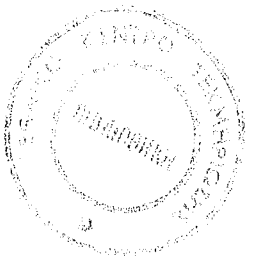
ΑΒΕΚΤ 00274

2714

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ



ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΜΕΘΟΔΟΙ ΑΞΙΟΛΟΓΗΣΗΣ ΤΗΣ ΠΡΟΒΛΕΠΤΙΚΗΣ ΙΚΑΝΟΤΗΤΑΣ ΓΡΑΜΜΙΚΩΝ ΜΟΝΤΕΛΩΝ

ΣΤΥΛΙΑΝΟΣ Α. ΨΑΡΑΚΗΣ

ΑΘΗΝΑ 1993

ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 1

Γενική Εισαγωγή. 1

ΚΕΦΑΛΑΙΟ 2

Ελεγχος της προβλεπτικής ικανότητας ενός γραμμικού μοντέλου με την χρησιμοποίηση της χ^2 κατανομής και συγκριτική αξιολόγηση δύο προβλεπτικών γραμμικών μοντέλων με την χρησιμοποίηση μιάς γενικευμένης μορφής της F κατανομής.

2.1 Εισαγωγή. 26

2.2 Ελεγχος για την αξιολόγηση της προβλεπτικής ικανότητας ενός γραμμικού μοντέλου με την χρησιμοποίηση της χ^2 κατανομης. 27

2.3 Συγκριτική αξιολόγηση της προβλεπτικής ικανότητας ενός γραμμικού μοντέλου με τη χρησιμοποίηση μιάς γενικευμένης μορφής της F κατανομής. 30

Προσδιορισμός της κατανομής του λόγου X/Y όταν X και Y ακολουθούν την διμεταβλητή γάμμα του Kibble. 44

ΚΕΦΑΛΑΙΟ 3

Μελέτη ιδιοτήτων των κατανομών folded normal, folded standard normal και προσδιορισμός του αθροίσματος ανεξάρτητων folded standard normal τυχαίων μεταβλητών.

3.1 Εισαγωγή. 53

3.2 Περίπτωση μεγάλου αριθμού παρατηρήσεων. 54

3.2.1 Ροπογεννήτρια της folded normal κατανομής. 55

3.2.2 Προσδιορισμός της κατανομής του αθροίσματος δύο ανεξάρτητων τυχαίων μεταβλητών που ακολουθούν την folded standard normal κατανομή. 57

3.2.3 Ευρεση της κατανομής της διαφοράς δύο ανεξάρτητων τυχαίων μεταβλητών που ακολουθούν την folded standard normal κατανομή. 60

3.2.4 Ανάπτυξη αλγορίθμου για την εύρεση τοσοστιαίων σημείων της folded normal κατανομής. 63

3.2.5 Η κατανομή του αθροίσματος n ανεξάρτητων folded standard normal τυχαίων μεταβλητών. 66

3.3 Περίπτωση μικρού αριθμού παρατηρήσεων. 72

Μελέτη της κατανομής folded t. 73

Σχέση της κατανομής folded t με την κατανομή folded normal. 78

Σχέσεις της folded t κατανομής με κατανομές που έχουν σχέση με παραγωγή τυχαίων αριθμών. 79

Περιοχές της folded t κατανομής. 84

ΚΕΦΑΛΑΙΟ 4

Οι διμεταβλητές μορφές της folded normal και folded t κατανομής.

4.1 Εισαγωγή 86

4.2 Η διμεταβλητή folded normal κατανομή. 87

Περιθώριες κατανομές της διμεταβλητής folded standard normal κατανομής. 94

Ροπογεννήτρια της διμεταβλητής folded standard normal κατανομής. 97

Συνάρτηση κατανομής της διμεταβλητής folded standard normal κατανομής. 99

4.3 Οι κατανομές διμεταβλητή folded t και διμεταβλητή folded standard t. 104

Σχέση της διμεταβλητής folded t με την διμεταβλητή folded standard normal κατανομή. 115

Σχέσεις της διμεταβλητής folded t με άλλες κατανομές. 116

Προσδιορισμός των περιθώριων κατανομών της διμεταβλητής folded t και της διμεταβλητής folded standard t. 117

Η συνάρτηση κατανομής της διμεταβλητής folded t. 118

ΚΕΦΑΛΑΙΟ 5

Άλλες μέθοδοι αξιολόγησης της προβλεπτικής ικανότητας γραμμικών μοντέλων.

5.1 Εισαγωγή. 123

5.2 Μία μη παραμετρική μέθοδος αξιολόγησης της προβλεπτικής ικανότητας ενός γραμμικού μοντέλου βασιζμενη στον έλεγχο σήματος της διωνυμικής κατανομής. 123

5.3 Μέθοδος για την συγκριτική αξιολόγηση δύο εναλλακτικών γραμικών μοντέλων που βασίζεται στον έλεγχο του συντελεστή συσχέτισης. 125

ΒΙΒΛΙΟΓΡΑΦΙΑ 129

ΠΑΡΑΡΤΗΜΑ I 148

ΠΑΡΑΡΤΗΜΑ II 169

ΠΑΡΑΡΤΗΜΑ III 178

ΠΑΡΑΡΤΗΜΑ IV 191

ΠΑΡΑΡΤΗΜΑ V 212

ΠΕΡΙΛΗΨΗ

Η επιλογή του κατάλληλου γραμμικού μοντέλου για την μελέτη ενός συγκεκριμένου προβλήματος έχει εξετασθεί από πολλές πλευρές στη διεθνή βιβλιογραφία Η έμφαση έχει δοθεί, κυρίως, στην αξιολόγηση της περιγραφικής ικανότητας των μοντέλων που είναι υποψήφια για χρησιμοποίηση.

Πρόσφατα, αναπτύσσονται μεθοδολογίες που αξιολογούν την προβλεπτική ικανότητα γραμμικών μοντέλων.

Η εργασία αυτή κινείται προς την τελευταία αυτή κατεύθυνση. Συγκεκριμένα, γίνεται προσπάθεια να αποτιμηθεί η προβλεπτική ικανότητα ενός μοντέλου με βάση την σύγκριση των παρατηρούμενων τιμών με τις αντίστοιχες αναμενόμενες τιμές της εξαρτημένης μεταβλητής σε μία ακολουθία χρονικών στιγμών που το υπό εξέταση μοντέλο έχει εφαρμοσθεί. Η αξιολόγηση στηρίζεται σε μία "βαθμολογία" του υπό κρίση μοντέλου κάθε χρονική στιγμή για την οποία υπάρχουν παρατηρήσεις. Η τελική "απόδοση" του μοντέλου εκφράζεται ως συνάρτηση της "βαθμολογίας" αυτής.

Για τον σκοπό αυτό κατασκευάζονται διάφορες στατιστικές συναρτήσεις (οι οποίες είναι συναρτήσεις της παρατηρούμενης και της αναμενόμενης τιμής της εξαρτημένης μεταβλητής).

Η προσέγγιση αυτή οδηγεί στην ανάγκη της μελέτης των κατανομών των στατιστικών αυτών συναρτήσεων ώστε να καταστεί δυνατή η εξαγωγή στατιστικών συμπερασμάτων. Μεταξύ των κατανομών που προκύπτουν και μελετώνται άλλες ορίζονται για

πρώτη φορά και άλλες έχουν μόνο αναφερθεί στη βιβλιογραφία χωρίς να έχουν μελετηθεί. Τέτοιες κατανομές είναι η folded-t (η οποία είναι η απόλυτη τιμή της t) , η bivariate folded-t και bivariate folded normal, καθώς επίσης και μία γενικευμένη μορφή της F κατανομής.

Στην κατεύθυνση που αναφέρθηκε αναπτύσσεται στατιστική συμπερασματολογία για την αξιολόγηση της προβλεπτικής ικανότητας ενός μοντέλου αλλά και συγκριτικής αξιολόγησης δύο μοντέλων.

ΚΕΦΑΛΑΙΟ 1

ΓΕΝΙΚΗ ΕΙΣΑΓΩΓΗ

Ενα σημαντικό πρόβλημα στην στατιστική και ειδικότερα στην ανάλυση δεδομένων είναι η επιλογή του καλύτερου μοντέλου το οποίο προσαρμόζεται στα υπό ανάλυση δεδομένα. Μερικές φορές η επιλογή αυτή ειναι αρκετά εύκολη αλλες φορές όμως είναι αρκετά δύσκολη.

Ενα πρόβλημα που επίσης εξετάζεται πολύ τα τελευταία χρόνια είναι η ανάπτυξη μεθοδολογιών για την αξιολόγηση της προβλεπτικής ικανότητας γραμμικών μοντέλων. Προς αυτή την κατεύθυνση κινείται η εργασία των Xekalaki-Katti (1984). Στην εργασία αυτή εισάγεται μία διαδικασία αξιολόγησης της προβλεπτικής ικανότητας ενός στοχαστικού γραμμικού μοντέλου αλλά και συγκριτικής αξιολόγησης δύο στοχαστικών γραμμικών μοντέλων.

Αναλυτικά:

Θεωρούμε το μοντέλο πρόβλεψης:

$$Y_t = X_t\beta + e_t$$

όπου

Y_t είναι διάνυσμα διαστάσεως $l_t \times 1$ των εξαρτημένων μεταβλητών

X_t πίνακας διαστάσεως $l_t \times m$ $(l_t > m, |X_t X_t| \neq 0)$ γνωστών συντελεστών

β διάνυσμα διαστάσεως m×1 συντελεστών παλινδρόμησης και

e_t διάνυσμα διαστάσεως $l_t \times 1$ με $E(e_t)=0$ και $V(e_t)=\sigma_t^2 I_t$ όπου I_t είναι ο $l_t \times l_t$ ταυτοτικός πίνακας.

Μία πρόβλεψη για την εξαρτημένη τυχαία μεταβλητή θα δίνεται από την στατιστική συνάρτηση:

$$\hat{Y}^{\circ}_{t+1} = X^{\circ}_{t+1} \hat{\beta}_t$$

όπου

$\hat{\beta}_t$ είναι ο εκτιμητής ελαχίστων τετραγώνων του β τον χρόνο t και

$\hat{X}^{\circ}_{t+1}$ είναι ένα 1×m διάνυσμα των τιμών των συντελεστών παλινδρόμησης την στιγμή t+1 t=0,1,...

$$\text{Έτσι } X_{t+1} = \begin{bmatrix} X_t \\ X^{\circ}_{t+1} \end{bmatrix} \quad \text{και} \quad Y_{t+1} = \begin{bmatrix} Y_t \\ Y^{\circ}_{t+1} \end{bmatrix}$$

Η μέθοδος αξιολόγησης των Xekalaki-Katti στηρίζεται στην κατασκευή ενός διαστήματος εμπιστοσύνης του οποίου το μήκος αυξάνεται η ελαττώνεται ανάλογα με το αν η πρόβλεψη είναι επιτυχής η μη επιτυχής. Επιτυχής θεωρείται η πρόβλεψη αν η πραγματική τιμή της εξαρτημένης μεταβλητής βρίσκεται μέσα σ'ένα "διάστημα πίστης" το οποίο κατασκευάζεται γύρω από την πρόβλεψη. Ακολούθως τα ενδεχόμενα, ανάλογα με το αν η πρόβλεψη είναι επιτυχής η μη επιτυχής, βαθμολογούνται με κάποιο βαθμολογικό κανόνα (scoring rule).Η τελική απόφαση βασίζεται στη μέση τιμή των αξιολογήσεων.

Πριν προχωρήσουμε όμως στην επεξεργασία του προβλήματος που εκθέσαμε καλό είναι να δούμε τις διάφορες μεθοδολογίες που έχουν αναπτυχθεί σχετικά με το πρόβλημα του πόσο καλά τα

δεδομένα περιγράφονται καλύτερα από ένα μοντέλο.

Για την επίλυση αυτού του προβλήματος πολλά κριτήρια εχουν κατά καιρους προταθεί στη βιβλιογραφία. Στην ενότητα αυτή αναφέρουμε μερικά απο τα πιό γνωστά κριτήρια επιλογής μοντέλων.

Ας θεωρήσουμε ένα στατιστικό μοντέλο με M παραμέτρους του οποίου την προσαρμογή στα δεδομένα θέλουμενα εξετάσουμε. Ενα από τα πιό γνωστά κριτήρια ειναι το:

Κριτήρια του Akaike (Akaike Information Criterion AIC) , Bayesian Information Criterion BIC):

Το κριτήριο του Akaike (1973,1974) βασίζεται στην ελαχιστοποίηση της παρακάτω ποσότητας:

$$AIC\ (M,n) = \ln(\hat{\sigma}_n^2) + 2M/n$$

όπου n ειναι το μέγεθος του δείγματος, $\hat{\sigma}_n^2$ είναι η εκτίμηση της διασποράς των καταλοίπων και M όπως προαναφέρθηκε είναι το πλήθος των παραμέτρων.

Σχετικά πιο πρόσφατα ο Akaike (1978,1979) όρισε μια Μπεϋζιανή επέκταση του AIC κριτηρίου το οποίο ονόμασε BIC (Bayesian Information Criterion) και το οποίο δίνεται από:

$$BIC(M,n) = n\ln(\hat{\sigma}_n^2) - (n-M)\ln\left(1-\frac{M}{n}\right) +$$

$$+ M\ln(n) + M_z \ln\left[\left(\frac{\sigma_z^2}{\hat{\sigma}_n^2} - 1\right)/M\right]$$

όπου $M, n, \hat{\sigma}_n^2$ ορίζονται όπως προηγουμένως και $\hat{\sigma}_z^2$ είναι η δειγματική διασπορά της σειράς των δεδομένων.

Με βάση το κριτήριο του Akaike και την επέκταση του έχουν

ορισθεί μία σειρα από άλλα χριτήρια τα οποία αναφέρουμε στην συνέχεια:

Το χριτήριο του Schwarz (Schwarz's Bayesian Information Criterion-SBIC).

Το χριτήριο του Schwarz (1978) είναι παρόμοιο με το BIC του Akaike χαι βασίζεται στην ελαχιστοποίηση της παραχάτω ποσότητας:

$$SBIC(M,n) = \ln(\hat{\sigma}_n^2) + M(\ln(n))/n$$

Το χριτήριο του Parzen (Parzen's criterion for autoregressive transfer functions CAT).

Ο Parzen (1977) πρότεινε το παραχάτω χριτήριο το οποίο ονόμασε CAT:

$$CAT(p) = \begin{cases} -(1+\frac{1}{n}), & p=0 \\ \frac{1}{n}\sum_{j=1}^{n}\frac{1}{\hat{\sigma}_j^2} - \frac{1}{\hat{\sigma}_p^2} & p=1,2,3 \end{cases}$$

όπου $\hat{\sigma}_j^2$ είναι η αμερόληπτη εχτιμήτρια του σ_n^2 όταν ένα AR(j) (Autoregressive σειρά τάξεως j) μοντέλο προσαρμόζεται στα δεδομένα χαι n είναι ο αριθμός των παρατηρήσεων. Η σειρά του p επιλέγεται όταν το CAT(p) γίνει ελάχιστο.

Το χριτήριο του Amemiya (Amemiya Prediction Criterion PC)

Ο Amemiya (1980) πρότεινε ένα χριτήριο παρόμοιο με εχείνο του Akaike το οποίο βασίζεται στην ελαχιστοποίηση της τοσότητας:

$$PC(M,n)= \hat{\sigma}^2_n \frac{n+M}{n-M}$$

Το κριτήριο του Hannan (Hannan's criterion HC)

Ο Hannan (1981) πρότεινε επίσης ενα κριτήριο το οποίο βασίζεται στην ελαχιστοποίηση της ποσότητας:

$$HC(M,n)= \ln(\hat{\sigma}^2_n) + M(2 \ln \ln(n))/n$$

Τέλος ένα παλαιότερο κριτήριο είναι:

Το κριτήριο του Theil (Theil's residual variance criterion RVC)

Το κριτήριο αυτό προτάθηκε το 1961 απο τον Theil και βασίζεται στην ελαχιστοποίηση της ποσότητας:

$$RVC(M,n)= \hat{\sigma}^2_n [n/(M-n)]$$

Ενας σημαντικός αριθμός επιστημόνων έχει ασχοληθεί με τα κριτήρια που αναφέραμε παραπάνω με τις ιδιότητες αλλά και με την διόρθωση τους κάτω από ορισμένες συνθήκες.

Ετσι ενδιαφέρον παρουσιάζει η εργασία του Chow (1981) ο οποίος δίνει μία βελτίωση του AIC και επίσης συγκρίνει το κριτήριο του Akaike με εκείνο του Schwarz.

Προς την ίδια κατεύθυνση βρίσκεται και η εργασία του Shibata(1984) όπου δίνεται ένας προσεγγιστικός κανόνας για την αποδοτικότητα του κριτηρίου του τελικού πρβλεπτικού σφάλματος (Final Prediction Error FPE) απ'όπου τελικά προέρχονται όλα τα παραπάνω κριτήρια που αναφέρθηκαν.

Ενδιαφέρον επίσης παρουσιάζει και η εργασία των Hurvich

και Tsai (1989) όπου δίνεται μία διόρθωση του κριτηρίου του Akaike στην περίπτωση που ο αριθμός των παρατηρήσεων είναι μικρός. Η διόρθωση του κριτηρίου που προτείνεται είναι:

$$AIC_c = AIC + \frac{2(M+1)(M+2)}{n-M-2}$$

δηλαδή το AIC_c είναι το άθροισμα του κριτηρίου του Akaike και ενός παράγοντα που εξαρτάται από το μέγεθος n του δείγματος καθώς επίσης και από το μέγεθος των παραμέτρων M του μοντέλου.

Στην εργασία των Smith και Spiegelhalter (1980) ορίζονται σφαιρικοί και τοπικοί Μπεϋζιανοί παράγοντες (Global and Local Bayes factors) και δίνεται ο ρόλος τους για την επιλογή μεταξύ γραμμικών μοντέλων. Επίσης δίνεται η σχέση των global Bayes factors με το κριτήριο του Schwarz καθώς επίσης και η σχέση των local Bayes factor με το κριτήριο του Akaike.

Οι Holmes και Huton (1989) στην εργασία τους κάνουν σύγκριση των κριτηρίων που αναφέρθηκαν στην αρχή της ενότητας και καταλήγουν στο συμπέρασμα ότι κάτω από ορισμένες συνθήκες το κριτήριο του Theil είναι το πιό αξιόπιστο.

Τέλος θα πρέπει να σημειώσουμε ότι ενδιαφέρον παρουσιάζουν και οι εργασίες των Zaman (1984) και Urbain (1989) οι οποίες μελετούν την σχέση των διαφόρων κριτηρίων που αναφέραμε παραπάνω με αλλες τεχνικές και δίνονται, στην πρώτη εργασία, μειονεκτήματα των διαφόρων κριτηρίων.

Επιλογή μοντέλων με μεθόδους που βασίζονται σε λάθη πρόβλεψης

Αν ο κύριος σκοπός ενός μοντέλου είναι η πρόβλεψη μελλοντικών τιμών τότε κριτήρια επιλογής μοντέλων μπορούν να βασισθούν στα λάθη πρόβλεψης.

Ετσι αν το λάθος τρόβλεψης βήματος 1 (1 step ahead forecast error) δίνεται από:

$$e_1 = Y_{n+1} - \hat{Y}_n(1)$$

τότε η σύγκριση στηρίζεται στα παρακάτω κριτήρια:

1. Μέσο εκατοστιαίο λάθος (Mean Percentage Error MPE)

$$MPE = \left(\frac{1}{M} \sum_{l=1}^{M} \frac{e_l}{Y_{n+1}} \right) 100\%$$

2. Μέσο τετραγωνικό σφάλμα (Mean Square Error MSE)

$$MSE = \frac{1}{M} \sum_{l=1}^{M} e_l^2$$

Το μέσο τετραγωνικό σφάλμα είναι ένας μεροληπτικός εκτιμητής της διασποράς των σφαλμάτων του γραμμικού μοντέλου.

Ενας αμερόληπτος εκτιμητής της διασποράς των σφαλμάτων που χρησιμεύει στην επιλογή μοντέλων είναι ο:

$$s^2 = \frac{1}{M-2} \sum_{l=1}^{M} e_l^2$$

Το καλύτερο μοντέλο είναι εκείνο που δίνει την μικρότερη

τιμή για το s^2.

3. Μέσο απόλυτο λάθος (Mean Absolute Error)

$$MAE = \frac{1}{M} \sum_{i=1}^{M} |e_i|$$

4. Μέσο απόλυτο εκατοστιαίο λάθος (Mean absolute percentage error MAPE)

$$MAPE = \left(\frac{1}{M} \sum_{i=1}^{M} \left| \frac{e_i}{Y_{n+1}} \right| \right) 100\%$$

5. Μέσο απόλυτο σχετικό λάθος (Mean Absolute Relative Deviation MARD)

$$MARD = \frac{1}{M} \left(\sum_{i=1}^{M} |e_i| \; / \; Y_i \right)$$

όπου $e_i = Y_i - \hat{Y}_i$ είναι το ένα βήμα μπροστά λάθος πρόβλεψης (one step ahead forecast error).

6. Συντελεστής προσδιορισμού (Coefficient of determination R^2).

$$R^2 = \frac{\sum_{i=1}^{M} (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{M} (Y_i - \bar{Y})^2}$$

Ο συντελεστής προσδιορισμού παίρνει τιμές στο διάστημα από 0 έως 1. Οπως ειναι γνωστο τιμές του R^2 κοντά στο 1 αποτελούν ένδειξη καλής προσαρμογής της ευθείας παλινδρόμησης στα δεδομένα του δείγματος. Αντίθετα τιμές του R^2 κοντά στο 0 αποτελούν ένδειξη κακής προσαρμογής. Στο σημείο αυτό θα πρέπει να σημειωθεί ότι ο συντελεστής προσδιορισμού δεν ενδείκνυται σαν το μόνο κριτήριο για την επιλογή του καλύτερου προβλεπτικού μοντέλου.

Ο Snee(1977) πρότεινε διάφορους τρόπους για την αξιολόγηση των μοντέλων:

1.Σύγκριση των συντελεστών πρόβλεψης($\hat{\beta}_i$) και των προβλέψεων του μοντέλου ($\hat{Y}_i$) με την φυσική θεωρία.

Η μέθοδος αυτή βασίζεται στο οτι αρνητικές προβλέψεις μιάς θεωρητικά θετικής ποσότητας ή συντελεστών με λάθος πρόσημο είναι ένδειξη μοντέλου με χαμηλή προβλεπτική ικανότητα.

2. <u>Συλλογή νέων δεδομένων για τον έλεγχο της προβλεπτικότητας του μοντέλου</u>.

Η μέθοδος αυτή στηρίζεται στην άποψη ότι η προβλεπτικότητα ενός μοντέλου μπορεί να ελεγχθεί με τη συλλογή δεδομένων τα οποία θα είναι δυνατόν να συγκριθούν με τις προβλέψεις του μοντέλου.

3.<u>Σύγκριση των αποτελεσμάτων με θεωρητικά και προσομοιομένα δεδομένα</u>.

Στην μέθοδο αυτή γίνεται σύγκριση των πραγματικών παρατηρήσεων με προσομοιομένες. Αν υπάρχει συμφωνία τότε μπορούμε να πούμε ότι το μοντέλο είναι αρκετά ικανοποιητικό.

4. Διαχωρισμός δεδομένων. (Data Splitting)

Από τα παραπάνω κριτήρια που αναπτύχθηκαν συμπεραίνουμε ότι η συλλογή καινούργιων δεδομένων αποτελεί μία από τις καλύτερες μεθόδους για την αξιολόγηση του μοντέλου. Πολλές φορές όμως, αφένός εξαιτίας της μορφής του μοντέλου και αφετέρου για πρακτικούς λόγους, δεν είναι δυνατο γίνει αυτό. Ετσι θα πρέπει να στηριχθούμε στα δεδομένα που έχουμε για την αξιολόγηση του μοντέλου. Στην περίπτωση αυτή μία διαδικασία η οποία προσομοιώνει την συλλογή νέων δεδομένων είναι αναγκαία. Η διαδικασία αυτή χωρίζει τα δεδομένα με κάποιο λογικό τρόπο σε δύο σύνολα. Το πρώτο σύνολο δεδομένων που καλούνται δείγμα προσαρμογής (fitting sample) χρησιμοποιούνται για την εκτίμηση των παραμέτρων του μοντέλου. Τα υπόλοπα δεδομένα τα οποία καλούνται δείγμα επιβεβαίωσης (validation sample) χρησιμοποιούνται για την μέτρηση της προβλεπτικής ικανότητας του μοντέλου.

Η μέθοδος διαχωρισμού των δεδομένων είναι πολύ γνωστή στην βιβλιογραφία και πολλοί συγγραφείς έχουν ασχοληθεί με αυτή. Χρησιμοτοιώντας την λογική αυτής της μεθόδου μία διαδικασία που επιτρέπει την αξιολόγηση μοντέλων και η οποία έχει προταθεί από τον Allen (1971) είναι η:

Προβλεπτικό άθροισμα τετραγώνων (Predictive Sum of Squares PRESS).

Η μέθοδος στηρίζεται στην ελαχιστοποίηση της ποσότητας:

$$PRESS = \sum_{i=1}^{n} (y_i - \hat{y}_{i,-i})^2 = \sum_{i=1}^{n} (e_{i,-i})^2$$

όπου $\hat{y}_{i,-i}$ είναι η προβλεπόμενη τιμή στην i περίοδο αν εξαιρέσουμε την παρατήρηση i. Στην μέθοδο αυτή θεωρούμε ένα σύνολο δεδομένων στα οποία παραμερίζουμε την πρώτη παρατήρηση του δείγματος και χρησιμοποιούμε τις υπόλοιπες για να εκτιμήσουμε τους συντελεστές του υποψηφίου μοντέλου. Επαναλαμβάνουμε την ίδια διαδικασία συνεχώς απομακρύνοντας κάθε φορά από μία παρατήρηση και έτσι το υποψήφιο μοντέλο προσαρμόζεται συνολικά n φορές όσες και ο συνολικός αριθμός παρατηρήσεων.

Στην βιβλιογραφία υπάρχουν αρκετές εργασίες γύρω από τον διαχωρισμό των δεδομένων και την στατιστική συνάρτηση PRESS. Οι πιό σημαντικές από αυτές είναι των Magee και Veall(1991) η οποία μελετάει κάποιες ιδιότητες του κριτηρίου PRESS ενώ η εργασία των Picard και Berk (1990) δίνει ιδιότητες και κάνει κριτική της μεθόδου του διαχωρισμού των δεδομένων.

Επιλογή συντελεστών παλινδρόμησης

Μία πολλή γνωστή μέθοδος για την προσαρμογή ενός μοντέλου σ'ένα σύνολο δεδομένων είναι η μέθοδος των ελαχίστων τετραγώνων. Εκτός όμως από το πρόβλημα της προσαρμογής των δεδομένων, στην πολλαπλή παλινδρόμηση υπάρχει και το πρόβλημα του κατά πόσον κάποιες από τις μεταβλητές παλινδρόμησης έχουν σημαντική προσφορά στο μοντέλο. Στην πολλαπλή παλινδρόμηση

παρέχεται η στατιστική συμπερασματολογία για τον καθορισμό του κατά πόσο μία μεταβλητή είναι σημαντική με τον έλεγχο της μηδενικής υπόθεσης

$H_0 : \beta_i = 0$ έναντι της εναλλακτικής

$H_1 : \beta_i \neq 0$ $i=0,1,\ldots,k$

Αν η H_0 δεν απορριφθεί για κάποια τιμή του i συμπεραίνουμε ότι δεν υπάρχουν στοιχεία ικανά να μας πείσουν ότι η αντίστοιχη μεταβλητή έχει συνεισφορά σημαντική στο μοντέλο. Στην περίπτωση αυτή ο όρος $\beta_i X_i$ διαγράφεται από το μοντέλο.

Για την επιλογή των συντελεστών παλινδρόμησης έχουν προταθεί πολλές διαδικασίες. Οι κυριώτερες από αυτές είναι:

1. Μέθοδος Αποκλεισμού Μεταβλητών (Backward Elimination Procedure).

Η μέθοδος αυτή ξεκινά περιλαμβάνοντας όλες τις μεταβλητές στο μοντέλο και σε κάθε βήμα αποκλείει μία μεταβλητή που δεν έχει σημαντική συνεισφορά σε αυτό. Η μέθοδος αποκλεισμού των μεταβλητών αποκλείει σε κάθε βήμα την μεταβλητή εκείνη που έχει το μεγαλύτερο παρατηρούμενο επίπεδο σημαντικότητας με την προυπόθεση ότι το παρατηρούμενο επίπεδο σημαντικότητας υπερβαίνει το επίπεδο σημαντικότητας που έχουμε προκαθορίσει.

2. Μέθοδος Προοδευτικής Προσθήκης Μεταβλητών (Forward Selection Procedure)

Η μέθοδος της προοδευτικής προσθήκης μεταβλητών χρησιμοποιεί την ακριβώς αντίθετη λογική από αυτήν της προήγουμενης μεθόδου. Εδώ το μοντέλο αναπτύσσεται με την

προσθήκη κάθε φορά μιάς ανεξάρτητης μεταβλητής. Το πλεονέκτημα της μεθόδου αυτής είναι ότι απαιτούνται λιγότερες πράξεις στον υπολογιστή. Η μέθοδος όμως αυτή παρουσιάζει μειονεκτήματα σε σχέση με την προηγούμενη από την ιδιαιτερότητα που παρατηρείται λόγω των μεγαλύτερων τιμών των συντελεστών συσχέτισης των εξαρτημένων μεταβλητών με τις ανεξάρτητες. Ενα άλλο μειονέκτημα είναι ότι υπάρχει η δυνατότητα μόνο προσθήκης ανεξάρτητων μεταβλητών οι οποίες από τη στιγμή που θα ενσωματωθούν στη διδικασία διατηρούνται έστω και αν η τιμή της στατιστικής συνάρτησης

$$T = \frac{\beta_i}{S_{\hat{\beta}_i}}$$

όπου β_i είναι η εκτιμήτρια ελαχίστων τετραγώνων του συντελεστή β_i της μεταβλητης X_i στο γενικό γραμμικό μοντέλο και $S_{\hat{\beta}_i}$ είναι η εκτιμούμενη τυπική απόκλιση της εκτιμήτριας $\hat{\beta}_i$ του συντελεστή παλινδρόμησης σε κάποιο βήμα δεν υπερβαίνει το χρίσιμο επίπεδο.

3.Μεθοδος της Βηματικής Παλινδρόμησης (Stepwise Regrssion)

Η μέθοδος αυτή είναι παρόμοια με τη μέθοδο της προοδευτικής προσθήκης μεταβλήτών. Η διαφορά των δύο μεθόδων έγκειται στο γεγονός ότι για κάθε διαδοχικό βήμα η υπόθεση:

$$H_0 : \beta_i = 0$$

ελέγχεται για όλες τις δυνατές ανεξάρτητες μεταβλητές ώστε να αποκλείονται εκείνες για τις οποίες οι τιμές της στατιστικής συνάρτησης $|T_i|$ είναι μικρότερη από ένα προκαθορισμένο χρίσιμο επίπεδο. Η επόμενη μεταβλητή προστίθεται με την ίδια διαδικασία

όπως στη μέθοδο της προοδευτικής προσθήκης μεταβλητών. Αυτή η βηματική επιλογή συνεχίζεται μέχρις ότου καταλήξουμε σ'ένα σύνολο μεταβλητών για το οποίο καμμία από τις μεταβλητές που περιέχει δεν έχουν τιμή για τη στατιστική συνάρτηση $|T_i|$ μικρότερη από κάποια συγκεκριμμένη κρίσιμη τιμή της μεταβλητής t και δεν υπάρχουν άλλες μεταβλητές που θα πρέπει να αξιολογηθούν για να περιληφθούν στο μοντέλο. Θα πρέπει εδώ να σημειώσουμε το ενδιαφέρον που παρουσιάζει η εργασία των Forsythe et. al. (1973) σχετικά με την προαναφερόμενη μέθοδο.

Μία μέθοδος πού δίνει τον αριθμό των μεταβλητών που θα πρέπει να εισαχθούν στο μοντέλο είναι και η μέθοδος του:

4. Πολλαπλού Συντελεστη Συσχέτισης (Multiple Correlation Coefficient)

Η μέθοδος αυτή προτάθηκε από τον Hocking (1976) και προβλέπει τον υπολογισμό της ποσότητας

$$R_p^2 = 1 - \frac{SSE_p}{SST}$$

για όλα τα δυνατά υποψήφια σύνολα μεταβλητών όπου SSE_p είναι το αθροισμα τετραγώνων των σφαλμάτων για τις p μεταβλητές και SST είναι το συνολικό άθροισμα τετραγώνων. Ενα σχεδιάγραμμα του μεγίστου R_p^2 ως προς το p δίνει μία καμπύλη που αρχικά αυξάνει και κατόπιν, για κάποια τιμή του p μένει σταθερή. Αυτή η τιμή του p είναι ο αριθμός των μεταβλητών του θα πρέπει να εισαχθούν στο μοντέλο.

Ο Προσαρτημένος πολλαπλός συντελεστης συσχέτισης είναι ο:

$$R_a^2\ (p)=1\ -\ \frac{n(1\ -\ R_p^2)}{n-p}$$

Το σύνολο των μεταβλητών που δίνει την μεγαλύτερη τιμή για τον προσαρτημένο πολλαπλό συντελεστή συσχέτισης είναι εκείνο που επιλέγεται.

5. Το κριτήριο C_p του Mallows

Ο Mallows(1973) πρότεινε μία εναλλακτική μέθοδο για την επιλογή του καλύτερου συνόλου μεταβλητών που θα πρέπει να περιληφθούν στο μοντέλο. Η επιλογή του καλύτερου συνόλου βασίζεται στην ελαχιστοποίηση της ποσότητας:

$$C_p = \frac{SSE_p}{s^2} - (n-2p)$$

όπου SSE_p είναι το άθροισμα τετραγώνων των καταλοίπων από μοντέλο που περιέχει p μεταβλητές από τις k υποψήφιες (συμπεριλαμβανομένης και της σταθεράς p=1,...k+1) και s^2 είναι το μέσο τετραγωνικό σφάλμα από το μοντέλο με k παραμέτρους.

Στην βιβλιογραφία μπορεί να δει κανείς πολλές εργασίες που αναφέρονται στο κριτήριο C_p καθώς και στην σχέση αυτού με τα κριτήρια των Akaike, Schwarz, PRESS που αναφέρθηκαν στην αρχή της ενότητας. Πρακάτω αναφέρουμε μερικές από αυτές:

Οι Kobayashi και Sakata (1990) αποδεικνύουν ότι το κριτήριο C_p είναι μεροληπτικό υπό την έννοια ότι η πιθανότητα για την επιλογή δύο γραμμικών μοντέλων δεν είναι ίση με 0.5 και αν ακόμη αυτά έχουν ίσο τετραγωνικό σφάλμα πρόβλεψης και δίνουν μία διόρθωση του κριτηρίου η οποία είναι χρήσιμη κάτω από

ορισμένες συνθήκες.

Ο Shibata (1981) δίνει επίσης μία μέθοδο για την επιλογή μεταβλητών στην περίπτωση που ο αριθμός των υπό έλεγχο μεταβλητών είναι απεριόριστος ή αυξάνει ανάλογα με το μέγεθος του δείγματος. Στην ίδια εργασία αποδεικνύεται ότι η μέθοδος αυτή είναι ασυμπτωτικά ισοδύναμη με τις μεθόδους του Akaike και την μέθοδο C_p.

Οι εργασίες των Nishii (1984), Yi και Judge (1988), Zhang(1992) δίνουν ασυμπτωτικές ιδιότητες των διαφόρων κριτηρίων. Ειδικότερα στην εργασία του Nishii αποδεικνύεται ότι τα κριτήρια επίλογης είναι ασυμπτωτικά ισοδύναμα κάτω από ειδικές συνθήκες. Στην εργασία του Zhang αποδεικνύεται ότι τα κριτήρια που αναφέρθηκαν είναι ασυμπτωτικά ισοδύναμα και μπορούν να θεωρηθούν ειδικές περιπτώσεις της

$$C(k,\lambda)=RSS(k)+\lambda k \qquad 0\le k\le K$$

όπου K είναι ο συνολικός αριθμός των συντελεστών παλινδρόμησης, k είναι ο αριθμός των συντελεστών που υπεισέρχονται στο μοντέλο. RSS(k) είναι το άθροισμα τετραγώνων των καταλοίπων και το λ δίνει κάποια ποινή στην περίπτωση κακής επιλογής (overfitting).

Στην εργασία του Amemiya (1980) έκτος του κριτηρίου που προτείνεται δίνεται και μία σύγκριση των κριτηρίων Akaike, Theil, Mallows και του Amemiya.

6. Το S_p κριτήριο

Παρόμοιο με το κριτήριο C_p είναι το S_p κριτήριο το οποίο

δόθηκε από τον Hocking (1976) αρχικά και στην τελική του μορφή απο τον Thompson (1978)[a,b]. Το κριτήριο αυτό βασίζεται στην ελαχιστοποίηση της τοσότητας:

$$S_p = \frac{SSE_p}{(n-p)(n-p-2)}$$

όπου SSE_p είναι το άθροισμα τετραγώνων των καταλοίπων από μοντέλο που περιέχει p μεταβλητές από τις k υποψήφιες (συμπεριλαμβανομένης και της σταθεράς p=1,...k+1). Σχετικά με αυτό το κριτήριο ενδιαφέρουσα είναι και η εργασία των Breiman και Freedman (1983).

7. Η τεχνική αμφικλινούς παλινδρόμησης (Ridge Regression)

Η τεχνική της αμφικλινούς παλινδρόμησης η οποία προτάθηκε από τους Hoerl και Kennard (1970)[a,b] δίνει έναν διαφορετικό τρόπο για την επιλογή μεταβλητών. Η μέθοδος στηρίζεται στην εκτίμηση των συντελεστών παλινδρόμησης β ενός γραμμικού μοντέλου $Y=X\beta+\varepsilon$ από την σχέση

$$\hat{\beta}^*(j) = [X'X+jI]X'Y \qquad j\geq 0$$

και στην ιδιότητα ότι υπάρχει πάντα κάποιο $j<0$ τέτοιο ώστε

$$E(L_1^2(j)] < E(L_1^2(0)] \text{ όπου } L_1^2(j)=(\hat{\beta}^*(j)-\beta)'(\hat{\beta}^*(j)-\beta).$$

Η διαδικασία επιλογής κατ'αρχήν διαγράφει εκείνες τις μεταβλητές για τις οποίες υπάρχει ελάχιστη προβλεπτική ικανότητα δηλαδή εκείνες οι οποίες δίνουν την μικρότερη απόλυτη τιμή για την $\hat{\beta}_i^*(j)$. Κατόπιν διαγράφονται οι μεταβλητές εκείνες των οποίων οι συντελεστές παλινδρόμησης τείνουν στο μηδέν όταν

το j αυξάνεται.Σαν μειονέκτημα της μεθόδου θα πρέπει να πούμε το ότι δεν αποσαφηνίζει πότε τελειώνει η διαδικασία. Ενδιαφέροντα σχόλια σχετικά με την διαδικασία της αμφικλινούς παλινδρόμησης και βιβλιογραφία σχετικά με αυτό το θέμα δίνονται στην εργασία των Draper και Nostrand (1979).

Αλλα κριτήρια επιλογής μοντέλων

Εκτός των κριτηρίων που αναπτύχθηκαν παραπάνω έχουν αναφερθεί στην βιβλιογραφία μία σειρά από άλλα κριτήρια τα οποία έχουν εφαρμογές κάτω από ειδικές συνθήκες και καλύπτουν μία μεγάλη περιοχή εφαρμογών από οικονομικά προβλήματα μέχρι προβλήματα μηχανικής. Παρακάτω αναφέρουμε τα πιό σημαντικά από αυτά.

Οι Edwards και Havranek (1987) προτείνουν έναν τρόπο επιλογής μοντέλων όταν έχουμε μία μεγάλη οικογένεια υποψηφίων μοντέλων. Το κριτήριο στηρίζεται στην άποψη ότι αν ένα μοντέλο απορρίπτεται τότε και κάθε υποσύνολό του είναι υποψήφιο για απόρριψη. Αν ένα μοντέλο γίνει δεκτό τότε κάθε υποσύνολό του είναι υποψήφιο για αποδοχή.

Οι Linhart και Zucchini (1986) προτείνουν μία γενική μεθοδολογία για την επιλογή μεταξύ δύο παραμετρικών μοντέλων βασισμένοι σε κριτήρια που μετρούν την διαφορά μεταξύ κατανομών. Η μέθοδος που προτείνουν είναι η εξής:

α. Επιλέγεται μια συναρτηση (discrepancy functional) $\Delta(G,F)$ που μετρά την διαφορά μεταξύ των δύο μοντέλων G και F.

β. Τα δεδομένα θεωρείται ότι προέρχονται από το μοντέλο F. Αυτό το μοντέλο είναι το σωστό. Η συναρτησιακή μορφή του F θεωρείται άγνωστη.

γ. Μέλη της παραμετρικής οικογένειας του μοντέλου G_θ χρησιμοποιούνται για να προσεγγίσουν την F. τότε τα παρατηρούμενα δεδομένα χρσιμοποιούνται για να εκτιμήσουν την παράμετρο θ. Η εκτιμήτρια $\hat{\theta}$ επιλέγεται έτσι ώστε να ελαχιστοποιεί την $\Delta(G_{\hat{\theta}}, F_n)$ όπου F_n είναι η εμπειρική συνάρτηση κατανομής. Μία τέτοια εκτιμήτρια καλείται ελάχιστη εκτιμητρια διαχωρισμού (minimum discrepancy estimator).

δ.Υπάρχουν διαφορετικές απόψεις για την μορφή της συνάρτησης διαφοράς (discrepancy) μεταξύ του παρατηρούμενου μοντέλου F και των μελών μιας παραμετρικής οικογένειας που θα πρέπει να χρησιμοποιηθεί:

(a) διαφορά ως προς την προσέγγιση $\Delta(G_{\theta_0}, F)$ όπου G_{θ_0} είναι το μέλος της παραμετρικής οικογένειας που είναι πιό κοντά στο F.

(b) διαφορά ως προς την εκτίμηση $\Delta(G_{\hat{\theta}}, G_{\theta_0})$ όπου $\hat{\theta}$ είναι η εκτιμήτρια της παραμέτρου θ και

(c) ολική διαφορά $\Delta(G_{\hat{\theta}}, F)$.

Σε μερικές περιπτώσεις η ολική διαφορά είναι το άθροισμα της διαφοράς ως προς την προσέγγιση και της διαφορά ως προς την εκτίμηση.

Στην ιδανική περίπτωση ο ερευνητής θα επιλέξει ένα μοντέλο

(την παραμετρική οικογένεια) στην βάση της ολικής διαφοράς. Αυτό όμως συνεπάγεται το άγνωστο μοντέλο F έτσι το κριτήριο θα είναι μία εκτιμήτρια της αναμενόμενης ολικής διαφοράς. Η επιλογή μοντέλου συνεπάγεται εκτίμηση του κριτηρίου για ποτιτλές παραμετρικές οικογένειες και επιλογή της οικογένειας που ελαχιστοποιεί το κριτήριο.

Για την επιλογή του μέτρου με το οποίο θα κρίνεται η διαφορά μεταξύ των μοντέλων υπάρχουν πολλές επιλογές. Οι πιό γνωστές και οι οποίες αναφέρονται στο βιβλίο των Linhart και Zucchini είναι των Kullback-Leibler (1951), Kolmogorov, Cramer-von Mises, chi-squared, και Gauss (άθροισμα διαφορών τετραγώνων).

Το μέτρο διαφοράς Kullback-Leibler (Kullback-Leibler Information Criterion KLIC 1951) ορίζεται ως εξής:

$$\Delta(G,F)=E_G\left[\log\frac{g(Y)}{f(Y|\theta)}\right]$$

όπου g,f είναι οι συναρτήσεις πυκνότητας των G,F και $E_G()$ δίνει την αναμενόμενη τιμή ως προς την σωστή κατανομή G.

Στο σημείο αυτό θα πρέπει να πούμε ότι το κριτήριο του Akaike προκύπτει ως μία προσέγγιστική αμερόληπτη εκτιμήτρια του KLIC.

Οι Pederson και Johnson (1990) ορίζουν μία ποσότητα λ^2 για την διαφορά μεταξύ του πραγματικού και του θεωρούμενου μοντέλου, βρίσκουν μία εκτιμήτρια της ποσότητας αυτή η οποία

βασίζεται στον Pearson X^2 έλεγχο καλής προσαρμογής, εξετάζουν τις στατιστικές ιδιότητες του και συζητούν την χρησιμότητά του.

Οι Lien και Vuong (1987) προτείνουν μια μέθοδο επιλογής μοντέλου βασιζόμενοι στον έλεγχο της μηδενικής υπόθεσης ότι τα μοντέλα είναι "KLIC" ισοδύναμα έναντι της εναλλακτικής ότι το ένα μοντέλο είναι πλησιέστερο στο αληθινό. Για τον έλεγχο της υπόθεσης αυτής προτείνονται απλές μέθοδοι επιλογής και προσαρτημένες στατιστικές συναρτήσεις πιθανοφάνειας (Likelihood Ratio statistics) οι οποίες είναι ασυμπτωτικά τυποποιημένες κανονικές κάτω από την μηδενική υπόθεση.

Σε μία μεταγενέστερη εργασία του ο Vuong(1989) χρησιμοποιόντας το "KLIC" κριτήριο για να μετρήσει την "διαφορά" ενός μοντέλου με το πραγματικό αναπτύσσει απλές στατιστικές συναρτήσεις πιθανοφάνειας για τον έλεγχο της υπόθεσης ότι τα υποψήφια μοντέλα είναι εξίσου κοντά στα πραγματικά δεδομένα έναντι της εναλλακτικής οτι το ένα μοντέλο είναι πιό κοντά στο πραγματικό. Οι έλεγχοι που αναπτύσονται είναι επιτυχείς στην περίπτωση που τα υποψήφια μοντέλα δεν είναι το ένα υποσύνολο του άλλου, (overlapping,ή nested and misspecified).

Ο Sawa (1978) δίνει μία διαδικασία για την επιλογή μοντέλων καθώς και τις ιδιότητες της. Επίσης δίνει τις διαφορές της διαδικασίας με αυτήν του Akaike. Οι διαδικασίες που προτείνονται βασίζονται στο κριτήριο KLIC.

Οι Swanepoel και Beer (1992) αναπτύσσουν δύο νέα κριτήρια επιλογής μοντέλων που βασίζονται στην τραχύτητα (roughness) των καταλοίπων. Το πρώτο κριτήριο χρησιμοποιεί μία "loss" συνάρτηση και το δεύτερο συμπεριλαμβάνει την εφαρμογή ενός ελέγχου υπόθεσης χρησιμοποιώντας την μεθοδολογία "bootstrap" (Efron 1979). Επίσης γίνεται σύγκριση των κριτηρίων που αναπτύσσονται με τα κριτήρια AIC, BIC, και της μεθόδου που δίνεται από τους Linhart και Zuchinni.

Οι Rao και Yu (1989) προτείνουν ένα κανόνα βάσει του οποίου γίνεται η επιλογή του καλύτερου μοντέλου στο πολλαπλό γραμμικό μοντέλο καθώς ο αριθμός των παρατηρήσεων γίνεται πολύ μεγάλος. Ο κανόνας δίνεται χωρίς να υποτεθεί κανονικότητα της κατανομής των λαθών.

Οι Pollak και Wales (1991) δίνουν μία διαφορετική προσέγγιση στο πρόβλημα της επιλογής μοντέλων βασίζοντας την προσέγγιση στην επικρατούσα διάταξη (dominance ordering) και στο κριτήριο της πιθανοφάνειας της επικρατούσας διάταξης (likelihood dominance criterion LDC) το οποίο γενικεύει την επικρατούσα διάταξη θεωρώντας μία σειρά από σύνθετα παραδεκτά παραμετρικά μεγέθη.

Ο Wu (1989) δίνει επίσης δύο κριτήρια επιλογης μεταβλητών χωρίς να υποθέσει, σε αντίθεση με τα άλλα κριτήρια, ότι τα σφάλματα έχουν την ίδια κατανομή και ακόμη ότι οι μέσοι τους

είναι ίσοι.

Οι εργασίες τέλος των Miller (1984), και Breiman και Spector (1992) δίνουν την δική τους άποψη για επιλογή των μεταβλητών παλινδρόμισης.

Εργασίες επίσης που δίνουν αντιμετωπίσεις του γενικότερου προβλήματος κάτω από ειδικές συνθήκες είναι αυτές των:

Klein και Brown (1984), (όταν υπάρχει ελάχιστη εκ των προτέρων πληροφορία), Poirier και Klepper (1981) (όπου συνδυάζονται κλασσικές και Μπεϋζιανές τεχνικές με την παρουσία panel δεδομένων (panel data sets)).Judge et. al. (1987) (όπου δίδεται ο ρόλος της "επεκταθείσας" (extended) διαδικασίας Stein στην επιλογή μοντέλων). Η εργασία επίσης του Lindley (1968) δίνει μια Μπεϋζιανή προσέγγιση στο πρόβλημα της επιλογής μεταβλητών.

Τέλος ενδιαφέρον παρουσιάζουν και οι εργασίες των Novotny και L.Mcdonald (1986), J.Mcdonald και Richards (1989), Pesaran (1987), Rissanen (1987), Hjorth (1989), Gauch (1988) και του Stone (1981).

Τέλος οι εργασίες του Thompson (1978a,1978b) καθώς και των Stone(1978), Hocking (1976,1983) και Shibata (1985) δίνουν μία κριτική των μεθόδων που υπάρχουν για την επιλογή μεταβλητών.

Στα προηγηθέντα είδαμε διάφορες μεθόδους και κριτήρια για την επιλογή μοντέλων που κυρίως αφορούν την περιγραφική μορφή του μοντέλου.

Οτως αναφέραμε όμως και στην αρχή της ενότητας το πρόβλημα

που θα μας απασχολήσει είναι να βρούμε τρόπους οι οποίοι βοηθούν στην αξιολόγηση της προβλεπτικής ικανότητας του μοντέλου. Φυσικά θα πρέπει να πούμε πως υπάρχει στενή σχέση μεταξύ του προβλήματος της επιλογής μοντέλων όσον αφορά την περιγραφική τους ικανότητα και αυτό της αξιολόγησης της προβλεπτικής τους ικανότητας (Weisberg 1985).

