

Δίνεται η ακόλουθη τριμηνιαία χρονοσειρά πωλήσεων smartphone (τεμάχια) μιας εταιρείας ηλεκτρονικών ειδών καθώς και οι εργάσιμες ημέρες (WD) των καταστημάτων της.

Τρίμηνο	Q1-21	Q2-21	Q3-21	Q4-21	Q1-22	Q2-22	Q3-22	Q4-22	Q1-23	Q2-23
Πωλήσεις	535	472	899	431	678	600	854	562		
WD	61	65	63	59	60	64	64	60	61	65

1.1. (1.0) Προσαρμόστε τη χρονοσειρά βάσει εργάσιμων ημερών και εν συνεχεία αποεποχικοποιήστε την θεωρώντας ως δείκτες εποχικότητας για τα Q1, Q2 και Q3 τις τιμές 1.15, 0.90 και 1.35 αντίστοιχα.

t	Y_t	WD_t	Y_t'	ΔE_t	Y_t''
1	535	61	545.52	1.15	474.37
2	472	65	451.67	0.90	501.85
3	899	63	887.58	1.35	657.47
4	431	59	454.38	0.60	757.29
5	678	60	702.86	1.15	611.18
6	600	64	583.13	0.90	647.92
7	854	64	829.98	1.35	614.80
8	562	60	582.61	0.60	971.01
9		61		1.15	
10		65		0.90	
Avg		62.2			

Προσαρμογή βάσει εργάσιμων ημερών: (σελ 37)
 $Y_t' = Y_t \cdot \frac{\text{Average}}{WD_t}$, $\text{Average} = \frac{\sum WD}{n} = 62.2 \leftarrow \text{ΛΑΘΟΣ}$
 μόνο 8 πρώτες

Αποεποχικοποίηση: (σελ 80) (?)
 τριμηνιαία εποχικότητας $\rightarrow n = 12/3 = 4$ περιοδοι
 $\text{avg}(\Delta E) = 1 \Rightarrow \frac{\sum \Delta E}{n} = 1 \Rightarrow$

$$\Delta E_1 + \Delta E_2 + \Delta E_3 + \Delta E_4 = 4 \Rightarrow \Delta E_4 = 4 - 3 \cdot 0.6$$

$$Y_t'' = \frac{Y_t'}{\Delta E_t}$$

1.2. (0.5) Δεδομένου ότι την περίοδο Q3-21 παρατηρείται ασυνέχεια λόγω προωθητικής ενέργειας, εξομαλύνετε την αποεποχικοποιημένη χρονοσειρά με κατάλληλες ενέργειες.

Εξομαλυνση ειδικού μεμονωτος σε αποεποχικοποιημένες τιμές: (σελ. 98)

$$D_t' = (t - t_0) \cdot \frac{D_{t_0+n} - D_{t_0}}{n+1} + D_{t_0} \Rightarrow D_3' = (3-2) \cdot \frac{D_4 - D_2}{2} + D_2 = \frac{D_2 + D_4}{2} = \frac{501.85 + 757.29}{2} = 629.57$$

1.3. (1.0) Προβλέψτε τις πωλήσεις της εταιρείας για τα πρώτα δύο τρίμηνα του 2023 χρησιμοποιώντας τη μέθοδο Holt με παραμέτρους εξομαλυνσης επιπέδου $\alpha=0.2$ και τάσης $\beta=0.1$. Στις προβλέψεις να λάβετε υπόψη την εποχικότητα της χρονοσειράς καθώς και τις εργάσιμες ημέρες της εταιρείας.

Holt Exponential Smoothing (σελ. 143)

▷ αρχικό επίπεδο S_0 : MO όλων / n πρώτων παρατηρήσεων

$$\text{επιλογή αυθαίρετα } n=4 \Rightarrow S_0 = \frac{\sum Y_t'}{4} = 597.75$$

▷ αρχική τάση T_0 : διαφορά $Y_1 - Y_2$ ή $(Y_n - Y_1)/(n-1)$

$$\text{πάλι επιλογή αυθαίρετα } n=8: T_0 = \frac{Y_8' - Y_1'}{7} = \frac{971.01 - 474.37}{7} = 70.95$$

Εξισώσεις: $e_t = y_t - \hat{y}_t$

$T_t = T_{t-1} + a \cdot b \cdot e_t$ $a=0.2, b=0.1$

$S_t = S_{t-1} + T_{t-1} + a \cdot e_t$

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

$m=1$, διότι θέλουμε αμέσως επόμενη περίοδο
 $F_t = S_{t-1} + T_{t-1}$

t	αποσποκ y_t	F_t	e_t	S_t	T_t
0				597.75	70.95
1	474.37	668.70	-194.33	629.83	67.06
2	501.85	696.89	-195.04	657.88	63.16
3	657.47				
4	757.29				
5	611.18				
6	647.92				
7	614.80				
8	971.01	851.31	119.70	880.05	50.79

