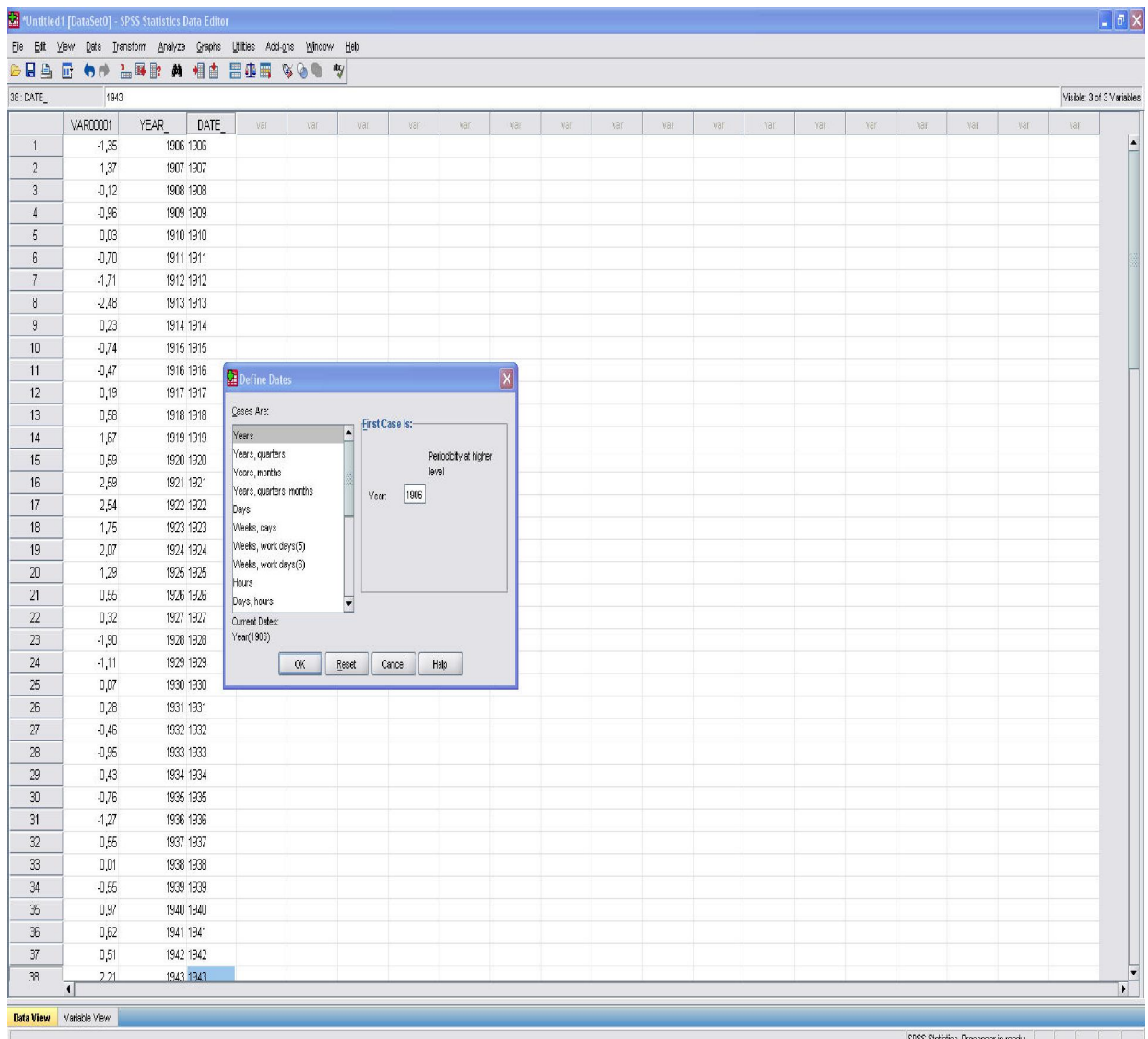


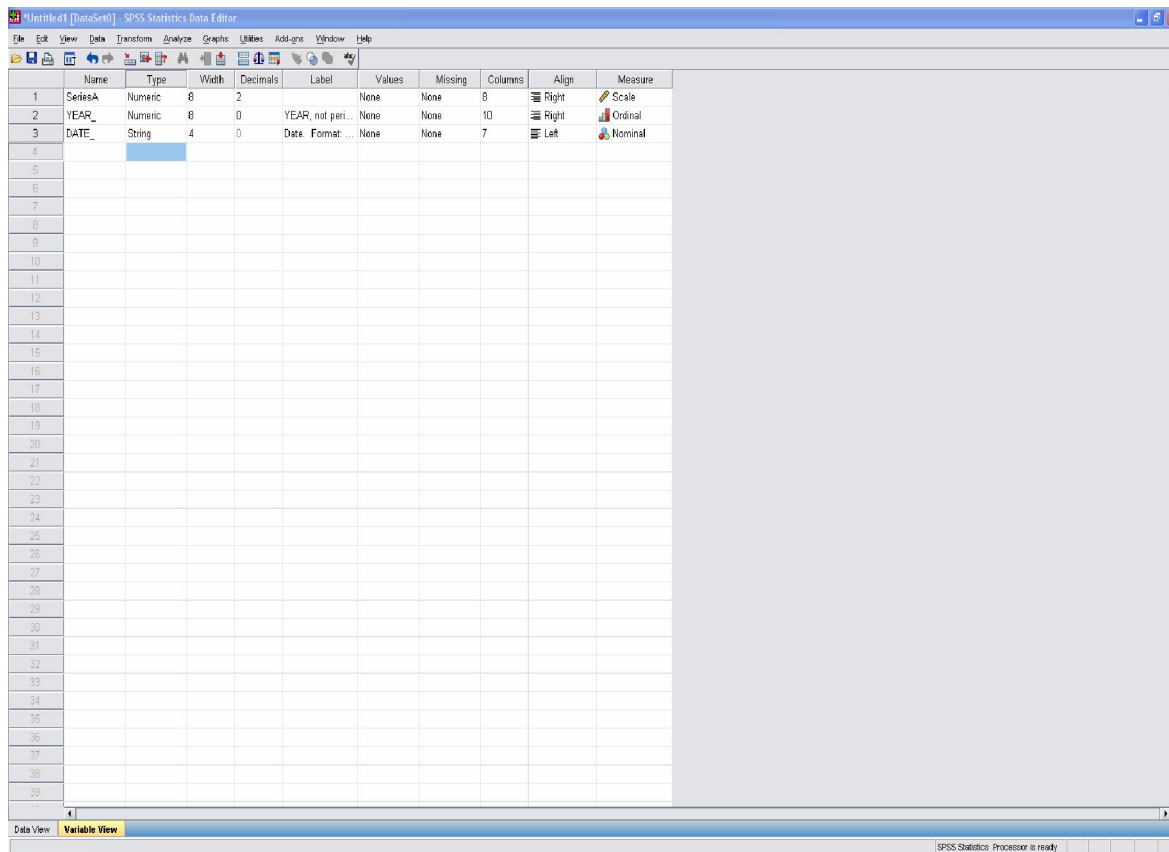
1. ΕΙΣΑΓΩΓΗ ΣΤΟ SPSS

Το SPSS είναι ένα στατιστικό πρόγραμμα γενικής στατιστικής ανάλυσης αρκετά εύκολο στη λειτουργία του. Για να πραγματοποιηθεί ανάλυση χρονοσειρών με τη βοήθεια του SPSS θα πρέπει απαραίτητα να έχει εγκατασταθεί μαζί με το πρόγραμμα και το πακέτο **Forecasting** ή σε πιο παλιές εκδόσεις του προγράμματος το πακέτο **Trends**. Στις σημειώσεις αυτές παρουσιάζονται οι βασικές λειτουργίες του SPSS 17 που αφορούν την ανάλυση χρονοσειρών με τη μεθοδολογία των Box και Jenkins, χρησιμοποιώντας ως παράδειγμα μια χρονοσειρά 100 παρατηρήσεων που καλύπτουν τη χρονική περίοδο 1906 – 2005.

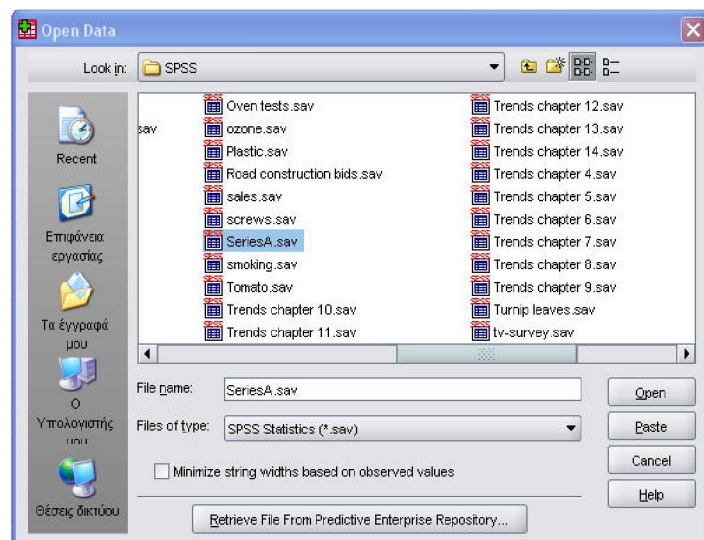
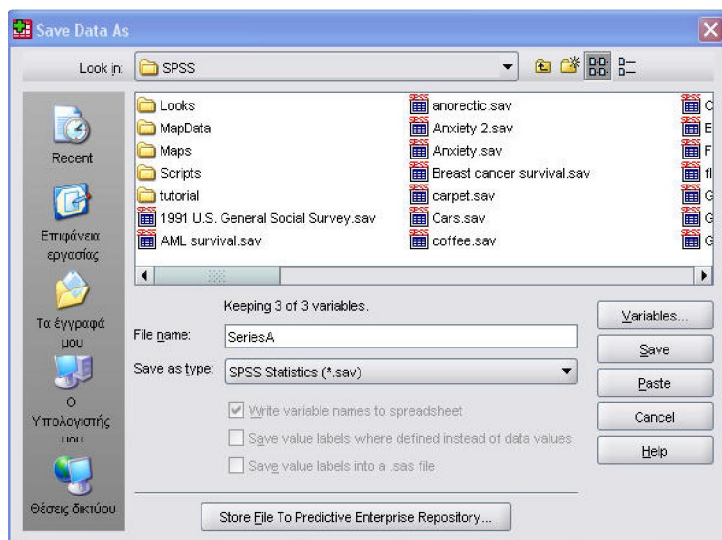
Εκτελώντας το SPSS 17 προκύπτει ένα κενό data file, στο οποίο εισάγονται τα δεδομένα της χρονοσειράς. Στη συνέχεια, μπορεί να επιλεγεί να δημιουργηθεί μια καινούρια μεταβλητή που να περιέχει την ημερομηνία που αντιστοιχεί η κάθε παρατήρηση της χρονοσειράς. Για να πραγματοποιηθεί αυτό επιλέγεται Data/Define Dates και καθορίζεται το χρονικό διάστημα αναφοράς της κάθε παρατήρησης της χρονοσειράς. Στην περίπτωση της χρονοσειράς που εξετάζεται επιλέγεται το έτος, δηλαδή years, αφού οι παρατηρήσεις της είναι σε ετήσια βάση, στο First case is εισάγεται το έτος που αντιστοιχεί στην πρώτη παρατήρηση (1906) και πατιέται OK.