Στην βιβλιογραφία μέθοδοι οι οποίες εκτιμούν την προβλεπτική ικανότητα ενός μοντέλου είναι αυτές των Copas (1983) όπου δίνεται μία μη Μπεϋζιανή προσέγγιση για πρόβλεψη και μοντελοποίηση , και Picard και Cook (1984) όπου δίνεται προσοχή σε μοντέλα που δημιουργούνται μέσω διαδικασιών επιλογής υποσυνόλων τα οποία είναι αρκετά δύσκολο να αξιολογηθούν μέσω κλασσικών μεθόδων. Η εργασία των Geisser και Eddy (1979) δίνει μία Μπεϋζιανή και επανειλημμένη χρησιμοποίηση του δείγματος (sample reuse) προσέγγιση στο πρόβλημα της επιλογής μοντέλων από την άποψη της προβλεπτικότητας χρησιμοποιώντας τα αποτελέσματα της εργασίας του Geisser (1975). Η εργασία του Atkinson (1969) δίνει μία προσέγγιση στο πρόβλημα της αξιολόγησης της προβλεπτικής ικανότητας χρησιμοποιόντας μία μέθοδο (tilde test) η οποία βασίζεται στο άθροισμα της διαφοράς των τετραγώνων των υπό εξέταση μοντέλων με ένα γραμμικό συνδυασμό τους. Οι Whitmore και Xekalaki (1990) ορίζουν τις "προβλεπτικές P-values" (predictive P-values) οι οποίες αποτελούν ένα μέτρο της προβλεπτικής ικανότητας ενός μοντέλου. Η εργασία των Martini και Spezzaferri (1984) δίνει μία διαδικασία η οποία επιλέγει το μοντέλο που δίνει την καλύτερη πρόβλεψη. Την δική του άποψη στο θέμα της προβλεπτικής

ικανότητας, σε σχέση με την μέθοδο της διασταυρωμένης επιβεβαίωσης χωρισμού δεδομένων (cross-validation), δίνει και ο Stone (1974).

Μία συγκεντρωτική βιβλιογραφία στο πρόβλημα της επιλογής μοντέλων είναι οι εργασία του Hamilton (1991).

Στα κεφάλαια που ακολουθούν παρουσιάζονται και μελετώνται διάφορες μέθοδοι οι οποίες δίνουν τρόπους αποτίμησης της προβλεπτικής ικανότητας ενός γραμμικού μοντέλου του οποίου η συναρτησιακή μορφή δεν είναι γνωστη.

Οι μέθοδοι που αναπτύσονται βασιζονται στην ιδέα της βαθμολόγησης της προβλεπτικής ικανότητας ενός η περισσοτέρων μοντέλων και στηριζόνται σε κανόνες αξιολόγησης που έχουν εισαχθεί από τους Xekalaki-Katti (1984). Στηριζομενοι στην ιδεα αυτη μελετάμε τις ιδιότητες των συναρτήσεων που χρησιμοποιούνται ως βαθμολογικοί κανόνες. Έτσι οδηγούμαστε στη μελέτη των κατανομών των συναρτήσεων αυτών που είτε είναι γνωστές (οπότε είναι δυνατή η σχετική στατιστική συμπεραματολογία και η λήψη απόφασης), είτε είναι άγνωστες οπότε τις μελετάμε και κατόπιν προχωράμε στην διαδικασία λήψης απόφασης με βάση τα παρατηρηθέντα δεδομένα.

ΚΕΦΑΛΑΙΟ 2

ΕΛΕΓΧΟΣ ΤΗΣ ΠΡΟΒΛΕΠΤΙΚΗΣ ΙΚΑΝΟΤΗΤΑΣ ΕΝΟΣ ΓΡΑΜΜΙΚΟΥ ΜΟΝΤΕΛΟΥ ΜΕ ΤΗΝ ΧΡΗΣΙΜΟΠΟΙΗΣΗ ΤΗΣ X^2 ΚΑΤΑΝΟΜΗΣ ΚΑΙ ΣΥΓΚΡΙΤΙΚΗ ΑΞΙΟΛΟΓΗΣΗ ΔΥΟ ΠΡΟΒΛΕΠΤΙΚΩΝ ΓΡΑΜΜΙΚΩΝ ΜΟΝΤΕΛΩΝ ΜΕ ΤΗΝ ΧΡΗΣΙΜΟΠΟΙΗΣΗ ΜΙΑΣ ΓΕΝΙΚΕΥΜΕΝΗΣ ΜΟΡΦΗΣ ΤΗΣ F ΚΑΤΑΝΟΜΗΣ

2.1 ΕΙΣΑΓΩΓΗ

Στην εργασία των Xekalaki-Katti (1984) εισάγεται μία διαδικασία αξιολόγησης της προβλεπτικής ικανότητας ενός γραμμικού μοντέλου αλλά και συγκριτικής αξιολόγησης δύο γραμμικών μοντέλων.

Η εργασία αυτή αντιμετώπισε την επίλυση ενός συγκεκριμένου προβλήματος. Το πρόβλημα τέθηκε από το υπουργείο Γεωργίας των ΗΠΑ από το οποίο ζητείτο να αξιολογηθεί ένα μοντέλο ή να γίνει σύγκριση δύο μοντέλων των οποίων όμως η συναρτησιακή μορφή δεν ήταν γνωστή και τα οποία προέβλεπαν την παραγωγή καλαμποκιού σε διάφορες περιοχές (Crop Reporting District (CRD)) της πολιτείας της Ινδιάνα των ΗΠΑ. Για το σκοπό αυτό δόθηκαν οι προβλεπόμενες τιμές της σοδειάς του καλαμποκιού από τα δύο μοντέλα καθώς επίσης και η πραγματική εσοδεία για μία σειρά ετών (1951-1980).

Με βάση τα δεδομένα αυτά και για την επίλυση του συγκεκριμένου προβλήματος δίνουμε τρόπους αποτίμησης της

προβλεπτικής ικανότητας γραμμικών μοντέλων.

Οι μέθοδοι που αναπτύσσονται βασίζονται στην ιδέα της βαθμολόγησης της προβλεπτικής ικανότητας ενός η περισσοτέρων γραμμικών μοντέλων και στηρίζονται σε κανόνες αξιολόγησης που δίνονται στην εργασία των Xekalaki-Katti(1984). Χρησιμοποιόντας τους κανόνες αξιολόγησης οδηγούμαστε σε κατανομές που είτε είναι γνωστές (οπότε είναι δυνατή η σχετική στατιστική συμπεραματολογία και η λήψη απόφασης), είτε είναι άγνωστες οπότε τις μελετούμε και κατόπιν προχωρούμε στην διαδικασία λήψης απόφασης.

Στην ενότητα που ακολουθεί δίνεται ένας τρόπος αξιολόγησης ενός μοντέλου ο οποίος χρησιμοποιεί σαν μέτρο την απόσταση της παρατηρούμενης από την προβλεπόμενη τιμή ενός μοντέλου κάθε χρονική στιγμή. Ακολούθως δίνεται ένας έλεγχος για την συγκριτική αξιολόγηση της προβλεπτικής ικανότητας δύο μοντέλων. Ο έλεγχος αυτός βασίζεται στον ορισμό μιάς νέας κατανομής η οποία είναι μία γενικευμένη μορφή της F κατανομής.

2.2 ΕΛΕΓΧΟΣ ΓΙΑ ΤΗΝ ΑΞΙΟΛΟΓΗΣΗ ΤΗΣ ΠΡΟΒΛΕΠΤΙΚΗΣ ΙΚΑΝΟΤΗΤΑΣ ΕΝΟΣ ΓΡΑΜΜΙΚΟΥ ΜΟΝΤΕΛΟΥ ΜΕ ΤΗΝ ΧΡΗΣΙΜΟΠΟΙΗΣΗ ΤΗΣ X^2 ΚΑΤΑΝΟΜΗΣ

Η προβλεπτική ικανότητα ενός μοντέλου μπορεί να μετρηθεί χρησιμοποιώντας σαν μέτρο την απόσταση της παρατηρούμενης από την προβλεπόμενη τιμή κάθε χρονική στιγμή.

Αναλυτικα:

Αν Y°_{t+1} είναι η παρατηρούμενη τιμή του μοντέλου την χρονική στιγμή t+1 και

$\hat{Y}^{\circ}_{t+1}$ είναι η προβλεπόμενη τιμή του μοντέλου την χρονική στιγμή t+1 τότε, επειδή είναι γνωστό ότι

$$\hat{Y}^{\circ}_{t+1} - Y^{\circ}_{t+1} \sim N(0,\ \sigma^2(1+ X^{\circ}_{t+1}(X_t X_t)^{-1} X'^{\circ}_{t+1})) \qquad (2.2.1)$$

μία στατιστική συνάρτηση που θα μπορούσε να χρησιμοποιηθεί για την αξιολόγηση μοντέλου ,ή μοντέλων, είναι η $|r_{t+1}|$ όπου r_{t+1} δίνεται από την σχέση

$$r_{t+1} = \frac{\hat{Y}^{\circ}_{t+1} - Y^{\circ}_{t+1}}{S_t \sqrt{(1+ X^{\circ}_{t+1}(X_t X_t)^{-1} X'^{\circ}_{t+1})}} \qquad (2.2.2)$$

και $S_t^2 = \dfrac{[\ Y_t - X_t\hat{\beta}_t]'\ [\ Y_t - X_t\hat{\beta}_t]}{[1_t - m]}.$ $t=0,1,2\ldots$

είναι η εκτίμηση του σ^2.

Λόγω της (2.2.1) είναι προφανές ότι η r_{t+1} ακολουθεί την N(0,1).

Έτσι μία βαθμολόγηση βασισμένη στο μέσο όρο των $|r_{t+1}|$ θα μπορούσε να είναι μία μέτρηση της προβλεπτικότητας του μοντέλου. Αυτό θ'αποτελέσει αντικείμενο έρευνας αργότερα.

Πριν προχωρήσουμε όμως στον προβληματισμό αυτό θα μπορούσαμε να θεωρήσουμε σαν βαθμολογικό κανόνα και το r_t^2 για κάθε χρονική στιγμή t. Στην περίπτωση αυτή ο μέσος όρος των r_t^2

$$R_n = \sum_{t=1}^{n} r_t^2 / n \qquad (2.2.3)$$

θα μπορούσε να είναι ο τελικός τρόπος μέτρησης της

προβλεπτικότητας του μοντέλου.

Πιό αναλυτικά:

Εχει δειχθεί (Brown et. al. 1975, Kendall et.al. 1963) ότι αν e_t είναι οι κανονικές μεταβλητές λάθους

με $E(e_t)=0$ και $V(e_t)=\sigma^2 I_t$ τότε οι ποσότητες

$$w_{t+1} = \frac{\hat{Y}^{\circ}_{t+1} - Y^{\circ}_{t+1}}{\sqrt{(1+ X^{\circ}_{t+1}(X_t X_t)^{-1} X^{\circ}_{t+1})}} \qquad t=0,1,2,...$$

είναι ανεξάρτητες κανονικές μεταβλητές με μέσο 0 και διασπορά σ^2. Τότε σύμφωνα με ένα χαρακτηρισμό (Kotlarski 1966) της κανονικής κατανομής από την t οι ποσότητες $r_{t+1}=w_{t+1}/s_t$ $t=0,1,2,...$ αποτελούν μία ακολουθία από ανεξάρτητες t μεταβλητές με l_t-m βαθμούς ελευθερίας $t=0,1,2,...$

Εται με βάση τις προυποθέσεις του μοντέλου που έχουμε θεωρήσει και για μεγάλο l_0. δηλαδή στην περίπτωση πού το δείγμα ειναι μεγάλο, οι μεταβλητές r_t $t=0,1,2,...$ θ'ακολουθούν προσεγγιστικά την standard normal κατανομή και θα είναι μεταξύ τους ανεξάρτητες.

Με βάση την παραπάνω παρατήρηση η στατιστική συνάρτηση

$$nR_n = \sum_{t=1}^{n} r_t^2$$

θ'ακολουθεί την χ^2 κατανομή με n βαθμούς ελευθερίας.

Ετσι τιμές του nR_n στην δεξιά ουρά της κατανομής χ^2_n θ'ατοτελεί ένδειξη μοντέλου το οποίο έχει χαμηλή προβλεπτική ικανότητα. Σχετικοί πίνακες δίνονται στην εργασία Psarakis-Panaretos (1990b). Αναλυτικοί πίνακες με εφαρμογή στα

δεδομένα που υπαρχουν στην εργασία Xekalaki-Katti (1984) και αναφερονται στην παραγωγή καλαμποκιού στις περιοχές CRD10-CRD100 (Crop District 10 έως Crop District 100) της πολιτείας της Ινδιάνα των ΗΠΑ για τα έτη 1951-1980 για το μοντέλο Α και 1956-1980 για το μοντέλο Β δίνονται στο παράρτημα Ι.(Οπως προαναφέρθηκε ήδη η συναρτησιακή μορφή των μοντέλων αυτών δεν είναι γνωστή)

Ο έλεγχος που αναπτύχθηκε παραπάνω έχει μελετηθεί, σε μία τροποποιημένη μορφή, από τους Box και Jenkins (1970) με το όνομα portmanteau lack of fit test όπου υπολογίζεται η ποσότητα:

$$Q = N \sum_{k=1}^{K} r_k^2$$

καθώς επίσης και από τον Spanos (1986) όπου λαμβάνεται υπ'όψη μόνο ένας συγκεκριμμένος αριθμός καταλοίπων.

2.3 ΣΥΓΚΡΙΤΙΚΗ ΑΞΙΟΛΟΓΗΣΗ ΤΗΣ ΠΡΟΒΛΕΠΤΙΚΗΣ ΙΚΑΝΟΤΗΤΑΣ ΔΥΟ ΓΡΑΜΜΙΚΩΝ ΜΟΝΤΕΛΩΝ ΜΕ ΤΗΝ ΧΡΗΣΙΜΟΠΟΙΗΣΗ ΜΙΑΣ ΓΕΝΙΚΕΥΜΕΝΗΣ ΜΟΡΦΗΣ ΤΗΣ F ΚΑΤΑΝΟΜΗΣ

Στην ενότητα αυτή δίνεται μία διαδικασία για την συγκριτική αξιολόγηση δύο μοντέλων.

Έστω δύο μοντέλα Α και Β για τα οποία έχουμε παρατηρήσεις για ένα αριθμό n_1 και n_2 ετών αντίστοιχα και από τα οποία θέλουμε να διαλέξουμε εκείνο που έχει την μεγαλύτερη προβλεπτική ικανότητα. Μία στατιστική συνάρτηση που θα μπορούσε να χρησιμοποιηθεί για την εξέταση της στατιστικής υπόθεσης :

H_o: Τα δύο μοντέλα είναι ισοδύναμα

έναντι της εναλλακτικής

H_1: Το μοντέλο Α έχει μεγαλύτερη προβλεπτική ικανότητα από το Β

θα μπορούσε να είναι ο λόγος των μέσων όρων των αποδόσεων (σκορ) που δίνονται στο κάθε μοντέλο κάθε φορά που αυτό χρησιμοποιείται. Ο λόγος αυτός δίνεται από την σχέση:

$$R_{n_1,n_2} = \frac{R_{n_1}}{R_{n_2}} = \frac{X}{Y} \qquad (2.3.1)$$

όπου τα R_{n_1}, R_{n_2} δίνονται απο την (2.2.3) για $n=n_1$ και $n=n_2$ αντίστοιχα.

Στην περίπτωση μεγάλου δείγματος η στατιστική συνάρτηση R_{n_1,n_2} θ'ακολουθεί την F κατανομή με n_1, n_2 βαθμούς ελευθερίας εφ'όσον ο αριθμητής και ο παρανομαστής του λόγου (2.3.1) είναι ανεξάρτητοι.

Έτσι τιμές του R_{n_1,n_2} στην δεξιά ουρά της F κατανομής θα αποτελούν ένδειξη ότι το μοντέλο Β έχει χαμηλότερη προβλεπτική ικανότητα από το μοντέλο Α.

Στην συγκεκριμένη όμως περίπτωση του προβλήματος που εξετάζουμε η υπόθεση της ανεξαρτησίας δεν φαίνεται να πληρούται. Εχει όμως αποδειχθεί (Kotlarski (1964)) ότι, κάτω από ορισμένες προυποθέσεις, ο λόγος X/Y ακολουθεί την κατανομή F έστω και αν οι X και Y είναι εξαρτημένες.

Συγκεκριμένα αν συμβολίσουμε με ψ το σύνολο των

διμεταβλητών κατανομών F(x,y) των οποίων οι τυχαίες μεταβλητές X και Y παίρνουν μόνο θετικές τιμές και δεν είναι αναγκαστικά ανεξάρτητες και των οποίων ο λόγος ακολουθεί την F κατανομή με παραμέτρους p_1 και p_2 τότε ισχύει το εξής θεώρημα:

Θεώρημα (Kotlarski (1964)):

Μία συνάρτηση κατανομής F(x,y) ανήκει στο σύνολο ψ αν και μόνο αν ο Mellin μετασχηματισμός της F(x,y)

$$h(u,v)=\int_0^{\infty}\int_0^{\infty} x^u y^v \, dF(x,y)$$

είναι της μορφής:

$$h(iu,iv)=[\Gamma(p_1+iu)/\Gamma(p_1)][\Gamma(p_2+iv)/\Gamma(p_2)]e^{\alpha(u,v)+i\beta(u,v)} \qquad (2.3.2)$$

όπου οι συναρτήσεις α(u,v) και β(u,v) θα πρέπει να πληρούν τις παρακάτω συνθηκες:

$$\alpha(u,-u)=0 \quad \beta(u,-u)=0$$

$$\alpha(-u,-v)=\alpha(u,v) \quad \beta(-u,-v)=\beta(u,v)$$

$$\alpha(u,v)\le\log[\Gamma(p_1)/\Gamma(p_1+iu)][\Gamma(p_1)/\Gamma(p_2+iv)]$$

όπου α(u,v),β(u,v) είναι συνεχείς πραγματικές συναρτήσεις στο (u,v).

Ειδικότερα για πραγματικές τιμές των u,v και θέτοντας v=-u στην (2.3.2) θα ισχύει:

$$h(u,-u) = [\Gamma(p_1+u)/\Gamma(p_1)]\ [\Gamma(p_2-u)/\Gamma(p_2)].$$

$$(-p < \mathrm{Re}\ u < p)$$

Στην συνέχεια εξετάζουμε αν τα συμπεράσματα του θεωρήματος του Kotlarski ισχύουν στην περίπτωση που η από κοινού κατανομή των X,Y είναι η διμεταβλητή Γάμμα του Kibble (1941).

Η κατανομή αυτή έχει συνάρτηση πυκνότητας:

$$f(x,y) = \frac{\varrho^{-(k-1)}}{\Gamma(k)\ (1-\varrho)^2}\ (xy)^{\frac{k-1}{2}}\ e^{-\frac{x+y}{1-\varrho^2}}\ .\ I_{k-1}\left(\frac{2\varrho\sqrt{xy}}{1-\varrho^2}\right)$$

$$= \frac{\varrho^{-(k-1)}}{\Gamma(k)\ (1-\varrho)^2}\ e^{-\frac{x+y}{1-\varrho^2}}\ \sum_{i=0}^{\infty}\left(\frac{\varrho}{1-\varrho^2}\right)^{k+2i-1} \times$$

$$\frac{x^{\frac{k-1}{2}+\frac{k-1}{2}+i}\ y^{\frac{k-1}{2}+\frac{k-1}{2}+i}}{\Gamma(i+1)\ \Gamma(i+k)}\ .$$

όπου

$$I_{k-1}\left(\frac{2\varrho\sqrt{xy}}{1-\varrho^2}\right) = \sum_{i=0}^{\infty}\left(\frac{2\varrho\sqrt{xy}}{2(1-\varrho^2)}\right)^{k-1+2i} \Big/ \left(\Gamma(i+1)\ \Gamma((i+k)\right)$$

$$= \sum_{i=0}^{\infty}\left(\frac{\varrho}{1-\varrho^2}\right)^{k+2i-1}\ (xy)^{\frac{k-1}{2}+i} \Big/ \left(\Gamma(i+1)\ \Gamma((i+k)\right).$$

$$\Rightarrow f(x,y) = \frac{1}{\Gamma(k)\ (1-\rho^2)^k}\ e^{-\frac{x+y}{1-\rho^2}} \sum_{i=0}^{\infty} \left(\frac{\rho}{1-\rho^2}\right)^{2i} \times \frac{(xy)^{k+i-1}}{\Gamma(i+1)\ \Gamma(i+k)} \Rightarrow$$

Δηλαδή τελικά, η μορφή της διμεταβλητής γάμμα κατανομής του Kibble δίνεται από την σχέση:

$$f(x,y) = \frac{e^{-\frac{x+y}{1-\rho^2}}}{\Gamma(k)\ (1-\rho^2)^k} \sum_{i=0}^{\infty} \frac{(\rho/(1-\rho^2)^{2i}}{\Gamma(i+1)\ \Gamma(i+k)} \cdot (xy)^{k-1+i} \cdot \quad (2.3.3)$$

Οι λόγοι που μας ενδιαφέρει η συγκεκριμένη αυτή μορφή της διμεταβλητής γάμμα για το πρόβλημά μας είναι οι εξής:

Είναι γνωστό (Kibble (1941) βλέπε και Patil (1984)) ότι ένα μοντέλο που οδηγεί σ'αυτήν την κατανομή είναι το εξής:

Εστω ότι οι τυχαίες μεταβλητές X_i, Y_i $i=1,2,\ldots,n$ ακολουθούν την τυποποιημένη κανονική κατανομή (οπότε η από κοινού κατανομή αυτών των τυχαίων μεταβλητών είναι η διμεταβλητή τυποποιημένη κανονική κατανομή) τότε η από κοινού κατανομή των

$$X = \sum_{i=1}^{n} X_i^2/2 \quad \text{και} \quad Y = \sum_{i=1}^{n} Y_i^2/2$$

είναι η διμεταβλητή γάμμα του Kibble (1941) με συνάρτηση πυκνότητας πιθανότητας την (2.3.3) και παράμετρο $k = n/2$.

Ας θεωρήσουμε την αντιστοιχία

$$X_i \longrightarrow r_i(A), \qquad Y_i \longrightarrow r_i(B)$$

όπου $r_i(M)$ συμβολίζει το σκορ του μοντέλου M, M=A.B όπως αυτό ορίζεται από την (2.2.3). Τότε είναι προφανές ότι οι μεταβλητές

$$R_{n_1}(A) / 2 = \sum_{i=1}^{n_1} r_i^2(A)/ 2 n_1 \qquad \text{και}$$

$$R_{n_2}(B) / 2 = \sum_{i=1}^{n_2} r_i^2(B)/ 2 n_2$$

αντιστοιχούν στις μεταβλητές X και Y αντίστοιχα και κατά συνέπεια έχουν από κοινού κατανομή την διμεταβλητή γάμμα του Kibble (1941). Επομένως η μελέτη του λόγου X/Y είναι κεντρικής σημασίας για την διερεύνηση του προβλήματος.

Στην συγκεκριμένη περίπτωση, προκειμένου να ισχύει το θεώρημα του Kotlarski θα πρέπει ο Mellin μετασχηματισμός της διμεταβλητής γάμμα του Kibble να έχει την μορφή

$$h(u,-u) = [\Gamma(k+u)/\Gamma(k)]\ [\Gamma(k-u)/\Gamma(k)]$$

Ο Mellin μετασχηματισμός της διμεταβλητής γάμμα του Kibble θα δίνεται από την σχέση:

$$h(u,v) = E(X^u Y^v) =$$

$$\frac{(1-\rho^2)^{-k}}{\Gamma(k)} \sum_{i=0}^{\infty} \frac{(\rho/(1-\rho^2))^{2i}}{\Gamma(i+1)\ \Gamma(i+k)} \int_0^{\infty}\int_0^{x} e^{-\frac{x+y}{1-\rho^2}} x^{u+k-1+i}\, y^{v+k+i-1} dxdy$$

Θέτοντας

$$I = \int_0^{\infty}\int_0^{\infty} e^{-\frac{x+y}{1-\rho^2}} x^{u+k-1+i}\, y^{v+k+i-1} dxdy$$

$$\text{Και} \quad \left.\begin{array}{l} x = r\cos^2\theta \\ y = r\sin^2\theta \end{array}\right\} \Rightarrow dxdy = 2r\sin\theta\cos\theta\, drd\theta \qquad 0<r<\infty \ , \quad 0<\theta<\pi/2.$$

Θα έχουμε

$$I = \int_0^{\infty}\int_0^{\pi/2} e^{-\frac{r(\cos^2\theta+\sin^2\theta)}{1-\varrho^2}} (r\cos^2\theta)^{u+k+i-1}(r\sin^2\theta)^{v+k+i-1} \times$$

$$2r\sin\theta\cos\theta drd\theta$$

$$= 2\int_0^{\infty}\int_0^{\pi/2} e^{-\frac{r}{1-\varrho^2}} r^{u+v+2k+2i-2+1} (\cos\theta)^{2(u+k+i-1)+1} \times$$

$$(\sin\theta)^{2(v+k+i-1)+1} drd\theta =$$

$$= 2\int_0^{\infty} e^{-\frac{r}{1-\varrho^2}} r^{u+v+2k+2i-1}\, dr \times$$

$$\int_0^{\pi/2} (\cos\theta)^{2(u+k+i)-1} (\sin\theta)^{2(v+k+i)-1}\, d\theta =$$

$$2\,\frac{\Gamma(u+v+2k+2i)}{(1-\varrho^2)^{-(u+v+2k+2i)}}\,\frac{1}{2}\,B(u+k+i,\ v+k+i)$$

$$\Rightarrow I = \frac{2\,\Gamma(u+v+2i+2k)}{(1-\varrho^2)^{-(u+v+2k+2i)}}\,\frac{1}{2}\,\frac{\Gamma(u+k+i)\,\Gamma(v+k+i)}{\Gamma(u+v+2k+2i)}$$

$$\Rightarrow I = \frac{\Gamma(u+k+i)\ \Gamma(v+k+i)}{(1-\varrho^2)^{-(u+v+2k+2i)}}$$

$$\Rightarrow h(u,v) = \frac{(1-\varrho^2)^{-k}}{\Gamma(k)} \sum_{i=0}^{\infty} \frac{\varrho^{2i}}{(1-\varrho^2)^{2i}} \frac{\Gamma(u+k+i)\ \Gamma(v+k+i)}{\Gamma(i+1)\ \Gamma(k+i)\ (1-\varrho^2)^{-(u+v+2k+2i)}}$$

$$= \frac{(1-\varrho^2)^{-k+u+v+2k}}{\Gamma(k)} \sum_{i=0}^{\infty} \frac{\varrho^{2i}\ \Gamma(u+k+i)\ \Gamma(v+k+i)}{i!\ \Gamma(k+i)}$$

$$= \frac{(1-\varrho^2)^{u+v+k}}{\Gamma(k)} \frac{\Gamma(k+u)\ \Gamma(k+v)}{\Gamma(k)} \times \sum_{i=0}^{\infty} \frac{(k+u)_{(i)}(k+v)_{(i)}}{k_{(i)}} \frac{\varrho^{2i}}{i!}$$

Επομένως

$$h(u,v) = \frac{\Gamma(k+u)}{\Gamma(k)} \frac{\Gamma(k+v)}{\Gamma(k)} \cdot (1-\varrho^2)^{u+v+k}\ {}_2F_1(k+u,\ k+v:\ k;\ \varrho^2)$$

όπου $${}_2F_1(a,b:c;z) = \sum_{r=0}^{\infty} \frac{a_{(r)}\, b_{(r)}}{c_{(r)}} \frac{z^r}{r!}$$

$$\Rightarrow h\ (0,0) = (1-\varrho^2)^k\ {}_2F_1(k,k:k;\varrho^2)$$

$$= (1-\varrho^2)^k\ {}_1F_0(k,:\ :\varrho^2)$$

$$= (1-\varrho^2)^k\ (1-\varrho^2)^{-k} = 1$$

Δηλαδή τελικά:

$$h(u,-u) = \frac{\Gamma(k+u)}{\Gamma(k)} \frac{\Gamma(k-u)}{\Gamma(k)} (1-\rho^2)^k \, {}_2F_1(k+u,k-u;\ k;\ \rho^2).$$

Από την παραπάνω σχέση καταλήγουμε στο συμπέρασμα ότι ο Mellin μετασχηματισμός της διμεταβητής γάμμα του Kibble δεν έχει την απαιτούμενη μορφή έτσι ώστε ο λόγος (2.3.1) να ακολουθεί την F κατανομή.

Στην συνέχεια εξετάζουμε το ενδεχόμενο χρησιμοποίησης μιάς εναλλακτικής μορφής της διμεταβλητής γάμμα, η οποία πιθανόν, να οδηγεί στο επιθυμητό αποτέλεσμα. Δηλαδή ο Mellin μετασχηματισμός της κατανομής αυτής να ικανοποιεί το θεώρημα του Kotlerski (1964).

Ως τέτοια εναλλακτική μορφή θεωρούμε την διμεταβλητή γάμμα τύπου II (Type II bivariate Gamma distribution Jensen (1970 a,b)).

Η συνάρτηση πυκνότητας πιθανότητας αυτής της κατανομής δίνεται από την σχέση:

$$f(x_2,x_2) = \frac{x_1^{\frac{n}{2}-1} \quad x_2^{\frac{n}{2}-1} \quad e^{-(x_1+x_2)/2}}{2^n\,[\Gamma(n/2)]^2} \sum_{j=0}^{\infty} j' \left[\frac{\Gamma(n/2)}{\Gamma(j+n/2)}\right]^2 \times$$

$$Cj. \sum_{r=0}^{j} \frac{\Gamma\left(\frac{n}{2}+j+1\right)\left(\frac{x_1}{2}\right)^r (-1)^r}{\Gamma\left(1+\frac{n}{2}+1\right) r!} \frac{1}{(j-r)!} \sum_{l=0}^{j} \frac{\Gamma\left(\frac{n}{2}+j+1\right)\left(\frac{-x_2}{2}\right)^l}{\Gamma\left(1+\frac{n}{2}\, r^1\right)!\ 1!-(j-1)!}$$

$$= \frac{\sum_{j=0}^{\infty}}{2^n\left[\Gamma(n_2)\right]^2} C_j \frac{\left(j!\right)^2}{\left[\left(\frac{n}{2}\right)_{(j)}\right]^2} \times A\ x_1^{\frac{n}{2}+r-1}\ x_2^{\frac{n}{2}+l-1}\ e^{-(x_1+x_2)/2}$$

όπου $C_j = \sum_{\Sigma j_i = j} \prod_{i=1}^{n} \left\{ \varrho_i^{2j_i} \frac{\Gamma(j_i+\frac{1}{2})}{\sqrt{\pi}\ \Gamma(j_i+1)} \right\}$ και

$$A = \sum_{r=0}^{j} \sum_{l=0}^{j} \frac{\left[\Gamma\left(\frac{n}{2}+j+1\right)\right]^2 (-1/2)^{r+l}}{r!l!\Gamma(r+\frac{n}{2}+1)\ \Gamma(l+\frac{n}{2}+1)(j-l)!(j-r)!}$$

Ο Mellin μετασχηματισμός της $f(x_1,x_2)$ είναι

$$h(u,v) = E(X^uY^v) = \int_0^{\infty} \int_0^{\infty} x^u y^v f(x,y)\ dxdy =$$

$$\frac{1}{2^n\left[\Gamma\left(\frac{n}{2}\right)\right]^2} \sum_{j=0}^{\infty} C_j \frac{(j!)^2}{\left[\left(\frac{n}{2}\right)_{(j)}\right]^2} A \times$$

$$\underbrace{\int_0^{\infty} \int_0^{\infty} x^{\frac{n}{2}+r+u-1}\ j^{\frac{n}{2}+l+v-1}\ e^{-(x+y)/2}\ dxdy}_{\downarrow}$$

$$I(r,l)$$

Κάνοντας τον μετασχηματισμό

$x=p\cos^2\theta$, $y=\varrho\sin^2\theta$ $(J=2p\sin\theta\cos\theta)$, $0 < \varrho < \infty$ $0 < \theta < \pi/2$

Εχουμε

$$I(r,l) =$$

$$2\int_0^{\pi/2}\int_0^{\infty} \varrho^{n+r+l+u+v-2+1}(\cos\theta)^{n+2r+2u-2+1}(\sin\theta)^{n+2l+2v-2+1}e^{-\varrho/2}$$

$$=2\int_0^{\infty}\varrho^{n+r+l+u+v-1}e^{-\varrho/2}d\varrho\int_0^{\pi/2}(\cos\theta)^{2(\frac{n}{2}+r+u)-1}(\sin\theta)^{2(\frac{n}{2}+l+v)-1}d\theta.$$

Ομως $\int_0^{\infty}\varrho^{n+r+l+u+v-1}e^{-\varrho/2}d\varrho = \frac{\Gamma(n+r+l+u+v)}{(1/2)^{n+r+l+u+v}}$ και

$$\int_0^{\pi/2}(\cos\theta)^{2(\frac{n}{2}+r+u)-1}(\sin\theta)^{2(\frac{n}{2}+l+v)-1}d\theta = \frac{1}{2}B(\tfrac{n}{2}+r+u,\ \tfrac{n}{2}+l+v)$$

Ετσι $I(r,l) = 2^{n+r+l+u+v}\,\Gamma(\frac{n}{2}+r+u)\;\;\Gamma(\frac{n}{2}+l+v)$

Επομενως

$$h(u,v) =$$

$$\cdot\ \frac{2^{u+v}}{\left[\Gamma(\frac{n}{2})\right]^2}\sum_{j=0}^{x}C_j\ \frac{(j!)^2}{\left[\left(\frac{n}{2}\right)_{(j)}\right]^2}\times$$

$$\times\sum_{r=0}^{j}\sum_{l=0}^{j}\frac{2^{r+l}\Gamma(\frac{n}{2}+r+u)\Gamma(\frac{n}{2}+l+v)\left[\Gamma(\frac{n}{2}+j+1)\right]^2(-\frac{1}{2})^{r+l}}{r!\,l!\,(j-l)!\,(j-r)!\,\Gamma(\frac{n}{2}+r+1)\Gamma(\frac{n}{2}+l+1)}$$

$$=\frac{2^{u+v}}{\left[\Gamma(\frac{n}{2})\right]^2}\sum_{j=0}^{x}C_j\frac{(j!)^2}{\left[\left(\frac{n}{2}\right)_{(j)}\right]^2}\left[\Gamma(\tfrac{n}{2}+j+1)\right]^2\times$$

$$\times \sum_{r=0}^{j} \frac{(-1)^r \Gamma(\frac{n}{2}+r+u)}{\Gamma(\frac{n}{2}+r+1)\, r!\,(j-r)!} \sum_{l=0}^{j} (-1)^l \frac{\Gamma(\frac{n}{2}+1+v)}{l!\,(j-l)!}$$

$$= 2^{u+v} \sum_{j=0}^{\infty} C_j \left(\frac{n}{2}+j\right)^2 \sum_{r=0}^{j} \binom{j}{r} \frac{\Gamma(\frac{n}{2}+r+u)}{\Gamma(\frac{n}{2}+r+1)} (-1)^r \sum_{l=0}^{j} \binom{j}{l} \frac{\Gamma(\frac{n}{2}+1+u)}{\Gamma(\frac{n}{2}+1+1)} (-1)^l$$

Αλλά

$$C_j = \sum_{j_1+\ldots+j_n=j} \prod_{i=1}^{n} \left\{ \varrho_i^{2j_i} \frac{\Gamma(j_i+\frac{1}{2})}{\sqrt{\pi}\,\Gamma(j_i+1)} \right\}$$

$$= \sum_{\sum j_i = j} \frac{1}{\sqrt{\pi}} \prod_{i=1}^{n} \frac{\varrho_i^{2j_i}}{(\frac{1}{2})_{(1/2+j_i)}} =$$

$$= \sum_{\sum j_i = j} \frac{1}{\sqrt{\pi}} \frac{1}{(\frac{1}{2})_{(1/2)}} \prod_{i=1}^{n} \frac{\varrho_i^{2j_i}}{1(j_i)}$$

$$\left(\text{και επειδή } (\tfrac{1}{2})_{(1/2)} = \frac{\Gamma(1/2+1/2)}{\Gamma(1/2)} = \frac{\Gamma(1)}{\Gamma(1/2)} = \frac{1}{\sqrt{\pi}} \right)$$

$$= \sum_{\sum j_i = j} \frac{\varrho_i^{2j_i}}{j_1!} \cdots \frac{\varrho_n^{2j_n}}{j_n!} = \frac{(\varrho_i^2+\ldots+\varrho_n^2)^j}{j!} .$$

Επίσης $\displaystyle\sum_{r=0}^{j} \binom{j}{r} (-1)^r \frac{\Gamma(\frac{n}{2}+r+u)}{\Gamma(\frac{n}{2}+r+1)}$

$$= j! \sum_{r=0}^{j} \frac{\Gamma(\frac{n}{2}+r+u)}{\Gamma(\frac{n}{2}+r+1)} \frac{1}{1(j-r)} \frac{(-1)^r}{r!}$$

$$= \frac{\Gamma(\frac{n}{2}+u)}{\Gamma(\frac{n}{2}+1)} \sum_{r=0}^{j} \frac{(\frac{n}{2}+u)_{(r)}}{(\frac{n}{2}+1)_{(r)}(j+1)_{(-r)}} \frac{(-1)^r}{r!}$$

$$\frac{\Gamma(\frac{n}{2}+u)}{\Gamma(\frac{n}{2}+1)} \sum_{r=0}^{j} \frac{(\frac{n}{2}+u)_{(r)}\ (-j)_{(r)}}{(\frac{n}{2}+1)_{(r)}} \frac{1}{r!} .$$

Είναι όμως γνωστό ότι:

$${}_mF_n\ (a_1,\ldots,a_m;\ b_1,\ldots,b_n;\ z) = \sum_{r=0}^{x} \frac{(a_1)_{(r)} \ldots (a_m)_{(r)}}{(b_1)_{(r)} \ldots (b_n)_{(r)}} \frac{z^r}{r!}$$

Επομένως

$$h(u,v) = \frac{\Gamma(\frac{n}{2}+u)}{\Gamma(\frac{n}{2}+1)}\ {}_2F_1(-j,\ \tfrac{n}{2}+u;\ \tfrac{n}{2}+1;\ 1)$$

$$\frac{\Gamma(\frac{n}{2}+u)}{\Gamma(\frac{n}{2}+1)} \frac{\Gamma(\frac{n}{2}+1) \ \ldots\ \Gamma(j-u+1)}{\Gamma(\frac{n}{2}+j+1) \ \ldots\ \Gamma(1-u)}$$

$$= \frac{\Gamma(\frac{n}{2}+u)}{\Gamma(\frac{n}{2}+1)} \frac{\Gamma(\frac{n}{2}+1) \ \ldots\ \Gamma(j-u+1)}{\Gamma(\frac{n}{2}+j+1)\ldots\Gamma(1-u)} = \frac{\Gamma(\frac{n}{2}+u)\ \ \Gamma(j-u+1)}{\Gamma(\frac{n}{2}+j+1)\ \Gamma(1-u)}$$

Ετσι

$$h(u,v) = \frac{2^{u+v}}{\left[\Gamma(n/2)\right]^2} \times$$

$$\sum_{j=0}^{\infty} \frac{(\varrho_1^2+\ldots+\varrho_n^2)^j}{j!} \frac{\left[\Gamma(\frac{n}{2}+j+1)\right]^2}{\left[\Gamma(\frac{n}{2})_{(j)}\right]^2} \frac{\Gamma(\frac{n}{2}+u)\ \ \Gamma(j-u+1)}{\Gamma(\frac{n}{2}+j+1)\ \Gamma(1-u)} \frac{\Gamma(\frac{n}{2}+v)\ \ \Gamma(j-v+1)}{\Gamma(\frac{n}{2}+j+1)\Gamma(1-v)}$$

$$= 2^{u+v} \frac{\Gamma(\frac{n}{2}+v)}{\Gamma(\frac{n}{2})} \cdot \frac{\Gamma(\frac{n}{2}+u)}{\Gamma(\frac{n}{2})} \sum_{j=0}^{\infty} \frac{(\varrho_1^2+\ldots+\varrho_n^2)^j}{j!} \frac{(1-u)_{(j)}\ \ (1-v)_{(j)}}{(\frac{n}{2})_{(j)}\ \ (\frac{n}{2})_{(j)}}$$

Δηλαδή τελικά

$$h(u,v) = 2^{u+v} \frac{\Gamma(\frac{n}{2}+v)}{\Gamma(\frac{n}{2})} \frac{\Gamma(\frac{n}{2}+u)}{\Gamma(\frac{n}{2})} {}_2F_2\ (1-u, 1-v;\ \tfrac{n}{2},\ \tfrac{n}{2}\ ;\varrho_1^2+\ldots+\varrho_n^2)$$

Το τελικο αποτέλεσμα δειχνει ότι και για την διμεταβλητή γάμμα κατανομής τύπου II ο Mellin μετασχηματισμός δεν ικανοποιεί την απαίτηση του θεωρήματος του Kotlarski.

Οι προηγηθείσες προσπάθειες μας οδηγούν στο συμπέρασμα ότι δεν υπάρχει, ενδεχομένως, κάποια μορφή διμεταβλητής γάμμα κατανομής για την οποία να ισχύει το θεώρημα του Kotlarski, ώστε ο λόγος X/Y να ακολουθεί την κατανομή F , με X και Y εξαρτημένες.

Η διαπίστωση αυτή μας αναγκάζει να στραφούμε προς τον αναλυτικό προσδιορισμό της κατανομής του λόγου X/Y.

ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΤΗΣ ΚΑΤΑΝΟΜΗΣ ΤΟΥ ΛΟΓΟΥ X/Y ΟΤΑΝ X ΚΑΙ Y ΑΚΟΛΟΥΘΟΥΝ ΤΗΝ ΔΙΜΕΤΑΒΛΗΤΗ ΓΑΜΜΑ ΤΟΥ ΚΙΒΒΛΕ (1941)

Οπως είναι γνωστό αν X,Y είναι εξαρτημένες τυχαίες μεταβλητές και αν $Z = \frac{X}{Y}$ τότε η συνάρτηση κατανομής της Z θα δίνεται από την:

$$F_Z(z) = P\left(\frac{X}{Y} \le z\right) = \int_0^{\infty} \underbrace{P(X \le zY \mid Y=y)}_{\downarrow\ F_{X/Y=y}(zy)} f_Y(y)\,dy \ .$$

και η συνάρτηση πυκνότητας από την

$$f_z(z) = \int_0^{\infty} f_{X/Y=y}(zy)\cdot y \cdot f_Y(y)\,dy \ .=$$

$$\int_0^{\infty} \frac{f_{X,Y}(zy,y)}{f_Y(y)} \cdot y \cdot f_Y(y)\,dy$$

$$\Rightarrow \boxed{f_z(z) = \int_0^{\infty} y\, f_{X,Y}(zy,y)\,dy}$$

Εκείνο που μας ενδιαφέρει τώρα, όπως προαναφέραμε, είναι να βρούμε τη συνάρτηση πυκνότητας πιθανότητας του πηλίκου X/Y έτσι ώστε κατόπιν να προχωρήσουμε στον έλεγχο για την σύγκριση των δύο μοντέλων.

Θα έχουμε:

$$f_{X/Y}(z) = \int_0^{\infty} y\, f_{X,Y}(zy,y)dy = \frac{(1-\varrho^2)^{-k}}{\Gamma(k)} \sum_{i=0}^{\infty} \frac{(\varrho/(1-\varrho^2)^{2i}}{\Gamma(i+1)\Gamma(i+k)} \times$$

$$\times \int_0^{\infty} e^{-\frac{zy+y}{1-\varrho^2}} z^{k-1+i}\, y^{2(k-1+i)+i} dy =$$

$$= \frac{(1-\varrho^2)^{-k}\, z^{k-1}}{\Gamma(k)} \sum_{i=0}^{\infty} \frac{[\varrho/(1-\varrho^2)]^{2i}}{\Gamma(i+k)} \frac{z^i}{i!} \int_0^{\infty} e^{-\frac{z+1}{1-\varrho^2}\cdot y}\, y^{2(k+i)-1} dy$$

Ομως $$\int_0^{\infty} e^{-\frac{z+1}{1-\varrho^2}\cdot y}\, y^{2(k+i)-1} dy = \left(\frac{z+1}{1-\varrho^2}\right)^{-2(k+i)} \Gamma(2k+2i)$$

$$= \frac{(1-\varrho^2)^{-k}}{\Gamma(k)}\, z^{k-1} \sum_{i=0}^{\infty} \frac{\varrho^{2i}(1-\varrho^2)^{-2i}}{i!\Gamma(i+k)} \frac{\Gamma(2k+2i)\ (z+1)^{-2k-2i}}{(1-\varrho^2)^{-2k-2i}}\, z^i$$

$$= \frac{(1-\varrho^2)^{k}}{\Gamma(k)}\, z^{k-1}\ (z+1)^{-2k} \sum_{i=0}^{\infty} \frac{\Gamma(2k+2i)}{\Gamma(i+k)} \frac{[\varrho^2(z+1)^{-2}]^i}{i!}\, z^i$$

Αλλά $$\frac{\Gamma(2k+2i)}{\Gamma(k)\Gamma(i+k)} = \frac{\Gamma(2k+2i)\ \Gamma(2k)\ \Gamma(k)}{\Gamma(k)\ \Gamma(k+i)\ \Gamma(2k)\ \Gamma(k)}$$

$$= \frac{(2k)_{(2i)}}{k_{(i)}} \cdot \left(B(k,k)\right)^{-1} = \left(B(k,k)\right)^{-1} \frac{2^{2i}\left(\frac{2k}{2}\right)_{(i)}\left(\frac{2k+1}{2}\right)_{(i)}}{k_{(i)}}$$

Εδώ έγινε χρήση των σχέσεων:

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\ \Gamma(\beta)}{\Gamma(\alpha+\beta)}\ , \qquad \gamma\iota\alpha \quad \alpha=\beta=k$$

και

$$\alpha_{(mn)} = n^{nm} \left(\frac{\alpha}{n}\right)_{(m)} \left(\frac{\alpha+1}{n}\right)_{(m)} \cdots \left(\frac{\alpha+n-1}{n}\right)_{(m)}$$

$$\gamma\iota\alpha \quad \alpha=2k\ , \quad m=i\ , \quad n=2\ .$$

Επομένως: $$\frac{\Gamma(2k+2i)}{\Gamma(k)\Gamma(i+k)} = \frac{2^{2i}\left(\frac{2k+1}{2}\right)_{(i)}}{B(k,k)} \qquad (2.3.4)$$

$$\Longrightarrow f_{X/Y}(z) = (1-p^2)^k \ \frac{z^{k-1}(1+z)^{-2k}}{B(k,k)} \sum_{i=0}^{\alpha} \left(\frac{2k+1}{2}\right)_{(i)} \frac{\left[4\varrho^2(z+1)^{-2}\right]^i z^i}{i!}$$

$$= (1-p^2)^k \ \frac{z^{k-1}(1+z)^{-2k}}{B(k,k)} \ {}_1F_0\left(\frac{2k+1}{2},;; \frac{4\varrho^2 z}{(1+z)^2}\right)$$

$$= (1-\varrho^2)^k \ \frac{z^{k-1}(1+z)^{-2k}}{B(k,k)} \left(1 - 4\frac{\varrho^2 z}{(z+1)^2}\right)^{-\frac{2k+1}{2}}$$

$$\Longrightarrow \boxed{f_{X/Y}(z) = \frac{(1-\varrho^2)^k}{B(k,k)}\ z^{k-1}\ (1+z)^{-2k} \left(1 - \left(\frac{2\varrho}{z+1}\right)^2 z\right)^{-\frac{2k+1}{2}}} \qquad (2.3.5)$$

Η σχέση (2.3.5) προσδιορίζει την συνάρτηση κατανομής του λόγου X/Y όταν η από κοινού κατανομή των (X,Y) είναι η διμεταβλητη γαμμα κατανομη του Kibble

Στη περίπτωση τώρα που οι X, Y είναι ανεξάρτητες (οπότε ρ=0) η συνάρτηση πυκνότητας του πηλίκου X/Y θα δίνεται από την σχέση

$$f_{X/Y}(z) = \frac{1}{B(k,k)}\, z^{k-1}\,(1+z)^{-2k}$$

η οποία είναι η F με 2k και 2k βαθμούς ελευθερίας μια και όπως είναι γνωστό

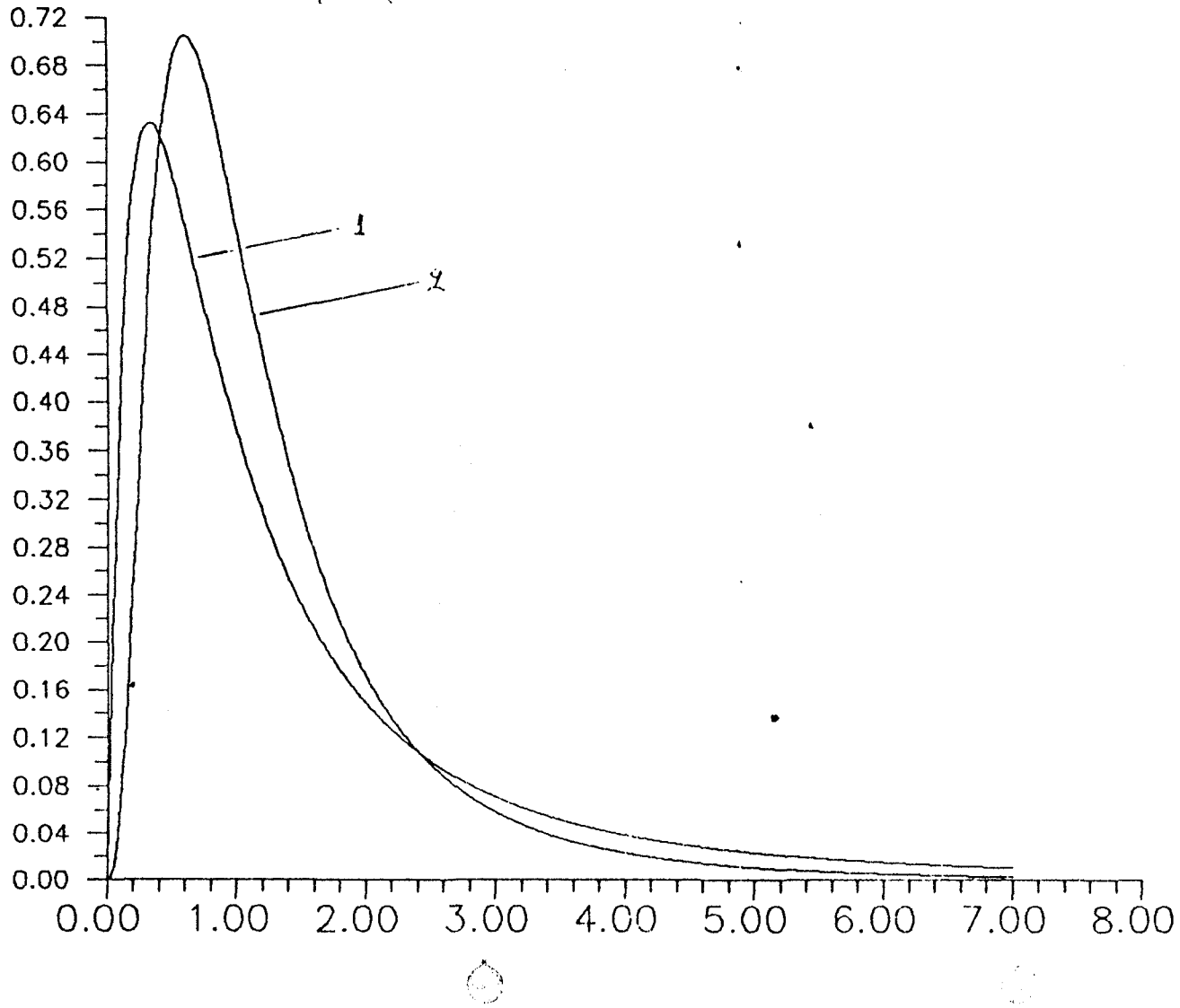
$$F_{\nu_1,\nu_2} \sim \frac{\nu_1^{\nu_1/2}\,\nu_2^{\nu_2/2}}{B\left(\frac{\nu_1}{2},\frac{\nu_2}{2}\right)}\, z^{\frac{\nu_1}{2}-1}\,(\nu_2+\nu_1 z)^{-\frac{\nu_1+\nu_2}{2}}$$

Η κατανομη (2.3.5) δεν έχει μελετηθεί μεχρι τώρα στην βιβλιογραφία και την ονομάζουμε γενικευμένη F (Generalised F distribution)

Στα σχήματα 1.2,και 3 δίνονται οι γραφικές παραστάσεις της (2.3.5) για διάφορες τιμές του k και του ρ.

