ΜΕΤΡΑ ΑΠΟΚΛΙΣΗΣ: ΑΚΡΙΒΕΙΑ ΑΣΥΜΠΤΩΤΙΚΩΝ ΚΑΤΑΝΟΜΩΝ

Σπυρίδων Κοθρούλας

Η παρούσα Διατριβη
εκπονηθηκε στο Τμημα Μαθηματικων
ως μερικη ικανοποιηση των απαιτησεων για την αποκτηση του
Μεταπτυχιακου Διπλωματος Ειδικευσης στη
Στατιστικη και Αναλογιστικα-Χρηματοοικονομικα Μαθηματικα

ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ

Τμήμα Μαθηματικών

Σάμος, Φεβρουάριος 2016

ΤΡΙΜΕΛΗΣ ΕΠΙΤΡΟΠΗ

Καραγρηγορίου Αλέξανδρος, Επιβλέπων Καθηγητής

Κουντζάκης Χρήστος, Μέλος

Ταχτσής Ελευθέριος, Μέλος

Περίληψη

Τα μέτρα απόστασης ή απόκλισης μπορεί να χρησιμοποιηθούν για τη μελέτη ελέγχων υποθέσεων. Για την εξαγωγή συμπερασμάτων γίνεται χρήση της X^2 κατανομής αφού η ασυμπτωτική κατανομή της φ-απόκλισης (κάτω από την μηδενική υπόθεση) είναι η X^2 με κατάλληλους βαθμούς ελευθερίας. Αυτό το ασυμπτωτικό αποτέλεσμα αποτελεί το θέμα μελέτης της παρούσης διατριβής. Συγκεκριμένα η διατριβή πραγματεύεται το βαθμό ακρίβειας της ασυμπτωτικής κατανομής.Η διερεύνηση επιτυγχάνεται με τη σύγκριση των ροπών της ασυμπτωτικής κατανομής και της πραγματικής κατανομής για πεπερασμένα μεγέθη. Έτσι λοιπόν για δεδομένη τιμή του n και δεδομένη φ προσεγγίζουμε με τη βοήθεια αναπτυγμάτων Taylor όσες περισσότερες ροπές της ακριβούς κατανομής είναι δυνατόν και τις συγκρίνουμε με τις αντίστοιχες ροπές της ασυμπτωτικής κατανομής. Στόχος μας είναι για κάθε συνάρτηση φ να προσδιορίσουμε το πεδίο τιμών της παραμέτρου λ ή α για το οποίο η απόσταση μεταξύ των ροπών που συγκρίθηκαν να είναι ελάχιστη. Η διερεύνηση αυτή βασίζεται στην υπόθεση ότι οι ροπές ορίζουν μονοσήμαντα την κατανομή.

Abstract

The distance or deviation measures can be used to study hypothesis testing .To draw conclusions made use of the X^2 distribution as the asymptotic distribution of φ-divergence (under the null hypothesis) is the X^2 with appropriate degrees of freedom. This asymptotic result is the subject matter of this dissertation. Specifically, the dissertation deals with the degree of precision of the asymptotic distribution. The Investigation is achieved by comparing the moments of the asymptotic distribution and the actual distribution of finite sizes.Thus for a given value of n and a given aircraft approach using blanks Taylor as many moments of the exact distribution is possible and compare them with the corresponding moments of the asymptotic distribution.Our goal is for every function φ to determine the range of λ or α for which the distance between the moments compared to be minimal. This will be based on the assumption that the moments uniquely define the distribution

Περιεχόμενα

Μέρος Ι
Τα Μέτρα Απόκλισης

Κεφάλαιο 0
Εισαγωγή

Στο Κεφάλαιο αυτό θα αναφερθούμε στην έννοια του μέτρου και θα επικεντρωθούμε στα μέτρα απόκλισης ή απόστασης (measures of divergence) όπως αυτά ορίζονται και αξιοποιούνται στη Μαθηματική Στατιστική και την Πιθανοθεωρία.

Σημαντικό ρόλο στα Μαθηματικά παίζουν μεταξύ πολλών άλλων εννοιών η μετρική αφού σε αυτό βασίζεται η θεμελίωση της ανάλυσης και του απειροστικού λογισμού. Τα μέτρα αποτελούν συναρτήσεις με πεδίο τιμών τον θετικό άξονα των πραγματικών αριθμών και χρησιμοποιούνται για να προσδιορίσουν την απόσταση μεταξύ σημείων ή συναρτήσεων ή σχέσεων στο επίπεδο. Οι βασικές ιδιότητες που πληρούν τα μέτρα είναι η μη αρνητικότητα, η συμμετρικότητα και η τριγωνική ιδιότητα όπως φαίνεται στον Ορισμό 0.1.

Ορισμός 0.1: Μια συνάρτηση d(x,y) που ορίζεται για όλα τα ζεύγη των σημείων x,y ενός χώρου Ω λέγεται μετρική αν για όλα τα σημεία x,y,z του Ω:

(α) $d(x,y) \geq 0$ και $d(x,y) = 0$ αν και μόνο αν x=y

(β) $d(x,y) = d(y,x)$ και

(γ) $d(x,z) \leq d(x,y) + d(y,z)$

Το σημαντικότερο μέτρο είναι η Ευκλείδια μετρική όπου $\Omega = \mathbb{R}^n$ και

$$d(x,y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} .$$

Ανάλογα μέτρα γνωστά ως μέτρα απόστασης ή απόκλισης μπορούν να οριστούν και στη Μαθηματική Στατιστική και την Πιθανοθεωρία. Ο ρόλος τους είναι εξίσου σημαντικός όπως και στα Μαθηματικά. Στη συγκεκριμένη περίπτωση τα **Μέτρα απόστασης αποτελούν εργαλεία καθορισμού της** απόστασης μεταξύ δύο πιθανοτήτων κατανομών, δηλαδή μετρούν πόσο «μακριά» είναι η μια κατανομή από την άλλη. Τα μέτρα αυτά μπορεί να αφορούν είτε 2 συναρτήσεις κατανομών F_1 και F_2 είτε 2 συναρτήσεις πυκνότητας ή μάζας πιθανότητας f_1 και f_2 ή p_1 και p_2.

Σημειώνεται ότι στη Στατιστική δεν μας ενδιαφέρει αν ικανοποιούνται όλες οι συνθήκες του ορισμού 0.1. Αρκεί η πρώτη να πληρείται.Χαρακτηριστικα παραδείγματα είναι τα εξής:

Για τη συμμετρία:

Χρησιμοποιώντας τον τύπο για την συνεχή περίπτωση της απόστασης K-L $\int f_1(x)\ln\left(\frac{f_1(x)}{f_2(x)}\right)$ έχουμε:

Έστω $\mathrm{f}_1 = 1$ και $\mathrm{f}_2 = \lambda e^{-\lambda x}$, τότε

$$\int_0^1 1\ln(\frac{1}{\lambda e^{-\lambda x}})dx = -\ln + \frac{\lambda}{2}$$

Έστω $\mathrm{f}_1 = \lambda e^{-\lambda x}$ και $\mathrm{f}_2 = 1$, τότε

$$\int_0^\infty \lambda e^{-\lambda x}\ln(\lambda e^{-\lambda x})dx = \ln\lambda - 1$$

Επομένως η ιδιότητα της συμμετρίας δεν πληρείται.

Για τη τριγωνική ιδιότητα:

Έστω $\mathrm{X} \sim \begin{cases} 0, p_x = 1/2 \\ 1, q_x = 1/2 \end{cases}$, $\mathrm{Y} \sim \begin{cases} 0, p_y = 1/10 \\ 1, q_y = 9/10 \end{cases}$ και $\mathrm{Z} \sim \begin{cases} 0, p_z = 1/4 \\ 1, q_z = 3/4 \end{cases}$

Χρησιμοποιώντας τον τύπο για την διακριτή περίπτωση της απόστασης K-L $\sum p\log\frac{p}{q}$ έχουμε:

$d(x,z) = 0.1438$

$d(z,y) = 0.092$

$d(x,y) = 0.51$

Επομένως δεν ισχύει η τριγωνική ιδιότητα $d(x,y) \le d(x,z) + d(z,y)$

Χαρακτηριστικό παράδειγμα της πρώτης περίπτωσης αποτελεί η γνωστή συνάρτηση Kolmogorov (βλέπε ενότητα 1.1) η οποία δίνεται από τον τύπο :

$$\mathrm{D}_{kolm}(F_1, F_2) = \max\left|F_1(x) - F_2(x)\right|$$

Στην περίπτωση αυτή η απόσταση ορίζεται να είναι η τιμή που αντιστοιχεί στο σημείο x* του πεδίου ορισμού που μεγιστοποιεί την απόλυτη διαφορά $F_1(x) - F_2(x)$ με x στο R. Η συνάρτηση αυτή έχει ευρεία χρήση στη Στατιστική Συμπερασματολογία και αποτελεί το βασικό στατιστικό εργαλείο της Απαραμετρικής (Μη Παραμετρικής) Στατιστικής για τη διερεύνηση της υπόθεσης αν η κατανομή ενός συνόλου δεδομένων ταυτίζεται με συγκεκριμένη (δεδομένη) κατανομή. Συγκεκριμένα αν $X_1, X_2, \ldots, X_n$ τυχαίο δείγμα από κατανομή με συνάρτηση κατανομής F και η ζητούμενη μηδενική υπόθεση του ελέγχου είναι

H_0: $F = F_0$

με F_0 γνωστή κατανομή τότε η απόσταση μεταξύ της F και της F_0 εκτιμάται από τη στατιστική συνάρτηση

$$\max_i |\hat{F}(x_i) - F_0(x_i)|$$

όπου $\hat{F}(\cdot)$ η εκτιμήτρια (εμπειρική συνάρτηση κατανομής) της F με βάση το τυχαίο δείγμα. Αν η απόσταση αυτή είναι πολύ μικρή τότε η μηδενική υπόθεση γίνεται αποδεκτή αλλιώς απορρίπτεται. Λόγω του ότι η κατανομή της πιο πάνω στατιστικής συνάρτησης κάτω από τη μηδενική υπόθεση δεν είναι γνωστή, η μελέτη του ελέγχου γίνεται με τη χρήση ειδικών πινάκων. Γενικεύσεις της πιο πάνω βασικής περίπτωσης υπάρχουν πολλές (Lilliefors, 1967; Smyrnov, 1949).

Χαρακτηριστικά παραδείγματα της δεύτερης περίπτωσης αποτελούν το Pearson chi-square test (Pearson, 1900) και η απόσταση των Kullback & Liebler (Kullback and Liebler, 1947). Η συνάρτηση του Pearson μπορεί να θεωρηθεί ως η πρώτη συνάρτηση στη Στατιστική για τον προσδιορισμό της απόστασης μεταξύ 2 συναρτήσεων μάζας πιθανότητας. Η ιδέα βασίζεται στη σύγκριση των παρατηρούμενων συχνοτήτων των κατηγοριών ή κλάσεων της κατανομής με τις αντίστοιχες συχνότητες που θα έπρεπε να παρατηρηθούν αν τα δεδομένα προέρχονταν από συγκεκριμένη (γνωστή) κατανομή. Όπως είναι προφανές αν η διαφορά αυτή είναι πολύ μικρή η μηδενική υπόθεση ότι τα δεδομένα προέρχονται από τη συγκεκριμένη κατανομή, γίνεται δεκτή. Σε διαφορετική περίπτωση η μηδενική υπόθεση απορρίπτεται. Στην περίπτωση αυτή η μελέτη του ελέγχου διευκολύνεται από το γεγονός ότι κάτω από τη μηδενική υπόθεση, η (στατιστική) συνάρτηση του Pearson ακολουθεί τη χι τετράγωνο κατανομή. Αν και η συνάρτηση του Pearson εξακολουθεί και έχει ευρύτατη χρήση, παραλλαγές και γενικεύσεις έχουν προταθεί πολλές (Neymann, 1949 κ.α).

Η απόσταση Kullback – Liebler αποτελεί μια από τις σημαντικότερες στατιστικές συναρτήσεις και ορίζεται, για τη συνεχή περίπτωση, ως εξής:

$$\mathrm{D}_{kull}(f_1, f_2) = \int f_1(x) \ln\left(\frac{f_1(x)}{f_2(x)}\right) dx.$$

Ο αντίστοιχος τύπος για τη διακριτή περίπτωση είναι

$$\mathrm{D}_{kull}(p_1, p_2) = \sum_x p_1(x) \ln\left(\frac{p_1(x)}{p_2(x)}\right).$$

Δεν είναι δύσκολο να παρατηρήσει κανείς ότι οι πιο πάνω ποσότητες είναι πάντα θετικές ενώ ισούνται με το 0 στην περίπτωση που οι 2 συναρτήσεις ταυτίζονται οπότε ο λογάριθμος μηδενίζεται.

Σημειώνεται ότι η απόσταση ΚΛ δεν πληροί τη συνθήκη του Ορισμού 0.1 αφού

$D_{kull}(f_1, f_2) \neq D_{kull}(f_2, f_1)$.

Η απόσταση Kullback-Liebler σχετίζεται άμεσα με τη έννοια της πιθανοφάνειας που αποτελεί κυρίαρχη έννοια της Στατιστικής Θεωρίας. Πράγματι αν χρειάζεται να διερευνηθεί κατά πόσον η συνάρτηση πυκνότητας πιθανότητας f ισούται με την f_1 ή την f_2, δηλαδή αν

H_0: f = f_1 vs. H_1: f = f_2

τότε συγκρίνουμε τις αποστάσεις $D_{kull}(f, f_1)$ και $D_{kull}(f, f_2)$. Όπως εύκολα φαίνεται, προκύπτει ότι

$$D_{kull}(f, f_1) - D_{kull}(f, f_2) = \int f(x) \ln\left(\frac{f_2(x)}{f_1(x)}\right) dx.$$

Αν η πιο πάνω ποσότητα είναι αρνητική τότε γίνεται δεκτή η μηδενική υπόθεση αλλιώς, αν είναι θετική, γίνεται δεκτή η εναλλακτική υπόθεση. Όπως γίνεται αντιληπτό η πιο πάνω ποσότητα βασίζεται αποκλειστικά στην ποσότητα $\ln\left(\frac{f_2(x)}{f_1(x)}\right)$ η οποία δεν είναι τίποτε άλλο παρά η στατιστική συνάρτηση λόγου πιθανοφανειών για τον πιο πάνω έλεγχο.

Ας υποθέσουμε ότι η ιδανική κατανομή είναι η f_1 και ότι αυτή εξαρτάται από μια άγνωστη παράμετρο θ (για σκοπούς συμβολισμού $f_{1,\theta}$). Τότε τίθεται το ερώτημα ποια είναι η ιδανική εκτιμήτρια για την άγνωστη παράμετρο. Για το σκοπό αυτό απαιτείται να προσδιοριστεί η τιμή της θ για την οποία ελαχιστοποιείται η απόσταση $K(f, f_{1,\theta})$ η οποία όμως ισοδυναμεί με την μεγιστοποίηση της ποσότητας

$$D^*_{kull}(f, f_{1,\theta}) = \int f(x) \ln\left(f_{1,\theta}(x)\right) dx \equiv E_f\left(\ln f_{1,\theta}(x)\right)$$

Με βάση το νόμο των μεγάλων αριθμών η μέση αυτή τιμή μπορεί να εκτιμηθεί από τον αντίστοιχο δειγματικό μέσο με βάση τυχαίο δείγμα $X_1, \ldots, X_n$:

$$D^*_{kull}(f, f_{1,\theta}) = \frac{1}{n}\sum_{i=1}^{n} \ln f_{1,\theta}(x_i) = \frac{1}{n}\ln\left(\prod_{i=1}^{n} f_{1,\theta}(x_i)\right)$$

που δεν είναι τίποτε άλλο παρά ο λογάριθμος της συνάρτησης πιθανοφάνειας η μεγιστοποίηση του οποίου οδηγεί στη γνωστή εκτιμήτρια μεγίστης πιθανοφάνειας. Με άλλα λόγια η εκτιμήτρια μεγίστης πιθανοφάνειας είναι εκείνη η τιμή της άγνωστης παραμέτρου που ελαχιστοποιεί την απόσταση Kullback-Liebler κάτι που καθιστά την συγκεκριμένη απόσταση μια από τις σημαντικότερες έννοιες της Μαθηματικής Στατιστικής. Μεταξύ των γενικεύσεων της απόστασης Kullback-Liebler περιλαμβάνονται και οι εξής, όπου είτε γενικεύεται η εσωτερική συνάρτηση (ln) είτε γενικεύεται η δύναμη της συνάρτησης εκτός του λογαρίθμου είτε και τα δύο μαζί. Για παράδειγμα αν φ μια συνάρτηση τις λεπτομέρειες της οποίας θα ορίσουμε αργότερα τότε η φ-απόκλιση ορίζεται ως εξής:

$$\Phi(f_1, f_2) = \int f_1(x)\varphi\left(\frac{f_1(x)}{f_2(x)}\right)dx$$

Προφανώς η KL απόσταση αποτελεί ειδική περίπτωση της φ-απόκλισης αν η συνάρτηση φ επιλεχθεί να είναι η συνάρτηση «λογάριθμος». Συνήθως η κάθε συνάρτηση φ που είναι δυνατόν να επιλέξει ο ερευνητής εμπεριέχει μια παράμετρο (index) με αποτέλεσμα για κάθε επιλογή της φ να δημιουργείται μια ολόκληρη οικογένεια μέτρων απόκλισης ως συνάρτηση της παραμέτρου (συνήθως συμβολίζεται με λ ή α). Μια από τις γνωστότερες συναρτήσεις φ είναι η

$$\varphi(x) = \varphi_{(\lambda)}(x) = \frac{1}{\lambda(\lambda+1)}(x^{\lambda+1} - x - \lambda(x-1)), \quad \lambda \neq 0,1$$

Στη συγκεκριμένη περίπτωση που αφορά το Power divergence των Cressie & Read (Cressie and Read, 1984; Read and Cressie, 1988) η παράμετρος συμβολίζεται με λ και παίρνει τιμές στο $\mathbb{R} - \{0,-1\}$.

Όπως και στην περίπτωση της KL απόστασης έτσι και η φ-απόκλιση μπορεί να χρησιμοποιηθεί για τη μελέτη ελέγχων υποθέσεων όπως αυτές που συζητήθηκαν νωρίτερα. Για την εξαγωγή συμπερασμάτων γίνεται χρήση της X^2 κατανομής αφού η ασυμπτωτική κατανομή της πιο πάνω στατιστικής συνάρτησης (κάτω από την μηδενική υπόθεση) είναι η X^2 με κατάλληλους βαθμούς ελευθερίας. Αυτό το ασυμπτωτικό αποτέλεσμα αποτελεί το θέμα μελέτης της παρούσης διατριβής. Συγκεκριμένα η διατριβή πραγματεύεται το βαθμό ακρίβειας της ασυμπτωτικής κατανομής. Η διερεύνηση επιτυγχάνεται με τη σύγκριση των ροπών της ασυμπτωτικής κατανομής και της πραγματικής κατανομής για πεπερασμένα μεγέθη. Έτσι λοιπόν για δεδομένη τιμή του n και δεδομένη φ προσεγγίζουμε με τη βοήθεια αναπτυγμάτων Taylor όσες περισσότερες ροπές της ακριβούς κατανομής είναι δυνατόν και τις συγκρίνουμε με τις αντίστοιχες ροπές της ασυμπτωτικής κατανομής. Στόχος μας είναι για κάθε συνάρτηση φ να προσδιορίσουμε το πεδίο τιμών της παραμέτρου λ ή α για το οποίο η απόσταση μεταξύ των ροπών που συγκρίθηκαν να είναι ελάχιστη. Η διερεύνηση αυτή βασίζεται στην υπόθεση ότι οι ροπές ορίζουν μονοσήμαντα την κατανομή.

Σημείωση. Χρειάζεται να αναφερθεί ότι το πιο πάνω σχόλιο για το μονοσήμαντο της κατανομής ισχύει σε δύο τουλάχιστον περιπτώσεις:

(α) Αν η τ.μ. X είναι διακριτή και το πεδίο τιμών της είναι πεπερασμένο (Theorem 10.1, Grinstead & Snell, 1997) και

(β) αν για τις ροπές ρ_j η σειρά : $\sum_{j=1}^{\infty} \frac{1}{\left(\rho_{2j}\right)^{\frac{1}{2j}}}$ αποκλίνει (Carleman, 1926)

Χαρακτηριστικό παράδειγμα αποτελεί η κατανομή lognormal. Πράγματι, όπως είναι γνωστό, αν Y ακολουθεί την N(0,1) τότε X = exp(Y) ακολουθεί τη lognormal με σ.π.π. f(x) και $E(X^r) = \exp(r^2/2)$ όπου $r \geq 1$ ακέραιος αριθμός. Ο Heyde (1963) παρατήρησε ότι τις ίδιες ροπές τις έχει και η τ.μ με συνάρτηση πυκνότητας

cg(x)=f(x)[1+1/2 sin(2πklogx)]

όπου k είναι οποιοσδήποτε θετικός ακέραιος και c>0 ώστε η g(x) να αποτελεί σ.π.π.

Ένα άλλο σημαντικό κριτήριο ώστε το σύνολο των ροπών α_j να καθορίζει μονοσήμαντα την σ.π.π. είναι το *Careleman Criterion* (Shohat & Tamarkin, 1943) σύμφωνα με το οποίο η κάτωθι σχέση αποτελεί απαραίτητη προϋπόθεση

$$\sum_{j=1}^{\infty} \frac{1}{(a_{2j})^{1/a_{2j}}} = \infty$$

Κεφάλαιο 1

Η έννοια του Μέτρου στη Στατιστική και την Πιθανοθεωρία

Στο κεφάλαιο αυτό θα αναφερθούμε στα κυριότερα μέτρα απόκλισης τα οποία συναντώνται στη Στατιστική και την Πιθανοθεωρία και θα καταγράψουμε πέρα από τους ορισμούς και κάποιες από τις κυριότερες ιδιότητές τους.

1.1 Μέτρα Απόκλισης: Ορισμός και Μέτρο Kolmogorov

Έστω Χ μια τυχαία μεταβλητή και έστω ότι η συνάρτηση κατανομής (ή κατανομή πιθανότητας) F_θ της Χ εξαρτάται από μια παράμετρο $\theta \in \Theta$. Επίσης ορίζουμε ότι η συνάρτηση πυκνότητας πιθανότητας είναι

$f_\theta(\mathrm{x}) = \dfrac{dF_\theta}{d\mu}(\mathrm{x})$ αν μ μέτρο Lebesgue

ή

$p_\theta(\mathrm{x}) = \Pr_\theta(\mathrm{X} = x)$ αν μ είναι μετρική (βλ ορισμό 0.1).