Πατώντας το κουμπί Variable View μπορούν να ονομαστούν οι μεταβλητές και να καθοριστεί ο αριθμός των δεκαδικών ψηφίων που θα περιλαμβάνει κάθε παρατήρηση. Χρησιμοποιείται το όνομα SeriesA για τη χρονοσειρά του παραδείγματος. Πατώντας το κουμπί Data View γίνεται επιστροφή στην προηγούμενη οθόνη.

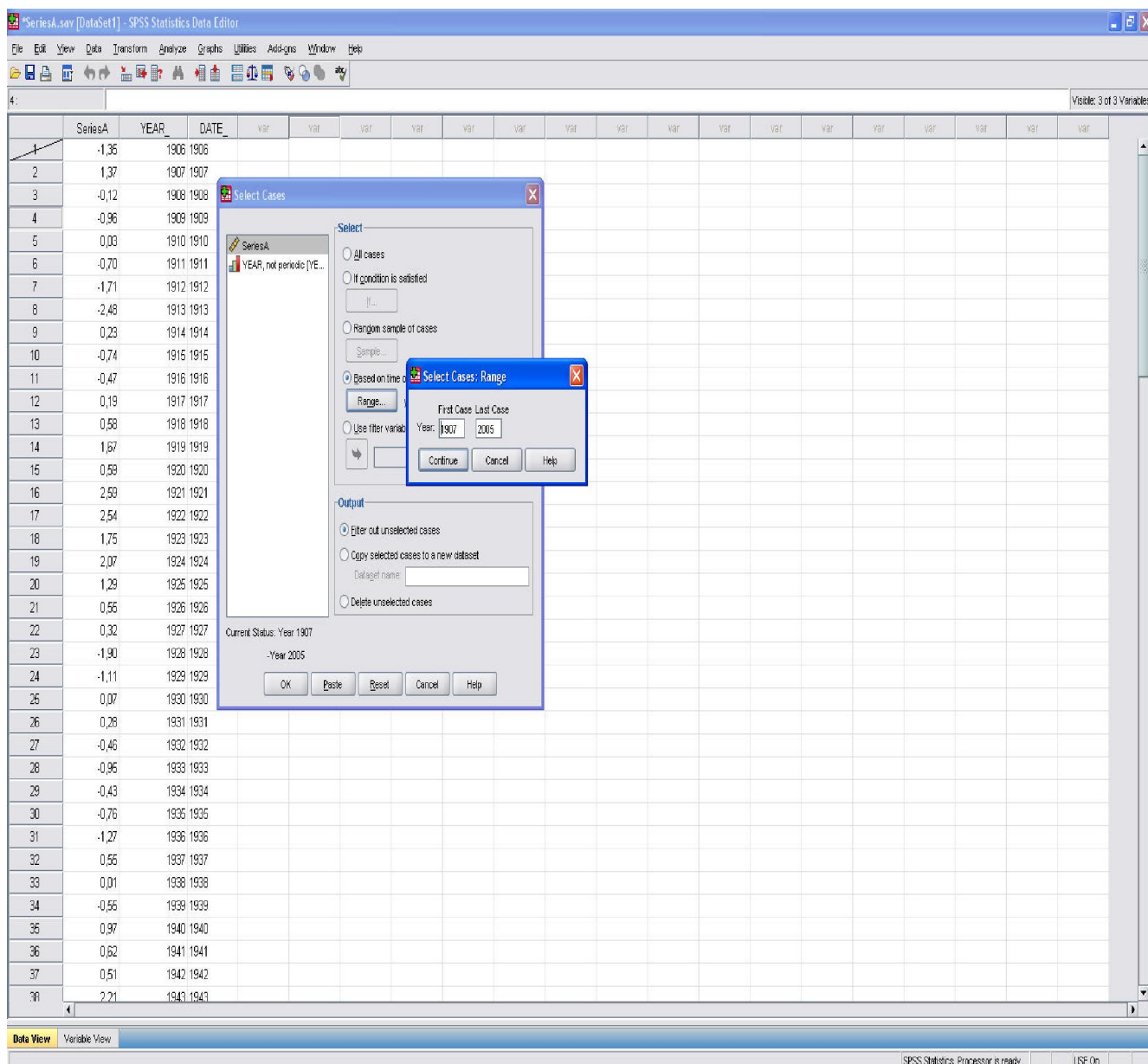


Για να αποθηκευτεί το data file που εργαζόμαστε επιλέγουμε File/Save as και διαλέγουμε ένα όνομα για το αρχείο μας, ενώ αν θέλουμε να ανοίξουμε ένα παλιότερο data file επιλέγουμε File/Open/Data και στη συνέχεια το όνομα του αρχείου. Στο παράδειγμα που εξετάζεται θα αποθηκευτεί το αρχείο με το όνομα SeriesA.



2. ΕΠΙΛΟΓΗ ΤΟΥ ΜΕΓΕΘΟΥΣ ΤΩΝ ΠΑΡΑΤΗΡΗΣΕΩΝ

Πολλές φορές χρειάζεται να εκτιμηθεί κάποιο υπόδειγμα με διαφορετικό αριθμό παρατηρήσεων, π.χ. με μια λιγότερη παρατήρηση. Για παράδειγμα έστω ότι επιθυμούμε στην ανάλυσή μας να μη συμπεριλαμβάνεται το έτος 1906, δηλαδή η παρατήρηση της χρονοσειράς που αναφέρεται στο πρώτο έτος. Στην περίπτωση αυτή επιλέγουμε Data/Select Cases. Στο πλαίσιο διαλόγου που ανοίγει επιλέγουμε Select Based on time or case range και πατάμε το κουμπί Range. Στη συνέχεια, πληκτρολογούμε το πρώτο έτος που θέλουμε να περιλαμβάνει το δείγμα μας, δηλαδή το 1907, και το τελευταίο, δηλαδή το 2005 και πατάμε Continue και Ok. Στον Data Editor του SPSS διαγράφεται ο αριθμός του δείγματος που αντιστοιχεί στην πρώτη παρατήρηση και το SPSS στη μετέπειτα ανάλυση θα αγνοεί το πρώτο έτος. Αν θέλουμε να ξαναχρησιμοποιήσουμε όλο το δείγμα επαναλαμβάνουμε τη προηγούμενη διαδικασία επιλέγοντας All Cases.



3. ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΣ ΜΙΑΣ ΧΡΟΝΟΣΕΙΡΑΣ

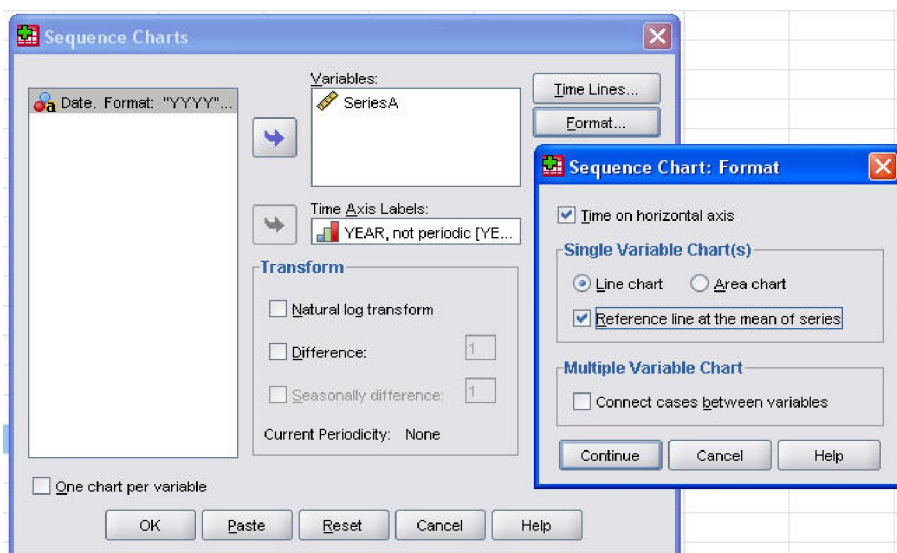
Μία χρήσιμη δυνατότητα του SPSS είναι η κατασκευή μετασχηματισμένων χρονοσειρών από τις αρχικές τιμές μιας χρονοσειράς. Τέτοιοι μετασχηματισμοί μπορεί να είναι η εφαρμογή διαφορών στην αρχική χρονοσειρά, οι χρονικές υστερήσεις, η εφαρμογή κάποιου κινητού μέσου κ.τ.λ. Για παράδειγμα, για να υπολογιστούν οι πρώτες διαφορές της χρονοσειράς SeriesA επιλέγουμε Transform/Create Time Series και στο πλαίσιο διαλόγου που προκύπτει επιλέγουμε στο Function την επιλογή Difference και στο Order πληκτρολογούμε τον αριθμό 1. Στη συνέχεια, επιλέγουμε την μεταβλητή SeriesA και την εισάγουμε στο πλαίσιο New Variable(s). Πατώντας το κουμπί OK δημιουργείται στον Data Editor η καινούρια μεταβλητή SeriesA_1 που αποτελείται από τις πρώτες διαφορές της χρονοσειράς SeriesA.

