

1.1 (0.4) Αναφέρετε τα ποιοτικά χαρακτηριστικά χρονοσειρών που αναγνωρίζετε στα δεδομένα του διπλανού διαγράμματος.

Τα ποιοτικά χαρακτηριστικά μιας χρονοσειράς είναι: (σελ 17)

1. Στασιμότητα: Οι στατιστικές ιδιότητες (μέσος όρος, διακύμανση, αυτοσυσχέτιση) παραμένουν σταθερές στο χρόνο. Η δεδομένη χρονοσειρά δεν είναι αυστηρά στατική λόγω του μεταβαλλοτού πλάτους και της ύπαρξης διαρθρωτικών αλλαγών.
2. Εποχιακότητα: Τακτικά, προβλέψιμα μοτίβα ή διακυμάνσεις που επαναλαμβάνονται σε καθορισμένα χρονικά διαστήματα (π.χ. μηνιαία, τριμηνιαία). Η δεδομένη χρονοσειρά παρουσιάζει εποχιακότητα, καθώς υπάρχουν σαφείς επαναλαμβανόμενες αιχμές περίπου μία φορά το χρόνο, που υποδηλώνουν εποχιακές επιδράσεις.
3. Κυκλικότητα: Μακροπρόθεσμα κυματοειδή μοτίβα που δεν έχουν σταθερή συχνότητα και συχνά επηρεάζονται από οικονομικούς ή συστημικούς παράγοντες. Στη δεδομένη χρονοσειρά, δεν υπάρχουν ενδείξεις μακροπρόθεσμων κύκλων πέραν της εποχικότητας.
4. Τάση: Μακροπρόθεσμη κίνηση ή κατεύθυνση των δεδομένων, είτε ανοδική είτε καθοδική. Στη δεδομένη χρονοσειρά, δεν παρατηρείται σταθερή ανοδική ή καθοδική τάση με την πάροδο του χρόνου· η τάση είναι ελάχιστη ή απουσιάζει.
5. Ασυνέχειες: Ξαφνικές, απότομες αλλαγές ή διακοπές στη χρονοσειρά λόγω εξωτερικών κλυδωνισμών ή διαρθρωτικών αλλαγών. Η δεδομένη χρονοσειρά έχει απότομες αιχμές, ιδίως γύρω στο 1991 και το 1992, υποδηλώνουν απότομα γεγονότα ή ακραίες τιμές.
6. Τυχαιότητα: Ανεξήγητη διακύμανση ή θόρυβος στα δεδομένα που δεν μπορεί να αποδοθεί σε άλλα στοιχεία. Η δεδομένη χρονοσειρά έχει ακανόνιστες διακυμάνσεις εκτός του εποχιακού προτύπου, που υποδηλώνουν στοχαστικό θόρυβο.

1.2 (0.5) Δίνεται η παρακάτω τριμηνιαία χρονοσειρά πωλήσεων (εκ. €) μίας εταιρείας. Όπως φαίνεται, τις περιόδους Q2, Q3 και Q4 2020 οι πωλήσεις της εταιρείας μειώθηκαν σημαντικά καθώς, λόγω COVID-19, τα καταστήματά της παρέμειναν κλειστά αρκετές ημέρες. Δώστε κριτικές προβλέψεις για το 2021 θεωρώντας πως τα καταστήματα της εταιρείας θα παραμείνουν κλειστά μόνο το τέταρτο τρίμηνο του έτους.

2018				2019				2020				2021			
Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
102	105	95	128	100	103	104	117	95	35	55	35	99	104	100	35
a	a	a	a	a	a	a	a	a	k	k	k	a	a	a	k