Στην πρώτη περίπτωση η Χ είναι τυχαία μεταβλητή με συνεχή κατανομή και στην δεύτερη περίπτωση είναι μια διακριτή τυχαία μεταβλητή. Αν h μια μετρήσιμη συνάρτηση τότε η αναμενόμενη τιμή της h(x) ορίζεται:

$E_\theta[h(x)] = \int_x h(x) \cdot f_\theta\left(x\right)$ dx αν μ είναι μέτρο Lebesgue

ή

$E_\theta[h(x)] = \sum_x h(x) \cdot p_\theta(x)$ αν μ είναι ένα μετρήσιμο μέτρο

Τα **Μέτρα απόστασης αποτελούν εργαλεία καθορισμού της** απόστασης μεταξύ δύο πιθανοτήτων κατανομών, δηλαδή μετρούν πόσο «μακριά» είναι η μια κατανομή από την άλλη. Τα μέτρα αυτά μπορεί να αφορούν είτε 2 συναρτήσεις κατανομών F_1 και F_2 είτε 2 συναρτήσεις πυκνότητας ή μάζας πιθανότητας f_1 και f_2 ή p_1 και p_2. Ένα κλασικό μέτρο απόκλισης μεταξύ 2 συναρτήσεων κατανομής είναι το μέτρο του Kolmogorov (Kolmogorov, 1933) το οποίο ορίζεται ως εξής:

Απόσταση Kolmogorov - Levy

Έστω $\mathrm{P}_{\theta_1}, \mathrm{P}_{\theta_2}$ δύο μέτρα πιθανότητας και $\mathrm{F}_{\theta_1}, F_{\theta_2}$ οι αντίστοιχες συναρτήσεις κατανομής.

Η απόσταση Kolmogorov μεταξύ F_{θ_1} και F_{θ_2} *ή* μεταξύ P_{θ_1} και P_{θ_2} δίνεται από τον τύπο:

$$\mathrm{K}_1(F_{\theta_1}, F_{\theta_2}) = \sup_{x \in R} \left| F_{\theta_1}(\mathrm{x}) - F_{\theta_2}(\mathrm{x}) \right|. \tag{1.1}$$

Το σημαντικότερο αποτέλεσμα της Πιθανοθεωρίας που σχετίζεται με την απόσταση Kolmogorov είναι το θεώρημα Glivenko-Cantelli το οποίο δίνει την συνέπεια της εμπειρικής συνάρτησης κατανομής.

Θεώρημα Glivenko-Cantelli (Glivenko, 1933 & Cantelli, 1933)

Αν

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I\{X_i < x\}$$

η εμπειρική συνάρτηση κατανομής για το τυχαίο δείγμα $X_1,\ldots X_n$ με σ.κ. F τότε

Η εμπειρική συνάρτηση κατανομής είναι ομοιόμορφα συνεπής εκτίμηση της πραγματικής συνάρτησης κατανομής:

$$\sup \left| F_n(x) - F(x) \to 0 \text{ σ.β} \right|$$

1.2 Μέτρα Απόκλισης Μεταξύ Δύο Πιθανοτήτων Κατανομών

Στην ενότητα αυτή θα οριστούν διάφορα μέτρα απόκλισης με πρώτο **το μέτρο των Kullback & Liebler** που ορίζεται ως εξής:

$$D_{kull}(\theta_1, \theta_2) = \int_x f_{\theta_1}(\mathrm{x}) \log\left(\frac{f_{\theta_1}(\mathrm{x})}{f_{\theta_2}(\mathrm{x})}\right) d\mu(\mathrm{x}) = \mathrm{E}_{\theta_1}\left[\log\left(\frac{f_{\theta_1}(X)}{f_{\theta_2}(X)}\right)\right] \tag{1.2}$$

O Jeffreys (1946) χρησιμοποίησε μια συμμετρική εκδοχή της (1.2)

$$J(\theta_1, \theta_2) = D_{kull}(\theta_1, \theta_2) + D_{kull}(\theta_2, \theta_1)$$

ως μέτρο απόκλισης μεταξύ δύο πιθανοτήτων κατανομών. Αυτό το μέτρο απόκλισης ονομάζεται J-απόκλιση και σε αντίθεση με το μέτρο των Kullback & Leibler ικανοποιεί τη συμμετρική ιδιότητα.

O Reniy (1961) παρουσίασε την πρώτη παραμετρική γενίκευση της (1.2) με την εισαγωγή ενός δείκτη (παραμέτρου) r:

$$D_r^1(\theta_1,\theta_2)=\frac{1}{r-1}\log\int_x f_{\theta_1}(\mathrm{x})^r\, f_{\theta_2}(\mathrm{x})^{1-r}\,d\mu(\mathrm{x})=\frac{1}{r-1}\log E_{\theta_1}\left[\left(\frac{f_{\theta_1}(X)}{f_{\theta_2}(X)}\right)^{r-1}\right],\ r>0,\ r\neq 1$$

Αργότερα οι Liese και Vajda (1987) το επέκτειναν για όλα τα $r \neq 1,0$

$$D_r^1(\theta_1,\theta_2)=\frac{1}{r(r-1)}\log\int_x f_{\theta_1}(\mathrm{x})^r\, f_{\theta_2}(\mathrm{x})^{1-r}\,d\mu(\mathrm{x})=\frac{1}{r(r-1)}\log E_{\theta_1}\left[\left(\frac{f_{\theta_1}(X)}{f_{\theta_2}(X)}\right)^{r-1}\right],\ r\neq 0,1 \qquad (1.3).$$

Για r = 1 και r = 0 μπορεί εύκολα να δειχθεί ότι

$$D_1^1(\theta_1,\theta_2)=\lim_{r\to 1} D_r^1(\theta_1,\theta_2)=D_{kull}(\theta_1,\theta_2)$$
$$D_0^1(\theta_1,\theta_2)=\lim_{r\to 0} D_r^1(\theta_1,\theta_2)=D_{kull}(\theta_2,\theta_1)$$

Το μέτρο απόκλισης $\mathrm{D}_{kull}(\theta_2,\theta_1)$ αναφέρεται και ως "ελάχιστη διάκριση πληροφοριών" μεταξύ των κατανομών πιθανότητας P_{θ_1} και P_{θ_2}.

Άλλες δύο παραμετρικές γενικεύσεις της (1.2) είναι οι **r-order** και **s-degree** μέτρα απόκλισης. Αυτά προτάθηκαν από τους Sharma και Mittal (1977) ως εξής :

$$D_r^s(\theta_1,\theta_2)=\frac{1}{s-1}\left[\left(\int_x f_{\theta_1}(\mathrm{x})^r\, f_{\theta_2}(\mathrm{x})^{1-r}\,d\mu(\mathrm{x})\right)^{\frac{s-1}{r-1}}-1\right]$$
$$=\frac{1}{s-1}\left[\left(E_{\theta_1}\left(\frac{f_{\theta_1}(\mathrm{X})}{f_{\theta_2}(\mathrm{X})}\right)^{r-1}\right)^{\frac{s-1}{r-1}}-1\right]\quad r,s\neq 1 \qquad (1.4)$$

$$D_1^s(\theta_1,\theta_2)=\frac{1}{s-1}\left[\exp((s-1)\int_x f_{\theta_1}(\mathrm{x})\log\left(\frac{f_{\theta_1}(\mathrm{x})}{f_{\theta_2}(\mathrm{x})}\right)d\mu(\mathrm{x}))-1\right]$$
$$=\frac{1}{s-1}\left[\exp\left((s-1)E_{\theta_1}\left(\log\frac{f_{\theta_1}(\mathrm{X})}{f_{\theta_2}(\mathrm{X})}\right)\right)-1\right],\ s\neq 1 \qquad (1.5)$$

Η οικογένεια **των μέτρων φ-απόκλισης** προτάθηκε ανεξάρτητα από τον Csiszar (1963) και τους Ali και Silvey (1966) και ορίζεται έτσι:

Ορισμός 1.1

Το μέτρο φ-απόκλισης μεταξύ των κατανομών πιθανότητας P_{θ_1} και P_{θ_2} δίνεται από τον τύπο

$$D_{\varphi}(P_{\theta_1}, P_{\theta_2}) = D_{\varphi}(\theta_1, \theta_2) = \int_{x} f_{\theta_2}(x)\ \varphi(\frac{f_{\theta_1}(x)}{f_{\theta_2}(x)}) d\mu(x) = E_{\theta_2}\left[\varphi(\frac{f_{\theta_1}(X)}{f_{\theta_2}(X)})\right], \varphi \in \Phi^* \qquad (1.6)$$

όπου Φ^* είναι η κλάση όλων των κυρτών συναρτήσεων φ(x), x ≥ 0 έτσι ώστε για x = 1, φ(1) = 0, 0φ(0 / 0) = 0 και 0φ(p/0) = $\lim_{u\to\infty}\varphi(u)/u$.

Το μέτρο απόκλισης Kullback-Leiber είναι η πιο γνωστή ειδική περίπτωση της οικογένειας των μέτρων φ-απόκλισης.

Σχόλιο:

Έστω φ ∈ Φ* είναι διαφορίσιμη στο x = 1 τότε η συνάρτηση ψ(x) = φ(x) - φ΄(1) (x − 1) ανήκει επίσης στη Φ* και έχει την επιπλέον ιδιότητα ψ΄(1) = 0. Από αυτή την ιδιότητα και μαζί με τη κυρτότητα συνεπάγεται ότι ψ(x) ≥ 0 ∀ x ≥ 0. Παρατηρήστε ότι τα 2 μέτρα συμπίπτουν

$$D_{\psi}(\theta_1, \theta_2) = \int_{x} f_{\theta_2}(x)\ (\varphi\left(\frac{f_{\theta_1}(x)}{f_{\theta_2}(x)}\right) - \varphi'(1)\left(\frac{f_{\theta_1}(x)}{f_{\theta_2}(x)} - 1\right)) d\mu(x) =$$

$$= \int_{x} f_{\theta_2}(x)\ \varphi\left(\frac{f_{\theta_1}(x)}{f_{\theta_2}(x)}\right) d\mu(x) = D_{\varphi}(\theta_1, \theta_2)$$

οπότε μπορούμε να θεωρήσουμε το σύνολο Φ^* να είναι ισοδύναμο με το $\Phi \equiv \Phi^* \cap \{\varphi : \varphi'(1) = 0\}$.

Το μέτρο απόκλισης Kullback-Leiber λαμβάνεται για : ψ(x) = xlog - x + 1 ή φ(x) = xlogx.

Η πιο σημαντική οικογένεια της φ-απόκλισης είναι η οικογένεια που μελετήθηκε από τους Cressie & Read (1984). Η **power divergence των Cressie & Read** δίνεται από :

$$I_{\lambda}(\theta_1, \theta_2) \equiv D_{\varphi(\lambda)}(\theta_1, \theta_2) = \frac{1}{\lambda(\lambda+1)}\left[\int_{x} \frac{f_{\theta_1}(x)^{\lambda+1}}{f_{\theta_2}(x)^{\lambda}} d\mu(x) - 1\right] = \frac{1}{\lambda(\lambda+1)}\left[(E_{\theta_1}(\frac{f_{\theta_1}(X)}{f_{\theta_2}(X)})^{\lambda}) - 1\right]$$

για $-\infty < \lambda < \infty$ ενώ δεν ορίζεται για λ = -1, λ=0.

Ωστόσο αν αξιοποιήσουμε τα συνεχή όρια του $\mathrm{I}_\lambda(\theta_1,\theta_2)$ όταν λ → -1 και λ → 0 τότε δεν είναι δύσκολο να δειχτεί ότι:

$$\lim_{\lambda\to 0}\mathrm{I}_\lambda(\theta_1,\theta_2)=D_{kull}(\theta_1,\theta_2)$$

και

$$\lim_{\lambda\to -1}\mathrm{I}_\lambda(\theta_1,\theta_2)=D_{kull}(\theta_2,\theta_1)$$

Μπορούμε να παρατηρήσουμε ότι η οικογένεια power divergence λαμβάνεται από την οικογένεια των μέτρων φ-απόκλισης (1.6) ως εξής:

$$\varphi(\mathrm{x})=\varphi_{(\lambda)}(\mathrm{x})=\frac{1}{\lambda(\lambda+1)}(\mathrm{x}^{\lambda+1}-x-\lambda(\mathrm{x}-1)),\lambda\neq 0,\lambda\neq 1$$
$$\varphi(\mathrm{x})=\varphi_{(0)}(\mathrm{x})=\lim_{\lambda\to 0}\varphi_{(\lambda)}(\mathrm{x})=x\log\mathrm{x}-x+1$$
$$\varphi(\mathrm{x})=\varphi_{(-1)}(\mathrm{x})=\lim_{\lambda\to -1}\varphi_{(\lambda)}(\mathrm{x})=-\log\mathrm{x}+x-1$$

Τα μέτρα απόκλισης των Renyi, Sharma και Mittal που δίνονται στις (1.3) και (1.4) όπως επίσης και το μέτρο Bhattacharyya (1943) που δίνεται από τον τύπο:

$$\mathrm{B}(\theta_1,\theta_2)=-\log\int_x\sqrt{f_{\theta_1}(\mathrm{x})f_{\theta_2}(\mathrm{x})}d\mu(\mathrm{x})$$

δεν είναι μέτρα απόκλισης. Μπορούμε όμως να παρατηρήσουμε ότι τα μέτρα αυτά μπορούν να γραφούν στην ακόλουθη μορφή :

$$D_\varphi^h(\theta_1,\theta_2)=h(D_\varphi(\theta_1,\theta_2))$$

όπου h είναι μια διαφορισίμη αύξουσα πραγματική συνάρτηση από το διάστημα $\left[0,\varphi(0)+\lim_{t\to\infty}\frac{\varphi(\mathrm{t})}{t}\right]$ στο $[0,\infty)$ με h(0)=0, $h'(0)>0$ και $\varphi\in\Phi^*$.

Η καινούργια οικογένεια αποκλίσεων ονομάζεται **οικογένεια μέτρων (h,φ)-απόκλισης**.

1.2.1 Ιδιότητες των μέτρων φ-απόκλισης

Στην ενότητα αυτή παρουσιάζουμε μερικές από τις πιο σημαντικές ιδιότητες των μέτρων φ-απόκλισης. Στα ακόλουθα υποθέτουμε την ύπαρξη της πρώτης παραγώγου της φ στο σημείο x = 1.

Πρόταση 1.1

Έστω $P_{\theta 1}$ και $P_{\theta 2}$ δύο κατανομές πιθανοτήτων και έστω $\varphi \in \Phi^*$ διαφορίσημη στο t=1.Τότε :

$0 \le D_{\varphi}(\theta_1, \theta_2) \le \varphi(0) + \lim_{r\to\infty} \frac{\varphi(r)}{r}$ όπου

$D_{\varphi}(\theta_1, \theta_2) = 0$ αν $P_{\theta_1} = P_{\theta_2}$ και (1.9)

$D_{\varphi}(\theta_1, \theta_2) = \varphi(0) + \lim_{r\to\infty} \frac{\varphi(r)}{r}$ αν $S_1 \cap S_2 = \varnothing$ (1.10).

Επίσης αν φ είναι αυστηρά κυρτή στο t=1 τότε η (1.9) ισχύει ανν $P_{\theta_1} = P_{\theta_2}$.

Επιπλέον αν $\varphi(0) + \lim_{r\to\infty} \frac{\varphi(r)}{r} < \infty$, η (1.10) ισχύει ανν $S_1 \cap S_2 = \varnothing$.

Πρόταση 1.2

Ας υποθέσουμε ότι οι κατανομές πιθανότητας $\{P_\theta\}$ $\theta \in \Theta$ είναι στη πραγματική ευθεία $\theta \in (a,b) \subset R$ και έστω P_θ είναι απόλυτα συνεχείς σε σχέση με ένα σ-πεπερασμένο μέτρο μ (μέτρο Lebesgue ή μέτρο μέτρησης).Υποθέτουμε ότι οι συναρτήσεις πυκνότητας ή οι συναρτήσεις μαζας πιθανότητας έχουν μονότονο λόγο πιθανότητας x.Αν $a < \theta_1 < \theta_2 < \theta_3 < b$ και φ συνεχής συνάρτηση τότε :

$D_{\varphi}(\theta_1, \theta_2) \le D_{\varphi}(\theta_1, \theta_3), \varphi \in \Phi^*$.

Σχόλιο:

Είναι προφανές ότι εάν h είναι διαφορίσιμη αύξουσα συνάρτηση, το (h, φ) μέτρο απόκλισης ικανοποιεί επίσης τις προτάσεις 1.1 και 1.2.

Σχόλιο:

Αν φ∈Φ* η οποία είναι αυστηρά κυρτή στο x =1 , η αντίστοιχη φ-απόκλιση είνα μια αντανακλαστική απόσταση (reflexive distance) στο χώρο $P=\{P_\theta\}$ θ∈Θ. Αυτό είναι δυνατό να ορίσει ένα νέο μέτρο απόκλισης, βασισμένο σε μια δοθείσα φ-απόκλιση με τέτοιο τρόπο ώστε το νέο μέτρο απόκλισης θα είναι όχι μόνο αυτανακλαστικό αλλά και συμμετρικό. Αυτό είναι εφικτό εάν λάβουμε υπόψη το μέτρο απόκλισης να σχετίζεται με τη συνάρτηση : φ(t) = φ(t) + t φ(1/t).

1.3 Άλλα μέτρα απόκλισης

Στην ενότητα αυτή παρουσιάζονται και άλλα σημαντικά μέτρα απόκλισης μεταξύ δύο πιθανοτήτων κατανομών που δεν είναι, σε γενικές γραμμές, ειδικές περιπτώσεις των φ - απόκλισης μέτρων. Θεωρούμε δύο ομάδες μέτρων. Η πρώτη αντιστοιχεί στα μέτρα Rφ-απόκλισης που θεσπίστηκαν από τους Burbea και Rao (1982a, b, c) και η δεύτερη εκείνη που αντιστοιχεί στις Bregman αποστάσεις που μελετήθηκε από τον Bregman (1967). Τέλος σημαντικά εργαλεία στη Στατιστική Θεωρία Πληροφορίας θεωρούνται και τα μέτρα Εντροπίας. Η εισαγωγή τους, είναι απαραίτητη για τον ορισμό των Rφ - απόκλισης μέτρων. Στην ενότητα αυτή υποθέτουμε, όπως και στην προηγούμενη, ότι το ολοκλήρωμα στους ορισμούς υπάρχει.

1.3.1 Εντροπία

Έστω X μια τυχαία μεταβλητή με κατανομή πιθανότητας P_θ.Το πρώτο μέτρο εντοπίας ήταν του Shannon (1948):

$$\mathrm{H}(\mathrm{x}) \equiv \mathrm{H}(\mathrm{P}_\theta) \equiv \mathrm{H}(\theta) = -\int_x f_\theta(\mathrm{x}) \log f_\theta(\mathrm{x}) d\mu(\mathrm{x}) = \mathrm{E}_\theta\left[-\log f_\theta(X)\right]$$

Η απόκλιση Kullback-Leiber σχετίζεται με την εντροπία του Shannon. Αν υποθέσουμε ότι η πιθανότητα κατανομής $P_{\theta 2}$ είναι η ομοιόμορφη κατανομή τότε έχουμε :

$$D_{kull}(\theta_1, \theta_2) = \mathrm{H}(\mathrm{P}_{\theta_2}) - \mathrm{H}(\mathrm{P}_{\theta_1}).$$

Γενίκευση της εντροπίας του Shannon προτάθηκε από τον Renyi (1961) :

$$H_r^1(\theta) = \frac{1}{1-r} \log \int_x f_\theta(\mathrm{x})^r d\mu(\mathrm{x}) = \frac{1}{1-r} \log \mathrm{E}_\theta\left[f_\theta(X)^{r-1}\right], r > 0, r \neq 1$$

Οι Liese και Vajda (1987) επέκτειναν την εντροπία Renyi για όλα τα $\mathrm{r} \in \mathrm{R}\text{-}\{0,1\}$:

$$H_r^1(\theta) = \frac{1}{r(1-r)} \log \int_x f_\theta(\mathrm{x})^r d\mu(\mathrm{x}) = \frac{1}{r(1-r)} \log \mathrm{E}_\theta\left[f_\theta(X)^{r-1}\right], r \neq 0,1 \qquad (1.21)$$

Η (1.21) θα αναφέρεται ως εντροπία Renyi. Αν και δεν ορίζεται για $\mathrm{r} = -1$ ή $\mathrm{r} = 0$ αν πάρουμε τα συνεχή όρια του $\mathrm{H}_r^1(\theta)$ όταν $\mathrm{r} \to 1$ και $\mathrm{r} \to 0$ τότε $\mathrm{H}_r^1(\theta)$ είναι συνεχής στο r. Δεν είναι δύσκολο να αποδειχθεί ότι:

$$\lim_{r\to1} \mathrm{H}_r^1(\theta) = \mathrm{H}(\theta) \text{ και } \lim_{r\to0} \mathrm{H}_r^1(\theta) = \int_x \log f_\theta(\mathrm{x}) d\mu(\mathrm{x})$$

Για να έχουμε ένα συστηματικό τρόπο μελέτης των διαφόρων μέτρων της εντροπίας οι Burbea και Rao εισήγαγαν τις λεγόμενες φ-εντροπίες :

$$\mathrm{H}_{\varphi}(\mathrm{x}) = \mathrm{H}_{\varphi}(\mathrm{P}_{\theta}) = \mathrm{H}_{\varphi}(\theta) = \int_{x} \varphi(f_{\theta}(\mathrm{x}) d\mu(\mathrm{x}) \qquad (1.22)$$

όπου $\varphi:(0,\infty) \rightarrow \mathrm{R}$ είναι μια συνεχής κοίλη συνάρτηση και $\varphi(0) = \lim_{t\downarrow 0}\varphi(t) \in (-\infty,\infty)$.

Με τις φ-εντροπίες συναντάμε το ίδιο πρόβλημα που συναντήσαμε με τις φ-αποκλίσεις. Έτσι μερικά σημαντικά μέτρα εντροπίας δεν γράφονται σαν φ-εντροπίες. Για το λόγο αυτό ο Salicrú et. al (1993) όρισε τις **(h, φ) εντροπίες** ως εξής:

$$\mathrm{H}_{\varphi}^{h}(\mathrm{x}) \equiv \mathrm{H}_{\varphi}^{h}(\mathrm{P}_{\theta}) \equiv \mathrm{H}_{\varphi}^{h}(\theta) = h\left(\int_{x} \varphi\left(f_{\theta}(\mathrm{x})\right) d\mu(\mathrm{x}) \right) \qquad (1.23)$$

όπου

$\varphi : (0,\infty) \rightarrow R$ είναι κύλη και h: $\mathrm{R} \rightarrow \mathrm{R}$ είναι διαφορίσιμη και αύξουσα ή

$\varphi : (0,\infty) \rightarrow R$ είναι κυρτή και h: $\mathrm{R} \rightarrow \mathrm{R}$ είναι διαφορίσιμη και φθίνουσα.

