριστάνει μια Γάμμα κατανομή, $\Gamma(2,1)$ για 200 σημεία . Η συνεχής γραμμή παριστάνει την DMPLE με $\alpha = 7$.

Σχήμα 3.6. Η διακεκομμένη γραμμή παριστάνει μια Γάμμα κατανομή, $\Gamma(2,1)$ για 300 σημεία . Η συνεχής γραμμή παριστάνει την DMPLE με $\alpha = 13$.

Σχήμα 3.7. Η διακεκομένη γραμμή παριστάνει την δίκορφη κανονική κατανομή $\frac{1}{2}N(0,1)+\frac{1}{2}N(3,1/4)$ για 100 σημεία . Η συνεχής γραμμή παριστάνει την DMPLE με $\alpha = 6$.

Σχήμα 3.8. Η διακεκομένη γραμμή παριστάνει την δίκορφη κανονική κατανομή $\frac{1}{2}N(0,1)+\frac{1}{2}N(3,1/4)$ για 200 σημεία . Η συνεχής γραμμή παριστάνει την DMPLE με $\alpha = 15$.

Σχήμα 3.9. Η διακεκομένη γραμμή παριστάνει την δίκορφη κανονική κατανομή $\frac{1}{2}N(0,1)+\frac{1}{2}N(3,1/4)$ για 300 σημεία . Η συνεχής γραμμή παριστάνει την DMPLE με $\alpha = 7$.

Κεφάλαιο 4

L_2 συνέπεια της ρίζας της πυκνότητας

Εστω $X_1, X_2, ..., X_n$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με συνάρτηση κατανομής F και συνάρτηση πυκνότητας f, $v = f^{1/2}$, που ικανοποιεί τις ακόλουθες υποθέσεις:

A1. $v \in H^1(-\infty, +\infty)$,

A2. $E_f |X|^\tau < +\infty$, για $\tau > 1$.

Συμβολίζουμε με v_n, την MPLE της ρίζας της συνάρτησης πυκνότητας και με $f_n = v_n^2$, την πρώτη MPLE των Good και Gaskins. Το 1982 ο Κλωνιάς [7] έδειξε συνέπεια της f_n ως πρός τη μετρική Hellinger και ως πρός τις νόρμες $L_1, L_2, \sup_R$ και Sobolev. Εμείς εδώ θα περιοριστούμε στην απόδειξη της συνέπειας της f_n ως πρός την απόσταση Hellinger, η οποία αποτελεί το κλειδί για την απόδειξη των άλλων μορφών σύγκλισης. Αρχικά θα αναφέρουμε το Λήμμα 5.1, από το βιβλίο του Billingsley [12], το οποίο θα χρειαστούμε στην απόδειξη του Θεωρήματος 5.1 που δίνει την ταχύτητα σύγκλισης της ομοιόμορφης εμπειρικής συνάρτησης κατανομής ως πρός μια συγκεκριμένη μετρική. Το θεώρημα αυτό, από το άρθρο του Κλωνιά [7], είναι χρήσιμο για την απόδειξη της L_2 συνέπειας της v_n.

• **Λήμμα 4.1** : Εστω $Z_1, ...Z_n$ ανεξάρτητα και ισόνομα στοιχεία ενός μετρικού χώρου, με $\|X\|$ συμβολίζουμε τη απόσταση του X από το μηδενικό στοιχείο του χώρου. Θεωρούμε $S_n \equiv Z_1 + ... + Z_n$. Για όλα τα $\lambda > c$ και τους θετικούς ακέραιους $n \geq m$ για τους οποίους:

$$0 < \delta_{mn} \equiv \min_{m \leq i \leq n} P(\|S_n - S_i\| < c),$$

έχουμε:

$$P(\max_{m \leq i \leq n} \|S_i\| \geq \lambda) \leq P(\|S_n\| \geq \lambda - c)/\delta_{mn}.$$

Ορίζουμε τη μετρική:

$$\left\| \frac{f-g}{q} \right\| \equiv \sup_{t \in (0,1)} \frac{|f(t) - g(t)|}{q(t)},$$

με $q(t) \equiv \{t(1-t)\}^p, p \in (1/2, 1)$.

Η (4.1) που ακολουθεί για $s = 0$, έπεται από το Θεώρημα 4 του Wellner [14].

• **Θεώρημα 4.1 :** Εστω $\xi_1, ...\xi_n$ ανεξάρτητες και ισόνομες τυχαίες μεταβλητές $U(0,1)$. Τότε για s και p τέτοια ώστε $s > 0$, $1/2 < p < 1$ και $s + p < 1$, έχουμε:

$$\lim_{n\to\infty} n^s \left\| \frac{\Gamma_n - I}{\{I(1-I)\}^p} \right\| = 0, a.s. \tag{4.1}$$

όπου Γ_n είναι η εμπειρική συνάρτηση κατανομής των $\xi_1, ...\xi_n$ και I η ταυτοτική συνάρτηση.

Απόδειξη : Εστω $A_n \equiv \{n^s \sup_{0<t\leq\vartheta} |\Gamma_n(t) - t| / t^p \geq \varepsilon\}$ για δοσμένα $0 < \vartheta < 1$ και $\varepsilon > 0$. Ορίζουμε

$$Y_i \equiv h_\vartheta(\xi_i) - \int_0^{\xi_i} (1-t)^{-1} h_\vartheta(t) dt$$

με $h_\vartheta(t) = t^{-p}\mathbf{1}(t \in [0, \vartheta])$. Τότε οι Y_i είναι ανεξάρτητες και ισόνομες τυχαίες μεταβλητές με μέση τιμή μηδέν και $E|Y_1|^\tau < +\infty$ για κάθε $\tau < 1/p$. Χρησιμοποιώντας την τελευταία ανισότητα από τη σελίδα 484 του Wellner [14] (η ανιοότητα αυτή οφείλεται στους Bahr και Esseen [16]), για $c = \varepsilon n^{-s}$ και $\gamma = \tau$, παίρνουμε ότι

$$P(A_n) \leq 2^{\tau+1}\varepsilon^{-\tau} E|Y_1|^\tau n^{-(\tau - s\tau - 1)} \tag{4.2}$$