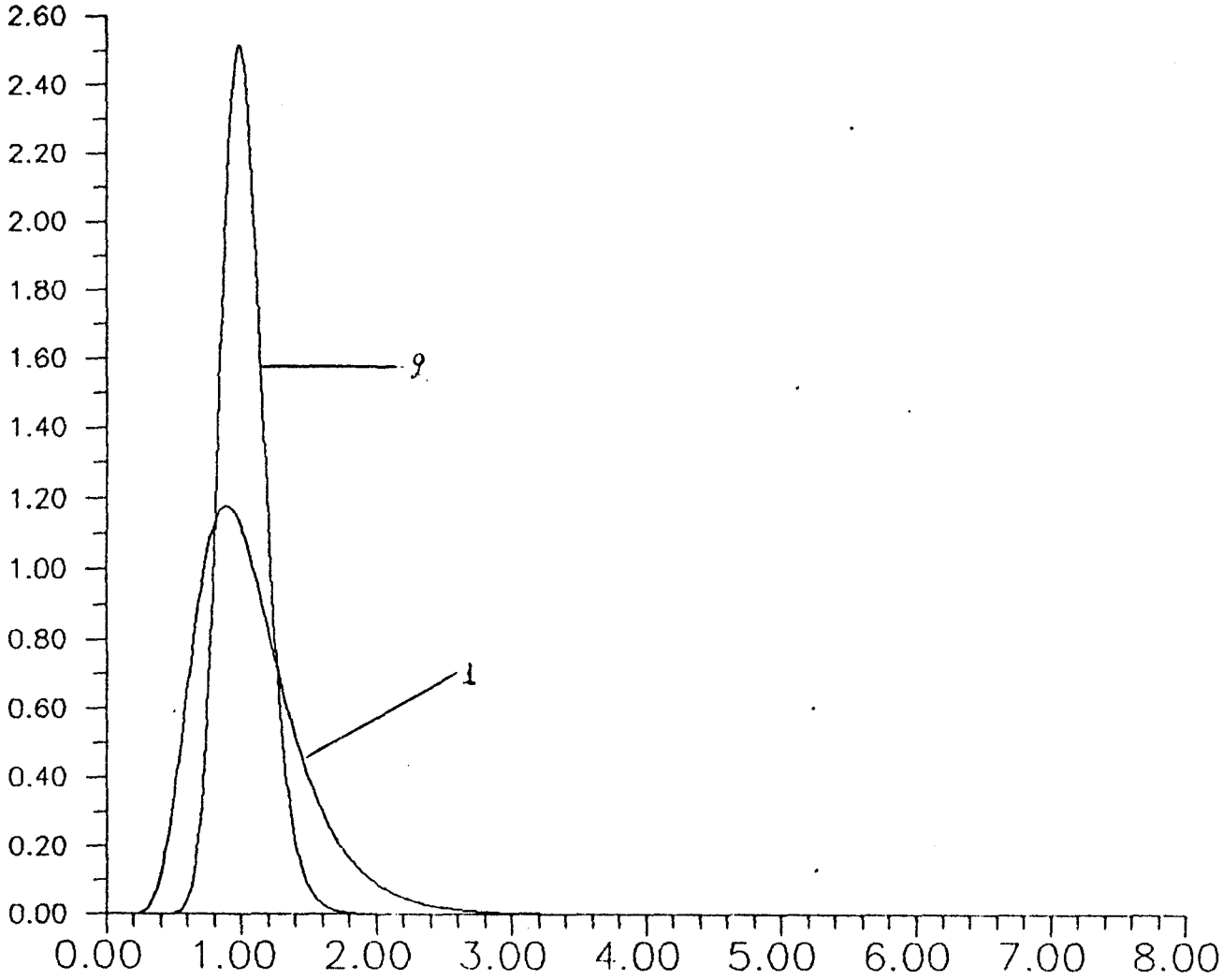
Σχήμα 1

Συνάρτηση πυκνότητας πιθανότητας
γενικευμένης F κατανομής
1 = (k = 2 ϱ = 0.2) 2 = (k = 2 ϱ = 0.9)



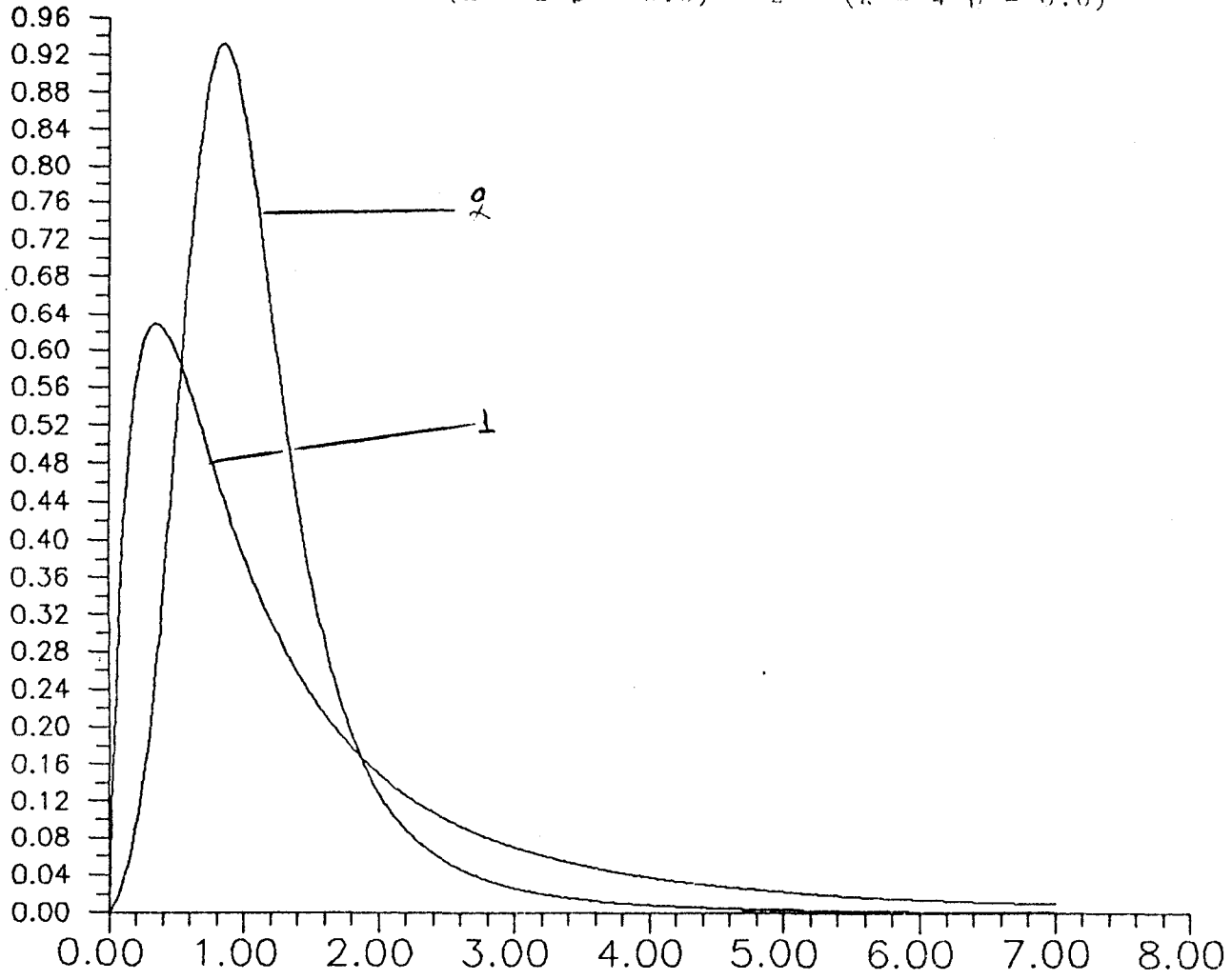
Σχήμα 2

Συνάρτηση πυκνότητας πιθανότητας
γενικευμένης F κατανομής
1 = (k = 15 ρ = 0.2) 2 = (k = 15 ρ = 0.9)



Σχήμα 3

Συνάρτηση πυκνότητας πιθανότητας
γενικευμένης F κατανομής
1 = (k = 2 ρ = 0.0) 2 = (k = 4 ρ = 0.0)



Θα αποδείξουμε στην συνέχεια ότι η (2.3.5) είναι μία καλά ορισμένη κατανομή. Οτι δηλαδή

$$\int_0^{\infty} f_{X/Y}(z)dz = 1$$

Είναι

$$\int_0^{\infty} f_{X/Y}(z)dz = \frac{(1-\varrho^2)^k}{B(k,k)} \underbrace{\int_0^{\infty} z^{k-1}(1+z)^{-2k} \left(1 - \frac{4\varrho^2 z}{(z+1)^2}\right)^{-(2k+1)/2} dz}_{I}$$

Αλλά

$$I = \sum_{r=0}^{\infty} \left(\frac{2k+1}{2}\right)_{(r)} \frac{(4\varrho^2)^r}{r!} \int_0^{\infty} z^{r+k-1} (1+z)^{2r-2k} dz$$

και

$$I_r = \int_0^{\infty} z^{r+k-1} (1+z)^{2r-2k} dz = \frac{\Gamma(k+r)\Gamma(k+r)}{\Gamma(2k+2r)}$$

Ετσι

$$I = \sum_{r=0}^{\infty} \left(\frac{2k+1}{2}\right)_{(r)} \frac{(4\varrho^2)^r}{r!} \frac{\Gamma(k+r)\Gamma(k+r)}{\Gamma(2k+2r)}$$

και λόγω της (2.3.4) θα έχουμε

$$I = \sum_{r=0}^{\infty} \frac{B(k,k)\ \Gamma(2k+2r)\ \ 2^{-2r}\ \ 4^{r}\rho^{2r}\ \Gamma(k+r)\ \Gamma(k+r)}{\Gamma(k)\ \Gamma(k+r)\qquad\quad r!\qquad\qquad\quad \Gamma(2k+2r)}$$

Επομένως

$$I = B(k,k) \sum_{r=0}^{\infty} k_{(r)} \frac{\rho^{2r}}{r!} = B(k,k)\ (1-\rho^{2})^{-k}$$

και τελικά

$$\int_{0}^{\infty} f_{X/Y}(z)dz = 1$$

Εκατοστιαία σημεία της κατανομής για διάφορες τιμές της παραμέτρου k (k=1 30(1)) και του συντελεστή συσχέτισης ρ (ρ=0.0 0.9(0.1)) δίνονται στο παράρτημα II.

Εχοντας κατασκευάσει τους πίνακες της γενικευμένης F κατανομής μπορούμε να προχωρήσουμε σε στατιστική συμπερασματολογία για την επίλυση διαφόρων πρακτικών προβλημάτων.

ΚΕΦΑΛΑΙΟ 3

ΜΕΛΕΤΗ ΙΔΙΟΤΗΤΩΝ ΤΩΝ ΚΑΤΑΝΟΜΩΝ FOLDED NORMAL, FOLDED STANDARD NORMAL ΚΑΙ ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΤΟΥ ΑΘΡΟΙΣΜΑΤΟΣ FOLDED STANDARD NORMAL ΤΥΧΑΙΩΝ ΜΕΤΑΒΛΗΤΩΝ

3.1 ΕΙΣΑΓΩΓΗ

Στο κεφάλαιο αυτό προχωρούμε στην εύρεση τρόπων αξιολόγησης μοντέλων χρησιμοποιόντας τον βαθμολογικό κανόνα $|r_{t+1}|$ που ορίσαμε στο κεφάλαιο 2 ο οποίος δίνεται από την:

$$|r_{t+1}| = \frac{|\hat{Y}^{\circ}_{t+1} - Y^{\circ}_{t+1}|}{S_t \sqrt{(1+ X^{\circ}_{t+1}(X_t X_t)^{-1} X^{\circ}_{t+1})}} \qquad (3.1.1)$$

Διακρίνουμε δύο περιπτώσεις. Την περίπτωση στην οποία έχουμε μεγάλο αριθμό παρατηρήσεων και την περίπτωση που έχουμε μικρό αριθμό παρατηρήσεων.Και στις δύο περιπτώσεις γίνεται προσέγγιση της κατανομής που ακολουθεί ο βαθμολογικός κανόνας (3.1.1) και κατόπιν γίνεται προσπάθεια μελέτης της κατανομής του αθροίσματος τυχαίων μεταβλητών που ακολουθούν την κατανομή αυτή πράγμα που θα βοηθήσει στην λήψη απόφασης για την αξιολόγηση της προβλεπτικής ικανότητας του υπό μελέτη μοντέλου.

3.2 ΠΕΡΙΠΤΩΣΗ ΜΕΓΑΛΟΥ ΑΡΙΘΜΟΥ ΠΑΡΑΤΗΡΗΣΕΩΝ

Σ'αυτή την περίπτωση η κατανομή της (3.1.1) είναι, κατά προσέγγιση, η απόλυτη τιμή της standard normal. Η κατανομή αυτή είναι ειδική περίπτωση της folded normal κατανομής η οποία έχει μελετηθεί από τους Leone, Nelson και Nottingham (1961) και της οποίας η συνάρτηση πυκνότητας πιθανότητας δίνεται από την σχέση:

$$f_X(x) = e^{-\frac{(x-\mu)^2}{2\sigma^2}} + e^{-\frac{(x+\mu)^2}{2\sigma^2}} \qquad x>0$$

Την κατανομή αυτή ονομάζουμε folded standard normal . Προκύπτει από την folded normal θέτοντας $\mu=0$ και $\sigma^2=1$. Η συνάρτηση πυκνότητας αυτής της κατανομής θα δίνεται από την παρακάτω σχέση:

$$F_{|z|}(z) = \sqrt{2/\pi}\ e^{-z^2/2} \qquad z > 0$$

με μέση τιμή $\mu_{SN_f} = \sqrt{2/\pi}$ και

διασπορά $\sigma^2_{SN_f} = 1 - 2/\pi$

(Οι ιδιότητες και οι εφαρμογές της folded normal κατανομής αναλύονται στις εργασίες των Elandt (1961), Johnson (1962), Gilbert και Mosteller (1966), Risvi (1971), Sundberg (1974), Nelson (1980) και Sinha (1983))

Έχοντας καθορίσει την κατανομή της (3.1.1) ως folded standard normal θα μελετήσουμε τις ιδιότητες της.

Αυτό θα μας βοηθήσει στην συνέχεια να προσδιορίσουμε την κατανομή του αθροίσματος ανεξάρτητων τυχαίων μεταβλητών που κατανέμονται όπως η (3.1.1). Αυτό θα μας δώσει την δυνατότητα να προχωρήσουμε σε στατιστική συμπερασματολογία στο πρόβλημα που μας απασχολεί.

3.2.1 ΡΟΠΟΓΕΝΝΗΤΡΙΑ ΤΗΣ FOLDED NORMAL ΚΑΤΑΝΟΜΗΣ

Οπως είναι γνωστό η συνάρτηση πυκνότητας της folded normal κατανομής δίνεται από την σχέση

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}\int_0^{\infty}\left[e^{-\frac{(x-u)^2}{2\sigma^2}}+e^{-\frac{(x+u)^2}{2\sigma^2}}\right]dx$$

Η ροπογεννήτρια συνάρτηση δίδεται από την σχέση:

$$M_X(t)=\frac{1}{\sqrt{2\pi}\,\sigma}\int_0^{\infty}e^{tx}\left[e^{-\frac{(x-u)^2}{2\sigma^2}}+e^{-\frac{(x+u)^2}{2\sigma^2}}\right]dx=$$

$$=\frac{1}{\sqrt{2\pi}\,\sigma}\int_0^{\infty}e^{tx}e^{-\frac{(x-u)^2}{2\sigma^2}}dx+\frac{1}{\sqrt{2\pi}\,\sigma}\int_0^{\infty}e^{tx}e^{-\frac{(x+u)^2}{2\sigma^2}}dx$$

θέτουμε

$$A=\frac{1}{\sqrt{2\pi}\,\sigma}\int_0^{\infty}e^{tx}e^{-\frac{(x-u)^2}{2\sigma^2}}dx \quad \text{και}$$

$$B = \frac{1}{\sqrt{2\pi}\,\sigma} \int_0^{\infty} e^{tx}\, e^{-\frac{(x+\mu)^2}{2\sigma^2}}\, dx$$

Είναι

$$A = \frac{1}{\sqrt{2\pi}\,\sigma} \int_0^{\infty} e^{-\frac{(x-\mu)^2+2tx\sigma^2}{2\sigma^2}}\, dx =$$

$$\frac{1}{\sqrt{2\pi}\,\sigma} \int_0^{\infty} e^{-\frac{[x-(\mu+t\sigma^2)]^2-(2t\mu+t^2\sigma^2)\sigma^2}{2\sigma^2}}\, dx =$$

$$e^{\mu t+\frac{1}{2}\sigma^2 t^2} \int_0^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{[x-(\mu+t\sigma^2)]^2}{2\sigma^2}}\, dx \qquad (3.2.1.1)$$

Θέτοντας στην (3.2.1.Γ) $u=\frac{x-(\mu+t\sigma^2)}{\sigma}$ θα έχουμε ότι το δεξί μέλος της (3.2.1.1) είναι ίσο με την παράσταση

$$e^{\mu t+\frac{1}{2}\sigma^2 t^2} \int_{-\frac{\mu}{\sigma}-t\sigma}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{u^2}{2}}\, du\ .$$

Επομένως

$$A = e^{\mu t+\frac{1}{2}\sigma^2 t^2} \left[1-\Phi\left(-\frac{\mu}{\sigma}-t\sigma\right)\right] \qquad (3.2.1.2)$$

Με παρόμοιο τρόπο προσδιορίζουμε το ολοκλήρωμα B

$$B = e^{-\mu t+\frac{1}{2}\sigma^2 t^2}\left[1-\Phi\left(+\frac{\mu}{\sigma}-t\sigma\right)\right] \qquad (3.2.1.3)$$

Η ροπογεννήτρια συνάρτηση θα δίνεται τελικά απο την σχέση

$$M_X(t)= e^{\mu t+\frac{1}{2}\sigma^2 t^2}\left[1-\Phi\left(-\frac{\mu}{\sigma}-t\sigma\right)\right]+$$

$$+e^{-\mu t+\frac{1}{2}\sigma^2 t^2}\left[1-\Phi\left(+\frac{\mu}{\sigma}-t\sigma\right)\right] \qquad (3.2.1.4)$$

Εύκολα μπορεί κανείς να επαληθεύσει το παραπάνω αποτελέσμα βρίσκοντας την μέση τιμή της folded normal μέσω της ροπογεννήτριας.

Στην (3.2.1.4) θέτοντας μ=0 και σ=1 βρίσκουμε εύκολα την ροπογεννήτρια της folded standard normal η οποία θα δίνεται από την σχέση:

$$M_{SF}(t) = 2\, e^{\frac{1}{2}\sigma^2 t^2}\left[1-\Phi\left(-t\right)\right] \qquad (3.2.1.5)$$

3.2.2 ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΤΗΣ ΚΑΤΑΝΟΜΗΣ ΤΟΥ ΑΘΡΟΙΣΜΑΤΟΣ ΔΥΟ ΑΝΕΞΑΡΤΗΤΩΝ ΤΥΧΑΙΩΝ ΜΕΤΑΒΛΗΤΩΝ ΠΟΥ ΑΚΟΛΟΥΘΟΥΝ ΤΗΝ FOLDED STANDARD NORMAL ΚΑΤΑΝΟΜΗ

Εστω X,Y δυο ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν την folded standard normal κατανομή. Θέλουμε να βρούμε την κατανομή

της Z=X+Y. Χρησιμοποιούμε την μέθοδο της συνέλιξης.

Οπως είναι γνωστό

$$f_Z = f_X * f_Y \longrightarrow f_Z = \int_0^Z f_Y(z-x)\, f_X(x)\, dx =$$

$$= \int_0^Z \frac{2}{\sqrt{2\pi}} e^{-\frac{(z-x)^2}{2}} \frac{2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \qquad (3.2.2.1)$$

μετά από πράξεις η (3.2.2.1) γίνεται

$$\frac{4}{2\pi} e^{-\frac{z^2}{4}} \int_0^Z e^{-(x-z/2)^2} dx \qquad (3.2.2.2)$$

εφαρμόζοντας τώρα τον μετασχηματισμό u= x - z/2 η (3.2.2.2) θα γίνει

$$\frac{4}{2\pi} e^{-\frac{z^2}{4}} \int_{-\frac{z}{2}}^{\frac{z}{2}} e^{-u^2} du = 2 \frac{4}{2\pi} e^{-\frac{z^2}{4}} \int_0^{\frac{z}{2}} e^{-u^2} du \qquad (3.2.2.3)$$

και σύμφωνα με το βιβλίο "Tables of integrals series and products " by I.S. Gradshteyn and I.M. Ryzhik (1980) ολοκλήρωμα 3.221 η (3.2.2.3) θα γίνει:

$$\frac{2}{\sqrt{\pi}} e^{-\frac{z^2}{4}} \Phi^*\left(\frac{z}{2}\right) \qquad (3.2.2.4)$$

όπου $\Phi^*(\frac{z}{2}) = \int_0^{\frac{z}{2}} e^{-t^2} dt$.

Για να επαληθεύσουμε ότι η (3.2.2.5) είναι πράγματι συνάρτηση πυκνότητας πιθανότητας θα πρέπει να αποδειχθεί ότι:

$$\int_0^{\infty} \frac{2}{\sqrt{\pi}} e^{-\frac{z^2}{4}} \Phi^*(\frac{z}{2}) = 1. \qquad (3.2.2.6)$$

Είναι

$$\int_0^{\infty} \frac{2}{\sqrt{\pi}} e^{-\frac{z^2}{4}} \left[\Phi^*(\frac{z}{2}) + 1 - \Phi^*(\frac{z}{2}) \right] = \int_0^{\infty} \frac{2}{\sqrt{\pi}} e^{-\frac{z^2}{4}}$$

άρα

$$\int_0^{\infty} \frac{2}{\sqrt{\pi}} e^{-\frac{z^2}{4}} \left[\Phi^*(\frac{z}{2}) \right] + \int_0^{\infty} \frac{2}{\sqrt{\pi}} e^{-\frac{z^2}{4}} \left[1 - \Phi^*(\frac{z}{2}) \right] =$$

$$\int_0^{\infty} \frac{2}{\sqrt{\pi}} e^{-\frac{z^2}{4}}$$

όμως

$$\int_0^{\infty} \frac{2}{\sqrt{\pi}} e^{-\frac{z^2}{4}} = 2 \qquad (3.2.2.7)$$ (Gradshteyn & Ryzhik (1980) ολοκλήρωμα 3.221-3 για q=1/2)

Προσδιορισμός του ολοκληρώματος:

$$\int_0^{\infty} \frac{2}{\sqrt{\pi}} e^{-\frac{z^2}{4}} \left[1 - \Phi^*\left(\frac{z}{2}\right)\right] \quad (3.2.2.8).$$

Το ολοκλήρωμα αυτό σύμφωνα με το βιβλίο των Gradshteyn and Ryzhik (ολοκλ. 6.286) για ν=1, $\mu^2=-1/4$, β=1/2 είναι ίσο με 1.

Λόγω των (3.2.2.7) , (3.2.2.8) η (3.2.2.6) θα είναι όντως ίση με 1. ο.ε.δ.

Αποδείξαμε ότι αν X,Y είναι δύο τυχαίες μεταβλητές που ακολουθούν την folded standard normal κατανομή τότε η συνάρτηση πυκνότητας της Z=X+Y θα δίνεται απο τον τυπο:

$$f_Z(z) = \frac{2}{\sqrt{\pi}} e^{-\frac{z^2}{4}} \Phi^*\left(\frac{z}{2}\right) \quad (3.2.2.9)$$

$$\text{όπου} \quad \Phi^*\left(\frac{z}{2}\right) = \frac{2}{\sqrt{\pi}} \int_0^{\frac{z}{2}} e^{-t^2} dt \quad (3.2.2.10)$$

Ετσί η (3.2.2.9) λόγω της (3.2.2.10) γίνεται:

$$f_Z(z) = \frac{4}{\sqrt{\pi}} e^{-\frac{z^2}{4}} \left[\Phi\left(\frac{z\sqrt{2}}{2} - \Phi(0)\right)\right] \quad (3.2.2.11)$$

3.2.3 ΕΥΡΕΣΗ ΤΗΣ ΚΑΤΑΝΟΜΗΣ ΤΗΣ ΔΙΑΦΟΡΑΣ ΔΥΟ ΤΥΧΑΙΩΝ ΑΝΕΞΑΡΤΗΤΩΝ ΜΕΤΑΒΛΗΤΩΝ ΠΟΥ ΑΚΟΛΟΥΘΟΥΝ ΤΗΝ FOLDED STANDARD NORMAL ΚΑΤΑΝΟΜΗ

Εστω X,Y δύο ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν την folded standard normal κατανομή. θέλουμε να βρούμε την κατανομή της Z=X-Y. Ακολουθούμε πάλι την μέθοδο της συνέλιξης.Ισχύει ότι

$f_Z = f_X * f_Y \longrightarrow$

$$f_Z = \begin{cases} \int_0^{\infty} f_X(z+y)\, f_Y(y)\, dy & z \geq 0 \quad (a) \\ \int_{-z}^{\infty} f_Y(z+y)\, f_Y(y)\, dy & z < 0 \quad (b) \end{cases} \qquad (3.2.3.1)$$

Ετσι η (3.2.3.1a) γίνεται

$$= \int_0^{\infty} \frac{2}{\sqrt{2\pi}}\, e^{-\frac{(z+y)^2}{2}}\, \frac{2}{\sqrt{2\pi}}\, e^{-\frac{y^2}{2}}\, dy \qquad (3.2.3.2)$$

και μετά από πράξεις η (3.2.3.2) γίνεται

$$\int_0^{\infty} \frac{4}{2\pi}\, e^{-\frac{z^2}{4}}\, e^{-(y+z/2)^2} \qquad (3.2.3.3)$$

θέτοντας

$u=(y+z)/2$ θα έχουμε

$$(3.2.3.3) = \frac{4}{2\pi}\, e^{-\frac{z^2}{4}} \int_{z/2}^{\infty} e^{-u^2}\, du =$$

$$\frac{4}{2\pi}\, e^{-\frac{z^2}{4}} \left[\int_0^{\infty} e^{-u^2}\, du - \int_0^{z/2} e^{-u^2}\, du \right] =$$

$$\frac{4}{2\pi}\, e^{-\frac{z^2}{4}} \left[\frac{\sqrt{\pi}}{2} - \frac{\sqrt{\pi}}{2}\, \Phi^*\left(\frac{z}{2}\right) \right] =$$

$$\frac{4}{2\pi} e^{-\frac{z^2}{4}} \frac{\sqrt{\pi}}{2} \left[1- \Phi^*(\frac{z}{2})\right] \qquad (3.2.3.4)$$

Με όμοιο τρόπο η (3.2.3.1b) γίνεται

$$\frac{4}{2\pi} e^{-\frac{z^2}{4}} \left[\int_0^{\infty} e^{-u^2} du - \int_0^{-z/2} e^{-u^2} du \right] =$$

$$\frac{4}{2\pi} e^{-\frac{z^2}{4}} \left[\frac{\sqrt{\pi}}{2} - \frac{\sqrt{\pi}}{2} \Phi^*(-\frac{z}{2}) \right] =$$

$$\frac{4}{2\pi} e^{-\frac{z^2}{4}} \frac{\sqrt{\pi}}{2} \left[1- \Phi^*(-\frac{z}{2}) \right] \qquad (3.2.3.5)$$

Οι σχέσεις (3.2.3.4), (3.2.3.5) δίνουν την συνάρτηση πυκνότητας πιθανότητας της διαφοράς δυο τυχαίων ανεξάρτητων μεταβλητών όταν το $z>0$ και οταν $z<0$, αντίστοιχα.

Αποδείξαμε λοιπόν ότι αν X,Y είναι δύο τυχαίες μεταβλητές που ακολουθούν την folded standard normal κατανομή τότε η συνάρτηση πυκνότητας της τυχαίας μεταβλητής $Z=X-Y$ θα δίνεται απο τον τύπο:

$$f_z(z) = \begin{cases} \frac{1}{\sqrt{\pi}} e^{-\frac{z^2}{4}} \left[1- \Phi^* (\frac{z}{2}) \right] & \text{όταν } z > 0 \text{ και} \\ \frac{1}{\sqrt{\pi}} e^{-\frac{z^2}{4}} \left[1- \Phi^* (- \frac{z}{2}) \right] & \text{οταν } z < 0 \end{cases}$$

3.2.4 ΑΝΑΠΤΥΞΗ ΑΛΓΟΡΙΘΜΟΥ ΓΙΑ ΤΗΝ ΕΥΡΕΣΗ ΠΟΣΟΣΤΙΑΙΩΝ ΣΗΜΕΙΩΝ ΤΗΣ FOLDED NORMAL ΚΑΤΑΝΟΜΗΣ

Στην ενότητα αυτή αναπτυσσεται αναπτύσσεται ένας αλγόριθμος για την επίλυση ολοκληρωτικών εξισώσεων όπου δοθέντος ενός k προσπαθούμε να βρούμε το u για το οποίο ισχύει ότι:

$$\int_0^u \frac{1}{\sqrt{2\pi}\,\sigma}\left[e^{-\frac{(x-\mu)^2}{2\sigma^2}}+e^{-\frac{(x+\mu)^2}{2\sigma^2}}\right]dx=k \qquad x>0 \qquad (3.2.4.1)$$

Θέτοντας $\frac{x-\mu}{\sigma}=v$ στην (3.2.4.1) και κάνοντας πράξεις καταλήγουμε στην σχέση:

$$\underbrace{\frac{1}{\sqrt{2\pi}}\left[\int_{-\frac{\mu}{\sigma}}^{\frac{u-\mu}{\sigma}} e^{-\frac{v^2}{2}}dv\right]}_{A}+\underbrace{\frac{1}{\sqrt{2\pi}}\left[\int_{-\frac{\mu}{\sigma}}^{\frac{u-\mu}{\sigma}} e^{-\frac{\left[v+\frac{2\mu}{\sigma}\right]^2}{2}}dv\right]}_{B}=k$$

Το ολοκλήρωμα B θέτοντας $v+\frac{2\mu}{\sigma}=v'$ γράφεται ως εξής:

$$B=\frac{1}{\sqrt{2\pi}}\int_{\frac{\mu}{\sigma}}^{\frac{u+\mu}{\sigma}} e^{-\frac{v'^2}{2}}dv'=B'$$

$$A+B'=k \Rightarrow \Phi\left(\frac{u-\mu}{\sigma}\right)-\Phi\left(-\frac{\mu}{\sigma}\right)+\Phi\left(\frac{u+\mu}{\sigma}\right)-\Phi\left(\frac{\mu}{\sigma}\right)=k \Rightarrow$$

$$\Phi\left(\frac{u-\mu}{\sigma}\right)+\Phi\left(\frac{u+\mu}{\sigma}\right)=k+1 \Rightarrow$$

$$\Phi\left(\frac{u}{\sigma}-\frac{\mu}{\sigma}\right)+\Phi\left(\frac{u}{\sigma}+\frac{\mu}{\sigma}\right)=k+1 \quad (3.2.4.2)$$

Θέτοντας στην (3.2.4.2) $u/\sigma = u'$ θα έχουμε:

$$\Phi\left(u'-\frac{\mu}{\sigma}\right)+\Phi\left(u'+\frac{\mu}{\sigma}\right)=k+1 \quad (3.2.4.3)$$

Θα πρέπει στη συνέχεια ξέροντας το μ,σ,k να βρούμε το u' για το οποίο ισχύει η (3.2.4.3).

Ο αλγόριθμος που θα χρησιμοποιηθεί είναι ο εξής:
Ξεκινάμε απο κάποιο u'_0 . Η σχέση (3.2.4.3) είναι ουσιαστικά ισοδύναμη με την σχέση:

$$\int_0^{u'-\frac{\mu}{\sigma}} \exp(-z^2/2)dz + \int_0^{u'+\frac{\mu}{\sigma}} \exp(-z^2/2)dz = k\sqrt{2\pi} \Rightarrow$$

$$\int_0^{u'_0+\Delta u'-\frac{\mu}{\sigma}} \exp(-z^2/2)dz+\int_0^{u'_0+\Delta u'+\frac{\mu}{\sigma}} \exp(-z^2/2)dz= k\sqrt{2\pi}+\Delta\varepsilon\Rightarrow$$

$$\int_0^{u'_0-\frac{\mu}{\sigma}} \exp(-z^2/2)dz + \int_{u'_0-\frac{\mu}{\sigma}}^{u'_0+\Delta u'-\frac{\mu}{\sigma}} \exp(-z^2/2)dz +$$

$$\int_0^{u'_0+\frac{\mu}{\sigma}} \exp(-z^2/2)dz + \int_{u'_0+\frac{\mu}{\sigma}}^{u'_0+\Delta u'+\frac{\mu}{\sigma}} \exp(-z^2/2)dz =$$

$k\sqrt{2\pi}+\Delta\varepsilon$ (3.2.4.4).

Ομως

$$\int_{u_0'-\frac{\mu}{\sigma}}^{u_0'+\Delta u'-\frac{\mu}{\sigma}} \exp(-z^2/2)dz = \exp\left(- \frac{(u_0' - \frac{\mu}{\sigma})^2}{2} \right) \Delta u' \quad (3.2.4.5)$$

και

$$\int_{u_0'+\frac{\mu}{\sigma}}^{u_0'+\Delta u'+\frac{\mu}{\sigma}} \exp(-z^2/2)dz = \exp\left(- \frac{(u_0' + \frac{\mu}{\sigma})^2}{2} \right) \Delta u' \quad (3.2.4.6)$$

Η (3.2.4.4) λόγω των (3.2.4.5),(3.2.4.6) θα γίνει:

$$\left[\exp\left(\frac{(u_0' - \frac{\mu}{\sigma})^2}{2} \right) + \exp\left(\frac{(u_0' + \frac{\mu}{\sigma})^2}{2} \right) \right] \Delta u' =$$

$$k\sqrt{2\pi} + \Delta\varepsilon - \int_0^{u_0'-\frac{\mu}{\sigma}} \exp(-z^2/2)dz - \int_0^{u_0'+\frac{\mu}{\sigma}} \exp(-z^2/2)dz \quad (3.2.4.7)$$

Απο την (3.2.4.7) μπορούμε να βρούμε το $\Delta u'$. Θέτουμε τώρα $u_0'' = u_0' + \Delta u'$ και επαναλαμβάνομε την παραπάνω διαδικασία. Έτσι μετά από κάποιες επαναλήψεις το Δu θα τείνει στο μηδέν και έτσι θα βρεθεί η ζητούμενη τιμή του u'.

Σημείωση:

Στην βιβλιογραφία (Leone et. al. (1961), Nelson(1980)) δίνονται μόνο περιοχές της folded normal κατανομής.

3.2.5 Η ΚΑΤΑΝΟΜΗ ΤΟΥ ΑΘΡΟΙΣΜΑΤΟΣ Ν ΑΝΕΞΑΡΤΗΤΩΝ FOLDED STANDARD NORMAL ΤΥΧΑΙΩΝ ΜΕΤΑΒΛΗΤΩΝ

Το άθροισμα ανεξάρτητων τυχαίων μεταβλητών που ακολουθούν την folded standard normal κατανομή έχει προσδιορισθεί από τον Hong-Ching Zhang (1986) (βλέπε επίσης Anderson, Fang (1990) "Statistical Inference in Elliptically Contoured and Related Distributions" απο τον Hong-Ching Zhang στην εργασία "The Distribution of the Squared Sum of the Independent Cauchy Variables").

Στην εργασία η συνάρτηση πυκνότητας του αθροίσματος ανεξάρτητων τυχαίων μεταβλητών που ακολουθούν την standard folded normal δίνεται απο την σχέση:

$$f_{|x_1|+\ldots|x_n|}(x) =$$

$$\frac{1}{2\pi} e^{-\frac{x^2}{2n}} \sum_{l=0}^{1} \left(\sqrt{2/\pi}\right)^{l} \sum_{r=0}^{\left[\frac{n-l}{2}\right]} \sum_{k=0}^{x} \alpha_{krl} \sum_{m=0}^{r+k} \frac{(-r-k)_m}{\left(\frac{l}{2}+1\right)_m m!} \left(\frac{x^2}{2n}\right)^m$$

όπου βέβαια $x>0$ και

$$\alpha_{kri} = \binom{n}{2r+l} \left(-\frac{2}{\pi}\right)^{r} \left(\frac{2}{n}\right)^{r+k+\frac{1}{2}+1} \Gamma\left(r+k+\frac{l}{2}+1\right) \times$$

$$\times \sum \frac{1}{\prod_{j=1}^{2r+l} (2k_j+1)k_j!\, 2^{k_j}}$$

$$(\alpha)_m = \alpha(\alpha+1)\ldots(\alpha+m+1)$$

Απόδειξη (Hong-Ching-Zhang):

Αποδεικνύεται ότι η χαρακτηριστική συνάρτηση του αθροίσματος των ανεξαρτήτων folded standard normal τυχαίων μεταβλητών δίνεται από την σχέση

$$\varphi(t) = \left(\sqrt{2/\pi}\right)^n e^{-\frac{t^2}{2}n}\left[\sqrt{\pi/2} + i\sum_{k=0}^{x}\frac{t^{2k+1}}{(2k+1)k!2^k}\right]^n$$

η οποία μπορεί να γραφεί ως

$$\varphi(t) = \left(\sqrt{2/\pi}\right)^n e^{-\frac{t^2}{2}n}\sum_{m=0}^{n}\binom{n}{m}\left(\sqrt{\pi/2}\right)^{n-m} i^m \times$$

$$\times\left(\sum_{k=0}^{\infty}\frac{t^{2k+1}}{(2k+1)k!2^k}\right)^m =$$

$$e^{-\frac{t^2}{2}n}\sum_{r=0}^{[\frac{n}{2}]}\binom{n}{2r}\left(-\frac{2}{\pi}\right)^r\left(\sum_{k=0}^{\infty}\frac{t^{2k+1}}{(2k+1)k!2^k}\right)^{2r} +$$

$$+i\left(\sqrt{2/\pi}\right)e^{-\frac{t^2}{2}n}\sum_{r=0}^{[\frac{n-1}{2}]}\binom{n}{2r+1}\left(-\frac{2}{\pi}\right)^r\left(\sum_{k=0}^{x}\frac{t^{2k+1}}{(2k+1)k!2^k}\right)^{2r+1} =$$

$$=\varphi_0(t)+i\left(\sqrt{2/\pi}\right)\varphi_1(t)$$

Ομως ισχύει:

$$\sum_{k=0}^{\infty}\frac{t^{2k+1}}{(2k+1)k!2^k}=\frac{2}{|t|}\sum_{k=0}^{\infty}\frac{|t^2|^{k+1}}{2(2k+1)k!2^k}$$

$$\leq\frac{2}{|t|}\sum_{k+1=1}^{x}\frac{|t^2|^{k+1}}{2^{k+1}(k+1)!}\leq\frac{2}{|t|}e^{\frac{t^2}{2}}$$

τότε για l=0,1 και $|t|>1$ θα έχουμε

$$e^{-\frac{t^2}{2}n}\left(\sum_{k=0}^{x}\frac{|t|^{2k+1}}{(2k+1)k!2^k}\right)^{2r+1}$$

$$\leq\frac{2^n}{|t|^{2r+1}}e^{-\frac{t^2}{2}(n-2r+1)}\qquad 2r+1<n$$

$$\leq\frac{2^n}{|t|^n}\qquad 2r+1=n\qquad (3.2.5.1)$$

Επειδή η φ(t) είναι συνεχής, και λόγω της προηγούμενης σχέσης, είναι γνωστό ότι $\int_{-x}^{+\infty}|\varphi_l(t)|dt<x$, l=0,1 και γι'αυτό $\int_{-\infty}^{+\infty}|\varphi(t)|dt<\infty$. Ετσι μπορούμε να υπολογίσουμε την πυκνότητα του αθροίσματος $|X_1|+\ldots+|X_n|$ αντιστρέφοντας την χαρακτηριστική συνάρτηση.

$$\varphi_1(t) = e^{-\frac{t^2}{2}n} \sum_{r=0}^{[\frac{n-1}{2}]} \binom{n}{2r+1} \left(-\frac{2}{\pi}\right)^r \left(\sum_{k=0}^{\infty} \frac{t^{2k+1}}{(2k+1)k!2^k} \right)^{2r+1} =$$

$$e^{-\frac{t^2}{2}n} \sum_{r=0}^{[\frac{n-1}{2}]} \binom{n}{2r+1} \left(-\frac{2}{\pi}\right)^r t^{2r+1} \times$$

$$\times \sum_{k=0}^{\infty} \sum_{k_1+\ldots k_{2r+1}=k} \frac{1}{\prod_{j=1}^{2r+1} (2k_j+1)k_j!2^{k_j}} t^{2k} =$$

$$e^{-\frac{t^2}{2}n} \sum_{r=0}^{[\frac{n-1}{2}]} \binom{n}{2r+1} \left(-\frac{2}{\pi}\right)^r t^{2r+1} \sum_{k=0}^{\infty} b_{kr1} t^{2k}$$

όπου

$$b_{kr1} = \sum_{k_1+\ldots k_{2r+1}=k} \frac{1}{\prod_{j=1}^{2r+1} (2k_j+1)k_j!2^{k_j}} \leq \frac{1}{2^k} \frac{(2r+1)^{k+2r+1}}{(k+2r+1)!}$$

Η μορφή αυτή δείχνει ότι η σειρά συγκλίνει ομοιόμορφα σε κάθε πεπερασμένο διάστημα. Η σειρά επίσης λόγω της (3.2.5.1) μπορεί να γίνει οσοδήποτε μικρή όταν το t γίνει αρκετά μεγάλο. Έτσι η σειρά συγκλίνει ομοιόμορφα στο $[0,\infty)$. Επειδή $\varphi_0(t)$ είναι άρτια και $\varphi_1(t)$ είναι περιττή θα έχουμε :

$$f_{|x_1|+\ldots+|x_2|}(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \varphi(t)dt$$

$$\frac{1}{\pi}\int_0^{+\infty} \cos tx\varphi_0(t)dt + \frac{1}{\pi}\sqrt{2/\pi}\int_0^{+\infty} \sin tx\varphi_1(t)dt$$

και λόγω του Gradshteyn & Ryzhik (1980) ολοκλήρωμα 3.952 θα έχουμε:

$$\frac{1}{\pi}\int_0^{+\infty} \cos tx\varphi_0(t)dt = \frac{1}{\pi}\sum_{r=0}^{[\frac{n}{2}]}\binom{n}{2r}\left(-\frac{2}{\pi}\right)^r \times$$

$$\times \sum_{k=0}^{\infty} b_{kr0}\int_0^{+\infty} \cos tx\, e^{-\frac{t^2}{2}n}\, t^{2r+2k}\, dt$$

$$\frac{1}{\pi}\sqrt{2/\pi}\int_0^{+\infty} \sin tx\varphi_1(t)dt = \frac{1}{\pi}\sqrt{2/\pi}\sum_{r=0}^{[\frac{n-1}{2}]}\binom{n}{2r+1}\left(-\frac{2}{\pi}\right)^r \times$$

$$\times \sum_{k=0}^{r} b_{kr1}\int_0^{+\infty} \sin tx\, e^{-\frac{t^2}{2}n}\, t^{2r+2k+1}\, dt$$

όπου

$$\int_0^{+\infty} \cos tx\, e^{-\frac{t^2}{2}n}\, t^{2r+2k}\, dt = \frac{\Gamma\left(\frac{2r+2k+1}{2}\right)}{2\,(\frac{n}{2})^{r+k+1/2}}\, F_1\left(\frac{2r+2k+1}{2}; \frac{1}{2}; -\frac{x^2}{2n}\right)$$

$$= \frac{\Gamma\left(r+k+\frac{1}{2}\right)}{2\ \left(\frac{n}{2}\right)^{r+k+1/2}}\ e^{-\frac{x^2}{2n}} \sum_{m=0}^{r+k} \frac{(-r-k)_m}{\left(\frac{1}{2}+1\right)_m m!}\ \left(\frac{x^2}{2n}\right)^m$$

$$\int_0^{+\infty} \sin t\ x\ e^{-\frac{t^2}{2}n}\ t^{2r+2k+1}\ dt =$$

$$x\ e^{-\frac{x^2}{2n}}\ \frac{\Gamma\left(\frac{2r+2k+3}{2}\right)}{2\ \left(\frac{n}{2}\right)^{r+k+3/2}}\ F_1\left(1-\frac{2r+2k+2}{2}\ ;\frac{3}{2};\ \frac{x^2}{2n}\right) =$$

$$\frac{\Gamma\left(r+k+1+\frac{1}{2}\right)}{2\ \left(\frac{n}{2}\right)^{r+k+1/2}}\ x\ e^{-\frac{x^2}{2n}} \sum_{m=0}^{r+k} \frac{(-r-k)_m}{\left(\frac{1}{2}+1\right)_m m!}\ \left(\frac{x^2}{2n}\right)^m .$$

Εστω

$$\alpha_{kri} = \binom{n}{2r+1} \left(-\frac{2}{\pi}\right)^r b_{kri} \left(\frac{2}{n}\right)^{r+k+\frac{1}{2}+1} \Gamma\left(r+k+\frac{1}{2}+1\right)$$

θα έχουμε:

$$f_{|x_1|+\ldots+|x_2|}(x) = \frac{1}{2\pi} \sum_{r=0}^{\left[\frac{n}{2}\right]} \sum_{k=0}^{\infty} \alpha_{kr0}\ e^{-\frac{x^2}{2n}} \sum_{m=0}^{r+k} \frac{(-r-k)_m}{\left(\frac{1}{2}\right)_m m!}\ \left(\frac{x^2}{2n}\right)^m$$

$$+\frac{1}{2\pi}\sqrt{2/\pi}\sum_{r=0}^{[\frac{n-1}{2}]}\sum_{k=0}^{x}\sigma_{kr1}\,e^{-\frac{x^2}{2n}}\sum_{m=0}^{r+k}\frac{(-r-k)_m}{(\frac{1}{2}+1)_m\,m!}\,(\frac{x^2}{2n})^m$$

Το τελευταίο άθροισμα μας δίνει το ζητούμενο αποτέλεσμα.

Ο καθορισμός της κατανομής του αθροίσματος folded normal τυχαίων μεταβλητών ολοκληρώνει την μελέτη της προσέγγισης αξιολόγησης με βάση τον κανόνα (3.1.1). Προκειμένου να χρησιμοποιηθούν τα αποτελέσματα αυτά σε πρακτικά προβλήματα απαιτείται ο προσδιορισμός των περιοχών πιθανότητας της κατανομής του αθροίσματος folded standard normal τυχαίων μεταβλητών (που είναι η κατανομή της στατιστικής συνάρτησης που μας ενδιαφέρει) ώστε να καταστεί δυνατόν να γίνουν στατιστικοί έλεγχοι υποθέσεων.

Η πινακοποίηση αυτή είναι αρκετά δύσκολη λόγω της πολύπλοκης μορφής της συνάρτησης πυκνότητας πιθανότητας και παραμένει ένα ανοικτό πρόβλημα.

3.3 ΠΕΡΙΠΤΩΣΗ ΜΙΚΡΟΥ ΑΡΙΘΜΟΥ ΠΑΡΑΤΗΡΗΣΕΩΝ

Οπως και στην προηγουμένως έτσι και στην περίπτωση αυτή προσπαθούμε να προσδιορίσουμε την κατανομή που θ'ακολουθεί ο βαθμολογικός κανόνας (3.1.1). Η κατανομή αυτή είναι η απόλυτη τιμή της κατανομής της t με l_t-m βαθμούς ελευθερίας. Την κατανομή αυτή την ονομάζουμε folded-t και την μελετάμε ιδιαίτερα αφού αφ'ενός μεν δεν έχει μελετηθεί μέχρι τώρα στην βιβλιογραφία, αφ'ετέρου η γνώση των ιδιοτήτων της θα μας βοηθήσει να προχωρήσουμε σε στατιστική συμπερασματολογία.

ΜΕΛΕΤΗ ΤΗΣ ΚΑΤΑΝΟΜΗΣ FOLDED-t

Οπως είναι γνωστό η συνάρτηση πυκνότητας μιας τυχαίας μεταβλητής T που ακολουθεί την t κατανομή με ν βαθμούς ελευθερίας δίνεται απο την σχέση

$$f_T(x) = \frac{\Gamma\left[\frac{\nu+1}{2}\right]}{\Gamma\left[\frac{\nu}{2}\right]\sqrt{\nu\pi}}\left[1+\frac{x^2}{\nu}\right]^{-\frac{\nu+1}{2}} \quad \nu\in N, \quad x\in R \qquad (3.3.1)$$

Εστω $F_T(x)$ η συνάρτηση κατανομής. Μας ενδιαφέρει να προσδιορίσουμε την συνάρτηση πυκνότητας της τυχαίας μεταβλητής $W=|T|$.