The screenshot displays the SPSS Statistics Data Editor window with a dataset named 'SeriesA.sav'. The data table has columns for 'SeriesA', 'YEAR', 'DATE', and 'SeriesA_1'. The 'SeriesA_1' column contains the first differences of the 'SeriesA' values. A 'Create Time Series' dialog box is open, showing the 'Variable->New name' field set to 'SeriesA_1=DIFF(SeriesA_1)'. The 'Name and Function' section shows the 'Function' set to 'Difference' and the 'Order' set to 1. The 'Current Periodicity' is set to 'None'. The dialog box has buttons for 'OK', 'Paste', 'Reset', 'Cancel', and 'Help'.

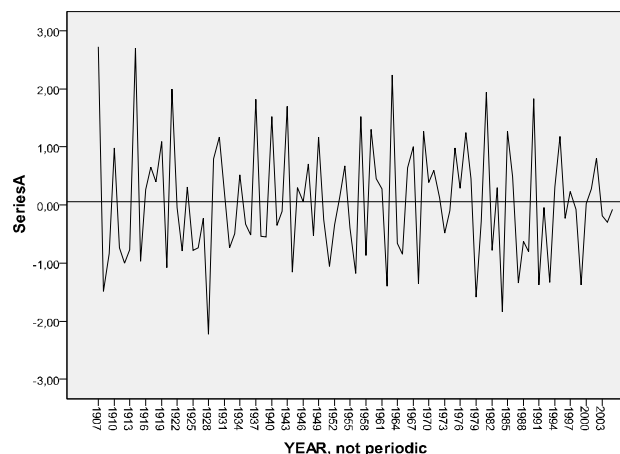
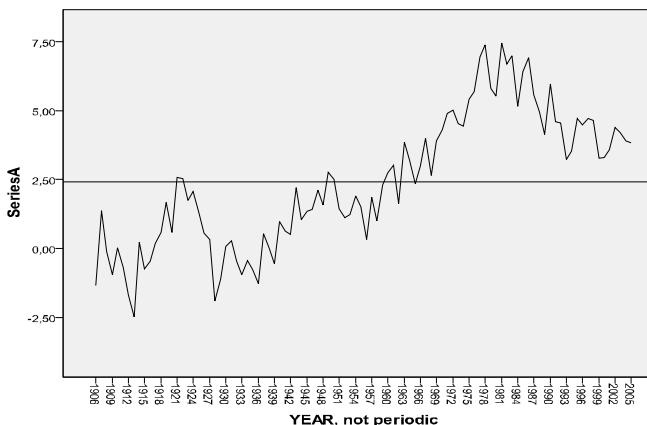
	SeriesA	YEAR	DATE	SeriesA_1
1	-1,35	1906	1906	.
2	1,37	1907	1907	2,72
3	-0,12	1908	1908	-1,49
4	-0,96	1909	1909	-0,84
5	0,03	1910	1910	0,98
6	-0,70	1911	1911	-0,73
7	-1,71	1912	1912	-1,00
8	-2,48	1913	1913	-0,77
9	0,23	1914	1914	2,71
10	-0,74	1915	1915	-0,97
11	-0,47	1916	1916	0,27
12	0,19	1917	1917	0,65
13	0,58	1918	1918	0,39
14	1,67	1919	1919	1,09
15	0,59	1920	1920	-1,08
16	2,59	1921	1921	2,00
17	2,54	1922	1922	-0,05
18	1,75	1923	1923	-0,79
19	2,07	1924	1924	0,31
20	1,29	1925	1925	-0,78
21	0,55	1926	1926	-0,74
22	0,32	1927	1927	-0,23
23	-1,90	1928	1928	-2,22
24	-1,11	1929	1929	0,80
25	0,07	1930	1930	1,17
26	0,28	1931	1931	0,21
27	-0,46	1932	1932	-0,74
28	-0,95	1933	1933	-0,50
29	-0,43	1934	1934	0,52
30	-0,76	1935	1935	-0,33
31	-1,27	1936	1936	-0,51
32	0,55	1937	1937	1,82
33	0,01	1938	1938	-0,54
34	-0,55	1939	1939	-0,55
35	0,97	1940	1940	1,52
36	0,62	1941	1941	-0,36
37	0,51	1942	1942	-0,11
38	2,21	1943	1943	1,70

4. ΚΑΤΑΣΚΕΥΗ ΤΟΥ ΔΙΑΓΡΑΜΜΑΤΟΣ ΜΙΑΣ ΧΡΟΝΟΣΕΙΡΑΣ

Για να κατασκευάσουμε το διάγραμμα μιας χρονοσειράς επιλέγουμε Analyze/Forecasting/Sequence Charts και μετακινούμε στο Variables τη μεταβλητή τις παρατηρήσεις της οποίας επιθυμούμε να απεικονίσουμε, δηλαδή τη μεταβλητή SeriesA, και στο Time Axis Labels τη μεταβλητή YEAR που συμβολίζει τη χρονική περίοδο. Το SPSS επιπλέον μας παρέχει τη δυνατότητα να σχεδιάσουμε τη χρονοσειρά έχοντας εφαρμόσει λογαριθμικό μετασχηματισμό στα δεδομένα ή μετασχηματισμό σε πρώτες διαφορές, διαλέγοντας τις αντίστοιχες επιλογές. Επίσης, αν θέλουμε να σχεδιάσουμε περισσότερες από μια χρονοσειρές μπορούμε να επιλέξουμε αν θα σχεδιαστούν όλες μαζί ή κάθε μία χωριστά. Αν θέλουμε στο διάγραμμα να περιλαμβάνεται και ο μέσος της χρονοσειράς μπορούμε να πατήσουμε το κουμπί Format και να διαλέξουμε Reference line at mean of series και Continue.



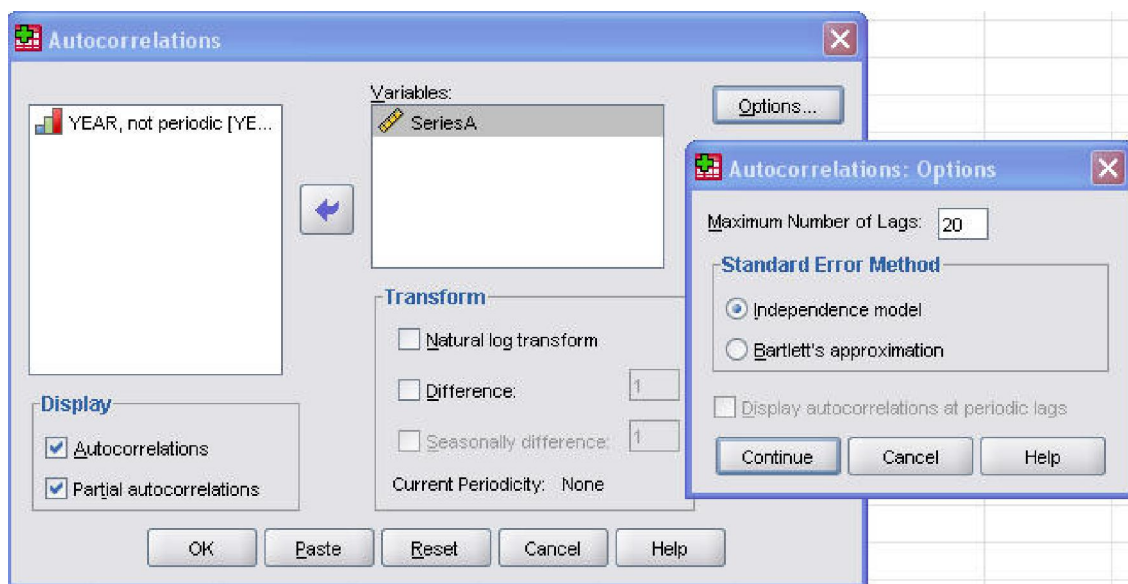
Πατώντας OK στο SPSS Statistics Viewer προκύπτει το διάγραμμα της χρονοσειράς από το οποίο διαπιστώνεται ότι η εξεταζόμενη χρονοσειρά είναι μη στάσιμη διότι οι τιμές της δεν κινούνται γύρω από το μέσο όρο της. Επαναλαμβάνοντας την προηγούμενη διαδικασία εφαρμόζοντας μετασχηματισμό σε πρώτες διαφορές για τις τιμές της χρονοσειράς διαπιστώνεται ότι η χρονοσειρά σε πρώτες διαφορές γίνεται στάσιμη.



Transforms: difference(1)

5. ΕΚΤΙΜΗΣΗ ΤΩΝ ΑΥΤΟΣΥΣΧΕΤΙΣΕΩΝ ΚΑΙ ΜΕΡΙΚΩΝ ΑΥΤΟΣΥΣΧΕΤΙΣΕΩΝ

Για να υπολογίσουμε τις αυτοσυσχετίσεις και μερικές αυτοσυσχετίσεις της χρονοσειράς επιλέγουμε **Analyze/Forecasting/Autocorrelations** και μετακινούμε στο **Variables** τη μεταβλητή **SeriesA** που περιέχει τις τιμές της χρονοσειράς. Στο **Display** αφήνουμε ενεργοποιημένα τα **Autocorrelations** και **Partial autocorrelations**. Αν θέλουμε μπορούμε να υπολογίσουμε τις αυτοσυσχετίσεις και μερικές αυτοσυσχετίσεις της χρονοσειράς έχοντας εφαρμόσει λογαριθμικό μετασχηματισμό στις παρατηρήσεις ή έχοντας εφαρμόσει διαφορές. Πατώντας το κουμπί **Options** μπορούμε να διαλέξουμε τον αριθμό των αυτοσυσχετίσεων και μερικών αυτοσυσχετίσεων που θέλουμε να υπολογιστούν. Συνήθως επιλέγουμε τον αριθμό 20 και επιστρέφουμε στην προηγούμενη οθόνη πατώντας **Continue**.