Ιδιότητες Εντροπίας

i) Η εντροπία του Shanon του X μπορεί να είναι αρνητική.

ii) Έστω $\varphi = (\varphi_1,.......\varphi_n)$ ομαλή bijection ("1-1" και επί) στο R^n και θεωρούμε $\mathrm{Y} = \varphi(\mathrm{x})$, τότε :

$\mathrm{H(Y)} = \mathrm{H(X)} - \int_{R^n} f(\mathrm{x})\log\left|J(\varphi(\mathrm{x}))\right|dx$ όπου $\mathrm{J(y)} = \det(\frac{d\psi_i}{dy_j}(y))$ $\mathrm{i,j} = 1,.....,\mathrm{n}$.

iii) Αν $\varphi = (\varphi_1,.....,\varphi_n)$ είναι γραμμικός μετασχηματισμός με $\varphi_i(\mathrm{x}) = \sum_{j=1}^{n} \alpha_{ij}x_j, i = 1,...,n$ τότε :

$H(Y) = H(X) + \log\left|\det(\mathrm{A})\right|$, όπου $\mathrm{A} = (\alpha_{ij}), \mathrm{i, j} = 1,....,\mathrm{n}$.

iv) Ισχύει :

$$\mathrm{H(X)} = -\int_{R^n} f_1(\mathrm{x})\log f_1(\mathrm{x})\,\mathrm{dx} \leq -\int_{R^m} f_1(\mathrm{x})\log f_2(\mathrm{x})\,\mathrm{dx}$$

Αν $f_1(\mathrm{x}) = f_2(\mathrm{x})$ τότε η ανισότητα ονομάζεται Gibbs's Lemma για συνεχής τυχαίες μεταβλητές.

1.3.2 Burbea και Rao μέτρα απόκλισης

Ο Pardo εισήγαγε την R_φ^h απόκλιση μεταξύ δύο κατανομών πιθανοτήτων P_{θ_1} και P_{θ_2} ως γενίκευση της (h,φ) εντροπίας:

$$R_\varphi^h(P_{\theta_1}, P_{\theta_2}) \equiv R_\varphi^h(\theta_1, \theta_2) \equiv H_\varphi^h\left(\frac{P_{\theta_1} + P_{\theta_2}}{2}\right) - \frac{H_\varphi^h(P_{\theta_1}) + H_\varphi^h(P_{\theta_2})}{2}$$

Για h(x) = x έχουμε την R_φ – απόκλιση του Burbea και Rao

Για φ(x) = xlogx έχουμε την ακτίνα πληροφοριών του Sibson.

Μια σημαντική οικογένεια από R_φ-αποκλίσεις βασίζεται στις φ_α-εντροπίες. Αυτή η οικογένεια εντροπιών περιέχει την οικογένεια των συναρτήσεων:

$$\varphi_\alpha(x) = \begin{cases} -x \log x, & \alpha = 1 \\ (1-\alpha)^{-1}\ (x^\alpha - x), & \alpha \neq 1 \end{cases}$$

Άλλη σημαντική οικογένεια μέτρων R_φ – απόκλισης είναι αν θεωρήσουμε τη Bose-Einstein (1984) εντροπία που εισήγαγε ο Burbea και για την οποία η φ δίνεται από τον τύπο:

$$\varphi(x) = \frac{x^5 - (1+x)^5 + 1 + (s-1)^{-1}(2^s - 2)\,x}{(s-2)},$$

όπου για s = 2 είναι συνεχής.

Άλλη σημαντική οικογένεια είναι η Fermi-Dirac εντροπία που εισήγαγε ο Kapur (1972) και η συνάρτηση φ δίνεται από τον τύπο:

$$\varphi(x) = \frac{x^5 + (1-x)^5 - 1}{(1-s)},$$

όπου για s = 1 είναι συνεχής.

Στο πλαίσιο αυτό οι Pardo και Vadja (1997) διαπίστωσαν ότι :

$$\frac{1}{t}\varphi(\mathrm{t}\frac{u+v}{2})-\frac{\varphi(tu)+\varphi(tv)}{2t}=\varphi(\frac{u+v}{2})-\frac{\varphi(u)+\varphi(v)}{2}$$

Ισχύει $\forall$ t,u,v $\Rightarrow D_{\varphi}(\mathrm{P}_{\theta_1},\mathrm{P}_{\theta_2})=R_{\varphi}(\mathrm{P}_{\theta_1},\mathrm{P}_{\theta_2})$

όπου $D_{\varphi}(\mathrm{P}_{\theta_1},\mathrm{P}_{\theta_2})$ είναι φ-απόκλιση μεταξύ $\mathrm{P}_{\theta_1},\mathrm{P}_{\theta_2}$ για $\varphi(\mathrm{x})=\varphi\left(\frac{x+1}{2}\right)-\frac{\varphi(\mathrm{x})+\varphi(1)}{2}$.

1.3.3 Bregman απόκλιση

Ο Bregman εισήγαγε την οικογένεια αποκλίσεων

$$\mathrm{B}_{\varphi}(\theta_1,\theta_2)=\int_{x}\varphi(f_{\theta_1}(\mathrm{x}))-\varphi(f_{\theta_2}(\mathrm{x}))-\varphi'(f_{\theta_2}(\mathrm{x}))\ (f_{\theta_1}(\mathrm{x})-f_{\theta_2}(\mathrm{x}))d\mu(\mathrm{x})$$

για κάθε διαφορίσιμη κυρτή συνάρτηση φ : $(0,\infty)\to \mathrm{R}$ με $\varphi(0)=\lim_{t\to 0}\varphi(t)\ \in(-\infty,\infty)$.
Παρατηρούμε ότι για φ(t) = tlogt , $\mathrm{B}_{\varphi}(\theta_1,\theta_2)$ είναι η Kullback-Leibler απόκλιση και για $\varphi(t)=t^2$ και διακριτή πιθανότητα κατανομής είναι η Ευκλείδια απόσταση.

1.4 Απόκλιση μεταξύ k πληθυσμών

Τα μέτρα απόκλισης που είδαμε προηγουμένως σχεδιάστηκαν για 2 πιθανότητες κατανομών. Μια γενική κατηγορία των μέτρων απόκλισης που ονομάζεται f-ανομοιομορφία μεταξύ κ πληθυσμών ορίστηκε από τους Gyorfi και Nemetz (1987) ως εξής :

$$D_F(\theta_1,......,\theta_\kappa)=\int_{x}F(f_{\theta_1}(\mathrm{x}),...,f_{\theta_\kappa}(\mathrm{x}))\,\mathrm{d}\,\mu(\mathrm{x})$$

όπου f είναι μια συνεχής, ομοιογενής (Μια $\mathrm{f}=\mathrm{f}(\mathrm{x},\mathrm{y})$ ορισμένη σε ένα υποσύνολο D του R^2 λέγεται ομογενής αν για κάθε (x,y) και για κάθε t > 0 ισχύει ότι (tx,ty) $\in$ D),
κυρτή συνάρτηση που ορίζεται στο σύνολο
$\mathrm{S}=\{(\mathrm{s}_1,...,s_k):0\le \mathrm{s}_i<\infty, i=1,....,k\}$.

Άλλη οικογένεια που ορίστηκε από τους Burbea και Rao (1982) είναι η φ-Jensen:
Η φ-Jensen διαφορά μεταξύ των πιθανοτήτων κατανομών $\mathrm{P}_{\theta_i}, i = 1,...,k$ είναι:

$$\mathrm{R}_{\varphi}(\mathrm{P}_{\theta_1},....,\mathrm{P}_{\theta_\kappa}) = \mathrm{H}_{\varphi}(\lambda_1 \mathrm{P}_{\theta_1} + ... + \lambda_\kappa \mathrm{P}_{\theta_\kappa}) - \sum_{i=1}^{k} \lambda_i \mathrm{H}_{\varphi}(\mathrm{P}_{\theta_i})$$

όπου $\sum_{i=1}^{k} \lambda_i = 1$ και H_{φ} είναι φ-εντροπία. Στην περίπτωση που φ(t) = -tlogt έχουμε την λεγόμενη πληροφορία Radius για k πιθανότητες κατανομών.

1.5 Φ-διαφορές

Η φ-διαφορά εμφανίστηκε από τον Lindsay όπου διαπίστωσε ότι για την φ-απόκλιση απαιτείται η διαφορισιμότητα της φ. Επειδή οι ιδιότητες αυτές δεν ικανοποιούνται στο (0, ∞) η φ-διαφορά είναι επέκταση της φ-απόκλισης.

Ορισμός 1.2

Η φ-διαφορά μεταξύ 2 πιθανοτήτων κατανομών P_{θ_1} και P_{θ_2} ορίζεται ως εξής :

$$D_{\varphi}(\mathrm{P}_{\theta_1}, \mathrm{P}_{\theta_2}) = \mathrm{D}_{\varphi}(\theta_1, \theta_2) = \int_{x} f_{\theta_2}(\mathrm{x}) \varphi\left(\frac{f_{\theta_1}(\mathrm{x})}{f_{\theta_2}(\mathrm{x})}\right) d\mu(\mathrm{x})$$

όπου η συνάρτηση φ:$(0,\infty) \to [0,\infty)$ είναι συνεχής, φθίνουσα στο (0,1) και αύξουσα στο $(1,\infty)$ με φ(1) = 0. Η τιμή στο φ(0) $\in (0,\infty)$ ορίζεται από τη συνεχή επέκταση.

Σχόλιο:

Η κλάση της φ-διαφοράς περιέχει όλες τις φ-αποκλίσεις με φ:$(0,\infty) \to (0,\infty)$ είναι κυρτή και ισούται με το μηδέν μόνο στο ένα. Από την υποτιθέμενη κυρτότητα και φ(1) = 0 συνεπάγεται ότι

$$\frac{\varphi(t) - \varphi(1)}{t-1} = \frac{\varphi(t)}{t-1}$$

είναι μη συνεχής για t > 0.

Κεφάλαιο 2

Έλεγχοι καλής προσαρμογής: Βελτιστοποίηση της Φ-απόκλισης

2.1 Εισαγωγή: Καλή προσαρμογή - Απλή μηδενική υπόθεση

Το πρόβλημα της καλής προσαρμογής για τον έλεγχο $\mathrm{H}_0 : \mathrm{F} = \mathrm{F}_0,$ όπου F_0 γνωστή κατανομή, αντιμετωπίζεται με το διαχωρισμό του φάσματος των τιμών σε ξένα μεταξύ τους διαστήματα και δοκιμάζοντας την υπόθεση $\mathrm{H}_0 : \rho = \rho^0$ για το διάνυσμα των παραμέτρων ρ της πολυμεταβλητής κατανομής που προκύπτει.

Έστω ότι P=$\{\mathrm{E}_i\}, \mathrm{i} = 1,....., \mathrm{M}$ είναι μια διαμέριση της πραγματικής ευθείας R σε M διαστήματα. Έστω ρ = $(\rho_1,....., \rho_\mathrm{M})^\mathrm{T}$ οι πραγματικές και $\rho^0 = (\rho_1^0,....., \rho_\mathrm{M}^0)^\mathrm{T}$ οι υποθετικές πιθανότητες των διαστημάτων $\mathrm{E}_i, i = 1,...., M$ με τέτοιο τρόπο ώστε :

$$\rho_i = \mathrm{Pr}_F(\mathrm{E}_i), \mathrm{i} = 1,..., \mathrm{M}$$

και

$$\rho_i^0 = \mathrm{Pr}_{Fo}(E_i) = \int_{E_i} dF_0 \quad \mathrm{i} = 1,...., \mathrm{M}.$$

Έστω $\mathrm{Y}_1,....., Y_n$ τυχαίο δείγμα από την F έστω $\mathrm{N}_i = \sum_{j=1}^{n} I_{E_i}(Y_j),$ με $I_{E_i}(Y_j) = 1$ αν $Y_j \in \mathrm{E}_i$ και μηδέν αλλιώς οι απόλυτες και $\hat{\rho}=(\hat{\rho}_1,...., \hat{\rho}_\mathrm{M})$ με $\hat{\rho}_i = \dfrac{N_i}{n}, i = 1,..., \mathrm{M}$ οι σχετικές συχνότητες στα διαστήματα αντίστοιχα.

Αν θέλουμε να ελέγξουμε την απλή μηδενική υπόθεση $\mathrm{H}_0 : \rho = \rho^0$ (2.1)

το πιο σύνηθες τεστ είναι του Pearson X^2 :

$$\mathrm{X}^2 \equiv \sum_{i=1}^{M} \frac{(N_i - np_i^0)^2}{np_i^o} \tag{2.2}$$

και το test του λόγου πιθανοφανειών G^2 :

$$\mathrm{G}^2 \equiv 2\sum_{i=1}^{M} N_i \log \frac{N_i}{np_i^0} \tag{2.3}$$

Αυτά τα δύο στατιστικά τεστ είναι ειδικές περιπτώσεις της οικογένειας Power divergence (δυναμική απόκλιση) των Cressie και Read (1984) και δίνεται :

$$\mathrm{T}_n^{\lambda}(\hat{\rho},\rho^0)=\frac{2n}{\lambda(\lambda+1)}\sum_{i=1}^{M}\hat{p}_i((\frac{\hat{p}_i}{p_i^0})^{\lambda}-1)=\frac{2}{\lambda(\lambda+1)}\sum_{i=1}^{M}N_i((\frac{N_i}{np_i^0})^{\lambda}-1) \qquad (2.4)$$

όπου $-\infty<\lambda<\infty$ με
$\mathrm{T}_n^{0}(\hat{\rho},\rho^0)$ και $\mathrm{T}_n^{1}(\hat{\rho},\rho^0)$ να ορίζονται από τα όρια $\mathrm{T}_n^{\lambda}(\hat{\rho},\rho^0)$ όταν $\lambda\to 0$ και $\lambda\to 1$ αντίστοιχα.

Ειδικές τιμές του λ στην (2.4) αντιστοιχούν σε γνωστά στατιστικά τεστ.

X^2 test : $\mathrm{X}^2(\lambda=1)$
G^2 λόγο πιθανοφανειών : $\mathrm{G}^2(\lambda=0)$
Λόγο πιθανοφάνειας Kullback : λ = -1
Neyman-τροποποιημένο X^2 test : λ = -2
Gressie-Read στατιστική δοκιμή : λ = $\frac{2}{3}$

Αν και τα στατιστικά τεστ δυναμικής-απόκλισης είναι μια σημαντική ευέλικτη οικογένεια είναι δυνατόν να εξεταστεί μια πιο γενική οικογένεια στατιστικών ελέγχων για τη μελέτη της (2.1) η οποία περιέχει τη (2.4) ως μια ειδική περίπτωση των στατιστικών δοκιμών φ-απόκλισης test όπου ορίζεται ως εξής:

$$\mathrm{T}_n^{\varphi}(\hat{\rho},\rho^0)=\frac{2n}{\varphi''(1)}\sum_{i=1}^{M}p_i^0\varphi\left(\frac{\hat{p}_i}{p_i^0}\right),\ \ \varphi\in\Phi^* \qquad (2.5)$$

όπου φ(x) είναι συνεχής διαφορίσιμη $\forall x>0$ και η δεύτερη παράγωγος $\varphi''(1)\neq 0$.

Η πιο γενική περίπτωση μέτρων απόκλισης είναι η φ_β-απόκλιση η οποία προτάθηκε από τους Mattheou & Karagrigoriou (2010) και ορίζεται ως εξής:

$$T_n^{\varphi}(\hat{p}, p^0) = \frac{2n}{\varphi''(1)} D_{\varphi}(\hat{p}, \rho^0) \ ,$$

όπου

$$D_{\varphi}(\hat{\rho}, \rho^0) = \sum_{j=1}^{M} (\rho_j^0)^{\beta} \varphi(\frac{\hat{\rho}_j}{\rho_j^0})$$

Προφανώς για β=1 προκύπτει η στατιστική συνάρτηση (2.5).

Οι Cressie και Read απέδειξαν την ασυμπτωτική κατανομή της δυναμικής-απόκλισης test $T_n^{\lambda}(\hat{\rho}, \rho^0)$ κατω από $H_0 : \rho = \rho^0, \forall \lambda \in R$ και οι Zografos et. al (1991) επέκτεινε το αποτέλεσμα στην οικογένεια $T_n^{\varphi}(\hat{\rho}, \rho^0)$ κατω από $H_0 : \rho = \rho^0, \forall \varphi \in \Phi^*$. Το αντίστοιχο αποτέλεσμα για την φ_{β}-απόκλιση οφείλεται στους Mattheou & Karagrigoriou (2010). Αυτό θα δούμε στην ενότητα (2.2) όχι μόνο κάτω από μηδενικές υποθέσεις αλλά και κάτω από συνεχείς ενναλακτικές υποθέσεις.

2.2 Ασυμπτωτική κατανομή Φ-αποκλίσεων

Σ΄αυτή την ενότητα δίνεται η ασυμπτωτική κατανομή της φ-απόκλισης $T_n^{\varphi}(\hat{\rho}, \rho^0)$ κάτω από τη μηδενική υπόθεση $H_0 : \rho = \rho^0$ έναντι της εναλλακτικής υπόθεσης $H_1 : \rho = \rho^* \neq \rho^0$.

Ο Pearson απέδειξε ότι η στατιστική συνάρτηση (2.2) ικανοποιεί τη σχέση

$$X^2 \to X_{M-1}^2$$

όπου X_k^2 είναι η χι-τετράφωνο κατανομή με k β.ε.

Υπενθυμίζεται ότι το στατιστικό test $T_n^{\lambda}(\hat{\rho}, \rho^0)$ συμπίπτει με το στατιστικό τεστ X^2 για λ =1.

Οι Cressie και Read απέδειξαν ότι:

$$T_n^{\lambda}(\hat{\rho}, \rho^0) \xrightarrow[n\to\infty]{L} X_{M-1}^2$$

κάτω από $H_0 : \rho = \rho^0, \forall \lambda \in R$.

Τέλος οι Zografos et. al (1990) απέδειξαν ότι:

$$\mathrm{T}_n^{\varphi}(\hat{\rho},\rho^0) \xrightarrow[n\to\infty]{L} X_{M-1}^2, \forall \varphi \in \Phi*$$

κάτω από $\mathrm{H}_0 : \rho = \rho^0$.

Τα πιο πάνω αποτελέσματα συνοψίζονται στο κάτωθι θεώρημα:

Θεώρημα 2.1

α)Κάτω από τη μηδενική υπόθεση $\mathrm{H}_0 : \rho = \rho^0 = (\rho_1^0,....,\rho_{\mathrm{M}}^0)^{\mathrm{T}}$ η ασυμπτωτική κατανομή της φ-απόκλισης $\mathrm{T}_n^{\lambda}(\hat{\rho},\rho^0)$ είναι η X^2 με M-1 βαθμούς ελευθερίας.

(β) Κάτω από τη μηδενική υπόθεση $\mathrm{H}_0 : \rho = \rho^0$ η ασυμπτωτική κατανομή της φ-απόκλισης στατιστικού τεστ $\mathrm{T}_n^{\varphi}(\hat{\rho},\rho^0)$ είναι η X^2 με M-1 βαθμούς ελευθερίας.

γ)Κάτω από τη μηδενική υπόθεση $\mathrm{H}_0 : \rho = \rho^0 = (\rho_1^0,....,\rho_{\mathrm{M}}^0)^{\mathrm{T}}$ η ασυμπτωτική κατανομή της φ_{β}-απόκλισης $\mathrm{T}_n^{\varphi}(\hat{\rho},\rho^0)$ είναι η X^2 με M-1 βαθμούς ελευθερίας.

Σαν απόρροια του θεωρήματος προκύπτει η ασυμπτωτική κατανομή όλων των ειδικών περιπτώσεων που εμπίπτουν στις πιο πάνω οικογένειες στατιστικών συναρτήσεων καθώς και στην οικογένεια των h-φ αποκλίσεων. Έτσι ενδεικτικά έχουμε

Σχόλιο:

α) Στην περίπτωση του Kullback-Leiber έχουμε :

$$\mathrm{T}_n^0(\hat{\rho},\rho^0) = 2n\mathrm{D}_{Kull}(\hat{\rho},\rho^0) \xrightarrow[n\to\infty]{L} X_{M-1}^2 \text{ (likelihood ratio test)}$$

και

$$\mathrm{T}_n^0(\rho^0,\hat{\rho}) = 2n\mathrm{D}_{Kull}(\rho^0,\hat{\rho}) \xrightarrow[n\to\infty]{L} X_{M-1}^2 \text{ (modified likelihood ratio test)}$$

b) Στην περίπτωση της (h,φ)– απόκλισης :

$$\mathrm{T}_n^{\varphi,h}(\hat{\rho},\rho^0)=\frac{2n}{h'(0)\varphi''(1)}D_\varphi^h(\hat{\rho},\rho^0)$$

και

$$\mathrm{T}_n^{\varphi,h}(\rho^0,\hat{\rho})=\frac{2n}{h'(0)\varphi''(1)}D_\varphi^h(\rho^0,\hat{\rho})$$

είναι η X^2 με M-1 βαθμούς ελευθερίας.

Με βάση το Θεώρημα (2.1) μπορεί κανείς να χρησιμοποιήσει το 100(1-α) εκατοστημόριο, $\mathrm{X}^2_{\mathrm{M}-1,\alpha}$ της X^2 με M-1 βαθμούς ελευθερίας που ορίζεται από την εξίσωση:
$\Pr(\mathrm{X}^2_{M-1}\geq X^2_{M-1,a})=a$ για να αποφανθεί για τον έλεγχο:

Απορρίπτω την H_0 σε επίπεδο σημαντικότητας α, αν $\mathrm{T}_n^{\varphi}(\hat{\rho},\rho^0)>\mathrm{X}^2_{\mathrm{M}-1,\alpha}$ (2.6)

ενώ διαφορετικά την αποδέχομαι.

Η σημαντικότητα ενός ελέγχου καθορίζεται με βάση την ισχύ του, δηλ. με βάση το πόσο καλά αναγνωρίζει η στατιστική συνάρτηση την εναλλακτική υπόθεση. Το πιο κάτω θεώρημα δίνει την ισχύ για τον έλεγχο που βασίζεται στην φ-απόκλιση.