Είναι εύκολο να δει κανείς ότι:

$$F_W(x) = F_T(x)-F_T(-x), \quad x>0 \qquad (3.3.2)$$

όπου λόγω της συμμετρίας της (3.3.1) η (3.3.2) θα γίνει

$$F_W(x) = 2F_T(x)-1, \qquad x>0$$

και η συνάρτηση πυκνότητας θα είναι

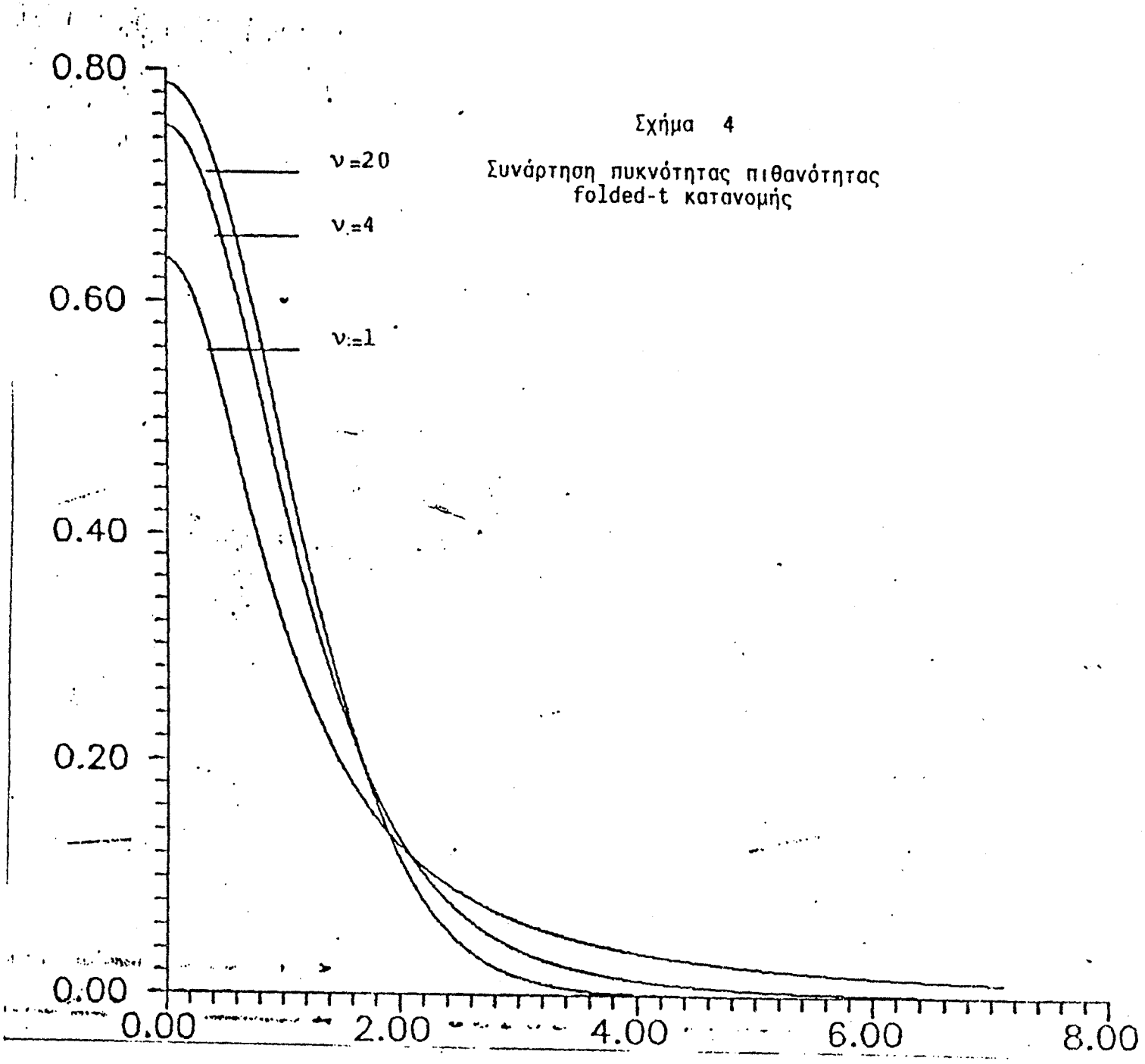
$$f_W(x) = 2f_T(x) \qquad x>0$$

Η παραπάνω οδηγεί στον παρακάτω ορισμό:

Ορισμός: Μία μη-αρνητική πραγματική τυχαία μεταβλητή X θα λέμε οτι θα ακολουθεί την folded t κατανομή με ν βαθμούς ελευθερίας αν η συνάρτηση πυκνότητας πιθανότητας της δίνεται από την:

$$f_X(x) = 2\,\frac{\Gamma\left[\frac{\nu+1}{2}\right]}{\Gamma\left[\frac{\nu}{2}\right]\sqrt{\nu\pi}}\left[1+\frac{x^2}{\nu}\right]^{-\frac{\nu+1}{2}} \qquad \nu\in N,\ x>0$$

Η γραφική παράσταση της συνάρτησης πυκνότητας της folded-t κατανομής για διάφορες τιμές του ν δίνεται στο σχήμα 4.



Σχήμα 4

Συνάρτηση πυκνότητας πιθανότητας
folded-t κατανομής

Θεώρημα 1: Εστω X μια μη αρνητική τυχαία μεταβλητή που ακολουθεί την folded t κατανομή με ν βαθμούς ελευθερίας. Τότε:

$$E(X) = \begin{cases} 2\sqrt{\nu/\pi}\ \dfrac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)(\nu-1)} & \text{για } \nu>1 \\ \infty & \text{για } \nu=1 \end{cases}$$

και

$$V(X) = \begin{cases} \dfrac{\nu}{\nu-2} - \dfrac{4\nu}{\pi(\nu-1)^2}\left[\dfrac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}\right]^2 & \text{για } \nu > 2 \\ \infty & \text{για } \nu \le 2 \end{cases}$$

Απόδειξη

Για την μέση τιμή έχουμε

$$E(X) = 2\,\frac{\Gamma\left[\frac{\nu+1}{2}\right]}{\Gamma\left[\frac{\nu}{2}\right]\sqrt{\nu\pi}} \int_0^x x\left[1+\frac{x^2}{\nu}\right]^{-\frac{\nu+1}{2}} dx =$$

$$\frac{\Gamma\left[\frac{\nu+1}{2}\right]}{\Gamma\left[\frac{\nu}{2}\right]\sqrt{\nu\pi}} \int_0^{\infty} \left[1+\frac{x^2}{\nu}\right]^{-\frac{\nu+1}{2}} dx^2 =$$

$$\frac{v\,\Gamma\left[\frac{v+1}{2}\right]}{\Gamma\left[\frac{v}{2}\right]\sqrt{v\pi}\left[-\frac{v}{2}+\frac{1}{2}\right]}\int_0^{\infty} d\left[1+\frac{x^2}{v}\right]^{-\frac{v}{2}+\frac{1}{2}} =$$

$$\frac{2v\,\Gamma\left[\frac{v+1}{2}\right]}{\Gamma\left[\frac{v}{2}\right]\sqrt{v\pi}\;(1-v)}\left[1+\frac{x^2}{v}\right]^{-\frac{v}{2}+\frac{1}{2}}\Bigg|_0^{\infty}$$

Αν $v>1 \Rightarrow -\frac{v}{2}+\frac{1}{2}<0 \Rightarrow \lim_{x\to\infty}\left[1+\frac{x^2}{v}\right]^{-\frac{v}{2}+\frac{1}{2}}=0$

Αν $v=1$ τότε θα έχουμε:

$$E(X)= 2\,\frac{1}{\Gamma\left[\frac{1}{2}\right]\sqrt{\pi}}\int_0^{\infty} x\left[1+x^2\right]^{-1}dx =$$

$$\frac{1}{\Gamma\left[\frac{1}{2}\right]\sqrt{\pi}}\int_0^{\infty}\frac{dx^2}{1+x^2} = \frac{1}{\Gamma\left[\frac{1}{2}\right]\sqrt{\pi}}\ln\left(1+x^2\right)\Big|_0^{\infty} = \infty$$

Επειδή $X=|T|$ όπου T είναι κάποια t τυχαία μεταβλητή θα έχουμε για την διασπορά:

$$V(X) = E(X^2) - \left[E(X)\right]^2 = E(T^2) - \left[E(T)\right]^2 \qquad (3.3.3)$$

όμως $E(T^2) = V(T) = \frac{v}{v-2} \qquad v > 2 \qquad (3.3.4)$

Η (3.3.3) μαζί με την (3.3.4) μας δίνει την διασπορά της folded-t.

ΣΧΕΣΗ ΤΗΣ ΚΑΤΑΝΟΜΗΣ FOLDED-t ΜΕ ΤΗΝ ΚΑΤΑΝΟΜΗ FOLDED NORMAL

Μας είναι αρκετά γνωστή η σχέση της t κατανομής με την τυποποιημενή κανονική κατανομή. Ετσι αν T_v ακολουθεί την t_v και αν Z ακολουθεί την N(0,1) τότε

$$\lim_{v \to x} F_{T_v}(x) = F_Z(x)$$

Είναι ενδιαφέρον σ'αυτό το σημείο να ελέγξουμε αν παρόμοια σχέση υπάρχει και μεταξύ της folded normal και της folded t.

Στα προηγούμενα αναφερθήκαμε στον ορισμό της folded normal και της folded standard normal των οποίων οι συναρτήσεις πυκνότητας πιθανότητας δίνονται από τις παρακάτω σχέσεις:

$$f_{|X|}(x) = \frac{1}{\sigma\sqrt{2\pi}} \left[\exp\left(\frac{x-\mu}{\sigma} \right)^2 + \left(\exp\left(\frac{x+\mu}{\sigma} \right)^2 \right] \right.$$

Θέτοντας στην παραπάνω $\mu=0$ και $\sigma=1$ θα έχουμε την συνάρτηση πυκνότητας της folded standard normal κατανομής

$$f_{|z|}(z) = \sqrt{2/\pi}\ e^{-z^2/2} \qquad z > 0 .$$

Για το όριο της συνάρτησης κατανομής της folded t θα έχουμε:

$$\lim_{v \to \infty} F_{|T_v|}(x) = \lim_{v \to x} (2 F_{T_v}(x) - 1) = F_{|z|}(x)$$

ότου $F_{|z|}(x)$ είναι η συνάρτηση κατανομής της standard folded normal κατανομής έτσι

$$\lim_{\nu \to \infty} F_{|T_\nu|}(x) = F_{|z|}(x)$$

που αποδεικνύει οτι η οριακή περίπτωση της folded-t είναι η folded standard normal.

ΣΧΕΣΕΙΣ ΤΗΣ FOLDED-t ΚΑΤΑΝΟΜΗΣ ΜΕ ΚΑΤΑΝΟΜΕΣ ΠΟΥ ΕΧΟΥΝ ΣΧΕΣΗ ΜΕ ΠΑΡΑΓΩΓΗ ΤΥΧΑΙΩΝ ΑΡΙΘΜΩΝ (GENERATION OF RANDOM NUNBERS)

Στην ενότητα αυτή δίνονται σχέσεις της κατανομής folded-t με άλλες κατανομές που έχουν σχέση με παραγωγή (generation) τυχαίων αριθμών. Από τ'αποτελέσματα που βρίσκουμε θα δούμε οτι η folded-t σχετίζεται με πολύ γνωστές κατανομές όπως η F κατανομη, η χ^2- κατανομή, η βήτα κατανομή δευτέρου τύπου, η folded normal κατανομή και η half Cauchy κατανομή. Οι σχέσεις αυτές δίνονται στα παρακάτω θεωρήματα:

Θεώρημα 1: Εστω Z μία μη αρνητική πραγματική τυχαία μεταβλητή που ακολουθεί την folded standard normal κατανομή. Εστω X μία άλλη μη αρνητική πραγματική τυχαία μεταβλητή που ακολουθεί την χ^2 κατανομή με ν βαθμούς ελευθερίας. Τότε η τυχαία μεταβλητή

$$T = Z / \sqrt{X/\nu}$$

ακολουθεί την folded t κατανομή.

Απόδειξη: Είναι γνωστό ότι αν Z_0 είναι μία τυχαία μεταβλητή που ακολουθεί την standard nornal κατανομή τότε η τυχαία μεταβλητή $T_0 = Z_0/\sqrt{X/\nu}$ είναι μία t τυχαία μεταβλητή με ν βαθμούς ελευθερίας. Τότε

$$T= Z/\sqrt{X/\nu} = |Z_0|/\sqrt{X/\nu} = |T_0|$$

απ'όπου και το αποτέλεσμα.

Θεώρημα 2: Εστω X, Y ανεξάρτητες μη αρνητικές τυχαίες μεταβλητές που ακολουθούν την chi κατανομη με παραμέτρους 1 και ν και συνάρτηση πυκνότητας πιθανότητας που δίνεται από τις σχέσεις

$$f_X(x) = \frac{2^{1/2}}{\Gamma(1/2)} \exp(-x^2/2) \qquad x>0 \qquad (3.3.5)$$

και

$$f_Y(y) = \frac{1}{2^{\frac{\nu}{2}-1}\Gamma(\nu/2)} y^{\nu-1} \exp(-y^2/2) \quad y>0$$

αντίστοιχα (βλέπε π.χ. Patil(1984). Τότε η τυχαία μεταβλητή

$$Z=\nu^{1/2} X/Y$$

θα έχει την folded t κατανομή με ν βαθμούς ελευθερίας.

Απόδειξη: Είναι γνωστό ότι η standard folded normal κατανομή είναι η chi κατανομή με 1 βαθμό ελευθερίας. Αν τώρα η W ακολουθεί την χ^2κατανομή με ν βαθμούς ελευθερίας τότε $Y=\sqrt{W}$. Το αποτέλεσμα είναι άμεση συνέπεια του πρώτου θεωρήματος.

Θεώρημα 3: Εστω Z_1 , Z_2 δύο ανεξάρτητες μη αρνητικές πραγματικές τυχαίες μεταβλητές που ακολουθούν την folded standard normal κατανομή. Τότε η

$$Y = Z_1 / Z_2$$

ακολουθεί την folded t με 1 βαθμό ελευθερίας.

Οπως είπαμε παραπάνω η folded standard normal κατανομή είναι μία ειδική περίπτωση της chi κατανομής. Γι'αυτό το λόγο στηριζόμενοι στο παραπάνω θεώρημα καταλήγουμε στο ακόλουθο θεώρημα:

Θεώρημα 4: Εστω Z_1 , Z_2 ανεξάρτητες chi μεταβλητές με παράμετρο 1 και συνάρτηση πυκνότητας πιθανότητας που δίνεται απο την (3.3.5) . Τότε η

$$Y= Z_1 / Z_2$$

ακολουθεί την folded t κατανομή με 1 βαθμό ελευθερίας.

Θεώρημα 5: Εστω X μία folded t μεταβλητή και Y μια άλλη τυχαία μεταβλητή που ακολουθεί την standard half- Cauchy κατανομή με συνάρτηση πυκνότητας πιθανότητας που δίνεται από την

$$f_Y(y) = \frac{2}{\pi(1+x^2)} \quad x>0$$

(Mijnheer (1968))

Τότε $X \stackrel{d}{=} Y$.

Δηλαδή η folded t κατανομή με ν=1 βαθμούς ελευθερίας είναι η half-Cauchy κατανομή.

Θεώρημα 6: Εστω X μία folded t τυχαία μεταβλητή. Τότε η τυχαία μεταβλητή

$$Y = X^2$$

ακολουθεί την F κατανομή με 1 και ν βαθμούς ελευθερίας και συνάρτηση πυκνότητας πιθανότητας που δίνεται από την σχέση

$$f(x) = \frac{1}{B(1/2,\ \nu/2)} x^{-1/2} \nu^{\nu/2} \quad (\nu+x)^{-(\nu+1)/2} \quad x>0$$

Θεώρημα 7: Εστω X είναι μια folded t τυχαία μεταβλητή. Τότε η τυχαία μεταβλητή

$$Y=X^2/\nu$$

ακολουθεί την βήτα κατανομή δυτέρου τύπου (beta prime η Pearson type VI κατανομη) με παραμέτρους 1/2 και ν/2 και συνάρτηση πυκνότητας πιθανότητας που δίνεται από τον τύπο

$$f(x) = \frac{1}{B(1/2,\ \nu/2)} x^{-1/2} (1+x)^{-(\nu+1)/2}, \quad x>0$$

(βλέπε π.χ. Patil(1984))

Θεώρημα 8: Εστω X να είναι μία τυχαία μεταβλητή που ακολουθεί την F κατανομή με ν και ν βαθμούς ελευθερίας και

συνάρτηση πυκνότητας πιθανότητας που δίνεται από την

$$f_X(x) = \frac{1}{B(v/2,\ v/2)}\ x^{v/2-1}\ (1+x)^{-v} \qquad ,\quad x>0$$

Τότε η τυχαία μεταβλητή

$$Y = \frac{v^{1/2}}{2}\ |X^{1/2} - X^{-1/2}|$$

ακολουθεί την folded t κατανομή με ν βαθμούς ελευθερίας.

Θεώρημα 9: Εστω X_1 , X_2 πραγματικές τυχαίες μεταβλητές που η κάθε μία ακολουθεί την χ^2 κατανομή με ν βαθμούς ελευθερίας και συνάρτηση πυκνότητας πιθανότητας που δίνεται από την σχέση

$$f_X(x) = \frac{2^{-v/2}}{\Gamma(v/2)}\ x^{v/2-1}\ \exp(-x/2) \qquad x>0$$

Τότε η τυχαία μεταβλητή

$$Y = \frac{v^{1/2}}{2}\ \frac{|X_1 - X_2|}{(X_1X_2)^{1/2}}$$

ακολουθεί την folded t κατανομή με ν βαθμούς ελευθερίας.

Θεώρημα 10: Εστω X μία πραγματική τυχαία μεταβλητή που ακολουθεί την γενικευμένη Cauchy κατανομή με παραμέτρους α , β=1, γ=2 και δ=(ν+1)/2 και συνάρτηση πυκνότητας πιθανότητας που δίνεται από την

$$f(x)= \frac{1}{B(1/2,\ \nu-1/2\ )}\ [1+(x-\alpha)^2]^{-(\nu+1)/2} \quad , \quad x,\alpha \in R$$

(βλέπε π.χ. Patil(1984))

Τότε η τυχαία μεταβλητή

$$Y = \nu^{1/2}|x-\alpha|$$

ακολουθεί την folded t κατανομή με (ν+1)/2 βαθμούς ελευθερίας.

ΠΕΡΙΟΧΕΣ ΤΗΣ FOLDED T ΚΑΤΑΝΟΜΗΣ

Για την αντιμετώπιση πρακτικών προβλημάτων θα είναι χρήσιμο να προσδιορίσουμε τις πιθανότητες που είναι συνδεδεμένες με την folded t κατανομή.

Για μικρές τιμές του ν ($\nu \leq 30$) μπορούμε να βρουμε περιοχες της κατανομης folded t κανοντας μια τροποποιηση των πινακων που αναφερονται στην t κατανομη βασιζομενοι βεβαια στην συναρτηση πυκνοτητας της κατανομης.

Για μεγάλες τιμές του ν μπορούμε να χρησιμοποιήσουμε τους πίνακες που δίνονται απο τον Leone et. al. (1961) της folded standard normal κατανομής (διότι όπως ήδη αποδείξαμε η οριακή περίπτωση της folded t είναι η folded standard normal).

Η εκτίμηση των ολοκληρωμάτων της folded t κατανομής της μορφής:

$$f_\nu(t) = 2\frac{\Gamma\left[\frac{\nu+1}{2}\right]}{\Gamma\left[\frac{\nu}{2}\right]\sqrt{\nu\pi}}\left[1+\frac{t^2}{\nu}\right]^{-\frac{\nu+1}{2}} \qquad t>0$$

μπορεί να γίνει με μία μετατροπή του αλγορίθμου του Cooper (1968) για τις περιοχές της t κατανομής (βλέπε ακομη Griffiths and Hill (1985)).

Με την ίδια λογική κατασκευάζονται και πίνακες που δίνουν τα εκατοστιαία σημεία της folded t κατανομης.

Πίνακες που δίνουν τις περιοχές καθώς και τα εκατοστιαία σημεία της folded t κατανομής δίνονται στην εργασία Psarakis-Panaretos (1990) και στο παράρτημα III.

Για την επίλυση πρακτικών προβλήματων είναι απαραίτητο να προχωρήσουμε στην εύρεση του αθροίσματος κατανομών που ακολουθούν την folded t κατανομή. Στην αντίστοιχη περίπτωση της t έχουν βρεθεί αθροίσματα κατανομών κάτω από ειδικές περιπτώσεις (βλέπε Walker & Saw (1982)).

Η εύρεση τέτοιων αθροισμάτων είναι μία πολύπλοκη διαδικασία η οποία μπορεί να επιλυθεί με αριθμητική ολοκλήρωση.

ΚΕΦΑΛΑΙΟ 4

ΟΙ ΔΙΜΕΤΑΒΛΗΤΕΣ ΜΟΡΦΕΣ ΤΗΣ FOLDED NORMAL ΚΑΙ ΤΗΣ FOLDED T ΚΑΤΑΝΟΜΗΣ

4.1 ΕΙΣΑΓΩΓΗ

Θεωρούμε και πάλι τον βαθμολογικό κανόνα (3.1.1)

$$|r_{t+1}| = \frac{|\hat{Y}^{\circ}_{t+1} - Y^{\circ}_{t+1}|}{S_t \sqrt{(1+ X^{\circ}_{t+1}(X_t' X_t)^{-1} X^{\circ\prime}_{t+1})}}$$

Οπως είναι γνωστό, η διαφορά

$$\hat{Y}^{\circ}_{t+1}(M) - Y^{\circ}_{t+1} = e_t(M),$$

όπου Μ δηλώνει το υπό εξέταση μοντέλο Α ή Β, ακολουθεί την κανονική κατανομή με μεσή τιμή 0 και διασπορά $\sigma^2(M)$. Στην ενοτητα αυτη θα μας απασχολησει η εύρεση της απο κοινού κατανομής των $|e_t(A)|$ και $|e_t(B)|$ η οποία θα μας χρησιμεύσει σε στατιστική συμπερασματολογία σχετική με την σύγκριση της προβλεπτικής ικανότητας δυομοντέλων. Ετσι, καταλήγουμε στον ορισμό κατανομών οι οποίες δεν έχουν μελετηθεί ακόμη στην βιβλιογραφία. Αυτές οι κατανομές είναι οι διμεταβλητές γενικεύσεις των folded normal και folded t τις οποίες μελετήσαμε στο προηγούμενο κεφάλαιο. Η ανάγκη της μελέτης των

ιδιοτήτων των κατανομών αυτών είναι προφανής. Έτσι στην συνέχεια μελετώνται βασικές ιδιότητές τους καθώς και η σχέση που τις συνδέει όπως επίσης και κάποιες ενδιαφέρουσες ειδικές περιπτώσεις αυτών.

4.2 Η ΔΙΜΕΤΑΒΛΗΤΗ FOLDED NORMAL ΚΑΤΑΝΟΜΗ

Θεωρούμε δύο γραμμικά μοντέλα Α και Β της μορφής:

$$Y_t = X_t(M)\ \beta_M + \varepsilon_t(M) \quad (M=A,B)$$

όπου όπως έχουμε αναφέρει στην ενότητα 1,

Y_t είναι διάνυσμα διαστάσεως $l_t \times 1$ των εξαρτημένων μεταβλητών,

$X_t(M)$ πίνακας διαστάσεως $l_t \times m_M$ $(l_t \geq m_M | X_t(M)' X_t(M)| \neq 0)$ γνωστών συντελεστών,

β_M διάνυσμα διαστάσεως $m_M \times 1$ συντελεστών παλινδρόμησης και

$\varepsilon_t(M)$ διάνυσμα διαστάσεως $l_t \times 1$ τυχαίων μεταβλητών που ακολουθούν την κανονική κατανομή με

$E(\varepsilon_t(M))=0$ και

$V(\varepsilon_t(M))=\sigma^2(M)\ I_t$ $(\sigma^2(M) < \infty)$ όπου I_t είναι ο $l_t \times l_t$ ταυτοτικός πίνακας.

Μία πρόβλεψη για την εξαρτημένη τυχαία μεταβλητή για την χρονική στιγμή t+1 θα δίνεται από την στατιστική συνάρτηση:

$$\hat{Y}^0_{t+1}(M) = X^0_{t+1}(M)\ \hat{\beta}_t(M)$$

όπου $\hat{\beta}_t(M)$ είναι η εκτιμήτρια ελαχίστων τετραγώνων του β_M την χρονική στιγμή t, και $X^0_{t+1}(M)$ είναι ένα $l_t \times m_M$ διάνυσμα των τιμών

των συντελεστών παλινδρόμησης την στιγμή t+1. Εστω $Y^{0}_{t+1}(M)$ η παρατηρούμενη τιμή της εξαρτημένης τυχαίας μεταβλητής την στιγμή t+1. Τότε, για δύο μοντέλα Α και Β τα οποία θέλουμε να συγκρίνουμε θα ισχύει ότι η μεταβλητή ως προς την προβλεπτική ικανότητα:

$$\hat{Y}^{\circ}_{t+1}(M) - Y^{\circ}_{t+1}(M) = e_{t+1}(M) \qquad (4.2.1)$$

ακολουθεί την κανονική κατανομή με μέση τιμή 0 και διασπορά σ_M όπου

$$\sigma^2_M = \sigma^2(M)\left[1 + X^{0}_{t+1}(M)\,(X'_t(M)X_t(M))^{-1}X'^{0}_{t+1}(M)\right] \qquad (4.2.2)$$

όπου M=A ή M=B.

Ετσι η απόλυτη τιμή της (4.2.1) η οποία θα είναι η

$$|\hat{Y}^{\circ}_{t+1}(M) - Y^{\circ}_{t+1}(M)| = |e_{t+1}(M)|$$

ακολουθεί την folded normal κατανομή με

μέση τιμή $\mu_f = \sqrt{2/\pi}\ \sigma_M$ και διασπορά $\sigma^2_f = \sigma^2_M(1-2/\pi)$ (Leone et.al. 1961).

Ειναι προφανές ότι για τον έλεγχο της υπόθεσης

$$H_0: E|e_t(A)| = E|e_t(B)|$$

απαιτείται η μελέτη της από κοινού κατανομής των $|e_t(A)|$ και $|e_t(B)|$. Η κατανομή αυτή είναι η διμεταβλητή μορφή της folded normal.

Πράγματι αν X,Y είναι δύο μεταβλητές που ακολουθουν την κανονική κατανομή τότε η από κοινού κατανομή τους είναι η διμεταβλητή κανονική της οποίας η συνάρτηση πυκνότητας, όπως

είναι γνωστό δίνεται από την σχέση

$$f_{X,Y}(x,y)=$$

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\varrho}}\exp\left(-\frac{1}{2(1-\varrho^2)}\left(\frac{(x-\mu_1)^2}{\sigma_1^2}-2\varrho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right)\right)$$

Εστω $F_{X,Y}(x,y)$ η αντίστοιχη συνάρτηση κατανομής των X,Y. Η από κοινού συνάρτηση κατανομής των τυχαίων μεταβλητών $|X|,|Y|$ θα είναι:

$$F_{|X|,|Y|}(x.y)= P(|X|\le x,\ |Y|\le y)= P(-x\le X\le x,\ -y\le Y\le y)=$$

$$F_{X,Y}(x,y)+F_{X,Y}(-x,-y)-F_{X,Y}(-x,y)-F_{X,Y}(x,-y)\quad x,y>0$$

Αν θεωρήσουμε τις αντίστοιχες συναρτήσεις πυκνότητας θα έχουμε:

$$f_{|X|,|Y|}(x,y)=f_{X,Y}(x,y)+f_{X,Y}(-x,-y)+f_{X,Y}(-x,y)+f_{X,Y}(x,-y)$$

Η σχέση αυτή οδηγεί στον ακόλουθο ορισμό:

Ορισμός: Η από κοινού κατανομή δύο μη αρνητικών πραγματικών τυχαίων μεταβλητών $Z=|X|$ και $W=|Y|$ θα λέμε οτι είναι η διμεταβλητή folded normal κατανομή αν η συνάρτηση πυκνότητας τους δίνεται από την σχέση

$$f_{Z,W}(x,y) = f_{|X|,|Y|}(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\varrho}}\times$$

$$\left\{ \exp\left[-\frac{1}{2(1-\rho^2)} \left(\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x-\mu_1)(y-\mu_2)}{\sigma_1 \sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2} \right) \right] \right.$$

$$+ \exp\left[-\frac{1}{2(1-\rho^2)} \left(\frac{(x+\mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x+\mu_1)(y+\mu_2)}{\sigma_1 \sigma_2} + \frac{(y+\mu_2)^2}{\sigma_2^2} \right) \right]$$

$$+ \exp\left[-\frac{1}{2(1-\rho^2)} \left(\frac{(x+\mu_1)^2}{\sigma_1^2} + 2\rho \frac{(x+\mu_1)(y-\mu_2)}{\sigma_1 \sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2} \right) \right]$$

$$\left. + \exp\left[-\frac{1}{2(1-\rho^2)} \left(\frac{(x-\mu_1)^2}{\sigma_1^2} + 2\rho \frac{(x-\mu_1)(y+\mu_2)}{\sigma_1 \sigma_2} + \frac{(y+\mu_2)^2}{\sigma_2^2} \right) \right] \right\}$$

όπου $x>0$ and $y>0$. (4.2.3)

Αν στην (4.2.3) θέσουμε $\mu_1=\mu_2=0$ παίρνουμε την συνάρτηση πυκνότητας της διμεταβλητής half normal κατανομής (βλέπε και Patil (1984).

Αν $\mu_1=\mu_2=0$ και $\sigma_1=\sigma_2=1$ τότε η (4.2.3) οδηγεί στην folded μορφή της διμεταβλητής standard normal κατανομής η οποία στην συνέχεια θ'αναφέρεται ως <u>διμεταβλητή folded standard normal</u> κατανομή της οποίας η συνάρτηση πυκνότητας θα δίνεται από την σχέση

$$f_{X,Y}(x,y)=\frac{1}{\tau\sqrt{1-\rho^2}}\left\{ \exp\left[-\frac{x^2+y^2-2\rho x y}{2(1-\rho^2)} \right] + \right.$$

$$\exp\left\{-\frac{x^2+y^2+2\rho x\,y}{2(1-\rho^2)}\right\}\Bigg\}$$

$$-1\le \rho \le 1 \qquad x,y \ge 0 \qquad (4.2.4)$$

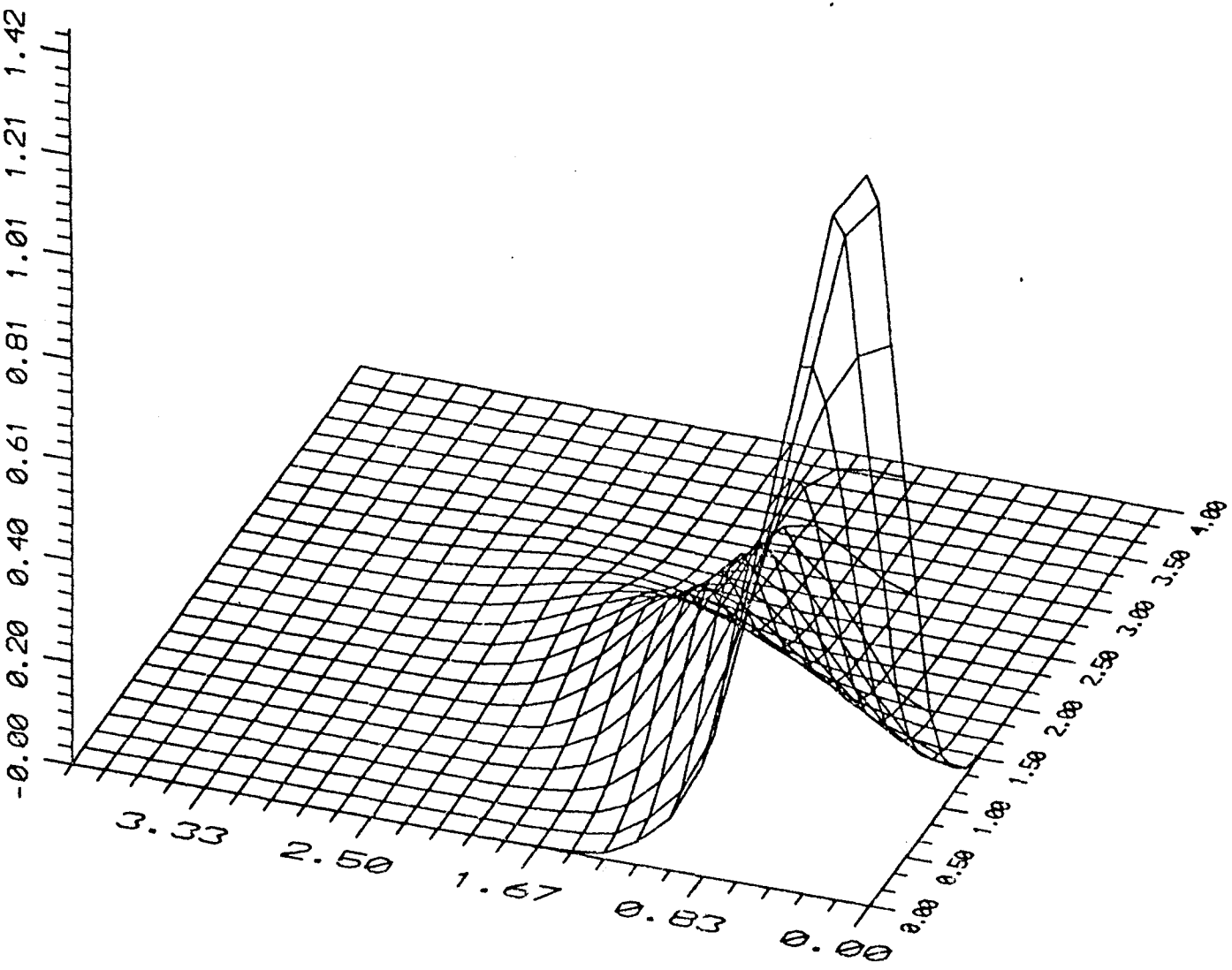
Στα σχήματα 5 και 6 δίνονται οι γραφικές παραστάσεις της συνάρτησης πυκνότητας διμεταβλητής folded standard normal για $\rho=0.2$ και 0.9.

Σχήμα 5

Συνάρτηση πυκνότητας πιθανότητας διπεταβλητής folded standard normal κατανομής για ρ=0.2

Σχήμα 6

Συνάρτηση πυκνότητας πιθανότητας
διμεταβλητής folded standard normal κατανομής για ρ=0.9



Για την διμεταβλητή folded standard normal κατανομή έχουμε ότι

$$E(X|Y) = |\rho|y$$

$$V(X|Y) = (1-\rho^2)$$

(Βλ. π.χ. Patil (1984) pp.34-35 όπου $\sigma_1=1$ και $\sigma_2=1$).

Ουσιαστικά η συνάρτηση πυκνότητας της διμεταβλητής folded standard normal προκύπτει από την διμεταβλητή half normal κατανομη θέτοντας $\sigma_1=1$ και $\sigma_2=1$ (Patil 1984). Αυτή η συνάρτηση πυκνοτητας έχει επίσης χρησιμοποιηθεί στο paper Abdel-Hameed and Allan R. Sampson (1978) .Εξαιτίας των παραπάνω δεν είναι αναγκαίο να αποδείξουμε ότι η διμεταβλητή folded standard normal κατανομή που δίνεται στην (4.2.4) είναι μία συνάρτηση πυκνότητας.

ΠΕΡΙΘΩΡΙΕΣ ΚΑΤΑΝΟΜΕΣ ΤΗΣ ΔΙΜΕΤΑΒΛΗΤΗΣ FOLDED STANDARD NORMAL ΚΑΤΑΝΟΜΗΣ

Προκειμένου να ορίσουμε τις περιθώριες κατανομές της διμεταβλητής folded standard normal χρειάζεται.να υπολογίσουμε το ακόλουθο ολοκλήρωμα:

$$\int_0^x f(x,y)\,dy = \frac{1}{\pi\sqrt{1-\rho^2}}\left[\int_0^x \exp\left\{-\frac{1}{2(1-\rho^2)}(x^2-2\rho xy+y^2)\right\}dy\right.$$

$$\left.+\int_0^x \exp\left\{-\frac{1}{2(1-\rho^2)}(x^2+2\rho x\,y+y^2)\right\}dy\right] =$$

$$= \frac{1}{\pi\sqrt{1-\varrho^2}} \exp\left(-\frac{x^2}{2(1-\varrho^2)}\right)\left[\underbrace{\int_0^{\infty} \exp\left\{-\frac{1}{2(1-\varrho^2)}(-2\varrho x\, y + y^2)\right\} dy}_{A}\right.$$

$$\left. + \underbrace{\int_0^{\infty} \exp\left\{-\frac{1}{2(1-\varrho^2)}(2\varrho x\, y + y^2)\right\} dy}_{B}\right] \qquad (4.2.5)$$

Είναι

$$A = \int_0^{\infty} \exp\left\{\left\{-\frac{y^2}{2(1-\varrho^2)}\right\} - \left\{-\frac{2\varrho x\, y}{2(1-\varrho^2)}\right\}\right\} dy$$

θέτοντας $\beta = \frac{1-\varrho^2}{2}$ και $\gamma = -\frac{\varrho x}{1-\varrho^2}$ θα έχουμε σύμφωνα με τους Gradshteyn & Ryzhik (1980) ολοκλήρωμα 3.222 (2)

$$A = \sqrt{\pi \frac{1-\varrho^2}{2}} \exp\left(\frac{(1-\varrho^2)\ \varrho^2 x^2}{2\ (1-\varrho^2)^2}\right)\left[1 - \Phi^*\left(-\frac{\varrho x}{1-\varrho^2}\sqrt{\frac{1-\varrho^2}{2}}\right)\right]$$

θέτοντας $\beta = \frac{1-\varrho^2}{2}$, $\gamma = \frac{\varrho x}{1-\varrho^2}$

έχουμε

$$B = \sqrt{\pi \frac{1-\varrho^2}{2}} \exp\left(\frac{(1-\varrho^2)\ \varrho^2 x^2}{2\ (1-\varrho^2)^2}\right)\left[1 - \Phi^*\left(\frac{\varrho x}{1-\varrho^2}\sqrt{\frac{1-\varrho^2}{2}}\right)\right]$$

Έτσι η (4.2.5) γίνεται

$$\int_0^x f(x\ ,y)\ dy =$$

$$\frac{1}{\pi\sqrt{1-\varrho^2}}\exp\left(-\frac{x^2}{2(1-\varrho^2)}\right)\sqrt{\pi\,\frac{1-\varrho^2}{2}}\ \exp\left(\frac{\varrho^2x^2}{2(1-\varrho^2)}\right)\times$$

$$\left[1-\Phi^*\left(-\frac{\varrho x}{1-\varrho^2}\sqrt{\frac{1-\varrho^2}{2}}\right)+1-\Phi^*\left(\frac{\varrho x}{1-\varrho^2}\sqrt{\frac{1-\varrho^2}{2}}\right)\right] \quad (4.2.6)$$

χάνοντας πράξεις στην (4.2.6) θα έχουμε:

$$\int_0^\infty f(x\cdot,y)\ dy= \frac{1}{\sqrt{2\pi}}\ \exp(-\frac{x^2}{2})\ \times$$

$$\left(2-\Phi^*\left(-\frac{\varrho x}{2\sqrt{1-\varrho^2}}\right)-\Phi^*\left(\frac{\varrho x}{2\sqrt{1-\varrho^2}}\right)\right) \quad (4.2.7)$$

όπου $\Phi^*(x)= \frac{2}{\sqrt{\pi}}\int_0^x \exp(-t^2)\ dt$

Έτσι

$$\Phi^*\left(-\frac{\varrho x}{2\sqrt{1-\varrho^2}}\right) = 2\left(\Phi\left(-\frac{\varrho x}{2\sqrt{2}\sqrt{1-\varrho^2}}\right)-\Phi\ (0)\right) \quad (4.2.8)$$

και

$$\Phi^*\left(\frac{\varrho x}{2\sqrt{1-\varrho^2}}\right) = 2\left[\Phi\left(\frac{\varrho x}{2\sqrt{2}\sqrt{1-\varrho^2}}\right) - \Phi(0)\right] \qquad (4.2.9)$$

Τελικά η (4.2.7) λόγω των (4.2.8),(4.2.9) θα γίνει

$$\int_0^{\infty} f(x,y)\,dy = \sqrt{\frac{2}{\pi}}\exp\left(-\frac{x^2}{2}\right)$$

που είναι η συνάρτηση πυκνότητας της folded standard normal κατανομής.

Ακολουθώντας την ίδια μέθοδο και για την μεταβλητή Y καταλήγουμε στο ότι θα ακολουθεί και αυτή την folded standard normal κατανομη.

Ετσι $E(X)=E(Y)=\sqrt{2/\pi}$ και $V(X)=V(Y)=1-2/\pi$

Κατά τον ίδιο τρόπο μπορεί κανείς εύκολα να δει ότι οι περιθώριες κατανομές της διμεταβλητής folded normal είναι οι folded normal κατανομές.

ΡΟΠΟΓΕΝΝΗΤΡΙΑ ΤΗΣ ΔΙΜΕΤΑΒΛΗΤΗΣ FOLDED STANDARD NORMAL ΚΑΤΑΝΟΜΗΣ

Για τον καθορισμό της m.g.f. ακολουθούμε τον Tallis M.G.(1961) Σύμφωμα με την εργασία αυτή η m.g.f m(t) της truncated multi-normal κατανομής δίνεται απο τον τύπο:

$$am=\exp^{T}\Phi_n(b_s,R) \qquad (4.2.10)$$

όπου $T=\frac{1}{2}t'Rt$ $\Phi_n(b_s,R)$ είναι η πολυμεταβλητή κανονική κατανομή με πίνακα συσχέτισης R και

$a=P(W_1>\alpha_1,W_2>\alpha_2,\ldots,W_n>\alpha_n)=\Phi_n(\alpha_s,R)$.

$b_s=\alpha_s-\zeta_s$ και $\zeta_s=\rho_{su}t_u$ $s=1,2,\ldots,n$.

Για την εύρεση της ροπογεννήτριας αυτής παρατηρούμε ότι

$$m(t)= a^{-1}\ {}^{(n)}\int_{\alpha_s}^{\infty} \exp(t'w)\ \varphi_n(w_s;R)\ dw_s =$$

$$=a^{-1}(2\pi)^{-n/2}|R|^{-1/2}\ {}^{(n)}\int_{\alpha_s}^{\infty} \exp\left[-\frac{1}{2}\{w'R^{-1}w-2t'w\}\right]dw_s \quad (4.2.11)$$

όπου t είναι διάνυσμα στήλη των t_s και w είναι το διάνυσμα στήλη των w_s και

${}^{(n)}\int(\ )\,dw_s$ είναι η συντομογραφία του

$$\int\int \cdots \int(\quad)dw_1dw_2\cdots dw_n \ .$$

Απο από την ταυτότητα

$$-\frac{1}{2}\{w'R^{-1}w-2t'w\}\equiv\frac{1}{2}t'Rt-\frac{1}{2}(w-\zeta)'R^{-1}(w-\zeta) \quad (4.2.12)$$

όπου $\zeta=Rt$

Έτσι η (4.2.11) λόγω της (4.2.12) γίνεται

$m(t)$ =

$$a^{-1}(2\pi)^{-n/2}|R|^{-1/2}e^{T}\ {}^{(n)}\int_{\alpha_s}^{\infty}\exp\{-\frac{1}{2}(w-\zeta)'R^{-1}(w-\zeta)dw_s \ .$$

όπου θέτοντας $X_s=W_s-\zeta_s$ έχουμε την (4.2.10).

Εφαρμόζοντας τα παραπάνω στην περίπτωση της διμεταβλητής folded standard normal θα έχουμε, θέτοντας $\alpha_s=0$ και $a=P(W>0,W>0)=1$

$$m(t)=2e^{T}\,\Phi_2(b_s;R)+2e^{T'}\,\Phi_2(b_s;R')$$

όπου $T=t_1^2+2\rho t_1 t_2+t_2^2$ $\quad T'=t_1^2-2\rho t_1 t_2+t_2^2$ $\quad b_1=-\xi_1=-\rho t_1-t_2$

$b_2=-t_1-\rho t_2$ $\quad R=\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}$ $\quad R'=\begin{pmatrix}1 & -\rho\\ -\rho & 1\end{pmatrix}$

ΣΥΝΑΡΤΗΣΗ ΚΑΤΑΝΟΜΗΣ ΤΗΣ ΔΙΜΕΤΑΒΛΗΤΗΣ FOLDED STANDARD NORMAL ΚΑΤΑΝΟΜΗΣ

Στην ενότητα αυτή γίνεται μία προσπάθεια προσδιορισμού της συνάρτησης κατανομής της διμεταβλητής folded standard normal κατανομής με αναφορά στα σχετικά αποτελέσματα του Tong (1990).

Στο βιβλίο αυτό δίνονται οι συναρτήσεις κατανομών της πολυμεταβλητής κανονικής και της πολυμεταβλητής t και ακολούθως γίνεται πινακοποίηση αυτών των κατανομών στην περίπτωση που οι τυχαίες μεταβλητές είναι exchangeable.

Συγκεκριμένα στην περίπτωση της πολυμεταβλητής κανονικής αποδεικνύεται το εξής θεώρημα

<u>θεώρημα</u>: (Tong(1990))

Οι παρακάτω προτάσεις είναι ισοδύναμες:

a) $X_1, \quad X_n$ είναι exchangeable κανονικές μεταβλητές.

b) $X=(X_1X_2\ldots X_n)$ ακολουθούν την $N(\mu_n,\Sigma)$ κατανομή όπου οι $X_1\ldots X_n$ έχουν κοινο μέσο μ, κοινή διασπορά σ^2 και κοινό συντελεστή συσχέτισης $\rho\in[0,1]$

c) Για $\rho=1$ X και $(\sigma Z_0+\mu,\ldots\sigma Z_0+\mu)$ έχουν την ίδια κατανομή όπου

Z_0 είναι N(0,1) τ.μ.. Για $\rho\in[0,1)$ η από κοινού συνάρτηση πυκνότητας της X θα δίνεται από την σχέση

$$f(x)=\int_{-\infty}^{\infty}\left[(\sigma\sqrt{1-\rho})^n\prod_{i=1}^{2}\varphi(u_i+\lambda z)\right]\varphi(z)dz$$

όπου $\varphi(z)$ είναι N(0,1) και

$$\lambda=\sqrt{\frac{\rho}{1-\rho}}\ , \qquad u_i=\frac{x_i-\mu}{\sigma\sqrt{1-\rho}} \quad (i=1,\ldots n).$$

c') Για $\rho\in[0,1)$ η από κοινού συνάρτηση πυκνότητας της X θα δίνεται από την σχέση

$$f(x)=c_n(\rho)\ \exp\left\{-\frac{1}{2}\left[\sum_{i=1}^{n}\frac{(x_i-\mu)^2}{\sigma^2(1-\rho)}-\frac{\rho\left(\sum_{i=1}^{n}(x_i-\mu)/\sigma\right)^2}{1+(n-1)\rho}\right]\right.$$

όπου $c_n(\rho)=(\sqrt{2\pi}\ \sigma)^{-n}(1-\rho)^{-(n-1)/2}[1+(n-1)\rho]^{-1/2}$.

d) $X=(X_1,\ldots X_n)$ και

$$(\ \sigma(\sqrt{1-\rho}\ Z_1+\sqrt{\rho}\ Z_0)+\mu,\ \ldots\ldots\sigma(\sqrt{1-\rho}\ Z_n+\sqrt{\rho}\ Z_0)+\mu\ )$$

έχουν την ίδια κατανομή όπου Z_0 Z_1 Z_n είναι τυπικές κανονικές τυχαίες μεταβλητές ,$\mu\in R$, $\sigma>0$ και $\rho\in[0,1]$.

Χρησιμοποιώντας το θεώρημα αυτό, με την προυπόθεση ότι οι τυχαίες μεταβλητές είναι exchangeable (ούτως ώστε να μπορεί να γίνει ο μετασχηματισμός που αναφέρεται στο παρατάνω θεώρημα

(a<=>d)) έχουμε το θεώρημα που ακολουθεί

Θεώρημα: (Tong(1990))

Αν $X_1, X_2, \ldots X_n$ είναι exchangeable κανονικές τυχαίες μεταβλητές με κοινό μέσο μ ,κοινή διασπορά σ^2 και συντελεστή συσχέτισης $\rho \in [0,1)$ η από κοινού συνάρτηση κατανομής των X_i. $i=1,\ldots,n$ δίνεται από την σχέση

$$F(x)=\int_{-\infty}^{\infty} \prod_{i=1}^{n} \Phi\left(\frac{\sqrt{\rho}\, z + a_i}{\sqrt{1-\rho}} \right) \varphi(z)\,dz$$

όπου $a_i=(x_i-\mu)/\sigma \quad (i=1,\ldots,n)$

και $\varphi(z)= \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad -\infty < z < \infty, \quad \Phi(z)= \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$

είναι η συνάρτηση πυκνότητας και η συνάρτηση κατανομής αντίστοιχα της N(0,1).

Για την συνάρτηση αυτή έχουν ήδη κατασκευαστεί από τον Tong πίνακες οι οποίοι δίνουν περιοχές της κατανομής .

Με παρόμοιο τρόπο μπορεί να βρεθεί η συνάρτηση κατανομής της πολυμεταβλητής folded normal. Αναλυτικά ισχύει ότι

Θεώρημα (Tong(1990)):

Αν $X_1 X_2 \ldots, X_n$ είναι exchangeable κανονικές μεταβλητές με κοινό μέσο 0 κοινή διασπορά σ^2 και κοινό συντελεστή συσχέτισης $\rho \in [0,1)$, η από κοινού συνάρτηση κατανομής της $|X|=(|X_1|,|X_2|\ldots|X_n|)$ δίνεται από την σχέση

$$F_{|X|}(x)=\int\prod_{i=1}^{n}\left[\Phi\left(\left(\frac{x_i}{\sigma}+\sqrt{\varrho}\ z\right)\Big/\sqrt{1-\varrho}\right)-\right.$$

$$\left.\Phi\left(\left(-\frac{x_i}{\sigma}+\sqrt{\varrho}\ z\right)\Big/\sqrt{1-\varrho}\right)\right]\varphi(z)\ dz$$

για την οποία κατασκευάζονται πίνακες οι οποίοι δίνουν περιοχές της κατανομής.

Με βάση τα μέχρι τώρα εκτεθέντα προσπαθούμε να βρούμε την συνάρτηση κατανομής της διμεταβλητής folded standard normal κατανομής.