για κάθε τ τέτοιο ώστε $1 < \tau < 1/p < 2$. Αν $s < 1 - \tau^{-1}$, τότε

$$\sum_{n=1}^{\infty} P(A_{[n^\eta]}) < +\infty, \tag{4.3}$$

αν $\eta(\tau - s\tau - 1) > 1$. Στη συνέχεια δεχόμαστε ότι $\eta > (\tau - s\tau - 1)^{-1}$ και ακόμα ότι $\eta \geq 1$. Εστω k τέτοιο ώστε: $n^\eta < k \leq (n+1)^\eta$. Ορίζουμε

$$S_k \equiv \sum_{i=1}^{k} Q_i$$

όπου $Q_i(t) \equiv \mathbf{1}(\xi_i \in [0,t]) - t$ και

$$\left\| \frac{z}{q} \right\|_0^\vartheta \equiv \sup_{0<t\leq\vartheta} |z(t)| / q(t).$$

με $q(t) = t^p$. Τότε

$$\begin{aligned}
\max_{n^\eta < k \leq (n+1)^\eta} \left\| \frac{\Gamma_k - I}{q} \right\|_0^\vartheta &= \max_k \left\{ \sup_{0<t\leq\theta} \frac{|\Gamma_k(t) - I(t)|}{q(t)} \right\} \\
&= \max_k \left\{ \sup_{0<t\leq\theta} \frac{\left|\frac{1}{k}\sum_{i=1}^k 1(\xi_i \leq t) - t\right|}{q(t)} \right\} \\
&= \max_k \left\{ \sup_{0<t\leq\theta} \frac{\left|\sum_{i=1}^k [1(\xi_i \leq t) - kt]\right|}{kq(t)} \right\} \qquad (4.4) \\
&= \max_k \left\{ \sup_{0<t\leq\theta} \frac{|S_k|}{kq(t)} \right\} = \max_k \left\{ \left\| \frac{S_k}{q} \right\|_0^\vartheta \frac{1}{k} \right\} \\
&< \max_k \left\{ \frac{1}{n^\eta} \left\| \frac{S_k}{q} \right\|_0^\vartheta \right\}.
\end{aligned}$$

Αρα από το Λήμμα 4.1,

$$\begin{aligned} P\left(\max_k n^s \left\|\frac{\Gamma_k - I}{q}\right\|_0^{\vartheta} > 2\lambda\right) &\leq P\left(\max_k n^s n^{-\eta} \left\|\frac{S_k}{q}\right\|_0^{\vartheta} > 2\lambda\right) \\ &= P\left(\max_k \left\|\frac{S_k}{q}\right\|_0^{\vartheta} \geq 2\lambda n^{\eta-s}\right) \\ &\leq P\left(\left\|\frac{S_{[(n+1)^\eta]}}{q}\right\|_0^{\vartheta} \geq 2\lambda n^{\eta-s} - \lambda n^{\eta-s}\right) / c_{\lambda,n} \\ &= P\left(\left\|\frac{S_{[(n+1)^\eta]}}{q}\right\|_0^{\vartheta} \geq \lambda n^{\eta-s}\right) / c_{\lambda,n}, \end{aligned} \tag{4.5}$$

με

$$c_{\lambda,n} = \min_k P\left(\left\|(S_{[(n+1)^\eta]} - S_k)/q\right\|_0^{\vartheta} < \lambda n^{\eta-s}\right).$$

Ομως

$$\left\|\frac{S_k}{q}\right\|_0^{\vartheta} = k\left\|\frac{\Gamma_k - I}{q}\right\|_0^{\vartheta} \Rightarrow \left\|\frac{S_{[(n+1)^\eta]}}{q}\right\|_0^{\vartheta} = [(n+1)^\eta]\left\|\frac{\Gamma_{[(n+1)^\eta]} - I}{q}\right\|_0^{\vartheta}.$$

Αρα η σχέση (4.5) γίνεται

$$\begin{aligned} P\left(\max_k n^s \left\|\frac{\Gamma_k - I}{q}\right\|_0^{\vartheta} > 2\lambda\right) &\leq P\left([(n+1)^\eta]\left\|\frac{\Gamma_{[(n+1)^\eta]} - I}{q}\right\|_0^{\vartheta} \geq \lambda n^{\eta-s}\right) / c_{\lambda,n} \\ &\leq P\left([(n+1)^\eta]^s\left\|\frac{\Gamma_{[(n+1)^\eta]} - I}{q}\right\|_0^{\vartheta} \geq \lambda 2^{-\eta(1-s)}\right) / c_{\lambda,n} \\ &= P\left(A_{[(n+1)^\eta]}\right) / c_{\lambda,n} \end{aligned} \tag{4.6}$$

αρκεί για την τελευταία ισότητα να πάρουμε για $\lambda = 2^{\eta(1-s)}\varepsilon$ και για την τελευταία ανισότητα να παρατηρήσουμε ότι

$$\frac{\lambda n^{\eta-s}}{[(n+1)^\eta]^{1-s}} \geq \frac{\lambda n^{s(\eta-1)}}{(1+\frac{1}{n})^{\eta(1-s)}} \geq \lambda 2^{-\eta(1-s)},$$

αφού $n \geq 1$ και άρα $\frac{1}{n} + 1 \leq 2 \Rightarrow (1+\frac{1}{n})^{\eta(1-s)} \leq 2^{\eta(1-s)}$. Αθροίζοντας πάνω σε όλα τα n τη σχέση (4.6) παίρνουμε

$$\sum_n P\left(\max_{n^\eta < k \leq (n+1)^\eta} n^s \left\|\frac{\Gamma_k - I}{q}\right\|_0^{\vartheta} \geq 2\lambda\right) \leq \sum_n P(A_{[(n+1)^\eta]})/c_{\lambda,n}. \tag{4.7}$$

