



ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ

ΜΑΘΗΤΕΣ Α' ΒΑΘΜΙΑΣ ΚΑΙ Β' ΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ. ΑΝΑΛΥΣΗ ΧΡΟΝΟΛΟΓΙΚΩΝ ΣΕΙΡΩΝ ΚΑΙ ΣΥΓΚΡΙΣΗ ΜΕΘΟΔΩΝ ΠΡΟΒΛΕΨΗΣ

Αθανάσιος Πλούτ. Ντάρδας

ΕΡΓΑΣΙΑ

Που υποβλήθηκε στο Τμήμα Στατιστικής
του Οικονομικού Πανεπιστημίου Αθηνών
ως μέρος των απαιτήσεων για την απόκτηση

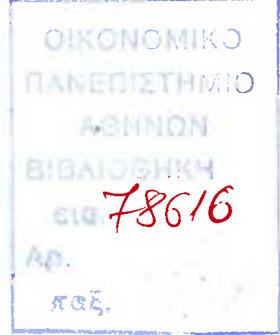
Μεταπτυχιακού Διπλώματος

Συμπληρωματικής Ειδίκευσης στη Στατιστική
Μερικής Παρακολούθησης (Part-time)

Αθήνα
Νοέμβριος 2005







ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

ΜΑΘΗΤΕΣ Α' ΒΘΜΙΑΣ ΚΑΙ Β' ΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ. ΑΝΑΛΥΣΗ ΧΡΟΝΟΛΟΓΙΚΩΝ ΣΕΙΡΩΝ ΚΑΙ ΣΥΓΚΡΙΣΗ ΜΕΘΟΔΩΝ ΠΡΟΒΛΕΨΗΣ

Αθανάσιος Πλουτ. Ντάρδας

ΕΡΓΑΣΙΑ

Που υποβλήθηκε στο Τμήμα Στατιστικής
του Οικονομικού Πανεπιστημίου Αθηνών
ως μέρος των απαιτήσεων για την απόκτηση

Μεταπτυχιακού Διπλώματος

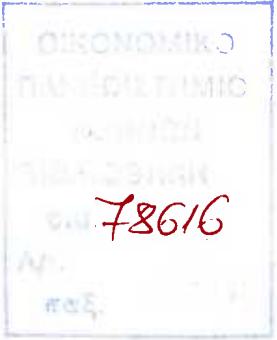
Συμπληρωματικής Ειδίκευσης στη Στατιστική

Μερικής Παρακολούθησης (Part-time)



Αθήνα
Απρίλιος 2005





ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

Εργασία που υποβλήθηκε ως μέρος των απαιτήσεων για την απόκτηση
Μεταπτυχιακού Διπλώματος Συμπληρωματικής Ειδίκευσης στη Στατιστική
Μερικής Παρακολούθησης (Part-time)

ΜΑΘΗΤΕΣ Α' ΒΑΘΜΙΑΣ ΚΑΙ Β' ΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ. ΑΝΑΛΥΣΗ ΧΡΟΝΟΛΟΓΙΚΩΝ ΣΕΙΡΩΝ ΚΑΙ ΣΥΓΚΡΙΣΗ ΜΕΘΟΔΩΝ ΠΡΟΒΛΕΨΗΣ

Αθανάσιος Πλουτ. Ντάρδας

Υπεύθυνο μέλος ΔΕΠ:

Αλ. Λειβαδά

Επίκουρη Καθηγήτρια

Ο Διευθυντής Μεταπτυχιακών Σπουδών

Μιχαήλ Λαζάνης

Καθηγητής



ΑΦΙΕΡΩΣΗ

Στον Χρήστο και τον Πλούταρχο



ΕΥΧΑΡΙΣΤΙΕΣ

Ευχαριστώ πολύ την κ. Λειβαδά που επέβλεψε αυτή την εργασία. Ευχαριστώ τους γονείς μου για την συμπαράσταση τους και την αρραβωνιαστικιά μου Μυρτώ για την υπομονή και παρότρυνση της να τελειώσω αυτή την εργασία .

Τέλος ευχαριστώ τον φίλο μου Γιάννη που ξόδεψε αρκετό χρόνο για να διορθώσει τον σχεδόν κατεστραμμένο υπολογιστή μου



ΒΙΟΓΡΑΦΙΚΟ ΣΗΜΕΙΩΜΑ

Γεννήθηκα και μεγάλωσα στην Ελευσίνα, όπου και τελείωσα την βασική μου εκπαίδευση το 1990.

Σπούδασα Μαθηματικά στο Πανεπιστήμιο Πατρών από το 1990 μέχρι το 1995.

Το 1996-97 έκανα μεταπτυχιακές σπουδές στο πανεπιστήμιο του Κέντ στο Κάντερμπουρυ (University of Kent at Canterbury) αποκτώντας μεταπτυχιακό δίπλωμα Σπουδών Επιχειρήσεων (Master of Business Studies).

Μετά την ολοκλήρωση της στρατιωτικής μου θητείας το 1999 εργάζομαι σαν καθηγητής μαθηματικών παραδίδοντας μαθήματα .





ABSTRACT

Athanasiros Ntardas

STUDENTS OF PRIMARY AND SECONDARY EDUCATION. TIME SERIES ANALYSIS AND COMPARISON OF FORECASTING METHODS

April 2005

In this dissertation we are going to analyse time series techniques to study the evolution of the number of students of primary and secondary education in Greece from 1970 to 1996 and compare the results of alternative forecasting methods for this occasion.

We examine the trend of the above time series and trace the reason for their behaviour along with other interesting elements. Then we apply three alternative methods of forecasting, Decomposition, Exponential smoothing and ARIMA methods and we compare their results.





ΠΕΡΙΛΗΨΗ

Αθανάσιος Ντάρδας

ΜΑΘΗΤΕΣ Α' ΒΑΘΜΙΑΣ ΚΑΙ Β' ΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ.

ΑΝΑΛΥΣΗ ΧΡΟΝΟΛΟΓΙΚΩΝ ΣΕΙΡΩΝ ΚΑΙ ΣΥΓΚΡΙΣΗ ΜΕΘΟΔΩΝ
ΠΡΟΒΛΕΨΗΣ

Απρίλιος 2005

Η εργασία αυτή θα ασχοληθεί με την ανάλυση των χρονολογικών σειρών του αριθμού των μαθητών και μαθητριών της πρωτοβάθμιας και δευτεροβάθμιας εκπαίδευσης της Ελλάδος από το 1970 μέχρι το 1996.

Αρχικά θα προσπαθήσουμε να προσδιορίσουμε την τάση της κάθε χρονολογικής σειράς και να εντοπίσουμε τους λόγους της συμπεριφοράς της καθώς και μερικά άλλα ενδιαφέροντα στοιχεία.

Στη συνέχεια θα εφαρμόσουμε και θα συγκρίνουμε τρεις διαφορετικές μεθόδους πρόβλεψης: μέθοδο Αποσύνθεσης συνιστωσών χρονολογικών σειρών, την Εκθετική Εξομάλυνση και την Μέθοδο ARIMA





ΚΑΤΑΛΟΓΟΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΚΕΦΑΛΑΙΟ 1 :ΧΡΟΝΟΛΟΓΙΚΕΣ ΣΕΙΡΕΣ ΚΑΙ ΜΕΘΟΔΟΙ ΑΝΑΛΥΣΗΣ.....	1
1.1 ΧΡΟΝΟΛΟΓΙΚΕΣ ΣΕΙΡΕΣ ΕΞΕΛΙΞΗΣ ΜΑΘΗΤΩΝ.....	1
1.2 ΠΑΛΙΝΔΡΟΜΗΣΗ	5
1.2.1 <i>Πολλαπλή παλινδρόμηση</i>	7
1.2.2 <i>Μοντέλα που μετατρέπονται σε γραμμικά</i>	10
1.3 ΕΚΘΕΤΙΚΗ ΕΞΟΜΑΛΥΝΣΗ	10
1.3.1 <i>Εκθετική εξομάλυνση των Holt</i>	11
1.4 ΜΕΘΟΔΟΙ ARIMA	12
1.5 ΜΕΤΡΑ ΑΚΡΙΒΕΙΑΣ	14
1.5.1 Μέσο Τετραγωνικό Σφάλμα (MSE ή MSD)	14
1.5.2 Μέσο Απόλυτο Σφάλμα (MAD)	15
1.5.2 Μέσο απόλυτο ποσοστιαίο σφάλμα (MAPE)	15
ΚΕΦΑΛΑΙΟ 2 :ΠΡΩΤΟΒΑΘΜΙΑ ΚΑΙ ΔΕΥΤΕΡΟΒΑΘΜΙΑ ΕΚΠΑΙΔΕΥΣΗ 1970/71-1996/97	17
2.1 ΠΡΩΤΟΒΑΘΜΙΑ ΕΚΠΑΙΔΕΥΣΗ	17
2.2 ΔΕΥΤΕΡΟΒΑΘΜΙΑ ΕΚΠΑΙΔΕΥΣΗ	17
ΚΕΦΑΛΑΙΟ 3 :Η ΠΡΩΤΟΒΑΘΜΙΑ ΕΚΠΑΙΔΕΥΣΗ	21
3.1 ΓΕΝΙΚΑ ΓΙΑ ΤΟ ΔΗΜΟΤΙΚΟ	21
3.2 ΤΑΣΗ ΔΙΑΦΟΡΑΣ ΜΑΘΗΤΩΝ-ΜΑΘΗΤΡΙΩΝ	23
3.3 ΕΚΤΙΜΗΣΗ ΤΑΣΗΣ ΜΑΘΗΤΩΝ ΔΗΜΟΤΙΚΟΥ	25
3.4 ΕΚΘΕΤΙΚΗ ΕΞΟΜΑΛΥΝΣΗ	28
3.5 ΥΠΟΔΕΙΓΜΑΤΑ ARIMA	30
3.6 ΣΥΓΚΡΙΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ	43
ΚΕΦΑΛΑΙΟ 4 :ΓΕΝΙΚΗ ΕΙΚΟΝΑ ΤΗΣ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ	45
4.1 ΣΥΝΟΛΟ ΜΑΘΗΤΩΝ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ	45
4.2 ΑΡΡΕΝΕΣ ΜΑΘΗΤΕΣ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ	47
4.3 ΜΑΘΗΤΡΙΕΣ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ	49
ΚΕΦΑΛΑΙΟ 5 :ΚΑΤΩΤΕΡΗ ΔΕΥΤΕΡΟΒΑΘΜΙΑ ΕΚΠΑΙΔΕΥΣΗ.....	55
5.1 ΓΕΝΙΚΗ ΕΙΚΟΝΑ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ,	55
5.2 ΕΚΤΙΜΗΣΗ ΤΑΣΗΣ ΜΑΘΗΤΡΙΩΝ	58
5.3 ΔΙΠΛΗ ΕΚΘΕΤΙΚΗ ΕΞΟΜΑΛΥΝΣΗ ΓΙΑ ΤΙΣ ΜΑΘΗΤΡΙΕΣ	61
5.4 ΥΠΟΔΕΙΓΜΑΤΑ ARIMA ΓΙΑ ΤΙΣ ΜΑΘΗΤΡΙΕΣ	63
5.5 ΣΥΜΠΕΡΑΣΜΑΤΑ ΓΙΑ ΤΙΣ ΜΑΘΗΤΡΙΕΣ	70
5.6 ΑΝΑΛΥΣΗ ΑΡΡΕΝΩΝ ΜΑΘΗΤΩΝ ΚΑΤΩΤΕΡΗΣ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ-ΤΑΣΗ.....	71
5.7 ΕΚΘΕΤΙΚΗ ΕΞΟΜΑΛΥΝΣΗ ΜΑΘΗΤΩΝ	74
5.8 ΥΠΟΔΕΙΓΜΑΤΑ ARIMA ΓΙΑ ΤΟΥΣ ΑΡΡΕΝΕΣ ΜΑΘΗΤΕΣ	77
5.9 ΣΥΜΠΕΡΑΣΜΑΤΑ ΓΙΑ ΤΟΥΣ ΑΡΡΕΝΕΣ ΜΑΘΗΤΕΣ	84
ΚΕΦΑΛΑΙΟ 6 :ΑΝΩΤΕΡΗ ΔΕΥΤΕΡΟΒΑΘΜΙΑ ΕΚΠΑΙΔΕΥΣΗ	85
6.1 ΓΕΝΙΚΗ ΕΙΚΟΝΑ ΑΝΩΤΕΡΗΣ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗ	85
6.2 ΜΑΘΗΤΡΙΕΣ ΑΝΩΤΕΡΗΣ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ-ΤΑΣΗ.....	89
6.3 ΔΙΠΛΗ ΕΚΘΕΤΙΚΗ ΕΞΟΜΑΛΥΝΣΗ ΜΑΘΗΤΡΙΩΝ	91
6.4 ΥΠΟΔΕΙΓΜΑΤΑ ARIMA ΜΑΘΗΤΡΙΩΝ	93
6.5 ΣΥΜΠΕΡΑΣΜΑΤΑ ΓΙΑ ΤΙΣ ΜΑΘΗΤΡΙΕΣ	99
6.6 ΤΑΣΗ ΑΡΡΕΝΩΝ ΜΑΘΗΤΩΝ	100
6.7 ΔΙΠΛΗ ΕΚΘΕΤΙΚΗ ΕΞΟΜΑΛΥΝΣΗ ΑΡΡΕΝΩΝ ΜΑΘΗΤΩΝ	103
6.8 ΥΠΟΔΕΙΓΜΑΤΑ ARIMA ΑΡΡΕΝΩΝ ΜΑΘΗΤΩΝ	104
6.8 ΣΥΜΠΕΡΑΣΜΑΤΑ ΓΙΑ ΤΟΥΣ ΜΑΘΗΤΕΣ	111
ΚΕΦΑΛΑΙΟ 7 :ΣΥΜΠΕΡΑΣΜΑΤΑ.....	113
7.1 ΔΗΜΟΤΙΚΟ	113
7.2 ΔΕΥΤΕΡΟΒΑΘΜΙΑ ΕΚΠΑΙΔΕΥΣΗ	113
7.2.1 <i>Κατώτερη δευτεροβάθμια εκπαίδευση.</i>	114
7.2.2 <i>Ανώτερη δευτεροβάθμια εκπαίδευση.</i>	115
7.3 ΣΤΑΤΙΣΤΙΚΑ ΣΥΜΠΕΡΑΣΜΑΤΑ	116

ΒΙΒΛΙΟΓΡΑΦΙΑ	121
ΠΑΡΑΡΤΗΜΑ	123
ΠΙΝΑΚΑΣ ΔΕΔΟΜΕΝΩΝ	123
MCRO ΥΠΟΛΟΓΙΣΜΟΥ ΜΕΤΡΩΝ ΑΚΡΙΒΕΙΑΣ.....	123
MACRO ΥΠΟΛΟΓΙΣΜΟΥ R-SQ KAI R-SQ(ADJ).....	124
MACRO ΥΠΟΛΟΓΙΜΟΥ ΣΥΝΤΕΛΕΣΤΗ DURBIN-WATSON	124

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

<u>Πίνακας 1.1:</u> Υποδείγματα που μετασχηματίζονται σε γραμμικά	10
<u>Πίνακας 3.1:</u> Πραγματικές τιμές και προβλέψεις του μοντέλου της παλινδρόμησης.....	27
<u>Πίνακας 3.2:</u> Μέτρα ακριβείας του υποδείγματος της παλινδρόμησης	27
<u>Πίνακας 3.3:</u> Πραγματικές τιμές και προβλέψεις Διπλής Εκθετικής Εξομάλυνσης.....	29
<u>Πίνακας 3.4 :</u> Μέτρα ακριβείας Διπλής Εκθετικής Εξομάλυνσης.....	30
<u>Πίνακας 3.5:</u> Πραγματικές τιμές και προβλέψεις υποδείγματος ARIMA(1,1,0).....	35
<u>Πίνακας 3.6:</u> Μέτρα ακριβείας ARIMA (1,1,0)	35
<u>Πίνακας 3.7:</u> Πραγματικές τιμές και προβλέψεις υποδείγματος ARIMA(1,1,0).....	38
<u>Πίνακας 3.8:</u> Μέτρα ακριβείας ARIMA(1,1,0)	38
<u>Πίνακας 3.9:</u> Πραγματικές τιμές και προβλέψεις υποδείγματος ARIMA(0,2,1).....	42
<u>Πίνακας 3.10:</u> Μέτρα ακριβείας ARIMA (0,2,1)	42
<u>Πίνακας 3.11:</u> Συγκριτικά αποτελέσματα μέτρων ακριβείας των μοντέλων που εξετάστηκαν.....	43
<u>Πίνακας 4.1:</u> Ποσοστά μαθητών και μαθητριών που συνεχίζουν στη δευτεροβάθμια εκπαίδευση.	53
<u>Πίνακας 5.1:</u> Πραγματικές τιμές και προβλέψεις των υποδειγμάτων της τάσης	61
<u>Πίνακας 5.2</u> Μέτρα ακριβείας των υποδειγμάτων της τάσης	61
<u>Πίνακας 5.3:</u> Πραγματικές τιμές και προβλέψεις της Διπλής Εκθετικής Εξομάλυνσης.....	62
<u>Πίνακας 5.4:</u> Μέτρα ακριβείας του υποδείγματος της Διπλής Εκθετικής Εξομάλυνσης	62
<u>Πίνακας 5.5:</u> Πραγματικές τιμές και προβλέψεις των υποδειγμάτων ARIMA.....	69
<u>Πίνακας 5.6:</u> Μέτρα ακριβείας των υποδειγμάτων ARIMA	69
<u>Πίνακας 5.7:</u> Συγκριτικά αποτελέσματα μέτρων ακριβείας των υποδειγμάτων για τις μαθήτριες.....	70
<u>Πίνακας 5.8:</u> Πραγματικές τιμές και προβλέψεις των υποδειγμάτων της τάσης	73
<u>Πίνακας 5.9:</u> Μέτρα ακριβείας των υποδειγμάτων της τάσης	73
<u>Πίνακας 5.10:</u> Πραγματικές τιμές και προβλέψεις των υποδειγμάτων της εκθετικής εξομάλυνσης	76
<u>Πίνακας 5.11:</u> Μέτρα ακριβείας των υποδειγμάτων της εκθετικής εξομάλυνσης	76
<u>Πίνακας 5.12:</u> Πραγματικές τιμές και προβλέψεις των υποδειγμάτων ARIMA	83
<u>Πίνακας 5.13:</u> Μέτρα ακριβείας των υποδειγμάτων ARIMA.....	83
<u>Πίνακας 5.14:</u> Σύγκριση μέτρων ακριβείας των υποδειγμάτων για τους μαθητές.....	84
<u>Πίνακας 6.1:</u> Πραγματικές τιμές και προβλέψεις των υποδειγμάτων της τάσης μαθητριών	91
<u>Πίνακας 6.2:</u> Μέτρα ακριβείας των υποδειγμάτων της τάσης	91
<u>Πίνακας 6.3:</u> Πραγματικές τιμές και προβλέψεις των υποδειγμάτων Διπλής Εκθετικής Εξομάλυνσης	92
<u>Πίνακας 6.4:</u> Μέτρα ακριβείας των υποδειγμάτων της Δ.Ε.Ε	92
<u>Πίνακας 6.5:</u> Πραγματικές τιμές και προβλέψεις των υποδειγμάτων ARIMA μαθητριών	98
<u>Πίνακας 6.6:</u> Μέτρα ακριβείας των υποδειγμάτων ARIMA	98
<u>Πίνακας 6.7:</u> Συγκριτικά αποτελέσματα υποδειγμάτων για τις μαθήτριες	99
<u>Πίνακας 6.8:</u> Συντελεστές προσδιορισμού των υποδειγμάτων τάσης	101
<u>Πίνακας 6.9:</u> Πραγματικές τιμές και προβλέψεις των υποδειγμάτων της τάσης μαθητών	101
<u>Πίνακας 6.10:</u> Μέτρα ακριβείας των υποδειγμάτων της τάσης	102
<u>Πίνακας 6.11:</u> Πραγματικές τιμές και προβλέψεις της ΔΕΕ.....	104
<u>Πίνακας 6.12:</u> Μέτρα ακριβείας της Δ.Ε.Ε	104
<u>Πίνακας 6.13:</u> Πραγματικές τιμές και προβλέψεις των υποδειγμάτων ARIMA μαθητών	110
<u>Πίνακας 6.14:</u> Μέτρα ακριβείας των υποδειγμάτων ARIMA	110
<u>Πίνακας 6.15:</u> Σύγκριση μέτρων ακριβείας για τους μαθητές ανώτερης δευτεροβάθμιας εκπαίδευσης	111
<u>Πίνακας 7.1:</u> Σύγκριση των μεθόδων με βάση τα μέτρα ακριβείας.....	118
<u>Πίνακας 7.2:</u> Σύγκριση μεθόδων πρόβλεψης με βάση τα μέτρα ακριβείας.....	119



ΚΑΤΑΛΟΓΟΣ ΓΡΑΦΗΜΑΤΩΝ

ΓΡΑΦΗΜΑ 1.1: Εγγεγραμμένοι μαθητές Πρωτοβάθμιας και Δευτεροβάθμιας εκπαίδευσης	2
ΓΡΑΦΗΜΑ 1.2: Δευτεροβάθμιας τάση του συνόλου των μαθητών	2
ΓΡΑΦΗΜΑ 1.3: Αριθμός Μαθητών Δημοτικού από το 1970/71 – 1996/97	4
ΓΡΑΦΗΜΑ 1.4: Αριθμός μαθητών Κατώτερης Δευτεροβάθμιας Εκπαίδευσης (ISCED 3)	4
ΓΡΑΦΗΜΑ 1.5: Αριθμός μαθητών Ανώτερης Δευτεροβάθμιας Εκπαίδευσης (ISCED 4)	5
ΓΡΑΦΗΜΑ 3.1: Αριθμός μαθητών Δημοτικού από το 1970/71 – 1996/97	21
ΓΡΑΦΗΜΑ 3.2: Διαφορά αριθμού αγοριών-κοριτσιών του Δημοτικού σχολείου	22
ΓΡΑΦΗΜΑ 3.3: Σύγκριση αριθμού αγοριών – κοριτσιών στο δημοτικό	22
ΓΡΑΦΗΜΑ 3.4: Τάση της διαφοράς αγοριών- κοριτσιών για τα έτη 1970/71 – 1996/97	24
ΓΡΑΦΗΜΑ 3.5: Τάση του αριθμού των μαθητών δημοτικού 1970/71-1993/94	27
ΓΡΑΦΗΜΑ 3.6: Προβλέψεις και πραγματικές τιμές για τα έτη 93/94 – 96/97	28
ΓΡΑΦΗΜΑ 3.7: Εκθετική Εξομάλυνση του Holt. Προσαρμογή και Πρόβλεψη	29
ΓΡΑΦΗΜΑ 3.8: Σύγκριση Πραγματικών Τιμών με τις Προβλέψεις της Εξομάλυνσης Holt	30
ΓΡΑΦΗΜΑ 3.9: Συντελεστές Αυτοσυσχέτισης για μαθητές Δημοτικού	31
ΓΡΑΦΗΜΑ 3.10: Μερικές αυτοσυσχετίσεις για τους Μαθητές Δημοτικού	32
ΓΡΑΦΗΜΑ 3.11: Πρώτες Διαφορές της χρονολογικής σειράς Μαθητών Δημοτικού	32
ΓΡΑΦΗΜΑ 3.12: Αυτοσυσχέτισεις των πρώτων διαφορών για τους μαθητές Δημοτικού	33
ΓΡΑΦΗΜΑ 3.13: Μερικές αυτοσυσχετίσεις των πρώτων διαφορών των μαθητών Δημοτικού	33
ΓΡΑΦΗΜΑ 3.14: Αυτοσυσχέτιση των καταλοίπων υποδείγματος ARIMA (1,1,0)	34
ΓΡΑΦΗΜΑ 3.15: Μερική Αυτοσυσχέτιση καταλοίπων μοντέλου ARIMA (1,1,0)	35
ΓΡΑΦΗΜΑ 3.16: Πραγματικές τιμές και πρόβλεψη υποδείγματος ARIMA(1,1,0)	36
ΓΡΑΦΗΜΑ 3.17: Αυτοσυσχέτιση καταλοίπων υποδείγματος ARIMA (1,1,1)	37
ΓΡΑΦΗΜΑ 3.18: Μερική αυτοσυσχέτιση καταλοίπων APIMA (1,1,1)	38
ΓΡΑΦΗΜΑ 3.19: Πραγματικές τιμές και πρόβλεψη υποδείγματος ARIMA (1,1,1)	39
ΓΡΑΦΗΜΑ 3.20: Αυτοσυσχέτιση δεύτερων διαφορών μαθητών/τριών Δημοτικού	39
ΓΡΑΦΗΜΑ 3.21: Μερική αυτοσυσχέτιση δεύτερων διαφορών μαθητών/τριών δημοτικού	40
ΓΡΑΦΗΜΑ 3.22: Αυτοσυσχέτιση καταλοίπων υποδείγματος ARIMA(0,2,1) για το Δημοτικό	41
ΓΡΑΦΗΜΑ 3.23: Μερική αυτοσυσχέτιση καταλοίπων ARIMA(0,2,1)για το Δημοτικό	41
ΓΡΑΦΗΜΑ 3.24: Πραγματικές τιμές και πρόβλεψη υποδείγματος ARIMA(0,2,1)	42
ΓΡΑΦΗΜΑ 4.1: Σύνολο μαθητών /τριών Δευτεροβάθμιας εκπαίδευσης από το 1970/71 - 1996/97	45
ΓΡΑΦΗΜΑ 4.2: Τάση της χρονολογικής σειράς μαθητών/τριών Δευτεροβάθμιας Εκπαίδευσης	46
ΓΡΑΦΗΜΑ 4.3: Άρενες Μαθητές Δευτεροβάθμιας εκπαίδευσης από 1970/71 –1996/97	47
ΓΡΑΦΗΜΑ 4.4: Τάση μαθητών Δευτεροβάθμιας Εκπαίδευσης 1970/71 – 1996/97	48
ΓΡΑΦΗΜΑ 4.5: Χρονολογική σειρά μαθητριών δευτεροβάθμιας εκπαίδευσης από το 1970/71 – 1996/97	49
ΓΡΑΦΗΜΑ 4.6: Τάση της σειράς των μαθητριών δευτεροβάθμιας εκπαίδευσης	50
ΓΡΑΦΗΜΑ 4.7: Μαθητές του δημοτικού και αντίστοιχοι δευτεροβάθμιας 6 χρόνια μετά	52
ΓΡΑΦΗΜΑ 5.1: Σύνολο μαθητών/τριών κατώτερης δευτεροβάθμιας εκπαίδευσης 1970/71- 1996/97	55
ΓΡΑΦΗΜΑ 5.2: Αριθμός μαθητών – μαθητριών κατώτερης δευτεροβάθμιας εκπαίδευσης	55
ΓΡΑΦΗΜΑ 5.3 Η διαφορά μαθητών – μαθητριών κατώτερης δευτεροβάθμιας εκπαίδευσης 1970/71 – 1996/97	56
ΓΡΑΦΗΜΑ 5.4: Λογαριθμική τάση της διαφοράς – αγοριών κοριτσιών	58
ΓΡΑΦΗΜΑ 5.5: Ανάλυση τάσης για τον αριθμό μαθητριών κατώτερης δευτεροβάθμιας εκπαίδευσης από το 1970/71 – 1993/94	60
ΓΡΑΦΗΜΑ 5.6: Προβλέψεις των υποδειγμάτων τάσης για τα έτης 1994/95 – 1996/97	60
ΓΡΑΦΗΜΑ 5.7 : Διπλή εκθετική εξομάλυνση, προσαρμογή και πρόβλεψη για τον αριθμό κοριτσιών στην κατώτερη δευτεροβάθμια εκπαίδευση	62
ΓΡΑΦΗΜΑ 5.8: Προβλέψεις της εκθετικής εξομάλυνσης για τα έτη 1995/95 – 1996/97	63
ΓΡΑΦΗΜΑ 5.9: Συντελεστές αυτοσυσχέτισης αριθμού μαθητριών για τα έτη 1970/71 – 1993/94	63
ΓΡΑΦΗΜΑ 5.10: Συντελεστές μερικής αυτοσυσχέτισης μαθητριών κατώτερης δευτεροβάθμιας εκπαίδευσης	64
ΓΡΑΦΗΜΑ 5.11: Αυτοσυσχέτιση πρώτων διαφορών για τις μαθήτριες	64
ΓΡΑΦΗΜΑ 5.12: Συντελεστές μερικής αυτοσυσχέτισης πρώτων διαφορών για τις μαθήτριες κατώτερης δευτεροβάθμιας εκπαίδευσης	65
ΓΡΑΦΗΜΑ 5.13: Συντελεστές αυτοσυσχέτισης δεύτερων διαφορών για τις μαθήτριες	65
ΓΡΑΦΗΜΑ 5.14: Συντελεστές μερικής αυτοσυσχέτισης δεύτερων διαφορών των μαθητριών κατώτερης δευτεροβάθμιας εκπαίδευσης	66
ΓΡΑΦΗΜΑ 5.15: Συντελεστές αυτοσυσχέτισης καταλοίπων υποδείγματος ARIMA (0,1,1)	67
ΓΡΑΦΗΜΑ 5.16: Συντελεστές μερικής αυτοσυσχέτισης καταλοίπων ARIMA (0,1,1)	68



<u>ΓΡΑΦΗΜΑ 5.17:</u> Αυτοσυσχέτιση καταλοίπων υποδείγματος ARIMA (0,2,2) για τις μαθήτριες.....	68
<u>ΓΡΑΦΗΜΑ 5.18:</u> Συντελεστές μερικής αυτοσυσχέτισης καταλοίπων υποδείγματος ARIMA (0,2,2)..	69
<u>ΓΡΑΦΗΜΑ 5.19:</u> Προβλέψεις και πραγματικές τιμές των υποδειγμάτων ARIMA.....	70
<u>ΓΡΑΦΗΜΑ 5.20:</u> Προσαρμογή των δύο υποδειγμάτων της τάσης για τους άρρενες μαθητές κατώτερης δευτεροβάθμιας εκπαίδευσης	72
<u>ΓΡΑΦΗΜΑ 5.21:</u> Πραγματικές τιμές και προβλέψεις των δύο υποδειγμάτων της τάσης για τους άρρενες μαθητές κατώτερης δευτεροβάθμιας εκπαίδευσης.....	74
<u>ΓΡΑΦΗΜΑ 5.22:</u> Πραγματικές και προσαρμοσμένες τιμές της διπλής εκθετικής εξομάλυνσης για τους μαθητές	75
<u>ΓΡΑΦΗΜΑ 5.23:</u> Πραγματικές και προσαρμοσμένες τιμές της απλής εκθετικής εξομάλυνσης	75
<u>ΓΡΑΦΗΜΑ 5.24:</u> Πραγματικές τιμές και προβλέψεις της Απλής και Διπλής Εκθετικής Εξομάλυνσης για τους άρρενες μαθητές από το 1993/94 – 1996/97	76
<u>ΓΡΑΦΗΜΑ 5.25:</u> Συντελεστές αυτοσυσχέτισης για τους άρρενες μαθητές κατώτερης δευτεροβάθμιας εκπαίδευσης	77
<u>ΓΡΑΦΗΜΑ 5.26:</u> Μερική αυτοσυσχέτιση χρονολογικής σειράς αρρένων μαθητών κατώτερης δευτεροβάθμιας εκπαίδευσης	78
<u>ΓΡΑΦΗΜΑ 5.27:</u> Αυτοσυσχέτιση πρώτων διαφορών των αρρένων μαθητών	78
<u>ΓΡΑΦΗΜΑ 5.28:</u> Μερική αυτοσυσχέτιση πρώτων διαφορών αρρένων μαθητών	79
<u>ΓΡΑΦΗΜΑ 5.29:</u> Αυτοσυσχέτιση καταλοίπων υποδείγματος ARIMA (1,0,0) για τους μαθητές κατώτερης δευτεροβάθμιας εκπαίδευσης	81
<u>ΓΡΑΦΗΜΑ 5.30:</u> Μερική αυτοσυσχέτιση καταλοίπων υποδείγματος ARIMA (1,0,0)	81
<u>ΓΡΑΦΗΜΑ 5.31:</u> Αυτοσυσχέτιση καταλοίπων υποδείγματος ARIMA (!,1,1).....	82
<u>ΓΡΑΦΗΜΑ 5.32:</u> Μερική αυτοσυσχέτιση καταλοίπων ARIMA(1,1,1)	82
<u>ΓΡΑΦΗΜΑ 5.33:</u> Πραγματικές τιμές και προβλέψεις των υποδειγμάτων ARIMA για τα έτη 1993/94 – 1996/97	83
<u>ΓΡΑΦΗΜΑ 6.1:</u> Αριθμός συνόλου μαθητών και μαθητριών ανώτερης δευτεροβάθμιας εκπαίδευσης από το 1970/71 – 1996/97	85
<u>ΓΡΑΦΗΜΑ 6.2:</u> Εξέλιξη αριθμού μαθητών – μαθητριών για τα έτη 1970/71 – 1996/97	86
<u>ΓΡΑΦΗΜΑ 6.3:</u> Εξέλιξη διαφοράς μαθητών – μαθητριών από το 1970/71 – 1996/97	86
<u>ΓΡΑΦΗΜΑ 6.4:</u> Προσαρμογή του λογαριθμικού υποδείγματος στην διαφορά μαθητών μαθητριών ανώτερης δευτεροβάθμιας εκπαίδευσης	88
<u>ΓΡΑΦΗΜΑ 6.5:</u> Προσαρμογή υποδειγμάτων τάσης για τις μαθήτριες ανώτερης δευτεροβάθμιας εκπαίδευσης	90
<u>ΓΡΑΦΗΜΑ 6.6:</u> Προσαρμογή της Διπλής Εκθετικής Εξομάλυνσης για τις μαθήτριες ανώτερης δευτεροβάθμιας εκπαίδευση από το 1970/71 – 1993/94	92
<u>ΓΡΑΦΗΜΑ 6.7:</u> Πραγματικές τιμές και Προβλέψεις υποδειγμάτων $Y_t = a + bt$, $\text{Log } Y_t = a + bt$ και Διπλής Εκθετικής Εξομάλυνσης.....	93
<u>ΓΡΑΦΗΜΑ 6.8:</u> Συντελεστές αυτοσυσχέτισης για τις μαθήτριες ανώτερης δευτεροβάθμιας εκπαίδευσης	93
<u>ΓΡΑΦΗΜΑ 6.9:</u> Συντελεστές μερικής αυτοσυσχέτισης μαθητριών	94
<u>ΓΡΑΦΗΜΑ 6.10:</u> Συντελεστές αυτοσυσχέτισης πρώτων διαφορών για τις μαθήτριες	94
<u>ΓΡΑΦΗΜΑ 6.11:</u> Συντελεστές μερικής αυτοσυσχέτισης πρώτων διαφορών	95
<u>ΓΡΑΦΗΜΑ 6.12:</u> Συντελεστές αυτοσυσχέτισης καταλοίπων υποδείγματος ARIMA (1,0,0)	96
<u>ΓΡΑΦΗΜΑ 6.13:</u> Μερική αυτοσυσχέτιση καταλοίπων υποδείγματος ARIMA(1,0,0)	97
<u>ΓΡΑΦΗΜΑ 6.14:</u> Συντελεστές συσχέτισης καταλοίπων υποδείγματος ARIMA (0,1,5).....	97
<u>ΓΡΑΦΗΜΑ 6.15:</u> Μερική αυτοσυσχέτιση των καταλοίπων του υποδείγματος ARIMA(0,1,5).....	98
<u>ΓΡΑΦΗΜΑ 6.16:</u> Προβλέψεις των υποδειγμάτων ARIMA και πραγματικές τιμές των μαθητριών ανώτερης δευτεροβάθμιας εκπαίδευσης	99
<u>ΓΡΑΦΗΜΑ 6.17:</u> Προσαρμογή υποδειγμάτων τάσης στους μαθητές ανώτερης δευτεροβάθμιας εκπαίδευσης	101
<u>ΓΡΑΦΗΜΑ 6.18:</u> Προβλέψεις και πραγματικές τιμές των υποδειγμάτων της τάσης	102
<u>ΓΡΑΦΗΜΑ 6.19:</u> Προσαρμογή υποδειγμάτος της Διπλής Εκθετικής Εξομάλυνσης για τους άρρενες μαθητές	103
<u>ΓΡΑΦΗΜΑ 6.20:</u> Συντελεστές αυτοσυσχέτισης μαθητών ανώτερης δευτεροβάθμιας εκπαίδευσης....	105
<u>ΓΡΑΦΗΜΑ 6.21:</u> Συντελεστές μερικής αυτοσυσχέτισης μαθητών	105
<u>ΓΡΑΦΗΜΑ 6.22:</u> Αυτοσυσχέτιση πρώτων διαφορών της χρονολογικής σειράς μαθητών ανώτερης δευτεροβάθμιας εκπαίδευσης	106
<u>ΓΡΑΦΗΜΑ 6.23:</u> Μερική αυτοσυσχέτιση πρώτων διαφορών των μαθητών.....	106
<u>ΓΡΑΦΗΜΑ 6.24:</u> Αυτοσυσχέτιση καταλοίπων υποδείγματος ARIMA (1,0,0) για τους μαθητές	108
<u>ΓΡΑΦΗΜΑ 6.25:</u> Μερική αυτοσυσχέτιση καταλοίπων ARIMA (1,0,0)	108
<u>ΓΡΑΦΗΜΑ 6.26:</u> Συντελεστές αυτοσυσχέτισης καταλοίπων υποδείγματος ARIMA (0,1,4) μαθητών ανώτερης δευτεροβάθμιας εκπαίδευσης	109

<u>ΓΡΑΦΗΜΑ 6.27:</u> Συντελεστές μερικής αυτοσυσχέτισης καταλοίπων υποδείγματος ARIMA (0,1,4)	109
<u>ΓΡΑΦΗΜΑ 6.28:</u> Πρόβλεψη υποδειγμάτων ARIMA και πραγματικές τιμές.....	110



ΚΕΦΑΛΑΙΟ 1 :ΧΡΟΝΟΛΟΓΙΚΕΣ ΣΕΙΡΕΣ. ΜΕΘΟΔΟΙ

ΑΝΑΛΥΣΗΣ ΚΑΙ ΠΡΟΒΛΕΨΗΣ

1.1 ΧΡΟΝΟΛΟΓΙΚΕΣ ΣΕΙΡΕΣ ΕΞΕΛΙΞΗΣ ΜΑΘΗΤΩΝ ΠΡΩΤΟΒΑΘΜΙΑΣ ΚΑΙ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ.

Η μείωση των μαθητών στην πρωτοβάθμια και στην δευτεροβάθμια εκπαίδευση (στα γυμνάσια ιδιαίτερα) είναι κάτι παραπάνω από αισθητή ιδιαίτερα τα τελευταία χρόνια. Στην επαρχία κυρίως αλλά και στα αστικά κέντρα μεγάλος είναι ο αριθμός των σχολείων που κλείνουν λόγω της έλλειψης μαθητών.

Η γνώση του αριθμού των μαθητών είναι σίγουρα ζωτικής σημασίας για τον εκπαιδευτικό προγραμματισμό ενός κράτους. Ο επαγγελματικός προσανατολισμός, η δημιουργία σχολικών κτηρίων, οι προσλήψεις δασκάλων και καθηγητών είναι μερικές από τις ενέργειες που επηρεάζονται άμεσα από αυτό.

Το Γράφημα 1.1 παρουσιάζει την διαχρονική εξέλιξη του αριθμού των μαθητών της πρωτοβάθμιας και της δευτεροβάθμιας εκπαίδευσης από το σχολικό έτος 1970/71 μέχρι το 1996/97. Όπως παρατηρούμε από το έτος 1986 έχει αρχίσει να φθίνει ο αριθμός των μαθητών και με εξαίρεση τα έτη 1992 και 1993 η φθίνουσα πορεία συνεχίζεται μέχρι το 1996.

• Τάση

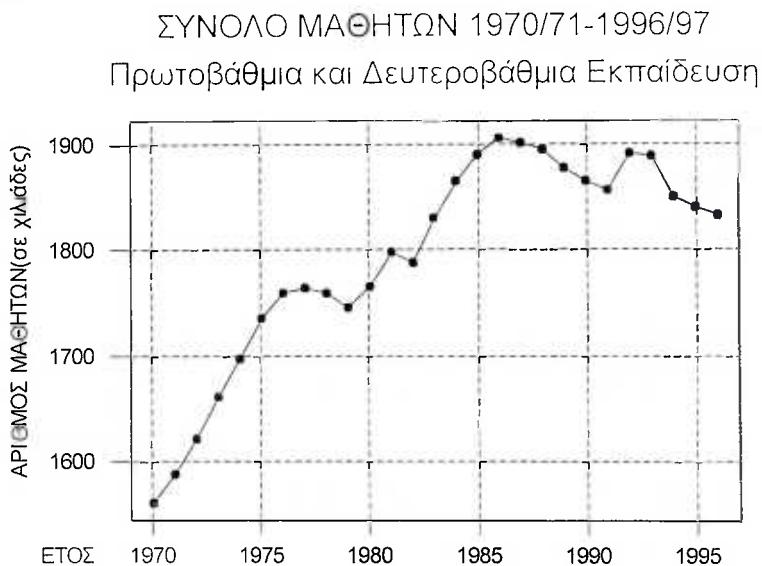
Στην ανάλυση των χρονολογικών σειρών οι μετρήσεις είναι δυνατόν να πραγματοποιούνται κάθε ώρα, μέρα, μήνα ή χρόνο ή ακόμα σε οποιοδήποτε τακτικό χρονικό διάστημα. Η βαθμιαία κίνηση ανοδική ή καθοδική της σειράς καθ' όλο το χρονικό διάστημα που συνήθως οφείλεται σε μακροχρόνιους παράγοντες και που αποτυπώνεται στην σειρά λέγεται τάση. Στο Γράφημα 1.1, ο αριθμός των μαθητών παρουσιάζει αυξητική τάση μέχρι το 1986, ενώ στη συνέχεια εμφανίζεται πτωτικός.

• Εποχικότητα

Ενώ η τάση αναγνωρίζεται αναλύοντας πολύχρονες κινήσεις, πολλές σειρές εμφανίζουν ένα συγκεκριμένο “μοτίβο” μεταβλητότητας μέσα σε περίοδο ενός χρόνου. Για παράδειγμα οι πωλήσεις κλιματιστικών παρουσιάζουν πτώση κατά τους φθινοπωρινούς και χειμερινούς μήνες και κορύφωση την άνοιξη και το καλοκαίρι. Η

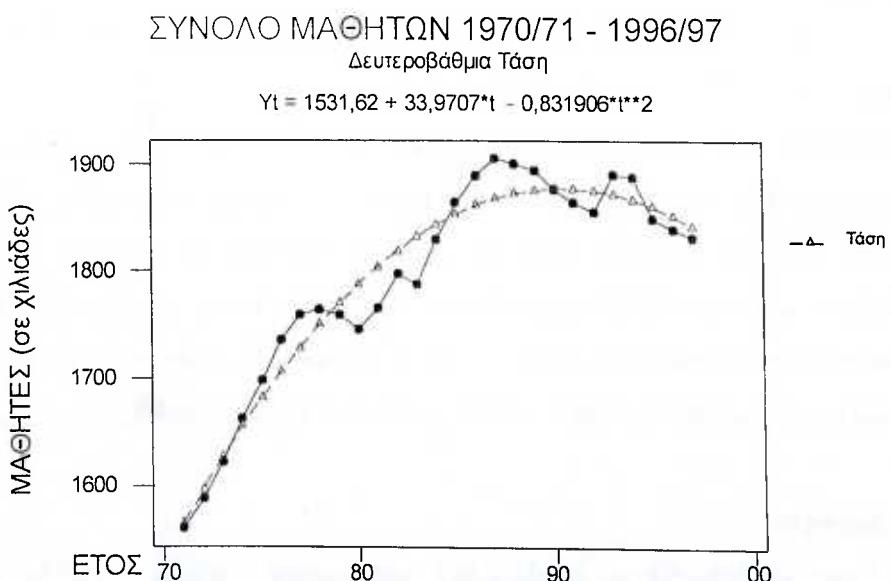


συνιστώσα αυτή ονομάζεται εποχικότητα. Στο Γράφημα 1.1 δεν παρουσιάζεται αυτή η συνιστώσα διότι τα δεδομένα προέρχονται από τιμές που αντιστοιχούν στο έτος.



ΠΗΓΗ:EUROSTAT

ΓΡΑΦΗΜΑ 1.1: Εγγεγραμμένοι μαθητές Πρωτοβάθμιας και Δευτεροβάθμιας εκπαίδευσης



ΓΡΑΦΗΜΑ 1.2: Δευτεροβάθμιας τάση του συνόλου των μαθητών

• Κυκλικότητα

Η κυκλικότητα είναι έννοια παρόμοια με την εποχικότητα με την διαφορά ότι εμφανίζεται σε χρονικά διαστήματα μεγαλύτερα του έτους και δεν είναι ομοιόμορφή. Αναγνωρίζεται παρατηρώντας εναλλασσόμενες ακολουθίες σημείων κάτω και πάνω από την υποτιθέμενη τάση (Γράφημα 1.2). Στο παράδειγμα μας θα μπορούσαμε να υποθέσουμε ότι εμφανίζονται τέσσερις κύκλοι. Από το 70-80, από το 80-83, από το 83-92 και το 92-96.

• Τυχαιότητα

Ο τυχαίος παράγοντας των χρονοσειρών είναι οι αποκλίσεις των πραγματικών τιμών από τις αναμενόμενες τιμές που προέρχονται από τον συνδυασμό των άλλων τριών παραγόντων (Τάση, Εποχικότητα, Κυκλικότητα).