Θεώρημα 2.2

Έστω $\rho^*=(\rho_1^*,...,\rho_{\mathrm{M}}^*)^{\mathrm{T}}$ είναι μια πιθανότητα κατανομής με $\rho^*\neq\rho^0$. Η ισχύς του τεστ με απόφαση που δίνεται από τη (2.6) στο $\rho^*=(\rho_1^*,...,\rho_{\mathrm{M}}^*)^{\mathrm{T}}$ είναι:

$$\mathrm{B}_{n,\varphi}(\rho_1^*,...,\rho_{\mathrm{M}}^*)=1-\Phi_n\Big(\frac{1}{\sigma_1(\rho^*)}\Big(\frac{\varphi''(1)}{2\sqrt{n}}\mathrm{X}^2_{\mathrm{M}-1,\alpha}-\sqrt{n}D_\varphi(\rho^*,\rho^0)\Big)\Big),$$

όπου Φ_n τείνει ομοιόμορφα στη τυπική κανονική συνάρτηση κατανομής και:

$$\sigma_1^2(\rho^*) = \sum_{i=1}^{M} p_i^* (\varphi'(\frac{\rho_i^*}{\rho_\iota^0}))^2 - (\sum_{i=1}^{M} p_i^* \varphi'(\frac{\rho_i^*}{\rho_\iota^0}))^2 \tag{2.7}$$

Σχόλιο:

Με τον ίδιο τρόπο όπως πριν μπορούμε να υποθέσουμε ότι :

$$\sqrt{n}(D_\varphi(\rho^0, \hat{\rho}) - D_\varphi(\hat{\rho}, \rho^0)) \xrightarrow[n\to\infty]{L} N(o, \sigma_2(\rho^*))$$

$$\sigma_2(\rho^*) = \sum_{i=1}^{M} p_i^* s_i^2 - (\sum_{i=1}^{M} p_i^* s_i^2)^2$$

με

$$s_i = \varphi(\frac{\rho_i^0}{p_i^*}) - \frac{\rho_i^0}{p_i^*}\varphi'(\frac{\rho_i^0}{p_i^*}), \mathrm{i} = 1, \ldots, \mathrm{M}$$

Σχόλιο

α) Στην περίπτωση του Kullback-Leiber μέτρο απόκλισης έχουμε :

$$\sigma_1^2(\rho^*) = \sum_{i=1}^{M} p_i^* (\log\frac{p_i^*}{p_i^0})^2 - (\sum_{i=1}^{M} p_i^* \log\frac{p_i^*}{p_i^0})^2$$

$$\sigma_2^2(\rho^*) = \sum_{i=1}^{M} \frac{(p_i^0)^2}{p_i^*} - 1$$

b) Στην περίπτωση (h,φ)-απόκλιση :

$$\sigma_1^2(\rho^*) = \sum_{i=1}^{M} p_i^* (\mathrm{h}'(\mathrm{D}_\varphi(\rho^*, \rho^0))\varphi'(\frac{p_i^*}{p_i^0}))^2 - (\sum_{i=1}^{M} p_i^* \,\mathrm{h}'(\mathrm{D}_\varphi(\rho^*, \rho^0))\varphi'(\frac{p_i^*}{p_i^0}))^2$$

και

$$\sigma_2^2(\rho^*) = \sum_{i=1}^{M} p_i^* ((\mathrm{h}'(\mathrm{D}_\varphi(\rho^0, \rho^*)))(\varphi(\frac{p_i^0}{p_i^*}) - \frac{p_i^0}{p_i^*}\varphi'(\frac{p_i^0}{p_i^*})))^2 - (\sum_{i=1}^{M} p_i^* (\mathrm{h}'(\mathrm{D}_\varphi(\rho^0, \rho^*)))(\varphi(\frac{p_i^0}{p_i^*}) - \varphi'(\frac{p_i^0}{p_i^*})))^2$$

2.3 Βελτιστοποίηση των φ-αποκλίσεων

Στην προηγούμενη ενότητα μελετήσαμε την οικογένεια της φ-απόκλισης $\mathrm{T}_n^{\varphi}(\hat{\rho}, \rho^0)$ για το πρόβλημα της καλής προσαρμογής. Ας συμβολίσουμε με $\mathrm{F}_{\mathrm{T}_n^{\varphi}(\hat{\rho}, \rho^0)}(\mathrm{t})$ την ακριβή κατανομή της $\mathrm{T}_n^{\varphi}(\hat{\rho}, \rho^0)$ για σταθερό φ. Τότε με βάση τα προηγούμενα θεωρήματα έχουμε

$$\mathrm{F}_{\mathrm{T}_n^{\varphi}(\hat{\rho}, \rho^0)}(\mathrm{t}) = F_{X^2_{M-1}}(\mathrm{t}) + o(1) \text{ καθώς } \mathrm{n} \to \infty \tag{2.10}$$

κάτω από την μηδενική υπόθεση

$$\mathrm{H}_0 : \rho = \rho^0. \tag{2.11}$$

Σε αυτή την ενότητα θα δούμε κάποια κριτήρια για να διαλέξουμε την καλύτερη συνάρτηση φ. Συγκεκριμένα ερευνούμε το κριτήριο με βάση την ταχύτητα της σύγκλισης της ακριβούς ροπής κ-τάξης της $\mathrm{T}_n^{\varphi}(\hat{\rho}, \rho^0)$ στην αντίστοιχη ασυμπτωτική ροπή για όσο το δυνατόν περισσότερα κ είναι εφικτό.
Η ακριβής κατανομή για κάθε μέλος της οικογένειας φ-απόκλισης στατιστικού τεστ $\mathrm{T}_n^{\varphi}(\hat{\rho}, \rho^0)$ διαφέρει, σύμφωνα με τη (2.10) από το X^2 κατά o(1).
Για τις ανάγκες των αποδείξεων υποθέτουμε ότι $\varphi \in \Phi^*$ είναι 4 φορές συνεχώς διαφορίσιμη στη γειτονιά του 1 και $\varphi''(1) \neq 0$.

2.4 Ακριβείς και Ασυμπτωτικές ροπές : Σύγκριση

Η ταχύτητα σύγκλισης των ακριβών ροπών στις ασυμπτωτικές ροπές στην οικογένεια των φ-αποκλίσεων στατιστικών τεστ, μας δίνει πληροφορίες σχετικά με τη ταχύτητα της σύγκλισης της ακριβούς κατανομής στην ασυμπτωτική κατανομή.

Στις αποδείξεις που ακολουθούν θα μελετηθούν τα αναπτύγματα Taylor των τριών πρώτων ακριβών ροπών της φ-απόκλισης $\mathrm{T}_n^{\varphi}(\hat{\rho}, \rho^0)$ και θα συγκριθούν με τις αντίστοιχες ροπές της X^2 κατανομής με Μ-1 βαθμούς ελευθερίας.Τα μεγέθη των όρων διόρθωσης μας δίνουν πληροφορίες σχετικά με τα σφάλματα που κάνουμε όταν χρησιμοποιούμε την ασυμπτωτική κατανομή αντί της ακριβούς κατανομής .

Στην ενότητα 2.4.1 εξετάζεται η περίπτωση της φ-απόκλισης και στην ενότητα 2.4.2 της φ_β-απόκλισης.

2.4.1 Βελτιστοποίηση φ-απόκλισης

Έστω

$$\mu_\beta(T_n^{\varphi}(\hat{\rho},\rho^0)) = E[(T_n^{\varphi}(\hat{\rho},\rho^0))^{\beta}] \text{ ,β = 1,2,3}$$

και υποθέτουμε ότι :

$$\mu_\beta(T_n^{\varphi}(\hat{\rho},\rho^0)) = E[(X_{M-1}^2)^{\beta}] + \frac{1}{n} f_\varphi^{\beta} + O(n^{-\frac{3}{2}}) \text{ ,β = 1,2,3}$$

Υπενθυμίζεται ότι:

$E[(X_{M-1}^2)^{\beta}] = M-1, M^2-1$ ή M^3+3M^2-M-3 αν β = 1,2,3 αντίστοιχα.

Τότε f_φ^{β} ελέγχει την ταχύτητα με την οποία οι 3 πρώτες ακριβείς ροπές της φ-απόκλισης $T_n^{\varphi}(\hat{\rho},\rho^0)$ συγκλίνουν στις 3 πρώτες ροπές της X^2 κατανομής με Μ-1 βαθμούς ελευθερίας. Η συνάρτηση φ για την οποία

$$f_\varphi^{\beta} = 0, \beta = 1,2,3$$

θα είναι η βέλτιστη.

Πρόταση 2.4

Για την ροπή πρώτης τάξης ισχύει ότι:

$E[T_n^{\varphi}(\hat{\rho},\rho^0)] = M-1+\frac{1}{n} f_\varphi^{1} + O(n^{-\frac{3}{2}})$ όπου :

$$f_\varphi^1 = \frac{\varphi'''(1)}{3\varphi''(1)}(2-3M+S) + \frac{\varphi^{IV}(1)}{4\varphi''(1)}(1-2M+S) \quad (2.13)$$

και $S=\sum_{j=1}^{M}(p_j^0)^{-1}$

Απόδειξη 2.4

Έστω $\mathrm{W}_j = \frac{1}{\sqrt{n}}(N_j - np_j^0),\ j = 1,....,M.$

Μιας τέταρτης τάξης έκφραση Taylor του $\mathrm{D}_\varphi(\hat{\mathrm{p}},\rho^0)$ γύρω από το ρ^0 δίνει :

$$\mathrm{D}_\varphi(\mathrm{p}\wedge,\rho^0) = \sum_{j=1}^{M}\left(\frac{\mathrm{dD}_\varphi(\mathrm{p},\rho^0)}{dp_j}\right)_{p=p^0}\frac{w_j}{\sqrt{n}} + \frac{1}{2!}\sum_{j=1}^{M}\left(\frac{\mathrm{d}^2\mathrm{D}_\varphi(\mathrm{p},\rho^0)}{dp_j^2}\right)_{p=p^0}\frac{w_j^2}{n} + \frac{1}{3!}\sum_{j=1}^{M}\left(\frac{\mathrm{d}^3\mathrm{D}_\varphi(\mathrm{p},\rho^0)}{dp_j^3}\right)_{p=p^0}\frac{w_j^3}{n\sqrt{n}}$$

$$+\frac{1}{4!}\sum_{j=1}^{M}\left(\frac{\mathrm{d}^4\mathrm{D}_\varphi(\mathrm{p},\rho^0)}{dp_j^4}\right)_{p=p^0}\frac{w_j^4}{n^2} + O_P(n^{-\frac{5}{2}}).$$

Αλλά

$$\left(\frac{\mathrm{dD}_\varphi(\mathrm{p},\rho^0)}{dp_j}\right)_{p=p^0} = \left(\varphi'\left(\frac{p_j}{p_j^0}\right)\right)_{p=p^0} = \varphi'(1)$$

$$\left(\frac{\mathrm{d}^2\mathrm{D}_\varphi(\mathrm{p},\rho^0)}{dp_j^2}\right)_{p=p^0} = \left(\frac{1}{p_j^0}\varphi''\left(\frac{p_j}{p_j^0}\right)\right)_{p=p^0} = \frac{1}{p_j^0}\varphi''(1)$$

$$\left(\frac{\mathrm{d}^3\mathrm{D}_\varphi(\mathrm{p},\rho^0)}{dp_j^3}\right)_{p=p^0} = \left(\frac{1}{(p_j^0)^2}\varphi'''\left(\frac{p_j}{p_j^0}\right)\right)_{p=p^0} = \frac{1}{(p_j^0)^2}\varphi'''(1)$$

$$\left(\frac{\mathrm{d}^4\mathrm{D}_\varphi(\mathrm{p},\rho^0)}{dp_j^4}\right)_{p=p^0} = \left(\frac{1}{(p_j^0)^3}\varphi^{IV}\left(\frac{p_j}{p_j^0}\right)\right)_{p=p^0} = \frac{1}{(p_j^0)^3}\varphi^{IV}(1)$$

Τότε

$$\mathrm{T}_n^\varphi(\mathrm{p}\wedge,\mathrm{p}^0) = \frac{2n}{\varphi''(1)}\mathrm{D}_\varphi(\mathrm{p}\wedge,\rho^0) = \sum_{j=1}^{M}\frac{w_j^2}{p_j^0} + \frac{\varphi'''(1)}{3\sqrt{n}\varphi''(1)}\sum_{j=1}^{M}\frac{w_j^3}{(p_j^0)^2} + \frac{\varphi^{IV}(1)}{12n\varphi''(1)}\sum_{j=1}^{M}\frac{w_j^4}{(p_j^0)^3} + O_P(n^{-\frac{3}{2}}).(2.14)$$

Από τη (4.9) μπορούμε να γράψουμε:

$$\mathrm{E}[\mathrm{T}_n^\varphi(\mathrm{p}\wedge,\mathrm{p}^0)] = \sum_{j=1}^{M}\frac{\mathrm{E}[w_j^2]}{p_j^0} + \frac{\varphi'''(1)}{3\sqrt{n}\varphi''(1)}\sum_{j=1}^{M}\frac{\mathrm{E}[w_j^3]}{(p_j^0)^2} + \frac{\varphi^{IV}(1)}{12n\varphi''(1)}\sum_{j=1}^{M}\frac{\mathrm{E}[w_j^4]}{(p_j^0)^3} + O(n^{-\frac{3}{2}}).\ (2.15)$$

Αφού $\mathrm{E}[O_P(n^{-\frac{3}{2}})] = O(n^{-\frac{3}{2}})$.

Η ροπή της συνάρτησης μιας πολαπλής τυχαίας μεταβλητής , $\mathrm{N} = (\mathrm{N}_1,....,\mathrm{N}_\mathrm{M})^\mathrm{T}$ με παραμέτρους n και p^0 είναι :

$\mathrm{M}_\mathrm{N}(t) = E[\exp(t^T N)] = (\mathrm{p}_1^0\exp(t_1),......,\mathrm{p}_M^0\exp(t_M))^n$ με $\mathrm{t} = (\mathrm{t}_1,...,t_M)^T$

Η ροπή της συνάρτησης Μ-διάστασης της τυχαίας μεταβλητής $W = \frac{1}{\sqrt{n}}(N - np^0)$ δίνεται από :

$$\mathrm{M}_W(t) = E[\exp(t^T \mathrm{W})] = E[\exp(t^T(N/\sqrt{n} - \sqrt{n}\,p^0))]$$

$$= \exp(-\sqrt{n}t^T p^0)E[\exp(t^T(N/\sqrt{n})] \ (2.16)$$

$= \exp(-\sqrt{n}t^T p^0)M(t/\sqrt{n})$, (2.16)και η α-ροπή της W_j σχετικά με τη προέλευση από :

$\mathrm{E}[\mathrm{W}_j^a] = (\frac{d^a M_W(t)}{dt_j^a})_{t=0}$ (2.17) για j = 1,...M και α = 1,2,.....

Από τις (2.16) και (2.17) έχουμε :

$$E[W_j^2] = -(\mathrm{p}_j^0)^2 + p_j^0$$

$$E[W_j^3] = \mathrm{n}^{-\frac{1}{2}}(2(\mathrm{p}_j^0)^3 - 3(\mathrm{p}_j^0)^2 + p_j^0)$$

$$E[W_j^4] = 3(\mathrm{p}_j^0)^4 - 6(\mathrm{p}_j^0)^3 + 3(\mathrm{p}_j^0)^2 + n^{-1}(-6(\mathrm{p}_j^0)^4 + 12(\mathrm{p}_j^0)^3 - 7(\mathrm{p}_j^0)^2 + p_j^0),$$

και αντικαθιστώντας αυτές τις εκφράσεις στην (2.15) η απόδειξη ολοκληρώθηκε.

Πρόταση 2.5

Για την ροπή δεύτερης τάξης ισχύει ότι:

$\mathrm{E}[\mathrm{T}_n^{\varphi}(\hat{\rho}, \rho^0)^2] = \mathrm{M}^2 - 1 + \frac{1}{n} f_{\varphi}^2 + O(\mathrm{n}^{-\frac{3}{2}})$ όπου :

$$\mathrm{f}_{\varphi}^2 = (2 - 2\mathrm{M} - \mathrm{M}^2 + S) + \frac{2\varphi'''(1)}{3\varphi''(1)}(10 - 13\mathrm{M} - 6M^2(M+8)S) + \frac{1}{3}(\varphi'''(1))^2(4 - 6\mathrm{M} - 3\mathrm{M}^2 + 5S) +$$

$$\frac{\varphi^{IV}(1)}{2\varphi''(1)}(3 - 5\mathrm{M} - 2M^2 + (M+3)S) \ (2.18)$$

και

$$\mathrm{S} = \sum_{j=1}^{M}(\mathrm{p}_j^0)^{-1}$$

Απόδειξη 2.5

Τετραγωνίζοντας και παίρνοντας τις εκφράσεις της (2.14) παίρνουμε :

$$\mathrm{E}[\mathrm{T}_n^{\varphi}(\hat{\mathrm{p}}, \mathrm{p}^0)^2] = \sum_{j=1}^{M}\frac{\mathrm{E}[w_j^4]}{(p_j^0)^2} + \sum_{j\neq i}^{M}\frac{E[\mathrm{W}_j^2 W_i^2]}{p_j^0 p_i^0} + \frac{2\varphi'''(1)}{3\sqrt{n}\varphi''(1)}(\sum_{j=1}^{M}\frac{\mathrm{E}[w_j^5]}{(p_j^0)^3} + \sum_{j\neq i}^{M}\frac{E[\mathrm{W}_j^2 W_i^3]}{p_j^0(p_i^0)^2}) +$$

$$\frac{1}{n}(\frac{\varphi^{IV}(1)}{6\varphi''(1)}(\sum_{j=1}^{M}\frac{\mathrm{E}[w_j^6]}{(p_j^0)^4} + \sum_{j\neq i}^{M}\frac{E[\mathrm{W}_j^2 W_i^4]}{p_j^0(p_i^0)^3}) + (\frac{\varphi'''(1)}{3\varphi''(1)})^2(\sum_{j=1}^{M}\frac{\mathrm{E}[w_j^6]}{(p_j^0)^4} + \sum_{j\neq i}^{M}\frac{E[\mathrm{W}_j^3 W_i^3]}{(p_j^0)^2(p_i^0)^2})) + O(n^{-\frac{3}{2}}) \ (2.19)$$

Από τη (2.16) έχουμε :

$E[W_j^{\alpha} W_i^{b}] = (\frac{d^{a+b} M_W(t)}{dt_j^a dt_i^b})_{t=0}$ (2.20) για j,i=1,...,M και a,b=1,2,...... Τότε :

$$E[W_j^5] = n^{-1/2}(-20(p_j^0)^5 + 50(p_j^0)^4 - 40(p_j^0)^3 + 10(p_j^0)^2)$$
$$+n^{-3/2}(24(p_j^0)^5 - 60(p_j^0)^4 + 50(p_j^0)^3 - 15(p_j^0)^2 + p_j^0), \text{ και}$$
$$E[W_j^6] = -15(p_j^0)^6 + 45(p_j^0)^5 - 45(p_j^0)^4 + 15(p_j^0)^3 + n^{-1}(130(p_j^0)^6 - 390(p_j^0)^5 + 415(p_j^0)^4$$
$$-180(p_j^0)^3 + 25(p_j^0)^2) + O(n^{-2}).$$

Για $j \neq i$ και τα δυο σταθερά ορίζουμε , $p_{ab} = (p_j^0)^a (p_j^0)^b$ τότε

$$E[W_j^2\ W_i^2] = 3p_{22} - p_{21} - p_{12} + p_{11} + n^{-1}(-6p_{22} + 2p_{21} + 2p_{12} - p_{11}),$$
$$E[W_j^2\ W_i^3] = n^{-1/2}(-20p_{23} + 5p_{13} + 15p_{22} - 6p_{12} - p_{21} + p_{11}),$$
$$E[W_j^2\ W_i^4] = -15p_{24} + 18p_{23} + 3p_{14} - 3p_{22} - 6p_{13} + 3p_{12} + \frac{1}{n}(130p_{24} - 156p_{23} - 26p_{14} + 41p_{22} + 42p_{13}$$
$$-p_{21} - 17p_{12} + p_{11}) + O(n^{-2})$$

$$E[W_j^3\ W_i^3] = -15p_{33} + 9p_{32} + 9p_{23} - 9p_{22} + n^{-1}(130p_{33} - 78p_{32} - 78p_{23} + 63p_{22} + 5p_{31}$$
$$+5p_{13} - 6p_{21} - 6p_{12} + p_{11}) + O(n^{-2}),$$

και αντικαθιστώντας αυτές τις εκφράσεις στην (2.19) και απλουστεύοντας έχουμε :

$$E[T_n^{\varphi}(p^{\wedge}, p^0)^2] = M^2 - 1 + \frac{1}{n} f_{\varphi}^2 + O(n^{-3/2}), \text{ όπου}$$
$$f_{\varphi}^2 = (2 - 2M - M^2 + S) + \frac{2\varphi'''(1)}{3\varphi''(1)}(10 - 13M - 6M^2 + (M+8)S)$$
$$+\frac{1}{3}(\frac{2\varphi'''(1)}{\varphi''(1)})^2(4 - 6M - 3M^2 + 5S) + \frac{\varphi^{IV}(1)}{2\varphi''(1)}(3 - 5M - 2M^2 + (M+3)S).$$

Για να αποκτήσουμε την προηγούμενη έκφραση πρέπει να χρησιμοποιήσουμε τις ακόλουθες σχέσεις:

$$\sum_{j=1}^{M}(p_j^0)^2 + \sum_{j\neq i}^{M} p_j^0 p_i^0 = 1\ ,\ \sum_{j\neq i}^{M} p_j^0 = M - 1 \text{ και } \sum_{j\neq i}^{M}\frac{p_j^0}{p_i^0} = S - M.$$