Έχουμε $S_8 = 880.05$ και $T_8 = 50.79$

$F_9'' = S_8 + 1 \cdot S_8 = 930.85$

$F_{10}'' = S_8 + 2 \cdot S_8 = 981.64$

Ενοχικον + επχασιμίσ:

$F_9 = DE \cdot (WDW / Avg) \cdot F_9'' = 1.15 \cdot 61 / 62.2 \cdot 930.85 = 1,049.82$

$F_{10} = DE \cdot (WDW / Avg) \cdot F_{10}'' = 0.9 \cdot 65 / 62.2 \cdot 981.64 = 923.25$

1.4. (0.5) Η εταιρεία σκοπεύει το Q2-23 να πραγματοποιήσει μία διαφημιστική καμπάνια παρόμοιας εμβέλειας του Q3-21. Δώστε τις κριτικές σας στις προβλέψεις για το διάστημα πρόβλεψης του ερωτήματος 1.3.

Έχουμε $D_3 = 657.47$ και $D_3' = 629.57 \Rightarrow Impact_s = \frac{D_3 - D_3'}{D_3'} \cdot 100\% = 4.43\%$

$F_9' = F_9 \cdot (1 + Impact) = F_9 \cdot 1.0443 = 1,096.33$

$F_9 = F_9' \cdot \frac{WD}{Avg}$

ΘΕΜΑ 2ο (2.0 μονάδες)

Δίνεται η ακόλουθη ετήσια χρονοσειρά, οι τιμές τις οποίες παρουσιάζουν τα έσοδα (χιλ. Ευρώ) ενός οργανισμού.

Έτος	2015	2016	2017	2018	2019	2020	2021	2022
Έσοδα	254	244	222	430	340	240	330	190

$\rightarrow \hat{y}_t = 281.25 + \text{ΣΤΑΣΙΜΗ}$

2.1. (1.0) Θεωρώντας ότι η χρονοσειρά είναι στάσιμη και ο συντελεστής αυτοσυσχέτισης για υστέρηση ενός έτους ισούται με -0.159, δημιουργήστε προβλέψεις για τα επόμενα δύο έτη χρησιμοποιώντας το μοντέλο ARIMA(1,0,0).

Στάσιμη με $ACF_{k=1} = -0.159 \sim \text{ARIMA}(1,0,0)$ για 2 έτη ($k=1 \rightarrow ACF = PACF$) ισού με $AR(1)$

$\phi_1 = \rho_1 = -0.159, \mu = y_t = 281.25$

$c = \mu(1 - \phi_1) = 281.25 \cdot 1.159 = 325.97$

$(1 - \phi_1 B)y_t = c \Rightarrow y_t - \phi_1 y_{t-1} = c \Rightarrow$

$F_t = c + \phi_1 y_{t-1} \Rightarrow F_t = 325.97 - 0.159 y_{t-1}$

$F_t = 325.97 - 0.159 y_{t-1}$

t=9: $F_9 = 325.97 - 0.159 \cdot 190 = 295.76$

t=10: $F_{10} = 325.97 - 0.159 \cdot 295.76 = 278.94$

2.2. (1.0) Στην παρούσα φάση ο οργανισμός προβλέπει τα έσοδα της δραστηριότητάς του χρησιμοποιώντας το μοντέλο ARIMA(0,1,0). Θα προτείνατε τη διατήρηση της εν λόγω προσέγγισης ή την αντικατάστασή της από το μοντέλο που εκτιμήσατε στο προηγούμενο ερώτημα; Αιτιολογήστε.

μοντέλο ARIMA (0,1,0) είναι το Random Walk: $(1-B)y_t = c \rightarrow y_t = y_{t-1} + c$ με $c=0$, διότι $n=N=0$

Γενικά: ▸ διαφοροισή 1^{ης} τάξης χρησιμοποιείται για αναλογική τάση

▸ αυτοαυτοανόρθωση θεωρεί περίπου γραμμική σχέση y_t και για $d=0, c \neq 0$ τείνει σε μ

Τα δεδομένα μας είναι σείσιμα $\sim$ ARIMA (1,0,0)

Το μοντέλο ARIMA(1,0,0) λαμβάνει υπόψη τη στασιμότητα και τη χαμηλή αρνητική αυτοσυσχέτιση της σειράς. Δεδομένου ότι η χρονοσειρά είναι στάσιμη, το ARIMA(1,0,0) είναι πιο κατάλληλο για πρόβλεψη από το ARIMA(0,1,0), το οποίο μοντελοποιεί μια μη στάσιμη διαδικασία τυχαίου περιπάτου. προτείνεται η αντικατάσταση του υπάρχοντος μοντέλου με το ARIMA(1,0,0).