Θέλουμε τώρα να βρούμε, για ποιά τιμή του η, το $c_{\lambda,n}$ είναι φραγμένο μακριά από το μηδέν, ώστε μέσω της σχέσης (4.3) να έχουμε τελικά ότι η σειρά

$$\sum_n P(A_{[(n+1)^\eta]})/c_{\lambda,n}$$

συγκλίνει. Εχουμε λοιπόν

$$\begin{aligned} 1 - c_{\lambda,n} &= 1 - \min_k P\left(\left\|\frac{S_{[(n+1)^\eta]} - S_k}{q}\right\|_0^{\vartheta} < \lambda n^{\eta-s}\right) \\ &= \max_k P\left(\left\|\frac{S_{[(n+1)^\eta]} - S_k}{q}\right\|_0^{\vartheta} \geq \lambda n^{\eta-s}\right). \end{aligned} \tag{4.8}$$

Ομως $S_{[(n+1)^\eta]}(t) - S_k(t)$ είναι ισόνομο με το

$$\sum_{i=1}^{[(n+1)^\eta]-k} \mathbf{1}(\xi_i \leq t) - t\{[(n+1)^\eta] - k\} = \{[(n+1)^\eta] - k\}\{\Gamma_{[(n+1)^\eta]-k}(t) - t\}.$$

Αρα

$$\begin{aligned} 1 - c_{\lambda,n} &= \max_k P\left(\left\| \frac{\{[(n+1)^\eta] - k\}\{\Gamma_{[(n+1)^\eta]-k} - I\}}{q} \right\|_0^\vartheta \geq \lambda n^{\eta - s} \right) \\ &= \max_k P\left(\left\| \frac{\Gamma_{[(n+1)^\eta]-k} - I}{q} \right\|_0^\vartheta \geq \frac{\lambda n^{\eta-s}}{[(n+1)^\eta] - k} \right) \\ &= \max_k P\left(\left\| \frac{\Gamma_{[(n+1)^\eta]-k} - I}{q} \right\|_0^\vartheta \geq \frac{2^{\eta(1-s)}\varepsilon n^{\eta-s}}{[(n+1)^\eta] - k} \right), \end{aligned} \tag{4.9}$$

όπου

$$\begin{aligned} & P\left(\left\| \frac{\Gamma_{[(n+1)^\eta]-k} - I}{q} \right\|_0^\vartheta \geq \frac{2^{\eta(1-s)}\varepsilon n^{\eta-s}}{[(n+1)^\eta] - k} \right) \\ = \ & P\left(\{[(n+1)^\eta] - k\}^s \sup_{0 < t \leq \vartheta} \left| \frac{\Gamma_{[(n+1)^\eta]-k}(t) - t}{t^p} \right| \geq \frac{2^{\eta(1-s)}\varepsilon n^{\eta-s}}{\{[(n+1)^\eta] - k\}^{1-s}} \right) \\ = \ & P(A_{[(n+1)^\eta]-k}), \end{aligned} \tag{4.10}$$

όπου τώρα σαν ε στον ορισμό του A_n έχουμε πάρει το $\frac{2^{\eta(1-s)}\varepsilon n^{\eta-s}}{\{[(1+n)^\eta]-k\}^{1-s}}$. Συνεπώς η σχέση (4.9) μέσω των σχέσεων (4.10) και (4.2) γίνεται:

$$\begin{aligned} 1 - c_{\lambda,n} &= \max_k P(A_{[(n+1)^\eta]-k}) \\ &\leq \max_k 2^{\tau+1}\left(\frac{2^{\eta(1-s)}\varepsilon n^{\eta-s}}{\{[(n+1)^\eta] - k\}^{1-s}}\right)^{-\tau} E\,|Y_1|^\tau \{[(n+1)^\eta] - k\}^{-(\tau-1-s\tau)} \\ &= \max_k 2^{\tau+1-\eta\tau(1-s)}\varepsilon^{-\tau} n^{-\tau(\eta-s)} E\,|Y_1|^\tau \{[(n+1)^\eta] - k\} \\ &= 2^{\tau+1-\eta\tau(1-s)}\varepsilon^{-\tau} n^{-\tau(\eta-s)} E\,|Y_1|^\tau \{[(n+1)^\eta] - [n^\eta] - 1\} \\ &< 2^{\tau+1-\eta\tau(1-s)}\varepsilon^{-\tau} E\,|Y_1|^\tau \{(n+1)^\eta - n^\eta\}/n^{\tau(\eta-s)} \\ &< 2^{\tau+1-\eta\tau(1-s)}\varepsilon^{-\tau} E\,|Y_1|^\tau\, \eta(n+1)^{\eta-1}/n^{\tau(\eta-s)}. \end{aligned} \tag{4.11}$$

Από τη σχέση (4.11) βλέπουμε ότι, για να είναι το $c_{\lambda,n}$ φραγμένο μακριά από το μηδέν θα πρέπει ο τελευταίος όρος στην παραπάνω σχέση να είναι φραγμένος από ένα αριθμό γνήσια μικρότερος της μονάδας για μεγάλα n, άρα θα πρέπει $\tau(\eta - s) > \eta - 1$ ή ισοδύναμα

$$\eta > \frac{s\tau - 1}{\tau - 1}.$$

Αυτό όμως ισχύει μια που $\eta \geq 1 > \frac{s\tau-1}{\tau-1}$. Πράγματι εξ' υποθέσεως $s < 1$, συνεπώς $s\tau < \tau \Rightarrow s\tau - 1 < \tau - 1 \Rightarrow \frac{s\tau-1}{\tau-1} < 1$. Αρα, για $\eta > \max\{1, \frac{1}{\tau-1-s\tau}\}$ έχουμε ότι

$$\sum_n P(A_{[(n+1)^\eta]})/c_{\lambda,n} < +\infty$$

και άρα από την σχέση (4.7)

$$\sum_n P\left(\max_k n^s \left\|\frac{\Gamma_k - I}{q}\right\|_0^{\vartheta} \geq 2\lambda\right) < +\infty.$$