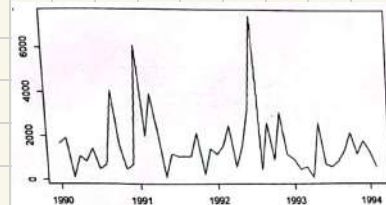
κριτική πρόβλεψη: (σελ 112)

$$\text{πρόβλεψη } Q1-2021 = \text{Avg}(Q1: \text{ανοίχτα}) = (102 + 100 + 95) / 3 = 99$$

$$\text{πρόβλεψη } Q2-2021 = \text{Avg}(Q2: \text{ανοίχτα}) = (105 + 103) / 2 = 104$$

$$\text{πρόβλεψη } Q3-2021 = \text{Avg}(Q3: \text{ανοίχτα}) = (95 + 104) / 2 = 99.5 \approx 100$$

$$\text{πρόβλεψη } Q4-2021 = \text{Avg}(Q4: \text{κλειστά}) = 35$$



1.3 (0.3) Δίνεται η παρακάτω ετήσια χρονοσειρά πωλήσεων. Προβείτε σε κατάλληλες ενέργειες για να συμπληρώσετε τις τιμές που λείπουν (NA). Αιτιολογήστε.

Sales 91 (NA) 164 161 107 153 (NA) 101 105 176

$$(91 + 164) / 2 = 127.5$$

$$(153 + 101) / 2 = 127$$

όχι τάση

όχι εποχικότητα

} άγνωστες τιμές = γραμμική παρεμβολή

(υποθέτει σταθερό ρυθμό μεταβολής)

$$a \quad \underbrace{x \quad x \quad x}_{n=3} \quad b \quad \sim \quad y_t = y_{t_0} + (t - t_0) \cdot \frac{y_{t_0+n+1} - y_{t_0}}{n+1} \quad \text{εδώ } n=1 \Rightarrow y_t = y_{t_0} + \frac{y_{t_0+2} - y_{t_0}}{2} = \frac{y_{t_0} + y_{t_0+2}}{2}$$

1.4 (0.3) Για μία τριμηνιαία εποχιακή χρονοσειρά υπολογίστηκαν οι παρακάτω δείκτες εποχικότητας (ΔE). Υπολογίστε την τιμή του λόγου εποχικότητας που λείπει. Αιτιολογήστε.

ΔE : 1.2, 0.8, NA, 0.9

τριμηνιαία εποχικότητας $\rightarrow n=12/3=4$ περιόδοι

$$\text{avg}(\Delta E) = 1 \Rightarrow \frac{\sum \Delta E}{n} = 1 \Rightarrow \Delta E_1 + \Delta E_2 + \Delta E_3 + \Delta E_4 = 4 \Rightarrow \Delta E_3 = 4 - 2.9 = 1.1$$

ΘΕΜΑ 2ο (2.5 μονάδες)

Ημέρα	1	2	3	4	5	6	7	8	9	10	11	12	13
Τιμή	1	0	0	4	3	0	5	0	2	0	0	3	2

2.1 (1.0) Δώστε τις προβλέψεις σας για τις επόμενες τρεις ημέρες χρησιμοποιώντας τη μέθοδο του Croston.

Croston: (σελ 227)

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

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

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

$$S_0^{\text{dem}} = \bar{y}_{\text{dem}} = 2.86$$

$$S_0^{\text{int}} = \bar{y}_{\text{int}} = 1.86$$

$$F_{\text{Croston}} = \frac{2.86}{1.87} = 1.53$$

$$\Rightarrow F_{11} = F_{15} = F_{16} = 1.53$$

Demands

y_t	F_t	e_t	S_t
-	-	-	2.86
1	2.86	-1.86	2.77
4	2.77	1.23	2.83
3	2.83	0.17	2.84
5	2.84	2.16	2.95
2	2.95	0.95	2.90
3	2.90	0.1	2.91
2	2.91	-0.91	2.86
	2.86		

Intervals

y_t	F_t	e_t	S_t
-	-	-	1.86
1	1.86	-0.86	1.82
3	1.82	1.18	1.88
1	1.88	-0.88	1.84
2	1.84	0.16	1.85
2	1.85	0.15	1.86
3	1.86	1.14	1.92
1	1.92	-0.92	1.87
	1.87		

2.2 (0.5) Σύμφωνα με τους Syntetos και Boylan, η μέθοδος SBA μπορεί να παράξει πιο αμερόληπτες προβλέψεις σε σχέση με τη μέθοδο του Croston. Αξιολογήστε τον εν λόγω ισχυρισμό θεωρώντας ότι η πραγματική ζήτηση στο προηγούμενο ερώτημα διαμορφώθηκε ως $\{0, 0, 1\}$.