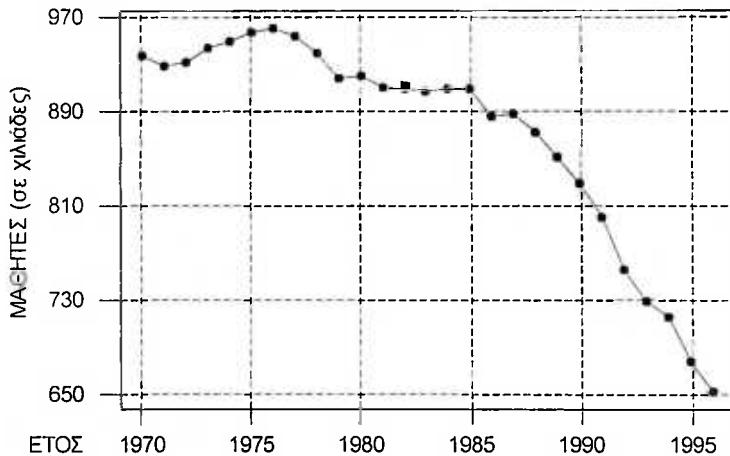
Στα Γραφήματα 1.1 και 1.2 έχουμε μια γενική εικόνα των μαθητών της πρωτοβάθμιας και δευτεροβάθμιας εκπαίδευσης. Ωστόσο η εικόνα αυτή είναι πολύ γενική για αναλύσεις και συμπεράσματα. Είναι σίγουρα προτιμότερο να την χωρίσουμε στα επί μέρους κομμάτια που αποτελούν αυτήν την σειρά. Σε μαθητές Δημοτικού, Γυμνασίου και Λυκείου και σε αγόρια – κορίτσια όπου αυτό απαιτείται.

Τα Γραφήματα 1.3, 1.4 και 1.5 δείχνουν διαφορετική εικόνα να καταγράφεται σε κάθε βαθμίδα εκπαίδευσης από αυτήν που εμφανίζεται στο σύνολο. Πιο συγκεκριμένα για την διαχρονική εξέλιξη του αριθμού των μαθητών Δημοτικού έχουμε μια διαρκή πτωτική πορεία από το έτος 1976/77 μέχρι το έτος 1996/97. Οι μαθητές της Κατώτερης Δευτεροβάθμιας Εκπαίδευσης έχουν αντίθετα αυξητική πορεία μέχρι το 1981/82 και συνεχίζεται με μείωση μέχρι το τελευταίο έτος που εξετάζεται. Αντίθετα με τα προηγούμενα ο αριθμός των μαθητών Ανώτερης Δευτεροβάθμιας Εκπαίδευσης αυξάνει διαρκώς, με εξαίρεση το έτος 1978/79 μέχρι το τέλος

Όλες οι παραπάνω χρονολογικές σειρές (Γραφήματα 1.3, 1.4, 1.5) εμφανίζουν μόνο την τάση και την άρρυθμη συνιστώσα, από τις συνιστώσες που αναλύσαμε, καθώς αφ' ενός η φύση των δεδομένων είναι τέτοια με την οποία δεν έχουμε εποχικότητα αφετέρου είναι δυσδιάκριτη κυκλικότητα μέσα σε αυτό το σχετικά μικρό χρονικό διάστημα. Με βάση τις παραπάνω διαπιστώσεις η ανάλυση τους θα γίνει με μεθόδους, Παλινδρόμησης, Εκθετικής Εξομάλυνσης και μεθόδους ARIMA. Με αυτές, προσαρμόζοντας το κατάλληλο μοντέλο, θα εκτιμήσουμε την τάση (μόνο με

την παλινδρόμηση) και θα προσπαθήσουμε δώσουμε κάποιες προβλέψεις για 3 έτη μετά την τελευταία γνωστή τιμή εκτιμώντας την αποτελεσματικότητα των μεθόδων.

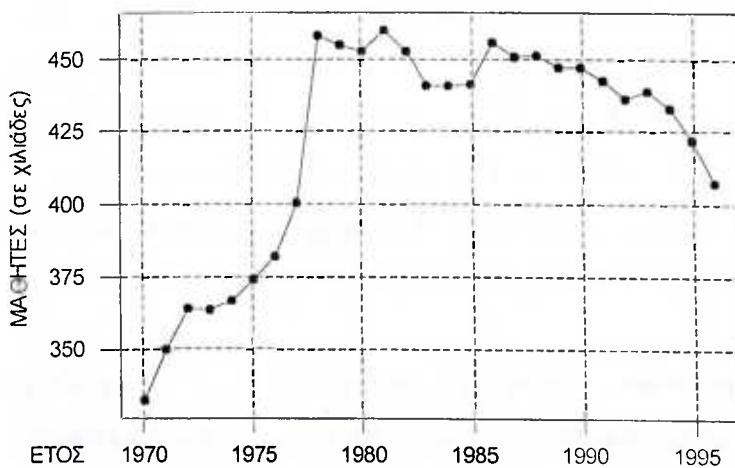
ΑΡΙΘΜΟΣ ΜΑΘΗΤΩΝ ΔΗΜΟΤΙΚΟΥ 1970/71-1996/97



ΠΗΓΗ: Eurostat

ΓΡΑΦΗΜΑ 1.3: Αριθμός Μαθητών Δημοτικού από το 1970/71 – 1996/97

ΜΑΘΗΤΕΣ ΚΑΤΩΤΕΡΗΣ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ 1970/71 - 1996/97

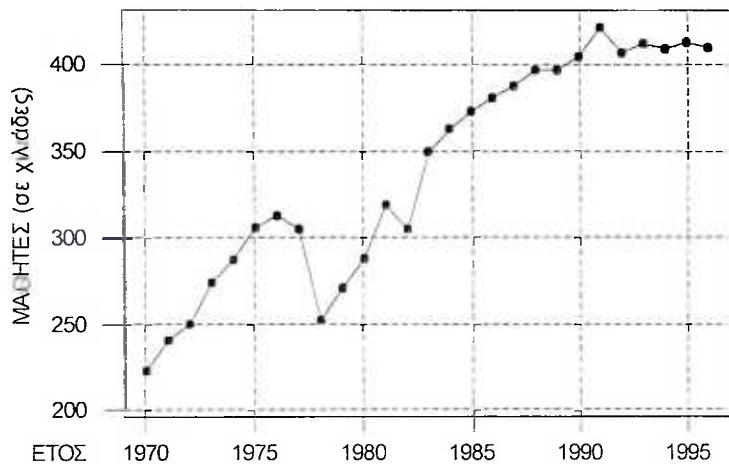


ΠΗΓΗ: Eurostat

ΓΡΑΦΗΜΑ 1.4: Αριθμός μαθητών Κατώτερης Δευτεροβάθμιας Εκπαίδευσης (ISCED 3)

ΜΑΘΗΤΕΣ ΑΝΩΤΕΡΗΣ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ

1970/71 - 1996/97



ΠΗΓΗ: Eurostat

ΓΡΑΦΗΜΑ 1.5: Αριθμός μαθητών Ανώτερης Δευτεροβάθμιας Εκπαίδευσης (ISCED 4)

1.2 ΠΑΛΙΝΔΡΟΜΗΣΗ

Η παλινδρόμηση στην γενική της μορφή προσπαθεί να εξηγήσει την σχέση δύο η περισσότερων μεταβλητών. Η πιο απλή περίπτωση είναι η γραμμική σχέση δύο μεταβλητών, μιας ανεξάρτητης (x) και μίας εξαρτημένης (y) η οποία εκφράζεται από το μοντέλο:

$$y = \beta_0 + \beta_1 x + \varepsilon$$

Στην περίπτωση μας η ανεξάρτητη μεταβλητή μπορεί να θεωρηθεί ο χρόνος t . Έτσι το μοντέλο μπορεί να γραφεί ανάλογα για την περίπτωση μας:

$$y = \beta_0 + \beta_1 t + \varepsilon$$

Η εκτίμηση των β_0 και β_1 γίνεται προσπαθώντας να ελαχιστοποιήσουμε το τετράγωνο του σφάλματος ε που είναι η διαφορά των πραγματικών τιμών της y από τις τιμές που δίνει η ευθεία γραμμή $\hat{y} = \beta_0 + \beta_1 t$ δηλ. την ελαχιστοποίηση του παρακάτω τύπου:

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Η μέθοδος αυτή λέγεται κριτήριο ελαχίστων τετραγώνων.

Η εκτίμηση των συντελεστών β_0 και β_1 , που αποτελούν το σημείο τομής του άξονα για από την ευθεία παλινδρόμησης και την κλίση της ευθείας παλινδρόμησης αντίστοιχα, και προκύπτουν από την ελαχιστοποίηση του τετραγωνικού σφάλματος, γίνεται από τους τύπους.

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{t}$$

$$\hat{\beta}_1 = \frac{n \sum_{i=1}^n y_i t_i - \sum_{i=1}^n y_i \sum_{i=1}^n t_i}{n \sum_{i=1}^n t_i^2 - \left(\sum_{i=1}^n t_i \right)^2}$$

Η στατιστική σημαντικότητα της β_1 και κατά συνέπεια η προσαρμογή της γραμμής παλινδρόμησης (δεδομένου ότι αν $\beta_1=0$ δεν έχουμε γραμμή) γίνεται με τον έλεγχο υπόθεσης:

$$H_0: \beta_1 = 0$$

$$H_1: \beta_1 \neq 0$$

Ο έλεγχος της παραπάνω υπόθεσης γίνεται με την χρήση της κατανομής t-student και η ελεγχοσυνάρτηση η οποία μας δίνει την τιμή για την σύγκριση δίνεται από τον τύπο:

$$t = \frac{\hat{\beta}_1}{S_{\hat{\beta}_1}}$$

όπου

$$S_{\hat{\beta}_1} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2 / (n-2)}{\sum_{i=1}^n (t_i - \bar{t})^2}}$$

Αν η τιμή της t είναι μεγαλύτερη από την κρίσιμη τιμή $t_{n,a}$ που λαμβάνεται από τους πίνακες της κατανομής t-student, με n να είναι ο αριθμός των παρατηρήσεων και a το επίπεδο εμπιστοσύνης, απορρίπτεται η μηδενική υπόθεση.

Ο έλεγχος της στατιστικής σημαντικότητας της παλινδρόμησης πραγματοποιείται επίσης με την κατανομή F. Χρησιμοποιώντας την ελεγχοσυνάρτηση:

$$F = \frac{\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{1}}{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{(n-2)}}$$

Η σύγκριση της τιμής που δίνει η ελεγχοσυνάρτηση γίνεται με την τιμή $F_{1,n-2,a}$ η οποία υπολογίζεται από τον πίνακα της F συνάρτησης. Αν $F > F_{1,n-2,a}$ τότε με ασφάλεια μπορούμε να απορρίψουμε την υπόθεση $\beta_1=0$, και να συνεχίσουμε τις εργασίες μας.

Ένας τρόπος για την επιλογή του πιο κατάλληλου μοντέλου παλινδρόμησης είναι η τυπική απόκλιση ή τυπικό σφάλμα της παλινδρόμησης. Η μαθηματική εξίσωση για τον υπολογισμό είναι:

$$S_y = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{(n-2)}}$$

Ένα κάλο μοντέλο θα πρέπει να έχει χαμηλή τιμή που σημαίνει ότι οι τιμές \hat{y}_i δεν αποκλίνουν πολύ από τις πραγματικές τιμές y_i .

Η καλή προσαρμογή του μοντέλου στα δεδομένα μπορεί επίσης να πραγματοποιηθεί με τον συντελεστή προσδιορισμού R^2 .

$$R^2 = \frac{\text{εξηγούμενη από το μοντέλο μεταβλητήτα}}{\text{συνολική ματαβλητήτα των δεδομένων}} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Όπως θα ανέμενε κανείς όσο μεγαλύτερη η μεταβλητήτα που εξηγείται από το μοντέλο τόσο καλύτερα αυτό “ταιριάζει” στα δεδομένα και κατά συνέπεια όσο πιο υψηλή η τιμή τόσο το καλύτερο.

1.2.1 Πολλαπλή παλινδρόμηση

Για να περιγραφεί η σχέση μιας εξαρτημένης μεταβλητής με περισσότερες από μία ανεξάρτητης, όπως είναι φυσικό δεν αρκεί η απλή παλινδρόμηση. Στις χρονολογικές σειρές αντίστοιχα η τάση δεν είναι πάντα ευθύγραμμη και κατά συνέπεια το απλό μοντέλο δεν είναι ικανό να εξηγήσει τα διαφορετικά είδη τάσεων διαφορετικών χρονολογικών σειρών.

Στην γενική περίπτωση της εξαρτημένης με τις ανεξάρτητες μεταβλητές το μοντέλο μπορεί να γραφεί με την μορφή:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon$$

Στην τάση των χρονολογικών σειρών δεν έχουμε περισσότερες από μία ανεξάρτητη μεταβλητή, υπάρχει όμως η περίπτωση μοντέλου δευτεροβάθμιας ή και τριτοβάθμιας τάσης όπου η ανεξάρτητη χρονική μεταβλητή χρησιμοποιείται παραπάνω από μία φορά :

$$y = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3$$

Η περίπτωση είναι ίδια με την γενική περίπτωση της παλινδρόμησης αφού θα μπορούσαμε να θεωρήσουμε τα t , t^2 , t^3 διαφορετικές ανεξάρτητες μεταβλητές x_1 , x_2 , x_3 .

Ο υπολογισμός των συντελεστών αποτελεί μια επίπονη εργασία η οποία διευκολύνεται από την χρήση των στατιστικών προγραμμάτων. Η τύποι για τον υπολογισμό των συντελεστών δεν παρουσιάζονται γιατί είναι πέρα από τους σκοπούς αυτής της εργασίας.

Η στατιστική σημαντικότητα της γραμμής παλινδρόμησης εξετάζεται από τον έλεγχο υποθέσεων:

$$H_0: \beta_1 = \beta_2 = \beta_3 = \dots = \beta_k = 0$$

$$H_1: \beta_i \neq 0 \text{ για κάποιο } i$$

Η στατιστική συνάρτηση που χρησιμοποιείται για τον έλεγχο αυτό είναι :

$$F = \frac{\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{k}}{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{(n - k - 1)}}$$

όπου κ οι βαθμοί ελευθερίας του μοντέλου της παλινδρόμησης με κ ανεξάρτητες μεταβλητές και $n-k-1$ οι βαθμοί ελευθερίας των λαθών με n τον αριθμό παρατηρήσεων που διαθέτουμε. Η τιμή αυτή συγκρίνεται με την τιμή $F_{k,n-k-1,a}$ από τους πίνακες της F κατανομής. Αν $F_{k,n-k-1,a} > F$ τότε απορρίπτεται η μηδενική υπόθεση. Όπως είδαμε στο απλό μοντέλο $k=1$ και $n-k-1=n-2$.

Ο έλεγχος για κάθε συντελεστή ξεχωριστά είναι:

$$H_0: \beta_i = 0$$

$$H_1: \beta_i \neq 0$$



και πραγματοποιείται από την στατιστική συνάρτηση:

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

αν $|T| > t_{n-k-1,a/2}$ τότε απορρίπτεται η μηδενική υπόθεση.

Για τις συγκρίσεις μεταξύ διαφορετικών μοντέλων μπορεί να χρησιμοποιηθεί όπως και στην απλή παλινδρόμηση η τυπική απόκλιση του μοντέλου S_y που δίνεται από τον τύπο:

$$S_y = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{(n - k - 1)}}$$

Με n τον αριθμό των δεδομένων και k τον αριθμό των ανεξάρτητων μεταβλητών που γρηγοριούνται στο μοντέλο.

Επίσης μπορεί να χρησιμοποιηθεί και ο συντελεστής προσδιορισμού:

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Επειδή ωστόσο η τιμή του επηρεάζεται από τον αριθμό των ανεξάρτητων μεταβλητών που χρησιμοποιούνται, πιο ακριβής είναι ο προσαρμοσμένος συντελεστής προσδιορισμού R^2_{adj} :

$$R^2_{adj} = 1 - [(1 - R^2) \frac{n-1}{n-2}]$$

1.2.2 Μοντέλα που μετατρέπονται σε γραμμικά.

Πολλές φορές η σχέση δύο ή περισσότερων μεταβλητών δεν είναι γραμμική αλλά με κάποιο μετασχηματισμό μετατρέπεται σε γραμμική όπου ισχύουν όσα αναφέραμε παραπάνω. Τέτοιοι μετασχηματισμοί είναι κυρίως από εκθετικά μοντέλα που περιγράφουν την σχέση δύο ή περισσότερων μεταβλητών και στη συγκεκριμένη περίπτωση της εξαρτημένης μεταβλητής για με τον χρόνο t. Τέτοια μοντέλα εκφράζουν κυρίως τάσεις εκθετικής μορφής:

	ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΣ	ΓΡΑΜΜΙΚΟ ΜΟΝΤΕΛΟ
$y = \beta_0 \cdot \beta_1^{t^1} \cdot \beta_2^{t^2} \cdots \beta_n^{t^n}$	$y' = \log y$	$y' = \log \beta_0 + \log \beta_1(t) + \log \beta_2(t^2) + \dots + \log \beta_n(t^n)$
$y = k \cdot (\beta_0)^{t^1}$	$y' = \log(\log y)$	$y' = \log(\log k) + \log(\log \beta_0) + \log \beta_1(t)$
$y = \frac{k}{1 + e^{\beta_0 + \beta_1 t}}$	$y' = \log(1/y)$	$y' = \log(1/k) + \log \beta_0 + \log \beta_1(t)$
$y = \beta_0 + \beta_1 \log t$	$t' = \log t$	$y = \beta_0 + \beta_1 t'$
$y = \beta_0(t)^{\beta_1}$	$y' = \log y, t' = \log t$	$y' = \log \beta_0 + \beta_1 t'$

Πίνακας 1.1: Υποδειγματα που μετασχηματίζονται σε γραμμικά

1.3 ΕΚΘΕΤΙΚΗ ΕΞΟΜΑΛΥΝΣΗ

Οι μέθοδοι εκθετικής εξομάλυνσης είναι μία διαδικασία με την οποία απομακρύνουμε τις τυχαίες κυμάνσεις των δεδομένων χρησιμοποιώντας σταθμισμένους μέσους όρους των δεδομένων, έτσι ώστε να εμφανιστεί το κρυμμένο σχήμα τους, δίνοντας όμως μεγαλύτερη βαρύτητα στις πρόσφατες παρατηρήσεις. Αυτό γίνεται κυρίως για να μπορεί το μοντέλο να “αντιληφθεί” πιο εύκολα τις αλλαγές που εμφανίζονται. Η λέξη εκθετική αναφέρεται στο γεγονός ότι οι συντελεστές στάθμισης είναι άνισοι και μειώνονται εκθετικά ξεκινώντας από τις πιο πρόσφατες παρατηρήσεις στις πιο παλιές.

Οι πιο γνωστές από τις μεθόδους εκθετικής εξομάλυνσης είναι:

1. Η Απλή Εκθετική Εξομάλυνση
2. Η Γραμμική Εκθετική Εξομάλυνση του Holt, ή και Διπλή Εκθετική Εξομάλυνση.
3. Η Γραμμική Εκθετική Εξομάλυνση των Holt-Winters.

Η πρώτη αναφερόμενη μέθοδος εφαρμόζεται κυρίως σε σειρές που δεν εμφανίζουν τάση και εποχικότητα και για αυτό τον λόγο δεν μπορεί να δώσει φερέγγυα αποτελέσματα στις χρονολογικές σειρές που εξετάζουμε.

Η δεύτερη μέθοδος έχει εφαρμογή σε σειρές που παρουσιάζουν μόνο τάση και όχι εποχική συνιστώσα και είναι εκείνη με την οποία θα αναλύσουμε τα δεδομένα και θα προχωρήσουμε σε προβλέψεις.

Η τελευταία και πιο σύνθετη μέθοδος δίνει προβλέψεις λαμβάνοντας υπόψη και την συνιστώσα της τάσης και την εποχική. Η έλλειψη εποχικότητας στα δεδομένα, μας αποτρέπει από την χρήση της.

1.3.1 Εκθετική εξομάλυνση του Holt.

Η μέθοδος αυτή συνοψίζεται σε τρεις εξισώσεις:

$$S_t = \alpha Y_t + (1-\alpha)(S_{t-1} + T_{t-1})$$

$$T_t = \gamma(S_t - S_{t-1}) + (1-\gamma)T_{t-1}$$

$$F_{t+m} = S_t + T_t m$$

S_t = επίπεδο την χρονική τιμή t .

α = συντελεστής στάθμισης με τιμές μεταξύ 0 και 1. Όσο μεγαλύτερη η τιμή του τόσο μεγαλύτερη έμφαση δίνεται στις χρονικά πιο πρόσφατες παρατηρήσεις.

T_t = εξομαλυσμένη τάση για την χρονική στιγμή t .

γ = συντελεστής ανάλογος του α .

F_{t+m} = πρόβλεψη τη περιόδους μετά την χρονική στιγμή t

Ο υπολογισμός των σταθμίσεων α και γ γίνεται ελαχιστοποιώντας το μέσο τετραγωνικό σφάλμα(Mean Square Error) κατά την προσαρμογή στα δεδομένα:

$$MSE = \frac{\sum_{i=1}^n (Y_i - F_i)^2}{n}$$

Η ελαχιστοποίηση αυτή γίνεται με την διαδικασία της δοκιμής διαφόρων τιμών. Είναι, ωστόσο, δυνατόν να υπολογιστούν οι βέλτιστες τιμές απευθείας από το στατιστικό πακέτο.

1.4 ΜΕΘΟΔΟΙ ARMA

Με τις μεθόδους αυτές προσπαθούμε να εκμεταλλευτούμε τη σχέση που μπορεί να εμφανίζεται μεταξύ τιμών της ίδιας μεταβλητής αλλά σε διαφορετικές χρονικές περιόδους. Η σχέση αυτή στις χρονολογικές σειρές λέγεται αυτοσυσχέτιση και είναι πολύ σημαντική για τον καθορισμό του κατάλληλου μοντέλου, υπολογίζεται από τον τύπο:

$$r_k = \frac{\sum_{t=k+1}^n (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2}$$

όπου $y_t = \eta$ τιμή της χρονοσειράς την χρονική στιγμή t

$\bar{y} = \eta$ μέση τιμή της y

$y_{t-k} = \eta$ τιμή της χρονοσειράς με χρονική υστέρηση k περιόδων από την στιγμή t .

Αν η τιμή της r_k είναι μη μηδενική τότε υπάρχει σειριακή συσχέτιση κ τάξης.

Πέρα από την αυτοσυσχέτιση υπάρχει και η “μερική αυτοσυσχέτιση”. Η μερική αυτοσυσχέτιση κ τάξης στην πραγματικότητα είναι ο συντελεστής φ_k της παλινδρόμησης της μεταβλητής y_{t-k} , στην εξίσωση:

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_k y_{t-k} + \varepsilon$$

1

Η μερική αυτοσυσχέτιση μετρά την συσχέτιση των τιμών της y που διαφέρουν κ χρονικές περιόδους αφού όμως αφαιρεθεί η συσχέτιση των ενδιάμεσων χρονικών υστερήσεων $k-1, k-2, k-3, \dots, 1$.

Τα μοντέλα ARMA (αυτοπαλίνδρομα υποδείγματα κινητών μέσων) αποτελούνται από δύο μέρη. Τα υποδείγματα AR και τα υποδείγματα MA. Το πρώτο μέρος (AR ή

αυτοπαλίνδρομο) βασίζεται στην παλινδρόμηση μεταξύ της μεταβλητής και των χρονικών υστερήσεων της. Ένα AR(k) αυτοπαλίνδρομο μοντέλο κ τάξης φαίνεται στην εξίσωση:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_k y_{t-k} + \varepsilon_t$$

Το δεύτερο μέρος το MA (κινητοί μέσοι) υπολογίζονται από το μοντέλο:

$$y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_k \varepsilon_{t-k}$$

όπου ε_t = το σφάλμα την χρονική στιγμή t
και $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-k}$ οι τιμές του τις προηγούμενες χρονικές στιγμές μέχρι κ περιόδους πριν. Το μοντέλο αυτό εξετάζει την εξάρτηση της μεταβλητής γ με τις προηγούμενες τιμές του σφάλματος.

Τα μοντέλα αυτά μπορούν να είναι AR(p) μόνο, MA(q) ή και μεικτά AR(p)MA(q). Ωστόσο τα παραπάνω δεν αρκούν για να εφαρμόσουμε την μέθοδο σε οποιαδήποτε σειρά. Μια σημαντική προϋπόθεση είναι η σειρά να είναι στάσιμη, δηλαδή ο μέσος και η διακύμανση να παραμένουν σταθερές κατά την διάρκεια των χρονικών περιόδων. Αν μια σειρά δεν είναι στάσιμη μπορεί να μετατραπεί, παίρνοντας διαδοχικά τις πρώτες διαφορές ($\Delta y_t = y_t - y_{t-1}$), τις δεύτερες ($\Delta^2 y_t = \Delta y_t - \Delta y_{t-1}$), τις τρίτες ($\Delta^3 y_t = \Delta^2 y_t - \Delta^2 y_{t-1}$) κοκ μέχρι να γίνει στάσιμη. Συνήθως παίρνουμε πρώτες διαφορές αν η τάση που παρουσιάζουν τα δεδομένα είναι γραμμική, δεύτερες διαφορές αν είναι δευτεροβάθμια κοκ. Η διαδικασία αυτή ονομάζεται ολοκλήρωση (intergration) κ βαθμού ανάλογα με τις διαφορές που έχουμε πάρει και το μοντέλο ονομάζεται ARIMA(p,k,q). Κατά συνέπεια η μέθοδος μπορεί να εφαρμοστεί σε οποιαδήποτε σειρά.

Αφού εξασφαλίσουμε ότι η σειρά είναι στάσιμη θα πρέπει να προσδιορίσουμε τις τιμές των p και q ώστε να εφαρμόσουμε το υπόδειγμα. Η επιλογή των κατάλληλων τιμών γίνεται συνήθως εξετάζοντας τις αυτοσυσχετίσεις και τις μερικές αυτοσυσχετίσεις της στάσιμης χρονοσειράς. Σαν γενικός κανόνας, όταν η αυτοσυσχέτιση μειώνεται εκθετικά στο 0, το υπόδειγμα είναι AR και η τάξη p καθορίζεται από τον αριθμό των μερικών αυτοσυσχετίσεων που είναι στατιστικά διαφορετικές από το 0. Αν οι μερικές αυτοσυσχετίσεις μειώνονται εικθετικά τότε

μιλάμε για υπόδειγμα MA του οποίου η τάξη q καθορίζεται από τις τιμές της αυτοσυσχέτισης που είναι στατιστικά σημαντικά διαφορετικές από το 0.

Η καλή προσαρμογή ενός μοντέλου ARMA, προϋποθέτει τα λάθη $\varepsilon_1, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots$ να είναι ασυσχέτιστα. Ο έλεγχος αυτός είναι δυνατός με δύο τρόπους. Πρώτον με των υπολογισμό και γραφική αναπαράσταση των αυτοσυσχέτισεων των καταλοίπων (residuals) ή με την τιμή της συνάρτησης Durbin-Watson. Τιμές αυτής μεταξύ του 1,5 και 2,5 δείχνουν έλλειψη συσχέτισης μεταξύ των καταλοίπων του μοντέλου. Υπολογίζεται από τον τύπο:

$$DW = \frac{\sum_{t=2}^n (\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^n \hat{\varepsilon}_t^2}$$

Η επιλογή του καλύτερου υποδείγματος ARMA γίνεται με τα κριτήρια των Akaike και Schwarz. Όσο μικρότερη η τιμή των παραπάνω τόσο καλύτερη είναι η προσαρμογή του υποδείγματος στα δεδομένα. Το Schwarz προτιμάται από το Akaike γιατί δεν επιτρέπεται από τον αριθμό των μεταβλητών που χρησιμοποιούμε στα υποδείγματα για τον υπολογισμό.

1.5 ΜΕΤΡΑ ΑΚΡΙΒΕΙΑΣ

Η σύγκριση της αποτελεσματικότητας των μοντέλων όσον αφορά τις προβλέψεις θα γίνει με τρία συγκεκριμένα μέτρα. Αυτά είναι MSE (μέσο τετραγωνικό σφάλμα), MAD (μέσο απόλυτο σφάλμα) και MAPE (Μέσο απόλυτο ποσοστιαίο σφάλμα).

1.5.1 Μέσο Τετραγωνικό Σφάλμα (MSE ή MSD)

$$MSE = \frac{\sum (y_i - f_i)^2}{n}$$

Το μέτρο αυτό υπολογίζει την μέση τιμή της τετραγωνικής απόκλισης της πρόβλεψης από την πραγματική τιμή. Με την μέθοδο αυτή δίνεται μεγαλύτερο βάρος σε μεγάλες αποκλίσεις από τις πραγματικές τιμές, μετατρέποντας σε ένα ευαίσθητο μέτρο σε μεγάλα σφάλματα. Αυτό είναι σημαντικό στοιχείο διότι σίγουρα είναι προτιμητέα υποδείγματα με σταθερά μετρίου μεγέθους σφάλματα από υποδείγματα με μικρότερα λάθη και περιστασιακά, έστω, μεγάλα. Μειονέκτημα του είναι δεν υπολογίζεται από την πραγματική τιμή.

ίδιες μονάδες με τα δεδομένα το οποίο ωστόσο λύνεται αν υπολογίσουμε την τετραγωνική του ρίζα.

1.5.2 Μέσο Απόλυτο Σφάλμα (MAD)

$$MAD = \frac{\sum |y_i - f_i|}{n}$$

Το μέσο απόλυτο σφάλμα δεν δίνει ιδιαίτερη βαρύτητα σε μικρά η μεγάλα λάθη, απλά υπολογίζει την μέση τιμή της απόλυτης τιμής τους. Ένα από τα πλεονεκτήματα του είναι ότι μετράται σε ίδιες μονάδες με τα δεδομένα.

1.5.2 Μέσο απόλυτο ποσοστιαίο σφάλμα (MAPE)

$$MAPE = \frac{\sum \left| \frac{y_i - f_i}{y_i} \right|}{n} \cdot 100$$

Το μέτρο αυτό είναι καθαρός αριθμός και κατά συνέπεια μπορεί να χρησιμοποιηθεί για τη σύγκριση δύο εντελώς διαφορετικών μοντέλων.

ΚΕΦΑΛΑΙΟ 2 :ΜΟΡΦΗ ΠΡΩΤΟΒΑΘΜΙΑΣ ΚΑΙ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ 1970/71-1996/97

2.1 ΠΡΩΤΟΒΑΘΜΙΑ ΕΚΠΑΙΔΕΥΣΗ

Η πρωτοβάθμια εκπαίδευση στην Ελλάδα κατά τα έτη στα οποία θα την μελετήσουμε αποτελείτο από τα δημοτικά σχολεία και η παρακολούθηση ήταν όπως και εξακολουθεί να είναι υποχρεωτική. Η διάρκεια της είναι 6 έτη και τα παιδιά εγγράφονται στη πρώτη τάξη από την ηλικία 5 ετών και 6 μηνών.

Στην πρωτοβάθμια εκπαίδευση μπορεί να τοποθετηθεί και το νηπιαγωγείο με την διαφορά ότι η φοίτηση του δεν ήταν υποχρεωτική.

Για την ίδρυση Δημοτικού σχολείου απαιτούνται τουλάχιστον 15 παιδιά αλλά όπου υπαγορεύονται εθνικοί ή γεωγραφικοί λόγοι αυτά μπορούν να ιδρυθούν μέχρι και με 7 μαθητές. Σε περιπτώσεις σχολείων με πολλά παιδιά θα πρέπει να αντιστοιχούν 30 παιδιά σε κάθε δάσκαλο

2.2 ΔΕΥΤΕΡΟΒΑΘΜΙΑ ΕΚΠΑΙΔΕΥΣΗ

Στην δευτεροβάθμια εκπαίδευση φοιτούν μαθητές που περάτωσαν το δημοτικό. Η ελάχιστη ηλικία για μαθητές της δευτεροβάθμιας είναι τα 11 χρόνια. Η δευτεροβάθμια εκπαίδευση διακρίνεται στην γενική εκπαίδευση και στην τεχνική επαγγελματική και εκκλησιαστική εκπαίδευση. Η Eurostat διακρίνει και τους δύο αυτούς τύπους της εκπαίδευσης σε δύο επιμέρους κατηγορίες: κατώτερη δευτεροβάθμια εκπαίδευση (lower secondary education –ISCED 2) η οποία είναι η εκπαίδευση που ακολουθεί μετά το δημοτικό και η ανώτερη δευτεροβάθμια εκπαίδευση (upper secondary education-ISCED 3) η οποία ακολουθεί μετά την ολοκλήρωση της κατώτερης. Στην συγκεκριμένη εργασία για συντομία θα αναφερόμαστε σε γυμνάσιο και λύκειο εννοώντας την κατώτερη δευτεροβάθμια και ανώτερη δευτεροβάθμια εκπαίδευση αντίστοιχα.

Η γενική εκπαίδευση περιλαμβάνει:

Α) Γα γυμνάσια. Τα παρακολουθούν τα παιδιά μετά το τέλος της πρωτοβάθμιας εκπαίδευσης. Υπάρχουν τα ημερήσια ή εσπερινά και τα χρόνια φοίτησης είναι 3 η 4

αντίστοιχα. Η φοίτηση στα γυμνάσια είναι υποχρεωτική. Ανήκουν στην κατώτερη δευτεροβάθμια εκπαίδευση (ISCED 2).

Β) Τα γενικά λύκεια. Υπάρχουν ημερήσια ή εσπερινά και τα χρόνια φοίτησης είναι 3 ή 4 αντίστοιχα. Στα γενικά λύκεια μπορούν να εγγραφούν χωρίς εξετάσεις τα παιδιά μετά την ολοκλήρωση της φοίτησης τους στο γυμνάσιο. Τα γενικά λύκεια με τη σειρά τους διακρίνονται σε γενικής κατεύθυνσης και σε πειραματικά-κλασικά. Ανήκουν στην ανώτερη δευτεροβάθμια εκπαίδευση (ISCED 3).

Γ) Τα ενιαία πολυκλαδικά. Υπάρχουν μόνο ημερήσια, είναι τριετούς φοίτησης και σ' αυτά μπορούν να εγγραφούν όσοι έχουν ολοκληρώσει την γυμνασιακή τους εκπαίδευση. Ανήκουν και αυτά στην ανώτερη δευτεροβάθμια εκπαίδευση (ISCED 3).

Η δευτεροβάθμια τεχνική-επαγγελματική και εκκλησιαστική εκπαίδευση αποτελείται από:

Α) Τα τεχνικά επαγγελματικά λύκεια. Υπάρχουν ημερήσια και εσπερινά με έτη φοίτησης 3 και 4 χρόνια αντίστοιχα. Στα λύκεια αυτά μπορούν να εγγραφούν χωρίς εξετάσεις όσοι έχουν ολοκληρώσει την φοίτηση τους στο γυμνάσιο. Ανήκει στην ανώτερη δευτεροβάθμια εκπαίδευση (ISCED 3).

Β) Τις τεχνικές-επαγγελματικές σχολές. Σ' αυτές εγγράφονται χωρίς εξετάσεις όσοι ολοκληρώνουν τη φοίτησή τους στο γυμνάσιο. Τα χρόνια φοίτησής τους είναι 2 για τις ημερήσιες και 3 για τις εσπερινές. Ανήκουν στην ανώτερη δευτεροβάθμια εκπαίδευση (ISCED 3)

Γ) Τα εκκλησιαστικά φροντιστήρια. Σ' αυτά εγγράφονται, μετά από εξετάσεις, μαθητές που έχουν απολυτήριο δημοτικού και έχουν συμπληρώσει το 16 έτος της ηλικίας τους. Τα φροντιστήρια αυτά διαρκούν 3 χρόνια. Ανήκουν στην κατώτερη δευτεροβάθμια εκπαίδευση (ISCED 2).

Δ) Τα εκκλησιαστικά γυμνάσια.. Αυτά διαρκούν 3 έτη και στην Α τάξη τους μπορούν να εγγραφούν όσοι μαθητές έχουν απολυτήριο δημοτικού. Ανήκουν στην κατώτερη δευτεροβάθμια εκπαίδευση (ISCED 2).

Ε) Τα εκκλησιαστικά λύκεια. Είναι τετραετή και στην Α τάξη εγγράφονται χωρίς εξετάσεις όσοι έχουν ολοκληρώσει την φοίτησή τους στα εκκλησιαστικά φροντιστήρια, στα γυμνάσια γενικής κατεύθυνσης και στα εκκλησιαστικά γυμνάσια. Κατατάσσονται στην ανώτερη δευτεροβάθμια εκπαίδευση (ISCED 3)

Στ) Τις αδιαβάθμητες τεχνικές-επαγγελματικές σχολές. Εποπτεύονται από τον Οργανισμό Απασχόλησης Εργατικού Δυναμικού. Διαρκούν 2 χρόνια και σ' αυτές εγγράφονται όσοι μαθητές έχουν τελειώσει το γυμνάσιο. Κατατάσσονται στην ανώτερη δευτεροβάθμια εκπαίδευση (ISCED 3)

Η μορφή της δευτεροβάθμιας εκπαίδευσης ωστόσο δεν ήταν έτσι από την αρχή των ετών που εξετάζουμε. Από τα έτη 1970/71 έως 1976/76 η δευτεροβάθμια εκπαίδευση αποτελείτο από την μέση γενική εκπαίδευση και την τεχνική επαγγελματική εκπαίδευση.

Στην μέση γενική εκπαίδευση ανήκαν τα γυμνάσια εξαετή η επταετή ανάλογα αν ήταν ημερήσια ή εσπερινά και οι μαθητές χρειάζονταν εισιτήριες εξετάσεις για να εγγραφούν. Αυτά ακολουθούσαν δύο τριετής κύκλους σπουδών. Τον κατώτερο ο οποίος και κατατάσσεται στην κατηγορία ISCED 3 και είναι οι 3 πρώτες τάξεις και στον ανώτερο, ο οποίος είναι οι 3 τελευταίες τάξεις του ημερήσιου ή οι τέσσερις του εσπερινού. Οι μαθητές αυτού του κύκλου κατατάσσονται στην κατηγορία ISCED 4.

Ο διαχωρισμός γυμνασίου-λυκείου έγινε το σχολικό έτος 1976/77 και η φοίτησή του ήταν προαιρετική μέχρι το έτος 1980/81, οπότε και γίνεται υποχρεωτική. Οι εισιτήριες εξετάσεις για το γυμνάσιο καταργήθηκαν το σχολικό έτος 1977/78 αλλά καθιερώθηκαν από την ίδια χρονιά για το Λύκειο και τελικά καταργήθηκαν και αυτές το σχολικό έτος 1981/82.

Στην τεχνική-επαγγελματική εκπαίδευση ανήκαν οι σχολές στις οποίες οι μαθητές εξειδικεύονταν σε διάφορες ειδικότητες. Η εγγραφή γινόταν χωρίς εξετάσεις.. Ανάλογα με τους κύκλους σπουδών κατατάσσονται στις βαθμίδες ISCED 3 ή ISCED 4.



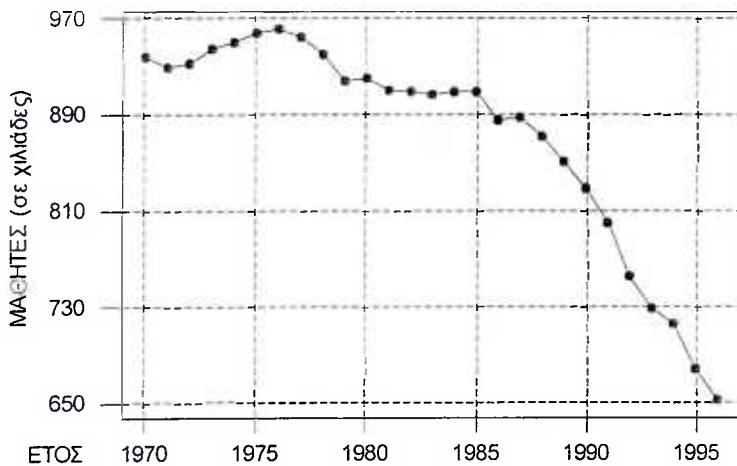
ΚΕΦΑΛΑΙΟ 3 :Η ΠΡΩΤΟΒΑΘΜΙΑ ΕΚΠΑΙΔΕΥΣΗ

3.1 ΓΕΝΙΚΗ ΕΙΚΟΝΑ ΤΟΥ ΑΡΙΘΜΟΥ ΜΑΘΗΤΩΝ ΣΤΟ ΔΗΜΟΤΙΚΟ ΣΧΟΛΕΙΟ

Ο πληθυσμός των μαθητών του δημοτικού σχολείου εχει αρχίσει να μειώνεται αισθητά από τα τέλη της δεκαετίας 70, όπως φαίνεται από το Γραφημα 3.1. Μετά το 1976, οπότε είχαμε την υψηλότερη τιμή των ετών που μελετάμε, η πτώση είναι συνεχής και ραγδαία και προφανώς συνεχίζεται μέχρι τις μέρες μας.

Το δημοτικό σχολείο αντικατοπτρίζει απόλυτα το δημογραφικό πρόβλημα της χώρας και αυτό γιατί η φοίτηση είναι υποχρεωτική από το σύνταγμα και ισχύει από την αρχή των ετών που εξετάζουμε. Έτσι, η μείωση των μαθητών δεν μπορεί να αποδοθεί σε λόγους εγκατάλειψης ή μη παρακολούθησης.

ΜΑΘΗΤΕΣ ΔΗΜΟΤΙΚΟΥ ΓΙΑ ΤΑ ΕΤΗ 1970/71-1996/97



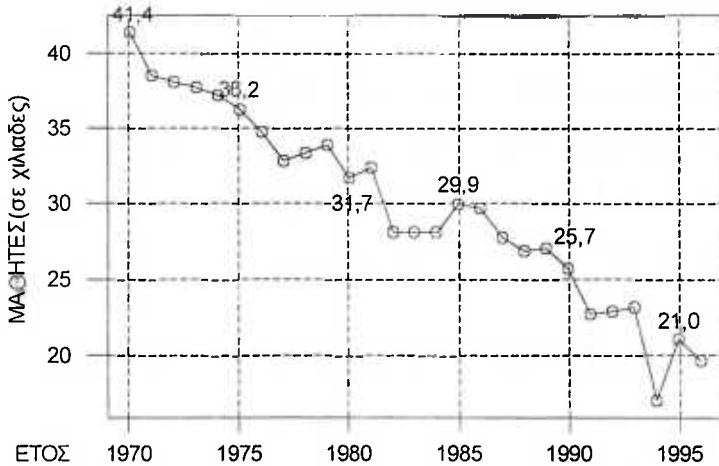
ΠΗΓΗ: Eurostat

ΓΡΑΦΗΜΑ 3.1: Αριθμός μαθητών Δημοτικού από το 1970/71 – 1996/97

Τα Γραφήματα 3.2 και 3.3 στην επόμενη σελίδα φανερώνουν ότι υπάρχει διαφορά στους αριθμούς των εγγεγραμμένων μαθητών των δύο φύλων στα δημοτικά σχολεία, η οποία μειώνεται με σταθερό ρυθμό από το 1970. Η διαφορά είναι αρκετά μεγάλη για να αποδοθεί μόνο στις γεννήσεις περισσότερων αγοριών (105 αγόρια για κάθε 100 κορίτσια) και πρέπει να έχει σχέση με την κοινωνική προκατάληψη σχετικά με την χρησιμότητα της εκπαίδευσης των κοριτσιών που κυριαρχούσε στην ελληνική

κοινωνία και αποβαλλόταν με την αστικοποίηση του πληθυσμού. Την πτωτική όμως τάση της διαφοράς των εγγεγραμμένων μαθητών των δύο φύλων ακολουθεί και μια φθίνουσα πορεία του αριθμού των εγγεγραμμένων μαθητών, κοριτσιών και αγοριών σαν απόρροια της μείωσης γεννήσεων.

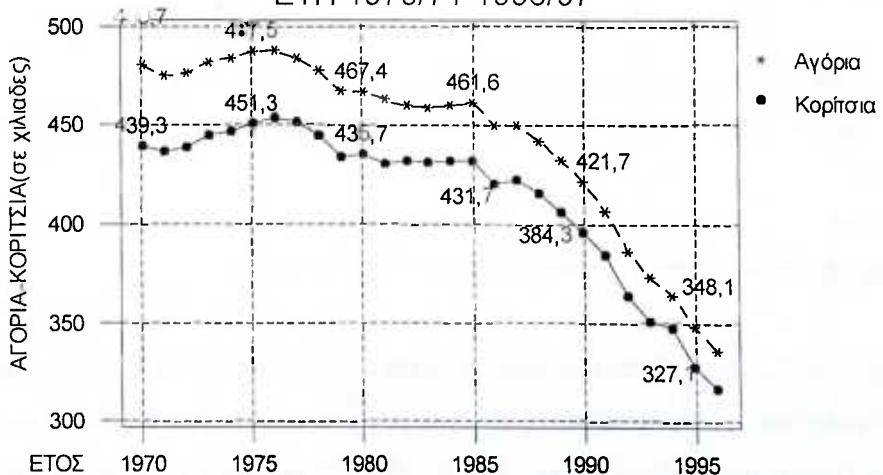
ΔΙΑΧΡΟΝΙΚΗ ΕΞΕΛΙΞΗ ΔΙΑΦΟΡΑΣ ΑΓΟΡΙΩΝ ΚΟΡΙΤΣΙΩΝ ΣΤΟ ΔΗΜΟΤΙΚΟ ΣΧΟΛΕΙΟ



ΠΗΓΗ: Eurostat

ΓΡΑΦΗΜΑ 3.2: Διαφορά αριθμού αγοριών-κοριτσιών του Δημοτικού σχολείου

ΑΡΙΘΜΟΙ ΑΓΟΡΙΩΝ ΚΑΙ ΚΟΡΙΤΣΙΩΝ ΠΟΥ ΕΓΓΡΑΦΟΝΤΑΙ ΣΤΟ ΔΗΜΟΤΙΚΟ ΣΧΟΛΕΙΟ ΕΤΗ 1970/71-1996/97



ΠΗΓΗ: Eurostat

ΓΡΑΦΗΜΑ 3.3: Σύγκριση αριθμού αγοριών - κοριτσιών στο δημοτικό

Το Γράφημα 3.2 δείχνει ότι ενώ το 1970 είχαμε 41,4 χιλιάδες περισσότερα αγόρια το 1996 είχαμε μόλις 19,6 χιλιάδες περισσότερα. Η πτωτική αυτή πορεία φαίνεται να ακολουθεί μια ευθεία γραμμή την οποία θα προσπαθήσουμε να υπολογίσουμε με την μέθοδο των ελαχίστων τετραγώνων.