Πρόταση 2.6

Για την ροπή τρίτης τάξης ισχύει ότι:

$$\mathrm{E}[\mathrm{T}_n^{\varphi}(\hat{\rho},\rho^0)^3]=\mathrm{M}^3+3\mathrm{M}^2-\mathrm{M}-3+\frac{1}{n}f_{\varphi}^3+O(\mathrm{n}^{-\frac{3}{2}}) \text{ όπου :}$$

$$f_{\varphi}^3=(26-24M-21M^2-3M^3+(19+3M)S)+\frac{\varphi'''(1)}{\varphi''(1)}(70-81M-64M^2-9M^3+(65+18M$$

$$+M^2)S)+(\frac{\varphi'''(1)}{\varphi''(1)})^2(20-26M-21M^2-3M^3+(25+5M)S)+\frac{3\varphi^{IV}(1)}{4\varphi''(1)}(15-22M-15M^2-2M^3$$

$$+(15+8M+M^2)S) \quad (2.21)$$

και

$$\mathrm{S}=\sum_{j=1}^{M}(\mathrm{p}_j^0)^{-1}$$

Απόδειξη 2.6

Λαμβάνοντας τις εκφράσεις από τη (2.14) έχουμε :

$$\mathrm{E}[\mathrm{T}_n^{\varphi}(\hat{\mathrm{p}},\mathrm{p}^0)^3]=\sum_{j=1}^{M}\frac{\mathrm{E}[w_j^6]}{(p_j^0)^3}+3\sum_{j\neq i}^{M}\frac{\mathrm{E}[w_j^4\,\mathrm{w}_i^2]}{(p_j^0)^2\,p_i^0}+\sum_{j\neq i\neq k}^{M}\frac{\mathrm{E}[w_j^2\,\mathrm{w}_i^2\mathrm{w}_k^2]}{p_j^0 p_i^0 p_k^0}+\frac{\varphi'''(1)}{\sqrt{n}\varphi''(1)}(\sum_{j=1}^{M}\frac{\mathrm{E}[w_j^7]}{(p_j^0)^4}+\sum_{j\neq i}^{M}\frac{\mathrm{E}[w_j^4\,\mathrm{w}_i^3]}{(p_j^0)^2(p_i^0)^2}$$

$$+2\sum_{j\neq i}^{M}\frac{\mathrm{E}[w_j^2\,\mathrm{w}_i^5]}{p_j^0(p_i^0)^3}+\frac{\mathrm{E}[w_j^2\,\mathrm{w}_i^2\mathrm{w}_k^3]}{p_j^0 p_i^0(p_k^0)^2})+\frac{1}{n}(\frac{1}{3}(\frac{\varphi'''(1)}{\varphi''(1)})^2(\sum_{j=1}^{M}\frac{\mathrm{E}[w_j^8]}{(p_j^0)^5}+\sum_{j\neq i}^{M}\frac{\mathrm{E}[w_j^6\,\mathrm{w}_i^2]}{(p_j^0)^4\,p_i^0}+2\sum_{j\neq i}^{M}\frac{\mathrm{E}[w_j^3\,\mathrm{w}_i^5]}{(p_j^0)^2(p_i^0)^3}$$

$$+\sum_{j\neq i\neq k}^{M}\frac{\mathrm{E}[w_j^3\,\mathrm{w}_i^3\mathrm{w}_k^2]}{(p_j^0)^2(p_i^0)^2\,p_k^0}+\frac{\varphi^{IV}(1)}{4\varphi''(1)}(\sum_{j=1}^{M}\frac{\mathrm{E}[w_j^8]}{(p_j^0)^5}+\sum_{j\neq i}^{M}\frac{\mathrm{E}[w_j^4\,\mathrm{w}_i^4]}{(p_j^0)^2(p_i^0)^3}+2\sum_{j\neq i}^{M}\frac{\mathrm{E}[w_j^2\,\mathrm{w}_i^6]}{p_j^0(p_i^0)^4}+\sum_{j\neq i\neq k}^{M}\frac{\mathrm{E}[w_j^2\,\mathrm{w}_i^2\mathrm{w}_k^4]}{p_j^0 p_i^0(p_k^0)^3}))$$

$$+O(n^{-3/2}). \qquad (2.22)$$

Χρησιμοποιώντας ξανά τη (2.16) παίρνουμε:

$$\mathrm{E}[\mathrm{w}_j^7]=n^{-1/2}(210(p_j^0)^7-735(p_j^0)^6+945(p_j^0)^5-525(p_j^0)^4+105(p_j^0)^3)+O(n^{-3/2}).$$

$$\mathrm{E}[\mathrm{w}_j^8]=105(p_j^0)^8-420(p_j^0)^7+630(p_j^0)^6-420(p_j^0)^5+105(p_j^0)^4+O(n^{-1}).$$

Για $\mathrm{j}\neq \mathrm{i}\neq \mathrm{k}$ όλα σταθερά ορίζουμε $\mathrm{p}_{ab}=(p_j^0)^a(p_i^0)^b$ και $\mathrm{p}_{abc}=(p_j^0)^a(p_i^0)^b(p_k^0)^c$.

Τότε

$$\mathrm{E}[\mathrm{w}_j^4\,\mathrm{w}_i^3]=\mathrm{n}^{-1/2}(210p_{43}-105p_{42}-210p_{33}+3p_{41}+144p_{32}+36p_{23}-6p_{31}-39p_{22}+3p_{21})+O(n^{-3/2}).$$

$$\mathrm{E}[\mathrm{w}_j^4\,\mathrm{w}_i^4]=105p_{44}-90(p_{43}+p_{34})+9(p_{24}+3p_{42})+108p_{33}-18(p_{32}+p_{23})+9p_{22}+O(n^{-1}).$$

$$\mathrm{E}[\mathrm{w}_j^5\,\mathrm{w}_i^2] = \mathrm{n}^{-1/2}(210\,p_{52} - 35\,p_{51} - 350\,p_{42} + 80\,p_{41} + 150\,p_{32} - 55\,p_{31} - 10\,p_{22} + 10\,p_{21}) + O(n^{-3/2}),$$
$$\mathrm{E}[\mathrm{w}_j^5\,\mathrm{w}_i^3] = 105\,p_{53} - 45\,p_{52} - 150\,p_{43} + 90\,p_{42} + 45\,p_{33} - 45\,p_{32} + O(n^{-1}),$$
$$\mathrm{E}[\mathrm{w}_j^6\,\mathrm{w}_i^2] = 105\,p_{62} - 15\,p_{61} - 225\,p_{52} + 45\,p_{51} + 135\,p_{42} - 45\,p_{41} - 15\,p_{32} + 15\,p_{31} + O(n^{-1}),$$
$$\mathrm{E}[\mathrm{w}_j^2\,\mathrm{w}_i^2\mathrm{w}_k^2] = -15\,p_{222} + 3(p_{122} + p_{212} + p_{221}) - (p_{112} + p_{121} + p_{211}) + p_{111} + n^{-1}(130\,p_{222} - 26(p_{122} + p_{212} + p_{221})$$
$$+7(p_{112} + p_{121} + p_{211}) - 3\,p_{111}) + O(n^{-2}),$$
$$\mathrm{E}[\mathrm{w}_j^2\,\mathrm{w}_i^2\mathrm{w}_k^3] = n^{-1/2}(210\,p_{223} - 105\,p_{222} - 35(p_{123} + p_{213}) + 8\,p_{113} + 24(p_{112} + p_{212}) + 3\,p_{221} - 9\,p_{112}$$
$$-(p_{211} + p_{121}) + p_{111}) + O(n^{-3/2}),$$
$$\mathrm{E}[\mathrm{w}_j^2\,\mathrm{w}_i^3\mathrm{w}_k^3] = 105\,p_{233} - 15\,p_{133} - 45(p_{223} + p_{232}) + 9(p_{132} + p_{123}) + 27\,p_{222} - 9\,p_{122} + O(n^{-1}),$$
$$\mathrm{E}[\mathrm{w}_j^2\,\mathrm{w}_i^2\mathrm{w}_k^4] = 105\,p_{224} - 15(p_{124} + p_{214}) - 90\,p_{223} + 3\,p_{114} + 18(p_{123} + p_{213}) + 9\,p_{222} - 6\,p_{113}$$
$$-3(p_{122} + p_{212}) + 3\,p_{112} + O(n^{-1}),$$

και από τη (2.22) έχουμε :

$$\mathrm{E}[\mathrm{T}_n^{\varphi}(\hat{\mathrm{p}},\mathrm{p}^0)^3] = M^3 + 3M^2 - M - 3 + \frac{1}{n} f_{\varphi}^3 + O(n^{-3/2})$$

όπου

$$f_{\varphi}^3 = (26 - 24M - 21M^2 - 3M^3 + (19 + 3M)S) + \frac{\varphi'''(1)}{\varphi''(1)}(70 - 81M - 64M^2 - 9M^3 + (65 + 18M$$
$$+M^2)S) + (\frac{\varphi'''(1)}{\varphi''(1)})^2(20 - 26M - 21M^2 - 3M^3 + (25 + 5M)S) + \frac{3\varphi^{IV}(1)}{4\varphi''(1)}(15 - 22M - 15M^2 - 2M^3$$
$$+(15 + 8M + M^2)S).$$

Για το περιεχόμενο της προηγούμενης έκφρασης πρέπει να χρησιμοποιήσουμε τις ακόλουθες σχέσεις

$$\sum_{j\neq i}^{M} \frac{(p_i^0)^2}{p_i^0} + \sum_{j\neq i\neq k}^{M} \frac{p_j^0 p_i^0}{p_k^0} = S - 2M + 1 \text{ και}$$
$$\sum_{j=1}^{M} (p_j^0)^3 + 3\sum_{j\neq i}^{M} (p_j^0)^2\, p_i^0 + \sum_{j\neq i\neq k}^{M} p_j^0 p_i^0 p_k^0 = 1.$$

Σχόλιο:

Είναι σαφές ότι $\mathrm{f}_{\varphi}^{i}, \mathrm{i} = 1,2,3$ ελέγχει την ταχύτητα με την οποία οι τρεις πρώτες ακριβείς ροπές της φ-απόκλισης $\mathrm{T}_n^{\varphi}(p\wedge, p^0)$ συγκλίνουν στις τρεις πρώτες ροπές της τυχαίας μεταβλητής X^2 με M-1 βαθμούς ελευθερίας.

Ας θεωρήσουμε μια συνάρτηση $\varphi \in \Phi^*$ που εξαρτάται από μια παράμετρο ή δείκτη "α" και την οποία θα τη δηλώνουμε με $\varphi \equiv \varphi_{\alpha}$.

Κάτω από τη μηδενική υπόθεση $H_0 : \rho = \rho^0 = (\frac{1}{M},...,\frac{1}{M})^T$ έχουμε

$S = \sum_{j=1}^{M} (p_j^0)^{-1} = M^2$

Ωστόσο για αυξανόμενο M, οι ρίζες της εξίσωσης $f_{\varphi\alpha}^{i} = 0$ συγλίνουν στις ρίζες της εξίσωσης $4\varphi_{\alpha}'''(1) + 3\varphi_{\alpha}^{IV}(1) = 0,$ για $\varphi'''(\alpha) \neq 0$ (2.23).

Ως εκ τούτου η (2.13) παίρνει τη μορφή:

$$4\varphi_{\alpha}'''(1)(\frac{2-3M+M^2}{1-2M+M^2}) + 3\varphi_{\alpha}^{IV}(1) = 0$$

Παρατηρούμε ότι οι ρίζες της τελευταίας εξίσωσης συγκλίνουν στις ρίζες της εξίσωσης

$$4\varphi_{\alpha}'''(1) + 3\varphi_{\alpha}^{IV}(1) = 0, \text{ για } \varphi'''(\alpha) \neq 0 \qquad (2.23)$$

καθώς $M \to \infty$.

Αν θεωρήσουμε την power-divergence με $\varphi \equiv \varphi(\lambda)$, οι ρίζες της εξίσωσης (2.23) είναι λ = 1 και $\lambda = \frac{2}{3}$ οι οποίες αποτελούν τις ιδανικές τιμές του λ σύμφωνα με τους Read & Cressie (1988).

2.4.2 Βελτιστοποίηση φ_β-απόκλισης

Στην ενότητα αυτή γενικεύονται οι προτάσεις 2.4 και 2.5 για την περίπτωση της φ_β-απόκλισης η οποία ορίζεται ως εξής:

$$T_n^{\varphi}(\hat{p}, p^0) = \frac{2n}{\varphi''(1)} D_{\varphi}(\hat{p}, \rho^0) \text{ ,όπου}$$

$$D_{\varphi}(\hat{\rho}, \rho^0) = \sum_{j=1}^{M} (\rho_j^0)^{\beta} \varphi(\frac{\hat{\rho}_j}{\rho_j^0})$$

Για β=1 προκύπτει η συνήθης φ-απόκλιση της προηγούμενης ενότητας. Σημειώνεται ότι λόγω της πολυπλοκότητας των πράξεων δεν είναι πρακτικά εύχρηστη η γενίκευση της Πρότασης 2.6.

Γενίκευση της Πρότασης 2.4

$$D_{\varphi}(\hat{\rho},\rho^0)=\sum_{j=1}^{M}(\rho_j^0)^{\beta}\varphi(\frac{\hat{\rho}_j}{\rho_j^0})$$

Έστω $\mathrm{W}_j=\frac{1}{\sqrt{n}}(N_j-np_j^0),\ j=1,....,M.$

Μιας τέταρτης τάξης έκφραση Taylor του $\mathrm{D}_{\varphi}(\hat{\mathrm{p}},\rho^0)$ γύρω από το ρ^0 δίνει :

$$\mathrm{D}_{\varphi}(\hat{\mathrm{p}},\rho^0)=\sum_{j=1}^{M}(\frac{\mathrm{dD}_{\varphi}(\mathrm{p},\rho^0)}{dp_j})_{p=p^0}\frac{w_j}{\sqrt{n}}+\frac{1}{2!}\sum_{j=1}^{M}(\frac{\mathrm{d}^2\mathrm{D}_{\varphi}(\mathrm{p},\rho^0)}{dp_j^2})_{p=p^0}\frac{w_j^2}{n}+\frac{1}{3!}\sum_{j=1}^{M}(\frac{\mathrm{d}^3\mathrm{D}_{\varphi}(\mathrm{p},\rho^0)}{dp_j^3})_{p=p^0}\frac{w_j^3}{n\sqrt{n}}$$
$$+\frac{1}{4!}\sum_{j=1}^{M}(\frac{\mathrm{d}^4\mathrm{D}_{\varphi}(\mathrm{p},\rho^0)}{dp_j^4})_{p=p^0}\frac{w_j^4}{n^2}+O_P(n^{-\frac{5}{2}}).$$

Αλλά

$$(\frac{\mathrm{dD}_{\varphi}(\mathrm{p},\rho^0)}{dp_j})_{p=p^0}=(\rho_j^0)^{\beta-1}\varphi'(1)$$

$$(\frac{\mathrm{d}^2\mathrm{D}_{\varphi}(\mathrm{p},\rho^0)}{dp_j^2})_{p=p^0}=(\rho_j^0)^{\beta-2}\varphi''(1)$$

$$(\frac{\mathrm{d}^3\mathrm{D}_{\varphi}(\mathrm{p},\rho^0)}{dp_j^3})_{p=p^0}=(\rho_j^0)^{\beta-3}\varphi'''(1)$$

$$(\frac{\mathrm{d}^4\mathrm{D}_{\varphi}(\mathrm{p},\rho^0)}{dp_j^4})_{p=p^0}=(\rho_j^0)^{\beta-4}\varphi^{IV}(1)$$

Τότε

$$\mathrm{T}_n^{\varphi}(\hat{\mathrm{p}},\mathrm{p}^0)=\frac{2n}{\varphi''(1)}\mathrm{D}_{\varphi}(\mathrm{p}\wedge,\rho^0)=\frac{2n\varphi'(1)}{\varphi''(1)\sqrt{n}}\sum_{j=1}^{M}(\rho_j^0)^{\beta-1}w_j+\sum_{j=1}^{M}(\rho_j^0)^{\beta-2}w_j^2+\frac{\varphi'''(1)}{3\sqrt{n}\varphi''(1)}\sum_{j=1}^{M}(\rho_j^0)^{\beta-3}w_j^3+$$
$$\frac{\varphi^{IV}(1)}{12n\varphi''(1)}\sum_{j=1}^{M}(\rho_j^0)^{\beta-4}w_j^4+O_P(n^{-\frac{3}{2}})\quad(3.1)$$

Άρα:

$$E[T_n^{\varphi}(\hat{p},p^0)]=\frac{2n\varphi'(1)}{\varphi''(1)\sqrt{n}}\sum_{j=1}^{M}(\rho_j^0)^{\beta-1}E(w_j)+\sum_{j=1}^{M}(\rho_j^0)^{\beta-2}E(w_j^2)+\frac{\varphi'''(1)}{3\sqrt{n}\varphi''(1)}\sum_{j=1}^{M}(\rho_j^0)^{\beta-3}E(w_j^3)+$$
$$\frac{\varphi^{IV}(1)}{12n\varphi''(1)}\sum_{j=1}^{M}(\rho_j^0)^{\beta-4}E(w_j^4)+O(n^{-\frac{3}{2}})=\sum_{j=1}^{M}(\rho_j^0)^{\beta-2}E(w_j^2)+\frac{\varphi'''(1)}{3\sqrt{n}\varphi''(1)}\sum_{j=1}^{M}(\rho_j^0)^{\beta-3}E(w_j^3)+$$
$$\frac{\varphi^{IV}(1)}{12n\varphi''(1)}\sum_{j=1}^{M}(\rho_j^0)^{\beta-4}E(w_j^4)+O(n^{-\frac{3}{2}})$$ (αφού ο πρώτος όρος ισούται με μηδέν)

Αφού $E[O_P(n^{-\frac{3}{2}})]=O(n^{-\frac{3}{2}})$.

Η ροπή της συνάρτησης μιας πολαπλής τυχαίας μεταβλητής , $N=(N_1,....,N_M)^T$ με παραμέτρους n και p^0 είναι :

$$M_N(t)=E[\exp(t^TN)]=(p_1^0\exp(t_1),......,p_M^0\exp(t_M))^n \text{ με } t=(t_1,...,t_M)^T$$

Η ροπή της συνάρτησης Μ-διάστασης της τυχαίας μεταβλητής $W=\frac{1}{\sqrt{n}}(N-np^0)$ δίνεται από :

$$M_W(t)=E[\exp(t^T W)]=E[\exp(t^T(N/\sqrt{n}-\sqrt{n}p^0))]$$
$$=\exp(-\sqrt{n}t^Tp^0)E[\exp(t^T(N/\sqrt{n})]\ (4.11)$$
$$=\exp(-\sqrt{n}t^Tp^0)M(t/\sqrt{n}),\ (4.11)$$ και η α-ροπή της W_j σχετικά με τη προέλευση από :

$$E[W_j^a]=\left(\frac{d^aM_W(t)}{dt_j^a}\right)_{t=0}\ (2.17)$$ για j = 1,...,M και α = 1,2,.....

Από τις (2.16) και (2.17) έχουμε :

$$E[W_j^2]=-(p_j^0)^2+p_j^0$$

$$E[W_j^3]=n^{-\frac{1}{2}}(2(p_j^0)^3-3(p_j^0)^2+p_j^0)$$

$$E[W_j^4]=3(p_j^0)^4-6(p_j^0)^3+3(p_j^0)^2+n^{-1}(-6(p_j^0)^4+12(p_j^0)^3-7(p_j^0)^2+p_j^0)$$

Άρα:

$$E[T_n^{\varphi}(\hat{p},p^0)]=\sum_{j=1}^{M}(\rho_j^0)^{\beta-2}[-(p_j^0)^2+p_j^0]+\frac{\varphi'''(1)}{3\sqrt{n}\varphi''(1)}\sum_{j=1}^{M}(\rho_j^0)^{\beta-3}[n^{-\frac{1}{2}}(2(p_j^0)^3-3(p_j^0)^2+p_j^0)]+$$
$$\frac{\varphi^{IV}(1)}{12n\varphi''(1)}\sum_{j=1}^{M}(\rho_j^0)^{\beta-4}[3(p_j^0)^4-6(p_j^0)^3+3(p_j^0)^2+n^{-1}(-6(p_j^0)^4+12(p_j^0)^3-7(p_j^0)^2+p_j^0)]+O(n^{-\frac{3}{2}})=$$
$$=\sum_{j=1}^{M}-(\rho_j^0)^{\beta}+\sum_{j=1}^{M}(\rho_j^0)^{\beta-1}+\frac{1}{n}\left[\frac{\varphi'''(1)}{3\varphi''(1)}(2\sum_{j=1}^{M}(\rho_j^0)^{\beta}-3\sum_{j=1}^{M}(\rho_j^0)^{\beta-1}+\sum_{j=1}^{M}(\rho_j^0)^{\beta-2})+\right.$$

$$+\frac{\varphi^{IV}(1)}{4\varphi''(1)}(\sum_{j=1}^{M}(\rho_j^0)^{\beta}-2\sum_{j=1}^{M}(\rho_j^0)^{\beta-1}+\sum_{j=1}^{M}(\rho_j^0)^{\beta-2})]+O(n^{-\frac{3}{2}})=$$

$$=\mathrm{M}_1-\mathrm{M}_0+\frac{1}{n}[\frac{\varphi'''(1)}{3\varphi''(1)}(2\mathrm{M}_0-3\mathrm{M}_1+\mathrm{M}_2)+\frac{\varphi^{IV}(1)}{4\varphi''(1)}(\mathrm{M}_0-2\mathrm{M}_1+\mathrm{M}_2)]+O(n^{-\frac{3}{2}})$$

Όπου

$$\sum_{j=1}^{M}(\rho_j^0)^{\beta}=\mathrm{M}_0$$

$$\sum_{j=1}^{M}(\rho_j^0)^{\beta-1}=\mathrm{M}_1$$

$$\sum_{j=1}^{M}(\rho_j^0)^{\beta-2}=\mathrm{M}_2$$

Γενίκευση της Πρότασης 2.5

Τετραγωνίζοντας και παίρνοντας τις εκφράσεις της (3.1) παίρνουμε :

$$\mathrm{E}[\mathrm{T}_n^{\varphi}(\hat{\mathrm{p}},\mathrm{p}^0)^2]=\sum_{j=1}^{M}(p_j^0)^{2\beta-4}E[w_j^4]+\sum_{j\neq i}^{M}(p_j^0)^{\beta-2}(p_i^0)^{\beta-2}E[w_j^2w_i^2]+$$

$$+\frac{2\varphi'''(1)}{3\sqrt{n}\varphi''(1)}(\sum_{j=1}^{M}(p_j^0)^{2\beta-5}E[w_j^5]+\sum_{j\neq i}^{M}(p_j^0)^{\beta-2}(p_i^0)^{\beta-3}E[w_j^2w_i^3]+$$

$$+\frac{1}{n}(\frac{\varphi^{IV}(1)}{6\varphi''(1)}(\sum_{j=1}^{M}(p_j^0)^{2\beta-6}E[w_j^6]+\sum_{j\neq i}^{M}(p_j^0)^{\beta-2}(p_i^0)^{\beta-4}E[w_j^2w_i^4])+$$

$$+(\frac{\varphi'''(1)}{3\varphi''(1)})^2(\sum_{j=1}^{M}(p_j^0)^{2\beta-6}E[w_j^6]+\sum_{j\neq i}^{M}(p_j^0)^{\beta-3}(p_i^0)^{\beta-3}E[w_j^3w_i^3]))+O(n^{-\frac{3}{2}})\quad(2.19)$$

Από τη (2.16) έχουμε :