Συνεπώς άπο το Λήμμα Borel Cantelli [12],

$$n^s \sup_{0<t\leq\vartheta} \frac{|\Gamma_n(t) - t|}{t^p} \to 0, a.s \tag{4.12}$$

καθώς $n \to +\infty$. Ομως

$$n^s \sup_{0<t<\vartheta} \frac{|\Gamma_n(t) - t|}{t^p(1-t)^p} \leq n^s \sup_{0<t<\vartheta} \frac{|\Gamma_n(t) - t|}{t^p} + n^s \sup_{0<t<\vartheta} \frac{|\Gamma_n(t) - t|}{(1-t)^p}.$$

Στην παραπάνω σχέση και οι δύο τελευταίοι όροι τείνουν στο μηδέν σχεδόν βεβαίως καθώς το $n \to +\infty$. Πράγματι για τον πρώτο όρο μόλις το αποδείξαμε και το όριο για τον δεύτερο όρο δίνεται από το Θεώρημα B στην σελίδα 62 του Serfling [15] (για το ορισμό του D_n βλέπε σελίδα 59 του ιδίου), μια που $(1-t)^{-1} < (1-\vartheta)^{-1}$. Για το διάστημα $[\vartheta, 1]$ θα έχουμε και πάλι ότι

$$n^s \sup_{\vartheta<t<1} \frac{|\Gamma_n(t) - t|}{t^p(1-t)^p} \leq n^s \sup_{\vartheta<t<1} \frac{|\Gamma_n(t) - t|}{t^p} + n^s \sup_{\vartheta<t<1} \frac{|\Gamma_n(t) - t|}{(1-t)^p}$$

όπου στην περίπτωση αυτή, το όριο του πρώτου όρου δίνεται και πάλι από τον Serfling [15] μια που $t^{-1} < \vartheta^{-1}$, ενώ το όριο του δεύτερου προκύπτει από τη σχέση (4.12). □

Για να δείξουμε την L_2 συνέπεια της ρίζας της συνάρτησης πυκνότητας είναι απαραίτητο να αφήσουμε την παράμετρο α του όρου ποινής να εξαρτάται από το μέγεθος του δείγματος, δηλαδή $\alpha = \alpha_n$, με τέτοιο τρόπο ώστε: $\alpha_n \to +\infty$ και $\frac{\alpha_n}{n} \to 0$, κάθως το $n \to +\infty$. Επιλέγουμε λοιπόν $\alpha_n = c_n n^t, t \in (0,1)$ και $c_n > 0$ τέτοιο ώστε $c_n \to c > 0$ κάθως το $n \to +\infty$. Γράφουμε με λ_n τον πολλαπλασιαστή Lagrange που αντιστοιχεί στον περιορισμό ότι η L_2 νόρμα του v είναι ίση με τη μονάδα, για τον οποίο ισχύει ότι $\lambda_n < n$. Πράγματι, στη Πρόταση 2.1.1 ορίσαμε τη νόρμα

$$\|u\|_\lambda^2 = 2\alpha \left\|u'\right\|_2^2 + \lambda \|u\|_2^2$$

και δείξαμε στη σχέση (2.22) ότι $\|u_\lambda\|_\lambda^2 = \frac{n}{2}$. Στην Πρόταση 2.1.2 είδαμε ότι για κατάλληλη επιλογή του λ, μπορούμε να έχουμε ότι $\|u\|_2^2 > 1$. Συνεπώς $\lambda_n < n$.

Το παρακάτω Λήμμα δίνει ένα άνω φράγμα για την απόσταση Hellinger μεταξύ δύο συναρτήσεων πυκνότητας, το οποίο χρησιμοποιούμε κυρίως στο Θεώρημα 4.2.

- **Λήμμα 4.2 :** Αν f και g είναι πυκνότητες, τότε

$$\left\|f^{1/2} - g^{1/2}\right\|_2^2 \leq \int f \log(f/g).$$

Απόδειξη : Εστω f και g συναρτήσεις πυκνότητας. Θυμίζουμε ότι, $1 - \frac{1}{x} \leq \log x$, $\forall x > 0$. Αρα

$$\int f \log\left(\frac{f}{g}\right) = \int f \log f - \int f \log g = 2\int f \log f^{1/2} - 2\int f \log g^{1/2}$$

$$
\begin{aligned}
&= 2\int f \log\left(\frac{f^{1/2}}{g^{1/2}}\right) \geq 2\int f\left(1-\frac{g^{1/2}}{f^{1/2}}\right)\\
&= 2\int f - 2\int f^{1/2}g^{1/2} = 2 - 2\int f^{1/2}g^{1/2}\\
&= \int f + \int g - 2\int f^{1/2}g^{1/2} = \int (f^{1/2}-g^{1/2})^2\\
&= \left\|f^{1/2}-g^{1/2}\right\|_2^2 . \square
\end{aligned}
$$

• **Λήμμα 4.3 :** Κάτω από την υπόθεση Α1 έχουμε:

$$
\begin{aligned}
&(\alpha_n/n)(\|v_n'\|_2^2 - \|v'\|_2^2) + 2\int \log v d(F_n - F)\\
\leq\ & 2\int \log v_n dF_n - 2\int \log v dF \qquad (4.13)\\
\leq\ & 2\int \log v_n d(F_n - F).
\end{aligned}
$$

Απόδειξη : Στην απόδειξη θα στηριχτούμε στο ότι το v_n είναι η MPLE της ρίζας της συνάρτησης πυκνότητας και στο Λήμμα 4.2. Εχουμε λοιπόν,

$$
\begin{aligned}
&2\int \log v d(F_n - F) + 2\int \log v dF - (\alpha_n/n)\left\|v'\right\|_2^2\\
=\ & \left(2\sum_{i=1}^{n} \log v(X_i) - \alpha_n \left\|v'\right\|_2^2\right)/n\\
\leq\ & \left(2\sum_{i=1}^{n} \log v_n(X_i) - \alpha_n \left\|v_n'\right\|_2^2\right)/n\\
=\ & 2\int \log v_n dF + 2\int \log v_n d(F_n - F) - (\alpha_n/n)\left\|v_n'\right\|_2^2\\
\leq\ & 2\int \log v dF + 2\int \log v_n d(F_n - F) - (\alpha_n/n)\left\|v_n'\right\|_2^2 .
\end{aligned}
$$