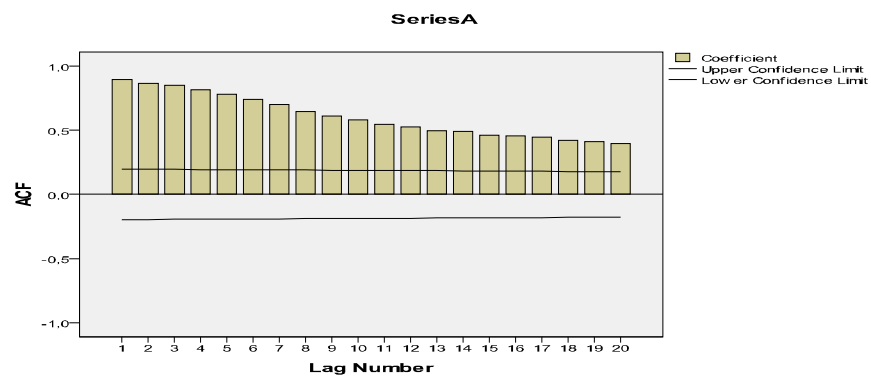
Επιλέγοντας **OK** στο **SPSS Statistics Viewer** προκύπτουν τα διάγραμμα αυτοσυσχετίσεων (**ACF**) και μερικών αυτοσυσχετίσεων (**PACF**) της χρονοσειράς στα οποία εκτός από τους συντελεστές αυτοσυσχέτισης απεικονίζεται και ένα 95% διάστημα εμπιστοσύνης που μας βοηθάει στην επιλογή των σημαντικών αυτοσυσχετίσεων. Όσες αυτοσυσχετίσεις βρίσκονται εκτός των ορίων αυτού του διαστήματος θεωρούνται σημαντικές. Επιπρόσθετα, εμφανίζονται οι τιμές των δειγματικών αυτοσυσχετίσεων και μερικών αυτοσυσχετίσεων της χρονοσειράς. Στον πίνακα των αυτοσυσχετίσεων το **SPSS** δίνει τις τιμές της στατιστικής **Q** των **Box** και **Ljung** με τα αντίστοιχα **P-Values** για τον έλεγχο της μηδενικής υπόθεσης ότι η χρονοσειρά είναι λευκός θόρυβος. Παρατηρούμε ότι οι αυτοσυσχετίσεις της χρονοσειράς φθίνουν με αργό ρυθμό και δεν συγκλίνουν στο μηδέν οπότε μπορεί να θεωρηθεί μη στάσιμη. Από τις τιμές της στατιστικής **Q** συμπεραίνεται ότι οι τιμές της χρονοσειράς δεν αποτελούν λευκό θόρυβο, δηλαδή συσχετίζονται μεταξύ τους.

Autocorrelations

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,896	,099	82,790	1	,000
2	,868	,098	161,224	2	,000
3	,852	,098	237,632	3	,000
4	,817	,097	308,557	4	,000
5	,781	,097	373,996	5	,000
6	,743	,096	433,820	6	,000
7	,703	,095	488,000	7	,000
8	,646	,095	534,338	8	,000
9	,613	,094	576,509	9	,000
10	,583	,094	615,012	10	,000
11	,546	,093	649,192	11	,000
12	,529	,093	681,573	12	,000
13	,499	,092	710,727	13	,000
14	,494	,092	739,681	14	,000
15	,464	,091	765,516	15	,000
16	,458	,091	790,942	16	,000
17	,447	,090	815,482	17	,000
18	,424	,090	837,808	18	,000
19	,409	,089	858,913	19	,000
20	,398	,089	879,061	20	,000

a. The underlying process assumed is independence (white noise).

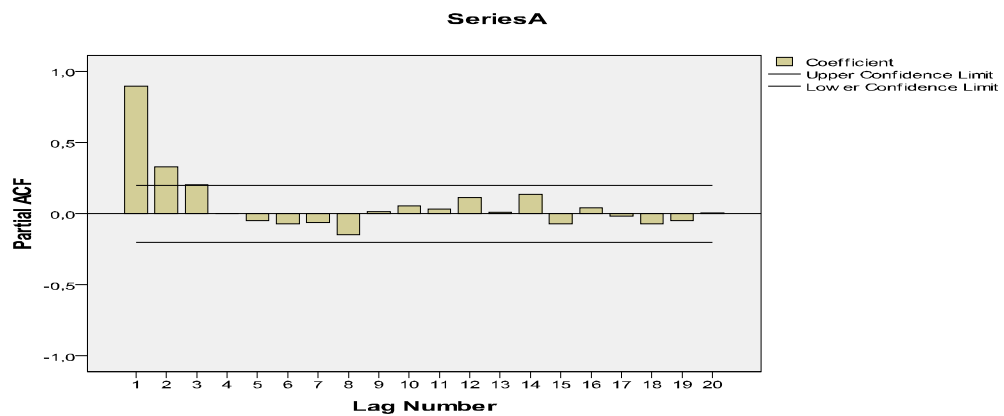
b. Based on the asymptotic chi-square approximation.



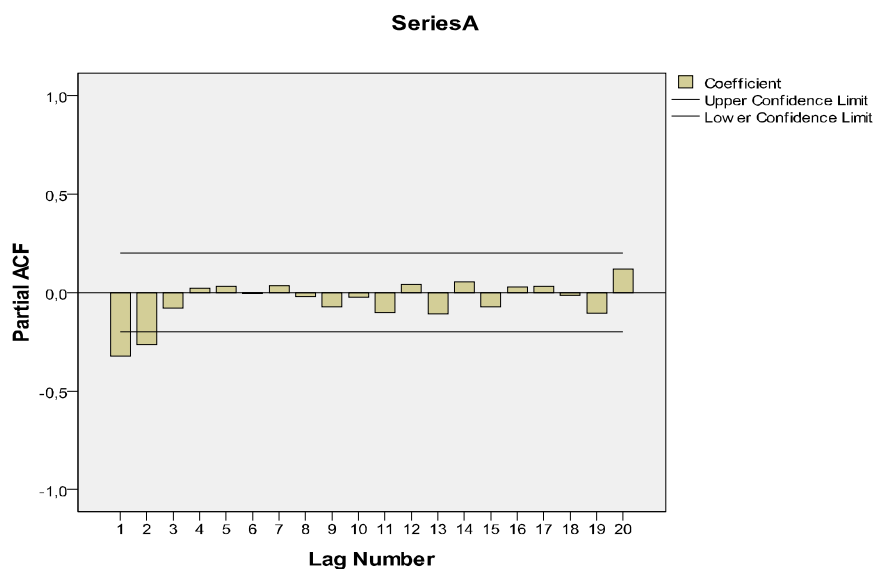
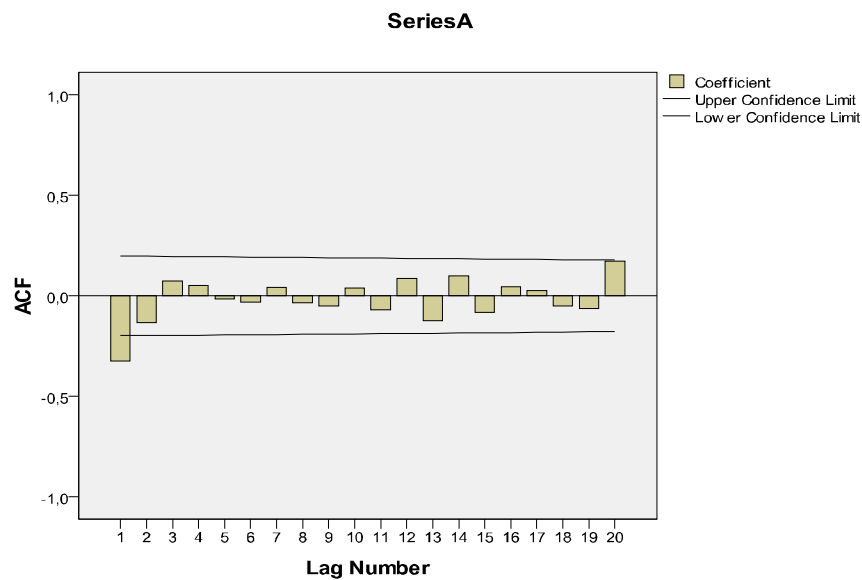
Partial Autocorrelations

Series:SeriesA

Lag	Partial Autocorrelation	Std. Error
1	,896	,100
2	,329	,100
3	,202	,100
4	-,001	,100
5	-,051	,100
6	-,073	,100
7	-,063	,100
8	-,146	,100
9	,013	,100
10	,052	,100
11	,032	,100
12	,113	,100
13	,008	,100
14	,133	,100
15	-,070	,100
16	,041	,100
17	-,018	,100
18	-,071	,100
19	-,050	,100
20	,005	,100