$E[\mathrm{W}_j^{\alpha}W_i^{b}]=(\frac{d^{a+b}M_W(t)}{dt_j^a dt_i^b})_{t=0}$ (2.20) για j,i=1,...,M και a,b=1,2,...... Τότε :

$$\mathrm{E}[\mathrm{W}_j^5]=n^{-1/2}(-20(p_j^0)^5+50(p_j^0)^4-40(p_j^0)^3+10(p_j^0)^2)$$

$$+n^{-3/2}(24(p_j^0)^5-60(p_j^0)^4+50(p_j^0)^3-15(p_j^0)^2+p_j^0),\ \text{και}$$

$$\mathrm{E}[\mathrm{W}_j^6]=-15(p_j^0)^6+45(p_j^0)^5-45(p_j^0)^4+15(p_j^0)^3+n^{-1}(130(p_j^0)^6-390(p_j^0)^5+415(p_j^0)^4$$

$$-180(p_j^0)^3+25(p_j^0)^2)+O(n^{-2}).$$

Για $j \neq i$ και τα δυο σταθερά ορίζουμε , $p_{ab} = (p_j^0)^a (p_j^0)^b$ τότε

$E[W_j^2 \, W_i^2] = 3p_{22} - p_{21} - p_{12} + p_{11} + n^{-1}(-6p_{22} + 2p_{21} + 2p_{12} - p_{11}),$

$E[W_j^2 \, W_i^3] = n^{-1/2}(-20p_{23} + 5p_{13} + 15p_{22} - 6p_{12} - p_{21} + p_{11}),$

$E[W_j^2 \, W_i^4] = -15p_{24} + 18p_{23} + 3p_{14} - 3p_{22} - 6p_{13} + 3p_{12} + \frac{1}{n}(130p_{24} - 156p_{23} - 26p_{14} + 41p_{22} + 42p_{13}$

$- p_{21} - 17p_{12} + p_{11}) + O(n^{-2})$

$E[W_j^3 \, W_i^3] = -15p_{33} + 9p_{32} + 9p_{23} - 9p_{22} + n^{-1}(130p_{33} - 78p_{32} - 78p_{23} + 63p_{22} + 5p_{31}$

$+5p_{13} - 6p_{21} - 6p_{12} + p_{11}) + O(n^{-2})$

Αντικαθιστώντας αυτές τις εκφράσεις στην (2.19) έχουμε :

$$
\begin{aligned}
E[T_n^{\varphi}(\hat{\rho},\rho^0)]^2 &= 3\sum_{j=1}^{M}(\rho_j^0)^{2\beta} - 6\sum_{j=1}^{M}(\rho_j^0)^{2\beta-1} + 3\sum_{j=1}^{M}(\rho_j^0)^{2\beta-2} + \\
&+n^{-1}(-6\sum_{j=1}^{M}(\rho_j^0)^{2\beta} + 12\sum_{j=1}^{M}(\rho_j^0)^{2\beta-1} - 7\sum_{j=1}^{M}(\rho_j^0)^{2\beta-2} + \sum_{j=1}^{M}(\rho_j^0)^{2\beta-3}) \\
&+3\sum_{j\neq i}^{M}(\rho_j^0)^{\beta}(\rho_i^0)^{\beta} - \sum_{j\neq i}^{M}(\rho_j^0)^{\beta}(\rho_i^0)^{\beta-1} - \sum_{j\neq i}^{M}(\rho_j^0)^{\beta-1}(\rho_i^0)^{\beta} + \sum_{j\neq i}^{M}(\rho_j^0)^{\beta-1}(\rho_i^0)^{\beta-1} \\
&+n^{-1}(-6\sum_{j\neq i}^{M}(\rho_j^0)^{\beta}(\rho_i^0)^{\beta} + 2\sum_{j\neq i}^{M}(\rho_j^0)^{\beta}(\rho_i^0)^{\beta-1} + 2\sum_{j\neq i}^{M}(\rho_j^0)^{\beta-1}(\rho_i^0)^{\beta} - \sum_{j\neq i}^{M}(\rho_j^0)^{\beta-1}(\rho_i^0)^{\beta-1}) \\
&+\frac{2\varphi'''(1)}{3n\varphi''(1)}(-20\sum_{j=1}^{M}(\rho_j^0)^{2\beta} + 50\sum_{j=1}^{M}(\rho_j^0)^{2\beta-1} - 40\sum_{j=1}^{M}(\rho_j^0)^{2\beta-2} + 10\sum_{j=1}^{M}(\rho_j^0)^{2\beta-3} - 20\sum_{j\neq i}^{M}(\rho_j^0)^{\beta}(\rho_i^0)^{\beta} + \\
&+5\sum_{j\neq i}^{M}(\rho_j^0)^{\beta-1}(\rho_i^0)^{\beta} + 15\sum_{j\neq i}^{M}(\rho_j^0)^{\beta}(\rho_i^0)^{\beta-1} - 6\sum_{j\neq i}^{M}(\rho_j^0)^{\beta-1}(\rho_i^0)^{\beta-1} - \sum_{j\neq i}^{M}(\rho_j^0)^{\beta}(\rho_i^0)^{\beta-2} + \sum_{j\neq i}^{M}(\rho_j^0)^{\beta-1}(\rho_i^0)^{\beta-2}) \\
&+\frac{1}{n}[\frac{\varphi^{IV}(1)}{6\varphi''(1)}(-15\sum_{j=1}^{M}(\rho_j^0)^{2\beta} + 45\sum_{j=1}^{M}(\rho_j^0)^{2\beta-1} - 45\sum_{j=1}^{M}(\rho_j^0)^{2\beta-2} + 15\sum_{j=1}^{M}(\rho_j^0)^{2\beta-3} - 15\sum_{j\neq i}^{M}(\rho_j^0)^{\beta}(\rho_i^0)^{\beta+2} + \\
&+18\sum_{j\neq i}^{M}(\rho_j^0)^{\beta}(\rho_i^0)^{\beta+1} + 3\sum_{j\neq i}^{M}(\rho_j^0)^{\beta-1}(\rho_i^0)^{\beta} - 3\sum_{j\neq i}^{M}(\rho_j^0)^{\beta}(\rho_i^0)^{\beta} - 6\sum_{j\neq i}^{M}(\rho_j^0)^{\beta-1}(\rho_i^0)^{\beta+1} + 3\sum_{j\neq i}^{M}(\rho_j^0)^{\beta-1}(\rho_i^0)^{\beta}) \\
&+(\frac{\varphi'''(1)}{3\varphi''(1)})^2(-15\sum_{j=1}^{M}(\rho_j^0)^{2\beta} + 45\sum_{j=1}^{M}(\rho_j^0)^{2\beta-1} - 45\sum_{j=1}^{M}(\rho_j^0)^{2\beta-2} + 15\sum_{j=1}^{M}(\rho_j^0)^{2\beta-3} - 15\sum_{j\neq i}^{M}(\rho_j^0)^{\beta}(\rho_i^0)^{\beta} + \\
&+9\sum_{j\neq i}^{M}(\rho_j^0)^{\beta}(\rho_i^0)^{\beta-1} + 9\sum_{j\neq i}^{M}(\rho_j^0)^{\beta-1}(\rho_i^0)^{\beta} - 9\sum_{j\neq i}^{M}(\rho_j^0)^{\beta-1}(\rho_i^0)^{\beta-1})] + O(n^{-3/2})
\end{aligned}
$$

Μέρος ΙΙ

Εφαρμογές και Παραδείγματα

Κεφάλαιο 3

Ειδικές φ-αποκλίσεις

Για την αξιολόγηση της αποτελεσματικότητας των προτεινομένων μεθόδων γίνεται στην παρούσα ενότητα ενδεικτικά χρήση τριών βασικών μέτρων απόκλισης. Συγκεκριμένα χρησιμοποιούνται τα εξής 3 μέτρα απόκλισης :

$$\text{Rukhin: } \varphi(\mathrm{x}) = \frac{(1-x)^2}{2(a+(1-a)x)},\ 0 \le a \le 1$$

$$\text{Cressie and Read: } \varphi(\mathrm{x}) = \frac{x^{\alpha+1} - x - \alpha(x-1)}{\alpha(\alpha+1)},\ a \ne 0,\text{-}1$$

$$\text{Lin: } \varphi(\mathrm{x}) = \frac{\alpha x \log x - (ax+1-a)\log(ax+1-a)}{a(1-a)},\ a \ne 0,1$$

Στόχος των προσομοιώσεων είναι ο καθορισμός του εύρους (ενδεχομένως διαφορετικού για κάθε μέτρο) των τιμών του δείκτη α ώστε οι αποκλίσεις (σφάλματα) μεταξύ των πραγματικών και των ασυμπτωτικών τριών πρώτων ροπών. Ως εκ τούτου για κάθε ένα από τα τρία μέτρα απόκλισης και κάθε μια από τις τρεις πρώτες ροπές προσδιορίζονται οι τιμές (ενδεχομένως διαφορετικές) των δεικτών α που ελαχιστοποιούν τις αποκλίσεις ΕΤ1, ΕΤ2 και ΕΤ3, όπου

ΕΤi= Πραγματική – Ασυμπτωτική Ροπή i τάξης, i=1, 2, 3.

Λαμβάνοντας υπόψη, για κάθε μέτρο ξεχωριστά, τις 3 (το πολύ) βέλτιστες τιμές του δείκτη α, καθορίζεται το εύρος των τιμών του δείκτη που αντιστοιχεί στις από κοινού ελάχιστες αποκλίσεις.

Η αξιολόγηση μελετάται για διαφορετικά μεγέθη δείγματος, δηλ., μικρά, μεσαία και μεγάλα. Έτσι για κάθε μέτρο οι υπολογισμοί αφορούν

n=20 και 30 (μικρά μεγέθη)

n= 50 και 100 (μεσαία μεγέθη) και

n= 300 (μεγάλα μεγέθη)

n= 1000 (μεγάλα μεγέθη).

Σε κάθε περίπτωση αναμένεται τα σφάλματα να μειώνονται και να τείνουν στο μηδέν όσο το μέγεθος του δείγματος αυξάνεται. Αυτό αναμένεται να συμβαίνει για κάθε μέτρο ξεχωριστά και για κάθε ροπή ξεχωριστά. Σημειώνεται ότι η πολυπλοκότητα των τύπων δεν επιτρέπει σε καμιά περίπτωση τη μελέτη ροπών μεγαλύτερης τάξης από την Τρίτη. Σε κάποιες μάλιστα περιπτώσεις ακόμα και η διερεύνηση της $3^{ης}$ ροπής καθίσταται δύσκολη (περίπτωση ισοπίθανων κατηγοριών με M=4, βλέπε Πίνακα 3). Σε οποιαδήποτε περίπτωση αν οι πραγματικές και ασυμπτωτικές ροπές των τριών πρώτων τάξεων είναι σχεδόν ίσες η εγκυρότητα της ασυμπτωτικής κατανομής μπορεί να θεωρηθεί ότι διασφαλίζεται. Πέραν δε αυτού η σχεδόν πλήρης ταύτιση των ροπών ακόμα και για κάποια σχετικά μικρά μεγέθη δειγμάτων (που άλλωστε είναι αυτά που συχνά συναντώνται στην πράξη) δίνει τη δυνατότητα στον ερευνητή την ανεξέλεγκτη χρήση της ασυμπτωτικής κατανομής χωρίς να τίθεται εν αμφιβόλω η εγκυρότητα των αποτελεσμάτων του.

Σημειώνεται ότι το θέμα που εξετάζεται στην παρούσα διατριβή έχει την εξής ιδιαιτερότητα: Οποιαδήποτε κατανομή και να πρέπει να μελετηθεί, το πεδίο τιμών της κωδικοποιείται με τέτοιο τρόπο ώστε να δημιουργούνται M κατηγορίες (διαστήματα) η ένωση των οποίων να ταυτίζεται με το πεδίο τιμών της κατανομής. Με άλλα λόγια οποιαδήποτε κατανομή μετατρέπεται σε πολυωνυμική κατανομή με τόσες υποκατηγορίες όσες επιθυμεί ο ερευνητής. Είναι ευνόητο ότι ο αριθμός των κατηγοριών είναι άρρηκτα συνδεδεμένος με το μέγεθος του δείγματος, δηλαδή τον αριθμό των πειραμάτων της πολυωνυμικής κατανομής. Έτσι αν το μέγεθος είναι αρκετά μικρό τότε ο αριθμός των διαστημάτων M θα πρέπει να μην είναι υπερβολικά μεγάλος διότι τότε ενδέχεται ο ερευνητής να καταλήξει να έχει ελάχιστες έως και καθόλου παρατηρήσεις σε κάποια από τα M διαστήματα. Ταυτόχρονα όμως αν το μέγεθος του δείγματος είναι αρκετά μεγάλο τότε το M δεν θα πρέπει να είναι υπερβολικά μικρό διότι τότε η αρχική κατανομή δεν θα έχει τεμαχιστεί σε ικανοποιητικό βαθμό ώστε να αναδειχθεί η καταλληλότητα της προτεινομένης μεθοδολογίας. Σε οποιαδήποτε περίπτωση για την αξιολόγηση της μεθόδου αρκεί κανείς να μελετήσει μια σειρά διαφορετικών πολυωνυμικών κατανομών. Εδώ οι κατανομές που εξετάζονται είναι οι εξής:

<u>Α. Ισοπίθανες κατηγορίες με M=2, 3 και 4 διαστήματα:</u>

(α) $p_1^0 = p_2^0 = 1/2$

(β) $p_1=p_2=p_3=1/3$

(γ) $p_1=p_2= p_3=p_4 =1/4$

<u>Β. Συμμετρικές κατανομές με M= 3 διαστήματα:</u>

(δ) $p_1=p_3=1/4$ & $p_2=1/2$ *

<u>Γ. Λοξές κατανομές με M = 2 και 3 διαστήματα:</u>

(ε) $p_1=1/4$ και $p_2=3/4$

(στ) p_1=1/8 p_2=1/4 και p_3=5/8 *

* Τα αποτελέσματα ισχύουν για οποιαδήποτε διάταξη των πιθανοτήτων.

Σχετικά με τις ροπές της ασυμπτωτική κατανομής υπενθυμίζεται ότι ανεξαρτήτως του μέτρου απόκλισης που χρησιμοποιείται η ασυμπτωτική κατανομή είναι X^2_{M-1} με M=2, 3 ή 4 ανάλογα την κατηγοριοποίηση που χρησιμοποιείται. Όπως γνωρίζουμε αν W ακολουθεί την X^2_k τότε

$$\mathrm{EW}^m = 2^m \frac{\Gamma\left(m+\frac{k}{2}\right)}{\Gamma\left(\frac{k}{2}\right)}.$$

Έτσι $\mathrm{E}\left(\mathrm{X}^2_{M-1}\right) = M-1$, $\mathrm{E}\left(\mathrm{X}^2_{M-1}\right)^2 = M^2-1$ και $\mathrm{E}\left(\mathrm{X}^2_{M-1}\right)^3 = \left(M^2-1\right)\left(M+3\right)$ που για M=2, 3 και 4 δίνουν τις αντίστοιχες τιμές των ασυμπτωτικών ροπών για τις δικές μας ανάγκες.

Στην επόμενη ενότητα υπολογίζονται οι τύποι των τριών ροπών για τα τρία μέτρα απόκλισης που έχουν επιλεχθεί να μελετηθούν. Τα αποτελέσματα των εφαρμογών παρουσιάζονται στην ενότητα 3.2.

3.1 Οι ροπές των μέτρων απόκλισης

Με τη χρήση του προγράμματος Matlab προσδιορίζονται οι τέσσερις παραγώγοι των μέτρων απόκλισης των Rukhin, Cressie & Read και Lin και αποδεικνύονται οι προτάσεις 2.4, 2.5 και 2.6 για τα συγκεκριμένα μέτρα απόστασης. Σε όλες τις περιπτώσεις ισχύει ότι $S = \sum_{j=1}^{M} (p_j^0)^{-1}$.

Μέτρο του Rukhin

Πρώτη παράγωγος: $\varphi'(\chi) = \frac{2\chi-2}{2\alpha-2\chi(\alpha-1)} + \frac{(2\alpha-2)(\chi-1)^2}{(2\alpha-2\chi(\alpha-1))^2}$

Δεύτερη παράγωγος: $\varphi''(\chi) = \frac{2}{2\alpha-2\chi(\alpha-1)} + \frac{2(2\alpha-2)^2(\chi-1)^2}{(2\alpha-2\chi(\alpha-1))^3} + \frac{2(2\alpha-2)(2\chi-2)}{(2\alpha-2\chi(\alpha-1))^2}$,

με $\varphi''(1) = 1$

Τρίτη παράγωγος: $\varphi'''(\chi) = \frac{2(2\alpha-2)}{(2\alpha-2\chi(\alpha-1))^2} + \frac{2(4\alpha-4)}{(2\alpha-2\chi(\alpha-1))^2} + \frac{2(2\alpha-2)^2(2\chi-2)}{(2\alpha-2\chi(\alpha-1))^3} +$

$\frac{6(2\alpha-2)^3(\chi-1)^2}{(2\alpha-2\chi(\alpha-1))^4} + \frac{2(2\alpha-2)(4\alpha-4)(2\chi-2)}{(2\alpha-2\chi(\alpha-1))^3}$, με $\varphi'''(1) = 3(\alpha-1)$

Τέταρτη παράγωγος: $\varphi^{IV}(\chi) = \frac{4(2\alpha-2)^2}{(2\alpha-2\chi(\alpha-1))^3} + \frac{2(4\alpha-4)^2}{(2\alpha-2\chi(\alpha-1))^3} + \frac{12(2\alpha-2)^3(2\chi-2)}{(2\alpha-2\chi(\alpha-1))^4} +$

$\frac{24(2\alpha-2)^4(\chi-1)^2}{(2\alpha-2\chi(\alpha-1))^5} + \frac{2(2\alpha-2)(4\alpha-4)}{(2\alpha-2\chi(\alpha-1))^3} + \frac{2(2\alpha-2)(8\alpha-8)}{(2\alpha-2\chi(\alpha-1))^3} + \frac{3(2\alpha-2)(4\alpha-4)^2(2\chi-2)}{(2\alpha-2\chi(\alpha-1))^4}$,

με $\varphi^{IV}(1) = 12\alpha^2 - 18\alpha + 12$

Για τη 2.4 πρόταση

$E[T_n^{\varphi}(\rho^\wedge, \rho^0)] = M - 1 + \frac{1}{n} f_{\varphi}^1 + O(n^{-3/2})$, όπου

$f_{\varphi}^1 = (a-1)(2-3M+S) + \frac{(12a^2-18a+12)}{4}(1-2M+S).$

Για τη 2.5 πρόταση

$E[T_n^{\varphi}(\rho^\wedge, \rho^0)]^2 = M^2 - 1 + \frac{1}{n} f_{\varphi}^2 + O(n^{-3/2})$, όπου

$f_{\varphi}^2 = (2-2M-M^2+S) + 2(a-1)(10-13M-6M^2(M+8)S) + \frac{1}{3}(3(\alpha-1))^2(4-6M-3M^2+5S) +$

$\frac{(12a^2-18a+12)}{2}(3-5M-2M^2+(M+3)S).$

Για τη 2.6 πρόταση

$E[T_n^{\varphi}(\rho^\wedge, \rho^0)]^3 = M^3 + 3M^2 - M - 3 + \frac{1}{n} f_{\varphi}^3 + O(n^{-3/2})$, όπου

$f_{\varphi}^3 = (26-24M-21M^2-3M^3+(19-3M)S) + 3(a-1)(70-81M-64M^2-9M^3+(65+18M+M^2)S)$

$+(3(\alpha-1))^2(20-26M-21M^2-3M^3+(25+5M)S)$

$+\frac{3(12a^2-18a+12)}{4}(15-22M-15M^2-2M^3+(15+8M+M^2)S).$

Μέτρο των Cressie and Read

Πρώτη παράγωγος: $\varphi'(\chi) = \dfrac{\alpha\chi^{\alpha} + \chi^{\alpha} - 1 - \alpha}{\alpha^2 + \alpha}$

Δεύτερη παράγωγος: $\varphi''(\chi) = \dfrac{\alpha^2\chi^{\alpha-1} + \alpha\chi^{\alpha-1}}{\alpha^2 + \alpha}$, με $\varphi''(1) = 1$

Τρίτη παράγωγος: $\varphi'''(\chi) = \dfrac{\alpha^3\chi^{\alpha-2} + \alpha^2\chi^{\alpha-2} - 2\alpha\chi^{\alpha-2}}{\alpha^2 + \alpha}$, με $\varphi'''(1) = \alpha - 1$

Τέταρτη παράγωγος: $\varphi^{IV}(\chi) = \dfrac{(\alpha^4 - 2\alpha^3)\chi^{\alpha-3} + (\alpha^3 - 2\alpha^2)\chi^{\alpha-3} + (-2\alpha^2 + 4\alpha)\chi^{\alpha-3}}{\alpha^2 + \alpha}$,

με $\varphi^{IV}(1) = (\alpha - 1)(\alpha - 2)$

Για τη 2.4 πρόταση

$E[T_n^{\varphi}(\rho\wedge, \rho^0)] = M - 1 + \dfrac{1}{n} f_{\varphi}^1 + O(n^{-3/2})$, όπου

$f_{\varphi}^1 = \dfrac{(a-1)}{3}(2 - 3M + S) + \dfrac{(\alpha-1)(\alpha-2)}{4}(1 - 2M + S)$.

Για τη 2.5 πρόταση

$E[T_n^{\varphi}(\rho\wedge, \rho^0)]^2 = M^2 - 1 + \dfrac{1}{n} f_{\varphi}^2 + O(n^{-3/2})$, όπου

$f_{\varphi}^2 = (2 - 2M - M^2 + S) + \dfrac{2(a-1)}{3}(10 - 13M - 6M^2(M+8)S) + \dfrac{1}{3}((\alpha-1))^2(4 - 6M - 3M^2 + 5S) +$
$\dfrac{(\alpha-1)(\alpha-2)}{2}(3 - 5M - 2M^2 + (M+3)S)$.

Για τη 2.6 πρόταση

$E[T_n^{\varphi}(\rho\wedge, \rho^0)]^3 = M^3 + 3M^2 - M - 3 + \dfrac{1}{n} f_{\varphi}^3 + O(n^{-3/2})$, όπου

$f_{\varphi}^3 = (26 - 24M - 21M^2 - 3M^3 + (19 - 3M)S) + (a-1)(70 - 81M - 64M^2 - 9M^3 + (65 + 18M + M^2)S)$
$+((\alpha-1))^2(20 - 26M - 21M^2 - 3M^3 + (25 + 5M)S)$
$+\dfrac{3(\alpha-1)(\alpha-2)}{4}(15 - 22M - 15M^2 - 2M^3 + (15 + 8M + M^2)S)$.