SBA: χρονοποίηση της Croston (σελ 232)

$$F_{\text{SBA}} = \left(1 - \frac{\alpha}{2}\right) \cdot F_{\text{Croston}} = 97.5\% \cdot F_{\text{Croston}}$$

$$\text{Μέτρηση bias: } ME = \frac{1}{n} \sum (y_t - F_t) = \frac{1}{n} \sum e_t$$

$$MA_{\text{Croston}} = \frac{1}{3} \sum e_{t, \text{Croston}} = -1.20$$

$$MA_{\text{SBA}} = \frac{1}{3} \sum e_{t, \text{SBA}} = -1.16$$

$\triangleright ME < 0 \Rightarrow$ δηλώνει αυστοδοξία / υπερεκτίμηση

$\triangleright |ME_{\text{SBA}}| < |ME_{\text{Croston}}|$: SBA μικρότερη προκατάληψη

2.3 (1.0) Εφαρμόστε τη μέθοδο ADIDA χρησιμοποιώντας κατάλληλο επίπεδο συνάθροισης, μέθοδο διαχωρισμού και μέθοδο προέκτασης. Ποια η σχετική ακρίβεια της μεθόδου σε σχέση με τη μέθοδο του Croston;

ADIDA (σελ 235)

D_t	D_t^m	$KMO(3)$	F_t	e_t
1	-	-		
0				
0	4			
4				
3				
0	8			
5				
0				
2	2			
0				
0		$(4+8+2)/3$		
3	5	4.67		
2				
0		$(8+2+5)/3$	1.67	-1.67
1		5	1.67	-0.67
0			1.67	-1.67

▷ ενιμεδο συναθροισμός (aggregation level): $m=3$ ημέρες

Έχουμε $n=13$ παρατηρήσεις, άρα θα αγνοήσουμε τις πρώτες $n \bmod m = 13 \bmod 3 = 1$ παρατηρήσεις και η νέα χρονοσειρά θα αποτελείται από $n' = n/m = 13/3 = 4$ παρατηρήσεις

▷ μέθοδος πρόβλεψης (forecasting method): $KMO(3)$

▷ μέθοδος διαχωρισμού (disaggregation method): όχι ενοχικίζηση
⇒ equal weights

$$\text{Σύγκριση ακριβείας: } MSE = \frac{1}{n} \sum (y_t - \hat{y}_t)^2 = \frac{1}{n} \sum e_t^2$$

$$MSE_{croston} = \frac{1}{3} \sum e_t^{croston^2} = 1.65 \quad MAE_{croston} = \frac{1}{n} \sum |e_t| = 1.197$$

$$MSE_{ADIDA} = \frac{1}{3} \sum e_t^{ADIDA^2} = 2.01 \quad MAE_{ADIDA} = \frac{1}{n} \sum |e_t| = 1.337$$

$$MSE_{ADIDA} > MSE_{croston} \Rightarrow \text{Croston πιο ακριβής}$$

$$\text{σχετική ακρίβεια} = \frac{MAE_{croston}}{MAE_{ADIDA}} = 0.895$$

Η Croston είναι περίπου $1/0.895 = 1.12$ φορές πιο ακριβής από ADIDA

ΘΕΜΑ 3ο (2.5 μονάδες)

Έτος	2014	2015	2016	2017	2018	2019	2020	2021	2022
Τιμή	373	399	481	492	567	599	679	708	823

ΟΧΙ ΣΤΑΣΙΜΗ

→ σταθερή ανοδική τάση

3.1 (1.5) Αξιολογήστε ποιοτικά τη στασιμότητα της χρονοσειράς και προβλέψτε τις τιμές της για τα έτη 2023 και 2024 εφαρμόζοντας το μοντέλο $ARMA(1,0)$. Κατά την εκτίμηση, αιτιολογήστε τη χρήση ή μη σταθεράς.

$ARMA(1,0)$: αυτοπαλινδρόμικο μοντέλο 1^{ης} τάξης χωρίς όρο κινητού μέσου, δηλαδή $AR(1)$.