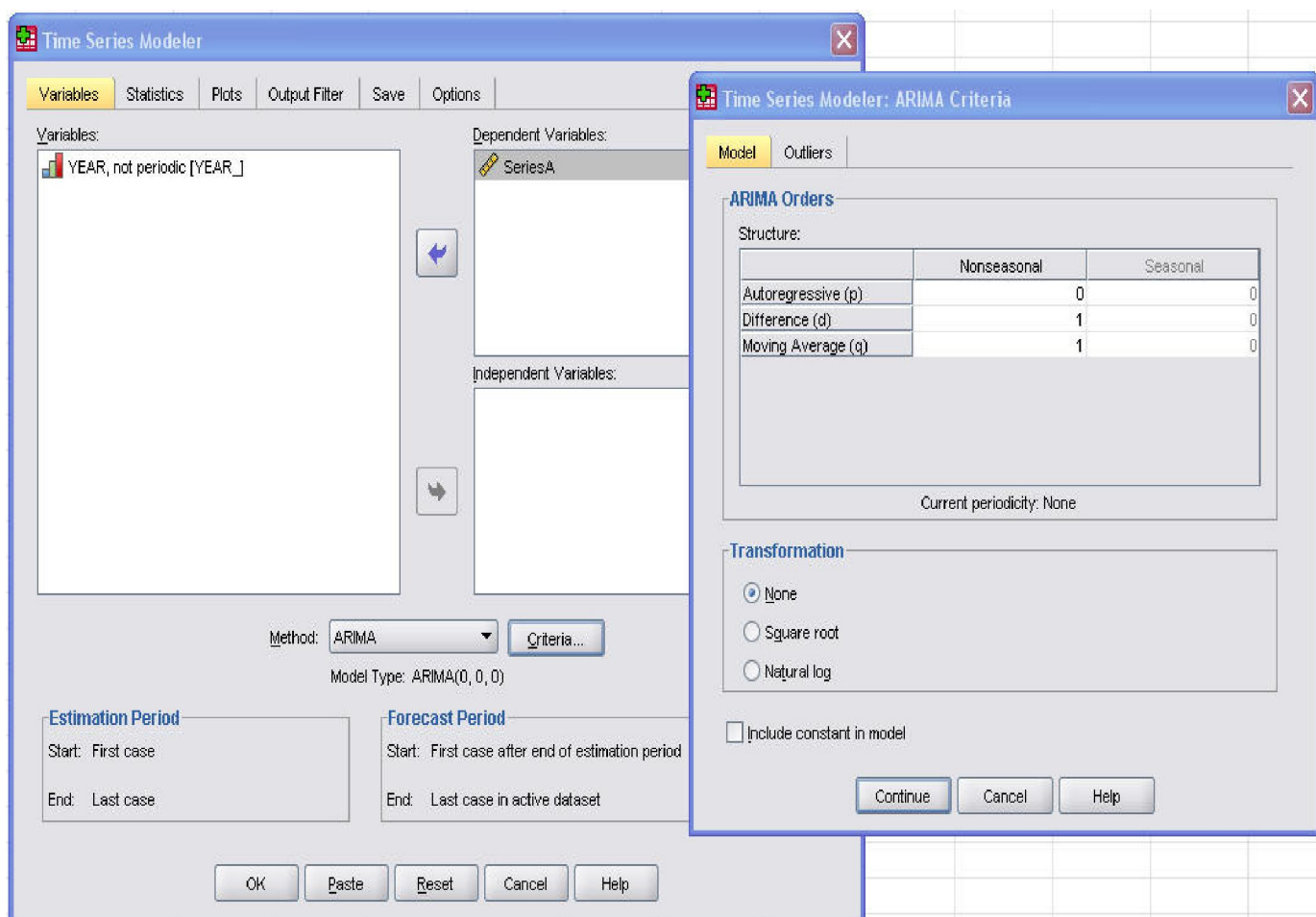


Στη συνέχεια υπολογίζουμε τις αυτοσυσχετίσεις και μερικές αυτοσυσχετίσεις της χρονοσειράς σε πρώτες διαφορές. Παρατηρούμε ότι οι αυτοσυσχετίσεις των παρατηρήσεων της χρονοσειράς σε πρώτες διαφορές φθίνουν γρήγορα και μηδενίζονται οπότε μπορούμε να θεωρήσουμε ότι η χρονοσειρά σε πρώτες διαφορές γίνεται στάσιμη. Επειδή υπάρχουν μόνο μια μη μηδενική αυτοσυσχέτιση και δύο μη μηδενικές μερικές αυτοσυσχετίσεις μπορεί να θεωρηθεί ότι η χρονοσειρά ταυτοποιείται ως μια διαδικασία κινητού μέσου πρώτης τάξης, δηλαδή $ARIMA(0, 1, 1)$.

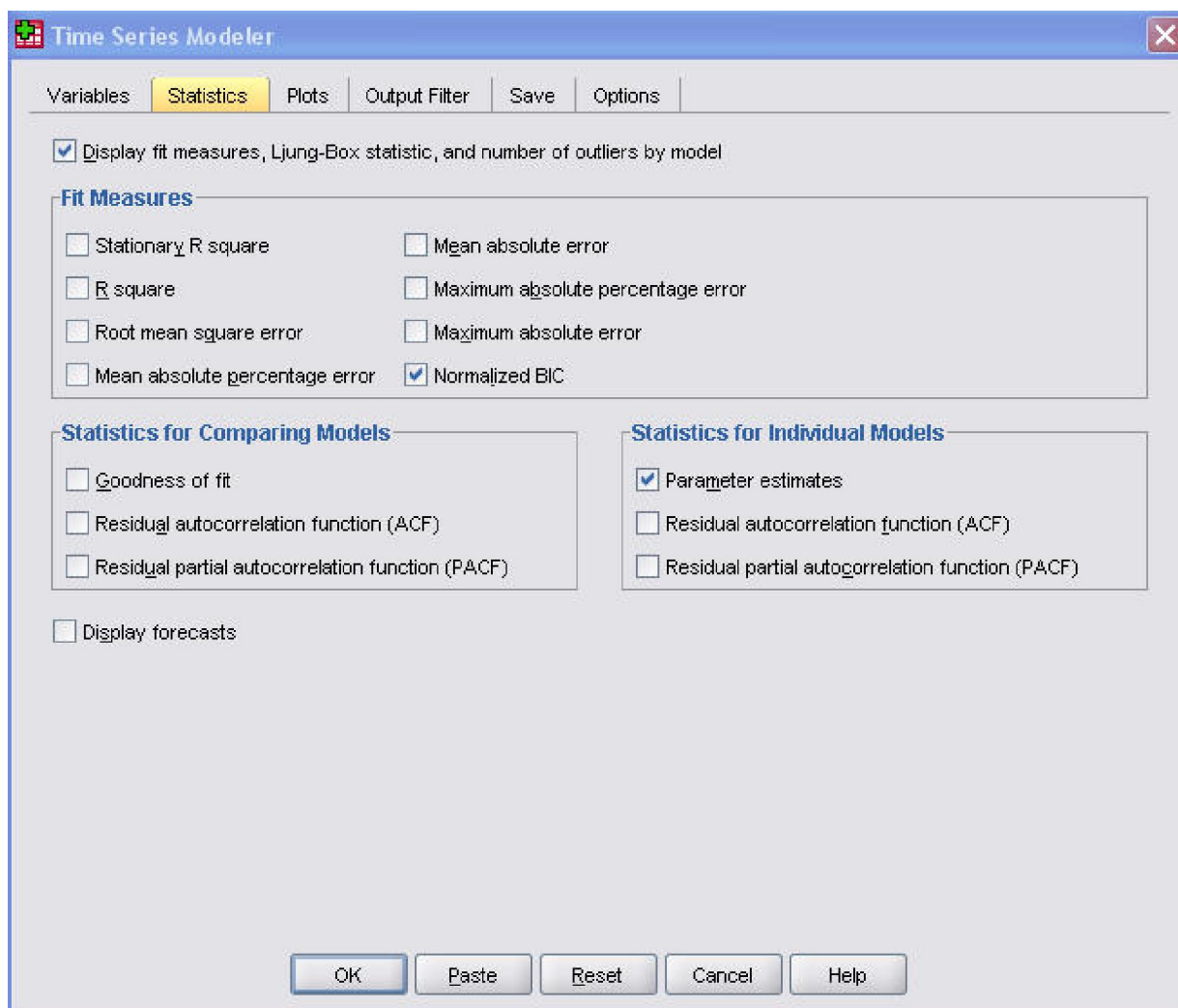


6. ΕΚΤΙΜΗΣΗ $ARIMA(p, d, q)$ ΥΠΟΔΕΙΓΜΑΤΩΝ

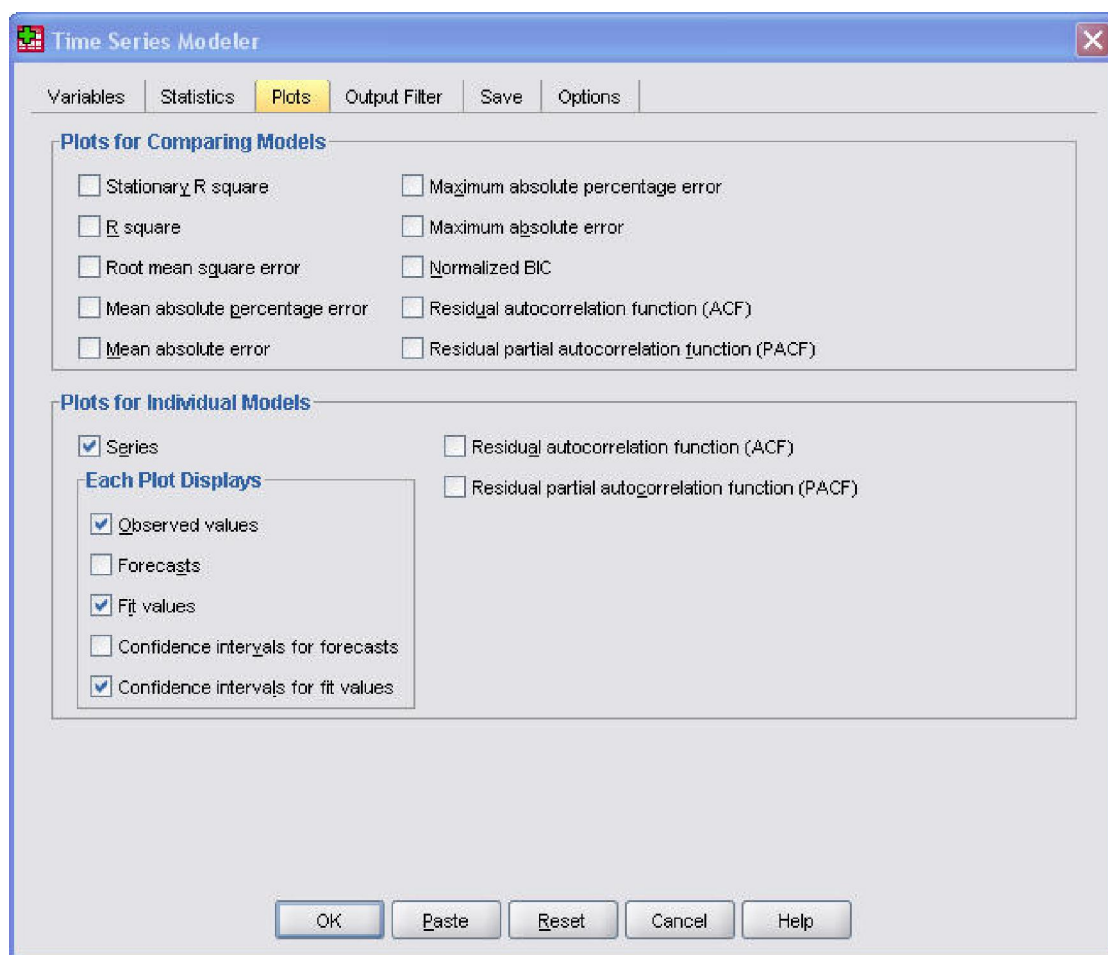
Για να εκτιμηθεί ένα $ARIMA(p, d, q)$ επιλέγουμε Analyze/Forecasting/Create Models και ανοίγουμε το παράθυρο του Time Series Modeler. Έστω ότι θέλουμε να εκτιμήσουμε ένα υπόδειγμα $ARIMA(0, 1, 1)$ για τη χρονοσειρά του παραδείγματός μας. Επιλέγουμε τη μεταβλητή SeriesA που περιέχει τις τιμές της χρονοσειράς και τη μετακινούμε στο πλαίσιο dependent Variable. Στην κυλιόμενη μπάρα Method επιλέγουμε ARIMA και πατάμε το κουμπί Criteria για να ανοίξει το παράθυρο Time Series Modeler: ARIMA Criteria. Στο πλαίσιο του ARIMA Orders Nonseasonal πληκτρολογούμε $p = 0$, $d = 1$ και $q = 1$. Στις επιλογές Transformation υπάρχει η δυνατότητα εκτίμησης του υποδείγματος μετασχηματίζοντας τις τιμές της χρονοσειράς με λογαρίθμους ή με τετραγωνικές ρίζες. Στην περίπτωση που εξετάζεται αφήνουμε την επιλογή None. Η επιλογή include constant in model αναφέρεται στο μέσο όρο της χρονοσειράς. Ειδικότερα, εάν η χρονοσειρά έχει μη μηδενικό μέσο όρο κατά τη διαδικασία της εκτίμησης, η εκτίμηση του υποδείγματος γίνεται σε αποκλίσεις από τον μέσο όρο οπότε θα πρέπει να επιλέγεται το include constant in model. Στην χρονοσειρά που εκτιμάμε εφαρμόζουμε μετασχηματισμό σε πρώτες διαφορές, δηλαδή ο μέσος όρος της χρονοσειράς θα είναι μηδέν οπότε δεν πρέπει να επιλεγεί το include constant in model. Αφού εισαχθούν οι προηγούμενες επιλογές πατάμε Continue και επιστρέφουμε στο παράθυρο του Time Series Modeler.