3.2 ΤΑΣΗ ΔΙΑΦΟΡΑΣ ΜΑΘΗΤΩΝ-ΜΑΘΗΤΡΙΩΝ

Πριν προχωρήσουμε στον υπολογισμό της τάσης των δεδομένων μας θα εξετάσουμε αν αυτά είναι προϊόν μιας τυχαίας διαδικασίας ή παρουσιάζουν πράγματι κάποια σχέση που μπορεί να μελετηθεί. Ο έλεγχος αυτός λέγεται έλεγχος τυχαιότητας και εξετάζει τις παρακάτω υποθέσεις:

H_0 : Τα δεδομένα τυχαία

H_1 : Τα δεδομένα μη τυχαία

Ο έλεγχος αυτός λέγεται έλεγχος Ροών και εξετάζει αν υπάρχουν συνεχόμενες εμφανίσεις (Ροές) παρατηρήσεων πάνω και κάτω από την διάμεσο τιμή, που σημαίνει έλλειψη τυχαιότητας ή αν οι παρατηρήσεις συγκεντρώνονται γύρω από τον μέσο με συνεχόμενα σκαμπανεβάσματα που σημαίνει τυχαιότητα. Το στατιστικό πακέτο Minitab έδωσε τα παρακάτω αποτελέσματα.

Median of Diff

Median of Diff = 29,600

Runs Test: Diff

```
Diff
K = 29,6000

The observed number of runs = 4
The expected number of runs = 14,4815
13 Observations above K 14 below
The test is significant at 0,0000
```

Η p-value=0,0001 απορρίπτει την μηδενική υπόθεση και κατά συνέπεια εξασφαλίζει την μη τυχαιότητα της χρονοσειράς προτρέποντας μας να συνεχίσουμε την ανάλυση μας απρόσκοπτα.

Εξετάζοντας την πρωτοβάθμια τάση για την Διαφορά των μαθητών του Δημοτικού , το στατιστικό πακέτο Minitab μας έδωσε τα παρακάτω αποτελέσματα

Regression Analysis: Diff versus Time

The regression equation is
 $\text{Diff} = 40,7 - 0,778 \text{ Time}$

Predictor	Coef	SE Coef	T	P
Constant	40,7162	0,5757	70,72	0,000
Time	-0,77814	0,03594	-21,65	0,000

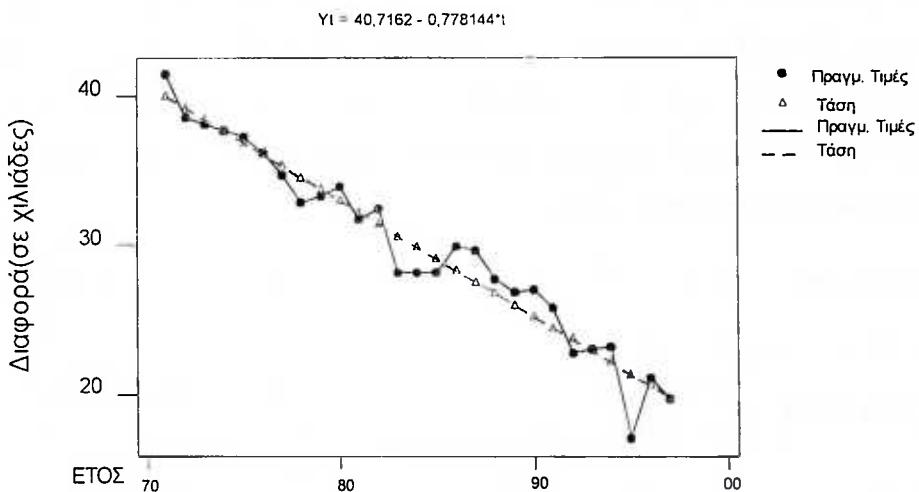
S = 1,454 R-Sq = 94,9% R-Sq(adj) = 94,7%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	991,82	991,82	468,86	0,000
Residual Error	25	52,88	2,12		
Total	26	1044,71			

Durbin-Watson statistic = 1,81

Ανάλυση Τάσης για την Διαφορά



ΓΡΑΦΗΜΑ 3.4: Τάση της διαφοράς αγοριών- κοριτσιών για τα έτη 1970/71 – 1996/97

Η εξίσωση μας δείχνει ότι η διαφορά μεταξύ αγοριών κοριτσιών μειώνεται κάθε χρόνο κατά 778 περίπου μαθητές ξεκινώντας από την αρχική διαφορά των 41 χιλιάδων. Με τον ρυθμό αυτό υπολογίζεται ότι η διαφορά αυτή θα εξαφανιστεί το 2022, δηλαδή 52 χρόνια μετά το 1970. Αυτό βέβαια δεν πρόκειται να πραγματοποιηθεί ποτέ γιατί στις ηλικίες μέχρι την εφηβεία τα αγόρια είναι πάντα περισσότερα από τα κορίτσια, αφού γενικά στις γεννήσεις αντιστοιχούν 105 αγόρια για κάθε 100 κορίτσια. Η εξισορρόπηση επέρχεται λόγω της υψηλότερης θνησιμότητας των ανδρών ειδικά από την περίοδο της εφηβείας και μετά. (Αναστασία

Κωστάκη, Δημομετρία, 1998). Κάποια χρονική στιγμή πιθανώς η διαφορά να σταθεροποιηθεί σε κάποια τιμή.

Όσον αφορά το στατιστικό κομμάτι της ανάλυσης, παρατηρούμε ότι ο συντελεστής βι που μας δίνει την εκτίμηση -0,778 είναι στατιστικά σημαντικός αφού η p-value=0 απορρίπτει την μηδενική υπόθεση. Στα ίδια συμπεράσματα καταλήγουμε και από το F-test. Που επιβεβαιώνει την σημαντικότητα της παλινδρόμησης.

Οι συντελεστές R^2 και R^2_{adj} δίνουν τιμές 94,4% και 94,1% που σημαίνει ότι περίπου 94% της μεταβλητότητας των δεδομένων εξηγείται από την γραμμή που υπολογίσαμε και κατά συνέπεια η ακρίβεια της μπορεί να χαρακτηριστεί υψηλή.

Σκοπός μας ωστόσο δεν είναι να εξετάσουμε την διαφορά μεταξύ αγοριών κοριτσιών αλλά την πτώση του αριθμού των μαθητών γενικά, γι' αυτό και στη συνέχεια θα ασχοληθούμε αποκλειστικά με αυτό.

3.3 ΕΚΤΙΜΗΣΗ ΤΑΣΗΣ ΜΑΘΗΤΩΝ ΔΗΜΟΤΙΚΟΥ

Από τα Γραφήματα 3.1 και 3.3 παρατηρούμε ότι το σχήμα της χρονολογικής σειράς είναι πανομοιότυπο για τα αγόρια και τα κορίτσια και για τον συνολικό αριθμό. Για τον λόγο αυτό αντί να ασχοληθούμε ξεχωριστά με την τάση του αριθμού αγοριών και κοριτσιών στο δημοτικό σχολείο θα εξετάσουμε την τάση στο σύνολο αφού αυτή θα αντικατοπτρίζει και την τάση για καθένα ξεχωριστά.

Η εκτίμηση της τάσης θα γίνει στις τιμές των 23 πρώτων ετών έτσι ώστε να χρησιμοποιηθούν τα επόμενα τρία χρόνια για έλεγχο της ικανότητας του μοντέλου για προβλέψεις.

Πριν από την κυρίως ανάλυση, ωστόσο θα πραγματοποιήσουμε τον έλεγχο τυχαιότητας και σε αυτή την χρονοσειρά. Τα αποτελέσματα είναι τα παρακάτω:

Median of Primary Total

Median of Primary Total = 892,80
MTB > Runs 892,8 'Primary Total'.

Runs Test: Primary Total

Primary

K = 892,8000

The observed number of runs = 4
The expected number of runs = 14,4815
13 Observations above K 14 below
The test is significant at 0,0000

Η τάση που “ταιριάζει” καλύτερα σ’ αυτά τα δεδομένα είναι της μορφής:

$$Y_t = 901,5 + 8,49t - 0,62t^2$$

Οι εκτιμήσεις του στατιστικού πακέτου για την τάση αυτή είναι:

Regression Analysis: Primary versus Time; Time**2

The regression equation is
Primary = 901 + 8,49 Time - 0,622 Time**2

Predictor	Coef	SE Coef	T	P
Constant	901,498	9,510	94,79	0,000
Time	8,491	1,753	4,84	0,000
Time**2	-0,62171	0,06807	-9,13	0,000

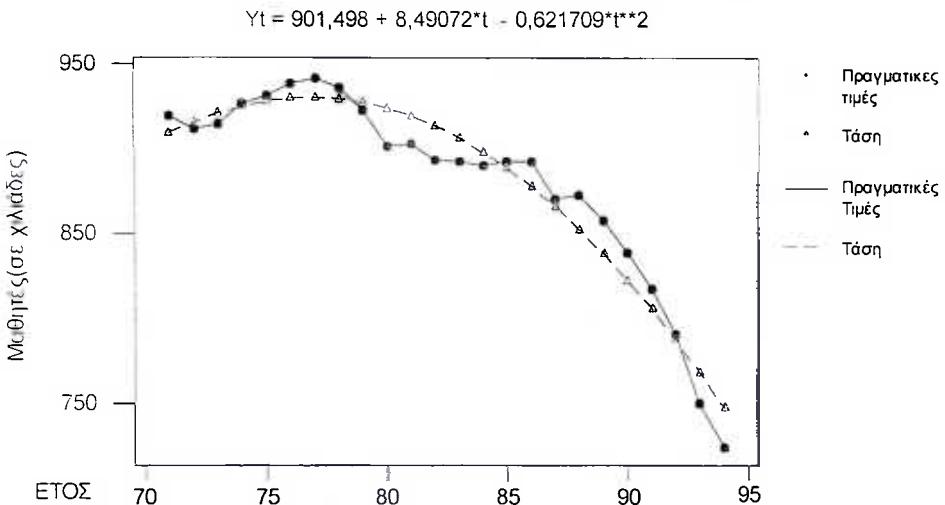
S = 14,25 R-Sq = 94,6% R-Sq(adj) = 94,0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	74141	37070	182,45	0,000
Residual Error	21	4267	203		
Total	23	78407			

Durbin-Watson statistic = 0,50

Ανάλυση Τάσης για το Σύνολο των Μαθητών Δημοτικού



ΓΡΑΦΗΜΑ 3.5: Τάση του αριθμού των μαθητών δημοτικού 1970/71-1993/94

Οι συντελεστές β_1 και β_2 των μεταβλητών t και t^2 είναι στατιστικά σημαντικοί, όπως επίσης και η σταθερά β_0 αφού η μηδενική υπόθεση που ελέγχει αν οι συντελεστές αυτοί είναι μηδενικοί απορρίπτεται σε επίπεδο σημαντικότητας 5%. Επίσης απορρίπτεται και η μη στατιστική σημαντικότητα της γραμμής παλινδρόμησης με τον έλεγχο F. Επίσης το ποσοστό της μεταβλητότητας των δεδομένων που εξηγεί η γραμμή αγγίζει το 94,6% και μπορεί να θεωρηθεί πολύ ικανοποιητική.

Θεωρώντας ότι η σειρά μας αποτελείται μόνο από τάση και τυχαία συνιστώσα μπορούμε να χρησιμοποιήσουμε την τάση αυτή για να δώσουμε πρόβλεψη για τα απόμενα τρία χρόνια. Η χρονική περίοδος, οι προβλέψεις και οι πραγματικές τιμές είναι:

ΠΙΝΑΚΑΣ ΠΡΑΓΜΑΤΙΚΩΝ ΚΑΙ ΠΡΟΒΛΕΠΟΜΕΝΩΝ ΤΙΜΩΝ

Πραγματικές Τιμές	Προβλέψεις
710,8	725,198
675,3	701,982
652,0	677,522

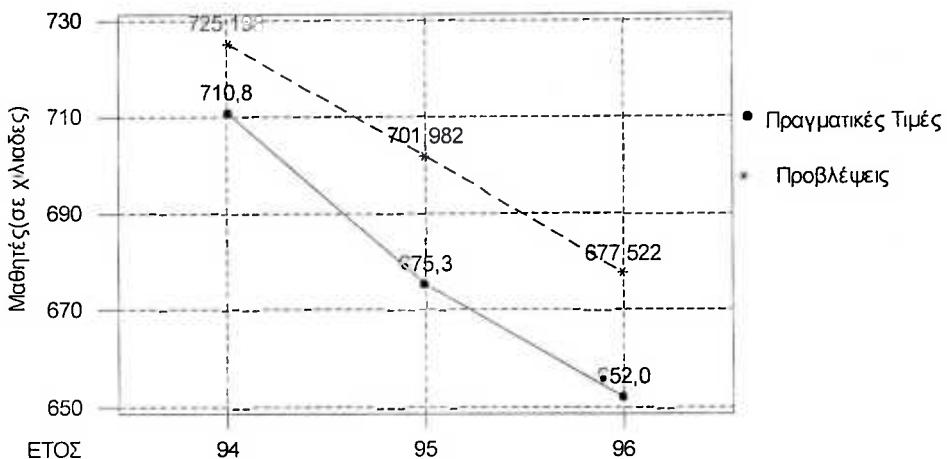
Πίνακας 3.1: Πραγματικές τιμές και προβλέψεις του μοντέλου της παλινδρόμησης

ΠΙΝΑΚΑΣ ΜΕΤΡΩΝ ΑΚΡΙΒΕΙΑΣ

MAPE	3,297
MAD	22,201
MSE	523,532

Πίνακας 3.2: Μέτρα ακρίβειας του υποδείγματος της παλινδρόμησης

Πρόβλεψη και Πραγματικές Τιμές



ΓΡΑΦΗΜΑ 3.6: Προβλέψεις και πραγματικές τιμές για τα έτη 93/94 – 96/97

Παρατηρούμε ότι οι προβλέψεις υποεκτιμούν τις μελλοντικές τιμές αφού η πτώση του αριθμού των μαθητών φαίνεται να είναι πιο ομαλή. Επιπλέον η αξιοπιστία των προβλέψεων είναι αμφισβητούμενη αφού η τιμή Durbin-Watson υποδεικνύει συσχέτιση των καταλοίπων. Η διαφορά μεταξύ των προβλεπόμενων και των πραγματικών τιμών αντικατοπτρίζονται στο Γράφημα 3.6.

3.4 ΕΚΘΕΤΙΚΗ ΕΞΟΜΑΛΥΝΣΗ

Όπως αναφέρθηκε και στο πρώτο κεφάλαιο η μορφή των δεδομένων είναι τέτοια ώστε μας επιτρέπει να χρησιμοποιήσουμε την μέθοδο της εκθετικής εξομάλυνσης του Holt η οποία εφαρμόζεται καλύτερα σε χρονολογικά δεδομένα που παρουσιάζουν μόνο τάση.

Με την μέθοδο της ελαχιστοποίησης του MSE για την προσαρμογή του μοντέλου, υπολογίσαμε τα α και γ αντίστοιχα έπειτα από μέθοδο διαδοχικών δοκίμων. Οι τιμές που μας έδωσαν ελάχιστη τιμή MSE ήταν $\alpha=0,797$ και $\gamma=0,812$ αντίστοιχα. Δίνεται δηλαδή μεγαλύτερη βαρύτητα στις πρόσφατες παρατηρήσεις όσον αφορά το επίπεδο των δεδομένων αλλά και όσον αφορά την τάση.

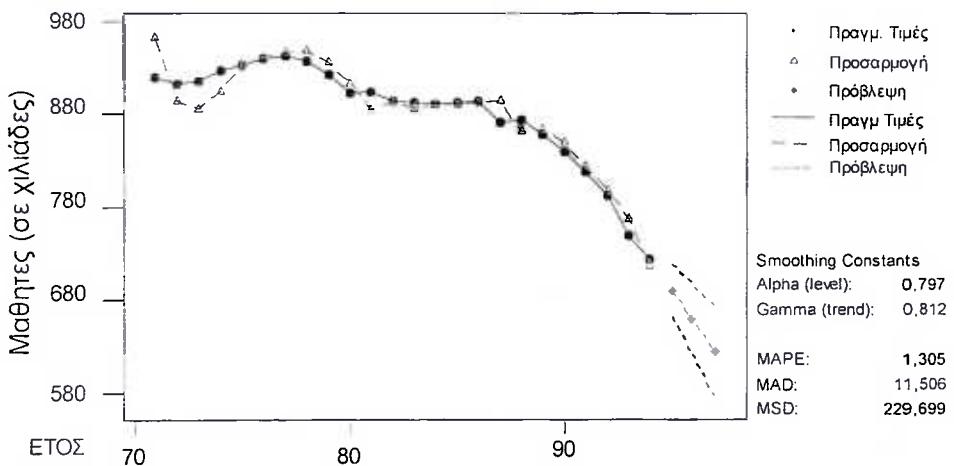
Double Exponential Smoothing

Data Primary Total
 Length 24,0000
 NMissing 0

Smoothing Constants
 Alpha (level): 0,797
 Gamma (trend): 0,812

Accuracy Measures
 MAPE: 1,305
 MAD: 11,506
 MSD: 229,699

Διπλή Εκθετική Εξομάλυνση για Δημοτικό



ΓΡΑΦΗΜΑ 3.7: Εκθετική Εξομάλυνση του Holt. Προσαρμογή και Πρόβλεψη

Οι προβλέψεις για τις τιμές των επόμενων τριών ετών είναι εμφανώς χειρότερες από τα μοντέλα της τάσης και εδώ φαίνεται να υπάρχει υποεκτίμηση των τιμών. Τα αποτελέσματα φαίνονται καλύτερα στη συνέχεια όπου παρουσιάζονται οι Πίνακες 3.3 και το Γράφημα 3.10.

ΠΙΝΑΚΑΣ ΠΡΑΓΜΑΤΙΚΩΝ ΚΑΙ ΠΡΟΒΛΕΠΟΜΕΝΩΝ ΤΙΜΩΝ

Πραγματικές Τιμές	Προβλέψεις
710,8	690,170
675,3	658,210
652,0	626,249

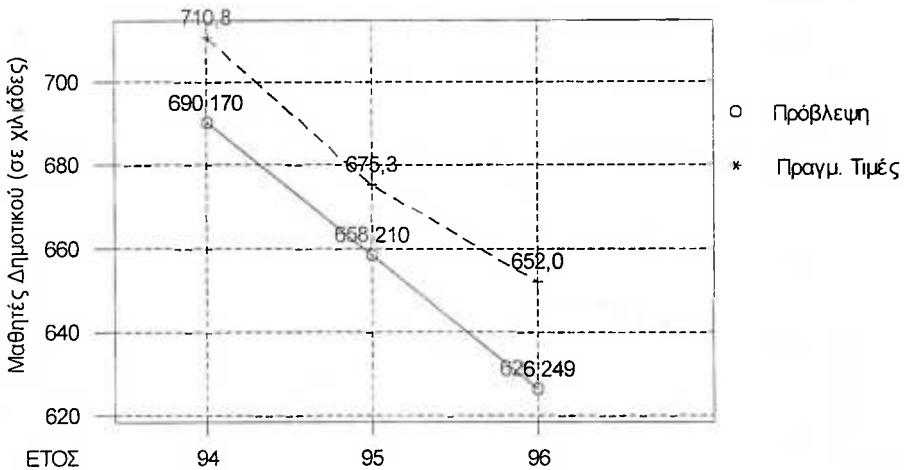
Πίνακας 3.3: Πραγματικές τιμές και προβλέψεις Διπλής Εκθετικής Εξομάλυνσης

ΠΙΝΑΚΑΣ ΜΕΤΡΩΝ ΑΚΡΙΒΕΙΑΣ

MAPE	3,23250
MAD	21,1570
MSE	460,259

Πίνακας 3.4: Μέτρα ακριβείας Διπλής Εκθετικής Εξομάλυνσης

Πρόβλεψη και Πραγματικές Τιμές στην Εκθετική Εξομάλυνση



ΓΡΑΦΗΜΑ 3.8: Σύγκριση Πραγματικών Τιμών με τις Προβλέψεις της Εξομάλυνσης Holt

3.5 ΥΠΟΔΕΙΓΜΑΤΑ ARIMA

Η προσαρμογή και ο έλεγχος των υποδειγμάτων αυτών θα γίνει με το στατιστικό πακέτο Eviews γιατί μας δίνει περισσότερες πληροφορίες για την προσαρμογή από το Minitab.

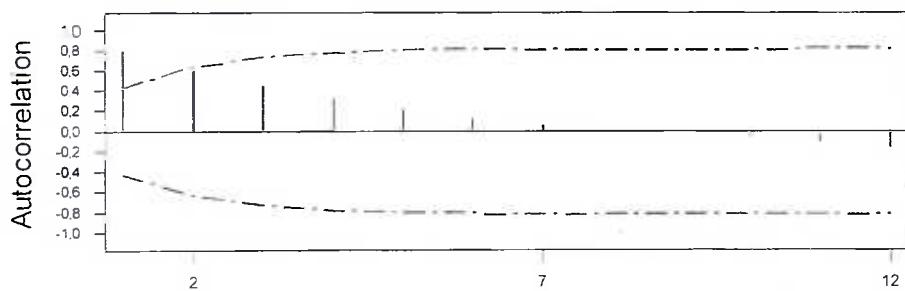
Αρχικά θα πρέπει να ελέγξουμε αν η σειρά είναι στάσιμη. Είναι φανερό από τα Γραφήματα η ύπαρξη τάσης στην χρονολογική σειρά. Κάτι τέτοιο ωστόσο εξετάζεται και με τα Γραφήματα αυτοσυσχετίσεων τα correlograms. Γενικά όταν οι συντελεστές αυτοσυσχέτισης ή μερικής αυτοσυσχέτισης μειώνονται με αργό ρυθμό είναι ένδειξη ότι η σειρά δεν είναι στάσιμη. Για την σειρά που εξετάζουμε έχουμε τα παρακάτω Γραφήματα 3.9 και 3.10.

Από το πρώτο γράφημα παρατηρούμε ότι οι αυτοσυσχετίσεις φθίνουν αργά σε σχέση με τον μικρό αριθμό παρατηρήσεων που έχουμε. Έτσι είναι απαραίτητο να πάρουμε

τουλάχιστον τις πρώτες διαφορές. Το Γράφημα 3.9 μερικής αυτοσυσχέτισης έχει μόνο μία τιμή υψηλή ενώ οι υπόλοιπες είναι στατιστικά ίσες με το μηδέν.

Οι πρώτες διαφορές της χρονολογικής σειράς δεν είναι ικανές να εξαφανίσουν της τάση. Αυτό μπορούμε να το δούμε στο Γράφημα 3.11 όπου η σειρά των πρώτων διαφορών εμφανίζει μια πτωτική τάση. Στο Γράφημα 3.12 που δείχνει τις αυτοσυσχετίσεις των πρώτων διαφορών θα μπορούσε να ισχυριστεί κανείς ότι φθίνουν εκθετικά προς το μηδέν και κατά συνέπεια έχει επιτευχθεί η στασιμότητα. Επιπλέον στο Γράφημα 3.13 των μερικών αυτοσυχετίσεων μόνο η πρώτη τιμή είναι στατιστικά σημαντική, γεγονός που μας οδηγεί σε υπόδειγμα AR(1) I(1)

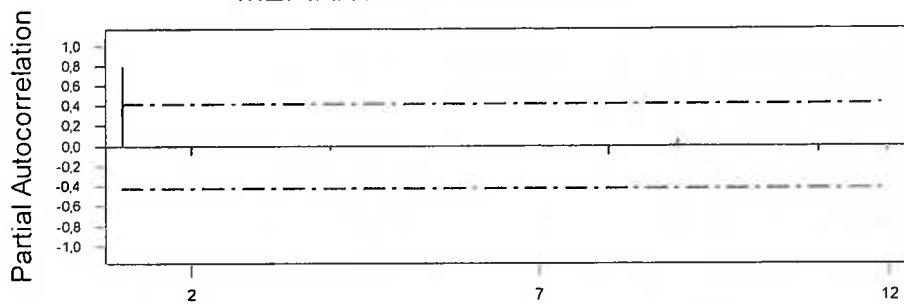
ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΓΙΑ ΤΟΥΣ ΜΑΘΗΤΕΣ ΔΗΜΟΤΙΚΟΥ



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0.80	3.90	17.22	8	-0.01	-0.02	38.82
2	0.60	1.95	27.44	9	-0.04	-0.09	38.88
3	0.45	1.27	33.44	10	-0.07	-0.16	39.07
4	0.32	0.85	36.60	11	-0.11	-0.28	39.64
5	0.21	0.55	38.06	12	-0.16	-0.40	40.94
6	0.13	0.33	38.64				
7	0.07	0.18	38.82				

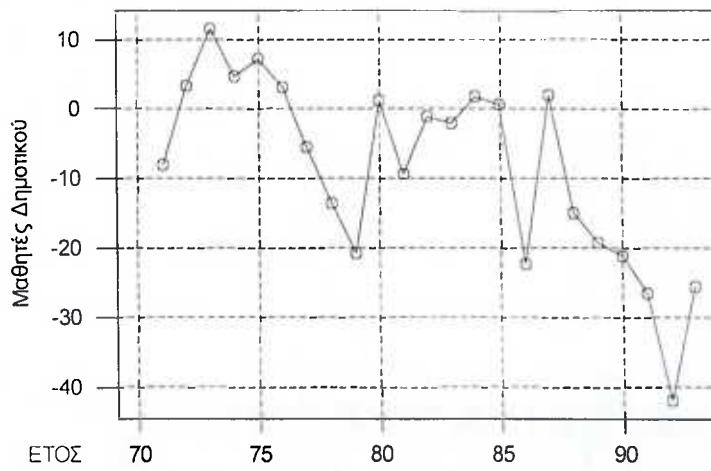
ΓΡΑΦΗΜΑ 3.9: Συντελεστές Αυτοσυσχέτισης για μαθητές Δημοτικού

ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ



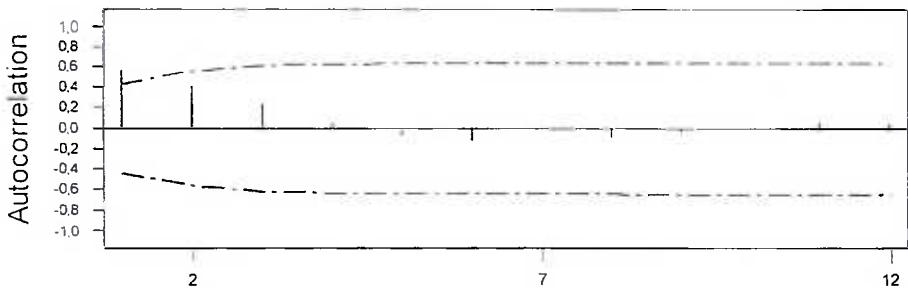
ΓΡΑΦΗΜΑ 3.10: Μερικές αυτοσυσχετίσεις για τους Μαθητές Δημοτικού

Πρώτες Διαφορές



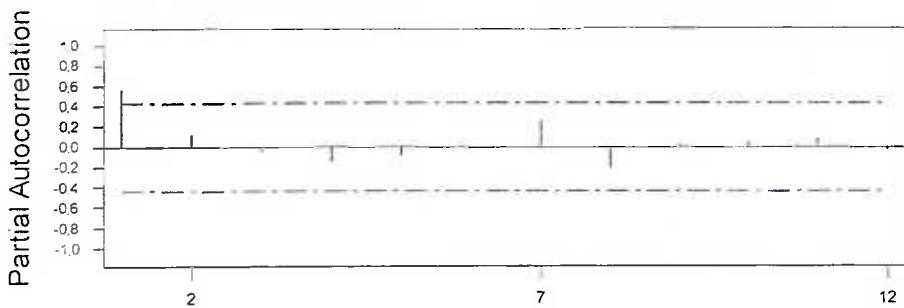
ΓΡΑΦΗΜΑ 3.11: Γιρώτες Διαφορές της χρονολογικής σειράς Μαθητών Δημοτικού

ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΓΙΑ ΤΙΣ ΠΡΩΤΕΣ ΔΙΑΦΟΡΕΣ



ΓΡΑΦΗΜΑ 3.12: Αυτοσυσχετίσεις των πρώτων διαφορών για τους μαθητές Δημοτικού

ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΓΙΑ ΤΙΣ ΠΡΩΤΕΣ ΔΙΑΦΟΡΕΣ



ΓΡΑΦΗΜΑ 3.13: Μερικές αυτοσυσχετίσεις των πρώτων διαφορών των μαθητών Δημοτικού

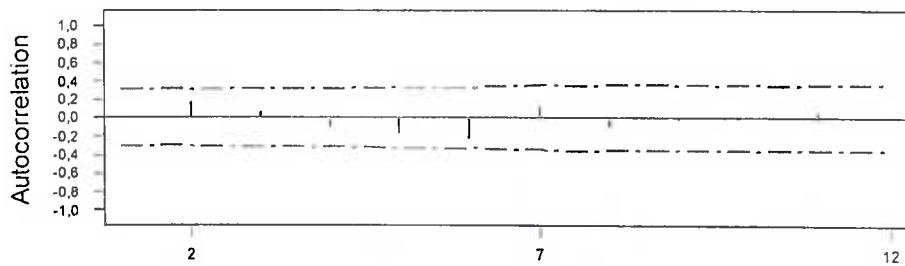
Ο υπολογισμός και η επεξεργασία των δεδομένων για την περίπτωση αυτή θα γίνει με το στατιστικό πρόγραμμα E-views γιατί μας δίνει περισσότερα στοιχεία για να συγκρίνουμε τα υποδείγματα που προκύπτουν. Έτσι για το πρώτο μοντέλο έχουμε:

Dependent Variable: DPRM
 Method: Least Squares
 Date: 03/07/04 Time: 08:42
 Sample(adjusted): 1972 1993
 Included observations: 22 after adjusting endpoints
 Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-9.887606	6.544984	-1.510715	0.1465
AR(1)	0.626287	0.185689	3.372767	0.0030
R-squared	0.362561	Mean dependent var	-8.554545	
Adjusted R-squared	0.330689	S.D. dependent var	13.83882	
S.E. of regression	11.32174	Akaike info criterion	7.777834	
Sum squared resid	2563.636	Schwarz criterion	7.877020	
Log likelihood	-83.55618	F-statistic	11.37556	
Durbin-Watson stat	2.240468	Prob(F-statistic)	0.003025	
Inverted AR Roots	.63			

Τα αποτελέσματα δίνουν: Τυπικό σφάλμα της Παλινδρόμησης (SE of regression) 11,32, Άθροισμα Τετραγώνου Καταλοίπων(Sum Square resid)=2563.636. Τέλος τα πιο σημαντικά κριτήρια για την προσαρμογή το Akaike και το Schwarz με τιμές 7,78 και 7,88 αντίστοιχα. Η τιμή Durbin-Watson είναι ένδειξη ότι τα κατάλοιπα είναι αυτοσυσχέτιστα. Κάτι τέτοιο φαίνεται και από τα Γραφήματα συσχέτισης 3.14 και αυτοσυσχέτισης 3.15 των καταλοίπων.

ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ

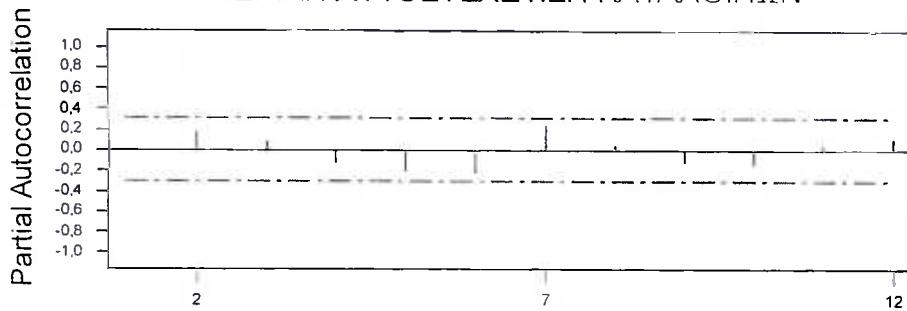


Lag	Corr	T	LBO	Lag	Corr	T	LBO
1	0.00	0.03	0.00	8	-0.10	-0.56	8.21
2	0.18	1.20	1.57	9	-0.03	-0.16	8.25
3	0.08	0.50	1.87	10	0.01	0.07	8.26
4	-0.10	-0.63	2.37	11	0.06	0.32	8.45
5	-0.17	-1.08	3.89	12	0.01	0.04	8.46
6	-0.24	-1.44	6.81				
7	0.13	0.74	7.69				

ΓΡΑΦΗΜΑ 3.14: Αυτοσυσχέτιση των καταλοίπων υποδείγματος ARIMA (1,1,0)



ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ



Lag	PAC	T	Lag	PAC	T
1	0,00	0,03	8	0,05	0,32
2	0,18	1,20	9	-0,12	-0,79
3	0,08	0,52	10	-0,15	-0,96
4	-0,14	-0,91	11	0,05	0,32
5	-0,22	-1,41	12	0,11	0,71
6	-0,23	-1,48			
7	0,24	1,59			

ΓΡΑΦΗΜΑ 3.15: Μερική Αυτοσυσχέτιση καταλοίπων μοντέλου ARIMA (1,1,0)

Η προβλεπτική ικανότητα του μοντέλου είναι καλύτερη από τα προηγούμενα. Πιο συγκεκριμένα δίνει τα παρακάτω αποτελέσματα:

ΠΙΝΑΚΑΣ ΠΡΑΓΜΑΤΙΚΩΝ ΚΑΙ ΠΡΟΒΛΕΠΟΜΕΝΩΝ ΤΙΜΩΝ

Πραγματικές Τιμές	Προβλέψεις
710,8	703,972
675,3	687,921
652,0	674,174

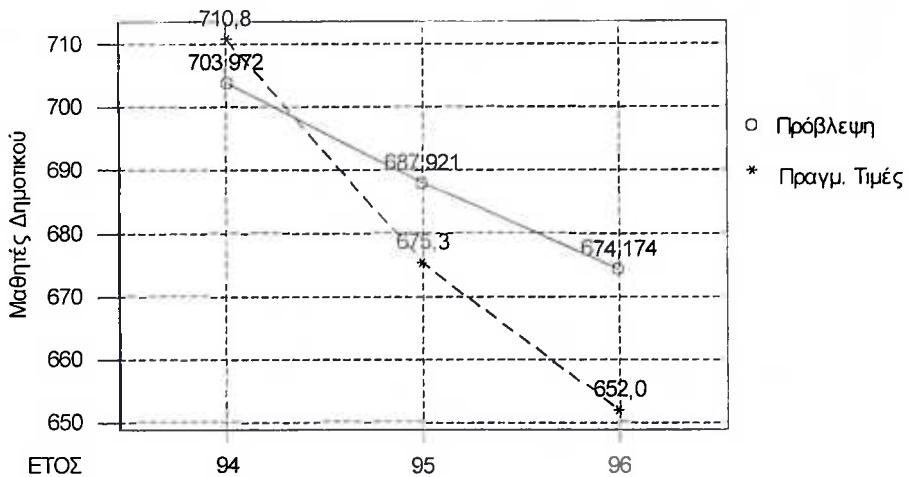
Πίνακας 3.5: Πραγματικές τιμές και προβλέψεις υποδείγματος ARIMA(1,1,0)

ΠΙΝΑΚΑΣ ΜΕΤΡΩΝ ΑΚΡΙΒΕΙΑΣ

<u>MAPE</u>	2,03124
<u>MAD</u>	13,8745
<u>MSE</u>	232,5369

Πίνακας 3.6: Μέτρα ακριβείας ARIMA (1,1,0)

ΠΡΑΓΜΑΤΙΚΕΣ ΤΙΜΕΣ ΚΑΙ ΠΡΟΒΛΕΨΗ ΓΙΑ ΤΑ ΕΤΗ 1994-1996



ΓΡΑΦΗΜΑ 3.16: Πραγματικές τιμές και πρόβλεψη υποδείγματος ARIMA(1,1,0)

Η επιλογή του παραπάνω υποδείγματος μας έδωσε κάποια αποτελέσματα τα οποία όσον αφορά την προβλεπτική τους ικανότητα μπορούν να χαρακτηριστούν ικανοποιητικά ωστόσο δεν ακολουθεί την πτωτική τάση με τον αναμενόμενο ρυθμό. Για τον λόγο αυτό επιλέξαμε το υπόδειγμα AR(1) I(1) MA(1) του οποίου τα αποτελέσματα εμφανίζονται στην επόμενη σελίδα.

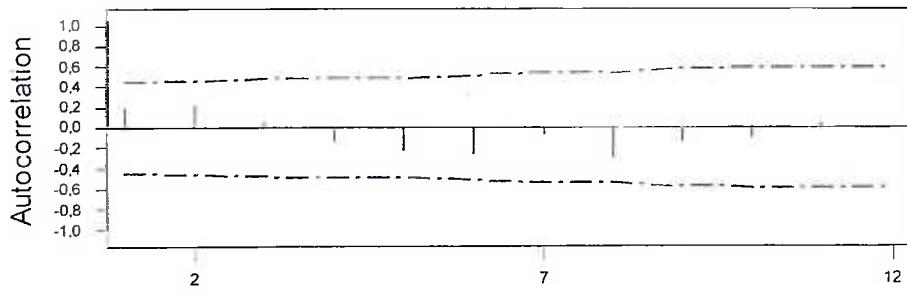
Το υπόδειγμα αυτό φαίνεται ότι έχει καλύτερη προσαρμογή στα δεδομένα από το προηγούμενο. Πιο συγκεκριμένα έχει υψηλότερους συντελεστές προσαρμογής R^2 και R^2_{adj} , οι οποίοι όμως δεν είναι θεωρούνται σαν κατάλληλο κριτήριο για την επιλογή μεταξύ υποδειγμάτων, έχει μικρότερο τυπικό σφάλμα παλινδρόμησης (9,85), μικρότερο άθροισμα τετραγώνων των καταλοίπων (1844,6) και το πιο σημαντικό μικρότερα Akaike και Schwarz με τιμές 7,54 και 7,69 αντίστοιχα.

Τα Γραφήματα 3.17 και 3.18 αυτοσυσχέτισης και μερικής αυτοσυσχέτισης των καταλοίπων δείχνουν ότι είναι πλήρως ασυσχέτιστα και κατά συνέπεια η προσαρμογή είναι καλή και μπορούμε να προχωρήσουμε στις προβλέψεις, οι οποίες είναι και το πιο κρίσιμο σημείο για τον έλεγχο αυτών των μοντέλων.

Dependent Variable: D(PRM)
 Method: Least Squares
 Date: 02/16/04 Time: 01:47
 Sample(adjusted): 1972 1993
 Included observations: 22 after adjusting endpoints
 Convergence achieved after 27 iterations
 Backcast: 1971

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-184.5527	1738.201	-0.106175	0.9166
AR(1)	0.991160	0.086335	11.48037	0.0000
MA(1)	-0.989670	0.000343	-2883.305	0.0000
R-squared	0.541349	Mean dependent var	-8.554545	
Adjusted R-squared	0.493070	S.D. dependent var	13.83882	
S.E. of regression	9.853108	Akaike info criterion	7.539575	
Sum squared resid	1844.591	Schwarz criterion	7.688353	
Log likelihood	-79.93532	F-statistic	11.21292	
Durbin-Watson stat	1.589066	Prob(F-statistic)	0.000608	
Inverted AR Roots	.99			
Inverted MA Roots	.99			

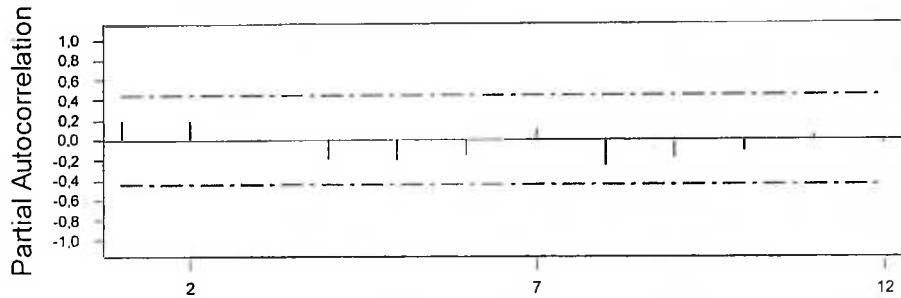
ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0,19	0,88	0,88	8	-0,31	-1,21	11,14
2	0,22	0,98	2,11	9	-0,15	-0,56	12,12
3	0,04	0,19	2,17	10	-0,11	-0,40	12,68
4	-0,15	-0,66	2,84	11	0,03	0,10	12,72
5	-0,24	-1,04	4,70	12	0,04	0,14	12,80
6	-0,27	-1,11	7,18				
7	-0,08	-0,32	7,41				

ΓΡΑΦΗΜΑ 3.17: Αυτοσυσχέτιση καταλοίπων υποδείγματος ARIMA (1,1,1)

ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ



Lag	PAC	T	Lag	PAC	T
1	0.19	0.88	8	-0.27	-1.26
2	0.19	0.88	9	-0.19	-0.90
3	-0.03	-0.12	10	-0.13	-0.59
4	-0.21	-0.99	11	0.05	0.24
5	-0.22	-1.02	12	-0.05	-0.23
6	-0.16	-0.76			
7	0.09	0.44			

ΓΡΑΦΗΜΑ 3.18: Μερική αυτοσυσχέτιση καταλοίπων ARIMA (1,1,1)

Ο επόμενος πίνακας επαληθεύει την καλύτερη προσέγγιση των δεδομένων. Όλες οι τιμές σύγκρισης της προβλεπτικότητας των υποδειγμάτων είναι ελαττωμένες σε σχέση με το προηγούμενο υπόδειγμα. Οι τιμές που δίνει το μοντέλο είναι όμως μικρότερες από τις πραγματικές αλλά δεν αποκλίνουν από τις πραγματικές όσο απέκλιναν οι τιμές που προέβλεπε το υπόδειγμα AR(1) I(1). Η εικόνα αυτών των αποτελεσμάτων αντικατοπτρίζεται καλύτερα στο Γράφημα 3.19.

ΠΙΝΑΚΑΣ ΠΡΑΓΜΑΤΙΚΩΝ ΚΑΙ ΠΡΟΒΛΕΠΟΜΕΝΩΝ ΤΙΜΩΝ

Πραγματικές Τιμές	Προβλέψεις
710,8	696,825
675,3	668,556
652,0	638,906

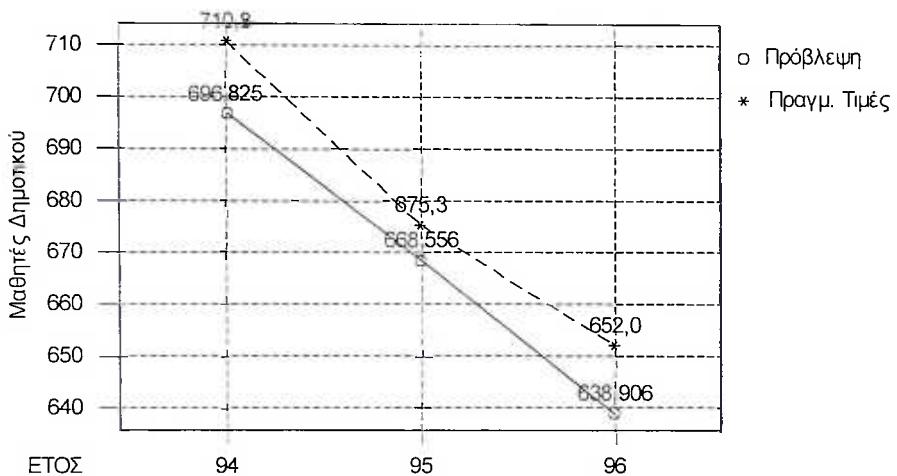
Πίνακας 3.7: Πραγματικές τιμές και προβλέψεις υποδείγματος ARIMA(1,1,0)

ΠΙΝΑΚΑΣ ΜΕΤΡΩΝ ΑΚΡΙΒΕΙΑΣ

MAPE	1,68793
MAD	11,2712
MSE	137,416

Πίνακας 3.8: Μέτρα ακριβείας ARIMA(1,1,0)

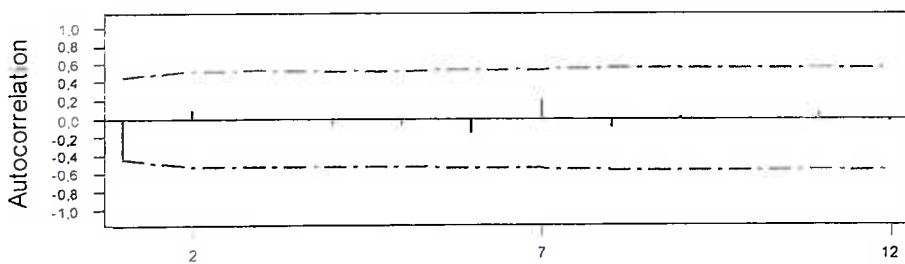
ΠΡΑΓΜΑΤΙΚΕΣ ΤΙΜΕΣ ΚΑΙ ΠΡΟΒΛΕΨΗ ΓΙΑ ΤΑ ΕΤΗ 1994-1996



ΓΡΑΦΗΜΑ 3.19: Πραγματικές τιμές και πρόβλεψη υποδείγματος ARIMA (1,1,1)

Η προσπάθεια για ένα καλύτερο μοντέλο, που θα δώσει και καλύτερες προβλέψεις δεν μπορεί να σταματήσει σ' αυτό το σημείο. Ειδικότερα όταν γνωρίζουμε ότι μπορεί να μην έχουμε επιτύχει απόλυτη στασιμότητα της σειρά και ίσως να χρειάζεται να πάρουμε ακόμα τις δεύτερες διαφορές της αρχικής σειράς.. Όπως θα δούμε στα επόμενα δύο Γραφήματα (3.20, 3.21) της αυτοσυσχέτισης και μερικής αυτοσυσχέτισης των δεύτερων διαφορών, η τάση εξαφανίζεται εντελώς με τις δεύτερες διαφορές και κατά συνέπεια ίσως μπορούν να δώσουν μοντέλα με καλύτερη ικανότητα μελλοντικής πρόβλεψης.