Μέτρο του Lin

Πρώτη παράγωγος: $\varphi'(\chi) = \dfrac{\log\chi - \log(a\chi + 1 - \alpha)}{1-\alpha}$

Δεύτερη παράγωγος: $\varphi''(\chi) = \dfrac{\dfrac{1}{\chi} - \dfrac{\alpha}{\alpha\chi + 1 - \alpha}}{1-\alpha}$, με $\varphi''(1) = 1$

Τρίτη παράγωγος: $\varphi'''(\chi) = \dfrac{-\dfrac{1}{\chi^2} + \dfrac{\alpha^2}{(\alpha\chi + 1 - \alpha)^2}}{1-\alpha}$, με $\varphi'''(1) = -\alpha - 1$

Τέταρτη παράγωγος: $\varphi^{IV}(\chi) = \dfrac{\dfrac{1}{\chi^4} - \dfrac{\alpha^3(2\alpha\chi + 2 - 2\alpha)}{(\alpha\chi + 1 - \alpha)^4}}{1-\alpha}$, με $\varphi^{IV}(1) = \dfrac{1-2a^3}{1-a}$

Για τη 2.4 πρόταση

$$\mathrm{E}[\mathrm{T}_n^{\varphi}(\rho\wedge, \rho^0)] = \mathrm{M} - 1 + \frac{1}{n} f_{\varphi}^1 + O(n^{-3/2}),$$

όπου

$$f_{\varphi}^1 = \frac{-\alpha - 1}{3}(2 - 3\mathrm{M} + \mathrm{S}) + \frac{1 - 2\alpha^3}{4 - 4\alpha}(1 - 2M + S).$$

Για τη 2.5 πρόταση

$$\mathrm{E}[\mathrm{T}_n^{\varphi}(\rho\wedge, \rho^0)]^2 = \mathrm{M}^2 - 1 + \frac{1}{n} f_{\varphi}^2 + O(n^{-3/2}),$$

όπου

$$f_{\varphi}^2 = (2 - 2\mathrm{M} - \mathrm{M}^2 + S) - \frac{2\alpha + 2}{3}(10 - 13\mathrm{M} - 6\,\mathrm{M}^2(\mathrm{M} + 8)\mathrm{S}) + \frac{1}{3}(-\alpha - 1)^2(4 - 6\mathrm{M} - 3\mathrm{M}^2 + 5S) +$$
$$\frac{1 - 2\alpha^3}{2 - 2\alpha}(3 - 5M - 2M^2 + (M + 3)S).$$

Για τη 2.6 πρόταση

$$\mathrm{E}[\mathrm{T}_n^{\varphi}(\rho\wedge, \rho^0)]^3 = \mathrm{M}^3 + 3\mathrm{M}^2 - \mathrm{M} - 3 + \frac{1}{n} f_{\varphi}^3 + O(n^{-3/2}),$$

όπου

$$f_{\varphi}^3 = (26 - 24\mathrm{M} - 21\mathrm{M}^2 - 3\mathrm{M}^3 + (19 - 3\mathrm{M})S) +$$
$$+ (-a - 1)(70 - 81\mathrm{M} - 64\,\mathrm{M}^2 - 9\mathrm{M}^3 + (65 + 18\,\mathrm{M} + \mathrm{M}^2)\mathrm{S})$$
$$+ (-\alpha - 1)^2(20 - 26\mathrm{M} - 21\mathrm{M}^2 - 3\mathrm{M}^3 + (25 + 5\mathrm{M})S)$$
$$+ \frac{3 - 6\alpha^3}{4 - 4\alpha}(15 - 22M - 15M^2 - 2\mathrm{M}^3 + (15 + 8M + \mathrm{M}^2)S).$$

3.2. Αποτελέσματα

Τα αποτελέσματα, για διάφορες τιμές του n του M και του S, αναγράφονται στους παρακάτω πίνακες. Για τα αποτελέσματα πέραν της τιμής του M και των πιθανοτήτων $p_j^0,\ j=1,2,...,M$ χρειάζεται να υπολογισθεί και η ποσότητα $S=\sum_{j=1}^{M}(p_j^0)^{-1}$ ενώ η επαναληπτική μέθοδος για την εύρεση του βέλτιστου δείκτη α ααπιτεί τον καθορισμό ενός αρχικού διαστήματος [a , b] εντός του οποίου βρίσκεται η επιθυμητή τιμή αλλά και ενός βήματος c με βάση το οποίο ο αλγόριθμος σαρώνει το αρχικό διάστημα [a, b]. Εδώ το c έχει καθοριστεί να είναι πάντοτε ίσο με 0.1. Οι τιμές των M, S, a & b δίνονται σε κάθε περίπτωση ξεχωριστά.

<u>Σχόλια επί των αποτελεσμάτων των Πινάκων – Περιπτώσεων 1-6.</u>

- Σε όλες τις περιπτώσεις επιβεβαιώνεται ότι όσο αυξάνεται το μέγεθος του δείγματος n μειώνεται η απόκλιση (σφάλμα) μεταξύ πραγματικής και ασυμπτωτικής ροπής 1^{ης}, 2^{ης} καθώς και 3^{ης} τάξης. Η μικρότερη τιμή σφάλματος που παρατηρήθηκε είναι ίση με 0 και παρατηρήθηκε για την απόσταση C&R και αφορά τη ροπή 1^{ης} τάξης. Η μέγιστη απόκλιση που παρατηρήθηκε είναι 72.09 μονάδες παρατηρήθηκε για μικρό δείγμα ίσο με n=20 και αφορά τη ροπή 3^{ης} τάξης για την περίπτωση 4 διαστημάτων.

- Από τα 3 μέτρα απόστασης που μελετήθηκαν το μέτρο C&R ΄΄συμπεριφέρεται΄΄ καλύτερα, σε όλες τις περιπτώσεις (ειδικότερα για τη τρίτη ροπή), απο τα άλλα 2 (Rukhin&Lin) ακόμη και για δείγματα μικρού μεγέθους (n=20 ή n=30).

- Παρατηρούμε ότι όταν το μέγεθος του δείγματος μεγαλώνει πολύ (n=1000) οι ροπές της ασυμπτωτικής κατανομής, και στα τρία μέτρα που μελετήθηκαν, στις πλείστες των περιπτώσεων είναι πολύ κοντά στο μηδέν.Έτσι μπορούμε να πούμε ότι οι πρώτες τρεις ροπές της ασυμπτωτικής κατανομής σχεδόν ταυτίζονται ή τείνουν να ταυτιστούν με τις αντίστοιχες ροπές της πραγματικής κατανομής.Αναφέρουμε ότι οι μέγιστες τιμές σφαλμάτων που παρατηρήθηκαν όταν n=1000 είναι 1.4418 και 1.08317 και εντοπίσθηκαν στην έκτη και πέμπτη περίπτωση αντίστοιχα για το μέτρο του Rukhin, ενώ σε όλες τις άλλες περιπτώσεις τα σφάλματα που παρατηρήθηκαν είναι μικρότερα της μονάδας.Τα συμπεράσματα αυτά καταδεικνύουν την εγκυρότητα της ασυμπτωτικής κατανομής για τις συγκεκριμένες ελεγχοσυναρτήσεις όπως ορίσθηκε στο θεώρημα 2.1.

- Για τις περιπτώσεις στις οποίες M=2 παρατηρούμε ότι τα σφάλματα είναι μικρότερα από τα αντίστοιχα των περιπτώσεων M=3 και M=4 (ειδικότερα στις τιμές της δεύτερης και τρίτης ροπής) όσον αφορά τα μέτρα του Rukhin και του Lin καθώς σε αυτά τα μέτρα παρατηρούνται οι μεγαλύτερες διαφοροποιήσεις σε σχέση με τον αριθμό Μ των υποδιαστημάτων.

- Για την περίπτωση της ροπής $1^{ης}$ τάξης τα σφάλματα ET1 είναι μικρότερα από 0.10 μονάδες σε όλες σχεδόν τις περιπτώσεις με την εξαίρεση κάποιων περιπτώσεων με μικρό μέγεθος δείγματος n=20 όπου το σφάλμα φτάνει μέχρι και 0.534 (ισοπίθανες κατηγορίες με M=4) που όμως και εκεί περιορίζεται κάτω από 0.10 όταν το μέγεθος του δείγματος γίνεται n=100.

- Για την περίπτωση της ροπής $2^{ης}$ τάξης παρατηρούνται περιπτώσεις με n=20 όπου το σφάλμα (απόκλιση) ET2 ισούται με 9.15 μονάδες (ισοπίθανες κατηγορίες με M=4) ή με 7.77 (τρίτη περίπτωση M=3) ή με 4.35 (συμμετρική κατανομή με M=3). Σε όλες όμως τις περιπτώσεις όταν το μέγεθος του δείγματος φτάνει το n=300 το σφάλμα ET2 δεν υπερβαίνει την τιμή 0.61.

- Όπως είναι αναμενόμενο για την ροπή $3^{ης}$ τάξης τα σφάλματα μπορεί σε κάποιες περιπτώσεις με μικρό μέγεθος δείγματος να είναι αρκετά μεγάλα. Για n=20, 30 και 100 οι μεγαλύτερες τιμές του σφάλματος που παρατηρήθηκαν είναι αντίστοιχα 72.09, 54,16 και 48.06 και αφορούν την περίπτωση των ισοπίθανων κατηγοριών με M=4. Γίνεται εύκολα αντιληπτό ότι όταν υπάρχουν αρκετές (3-4 ή περισσότερες) ισοπίθανες ή μη κατηγορίες η κατανομή τυχαίου δείγματος μικρού μεγέθους (π.χ. n=20) στις κατηγορίες μπορεί να μην αποτυπώνει με ικανοποιητική ακρίβεια την πραγματικότητα με αποτέλεσμα να οδηγεί συχνά σε στρεβλώσεις και κατά συνέπεια σημαντικές αποκλίσεις από την πραγματικότητα. Πρέπει πάντως να τονισθεί και αυτό αποτυπώνεται στις αναλύσεις που έγιναν ότι όσο το μέγεθος του δείγματος αυξάνεται παρατηρείται όλο και καλύτερη αποτύπωση της πραγματικότητας με αποτέλεσμα η απόκλιση να φθίνει και να φτάνει στα χαμηλότερά της επίπεδα όταν n=1000. Είναι χαρακτηριστικό ότι η μέγιστη απόκλιση που παρατηρήθηκε για n=300 είναι ίση με 4.806 και για n=1000 ίση με 1.4 και αφορά στην περίπτωση στο σφάλμα ET3 για τέσσερεις ισοπίθανες κατηγορίες.

- Για την περίπτωση δειγμάτων μεγέθους n=100 (μεσαίο μέγεθος) έχει σε κάθε μια από τις περιπτώσεις που μελετήθηκαν, προταθεί το ιδανικό εύρος τιμών του δείκτη α για το οποίο είναι οι αποκλίσεις και των τριών ροπών ταυτόχρονα βέλτιστη. Πρόκειται ουσιαστικά για το κοινό (και για τις τρεις ροπές) εύρος τιμών του δείκτη που τα 3 σφάλματα (αποκλίσεις) είναι ικανοποιητικά μικρά χωρίς να είναι αναγκαστικά και τα τρία ελάχιστα. Αυτό θεωρήθηκε αναγκαίο ώστε να μην επιλέγει ο ερευνητής μια και μοναδική τιμή του δείκτη για την οποία η μια απόκλιση αν είναι ελάχιστη και η άλλη όχι. Το προτεινόμενο εύρος δίνει ευελιξία στον ερευνητή να επιλέξει την τιμή εκείνη από το βέλτιστο εύρος που εξυπηρετεί καλύτερα τις ανάγκες του.Παρατηρούμε ότι το ιδανικό εύρος τιμών για το δείκτη α για τα 3 μέτρα είναι.

	Ελάχιστο	**Μέγιστο**
RUKHIN	**0.7**	**1**
C&R	**0.8**	**1.2**
LIN	**-1**	**-0.8**

βλέπε πίνακες 7 και 8 για περισσότερα.

- Ταυτόχρονα έχοντας μελετήσει 3 βασικά μέτρα και διαπιστώνοντας για τα συνήθη μεγέθη δείγματος τον τρόπο με τον οποίο αυτά συμπεριφέρονται, ο ενδιαφερόμενος ερευνητής αποκτά μια καλύτερη εικόνα για τη αξιοπιστία τους και έχει τη δυνατότητα συγκρίνοντας τα να επιλέξει εκείνο που ταιριάζει καλύτερα στο πρόβλημα που μελετά. Σημειώνεται ότι από ασυμπτωτικής άποψης τα 3 μέτρα είναι απολύτως ισοδύναμα αφού η ασυμπτωτική κατανομή και στις 3 περιπτώσεις για ίδιο M είναι ακριβώς η ίδια. Έτσι μόνο η μελέτη δειγμάτων πεπερασμένου μεγέθους είναι ικανή να αναδείξει τυχούσες διαφορές και διαφοροποιήσεις που μπορεί να αναδείξουν δυνατότητες ή αδυναμίες των μέτρων απόκλισης.

- Υπενθυμίζεται ότι τα τρία αυτά μέτρα ακριβώς λόγω του δείκτη από τον οποίο εξαρτώνται αποτελούν οικογένειες μέτρων (διαφορετικό μέτρο για διαφορετική τιμή του δείκτη) στις οποίες περιλαμβάνονται πολλά δημοφιλή μέτρα απόκλισης. Είναι χαρακτηριστική η περίπτωση του μέτρου των Cressie & Read που για διαφορετικές τιμές του δείκτη προκύπτουν μερικά από τα γνωστότερα μέτρα απόκλισης με ευρεία χρήση σε ελέγχους υποθέσεων και γενικά στη Στατιστική Συμπερασματολογία. Ανατρέχοντας στην ενότητα 2.1 της παρούσας διατριβής βλέπουμε ότι οι στατιστικές συναρτήσεις που προκύπτουν για διάφορες τιμές του δείκτη α είναι:

X^2 test : $\mathrm{X}^2(\alpha=1)$

G^2 λόγο πιθανοφανειών : $\mathrm{G}^2(\alpha=0)$

Λόγο πιθανοφάνειας Kullback : α = -1

Neyman-τροποποιημένο X^2 test : α = -2

Gressie-Read στατιστική δοκιμή : α = $\frac{2}{3}$

Υπενθυμίζεται ότι η πρώτη των ειδικών αυτών περιπτώσεων είναι η ελεγχοσυνάρτηση του Pearson (1900) η οποία αποτελεί ακόμα και σήμερα ένα από τα σημαντικότερα εργαλεία στην απαραμετρική στατιστική. Η $3^η$ ειδική περίπτωση αποτελεί μια από τις γνωστότερες παραμετρικές ελεγχοσυναρτήσεις όπου γίνεται σύγκριση δύο πιθανοφανειών για την επιλογή της μιας από τις δύο ως της ιδανικότερης κατανομής για τα δεδομένα. Η μέθοδος του λόγου πιθανοφανειών έχει τις ρίζες της στο Λήμμα των Neyman & Pearson που αποτελεί τη βάση της θεωρίας ελέγχου υποθέσεων. Τέλος η ειδική αυτή περίπτωση, και η $2^η$ και η $3^η$,σχετίζονται επίσης άμεσα με τη συνάρτηση πιθανοφάνειας και συνδέονται με την Εκτιμητική αφού ισοδυναμούν με τη δημοφιλέστερη μέθοδο εκτίμησης αυτή της μεγίστης πιθανοφάνειας.

Είναι σημαντικό να αναφέρουμε ότι το προτεινόμενο εύρος τιμών του δείκτη για το μέτρο C&R με βάση την παρούσα μελέτη (βλέπε πίνακες 7 και 8) είναι το διάστημα [0.80 , 1.2] με αποτέλεσμα η ελεγχοσυνάρτηση του Pearson να είναι η μόνη που εμπίπτει στο προτεινόμενο εύρος. Το γεγονός αυτό και μόνο δείχνει τη σημασία αλλά και τη σημαντικότητα της συγκεκριμένης συνάρτησης αφού με βάση την παρούσα μελέτη παρουσιάζει τα ελάχιστα σφάλματα (αποκλίσεις) όσον αφορά τις τρεις πρώτες ροπές. Το συμπέρασμα αυτό αποτελεί έναν πρόσθετο λόγο που δικαιολογεί την ευρεία ακόμα και στις μέρες μας χρήση της.

<u>Πίνακας 1: Τιμές του δείκτη που ελαχιστοποιεί το σφάλμα ET1,ET2 και ET3.Το σφάλμα σε παρένθεση.M=2 (1/2,1/2), S=4, b=0.1</u>

<u>rukhin</u> (0.8:b:1.2)

n	a->min|ET1|	a->min|ET2|	a->min|ET3|
20	0.8 (0.066)	1 (0.65)	0.8 (3.03)
30	0.8 (0.044)	1 (0.43333)	0.8 (2.02)
50	0.8 (0.0264)	1 (0.26)	0.8 (1.212)
100	0.8 (0.0132)	1 (0.13)	0.8 (0.606)
300	0.8 (0.0044)	1 (0.04333)	0.8 (0.202)
1000	0.8 (0.00132)	1 (0.013)	0.8 (0.0606)

<u>car</u> (0.8:b:1.2)

n	a->min|ET1|	a->min|ET2|	a->min|ET3|
20	1 (0)	1 (0.1)	0.8 (1.185)
30	1 (0)	1 (0.0666667)	0.8 (0.79)
50	1 (0)	1 (0.04)	0.8 (0.474)
100	1 (0)	1 (0.02)	0.8 (0.237)
300	1 (0)	1 (0.00666667)	0.8 (0.079)
1000	1 (0)	1 (0.002)	0.8 (0.0237)

<u>lin</u> (-1.2:b:-0.9)

n	a->min|ET1|	a->min|ET2|	a->min|ET3|
20	-0.9 (0.0161711)	-1 (0.0875)	-0.9 (0.197961)
30	-0.9 (0.0107807)	-1 (0.0583333)	-0.9 (0.131974)
50	-0.9 (0.00646842)	-1 (0.035)	-0.9 (0.0791842)
100	-0.9 (0.00323421)	-1 (0.0175)	-0.9 (0.0395921)
300	-0.9 (0.00107807)	-1 (0.00583333)	-0.9 (0.0131974)
1000	-0.9 (0.000323421)	-1 (0.00175)	-0.9 (0.00395921)

Πίνακας 2: Τιμές του δείκτη που ελαχιστοποιεί το σφάλμα ET1,ET2 και ET3.Το σφάλμα σε παρένθεση.M=2 (1/4,3/4), S=5.333333, b=0.1

rukhin (0.5:b:1.2)

n	a->min|ET1|	a->min|ET2|	a->min|ET3|
20	0.7 (0.134)	1 (1.71667)	0.7 (8.72667)
30	0.7 (0.0893333)	1 (1.14444)	0.7 (5.81778)
50	0.7 (0.0536)	1 (0.686667)	0.7 (3.49067)
100	0.7 (0.0268)	1 (0.343333)	0.7 (1.74533)
300	0.7 (0.00893333)	1 (0.114444)	0.7 (0.581778)
1000	0.7 (0.00268)	1 (0.0343333)	0.7 (0.174533)

car (0.8:b:1.2)

n	a->min|ET1|	a->min|ET2|	a->min|ET3|
20	1 (0)	1 (0.0333333)	1 (0.166666)
30	1 (0)	1 (0.0222222)	1 (0.111111)
50	1 (0)	1 (0.0133333)	1 (0.0666665)
100	1 (0)	1 (0.00666667)	1 (0.0333333)
300	1 (0)	1 (0.00222222)	1 (0.011111)
1000	1 (0)	1 (0.000666667)	1 (0.00333332)

lin (-1.2:b:-0.8)

n	a->min|ET1|	a->min|ET2|	a->min|ET3|
20	-0.8 (0.0283518)	-1 (0.404167)	-0.8 (2.30361)
30	-0.8 (0.0189012)	-1 (0.269444)	-0.8 (1.53574)
50	-0.8 (0.0113407)	-1 (0.161667)	-0.8 (0.921444)
100	-0.8 (0.00567037)	-1 (0.0808333)	-0.8 (0.460722)
300	-0.8 (0.00189012)	-1 (0.0269444)	-0.8 (0.153574)
1000	-0.8 (0.000567037)	-1 (0.00808333)	-0.8 (0.0460722)

Πίνακας 3: Τιμές του δείκτη που ελαχιστοποιεί το σφάλμα ET1,ET2 και ET3.Το σφάλμα σε παρένθεση.M=3 (1/3,1/3,1/3), S=9, b=0.1

Rukhin (0.8:b:1.2)

n	a->min|ET1|	a->min|ET2|	a->min|ET3|
20	0.8 (0.244)	1 (3.4)	0.8 (22.492)
30	0.8 (0.162667)	1 (2.26667)	0.8 (14.9947)
50	0.8 (0.0976)	1 (1.36)	0.8 (8.9968)
100	0.8 (0.0488)	1 (0.68)	0.8 (4.4984)
300	0.8 (0.0162667)	1 (0.226667)	0.8 (1.49947)
1000	0.8 (0.00488)	1 (0.068)	0.8 (0.44984)

car (0.8:b:1.2)

n	a->min|ET1|	a->min|ET2|	a->min|ET3|
20	1 (0)	1 (0.2)	1.2 (2.688)
30	1 (0)	1 (0.133333)	1.2 (1.792)
50	1 (0)	1 (0.08)	1.2 (1.0752)
100	1 (0)	1 (0.04)	1.2 (0.5376)
300	1 (0)	1 (0.0133333)	1.2 (0.1792)
1000	1 (0)	1 (0.004)	1.2 (0.05376)

lin (-1.2:b:-0.9)

n	a->min|ET1|	a->min|ET2|	a->min|ET3|
20	-0.9 (0.0613509)	-1 (0.7)	-0.9 (5.33053)
30	-0.9 (0.0409006)	-1 (0.466667)	-0.9 (3.55368)
50	-0.9 (0.0245404)	-1 (0.28)	-0.9 (2.13221)
100	-0.9 (0.0122702)	-1 (0.14)	-0.9 (1.06611)
300	-0.9 (0.00409006)	-1 (0.0466667)	-0.9 (0.355368)
1000	-0.9 (0.00122702)	-1 (0.014)	-0.9 (0.106611)