Στη συνέχεια πατάμε το κουμπί Statistics και μεταφερόμαστε στο παράθυρο Time Series Modeler:Statistics. Στις επιλογές Fit Measures υπάρχουν διαθέσιμα κάποια στατιστικά μέτρα που χρησιμεύουν στην αξιολόγηση της ερμηνευτικής ικανότητας του υποδείγματος που θα εκτιμηθεί και στην επιλογή του καλύτερου υποδείγματος όταν εκτιμώνται διαφορετικά υποδείγματα. Τα πιο αξιόπιστα από αυτά, για την ανάλυση χρονοσειρών, είναι τα μέτρα που ανήκουν στην κατηγορία των πληροφοριακών κριτηρίων όπως το AIC και το BIC τα οποία προσφέρονται σε παλιότερες εκδόσεις του SPSS. Μεταξύ πολλών εκτιμηθέντων υποδειγμάτων καλύτερο θεωρείται αυτό που έχει τη μικρότερη τιμή στα δύο προηγούμενα κριτήρια. Στο SPSS 17 υπάρχει μια τροποποιημένη εκδοχή του κριτηρίου BIC που ονομάζεται Normalized BIC την οποία και επιλέγουμε. Στις επιλογές Statistics for Comparing Models δεν επιλέγουμε τίποτα ενώ από τις επιλογές Statistics for Individuals Models επιλέγουμε Parameter estimates. Οι επιλογές Residual autocorrelation_function και Residual partial autocorrelation_function υπολογίζουν τις αυτοσυσχετίσεις και μερικές αυτοσυσχετίσεις των σφαλμάτων από το εκτιμηθέν υπόδειγμα. Οι αυτοσυσχετίσεις αυτές που μπορούν να υπολογιστούν και με άλλο τρόπο θα εξεταστούν αναλυτικά σε επόμενη ενότητα και για το λόγο αυτό προς το παρόν δεν θα επιλεγεί ο υπολογισμός τους.



Πατώντας το κουμπί plots μπορεί να επιλεγθεί η κατασκευή κάποιων διαγραμμάτων Αυτό που ενδιαφέρει εδώ είναι η κατασκευή ενός διαγράμματος που να απεικονίζει ταυτόχρονα τις πραγματικές τιμές της χρονοσειράς μαζί με τις εκτιμημένες τιμές της και τα κάτω και άνω όρια ενός διαστήματος εμπιστοσύνης εντός του οποίου βρίσκονται οι εκτιμήσεις. Το διάγραμμα αυτό προκύπτει διαλέγοντας από τις επιλογές του Plots for Individuals Models το Series και από τις επιλογές του Each Plot Displays τα Observed values, Fit values και Confidence intervals for fit values. Εάν επιλεγθούν και οι επιλογές Residual autocorrelation_function και Residual partial autocorrelation_function (δεν θα γίνει προς το παρόν) θα κατασκευαστούν και τα διαγράμματα των αυτοσυσχετίσεων και μερικών αυτοσυσχετίσεων των σφαλμάτων από το εκτιμηθέν υπόδειγμα.



Αφού πατηθεί το κουμπί Save στο πλαίσιο Variables του Save Variables επιλέγονται τα Predicted Values, Lower Confidence Limits, Upper Confidence Limits και Noise Residuals για να δημιουργηθούν στο data file τέσσερις νέες μεταβλητές που θα περιλαμβάνουν τις εκτιμημένες τιμές, το κάτω όριο του διαστήματος εμπιστοσύνης των εκτιμήσεων, το άνω όριο του διαστήματος εμπιστοσύνης των εκτιμήσεων και τα σφάλματα που θα προκύψουν από την εκτίμηση του υποδείγματος.

Time Series Modeler

Variables | Statistics | Plots | Output Filter | **Save** | Options

Save Variables

Variables:

Description	Save	Variable Name Prefix
Predicted Values	☒	Predicted
Lower Confidence Limits	☒	LCL
Upper Confidence Limits	☒	UCL
Noise Residuals	☒	NResidual

For each item you select, one variable is saved per dependent variable.

Export Model File

File:

Πατώντας OK στο SPSS Statistics Viewer προκύπτουν τα αποτελέσματα από την εκτίμηση του υποδείγματος. Πιο συγκεκριμένα:

Model Description

			Model Type
Model ID	SeriesA	Model_1	ARIMA(0,1,1)

Στο Model Description περιγράφεται ο τύπος του υποδείγματος που εκτιμήθηκε, δηλαδή ARIMA(0, 1, 1).

Model Statistics

Model	Number of Predictors	Model Fit statistics	Ljung-Box Q(18)			Number of Outliers
		Normalized BIC	Statistics	DF	Sig.	
SeriesA-Model_1	0	-,064	6,345	17	,991	0

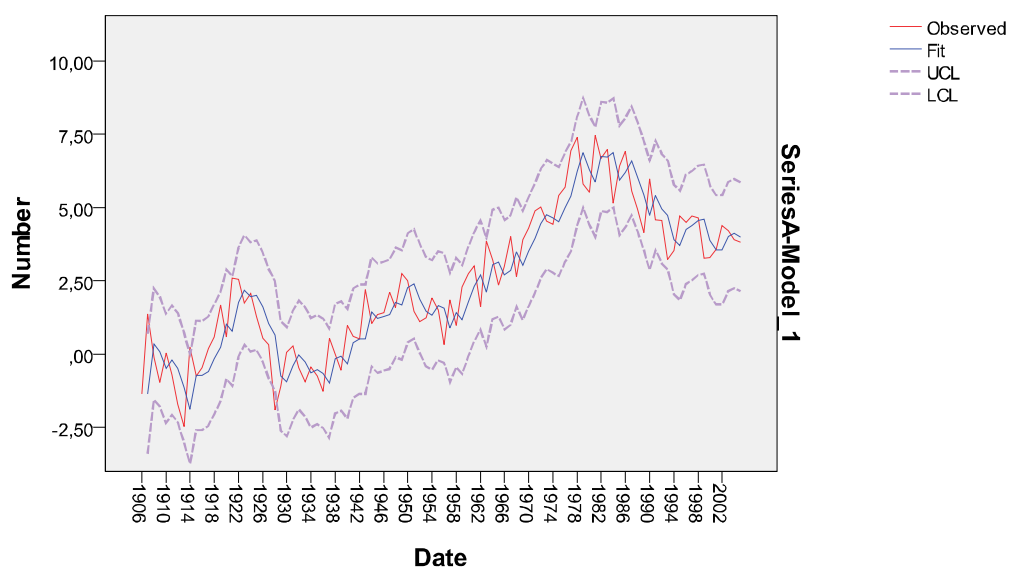
Στο Model Statistics περιλαμβάνονται η τιμή του πληροφοριακού κριτηρίου Normalized BIC και η τιμή της στατιστικής Q των Box και Ljung για 18 χρονικές υστερήσεις με το P-Value

για τον έλεγχο της μηδενικής υπόθεσης ότι η χρονοσειρά των σφαλμάτων είναι λευκός θόρυβος.

ARIMA Model Parameters

				Estimate	SE	t	Sig.
SeriesA-Model_1	SeriesA	No Transformation	Difference	1			
			MA				
			Lag 1	,455	,091	5,012	,000

Στο ARIMA Model Parameters περιλαμβάνονται οι τιμές των παραμέτρων του εκτιμηθέντος υποδείγματος, δηλαδή στην περίπτωση που εξετάζεται η τιμή του όρου του κινητού μέσου πρώτης τάξης μαζί με το τυπικό σφάλμα της εκτίμησης, την τιμή της t στατιστικής για τον έλεγχο της στατιστικής του σημαντικότητας και το P-Value του ελέγχου.



Στο προηγούμενο διάγραμμα απεικονίζονται οι τιμές της χρονοσειράς μαζί με τις εκτιμήσεις τους και τα κάτω και άνω όρια του διαστήματος εμπιστοσύνης των εκτιμήσεων. Επιπρόσθετα, στο data file έχουν δημιουργηθεί τέσσερις νέες μεταβλητές που περιλαμβάνουν τις εκτιμημένες τιμές, το κάτω όριο και το άνω όριο του διαστήματος εμπιστοσύνης των εκτιμήσεων και τα σφάλματα από την εκτίμηση του υποδείγματος.