ΘΕΜΑ 3ο (2.5 μονάδες)

Δίνεται η ακόλουθη τετραμηνιαία χρονοσειρά ζήτησης λιπασμάτων (τόνοι) ενός προμηθευτή.

P	P1-19	P2-19	P3-19	P1-20	P2-20	P3-20	P1-21	P2-21	P3-21	P1-22	P2-22	P3-22
Τιμή	15	1	0	7	0	0	9	1	0	11	0	0

→ ΕΠΟΧΙΚΟΤΗΤΑ

3.1 (1.0) Δώστε τις προβλέψεις σας για το επόμενο έτος χρησιμοποιώντας τη μέθοδο του Croston και αξιολογήστε την ακρίβειά τους αν η πραγματική ζήτηση διαμορφώθηκε στους 11,0,0 τόνους.

Croston: (σελ 227)

$$S_0 = \bar{y}_t \quad F_t = S_{t-1}$$

$$a = 0.05 \quad e_t = y_t - F_t$$

$$F_{\text{Croston}} = \frac{F_{\text{dem}}}{F_{\text{int}}} \quad S_t = S_{t-1} + a \cdot e_t$$

$$S_0^{\text{dem}} = \bar{y}_{\text{dem}} = 7.33$$

$$S_0^{\text{int}} = \bar{y}_{\text{int}} = 1.67$$

$$F_{\text{Croston}} = \frac{7.32}{1.67} = 4.37 = F_2 = F_3 = F_{14}$$

y_t	F_t	e_t	e_t^2
-------	-------	-------	---------

$$11 \quad 4.37 \quad 6.63 \quad 43.96$$

$$0 \quad 4.37 \quad -4.37 \quad 19.10$$

$$0 \quad 4.37 \quad -4.37 \quad 19.10$$

Demands

y_t	F_t	e_t	S_t
-	-	-	7.33
15	7.33	7.67	7.72
1			
7			
9			
1			
11	7.12	3.88	7.32
	7.32		

Intervals

y_t	F_t	e_t	S_t
-	-	-	1.67
1	1.67	-0.67	1.63
1			
2			
3			
1			
2	1.66	0.34	1.67
	1.67		

$$\text{ακρίβεια: } MAE = \frac{1}{n} \sum |y_t - F_t| = \frac{1}{n} \sum |e_t| = 5.11$$

$$MSE = \frac{1}{n} \sum (y_t - F_t)^2 = \frac{1}{n} \sum e_t^2 = 27.39$$

3.2 (1.5) Εφαρμόστε τη μέθοδο ADIDA αιτιολογώντας το επίπεδο συνάθροισης και τη μέθοδο διαχωρισμού που επιλέξατε και υπολογίστε τη σχετική ακρίβεια της μεθόδου σε σχέση με αυτή του Croston.

ADIDA (σελ 233)

▸ επίπεδο συνάθροισης (aggregation level): $m = 3$ ημέρες λόγω εποχικότητας

Έχουμε $n = 12$ παρατηρήσεις, άρα θα αγνοήσουμε τις πρώτες $n \bmod m = 0$ παρατηρήσεις και η νέα χρονοσειρά θα αποτελείται από $n' = n/m = 12/3 = 4$ παρατηρήσεις

▸ μέθοδος πρόβλεψης (forecasting method): KMO(3)

D_t	D_t^*	$KMO(3)$	w_i	F_t	e_t
15					
1	16				
0					
7					
0	7				
0					
9					
1	10				
0					
11		$(16+7+10)/3$			
0	11	11			
0					
11		$(9+10+10)/3$	951	887	213
0		933	51	047	-0.47
0		0	0	0	0

► μέθοδος διαχωρισμοί (disaggregation method):

ΕΠΟΧΙΚΟΤΗΤΑ $\Rightarrow$ Average weights

Συνολική ζήτηση P1: $15+7+9+11=42 \Rightarrow w_1 = 42/44=0.95$

Συνολική ζήτηση P2: $1+0+1+0=2 \Rightarrow w_2 = 2/44=0.05$

Συνολική ζήτηση P3: $0+0+0+0=0 \Rightarrow w_3 = 0/44=0$

Σχετική ακρίβεια:

$MAE_{crossin} = |ln \cdot \sum |e_i| = 5.12$

$MAE_{ADIDA} = 0.87$

σχετική ακρίβεια = $\frac{MAE_{crossin}}{MAE_{ADIDA}} = 5.88$

Η ADIDA είναι 5.88 φορές πιο ακριβής με βάση το μέσο απόλυτο σφάλμα (MAE)

ΘΕΜΑ 4ο (2.5 μονάδες)

Δίνεται η ακόλουθη ετήσια χρονοσειρά παραγωγής μπαμπακιού (τόνοι) μιας γεωργικής μονάδας.