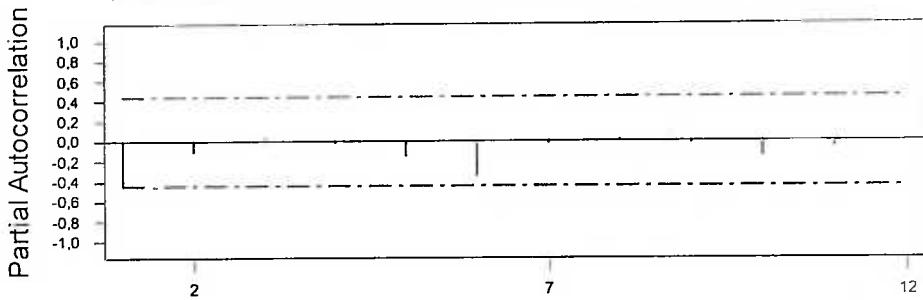
ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΔΕΥΤΕΡΩΝ ΔΙΑΦΟΡΩΝ



Lag	Corr	T	LBO	Lag	Corr	T	LBO
1	-0,44	-2,08	4,96	8	-0,09	-0,35	8,34
2	0,10	0,38	5,20	9	0,01	0,05	8,35
3	0,04	0,15	5,24	10	-0,01	-0,03	8,35
4	-0,08	-0,30	5,40	11	0,08	0,29	8,65
5	-0,08	-0,31	5,59	12	-0,04	-0,13	8,71
6	-0,17	-0,66	6,54				
7	0,20	0,78	8,01				

ΓΡΑΦΗΜΑ 3.20: Αυτοσυσχέτιση δεύτερων διαφορών μαθητών/τριων Δημοτικού

ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΔΕΥΤΕΡΩΝ ΔΙΑΦΟΡΩΝ



Lag	PAC	T	Lag	PAC	T
1	-0,44	-2,08	8	0,02	0,12
2	-0,13	-0,59	9	-0,03	-0,13
3	0,04	0,18	10	-0,16	-0,76
4	-0,04	-0,17	11	-0,07	-0,31
5	-0,16	-0,77	12	-0,01	-0,04
6	-0,36	-1,70			
7	-0,04	-0,18			

ΓΡΑΦΗΜΑ 3.21: Μερική αυτοσυσχέτιση δεύτερων διαφορών μαθητών/τριων δημοτικού

Το υπόδειγμα με τα καλύτερα αποτελέσματα ήταν το AR(0) I(2) MA(1)

Dependent Variable: D(PRM,2)

Method: Least Squares

Date: 03/14/04 Time: 23:39

Sample(adjusted): 1972 1993

Included observations: 22 after adjusting endpoints

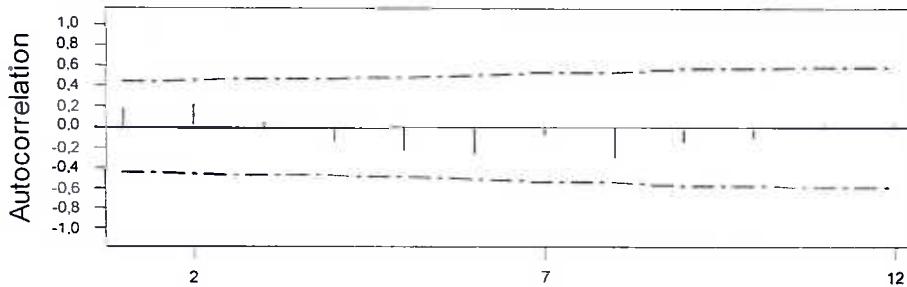
Convergence achieved after 22 iterations

Backcast: 1971

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.591973	0.319192	-4.987515	0.0001
MA(1)	-0.989815	0.000250	-3966.950	0.0000
R-squared	0.413067	Mean dependent var	-0.795455	
Adjusted R-squared	0.383720	S.D. dependent var	12.11616	
S.E. of regression	9.511612	Akaike info criterion	7.429412	
Sum squared resid	1809.415	Schwarz criterion	7.528597	
Log likelihood	-79.72353	F-statistic	14.07542	
Durbin-Watson stat	1.633388	Prob(F-statistic)	0.001256	
Inverted MA Roots	.99			

Το μοντέλο αυτό έχει καλύτερη προσαρμογή από τα αλλά με μικρότερες τιμές σε τυπικό σφάλμα παλινδρόμησης 9,5116, μικρότερο άθροισμα τετραγώνων των καταλοίπων 1809,415 και φυσικά μικρότερα Akaike και Schwarz κριτήρια με τιμές 7,43 και 7,53. Επίσης η τιμή Durbin-Watson και τα γραφήματα αυτοσυσχέτισης και μερικής αυτοσυσχέτισης (Γραφήματα 3.22 και 3.23) των καταλοίπων δείχνουν προσαρμογή του υποδείγματος χωρίς προβλήματα

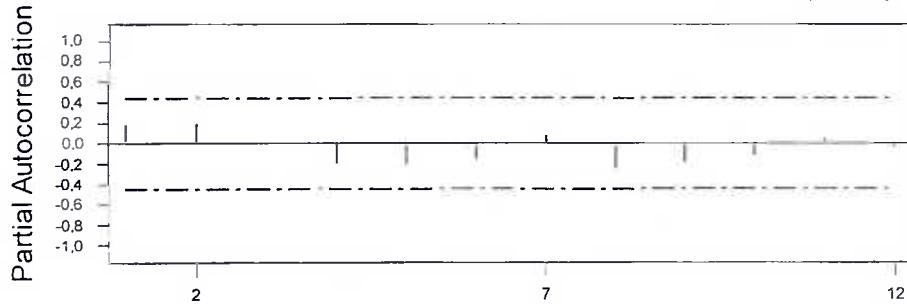
ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ ΥΠΟΔΕΙΓΜΑΤΟΣ ARIMA (0,2,1)



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0,17	0,82	0,76	8	-0,31	-1,19	10,79
2	0,21	0,98	1,98	9	-0,15	-0,54	11,70
3	0,04	0,19	2,03	10	-0,11	-0,40	12,24
4	-0,15	-0,66	2,71	11	0,03	0,10	12,28
5	-0,24	-1,04	4,55	12	0,04	0,13	12,35
6	-0,27	-1,11	6,97				
7	-0,08	-0,31	7,20				

ΓΡΑΦΗΜΑ 3.22: Αυτοσυσχέτιση καταλοίπων υποδείγματος ARIMA(0,2,1) για το Δημοτικό

ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ ARIMA(0,2,1)



Lag	PAC	T	Lag	PAC	T
1	0,17	0,82	8	-0,26	-1,22
2	0,19	0,89	9	-0,19	-0,91
3	-0,02	-0,10	10	-0,13	-0,62
4	-0,21	-0,99	11	0,05	0,22
5	-0,22	-1,04	12	-0,05	-0,23
6	-0,16	-0,76			
7	0,09	0,42			

ΓΡΑΦΗΜΑ 3.23: Μερική αυτοσυσχέτιση καταλοίπων ARIMA(0,2,1) για το Δημοτικό

Η ικανότητα του μοντέλου που εξετάσαμε δεν είναι βελτιωμένη σε σχέση με τα άλλα υποδείγματα ARIMA που εξετάσαμε πιο πάνω.

ΠΙΝΑΚΑΣ ΠΡΑΓΜΑΤΙΚΩΝ ΚΑΙ ΠΡΟΒΛΕΠΟΜΕΝΩΝ ΤΙΜΩΝ

Πραγματικές Τιμές	Προβλέψεις
710,8	696,341
675,3	667,391
652,0	636,848

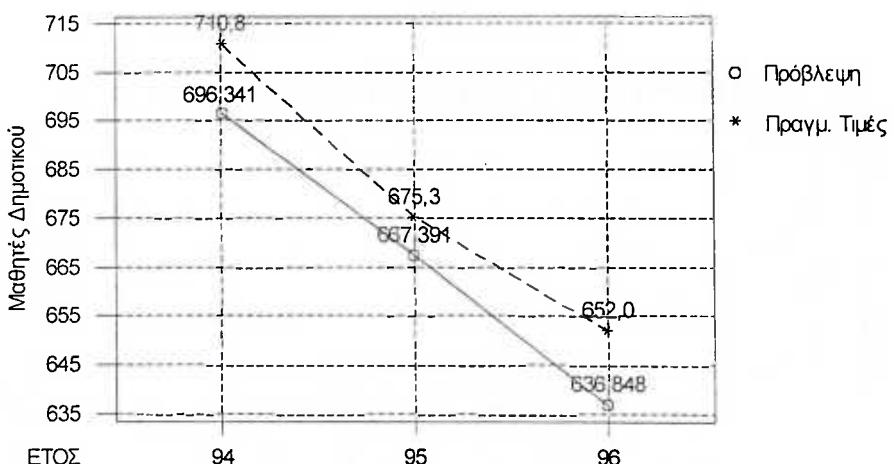
Πίνακας 3.9: Πραγματικές τιμές και προβλέψεις υποδείγματος ARIMA(0,2,1)

ΠΙΝΑΚΑΣ ΜΕΤΡΩΝ ΑΚΡΙΒΕΙΑΣ

MAPE	1,88024
MAD	12,5067
MSE	167,065

Πίνακας 3.10: Μέτρα ακριβείας ARIMA (0,2,1)

ΠΡΑΓΜΑΤΙΚΕΣ ΤΙΜΕΣ ΚΑΙ ΠΡΟΒΛΕΨΗ ΓΙΑ ΤΑ ΕΤΗ 1994-1996



ΓΡΑΦΗΜΑ 3.24: Πραγματικές τιμές και πρόβλεψη υποδείγματος ARIMA(0,2,1)

Και σ' αυτή την περίπτωση είναι εμφανής η υποτιμημένη εκτίμηση για τις τιμές των ετών 1994 – 1996. Αυτό το υπόδειγμα είναι το δεύτερο καλύτερο από τα ARIMA που σημαίνει ότι οι δεύτερες διαφορές βοήθησαν στην προσαρμογή δεν βοήθησαν ωστόσο στην πρόβλεψη. Συνοψίζουμε τα αποτελέσματα όλων των μοντέλων στον παρακάτω πίνακα

3.6 ΣΥΓΚΡΙΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ

Για την συγκεκριμένη χρονολογική σειρά η καλύτερη πρόβλεψη προέρχεται από το υπόδειγμα ARIMA (1,1,1) αλλά είναι ενδιαφέρον ότι και το δεύτερο καλύτερο μοντέλο είναι επίσης από το υπόδειγμα ARIMA(0,2,1). Τελικά οι πιο λανθασμένες σε μέγεθος προβλέψεις προέρχονται από το υπόδειγμα της απλής τάσης.

	MAPE	MAD	MSE
$Y_t=901+8,48t-0,622t^2$	3,297	22,201	523,532
Δ.Ε.Ε (Διπλή Εκθετική Εξομάλυνση)	1,305	11,506	229,699
ARIMA(1,1,0)	2,03124	13,8745	232,536
ARIMA(1,1,1)	1,6787	11,2712	137,416
ARIMA(0,2,1)	1,88024	12,5067	167,065

Πίνακας 3.11: Συγκριτικά αποτελέσματα μέτρων ακριβείας των μοντέλων που εξετάστηκαν

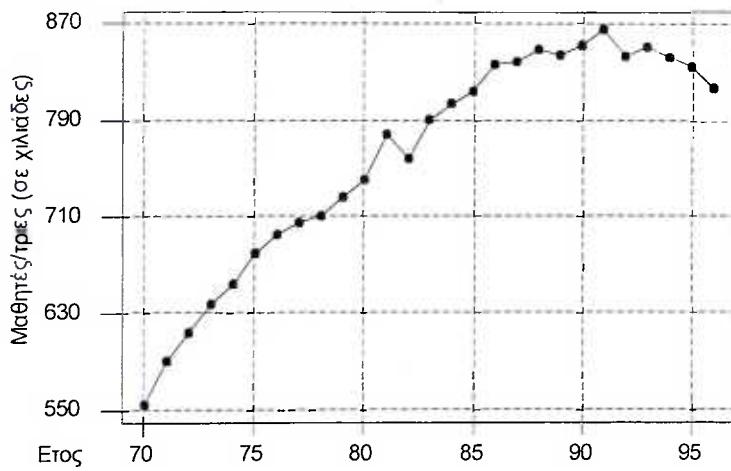
ΚΕΦΑΛΑΙΟ 4 :ΓΕΝΙΚΗ ΕΙΚΟΝΑ ΤΗΣ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ

4.1 ΣΥΝΟΛΟ ΜΑΘΗΤΩΝ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ

Η συνολική εικόνα του αριθμού των μαθητών της δευτεροβάθμιας εκπαίδευσης είναι μια διαρκής άνοδος του αριθμού των μαθητών από το 1970/71, η οποία προφανώς θα είχε αρχίσει νωρίτερα, και οφειλόταν στην συνειδητοποίηση αφ' ενός του πληθυσμού στα οφέλη της εκπαίδευσης και αφ ετέρου στη αύξηση του αριθμού των σχολείων που δημιουργούνται, άλλες εποχές με γρήγορους ρυθμούς άλλες με πιο αργούς.

Σημείο καμπής για την δευτεροβάθμια εκπαίδευσης είναι το έτος 1991/92 μετά το οποίο η μείωση των γεννήσεων αρχίζει να επηρεάζει αισθητά τους μαθητές στο γυμνάσιο και στο λύκειο.

ΣΥΝΟΛΟ ΜΑΘΗΤΩΝ/ΤΡΙΩΝ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ



ΠΗΓΗ: Eurostat

ΓΡΑΦΗΜΑ 4.1: Σύνολο μαθητών /τριων Δευτεροβάθμιας εκπαίδευσης από το 1970/71 - 1996/97

Η στατιστική ανάλυση που ακολουθεί δίνει ένα δευτεροβάθμιο υπόδειγμα ($Y_t=532+26,8t+0,542t^2$) για την τάση της χρονολογικής αυτής σειράς. Η προσαρμογή του υποδείγματος είναι πλήρως ικανοποιητική αφού ο συντελεστής προσδιορισμού είναι της τάξης του 98.5% που σημαίνει ότι το περίπου το 99% της μεταβλητότητας

των δεδομένων εξηγείται από το συγκεκριμένο υπόδειγμα. Επιπλέον όλοι οι συντελεστές του είναι στατιστικά σημαντικοί καθώς και συνολικά η παλινδρόμηση είναι επίσης στατιστικά σημαντική, όπως φαίνεται από τον F έλεγχο.

Regression Analysis: Secondary versus time; Time**2

The regression equation is
 $Secondary = 532 + 26,8 \text{ time} - 0,572 \text{ Time}^{**2}$

Predictor	Coef	SE Coef	T	P
Constant	531,630	7,429	71,56	0,000
time	26,820	1,223	21,93	0,000
Time**2	-0,57202	0,04239	-13,50	0,000

S = 11,93 R-Sq = 98,5% R-Sq(adj) = 98,3%

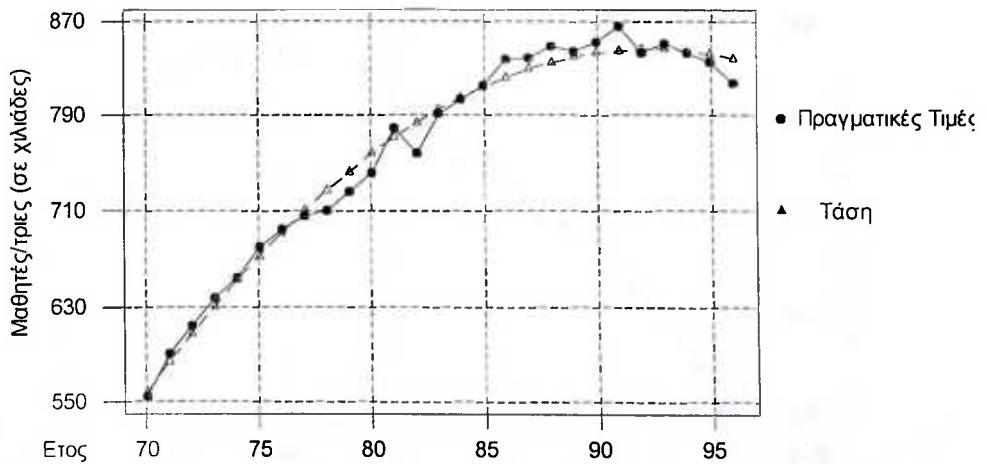
Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	217095	108548	763,13	0,000
Residual Error	24	3414	142		
Total	26	220509			

Source	DF	Seq SS
time	1	191190
Time**2	1	25905

Durbin-Watson statistic = 1,14

ΤΑΣΗ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ



ΓΡΑΦΗΜΑ 4.2: Τάση της χρονολογικής σειράς μαθητών/τριών Δευτεροβάθμιας Εκπαίδευσης

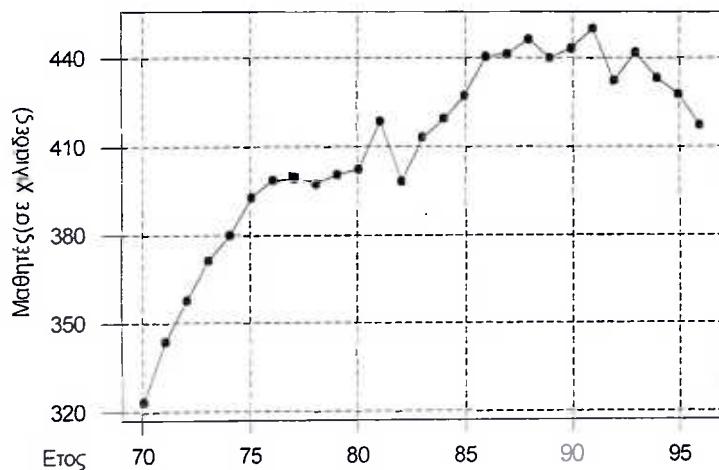
Παρουσιάζει ενδιαφέρον να εξετάσουμε τις χρονολογικές σειρές των αγοριών και των κοριτσιών ξεχωριστά

4.2 ΑΡΡΕΝΕΣ ΜΑΘΗΤΕΣ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ

Η εικόνα που εμφανίζεται και σ' αυτή την περίπτωση είναι ανάλογη της γενικής εικόνας που εμφανίζεται στην δευτεροβάθμια εκπαίδευση. Υπάρχει διαρκής αύξηση του αριθμού των μαθητών από το σχολικό έτος 1970/71 μέχρι ουσιαστικά και το 1986/87. Στην συνέχεια έχουμε σταθεροποίηση για λίγα έτη μέχρι το 1991/92 οπότε και εμφανίζεται η πτώση.

Η τάση που προσαρμόζεται καλύτερα στην σειρά αυτή είναι δευτεροβάθμια της μορφής $Y_t = b_0 + b_1 t + b_2 t^2$. Η προσαρμογή της δεν είναι τόσο ικανοποιητική όσο στο σύνολο των μαθητών

ΑΡΡΕΝΕΣ ΜΑΘΗΤΕΣ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ



ΠΗΓΗ: Eurostat

ΓΡΑΦΗΜΑ 4.3: Άρρενες Μαθητές Δευτεροβάθμιας εκπαίδευσης από 1970/71 – 1996/97

Regression Analysis: SecM versus time; Time**2

The regression equation is
 $SecM = 326 + 10.6 \text{ time} - 0.253 \text{ Time}^{**2}$

Predictor	Coeff	SE Coef	T	P
Constant	325.644	5.757	56.56	0.000
time	10.6344	0.9476	11.22	0.000
Time**2	-0.25283	0.03285	-7.70	0.000

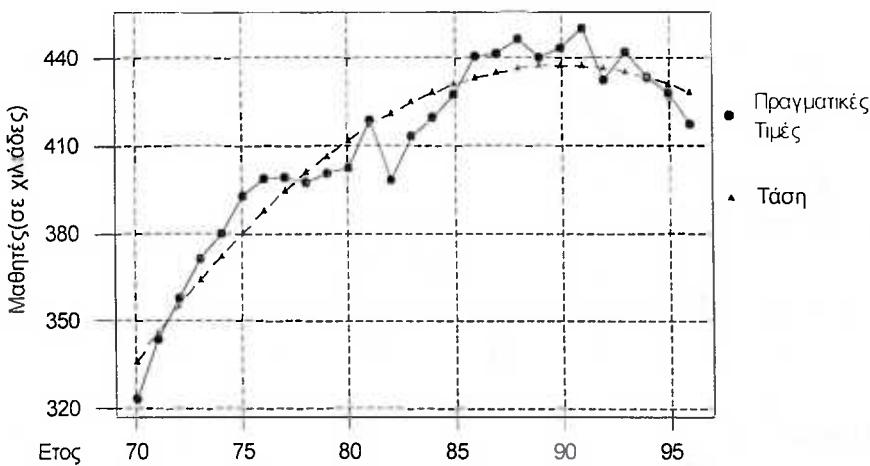
S = 9.242 R-Sq = 92.6% R-Sq(adj) = 92.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	25764	12882	150.81	0.000
Residual Error	24	2050	85		
Total	26	27814			

Durbin-Watson statistic = 0,95

ΤΑΣΗ ΑΡΡΕΝΩΝ ΜΑΘΗΤΩΝ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ



ΓΡΑΦΗΜΑ 4.4: Τάση μαθητών Δευτεροβάθμιας Εκπαίδευσης 1970/71 – 1996/97

Το υπόδειγμα εξηγεί το 92,5% της συνολικής μεταβλητότητας όπως συμπεραίνουμε από τον συντελεστή προσδιορισμού. Αυτό σημαίνει ότι ακολουθεί την “πορεία” της χρονολογικής σειράς σε πολύ υψηλό βαθμό.

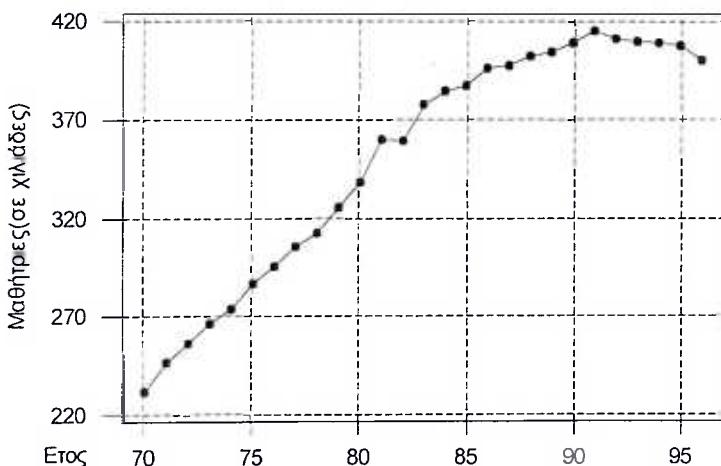
Όλοι οι συντελεστές είναι στατιστικά σημαντικοί αφού απορρίπτεται η υπόθεση της μηδενικότητας εμφανίζει μηδενική πιθανότητα. Σημαντική θεωρείται και η

παλινδρόμηση γενικά όπως προκύπτει από το F-test. Πρόβλημα εμφανίζεται στην τιμή του συντελεστή Durbin-Watson ο οποίος σημειώνει πιθανή συσχέτιση των καταλοίπων.

4.3 ΜΑΘΗΤΡΙΕΣ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ

Είναι πολύ αξιοπερίεργη η συμπεριφορά της χρονολογικής σειράς των μαθητριών δευτεροβάθμιας εκπαίδευσης γιατί ομοιάζει παρά πολύ στο σύνολο, πολύ περισσότερο από την χρονολογική σειρά των μαθητών. Η σειρά αυτή εξελίσσεται πολύ πιο ομαλά από τις δύο προηγούμενες σειρές. Όπως φαίνεται από το Γράφημα 4.5 από το 1970/71 έχουμε μια διαρκώς αυξανόμενη πορεία η οποία έχει την μέγιστη τιμή της το 1991/92. Στην συνέχεια υπάρχει η κοινή πτώση που εμφανίζεται και στους μαθητές και οφείλεται κατά πάσα πιθανότητα στην τραγική μείωση των γεννήσεων που είχε ξεκινήσει ήδη από την δεκαετία του 70 αλλά έγινε ιδιαίτερα έντονη την δεκαετία του 80 και εξακολούθει να συνεχίζεται και σήμερα.

ΜΑΘΗΤΡΙΕΣ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ



ΠΗΓΗ: Eurostat

ΓΡΑΦΗΜΑ 4.5: Χρονολογική σειρά μαθητριών δευτεροβάθμιας εκπαίδευσης από το 1970/71 – 1996/97

Η εύρεση της τάσης δεν είναι καθόλου δύσκολη υπόθεση και το υπόδειγμα ($Y_t=224+8.95t-0.0151t^2$) αγγίζει πολύ υψηλό ποσοστό προσαρμογής.

Regression Analysis: SecF versus time; Time**2; Time**3

The regression equation is
 $SecF = 224 + 8.95 \text{ time} + 0.315 \text{ Time}^{**2} - 0.0151 \text{ Time}^{**3}$

Predictor	Coeff	SE Coef	T	P
Constant	224.392	3.528	63.60	0.000
time	8.947	1.072	8.35	0.000
Time**2	0.31540	0.08807	3.58	0.002
Time**3	-0.015104	0.002070	-7.30	0.000

S = 3.963 R-Sq = 99.6% R-Sq(adj) = 99.6%

Analysis of Variance

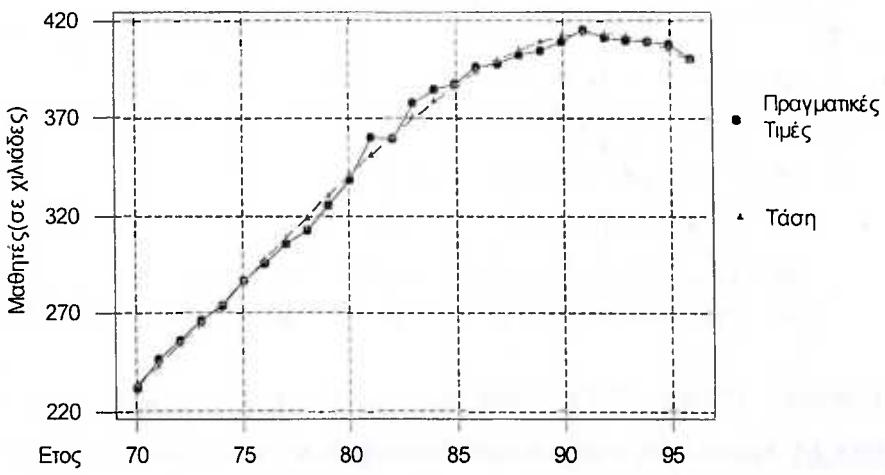
Source	DF	SS	MS	F	P
Regression	3	95005	31668	2016.82	0.000
Residual Error	23	361	16		
Total	26	95366			

Source	DF	Seq SS
time	1	86113
Time**2	1	8056
Time**3	1	836

Durbin-Watson statistic = 1.45

Η τριτοβάθμια τάση της μορφής $Y_t = b_0 + b_1 t + b_2 t^2 + b_3 t^3$ προσαρμόζεται σχεδόν απόλυτα στα δεδομένα. Πιο συγκεκριμένα “επεξηγεί” τη μεταβλητότητα της σειράς σε ποσοστό 99,6%, όπως αποκαλύπτουν οι συντελεστές προσδιορισμού.

ΤΑΣΗ ΜΑΘΗΤΡΙΩΝ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ



ΓΡΑΦΗΜΑ 4.6: Τάση της σειράς των μαθητριών δευτεροβάθμιας εκπαίδευσης

Θέμα μη σημαντικότητας των συντελεστών δεν τίθεται καθώς η υπόθεση της μηδενικότητας τους απορρίπτεται από τις p-τιμές των ελέγχων t, όπως απορρίπτεται

και η μη σημαντικότητα της παλινδρόμησης από την ρ-τιμή του ελέγχου F. Επιπλέον, τα κατάλοιπα μπορούμε να δεχτούμε ότι είναι ασυσχέτιστα αφού η τιμή της Durbin-Watson είναι περίπου 1,5.

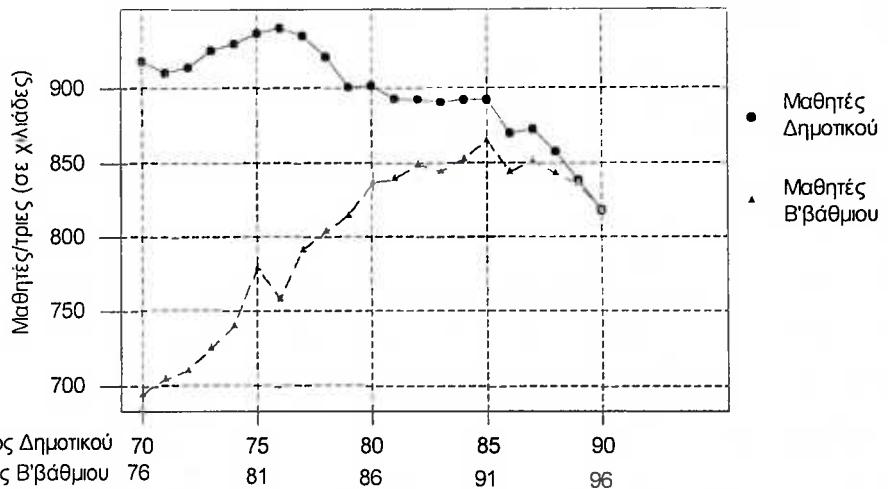
• Η είχε ενδιαφέρον πέρα από τις τάσεις των χρονολογικών σειρών να προσπαθήσουμε μέσα από τα δεδομένα μας να ανακαλύψουμε το ποσοστό των μαθητών του δημοτικού που συνέχισαν στην δευτεροβάθμια εκπαίδευση.

Γνωρίζοντας τον συνολικό αριθμό των μαθητών του δημοτικού μια δεδομένη χρονική στιγμή T_v , ύστερα από 6 χρόνια αυτός ο αριθμός θα πρέπει να βρίσκεται στην δευτεροβάθμια εκπαίδευση την χρονική στιγμή T_{v+6} . Δηλαδή οι μαθητές του δημοτικού κατά το σχολικό έτος 1970/71 θα πρέπει να είναι όλοι μαθητές δευτεροβάθμιας το έτος 1976/77. Η διαφορά των δύο αυτών θα είναι τα άτομα που εγκατέλειψαν την εκπαίδευση τους κατά την δευτή πορεία και αυτό είτε θα έγινε στο δημοτικό, είτε κατά την μετάβαση από το δημοτικό στη δευτεροβάθμια, είτε κατά την διάρκεια της δευτεροβάθμιας εκπαίδευσης.

Παρατηρούμε στο Γράφημα 4.7 την διαφορά μεταξύ των μαθητών/τριών του δημοτικού για το έτος 1970/71 και των μαθητών/τριών δευτεροβάθμιας το έτος 1976/77. Η τιμή της από 225 χιλιάδες αρχικά καταλήγει να γίνεται σχεδόν μηδενική στο τέλος. Όπως φαίνεται και από τον Πίνακα 4.1 το ποσοστό των μαθητών του δημοτικού του 1970/71 που συνέχισε έξι χρόνια μετά και βρέθηκε στο γυμνάσιο ή στο λύκειο ήταν περίπου 75%, ενώ σχεδόν όλοι οι μαθητές που βρίσκονταν στο δημοτικό το έτος 1990/91 βρέθηκαν στην δευτεροβάθμια εκπαίδευση. Βέβαια θα πρέπει να τονίσουμε ότι στα ποσοστά αυτά υπεισέρχονται και άτομα που μπορεί να εγγράφηκαν στην δευτεροβάθμια εκπαίδευση χωρίς να έχουν τελειώσει την πρωτοβάθμια εκπαίδευση της χώρας μας. Αυτό συμβαίνει κυρίως για τους μαθητές του δημοτικού από το σχολικό έτος 1984/85 και μετά, οι οποίοι συνέχισαν στην δευτεροβάθμια εκπαίδευση τα έτη 1990/91 οπότε με την κατάρρευση των ανατολικών καθεστώτων εισέρευσαν πολλοί μετανάστες.

Τα αγόρια είχαν από την αρχή υψηλότερο ποσοστό ατόμων που συνέχιζαν σε σχέση με τα κορίτσια. Έτσι οι μαθητές του δημοτικού το 1970/71 εγγράφηκαν έστω και σε μια τάξη του γυμνασίου κατά 83% ενώ οι μαθήτριες κατά 67%. Εδώ θα σημειώσουμε ότι θεωρούμε το ποσοστό εγκατάλειψης του δημοτικού σχεδόν αμελητέο.

ΜΑΘΗΤΕΣ ΔΗΜΟΤΙΚΟΥ ΚΑΙ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ



ΓΡΑΦΗΜΑ 4.7: Μαθητές του δημοτικού και αντίστοιχοι δευτεροβάθμιας 6 χρόνια μετά

Στο τέλος της περιόδου που εξετάζουμε οι μαθητές συνεχίζουν μετά έξι χρόνια κατά 99% περίπου ενώ οι μαθήτριες αυξημένες και κατά 1%. Και στα δύο αυτά ποσοστά πρέπει να έχει ενσωματωθεί και μέρος μεταναστών που τελείωσαν την πρωτοβάθμια εκπαίδευση στη χώρα τους και συνέχισαν την δευτεροβάθμια εδώ, καθώς επίσης και άτομα που επανέλαβαν την τάξη τους..

Επειδή η δευτεροβάθμιας εκπαίδευση δεν θεωρείται ενιαία και αλλά χωρίζεται σε κατώτερη δευτεροβάθμια και ανώτερη δευτεροβάθμια από τα οποία μόνο το γυμναστικό είναι υποχρεωτικό ως προς την παρακολούθηση, είναι καλύτερο να ξεχωρίσουμε τις δύο αυτές βαθμίδες της δευτεροβάθμιας εκπαίδευσης και να τις αναλύσουμε ξεχωριστά.

Σχολικό Έτος	Ποσοστό Συνόλου που συνεχίζει στη δευτεροβάθμια	Ποσοστό Αρρενών Μαθητών που συνεχίζει στη δευτεροβάθμια	Ποσοστό Μαθητριών που συνεχίζει στη δευτεροβάθμια
1970/71	75,5	83,0	67,3
1971/72	77,3	84,0	70,0
1972/73	77,6	83,4	71,3
1973/74	78,3	83,0	73,1
1974/75	79,5	83,1	75,6
1975/76	83,0	85,9	79,8
1976/77	80,5	81,6	79,3
1977/78	84,5	85,3	83,6
1978/79	87,1	87,8	86,5
1979/80	90,3	91,3	89,2
1980/81	92,7	94,3	90,9
1981/82	93,9	95,4	92,2
1982/83	95,1	97,0	93,1
1983/84	94,8	95,8	93,8
1984/85	95,5	96,3	94,6
1985/86	96,8	97,5	96,1
1986/87	96,8	96,0	97,7
1987/88	97,5	98,1	96,9
1988/89	98,2	98,0	98,5
1989/90	99,5	98,8	100,3
1990/91	100,0	99,0	101,0

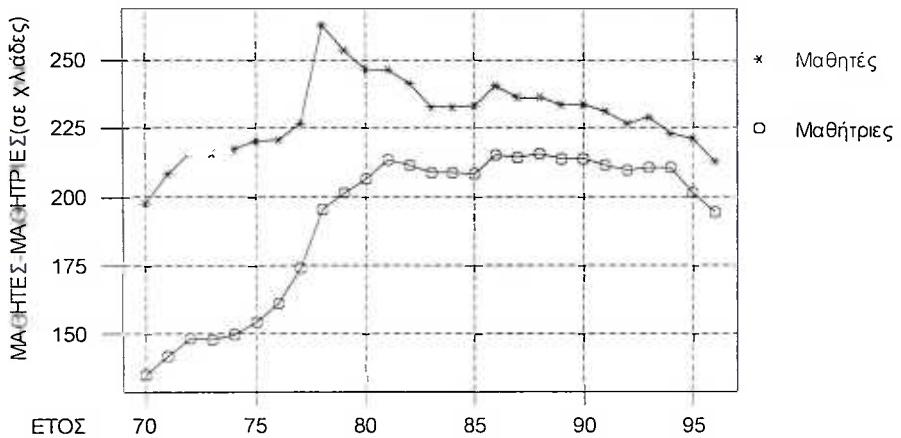
Πίνακας 4.1: Ποσοστά μαθητών και μαθητριών που συνεχίζουν στη δευτεροβάθμια εκπαίδευση.



ΚΕΦΑΛΑΙΟ 5 : ΚΑΤΩΤΕΡΗ ΔΕΥΤΕΡΟΒΑΘΜΙΑ ΕΚΠΑΙΔΕΥΣΗ

5.1 ΓΕΝΙΚΗ ΕΙΚΟΝΑ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ,

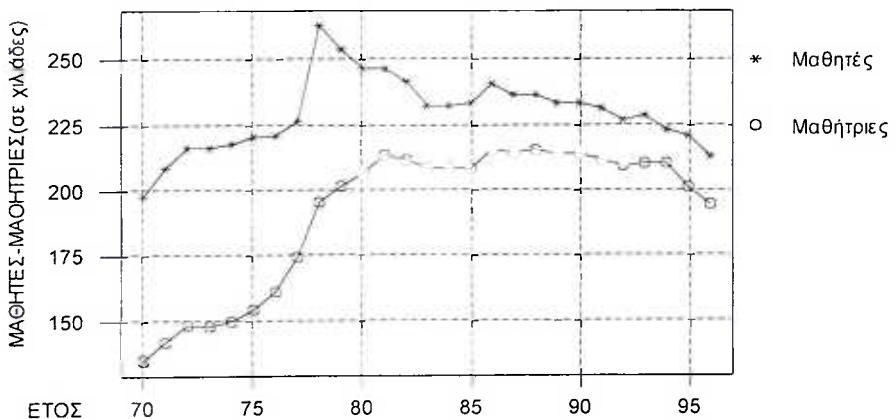
ΑΡΙΘΜΟΣ ΜΑΘΗΤΩΝ - ΜΑΘΗΤΡΙΩΝ 1970-1996



ΠΗΓΗ: Eurostat

ΓΡΑΦΗΜΑ 5.1: Σύνολο μαθητών/τριων κατώτερης δευτεροβάθμιας εκπαίδευσης 1970/71- 1996/97

ΑΡΙΘΜΟΣ ΜΑΘΗΤΩΝ - ΜΑΘΗΤΡΙΩΝ 1970-1996



ΠΗΓΗ: Eurostat

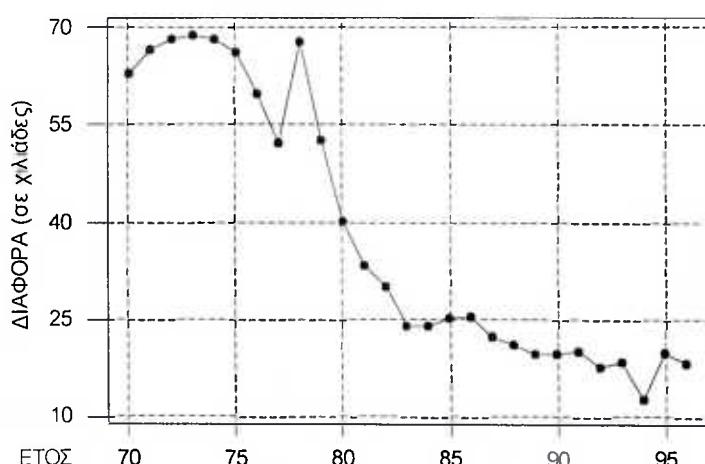
ΓΡΑΦΗΜΑ 5.2: Αριθμός μαθητών – μαθητριών κατώτερης δευτεροβάθμιας εκπαίδευσης

Ο συνολικός αριθμός των μαθητών και μαθητριών του Γυμνασίου παρουσιάζει μια περίεργη συμπεριφορά. Για τα έτη 1970 μέχρι 1978 εμφανίζει αυξητική τάση και στην συνέχεια αρχίζει μια διαρκής πτώση. Η ραγδαία αύξηση των 58 χιλιάδων μαθητών γυμνασίου του 1978 σε σχέση με το 1977 οφείλεται στον διαχωρισμό του

πρώην εξατάξιου γυμνασίου σε γυμνάσιο και λύκειο και συγχρόνως στην κατάργηση των εισιτήριων εξετάσεων για αυτό. Μετά το 1978 ωστόσο αρχίζει μια διαρκής πτώση του πληθυσμού των μαθητών γυμνασίου η οποία μπορεί να αποδοθεί στη συνεχή μείωση γεννήσεων και σαν συνέπεια στη μείωση των νεαρών ατόμων, μαθητών οποιασδήποτε βαθμίδας εκπαίδευσης. Στο Γράφημα 5.2 παρουσιάζεται η εξέλιξη του αριθμού των μαθητών και μαθητριών συγχρόνως.

Ίδια παράξενη αύξηση εμφανίζεται και στη διαφορά του αριθμού αγοριών και κοριτσιών. Έτσι ενώ μειώνεται ταχύτατα από τις αρχές της δεκαετίας του 70, από τα μέσα της δεκαετίας του 80 συνεχίζει μεν την πτώση αλλά με πολύ πιο αργούς ρυθμούς, και παραδόξως αυξάνεται κατά τα έτη 1995/96 και 1996/1997.

ΔΙΑΦΟΡΑ ΜΑΘΗΤΩΝ - ΜΑΘΗΤΡΙΩΝ



ΓΡΑΦΗΜΑ 5.3 Η διαφορά μαθητών – μαθητριών κατώτερης δευτεροβάθμιας εκπαίδευσης 1970/71 – 1996/97

Αρχικά εξετάζουμε την μη τυχαιότητα της σειράς που προέρχεται από την διαφορά μαθητών μαθητριών. Γίνεται δηλαδή ο έλεγχος:

$$H_0: \text{Τα δεδομένα τυχαία}$$

$$H_1: \text{Τα δεδομένα μη τυχαία}$$

Median of Diff

Median of Diff = 25,300
Runs Test: Diff

Diff
K = 25,3000

The observed number of runs = 2
The expected number of runs = 14,4815
13 Observations above K 14 below
The test is significant at 0,0000

Η p-τιμή = 0,0000 απορρίπτει την μηδενική υπόθεση, άρα προχωρούμε στην ανάλυση μας.

Το υπόδειγμα που εμφανίζει την καλύτερη προσαρμογή στην χρονολογική σειρά της διαφοράς αγοριών κοριτσιών είναι :

A) Regression Analysis: logDiff versus Log(Time); Log(Time)**2

The regression equation is
logDiff = 4,10 + 0,514 Log(Time) - 0,287 Log(Time)**2

Predictor	Coef	SE Coef	T	P
Constant	4,0964	0,1353	30,27	0,000
Log(Time)	0,5137	0,1458	3,52	0,002
Log(time)**2	-0,28708	0,03729	-7,70	0,000

S = 0,1568 R-Sq = 92,7% R-Sq(adj) = 92,1%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	7,4901	3,7450	152,38	0,000
Residual Error	24	0,5899	0,0246		
Total	26	8,0799			

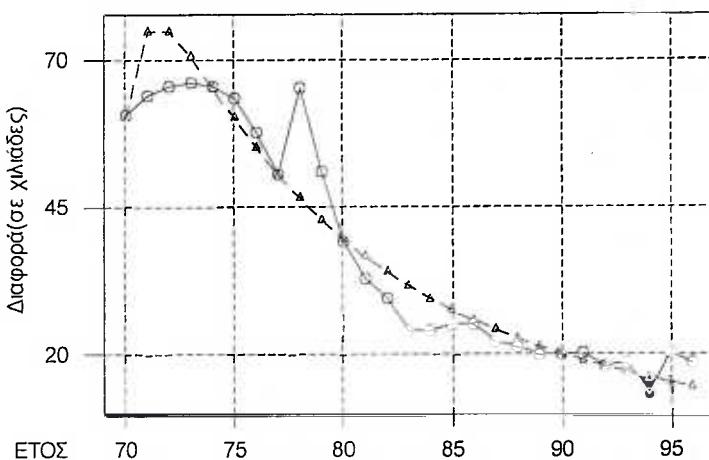
Durbin-Watson statistic = 1,13

Οι τιμές R-sq και R-sq(adj) για το μοντέλο που μπορούν να συγκριθούν με οποιοδήποτε μοντέλο αφού προέρχονται από τα μετασχηματισμένα δεδομένα στις πραγματικές τιμές, είναι αντίστοιχα: 91,81% και 91, 12%. Επιπλέον όλες οι παράμετροι είναι στατιστικά σημαντικές όπως και η παλινδρόμηση. Εξαίρεση αποτελεί η τιμή Durbin-Watson που εμφανίζεται αρκετά μικρή.