*SeriesA.sav [DataSet1] - SPSS Statistics Data Editor																			
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help																			
1: SeriesA -1,352 Visible: 7 of 7 Variables																			
	SeriesA	YEAR_	DATE_	Predicted_SeriesA_Model_1	UCL_SeriesA_Model_1	UCL_SeriesA_Model_1	NResidual_SeriesA_Model_1	VAR	VAR	VAR	VAR	VAR	VAR	VAR	VAR	VAR	VAR	VAR	VAR
1	-1,35	1906	1906																
2	1,37	1907	1907	-1,35	-3,40	0,70	2,72												
3	-0,12	1908	1908	0,35	-1,55	2,24	-0,46												
4	-0,96	1909	1909	0,09	-1,79	1,96	-1,04												
5	0,03	1910	1910	-0,48	-2,25	1,38	0,51												
6	-0,70	1911	1911	-0,20	-2,07	1,66	-0,50												
7	-1,71	1912	1912	-0,48	-2,34	1,39	-1,23												
8	-2,48	1913	1913	-1,15	-3,01	0,72	-1,33												
9	0,23	1914	1914	-1,87	-3,74	-0,01	2,10												
10	-0,74	1915	1915	-0,73	-2,59	1,14	-0,01												
11	-0,47	1916	1916	-0,73	-2,60	1,13	0,27												
12	0,19	1917	1917	-0,69	-2,45	1,28	0,77												
13	0,58	1918	1918	-0,17	-2,03	1,70	0,74												
14	1,67	1919	1919	0,24	-1,62	2,10	1,43												
15	0,59	1920	1920	1,02	-0,85	2,88	-0,43												
16	2,59	1921	1921	0,78	-1,08	2,65	1,80												
17	2,54	1922	1922	1,77	-0,10	3,63	0,77												
18	1,75	1923	1923	2,19	0,32	4,05	-0,44												
19	2,07	1924	1924	1,95	0,09	3,81	0,12												
20	1,29	1925	1925	2,01	0,15	3,88	-0,72												
21	0,55	1926	1926	1,62	-0,24	3,48	-1,07												
22	0,32	1927	1927	1,04	-0,83	2,90	-0,72												
23	-1,90	1928	1928	0,64	-1,22	2,51	-2,55												
24	-1,11	1929	1929	-0,74	-2,61	1,12	-0,36												
25	0,07	1930	1930	-0,94	-2,80	0,92	1,01												
26	0,28	1931	1931	-0,39	-2,26	1,47	0,67												
27	-0,46	1932	1932	-0,03	-1,89	1,84	-0,43												
28	-0,95	1933	1933	-0,26	-2,12	1,60	-0,89												
29	-0,43	1934	1934	-0,64	-2,50	1,23	0,21												
30	-0,76	1935	1935	-0,62	-2,39	1,34	-0,23												
31	-1,27	1936	1936	-0,65	-2,51	1,21	-0,62												
32	0,55	1937	1937	-0,99	-2,85	0,88	1,53												
33	0,01	1938	1938	-0,15	-2,02	1,71	0,16												
34	-0,55	1939	1939	-0,07	-1,93	1,80	-0,48												
35	0,97	1940	1940	-0,33	-2,19	1,54	1,30												
36	0,62	1941	1941	0,38	-1,48	2,25	0,24												

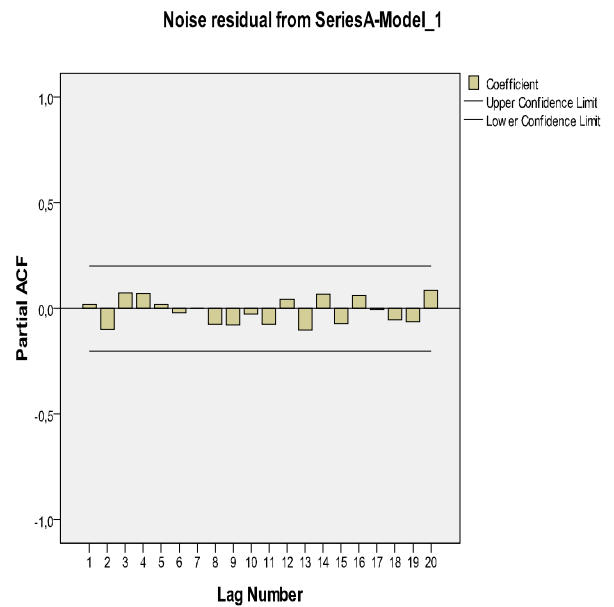
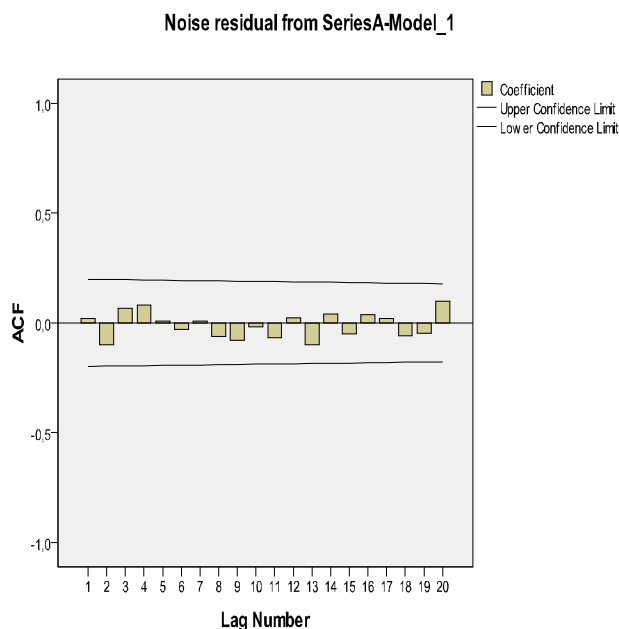
7. ΕΛΕΓΧΟΣ ΤΟΥ ΥΠΟΔΕΙΓΜΑΤΟΣ

Για να είναι το υπόδειγμα ικανοποιητικό πρέπει:

- Οι αυτοσυσχετίσεις και μερικές αυτοσυσχετίσεις της χρονοσειράς των σφαλμάτων να μην διαφέρουν σημαντικά από το 0. Μία ή δύο συσχετίσεις υψηλής τάξης μπορεί να υπερβαίνουν το 95% διάστημα εμπιστοσύνης αλλά αν η πρώτη ή η δεύτερη συσχέτιση είναι πολύ μεγάλη τότε η χρονοσειρά δεν είναι ταυτοποιημένη σωστά.
- Τα σφάλματα πρέπει να είναι λευκός θόρυβος. Ένα τεστ για αυτόν τον έλεγχο γίνεται με τη στατιστική Q των Box και Ljung. Εξετάζεται η τιμή της Q για αριθμό χρονικών υστερήσεων περίπου ίσο με το ένα τέταρτο του δείγματος (αλλά όχι περισσότερο από 50).

Ο υπολογισμός των αυτοσυσχετίσεων και μερικών αυτοσυσχετίσεων των σφαλμάτων, καθώς και των τιμών της στατιστικής Q γίνεται όπως δείχτηκε στην ενότητα 5, δηλαδή επιλέγεται Analyze/Forecasting/Autocorrelations και στο Variables μετακινείται η μεταβλητή NResidual_SeriesA_Model_1 που περιέχει τις τιμές των σφαλμάτων.

Από τα διαγράμματα των αυτοσυσχετίσεων και μερικών αυτοσυσχετίσεων διαπιστώνεται ότι δεν υπάρχουν συσχετίσεις που να υπερβαίνουν τα όρια των διαστημάτων εμπιστοσύνης. Επίσης, η υψηλή τιμή του p-value για τον έλεγχο με τη στατιστική Q των Box και Ljung οδηγεί στο συμπέρασμα ότι τα σφάλματα έχουν συμπεριφορά λευκού θορύβου.



Autocorrelations

Series: Noise residual from SeriesA-Model_1

Lag	Autocorrelation	Std. Error ^a	Box-Ljung Statistic		
			Value	df	Sig. ^b
1	,018	,099	,035	1	,853
2	-,099	,098	1,042	2	,594
3	,067	,098	1,515	3	,679
4	,081	,097	2,204	4	,698
5	,007	,097	2,209	5	,819
6	-,030	,096	2,308	6	,889
7	,007	,096	2,313	7	,941
8	-,062	,095	2,730	8	,950
9	-,080	,095	3,444	9	,944
10	-,019	,094	3,483	10	,968
11	-,067	,094	3,999	11	,970
12	,024	,093	4,062	12	,982
13	-,100	,093	5,231	13	,970
14	,041	,092	5,432	14	,979
15	-,050	,092	5,728	15	,984
16	,036	,091	5,882	16	,989
17	,020	,091	5,929	17	,994
18	-,058	,090	6,345	18	,995
19	-,048	,089	6,632	19	,996
20	,099	,089	7,864	20	,993

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Επιπρόσθετα, μια άλλη τεχνική που χρησιμοποιείται για τον έλεγχο της σωστής ταυτοποίησης της χρονοσειράς είναι η εκτίμηση ενός διευρυμένου υποδείγματος είτε από την πλευρά του αυτοπαλίνδρομου μέρους είτε από την πλευρά του μέρους των κινητών μέσων όρων. Για να είναι το εκτιμηθέν υπόδειγμα ικανοποιητικό θα πρέπει στα διευρυμένα υποδείγματα οι συντελεστές να είναι στατιστικά μη σημαντικοί. Στην περίπτωση της χρονοσειράς που εξετάζεται εκτιμήθηκε ένα υπόδειγμα $ARIMA(0, 1, 1)$ οπότε ως διευρυμένα υποδείγματα θα πρέπει να εκτιμηθούν το $ARIMA(1, 1, 1)$ και το $ARIMA(0, 1, 2)$. Εκτιμώντας τα υποδείγματα αυτό με τη διαδικασία που περιγράφηκε στην ενότητα 6 στο SPSS Statistics Viewer προκύπτουν οι παρακάτω πίνακες εκτιμήσεων:

ARIMA Model Parameters

					Estimate	SE	t	Sig.
SeriesA-Model_1	SeriesA	No Transformation	AR	Lag 1	,031	,225	,136	,892
			Difference	1				
			MA	Lag 1	,475	,199	2,394	,019

ARIMA Model Parameters

					Estimate	SE	t	Sig.
SeriesA-Model_1	SeriesA	No Transformation	Difference	1				
			MA	Lag 1	,435	,102	4,259	,000
				Lag 2	,029	,101	,281	,779

Από τα προηγούμενα αποτελέσματα διαπιστώνεται ότι στα διευρυμένα υποδείγματα οι πρόσθετοι όροι δεν είναι στατιστικά σημαντικοί. Κατά συνέπεια, η χρονοσειρά SeriesA που αναλύθηκε έχει ταυτοποιηθεί σωστά ως ARIMA(0, 1, 1) διαδικασία.

8. ΠΡΟΒΛΕΨΕΙΣ

Για να διενεργηθούν προβλέψεις για τις μελλοντικές τιμές της χρονοσειράς εκτιμάται κανονικά το υπόδειγμα όπως δείχτηκε στην ενότητα 6 αφού όμως προηγουμένως πατηθεί το κουμπί Options και στο πλαίσιο Forecast Period επιλεχθεί First case after end of estimation period through a specified date και εισαχθεί μέχρι πια χρονική περίοδο το SPSS θα υπολογίσει τις προβλέψεις. Ας υποτεθεί ότι χρειάζονται προβλέψεις μέχρι το έτος 2008, οπότε στο πλαίσιο πληκτρολογείται το 2008.