Στην τελευταία ανισότητα χρησιμοποιήσαμε ότι

$$
2\int \log v_n dF \leq 2\int \log v dF,
$$

το οποίο, για παράδειγμα, είναι άμεσο από το Λήμμα 5.2. Πράγματι

$$
2\int \log v dF - 2\int \log v_n dF = \int f \log\left(\frac{v}{v_n}\right)^2 = \int f \log\left(\frac{f}{f_n}\right) \geq 0
$$

$\square$

Στην απόδειξη για τη σύγκλιση της συνάρτησης πυκνότητας ως πρός την απόσταση Hellinger, θα εμφανιστούν τα ακόλουθα δύο συναρτησοειδή $\int \log v d(F_n - F)$ και $\int \log v_n d(F_n - F)$, που όπως θα δούμε στη συνέχεια συγκλίνουν στο μηδέν σχεδόν βεβαίως καθώς το n τείνει στο άπειρο. Για να το δούμε αυτό θα χρειαστούμε ένα φράγμα για το δεύτερο συναρτησοειδές που δίνεται από το ακόλουθο Λήμμα.

• **Λήμμα 4.4 :** Αν $E_f\,|X| < +\infty$, τότε

$$\left|\int \log v_n d(F_n - F)\right| \leq (\lambda_n/\alpha_n)^{1/2}\,\|(\Gamma_n - I)/h\| \int h \circ F, \tag{4.14}$$

όπου $h \geq 0$ με $\int h \circ F < +\infty$ και Γ_n είναι η ομοιόμορφη εμπειρική συνάρτηση κατανομής των $F(X_1), ..., F(X_n)$.

Απόδειξη : Πριν ξεκινήσουμε θα ήταν χρήσιμο να αναφέρουμε την γνωστή Πρόταση από την Θεωρία Πιθανοτήτων, σύμφωνα με την οποία αν η τυχαία μεταβλητή X έχει συνάρτηση κατανομής F συνεχή, τότε η τυχαία μεταβλητή $Y \doteq F(X)$ ακολουθεί μια ομοιόμορφη κατανομή $U(0,1)$. Από τη σχέση (2.21) του Κεφαλαίου 2 (με την μόνη διαφορά ότι αντί για 2α βάζουμε α) θα έχουμε για $x > X_{(n)} \equiv \max(X_1, ..., X_n)$ ότι

$$\log v_n(x) = \log\left\{\left(\frac{1}{4\lambda_n\alpha_n}\right)^{1/2} \sum_{i=1}^{n} \frac{\exp\{(\lambda_n/\alpha_n)^{1/2}X_i\}}{v_n(X_i)}\right\} - (\lambda_n/\alpha_n)^{1/2}x \tag{4.15}$$

με $F_n(x) \doteq \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(X_i \leq x) = 1$, ενώ για $x < X_{(1)} \equiv (X_1, ..., X_n)$

$$\log v_n(x) = \log\left\{\left(\frac{1}{4\lambda_n\alpha_n}\right)^{1/2} \sum_{i=1}^{n} \frac{\exp\{-(\lambda_n/\alpha_n)^{1/2}X_i\}}{v_n(X_i)}\right\} + (\lambda_n/\alpha_n)^{1/2}x, \tag{4.16}$$

με $F_n(x) \doteq \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(X_i \leq x) = 0$. Δηλαδή ο $\log v_n(x)$, είναι μια γραμμική συνάρτηση του x, για $x > X_{(n)}$ και για $x < X_{(1)}$. Ισχυρίζομαι ότι:

$$\lim_{x\to\pm\infty}[\{F_n(x) - F(x)\}\log v_n(x)] = 0.$$

Πράγματι, παρατηρούμε ότι $F_n(x) - F(x) = 1 - F(x)$ για $x > \mathrm{X}_{(n)}$. Αρα αρκεί να δείξουμε ότι: $\lim_{x\to+\infty}(1 - F(x))\log v_n(x) = 0$. Αφού ο $\log v_n(x)$ είναι της μορφής $ax + \beta$, αρκεί να δούμε ότι

$$\lim_{x\to+\infty}(1 - F(x))\,(ax + \beta) = 0.$$

Παρατηρούμε ότι $\lim\limits_{x\to+\infty}\beta\,(1 - F(x)) = 0$, μια που $\lim\limits_{x\to+\infty}F(x) = 1$. Για τον άλλο όρο έχουμε ότι:

$$\begin{aligned}\lim_{x\to+\infty} a\,(1 - F(x))\,x &= \lim_{x\to+\infty} ax\int_x^{+\infty} f(t)dt \\ &\leq \lim_{x\to+\infty} a\int_x^{+\infty} tf(t)dt = 0,\end{aligned}$$

μια που $E_f\,|X| < +\infty$. Ανάλογα δουλεύουμε στην περίπτωση που $x \to -\infty$. Αρα θα έχουμε

$$\begin{aligned}\left|\int \log v_n d(F_n - F)\right| &= \left|\{F_n(x) - F(x)\}\log v_n(x)\,|_{-\infty}^{+\infty} - \int (F_n - F)\frac{v_n'}{v_n}\right| \\ &= \left|-\int (F_n - F)\frac{v_n'}{v_n}\right| \\ &\leq \int |F_n - F|\left|\frac{v_n'}{v_n}\right|.\end{aligned} \tag{4.17}$$

Από τη σχέση (2.25) του Κεφαλαίου 2 σύμφωνα με την οποία

$$(v_n')^2 \leq \frac{\lambda_n}{\alpha_n}v_n^2,$$

η σχέση (4.17) θα γίνει

$$\left|\int \log v_n d(F_n - F)\right| \le (\frac{\lambda_n}{\alpha_n})^{1/2}\int |F_n - F|. \tag{4.18}$$