ΑΝΑΛΥΣΗ ΤΑΣΗΣ ΓΙΑ ΤΗΝ ΔΙΑΦΟΡΑ

$$\log Y_t = 4,0964 - 0,5137 \log t - 0,28708 \log t^{**2}$$



ΓΡΑΦΗΜΑ 5.4: Λογαριθμική τάση της διαφοράς – αγοριών κοριτσιών

Παρατηρώντας το Γράφημα 5.4, μπορούμε να αποδεχτούμε ότι το υπόδειγμα προσεγγίζει αρκετά καλά τα δεδομένα. Δεχόμενοι ότι η ανοδική τάση της διαφοράς για τα έτη 95 και 96 είναι προσωρινή και δεν μπορεί να συνεχίσει να είναι ανοδική το μοντέλο σημειώνει αυτή την μείωση. Σύμφωνα με το λογαριθμικό υπόδειγμα η διαφορά αναμένεται να μηδενιστεί ύστερα από 5 περίπου έτη

5.2 ΕΚΤΙΜΗΣΗ ΤΑΣΗΣ ΜΑΘΗΤΡΙΩΝ

Θα κάνουμε τον συνηθισμένο έλεγχο τυχαιότητας για την χρονολογική σειρά του αριθμού μαθητριών στο γυμνάσιο.

Median of LS-Females

Median of LS-Females = 208,35

Runs Test: LS-Females

LS-Female

K = 208,3500

The observed number of runs = 4
 The expected number of runs = 13,0000
 12 Observations above K 12 below
 The test is significant at 0,0002

Όπως θα περιμέναμε το τεστ απορρίπτει την τυχαιότητα της σειράς.

Τα υποδείγματα που εμφανίζουν την καλύτερη προσαρμογή στα δεδομένα είναι τα 2 παρακάτω:

A) Regression Analysis: LS-Females versus Time; Time**2

The regression equation is
 $LS\text{-Females} = 115 + 10,6 \text{ Time} - 0,277 \text{ Time}^{**2}$

Predictor	Coef	SE Coef	T	P
Constant	114,828	5,042	22,78	0,000
Time	10,5626	0,9292	11,37	0,000
Time**2	-0,27732	0,03609	-7,69	0,000

S = 7,557 R-Sq = 93,9% R-Sq(adj) = 93,3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	18523,1	9261,5	162,19	0,000
Residual Error	21	1199,2	57,1		
Total	23	19722,2			

Durbin-Watson statistic = 0,47

B) Regression Analysis: LogLSF versus Time; Time**2

The regression equation is
 $\text{LogLSF} = 4,80 + 0,0615 \text{ Time} - 0,00164 \text{ Time}^{**2}$

Predictor	Coef	SE Coef	T	P
Constant	4,80219	0,02668	179,96	0,000
Time	0,061529	0,004918	12,51	0,000
Time**2	-0,0016435	0,0001910	-8,61	0,000

S = 0,04000 R-Sq = 94,7% R-Sq(adj) = 94,2%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	0,59896	0,29948	187,22	0,000
Residual Error	21	0,03359	0,00160		
Total	23	0,63255			

Durbin-Watson statistic = 0,50

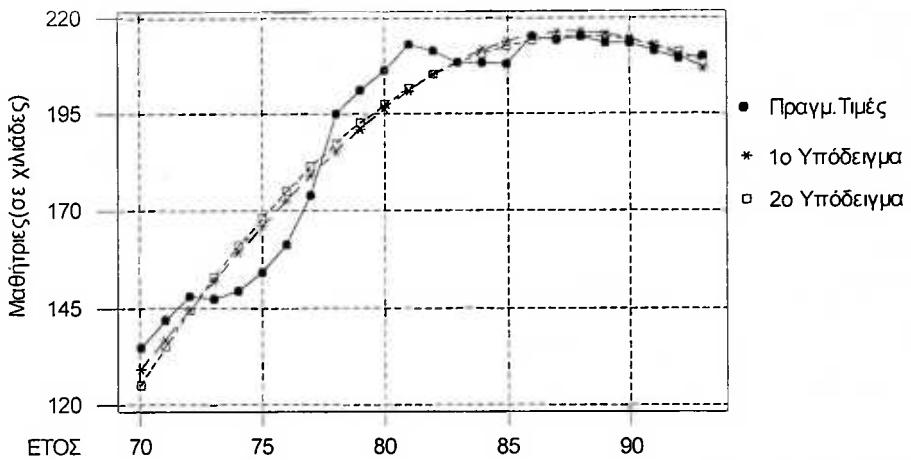
** R-sq = 94,7% και R-Sq(adj) = 94,2% συγκρίσιμες τιμές

Από τα παραπάνω υποδείγματα το λογαριθμικό εμφανίζει την καλύτερη προσαρμογή αν και οι διαφορές με τα άλλα είναι μικρές και οι τιμές του συντελεστή προσαρμογής είναι αρκετά υψηλές. Επίσης και στα δύο μοντέλα, οι συντελεστές είναι στατιστικά σημαντικοί αλλά οι τιμές των Durbin-Watson αρκετά χαμηλές

Τα Γραφήματα 5.5 και 5.6 επιβεβαιώνουν τα στατιστικά αποτελέσματα για τα μοντέλα. Η διαφορά τους είναι πολύ μικρή και πολύ σημαντική θα είναι η προβλεπτική τους ικανότητα για τα επόμενα 3 χρόνια για να επιλέξουμε κάποιο σαν καταλληλότερο.

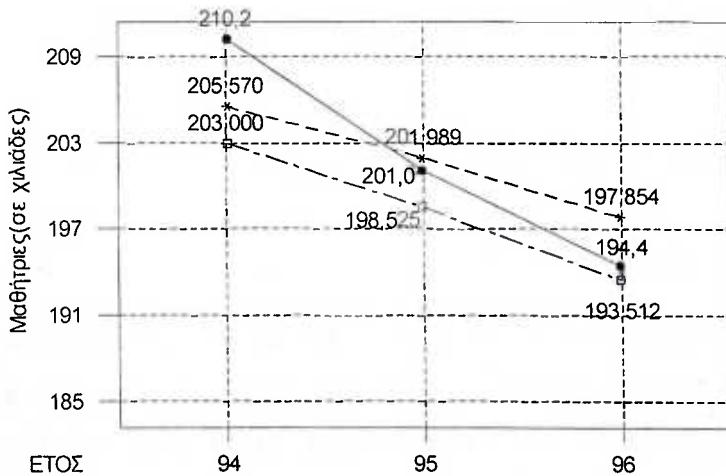


ΑΝΑΛΥΣΗ ΤΑΣΗΣ ΓΙΑ ΤΙΣ ΜΑΘΗΤΡΙΕΣ



ΓΡΑΦΗΜΑ 5.5: Ανάλυση τάσης για τον αριθμό μαθητριών κατώτερης δευτεροβάθμιας εκπαίδευσης από το 1970/71 – 1993/94

ΠΡΑΓΜΑΤΙΚΕΣ ΤΙΜΕΣ ΚΑΙ ΠΡΟΒΛΕΨΕΙΣ



ΓΡΑΦΗΜΑ 5.6: Προβλέψεις των υποδειγμάτων τάσης για τα έτης 1994/95 – 1996/97

ΠΙΝΑΚΑΣ ΠΡΑΓΜΑΤΙΚΩΝ ΤΙΜΩΝ ΚΑΙ ΠΡΟΒΛΕΠΟΜΕΝΩΝ

	Πραγμ. Τιμές	1 ^ο Υπόδειγμα $Y_t = b_0 + b_1 t + b_2 t^2$	2 ^ο Υπόδειγμα $\log Y_t = b_0 + b_1 t + b_2 t^2$
1994/95	210,2	205,570	203,000
1995/96	201,0	201,989	198,525
1996/97	194,4	197,854	193,512

Πίνακας 5.1: Πραγματικές τιμές και προβλέψεις των υποδειγμάτων της τάσης

ΠΙΝΑΚΑΣ ΜΕΤΡΩΝ ΑΚΡΙΒΕΙΑΣ

	1 ^ο Υπόδειγμα $Y_t = b_0 + b_1 t + b_2 t^2$	2 ^ο Υπόδειγμα $\log Y_t = b_0 + b_1 t + b_2 t^2$
MAPE	1,49592	1,75067
MAD	3,02443	3,52076
MSE	11,4491	19,5818

Πίνακας 5.2 Μέτρα ακριβείας των υποδειγμάτων της τάσης

Το υπόδειγμα της δευτεροβάθμιου υποδείγματος δίνει ελαφρώς καλύτερα αποτελέσματα από το λογαριθμικό υπόδειγμα. Μπορούμε δηλαδή να θεωρήσουμε ότι το 1^ο υπόδειγμα είναι το καλύτερο για να περιγράψει την τάση αυτής της σειράς.

5.3 ΔΙΠΛΗ ΕΚΘΕΤΙΚΗ ΕΞΟΜΑΛΥΝΣΗ ΓΙΑ ΤΙΣ ΜΑΘΗΤΡΙΕΣ

Το καλύτερο υπόδειγμα της εκθετικής εξομάλυνσης μετά από διαδοχικές δοκιμές, έτσι ώστε να ελαχιστοποιηθεί το MSD, έχει $\alpha=1$ και $\gamma=0,373$ που σημαίνει ότι μεγαλύτερη έμφαση δίνεται στις πρόσφατες παρατηρήσεις παρά στις παλιότερες.

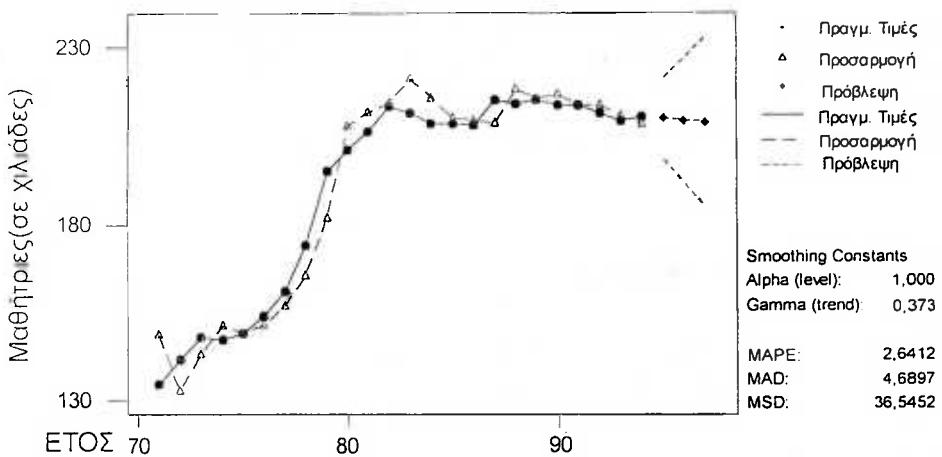
Double Exponential Smoothing

```
Data      LS-Females
Length   24,0000
NMissing 0
```

```
Smoothing Constants
Alpha (level): 1,000
Gamma (trend): 0,373
```

```
Accuracy Measures
MAPE: 2,6412
MAD: 4,6897
MSD: 36,5452
```

ΔΙΠΛΗ ΕΚΘΕΤΙΚΗ ΕΞΟΜΑΛΥΝΣΗ ΓΙΑ ΤΙΣ ΜΑΘΗΤΡΙΕΣ



ΓΡΑΦΗΜΑ 5.7 : Διπλή εκθετική εξομάλυνση, προσαρμογή και πρόβλεψη για τον αριθμό κοριτσιών στην κατώτερη δευτεροβάθμια εκπαίδευση

Οι προβλέψεις που δίνει το υπόδειγμα αυτό της εκθετικής εξομάλυνσης του Holt είναι χειρότερες από τις προβλέψεις που δίνουν τα δύο καλύτερα μοντέλα της τάσης. Είναι φανερή η αδυναμία του να ακολουθήσει την πτωτική πορεία που ακολουθούν τα πραγματικά δεδομένα. Έτσι αν και εμφανίζει πτωτική τάση η πρόβλεψη δεν έχει την κλίση που αναμένεται

ΠΙΝΑΚΑΣ ΠΡΑΓΜΑΤΙΚΩΝ ΤΙΜΩΝ ΚΑΙ ΠΡΟΒΛΕΠΟΜΕΝΩΝ

Ετος	Πραγμ. Τιμές	Δ.Ε.Ε
1994/95	210,2	209,879
1995/96	201,0	209,459
1996/97	194,4	209,038

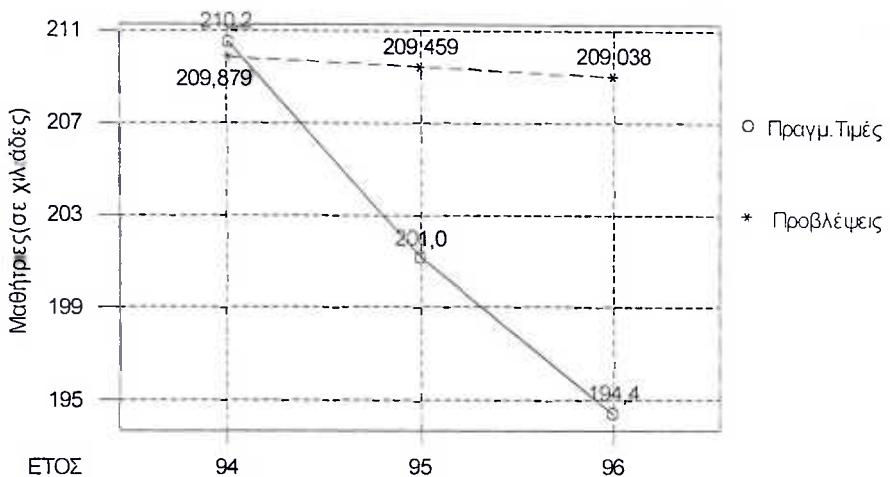
Πίνακας 5.3: Πραγματικές τιμές και προβλέψεις της Διπλής Εκθετικής Εξομάλυνσης

ΠΙΝΑΚΑΣ ΜΕΤΡΩΝ ΑΚΡΙΒΕΙΑΣ

	Δ.Ε.Ε
MAPE	3,73122
MAD	7,80575
MSE	95,3068

Πίνακας 5.4: Μέτρα ακριβείας του υποδείγματος της Διπλής Εκθετικής Εξομάλυνσης

ΠΡΑΓΜΑΤΙΚΕΣ ΤΙΜΕΣ ΚΑΙ ΠΡΟΒΛΕΨΕΙΣ ΤΗΣ ΕΚΟΕΤΙΚΗΣ ΕΞΟΜΑΛΥΝΣΗΣ HOLT

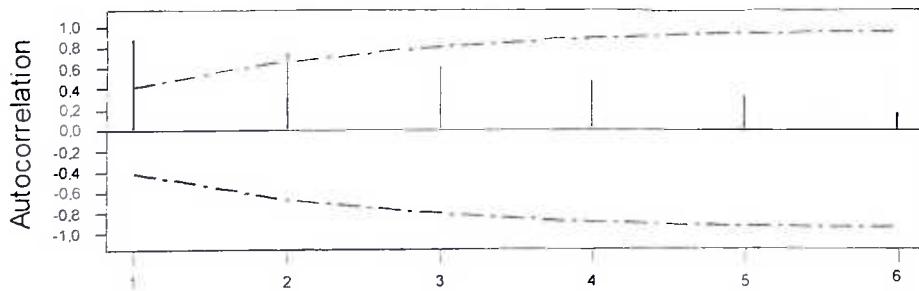


ΓΡΑΦΗΜΑ 5.8: Προβλέψεις της εκθετικής εξομάλυνσης για τα έτη 1995/95 – 1996/97

5.4 ΥΠΟΔΕΙΓΜΑΤΑ ARIMA ΓΙΑ ΤΙΣ ΜΑΘΗΤΡΙΕΣ

Από το Γράφημα 5.2 φαίνεται από ότι η σειρά δεν είναι στάσιμη. Αυτό υποδεικνύουν και τα γραφήματα 5.10 και 5.10 της αυτοσυσχέτισης και της μερικής αυτοσυσχέτισης της χρονολογικής σειράς

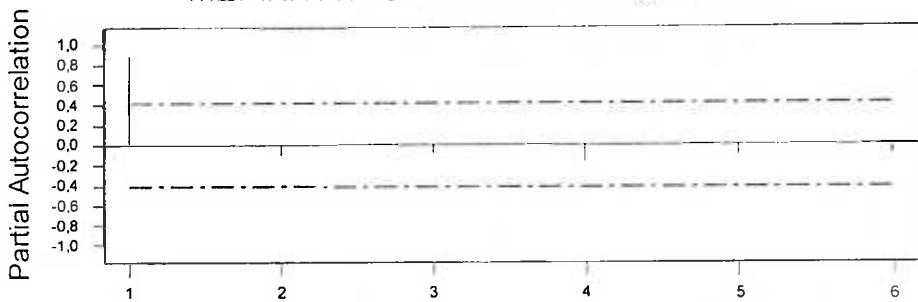
ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΑΡΙΘΜΟΥ ΜΑΘΗΤΡΙΩΝ



Lag	Corr	T	LBQ
1	0,89	4,35	21,35
2	0,76	2,33	37,85
3	0,63	1,60	49,71
4	0,48	1,10	56,85
5	0,32	0,70	60,21
6	0,17	0,36	61,18

ΓΡΑΦΗΜΑ 5.9: Συντελεστές αυτοσυσχέτισης αριθμού μαθητριών για τα έτη 1970/71 – 1993/94

ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΜΑΘΗΤΡΙΩΝ

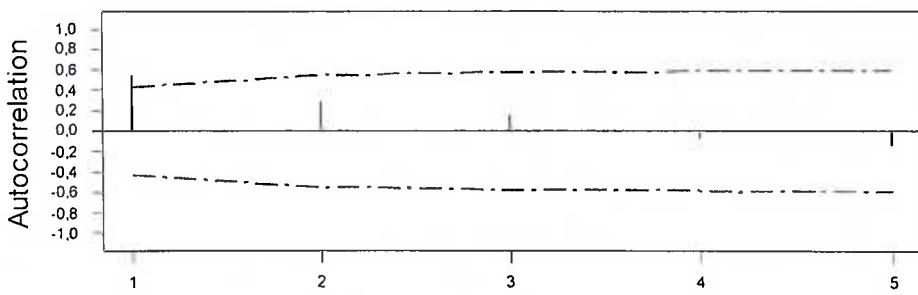


Lag	PAC	T
1	0,89	4,35
2	-0,11	-0,56
3	-0,10	-0,49
4	-0,19	-0,91
5	-0,12	-0,59
6	-0,10	-0,47

ΓΡΑΦΗΜΑ 5.10: Συντελεστές μερικής αυτοσυσχέτισης μαθητριών κατώτερης δευτεροβάθμιας εκπαίδευσης

Στο Γράφημα 5.10 .Οι συντελεστές αυτοσυσχέτισης μειώνονται μεν αλλά με πολύ αργό ρυθμό και κατά συνέπεια θα πρέπει να καταφύγουμε στις πρώτες ή δεύτερες διαφορές για να μπορέσουμε να υπολογίσουμε κάποιο καλό υπόδειγμα.

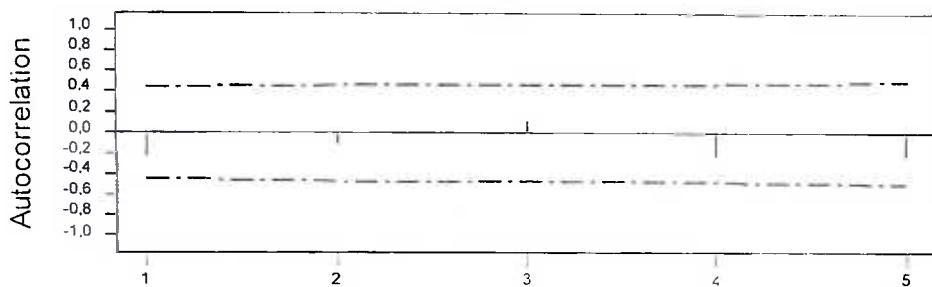
ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΠΡΩΤΩΝ ΔΙΑΦΟΡΩΝ



Lag	Corr	T	LBQ
1	0,55	2,63	7,88
2	0,29	1,09	10,16
3	0,16	0,57	10,88
4	-0,08	-0,29	11,08
5	-0,15	-0,53	11,79

ΓΡΑΦΗΜΑ 5.11: Αυτοσυσχέτιση πρώτων διαφορών για τις μαθήτριες

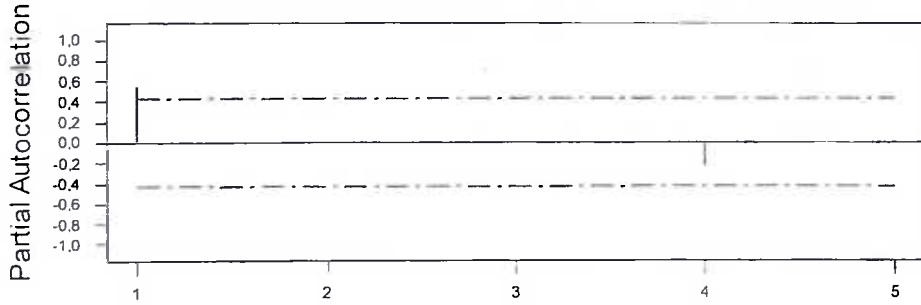
ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΔΕΥΤΕΡΩΝ ΔΙΑΦΟΡΩΝ



Lag	Corr	T	LBQ
1	-0,23	-1,08	1,32
2	-0,11	-0,49	1,65
3	0,12	0,52	2,04
4	-0,23	-0,98	3,52
5	-0,23	-0,96	5,15

ΓΡΑΦΗΜΑ 5.12: Συντελεστές μερικής αυτοσυσχέτισης πρώτων διαφορών για τις μαθήτριες κατώτερης δευτεροβάθμιας εκπαίδευσης

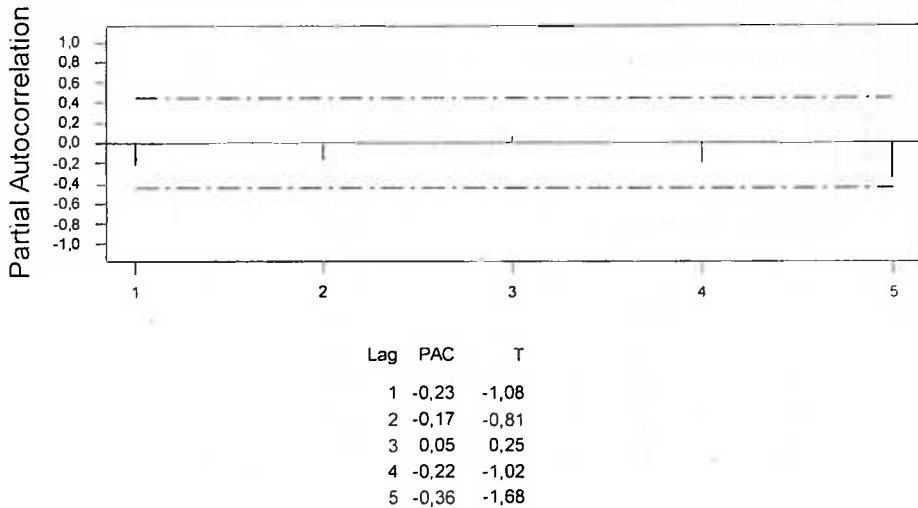
ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΠΡΩΤΩΝ ΔΙΑΦΟΡΩΝ



Lag	PAC	T
1	0,55	2,63
2	-0,02	-0,09
3	0,01	0,05
4	-0,24	-1,16
5	-0,01	-0,05

ΓΡΑΦΗΜΑ 5.13: Συντελεστές αυτοσυσχέτισης δεύτερων διαφορών για τις μαθήτριες

ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΔΕΥΤΕΡΩΝ ΔΙΑΦΟΡΩΝ



ΓΡΑΦΗΜΑ 5.14: Συντελεστές μερικής αυτοσυσχέτισης δεύτερων διαφορών των μαθητριών κατώτερης δευτεροβάθμιας εκπαίδευσης

Τα γραφήματα μπορούν να μας οδηγήσουν σε κάποια υποδείγματα. Οι συσχετίσεις και αυτοσυσχετίσεις των πρώτων διαφορών μας οδηγούν σε υπόδειγμα ARIMA (1,1,0) ενώ των δεύτερων διαφορών σε υπόδειγμα ARIMA (0,2,2).

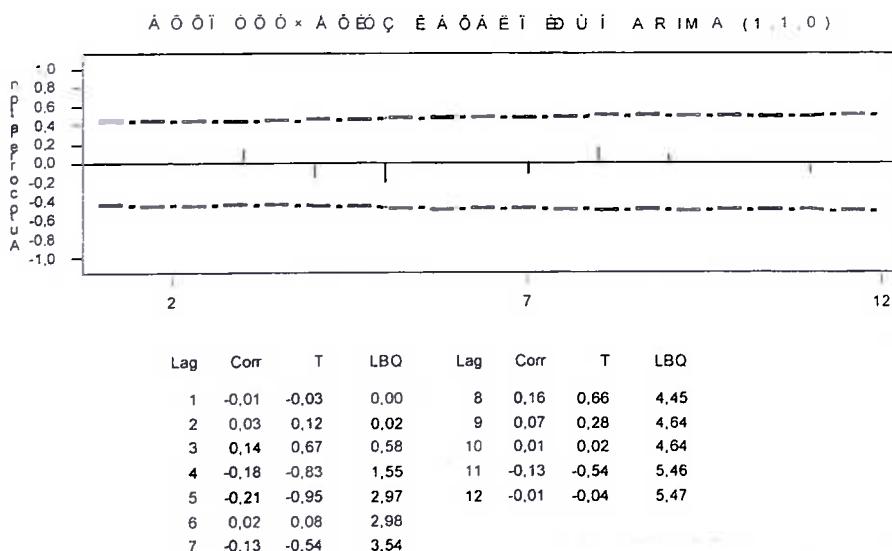
Dependent Variable: D(LSF,1)
 Method: Least Squares
 Date: 03/18/04 Time: 01:18
 Sample(adjusted): 1972 1993
 Included observations: 22 after adjusting endpoints
 Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.789923	2.370011	1.177177	0.2529
AR(1)	0.553764	0.184646	2.999050	0.0071
R-squared	0.310209	Mean dependent var	3.122727	
Adjusted R-squared	0.275720	S.D. dependent var	5.796545	
S.E. of regression	4.933130	Akaike info criterion	6.116332	
Sum squared resid	486.7154	Schwarz criterion	6.215518	
Log likelihood	-65.27966	F-statistic	8.994299	
Durbin-Watson stat	2.008782	Prob(F-statistic)	0.007091	
Inverted AR Roots	.55			

Dependent Variable: D(LSF,2)
 Method: Least Squares
 Date: 04/06/04 Time: 23:56
 Sample(adjusted): 1972 1993
 Included observations: 22 after adjusting endpoints
 Convergence achieved after 66 iterations
 Backcast: OFF (Roots of MA process too large for backcast)

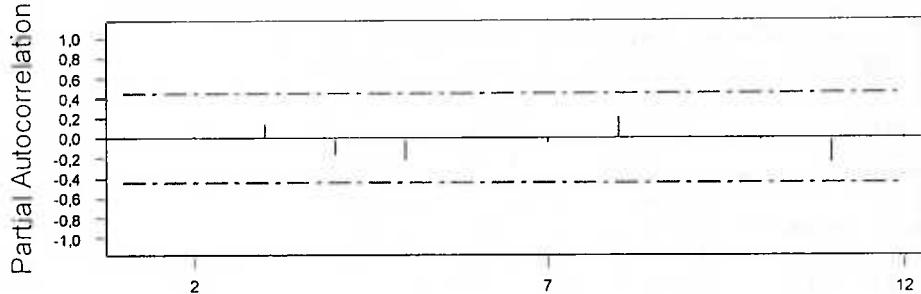
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.802448	0.459754	-1.745385	0.0971
MA(1)	-1.062677	0.311690	-3.409403	0.0029
MA(2)	-0.700617	0.346472	-2.022143	0.0575
R-squared	0.604459	Mean dependent var	-0.268182	
Adjusted R-squared	0.562823	S.D. dependent var	5.472214	
S.E. of regression	3.618193	Akaike info criterion	5.535950	
Sum squared resid	248.7351	Schwarz criterion	5.684729	
Log likelihood	-57.89546	F-statistic	14.51773	
Durbin-Watson stat	2.222143	Prob(F-statistic)	0.000149	
Inverted MA Roots	1.52		-.46	
Estimated MA process is noninvertible				

Το δεύτερο υπόδειγμα δείχνει να έχει καλύτερη προσαρμογή στα δεδομένα από το πρώτο. Εμφανίζει καλύτερο τυπικό σφάλμα παλινδρόμησης (S.E of regression) 3,62 έναντι 4,93, αθροισμα τετραγώνων καταλοίπων (Sum squared resid) 248,73 έναντι 486,71 και Akaike και Schwarz με τιμές 5,53 και 5,68 έναντι 6,11 και 6,21. Και στα δύο υποδείγματα φαίνεται ότι τα κατάλοιπα δεν είναι αυτοσυχετισμένα αφού οι τιμές της παραμέτρου Durbin-Watson είναι κοντά στην τιμή 2 αλλά και τα Γραφήματα 5.16 – 5.19 δείχνουν όλους του συντελεστές αυτοσυσχέτισης και μερικής αυτοσυσχέτισης μη σημαντικούς.



ΓΡΑΦΗΜΑ 5.15: Συντελεστές αυτοσυσχέτισης καταλοίπων υποδείγματος ARIMA (0,1,1)

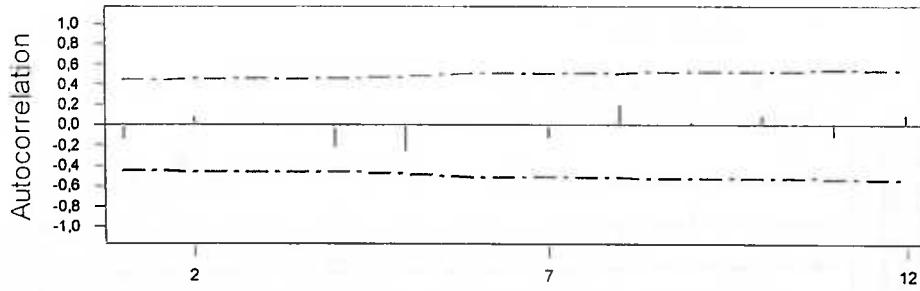
ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ ARIMA (1,1,0)



Lag	PAC	T	Lag	PAC	T
1	-0,01	-0,03	8	0,20	0,96
2	0,03	0,12	9	-0,00	-0,00
3	0,14	0,67	10	-0,03	-0,16
4	-0,18	-0,86	11	-0,25	-1,18
5	-0,23	-1,09	12	-0,02	-0,07
6	0,01	0,04			
7	-0,06	-0,27			

ΓΡΑΦΗΜΑ 5.16: Συντελεστές μερικής αυτοσυσχέτισης καταλοίπων ARIMA (0,1,1)

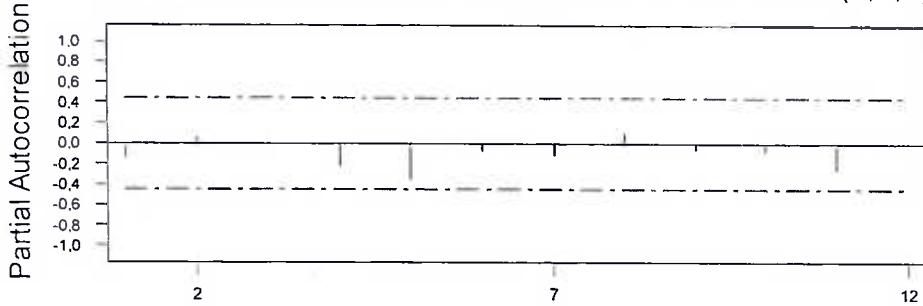
ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ ARIMA (0,2,2)



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	-0,15	-0,72	0,59	8	0,20	0,79	6,61
2	0,08	0,34	0,74	9	0,02	0,09	6,63
3	-0,00	-0,01	0,74	10	0,08	0,32	6,92
4	-0,23	-1,04	2,26	11	-0,13	-0,50	7,71
5	-0,27	-1,18	4,55	12	0,10	0,38	8,23
6	0,01	0,05	4,55				
7	-0,13	-0,54	5,17				

ΓΡΑΦΗΜΑ 5.17: Αυτοσυσχέτιση καταλοίπων υποδείγματος ARIMA (0,2,2) για τις μαθήτριες

ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ ARIMA (0,2,2)



Lag	PAC	T	Lag	PAC	T
1	-0.15	-0.72	8	0.10	0.46
2	0.05	0.25	9	-0.08	-0.36
3	0.02	0.08	10	-0.08	-0.38
4	-0.24	-1.10	11	-0.27	-1.28
5	-0.37	-1.73	12	0.01	0.03
6	-0.09	-0.42			
7	-0.13	-0.61			

ΓΡΑΦΗΜΑ 5.18: Συντελεστές μερικής αυτοσυσχέτισης καταλοίπων υποδείγματος ARIMA (0,2,2)

Η προσαρμογή των δύο μοντέλων είναι αρκετά καλή, με το δεύτερο μοντέλο να υπερτερεί. Η περιμέναμε ωστόσο η πρόβλεψη του να είναι καλύτερη από το υπόδειγμα ARIMA (0,2,2) αφού αυτό εμφανίζεται να είναι μη αντιστρέψιμο και κατά συνέπεια δεν δίνει τόσο αξιόπιστες προβλέψεις. Η πραγματικότητα ωστόσο είναι διαφορετική.

ΠΙΝΑΚΑΣ ΠΡΑΓΜΑΤΙΚΩΝ ΤΙΜΩΝ ΚΑΙ ΠΡΟΒΛΕΠΟΜΕΝΩΝ

	Πραγμ. Τιμές	ARIMA(1,1,0)	ARIMA(0,2,2)
1994/95	210,2	212,099	209,280
1995/96	201,0	214,340	206,337
1996/97	194,4	216,826	202,591

Πίνακας 5.5: Πραγματικές τιμές και προβλέψεις των υποδειγμάτων ARIMA

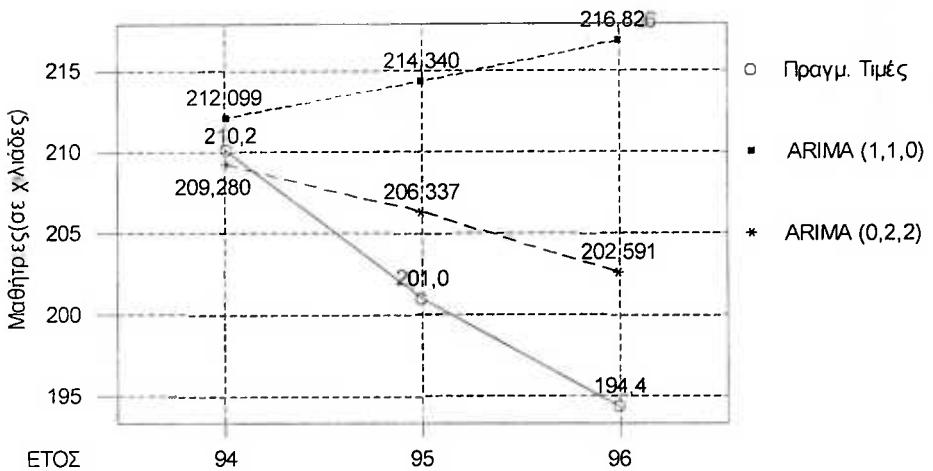
ΠΙΝΑΚΑΣ ΜΕΤΡΩΝ ΑΚΡΙΒΕΙΑΣ

	ARIMA(1,1,0)	ARIMA(0,2,2)
MAPE	5,82054	2,35646
MAD	12,5547	4,81607
MSE	228,156	32,1432

Πίνακας 5.6: Μέτρα ακριβείας των υποδειγμάτων ARIMA



ΠΡΑΓΜΑΤΙΚΕΣ ΤΙΜΕΣ ΚΑΙ ΠΡΟΒΛΕΨΕΙΣ ΤΩΝ ΥΠΟΔΕΙΓΜΑΤΩΝ ARIMA



ΓΡΑΦΗΜΑ 5.19: Προβλέψεις και πραγματικές τιμές των υποδειγμάτων ARIMA

Το πρώτο υπόδειγμα δίνει εκτιμήσεις μεγαλύτερες από τις πραγματικές και αποκλίνει σημαντικά αφού προβλέπει αύξηση του πληθυσμού των κοριτσιών ενώ αυτός παρουσίασε μείωση. Αντίθετα το δεύτερο υπόδειγμα δείχνει να ακολουθεί την πτωτική τάση των μαθητριών, δίνει μειούμενες τιμές, δεν δίνει όμως την μείωση στο πραγματικό της μέγεθος. Η απόκλιση του όμως από τις πραγματικές τιμές δεν μπορούμε να την θεωρήσουμε σημαντική.

5.5 ΣΥΜΠΕΡΑΣΜΑΤΑ ΓΙΑ ΤΙΣ ΜΑΘΗΤΡΙΕΣ

	MAPE	MAD	MSE
$Y_t = b_0 + b_1 t + b_2 t^2$	1,49592	3,02443	11,4491
$\log Y_t = b_0 + b_1 t + b_2 t^2$	1,75062	3,52076	19,5818
ΔΕΕ	3,73122	7,80575	95,3068
ARIMA(1,1,0)	5,82054	12,5542	228,156
ARIMA (0,2,2)	2,3564	4,81607	32,1432

Πίνακας 5.7: Συγκριτικά αποτελέσματα μέτρων ακριβείας των υποδειγμάτων για τις μαθήτριες

Αυτή τη φορά ένα απλό μοντέλο δίνει καλύτερα αποτελέσματα από πολυπλοκότερα υποδείγματα. Η απλή παραβολική τάση διηλαδή προβλέπει καλύτερα μέσα σε ένα διάστημα τριών ετών από ένα υπόδειγμα ARIMA αν και η διαφορά τους δεν είναι τόσο μεγάλη όσο στο προηγούμενο κεφάλαιο. Το εντυπωσιακό είναι ότι και το δεύτερο καλύτερο είναι και αυτό υπόδειγμα τάσης. Τα χειρότερα αποτελέσματα τα

παίρνουμε από το υπόδειγμα ARIMA (1,1,0) γεγονός που ίσως δείχνει ότι η σειρά δεν είναι αρκετά στάσιμη ώστε να εφαρμόσουμε τις απαιτούμενες μεθόδους.

5.6 ΑΝΑΛΥΣΗ ΑΡΡΕΝΩΝ ΜΑΘΗΤΩΝ ΚΑΤΩΤΕΡΗΣ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ ΕΚΠΑΙΔΕΥΣΗΣ-ΤΑΣΗ

Πρώτα θα εξασφαλίσουμε την τυχαιότητα της χρονολογικής σειράς για να μπορέσουμε να εξασφαλίσουμε την απρόσκοπτη ανάλυση της..

$$H_0: \text{Τα δεδομένα τυχαία}$$

$$H_1: \text{Τα δεδομένα μη τυχαία}$$

Median of LS-Males

Median of LS-Males = 232,50
MTB > Runs 232,5 'LS-Males'.

Runs Test: LS-Males

```
LS-Males
K = 232,5000
The observed number of runs = 5
The expected number of runs = 12,9167
11 Observations above K 13 below
The test is significant at 0,0009
```

Η χρονολογική σειρά των μαθητών του γυμνασίου παρουσιάζει αυξητική τάση από το 1970 μέχρι το 1978, όπου προφανώς λόγω της παιδαγωγικής μεταρρύθμισης 1976-1977 είχαμε και την μεγαλύτερη αύξηση στον αριθμό των μαθητών. Από την στιγμή εκείνη και μετά αρχίζει μια ραγδαία πτώση.

Τα επικρατέστερα υποδείγματα για την τάση της χρονολογικής σειράς είναι τα ακόλουθα:

A) Regression Analysis: LS-Males versus Time; Time**2

The regression equation is
LS-Males = 195 + 6,61 Time - 0,228 Time**2

Predictor	Coef	SE Coef	T	P
Constant	194,799	5,883	33,11	0,000
Time	6,607	1,084	6,09	0,000
Time**2	-0,22808	0,04211	-5,42	0,000

S = 8,818 R-Sq = 66,4% R-Sq(adj) = 63,2%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	3223,3	1611,7	20,73	0,000
Residual Error	21	1632,9	77,8		
Total	23	4856,2			

Durbin-Watson statistic = 1,04

B)Regression Analysis: LogLSM versus Time; Time**2

The regression equation is
 $\text{LogLSM} = 5,28 + 0,0292 \text{ Time} - 0,00100 \text{ Time}^{**2}$

Predictor	Coef	SE Coef	T	P
Constant	5,27900	0,02449	215,52	0,000
Time	0,029231	0,004514	6,48	0,000
Time**2	-0,0010024	0,0001753	-5,72	0,000

S = 0,03671 R-Sq = 69,4% R-Sq(adj) = 66,4%

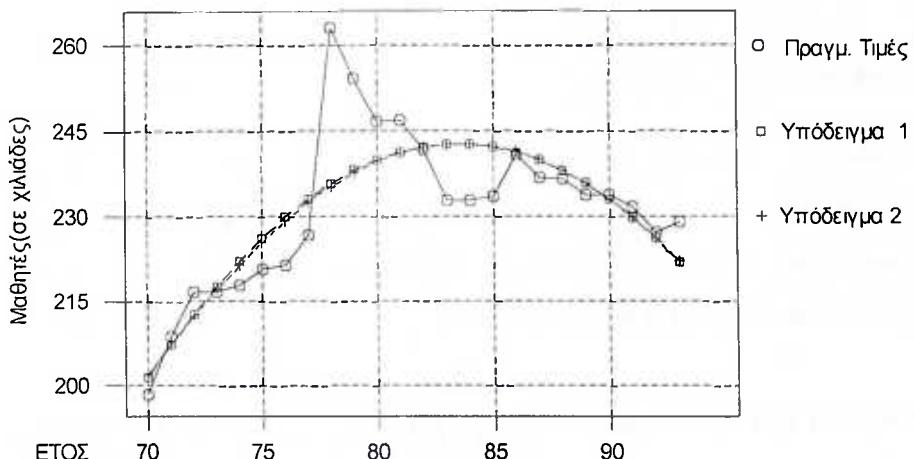
Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	0,064072	0,032036	23,77	0,000
Residual Error	21	0,028303	0,001348		
Total	23	0,092375			

Durbin-Watson statistic = 1,03

** R-sq = 65,96% και R-Sq(adj) = 62,72% συγκρίσιμες τιμές

ΑΝΑΛΥΣΗ ΤΑΣΗΣ ΓΙΑ ΤΟΥΣ ΜΑΟΗΤΕΣ ΓΥΜΝΑΣΙΟΥ



ΓΡΑΦΗΜΑ 5.20: Προσαρμογή των δύο υποδειγμάτων της τάσης για τους άρρενες μαθητές κατώτερις δευτεροβάθμιας εκπαίδευσης

Τα δύο υποδείγματα είναι στατιστικά σημαντικά και οι συντελεστές τους διαφορετικοί από το μηδέν. Το μεν πρώτο δευτεροβάθμιο εξηγεί 66,4% της μεταβλητής των δεδομένων ενώ το δεύτερο εξηγεί 65,96%. Ωστόσο όταν στον συντελεστή προσδιορισμού λαμβάνονται υπόψη ο αριθμός παρατηρήσεων σε σχέση με τον αριθμό ανεξάρτητων μεταβλητών όπως ο Προσαρμοσμένος Συντελεστής Προσδιορισμού R^2_{adj} , το πρώτο μοντέλο υστερεί σε προσαρμογή από το δεύτερο με τιμές 63,2% και 64,41% αντίστοιχα. Η διαφορά των δύο υποδειγμάτων δεν είναι

αρκετά μεγάλη κάτι που φαίνεται και στο Γράφημα 5.20 όπου οι τιμές τους σχεδόν συμπίπτουν. Προβληματική είναι για μια ακόμη φορά η τιμή Durbin-Watson..

Οι προβλέψεις των δύο υποδειγμάτων είναι επίσης πολύ ικανοποιητικές. Είναι η καλύτερη πρόβλεψη που έχουν πραγματοποιήσει τα μέχρι τώρα υποδείγματα.