Θα έχουμε:

$$F_{|X|}(x)=P(|X_1|\le x_1,\ |X_2|\le x_2)=$$

$$P(-x_1\le X_1\le x_1\ ,-x_2\le X_2\le x_2) \qquad (4\ell.13)$$

Ισχύει επίσης το εξής αποτέλεσμα:

Πόρισμα: (Tong(1990))

Εστω Z_0, Z_1, Z_2 ανεξάρτητες $N(0,1)$ τυχαίες μεταβλητές έστω

$\mu=(\mu_1,\mu_2)$ αυθαίρετο αλλά σταθερό και έστω $\Sigma=\begin{pmatrix}\sigma_1^2 & \sigma_{12}\\ \sigma_{12} & \sigma_2^2\end{pmatrix}$.

Εστω ότι για $|\varrho|<1$

$$X_1=\sigma_1(\sqrt{1-|\varrho|}\ Z_1+\sqrt{|\varrho|}\ Z_0)+\mu_1$$

$$X_2=\sigma_2(\sqrt{1-|\varrho|}\ Z_2+\ \delta_r\sqrt{|\varrho|}\ Z_0)+\mu_2 \qquad \text{οπου } \delta_r=\begin{cases} 1 & \text{αν } \varrho\geq 0 \\ -1 & \text{αν } \varrho<0 \end{cases}$$

Τότε το $(X_1,X_2)'$ θα ακολουθεί την διμεταβλητή κανονική κατανομή $N_2(\mu,\Sigma)$.

(Θα πρέπει να σημειωθεί ότι εδώ δεν απαιτείται οι τυχαίες μεταβλητές X_1,X_2 να είναι exchangeable).

Με βάση τον παραπάνω μετασχηματισμό η για $\mu=0$ θα γίνει:

$$F_{|X|}(x)= P\Big(\Big\{-x_1\leq\sigma_1(\sqrt{1-|\varrho|}\ Z_1+\sqrt{|\varrho|}\ Z_0)\leq x_1\Big\} \cap$$

$$\Big\{-x_2\leq\sigma_2(\sqrt{1-|\varrho|}\ Z_2+\delta_r\sqrt{|\varrho|}\ Z_0)\leq x_2\Big\}\Big) =$$

$$\int_{-\infty}^{+\infty} P\Big(\Big\{-\frac{x_1}{\sigma_1}-\sqrt{|\varrho|}z\leq\sqrt{1-|\varrho|}Z_1\leq\frac{x_1}{\sigma_1}-\sqrt{|\varrho|}\ z\Big\}\ \cap$$

$$\Big\{-\frac{x_2}{\sigma_2}-\delta_r\sqrt{|\varrho|}z\leq\sqrt{1-|\varrho|}Z_2\leq\frac{x_2}{\sigma_2}-\delta_r\sqrt{|\varrho|}\ z\Big\}\ \Big|\ Z_0=z\Big)\varphi(z)dz=$$

$$\int_{-\infty}^{x}\Big[\ \Phi\Big(\Big(\frac{x_1}{\sigma_1}-\sqrt{|\varrho|}z\Big)\Big/\sqrt{1-|\varrho|}\ \Big)-\Phi\Big(\Big(-\frac{x_1}{\sigma_1}-\sqrt{|\varrho|}z\Big)\Big/\sqrt{1-|\varrho|}\ \Big)\Big]\times$$

$$\left[\Phi\left(\left(\frac{x_2}{\sigma_2}-\delta_r\sqrt{|\rho|}z\right)\Big/\sqrt{1-|\rho|}\right)-\Phi\left(\left(-\frac{x_2}{\sigma_2}-\delta_r\sqrt{|\rho|}z\right)\Big/\sqrt{1-|\rho|}\right)\right]\varphi(z)dz \qquad (4.2.14)$$

4.2.14)

Η παραπάνω σχέση για $\sigma_i=1$ δίνει την συνάρτηση κατανομής της διμεταβλητής folded standard normal. Για την κατανομή αυτή είναι δυνατόν να χρησιμοποιηθούν οι πίνακες που δίνονται στον Tong(1990).

Τα παραπάνω ισχύουν γενικά για την διμεταβλητή περίπτωση (αφού στην περίπτωση αυτή δεν απαιτείται να είναι οι μεταβλητές exchangeable).

4.3 ΟΙ ΚΑΤΑΝΟΜΕΣ ΔΙΜΕΤΑΒΛΗΤΗ FOLDED T ΚΑΙ ΔΙΜΕΤΑΒΛΗΤΗ FOLDED STANDARD T

Στην περίπτωση που η σ_M^2 M= A,B δεν είναι γνωστή , ή το δείγμα που χρησιμοποιείται είναι μικρό η από κοινού κατανομή των

$$\left(\frac{|e_{t+1}(A)|}{\sigma_A}\ ,\ \frac{|e_{t+1}(B)|}{\sigma_B}\right)$$

(δηλ. η διμεταβλητή folded standard normal) δεν μπορεί να χρησιμοποιηθεί. Παρατηρώντας όμως ότι κάτω από την

$$H_0:\ E(|e_{t+1}(A)|) = E(|e_{t+1}(B)|)$$

ισχύει ότι

$$\sigma_A^2 = \sigma_B^2 = \sigma^2$$

μπορούμε να μελετήσουμε την παραπάνω από κοινού κατανομή όπως αυτή διαμορφώνεται αν το σ^2 αντικατασταθεί με μία εκτιμήτρια

του έστω S^2. Στην περίπτωση αυτή είναι προφανής η αναζήτηση ορισμού και μελέτης της από κοινού κατανομής των μεταβλητών $|X|/S$ και $|Y|/S$ όπου (X,Y) ακολουθεί την διμεταβλητή κανονική κατανομή με $\mu_x=\mu_y=0$, $\sigma_x^2=\sigma_y^2=1$, $\rho>0$ και S^2 είναι μία τυχαί μεταβλητή ανεξάρτητη των X,Y, τέτοια ώστε nS^2 ακολουθεί την χ_n^2 .

Την κατανομη αυτη ονομαζουμε <u>διμεταβλητη folded t κατανομή με n βαθμούς ελευθερίας</u>

Προκειμένου να δώσουμε τον μαθηματικό ορισμό της κατανομής αυτής να προσδιορίσουμε την συνάρτηση πυκνότητας πιθανότητας και να την μελετήσουμε κάνουμε χρήση της σχέσης που συνδέει την πολυμεταβλητή κανονική κατανομή με την πολυμεταβλητή κατανομή t.

Εστω $R=(\rho_{ij})$ να είναι ένας nxn συμμετρικός πίνακας τέτοιος ώστε να είναι είτε απόλυτα θετικός (positive definite είτε ημιαπόλυτα θετικός (positive semidefinite) και $\rho_{ii}=1$ $(i=1,...n)$.

Εστω οτι $Z=(Z_1,...Z_n)$ ακολουθεί την $N(0.R)$ κατανομή και S μία τυχαία μεταβλητή η οποία

(i) είναι ανεξάρτητη της Z και

(ii) $vS^2 \sim \chi^2(v)$ κατανομή.

Τότε, μια φυσική γενίκευση της Student μεταβλητής όπως είναι γνωστό (Tong (1990). Dunnet and Sobel (1954)) είναι η:

$$t=(t_1.t_2,....,t_n)=(Z_1/S.\ Z_2/S,....Z_n/S)$$

Για n=2 είναι εύκολο να αντικαταστήσουμε τα t_1,t_2 και ρ_{12} με το x,y,ρ και να πάρουμε την συνάρτηση πυκνότητας πιθανότητας της διμεταβλητής t κατανομής με n βαθμούς ελευθερίας η οποία

θα δίνεται από την σχεση (Dunnet and Sobel 1954):

$$f_{X,Y}(x,y)= \frac{1}{2\pi\sqrt{1-\rho^2}} \left[1+ \frac{x^2- 2\rho xy + y^2}{n(1-\rho^2)} \right]^{-\frac{n+2}{2}}$$

Για να βρούμε την folded μορφη της παραπάνω κατανομήςθα πρέπει να προσδιορίσουμε την απο κοινού κατανομή των τυχαίων μεταβλητών

$$|X|=|Z_1|/ S \ , \quad |Y|=|Z_2|/ S$$

όπου $|Z_1|, |Z_2|$ ακολουθούν την folded normal κατανομή.

Εχουμε:

$$F_{|x|,|y|}(x,y)= P(|X|\leq x, \ |Y|\leq y)= P(-x\leq X\leq x, \ -y\leq Y\leq y)=$$

$$F_{X,Y}(x,y)+F_{X,Y}(-x,-y)-F_{X,Y}(-x,y)-F_{X,Y}(x,-y) \quad x,y>0$$

Για τις αντίστοιχες συναρτήσεις πυκνότητας θα έχουμε:

$$f_{|x|,|y|}(x,y)=f_{X,Y}(x,y)+f_{X,Y}(-x,-y)+f_{X,Y}(-x,y)+f_{X,Y}(x,-y)$$

Η σχέση αυτή οδηγεί στον ακόλουθο ορισμό:

Ορισμός: Η από κοινού κατανομή δύο μη αρνητικών πραγματικών τυχαίων μεταβλητών

$$Z=|X| \text{ και } W=|Y|$$

θα λέμε ότι ειναι η διμεταβλητή folded t κατανομή με n βαθμούς ελευθερίας αν η συνάρτηση πυκνότητας δίνεται από την σχέση

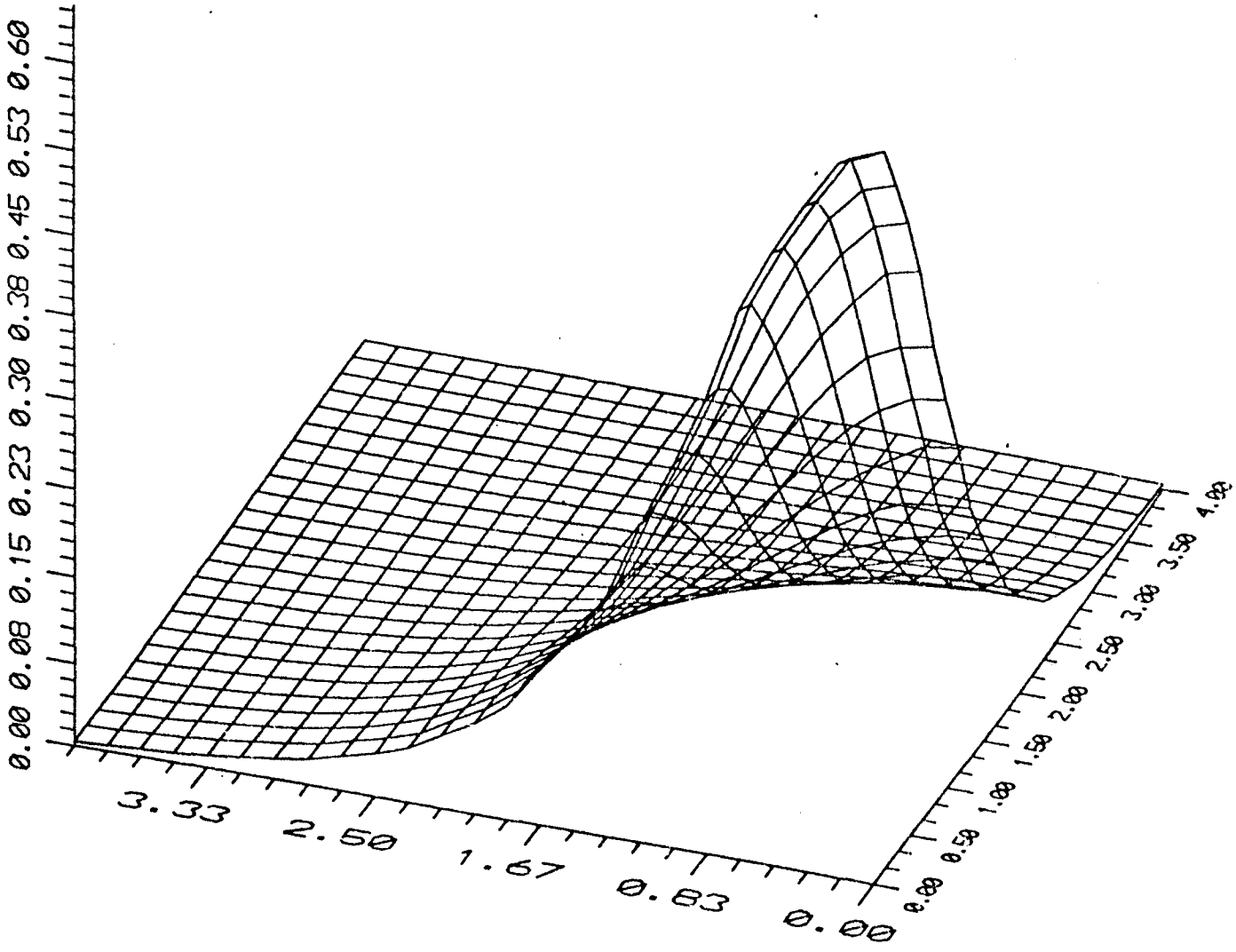
$$f_{Z,W}(x,y)= \frac{1}{\pi\sqrt{1-\rho^2}} \left(\left[1+ \frac{x^2- 2\rho xy + y^2}{n(1-\rho^2)} \right]^{-\frac{n+2}{2}} + \right.$$

$$+\left[1+\frac{x^2+2\rho xy+y^2}{n(1-\rho^2)}\right]^{-\frac{n+2}{2}}\Bigg\} \qquad x,y>0 \qquad (4.3.1)$$

Στα σχήματα που ακολουθούν 7,8,9 και 10 δίνονται οι γραφικες παραστάσεις της διμεταβλητής folded t κατανομής για α) n=4 , ρ=0.2 β) n=4 , ρ=0.9 γ) n=20 , ρ=0.2 και δ) n=20 , ρ=0.9.

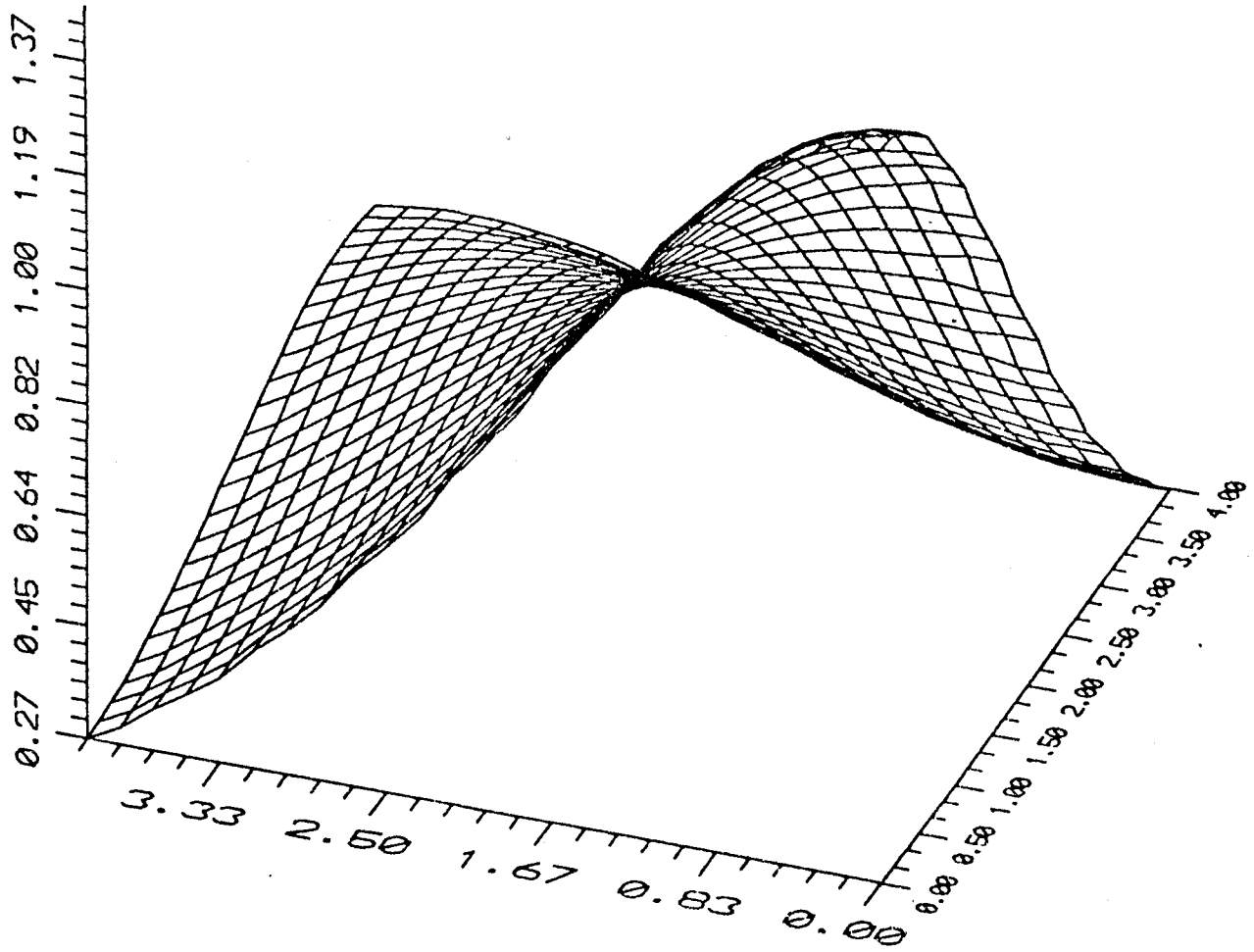
Σχήμα 7

Συνάρτηση πυκνότητας πιθανότητας
διμεταβλητής folded standard t κατανομής για n=4 ρ=0.2



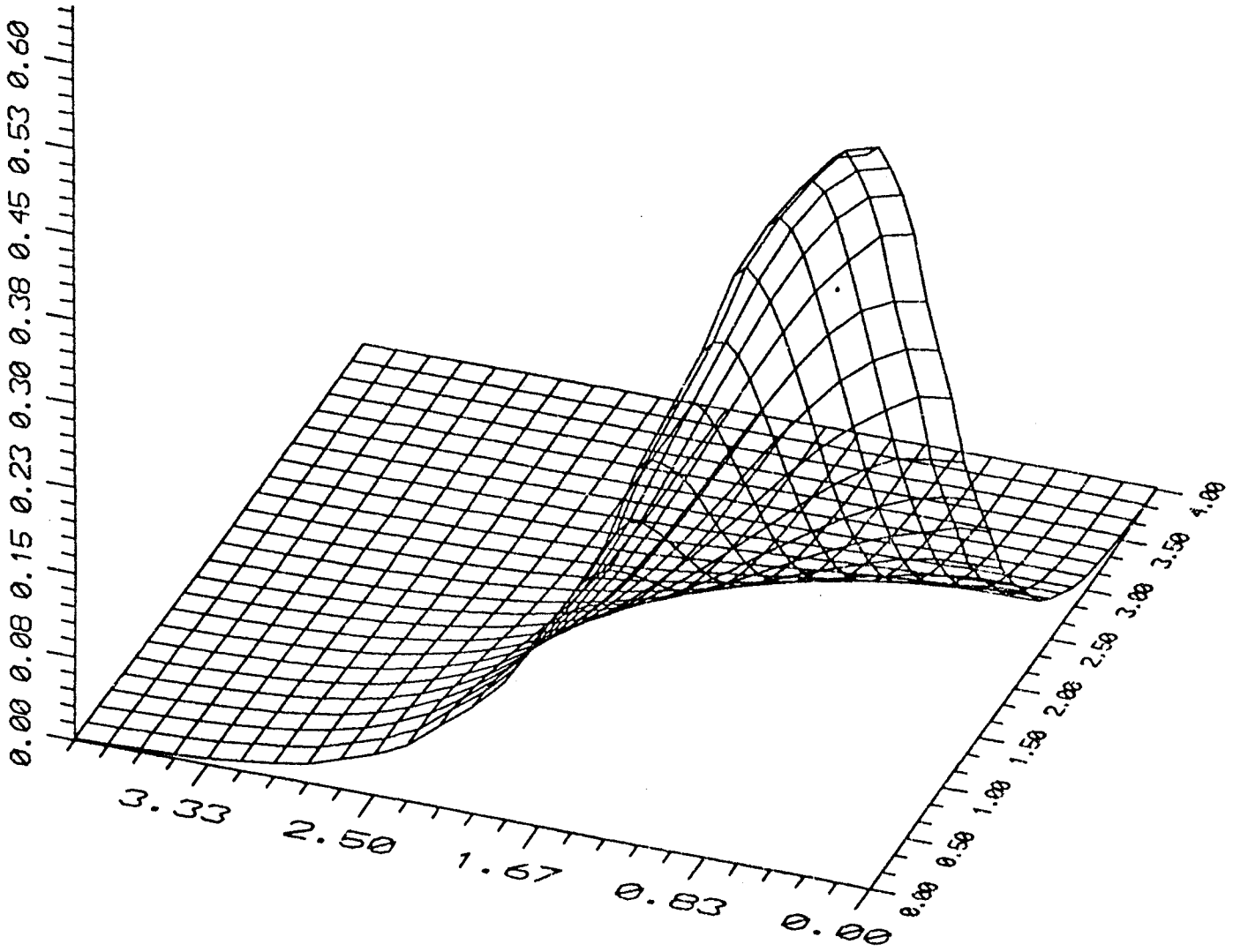
Σχήμα 8

Συνάρτηση πυκνότητας πιθανότητας
διμεταβλητής folded standard t κατανομής για n=4 ρ=0.9



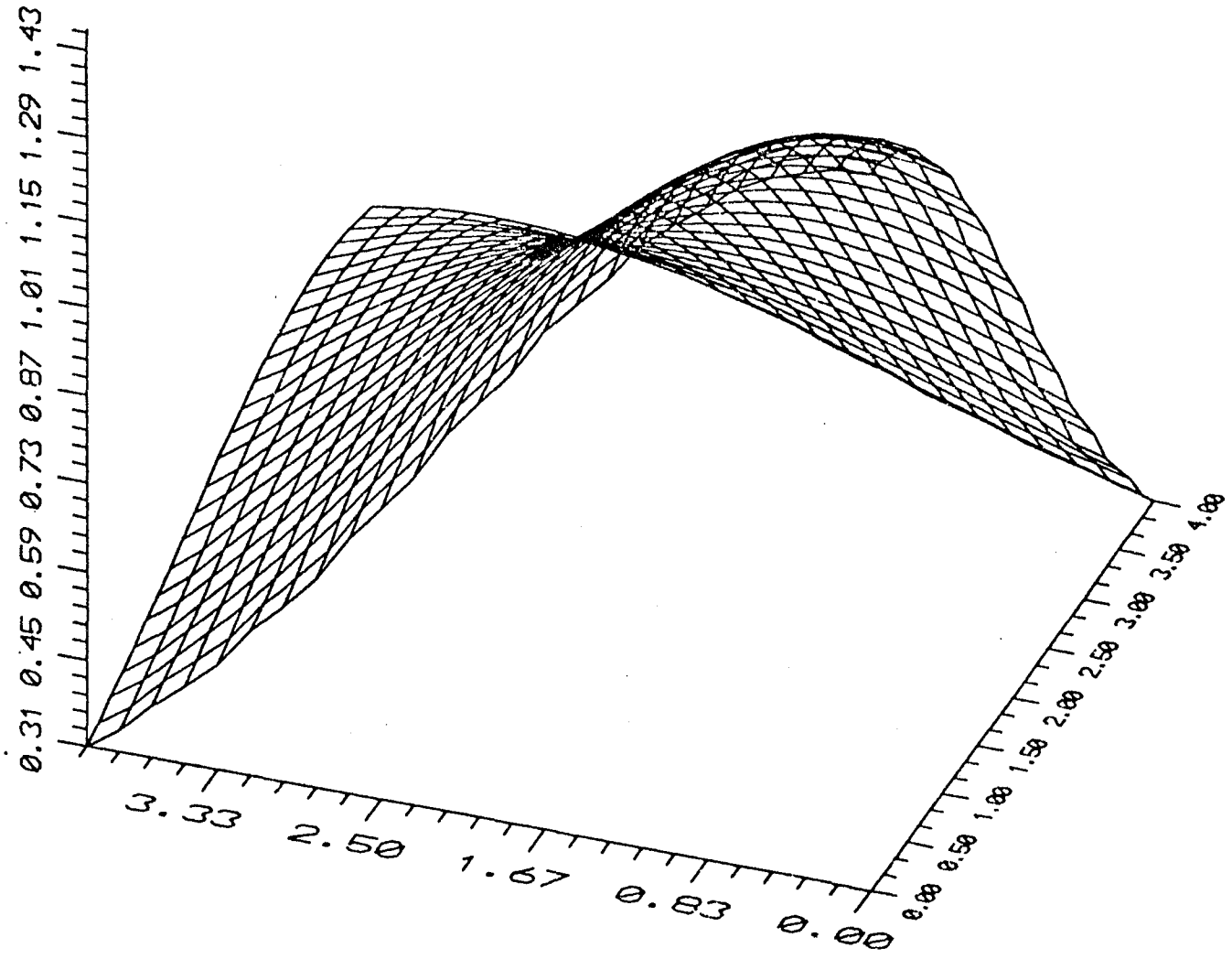
Σχήμα 9

Συνάρτηση πυκνότητας πιθανότητας
διμεταβλητής folded standard t κατανομής για n=20 ρ=0.2



Σχήμα 10

Συνάρτηση πυκνότητας πιθανότητας
διμεταβλητής folded standard t κατανομής για n=20 ρ=0.9



Θέτοντας ρ=0, ορίζουμε την διμεταβλητή folded standard t κατανομή με n βαθμούς ελευθερίας ως την κατανομή με συνάρτηση πυκνότητας πιθανότητας την

$$f_{X,Y}(x,y)= \frac{4}{2\pi} \left(1+ \frac{x^2+y^2}{n} \right)^{- \frac{n+2}{2}} \qquad x,y >0 \qquad (4.3.2)$$

Στην εργασία του F.E. Steffens (1969) αποδεικνύεται ότι:

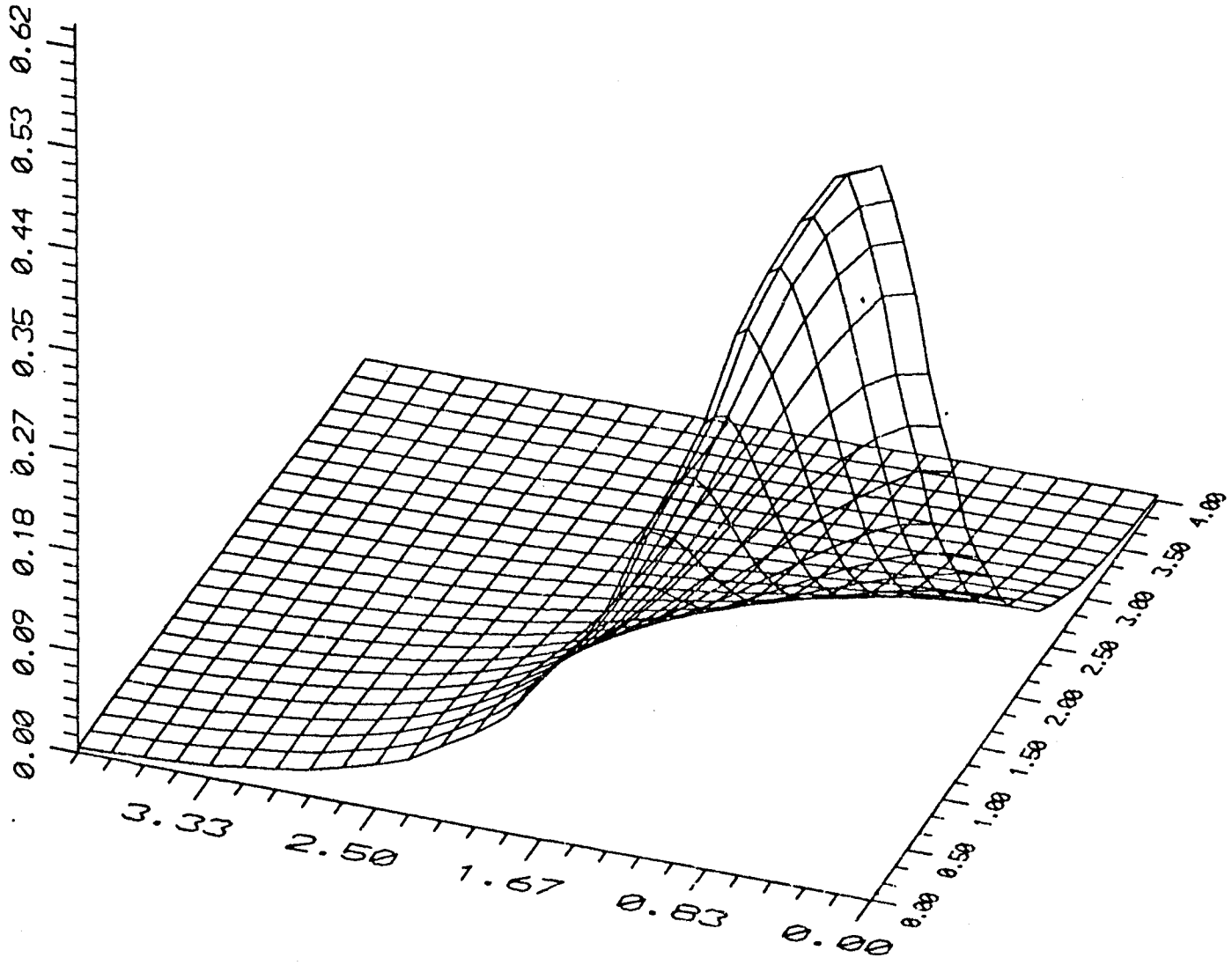
$$\int_0^{\infty} \int_0^{\infty} \frac{1}{2\pi} \left(1+ \frac{x^2+y^2}{n} \right)^{- \frac{n+2}{2}} = 1/4$$

επομένως η (4.3.2) είναι μία συνάρτηση πυκνότητας πιθανότητας.

Στα σχήματα 11,12 δίνονται οι γραφικες παραστάσεις της διμεταβλητής standard folded t για n=4 , n=20.

Σχήμα 11

Συνάρτηση πυκνότητας πιθανότητας
διμεταβλητής folded standard t κατανομής για n=4



Σχημα 12
Συνάρτηση πυκνότητας πιθανότητας
διμεταβλητής folded standard t κατανομής για n=20

ΣΧΕΣΗ ΤΗΣ ΔΙΜΕΤΑΒΛΗΤΗΣ FOLDED T ΜΕ ΤΗΝ ΔΙΜΕΤΑΒΛΗΤΗ FOLDED STANDARD NORMAL ΚΑΤΑΝΟΜΗ

Οπως είναι γνωστό για μεγάλους βαθμούς ελευθερίας η διμεταβλητή t συγκλίνει στην διμεταβλητή standard normal. Κάτι ανάλογο συμβαίνει και στην περίπτωση των folded κατανομών.

Πιο συγκεκριμένα όταν $n \longrightarrow \infty$ είναι γνωστό ότι ισχύει:

$$\lim_{n\longrightarrow\infty} \left[1+ \frac{x^2- 2\rho xy + y^2}{n(1-\rho^2)} \right]^{-\frac{n+2}{2}} =$$

$$\exp\left\{ - \frac{x^2+ y^2-2\rho x\, y}{2(1-\rho^2)} \right\} \quad (4.3.3)$$

και

$$\lim_{n\longrightarrow\infty} \left[1+ \frac{x^2+ 2\rho xy + y^2}{n(1-\rho^2)} \right]^{-\frac{n+2}{2}} =$$

$$\exp\left\{ - \frac{x^2+ y^2+2\rho x\, y}{2(1-\rho^2)} \right\} \quad (4.3.4)$$

Από τις (4.3.3).(4.3.4) είναι φανερό ότι η διμεταβλητή folded t κατανομή συγκλίνει στην διμεταβλητή folded standard normal κατανομή.

Επίσης επειδή

$$\lim_{n\to\infty}\left(1+\frac{x^2+y^2}{n}\right)^{-\frac{n+2}{2}}=\exp\left(-\frac{x^2+y^2}{2}\right)$$

η διμεταβλητή folded standard t θα συγκλίνει στην

$$\frac{2}{\pi}\exp\left(-\frac{x^2+y^2}{2}\right)$$

η οποία είναι η folded standard normal κατανομή για ρ=0. (Βλέπε και Y.L.Tong (1990))

ΣΧΕΣΕΙΣ ΤΗΣ ΔΙΜΕΤΑΒΛΗΤΗΣ FOLDED T ΜΕ ΑΛΛΕΣ ΚΑΤΑΝΟΜΕΣ

Η διμεταβλητή Cauchy κατανομή με παράμετρο c έχει συνάρτηση πυκνότητας πιθανότητας

$$f_{X,Y}(x,y)=\frac{1}{2\pi}\,c[c^2+x^2+y^2]^{-3/2} \qquad -\infty<x,y<\infty \qquad (4.3.5)$$

Έστω $F_{X,Y}(x,y)$ η αντίστοιχη συνάρτηση κατανομής των X,Y. Η από κοινού συνάρτηση κατανομής των τυχαίων μεταβλητών $|X|,|Y|$ θα είναι:

$$F_{|X|,|Y|}(x,y)=P(|X|\le x,\ |Y|\le y)=P(-x\le X\le x,\ -y\le Y\le y)=$$

$$F_{X,Y}(x,y)+F_{X,Y}(-x,-y)-F_{X,Y}(-x,y)-F_{X,Y}(x,-y) \quad x,y>0$$

Αν θεωρήσουμε τις αντίστοιχες συναρτήσεις πυκνότητας πιθανότητας θα έχουμε:

$$f_{|X|,|Y|}(x,y)=f_{X,Y}(x,y)+f_{X,Y}(-x,-y)+f_{X,Y}(-x.y)+f_{X,Y}(x,-y)$$

Η σχέση αυτή οδηγεί στον ακόλουθο ορισμό:

Ορισμός: Η από κοινού κατανομή δύο μη αρνητικών πραγματικών τυχαίων μεταβλητών $Z=|X|$ και $W=|Y|$ θα λέμε οτι είναι η διμεταβλητή half Cauchy κατανομή αν η από κοινού συνάρτηση πυκνότητας τους δίνεται από την σχέση

$$f_{Z,W}(x,y)=f_{|X|,|Y|}(x,y)$$

$$f_{Z,W}(x,y)=\frac{2}{\pi}\left(c\left(c^2+x^2+y^2\right)\right)^{-3/2} \qquad x,y>0 \qquad (4.3.6)$$

Επομενως η διμεταβλητή folded t για n=1 ρ=0 είναι η διμεταβλητή half Cauchy.

ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΤΩΝ ΠΕΡΙΘΩΡΙΩΝ ΚΑΤΑΝΟΜΩΝ ΤΗΣ ΔΙΜΕΤΑΒΛΗΤΗΣ FOLDED T ΚΑΙ ΤΗΣ ΔΙΜΕΤΑΒΛΗΤΗΣ FOLDED STANDARD T

Για την εύρεση των περιθωρίων αυτών κατανομών γίνεται χρήση της θεωρίας των elliptically contoured distributions.

Είναι γνωστό (Tong 1990 proposition 4.1.1) ότι οι περιθώριες κατανομες των elliptically contoured κατανομων είναι και αυτές elliptically contoured. Οπως είναι γνωστό σ'αυτήν την οικογένεια κατανομών ανήκουν η πολυμεταβλητή normal όπως επίσης η πολυμεταβλητή t και φυσικά οι αντίστοιχες διμεταβλητές κατανομές.

Με βάση την θεωρία η διμεταβλητη folded standard t

κατανομή που δίνεται από την σχέση

$$f_{X,Y}(x,y) = \frac{4}{2\pi}\left(1+\frac{x^2+y^2}{n}\right)^{-\frac{n+2}{2}} =$$

$$= 4\,\frac{\Gamma\left(\frac{n+2}{2}\right)}{(n\pi)^{2/2}\,\Gamma(n/2)}\left(1+\frac{x^2+y^2}{n}\right)^{-\frac{n+2}{2}} \qquad x,y>0$$

θα είναι elliptically contoured κατανομή. Επομένως από τις ιδιότητες της πολυμεταβλητής t Tong(1990) θα έχουμε ότι οι περιθώριες κατανομές θα προσδιορίζονται από τις σχέσεις της folded t κατανομής

$$f_x(x) = 2\,\frac{\Gamma\left(\frac{n+1}{2}\right)}{(n\pi)^{1/2}\,\Gamma(n/2)}\left(1+\frac{x^2}{2}\right)^{-\frac{n+1}{2}} \qquad x>0$$

και

$$f_y(x) = 2\,\frac{\Gamma\left(\frac{n+1}{2}\right)}{(n\pi)^{1/2}\,\Gamma(n/2)}\left(1+\frac{y^2}{2}\right)^{-\frac{n+1}{2}} \qquad y>0$$

Η ΣΥΝΑΡΤΗΣΗ ΚΑΤΑΝΟΜΗΣ ΤΗΣ ΔΙΜΕΤΑΒΛΗΤΗΣ FOLDED T

Ο Tong (1990) στο κεφάλαιο 9 της μονογραφίας του δίδει μια μορφή της συνάρτησης κατανομής της πολυμεταβλητής t ξεκινώντας από τον τύπο της συνάρτησης πυκνότητας όπως αυτή δόθηκε στην παράγραφο 3.3 (βλέπε ακόμη Dunett και Sobel (1955)) κάτω από την υπόθεση ότι οι τυχαίες μεταβλητές $t_1, t_2, \ldots t_n$ είναι

exchangeable.

Χρησιμοποιόντας την συνάρτηση κατανομής και κάνοντας αριθμητική ολοκλήρωση είναι δυνατόν να κατασκευάστούν πίνακες της πολυμεταβλητής t καθώς επίσης και πίνακες των μεταβλητών $|t_i|$.

Στην περίπτωση των δυο μεταβλητών θα μπορούσαμε, (υπο την προυπόθεση βέβαια ότι δεν απαιτείται να είναι exchangeable πράγμα το οποίο θα πρέπει να εξετάσουμε), να προχωρήσουμε στην εύρεση της συνάρτησης κατανομής της διμεταβλητης folded t. Για τον λόγο αυτόπροσπαθούμε να κατασκευάσουμε ένα μετασχηματισμό αντίστοιχο με αυτόν που χρησιμοποιεί ο Tong για την διμεταβλητή folded normal κατανομή.

θεωρούμε τον παρακάτω μετασχηματισμό:

$$t_1=(\sqrt{1-|\rho|}\ Z_1+\sqrt{|\rho|}\ Z_0)/\ S$$

και

$$t_2=(\sqrt{1-|\rho|}\ Z_2+\ \delta_r\sqrt{|\rho|}\ Z_0)/\ S \qquad \delta_r=\begin{cases} 1 & \text{αν } \rho\geq 0 \\ -1 & \text{αν } \rho<0 \end{cases}$$

όπου $Z_0\ Z_1\ Z_2$ ακολουθούν την $N(0,1)$ και νS^2 ακολουθεί την χ^2 με ν βαθμούς ελευθερίας .Τότε, με βάση το θεωρήμα 5.3.1 του Tong, στο οποίο ήδη αναφερθήκαμε παραπάνω (θεωρ, σελ.), και το γεγονός ότι η διμεταβλητή t προκύπτει από την $(Z_1/\ S, Z_2/\ S)$ έχουμε ότι (t_1,t_2) ακολουθεί την διμεταβλητή t κατανομή.

Σύμφωνα με τον Tong (1990) προσδιορίζουμε την αναλυτική μορφή της συνάρτησης κατανομής της διμεταβλητης folded t με ν βαθμούς ελευθερίας.

θα έχουμε

$$F_{|t_1|\ |t_2|}(\alpha,\alpha) =$$

$$P\left[\bigcap_{i=1}^{2}\{|t_i|\le\alpha\}\right]=P\left[(-\alpha\le t_1\le\alpha)\cap(-\alpha\le t_2\le\alpha)\right]=$$

$$EP\left[\left(Z_1\le\frac{\alpha s-\sqrt{|\varrho|}Z_0}{\sqrt{1-|\varrho|}}-Z_1\le\frac{-\alpha s-\sqrt{|\varrho|}Z_0}{\sqrt{1-|\varrho|}}\right)\cap\right.$$

$$\left.\left(Z_1\le\frac{\alpha s-\delta_r\sqrt{|\varrho|}Z_0}{\sqrt{1-|\varrho|}}-Z_1\le\frac{-\alpha s-\delta_r\sqrt{|\varrho|}Z_0}{\sqrt{1-|\varrho|}}\right)\ \middle|\ S=s\right]=$$

$$E\left\{EP\left[\left(Z_1\le\frac{\alpha s-\sqrt{|\varrho|}z}{\sqrt{1-|\varrho|}}-Z_1\le\frac{-\alpha s-\sqrt{|\varrho|}z}{\sqrt{1-|\varrho|}}\right)\times\right.\right.$$

$$\left.\left.\left(Z_1\le\frac{\alpha s-\delta_r\sqrt{|\varrho|}z}{\sqrt{1-|\varrho|}}-Z_1\le\frac{-\alpha s-\delta_r\sqrt{|\varrho|}z}{\sqrt{1-|\varrho|}}\right)\middle|(S,Z_0)=(s,z)\right]\middle|\ S=s\right\}=$$

$$\int_0^{\infty}\left\{\int_{-\infty}^{\infty}\left[\Phi\left(\frac{\alpha s-\sqrt{|\varrho|}z}{\sqrt{1-|\varrho|}}\right)-\Phi\left(\frac{-\alpha s-\sqrt{|\varrho|}z}{\sqrt{1-|\varrho|}}\right)\right]\times\right.$$

$$\left[\Phi\left(\frac{\alpha s-\delta_r\sqrt{|\varrho|}z}{\sqrt{1-|\varrho|}}\right)-\Phi\left(\frac{-\alpha s-\delta_r\sqrt{|\varrho|}z}{\sqrt{1-|\varrho|}}\right)\right]\varphi(z)dz\Bigg\}\ \psi_v(s)ds$$

όπου $\psi_v(s)$ είναι η συνάρτηση πυκνότητας της S.

Κάνοντας αλλαγή μεταβλητών στα ολοκλήρωματα και ειδικότερα θέτοντας:

$$u=(\sqrt{2}\ /\ 2)z \qquad \text{και} \qquad w=vs^2/\ 2$$

μετά από πράξεις θα έχουμε:

$$F_{|\varrho_1|\varrho_2|}(\alpha,\alpha)=\cancel{\gamma_2(v,\alpha)}=\int_0^\infty\left[\int_{-\infty}^{\infty}\frac{1}{\sqrt{\pi}\ \Gamma(v/2)}\left[\Phi\left(\sqrt{\frac{2|\varrho|}{1-|\varrho|}}\ u+\alpha\sqrt{\frac{2w}{v(1-|\varrho|)}}\right)-\right.\right.$$

$$\left.-\Phi\left(\sqrt{\frac{2|\varrho|}{1-|\varrho|}}\ u-\alpha\sqrt{\frac{2w}{v(1-|\varrho|)}}\right)\right]\times$$

$$\left[\Phi\left(\delta_r\sqrt{\frac{2|\varrho|}{1-|\varrho|}}\ u+\alpha\sqrt{\frac{2w}{v(1-|\varrho|)}}\right)-\right.$$

$$\left.\left.\Phi\left(\delta_r\sqrt{\frac{2|\varrho|}{1-|\varrho|}}\ u-\alpha\sqrt{\frac{2w}{v(1-|\varrho|)}}\right)\right]w^{v/2-1}\exp(-u\ )^2du\right]\exp(-w)dw$$

Το παραπάνω ολοκλήρωμα μπορεί να υπολογισθεί με αριθμητικές μεθόδους.

Στην περίπτωση κατά την οποία δεν ισχύουν τα παραπάνω θα πρέπει, για να εφαρμόσουμε την θεωρία του Tong του δίνεται στο κεφάλαιο 9 του βιβλίου του, να υποθέσουμε ότι οι τυχαίες

μεταβλητές είναι exchangeable ή ισοδύναμα ότι ο συντελεστής συσχέτισης ρε[0,1]. Αυτό γιατί σύμφωνα με το πρόβλημα 5.23 (Tong (1990) σελ 122) αναφέρεται οτι οι $t_1 t_2 \ldots t_n$ τυχαίες μεταβλητές ειναι exchangeable αν και μόνο αν ρε[0,1].

Η υπόθεση αυτή μπορεί να γίνει στα προβλήματα αξιολόγησης που μελετάμε με βάση βαθμολογικούς κανόνες, χωρίς να περιορίζεται ουσιαστικά η γενικότητα της μεθοδολογίας. Αυτό γιατί είναι φυσικό να μας ενδιαφέρει η σύγκριση μοντέλων με θετικά συσχετισμένη συμπεριφορά.

Εχοντας καταλήξει στον προσδιορισμό των διαφόρων μορφών της κατανομής της συνάρτησης στην οποία στηρίζεται ο κανόνας αξιολόγησης των εναλλακτικών μοντέλων και έχοντας διαθέσιμους τους πίνακες των περιοχών πιθανότητας της κατανομής αυτής στις διαφορες περιπτώσεις είναι δυνατόν να προχωρήσουμε σε στατιστική συμπερασματολογία για συγκεκριμένα πρακτικά προβλήματα.

ΚΕΦΑΛΑΙΟ 5

ΑΛΛΕΣ ΜΕΘΟΔΟΙ ΑΞΙΟΛΟΓΗΣΗΣ ΤΗΣ ΠΡΟΒΛΕΠΤΙΚΗΣ ΙΚΑΝΟΤΗΤΑΣ ΓΡΑΜΜΙΚΩΝ ΜΟΝΤΕΛΩΝ

5.1 ΕΙΣΑΓΩΓΗ

Στο κεφάλαιο αυτό περιγράφονται δύο ακόμα μέθοδοι αξιολόγησης γραμμικών μοντέλων. Η πρώτη είναι μη παραμετρική και αναφέρεται στην αξιολόγηση ενός μοντέλου ενώ η δεύτερη αξιολογεί συγκριτικά δύο εναλλακτικά γραμμικά μοντέλα.

5.2 ΜΙΑ ΜΗ ΠΑΡΑΜΕΤΡΙΚΗ ΜΕΘΟΔΟΣ ΑΞΙΟΛΟΓΗΣΗΣ ΤΗΣ ΠΡΟΒΛΕΠΤΙΚΗΣ ΙΚΑΝΟΤΗΤΑΣ ΕΝΟΣ ΓΡΑΜΜΙΚΟΥ ΜΟΝΤΕΛΟΥ ΒΑΣΙΖΟΜΕΝΗ ΣΤΟΝ ΕΛΕΓΧΟ ΣΗΜΑΤΟΣ ΤΗΣ ΔΙΩΝΥΜΙΚΗΣ ΚΑΤΑΝΟΜΗΣ

Στην ενότητα αυτή εισάγεται μία μη παραμετρική μέθοδος αξιολόγησης ενός γραμμικού μοντέλου όπως αυτό ορίσθηκε στο κεφάλαιο 2 ο οποίος βασίζεται στον έλεγχο σήματος της διωνυμικής κατανομής.

Θεωρούμε το 100(1-α)% διάστημα εμπιστοσύνης για την τιμή της εξαρτημένης μεταβλητής την χρονική στιγμή t+1 (Y^{0}_{t+1}) με την τιμή $\hat{Y}^{0}_{t+1}$ που προβλέπεται από το μοντέλο.

Ορίζουμε ένα βαθμολογικό κανόνα ο οποίος "βαθμολογεί" την απόδοση του μοντέλου με την τιμή 1 ή 0 ανάλογα με το αν η τιμή Y^{0}_{t+1} (όταν παρατηρηθεί) ανήκει ή δεν ανήκει, αντίστοιχα, στο διάστημα εμπιστοσύνης που κατασκευάζουμε.

Δηλαδή αν Z_t είναι ο "βαθμος" του μοντέλου την χρονική στιγμή

t έχουμε

$$Z_i = \begin{cases} 1 \text{ αν } \hat{Y}^{0}_{t+1} \in \hat{Y}^{0}_{t+1} \pm t_{l_t-m,\,1-a/2}\, S_{\hat{Y}^{0}_{t+1}} \\ 0 \quad \text{διαφορετικά} \end{cases}$$

όπου $S_{\hat{Y}^{0}_{t+1}} = S_t \sqrt{1+X^{0}_{t+1}\,(X_t^{'}X_t)^{-1}\,X^{0'}_{t+1}}$

Θεωρούμε επομένως n ανεξάρτητες δοκιμές $(Z_1, Z_2, \ldots, Z_n)$, όσες είναι και ο αριθμός των φορών που έχουμε παρατηρήσεις για το μοντέλο.

Εστω T ο αριθμός των επιτυχιών (T=# των $\{Z_i=1\}$)με πιθανότητα επιτυχίας $P(Z_i=1) = p$.

Υπό τις παραπάνω προυποθέσεις η τυχαία μεταβλητή T ακολουθεί την διωνυμική κατανομή με παραμέτρους n και p.

Ενας στατιστικός έλεγχος που οδηγεί στην αξιολόγηση ενός μοντέλου είναι ο:

$$H_0 : p \le 1/2 \qquad (5.2.1)$$
$$H_1 : p > 1/2$$

Κάτω από την μηδενική υπόθεση η T θ'ακολουθεί την διωνυμική κατανομή με παραμέτρους n και p=1/2.

Σύμφωνα με τον έλεγχο αυτό απορρίπτουμε την μηδενική σε επίπεδο σημαντικότητας a αν η τιμή της T ειναι μεγαλύτερη από μία τιμή t έτσι ώστε να ισχύει:

$$P(T>t \;/\; n,\; p=1/2)=a$$

ή ισοδύναμα αν

$P(T \leq t \ / \ n,\ p=1/2)=1-\alpha.$

Με δεδομένο το σύνολο δοκιμών n και το επίπεδο σημαντικότητας α όπως είναι γνωστό (Conover (1986)) η κρίσιμη τιμή της t για την περίπτωση μεγάλου δείγματος προσεγγίζεται απο τον τύπο:

$$t=1/2(n+w_{1-\alpha/2}\sqrt{n}) \qquad (5.2.2)$$

όπου $w_{1-\alpha/2}$ δίνεται από τους πίνακες τιμών της κανονικής κατανομής.

Για μικρο δείγμα η κρίσιμη τιμή του t δίνεται από τους πίνακες τιμών της διωνυμικής κατανομής.

Προκειμένου να εφαρμόσουμε την μέθοδο αυτή σε πραγματικά δεδομένα χρησιμοποιούμε τα δεδομένα Xekalaki-Katti (1984) για την παραγωγή καλαμποκιού στις περιοχές CRD10-CRD100 της πολιτείας της Ινδιάνα των ΗΠΑ για τα έτη 1963-1980.

Τα σχετικά αποτελέσματα δίνονται δίνονται στο παράρτημα IV. Απο τους πίνακες γίνεται φανερό ότι και τα δύο μοντέλα A και B που χρησιμοποιήθηκαν από το USDA (United States Department of Agriculture) έχουν ικανοποιητική προβλεπτική ικανότητα.