Είναι εύκολο να δούμε ότι

$$\int |F_n - F| = \int \frac{|\Gamma_n \circ F - F|}{h \circ F} h \circ F, \quad a.s$$

μια που

$$\Gamma_n \circ F(x) = \frac{1}{n}\sum_{i=1}^{n} 1(F(X_i) \le F(x))$$

και

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(X_i \le x) = \frac{1}{n}\sum_{i=1}^{n} 1(F(X_i) \le F(x)) \ a.s.$$

Αρα από τη σχέση (4.18) έχουμε ότι:

$$\begin{aligned}
\left|\int \log v_n d(F_n - F)\right| &\le (\frac{\lambda_n}{\alpha_n})^{1/2}\int \frac{|\Gamma_n(F(t)) - F(t)|}{h(F(t))} h(F(t))dt \\
&= (\frac{\lambda_n}{\alpha_n})^{1/2}\int_0^1 \frac{|\Gamma_n(x) - x|}{h(x)} h(x) dF^{-1}(x) \\
&\le (\frac{\lambda_n}{\alpha_n})^{1/2}\left\|\frac{\Gamma_n - I}{h}\right\|\int_0^1 \frac{h(x)}{f(F^{-1}(x))}dx \\
&= (\frac{\lambda_n}{\alpha_n})^{1/2}\left\|\frac{\Gamma_n - I}{h}\right\|\int_0^1 h(x) dF^{-1}(x)dx \\
&= (\frac{\lambda_n}{\alpha_n})^{1/2}\|(\Gamma_n - I)/h\|\int h \circ F. \square
\end{aligned}$$

• **Παρατήρηση 4.1** : Στην συνέχεια, χρησιμοποιούμε για συνάρτηση h την $(t(1-t))^p$, με $t \in (0,1)$ και $p \in (1/2, 1)$ και υποθέτουμε ότι:

$$\int \{F(1-F)\}^p < +\infty$$

ή την επαρκή της συνθήκη $E_f |X|^\tau < +\infty$ για κάποιο $\tau > 1/p > 1$, που είναι η υπόθεση A2. Ισχυριζόμαστε ότι αν $E_f |X|^\tau < +\infty$, για κάποιο $\tau > 1/p$, τότε $\int\{F(1-F)\}^p < +\infty$. Εστω λοιπόν ότι $E_f |X|^\tau < +\infty$, για κάποιο $\tau > 1/p$. Τότε: $|x|^\tau F(x)(1-F(x)) \to 0$, όταν $|x| \to +\infty$. Πράγματι, έστω ότι $x \to +\infty$. Τότε

$$\begin{aligned}
|x|^\tau \left(\int_{-\infty}^{x} f(t)dt\right)\left(\int_{x}^{+\infty} f(t)dt\right) &\le \left(\int_{-\infty}^{x} f(t)dt\right)\left(\int_{x}^{+\infty} |t|^\tau f(t)dt\right) \\
&\le \left(\int_{-\infty}^{+\infty} f(t)dt\right)\left(\int_{x}^{+\infty} |t|^\tau f(t)dt\right) \to 0,
\end{aligned}$$

καθώς $x \to +\infty$. Ανάλογα όταν $x \to -\infty$. Αρα υπάρχουν θετικές σταθερές c, m τέτοιες ώστε $|x|^\tau F(x)(1-F(x)) < c$ για κάθε $x \in M \equiv \{x \in R : |x| > m\}$. Συνεπώς

$$\begin{aligned}
\int_{-\infty}^{+\infty} (F(x)(1-F(x)))^p dx &= \int_{M^c} (F(x)(1-F(x)))^p dx + \int_M (F(x)(1-F(x)))^p dx \\
&\le \int_{M^c} (F(x)(1-F(x)))^p dx + c^p \int_M |x|^{-\tau p} dx < +\infty,
\end{aligned}$$

μια που $\tau > 1/p$.□

Το όριο των δύο συναρτησοειδών που είδαμε παραπάνω δίνεται από την ακόλουθη Πρόταση 4.1.

• **Πρόταση 4.1 :** Κάτω από τις υποθέσεις A1 και A2 και αν τα α_n επιλεγούν ώστε $\alpha_n \to +\infty$ και $\frac{\alpha_n}{n} \to 0, n \to +\infty$ έχουμε:

1. $n^d \int \log v d(F_n - F) \to 0$, a.s καθώς το $n \to +\infty$, για κάθε d τέτοιο ώστε $0 \leq d < 1/2$,
2. $n^d \int \log v_n d(F_n - F) \to 0$, a.s καθώς το $n \to +\infty$, για κάθε d τέτοιο ώστε
$$d < (t/2) - (2(\min(\tau, 2))^{-1} - 1)/2, \quad t \in (0,1), \quad \tau > 1.$$

Απόδειξη :

1. Παρατηρούμε ότι
$$\begin{aligned} 2\int \log v d(F_n - F) &= \int \log f dF_n - \int \log f dF \\ &= \frac{1}{n}\sum_{i=1}^{n}\left(\log f(X_i) - \int \log f dF\right). \end{aligned} \tag{4.19}$$
Το αποτέλεσμα τώρα έπεται από το νόμο των διαδοχικών λογαρίθμων (βλέπε [12]), μια που από την υπόθεση A2 έχουμε ότι $\int f \left|\log f\right|^2 < +\infty$ (βλέπε [9],[10]).