Τρίμηνο	2016	2017	2018	2019	2020	2021	2022
Τιμή	9	12	25	13	8	19	4

4.1 (1.5) Εφαρμόστε τη μέθοδο πρόβλεψης Theta για την πρόβλεψη της παραγωγής τα επόμενα δύο έτη.

Χρησιμοποιείτε δύο γραμμές Theta με παραμέτρους $\theta_1=0$ και $\theta_2=1.5$. Για την προέκταση της γραμμής με παράμετρο θ_2 χρησιμοποιείτε ως μέθοδο πρόβλεψης έναν ΚΜΟ(3). Δίνεται η εξίσωση της ευθείας γραμμικής παλινδρόμησης $Y=15.43-0.64 \cdot t$.

$\theta_1=0$ → LRL: Περιγράφει χρονοσειρά μέσω γραμμικής τάσης.

Προεκτείνεται μέσω αυτής γραμμικής παλινδρόμησης (LRL)

Εξίσωση $LRL_1: Y_t^{\theta=0} = a + bt = 15.43 - 0.64t$

$\theta_2=1.5$ → 15 φορές μεγαλύτερες συνεχείς διαφορές από και κληνδόντες σε σχέση με αρχική, ενισχύονται βραχυπρόθεσμη συμπεριφορά

$Y_t^{\theta=1.5} = \theta \cdot Y_t + (1-\theta) \cdot Y_t^{\theta=0} = 15 Y_t - 0.5 Y_t^{\theta=0}$

Προεκτείνεται με ΚΜΟ(3)

Βήμα 1: αποεποχικοποίηση (δεν χρειάζεται)

Βήμα 2: αποσύνθεση

t	y_t	$y_t^{0:0}$	$y_t^{0:15}$
1	9	14.79	6.11
2	12	14.15	10.93
3	25	13.51	30.75
4	13	12.87	13.07
5	8	12.23	5.89
6	19	11.59	22.71
7	4	10.95	0.53

$$y_t^{0:0} = a + bt = 15.43 - 0.64t$$

$$y_t^{0:15} = 15 y_t - 0.5 y_t^{0:0}$$

Βήμα 3: Πρόβλεψη

$\Theta = 0$: αντί γραμ παλινδρό

$$t=8: F_8^{0:0} = 10.31$$

$$t=9: F_9^{0:0} = 9.67$$

$\Theta = 15$: KMO (3)

$$t=8: F_8^{0:15} = \frac{5.89 + 22.71 + 0.53}{3} = 9.71$$

$$t=9: F_9^{0:15} = 9.71$$

Βήμα 4: συνδιασμός

t	$F_t^{0:0}$	$F_t^{0:15}$	F_t
8	10.31	9.71	10.01
9	9.67	9.71	9.69

$$F_t = \frac{F_t^{0:0} + F_t^{0:15}}{2}$$

Βήμα 5: ενοχοποίηση ~ δεν έχει

$$\text{Τελική πρόβλεψη: } \begin{cases} 2023: F_8 = 10.01 \\ 2024: F_9 = 9.69 \end{cases}$$

4.2 (1.0) Υπολογίστε το διάστημα εμπιστοσύνης των προβλέψεών σας θεωρώντας επίπεδο 90%.

(οεί 116)

t	y_t	$y_t^{0:0}$	$y_t^{0:15}$	F_t	e_t	e_t^2
1	9	14.79	6.11	10.45	-1.45	2.10
2	12	14.15	10.93	12.54	-0.54	0.29
3	25	13.51	30.75	22.13	2.87	8.25
4	13	12.87	13.07	12.97	0.03	0
5	8	12.23	5.89	9.06	-1.06	1.12
6	19	11.59	22.71	17.15	1.85	3.43
7	4	10.95	0.53	5.74	-1.74	3.02

$$90\% \rightarrow t = 1.645$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum e_t^2} = 1.61$$

$$F_i = F_i \pm t \cdot RMSE \cdot \sqrt{i-n}$$

$$F_8 = 10.01 \pm 1.645 \cdot 1.61 \cdot \sqrt{8-7} = 10.01 \pm 2.65$$

$$F_9 = 9.69 \pm 1.645 \cdot 1.61 \cdot \sqrt{9-7} = 9.69 \pm 3.74$$