ΠΙΝΑΚΑΣ ΠΡΑΓΜΑΤΙΚΩΝ ΤΙΜΩΝ ΚΑΙ ΠΡΟΒΛΕΠΟΜΕΝΩΝ

	Πραγμ. Τιμές	1 ^ο Υπόδειγμα $Y_t = b_0 + b_1 t + b_2 t^2$	2 ^ο Υπόδειγμα $\log Y_t = b_0 + b_1 t + b_2 t^2$
1994/95	222,9	221,995	222,107
1995/96	220,9	217,426	217,733
1996/97	212,7	212,401	213,019

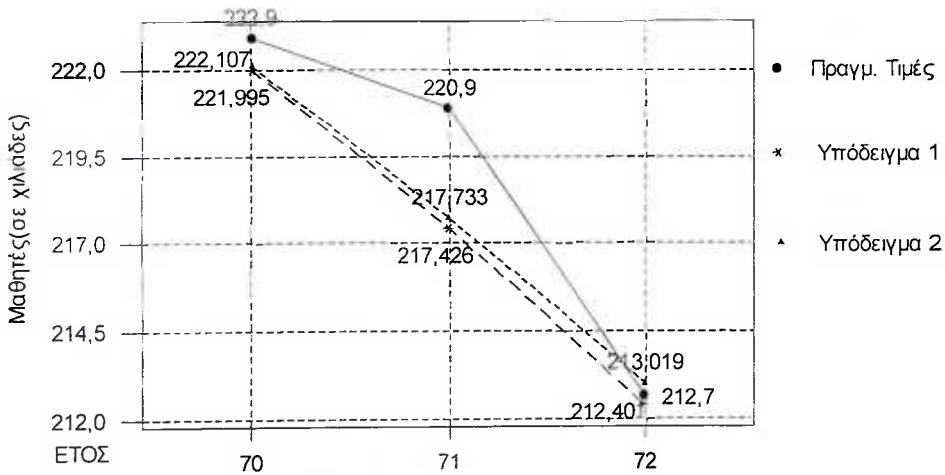
Πίνακας 5.8: Πραγματικές τιμές και προβλέψεις των υποδειγμάτων της τάσης

ΠΙΝΑΚΑΣ ΜΕΤΡΩΝ ΑΚΡΙΒΕΙΑΣ

	1 ^ο Υπόδειγμα $Y_t = b_0 + b_1 t + b_2 t^2$	2 ^ο Υπόδειγμα $\log Y_t = b_0 + b_1 t + b_2 t^2$
MAPE	0,715506	0,653701
MAD	1,55955	1,42618
MSD	4.32631	3,58606

Πίνακας 5.9: Μέτρα ακριβείας των υποδειγμάτων της τάσης

ΠΡΑΓΜΑΤΙΚΕΣ ΤΙΜΕΣ ΚΑΙ ΠΡΟΒΛΕΨΕΙΣ ΥΠΟΔΕΙΓΜΑΤΩΝ ΤΑΣΗΣ



ΓΡΑΦΗΜΑ 5.21: Πραγματικές τιμές και προβλέψεις των δύο υποδειγμάτων της τάσης για τους άρρενες μαθητές κατώτερης δευτεροβάθμιας εκπαίδευσης

Το λογαριθμικό υπόδειγμα δίνει ελαφρώς καλύτερες προβλέψεις από το απλό δευτεροβάθμιο. Είναι γενικά εντυπωσιακά μικρές και οι τρεις τιμές των μέτρων

αικρίβειας. Είναι οι μικρότερες από όσες έχουμε συναντήσει μέχρι τώρα στα υποδείγματα που εξετάσαμε.

5.7 ΕΚΘΕΤΙΚΗ ΕΞΟΜΑΛΥΝΣΗ ΜΑΘΗΤΩΝ

Προσπαθώντας να ελαχιστοποιήσουμε το μέσο τετραγωνικό σφάλμα (MSE) καταλήγουμε σε τιμή του $\alpha = 1$ και τιμή γ που προσεγγίζει το 0. Αυτό σημαίνει ότι ίσως το πιο κατάλληλο μοντέλο να μην αποδίδεται από την διπλή εκθετική εξομάλυνση αλλά άπό την απλή που εφαρμόζεται σε σειρές που δεν εμφανίζουν τάση. Παρακάτω αποτυπώνονται τα αποτελέσματα της εφαρμογής και των δύο μεθόδων.

Double Exponential Smoothing

```
Data      LS-Males
Length    24,0000
NMissing 0

Smoothing Constants
Alpha (level): 1,000000
Gamma (trend): 0,000001

Accuracy Measures
MAPE: 2,6128
MAD: 6,0687
MSD: 99,5500
```

Single Exponential Smoothing

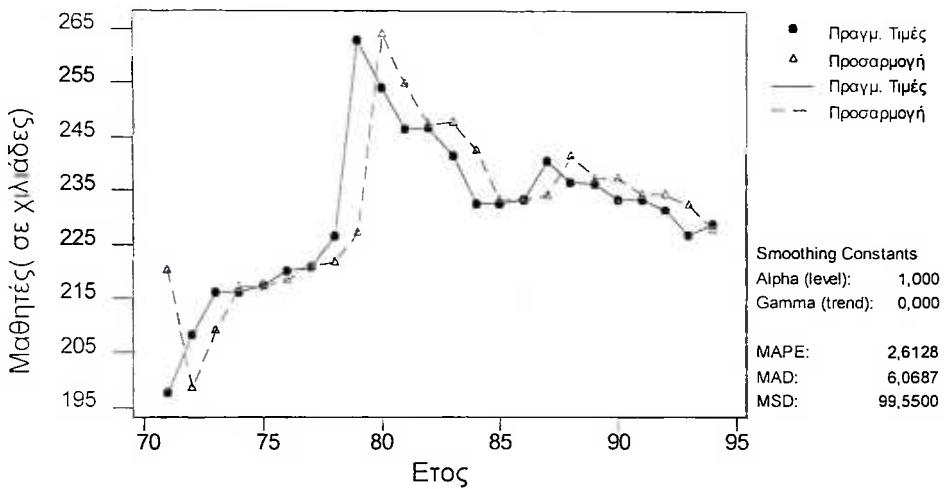
```
Data      LS-Males
Length    24,0000
NMissing 0

Smoothing Constant
Alpha: 1

Accuracy Measures
MAPE: 2,4010
MAD: 5,6194
MSD: 88,7596
```

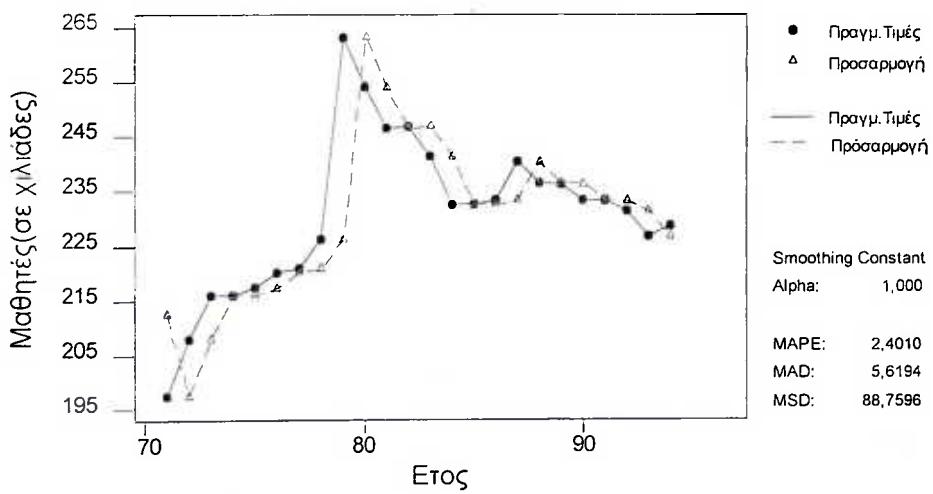
Η απλή εκθετική εξομάλυνση εμφανίζει καλύτερη προσαρμογή από την διπλή παρά το γεγονός ότι η σειρά εμφανίζει τάση. Στην πραγματικότητα και στις δύο περιπτώσεις αυτό που γίνεται είναι να εμφανίζεται σαν τιμή που προσαρμόζει το υπόδειγμα στην παρατήρηση Y_{t+1} η ακριβώς προηγούμενη παρατηρούμενη τιμή δηλ ηY_t . Αυτό στην περίπτωση της Διπλής εκθετικής εξομάλυνσης γίνεται με μία μικρή διαφορά επειδή η τιμή γ δεν είναι ακριβώς μηδεν, ενώ στην απλή εκθετική εξομάλυνση πραγματοποιείται επακριβώς.:

ΔΙΠΛΗ ΕΚΘΕΤΙΚΗ ΕΞΟΜΑΛΥΝΣΗ ΓΙΑ ΜΑΘΗΤΕΣ ΓΥΜΝΑΣΙΟΥ



ΓΡΑΦΗΜΑ 5.22: Πραγματικές και προσαρμοσμένες τιμές της διπλής εκθετικής εξομάλυνσης για τους μαθητές

ΑΠΛΗ ΕΚΘΕΤΙΚΗ ΕΞΟΜΑΛΥΝΣΗ ΓΙΑ ΜΑΘΗΤΕΣ ΓΥΜΝΑΣΙΟΥ



ΓΡΑΦΗΜΑ 5.23: Πραγματικές και προσαρμοσμένες τιμές της απλής εκθετικής εξομάλυνσης

Ο έλεγχος θα συνεχιστεί με την ικανότητα τους να δίνουν προβλέψεις :

ΠΙΝΑΚΑΣ ΠΡΑΓΜΑΤΙΚΩΝ ΤΙΜΩΝ ΚΑΙ ΠΡΟΒΛΕΠΟΜΕΝΩΝ

	Πραγμ. Τιμές	Απλή Εκθετική Εξομάλυνση	Διπλή Εκθετική Εξομάλυνση
1994/95	222,9	226,8	227,705
1995/96	220,9	226,8	228,610
1996/97	212,7	226,8	229,515

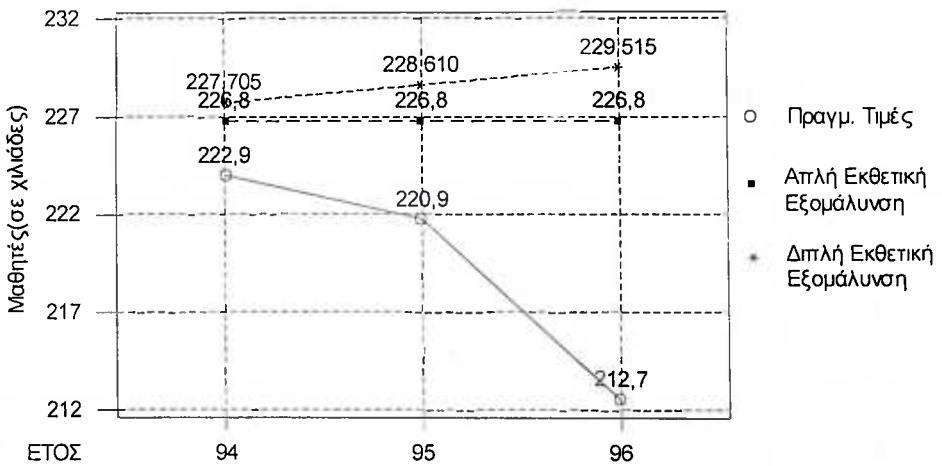
Πίνακας 5.10: Πραγματικές τιμές και προβλέψεις των υποδειγμάτων της εκθετικής εξομάλυνσης

ΠΙΝΑΚΑΣ ΜΕΤΡΩΝ ΑΚΡΙΒΕΙΑΣ

	Απλή Εκθετική Εξομάλυνση	Διπλή Εκθετική Εξομάλυνση
MAPE	3,51264	4,26965
MAD	7,96667	9,77657
MSD	82,9433	121,757

Πίνακας 5.11: Μέτρα ακριβείας των υποδειγμάτων της εκθετικής εξομάλυνσης

ΠΡΟΒΛΕΨΗ ΚΑΙ ΠΡΑΓΜΑΤΙΚΕΣ ΤΙΜΕΣ



ΓΡΑΦΗΜΑ 5.24: Πραγματικές τιμές και προβλέψεις της Απλής και Διπλής Εκθετικής Εξομάλυνσης για τους άρρενες μαθητές από το 1993/94 – 1996/97

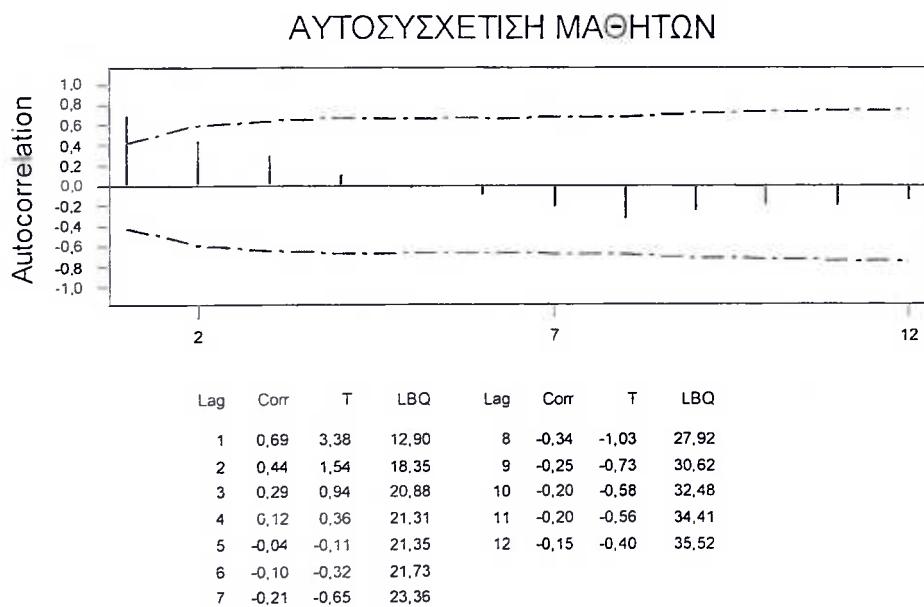
Η απλή εκθετική εξομάλυνση δίνει τρεις τιμές σταθερές αφού δεν λαμβάνει υπόψη την υπάρχουσα τάση. Η διπλή εκθετική εξομάλυνση αντιθέτως δίνει τιμές που αυξάνονται. Αυτό συμβαίνει γιατί ενώ για να καθορίσει το επίπεδο δίνει περισσότερη βαρύτητα στα πιο πρόσφατα δεδομένα αφού η τιμή του $\alpha=1$, ενώ για να καθορίσει την τάση δίνει βαρύτητα στις πιο παλιές παρατηρήσεις. Έτσι τα αποτελέσματα της

απλής εκθετικής εξομάλυνσης εμφανίζονται να έχουν μεγαλύτερη ακρίβεια από αυτά της διπλής εκθετικής εξομάλυνσης.

5.8 ΥΠΟΔΕΙΓΜΑΤΑ ARIMA ΓΙΑ ΤΟΥΣ ΑΡΡΕΝΕΣ ΜΑΘΗΤΕΣ

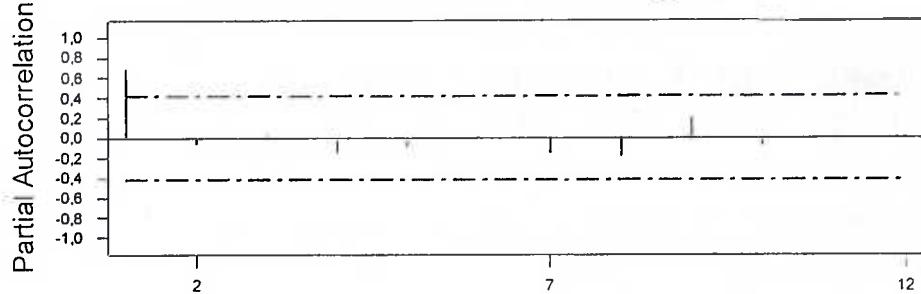
Θα ελέγξουμε αρχικά την στασιμότητα της σειράς αν και περιμένουμε ότι θα χρειαστεί να πάρουμε τουλάχιστον τις πρώτες διαφορές αφού υπάρχει δευτεροβάθμια τάση. Εξετάζουμε τα αυτοσυσχέτιση και την μερική αυτοσυσχέτιση της χρονολογικής σειράς.

Τα γραφήματα 5.25 και 5.26 επιβεβαιώνουν το γεγονός ότι η σειρά δεν είναι στάσιμη, κάτι που ήταν αναμενόμενο. Κατά συνέπεια θα πάρουμε τις πρώτες διαφορές και θα επανεξετάσουμε την σειρά..



ΓΡΑΦΗΜΑ 5.25: Συντελεστές αυτοσυσχέτισης για τους άρρενες μαθητές δευτεροβάθμιας εκπαίδευσης

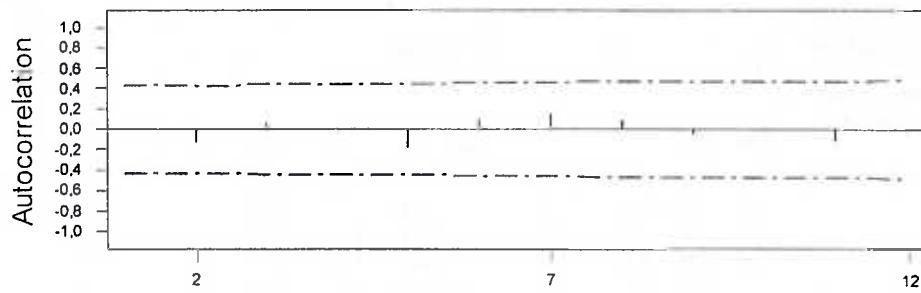
ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΜΑΘΗΤΩΝ



Lag	PAC	T	Lag	PAC	T
1	0,69	3,38	8	-0,19	-0,91
2	-0,07	-0,35	9	0,21	1,02
3	0,03	0,16	10	-0,09	-0,42
4	-0,15	-0,75	11	-0,05	-0,24
5	-0,10	-0,48	12	-0,02	-0,08
6	-0,01	-0,03			
7	-0,17	-0,83			

ΓΡΑΦΗΜΑ 5.26: Μερική αυτοσυσχέτιση χρονολογικής σειράς αρρένων μαθητών κατώτερης δευτεροβάθμιας εκπαίδευσης

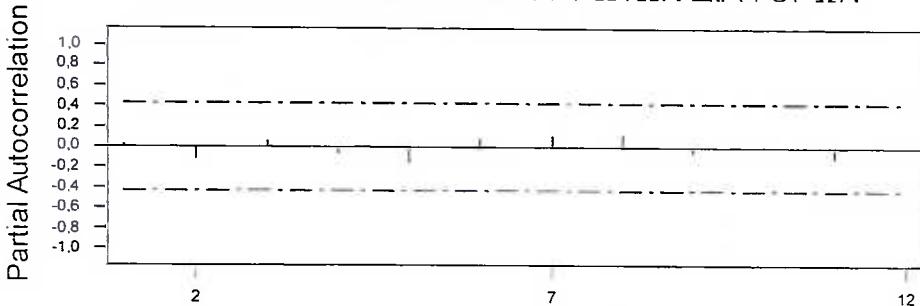
ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΠΡΩΤΩΝ ΔΙΑΦΟΡΩΝ



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0,02	0,08	0,01	8	0,09	0,39	3,24
2	-0,14	-0,67	0,53	9	-0,06	-0,25	3,37
3	0,06	0,28	0,63	10	-0,01	-0,06	3,38
4	-0,05	-0,22	0,70	11	-0,12	-0,53	4,09
5	-0,19	-0,87	1,81	12	-0,05	-0,21	4,21
6	0,10	0,44	2,13				
7	0,15	0,68	2,94				

ΓΡΑΦΗΜΑ 5.27: Αυτοσυσχέτιση πρώτων διαφορών των αρρένων μαθητών

ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΠΡΩΤΩΝ ΔΙΑΦΟΡΩΝ



Lag	PAC	T	Lag	PAC	T
1	0,02	0,08	8	0,13	0,62
2	-0,14	-0,57	9	-0,06	-0,29
3	0,06	0,31	10	-0,02	-0,10
4	-0,07	-0,34	11	-0,11	-0,54
5	-0,17	-0,82	12	0,01	0,02
6	0,09	0,43			
7	0,11	0,52			

ΓΡΑΦΗΜΑ 5.28: Μερική αυτοσυγχέτιση πρώτων διαφορών αρρένων μαθητών

Τα γραφήματα δείχνουν ότι οι πρώτες διαφορές είναι αρκετές για να μετατρέψουν τη σειρά σε στάσιμη. Τα γραφήματα αυτοσυγχέτισης και μερικής αυτοσυγχέτισης παρουσιάζουν την χρονολογική σειρά που προκύπτει από τις πρώτες διαφορές να είναι ασυγχέτιστη έτσι το υπόδειγμα που θα ταίριαζε σ' αυτή την περίπτωση θα ήταν ARIMA(0,1,0) ή και ακόμα χωρίς να πάρουμε διαφορές θα μπορούσαμε να θεωρήσουμε ότι ταιριάζει ένα μοντέλο της μορφής ARIMA(1,0,0).

Τα δύο υποδείγματα που δίνουν τελικά τα καλύτερα αποτελέσματα όσον αφορά την προσαρμογή είναι το ARIMA(1,0,0) και το ARIMA(1,1,1) και παρουσιάζονται στην επόμενη σελίδα.

Το δεύτερο υπόδειγμα εμφανίζει καλύτερα αποτελέσματα. Έχει μικρότερο τυπικό σφάλμα της παλινδρόμησης(SE of regression) 6,15 έναντι 8,16, άθροισμα τετραγώνων των καταλοίπων(Sum square resid) 719,8 έναντι 1396,9 και τις τιμές Akaike και Schwartz. Τέλος οι τιμές των Durbin-Watson και για τα δύο υποδείγματα δείχνουν ότι δεν υπάρχει πρόβλημα αυτοσυγχέτισης των καταλοίπων, κάτι που

φαίνεται και στα Γραφήματα 5.29, 5.30 5.31, και 5.32 των αυτοσυσχετίσεων και των μερικών αυτοσυσχετίσεων όπου όλοι οι συντελεστές είναι στατιστικά μηδενικοί.

:

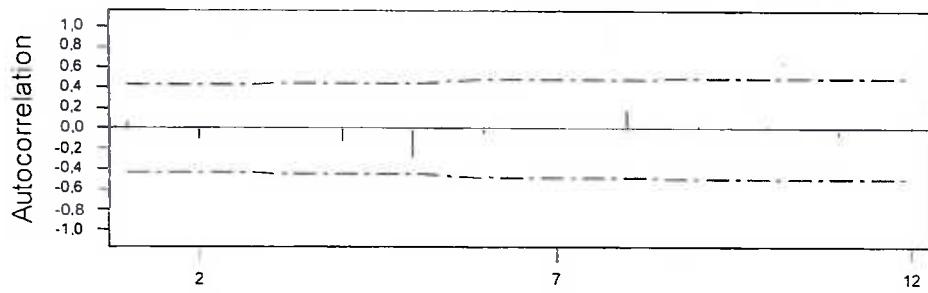
Dependent Variable: LSM
 Method: Least Squares
 Date: 04/22/04 Time: 11:34
 Sample(adjusted): 1971 1993
 Included observations: 23 after adjusting endpoints
 Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	235.2704	5.721647	41.11934	0.0000
AR(1)	0.689677	0.117097	5.889815	0.0000
R-squared	0.622912	Mean dependent var	232.2652	
Adjusted R-squared	0.604955	S.D. dependent var	12.97661	
S.E. of regression	8.156132	Akaike info criterion	7.118358	
Sum squared resid	1396.972	Schwarz criterion	7.217097	
Log likelihood	-79.86112	F-statistic	34.68993	
Durbin-Watson stat	1.875427	Prob(F-statistic)	0.000008	
Inverted AR Roots	.69			

Dependent Variable: D(LSM)
 Method: Least Squares
 Date: 04/14/04 Time: 00:56
 Sample(adjusted): 1972 1993
 Included observations: 22 after adjusting endpoints
 Convergence achieved after 80 iterations
 Backcast: OFF (Roots of MA process too large for backcast)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.903409	2.937529	-0.307540	0.7618
AR(1)	0.694428	0.119713	5.800789	0.0000
MA(1)	-1.564053	0.305549	-5.118821	0.0001
R-squared	0.594876	Mean dependent var	0.936364	
Adjusted R-squared	0.552232	S.D. dependent var	9.198061	
S.E. of regression	6.154927	Akaike info criterion	6.598507	
Sum squared resid	719.7794	Schwarz criterion	6.747285	
Log likelihood	-69.58357	F-statistic	13.94963	
Durbin-Watson stat	2.135904	Prob(F-statistic)	0.000187	
Inverted AR Roots	.69			
Inverted MA Roots	1.56			
Estimated MA process is noninvertible				

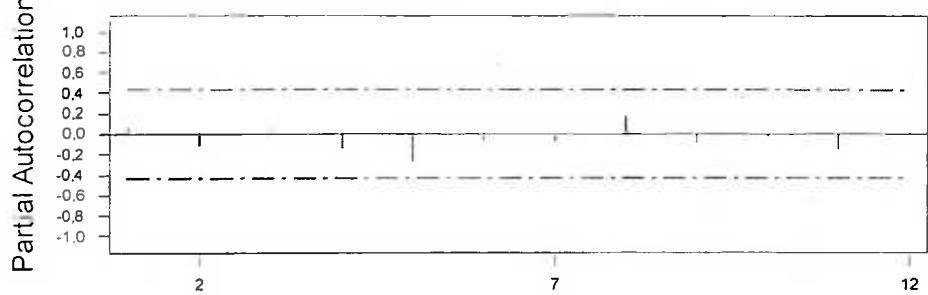
ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ ARIMA(1,0,0)



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0,06	0,30	0,10	8	0,18	0,76	5,26
2	-0,11	-0,55	0,46	9	0,01	0,06	5,27
3	0,01	0,06	0,46	10	0,02	0,09	5,29
4	-0,14	-0,64	1,02	11	-0,07	-0,31	5,55
5	-0,30	-1,40	3,92	12	-0,02	-0,09	5,58
6	-0,06	-0,26	4,04				
7	-0,01	-0,03	4,05				

ΓΡΑΦΗΜΑ 5.29: Αυτοσυσχέτιση καταλοίπων υποδείγματος ARIMA (1,0,0) για τους μαθητές κατώτερης δευτεροβάθμιας εκπαίδευσης

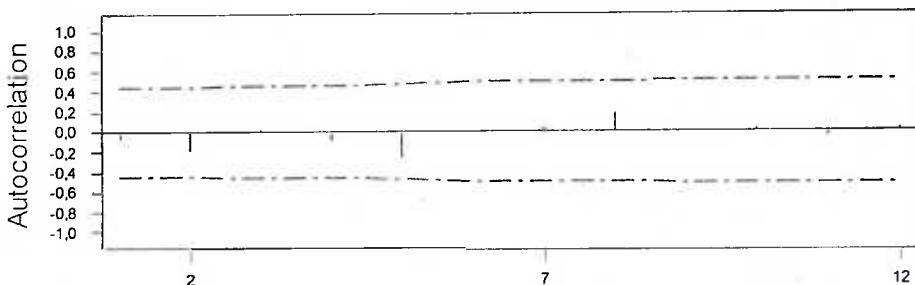
ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ ARIMA(1,0,0)



Lag	PAC	T	Lag	PAC	T
1	0,06	0,30	8	0,17	0,80
2	-0,12	-0,57	9	-0,10	-0,47
3	0,03	0,14	10	-0,04	-0,18
4	-0,16	-0,75	11	-0,16	-0,79
5	-0,29	-1,38	12	-0,00	-0,01
6	-0,07	-0,36			
7	-0,08	-0,40			

ΓΡΑΦΗΜΑ 5.30: Μερική αυτοσυσχέτιση καταλοίπων υποδείγματος ARIMA (1,0,0)

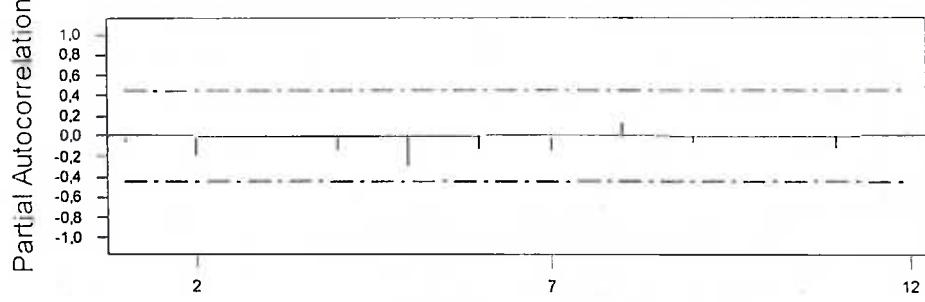
ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ ARIMA(1,1,1)



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	-0,08	-0,35	0,14	8	0,18	0,77	5,07
2	-0,19	-0,90	1,13	9	-0,00	-0,01	5,07
3	0,03	0,11	1,14	10	0,02	0,09	5,10
4	-0,10	-0,47	1,47	11	-0,06	-0,24	5,27
5	-0,27	-1,21	3,77	12	0,02	0,10	5,30
6	0,00	0,01	3,77				
7	0,03	0,13	3,80				

ΓΡΑΦΗΜΑ 5.31: Αυτοσυσχέτιση καταλοίπων υποδείγματος ARIMA (1,1,1)

ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ ARIMA(1,1,1)



Lag	PAC	T	Lag	PAC	T
1	-0,08	-0,35	8	0,12	0,56
2	-0,20	-0,94	9	-0,08	-0,37
3	-0,01	-0,03	10	-0,02	-0,09
4	-0,15	-0,70	11	-0,13	-0,61
5	-0,31	-1,46	12	0,03	0,15
6	-0,14	-0,64			
7	-0,15	-0,68			

ΓΡΑΦΗΜΑ 5.32: Μερική αυτοσυσχέτιση καταλοίπων ARIMA(1,1,1)

ΠΙΝΑΚΑΣ ΠΡΑΓΜΑΤΙΚΩΝ ΤΙΜΩΝ ΚΑΙ ΠΡΟΒΛΕΠΟΜΕΝΩΝ

	Πραγμ. Τιμές	ARIMA(1,0,0)	ARIMA(1,1,1)
1994/95	222,9	230,739	227,705
1995/96	220,9	232,145	228,610
1996/97	212,7	233,115	229,515

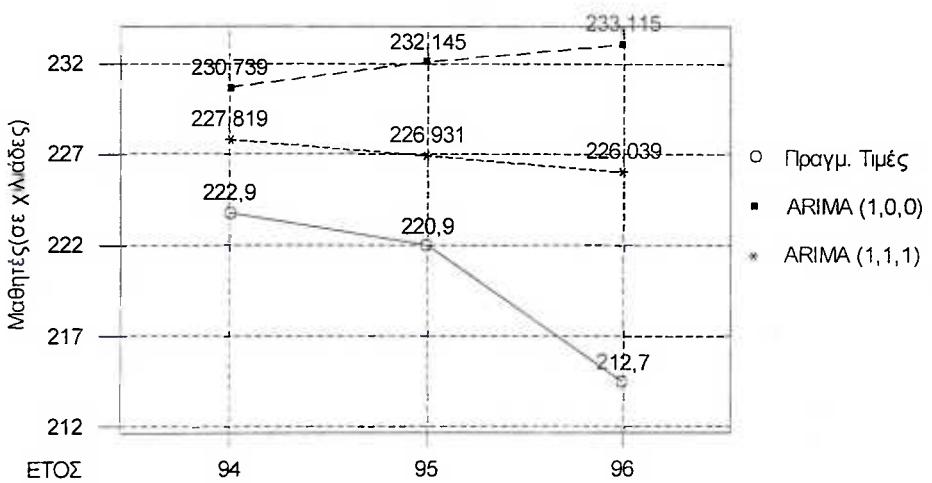
Πίνακας 5.12: Πραγματικές τιμές και προβλέψεις των υποδειγμάτων ARIMA

ΠΙΝΑΚΑΣ ΜΕΤΡΩΝ ΑΚΡΙΒΕΙΑΣ

	ARIMA(1,0,0)	ARIMA(1,1,1)
MAPE	5,66626	3,57271
MAD	13,1663	8,09643
MSD	201,558	79,4996

Πίνακας 5.13: Μέτρα ακριβείας των υποδειγμάτων ARIMA

ΠΡΟΒΛΕΨΗ ΚΑΙ ΠΡΑΓΜΑΤΙΚΕΣ ΤΙΜΕΣ



ΓΡΑΦΗΜΑ 5.33: Πραγματικές τιμές και προβλέψεις των υπόδειγμάτων ARIMA για τα έτη 1993/94 – 1996/97

Όπως και στην προσαρμογή, έτσι και στην πρόβλεψη το υπόδειγμα (1,1,1) δίνει καλύτερα αποτελέσματα ακολουθώντας μεν την πτωτική τάση, έχοντας ωστόσο “χάσει” το επίπεδο των πραγματικών τιμών. Έτσι έχουμε τιμές που είναι αυξημένες σε σχέση με τις πραγματικές. Το άλλο υπόδειγμα σφάλλει και στην εκτίμηση της τάσης των δεδομένων και στην εκτίμηση του επιπέδου τους. Έτσι δίνει τιμές πολύ πιο λανθασμένες από το υπόδειγμα (1,1,1) αλλά και από όλα τα υποδείγματα που

εξετάσαμε μέχρι τώρα. Ίσως έλλειψη στασιμότητας να είναι ο λόγος , αφού το υπόδειγμα (1,0,0) συμπεριφέρεται σαν να είναι στάσιμη η σειρά.

5.9 ΣΥΜΠΕΡΑΣΜΑΤΑ ΓΙΑ ΤΟΥΣ ΑΡΡΕΝΕΣ ΜΑΘΗΤΕΣ

	MAPE	MAD	MSE
$Y_t = b_0 + b_1 t + b_2 t^2$	0,715506	1,55955	4,32631
$\log Y_t = b_0 + b_1 t + b_2 t^2$	0,653701	1,42618	3,58606
Δ.Ε.Ε	3,51264	7,96667	82,9433
Α.Ε.Ε	4,26965	9,77657	121,757
ARIMA(1,0,0)	5,66626	13,1663	201,558
ARIMA(1,1,1)	3,57271	8,09643	79,4996

Πίνακας 5.14: Σύγκριση μέτρων ακριβείας των υποδειγμάτων για τους μαθητές

Για μια ακόμη φορά παρατηρούμε ότι τα υποδείγματα της τάσης δίνουν πολύ πιο αξιόπιστες προβλέψεις από τα άλλα τα πιο πολύπλοκα μοντέλα. Το αξιοπερίεργο είναι ότι η ακρίβεια και των δύο υποδειγμάτων στην περίπτωση που εξετάζουμε είναι η καλύτερη που έχουμε συναντήσει στα πιο ακριβή μοντέλα αφού τιμές της μέσης τετραγωνικής απόκλισης που να αγγίζουν το 4,3 και το 3,6 είναι πολύ μικρότερες από οποιεσδήποτε άλλες. Το ερώτημα βέβαια που τίθεται είναι κατά πόσο μπορούμε να βασιστούμε σε αυτά τα υποδείγματα για αξιόπιστες προβλέψεις αφού φαίνεται να έχουν συσχετισμένα κατάλοιπα.

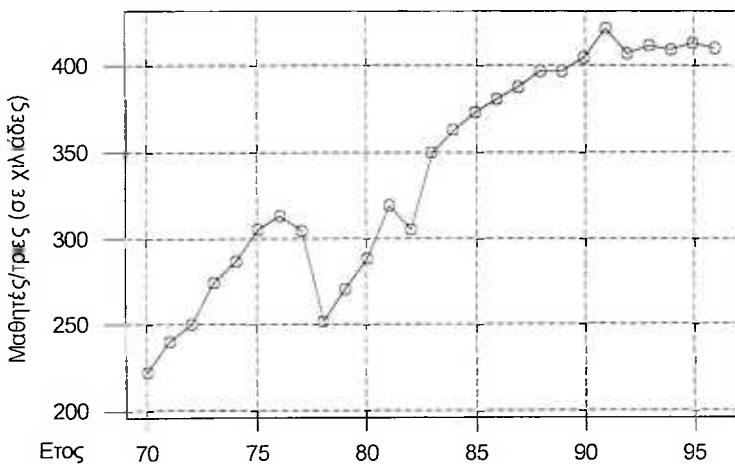
Τα υποδείγματα ARIMA στην περίπτωση των μαθητών του γυμνασίου δεν δίνουν πολύ καλύτερες λύσεις από τις εκθετικές εξομαλύνσεις και κυρίως την απλή εκθετική εξομάλυνση αφού υπερτερεί στο μέσο τετραγωνικό σφάλμα (MSD) αλλά υστερεί στο μέσο απόλυτο ποσοστιαίο σφάλμα(MAPE) και στο μέσο απόλυτο σφάλμα(MAD) χωρίς όμως πολύ μεγάλες διαφορές.

ΚΕΦΑΛΑΙΟ 6 :ΑΝΩΤΕΡΗ ΔΕΥΤΕΡΟΒΑΘΜΙΑ ΕΚΠΑΙΔΕΥΣΗ

6.1 ΓΕΝΙΚΗ ΕΙΚΟΝΑ ΑΝΩΤΕΡΗΣ ΔΕΥΤΕΡΟΒΑΘΜΙΑ ΕΚΠΑΙΔΕΥΣΗ

Η μορφή που εμφανίζει η χρονολογική σειρά των μαθητών και μαθητριών του Λυκείου, προκαλεί προβληματισμό. Ενώ η τάση είναι αυξητική μέχρι το 1976, με μία μικρή πτώση το 1977, το 1978 υπάρχει μια ραγδαία μείωση. Αυτό οφείλεται στην δημιουργία του Λυκείου η εγγραφή στο οποίο γινόταν μετά από εισιτήριες εξετάσεις ενώ στο εξατάξιο γυμνάσιο η μετάβαση από τον πρώτο κύκλο στον δεύτερο γινόταν χωρίς εξετάσεις.

ΜΑΘΗΤΕΣ ΛΥΚΕΙΟΥ 1970/71 - 1996/97

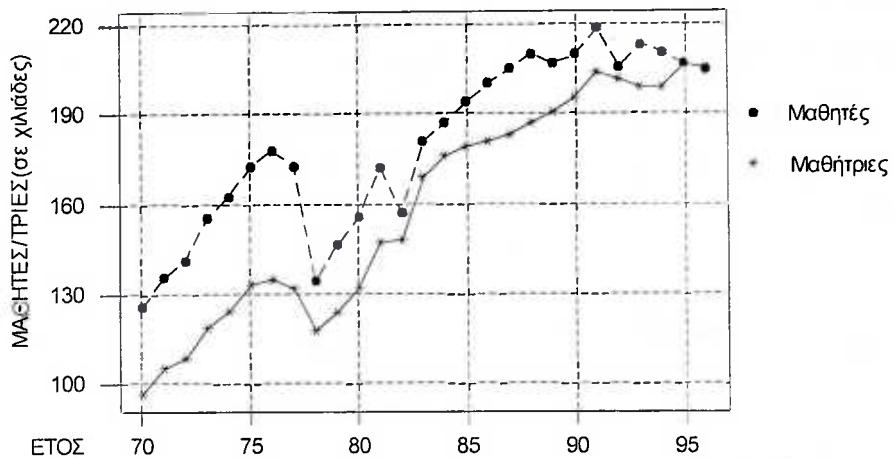


ΠΗΓΗ: Eurostat

ΓΡΑΦΗΜΑ 6.1: Αριθμός συνόλου μαθητών και μαθητριών ανώτερης δευτεροβάθμιας εκπαίδευσης από το 1970/71 – 1996/97

Την ίδια γενική εικόνα παρουσιάζουν και οι επιμέρους χρονολογικές σειρές που αποτελούν το σύνολο των μαθητών, όπως φαίνεται από το γράφημα 5.2. Οι μαθητές και οι μαθήτριες εμφανίζουν αυξητική τάση από το 1970/71 μέχρι το 1976/1977. Η αύξηση των μαθητών είναι μεγαλύτερη, ωστόσο, παρά των μαθητριών. Μεγαλύτερη είναι και η πτώση του αριθμού των μαθητών σε σχέση με τον αριθμό των κοριτσιών το έτος 1978/79. Ενώ στη συνέχεια αυξάνονται συγχρόνως από την επόμενη χρονιά μέχρι το 1996/97 οπότε δείχνουν να σταθεροποιούνται.

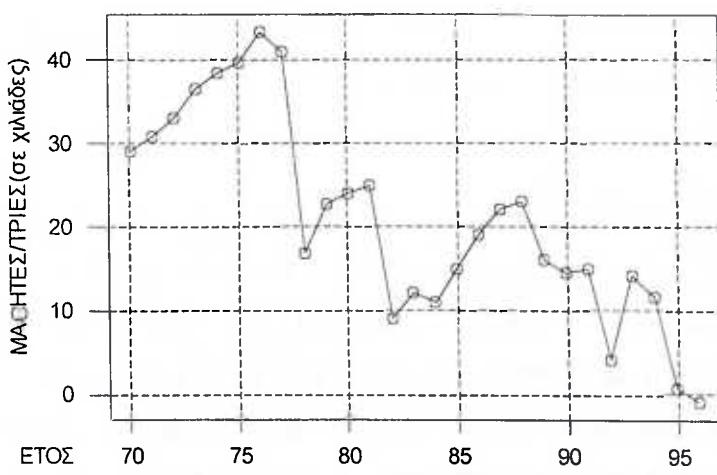
ΜΑΘΗΤΕΣ ΚΑΙ ΜΑΘΗΤΡΙΕΣ ΛΥΚΕΙΟΥ 1970/71 - 1996/1997



ΠΗΓΗ: Eurostat

ΓΡΑΦΗΜΑ 6.2: Εξέλιξη αριθμού μαθητών – μαθητριών για τα έτη 1970/71 – 1996/97

ΔΙΑΦΟΡΑ ΜΑΘΗΤΩΝ-ΜΑΘΗΤΡΙΩΝ ΛΥΚΕΙΟΥ



ΠΗΓΗ: Eurostat

ΓΡΑΦΗΜΑ 6.3: Εξέλιξη διαφοράς μαθητών – μαθητριών από το 1970/71 – 1996/97

Η διαφορά του αριθμού αγοριών και κοριτσιών εμφανίζει και αυτή μια σχετικά ασταθή αλλά πτωτική πορεία. Ξεκινάει από την τιμή των 29.100 ατόμων, αγγίζει την μέγιστη τιμή των 43600 το 1976/77 και παρουσιάζει μια ραγδαία πτώση το 1978/79 όπου μειώνεται στις 16,9 χιλιάδες και τελικά μετά από διαδοχικά σκαμπανεβάσματα καταλήγουμε στο έτος 1996/97 όπου υπερτερούν οι μαθήτριες των μαθητών κατά 900 περίπου άτομα. Φαίνεται ότι περισσότερα αγόρια αναζητούν διαφορετικές λύσεις μετά το γυμνάσιο σε σχέση με τα κορίτσια. Η προσπαθήσουμε να αποκαλύψουμε την μορφή της τάσης που κρύβει αυτή η χρονολογική σειρά.

Η τυχαιότητα των δεδομένων πρέπει να ελεγχθεί πριν την ανάλυση. Πραγματοποιείται πάλι ο έλεγχος:

H_0 : Τα δεδομένα τυχαία

H_1 : Τα δεδομένα μη τυχαία

Median of Diff

Median of Diff = 19,100
MTB > Runs 19,1 'Diff'.

Runs Test: Diff

Diff

```
K = 19,1000
The observed number of runs = 6
The expected number of runs = 14,4815
13 Observations above K 14 below
The test is significant at 0,0009
```

Η p-τιμή= 0,0009 απορρίπτει την μη τυχαιότητα των δεδομένων και για τον λόγο αυτό συνεχίζουμε την ανάλυση.

Regression Analysis: Diff versus log(Time); log(Time)**2

The regression equation is
Diff = 27,9 + 15,2 log(Time) - 6,76 log(Time)**2

Predictor	Coeff	SE Coef	T	P
Constant	27,900	5,965	4,68	0,000
log(Time)	15,208	6,426	2,37	0,026
log(Time)**2	-6,761	1,644	-4,11	0,000

S = 6,911 R-Sq = 70,9% R-Sq(adj) = 68,4%

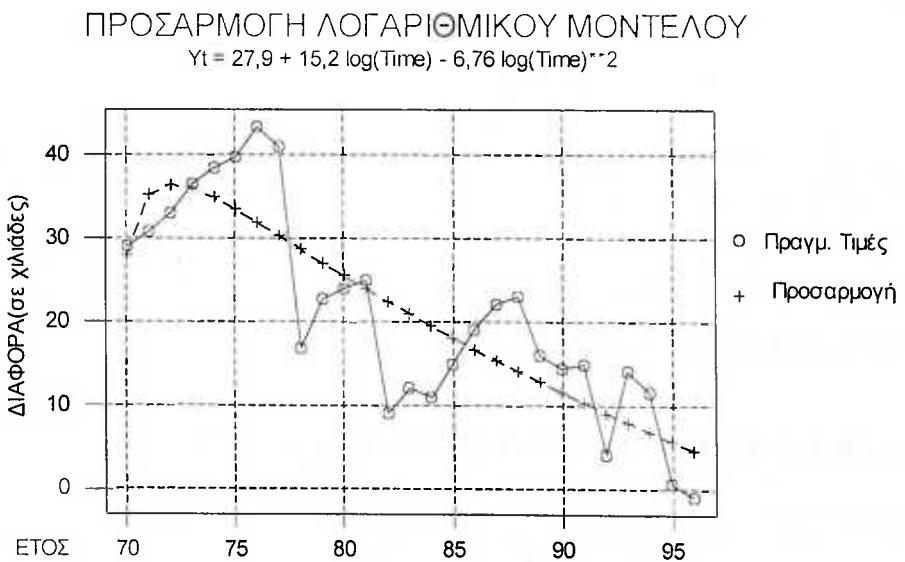
Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	2788,4	1394,2	29,19	0,000
Residual Error	24	1146,3	47,8		
Total	26	3934,7			

Durbin-Watson statistic = 1,16

Την καλύτερη προσαρμογή παρουσιάζει το παραπάνω λογαριθμικό υπόδειγμα, όπου έχουμε χρησιμοποιήσει τις λογαριθμικές τιμές του χρόνου και το τετράγωνο της σαν ανεξάρτητες μεταβλητές. Όπως είναι εμφανές και οι δύο συντελεστές τους είναι στατιστικά σημαντικοί αφού οι R -τιμές τους απορρίπτουν την υπόθεση της μηδενικότητας.