5.3 ΜΕΘΟΔΟΣ ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΤΙΚΗ ΑΞΙΟΛΟΓΗΣΗ ΔΥΟ ΕΝΑΛΛΑΚΤΙΚΩΝ ΓΡΑΜΜΙΚΩΝ ΜΟΝΤΕΛΩΝ ΠΟΥ ΒΑΣΙΖΕΤΑΙ ΣΤΟΝ ΕΛΕΓΧΟ ΤΟΥ ΣΥΝΤΕΛΕΣΤΗ ΣΥΣΧΕΤΙΣΗΣ

Στην ενότητα αυτή δίνεται ένας άλλος τρόπος αξιολόγησης μοντέλων που στηρίζεται στην απόλυτη τιμή της διαφοράς της προβλεπόμενης από την παρατηρούμενη τιμή για ένα μοντέλο την χρονική στιγμή t. Στην περίπτωση αυτή το καλύτερο μοντέλο θα

θεωρηθεί εκείνο που έχει τον μικρότερο τελικό βαθμό.

Θεωρούμε δυο γραμμικά μοντέλα A και B της μορφής:

$$Y_t = X_t(M)\ \beta_M + \varepsilon_t(M) \quad (M=A,B)$$

όπου όπως έχουμε αναφέρει στο κεφάλαιο 1,

Y_t είναι διάνυσμα διαστάσεως $l_t \times 1$ των εξαρτημένων μεταβλητών.

$X_t(M)$ πίνακας διαστάσεως $l_t \times m_M$ $(l_t \geq m_M | X_t(M)' X_t(M)| \neq 0)$ γνωστών συντελεστών.

β_M διάνυσμα διαστάσεως $m_M \times 1$ συντελεστών παλινδρόμησης και

$\varepsilon_t(M)$ διάνυσμα διαστάσεως $l_t \times 1$ τυχαίων μεταβλητών που ακολουθούν την κανονική κατανομή με $E(\varepsilon_t(M))=0$ και $V(\varepsilon_t(M))=\sigma^2(M)\ I_t (\sigma^2(M) < \infty)$ και I_t είναι ο $l_t \times l_t$ ταυτοτικός πίνακας.

Μία πρόβλεψη για την εξαρτημένη τυχαία μεταβλητή για την χρονική στιγμή t+1 θα δίνεται από την στατιστική συνάρτηση:

$$\hat{Y}^{0}_{t+1}(M) = X^{0}_{t+1}(M)\ \hat{\beta}_t(M)$$

οπου $\hat{\beta}_t(M)$ είναι η εκτιμήτρια ελαχίστων τετραγώνων του β_M την χρονική στιγμή t, και $X^{0}_{t+1}(M)$ είναι ένα $l_t \times m_M$ διάνυσμα των τιμών των συντελεστών παλινδρόμησης την στιγμή t+1. Εστω $Y^{0}_{t+1}(M)$ είναι η παρατηρούμενη τιμή της εξαρτημένης τυχαίας μεταβλητής την στιγμή t+1. Τότε για δύο μοντέλα A και B τα οποία θέλουμε να συγκρίνουμε θα ισχύει ότι η τυχαία μεταβλητή

$$\hat{Y}^{0}_{t+1}(M) - Y^{0}_{t+1}(M) = e_{t+1}(M) \qquad (5.3.1)$$

θ'ακολουθεί την κανονική κατανομή με μέση τιμή 0 και διασπορά σ_M^2 όπου

$$\sigma_M^2 = \sigma^2(M)\left[1 + X^{o}_{t+1}(M)\,(X'_t(M)X_t(M))^{-1}X'^{o}_{t+1}(M)\right] \quad \text{με } M=A \text{ ή } M=B \qquad (5.3.2)$$

Έτσι η απόλυτη τιμή της (5.3.1) η οποία είναι η

$$|\hat{Y}^{o}_{t+1}(M) - Y^{o}_{t+1}(M)| = |e_{t+1}(M)|$$

θα ακολουθεί την folded normal κατανομή με

μέση τιμή $\mu_f = \sqrt{2/\pi}\,\sigma_M$ και διασπορά $\sigma_f^2 = \sigma_M^2(1-2/\pi)$

(Leone et.al. 1961).

Με άλλα λόγια:

$$E(|e_t(A)|) = \sqrt{2/\pi}\,\sigma_A \quad \text{και}$$
$$E(|e_t(B)|) = \sqrt{2/\pi}\,\sigma_B \qquad (5.3.3)$$

Ενας έλεγχος υποθέσεων που θα μπορούσε να γίνει είναι:

$$H_0: E(|e_t(A)|) = E(|e_t(B)|)$$
$$H_1: E(|e_t(A)|) < E(|e_t(B)|) \qquad (5.3.4)$$

Εξ'αιτίας της (5.3.3) η (5.3.4) είναι ισοδύναμη με

$$H_0: \sigma_A^2 = \sigma_B^2$$
$$H_1: \sigma_B^2 < \sigma_B^2 \qquad (5.3.5)$$

Σ'αυτό το σημείο θα πρέπει να σημειώσουμε ότι

$$E\left\{ (|e_t(A)|+|e_t(B)|) \ , \ (|e_t(A)|-|e_t(B)|)\right\} =$$

$$E\ (e_t^2(A) - e_t^2(B)) = \sigma_A^2 - \sigma_B^2 \qquad (5.3.6)$$

Αν θέσουμε

$$R_{t+} = |e_t(A)|+|e_t(B)| \text{ και } R_{t-} = |e_t(A)|-|e_t(B)| \qquad (5.3.7)$$

θα έχουμε

$$Cov.(R_{t+}, R_{t-}) = E\left[(|e_t(A)|+|e_t(B)|)(|e_t(A)|-|e_t(B)|)\right] -$$

$$- E\left[(|e_t(A)|+|e_t(B)|)\right] E\left[(|e_t(A)|-|e_t(B)|)\right] \qquad (5.3.8)$$

Εξ'αιτίας των (5.3.3),(5.3.6) και (5.3.7) η (5.3.8) θα γίνει

$$Cov.(R_{t+}, R_{t-}) = \sigma_A^2 - \sigma_B^2 -$$

$$-\left[\sqrt{2/\pi}\ \sigma_A+\sqrt{2/\pi}\ \sigma_B\right]\left[\sqrt{2/\pi}\ \sigma_A-\sqrt{2/\pi}\ \sigma_B\right]$$

και επομένως

$$Cov.(R_{t+}, R_{t-}) = (\sigma_A^2 - \sigma_B^2)\left(1 - \frac{2}{\pi}\right) \qquad (5.3.9)$$

Λόγω της (5.3.9) ο έλεγχος υποθέσεων (5.3.5) θα είναι ισοδύναμη με τον έλεγχο

$$H_0: Cov.(R_{t+}, R_{t-}) = 0$$

$$H_1: Cov.(R_{t+}, R_{t-}) < 0 \qquad (5.3.10)$$

ο οποίος είναι ισοδύναμος με τον

$$H_0: \rho_{R_{t+}\,R_{t-}} = 0$$
$$H_1: \rho_{R_{t+}\,R_{t-}} < 0 \qquad (5.3.11)$$

Μία στατιστική συνάρτηση που θα μπορούσε να χρησιμοποιηθεί για τον παραπάνω έλεγχο είναι η συνάρτηση

$$\tilde{X} = \frac{\left[\frac{\sum R_{t+} R_{t-}}{n} - \overline{R_{t+} R_{t-}} \right]}{\sqrt{\frac{\sum \left(R_{t+} - \overline{R_{t+}} \right)^2}{n} \; \frac{\sum \left(R_{t-} - \overline{R_{t-}} \right)^2}{n}}}$$

$$\text{οπου} \quad \overline{R}_{t+} = \left(\sum_{t=1}^{n} |e_t(A)| + |e_t(B)| \right) / n$$

Κάτω από την μηδενική υπόθεση η ασυμπτωτική κατανομή του $\sqrt{n}\ \tilde{R}$ είναι η κανονική με μέση τιμή 0 και διασπορά $Var(R_{t+}\,R_{t-}) / Var(R_{t+})Var(R_{t-})$ (Lehmann (1966).

Ετσι, τιμές του $Z_R = \sqrt{n}\ \tilde{R}$ στην αριστερή ουρά της κανονικής κατανομής θα οδηγούν σε απόρριψη της μηδενικής υπόθεσης. Αυτό συνεπάγεται ότι το μοντέλο Α περιγράφει καλύτερα τα υπό ανάλυση δεδομένα από το μοντέλο Β. Στο παράρτημα V γίνεται εφαρμογή της παραπάνω μεθόδου πάνω σε δεδομένα από την συγκομιδή καλαμποκιού στις περιοχές CRD10(Crop Reporting District 10) έως και CRD100(Crop Reporting District 100) της πολιτείας της Ινδιάνας των ΗΠΑ για τα έτη 1963-1980. Τα στοιχεία αυτά περιέχονται στην εργασία των Xekalaki-Katti (1984).

ΒΙΒΛΙΟΓΡΑΦΙΑ

Akaike,H. (1973) "Information theory and an extension of the maximum likelihood principle." *Proc. 2nd Inter. Symposium on Information Theory 267-281 Budapest*

Akaike,H. (1974) "A new look at the statistical model identification." *IEEE Transaction on Automatic Control 19(6) 716-723*

Akaike,H. (1978) "A bayesian analysis of the minimum AIC procedure" *Ann. Inst. Statist. Math. 30A 9-14*

Akaike,H. (1979 "A bayesian extension of the minimum AIC procedure of autoregressive model fitting." *Biometrika 66 237-242*

Allen,D.M. (1971) "Mean square error of preiction as a criterion for selecting variables." *Technometrics 16 125-127*

Amemiya,T. (1980) " Selection of Regressors." *International Economic Review 21(2) 331-354*

Anderson, T.W. Fang,K.T. (1990) "Statistical inference in elliptically contoured and related distributions" *Academic Press*

Atkinson,A. (1969) "A test for discriminating between models." *Biometrika 56(2) 337-347*

Breiman,L. Freedman, D.(1983)" How many variables should be entered in a regression equation?" *JASA 78(381) 131-136*

Breiman,L. Spector,P. (1992) "Submodel selection and evaluation in regression. The X-random case." *International Statistical Review 60(3) 291-319*

Brown,R.L. Durbin,J. Evans,J.M. (1975) "Techniques for testing the constancy of regression relationships over time." *J.R.S.S. B 24*

149-192

Conover,W.(1980) "Practical nonparametric statistics" *John Wiley and Sons. 2nd edition*

Cooper,B.E. (1968) "The integral of Student's t-distribution" *J.R.S.S. Algorithm AS 3*

Copas,J.B. (1983) "Regression prediction and shrinkage" *J.R.S.S.B 45(3 311-354*

Chow,G.C. (1981) "A comparison of the information and posterior probability criterion for model selection." *J. of econometrics 16 21-33*

Daniel, C. (1959) "Use of half normal plots in interpreting factorial two level experiments." *Technometrics1(4) 311-341*

Draper,N.R. Craig Van Nostrand (1979) :" Ridge regression and James-Stein Estimation: Review and comments." *Technometrics 21(4 451-466*

Dunnett,C. and Sobel,M. (1955) "Approximations to the probability integral and certain percentage point of a multivariate analogue of students t distribution." *Biometrika 258-26*

Edwards,D. Havranek,T. (1987) "A fast model selection procedures for large families of models." *JASA 82(397) 205-213*

Elandt,R.C. (1961) "The folded normal distribution:Two methods of estimating parameters from moments." *Technometrics 3(4) 551-562*

Forsytne,A.B.Engleman,L. Jennrich,R. May,P.R.A.(1973) "A stopping rul for variable selection in multiple regression." *JASA 68(341) 75-*

Gaush,H.G. (1988) "Models selection and validation for yield trials with interaction." *Biometrics 44 705-715*

Geisser,S. (1975) "The predictive sample reuse method with applications." *JASA 70(350) 320-328*

Geisser,S. Eddy,W. (1979) "A predictive approach to model selection." *JASA 74(365) 153-160*

Gradshteyn,I.S. Ryzhik,I.M. (1980) "Table of Integrals, series, and products." *4th edition Academic Press.*

Griffiths.P. Hill,I.D. (1985) "Applied Statistics algorithms" *Ellis Horwood Limited Chichester*

Hameed,A. Sampson,A.R. (1978) "Positive dependence of the bivariate and trivariate absolute normal t,x^2 and F distributions." *Annals of Statistics 6(6) 1360-1368*

Hamilton,M.A.(1991) "Model validation:An annotated bibliography." *Comm. Statist. Theory. Meth. 20(7) 2207-2266*

Hannan,E.J. Quinn,B.G. (1979) "The determination of the order of an autoregression" *J.R.S.S. B 41(2) 190-195*

Hjorth,U. (1989) "On model selection in the computer age" *J of Statistical Planning and inference 23 101-115*

Hocking,R.R. (1976) "The analysis and selection of variables in linear regression" *Biometrics 32 1-49*

Hocking,R.R. (1983) "Developments in linear regression methodology :1959-1982" *Technometrics 25(3) 219-249*

Hoerl,A.E. Kennard,R.W.(1970a) " Ridge Regression: Biased estimation for non-orthogonal polynomials." *Technometrics 12 55-68*

Hoerl,A.E. Kennard,R.W.(1970a) " Ridge Regression: Application to non-orthogonal polynomials." *Technometrics 12 69-72*

Holmes,J. Hutton,P. (1989) "Optimal model selection when the true relationship is weak and occurs with adelay." *Economics Letters 28 47-51*

Hurvich,C. C.Ling Tsai(1989) "Regression and time series model selection in small samples." *Biometrika 76(2) 297-307*

Jacob,J. Shiryaev,A..N.(1987) "Limit theorems for stochastic processes" *Springer-Verlag*

Jensen, D.R. (1970a) "The joint distribution of traces of Wishart matrices and some applications." *Ann. Math. Statist. 41 133-145*

Jensen, D.R. (1970b) "A generalization of the multivariate Rayleigh distribution." *Sankhya Ser. A 32 193-208*

Johnson,N.L. (1962) "The folded normal distribution:Accuracy of estimation by maximum likelihood." *Technometrics 4(2) 249-256*

Judge,G. Yi, G. Yansey,T. Terasvirta,T.(1987) "The extended Stein procedure for simultaneous model selection and parameter estimation" *Journal of Econometrics 35 375-391*

Kendall,M. Stuart,A. Ord,J.K.(1983) "The advanced theory os statistics." *Vol. # (4th edition) Griffin London*

Keramidas,E. Lee,J. (1990) "Forecasting technological substitutions with concurrent short time series ." *JASA 85(411) 625-632*

Kibble,W.F. (1941) "A two variate gamma type distribution" *Sankya 5 137-150*

Klein,R.W. Brown,S.J. (1984) "Model selection when there is "minimal prior information." *Econometrika 52(5) 1291-1312*

Kobayashi,M. Sakata,S. (199) "Mallows C_p criterion and unbiasedness of model selection." *J. of Econometrics 45 385-395*

Kotlarski,I. (1964) " On bivariate random variables where the quotient of their coordinates follows some known distribution." *Ann. of. Math. Stat. 1673-1684*

Kotlarski,I. (1966) " On characterizing the normal distribution by student's law." *Biometrika 53 603-606*

Kullback,S. Leibler,I. (1951) "On information and sufficiency." *Ann.of Math. Stat. 79-86*

Lehmann,E.L. (1986) "Testing Statistical Hypotheses" Second Edition *John Wiley and Sons.*

Leone, F.C. Nelson,L.C. Nottingham,R.B.(1961) "The Folded normal distribution." *Technometrics 3(4) 543-550*

Lien,D. Vuong,Q.(1987) "Selecting the best linear regression model. A classical approach." *J. of Econometrics 35 3-23*

Lindley,D. (1968) "The choice of variables in multiple regression." *J.R.S.S A 31-66*

Linhart,H. Zucchini,W. (1986) "Model Selection" *John Wiley and Sons*

Mallows,C.L. (1973) "Some comments on C_p." *Technometrics 15(4) 661-675*

Magee,L. Veall,M. (1991) "Selecting regressors for prediction using PRESS and white t statistics." *J. of Business and Economics Statistics 9(1) 91-96*

Martini,A.S. Spezzaferri,F. (1984) "A predictive model selection criterion." *J.R.S.S. B 46(2) 296-303*

Mijnheer,J.L.(1968) "Steekproven uit de halve Cauchy verdeling" *Statistika Neerlandica 22 97-101*

Miller,A.J. (1984) "Selection of subsets of regression variables" *J.R.S.S. A 147(3) 389-425*

Nelson, L.C." (1980) "The Folded normal distribution." *J.of Quality Technology 12(4) 236-238*

Nelson,W. (1984) "Reconstruction of a Gaussian mean from folded data" *IEEE Transactions on Aerospace and Electronic Systems 20 112-118*

Niki,N.(1979) "Multi-folding the normal distributions and mutual transformation between uniform and normal random variables." *Ann. Inst. Statist. Math. 31 A 125-140*

Nishii,R. (1984) "Asymptotic properties of criteria for selection of variables in multiple regression." *The Annals of Statistics*

12(2) 758-765

Nordheim,E. O'Malley,D. Chow,S.C. (1984) "On the performance of a likelihood ratio test for genetic linkage." *Biometrics 40 785-790*

Novotny,T.J. Mcdonald,L.L. (1986) "Model selection using discriminant analysis." *Journal of Applied Statistics 13(2) 159-165*

Parzen,E. (1977) "Multiple time series modeling: Determining the order of approximating autoregressive schemes." *Multivariate Analysis IV (Ed. P.Krishnaiah) 283-295*

Patil,G.P. Boswell,M.T. Ratnaparkhi,M.V.(1984) "Dictionary and classified bibliography of statistical distributions in scientific work." *V.2 Continuous Univariate Models.International Co-operative Publishing House USA*

Patil,G.P. Boswell,M.T. Ratnaparkhi,M.V.Roux,J.J.J.(1984)"Dictionary and classified bibliography of statistical distributions in scientific work." *V.3 Multivariate Models International Co-operative Publishing House USA*

Pederson,S.P. Johnson,M.E. (1990) "Estimating model discrepancy." *Technometrics 32(3) 303-314*

Pesaran, M. (1973) " On the general problem of model selection." *The Review of Economic Studies 153-171*

Picard,R. Berk,K. (1990) "Data Splitting." *American Statistician 44(2) 140-147*

Picard,R. Cook,D. (1984) "Cross-validation of regression models." *JASA 79 (387) 575-583*

Poirier,D. Klepeer,S. (1981) "Model occurence and model selection in panel data sets." *J. of Econometrics 17 333-350*

Pollak,R. Wales,T. (1991) "The likelihood dominance criterion. A new approach to model selection." *J. of Econometrics 47 227-242*

Psarakis,S. Panaretos,J. (1990a) "The folded normal distribution." *Comm. Stat, Theory and Methods 22(7) 2717-2734*

Psarakis,S. Panaretos,J. (1990b) "On an approach to model evaluation and selection." *XV Biometric Conference Budapest*

Rao,C.R. Wu,Y. (1989) "A strongly consistent procedure for model selection in a regression problem." *Biometrika 76(2) 369-374*

Rissanen, J. (1987) "Stochastic complexity." *J.R.S.S. B 49(3) 223-239 and 252-265*

Risvi,M.H. (1971) "Some selection problems involving folded normal distributions." *Technometrics 13(2) 355-369*

Sawa,T. (1978) "Information criteria for discriminating among alternatives regression models." *Econometrika 46(6) 273-1291*

Schwarz,G. (1978) "Estimating the dimension of a model" *Annals of Statistics 6 461-464*

Shibata,R. (1980) "Asymptotically efficient selection of the order of the model for estimating parameters of a linear process." *The Annals of Statistics 8(1) 147-164*

Shibata,R. (1981) "An optimal selection of regression variables." *Biometrika 68(1) 45-54*

Shibata,R. (1984) "Approximate efficient of a selection procedure for the number of regression variables." *Biometrika 71(1) 43-49*

Shibata,R. (1985) "Various model selection techniques in time series analysis" 179-187 *Handbook of Statistics V.5 E.J.Hannan, P.R.Krishnaiah, M.M.Rao eds. Elsevie Science Publishers B.V.*

Sinha, S.K. (1983) "Folded normal distribution-A Bayesian approash." *J. of the Indian Statistical Accosiation 21 31-34*

Sundberg,R. (1974) "On estimatin and testing for the folded normal distribution." *Communications in Statistics 3 55-72*

Smith,A.F.M. Spiegelhalter,D.J. "Bayes factors and choice criteria for linear models." *J.R.S.S. B 42(2) 213-220*

Snee,R.D. (1977) "Validation of Regression Models: Methods and Examples." *Technometrics 19(4) 415-428*

Stepffens,F.E. (1969) "Critical values for bivariate student t-tests." *JASA 64 637-646*

Stone,M. (1974) "Cross-Validatory Choise and Assesment in Statistical Prediction." *J.R.S.S. A 111-147*

Stone,M. (1978) "Cross-Validation: A Review." *Math. Operations Furch Statist. Ser. Statistics 9(1) 127-139*

Stone,C. (1981) "Admissible selection of an accurate and parsimoniou normal linear regression model." *The Annals of Statistics 9(3) 475-485*

Swanepoel,J.W.H. C.F.De Beer (1992) "Some new model selection criteria in simple regression." *J. Stat. Comput. Simul. 42 137-149*

Tallis, G.M.(1961) "The moment Generating Function of the Truncated Multi-Normal Distribution" *J.R.S.S. 223-229*

Thompson,M. (1978a) "Selection of variables in multiple regression. Part 1" *International Statistical Review 46 1-19*

Thompson,M. (1978b) "Selection of variables in multiple regression. Part II " *International Statistical Review 46 126-46*

Tong,Y.L. (1990) "The multivariate normal distribution." *Springer Verlag*

Urbain,J.D. (1989) "Model selection criteria and granger causality tests." An empirical note." *Econom. Letters 29 317-320*

Yi,G. Judge,G. (1988) "Statistical model selection criteria." *Economics Letters 28 47-51*

Vuong,Q.H.(1989) "Likelihood ratio tests for selection and

non-nested hypotheses." *Econometrica 57(2) 307-333*

Walker, G.A. Saw,J.G. (1978) "The distribution of linear combinations of t-variables." *JASA 73 876-878*

Weisberg,S. (1985) "Applied linear regression." *John Wiley and Sons. Second Edition*

Whitmore,G.A. Xekalaki,E. (1990) "P-values as measures of predictive validity." *Biometrical Journal 32 977-983*

Wu,Y. (1989) "On a class of model selection procedures." *Comm. Stat. Theory and Meth. 18(9) 3267-3287*

Xekalaki,E. Katti,K. (1984) "Rating models based on concordance with observed data." *Biometrical Journal 26 173-184.*

Zahn,D. (1975) "Modifications of and revised critical values for the half normal plot" *Technometrics 17(2) 189-200*

Zahn,D. (1975) "An empirical study of the half normal plot." *Technometrics 17(2) 201-211*

Zaman,A. (1984) "Avoiding model selection by the use of shrinkage techniques." *J. of Econometrics 25 73-85*

Zhang,P. (1992) "On the distibutional properties of model selection criteria." *JASA 87(419) 732-737*

Zhang,Hong-Ching (1990) "The distribution of the squared sum of independent Cauchy variables." *In "Statistical inference in elliptically contoured and related distributions" (Anderson,T.W. Fang,K. eds.) 193-200*

ΠΡΟΣΘΕΤΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

Παρακάτω δίδεται μία σειρά εργασιών που έχουν γενικότερη σχέση με το πρόβλημα της επιλογής και αξιολόγησης μοντέλων.

Abraham,B. Ledolter,J. (1986) "Forecast functions implied by autoregressive integrating moving average models and othes relate forecast procedures." *Inter. Statist. Review 54(1) 51-66*

Alatiqi,I.M. Meziou,A.M. Gasmelseed,G.A. (1989) "Modelling ,simulation and sensitivity analysis of steam-methane reformers." *Int. J. Hydrogen Energy 14(4) 241-256*

Albers,W. (1978) "Testing the mean of a normal population under dependence." *The Annals of Statistics 6 1337-1344*

Anderson,R.L. (1954) "The problem of autocorrelation in regression analysis." *JASA 113-129*

Anscombe,F.J. (1941) "Examination of residuals" *Fourth Berkeley Symposium 1-36*

Ansley,C.F. Spivey,W.A. Wrobelski,W.J. (1977) "A class of transformations for Box-Jekins seasonal models." *Applied Statistics 26(2) 173-178*

Anthes,R.A. (1983) "Regional models of the atmosphere in middle latitudes." *Montnly Weather Review V.III 1306-1335*

Armstrong,J.S. (1988) "Research needs in forecasting.". *International J of Forecasting. 4 449-465*

Atkinson,A.C.(1978) "Posterior probabilities for choosing a regression model." *Biometrika 65(1) 39-48*

Atkinson,A.C.(1982) "Regression diagnostics transformation and constructed variables." *J.R.S.S B 44(1) 1-36*

Azzabini,A. Bowman,A.W. Hardle,W. (1989) "On the use of nonparametric regression for model checking." *Biometrika 76(1) 1-11*

Belsey,D. (1988) "Modeling and forecasting reliability." *Internationa J. of Forecasting 4 427-447*

Berk,K. (1978) "Comparing subset regression procedures." *Technometrics 20(1) 1-6*

Bhensalli,R.J. (1986) "Asymtotically efficient selection of the order by the criterion autoregressive tranfer function." *The Annals of Statistics 14(1) 315-325*

Box,G.E.P. Pierce,D. (1970) "Distribution of residual autocorrelation in autoregressive integrated moving average time series models." *JASA 65(332) 1509-1526*

Box,G.E.P. Tiao,G.C. (1976) "Comparison of forecast and actuality." *Appl.Statistics 25(3) 195-200*

Bush,R.R. Mosteller,F.(1988) "A comparison of eight models." *Readings in mathematical social science Eds. Lazarsfeld,P.F. Henry,N.W. pp.335-349 Cambridge Massachusetts:the M.I.T. Press*

Butler.r. Rothman,E. (1980) "Predictive intervals based on reuse of th sample." *JASA 75(372) 881-889*

Chatfield,C. (1988) "What is the 'best' method of forecasting." *J. of Applied Statistics 15(1) 19-38*

Chon-Fu Cuen (1983) "Score-tests for regression models." *JASA 78(381) 158-161*

Cleamen,R. (1989) "Combining forecasts: Areview and annotated bibliography." *International J.of Forecasting 5 559-583*

Cohen,A. Sackrowitz,H.B. (1967) "An approash to inference following model selection with applications to transformation based and adaptive inference." *JASA 82(400) 1123-1130*

Cox,D.R. Snell,E.J. (1968) "A general definition of residuals." *JRSS B 248-275*

Creutin ,J.D. Obled,C. (1982) "Objectives alalyses and techniques for ainfall fields: an objective comparison." *Water Resources Reseai 18(2) 413-431*

Critcley,F. Ford,I. (1985) "Interval estimation in discrimination:The multivariate normal equal covariance case." *Biometrika 72(1) 109-116*

Davies,B.N. Newbold,P. (1980) "Forecasting with mispecified models." *Applied Statistics 29(1) 87-92*

Dawid,A. (1982) "The well calibrated bayesian." *JASA 77(179) 605-613*

Dragosevic,M. Stanvovic,S. (1989) "A generalized least squares method for frequency estimation." *IEEE Trans. on Acoustics 37(6) 805-8*

Duan,N. K.Chen Li (1987) "Distribution free and link free estimation for the sample selection model." *J. of Econometrics 35 25-35*

Durbin,J. Watson,G.S. (1950) "Testing for serial correlation in least squares regression I." *Biometrika 37 409-428*

Durbin,J. Watson,G.S. (1951) "Testing for serial correlation in least squares regression II." *Biometrika 38 159-178*

Durbin,J. (1969) "Tests for serial correlation in in regression analysis based on the periodogram of least-squares residuals." *Biometrika 56 1-15*

Durbin,J. (1970) "Testing for serial correlation in least-squares regression when some of the regressors are laggged dependent variables." *Econometrika 38(3) 410-421*

Durbin,J. (1970) "An alternative to the bounds test for testing for serial correlationin least squares regression." *Econometrika 38(3) 422-429*

Ebbeler,D. (1975) "On the probability of correct model selection using the maximum R^2 choice criterion." *Intrern. Econ. Review 16(2)*

Ehsenberg,A.S.C. (1982) "How good is best." *J.R.S.S. A 145(3) 364-366*

Findley,D.(1991) "Counterexamles to parsimony and BIC." *Ann. of The Inst of Math. Stat. 43(3) 505-514*

Fraker.I. Moffit,R. (1988) "The effect of food stamps on labor supply." *J. of Public Economics 35 25-56*

Fraser,D.A.S. Haq,M.S. (1969) "Structural p-robability and prediction for the multivariate model." *J.R.S.S. 317-331*

Gasride.M.J.(1971) "Some computational procedures for the best subset problem." *Applied Statistics 20 8-15*

Geisser,S. (1963) "Multivariate analysis of variance for a special covariance case." *JASA 56 660-669*

Gentleman,S.F. Wilk,M.B. (1975) "Detecting outliers in a two way table: I Statistical bahavior of residuals." *17(1) 1-44*

Gentleman,S.F. (1975) "Detecting outliers II. Supplementing the direct analysis of residuals." *Biometrics 31 387-410*

Golup,G. (1965) "Numerical methods for solving linear least squares problems." *Numerische Mathematik 7 206-216*

Gomez,G.L. Tintner,G. (1980)"The application of diffusion processes in problems of developmental economic planning:A case study (Colombia)" *Studies in Economic Theory and Practise.North Holland Publ.177-194* Goodman, L. (1964) "The analysis of persistence in a chain of multiple events." *Biometrika 51 405-411*

Gooijer,J.G. Abraham,B. Gould,A. Robinson,L. (1985) "Methods for

determining the order of an autoregressive moving average proces: A survey." *Inter.Stat. Review 53(3) 301-329*

Hahn,G. (1972) "Simultaneous prediction intervals for a regression model." *Technometrics 14(1) 203-214*

Hamouta,E.M. Leone,F. (1977) "The robustness of efficiency of adjustec trimmed estimators in linear regression." *Technometrics 19(1) 19-34*

Hanmoun,E.J. (19) "Testing for serial correlation in least squares regression." *Biometrika 44 57-66*

Hannan,J. Terrell,R.D. (19) "Testing for serial correlation after least squares regression." *Econometrika 36(1) 133-150*

Helm,R. (1974) "The average estimated variance criterion for the selection of variable problem in general linear models." *Technometrics 16(2) 261-273*

Henderson,H. Velleman,P. (1981) "Building multiple regression models interactively." *Biometrics 37 391-411*

Hocking,R.R. Leslie,R.N. (1967) "Selection of the best subset in regression analysis." *Technometrics 9(4) 531-540*

Hurvich,C. C.Ling Tsai (1990) "The impact of model selection on inference in linear regression." *The American Statistician 40(3 214-217*

Jennrich,R. Oman,S. (1986) "How much does Stein estimation help in multiple linear regression." *Technometrics 28(2) 113-121*

Joe,H. (1990) "Extended use of paired comparison models with applications to chess ranking." *Appl.Statistics 39(1) 85-93*

Judge,G. Yi,G. Vansey,T.(1987) "The extended Stein procedure for simultaneous model selection and parameter estimation." *J. of Econometrics. 35 333-339*

Kailath,T. (1974) "A view of three decades of linear filtering theory *IEEE Tranc. on Information theory 20(2) 146-181*

Kohn,R. Ansley,C. (1989) "A fast algorithm for signal etraction influence and cross-validation in state space models." *Biometrika 76(1) 65-79*

Koreisha,S. Dukkila,T. (1990) "Linear methods for estimating ARMA and regression models with serial correlation." *Comm. Stat. Sim. and Comp. 19(1) 71-102*

Krasker,N.S.(1981) "The role of bounded-influence estimation in model selection." *J. of Econometrics 16 131-138*

Lampard,D.G. (1968) "A stochastic process whose successive intervals between events form a first orded Markov chain-I" *J. Appl. Prob. 5 648-668*

Landry,M. Malouin,J. Oral,M. (1983) "Model validation in operation research." *European J. of Operational Research 14 207-220*

Landwehr,J. Regibon,D. Shoemaker,A. (1984) "Graphical methods for assessing logistic regression models." *JASA 79(385) 61-80*

Lehmann,E.L. (1990) "Model specification:The views of Fisher and Neyman and later developments." *Statistical Science 5(2) 160-168*

Lindley,D. (1982) "Scoring rules and the inevitability of probability *Intern. Statist. Review 50 1-26*

Little,R.J.A. (1979) "Maximum likelihood inference for multiple regression with missing values:A simulation study." *J.R.S.S. B 41(1) 76-87*

Makridakis,S. Hibon,M. (1979) "Accuracy of forecasting .An empirical investigation." *J.R.S.S. A 142 97-145*

Mcdonald,G.C. (1979) "Non parametric selection procedures applied to state traffic atality rules." *Technometrics 21(4) 515-523*

Mcdonald,J. Rishards,D. (1989) "Model selection: Some generalized distributions." *Comm. Stat. Theory and Methods 16(4) 1049-1074*

Meeden,G. (1979) "Comparing two probability appraisers." *JASA 74(366) 299-302*

Mitchell,A. Krzonowski,N. (1985) The Mahalanobis distance and ellipti distributions." *Biometrika 72(2) 464-467*

Morris,C. (1979) "A finite selection model for experimental design of the health insurance study." *J. of Econometrics 11 43-61*

Newbold,P. Granger,C.W.J. (1974) "Experience with forecasting univariate time series and the combination of forecasts." *J.R.S.S. A 1974 137*

131-164

Newbold,P.(1981) "Some recent developments in time series analysis I. *Intenational Statistical Review 49*

Newbold,P.(1984) "Some recent developments in time series analysis II *Intenational.Statistical Review 52*

Newbold,P.(1988) "Some recent developments in time series analysis III." *Intenational Statistical Review 56*

Pagan,A. "On the role of simulation in the statistical evaluation of econometric models." *J. of Econometrics 40 125-139*

Pielman,T. Pfaffenberg,R. (1989) "Efficiency of ordinary least squares for linear models with autocorrelation." *JASA 84(405) 248*

Pino,F. Morretin,P. Mentz,P. (1987) "Modeling and forecasting linear combinations of time series." *Inter. Stat. Review 55(3) 295-313*

Pregibon,D. (1980) "Goodness at link tests for generalized linear models." *Applied Statistics 29(1) 15-24*

Press.S.J. (1969) "On serial correlation." The Annals of Mathematical *Statistics 40(1) 188-196*

Pukkila,T. Koreisha,S. Kallinen,A. (1990) "The identification of ARMA models." *Biometrika 77(3) 537-548*

Rayner,J.C.W. Best,D.J. (1990) "Smooth tests of goodness of fit.An overview." *Intern. Statist. Review 58(1) 1-17*

Reckhow,K.H. Clements,J.T. Dodd,R.C. (1990) "Statistical evaluation o mechanistic water-quality models." *J. of Enviromental Engineeri 116 250-268*

Rencher,A. Pun,E. (1980) "Inflation of R^2 in best subset regression." *Technometrics 22(1) 49-53*

Roecher,E. (1991) "Prediction error and its estimation for subset selected models." *Technometrics 33(4) 459-468*

Sargan,J.D. Bhargawa,A. (19) "Testing residuals from least squares regression for being generated by the Gaussian random walk." *Econometrika 51(1) 153-174*

Sastri,T. (1989) "Multipass seasonal adjustment filter." Management *Science 33(1) 100-123*

Scheffe H. (1970) "Practical solutions of the Behrens-Fisher problem. *JASA 65 1501-1508*

Shively,T. Ansley,C. Kohn,R. (199) "Fast evaluation of the distributi of the Durbin-Watson and othe invariant test statistics in time series regression." *JASA 85(411) 676-685*

Simmons,J. (1989) "A probabilistic model of squash: Strtegies and application." *Appl. Statistics 38(1) 95-110*

Simonoff,J. (1988) "Regressions diagnostics to detect nonrandom missigness in linear regression." *Technometrics 30(2) 205-214*

Stine,R.(1985) "Bootstrap prediction intervals for regression." *JASA 80(329) 1026-1031*

Swanepoel,J.W.H. C.F.De Beer (1992) "Some new model selection criteri

in simple regression." *J. Statist. Comput. Simul. 42 137-149*

Tong,H. (1975) "Autoregressive model fitting with noisy data by Akaike's information criterion." *IEEE Trans. on Inform. Theory 476-480*

Tryfos,P. Blackmore,R. (1985) "Forecasting records." *JASA 80(389) 46-50*

Tsay,R.S. (1984) "Order selection in non stationary autoregressive models." *The Annals of Statistics 12(4) 1425-1433*

Van Der Laaan,P. Verdooren,L.R. (1989) "Selection of populations: An overview and some recent results." *Biom. J. 4 383-420*

Winkler,R. (1969) "Scoring rules and the evaluation of probability assessors." *JASA 1073-1078*

Winkler,R. (1971) "Probabilistic prediction: Some experimental results." *JASA 66(366) 675-685*

Winkler,R. Makridakis,S. (1983) "The Combinations of Forecasts." *J.R.S.S. A 150-157*

ΠΑΡΑΡΤΗΜΑ Ι

ΑΠΟΤΕΛΕΣΜΑΤΑ ΜΕΘΟΔΟΥ ΑΞΙΟΛΟΓΗΣΗΣ ΤΗΣ ΠΡΟΒΛΕΠΤΙΚΗΣ ΙΚΑΝΟΤΗΤΑΣ ΕΝΟΣ ΓΡΑΜΜΙΚΟΥ ΜΟΝΤΕΛΟΥ ΜΕ ΤΗΝ ΧΡΗΣΙΜΟΠΟΙΗΣΗ ΤΗΣ X^2 ΚΑΤΑΝΟΜΗΣ

Y_t^0 η παρατηρούμενη τιμή του μοντέλου την χρονική στιγμή t

$\hat{Y}_t^0(A)$ η προβλεπόμενη τιμή του μοντέλου Α την χρονική στιγμή t

$\hat{Y}_t^0(B)$ η προβλεπόμενη τιμή του μοντέλου Β την χρονική στιγμή t

ΠΙΝΑΚΑΣ Ι

CRD10

MODEL A

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0(A)$	r_t^2
1	1951	30.4	31.3	0.1559
2	1952	32.1	30.4	0.1441
3	1953	30.1	32.1	0.1818
4	1954	38.7	30.7	4.336
5	1955	31.8	36.6	0.7555
6	1956	39.0	33.4	1.996
7	1957	39.7	39	0.02957
8	1958	43.0	36	1.497
9	1959	42.0	38.3	0.7298
10	1960	48.2	41.9	1.919
11	1961	50.7	41.3	7.134
12	1962	53.8	44.9	2.079
13	1963	57.7	32.7	1.389
14	1964	48.9	62.9	1.326
15	1965	61.2	49.1	1.167
16	1966	51.4	59.2	0.8193
17	1967	57.4	44.2	3.197
18	1968	58.5	58.9	3.398E-3
19	1969	66.6	56.2	1.79
20	1971	71.6	47.0	8.216
21	1972	69.6	65.2	0.3996
22	1973	67.5	66.3	0.02852
23	1974	45.5	64.5	10.95
24	1975	65.2	59.4	0.6883
25	1976	70.6	49.3	8.304
26	1977	66.1	47.9	1.281
27	1978	66.9	67.0	2.248E-4
28	1979	73.1	70.0	0.216
29	1980	49.3	63.2	3.685

Σύνολο r_t^2 :64.28
Βαθμοί ελευθερίας :29
p-value :1.76E-4

ΠΙΝΑΚΑΣ ΙΙ

CRD20

MODEL A

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0(A)$	r_t^2
1	1951	36.0	32.7	1.9169
2	1952	36.7	36.0	0.02189
3	1953	33.7	36.2	0.3547
4	1954	41.4	42.9	0.147
5	1955	37.8	33.1	0.7742
6	1956	40.1	39.4	0.03366
7	1957	40.4	40.0	0.01985
8	1958	42.4	43.1	0.01694
9	1959	39.1	39.5	0.01188
10	1960	44.3	43.0	0.1621
11	1961	47.7	43.2	2.544
12	1962	49.6	43.5	3.341
13	1963	52.3	48.2	0.1075
14	1964	44.6	56.6	2.481
15	1965	56.8	42.6	4.667
16	1966	46.8	49.3	0.1718
17	1967	50.9	42.2	2.258
18	1968	55.5	50.5	0.8237
19	1969	59.5	60.9	0.05717
20	1971	62.8	56.3	0.9212
21	1972	61.8	60.6	0.08711
22	1973	62.3	58.9	0.7158
23	1974	38.9	59.0	31.24
24	1975	62.4	64.1	0.03564
25	1976	67.6	57.8	3.146
26	1977	62.6	63.9	0.05361
27	1978	61.7	65.4	0.541
28	1979	68.4	67.3	0.05361
29	1980	54.0	63.7	5.184

Σύνολο r^2 :61.9

Βαθμοί ελευθερίας :29

p-value :3.56E-4

ΠΙΝΑΚΑΣ ΙΙΙ

CRD30

MODEL A

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0(A)$	r_t^2
1	1951	32.6	39.5	6.919
2	1952	36.3	33.5	0.2979
3	1953	33.4	44.7	2.335
4	1954	40.6	56.2	1.175
5	1955	38.3	38	3.165E-3
6	1956	38.1	40.8	1.834
7	1957	34.8	39.4	1.834
8	1958	42.1	41.3	0.01295
9	1959	37.3	32.3	1.677
10	1960	38.6	39.6	0.06103
11	1961	45.9	42.7	0.7064
12	1962	45.5	42.3	0.6399
13	1963	51.2	42.8	0.2405
14	1964	39.3	54.8	2.465
15	1965	52.6	32.8	5.926
16	1966	48.3	51.0	0.09863
17	1967	46.4	44.7	0.05661
18	1968	54.2	46.9	1.094
19	1969	52.8	54.4	0.0385
20	1971	55.1	57.7	0.0695
21	1972	59.3	57.1	0.176
22	1973	58.1	59.9	0.1358
23	1974	35.3	49.9	7.421
24	1975	52.4	50.7	0.06022
25	1976	62.8	58.1	0.6089
26	1977	64.9	60.7	0.6919
27	1978	60.9	60.3	0.01775
28	1979	67.5	66.4	0.07174
29	1980	65.1	65.1	0

Σύνολο r_t^2 :35.37

Βαθμοι ελευθερίας :29

p-value :0.193

ΠΙΝΑΚΑΣ IV

CRD40

MODEL A

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0(A)$	r_t^2
1	1951	34.2	37.6	1.195
2	1952	31.8	29.7	0.4317
3	1953	34.2	38.4	1.61
4	1954	35.3	35.1	2.561E-3
5	1955	33.6	29.9	0.9244
6	1956	40.6	40.0	0.05448
7	1957	35.7	39.5	2.861
8	1958	39.9	21.6	8.015
9	1959	40.8	38.8	0.4737
10	1960	44.7	42.5	0.4885
11	1961	46.3	42.5	1.803
12	1962	55.8	42.9	14.4
13	1963	60.1	72.5	0.7521
14	1964	48.1	62.2	2.746
15	1965	64.4	46.4	6.947
16	1966	48.9	53.5	0.5392
17	1967	57.0	52.9	0.5181
18	1968	56.4	49.6	1.511
19	1969	65.8	55.4	4.157
20	1971	67.7	74.3	0.7936
21	1972	69.6	64.8	0.8173
22	1973	67.5	60.8	1.758
23	1974	43.0	50.1	1.445
24	1975	69.1	53.5	8.392
25	1976	70.8	74.7	0.4125
26	1977	63.8	43.5	3.973
27	1978	73.1	79.2	1.192
27	1979	74.7	69.9	0.6684
29	1980	59.1	71.6	4.958

Σύνολο r_t^2 :73.84

Βαθμοί ελευθερίας :29

p-value :1E-5

ΠΙΝΑΚΑΣ V

CRD50

MODEL A

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0(A)$	r_t^2
1	1951	38.0	39.2	0.9054
2	1952	33.0	32.6	0.01003
3	1953	37.1	41.1	0.6995
4	1954	37.1	38.0	0.04743
5	1955	38.0	31.7	1.936
6	1956	40.2	41.1	0.0873
7	1957	38.8	42.3	1.447
8	1958	39.6	29.2	0.5827
9	1959	42.1	39.1	0.9662
10	1960	44.4	42.8	0.2457
11	1961	46.5	43.7	1.109
12	1962	55.4	44.6	7.287
13	1963	56.4	68.3	0.6115
14	1964	46.2	60.3	3.163
15	1965	62.1	36.7	14.62
16	1966	50.9	54.3	0.2398
17	1967	52.0	50.1	0.06496
18	1968	56.8	55.3	0.05467
19	1969	67.5	55.4	3.767
20	1971	65.9	75.9	1.47
21	1972	67.3	62.0	0.8253
22	1973	66.5	72.7	0.5288
23	1974	47.5	64.7	7.137
24	1975	64.6	65.4	0.02333
25	1976	72.5	75.3	0.2008
26	1977	64.1	53.3	1.447
27	1978	73.7	64.0	3.043
28	1979	73.7	74.4	9.501E-3
29	1980	67.7	66.4	0.07354

Σύνολο r_t^2 :51.79

Βαθμοί ελευθερίας :29

p-value :5.753E-3

ΠΙΝΑΚΑΣ VI

CRD60

MODEL A

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0(A)$	r_t^2
1	1951	35.0	36.0	0.09466
2	1952	28.5	32.2	1.481
3	1953	34.7	37.3	0.5431
4	1954	38.9	37.6	0.04229
5	1955	38.3	37.1	0.08221
6	1956	37.6	40.9	1.728
7	1957	36.8	39.1	0.9337
8	1958	36.2	35.7	8.431E-3
9	1959	37.6	33.7	2.151
10	1960	43.4	40.4	0.9275
11	1961	45.9	44.0	0.6054
12	1962	53.8	23.4	13.69
13	1963	52.5	128.6	9.86
14	1964	39.6	58.2	4.421
15	1965	51.2	32.7	4.806
16	1966	47.3	47.4	2.014E-4
17	1967	46.9	42.0	0.4944
18	1968	53.7	53.4	2.511E-3
19	1969	59.4	63.7	0.4927
20	1971	57.9	62.0	0.2965
21	1972	62.1	53.8	2.357
22	1973	61.7	65.3	0.4775
23	1974	47.4	59.2	3.849
24	1975	65.7	58.6	0.4202
25	1976	65.7	69.4	0.5613
26	1977	59.5	50.6	2.801
27	1978	68.0	67.4	0.01911
28	1979	70.3	49.6	5.157
29	1980	70.6	70.7	4.521E-4

Σύνολο r_t^2 : 58.3

Βαθμοί ελευθερίας :29

p-value :0.001

ΠΙΝΑΚΑΣ VII

CRD70

MODEL A

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0(A)$	r_t^2
1	1951	29.9	29.2	0.05496
2	1952	26.7	26.0	0.03201
3	1953	25.7	27.0	0.2103
4	1954	22.7	27.0	2.378
5	1955	31.9	21.5	8.122
6	1956	38.3	32.9	2.828
7	1957	32.2	36.1	1.349
8	1958	34.8	27.8	0.244
9	1959	35.3	33.0	0.2751
10	1960	37.0	34.1	0.7177
11	1961	41.8	35.5	3.985
12	1962	46.3	37.1	8.02
13	1963	52.6	49.4	0.06725
14	1964	48.4	54.5	0.6119
15	1965	60.9	45.6	5.896
16	1966	47.2	59.6	3.958
17	1967	62.0	54.1	1.378
18	1968	52.6	55.9	0.2545
19	1969	61.7	53.4	2.606
20	1971	60.1	58.8	0.03566
21	1972	65.8	55.9	3.585
22	1973	61.5	59.4	0.1818
23	1974	49.2	51.2	0.04136
24	1975	63.3	58.1	1.262
25	1976	70.3	66.5	0.5117
26	1977	66.0	62.0	0.795
27	1978	68.2	68.5	4.936E-3
28	1979	66.0	62.7	0.3291
29	1980	61.8	66.7	1.005

Σύνολο r_t^2 :50.74

Βαθμοί ελευθερίας :29

p-value :7.508E-3

ΠΙΝΑΚΑΣ VIII

CRDθO

MODEL A

Δείκτης	Ετος	Y_t^o	$\hat{Y}_t^o(A)$	r_t^2
1	1951	27.2	27.2	0
2	1952	21.5	23.5	0.2571
3	1953	25.9	28.6	0.8785
4	1954	20.8	23.8	1.208
5	1955	31.0	25.2	1.862
6	1956	35.2	29.5	5.337
7	1957	30.6	30.4	4.352E-3
8	1958	35.6	32.8	0.1054
9	1959	34.0	32.8	0.1406
10	1960	32.5	31.9	0.05074
11	1961	42.6	35.7	8.102
12	1962	41.5	35.6	4.338
13	1963	47.6	36.2	1.432
14	1964	37.6	44.6	2.017
15	1965	52.6	32.5	13.44
16	1966	44.6	52.6	1.754
17	1967	54.5	51.8	0.2135
18	1968	48.5	47.3	0.05046
19	1969	57.9	50.4	2.79
20	1971	52.8	58.5	0.8682
21	1972	60.5	53.5	1.904
22	1973	57.3	54.9	0.2829
23	1974	49.3	53.9	0.3798
24	1975	42.8	53.3	6.359
25	1976	64.7	54.0	4.127
26	1977	60.4	52.1	2.35
27	1978	62.2	63.0	0.02051
28	1979	60.4	58.3	0.1524
29	1980	54.6	60.1	1.116

Σύνολο r_t^2	:61.59
Βαθμοί ελευθερίας	:29
p-value	:3.9E-4

ΠΙΝΑΚΑΣ IX

CRD90

MODEL A

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0(A)$	r_t^2
1	1951	25.9	28.9	0.6206
2	1952	24.8	28.4	0.6896
3	1953	30.5	32.1	0.1962
4	1954	24.5	24.7	2.241E-3
5	1955	30.5	27.4	0.3675
6	1956	34.7	34.1	0.05098
7	1957	33.2	34.8	0.3861
8	1958	34.0	22.9	2.148
9	1959	35.0	34.3	0.05383
10	1960	34.4	35.4	0.1614
11	1961	43.5	37.1	6.973
12	1962	47.7	38.0	12.78
13	1963	46.4	49.4	0.1008
14	1964	40.1	41.8	0.08143
15	1965	51.6	33.0	15.27
16	1966	50.4	47.6	0.2597
17	1967	53.6	49.2	0.6037
18	1968	49.6	54.3	0.7509
19	1969	57.8	49.5	4.186
20	1971	54.5	59.3	0.7451
21	1972	56.5	53.4	0.4112
22	1973	50.6	51.4	0.01528
23	1974	47.7	54.3	1.668
24	1975	47.9	48.2	5.299E-3
25	1976	63.5	50.4	10.4
26	1977	63.7	54.6	3.924
27	1978	64.3	61.5	0.4276
28	1979	63.0	62.4	0.02001
29	1980	59.6	60.5	0.03855