The screenshot shows the 'Time Series Modeler' dialog box with the 'Options' tab selected. The 'Forecast Period' section has two radio buttons: 'First case after end of estimation period through last case in active dataset' (unselected) and 'First case after end of estimation period through a specified date' (selected). Below the selected option is a 'Date:' label and a text box containing 'Year' and '2008'. The 'User-Missing Values' section has two radio buttons: 'Treat as invalid' (selected) and 'Treat as valid' (unselected). To the right of these are three input fields: 'Confidence Interval Width (%)' with the value '95', 'Prefix for Model Identifiers in Output' with the value 'Model', and 'Maximum Number of Lags Shown in ACF and PACF Output' with the value '24'. At the bottom are buttons for 'OK', 'Paste', 'Reset', 'Cancel', and 'Help'.

Εάν πατηθεί το κουμπί plots μπορεί να επιλεχθεί στο διάγραμμα με τις πραγματικές τιμές της χρονοσειράς, τις εκτιμήσεις και τα διαστήματα εμπιστοσύνης που έχει αναφερθεί να περιληφθούν και οι προβλέψεις.

Time Series Modeler

Variables | Statistics | **Plots** | Output Filter | Save | Options

Plots for Comparing Models

☐ Stationary R square ☐ Maximum absolute percentage error
☐ R square ☐ Maximum absolute error
☐ Root mean square error ☐ Normalized BIC
☐ Mean absolute percentage error ☐ Residual autocorrelation function (ACF)
☐ Mean absolute error ☐ Residual partial autocorrelation function (PACF)

Plots for Individual Models

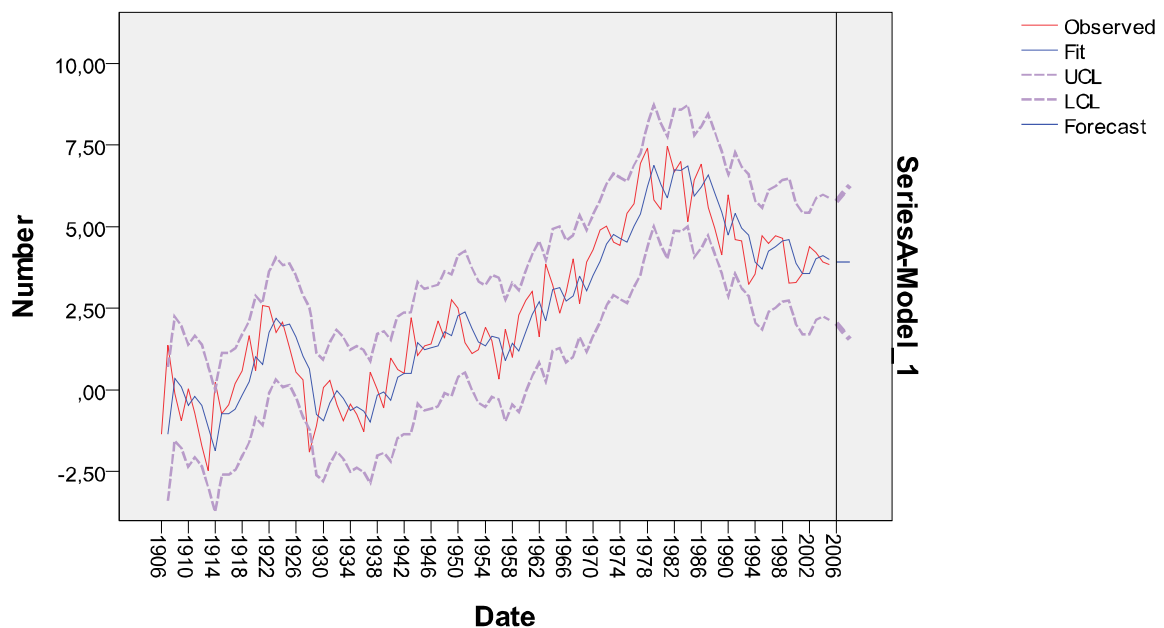
☒ Series ☐ Residual autocorrelation function (ACF)
☐ Residual partial autocorrelation function (PACF)

Each Plot Displays

☒ Observed values
☒ Forecasts
☒ Fit values
☒ Confidence intervals for forecasts
☒ Confidence intervals for fit values

OK Paste Reset Cancel Help

Πατώντας OK προκύπτουν στο SPSS Statistics Viewer τα γνωστά αποτελέσματα με τη διαφορά ότι στο διάγραμμα περιλαμβάνονται και οι προβλέψεις ενώ στο data file θα δημιουργηθούν οι μεταβλητές που όμως θα περιλαμβάνουν και τις προβλέψεις μέχρι το 2008.



SeriesA.sav [DataSet1] - SPSS Statistics Data Editor																				
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help																				
1: SeriesA -1,362																				
	SeriesA	YEAR_	DATE_	Predicted_SeriesA_Model_1	LCL_SeriesA_Model_1	UCL_SeriesA_Model_1	NResidual_SeriesA_Model_1	Var1	Var2	Var3	Var4	Var5	Var6	Var7	Var8	Var9	Var10	Var11	Var12	Var13
71	5,70	1976	1976	5,00	3,14	6,87	0,69													
72	6,94	1977	1977	5,38	3,52	7,25	1,56													
73	7,40	1978	1978	6,23	4,37	8,10	1,17													
74	5,81	1979	1979	6,87	5,00	8,73	-1,06													
75	5,52	1980	1980	6,29	4,43	8,16	-0,77													
76	7,46	1981	1981	5,87	4,01	7,74	1,59													
77	6,69	1982	1982	6,74	4,88	8,60	-0,05													
78	6,99	1983	1983	6,71	4,86	8,58	0,28													
79	5,15	1984	1984	6,86	5,00	8,73	-1,71													
80	6,42	1985	1985	5,93	4,07	7,80	0,49													
81	6,91	1986	1986	6,20	4,33	8,06	0,71													
82	5,57	1987	1987	6,59	4,72	8,45	-1,01													
83	4,95	1988	1988	6,03	4,17	7,90	-1,09													
84	4,14	1989	1989	5,44	3,58	7,31	-1,30													
85	5,97	1990	1990	4,73	2,87	6,60	1,24													
86	4,59	1991	1991	5,41	3,54	7,27	-0,82													
87	4,55	1992	1992	4,96	3,10	6,83	-0,41													
88	3,23	1993	1993	4,74	2,88	6,60	-1,51													
89	3,53	1994	1994	3,92	2,05	5,78	-0,38													
90	4,71	1995	1995	3,71	1,84	5,57	1,01													
91	4,49	1996	1996	4,26	2,39	6,12	0,23													
92	4,72	1997	1997	4,38	2,52	6,24	0,34													
93	4,64	1998	1998	4,56	2,70	6,43	0,08													
94	3,27	1999	1999	4,60	2,74	6,47	-1,34													
95	3,29	2000	2000	3,88	2,01	5,74	-0,58													
96	3,57	2001	2001	3,56	1,69	5,42	0,01													
97	4,38	2002	2002	3,57	1,70	5,43	0,82													
98	4,20	2003	2003	4,01	2,15	5,87	0,19													
99	3,91	2004	2004	4,11	2,25	5,98	-0,21													
100	3,83	2005	2005	4,00	2,14	5,86	-0,17													
101	.	2006	2006	3,91	2,04	5,77	.													
102	.	2007	2007	3,91	1,78	6,03	.													
103	.	2008	2008	3,91	1,55	6,26	.													
104																				
105																				
106																				

Όπως διαπιστώνεται, οι τιμές των προβλέψεων για τα έτη 2006, 2007 και 2008 είναι ίδιες, δηλαδή η τιμή 3,91. Αυτό οφείλεται στο γεγονός πως με ένα υπόδειγμα $ARIMA(0, 1, 1)$ μπορούν να υπολογιστούν προβλέψεις μόνοι για μια μελλοντική περίοδο.

9. Η ΧΡΟΝΟΣΕΙΡΑ SeriesA

ΕΤΟΣ	ΤΙΜΗ	ΕΤΟΣ	ΤΙΜΗ	ΕΤΟΣ	ΤΙΜΗ
1 1906	-1,352	41 1946	1,408	81 1986	6,912
2 1907	1,372	42 1947	2,113	82 1987	5,573
3 1908	-0,118	43 1948	1,584	83 1988	4,947
4 1909	-0,955	44 1949	2,758	84 1989	4,140
5 1910	0,029	45 1950	2,504	85 1990	5,972
6 1911	-0,703	46 1951	1,451	86 1991	4,592
7 1912	-1,707	47 1952	1,109	87 1992	4,554
8 1913	-2,476	48 1953	1,238	88 1993	3,227
9 1914	0,231	49 1954	1,908	89 1994	3,533
10 1915	-0,736	50 1955	1,510	90 1995	4,714
11 1916	-0,465	51 1956	0,334	91 1996	4,485
12 1917	0,185	52 1957	1,852	92 1997	4,716
13 1918	0,579	53 1958	0,988	93 1998	4,639
14 1919	1,667	54 1959	2,289	94 1999	3,268
15 1920	0,590	55 1960	2,745	95 2000	3,293
16 1921	2,586	56 1961	3,018	96 2001	3,571
17 1922	2,540	57 1962	1,620	97 2002	4,381
18 1923	1,752	58 1963	3,858	98 2003	4,200
19 1924	2,066	59 1964	3,203	99 2004	3,906
20 1925	1,290	60 1965	2,353	100 2005	3,828
21 1926	0,548	61 1966	2,997		
22 1927	0,318	62 1967	4,004		
23 1928	-1,902	63 1968	2,646		
24 1929	-1,105	64 1969	3,912		
25 1930	0,067	65 1970	4,296		
26 1931	0,280	66 1971	4,894		
27 1932	-0,455	67 1972	5,017		
28 1933	-0,953	68 1973	4,531		
29 1934	-0,430	69 1974	4,428		
30 1935	-0,756	70 1975	5,405		
31 1936	-1,271	71 1976	5,698		
32 1937	0,545	72 1977	6,942		
33 1938	0,006	73 1978	7,398		
34 1939	-0,546	74 1979	5,811		
35 1940	0,973	75 1980	5,524		
36 1941	0,618	76 1981	7,464		
37 1942	0,512	77 1982	6,688		
38 1943	2,212	78 1983	6,990		
39 1944	1,053	79 1984	5,154		
40 1945	1,349	80 1985	6,422		