2. Από το Λήμμα 4.4 έχουμε ότι
$$n^d \left|\int \log v_n d(F_n - F)\right| \leq n^d \left(\frac{\lambda_n}{\alpha_n}\right)^{1/2} \left\|\frac{\Gamma_n - I}{h}\right\| \int h \circ F.$$
Ομως $\lambda_n < n$, άρα
$$n^d \left|\int \log v_n d(F_n - F)\right| < n^d (n/\alpha_n)^{1/2} \left\|\frac{\Gamma_n - I}{h}\right\| \int h \circ F.$$
Αν πάρουμε για $h(t) = \{t(1-t)\}^p$, με $p > 1/\tau$ η παραπάνω σχέση θα γίνει
$$n^d \left|\int \log v_n d(F_n - F)\right| < n^d (n/\alpha_n)^{1/2} \left\|\frac{\Gamma_n - I}{\{I(1-I)\}^p}\right\| \int \{F(1-F)\}^p.$$
Από το Θεώρημα 4.1 για $s = d + (1-t)/2$ θα έχουμε ότι
$$n^s \left\|\frac{\Gamma_n - I}{\{I(1-I)\}^p}\right\| \to 0 \;\; a.s, \quad , n \to +\infty.$$
για p κοντά στο $1/\tau$ ώστε $p + s < 1$. Συνεπώς θα έχουμε το ζητούμενο αποτέλεσμα μια που από την Παρατήρηση 4.1 $\int \{F(1-F)\}^p < +\infty$.□

Είμαστε τώρα σε θέση να δούμε την L_2 συνέπεια του $v_n = f_n^{1/2}$, δηλαδή την Hellinger συνέπεια της f_n.

• **Θεώρημα 4.2 :** Κάτω από τις υποθέσεις A1 και A2, που διατυπώσαμε στην αρχή του Κεφαλαίου και για $d < (1-t)/2$, $d < t/4 - (2(\tau \wedge 2)^{-1} - 1)/4$ έχουμε ότι:

$$\lim_{n\to\infty} n^d \|v_n - v\|_2 = 0, a.s.$$

Απόδειξη : Αρχικά από το Λήμμα 4.2 έχουμε:

$$\begin{aligned}
\|v_n - v\|_2^2 &= \left\| f_n^{1/2} - f^{1/2} \right\|_2^2 \leq \int f \log \frac{f}{f_n} \\
&= 2\int f \log \frac{v}{v_n} = 2\int \log \frac{v}{v_n} dF \\
&= 2\int \log v dF - 2\int \log v_n dF \\
&= 2\int \log v_n d(F_n - F) - 2\int \log v_n dF_n + 2\int \log v dF \\
&\leq 2\int \log v_n d(F_n - F) - 2\int \log v d(F_n - F) - (\alpha_n/n)(\|v_n'\|_2^2 - \|v'\|_2^2) \\
&\leq 2\int \log v_n d(F_n - F) - 2\int \log v d(F_n - F) + (\alpha_n/n)\|v'\|_2^2, \qquad (4.20)
\end{aligned}$$

όπου η προτελευταία ανισότητα έπεται από την πρώτη ανισότητα του Λήμματος 4.3 Μένει να δούμε ότι κάθε ένας από τους όρους αυτούς τείνει στο μηδέν αν τον πολλαπλασιάσουμε με n^d, καθώς το n τείνει στο άπειρο. Είναι εύκολο να δούμε ότι $n^d\alpha_n/n \to 0$ καθώς το n τείνει στο άπειρο. Πράγματι $\alpha_n = c_n n^t$ με $t \in (0,1)$ και $c_n \to c$, άρα $n^d\alpha_n/n = n^d c_n n^t/n = c_n n^{d+t-1}$ οπου $d + t - 1 < 0$, αφού $d < \frac{1-t}{2} < 1 - t$. Το όριο των δύο άλλων όρων δίνεται από την Πρόταση 4.1, μια που το d, από τις υποθέσεις του θεωρήματος, ικανοποιεί τις συνθήκες της πρότασης αυτής. □