Σύνολο r_t^2 :63.34

Βαθμοί ελευθερίας :29

p-value :2.33E-4

ΠΙΝΑΚΑΣ Χ

CRD100

MODEL A

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0(A)$	r_t^2
1	1951	33.3	36.4	1.323
2	1952	31.4	30.2	0.1274
3	1953	32.3	36.4	1.676
4	1954	34.8	34.3	9.084E-3
5	1955	35.1	30.0	1.603
6	1956	38.9	38.3	0.04454
7	1957	37.0	39.6	1.187
8	1958	39.5	24.4	1.835
9	1959	39.2	37.7	0.2914
10	1960	42.7	39.2	1.43
11	1961	46.4	41.8	2.486
12	1962	51.5	44.6	4.897
13	1963	54.6	52.0	0.05308
14	1964	45.2	59.9	3.868
15	1965	59.0	44.1	4.636
16	1966	49.0	59.1	2.983
17	1967	54.0	49.0	0.6973
18	1968	55.2	53.3	0.1281
19	1969	62.8	55.0	2.278
20	1971	63.4	71.2	1.326
21	1972	65.3	62.8	0.2524
22	1973	64.0	63.2	0.03199
23	1974	44.6	53.9	2.766
24	1975	61.5	46.3	10.47
25	1976	69.0	67.5	0.07445
26	1977	64.0	57.0	1.25
27	1978	67.8	74.5	2.164
28	1979	70.3	68.9	0.07489
29	1980	60.3	65.4	1.289

Σύνολο r_t^2:51.25

Βαθμοί ελευθερίας :29

p-value :6.601E-3

ΠΙΝΑΚΑΣ Ι

CRD10

MODEL B

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0$	r_t^2
1	1956	39.0	32.7	1.768
2	1957	39.7	41.3	0.1692
3	1958	43.0	42.5	0.01823
4	1959	42.0	43.0	0.07437
5	1960	48.2	44.5	1.233
6	1961	50.7	48.1	0.6355
7	1962	53.8	50.3	1.106
8	1963	57.7	54.3	1.144
9	1964	48.9	52.9	1.484
10	1965	61.2	53.6	5.337
11	1966	51.4	57.0	2.836
12	1967	57.4	62.5	2.16
13	1968	58.5	60.8	0.4282
14	1969	66.6	62.3	1.535
15	1971	71.6	69.2	0.466
16	1972	69.6	69.4	3.501E-3
17	1973	67.5	70.1	0.5783
18	1974	45.5	69.6	49.5
19	1975	65.2	69.3	0.6665
20	1976	70.6	71.1	0.01015
21	1977	66.1	71.1	0.5098
22	1978	66.9	71.6	0.9653
23	1979	73.1	76.0	0.3613
24	1980	49.3	71.2	19.77

Σύνολο r_t^2 :92.8

Βαθμοί ελευθερίας :24

p-value :0.000

ΠΙΝΑΚΑΣ ΙΙ

CRD20

MODEL B

Δείκτης	Έτος	Y_t^0	$\hat{Y}_t^0$	r_t^2
1	1956	40.1	38.9	0.071
2	1957	40.4	42.8	0.4026
3	1958	42.4	47.9	2.244
4	1959	39.1	38.1	0.07339
5	1960	44.3	44.5	3.615E-3
6	1961	47.7	46.6	0.1198
7	1962	49.6	45.3	1.839
8	1963	52.3	50.8	0.2282
9	1964	44.6	48.3	1.514
10	1965	56.8	48.9	6.505
11	1966	46.8	48.3	0.2073
12	1967	50.9	51.4	0.02142
13	1968	55.5	53.4	0.2461
14	1969	59.5	56.1	1.168
15	1971	62.8	59.9	0.8795
16	1972	61.8	59.9	0.3815
17	1973	62.3	61.4	0.08092
18	1974	38.9	57.5	32.88
19	1975	62.4	60.5	0.2145
20	1976	67.6	62.1	1.895
21	1977	62.6	61.4	0.07918
22	1978	61.7	63.3	0.1647
23	1979	68.4	67.0	0.1251
24	1980	54.0	66.2	8.252

Σύνολο r_t^2 :59.6

Βαθμοί ελευθερίας :24

p-value :7.3E-5

ΠΙΝΑΚΑΣ ΙΙΙ

CRD30

MODEL B

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0$	r_t^2
1	1956	38.1	41.2	0.3949
2	1957	34.8	39.3	0.93016
3	1958	42.1	45.5	0.5274
4	1959	37.3	34.8	0.31112
5	1960	38.6	38.8	2.3356E-3
6	1961	45.9	43.3	0.42529
7	1962	45.5	42.9	0.45527
8	1963	51.2	48.6	0.42592
9	1964	39.3	45.5	2.6918
10	1965	52.6	45.4	3.5739
11	1966	48.3	48.6	6.0247E-3
12	1967	46.4	48.0	0.16793
13	1968	54.2	48.0	2.7323
14	1969	52.8	50.8	0.27545
15	1971	55.1	54.4	0.034732
16	1972	59.3	55.6	1.0036
17	1973	58.1	57.3	0.046335
18	1974	35.3	51.1	16.743
19	1975	52.4	48.8	0.66892
20	1976	62.8	60.8	0.21563
21	1977	64.9	58.7	2.2155
22	1978	60.9	59.1	0.18245
23	1979	67.5	62.8	1.2723
24	1980	65.1	64.1	0.052171

Σύνολο r_t^2 :35.354
Βαθμοί ελευθερίας :24
p-value :0.063

ΠΙΝΑΚΑΣ IV

CRD40

MODEL B

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0$	r_t^2
1	1956	40.6	35.6	1.351
2	1957	35.7	38.5	0.4614
3	1958	39.9	41.8	0.1916
4	1959	40.8	36.8	1.06
5	1960	44.8	39.9	1.774
6	1961	46.3	42.3	1.209
7	1962	55.8	47.9	4.288
8	1963	60.1	47.5	11.21
9	1964	48.1	50.6	0.3216
10	1965	64.4	52.5	7.783
11	1966	48.9	53.3	0.8067
12	1967	57.0	56.6	7.267E-3
13	1968	56.4	58.1	0.1408
14	1969	65.8	61.2	1.063
15	1971	67.7	67.3	7.59E-3
16	1972	69.6	65.2	0.9883
17	1973	67.5	66.0	0.1
18	1974	43.0	64.9	23.2
19	1975	69.1	66.0	0.3273
20	1976	70.8	67.8	0.3191
21	1977	63.8	69.2	1.059
22	1978	73.1	74.4	0.0611
23	1979	74.7	80.7	1.154
24	1980	59.1	73.6	7.885

Σύνολο r_t^2 :66.76

Βαθμοί ελευθερίας :24

p-value :7E-6

ΠΙΝΑΚΑΣ V

CRD50

MODEL B

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0$	r_t^2
1	1956	40.2	38.9	0.0975
2	1957	38.8	41.8	0.6011
3	1958	39.6	49.1	4.929
4	1959	42.1	33.8	4.077
5	1960	44.4	39.6	1.397
6	1961	46.5	42.1	1.14
7	1962	55.4	48.5	2.698
8	1963	56.4	49.4	2.71
9	1964	46.2	50.6	1.123
10	1965	62.1	49.6	8.769
11	1966	50.9	52.7	0.1381
12	1967	52.0	54.4	0.255
13	1968	56.8	52.2	1.064
14	1969	67.5	58.8	3.955
15	1971	65.9	62.8	0.4682
16	1972	67.3	60.6	2.162
17	1973	68.5	64.8	0.6571
18	1974	47.5	63.2	10.88
19	1975	64.6	60.3	0.6733
20	1976	72.5	67.9	0.854
21	1977	64.1	67.4	0.4491
22	1978	73.7	70.8	0.35
23	1979	73.7	80.7	1.573
24	1980	67.7	67.4	3.404E-3

Σύνολο r_t^2: 51.03

Βαθμοί ελευθερίας :24

p-value : 1.045E-3

ΠΙΝΑΚΑΣ VI

CRD60

MODEL B

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0$	r_t^2
1	1956	37.6	38.2	0.01496
2	1957	36.8	38.1	0.08157
3	1958	36.2	45.6	4.045
4	1959	37.5	35.2	0.2621
5	1960	43.4	37.1	2.167
6	1961	45.9	41.8	0.8987
7	1962	53.8	49.6	0.7882
8	1963	52.5	43.8	4.27
9	1964	39.6	45.4	1.693
10	1965	51.2	43.2	3.144
11	1966	47.3	45.0	0.2441
12	1967	46.9	46.6	4.442E-3
13	1968	53.7	50.0	0.7386
14	1969	59.4	54.4	1.347
15	1971	57.9	55.4	0.336
16	1972	62.1	54.5	3.146
17	1973	61.7	59.2	0.3328
18	1974	47.4	55.9	3.635
19	1975	65.7	56.0	4.486E-3
20	1976	65.7	59.2	2.253
21	1977	59.5	59.4	5.127E-4
22	1978	68.0	64.4	0.6924
23	1979	70.3	77.9	2.146
24	1980	70.6	67.4	0.5451

Σύνολο r_t^2 :32.79

Βαθμοί ελευθερίας :24

p-value :0.108

ΠΙΝΑΚΑΣ VII

CRD70

MODEL B

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0$	r_t^2
1	1956	38.3	36.3	0.2201
2	1957	32.2	39.7	3.787
3	1958	34.8	39.1	1.214
4	1959	35.3	36.4	0.08552
5	1960	37.0	39.7	0.5529
6	1961	41.8	41.0	0.05093
7	1962	46.3	42.2	1.436
8	1963	52.6	46.2	3.492
9	1964	48.4	47.2	0.1165
10	1965	60.9	50.4	9.357
11	1966	47.2	51.2	1.064
12	1967	62.0	60.0	0.258
13	1968	52.6	53.9	0.1158
14	1969	61.7	56.1	2.274
15	1971	60.1	64.0	1.077
16	1972	65.8	63.0	0.5565
17	1973	61.5	62.8	0.1183
18	1974	49.2	65.7	20.32
19	1975	63.3	63.8	0.01256
20	1976	70.3	69.2	0.06415
21	1977	66.0	65.4	0.01871
22	1978	68.2	68.2	0
23	1979	66.0	72.7	2.596
24	1980	61.8	64.0	0.2292

Σύνολο r_t^2 :49.01

Βαθμοί ελευθερίας :24

p-value :1.869E-3

ΠΙΝΑΚΑΣ VIII

CRDBO

MODEL B

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0$	r_t^2
1	1956	35.2	34.0	0.09906
2	1957	30.6	35.3	1.92
3	1958	35.6	45.1	5.097
4	1959	34.0	33.0	0.081
5	1960	32.5	35.9	0.9788
6	1961	42.6	37.8	2.113
7	1962	41.5	39.8	0.2625
8	1963	47.6	43.9	1.311
9	1964	37.6	42.1	3.955
10	1965	52.6	44.6	5.916
11	1966	44.6	44.7	7.625E-4
12	1967	54.5	52.9	0.1958
13	1968	48.5	48.1	0.01319
14	1969	57.9	51.5	3.546
15	1971	52.8	58.2	2.383
16	1972	60.5	54.7	2.598
17	1973	57.3	57.5	2.934E-3
18	1974	49.3	56.7	3.782
19	1975	42.8	55.3	10.9
20	1976	64.7	60.4	1.097
21	1977	60.4	56.8	0.7517
22	1978	62.2	64.8	0.3934
23	1979	60.4	70.0	4.669
24	1980	54.6	56.4	0.1645

Σύνολο r_t^2 :52.23

Βαθμοί ελευθερίας :24

p-value :7.3E-4

ΠΙΝΑΚΑΣ IX

CRD90

MODEL B

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0$	r_t^2
1	1956	34.7	33.6	0.06879
2	1957	33.2	35.6	0.4415
3	1958	34.0	36.6	0.5664
4	1959	35.0	31.2	1.351
5	1960	34.4	36.4	0.3961
6	1961	43.5	37.7	3.492
7	1962	47.7	41.7	3.342
8	1963	46.4	45.9	0.02149
9	1964	40.1	42.1	0.3807
10	1965	51.6	48.9	0.5682
11	1966	50.4	44.0	4.17
12	1967	53.6	53.4	3.373E-4
13	1968	49.6	48.0	0.234
14	1969	57.6	49.2	7.011
15	1971	54.5	58.4	1.302
16	1972	56.5	54.9	0.2177
17	1973	50.8	56.3	2.492
18	1974	47.7	56.6	4.007
19	1975	47.9	53.2	2.012
20	1976	63.5	58.8	1.641
21	1977	63.7	55.4	4.737
22	1978	64.3	61.5	0.5351
23	1979	63.0	64.4	0.1374
24	1980	59.6	59.0	0.02111

Σύνολο r_t^2 :39.15

Βαθμοί ελευθερίας :24

p-value :0.0263

ΠΙΝΑΚΑΣ Χ

CRD100

MODEL B

Δείκτης	Ετος	Y_t^o	$\hat{Y}_t^o$	r_t^2
1	1956	38.9	37.9	0.0748
2	1957	37.0	41.1	1.725
3	1958	39.5	40.9	0.2079
4	1959	39.2	37.8	0.2421
5	1960	42.7	42.0	0.06355
6	1961	46.4	43.3	1.357
7	1962	51.5	44.8	6.329
8	1963	54.6	48.5	4.394
9	1964	45.2	49.1	1.714
10	1965	59.0	49.8	8.937
11	1966	49.0	51.6	0.6015
12	1967	54.0	56.6	0.5896
13	1968	55.2	54.8	0.01399
14	1969	62.8	57.1	2.987
15	1971	63.4	63.0	0.01464
16	1972	65.3	62.6	0.699
17	1973	64.0	64.4	0.01456
18	1974	44.6	65.0	41.76
19	1975	61.5	62.2	0.02379
20	1976	69.0	66.9	0.2256
21	1977	64.0	57.0	0.0327
22	1978	67.8	74.5	5.329E-4
23	1979	70.3	68.9	2.243E-3
24	1980	60.3	65.4	1.911

Σύνολο r_t^2 :73.92
Βαθμοί ελευθερίας :24
p-value :1E-6

ΠΑΡΑΡΤΗΜΑ ΙΙ

Εκατοστιαία σημεία της γενικευμένης F κατανομής
με συνάρτηση πυκνότητας πιθανότητας

$$f_z(z) = \frac{(1-\rho^2)^k}{B(k,k)}\, z^{k-1}\, (1+z)^{-2k} \left[1 - \left(\frac{2\rho}{z+1} \right)^2 z \right]^{-\frac{2k+1}{2}}$$

Εκατοστιαία σημεία της γενικευμένης F κατανομής

k	p	ρ: 0.0	0.1	0.2	0.3	0.4
1	0.75	3.000	2.990	2.940	2.87	2.76
	0.90	9.000	8.930	8.720	8.36	7.85
	0.95	19.00	18.80	18.3	17.4	16.27
	0.99	99.00	98.10	95.2	90.3	83.5
2	0.75	2.065	2.06	2.035	2.00	1.95
	0.90	4.11	4.08	4.01	3.88	3.71
	0.95	6.39	6.34	6.20	5.97	5.64
	0.99	15.98	15.84	15.42	14.71	13.72
3	0.75	1.785	1.775	1.765	1.74	1.70
	0.90	3.055	3.04	3.00	2.92	2.81
	0.95	4.284	4.26	4.18	4.04	3.85
	0.99	8.47	8.40	8.20	7.87	7.40
4	0.75	1.64	1.635	1.63	1.605	1.575
	0.90	2.59	2.58	2.55	2.49	2.41
	0.95	3.44	3.42	3.36	3.27	3.145
	0.99	6.03	5.99	5.86	5.64	5.34
5	0.75	1.551	1.548	1.535	1.52	1.49
	0.90	2.32	2.31	2.29	2.24	2.18
	0.95	2.98	2.96	2.92	2.84	2.74
	0.99	4.85	4.82	4.73	4.57	4.34
6	0.75	1.49	1.489	1.477	1.465	1.445
	0.90	2.15	2.14	2.12	2.08	2.02
	0.95	2.687	2.675	2.65	2.57	2.485
	0.99	4.155	4.13	4.06	3.93	3.75
7	0.75	1.445	1.44	1.435	1.42	1.405
	0.90	2.02	2.01	2.00	1.96	1.91
	0.95	2.49	2.47	2.44	2.39	2.31
	0.99	3.7	3.68	3.62	3.51	3.36
8	0.75	1.41	1.408	1.40	1.39	1.37
	0.90	1.93	1.92	1.90	1.87	1.83
	0.95	2.335	2.325	2.29	2.25	2.18
	0.99	3.37	3.36	3.30	3.21	3.08

Εκατοστιαία σημεία της γενικευμένης F κατανομής (συνέχεια)

k	p	ρ: 0.5	0.6	0.7	0.8	0.9
1	0.75	2.62	2.44	2.24	1.98	1.645
	0.90	7.2	6.4	5.45	4.33	3.02
	0.95	14.73	12.84	10.60	8.02	5.04
	0.99	74.8	64.1	51.7	36.7	20.4
2	0.75	1.88	1.795	1.69	1.56	1.38
	0.90	3.46	3.2	2.85	2.44	1.93
	0.95	5.22	4.7	4.07	3.34	2.46
	0.99	12.45	10.90	9.05	6.91	4.45
3	0.75	1.655	1.595	1.515	1.42	1.29
	0.90	2.67	2.49	2.27	2.00	1.66
	0.95	3.61	3.31	2.945	2.51	1.97
	0.99	6.8	6.05	5.17	4.13	2.91
4	0.75	1.535	1.486	1.435	1.347	1.245
	0.90	2.3	2.17	2.00	1.8	1.53
	0.95	2.96	2.74	2.48	2.16	1.76
	0.99	4.95	4.47	3.89	3.2	2.38
5	0.75	1.465	1.424	1.37	1.305	1.215
	0.90	2.09	1.96	1.84	1.67	1.46
	0.95	2.6	2.43	2.22	1.965	1.64
	0.99	4.05	3.69	3.25	2.73	2.11
6	0.75	1.415	1.378	1.335	1.275	1.192
	0.90	1.95	1.85	1.74	1.59	1.41
	0.95	2.37	2.23	2.06	1.835	1.56
	0.99	3.52	3.23	2.88	2.46	1.94
7	0.75	1.38	1.345	1.32	1.25	1.175
	0.90	1.85	1.76	1.66	1.54	1.37
	0.95	2.21	2.09	1.935	1.75	1.51
	0.99	3.16	2.92	2.62	2.27	1.82
8	0.75	1.35	1.32	1.28	1.23	1.165
	0.90	1.77	1.70	1.61	1.49	1.34
	0.95	2.1	1.985	1.85	1.675	1.46
	0.99	2.91	2.7	2.45	2.14	1.75

Εκατοστιαία σημεία της γενικευμένης F κατανομής (συνέχεια)

k	p	ρ: 0.0	0.1	0.2	0.3	0.4
9	0.75	1.38	1.378	1.375	1.365	1.345
	0.90	1.85	1.849	1.83	1.80	1.76
	0.95	2.22	2.21	2.19	2.14	2.18
	0.99	3.13	3.12	3.07	2.99	2.87
10	0.75	1.36	1.358	1.35	1.34	1.325
	0.90	1.79	1.785	1.775	1.75	1.710
	0.95	2.125	2.115	2.095	2.055	2.00
	0.99	2.94	2.93	2.88	2.81	2.705
11	0.75	1.337	1.336	1.333	1.321	1.308
	0.90	1.745	1.74	1.725	1.705	1.67
	0.95	2.05	2.04	2.02	1.983	1.935
	0.99	2.785	2.775	2.735	2.67	2.575
12	0.75	1.322	1.32	1.314	1.305	1.292
	0.90	1.705	1.70	1.685	1.665	1.63
	0.95	1.983	1.977	1.955	1.925	1.876
	0.99	2.66	2.65	2.61	2.55	2.465
13	0.75	1.308	1.305	1.299	1.29	1.279
	0.90	1.665	1.664	1.65	1.63	1.60
	0.95	1.93	1.922	1.905	1.875	1.83
	0.99	2.555	2.545	2.51	2.455	2.375
14	0.75	1.294	1.292	1.287	1.279	1.267
	0.90	1.635	1.63	1.62	1.6	1.57
	0.95	1.884	1.876	1.86	1.83	1.787
	0.99	2.465	2.455	2.425	2.37	2.295
15	0.75	1.287	1.282	1.278	1.267	1.257
	0.90	1.605	1.604	1.59	1.575	1.546
	0.95	1.843	1.835	1.82	1.794	1.752
	0.99	2.39	2.38	2.35	2.3	2.23

Εκατοστιαία σημεία της γενικευμένης F κατανομής (συνέχεια)

k	p	ρ: 0.5	0.6	0.7	0.8	0.9
9	0.75	1.325	1.296	1.26	1.215	1.153
	0.90	1.71	1.64	1.56	1.455	1.315
	0.95	2.00	1.95	1.775	1.63	1.427
	0.99	2.72	2.53	2.31	2.03	1.68
10	0.75	1.305	1.277	1.245	1.203	1.143
	0.90	1.665	1.6	1.525	1.425	1.295
	0.95	1.93	1.837	1.725	1.585	1.4
	0.99	2.565	2.405	2.2	1.95	1.63
11	0.75	1.287	1.264	1.233	1.193	1.136
	0.90	1.62	1.565	1.49	1.4	1.277
	0.95	1.87	1.783	1.677	1.55	1.375
	0.99	2.45	2.3	2.11	1.88	1.59
12	0.75	1.274	1.25	1.221	1.183	1.13
	0.90	1.59	1.535	1.465	1.38	1.265
	0.95	1.815	1.735	1.635	1.515	1.355
	0.99	2.35	2.21	2.04	1.825	1.555
13	0.75	1.263	1.239	1.213	1.177	1.124
	0.90	1.56	1.51	1.44	1.36	1.253
	0.95	1.775	1.697	1.605	1.49	1.338
	0.99	2.27	2.135	1.975	1.78	1.525
14	0.75	1.25	1.227	1.203	1.167	1.117
	0.90	1.53	1.485	1.423	1.345	1.24
	0.95	1.733	1.663	1.577	1.47	1.324
	0.99	2.195	2.075	1.925	1.74	1.497
15	0.75	1.242	1.222	1.195	1.162	1.115
	0.90	1.51	1.465	1.405	1.33	1.31
	0.95	1.7	1.63	1.552	1.453	1.31
	0.99	2.135	2.025	1.88	1.705	1.475

Εκατοστιαία σημεία της γενικευμένης F κατανομής (συνέχεια)

k	p	ρ: 0.0	0.1	0.2	0.3	0.4
16	0.75	1.272	1.271	1.267	1.259	1.247
	0.90	1.585	1.58	1.57	1.55	1.525
	0.95	1.805	1.798	1.783	1.757	1.72
	0.99	2.32	2.31	2.285	2.235	2.17
17	0.75	1.263	1.262	1.257	1.247	1.239
	0.90	1.56	1.553	1.546	1.53	1.505
	0.95	1.775	1.767	1.753	1.727	1.697
	0.99	2.26	2.25	2.225	2.18	2.117
18	0.75	1.255	1.253	1.247	1.243	1.232
	0.90	1.54	1.535	1.525	1.510	1.486
	0.95	1.745	1.74	1.723	1.697	1.667
	0.99	2.208	2.195	2.172	2.13	2.07
19	0.75	1.247	1.246	1.243	1.233	1.223
	0.90	1.52	1.519	1.51	1.495	1.471
	0.95	1.717	1.711	1.697	1.676	1.644
	0.99	2.16	2.15	2.127	2.086	2.03
20	0.75	1.239	1.238	1.233	1.227	1.217
	0.90	1.505	1.504	1.495	1.48	1.456
	0.95	1.695	1.69	1.676	1.653	1.624
	0.99	2.115	2.105	2.085	2.046	1.994
21	0.75	1.233	1.231	1.229	1.222	1.212
	0.90	1.49	1.489	1.48	1.465	1.44
	0.95	1.672	1.667	1.654	1.633	1.604
	0.99	2.075	2.07	2.049	2.01	1.956
22	0.75	1.227	1.226	1.223	1.216	1.208
	0.90	1.475	1.474	1.466	1.451	1.43
	0.95	1.654	1.647	1.635	1.613	1.584
	0.99	2.04	2.034	2.01	1.976	1.925
23	0.75	1.223	1.221	1.218	1.212	1.203
	0.90	1.465	1.460	1.455	1.440	1.567
	0.95	1.633	1.629	1.617	1.597	1.567
	0.99	2.005	2.000	1.98	1.946	1.897

Εκατοστιαία σημεία της γενικευμένης F κατανομής (συνέχεια)

k	p	ρ: 0.5	0.6	0.7	0.8	0.9
16	0.75	1.232	1.212	1.187	1.157	1.112
	0.90	1.49	1.445	1.39	1.32	1.225
	0.95	1.675	1.61	1.527	1.427	1.297
	0.99	2.08	1.975	1.84	1.675	1.46
17	0.75	1.223	1.207	1.182	1.153	1.103
	0.90	1.471	1.43	1.376	1.307	1.216
	0.95	1.644	1.582	1.508	1.414	1.287
	0.99	2.035	1.935	1.805	1.645	1.437
18	0.75	1.217	1.197	1.177	1.143	1.103
	0.90	1.455	1.415	1.364	1.297	1.207
	0.95	1.620	1.563	1.493	1.397	1.277
	0.99	1.99	1.895	1.773	1.62	1.418
19	0.75	1.213	1.193	1.173	1.143	1.103
	0.90	1.44	1.402	1.351	1.287	1.203
	0.95	1.59	1.543	1.472	1.387	1.27
	0.99	1.955	1.86	1.744	1.599	1.41
20	0.75	1.203	1.187	1.167	1.137	1.097
	0.90	1.426	1.39	1.341	1.28	1.197
	0.95	1.576	1.527	1.46	1.375	1.262
	0.99	1.92	1.83	1.72	1.58	1.395
21	0.75	1.199	1.183	1.162	1.133	1.097
	0.90	1.415	1.377	1.331	1.274	1.193
	0.95	1.564	1.511	1.447	1.362	1.254
	0.99	1.89	1.801	1.695	1.56	1.384
22	0.75	1.194	1.179	1.157	1.132	1.093
	0.90	1.404	1.379	1.323	1.353	1.187
	0.95	1.549	1.498	1.434	1.353	1.247
	0.99	1.86	1.775	1.675	1.544	1.374
23	0.75	1.187	1.172	1.154	1.127	1.093
	0.90	1.391	1.358	1.315	1.259	1.183
	0.95	1.531	1.484	1.424	1.344	1.242
	0.99	1.835	1.754	1.654	1.53	1.364

Εκατοστιαία σημεία της γενικευμένης F κατανομης(συνέχεια)

k	p	ρ: 0.0	0.1	0.2	0.3	0.4
24	0.75	1.217	1.215	1.211	1.205	1.197
	0.90	1.454	1.450	1.442	1.428	1.408
	0.95	1.615	1.612	1.6	1.581	1.553
	0.99	1.978	1.972	1.952	1.918	1.872
25	0.75	1.212	1.211	1.207	1.201	1.193
	0.90	1.442	1.44	1.432	1.418	1.4
	0.95	1.6	1.596	1.585	1.566	1.54
	0.99	1.95	1.944	1.924	1.892	1.848
26	0.75	1.208	1.205	1.203	1.197	1.189
	0.90	1.432	1.43	1.422	1.408	1.39
	0.95	1.585	1.581	1.57	1.552	1.526
	0.99	1.924	1.918	1.90	1.868	1.824
27	0.75	1.203	1.202	1.199	1.193	1.183
	0.90	1.422	1.42	1.412	1.4	1.382
	0.95	1.57	1.566	1.558	1.54	1.514
	0.99	1.9	1.894	1.876	1.846	1.804
28	0.75	1.199	1.198	1.196	1.189	1.181
	0.90	1.412	1.410	1.402	1.39	1.372
	0.95	1.558	1.556	1.544	1.528	1.502
	0.99	1.878	1.872	1.854	1.826	1.784
29	0.75	1.197	1.194	1.191	1.185	1.178
	0.90	1.404	1.402	1.394	1.382	1.366
	0.95	1.546	1.543	1.532	1.516	1.492
	0.99	1.856	1.852	1.834	1.806	1.766
30	0.75	1.192	1.191	1.187	1.182	1.174
	0.90	1.396	1.394	1.386	1.375	1.358
	0.95	1.534	1.531	1.522	1.505	1.482
	0.99	1.838	1.832	1.816	1.788	1.748

Εκατοστιαία σημεία της γενικευμένης F κατανομης(συνέχεια)

k	p	ρ: 0.5	0.6	0.7	0.8	0.9
24	0.75	1.185	1.17	1.151	1.125	1.089
	0.90	1.382	1.35	1.306	1.252	1.178
	0.95	1.516	1.469	1.412	1.336	1.236
	0.99	1.810	1.732	1.634	1.512	1.352
25	0.75	1.181	1.166	1.147	1.123	1.088
	0.90	1.374	1.34	1.3	1.246	1.174
	0.95	1.504	1.458	1.401	1.328	1.229
	0.99	1.788	1.712	1.618	1.5	1.344
26	0.75	1.177	1.161	1.144	1.119	1.086
	0.90	1.366	1.344	1.292	1.240	1.17
	0.95	1.491	1.447	1.390	1.320	1.224
	0.99	1.766	1.694	1.602	1.488	1.336
27	0.75	1.174	1.159	1.141	1.117	1.081
	0.90	1.356	1.326	1.286	1.236	1.166
	0.95	1.48	1.437	1.383	1.314	1.22
	0.99	1.748	1.676	1.588	1.476	1.328
28	0.75	1.17	1.156	1.138	1.115	1.081
	0.90	1.35	1.32	1.28	1.23	1.163
	0.95	1.47	1.426	1.374	1.307	1.215
	0.99	1.73	1.66	1.574	1.464	1.32
29	0.75	1.167	1.153	1.136	1.113	1.081
	0.90	1.342	1.312	1.274	1.226	1.16
	0.95	1.459	1.418	1.367	1.302	1.211
	0.99	1.712	1.645	1.561	1.455	1.314
30	0.75	1.164	1.15	1.131	1.111	1.079
	0.90	1.336	1.306	1.27	1.222	1.157
	0.95	1.45	1.41	1.359	1.296	1.207
	0.99	1.696	1.632	1.55	1.446	1.308

ΠΑΡΑΡΤΗΜΑ ΙΙΙ

Περιοχές και εκατοστιαία σημεία της
folded t κατανομής με συνάρτηση πυκνότητας πιθανότητας

$$f_X(x) = 2\,\frac{\Gamma\left[\frac{\nu+1}{2}\right]}{\Gamma\left[\frac{\nu}{2}\right]\sqrt{\nu\pi}}\left[1+\frac{x^2}{\nu}\right]^{-\frac{\nu+1}{2}} \qquad \nu\in\mathbb{N}\ ,\ x>0$$

Περιοχές πιθανοτήτων της folded t κατανομής

t_f \ ν	1	2	3	4	5	6
0.0	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
0.1	0.06345	0.07534	0.07334	0.07484	0.07576	0.07637
0.2	0.12566	0.14002	0.14572	0.14876	0.15063	0.1519
0.3	0.18554	0.20751	0.21623	0.22087	0.22375	0.22570
0.4	0.24223	0.27216	0.28406	0.29040	0.29432	0.29699
0.5	0.29516	0.33333	0.34855	0.35667	0.36170	0.36511
0.6	0.34404	0.39056	0.40919	0.41915	0.42533	0.42954
0.7	0.38880	0.44360	0.46567	0.47749	0.48485	0.48985
0.8	0.42955	0.49236	0.51780	0.53147	0.53998	0.54578
0.9	0.46652	0.53689	0.56554	0.58099	0.59062	0.59720
1.0	0.50000	0.57735	0.60899	0.62609	0.63678	0.64408
1.1	0.53029	0.61396	0.64831	0.66691	0.67854	0.68650
1.2	0.55771	0.64699	0.68373	0.70364	0.71610	0.72463
1.3	0.58257	0.67675	0.71553	0.73654	0.74969	0.75869
1.4	0.60513	0.70352	0.74399	0.76589	0.77959	0.78895
1.5	0.62566	0.72760	0.76941	0.79200	0.80609	0.81571
1.6	0.64438	0.74926	0.79209	0.81515	0.82950	0.83928
1.7	0.66149	0.76876	0.81230	0.83564	0.85012	0.85996
1.8	0.67717	0.78633	0.83032	0.85376	0.86824	0.87804
1.9	0.69157	0.80218	0.84636	0.86976	0.88413	0.89383
2.0	0.70483	0.81649	0.86067	0.88388	0.89806	0.90757
2.1	0.71707	0.82945	0.87343	0.89634	0.91024	0.91951
2.2	0.72840	0.84119	0.88482	0.90734	0.92090	0.92989
2.3	0.73890	0.85185	0.89501	0.91706	0.93022	0.93889
2.4	0.74866	0.86154	0.90412	0.92564	0.93837	0.94670
2.5	0.75776	0.87038	0.91229	0.93323	0.94550	0.95347
2.6	0.76624	0.87845	0.91962	0.93995	0.95175	0.95933
2.7	0.77418	0.88584	0.92621	0.94590	0.95721	0.96441
2.8	0.78162	0.89260	0.93214	0.95118	0.96200	0.96883
2.9	0.78860	0.89881	0.93749	0.95588	0.96620	0.97266
3.0	0.79516	0.90453	0.94233	0.96005	0.96990	0.97599
3.1	0.80134	0.90979	0.94670	0.96377	0.97314	0.97888
3.2	0.80717	0.91465	0.95066	0.96709	0.97600	0.98139
3.3	0.81268	0.91915	0.95426	0.97006	0.97852	0.98359
3.4	0.81789	0.92331	0.95753	0.97272	0.98074	0.98550
3.5	0.82282	0.92717	0.96051	0.97510	0.98271	0.98717
3.6	0.82750	0.93075	0.96323	0.97724	0.98445	0.98863
3.7	0.83195	0.93409	0.96572	0.97916	0.98600	0.98991
3.8	0.83618	0.93720	0.96799	0.98089	0.98737	0.99103
3.9	0.84020	0.94010	0.97007	0.98245	0.98859	0.99201

Περιοχές πιθανοτήτων της folded t κατανομης (συνέχεια)

$t_z \backslash \nu$	1	2	3	4	5	6
4.0	0.84404	0.94280	0.97199	0.98386	0.98967	0.9928
4.1	0.84770	0.94534	0.97374	0.98514	0.99064	0.9936
4.2	0.85119	0.94771	0.97536	0.98630	0.99151	0.9943
4.3	0.85453	0.94994	0.97685	0.98735	0.99228	0.9949
4.4	0.85773	0.95203	0.97823	0.98830	0.99297	0.9954
4.5	0.86079	0.95399	0.97950	0.98917	0.99360	0.9958
4.6	0.86372	0.95584	0.98068	0.98996	0.99415	0.9963
4.7	0.86653	0.95758	0.98177	0.99069	0.99466	0.9966
4.8	0.86924	0.95923	0.98279	0.99135	0.99511	0.9969
4.9	0.87183	0.96078	0.98373	0.99195	0.99552	0.9972
5.0	0.87433	0.96225	0.98460	0.99250	0.99589	0.9975
5.1	0.87673	0.96363	0.98542	0.99301	0.99622	0.9977
5.2	0.87904	0.96495	0.98618	0.99348	0.99653	0.9979
5.3	0.88127	0.96619	0.98688	0.99391	0.99680	0.9981
5.4	0.88342	0.96737	0.98755	0.99430	0.99705	0.9983
5.5	0.88550	0.96849	0.98817	0.99467	0.99728	0.9984
5.6	0.88750	0.96956	0.98874	0.99500	0.99749	0.9986
5.7	0.88943	0.97057	0.98929	0.99531	0.99768	0.9987
5.8	0.89130	0.97153	0.98980	0.99560	0.99785	0.9988
5.9	0.89311	0.97245	0.99027	0.99587	0.99800	0.9989
6.0	0.89486	0.97332	0.99072	0.99611	0.99815	0.9990

Περιοχές πιθανοτήτων της folded t κατανομης (συνέχεια)

$t_c \backslash \nu$	7	8	9	10	11	12
0.0	0.00000	0.00000	0.00000	0.00000	0.00000	0.0000
0.1	0.07685	0.07719	0.07746	0.07767	0.07785	0.0780
0.2	0.15283	0.15352	0.15407	0.15451	0.15486	0.1551
0.3	0.22710	0.22817	0.22900	0.22967	0.23022	0.2306
0.4	0.29891	0.30037	0.30151	0.30243	0.30318	0.3038
0.5	0.36759	0.36946	0.37092	0.37210	0.37307	0.3738
0.6	0.43258	0.43488	0.43669	0.43814	0.43933	0.4403
0.7	0.49348	0.49622	0.49838	0.50011	0.50153	0.5027
0.8	0.54999	0.55318	0.55568	0.55769	0.55935	0.5607
0.9	0.60197	0.60559	0.60843	0.61072	0.61260	0.6141
1.0	0.64938	0.65340	0.65656	0.65910	0.66119	0.6629
1.1	0.69228	0.69667	0.70011	0.70289	0.70517	0.7070
1.2	0.73082	0.73553	0.73922	0.74220	0.74465	0.7467
1.3	0.76523	0.77019	0.77409	0.77723	0.77981	0.7819
1.4	0.79575	0.80092	0.80497	0.80823	0.81091	0.8131
1.5	0.82270	0.82799	0.83214	0.83549	0.83824	0.8405
1.6	0.84636	0.85173	0.85593	0.85931	0.86209	0.8644
1.7	0.86707	0.87244	0.87665	0.88003	0.88280	0.8851
1.8	0.88511	0.89044	0.89460	0.89794	0.90068	0.9029
1.9	0.90079	0.90603	0.91011	0.91337	0.91604	0.9182
2.0	0.91438	0.91948	0.92344	0.92661	0.92919	0.9313
2.1	0.92612	0.93106	0.93488	0.93792	0.94039	0.9424
2.2	0.93626	0.94100	0.94465	0.94755	0.94991	0.9518
2.3	0.94500	0.94952	0.95300	0.95574	0.95796	0.9598
2.4	0.95253	0.95682	0.96010	0.96268	0.96476	0.9664
2.5	0.95900	0.96305	0.96613	0.96855	0.97049	0.9720
2.6	0.96457	0.96838	0.97126	0.97350	0.97530	0.9767
2.7	0.96936	0.97292	0.97560	0.97768	0.97934	0.9806
2.8	0.97347	0.97680	0.97928	0.98120	0.98272	0.9839
2.9	0.97701	0.98010	0.98240	0.98416	0.98555	0.9866
3.0	0.98005	0.98292	0.98504	0.98665	0.98792	0.9889
3.1	0.98267	0.98533	0.98727	0.98874	0.98989	0.9908
3.2	0.98493	0.98738	0.98916	0.99050	0.99154	0.9923
3.3	0.98687	0.98914	0.99077	0.99198	0.99292	0.9936
3.4	0.98855	0.99063	0.99212	0.99322	0.99407	0.9947
3.5	0.99000	0.99192	0.99327	0.99427	0.99502	0.9956
3.6	0.99126	0.99301	0.99425	0.99515	0.99583	0.9963
3.7	0.99234	0.99395	0.99507	0.99589	0.99649	0.9969
3.8	0.99328	0.99476	0.99578	0.99651	0.99705	0.9974
3.9	0.99410	0.99545	0.99638	0.99703	0.99752	0.9978

Περιοχές πιθανοτήτων της folded t κατανομής (συνέχεια)

t_c \ ν	7	8	9	10	11	12
4.0	0.99480	0.99605	0.99688	0.99748	0.99791	0.99823
4.1	0.99542	0.99656	0.99732	0.99785	0.99824	0.99852
4.2	0.99596	0.99700	0.99769	0.99817	0.99851	0.99876
4.3	0.99643	0.99738	0.99800	0.99843	0.99874	0.99896
4.4	0.99684	0.99771	0.99827	0.99866	0.99893	0.99913
4.5	0.99720	0.99799	0.99851	0.99885	0.99909	0.99927
4.6	0.99751	0.99824	0.99870	0.99901	0.99923	0.99938
4.7	0.99779	0.99845	0.99887	0.99915	0.99934	0.99948
4.8	0.99803	0.99864	0.99902	0.99927	0.99944	0.99956
4.9	0.99824	0.99880	0.99915	0.99937	0.99952	0.99963
5.0	0.99843	0.99894	0.99926	0.99946	0.99959	0.99969
5.1	0.99860	0.99907	0.99935	0.99953	0.99965	0.99973
5.2	0.99874	0.99917	0.99943	0.99959	0.99970	0.99977
5.3	0.99887	0.99927	0.99950	0.99965	0.99974	0.99981
5.4	0.99899	0.99935	0.99956	0.99969	0.99978	0.99983
5.5	0.99909	0.99942	0.99961	0.99973	0.99981	0.99986
5.6	0.99918	0.99948	0.99966	0.99977	0.99983	0.99988
5.7	0.99926	0.99954	0.99970	0.99980	0.99986	0.99990
5.8	0.99933	0.99959	0.99974	0.99982	0.99988	0.99991
5.9	0.99940	0.99963	0.99977	0.99984	0.99989	0.99992
6.0	0.99945	0.99967	0.99979	0.99986	0.99991	0.99993

Περιοχές πιθανοτήτων της folded t κατανομης (συνέχεια)

t_f \ ν	13	14	15	16	17	18
0.0	0.00000	0.00000	0.00000	0.00000	0.00000	0.0000
0.1	0.07812	0.07823	0.07833	0.07841	0.07848	0.0785
0.2	0.15542	0.15564	0.15583	0.15599	0.15614	0.1562
0.3	0.23107	0.23141	0.23170	0.23196	0.23218	0.2323
0.4	0.30434	0.30480	0.30520	0.30555	0.30586	0.3061
0.5	0.37456	0.37515	0.37566	0.37611	0.37651	0.3768
0.6	0.44117	0.44190	0.44253	0.44309	0.44357	0.4440
0.7	0.50373	0.50460	0.50535	0.50601	0.50660	0.5071
0.8	0.56191	0.56292	0.56380	0.56457	0.56525	0.5658
0.9	0.61551	0.61665	0.61765	0.61853	0.61930	0.6199
1.0	0.66443	0.66571	0.66682	0.66780	0.66866	0.6694
1.1	0.70871	0.71011	0.71132	0.71239	0.71333	0.7141
1.2	0.74844	0.74994	0.75125	0.75239	0.75340	0.7543
1.3	0.78382	0.78540	0.78677	0.78798	0.78904	0.7899
1.4	0.81507	0.81671	0.81814	0.81939	0.82050	0.8214
1.5	0.84249	0.84417	0.84563	0.84691	0.84804	0.8490
1.6	0.86639	0.86808	0.86955	0.87084	0.87198	0.8729
1.7	0.88708	0.88877	0.89023	0.89151	0.89264	0.8936
1.8	0.90490	0.90655	0.90799	0.90925	0.91036	0.9113
1.9	0.92015	0.92177	0.92316	0.92439	0.92546	0.9264
2.0	0.93315	0.93471	0.93605	0.93722	0.93826	0.9391
2.1	0.94418	0.94566	0.94694	0.94805	0.94904	0.9499
2.2	0.95350	0.95489	0.95610	0.95715	0.95807	0.9588
2.3	0.96134	0.96264	0.96377	0.96475	0.96561	0.9663
2.4	0.96791	0.96913	0.97017	0.97108	0.97187	0.9725
2.5	0.97341	0.97453	0.97549	0.97632	0.97705	0.9776
2.6	0.97799	0.97902	0.97990	0.98065	0.98131	0.9819
2.7	0.98180	0.98274	0.98354	0.98422	0.98482	0.9853
2.8	0.98497	0.98581	0.98654	0.98715	0.98769	0.9881
2.9	0.98759	0.98835	0.98900	0.98955	0.99003	0.9904
3.0	0.98976	0.99044	0.99102	0.99152	0.99194	0.9923
3.1	0.99155	0.99216	0.99268	0.99312	0.99349	0.9938
3.2	0.99303	0.99357	0.99403	0.99442	0.99475	0.9950
3.3	0.99425	0.99473	0.99514	0.99548	0.99576	0.9960
3.4	0.99525	0.99568	0.99604	0.99633	0.99659	0.9968
3.5	0.99608	0.99646	0.99677	0.99703	0.99725	0.9974
3.6	0.99676	0.99710	0.99737	0.99760	0.99779	0.9979
3.7	0.99732	0.99762	0.99786	0.99805	0.99822	0.9983
3.8	0.99779	0.99804	0.99825	0.99842	0.99856	0.9986
3.9	0.99817	0.99839	0.99857	0.99872	0.99884	0.9989

Περιοχές πιθανοτήτων της folded t κατανομης (συνέχεια)

t_f \ ν	13	14	15	16	17	18
4.0	0.99848	0.99868	0.99884	0.99896	0.99907	0.9991
4.1	0.99874	0.99891	0.99905	0.99916	0.99925	0.9993
4.2	0.99896	0.99910	0.99922	0.99932	0.99939	0.9994
4.3	0.99913	0.99926	0.99936	0.99944	0.99951	0.9995
4.4	0.99928	0.99939	0.99948	0.99955	0.99960	0.9996
4.5	0.99940	0.99950	0.99957	0.99963	0.99968	0.9997
4.6	0.99950	0.99958	0.99965	0.99970	0.99974	0.9997
4.7	0.99958	0.99965	0.99971	0.99975	0.99979	0.9998
4.8	0.99965	0.99971	0.99976	0.99980	0.99983	0.9998
4.9	0.99970	0.99976	0.99980	0.99983	0.99986	0.9998
5.0	0.99975	0.99980	0.99984	0.99986	0.99989	0.9999
5.1	0.99979	0.99983	0.99986	0.99989	0.99991	0.9999
5.2	0.99982	0.99986	0.99989	0.99991	0.99992	0.9999
5.3	0.99985	0.99988	0.99991	0.99992	0.99994	0.9999
5.4	0.99987	0.99990	0.99992	0.99994	0.99995	0.9999
5.5	0.99989	0.99992	0.99993	0.99995	0.99996	0.9999
5.6	0.99991	0.99993	0.99994	0.99996	0.99996	0.9999
5.7	0.99992	0.99994	0.99995	0.99996	0.99997	0.9999
5.8	0.99993	0.99995	0.99996	0.99997	0.99997	0.9999
5.9	0.99994	0.99996	0.99997	0.99997	0.99998	0.9999
6.0	0.99995	0.99996	0.99997	0.99998	0.99998	0.9999

Περιοχές τιθανοτήτων της folded t κατανομης (συνέχεια)

$_f$\ ν	19	20	21	22	23	24
.0	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
.1	0.07860	0.07866	0.07807	0.07874	0.08789	0.07882
.2	0.15639	0.15649	0.15659	0.15668	0.15676	0.15683
.3	0.23256	0.23272	0.23287	0.23300	0.23312	0.23323
.4	0.30638	0.30660	0.30680	0.30698	0.30715	0.30730
.5	0.37718	0.37746	0.37772	0.37796	0.37817	0.37837
.6	0.44440	0.44475	0.44507	0.44536	0.44562	0.44587
.7	0.50758	0.50800	0.50839	0.50873	0.50905	0.50934
.8	0.56640	0.56688	0.56733	0.56773	0.56810	0.56844
.9	0.62061	0.62117	0.62167	0.62213	0.62255	0.62293
.0	0.67012	0.67074	0.67130	0.67181	0.67228	0.67271
.1	0.71492	0.71560	0.71621	0.71677	0.71728	0.71775
.2	0.75511	0.75583	0.75649	0.75709	0.75764	0.75814
.3	0.79084	0.79161	0.79231	0.79294	0.79352	0.79405
.4	0.82236	0.82316	0.82388	0.82454	0.82514	0.82569
.5	0.84995	0.85076	0.85150	0.85217	0.85278	0.85334
.6	0.87390	0.87472	0.87546	0.87613	0.87675	0.87731
.7	0.89455	0.89536	0.89610	0.89677	0.89738	0.89794
.8	0.91223	0.91303	0.91375	0.91440	0.91500	0.91555
.9	0.92726	0.92805	0.92874	0.92936	0.92995	0.93048
.0	0.93999	0.94073	0.94139	0.94200	0.94255	0.94306
.1	0.95068	0.95138	0.95201	0.95258	0.95310	0.95357
.2	0.95961	0.96027	0.96085	0.96139	0.96187	0.96232
.3	0.96704	0.96765	0.96819	0.96869	0.96914	0.96955
.4	0.97319	0.97375	0.97425	0.97470	0.97511	0.97549
.5	0.97825	0.97876	0.97922	0.97963	0.98000	0.98034
.6	0.98241	0.98287	0.98328	0.98365	0.98399	0.98429
.7	0.98581	0.98622	0.98659	0.98692	0.98722	0.98749
.8	0.98857	0.98894	0.98927	0.98956	0.98982	0.99007
.9	0.99082	0.99114	0.99143	0.99169	0.99192	0.99213
.0	0.99263	0.99292	0.99317	0.99340	0.99360	0.99379
.1	0.99410	0.99435	0.99457	0.99477	0.99495	0.99511
.2	0.99528	0.99550	0.99569	0.99586	0.99602	0.99615
.3	0.99623	0.99642	0.99659	0.99673	0.99686	0.99698
.4	0.99699	0.99715	0.99730	0.99742	0.99754	0.99764
.5	0.99760	0.99774	0.99786	0.99797	0.99807	0.99815
.6	0.99809	0.99821	0.99831	0.99840	0.99848	0.99856
.7	0.99848	0.99858	0.99867	0.99874	0.99881	0.99887
.8	0.99879	0.99887	0.99895	0.99901	0.99907	0.99911
.9	0.99903	0.99911	0.99917	0.99923	0.99927	0.99931

Περιοχές πιθανοτήτων της folded t κατανομης (συνέχεια)