Επίσης ο έλεγχος που εξετάζει την σημαντικότητα του υπόδειγματος συνολικά με τον έλεγχο F , απορρίπτει την μηδενική υπόθεση. Οι συντελεστές προσδιορισμού R^2 και R^2_{adj} δείχνουν ότι το υπόδειγμα εξηγεί το 70,9% ή το 68,4% της συνολικής μεταβλητότητας και το τυπικό σφάλμα είναι της τιμής 6,911. Για την χρονική στιγμή $t=1$ που αντιστοιχεί στο έτος 1970/71 το υπόδειγμα δίνει διαφορά 27,9 χιλιάδων ατόμων η οποία αυξάνει σταδιακά σε 36,4 χιλιάδες το 1972/73 για να αρχίσει η αποκλιμάκωση από εκείνη την χρονιά και να καταλήξει να μειωθεί στις 4,6 χιλιάδες το 1996/97. Όλα παρουσιάζονται στο Γράφημα 6.4



ΓΡΑΦΗΜΑ 6.4: Προσαρμογή του λογαριθμικού υπόδειγματος στην διαφορά μαθητών μαθητριών ανώτερης δευτεροβάθμιας εκπαίδευσης

6.2 ΜΑΘΗΤΡΙΕΣ ΑΝΩΤΕΡΗΣ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ-ΤΑΣΗ

Για να μελετηθεί η χρονολογική σειρά, αναγκαία είναι η μη τυχαιότητα της. Ήα πραγματοποιήσουμε τον συνηθισμένο έλεγχο:

H_0 : Τα δεδομένα τυχαία

H_1 : Τα δεδομένα μη τυχαία

Median of US-Female

Median of US-Female = 147,50

Runs Test: US-Female

US-Femal

K = 147,5000

The observed number of runs = 2
 The expected number of runs = 13,0000
 12 Observations above K 12 below
 The test is significant at 0,0000

Αφού η p- τιμή είναι σχεδόν μηδέν και κατά συνέπεια τα δεδομένα δεν είναι τυχαία μπορούμε να προχωρήσουμε στην ανάλυση μας.

Για την μορφή της τάσης καλύτερη προσαρμογή παρουσιάζουν δύο μοντέλα. Το ένα είναι το απλό γραμμικό της μορφής $Y_t = a + bt$ και το απλό εκθετικό της μορφής $Y_t = a^b t$ το οποίο μετασχηματίζεται στο $\log Y_t = a' + b' t$. Το πρώτο παρουσιάζει συντελεστές προσδιορισμού 94,1% και 93,8% αντίστοιχα, ενώ το δεύτερο οι συντελεστές προσδιορισμού, που προέρχονται από μετατροπή των λογαριθμικών τιμών του μοντέλου σε κανονικές τιμές, είναι 93,3% και 93% αντίστοιχα. Τα πλήρη αποτελέσματα είναι:

A) Regression Analysis: US-Female versus Time

The regression equation is
 US-Female = 93,9 + 4,77 Time

Predictor	Coef	SE Coef	T	P
Constant	93,903	3,645	25,76	0,000
Time	4,7737	0,2551	18,71	0,000

S = 8,651 R-Sq = 94,1% R-Sq(adj) = 93,8%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	26207	26207	350,14	0,000
Residual Error	22	1647	75		
Total	23	27853			

Durbin-Watson statistic = 0,62

B) Regression Analysis: US-Female versus Time

The regression equation is
 $\log(\text{US-Female}) = 2,00205 + 0,0138513 \text{ Time}$

Predictor	Coef	SE Coef	T	P
Constant	2,00205	0,01149	174,18	0,000
Time	0,0138513	0,0008044	17,22	0,000

S = 0,02728 R-Sq = 93,1% R-Sq(adj) = 92,8%

S = 0,0272792 R-Sq = 93,1 % R-Sq(adj) = 92,8 %

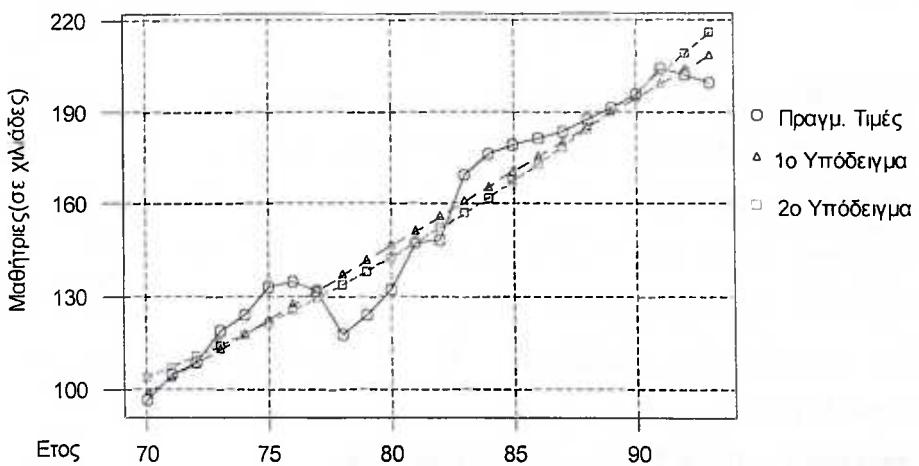
Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0,220636	0,220636	296,493	0,000
Error	22	0,016371	0,000744		
Total	23	0,237008			

Durbin-Watson statistic = 0,64

adjusted R_sqr = 93,3356
adjusted R_sqr(adj) = 93,0327

ΠΡΟΣΑΡΜΟΓΗ ΥΠΟΔΕΙΓΜΑΤΩΝ ΤΑΣΗΣ



ΓΡΑΦΗΜΑ 6.5: Προσαρμογή υποδειγμάτων τάσης για τις μαθήτριες ανώτερης δευτεροβάθμιας εκπαίδευσης

Και τα δύο υποδείγματα έχουν τους συντελεστές τους στατιστικά σημαντικούς αλλά και ο έλεγχος F επιβεβαιώνει την σημαντικότητα της παλινδρόμησης ως προς το ποσοστό της μεταβλητότητας που εξηγούν και τα δύο μοντέλα. Δεδομένου ότι το απλό γραμμικό μοντέλο εμφανίζει ελαφρά καλύτερη προσαρμογή στα δεδομένα, αξίζει να συνεχίσουμε στον έλεγχο της πρόβλεψης.

ΠΙΝΑΚΑΣ ΠΡΑΓΜΑΤΙΚΩΝ ΤΙΜΩΝ ΚΑΙ ΠΡΟΒΛΕΠΟΜΕΝΩΝ

Έτος	Πραγμ. Τιμές	1 ^ο Υπόδειγμα $Y_t = b_0 + b_1 t + b_2 t^2$	2 ^ο Υπόδειγμα $\log Y_t = b_0 + b_1 t + b_2 t^2$
1994/95	199,0	213,247	223,015
1995/96	206,3	218,020	230,242
1996/97	205,7	222,794	237,704

Πίνακας 6.1: Πραγματικές τιμές και προβλέψεις των υποδειγμάτων της τάσης μαθητριών

ΠΙΝΑΚΑΣ ΜΕΤΡΩΝ ΑΚΡΙΒΕΙΑΣ

	1 ^ο Υπόδειγμα $Y_t = b_0 + b_1 t + b_2 t^2$	2 ^ο Υπόδειγμα $\log Y_t = b_0 + b_1 t + b_2 t^2$
MAPE	5,30620	11,5436
MAD	11,4324	26,6536
MSD	135,739	724,728

Πίνακας 6.2: Μέτρα ακριβείας των υποδειγμάτων της τάσης

Στην περίπτωση αυτή και τα δύο υποδείγματα αδυνατούν να μας δώσουν πολύ ακριβή αποτελέσματα. Ειδικά το 2^ο υπόδειγμα απέχει πάρα πολύ από τις πραγματικές τιμές. Αυτό ίσως να οφείλεται στο γεγονός ότι την χρονική στιγμή που ξεκινούν οι προβλέψεις έχει αρχίσει να σταθεροποιείται ο αριθμός των μαθητριών και ίσως να ξεκινάει και μια πιωτική πορεία η οποία αν δεν εμφανιστεί αμέσως θα είναι σίγουρη μετά από μικρό χρονικό διάστημα όταν πλέον θα επιδρά και στην ανώτερη δευτεροβάθμια εκπαίδευση η πληθυσμιακή μείωση. Επιπλέον προβληματισμός μπορεί να προέλθει από τις χαμηλές τιμές των συντελεστών Durbin-Watson.

6.3 ΔΙΠΛΗ ΕΚΘΕΤΙΚΗ ΕΞΟΜΑΛΥΝΣΗ ΜΑΘΗΤΡΙΩΝ

Το υπόδειγμα που καταλήγουμε προσπαθώντας να ελαχιστοποιήσουμε το MSE δίνει τιμή του $\alpha=1$ και γ σχεδόν μηδέν. Όσον αφορά διλαδή το επίπεδο της χρονολογικής σειράς μεγαλύτερο βάρος δίνεται σε παλιότερες παρατηρήσεις ενώ όσον αφορά την τάση το βάρος δίνεται στις πιο πρόσφατες

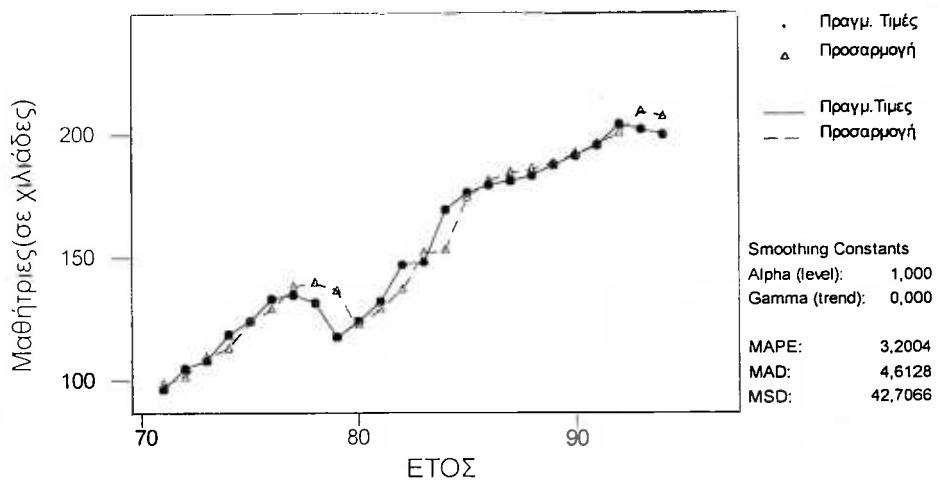
Double Exponential Smoothing

Data US-Female
Length 24,0000
NMissing 0

Smoothing Constants
Alpha (level): 1,0000
Gamma (trend): 0,0001

Accuracy Measures
MAPE: 3,2004
MAD: 4,6128
MSD: 42,7066

ΔΙΠΛΗ ΕΚΘΕΤΙΚΗ ΕΞΟΜΑΛΥΝΣΗ ΓΙΑ ΜΑΘΗΤΡΙΕΣ ΛΥΚΕΙΟΥ



ΓΡΑΦΗΜΑ 6.6: Προσαρμογή της Διπλής Εκθετικής Εξομάλυνσης για τις μαθήτριες ανώτερης δευτεροβάθμιας εκπαίδευση από το 1970/71 – 1993/94

Οι προβλέψεις που δίνει το υπόδειγμα αυτό και η ακρίβεια τους αποτυπώνονται παρακάτω.

ΠΙΝΑΚΑΣ ΠΡΑΓΜΑΤΙΚΩΝ ΤΙΜΩΝ ΚΑΙ ΠΡΟΒΛΕΠΟΜΕΝΩΝ

Ετος	Πραγμ. Τιμές	Διπλή Εκθετική Εξομάλυνση
1994/95	199,0	206,374
1995/96	206,3	211,147
1996/97	205,7	215,921

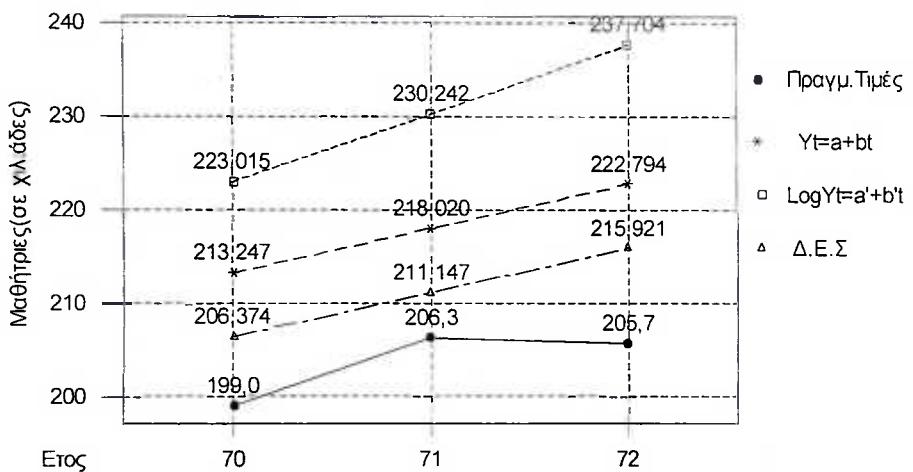
Πίνακας 6.3: Πραγματικές τιμές και προβλέψεις των υποδειγμάτων Διπλής Εκθετικής Εξομάλυνσης

ΠΙΝΑΚΑΣ ΜΕΤΡΩΝ ΑΚΡΙΒΕΙΑΣ

	Δ.Ε.Ε
MAPE	3,53399
MAD	7,48039
MSD	60,7744

Πίνακας 6.4: Μέτρα ακριβείας των υποδειγμάτων της Δ.Ε.Ε

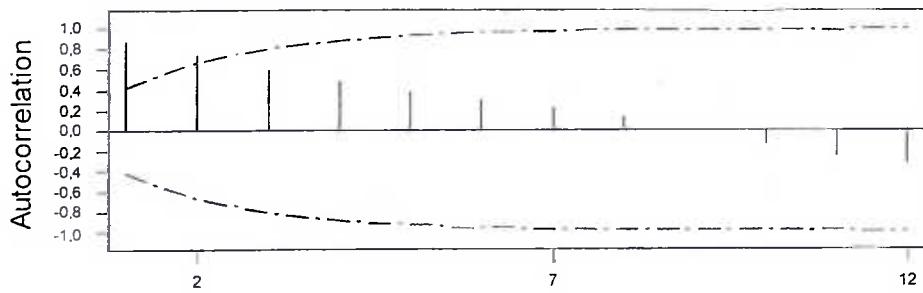
ΠΡΑΓΜΑΤΙΚΕΣ ΤΙΜΕΣ ΚΑΙ ΠΡΟΒΛΕΨΕΙΣ



ΓΡΑΦΗΜΑ 6.7: Πραγματικές τιμές και Προβλέψεις υποδειγμάτων $Y_t = a + b t$, $\log Y_t = a' + b' t$ και Διπλής Εκθετικής Εξομάλυνσης

6.4 ΥΠΟΔΕΙΓΜΑΤΑ ARIMA ΜΑΘΗΤΡΙΩΝ

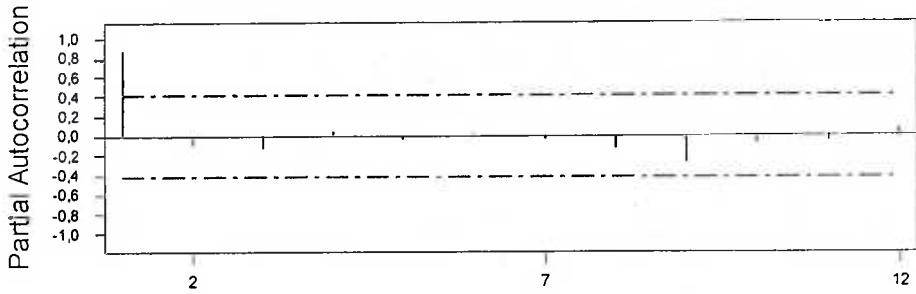
ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΜΑΘΗΤΡΙΩΝ Α.Δ.Ε



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0,88	4,30	20,93	8	0,13	0,27	64,97
2	0,75	2,30	36,79	9	-0,01	-0,03	64,98
3	0,60	1,54	47,54	10	-0,15	-0,32	65,99
4	0,49	1,14	54,90	11	-0,28	-0,58	69,63
5	0,38	0,84	59,57	12	-0,34	-0,71	75,71
6	0,30	0,64	62,61				
7	0,22	0,46	64,35				

ΓΡΑΦΗΜΑ 6.8: Συντελεστές αυτοσυσχέτισης για τις μαθήτριες ανώτερης δευτεροβάθμιας εκπαίδευσης

ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΜΑΘΗΤΡΙΩΝ Α.Δ.Ε

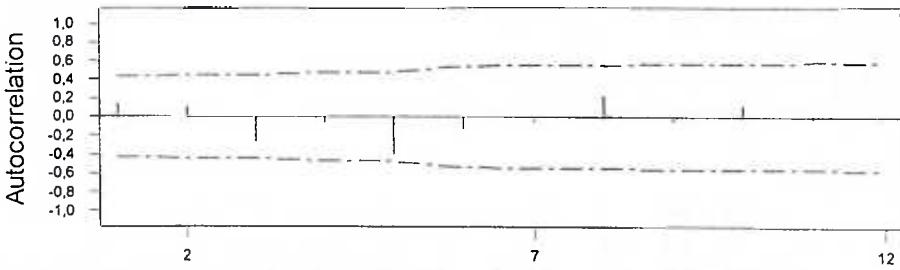


Lag	PAC	T	Lag	PAC	T
1	0.88	4.30	8	-0.14	-0.69
2	-0.10	-0.50	9	-0.29	-1.41
3	-0.14	-0.70	10	-0.09	-0.43
4	0.05	0.24	11	-0.07	-0.36
5	-0.05	-0.23	12	0.08	0.39
6	0.02	0.10			
7	-0.05	-0.24			

ΓΡΑΦΗΜΑ 6.9: Συντελεστές μερικής αυτοσυσχέτισης μαθητριών

Τα γραφήματα αυτοσυσχέτισης της χρονοσειράς παρουσιάζει μια σχετικά αργή, φθίνουσα πορεία προς το μηδέν ενώ η μερική αυτοσυσχέτιση έχει μόνο τον συντελεστή πρώτης τάξης διαφορετικό του μηδέν. Αυτό μπορεί να σημαίνει ότι χρειάζεται να πάρουμε τις πρώτες διαφορές ή ότι το κατάλληλο υπόδειγμα είναι το ARIMA(1,0,0). Έχει ενδιαφέρον να δούμε και τις μερικές συσχετίσεις και αυτοσυσχετίσεις των πρώτων διαφορών.

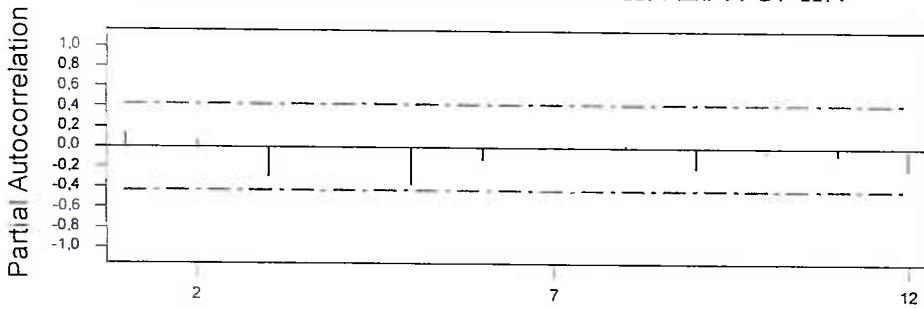
ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΠΡΩΤΩΝ ΔΙΑΦΟΡΩΝ



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0.14	0.65	0.48	8	0.22	0.83	11.22
2	0.09	0.44	0.71	9	-0.06	-0.21	11.36
3	-0.28	-1.31	2.98	10	0.11	0.41	11.91
4	-0.07	-0.30	3.12	11	-0.03	-0.10	11.94
5	-0.42	-1.80	8.63	12	-0.05	-0.18	12.06
6	-0.13	-0.51	9.24				
7	-0.06	-0.23	9.37				

ΓΡΑΦΗΜΑ 6.10: Συντελεστές αυτοσυσχέτισης πρώτων διαφορών για τις μαθήτριες

ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΠΡΩΤΩΝ ΔΙΑΦΟΡΩΝ



Lag	PAC	T	Lag	PAC	T
1	0,14	0,65	8	0,02	0,09
2	0,08	0,36	9	-0,22	-1,04
3	-0,31	-1,49	10	-0,06	-0,28
4	0,00	0,02	11	-0,09	-0,43
5	-0,39	-1,87	12	-0,23	-1,12
6	-0,13	-0,63			
7	0,00	0,01			

ΓΡΑΦΗΜΑ 6.11: Συντελεστές μερικής αυτοσυσχέτισης πρώτων διαφορών

Τα παραπάνω γραφήματα έχουν μετατρέψει σε στάσιμη την χρονολογική σειρά. Στην πραγματικότητα έχει μετατραπεί σε λευκό θόρυβο. Μόνο η αυτοσυσχέτιση και η μερική αυτοσυσχέτιση 5^{ου} βαθμού παρουσιάζει υψηλότερη τιμή, γεγονός που ίσως χρειαστεί να εκμεταλλευτούμε.

Τα δύο υποδείγματα που προκρίθηκαν έχοντας χαμηλότερα Akaike και Schwarz κριτήρια είναι τα παρακάτω:

Dependent Variable: USF
 Method: Least Squares
 Date: 04/27/04 Time: 16:26
 Sample(adjusted): 1971 1993
 Included observations: 23 after adjusting endpoints
 Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	254.1167	105.2729	2.413885	0.0250
AR(1)	0.956573	0.042405	22.55811	0.0000
R-squared	0.960368	Mean dependent var	156.0478	
Adjusted R-squared	0.958480	S.D. dependent var	33.35626	
S.E. of regression	6.796803	Akaike info criterion	6.753723	
Sum squared resid	970.1271	Schwarz criterion	6.852462	
Log likelihood	-75.66781	F-statistic	508.8684	
Durbin-Watson stat	1.675755	Prob(F-statistic)	0.000000	
Inverted AR Roots	.96			

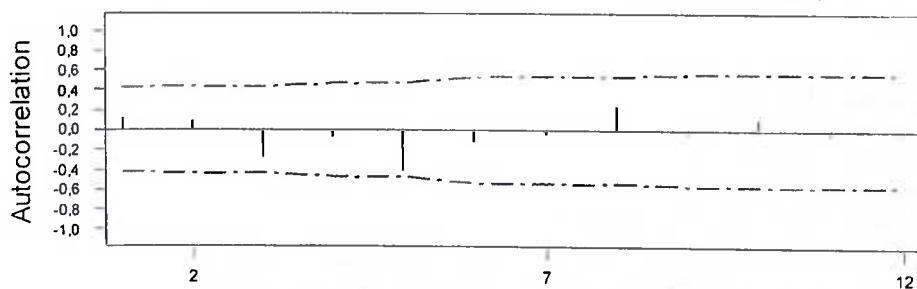
Dependent Variable: D(USF)
 Method: Least Squares
 Date: 05/02/04 Time: 23:37
 Sample(adjusted): 1971 1993
 Included observations: 23 after adjusting endpoints
 Convergence achieved after 12 iterations
 Backcast: 1966 1970

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.121639	0.835582	4.932659	0.0001
MA(5)	-0.881296	0.039421	-22.35600	0.0000
R-squared	0.366585	Mean dependent var	4.452174	
Adjusted R-squared	0.336423	S.D. dependent var	6.804335	
S.E. of regression	5.542829	Akaike info criterion	6.345828	
Sum squared resid	645.1821	Schwarz criterion	6.444567	
Log likelihood	-70.97703	F-statistic	12.15363	
Durbin-Watson stat	2.074363	Prob(F-statistic)	0.002203	
Inverted MA Roots	.98 -.79+.57i -.79-.57i	.30+.93i .30-.93i	-.79-.57i	

Τα υποδείγματα αυτά είναι τα ARIMA(1,0,0) και ARIMA(0,1,5) όπου όμως μόνο ο πέμπτος όρος των MA περιλαμβάνεται. Το δεύτερο υπόδειγμα έχει σαφώς καλύτερα αποτελέσματα όσον αφορά την προσαρμογή.

Οι αυτοσυσχετίσεις και μερικές αυτοσυσχετίσεις των καταλοίπων και των δύο υποδειγμάτων επιβεβαιώνουν την καλή τους προσαρμογή αφού σε όλα τα γραφήματα, οι συντελεστές είναι στατιστικά ίσοι με το μηδέν κάτι που επιβεβαιώνεται επίσης και από τις τιμές των Durbin-Watson που είναι 1,68 και 2,07 αντίστοιχα.

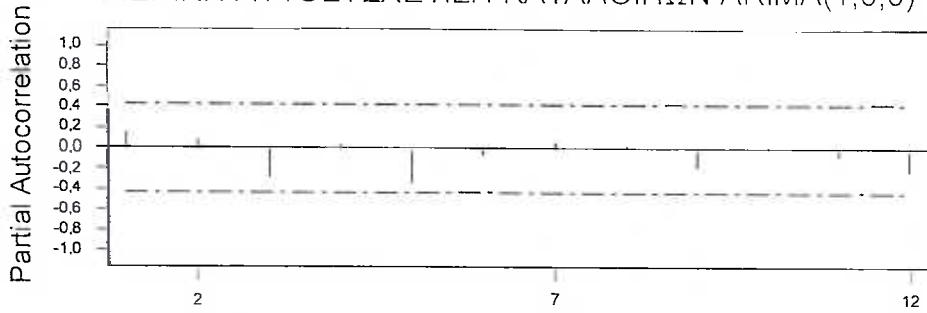
ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ ARIMA(1,0,0)



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	0.13	0.64	0.46	8	0.25	0.96	12.28
2	0.09	0.45	0.71	9	-0.06	-0.22	12.42
3	-0.29	-1.38	3.19	10	0.11	0.39	12.93
4	-0.07	-0.31	3.34	11	-0.04	-0.16	13.02
5	-0.42	-1.86	9.17	12	-0.04	-0.16	13.11
6	-0.13	-0.50	9.75				
7	-0.05	-0.19	9.84				

ΓΡΑΦΗΜΑ 6.12: Συντελεστές αυτοσυσχέτισης καταλοίπων υποδειγμάτος ARIMA (1,0,0)

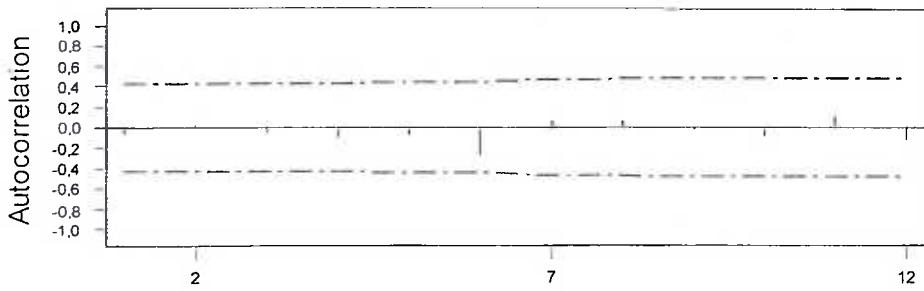
ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ ARIMA(1,0,0)



Lag	PAC	T	Lag	PAC	T
1	0,15	0,72	8	0,02	0,08
2	0,09	0,44	9	-0,20	-0,95
3	-0,30	-1,42	10	-0,02	-0,09
4	0,03	0,14	11	-0,09	-0,44
5	-0,35	-1,69	12	-0,23	-1,10
6	-0,08	-0,36			
7	0,05	0,25			

ΓΡΑΦΗΜΑ 6.13: Μερική αυτοσυσχέτιση καταλοίπων υποδείγματος ARIMA(1,0,0)

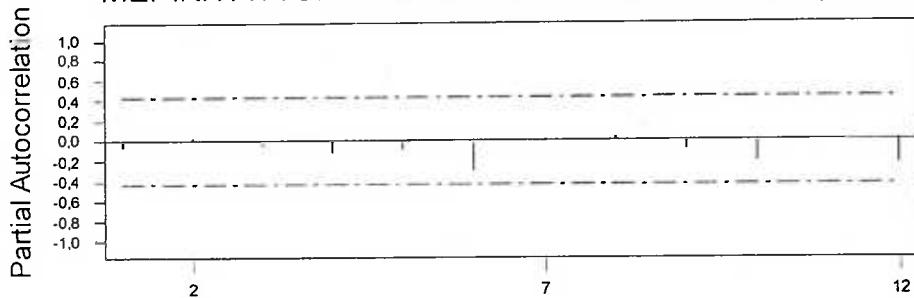
ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ ARIMA(0,1,5)



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	-0,08	-0,36	0,15	8	0,06	0,27	3,79
2	0,02	0,10	0,16	9	-0,03	-0,12	3,82
3	-0,06	-0,29	0,27	10	-0,10	-0,45	4,29
4	-0,11	-0,52	0,64	11	0,10	0,43	4,78
5	-0,09	-0,40	0,87	12	-0,13	-0,57	5,71
6	-0,28	-1,31	3,53				
7	0,06	0,24	3,64				

ΓΡΑΦΗΜΑ 6.14: Συντελεστές συσχέτισης καταλοίπων υποδείγματος ARIMA (0,1,5)

ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ ARIMA(0,1,5)



Lag	PAC	T	Lag	PAC	T
1	-0.08	-0.36	8	0.04	0.18
2	0.01	0.07	9	-0.09	-0.46
3	-0.06	-0.28	10	-0.22	-1.07
4	-0.12	-0.58	11	0.01	0.04
5	-0.10	-0.50	12	-0.25	-1.19
6	-0.31	-1.48			
7	-0.02	-0.12			

ΓΡΑΦΗΜΑ 6.15: Μερική αυτοσυσχέτιση των καταλοίπων του υποδείγματος ARIMA(0,1,5)

ΠΙΝΑΚΑΣ ΠΡΑΓΜΑΤΙΚΩΝ ΤΙΜΩΝ ΚΑΙ ΠΡΟΒΛΕΠΟΜΕΝΩΝ

	Πραγμ. Τιμές	ARIMA(1,0,0)	ARIMA(0,1,5)
1994/95	199,0	201,489	201,549
1995/96	206,3	203,775	204,141
1996/97	205,7	205,961	202,795

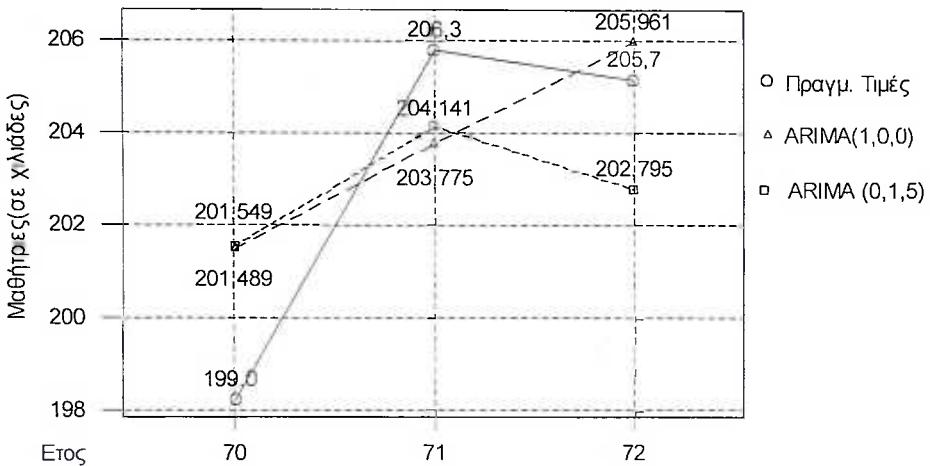
Πίνακας 6.5: Πραγματικές τιμές και προβλέψεις των υποδειγμάτων ARIMA μαθητριών

ΠΙΝΑΚΑΣ ΜΕΤΡΩΝ ΑΚΡΙΒΕΙΑΣ

	ARIMA(1,0,0)	ARIMA(0,1,5)
MAPE	0,867113	1,24656
MAD	1,75847	2,53767
MSD	4,21385	6,53256

Πίνακας 6.6: Μέτρα ακριβείας των υποδειγμάτων ARIMA

ΠΡΑΓΜΑΤΙΚΕΣ ΤΙΜΕΣ ΚΑΙ ΠΡΟΒΛΕΨΕΙΣ



ΓΡΑΦΗΜΑ 6.16: Προβλέψεις των υποδειγμάτων ARIMA και πραγματικές τιμές των μαθητριών ανώτερης δευτεροβάθμιας εκπαίδευσης

Τα δύο μοντέλα έχουν πολύ καλή πρόβλεψη για τα έτη που εξετάζουμε. Το ARIMA(1,0,0) έχει ελαφρώς καλύτερα αποτελέσματα από το ARIMA(0,1,5) αλλά το δεύτερο μοντέλο δείχνει να ακολουθεί καλύτερα τις αλλαγές της τάσης. Ενώ δηλαδή το πρώτο μοντέλο δίνει τιμές πιο κοντά στις πραγματικές παρατηρώντας το γράφημα 6.16 βλέπει κανείς ότι δίνει διαρκώς αυξανόμενες τιμές, ενώ το δεύτερο ακολουθεί την φορά των πραγματικών τιμών, αφού μετά την αύξηση που εμφανίζει η δεύτερη τιμή, η τρίτη παρουσιάζει πτώση όπως και στις πραγματικές τιμές.

6.5 ΣΥΜΠΕΡΑΣΜΑΤΑ ΓΙΑ ΤΙΣ ΜΑΘΗΤΡΙΕΣ

	MAPE	MAD	MSE
$Y_t = b_0 + b_1 t + b_2 t^2$	5,30620	11,4324	135,739
$\log Y_t = b_0 + b_1 t + b_2 t^2$	11,5436	26,6536	724,728
ΔΕΕ	3,53399	7,48039	60,7744
ARIMA(1,0,0)	0,867113	1,75847	4,21385
ARIMA (0,1,5)	1,24656	2,53767	6,53256

Πίνακας 6.7: Συγκριτικά αποτελέσματα υποδειγμάτων για τις μαθήτριες

Είναι πρώτη φορά στις μέχρι τώρα αναλύσεις των χρονολογικών σειρών των μαθητών και μαθητριών στις οποίες τα υποδείγματα της τάσης δίνουν τις χειρότερες εκτιμήσεις. Οι καλύτερες δίνονται από τα μοντέλα ARIMA και η εκθετική

εξομάλυνση δίνει σταθερά μέτριες εκτιμήσεις δηλαδή ανάμεσα στις καλύτερες και στις χειρότερες.

6.6 ΤΑΣΗ APPENΩΝ ΜΑΘΗΤΩΝ

Για το καλύτερο υπόδειγμα τάσης των μαθητών υπάρχουν 2 υποψήφια. Το απλό γραμμικό και το εκθετικό

A) Regression Analysis: US-Male versus Time

The regression equation is
US-Male = 132 + 3.59 Time

Predictor	Coef	SE Coef	T	P
Constant	131.908	5.527	23.87	0.000
Time	3.5857	0.3868	9.27	0.000

S = 13.12 R-Sq = 79.6% R-Sq(adj) = 78.7%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	14786	14786	85.92	0.000
Residual Error	22	3786	172		
Total	23	18572			

Durbin-Watson statistic = 0.91

B) Regression Analysis: Log(USM) versus Time

The regression equation is
Log(USM) = 2.13 + 0.00896 Time

Predictor	Coef	SE Coef	T	P
Constant	2.12972	0.01483	143.60	0.000
Time	0.008959	0.001038	8.63	0.000

S = 0.03520 R-Sq = 77.2% R-Sq(adj) = 76.2%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	0.092293	0.092293	74.50	0.000
Residual Error	22	0.027256	0.001239		
Total	23	0.119549			

Durbin-Watson statistic = 0.94

Data Display

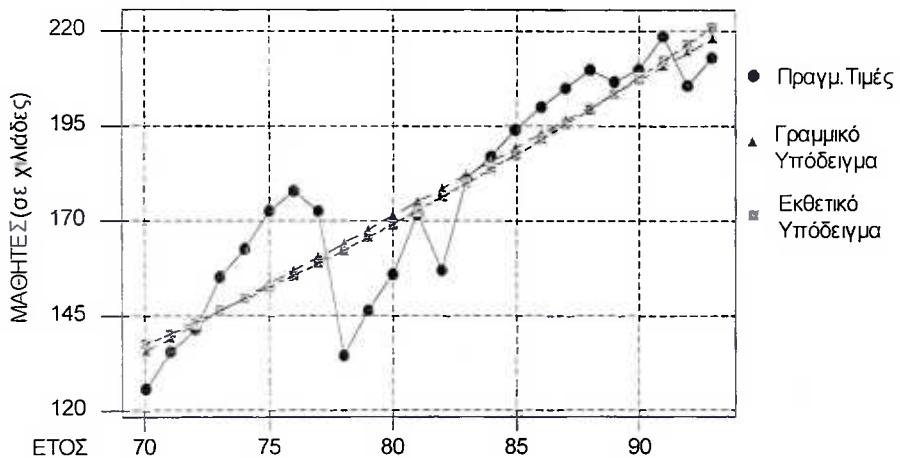
adupted R_sqr = 79.6965
adupted R_sqr(adj) = 78.7736

Κάθε ένα από τα υποδείγματα παρουσιάζει πλεονεκτήματα και μειονεκτήματα. Με μια πρώτη ματιά βλέπουμε ότι οι συντελεστές και των τριών υποδειγμάτων είναι στατιστικά σημαντικοί καθώς επίσης και τα F-tests απορρίπτουν την υπόθεση μη σημαντικής παλινδρόμησης. Και τα δύο έχουν μικρό συντελεστή Durbin-Watson. Η προσαρμογή όμως των υποδειγμάτων διαφέρει ελαφρώς.

	ΓΡΑΜΜΙΚΟ	ΕΚΘΕΤΙΚΟ
R^2	79,6%	79,7%
R^2_{adj}	78,7%	78,8%

Πίνακας 6.8: Συντελεστές προσδιορισμού των υποδειγμάτων τάσης

ΤΑΣΗ ΜΑΘΗΤΩΝ ΑΝΩΤΕΡΗΣ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ



ΓΡΑΦΗΜΑ 6.17: Προσαρμογή υποδειγμάτων τάσης στους μαθητές ανώτερης δευτεροβάθμιας εκπαίδευσης

Η διαφορά των υποδειγμάτων είναι στην ικανότητα να προβλέπουν τα τρία επόμενα χρόνια και κατά συνέπεια να ακολουθήσουν την μελλοντική τάση. Πιο συγκεκριμένα όπως φαίνεται από τους παρακάτω πίνακες η χειρότερη πρόβλεψη προέρχεται από το υπόδειγμα που προσαρμόζεται καλύτερα στα δεδομένα, δηλαδή το εκθετικό. Η καλύτερη πρόβλεψη προέρχεται από το τρίτο υπόδειγμα που έχει και την χειρότερη συγκριτικά προσαρμογή.

ΠΙΝΑΚΑΣ ΠΡΑΓΜΑΤΙΚΩΝ ΤΙΜΩΝ ΚΑΙ ΠΡΟΒΛΕΠΟΜΕΝΩΝ

Ετος	Πραγμ. Τιμές	1 ^ο Υπόδειγμα $Y_t = b_0 + b_1 t$	2 ^ο Υπόδειγμα $\log Y_t = b_0 + b_1 t$
1994/95	210.6	221.550	225.778
1995/96	207.0	225.136	230.483
1996/97	204.8	228.722	235.287

Πίνακας 6.9: Πραγματικές τιμές και προβλέψεις των υποδειγμάτων της τάσης μαθητών

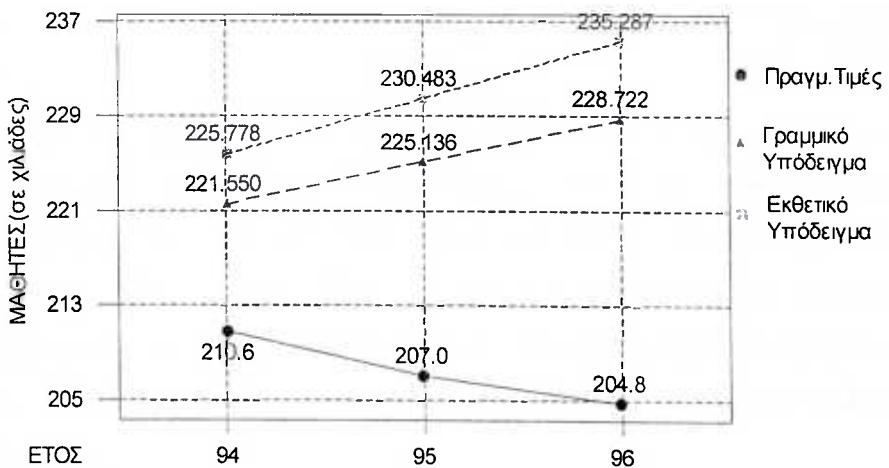
ΠΙΝΑΚΑΣ ΜΕΤΡΩΝ ΑΚΡΙΒΕΙΑΣ

	1 ^ο Υπόδειγμα $Y_t = b_0 + b_1 t$	2 ^ο Υπόδειγμα $\log Y_t = b_0 + b_1 t$
MAPE	7.81903	9.95616
MAD	17.6694	23.0493
MSE	340.359	570.429

Πίνακας 6.10: Μέτρα ακριβείας των υπόδειγμάτων της τάσης

Και στις τρεις περιπτώσεις πάντως τα μοντέλα δεν μπορούν να ανιχνεύσουν την πτώση που ακολουθεί την τελευταία τιμή της προσαρμογής και δίνουν διαρκώς αυξανόμενες τιμές. Το 3^ο υπόδειγμα εμφανίζει καλύτερα αποτελέσματα μεν αλλά και αυτό δίνει αποτελέσματα που υπερτιμούν τα πραγματικά και είναι διαρκώς αυξανόμενα. Καλύτερα παρουσιάζονται στο Γράφημα 6.18.

ΠΡΑΓΜΑΤΙΚΕΣ ΤΙΜΕΣ ΚΑΙ ΠΡΟΒΛΕΨΕΙΣ



ΓΡΑΦΗΜΑ 6.18: Προβλέψεις και πραγματικές τιμές των υπόδειγμάτων της τάσης

6.7 ΔΙΠΛΗ ΕΚΘΕΤΙΚΗ ΕΞΟΜΑΛΥΝΣΗ ΑΡΡΕΝΩΝ ΜΑΘΗΤΩΝ

Όπως αναφέραμε η επιλογή του καλύτερου υποδείγματος γίνεται με την ελαχιστοποίησης το μέσου τετραγωνικού σφάλματος (MSE). Με βάση αυτή την επιλογή μας το καλύτερο υπόδειγμα έχει $\alpha=0,823$ και γ σχεδόν μηδέν. Έχουμε δηλαδή ένα ακόμα μοντέλο παρόμοιο με αυτό που βρήκαμε για τις μαθήτριες του λυκείου. Πάντως η προσαρμογή του είναι σαφώς χειρότερη από το αντίστοιχο μοντέλο για τις μαθήτριες αφού αυτό που εξετάζουμε έχει μεγαλύτερο το MSD αλλά και τα άλλα μέτρα.

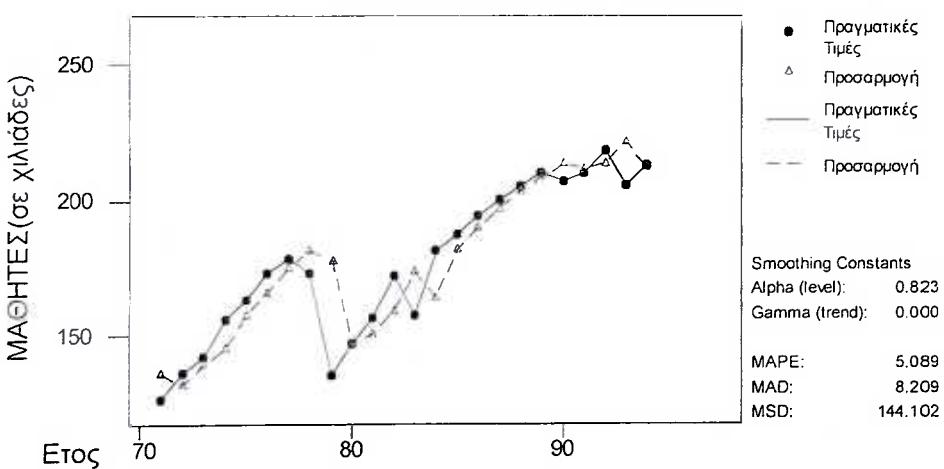
Double Exponential Smoothing

Data US-Male
Length 24.0000
NMissing 0

Smoothing Constants
Alpha (level): 0.82300
Gamma (trend): 0.00001

Accuracy Measures
MAPE: 5.089
MAD: 8.209
MSD: 144.102

ΔΙΠΛΗ ΕΚΘΕΤΙΚΗ ΕΞΟΜΑΛΥΝΣΗ ΜΑΘΗΤΩΝ Α.Δ.Ε



ΓΡΑΦΗΜΑ 6.19: Προσαρμογή υποδείγματος της Διπλής Εκθετικής Εξομάλυνσης για τους άρρενες μαθητές

Παρά το γεγονός ότι η μέθοδος της εκθετικής εξομάλυνσης θεωρείται πιο πολύπλοκη από την παλινδρόμηση, δεν δίνει καλύτερα αποτελέσματα από τα υποδείγματα αυτής.