μοντέλο $AR(1)$: $(1 - \phi_1)Y_t = c \Rightarrow Y_t - \phi_1 Y_{t-1} = c \Rightarrow F_t = c + \phi_1 Y_{t-1}$

$d=0$ $\begin{cases} c=0: \text{μακροπρόθεσμα οι προβλέψεις} \rightarrow 0, \text{δεν το θέλουμε} \\ c \neq 0: \text{μακροπρόθεσμα οι προβλέψεις} \rightarrow \text{μέση τιμή χρονοσειράς} \end{cases}$

Η σειρά ΔΕΝ είναι στασιμη $\Rightarrow$ εφαρμόζουμε διαφορίση 1^{ης} τάξης

Διαφ $\mu = \text{Avg}(Y_t') = 56.25$

Έτος	Y_t	Y_t'	$Y_t' - \mu$	$(Y_t' - \mu) \cdot (Y_{t-1}' - \mu)$	$(Y_t' - \mu)^2$
2014	373	-	-	-	-
2015	399	26	-30.25	-	915
2016	481	82	25.75	-778.94	663
2017	492	11	-45.25	-1165.19	2048
2018	567	75	18.75	-848.44	352
2019	599	32	-24.25	-454.69	588
2020	679	80	23.75	-575.94	564
2021	708	29	-27.25	-647.19	743
2022	823	115	58.75	-1600.94	3452
Sum			-	-6071.31	9324

← για υπολογισμό ϕ_1 $\otimes$ αναμένεται καλάντζωση μεταξύ θετικών λαμπητικών

$$\phi_1 = \rho_1 = \frac{\sum_{t=1}^n (Y_t' - \mu)(Y_{t-1}' - \mu)}{\sum_{t=1}^n (Y_t' - \mu)^2} = \frac{-6071.31}{9324} = -0.65 < 0 \quad \swarrow$$

$$c = \mu \cdot (1 - \phi_1) = 56.25 \cdot (1 + 0.65) = 92.87$$

↪ στασιμη χρονος

$$F_t = c + \phi_1 Y_{t-1} \rightarrow F_t = 42.87 - 0.65 Y_{t-1} \rightarrow \begin{cases} F_{2023} = 42.87 - 0.65 \cdot 823 = -442.08 & (*) \\ F_{2024} = 42.87 + 0.65 \cdot 442.08 = 380.22 \end{cases}$$

3.2 (1.0) Υπολογίστε την τιμή του κριτηρίου AIC για το μοντέλο του προηγούμενου ερωτήματος καθώς και για το μοντέλο ARIMA(1,1,0). Βάσει αποτελέσματος, ποιο μοντέλο θα επιλέγατε εν τέλει;

$d=1$: $\begin{cases} c=0: \text{μακροση προβλέψεις ακολ' ειδικά γραμμή} \\ c=0: \text{μακροση προβλέψεις ισούνται με σταθερά } \neq 0 \end{cases}$

Παράγοντες AR(p) Διαφόριση(d) Σταθερά Παράγοντες MA(q)

$$y_{2021} = 708, y_{2022} = 823 \Rightarrow \begin{cases} F_{2023} = 0.35 \cdot 708 + 0.65 \cdot 823 + 92.87 = 841.12 \\ F_{2024} = 0.35 \cdot 823 + 0.65 \cdot 841.12 = 922.21 \end{cases}$$

$$\begin{aligned} AIC &= -2 \log l + 2(p+q+P+Q+k+1) & \mu \varepsilon & \begin{cases} k=0, c=0 \\ k=1, c \neq 0 \end{cases} \\ AIC &= -2 \log l + 2(11+0+1+0+1) \end{aligned}$$

$$AIC = -2 \cdot \log L + 4 \text{ (και για τα 2)}$$

$$MSE = \frac{1}{n} \sum e_i^2$$

μέθοδος Theta: $y_t^{\theta} = \theta \cdot y_t + a_0 + b_0(t-1)$

Τρίμηνο	A 2021	B 2021	Γ 2021	Δ 2021	A 2022	B 2022	Γ 2022	Δ 2022	A 2023
Τιμή	22	33	39	28	25	39	54	37	24