$t_z \setminus \nu$	19	20	21	22	23	24
4.0	0.99923	0.99929	0.99935	0.99939	0.99943	0.99947
4.1	0.99939	0.99944	0.99948	0.99952	0.99956	0.99959
4.2	0.99951	0.99955	0.99959	0.99962	0.99965	0.99968
4.3	0.99961	0.99965	0.99968	0.99971	0.99973	0.99975
4.4	0.99969	0.99972	0.99975	0.99977	0.99979	0.99980
4.5	0.99975	0.99978	0.99980	0.99982	0.99983	0.99985
4.6	0.99980	0.99982	0.99984	0.99986	0.99987	0.99988
4.7	0.99984	0.99986	0.99987	0.99989	0.99990	0.99991
4.8	0.99987	0.99989	0.99990	0.99991	0.99992	0.99993
4.9	0.99990	0.99991	0.99992	0.99993	0.99994	0.99994
5.0	0.99992	0.99993	0.99994	0.99994	0.99995	0.99995
5.1	0.99993	0.99994	0.99995	0.99995	0.99996	0.99996
5.2	0.99994	0.99995	0.99996	0.99996	0.99997	0.99997
5.3	0.99995	0.99996	0.99997	0.99997	0.99997	0.99998
5.4	0.99996	0.99997	0.99997	0.99997	0.99998	0.99998
5.5	0.99997	0.99997	0.99998	0.99998	0.99998	0.99998
5.6	0.99997	0.99998	0.99998	0.99998	0.99998	0.99999
5.7	0.99998	0.99998	0.99998	0.99999	0.99999	0.99999
5.8	0.99998	0.99998	0.99999	0.99999	0.99999	0.99999
5.9	0.99998	0.99999	0.99999	0.99999	0.99999	0.99999
6.0	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999

Περιοχές πιθανοτήτων της folded t κατανομής (συνέχεια)

x \ ν	24	25	26	27	28	29	30
.0	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.0000
.1	0.07882	0.07885	0.07888	0.07891	0.07894	0.07896	0.0789
.2	0.15683	0.15690	0.15696	0.15702	0.15707	0.15712	0.1571
.3	0.23323	0.23334	0.23343	0.23352	0.23360	0.23368	0.2337
.4	0.30730	0.30744	0.30757	0.30769	0.30780	0.30791	0.3080
.5	0.37837	0.37855	0.37871	0.37887	0.37901	0.37915	0.3792
.6	0.44587	0.44609	0.44629	0.44649	0.44666	0.44683	0.4469
.7	0.50934	0.50960	0.50985	0.51008	0.51029	0.51049	0.5106
.8	0.56844	0.56875	0.56904	0.56930	0.56955	0.56978	0.5699
.9	0.62293	0.62329	0.62361	0.62392	0.62420	0.62446	0.6247
.0	0.67271	0.67310	0.67347	0.67381	0.67412	0.67441	0.6746
.1	0.71775	0.71818	0.71858	0.71895	0.71930	0.71962	0.7199
.2	0.75814	0.75861	0.75904	0.75943	0.75980	0.76014	0.7604
.3	0.79405	0.79454	0.79499	0.79541	0.79579	0.79616	0.7964
.4	0.82569	0.82619	0.82666	0.82710	0.82750	0.82787	0.8282
.5	0.85334	0.85386	0.85433	0.85478	0.85519	0.85557	0.8559
.6	0.87731	0.87783	0.87831	0.87876	0.87917	0.87956	0.8799
.7	0.89794	0.89845	0.89893	0.89937	0.89978	0.90016	0.9005
.8	0.91555	0.91605	0.91652	0.91695	0.91735	0.91772	0.9180
.9	0.93046	0.93097	0.93142	0.93183	0.93222	0.93258	0.9329
.0	0.94306	0.94352	0.94395	0.94434	0.94471	0.94505	0.9453
.1	0.95357	0.95401	0.95441	0.95479	0.95513	0.95545	0.9557
.2	0.96232	0.96273	0.96310	0.96345	0.96377	0.96407	0.9643
.3	0.96955	0.96992	0.97027	0.97059	0.97088	0.97116	0.971
.4	0.97549	0.97583	0.97614	0.97644	0.97671	0.97696	0.9771
.5	0.98034	0.98065	0.98094	0.98120	0.98144	0.98167	0.9818
.6	0.98429	0.98457	0.98483	0.98506	0.98528	0.98548	0.9856
.7	0.98749	0.98774	0.98796	0.98817	0.98837	0.98854	0.9887
.8	0.99007	0.99029	0.99049	0.99067	0.99084	0.99100	0.9911
.9	0.99213	0.99233	0.99250	0.99266	0.99281	0.99295	0.9930
.0	0.99379	0.99396	0.99411	0.99425	0.99438	0.99450	0.9946
.1	0.99511	0.99525	0.99538	0.99550	0.99562	0.99572	0.9958
.2	0.99615	0.99628	0.99639	0.99650	0.99659	0.99668	0.9967
.3	0.99698	0.99709	0.99719	0.99727	0.99735	0.99743	0.9975
.4	0.99764	0.99773	0.99781	0.99788	0.99795	0.99801	0.9980
.5	0.99815	0.99823	0.99830	0.99836	0.99842	0.99847	0.9985
.6	0.99856	0.99862	0.99868	0.99873	0.99878	0.99882	0.9988
.7	0.99887	0.99893	0.99898	0.99902	0.99906	0.99910	0.9991
.8	0.99912	0.99917	0.99921	0.99925	0.99928	0.99931	0.9993
.9	0.99932	0.99935	0.99939	0.99942	0.99945	0.99947	0.9994

Περιοχές πιθανοτήτων της folded t κατανομης (συνέχεια)

\ ν	24	25	26	27	28	29	30
0	0.99947	0.99950	0.99953	0.99955	0.99957	0.99959	0.999
1	0.99959	0.99961	0.99963	0.99966	0.99967	0.99969	0.999
2	0.99968	0.99970	0.99972	0.99973	0.99975	0.99976	0.999
3	0.99975	0.99977	0.99978	0.99980	0.99981	0.99982	0.999
4	0.99980	0.99982	0.99983	0.99984	0.99985	0.99986	0.999
5	0.99985	0.99986	0.99987	0.99988	0.99989	0.99989	0.999
6	0.99988	0.99989	0.99990	0.99991	0.99991	0.99992	0.999
7	0.99991	0.99991	0.99992	0.99993	0.99993	0.99994	0.999
8	0.99993	0.99993	0.99994	0.99994	0.99995	0.99995	0.999
9	0.99994	0.99995	0.99995	0.99996	0.99996	0.99996	0.999
0	0.99995	0.99996	0.99996	0.99996	0.99997	0.99997	0.999
1	0.99996	0.99997	0.99997	0.99997	0.99997	0.99998	0.999
2	0.99997	0.99997	0.99998	0.99998	0.99998	0.99998	0.999
3	0.99998	0.99998	0.99998	0.99998	0.99998	0.99998	0.999
4	0.99998	0.99998	0.99998	0.99998	0.99999	0.99999	0.999
5	0.99998	0.99998	0.99999	0.99999	0.99999	0.99999	0.999
6	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.999
7	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.999
8	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.999
9	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.999
∞	0.99999	0.99999	0.99999	0.99999	0.99999	0.99999	0.999

Εκατοστιαία σημεία της folded t κατανομης

ν	a=0.55	a=0.60	a=0.70	a=0.75	a=0.8
1	1.1708	1.3764	1.9626	2.4142	3.0777
2	0.9313	1.0606	1.3862	1.6036	1.8856
3	0.8661	0.9782	1.2498	1.4227	1.6377
4	0.8361	0.9407	1.1895	1.3445	1.5332
5	0.8188	0.9193	1.1557	1.3010	1.4759
6	0.8076	0.9055	1.1341	1.2734	1.4398
7	0.7997	0.8958	1.1191	1.2544	1.4149
8	0.7939	0.8887	1.1081	1.2404	1.3972
9	0.7894	0.8832	1.0997	1.2297	1.3836
10	0.7859	0.8788	1.0930	1.2213	1.3724
11	0.7830	0.8753	1.0876	1.2145	1.3635
12	0.7806	0.8724	1.0832	1.2089	1.3564
13	0.7786	0.8699	1.0795	1.2042	1.3504
14	0.7769	0.8678	1.0763	1.2002	1.3452
15	0.7754	0.8660	1.0735	1.1968	1.3406
16	0.7741	0.8645	1.0711	1.1938	1.3368
17	0.7730	0.8631	1.0690	1.1911	1.3334
18	0.7720	0.8618	1.0672	1.1888	1.3304
19	0.7711	0.8608	1.0655	1.1867	1.3277
20	0.7702	0.8598	1.0640	1.1848	1.3253
21	0.7695	0.8589	1.0627	1.1832	1.3232
22	0.7689	0.8581	1.0614	1.1816	1.3212
23	0.7682	0.8573	1.0603	1.1802	1.3195
24	0.7677	0.8567	1.0593	1.1790	1.3180
25	0.7672	0.8560	1.0584	1.1778	1.3166
26	0.7667	0.8555	1.0575	1.1767	1.3154
27	0.7663	0.8550	1.0567	1.1757	1.3143
28	0.7659	0.8545	1.0560	1.1748	1.3133
29	0.7655	0.8540	1.0553	1.1739	1.3125
30	0.7652	0.8536	1.0546	1.1731	1.3116
40	0.7626	0.8505	1.0500	1.1674	1.3062
60	0.7601	0.8447	1.0454	1.1617	1.2960
100	0.7576	0.8444	1.0409	1.1560	1.2866

Εκατοστιαία σημεία της folded t κατανομής (συνέχεια)

ν\	a=0.90	a=0.95	a=0.975	a=0.99	a=0.
1	6.3137	12.7062	25.4517	63.6567	127.
2	2.9199	4.3062	6.2053	9.9248	14.09
3	2.3534	3.1824	4.1765	5.8409	7.453
4	2.1325	2.7774	3.4954	4.6041	5.597
5	2.0156	2.5714	3.1644	4.0321	4.773
6	1.9437	2.4476	2.9696	3.7084	4.31
7	1.8951	2.3653	2.8420	3.5003	4.030
8	1.8600	2.3066	2.7522	3.3562	3.833
9	1.8336	2.2630	2.6857	3.2506	3.690
10	1.8129	2.2287	2.6344	3.1700	3.582
11	1.7963	2.2015	2.5937	3.1065	3.497
12	1.7827	2.1794	2.5606	3.0552	3.429
13	1.7713	2.1609	2.5332	3.0130	3.373
14	1.7617	2.1453	2.5101	2.9774	3.326
15	1.7535	2.1320	2.4904	2.9473	3.286
16	1.7463	2.1204	2.4734	2.9213	3.252
17	1.7400	2.1103	2.4586	2.8988	3.222
18	1.7345	2.1014	2.4455	2.8790	3.197
19	1.7295	2.0935	2.4340	2.8615	3.174
20	1.7251	2.0865	2.4236	2.8459	3.153
21	1.7211	2.0801	2.4144	2.8319	3.135
22	1.7175	2.0744	2.4065	2.8193	3.119
23	1.7143	2.0691	2.3845	2.8078	3.104
24	1.7113	2.0644	2.3915	2.7974	3.091
25	1.7085	2.0600	2.3851	2.7875	3.078
26	1.7060	2.0560	2.3793	2.7792	3.067
27	1.7037	2.0523	2.3739	2.7712	3.057
28	1.7015	2.0490	2.3690	2.7635	3.047
29	1.6995	2.0457	2.3643	2.7568	3.038
30	1.6976	2.0427	2.3600	2.7504	3.030
40	1.6842	2.0215	2.3294	2.7049	2.971
60	1.6710	2.0007	2.2995	2.6607	2.914
120	1.6580	1.9804	2.2703	2.6178	2.860

ΠΑΡΑΡΤΗΜΑ IV

ΑΠΟΤΕΛΕΣΜΑΤΑ ΜΙΑΣ ΜΗ ΠΑΡΑΜΕΤΡΙΚΗΣ ΜΕΘΟΔΟΥ ΑΞΙΟΛΟΓΗΣΗΣ ΤΗΣ ΠΡΟΒΛΕΠΤΙΚΗΣ ΙΚΑΝΟΤΗΤΑΣ ΓΡΑΜΜΙΚΩΝ ΜΟΝΤΕΛΩΝ

Y^0_{t+1} η παρατηρούμενη τιμή του μοντέλου την χρονική στιγμή t+1

$\hat{Y}^0_{t+1}$ η προβλεπόμενη τιμή του μοντέλου την χρονική στιγμή t+1

$$R = \begin{cases} 1 \text{ αν } \quad Y^0_{t+1} \in \hat{Y}^0_{t+1} \pm t_{1_t-m,\,1-a/2} \cdot S_{\hat{Y}^0_{t+1}} \\ 0 \text{ διαφορετικά} \end{cases}$$

όπου $S_{\hat{Y}^0_{t+1}} = S_t \sqrt{1+X^0_{t+1}(X_t X_t)^{-1} X^0_{t+1}}$

ΣΦΑΛΜΑ

ΠΙΝΑΚΑΣ Ι

CRD10

MODEL A

Δείκτης	Ετος	Y^0_{t+1}	$\hat{Y}^0_{t+1}$	$\hat{Y}^0_{t+1} \pm A^*$	R
1	1963	57.7	32.7	(-12.516,77.916)	1
2	1964	48.9	62.9	(37.121,88.678)	1
3	1965	61.2	49.1	(25.464,72.735)	1
4	1966	51.4	59.2	(41.931,76,468)	1
5	1967	57.4	44.2	(28.744,59.655)	1
6	1968	58.5	58.9	(44.582,73.217)	1
7	1969	66.6	56.2	(40.028,72.371)	1
8	1971	71.6	47.0	(29.197,64.802)	0
9	1972	69.6	65.2	(50.797,79.602)	1
10	1973	67.5	66.3	(51.632,80.697)	1
11	1974	45.5	64.5	(52.670,76.329)	0
12	1975	65.2	59.4	(45.026,73.773)	1
13	1976	70.6	49.3	(34.130,64.469)	0
14	1977	66.1	47.9	(32.036,73.763)	1
15	1978	66.9	67.0	(53.354,80.645)	1
16	1979	73.1	70.0	(56.375,83.624)	1
17	1980	49.3	63.2	(49.129,76.670)	1

όπου $A^* = t_{1_t-m,1-a/2}\, S_t\sqrt{1+X^0_{t+1}(X_t X_t)^{-1}X^0_{t+1}}$

ΠΙΝΑΚΑΣ ΙΙ

CRD20

MODEL A

Δείκτης	Ετος	Y^0_{t+1}	$\hat{Y}^0_{t+1}$	$\hat{Y}^0_{t+1} \pm A^*$	R
1	1963	52.3	48.2	(21.544,74.658)	1
2	1964	44.6	56.6	(40.444,72.755)	1
3	1965	56.8	42.6	(28.728,56.471)	0
4	1966	46.8	49.3	(36.623,61.796)	1
5	1967	50.9	42.2	(30.079,54.320)	1
6	1968	55.5	50.5	(40.154,60.845)	1
7	1969	59.5	60.9	(48.720,73.079)	1
8	1971	62.8	56.3	(42.251,70.348)	1
9	1972	61.8	60.6	(52.187,69.012)	1
10	1973	62.3	58.9	(50.603,67.196)	1
11	1974	38.9	59.0	(51.591,66.408)	0
12	1975	62.4	64.1	(45.585,82.614)	1
13	1976	67.6	57.8	(46.460,69.139)	1
14	1977	62.6	63.9	(52.396,75.403)	1
15	1978	61.7	65.4	(55.109,75.960)	1
16	1979	69.4	67.3	(58.494,76.105)	1
17	1980	54.0	63.7	(55.008,72.391)	0

όπου $A^* = t_{i_t - m, 1-a/2} \; S_t \sqrt{1 + X^0_{t+1} (X_t' X_t)^{-1} X^0_{t+1}}$

ΠΙΝΑΚΑΣ ΙΙΙ

CRDΞO

MODEL A

Δείκτης	Ετος	Y^0_{t+1}	$\hat{Y}^0_{t+1}$	$\hat{Y}^0_{t+1} \pm A^*$	R
1	1963	51.2	42.8	(6.281,79.318)	1
2	1964	39.3	54.8	(33.667,75.732)	1
3	1965	52.6	32.8	(15.635,49.964)	0
4	1966	48.3	51.0	(32.933,69.066)	1
5	1967	46.4	44.7	(29.742,59.657)	1
6	1968	54.2	46.9	(32.334,61.465)	1
7	1969	52.8	54.4	(37.438,71.361)	1
8	1971	55.1	57.7	(37.243,78.157)	1
9	1972	59.3	57.1	(46.250,67.949)	1
10	1973	58.1	59.9	(49.815,69.984)	1
11	1974	35.3	49.9	(38.859,60.940)	0
12	1975	52.4	50.7	(36.456,64.943)	1
13	1976	62.5	58.1	(45.738,70.461)	1
14	1977	64.9	60.7	(50.354,71.045)	1
15	1978	60.9	60.3	(51.086,69.513)	1
16	1979	67.5	66.4	(58.010,74.789)	1
17	1980	65.1	65.1	(57.010,73.159)	1

όπου $A^* = t_{t-m,1-a/2}\, S_t \sqrt{1+X^0_{t+1}(X_t'X_t)^{-1}X^0_{t+1}}$

ΠΙΝΑΚΑΣ IV

CRD40

MODEL A

Δείκτης	Ετος	Y^0_{t+1}	$\hat{Y}^0_{t+1}$	$\hat{Y}^0_{t+1} \pm A^*$	R
1	1963	60.1	72.5	(42.015,102.984)	1
2	1964	48.1	62.2	(44.465,80.534)	1
3	1965	64.4	46.4	(31.988,60.811)	0
4	1966	48.9	53.5	(40.336,66.663)	1
5	1967	57.0	52.9	(40.975,64.824)	1
6	1968	56.4	49.6	(38.058,61.141)	1
7	1969	65.8	55.4	(44.790,66.009)	1
8	1971	67.7	74.3	(56.932,89.667)	1
9	1972	69.6	64.8	(53.814,75.785)	1
10	1973	67.5	60.8	(50.368,71.231)	1
11	1974	43.0	50.1	(37.931,62.268)	1
12	1975	69.1	53.5	(42.428,64.571)	0
13	1976	70.8	74.7	(62.238,87.161)	1
14	1977	63.8	43.5	(22.634,64.365)	1
15	1978	73.1	79.2	(67.771,90.628)	1
16	1979	74.7	69.9	(57.907,81.892)	1
17	1980	59.1	71.6	(60.147,83.052)	0

όπου $A^* = t_{1_t - m, 1-a/2}\, S_t \sqrt{1+X^0_{t+1}(X_t X_t)^{-1} X^0_{t+1}}$

ΠΙΝΑΚΑΣ V

CRD50

MODEL A

Δείκτης	Ετος	Y^0_{t+1}	$\hat{Y}^0_{t+1}$	$\hat{Y}^0_{t+1} \pm A^*$	P
1	1963	56.4	66.3	(35.857,100.742)	1
2	1964	46.2	60.3	(43.488,77.111)	1
3	1965	62.1	36.7	(22.681,50.718)	0
4	1966	50.9	54.3	(39.708,68.891)	1
5	1967	52.0	50.1	(34.493,65.706)	1
6	1968	56.8	55.3	(42.129,68.470)	1
7	1969	67.5	55.4	(42.431,68.436)	1
8	1971	65.9	75.9	(58.792,93.007)	1
9	1972	67.3	62.0	(49.928,74.071)	1
10	1973	68.5	72.7	(60.776,84.623)	1
11	1974	47.5	64.7	(51.437,77.962)	0
12	1975	64.6	65.4	(54.632,76.167)	1
13	1976	72.5	75.3	(62.476,88.123)	1
14	1977	64.1	53.3	(34.907,71.697)	1
15	1978	73.7	64.0	(52.623,75.376)	1
16	1979	73.7	74.4	(59.730,89.069)	1
17	1980	67.7	66.4	(56.620,76.179)	1

όπου $A^* = t_{t-m,\,1-a/2}\, S_t \sqrt{1+X^0_{t+1}(X_t' X_t)^{-1} X^0_{t+1}}$

ΠΙΝΑΚΑΣ VI

CRDÉO

MODEL A

Δείκτης	Ετος	Y^0_{t+1}	$\hat{Y}^0_{t+1}$	$\hat{Y}^0_{t+1} \pm A^*$	R
1	1963	52.5	128.6	(76.931,80.268)	0
2	1964	39.6	58.2	(39.442,76.957)	1
3	1965	51.2	32.7	(14.891,50.508)	0
4	1966	47.3	47.4	(32.590,62.209)	1
5	1967	46.9	42.0	(27.410,56.589)	1
6	1968	53.7	53.4	(40.907,65.892)	1
7	1969	59.4	63.7	(50.957,76.442)	1
8	1971	57.9	62.0	(46.381,77.618)	1
9	1972	62.1	53.8	(50.912,73.2287)	1
10	1973	61.7	65.3	(54.545,76.054)	1
11	1974	47.4	59.2	(46.809,71.590)	1
12	1975	65.7	58.6	(49.402,67.797)	1
13	1976	65.7	69.4	(59.264,79.535)	1
14	1977	59.5	50.6	(39.705,61.494)	1
15	1978	68.0	67.4	(58.522,76.277)	1
16	1979	70.3	49.6	(30.979,68.220)	0
17	1980	70.6	70.7	(61.105,80.294)	1

$$\text{ΟΠΟΥ } A^* = t_{1_t-m,1-a/2}\, S_t \sqrt{1+X^0_{t+1}(X_t'X_t)^{-1}X^0_{t+1}}$$

ΠΙΝΑΚΑΣ VII

CRD70

MODEL A

Δείκτης	Ετος	Y^0_{t+1}	$\hat{Y}^0_{t+1}$	$\hat{Y}^0_{t+1} \pm A^*$	R
1	1963	52.6	49.4	(23.092,75.707)	1
2	1964	48.4	54.5	(37.965,71.034)	1
3	1965	60.9	45.6	(32.302,58.897)	0
4	1966	47.2	59.6	(46.502,72.697)	1
5	1967	62.0	54.1	(40.012,68.187)	1
6	1968	52.6	55.9	(42.250,69.549)	1
7	1969	61.7	53.4	(42.688,64.111)	1
8	1971	60.1	58.8	(44.519,73.080)	1
9	1972	65.8	55.9	(45.080,66.719)	1
10	1973	61.5	59.4	(49.233,69.566)	1
11	1974	49.2	51.2	(30.940,71.459)	1
12	1975	63.3	58.1	(48.583,67.616)	1
13	1976	70.3	66.5	(55.597,77.402)	1
14	1977	66.0	62.0	(52.808,71.191)	1
15	1978	66.2	68.5	(59.764,77.235)	1
16	1979	66.0	62.7	(50.949,74.450)	1
17	1980	61.8	66.7	(56.727,76.672)	1

όπου $A^* = t_{1_t-m,1-a/2}\, S_t \sqrt{1+X^0_{t+1}(X_t'X_t)^{-1}X^0_{t+1}}$

ΠΙΝΑΚΑΣ VIII

CRD80

MODEL A

Δείκτης	Ετος	Y^0_{t+1}	$\hat{Y}^0_{t+1}$	$\hat{Y}^0_{t+1} \pm A^*$	P
1	1963	47.6	36.2	(15.886,56.513)	1
2	1964	37.6	44.6	(31.162,58.037)	1
3	1965	52.6	32.5	(20.930,44.069)	0
4	1966	44.6	52.6	(39.904,65.295)	1
5	1967	54.5	51.8	(39.566,64.033)	1
6	1968	48.5	47.3	(35.234,59.365)	1
7	1969	57.9	50.4	(41.060,59.739)	1
8	1971	52.8	58.5	(45.810,71.189)	1
9	1972	60.5	53.5	(43.452,63.547)	1
10	1973	57.3	54.9	(45.585,64.214)	1
11	1974	49.3	53.9	(38.524,65.275)	1
12	1975	42.8	53.3	(44.739,61.860)	0
13	1976	64.7	54.0	(43.189,64.810)	1
14	1977	60.4	52.1	(41.006,63.193)	1
15	1978	62.2	63.0	(53.619,72.380)	1
16	1979	60.4	58.3	(47.314,69.287)	1
17	1980	54.6	60.1	(49.479,70.720)	1

όπου $A^* = t_{t-m,1-a/2}\, S_t \sqrt{1+X^0_{t+1}(X_t X_t)^{-1} X^0_{t+1}}$

ΠΙΝΑΚΑΣ IX

CRD90

MODEL A

Δείκτης	Ετος	Y^0_{t+1}	$\hat{Y}^0_{t+1}$	$\hat{Y}^0_{t+1} \pm A^*$	R
1	1963	46.4	49.4	(29.250,69.549)	1
2	1964	40.1	41.8	(29.167,54.432)	1
3	1965	51.6	33.0	(22.955,43.044)	0
4	1966	50.4	47.6	(36.054,59.145)	1
5	1967	53.6	49.2	(37.343.61.056)	1
6	1968	49.6	54.3	(42.983,65.616)	1
7	1969	57.8	49.5	(41.061,57.938)	1
8	1971	54.5	59.3	(47.764,70.835)	1
9	1972	56.5	53.4	(43.396,63.403)	1
10	1973	50.8	51.4	(41.380,61.419)	1
11	1974	47.7	54.3	(43.773,64.826)	1
12	1975	47.9	48.2	(39.726,56.673)	1
13	1976	63.5	50.4	(42.064,58.735)	0
14	1977	63.7	54.6	(45.180,64.019)	1
15	1978	64.3	61.5	(52.740.70.259)	1
16	1979	63.0	62.4	(53.734,71.065)	1
17	1980	59.6	60.5	(51.149,69.850)	1

όπου $A^* = t_{1_t-m,1-a/2}\, S_t \sqrt{1+X^0_{t+1}(X_t'X_t)^{-1}X^0_{t+1}}$

ΠΙΝΑΚΑΣ Χ

CRD100

MODEL A

Δείκτης	Ετος	Y^0_{t+1}	$\hat{Y}^0_{t+1}$	$\hat{Y}^0_{t+1} \pm A^*$	R
1	1963	54.6	52.0	(27.940,76.059)	1
2	1964	45.2	59.9	(44.052,75.747)	1
3	1965	59.0	44.1	(29.511,58.688)	0
4	1966	49.0	59.1	(46.811,71.388)	1
5	1967	54.0	49.0	(36.464,61.535)	1
6	1968	55.2	53.3	(42.221,64.378)	1
7	1969	62.8	55.0	(44.250,65.749)	1
8	1971	63.4	71.2	(57.148,85.251)	1
9	1972	65.3	62.8	(52.503,73.096)	1
10	1973	64.0	63.2	(53.965,72.434)	1
11	1974	44.6	53.9	(42.381,65.418)	1
12	1975	61.5	46.3	(36.641,55.958)	0
13	1976	69.0	67.5	(56.217,78.782)	1
14	1977	64.0	57.0	(44.172,69.827)	1
15	1978	67.6	74.5	(65.183,83.816)	1
16	1979	70.3	68.9	(58.449,79.350)	1
17	1980	60.3	65.4	(56.236,74.563)	1

όπου $A^* = t_{t-m,1-a/2}\, S_t \sqrt{1+X^0_{t+1}(X_t X_t)^{-1} X^0_{t+1}}$

ΠΙΝΑΚΑΣ Ι

CRD10

MODEL B

Δείκτης	Ετος	Y^{o}_{t+1}	$\hat{Y}^{o}_{t-1}$	$\hat{Y}^{o}_{t+1} \pm A^{*}$	R
1	1963	57.7	54.3	(47.765,60.834)	1
2	1964	48.9	52.9	(46.192,59.637)	1
3	1965	61.2	53.6	(46.859,60.340)	0
4	1966	51.4	57.0	(50.197,63.802)	1
5	1967	57.4	62.5	(55.443,69.556)	1
6	1968	58.5	60.8	(53.629,67.970)	1
7	1969	66.6	62.3	(55.228,69.371)	1
8	1971	71.6	69.2	(62.194,76.205)	1
9	1972	69.6	69.4	(62.529,76.270)	1
10	1973	67.5	70.1	(63.157,77.042)	1
11	1974	45.5	69.6	(62.651,76.548)	0
12	1975	65.2	69.3	(59.121,79.478)	1
13	1976	70.6	71.1	(61.048,81.151)	1
14	1977	66.1	69.6	(59.682,79.517)	1
15	1978	66.9	71.6	(61.929,81.270)	1
16	1979	73.1	76.0	(66.254,85.745)	1
17	1980	49.3	71.2	(61.258,81.141)	0

όπου $A^{*} = t_{1_t - m, 1-a/2}\ S_t \sqrt{1 + X^{o}_{t+1} (X_t X_t)^{-1} X^{o}_{t+1}}$

ΠΙΝΑΚΑΣ ΙΙ

CRD20

MODEL B

Δείκτης	Ετος	Y^0_{t+1}	$\hat{Y}^0_{t+1}$	$\hat{Y}^0_{t+1} \pm A^*$	R
1	1963	52.3	50.8	(44.344,57.255)	1
2	1964	44.6	48.3	(42.129,57.470)	1
3	1965	56.8	48.9	(42.553,55.246)	0
4	1966	46.8	48.3	(41.560,55.039)	1
5	1967	50.9	51.4	(44.421,58.378)	1
6	1968	55.5	53.4	(46.820,59.979)	1
7	1969	59.5	56.1	(49.690,62.509)	1
8	1971	62.6	59.9	(53.607,66.192)	1
9	1972	61.8	59.9	(53.646,66.153)	1
10	1973	62.3	61.4	(54.975,67.824)	1
11	1974	38.9	57.5	(50.919,64.080)	0
12	1975	62.4	60.5	(52.186,68.813)	1
13	1976	67.6	62.1	(54.009,70.190)	1
14	1977	62.6	61.4	(52.772,70.027)	1
15	1978	61.7	63.3	(55.329,71.270)	1
16	1979	68.4	67.0	(59.002,74.997)	1
17	1980	54.0	66.2	(57.627,74.772)	0

όπου $A^* = t_{1_t-m,1-a/2}\, S_t\sqrt{1+X^0_{t+1}(X_t X_t)^{-1}X^0_{t+1}}$

ΠΙΝΑΚΑΣ ΙΙΙ

CRD30

MODEL B

Δείκτης	Ετος	Y^0_{t+1}	$\dot{Y}^0_{t+1}$	$\hat{Y}^0_{t+1} \pm A^*$	R
1	1963	51.2	48.6	(40.423,56.776)	1
2	1964	39.3	45.5	(37.757,53.242)	1
3	1965	52.6	45.4	(37.608,53.191)	1
4	1966	48.3	48.6	(40.704,56.495)	1
5	1967	46.4	48.0	(40.035,55.964)	1
6	1968	54.2	48.0	(40.358,55.641)	1
7	1969	52.8	50.8	(43.035,58.554)	1
8	1971	55.1	54.4	(46.764,62.035)	1
9	1972	59.3	55.6	(48.100,63.099)	1
10	1973	58.1	57.3	(49.760.64.839)	1
11	1974	35.3	51.1	(43.274.58.925)	0
12	1975	52.4	48.8	(39.887,57.712)	1
13	1976	62.8	60.8	(52.086,69.513)	1
14	1977	64.9	58.7	(50.279.67.120)	1
15	1978	60.9	59.1	(50.587,67.612)	1
16	1979	67.5	62.8	(54.389,71.210)	1
17	1980	65.1	64.1	(55.268,72.931)	1

όπου $A^* = t_{1_t - m, 1-a/2}\, S_t \sqrt{1 + X^0_{t+1}(X_t X_t)^{-1} X^0_{t+1}}$

ΠΙΝΑΚΑΣ IV

CRD40

MODEL B

Δείκτης	Ετος	Y^0_{t+1}	$\hat{Y}^0_{t+1}$	$\hat{Y}^0_{t+1} \pm A^*$	E
1	1963	60.1	47.5	(39.775,55.224)	0
2	1964	48.1	50.6	(39.066,57.133)	1
3	1965	64.4	52.5	(43.773,61.226)	0
4	1966	48.9	53.3	(43.293,63.306)	1
5	1967	57.0	56.6	(47.028,66.171)	1
6	1968	56.4	58.1	(48.870,67.329)	1
7	1969	65.8	61.2	(52.122,70.277)	1
8	1971	67.7	67.3	(57.966,76.633)	1
9	1972	69.6	65.2	(56.212,74.187)	1
10	1973	67.5	66.0	(56.378,75.621)	1
11	1974	43.0	64.9	(55.684,74.115)	0
12	1975	69.1	66.0	(55.028,76.971)	1
13	1976	70.8	67.8	(57.054,78.545)	1
14	1977	63.8	69.2	(58.594,79.805)	1
15	1978	73.1	74.4	(63.776,85.023)	1
16	1979	74.7	80.7	(69.425,91.974)	1
17	1980	59.1	73.6	(63.183,84.016)	0

όπου $A^* = t_{1_t-m,1-a/2}\, S_t\sqrt{1+X^0_{t+1}(X_t X_t)^{-1}X^0_{t+1}}$

ΠΙΝΑΚΑΣ V

CRDΞO

MODEL B

Δείκτης	Ετος	Y^0_{t+1}	$\hat{Y}^0_{t+1}$	$\hat{Y}^0_{t+1} \pm A^*$	R
1	1963	56.4	49.4	(40.672,58.127)	1
2	1964	46.2	50.6	(42.094,59.105)	1
3	1965	62.1	49.6	(40.964,58.235)	0
4	1966	50.9	52.7	(42.804,62.595)	1
5	1967	52.0	54.4	(44.704,64.095)	1
6	1968	56.8	52.2	(43.113,61.286)	1
7	1969	67.5	58.8	(49.897,67.702)	1
8	1971	65.9	62.8	(53.590,72.009)	1
9	1972	67.3	60.6	(51.046,69.553)	1
10	1973	68.5	64.8	(55.540,74.059)	1
11	1974	47.5	63.2	(53.555,72.844)	0
12	1975	64.6	60.3	(49.688,70.911)	1
13	1976	72.5	67.9	(57.829,77.970)	1
14	1977	64.1	67.4	(57.444,77.355)	1
15	1978	73.7	70.8	(60.897,80.702)	1
16	1979	73.7	80.7	(69.433,91.966)	1
17	1980	67.7	67.4	(57.028,77.771)	1

όπου $A^* = t_{1_t - m, 1-a/2}\ S_t \sqrt{1 + X^0_{t+1} (X'_t X_t)^{-1} X^0_{t+1}}$

ΠΙΝΑΚΑΣ VI

CRDO

MODEL B

Δείκτης	Ετος	Y^0_{t+1}	$\hat{Y}^0_{t+1}$	$\hat{Y}^0_{t+1} \pm A^*$	R
1	1963	52.5	43.8	(35.173,52.426)	0
2	1964	39.6	45.4	(36.281,54.516)	1
3	1965	51.2	43.2	(33.983,52.416)	1
4	1966	47.3	45.0	(35.503,54.496)	1
5	1967	46.9	46.6	(37.428,55.771)	1
6	1968	53.7	50.0	(41.238,58.761)	1
7	1969	59.4	54.4	(45.462,63.157)	1
8	1971	57.9	55.4	(46.642,64.157)	1
9	1972	62.1	54.5	(45.807,63.192)	1
10	1973	61.7	59.2	(50.417,67.982)	1
11	1974	47.4	55.9	(46.872,64.927)	1
12	1975	65.7	56.0	(46.937,65.062)	1
13	1976	65.7	59.2	(50.445,67.954)	1
14	1977	59.5	59.4	(50.479,68.320)	1
15	1978	68.0	64.4	(55.623,73.176)	1
16	1979	70.3	77.9	(67.433,88.366)	1
17	1980	70.6	67.4	(58.663,76.136)	1

όπου $A^* = t_{1_t-m,\,1-a/2}\, S_t \sqrt{1+X^0_{t+1}(X_t X_t)^{-1} X^0_{t+1}}$

ΠΙΝΑΚΑΣ VII

CRD70

MODEL B

Δείκτης	Ετος	Y^0_{t+1}	$\hat{Y}^0_{t+1}$	$\hat{Y}^0_{t+1} \pm A^*$	R
1	1963	52.6	46.2	(39.187,53.212)	1
2	1964	48.4	47.2	(40.017,54.382)	1
3	1965	60.9	50.4	(43.397,57.402)	0
4	1966	47.2	51.2	(43.300,59.099)	1
5	1967	62.0	60.0	(51.987,68.012)	1
6	1968	52.6	53.9	(46.133,61.666)	1
7	1969	61.7	56.1	(48.560,63.639)	1
8	1971	60.1	64.0	(56.377,71.622)	1
9	1972	65.8	63.0	(55.393,70.606)	1
10	1973	61.5	62.8	(55.148,70.451)	1
11	1974	49.2	65.7	(58.293,73.106)	0
12	1975	63.3	63.6	(54.780,72.819)	1
13	1976	70.3	69.2	(60.427,77.972)	1
14	1977	66.0	65.4	(56.546,74.253)	1
15	1978	68.2	68.2	(59.594,76.805)	1
16	1979	66.0	72.7	(64.316,81.083)	1
17	1980	61.8	64.0	(54.742,73.257)	1

όπου $A^* = t_{1_t - m, 1-\alpha/2}\ S_t \sqrt{1+X^0_{t+1}(X_t X_t)^{-1} X^0_{t+1}}$

ΠΙΝΑΚΑΣ VIII

CRD80

MODEL B

Δείκτης	Ετος	Y^0_{t+1}	$\hat{Y}^0_{t+1}$	$\hat{Y}^0_{t+1} \pm A^*$	R
1	1963	47.6	43.9	(37.278,50.521)	1
2	1964	37.6	42.1	(35.414,48.785)	1
3	1965	52.6	44.6	(37.713.51.486)	0
4	1966	44.6	44.7	(37.312,52.087)	1
5	1967	54.5	52.9	(45.529,60.270)	1
6	1968	48.5	48.1	(41.01,55.188)	1
7	1969	57.9	51.5	(44.590,58.409)	1
8	1971	52.8	58.2	(51.097,65.302)	1
9	1972	60.5	54.7	(47.400,61.999)	1
10	1973	57.3	57.5	(50.016,64.983)	1
11	1974	49.3	56.7	(48.995,64.404)	1
12	1975	42.8	55.3	(47.640,62.959)	1
13	1976	64.7	60.4	(52.099,68.700)	1
14	1977	60.4	56.8	(48.412,65.187)	1
15	1978	62.2	64.8	(56.432,73.167)	1
16	1979	60.4	70.0	(61.028,78.961)	0
17	1980	54.6	56.4	(47.452,65.347)	1

όπου $A^* = t_{1_t - m, 1-a/2}\ S_t \sqrt{1+X^0_{t+1}(X_t' X_t)^{-1} X^0_{t+1}}$

ΠΙΝΑΚΑΣ IX

CRD=0

MODEL B

Δείκτης	Ετος	Y^0_{t+1}	$\hat{Y}^0_{t+1}$	$\hat{Y}^0_{t+1} \pm A^*$	R
1	1963	46.4	45.9	(38.887,52.912)	1
2	1964	40.1	42.1	(35.445,48.754)	1
3	1965	51.6	48.9	(41.560,56.239)	1
4	1966	50.4	44.0	(37.588,50.411)	1
5	1967	53.6	53.4	(46.365,60.434)	1
6	1968	49.6	48.0	(41.865,54.134)	1
7	1969	57.6	49.2	(42.582,55.817)	0
8	1971	54.5	58.4	(51.444,65.355)	1
9	1972	56.5	54.9	(47.929,61.870)	1
10	1973	50.8	56.3	(49.225,63.374)	1
11	1974	47.7	56.6	(46.608,63.591)	1
12	1975	47.9	53.2	(45.626,60.773)	1
13	1976	63.5	58.8	(51.371,66.228)	1
14	1977	63.7	55.4	(47.684,63.115)	0
15	1978	64.3	61.5	(53.761,69.238)	1
16	1979	63.0	64.4	(56.770,72.029)	1
17	1980	59.6	59.0	(50.665,67.334)	1

όπου $A^* = t_{1_t - m, 1-a/2} \; S_t \sqrt{1 + X^0_{t+1} (X_t' X_t)^{-1} X^0_{t+1}}$

ΠΙΝΑΚΑΣ Χ

CRD100

MODEL B

Δείκτης	Ετος	Y^0_{t+1}	$\hat{Y}^0_{t+1}$	$\hat{Y}^0_{t+1} \pm A^*$	R
1	1963	54.6	48.5	(42.525,54.474)	0
2	1964	45.2	49.1	(42.997,55.202)	1
3	1965	59.0	49.8	(43.504,56.095)	0
4	1966	49.0	51.6	(44.751,58.448)	1
5	1967	54.0	56.6	(46.692,63.507)	1
6	1968	55.2	54.8	(47.909,61.690)	1
7	1969	62.8	57.1	(50.388,63.811)	1
8	1971	63.4	63.0	(56.280,69.719)	1
9	1972	65.3	62.6	(56.042,69.157)	1
10	1973	64.0	64.4	(57.675,71.124)	1
11	1974	44.6	65.0	(58.601,71.398)	0
12	1975	61.5	62.2	(53.009,71.390)	1
13	1976	69.0	66.9	(57.995,75.844)	1
14	1977	64.0	63.2	(54.256,72.143)	1
15	1978	67.8	67.7	(59.030,76.369)	1
16	1979	70.3	70.5	(61.975,79.024)	1
17	1980	60.3	66.6	(57.468,75.731)	1

όπου $A^* = t_{1_t-m,1-a/2}\, S_t\sqrt{1+X^0_{t+1}(X_t X_t)^{-1}X^0_{t+1}}$

ΠΑΡΑΡΤΗΜΑ V

ΑΠΟΤΕΛΕΣΜΑΤΑ ΜΕΘΟΔΟΥ ΣΥΓΚΡΙΤΙΚΗΣ ΑΞΙΟΛΟΓΗΣΗΣ ΤΗΣ ΠΡΟΒΛΕΠΤΙΚΗΣ ΙΚΑΝΟΤΗΤ

ΔΥΟ ΓΡΑΜΜΙΚΩΝ ΜΟΝΤΕΛΩΝ ΒΑΣΙΣΜΕΝΗ ΣΤΟΝ ΕΛΕΓΧΟ ΤΟΥ ΣΥΝΤΕΛΕΣΤΗ ΣΥΣΧΕΤΙΣΗΣ

Y_t^0 η παρατηρούμενη τιμή του μοντέλου την χρονική στιγμή t

$\hat{Y}_t^0(A)$ η προβλεπόμενη τιμή του μοντέλου Α την χρονική στιγμή t

$\hat{Y}_t^0(B)$ η προβλεπόμενη τιμή του μοντέλου Β την χρονική στιγμή t

CRDIO-CRD100

ΠΙΝΑΚΑΣ Ι

CRD10

Δείκτης	Ετος	Y_t^o	$\hat{Y}_t^o(A)$	$\hat{Y}_t^o(B)$
1	1963	57.7	32.7	54.3
2	1964	48.9	62.9	52.9
3	1965	61.2	49.1	53.6
4	1966	51.4	59.2	57.0
5	1967	57.4	44.2	62.5
6	1968	58.5	58.9	60.8
7	1969	66.6	56.2	62.3
8	1971	71.6	47.0	69.2
9	1972	69.6	65.2	69.4
10	1973	67.5	66.3	70.1
11	1974	45.5	64.5	69.6
12	1975	65.2	59.4	69.3
13	1976	70.6	49.3	71.1
14	1977	66.1	47.9	71.1
15	1978	66.9	67.0	71.6
16	1979	73.1	70.0	76.0
17	1980	49.3	63.2	71.2

$R = -0.1303$

$Z_R = -0.3408$

ΠΙΝΑΚΑΣ ΙΙ

CRD20

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0(A)$	$\hat{Y}_t^0(B)$
1	1963	52.3	48.2	50.8
2	1964	44.6	56.6	48.3
3	1965	56.8	42.6	48.9
4	1966	46.8	49.3	48.3
5	1967	50.9	42.2	51.4
6	1968	55.5	50.5	53.4
7	1969	59.5	60.9	56.1
8	1971	62.8	56.3	59.9
9	1972	61.8	60.6	59.9
10	1973	62.3	58.9	61.4
11	1974	38.9	59.0	57.5
12	1975	62.4	64.1	60.5
13	1976	67.6	57.8	62.1
14	1977	62.6	63.9	61.4
15	1978	61.7	65.4	63.3
16	1979	68.4	67.3	67.0
17	1980	54.0	63.7	66.2

$R = 0.2137$

$Z_R = 0.5467$

ΠΙΝΑΚΑΣ III

CRD30

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0(A)$	$\hat{Y}_t^0(B)$
1	1963	51.2	42.8	48.6
2	1964	39.3	54.8	45.5
3	1965	52.6	32.8	45.4
4	1966	48.3	51.0	48.6
5	1967	46.4	44.7	48.0
6	1968	54.2	46.9	48.0
7	1969	52.8	54.4	50.8
8	1971	55.1	57.7	54.4
9	1972	59.3	57.1	55.6
10	1973	56.1	59.9	57.3
11	1974	35.3	49.9	51.1
12	1975	52.4	50.7	48.8
13	1976	62.8	58.1	60.8
14	1977	64.9	60.7	58.7
15	1978	60.9	60.3	59.1
16	1979	67.5	66.4	62.8
17	1980	65.1	65.1	64.1

$R = 0.5388$

$Z_R = 0.6895$

ΠΙΝΑΚΑΣ IV

CRD40

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0(A)$	$\hat{Y}_t^0(B)$
1	1963	60.1	72.5	47.5
2	1964	48.1	62.2	50.6
3	1965	64.4	46.4	52.5
4	1966	48.9	53.5	53.3
5	1967	57.0	52.9	56.6
6	1968	56.4	49.6	58.1
7	1969	65.8	55.4	61.2
8	1971	67.7	74.3	67.3
9	1972	69.6	64.8	65.2
10	1973	67.5	60.8	66.0
11	1974	43.0	50.1	64.9
12	1975	69.1	53.5	66.0
13	1976	70.8	74.7	67.8
14	1977	63.8	43.5	69.2
15	1978	73.1	79.2	74.4
16	1979	74.7	69.9	80.7
17	1980	59.1	71.6	73.6

$R = -0.1401$

$Z_R = -0.2097$

ΠΙΝΑΚΑΣ V

CRDΞO

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0(A)$	$\hat{Y}_t^0(B)$
1	1963	56.4	68.3	49.4
2	1964	46.2	60.3	50.6
3	1965	62.1	36.7	49.6
4	1966	50.9	54.3	52.7
5	1967	52.0	50.1	54.4
6	1968	56.8	55.3	52.2
7	1969	67.5	55.4	58.8
8	1971	65.9	75.9	62.8
9	1972	67.3	62.0	60.6
10	1973	68.5	72.7	64.8
11	1974	47.5	64.7	63.2
12	1975	64.6	65.4	60.3
13	1976	72.5	75.3	67.9
14	1977	64.1	53.3	67.4
15	1978	73.7	64.0	70.8
16	1979	73.7	74.4	80.7
17	1980	67.7	66.4	67.4

$R = 0.6979$

$Z_R = 1.0716$

ΠΙΝΑΚΑΣ VI

CRDCO

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0(A)$	$\hat{Y}_t^0(B)$
1	1963	52.5	128.6	43.8
2	1964	39.6	58.2	45.4
3	1965	51.2	32.7	43.2
4	1966	47.3	47.4	45.0
5	1967	46.9	42.0	46.6
6	1968	53.7	53.4	50.0
7	1969	59.4	63.7	54.4
8	1971	57.9	62.0	55.4
9	1972	62.1	53.8	54.5
10	1973	61.7	65.3	59.2
11	1974	47.4	59.2	55.9
12	1975	65.7	58.6	56.0
13	1976	65.7	69.4	59.2
14	1977	59.5	50.6	59.4
15	1978	66.0	67.4	64.4
16	1979	70.3	49.6	77.9
17	1980	70.6	70.7	67.4

$R = 0.9630$

$Z_R = 0.9523$

ΠΙΝΑΚΑΣ VII

CRD70

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0(A)$	$\hat{Y}_t^0(B)$
1	1963	52.6	49.4	46.2
2	1964	48.4	54.5	47.2
3	1965	60.9	45.6	50.4
4	1966	47.2	59.6	51.2
5	1967	62.0	54.1	60.0
6	1968	52.6	55.9	53.9
7	1969	61.7	53.4	56.1
8	1971	60.1	58.8	64.0
9	1972	65.8	55.9	63.0
10	1973	61.5	59.4	62.8
11	1974	49.2	51.2	65.7
12	1975	63.3	58.1	63.8
13	1976	70.3	66.5	69.2
14	1977	66.0	62.0	65.4
15	1978	66.2	65.5	68.2
16	1979	66.0	62.7	72.7
17	1980	61.8	66.7	64.0

$R = -0.0519$

$Z_R = -0.0841$

ΠΙΝΑΚΑΣ VIII

CRDE0

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0(A)$	$\hat{Y}_t^0(B)$
1	1963	47.6	36.2	43.9
2	1964	37.6	44.6	42.1
3	1965	52.6	32.5	44.6
4	1966	44.6	52.6	44.7
5	1967	54.5	51.8	52.9
6	1968	48.5	47.3	48.1
7	1969	57.9	50.4	51.5
8	1971	52.8	58.5	58.2
9	1972	60.5	53.5	54.7
10	1973	57.3	54.9	57.5
11	1974	49.3	53.9	56.7
12	1975	42.6	53.3	55.3
13	1976	64.7	54.0	60.4
14	1977	60.4	52.1	56.8
15	1978	62.2	63.0	64.8
16	1979	60.4	58.3	70.0
17	1980	54.6	60.1	56.4

$R = 0.3288$

$Z_R = 0.4789$

ΠΙΝΑΚΑΣ IX

CRD90

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0(A)$	$\hat{Y}_t^0(B)$
1	1963	46.4	49.4	45.9
2	1964	40.1	41.8	42.1
3	1965	51.6	33.0	48.9
4	1966	50.4	47.6	44.0
5	1967	53.6	49.2	53.4
6	1968	49.6	54.3	48.0
7	1969	57.6	49.5	49.2
8	1971	54.5	59.3	58.4
9	1972	56.5	53.4	54.9
10	1973	50.6	51.4	56.3
11	1974	47.7	54.3	56.6
12	1975	47.9	46.2	53.2
13	1976	63.5	50.4	56.8
14	1977	63.7	54.6	55.4
15	1978	64.3	61.5	61.5
16	1979	63.0	62.4	64.4
17	1980	59.6	60.5	59.0

$R = 0.4541$

$Z_R = 0.3484$

ΠΙΝΑΚΑΣ Χ

CRD100

Δείκτης	Ετος	Y_t^0	$\hat{Y}_t^0(A)$	$\hat{Y}_t^0(B)$
1	1963	54.6	52.0	48.5
2	1964	45.2	59.9	49.1
3	1965	59.0	44.1	49.8
4	1966	49.0	59.1	51.6
5	1967	54.0	49.0	56.6
6	1968	55.2	53.3	54.8
7	1969	62.8	55.0	57.1
8	1971	63.4	71.2	63.0
9	1972	65.3	62.8	62.6
10	1973	64.0	63.2	64.4
11	1974	44.6	53.9	65.0
12	1975	61.5	46.3	62.2
13	1976	69.0	67.5	66.9
14	1977	64.0	57.0	63.2
15	1978	67.8	74.5	67.7
16	1979	70.3	68.9	70.5
17	1980	60.3	65.4	66.6

$R = 0.0394$

$Z_R = 0.0641$