ΠΙΝΑΚΑΣ ΠΡΑΓΜΑΤΙΚΩΝ ΤΙΜΩΝ ΚΑΙ ΠΡΟΒΛΕΠΟΜΕΝΩΝ

	Πραγμ. Τιμές	<u>ΔΕΕ</u>
1994/95	210.6	216.582
1995/96	207.0	220.168
1996/97	204.8	223.753

Πίνακας 6.11: Πραγματικές τιμές και προβλέψεις της ΔΕΕ

ΠΙΝΑΚΑΣ ΜΕΤΡΩΝ ΑΚΡΙΒΕΙΑΣ

	<u>ΔΕΕ</u>
MAPE	5.73782
MAD	12.7011
MSD	189.468

Πίνακας 6.12: Μέτρα ακριβείας της Δ.Ε.Ε

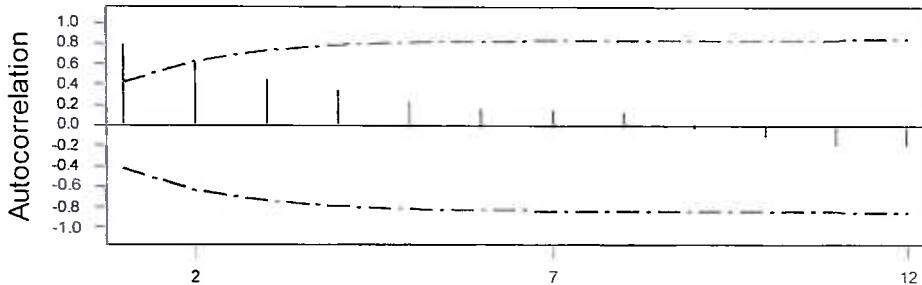
Παρατηρούμε παρόμοιο φαινόμενο με τα υποδείγματα της τάσης. Το υπόδειγμα προβλέπει αύξηση των αριθμών μαθητών στηριζόμενο στις τελευταίες παρατηρήσεις και στη γενική τάση και έτσι υπερεκτιμά τα αποτελέσματα. Η διαφορά του είναι μικρότερη από τα δύο πρώτα μοντέλα της τάσης αλλά μεγαλύτερη από το τρίτο.

6.8 ΥΠΟΔΕΙΓΜΑΤΑ ARIMA APPENΩΝ ΜΑΘΗΤΩΝ

Η εξεταζόμενη σειρά είναι φανερό από τα γραφήματα 6.17 αλλά και 6.19 ότι δεν είναι στάσιμη κάτι που εμφανίζεται και στα γραφήματα 6.20 και 6.21 των αυτοσυσχετίσεων και μερικών αυτοσυσχετίσεων.

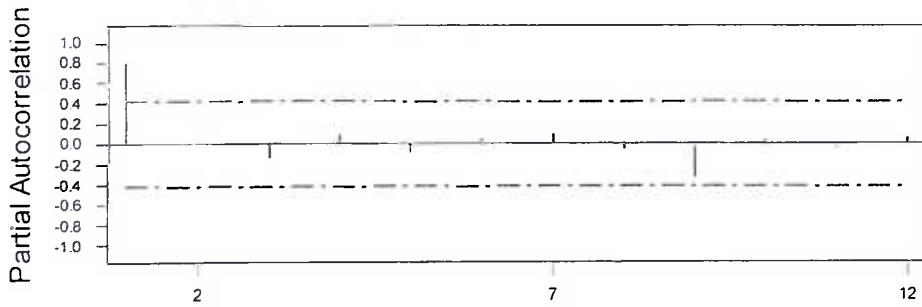
Από τα παραπάνω μπορούμε να συμπεράνουμε ότι η χρονολογική σειρά μπορεί να προσεγγισθεί είτε παίρνοντας τις πρώτες διαφορές και επανεξετάζοντας τους συντελεστές αυτοσυσχέτισης και μερικής αυτοσυσχέτισης, είτε με το μοντέλο ARIMA (1,0,0) αφού έχουμε μόνο τον πρώτο συντελεστή των μερικών αυτοσυσχετίσεων διαφορετικό από το μηδέν.

ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΜΑΘΗΤΩΝ Α.Δ.Ε



ΓΡΑΦΗΜΑ 6.20: Συντελεστές αυτοσυσχέτισης μαθητών ανώτερης δευτεροβάθμιας εκπαίδευσης

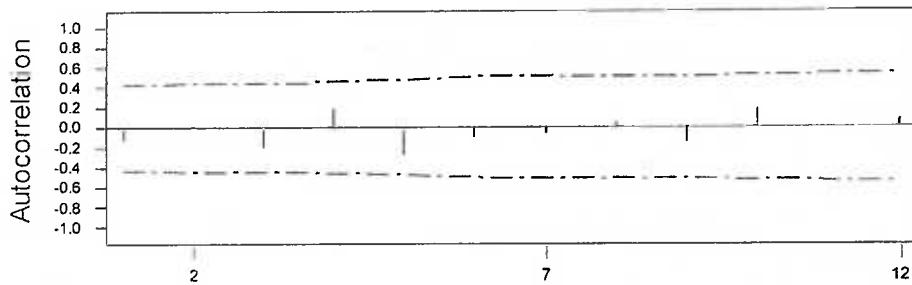
ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΜΑΘΗΤΩΝ Α.Δ.Ε



Lag	PAC	T	Lag	PAC	T
1	0.79	3.88	8	-0.06	-0.28
2	0.02	0.11	9	-0.35	-1.69
3	-0.15	-0.72	10	0.04	0.17
4	0.08	0.41	11	-0.05	-0.27
5	-0.10	-0.50	12	0.06	0.31
6	0.05	0.22			
7	0.09	0.46			

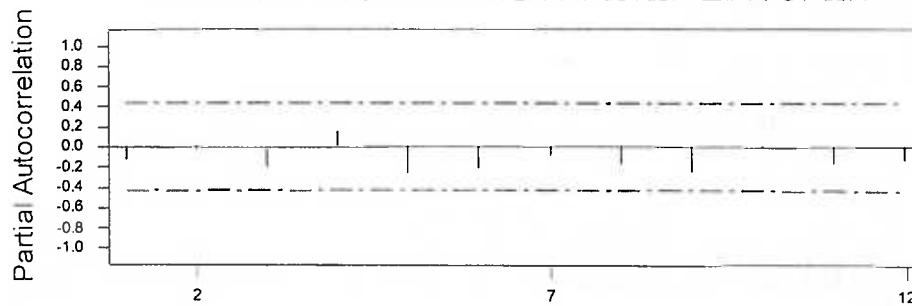
ΓΡΑΦΗΜΑ 6.21: Συντελεστές μερικής αυτοσυσχέτισης μαθητών

ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΠΡΩΤΩΝ ΔΙΑΦΟΡΩΝ



ΓΡΑΦΗΜΑ 6.22: Αυτοσυσχέτιση πρώτων διαφορών της χρονολογικής σειράς μαθητών ανώτερης δευτεροβάθμιας εκπαίδευσης

ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΠΡΩΤΩΝ ΔΙΑΦΟΡΩΝ



Lag	PAC	T	Lag	PAC	T
1	-0.13	-0.64	8	-0.18	-0.88
2	-0.03	-0.15	9	-0.25	-1.22
3	-0.21	-1.03	10	0.02	0.12
4	0.15	0.70	11	-0.17	-0.83
5	-0.28	-1.34	12	-0.13	-0.63
6	-0.23	-1.11			
7	-0.09	-0.45			

ΓΡΑΦΗΜΑ 6.23: Μερική αυτοσυσχέτιση πρώτων διαφορών των μαθητών

Τα Γραφήματα 6.22 και 6.23 δείχνουν ότι η έλλειψη στασιμότητας της σειράς έχει εξαλειφθεί, οπότε μπορούμε να προχωρήσουμε στην ανάλυση. Τα επικρατέστερα υποδείγματα με βάση την καλύτερη προσαρμογή τους είναι δύο:

Dependent Variable: USM
 Method: Least Squares
 Date: 05/13/04 Time: 06:30
 Sample(adjusted): 1971 1993
 Included observations: 23 after adjusting endpoints
 Convergence achieved after 3 iterations

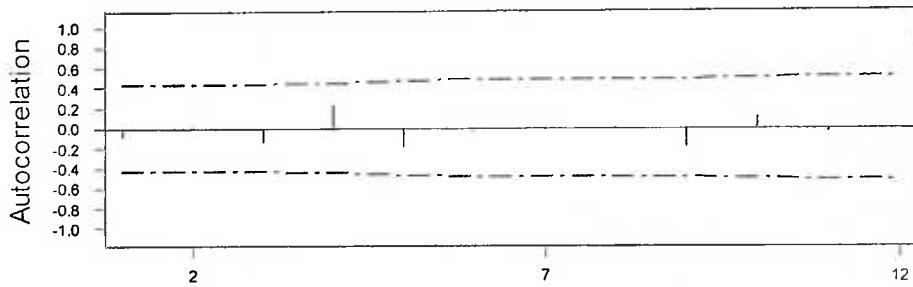
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	202.5188	25.91645	7.814295	0.0000
AR(1)	0.861189	0.092997	9.260390	0.0000
R-squared	0.803288	Mean dependent var	178.9435	
Adjusted R-squared	0.793920	S.D. dependent var	26.85396	
S.E. of regression	12.19063	Akaike info criterion	7.922153	
Sum squared resid	3120.838	Schwarz criterion	8.020891	
Log likelihood	-89.10476	F-statistic	85.75482	
Durbin-Watson stat	2.167026	Prob(F-statistic)	0.000000	
Inverted AR Roots	.86			

Dependent Variable: D(USM)
 Method: Least Squares
 Date: 05/12/04 Time: 22:58
 Sample(adjusted): 1971 1993
 Included observations: 23 after adjusting endpoints
 Convergence achieved after 10 iterations
 Backcast: 1967 1970

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MA(4)	0.885118	0.046620	18.98575	0.0000
R-squared	0.249566	Mean dependent var	3.800000	
Adjusted R-squared	0.249566	S.D. dependent var	12.52623	
S.E. of regression	10.85117	Akaike info criterion	7.648927	
Sum squared resid	2590.453	Schwarz criterion	7.698297	
Log likelihood	-86.96266	Durbin-Watson stat	2.090367	
Inverted MA Roots	.69+.69i	.69+.69i	-.69 -.69i	-.69 -.69i

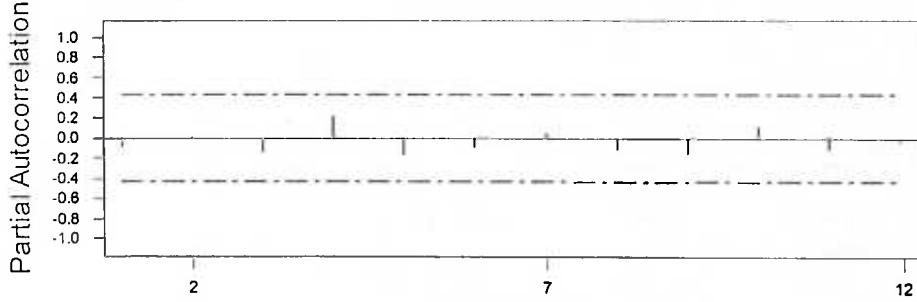
Καλύτερη προσαρμογή με βάση τα κριτήρια Akaike και Schwarz έχει το δεύτερο υπόδειγμα. Τα υποδείγματα έχουν συντελεστές Durbin-Watson κοντά στην τιμή 2 που σημαίνει ότι δεν υπάρχει συσχέτιση των καταλοίπων κάτι που θα εξεταστεί και με τα γραφήματα συσχέτισης και μερικής αυτοσυσχέτισης.

ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ ARIMA (1,0,0)



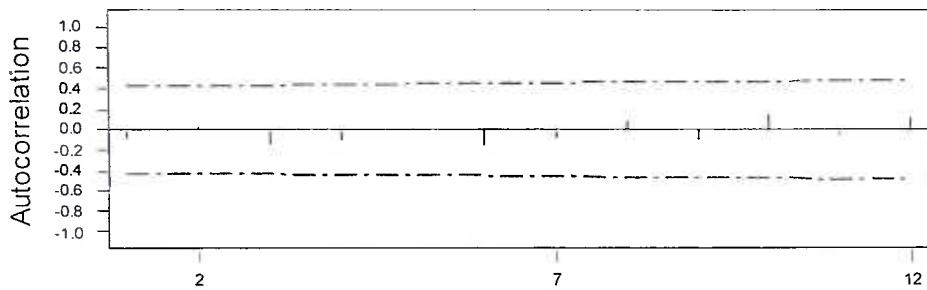
ΓΡΑΦΗΜΑ 6.24: Αυτοσυσχέτιση καταλοίπων υποδείγματος ARIMA (1,0,0) για τους μαθητές

ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ ARIMA(1,0,0)



ΓΡΑΦΗΜΑ 6.25: Μερική αυτοσυσχέτιση καταλοίπων ARIMA (1,0,0)

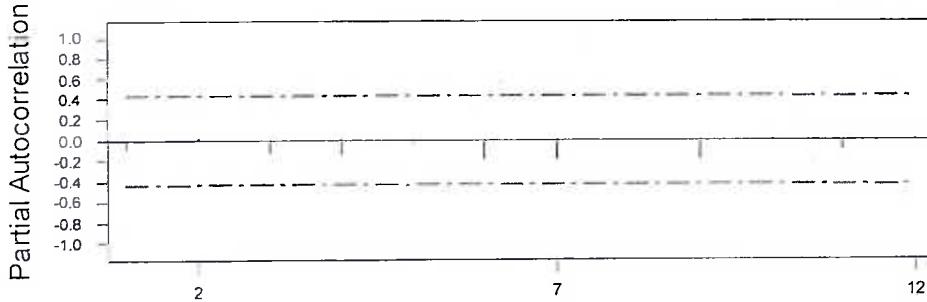
ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ ARIMA (0,1,4)



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	-0.10	-0.48	0.27	8	0.08	0.37	2.85
2	0.02	0.12	0.28	9	-0.11	-0.47	3.32
3	-0.15	-0.70	0.90	10	0.16	0.68	4.39
4	-0.12	-0.54	1.31	11	-0.06	-0.27	4.59
5	-0.02	-0.10	1.33	12	0.11	0.48	5.25
6	-0.17	-0.77	2.29				
7	-0.09	-0.41	2.59				

ΓΡΑΦΗΜΑ 6.26: Συντελεστές αυτοσυσχέτισης καταλοίπων υποδείγματος ARIMA (0,1,4) μαθητών ανώτερης δευτεροβάθμιας εκπαίδευσης

ΜΕΡΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ ΚΑΤΑΛΟΙΠΩΝ ARIMA(0,1,4)



Lag	PAC	T	Lag	PAC	T
1	-0.10	-0.48	8	0.00	0.02
2	0.01	0.07	9	-0.20	-0.98
3	-0.14	-0.69	10	-0.00	-0.01
4	-0.15	-0.72	11	-0.11	-0.54
5	-0.05	-0.23	12	-0.01	-0.05
6	-0.21	-1.00			
7	-0.20	-0.95			

ΓΡΑΦΗΜΑ 6.27: Συντελεστές μερικής αυτοσυσχέτισης καταλοίπων υποδείγματος ARIMA (0,1,4)

Τα κατάλοιπα είναι ασυσχέτιστα. άρα τα υποδείγματα έχουν καλή εφαρμογή στα δεδομένα και έτσι μπορούμε να συνεχίσουμε με τις προβλέψεις των υποδειγμάτων.

ΠΙΝΑΚΑΣ ΠΡΑΓΜΑΤΙΚΩΝ ΤΙΜΩΝ ΚΑΙ ΠΡΟΒΛΕΠΟΜΕΝΩΝ

Έτος	Πραγμ. Τιμές	ARIMA(1,0,0)	ARIMA(0,1,4)
1994/95	210.6	211.717	216.643
1995/96	207.0	210.440	227.444
1996/97	204.8	209.341	215.990

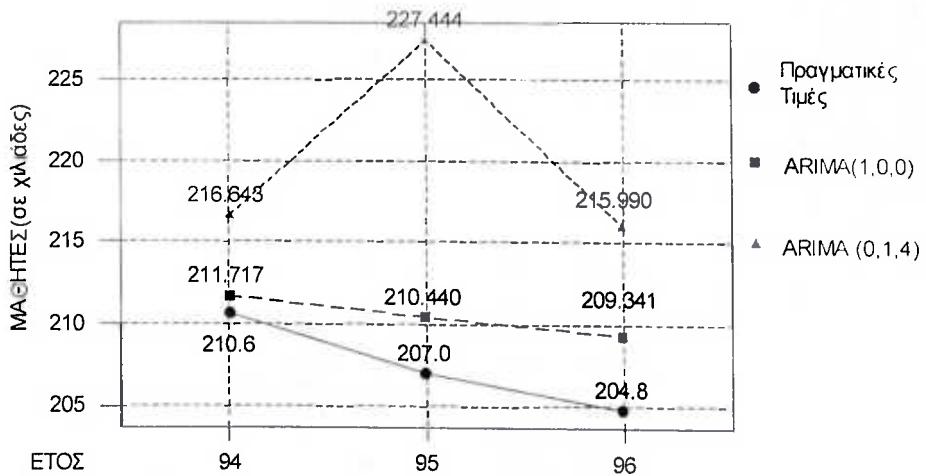
Πίνακας 6.13: Πραγματικές τιμές και προβλέψεις των υποδειγμάτων ARIMA μαθητών

ΠΙΝΑΚΑΣ ΜΕΤΡΩΝ ΑΚΡΙΒΕΙΑΣ

	ARIMA(1,0,0)	ARIMA(0,1,4)
MAPE	1.44392	5.65291
MAD	3.03289	12.5590
MSD	11.2349	193.231

Πίνακας 6.14: Μέτρα ακριβείας των υποδειγμάτων ARIMA

ΠΡΑΓΜΑΤΙΚΕΣ ΤΙΜΕΣ ΚΑΙ ΠΡΟΒΛΕΨΕΙΣ



ΓΡΑΦΗΜΑ 6.28: Πρόβλεψη υποδειγμάτων ARIMA και πραγματικές τιμές

Το υπόδειγμα ARIMA(1,0,0) δίνει καλύτερες προβλέψεις από το υπόδειγμα ARIMA(0,1,4). Οι τιμές και των δύο είναι μεγαλύτερες από τις πραγματικές ωστόσο το πρώτο υπόδειγμα δίνει τιμές πιο κοντά στις πραγματικές και διαρκώς μειούμενες ενώ το δεύτερο υπόδειγμα εκτός από μεγαλύτερες τιμές από τις πραγματικές, αδυνατεί να ακολουθήσει την πτωτική τάση.

6.8 ΣΥΜΠΕΡΑΣΜΑΤΑ ΓΙΑ ΤΟΥΣ ΜΑΘΗΤΕΣ

	MAPE	MAD	MSD
$Y_t = b_0 + b_1 t$	7.81903	17.6694	340.359
$\log Y_t = b_0 + b_1 t$	9.95616	23.0493	570.429
ΔΕΕ	5.73782	12.7011	189.468
ARIMA(1,0,0)	1.44392	3.03289	11.2349
ARIMA (0,1,4)	5.65291	12.5590	193.231

Πίνακας 6.15: Σύγκριση μέτρων ακριβείας για τους μαθητές ανώτερης δευτεροβάθμιας εκπαίδευσης

Το καλύτερο υπόδειγμα όσον αφορά την πρόβλεψη είναι το ARIMA(1,0,0) κάτι που ίσως δεν ήταν τόσο αναμενόμενο αφού η σειρά δεν είναι στάσιμη. Ενώ η Διπλή Εκθετική Εξομάλυνση δίνει μια σταθερά μέτρια πρόβλεψη, τα υποδείγματα της τάσης δίνουν τις χειρότερες προβλέψεις. Είναι χαρακτηριστικό ότι για τις μαθήτριες του Λυκείου αλλά και για τους Μαθητές του Λυκείου πιο καλή πρόβλεψη δίνουν τα υποδείγματα ARIMA. Ίσως σε καταστάσεις που είναι πιο ευμετάβλητες τα υποδείγματα αυτά να ανιχνεύουν τις αλλαγές και ανταποκρίνονται καλύτερα, ενώ τα υποδείγματα της τάσης δίνουν καλύτερες προβλέψεις όταν υπάρχει μεγαλύτερη σταθερότητα και λιγότερη αβεβαιότητα.



ΚΕΦΑΛΑΙΟ 7 :ΣΥΜΠΕΡΑΣΜΑΤΑ

7.1 ΔΗΜΟΤΙΚΟ

Το πρόβλημα που εμφανίζεται στην παιδεία και αφορά τους μαθητές των σχολείων είναι ιδιαίτερα έντονο. Όσον αφορά τα δημοτικά, ήδη από τα μέσα της δεκαετίας του 70 έχει αρχίσει μια διαρκής πτώση του αριθμού των μαθητών η οποία οφείλεται αποκλειστικά στην μείωση των γεννήσεων που εμφανίζεται σταδιακά από την δεκαετία του 60, αλλά υπήρξε ραγδαία από τις αρχές της δεκαετίας του 80. Έτσι παρά την αύξηση των σχολικών κτηρίων στην επικράτεια η πτώση του αριθμού των μαθητών υπήρξε εξίσου ραγδαία μετά τα μέσα του 80 όταν τα παιδιά που γεννιόνταν στις αρχές του 80, άρχισαν να πηγαίνουν στο δημοτικό.

Βασισμένοι στα δεδομένα της Eurostat ο πληθυσμός των μαθητών δεν φαίνεται να επηρεάζεται από την είσοδο χιλιάδων μεταναστών στην χώρα, η οποία παρατηρήθηκε ήδη από τις αρχές της δεκαετίας του 90. Βέβαια οι πρώτοι αυτοί μετανάστες δεν είχαν εγκατασταθεί ακόμα μόνιμα με τις οικογένειες τους αφού ήταν παράνομοι, άρα θα περιμένουμε αύξηση των μαθητών λόγω των μεταναστών μερικά χρόνια αργότερα και συγκεκριμένα μετά από το 1998 οπότε άρχισε η νομιμοποίησή τους, και θα πρέπει να συνεχίζεται μέχρι και σήμερα.

Η διαφορά μεταξύ αγοριών και κοριτσιών είναι μεγαλύτερη στα παλαιότερα έτη και μειώνεται όσο πλησιάζουμε στα πιο πρόσφατα. Με δεδομένο ότι η διαφορά του πληθυσμού αγοριών και κοριτσιών είναι σχεδόν σταθερή στο πέρασμα των ετών, αυτό αποτελεί μια ένδειξη της προκατάληψης που υπήρχε παλιότερα και στερούσε την στοιχειώδη ακόμα εκπαίδευση από τα θηλυκά μέλη των οικογενειών αφού δεν την θεωρούσαν τόσο απαραίτητη για τα κορίτσια όσο για τα αγόρια.

7.2 ΔΕΥΤΕΡΟΒΑΘΜΙΑ ΕΚΠΑΙΔΕΥΣΗ

Η συνολική μορφή που εμφανίζει η δευτεροβάθμια εκπαίδευση είναι μια διαρκής και ομαλή αύξηση του αριθμού των μαθητών μέχρι τις αρχές της δεκαετίας του 90, οπότε και αρχίζει η πτώση. Εκείνη τη χρονική στιγμή αρχίζει να επηρεάζει τον πληθυσμό των Γυμνασίων και Λυκείων η υπογεννητικότητα της χώρας.

Λογικά θα έπρεπε η μείωση των μαθητών να έχει ξεκινήσει τουλάχιστον από την δεκαετία του 80, λόγω της ελάττωσης του νεανικού πληθυσμού, όμως αυτό δεν εμφανίζεται γιατί ενώ παλαιότερα η παρακολούθηση της δευτεροβάθμιας εκπαίδευσης δεν γινόταν από μια μεγάλη μερίδα των απόφοιτων του δημοτικού, στα χρόνια πλέον του 1970 όπου η υποχρεωτική εκπαίδευση γίνεται 9ετής αλλά και η συνειδητοποίηση της κοινωνίας για την προσφορά της παιδείας στην βελτίωση των συνθηκών της ζωής δημιουργεί μια αύξηση στην δευτεροβάθμια εκπαίδευση η οποία υπερκαλύπτει την μείωση η οποία επέρχεται λόγω μείωσης των γεννήσεων.

Η εξέλιξη των μαθητών και των μαθητριών είναι παρόμοιες μεταξύ τους, με την διαφορά ότι η σειρά των κοριτσιών εμφανίζει πιο ομαλή εξέλιξη από αυτή των αγοριών. Οι σειρές αυτές αυξάνονται διαρκώς μέχρι το σχολικό έτος 1991/92, οπότε και αρχίζει η καθοδική πορεία τους. Οι σειρά των αγοριών πάντως εμφανίζεται πιο ασταθής με αρκετές αυξομειώσεις στην πορεία του χρόνου που εξετάζουμε.

Η δευτεροβάθμια εκπαίδευση χωρίζεται όμως σε δύο επιμέρους βαθμίδες. Την κατώτερη δευτεροβάθμια εκπαίδευση όπου το μεγαλύτερο ποσοστό των μαθητών είναι στα γυμνάσια και την ανώτερη δευτεροβάθμια εκπαίδευση όπου το μεγαλύτερο ποσοστό μαθητών φοιτούν στα λύκεια. Ήδη από το 1977 η παρακολούθηση του γυμνασίου είναι υποχρεωτική, ενώ του λυκείου εξακολουθεί να είναι προαιρετική

7.2.1 Κατώτερη δευτεροβάθμια εκπαίδευση.

Στην κατώτερη δευτεροβάθμια εκπαίδευση οι μαθητές και οι μαθήτριες αυξάνονται από το 1970 με μια έκρηξη την χρονιά 1978/79 οπότε και επιβλήθηκε η 9ετής εκπαίδευση. Από την χρονιά εκείνη και μετά υπάρχει μια σταθερή πορεία για μερικά έτη οπότε και αρχίζει η πτώση. Οι μαθητές αυξάνονται αργά μέχρι την ραγδαία αύξηση του 1978/79 και στη συνέχεια μειώνονται αργά αλλά συνεχόμενα. Αντίθετα οι μαθήτριες αυξάνονται συνεχώς από το 1970/71 μέχρι το 1981/82. Υπάρχει σταθερότητα μέχρι το 1994/95 και πτώση στα δύο τελευταία σχολικά έτη. Η ενσωμάτωση των μεταναστών που θα επέφερε και αύξηση του αριθμού των μαθητών δεν έχει γίνει ακόμα την περίοδο που εξετάζουμε ή είναι η πτώση μεγαλύτερη και μετριάζεται με την είσοδο τους στα σχολεία.



Η διαφορά των μαθητών και των μαθητριών, ενώ μειώθηκε σε σχέση με την δεκαετία του 70 εμφάνισε μια μικρή ανάκαμψη την χρονιά 1995/96. Η ανάκαμψη αύτη ίσως είναι τυχαία και παροδική γιατί η διαφορά υποχωρεί την αμέσως επόμενη χρονιά. Και εδώ παρατηρούμε το φαινόμενο που παρατηρήθηκε στο δημοτικό, να είναι μεγαλύτερη όσο απομακρυνόμαστε από τα πρόσφατα έτη

7.2.2 Ανώτερη δευτεροβάθμια εκπαίδευση.

Το σύνολο των μαθητών και μαθητριών των Λυκείων είναι διαρκώς αυξανόμενο με εξαίρεση την χρονική περίοδο 1978/79 οπότε σημείωσε μια απότομη πτώση που οφειλόταν στον διαχωρισμό του εξατάξιου γυμνασίου και την απαίτηση εισιτηρίων εξετάσεων για να εγγραφούν οι μαθητές σ' αυτό. Από την επόμενη χρονιά όμως και μέχρι το 1991/92 υπάρχει αύξηση των μαθητών. Μετά το έτος αυτό η σειρά φαίνεται ότι σταθεροποιείται σε ένα συγκεκριμένο επίπεδο. Μέχρι την στιγμή που εξετάζουμε ο αυξανόμενος αριθμός μαθητών που συνεχίζει στο Λύκειο υπερκαλύπτει την πτώση λόγω του αριθμού των γεννήσεων με αποτέλεσμα να μην παρατηρείται μείωση στην χρονοσειρά. Η σταθεροποίηση που παρατηρείται μετά το 1991/92 αποτελεί προμήνυμα της επερχόμενης υποχώρησης.

Οι μαθητές έχουν παρόμοια πορεία με το σύνολο. Αύξηση από το 1970 μέχρι το 1977, μια ραγδαία πτώση για το έτος 1978 και στη συνέχεια άνοδο μέχρι το 1991. Από το σημείο και αυτό αρχίζει η υποχώρηση. Στα αγόρια η μείωση του αριθμού τους έχει αρχίσει και γίνεται αισθητή από το 1992.

Οι μαθήτριες έχουν και αυτές συνεχή αύξηση με μια πτώση το 1978/79 αλλά η σειρά εξελίσσεται πιο ομαλά και η πτώση είναι πιο μικρή. Επιπλέον οι μαθήτριες έχουν αυξητική τάση μέχρι και το 1996/97 σε αντίθεση με τους μαθητές. Ίσως αυτό είναι ένδειξη στροφής των μαθητών σε άλλες λύσεις μετά το γυμνάσιο ή ότι ο αριθμός τους επηρεάζεται από την υπογεννητικότητα ενώ των μαθητριών την μείωση αυτή την υπερκαλύπτει η αυξανόμενη ανάγκη για εκπαίδευση.

Η διαφορά μεταξύ μαθητών και μαθητριών είναι κυμαινόμενη αλλά σε γενική μορφή καθοδική. Μπορεί να υποθέσει κάποιος ότι παρουσιάζει 3 ή 4 άνισους κύκλους. Από το 1970 μέχρι το 1978/79, από το 1978/79 μέχρι το 1982/83, από το 182/83 μέχρι το 1992/93 και από το 1992/93 μέχρι το 1996/97. Στις μεταβολές αυτές έχουμε διαρκείς

ανξομειώσεις με τελική κατάληξη την χρονιά 1996/97 οι μαθήτριες να είναι περισσότερες από τους μαθητές γεγονός που ή δείχνει ότι περισσότερα αγόρια αναζητούν διαφορετικές λύσεις το μέλλον τους. Η είναι εντελώς τυχαίο και προφανώς θα επανέλθουμε σε μια σταθερή κατάσταση στο μέλλον.

7.3 ΣΤΑΤΙΣΤΙΚΑ ΣΥΜΠΕΡΑΣΜΑΤΑ

Τα υποδείγματα που εξετάσαμε ήταν στην ουσία τριών ειδών. Υποδείγματα αποσύνθεσης των συνιστωσών, εκθετικής εξομάλυνσης (κυρίως διπλής) και υποδείγματα ARIMA. Στα υποδείγματα αποσύνθεσης συνιστωσών βασιστήκαμε στην ύπαρξη τάσης και της τυχαιότητας ως μοναδικών συνιστωσών των χρονολογικών σειρών. Οι τάσεις που προσδιορίστηκαν και πάνω στις οποίες βασίστηκαν οι προβλέψεις ήταν η δευτεροβάθμια για την διαχρονική εξέλιξη των συνόλου των μαθητών Πρωτοβάθμιας Εκπαίδευσης, η δευτεροβάθμια και η δευτεροβάθμια λογαριθμική για την διαχρονική εξέλιξη των αρρένων μαθητών Κατώτερης Δευτεροβάθμιας Εκπαίδευσης, των μαθητριών της ίδιας βαθμίδας, αλλά και των μαθητριών Ανώτερης Δευτεροβάθμιας Εκπαίδευσης και τέλος η απλή γραμμική και πρωτοβάθμια λογαριθμική για τους άρρενες μαθητές Ανώτερης Δευτεροβάθμιας Εκπαίδευσης.

Θα προσπαθήσουμε να εντοπίσουμε την απόδοση των παραπάνω μεθόδων στον τομέα της πρόβλεψης. Θα θεωρήσουμε δηλαδή τις 3 κατηγορίες και θα κατατάξουμε με βάση τα 3 μέτρα ακριβείας

Από τους Πίνακες 7.1 και 7.2 συμπεραίνουμε, ότι με βάση το μέτρο MAPE έχουμε το υπόδειγμα που βασίζεται στην ανάλυση συνιστωσών να παρουσιάζει καλύτερες προβλέψεις για τις χρονολογικές σειρές των αρρένων μαθητών και μαθητριών της Κατώτερης Δευτεροβάθμιας Εκπαίδευσης, ενώ για τις υπόλοιπες περιπτώσεις εμφανίζει τις χειρότερες. Η μέθοδος της Εκθετικής εξομάλυνσης, να παρουσιάζει την καλύτερη προβλεπτική ικανότητα για την Πρωτοβάθμια Εκπαίδευση, την χειρότερη για τις μαθήτριες Κατώτερης Δευτεροβάθμιας και για τις υπόλοιπες περιπτώσεις έχει την δεύτερη καλύτερη επίδοση. Η μέθοδος ARIMA ήταν αποτελεσματικότερη για τους άρρενες μαθητές και μαθήτριες της Ανώτερης Δευτεροβάθμιας Εκπαίδευσης, είχε την δεύτερη καλύτερη επίδοση για τις μαθήτριες Κατώτερης Δευτεροβάθμιας και

ήταν υποδεέστερη σε απόδοση των άλλων δύο μεθόδων, στους άρρενες μαθητές Κατώτερης Δευτεροβάθμιας.

Για το μέτρο της μέσης απόλυτης απόκλισης (MAD) έχουμε την μέθοδο της αποσύνθεσης συνιστωσών να έχει την ίδια απόδοση με το μέτρο MAPE. Η εκθετική εξομάλυνση να δίνει τις χειρότερες προβλέψεις για τις μαθήτριες Κατώτερης Δευτεροβάθμιας εκπαίδευσης ενώ στις υπόλοιπες περιπτώσεις τα αποτελέσματα είναι τα δεύτερα καλύτερα. Η μέθοδος ARIMA προσεγγίζει τις μελλοντικές τιμές καλύτερα από τις άλλες μεθόδους για το σύνολο των μαθητών της Πρωτοβάθμιας εκπαίδευσης καθώς επίσης και για τους άρρενες μαθητές και τις μαθήτριες της Ανώτερης Δευτεροβάθμιας. Για τις μαθήτριες της κατώτερης δευτεροβάθμιας υπερτερεί της Εκθετικής Εξομάλυνσης αλλά υστερεί της Αποσύνθεσης Συνιστωσών, ενώ για τους άρρενες μαθητές της ίδιας βαθμίδας υστερεί και των δύο προαναφερομένων.

Για το πιο σημαντικό ίσως μέτρο το MSE, έχουμε την Αποσύνθεση Συνιστωσών να έχει την ίδια ακριβώς εικόνα με τα προηγούμενα μέτρα. Η εκθετική εξομάλυνση αποκλίνει ως προς τις προβλέψεις της περισσότερο από τις άλλες δύο μεθόδους όταν προσεγγίζει τους άρρενες μαθητές και τις μαθήτριες της Κατώτερης Δευτεροβάθμιας εκπαίδευσης ενώ στις υπόλοιπες περιπτώσεις είναι η δεύτερη καλύτερη επιλογή. Η μέθοδος ARIMA τέλος παρουσιάζει τις καλύτερες προβλέψεις για την Πρωτοβάθμια Εκπαίδευση και την Ανώτερη δευτεροβάθμια, ενώ στην Κατώτερη Δευτεροβάθμια παρουσιάζει τις αμέσως καλύτερες επιδόσεις.

Μπορούμε να συμπεράνουμε κατά συνέπεια ότι τα υποδείγματα της τάσης μπορεί να δίνουν σε κάποιες περιπτώσεις καλύτερα αποτελέσματα, αλλά δεν δίνουν σε καμία περίπτωση μέτριες εκτιμήσεις όταν αποτυγχάνουν. Οι εκτιμήσεις τους είναι περισσότερο ακριβείς σε χρονολογικές σειρές που δεν παρουσιάζουν σχετικά ομαλή πορεία χωρίς ξαφνικές αλλαγές. Ωστόσο, από την πρέπει να αναφέρουμε ωστόσο ότι τα μοντέλα περιπτώσης που εξετάσαμε εμφανίζονται προβληματικά όσον αφορά την αυτοσυγχέτιση των καταλοίπων τους και για το λόγο αυτό οι προβλέψεις τους δεν μπορούν να θεωρηθούν απόλυτα φερέγγυες.

Αντίθετα τα υποδείγματα ARIMA εμφανίζουν αρκετές φορές τα καλύτερα αποτελέσματα, ωστόσο, τις δύο φορές που δεν είναι τα καλύτερα έχουν την δεύτερη καλύτερη επίδοση. Τα μοντέλα αυτά ανταποκρίνονται καλύτερα από τα προηγούμενα σε περιβάλλον “ασταθές”, αλλά δεν δίνουν καλύτερες προβλέψεις σε χρονολογικές σειρές που εξελίσσονται ομαλά.

Η εκθετική εξομάλυνση τέλος απέτυχε να δώσει τις καλύτερες προβλέψεις σε όποιο μέτρο και αν εξετάσουμε. Επειδή τις χειρότερες επιδόσεις τις έχει στο MSE, το οποίο δίνει μεγαλύτερη βαρύτητα στα μεγάλα σφάλματα από τα αλλά δύο μέτρα ακριβείας, μπορούμε να συμπεράνουμε ότι παρουσιάζει κάποιες πολύ μεγάλες αποκλίσεις που διογκώνουν το MSE και αρκετές μικρές.

		Αποσύνθεση συνιστωσών	Εκθετική εξομάλυνση	ARIMA
ΜΑΘΗΤΕΣ ΚΑΙ ΜΑΘΗΤΡΙΕΣ ΔΗΜΟΤΙΚΟΥ		Δευτεροβάθμια τάση		(1,1,1)
	MAPE	3,297	1,305	1,6787
	MAD	22,201	11,506	11,2712
	MSE	523,532	229,699	137,416
ΜΑΘΗΤΡΙΕΣ ΓΥΜΝΑΣΙΟΥ		Δευτεροβάθμια τάση		(0,2,2)
	MAPE	1,49592	3,73122	2,3564
	MAD	3,02443	7,80575	4,81607
	MSE	11,4491	95,3068	32,1432
ΜΑΘΗΤΕΣ ΓΥΜΝΑΣΙΟΥ		Λογαριθμική Β' βάθμια Τάση		(1,1,1)
	MAPE	0,653701	3,51264	3,57271
	MAD	1,42618	7,96667	8,09643
	MSE	3,58606	82,9433	79,4996
ΜΑΘΗΤΕΣ ΛΥΚΕΙΟΥ		Απλή Γραμμική τάση		(1,0,0)
	MAPE	7,81903	5,73782	1,44392
	MAD	17,6694	12,7011	3,03289
	MSE	340,395	189,468	11,2349
ΜΑΘΗΤΡΙΕΣ ΛΥΚΕΙΟΥ		Δευτεροβάθμια τάση		(1,0,0)
	MAPE	5,03620	3,53399	0,867113
	MAD	11,4324	7,48039	1,75847
	MSE	135,739	60,7744	4,21385

Πίνακας 7.1: Σύγκριση των μεθόδων με βάση τα μέτρα ακριβείας

		Αποσύνθεση συνιστωσών	Εκθετική εξομάλυνση	ARIMA
ΜΑΘΗΤΕΣ ΚΑΙ ΔΗΜΟΤΙΚΟΥ	MAPE	3	1	2
ΜΑΘΗΤΡΙΕΣ ΓΥΜΝΑΣΙΟΥ	MAD	3	2	1
ΜΑΘΗΤΕΣ ΓΥΜΝΑΣΙΟΥ	MSE	3	2	1
ΜΑΘΗΤΕΣ ΛΥΚΕΙΟΥ	MAPE	1	3	2
ΜΑΘΗΤΕΣ ΛΥΚΕΙΟΥ	MAD	1	3	2
ΜΑΘΗΤΕΣ ΛΥΚΕΙΟΥ	MSE	1	2	3
ΜΑΘΗΤΡΙΕΣ ΛΥΚΕΙΟΥ	MAPE	3	2	1
ΜΑΘΗΤΡΙΕΣ ΛΥΚΕΙΟΥ	MAD	3	2	1
ΜΑΘΗΤΡΙΕΣ ΛΥΚΕΙΟΥ	MSE	3	2	1

Πίνακας 7.2: Σύγκριση μεθόδων πρόβλεψης με βάση τα μέτρα ακριβείας

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Ζ. Γωργαντά Β (1987), *H προσέγγιση Box-Jenkins στην ανάλυση και πρόβλεψη χρονολογικών σειρών*, Κέντρο προγραμματισμού και οικονομικών μελετών, Αθήνα
2. Α. Κωστάκη (1998) *Δημομετρία*, Εκδόσεις Οικονομικού Πανεπιστημίου Αθηνών
3. Ι. Πανάρετον, και Ε. Ξεκαλάκη (1998), *Εισαγωγή στη στατιστική σκέψη (συμπλήρωμα)*, Αθήνα
4. Π.Θ Τζωρτζόπουλου (1980), *Ανάλυσις χρονολογικών σειρών*, Αθήνα
5. ΕΣΥΕ (1975, 1976, 1981, 1984, 1994), *Στατιστική της εκπαίδευσης*, Αθήνα
6. Box, G.E.P. and G.M. Jenkins (1976), *Time series analysis*, Rev. ed. Holden Day, San Francisco
7. Cryer, D.J and R.B. Miller (1994), *Statistics for business*, Duxbury Press, California
8. Makridakis, S and S.C. Wheelwright (1989), *Forecasting methods for management*, John Wiley & Sons, New York
9. Montgomery, D.C. and E.A. Peck (1992), *Introduction to linear regression analysis*, John Wiley & Sons, New York
10. Eurostat (1985) *Education and training*
11. Eurostat(1995-98), *Education across European Union*





ΠΑΡΑΡΤΗΜΑ

ΠΙΝΑΚΑΣ ΔΕΔΟΜΕΝΩΝ

Year	Primary Total	P- Females	P- Males	LS- Total	LS- Females	LS- Males	US- Total	US- Female	US- Male
1970/71	920	439,3	480,7	332,3	134,7	197,6	222,5	96,7	125,8
1971/72	911,9	436,7	475,2	349,7	141,6	208,1	240,4	104,8	135,6
1972/73	915,3	438,6	476,7	364	147,9	216,1	249,9	108,4	141,5
1973/74	926,9	444,6	482,3	363,4	147,3	216,1	274	118,7	155,3
1974/75	931,6	447,2	484,4	366,6	149,2	217,4	286,9	124,2	162,7
1975/76	938,8	451,3	487,5	374,1	154	220,1	305,5	132,9	172,6
1976/77	941,9	453,6	488,3	381,8	161	220,8	312,7	134,7	178
1977/78	936,4	451,8	484,6	400,5	174,1	226,4	304,4	131,7	172,7
1978/79	922,7	444,7	478	458,2	195,2	263	251,9	117,5	134,4
1979/80	901,9	434	467,9	455,3	201,3	254	270,3	123,8	146,5
1980/81	903,1	435,7	467,4	452,8	206,3	246,5	288	132	156
1981/82	893,8	430,7	463,1	460	213,3	246,7	319	147	172
1982/83	892,8	432,3	460,4	453	211,5	241,5	305	148	157
1983/84	890,8	431,3	459,4	441	208,5	232,5	350	169	181
1984/85	892,7	432,3	460,4	441	208,5	232,5	363	176	187
1985/86	893,3	431,7	461,6	441,5	208,2	233,3	373,1	179	194
1986/87	871	420,7	450,3	455,8	215,2	240,5	381,1	181	200,1
1987/88	873	422,6	450,3	450,8	214,3	236,5	388,1	183	205,1
1988/89	858,1	415,6	442,4	451,6	215,3	236,4	397,1	187,1	210,1
1989/90	838,9	405,9	432,9	447,1	213,7	233,4	397,4	190,7	206,8
1990/91	817,7	396	421,7	447,1	213,7	233,4	405,1	195,3	209,8
1991/92	791,2	384,3	407	442,8	211,4	231,4	422,2	203,6	218,6
1992/93	749,3	363,2	386,1	436,1	209,3	226,8	407,3	201,6	205,7
1993/94	723,7	350,3	373,4	439	210,3	228,7	412,3	199,1	213,2
1994/95	710,8	346,9	363,9	433	210,2	222,9	409,6	199	210,6
1995/96	675,3	327,1	348,1	421,9	201	220,9	413,2	206,3	207
1996/97	652,0	346,2	335,8	407,1	194,4	212,7	410,5	205,7	204,8

MACRO ΥΠΟΛΟΓΙΣΜΟΥ ΜΕΤΡΩΝ ΑΚΡΙΒΕΙΑΣ

MACRO

FORERROR K1 K2 K3

MCONSTANTS N M MSE MAD MAPE

MCOLUMN K1 K2 K3

LET N = COUNT (K1)

LET M= COUNT (K2)

IF N ~= M

NOTE **ERROR***"Variables have different length"

ELSE

LET MSE =(SSQ(K1-K2))/N

LET MAD = SUM(ABS(K1-K2))/N

LET MAPE = SUM(ABS((K1-K2)/K1)*100)/N

PRINT MAPE MAD MSE

```
SET K3  
MAPE MAD MSE  
END  
ENDIF  
ENDMACRO
```

MACRO ΥΠΟΛΟΓΙΣΜΟΥ R-sq KAI R-sq(adj)
MACRO

```
RSQ K1 K2 K3  
  
MCOLUMN K1 K2  
MCONSTANT N M1 M2 K3  
  
LET N= COUNT(K1)  
LET M1=1-(SSQ(K1-K2)/SSQ(K1-MEAN(K1)))  
  
LET M2 = 1-((1-M1)*(N-1)/(N-K3))  
NAME M1='R-sq=' M2='R-sq(adj)='
```

```
PRINT M1 M2  
  
ENDMACRO
```

MACRO ΥΠΟΛΟΓΙΜΟΥ ΣΥΝΤΕΛΕΣΤΗ DURBIN-WATSON
MACRO

```
DW K1  
  
MCOLUMN K1  
MCONSTANT M1  
  
LET M1=(SSQ(K1-LAG(K1)))/(SSQ(K1))  
  
NAME M1='D-W='  
  
PRINT M1  
  
ENDMACRO
```



Δυνατά