Κεφάλαιο 5

Παράρτημα

Στο παράρτημα αυτό δίνουμε τον κώδικα για το πρόγραμμα που υπολογίζει την γραμμική spline DMPLE, που παρουσιάστηκε στο κεφάλαιο 3.

```
c========================================================================
c This program performs nonparametric probability density
c function estimation by the Penalized Likelihood Method.
c
c Mania Kokologiannaki
c last update: 12/11/97
c========================================================================
INTEGER INIT, MAXIT, NOBS, NODE, HALF
INTEGER NMISS, NOUT
INTEGER N,IA,idum
REAL*8 ALPHA,EPS
REAL*8 PP,XM,ran1,gasdev
REAL*8 gamdev,poidev,bnldev,rr
REAL*8 s1,s2,hh
PARAMETER (EPS=0.0001,INIT=0,MAXIT=30)
PARAMETER (NODE=231)
PARAMETER (NOBS=300)
PARAMETER (HALF=50) c————————————————————————————
c EPS = Convergence criterion
c INIT = Initialization option
```

```
c MAXIT = Maximum number of iterations
c NODE = Number of mesh nodes for the discrete probability
c density estimate
c NOBS = Numbers of observation
c————————————————————————————————
REAL*8 BNDS(2),DENS(NODE),STAT(4),X(NOBS)
EXTERNAL DESPL, UMACH
print*,'dwse to alpha'
read*, ALPHA
print*,'dwse to diastima [a,b]'
read*,s1,s2
bnds(1)=s1
bnds(2)=s2*0.5d0+3.d0
hh=(bnds(2)-bnds(1))/dfloat(NODE-1)
print*,'dwse to arxiko seed (arnitikos akeraios)'
read*,idum
print*,'gia Gaussian dwse 1, gia Gamma dwse 2'
read*, ii
i=0
if(ii.eq.1)then
21 continue
rr=gasdev(idum)
if(rr.gt.bnds(1).and.rr.lt.bnds(2))then
i = i+1
X(i) = rr
if(i.gt.NOBS)goto 44
endif
goto 21
elseif(ii.eq.2)then
print*,'dwse to ia gia thn Gamma'
read*,ia
22 continue
```

```
rr=gamdev(ia,idum)
if(rr.gt.bnds(1).and.rr.lt.bnds(2))then
i = i+1
X(i) = rr
if(i.gt.NOBS)goto 44
endif
goto 22
endif
44 continue
do i=HALF+1,NOBS
x(i)=0.5d0*x(i)+3.d0
enddo
do i=1,NOBS
write(33,*)x(i)
enddo
CALL DESPL(NOBS,X,NODE,BNDS,INIT,ALPHA,MAXIT,EPS,DENS,
STAT,NMISS)
CALL UMACH (2, NOUT)
do i=0,NODE-1
qw = bnds(1) +dfloat(i)*hh
write(11,*)qw,dens(i+1)
enddo
write(12,*)'LOG-LIKELIHOOD=',stat(1)
write(12,*)'PENALTY=',stat(2)
write(12,*)'MEAN=',stat(3)
write(12,*)'VARIABLE=',stat(4)
END
c=================================================================
c This routine produce uniform random numbers inside the
c inderval (0,1)
FUNCTION ran1(idum)
INTEGER idum,IA,IM,IQ,IR,NTAB,NDIV
```

```
REAL*8 ran1,AM,EPS,RNMX
PARAMETER (IA=16807,IM=2147483647,AM=1./IM,IQ=127773,IR=2836,
@ NTAB=32,NDIV=1+(IM-1)/NTAB,EPS=1.2e-7,RNMX=1.-EPS)
INTEGER j,k,iv(NTAB),iy
SAVE iv,iy
DATA iv /NTAB*0/, iy /0/
if (idum.le.0.or.iy.eq.0) then
idum=max(-idum,1)
do 11 j=NTAB+8,1,-1
k=idum/IQ
idum=IA*(idum-k*IQ)-IR*k
if (idum.lt.0) idum=idum+IM
if (j.le.NTAB) iv(j)=idum
11 continue
iy=iv(1)
endif
k=idum/IQ
idum=IA*(idum-k*IQ)-IR*k
if (idum.lt.0) idum=idum+IM
j=1+iy/NDIV
iy=iv(j)
iv(j)=idum
ran1=min(AM*iy,RNMX)
return
END
c============================================================
c This routine produce standard normal random numbers
FUNCTION gasdev(idum)
INTEGER idum
REAL*8 gasdev
CU USES ran1
INTEGER iset
```

```
REAL*8 fac,gset,rsq,v1,v2,ran1
SAVE iset,gset
DATA iset/0/
if (iset.eq.0) then
1 v1=2.*ran1(idum)-1.
v2=2.*ran1(idum)-1.
rsq=v1**2+v2**2
if(rsq.ge.1..or.rsq.eq.0.)goto 1
fac=sqrt(-2.*log(rsq)/rsq)
gset=v1*fac
gasdev=v2*fac
iset=1
else
gasdev=gset
iset=0
endif
return
END
c=============================================================
c This routine produce random number following the gamma
c distribution with λ = 1
FUNCTION gamdev(ia,idum)
INTEGER ia,idum
REAL*8 gamdev
CU USES ran1
INTEGER j
REAL*8 am,e,s,v1,v2,x,y,ran1
if(ia.lt.1)pause 'bad argument in gamdev'
if(ia.lt.6)then
x=1.
do 11 j=1,ia
x=x*ran1(idum)
```

```
11 continue
x=-log(x)
else
1 v1=2.*ran1(idum)-1.
v2=2.*ran1(idum)-1.
if(v1**2+v2**2.gt.1.)goto 1
y=v2/v1
am=ia-1
s=sqrt(2.*am+1.)
x=s*y+am
if(x.le.0.)goto 1
e=(1.+y**2)*exp(am*log(x/am)-s*y)
if(ran1(idum).gt.e)goto 1
endif
gamdev=x
return
END
```

Βιβλιογραφία

[1] Good I.J. και Gaskins R.A., *Nonparametric roughness penalties for probability densities*, Biometrika, 58, 255-277 (1971).

[2] Thompson J.R., Tapia R.A., *Nonparametric Function Estimation, Modeling, and Simulation*, Society for Industrial and Applied Mathematics (siam), Philadelphia (1980).

[3] Scott D.W. ,Tapia R.A. και Thompson J.R. *Nonparametric probability density by discrete maximum penalized-likelihood criteria*, Ann.Statist. 8, 820-832 (1980).

[4] Silverman B.W., *On the estimation of a probability density function by the maximum penalized likelihood method*, Ann.Statist. 10, 795-810 (1982).

[5] Silverman B.W., *Density Estimation for Statistics and Data Analysis*, Chapman and Hall (1986).

[6] De Montricher G.F. , Tapia R.A. και Thompson J.R., *Nonparametric maximum likelihood estimation of probability densities by penalty function methods*, Ann. Statist. 3, 1329-1348 (1975).

[7] Klonias V.K., *Consistency of two nonparametric maximum panalized likelihood estimatorw of the probability density function*, Ann.Statist. 10, 811-824 (1982b).

[8] Klonias V.K., *On a classof nonparametric density and regression estimators*, Ann.Statist. 12, 1263-1284 (1984).

[9] Klonias V.K., *Conditions for finite absolute moments of the logarithm of the probability density function*, Tech. Report 335, Dept.of Math. Sciences, The Johns Hopkins University (1981a).

[10] Gijbels I., Klonias V.K., *Density estimation under the Koziol-Green model of censoring by penalized likelihood methods*, Canadian Journal of Statistics, 19, 23-38 (1991).

[11] Loeve M., *Probability Theory*, Van Nostrand, Princeton, N.j.

[12] Billingsley P., *Convergence of probability measures*, Wiley, New York (1968).

[13] *International Mathematical and Statistical Libraries Subroutines DESPL/DDESPL*, Houston, Texas.

[14] Wellner J.A., A law of the iterated logarithm for functions of order statistics, Ann.Statist. 5, 481-494 (σελίδα 484) (1977).

[15] Serfling R.J., Aproximation theorems of Mathematical Statistics, John Wiley Sons (1980).

[16] von Bahr B., Esseen C.G. Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$, Ann. Math. Statist. 36, 299-303 (1965).

[17] William H. Press, Saul A. Teukolsky, William T.Vetterling, Brian P. Flannery, *Numerical recipes in fortran*, Second Edition, Cambridge.

[18] Lions, J.L., Magenes, E. *Non-homogeneous boundary value problems and applications*, Translated from the French by P.Kenneth. New York, Springer-Verlag (1972).