4.1 (1.5) Εφαρμόστε τη μέθοδο Theta για την πρόβλεψη των Β, Γ και Δ τριμήνων του 2023, αποσυνθέτοντας την αρχική χρονοσειρά σε δύο γραμμές Theta με παραμέτρους $\theta_1 = 0$ και $\theta_2 = 2.5$. Χρησιμοποιήστε ως παράμετρο εξομάλυνσης για την Απλή Εκθετική Εξομάλυνση $\alpha = 0.5$ και αρχικό επίπεδο της επιλογής σας. Δίνεται η εξίσωση της ευθείας γραμμικής παλινδρόμησης που περνά βέλτιστα από τα αποεποχικοποιημένα δεδομένα: $Y = 1.2 \cdot t - 29.8$ (τιμές σε συνάρτηση με το χρόνο)

$$DE: 0.6 \quad 1.1 \quad 1.4 \quad 0.9 \Rightarrow \sum DE = n = 4$$

Προεκτινεται μέσω αυτής γραμμικής παλινδρόμησης (LRL)
 Εξίσωση $LRL_1 = Y_{1.0.0} = at + bt = 12t + 298$

$$y_t^{0=2.5} = \theta \cdot y_t + (1-\theta) \cdot y_t^{0=0} = 25y_t - 15 \cdot y_t^{0=0}$$

Προεκτείνεται με άλλη εκθ. εφοδιαστική, $a=0.5$ (σελ. 134)

Βήμα 1: ανοσοχρηματοδότηση

t	Y_t	SI_t	$Y_t' = Y_t / SI_t$
1	22	0.6	36.67
2	33	1.1	30
3	39	1.4	27.86
4	28	0.9	31.11
5	25	0.6	41.67
6	39	1.1	35.45
7	54	1.4	38.57
8	37	0.9	41.11
9	24	0.6	40

Βήμα 2: ανοσοβύθισση

t	Y_t'	$Y_t^{0.25}$	$Y_t^{0.25}$
1	36.67	31	45.17
2	30	32.2	26.7
3	27.86	33.4	19.55
4	31.11	34.6	25.88
5	41.67	35.8	50.48
6	35.45	37	33.13
7	38.57	38.2	39.13
8	41.11	39.4	43.68
9	40	40.6	39.10

$$Y_t^{0.25} = LRL_t = 12t + 29.8$$

$$Y_t^{0.25} = 25Y_t - 1.5Y_t^{0.25}$$

Βήμα 3: Πρόβλεψη

$\Theta=0$: αλληλ. χρημ. παροχρηματοδότηση

$$t=10: F_{10}^{0.25} = 41.8$$

$$t=11: F_{11}^{0.25} = 43.0$$

$$t=12: F_{12}^{0.25} = 44.2$$

$\Theta=2.5$: αλληλ. εκθ. εφοδιασμού

αρχικό ενδεκά:

$$S_0 = \frac{\sum_{i=1}^4 Y_i^{0.25}}{4} = \frac{45.17 + 26.7 + 19.55 + 25.88}{4} = 29.32$$

t	$Y_t^{0.25}$	$F_t^{0.25}$	e_t	S_t
0				29.32
1	45.17	29.32	15.85	37.25
2	26.7	37.25	-10.55	31.98
3	19.55	31.98	-12.43	25.76
4	25.88	25.76	0.12	25.82
5	50.48	25.82	24.66	38.15
6	33.13	38.15	-5.02	35.64
7	39.13	35.64	3.49	37.39
8	43.68	37.39	6.29	40.54
9	39.10	40.54	-1.44	39.82
		39.82		

$$F_t^{0.25} = S_{t-1}$$

$$e_t = Y_t^{0.25} - F_t^{0.25}$$

$$S_t = S_{t-1} + a \cdot e_t \rightarrow a = 0.5$$

$\Theta=2.5$

$$t=10: F_{10}^{0.25} = 39.82$$

$$t=11: F_{11}^{0.25} = 39.82$$

$$t=12: F_{12}^{0.25} = 39.82$$

Βήμα 4,5: συνδιασμός + ενοχοποίηση

t	$F_t^{0.25}$	$F_t^{0.25}$	F_t'	SI_t	F_t
10	41.8	39.82	40.81	1.1	44.89
11	43.0	39.82	41.41	1.4	57.98
12	44.2	39.82	42.01	0.9	37.81

$$F_t' = \frac{F_t^{0.25} + F_t^{0.25}}{2}$$

$$F_t = F_t' \cdot SI_t$$

Τελική πρόβλεψη:

$$B_{2023} = 44.89$$

$$F_{2023} = 57.97$$

$$D_{2023} = 37.81$$

4.2 (1.0) Υπολογίστε τα διαστήματα εμπιστοσύνης των προβλέψεων που παρήχθησαν με τη μέθοδο Theta για επίπεδο εμπιστοσύνης 80%.