Πίνακας 4: Τιμές του δείκτη που ελαχιστοποιεί το σφάλμα ΕΤ1,ΕΤ2 και ΕΤ3.Το σφάλμα σε παρένθεση.M=3 (1/4,1/2,1/4), S=10, b=0.1

rukhin (0.8:b:1.2)

n	a->min|ET1|	a->min|ET2|	a->min|ET3|
20	0.8 (0.3)	1 (4.35)	0.8 (29.376)
30	0.8 (0.2)	1 (2.9)	0.8 (19.584)
50	0.8 (0.12)	1 (1.74)	0.8 (11.7504)
100	0.8 (0.06)	1 (0.87)	0.8 (5.8752)
300	0.8 (0.02)	1 (0.29)	0.8 (1.9584)
1000	0.8 (0.006)	1 (0.087)	0.8 (0.58752)

car (0.8:b:1.2)

n	a->min|ET1|	a->min|ET2|	a->min|ET3|
20	1 (0)	1 (0.15)	1.2 (0.216)
30	1 (0)	1 (0.1)	1.2 (0.144)
50	1 (0)	1 (0.06)	1.2 (0.0864)
100	1 (0)	1 (0.03)	1.2 (0.0432)
300	1 (0)	1 (0.01)	1.2 (0.0144)
1000	1 (0)	1 (0.003)	1.2 (0.00432)

lin (-1.2:b:-0.9)

n	a->min|ET1|	a->min|ET2|	a->min|ET3|
20	-0.9 (0.0758553)	-1 (0.975)	-0.9 (8.43916)
30	-0.9 (0.0505702)	-1 (0.65)	-0.9 (5.6261)
50	-0.9 (0.0303421)	-1 (0.39)	-0.9 (3.37566)
100	-0.9 (0.0151711)	-1 (0.195)	-0.9 (1.68783)
300	-0.9 (0.00505702)	-1 (0.065)	-0.9 (0.562611)
1000	-0.9 (0.00151711)	-1 (0.0195)	-0.9 (0.168783)

Πίνακας 5: Τιμές του δείκτη που ελαχιστοποιεί το σφάλμα ET1,ET2 και ET3.Το σφάλμα σε παρένθεση.M=3 (1/8,1/4,5/8), S=13.6, b=0.1

rukhin (a=0.8:b:1.2)

n	a->min|ET1|	a->min|ET2|	a->min|ET3|
20	0.8 (0.5016)	1 (7.77)	0.8 (54.1584)
30	0.8 (0.3344)	1 (5.18)	0.8 (36.1056)
50	0.8 (0.20064)	1 (3.108)	0.8 (21.6634)
100	0.8 (0.10032)	1 (1.554)	0.8 (10.8317)
300	0.8 (0.03344)	1 (0.518)	0.8 (3.61056)
1000	0.8 (0.010032)	1 (0.1554)	0.8 (1.08317)

car (a=0.8:b:1.2)

n	a->min|ET1|	a->min|ET2|	a->min|ET3|
20	1 (0)	1 (0.03)	0.8 (0.1008)
30	1 (0)	1 (0.02)	0.8 (0.0672)
50	1 (0)	1 (0.012)	0.8 (0.04032)
100	1 (0)	1 (0.006)	0.8 (0.02016)
300	1 (0)	1 (0.002)	0.8 (0.00672)
1000	1 (0)	1 (0.0006)	0.8 (0.002016)

lin (-1.2:b:-0.9)

n	a->min|ET1|	a->min|ET2|	a->min|ET3|
20	-0.9 (0.128071)	-1 (1.965)	-0.9 (19.6302)
30	-0.9 (0.0853807)	-1 (1.31)	-0.9 (13.0868)
50	-0.9 (0.0512284)	-1 (0.786)	-0.9 (7.85209)
100	-0.9 (0.0256142)	-1 (0.393)	-0.9 (3.92605)
300	-0.9 (0.00853807)	-1 (0.131)	-0.9 (1.30868)
1000	-0.9 (0.00256142)	-1 (0.0393)	-0.9 (0.392605)

<u>Πίνακας 6: Τιμές του δείκτη που ελαχιστοποιεί το σφάλμα ΕΤ1,ΕΤ2 και ΕΤ3.Το σφάλμα σε παρένθεση.Μ=4 (1/4,1/4,1/4,1/4), S=16, b=0.1</u>

<u>rukhin (</u>0.8:b:1.2)

n	a->min|ET1|	a->min|ET2|	a->min|ET3|
20	0.8 (0.534)	1 (9.15)	0.8 (72.09)
30	0.8 (0.356)	1 (6.1)	0.8 (48.06)
50	0.8 (0.2136)	1 (3.66)	0.8 (28.836)
100	0.8 (0.1068)	1 (1.83)	0.8 (14.418)
300	0.8 (0.0356)	1 (0.61)	0.8 (4.806)
1000	0.8 (0.01068)	1 (0.183)	0.8 (1.4418)

<u>car (</u>0.8:b:1.2)

n	a->min|ET1|	a->min|ET2|	a->min|ET3|
20	1 (0)	1 (0.3)	1.2 (2.346)
30	1 (0)	1 (0.2)	1.2 (1.564)
50	1 (0)	1 (0.12)	1.2 (0.9384)
100	1 (0)	1 (0.06)	1.2 (0.4692)
300	1 (0)	1 (0.02)	1.2 (0.1564)
1000	1 (0)	1 (0.006)	1.2 (0.04692)

<u>lin (</u>-1.2:b:-0.9)

n	a->min|ET1|	a->min|ET2|	a->min|ET3|
20	-0.9 (0.135539)	-1 (2.0625)	-0.9 (19.491)
30	-0.9 (0.0903596)	-1 (1.375)	-0.9 (12.994)
50	-0.9 (0.0542158)	-1 (0.825)	-0.9 (7.79638)
100	-0.9 (0.0271079)	-1 (0.4125)	-0.9 (3.89819)
300	-0.9 (0.00903596)	-1 (0.1375)	-0.9 (1.2994)
1000	-0.9 (0.00271079)	-1 (0.04125)	-0.9 (0.389819)

Πίνακας 7: Εύρος βέλτιστων δεικτών α για ισοπίθανες κατανομές (ενδεικτικό δείγμα n=100)

		ΕΥΡΟΣ	ERRORS		
			ET1	ET2	ET3
1/2,1/2	RUKHIN	(0.955116,1.03254)	(0.0143872,0.0155199)	[-0.5,1]	(0.730654,0.849586)
	CaR	(0.920412,1.07351)	(-0.000170266,0.00021480)	[-0.5,0.5]	(-0.317878,-0.277445)
	LIN	(-1.08027,-0.92532)	(0.003356,0.0042318)	[-0.5,0.5]	(0.05238,0.144339)
1/3,1/3,1/3	RUKHIN	(0.94553,1.01105)	(0.0559984,0.0608987)	[-0.5,6.5]	(5.70774,6.53097)
	CaR	(0.984981,1.0267)	(-8.18711e-005,5.231e-005)	[-1,0.5]	(-0.642006,-0.634473)
	LIN	(-1.018,-0.919623)	(0.0127765,0.0155311)	[-0.5,3]	(1.15044,1.6081)
1/4,1/4,1/4, 1/4	RUKHIN	(0.999538,1.00634)	(0.13491,0.136247)	[-0.5,2]	(20.635,20.9322)
	CaR	(0.995465,1.01571)	(-3.3721e-005,1.180e-005)	[-2,0.5]	(-1.02754,-0.992173)
	LIN	(-1.00745,-0.95438)	(0.0306158,0.0342785)	[-0.5,6]	(4.67009,5.47479)

Πίνακας 8:Εύρος βέλτιστων δεικτών α για μη ισοπίθανες κατανομές(ενδεικτικό δείγμα n=100)

		ΕΥΡΟΣ	ERRORS		
			ET1	ET2	ET3
1/4,3/4	RUKHIN	(0.89648,1.03302)	(0.0307467,0.0366723)	[-0.5,3]	(2.37689,3.3410)
	CaR	(0.941783,1.05675)	(-6.0032e-05,0.00010627)	[-0.5,0.5]	(-0.008855,0.081499)
	LIN	(-1.08027,-0.89923)	(0.0042318,0.0070902)	[-0.5,1]	(0.144339,0.690381)
1/4,1/2,1/4	RUKHIN	(0.94553,1.01105)	(0.0685082,0.0762349)	[-0.5,9]	(7.43954,8.86868)
	CaR	(0.986732,1.02422)	(-5.3217e-005,3.5375e-005)	[-1,0.5]	(-0.373886,-0.332364)
	LIN	(-0.928948,-1.01762)	(0.0161592,0.019429)	[-0.5,3]	(1.86824,2.46338)
1/8,4/5,5/8	RUKHIN	(0.916447,1.01279)	(-0.885492,-0.868464)	[-0.5,15]	(13.7421,17.3102)
	CaR	(0.995465,1.01571)	(1.949e-006,6.52754e-006)	[-0.5,1]	(0.569382,0.68939)
	LIN	(-1.0167,-0.923252	(0.0270857,0.0334365)	[-0.5,4.5]	(4.22951,5.5356)

ΠΑΡΑΡΤΗΜΑ

Κώδικες MATLAB

Οι κώδικες που χρησιμοποιήθηκαν στο Matlab για τα συγκεκριμένα μέτρα απόστασης, για τις προτάσεις 2.4, 2.5, και 2.6 είναι οι ακόλουθοι:

Rukhin

```
function[]=rukhin4(M,S,n,b)
    cnt=1;%den mporw na valw a
    for  a=0.8:b:1.2

            a2(cnt)=a;
            f(cnt)= (a-1)*(2-3*M+S)+(1/4)*(12*(a^2)-18*a+12)*(1-2*M+S);

f2(cnt)=(2-2*M-(M^2)+S)+2*(a-1)*(10-13*M-6*(M^2)*(M+8)*S)+((1/3)*((3*(a
-1))^2)*(4-6*M-3*(M^2)+5*S))+((12*(a^2)-18*a+12)/2)*(3-5*M-2*(M^2)+(M+3
)*S);

f3(cnt)=(26-24*M-21*(M^2)-3*(M^3)+(19-3*M)*S)+3*(a-1)*(70-81*M-64*(M^2)
-9*(M^3)+(65+18*M+(M^2))*S)+((3*(a-1))^2)*(20-26*M-21*(M^2)-3*(M^3)+(25
+5*M)*S)+(3*(12*(a^2)-18*a+12)/4)*(15-22*M-15*(M^2)-2*(M^3)+(15+8*M+(M^
2))*S);

            ET0(cnt)=2;
            ET00(cnt)=-0.5;
            ET(cnt)=M-1+(1/n)*f(cnt)-2;
            ET2(cnt)=(M^2)-1+(1/n)*f2(cnt)-8;
            ET3(cnt)=(M^3)+3*(M^2)-M-3+(1/n)*f3(cnt)-48;
            disp(sprintf('a = %g   ET1 =  %g   ET2 = %g   ET3 = %g',
a2(cnt),ET(cnt),ET2(cnt),ET3(cnt)));
            cnt=cnt+1;
    end

    %M->min, I->index

    [M,I] = min(abs(ET(:)));
    disp(sprintf('min: ET1 = %g , a = %g',M,a2(I)));

    [M,I] = min(abs(ET2(:)));
    disp(sprintf('min: ET2 = %g , a = %g',M,a2(I)));

    [M,I] = min(abs(ET3(:)));
    disp(sprintf('min: ET3 = %g , a = %g',M,a2(I)));

    figure; % opens new figure window
    plot(a2,ET,'y');
    title(' a / ET '),xlabel('a'),ylabel('ET')
    hold on ;
plot(a2,ET2,'r');plot(a2,ET3,'b');plot(a2,ET0,'k');plot(a2,ET3,'b');
plot(a2,ET00,'k');legend('ET','ET2','ET3','ET0','ET00');; hold off;
end
```

Cressie and Read

```
function[]=car4(M,S,n,lower,upper,b)
    cnt=1;%den mporw na valw a
    for a=lower:b:upper
        if(a==-1||a==0)

        else
            a2(cnt)=a;
            f(cnt)=((a-1)/3)*(2-3*M+S)+(((a-1)*(a-2))/(4))*(1-2*M+S);

f2(cnt)=(2-2*M-(M^2)+S)+((2*(a-1))/3)*(10-13*M-6*(M^2)*(M+8)*S)+(1/3)*(
(a-1)^2)*(4-6*M-3*(M^2)+5*S)+(((a-1)*(a-2))/2)*(3-5*M-2*(M^2)+(M+3)*S);

f3(cnt)=(26-24*M-21*(M^2)-3*(M^3)+(19+3*M)*S)+(a-1)*(70-81*M-64*(M^2)-9
*(M^3)+(65+18*M+(M^2))*S)+((a-1)^2)*(20-26*M-21*(M^2)-3*(M^3)+(25+5*M)*
S)+((3*(a-1)*(a-2))/(4))*(15-22*M-15*(M^2)-2*(M^3)+(15+8*M+(M^2))*S);

            ET0(cnt)=0.5;
            ET00(cnt)=-2;
            ET(cnt)=M-1+(1/n)*f(cnt)-2;
            ET2(cnt)=(M^2)-1+(1/n)*f2(cnt)-8;
            ET3(cnt)=(M^3)+3*(M^2)-M-3+(1/n)*f3(cnt)-48;
            disp(sprintf('a = %g   ET1 =  %g   ET2 = %g   ET3 = %g',
a2(cnt),ET(cnt),ET2(cnt),ET3(cnt)));
            cnt=cnt+1;
        end
    end

    %M->min, I->index

    [M,I] = min(abs(ET(:)));
    disp(sprintf('min: ET1 = %g , a = %g',M,a2(I)));

    [M,I] = min(abs(ET2(:)));
    disp(sprintf('min: ET2 = %g , a = %g',M,a2(I)));

    [M,I] = min(abs(ET3(:)));
    disp(sprintf('min: ET3 = %g , a = %g',M,a2(I)));

    figure; % opens new figure window
    plot(a2,ET,'y');
    title(' a / ET '),xlabel('a'),ylabel('ET')
    hold on ; plot(a2,ET2,'r');plot(a2,ET3,'b');plot(a2,ET0,'k');
plot(a2,ET00,'k');legend('ET','ET2','ET3','ET0','ET00'); hold off;
end
```

Lin

```
function[]=lin4(M,S,n,lower,upper,b)
    cnt=1;%den mporw na valw a
    for a=lower:b:upper
        if(a==0||a==1)

        else
            a2(cnt)=a;

f(cnt)=((-a-1)/3)*(2-3*M+S)+((1-2*(a^3))/(4-4*a))*(1-2*M+S);

f2(cnt)=(2-2*M-(M^2)+S)-((2*a+2)/3)*(10-13*M-6*(M^2)*(M+8)*S)+(1/3)*((-
a-1)^2)*(4-6*M-3*(M^2)+5*S)+((1-2*(a^3))/(2-2*a))*(3-5*M-2*(M^2)+(M+3)*
S);

f3(cnt)=(26-24*M-21*(M^2)-3*(M^3)+(19+3*M)*S)-(a+1)*(70-81*M-64*(M^2)-9
*(M^3)+(65+18*M+(M^2))*S)+((-a-1)^2)*(20-26*M-21*(M^2)-3*(M^3)+(25+5*M)
*S)+((3-6*(a^3))/(4-4*a))*(15-22*M-15*(M^2)-2*(M^3)+(15+8*M+(M^2))*S);

            ET0(cnt)=6;
            ET00(cnt)=-0.5;
            ET(cnt)=M-1+(1/n)*f(cnt)-2;
            ET2(cnt)=(M^2)-1+(1/n)*f2(cnt)-8;
            ET3(cnt)=(M^3)+3*(M^2)-M-3+(1/n)*f3(cnt)-48;
            disp(sprintf('a = %g   ET1 =  %g   ET2 = %g   ET3 = %g',
a2(cnt),ET(cnt),ET2(cnt),ET3(cnt)));
            cnt=cnt+1;
        end
    end

    %M->min, I->index

    [M,I] = min(abs(ET(:)));
    disp(sprintf('min: ET1 = %g , a = %g',M,a2(I)));

    [M,I] = min(abs(ET2(:)));
    disp(sprintf('min: ET2 = %g , a = %g',M,a2(I)));

    [M,I] = min(abs(ET3(:)));
    disp(sprintf('min: ET3 = %g , a = %g',M,a2(I)));

    figure; % opens new figure window
    plot(a2,ET,'y');
    title(' a / ET '),xlabel('a'),ylabel('ET')

    hold on ; plot(a2,ET2,'r');plot(a2,ET0,'k');plot(a2,ET3,'b');
plot(a2,ET00,'k');legend('ET','ET2','ET3','ET0','ET00'); hold off;
end
```

REFERENCES

1. Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. *Second International Symposium on Information Theory.* (B. N. Petrov and F. Csaki, eds.), 267-281, Akademiai Kaido, Budapest.
2. Ali, S.M. and Silvey, S.D. (1966). A general class of coefficients of divergence of one distribution from another, *J. R. Statist. Soc.* **B 28**, 131–142.
3. Basu, A., Harris, I. R., Hjort, N. L. and Jones, M. C. (1998). Robust and efficient estimation by minimising a density power divergence. *Biometrika*, **85**, 549–559.
4. Basu, A., Shioya, H. and Park, C. (2011). *Statistical Inference: The Minimum Distance Approach*. Boca Raton: Chapman & Hall/CRC.
5. Burbea, J. and Rao, C. R. (1982a). On the convexity of some divergence measures based on entropy functions. *IEEE Transactions on Information Theory*, **28**, 489-495.
6. Burbea, J. and Rao, C. R. (1982b). On the convexity of higher order Jensen differences based on entropy functions. *IEEE Transactions on Information Theory*, **28**, 961-963.
7. Burbea, J. and Rao, C. R. (1982c). Entropy differential metric, distance and divergence measures in probability spaces: *A unified approach. J. of Multivariate Anal*., **12**, 575-596.
8. Cantelli, F. P. (1933). Sulla determinazione empirica delle leggi di probabilita. *Giorn. Ist. Ital. Attuari,* **4,** 421-424.
9. Cressie, N. and Read, T. R. C. (1984). Multinomial goodness-of-fit tests. *J. R. Statist. Soc.*, **B 46**, 440–454.
10. Cressie, N. and Read, T. R. C. (1988). *Goodness-of-Fit Statistics for Discrete Multivariate Data*, Springer Verlag, New York.
11. Csiszár, I. (1963). Eine Informationstheoretische Ungleichung und ihre Anwendung auf den Bewis der Ergodizitat on Markhoffschen Ketten. *Publication of the Mathematical Institute of the Hungarian Academy of Sciences*, **8**, 84-108.
12. Carleman, T. (1926). *Les fonctions quasi analytiques* (in French). Paris: Gauthier-Villars.
13. H.S. Chen, K. Lai, and Ying, Z. (2004). Goodness of fit tests and minimum power

divergence estimators for survival data, *Statist. Sinica* **14**, 231–248.

14. D'Agostino, R.B. and Stephens, M.A. (1986). *Goodness-of-Fit Techniques*. Marcel Dekker, New York.
15. Glivenko, V. (1933). Sulla determinazione empirica della legge di probabilita. *Giorn. Ist. Ital. Attuari* , **4**, 92-99.
16. Grinstead, C.M. and Snell, J. L. (1997). *Introduction to Probability*, 2nd revised edition, AMS.
17. Heyde, (1963) On a property of the lognormal distribution, *J. R. Stat. Soc. Ser. B Stat. Methodol*, **25**, 392-393.
18. Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society, Series A*, 186, 453-561.
19. Kagan, A. M. (1963) On the theory of Fisher's amount of information, *Sov. Math. Dokl*. **4**, 991–993.
20. Kapur, J. N. (1972). Measures of uncertainty, mathematical programming and physics. *Journal of Indian Society of Agriculture and Statistics*, **24**, 47-66.
21. Kolmogorov, A. N. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giornale dell' Istituto Italiano degli Attuari*, **4**, 83-91.
22. Kullback, S. and Leibler, R. (1951). On information and sufficiency. *Annals of Mathematical Statistics*, **22**, 79-86.
23. Lindsay, B. G. (1994). Efficiency versus robustness: the case for minimum Hellinger distance and related methods. *Ann. Statist*., **22**, 1081-1114.
24. Lee, S., Vonta, I, and Karagrigoriou, A. (2011). A maximum entropy type test of fit, *Comput. Statist. and Data Anal*, ***55***, *2635-2643.*
25. Liese, F. and Vajda, I. (1987). *Convex Statistical Distances*. Teubner, Leipzig.
26. Makrides, A., Karagrigoriou, A., and Vonta, F. (2013). Entropy type goodness of fit tests for heavy-tailed distributions, In *Statistical Models and Methods for Reliability and Survival Analysis*, Couallier, V., Gerville-Reache, L., Huber, C., Limnios N. and Mesbah, M. eds, Hermes-Wiley, 33-44.
27. Mattheou, K. and Karagrigoriou, A. (2010). A new family of divergence measures for tests of fit, *Austr. and N. Zealand J. of Statist.*, **52** (2), 187-200.
28. Matusita, K. (1967). On the notion of affinity of several distributions and some of

its applications, *Ann. Inst. Statist. Math*. **19**, 181–192.

29. Pardo, L. (2006). *Statistical Inference Based on Divergence Measures*. Chapman & Hall/CRC, Boca Raton.
30. Pearson, K. (1900). On the criterion that a given system of deviation from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophy Magazine* **50**, 157-172.
31. Rao, C. R. (1982). Diversity: Its measurement, decomposition, apportionment and analysis. *Sankhya A*, **44**, 1-22.
32. Renyi, A. (1961). On measures of entropy and information. *Proc. of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, **1**, 547 – 561.
33. Rukhin, A. L. (1994). Optimal estimator for the mixture parameter by the method of moments and information affinity. *Transactions 12th Prague Conference on Information Theory*, 214-219.
34. Salicru, M. Menendez, M. L. and Pardo, L. (1993). Asymptotic distribution of (h,φ)-entropies. *Commun. in Statist. Theory and Methods*, **22 (7)**, 2015-2031.
35. Serfling, R. J. (1980). *Approximations Theorems of Mathematical Statistics.* John Wiley, New York.
36. Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal* **27**, 379-423 & 623-656.
37. Shohat & Tamarkin, (1943). The Problem of Moments. *American Mathematical Society*.
38. Vonta, F. and Karagrigoriou, A. (2010). Generalized measures of divergence in survival analysis and reliability. *J. of Applied Prob*., **47 (1)** 216-234.
39. Vonta, F. and Karagrigoriou, A. (2011). Information measures in biostatistics and reliability, In *Mathematical and Statistical Models and Methods in Reliability*, Rykov, V. V., Balakrishnan, N., and Nikulin (eds), Birkhauser, Boston, 401- 413.
40. Vonta, I. and Karagrigoriou, A. (2013). Goodness of fit tests via φ-measures of divergence for censored data, *J. Statist. Comput. Simul*., **84 (5),** 946-963.
41. Vonta, F., Mattheou, K. and Karagrigoriou, A. (2012). On properties of the (Φ,α)-power divergence family, *Meth. & Comput. Appl. Prob*., **Vol. 14**, 335-356.

42. Zografos, K. (1993) Asymptotic properties of Φ -divergence statistic and applications in contingency tables. *Intern. J. of Math. & Statist. Sc.*, **2**, 5-21.

43. Zografos, K., Ferentinos, K. and Papaioannou, T. (1990). Φ–divergence statistics: Sampling properties, multinomial goodness of fit and divergence tests. *Communication in Statistics (Theory and Methods)*, **19** (5), 1785 – 1802.