t	Y_t	SI _t	Y_t'	$Y_t^{0=0}$	$Y_t^{0=25}$	F_t'	F_t	e_t
1	22	0.6	36.67	31	45.17	38.08	22.85	0.72
2	33	1.1	30	32.2	26.7	29.45	32.40	0.37
3	39	1.4	27.86	33.4	19.55	26.47	37.06	3.76
4	28	0.9	31.11	34.6	25.88	30.24	27.22	0.62
5	25	0.6	41.67	35.8	50.48	43.13	25.88	0.77
6	39	1.1	35.45	37	33.13	35.07	38.58	0.18
7	54	1.4	38.57	38.2	39.13	38.66	54.13	0.02
8	37	0.9	41.11	39.4	43.68	41.54	37.39	0.15
9	24	0.6	40	40.6	39.10	39.85	23.91	0.01
		Y_t/SI	x	y	Avg(x,y)	$F_t' \cdot SI$	6.60	

$$F_t' = \frac{F_t^{0=0} + F_t^{0=25}}{2}$$

$$F_t = F_t' \cdot SI_t$$

$$e^t = Y_t - F_t$$

$$RMSE = \sqrt{\frac{1}{9} \sum e^2} = \sqrt{\frac{6.60}{9}} = 0.86$$

$$80\% \rightarrow 1.28$$

$$F_i = F_i \pm 1.28 \cdot RMSE \cdot \sqrt{i-n}$$

$$F_0 = 44.89 \pm 1.28 \cdot 0.86 \cdot \sqrt{1} = 44.89 \pm 1.088 \rightarrow [43.802, 45.978] \quad (\text{σελ 116})$$

$$F_1 = 57.98 \pm 1.28 \cdot 0.86 \cdot \sqrt{2} = 57.98 \pm 1.539 \rightarrow [56.441, 59.519]$$

$$F_2 = 37.81 \pm 1.28 \cdot 0.86 \cdot \sqrt{3} = 37.81 \pm 1.884 \rightarrow [35.926, 39.694]$$

4.3 (1.0) Θεωρώντας ότι οι πραγματικές τιμές της χρονοσειράς για το διάστημα πρόβλεψης του πρώτου ερωτήματος είναι: {47, 68, 37}, αξιολογήστε τις προβλέψεις σας ως προς την ακρίβεια και προκατάληψή τους, καθώς και τα διαστήματα εμπιστοσύνης ως προς τον βαθμό κάλυψης που προσέφεραν, χρησιμοποιώντας κατάλληλες μετρικές.

Y_t	F_t	e_t	e_t^2	Y_t	διάστημα	εκτός
47	44.89	2.11	4.452	47	[43.802, 45.978]	όχι
68	57.98	10.02	100.4	68	[56.441, 59.519]	όχι
37	37.81	-0.81	0.656	37	[35.926, 39.694]	ναι
		11.32	105.51			

μεροληψία: $ME = \frac{1}{n} \sum (Y_i - F_i) = \frac{1}{n} \sum e_i = 3.77 > 0$ αναισθηδοφία

ακρίβεια: $MSE = \frac{1}{n} \sum (Y_i - F_i)^2 = \frac{1}{n} \sum e_i^2 = 35.17 \rightarrow$ μη ακριβείς προβλέψεις

βαθμός κάλυψης: δυο τιμές είναι εκτός ~ νέο διάστημα εμπιστοσύνης δ

$$F_i + 1.28 \cdot RMSE \cdot \sqrt{i-n} \geq Y_i \Rightarrow 57.97 + 8.085 \sqrt{1-9} \geq 68 \Rightarrow \delta \geq 8.34$$

$\delta \geq 8.34$ είναι αδύνατο, διότι η μέγιστη τιμή είναι 2.58