



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
Μονάδα Προβλέψεων & Στρατηγικής
Forecasting & Strategy Unit

Τεχνικές Προβλέψεων

Διακοπτόμενη Ζήτηση

Διακοπτόμενη Ζήτηση

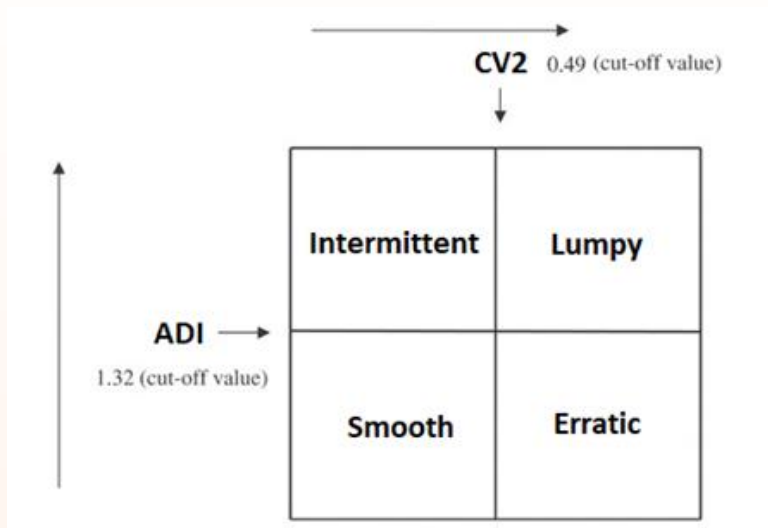
- Η διακοπτόμενη ζήτηση έχει την ιδιότητα να εμφανίζεται σποραδικά, δηλαδή να μεσολαβούν περίοδοι όπου αυτή είναι μηδενική.
- Όταν παρουσιαστεί ζήτηση, το μέγεθος αυτής ποικίλει και, σε πολλές περιπτώσεις, σε σημαντικό βαθμό.
- Συνήθως, διακοπτόμενη ζήτηση παρουσιάζεται σε μονάδες διαχείρισης αποθεμάτων (*stock keeping units*) και ανταλλακτικών (*spare parts*).
- Η διακοπτόμενη φύση των δεδομένων αυτών δημιουργεί σημαντικά προβλήματα τόσο στη διαδικασία της πρόβλεψης όσο και στον έλεγχο των αποθεμάτων.
- Η δυσκολία της πρόβλεψης δεν έγκειται μονάχα στην ασυνέχεια των ζητήσεων, αλλά και στη μεγάλη διακύμανση μεταξύ δύο διαδοχικών μη μηδενικών παρατηρήσεων.

Κατηγοριοποίηση Χρονοσειρών

Αν και με τον όρο «διακοπτόμενη ζήτηση» συνήθως αναφερόμαστε σε οποιαδήποτε χρονοσειρά παρουσιάζει **σποραδική ζήτηση** (εναλλαγή μη μηδενικής ζήτησης με μηδενική ζήτηση), οι χρονοσειρές μπορούν να κατηγοριοποιηθούν περαιτέρω ως εξής:

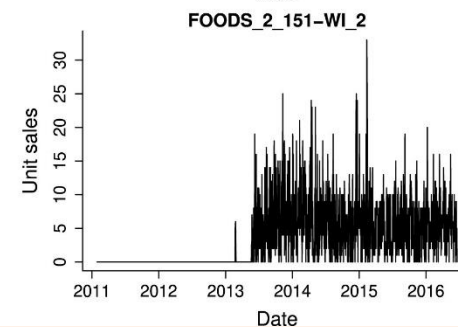
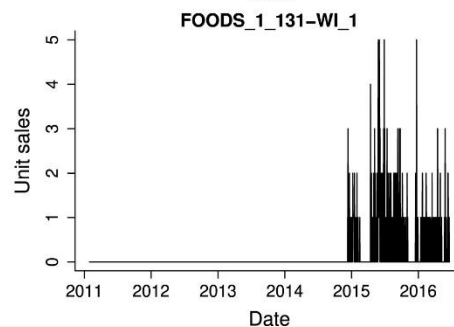
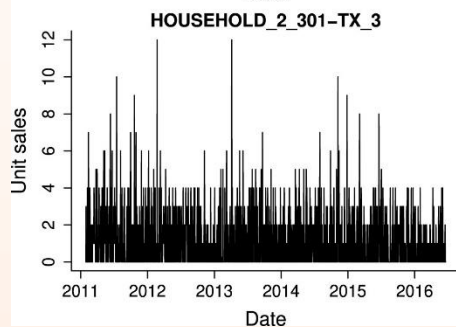
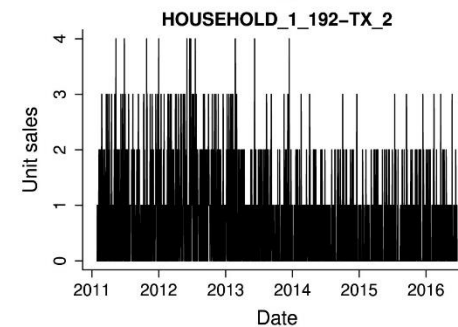
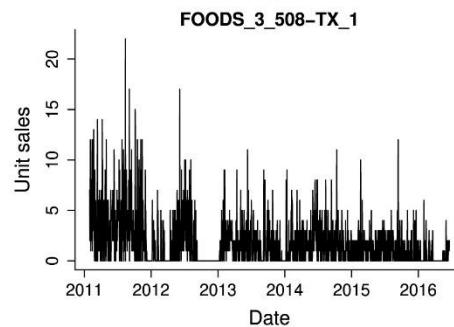
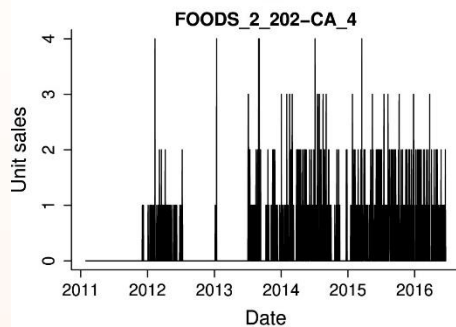
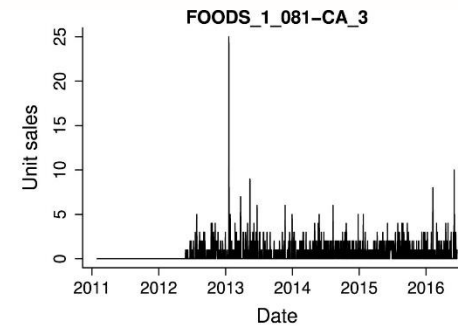
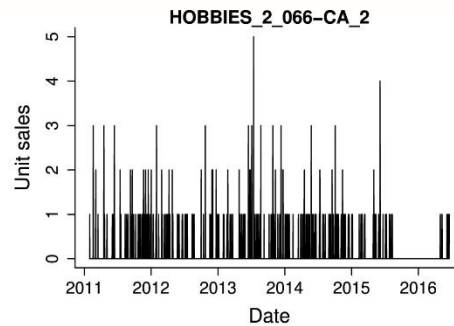
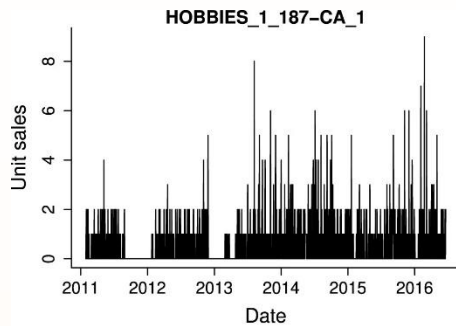
- **Smooth:** Ζήτηση χαμηλής διακύμανσης και σποραδικότητας
 - **Erratic:** Ζήτηση υψηλής διακύμανσης αλλά χαμηλής σποραδικότητας
 - **Intermittent:** Ζήτηση χαμηλής διακύμανσης αλλά υψηλής σποραδικότητας
 - **Lumpy:** Ζήτηση υψηλής διακύμανσης και σποραδικότητας
- Η «**σποραδικότητα**» της ζήτησης μετράτε από το δείκτη **ADI** (Average Demand Interval) ο οποίος παρουσιάζει πόσες περιόδοι μεσολαμβάν κατά μέσο όρο μέχρι να εμφανιστεί μη μηδενική ζήτηση
 - Η «**διακύμανση**» της ζήτησης μετράτε από το δείκτη **CV2** (squared Coefficient of Variation) ο οποίος παρουσιάζει τη διακύμανση της ζήτησης, όταν αυτή εμφανίζεται

Κατηγοριοποίηση Χρονοσειρών

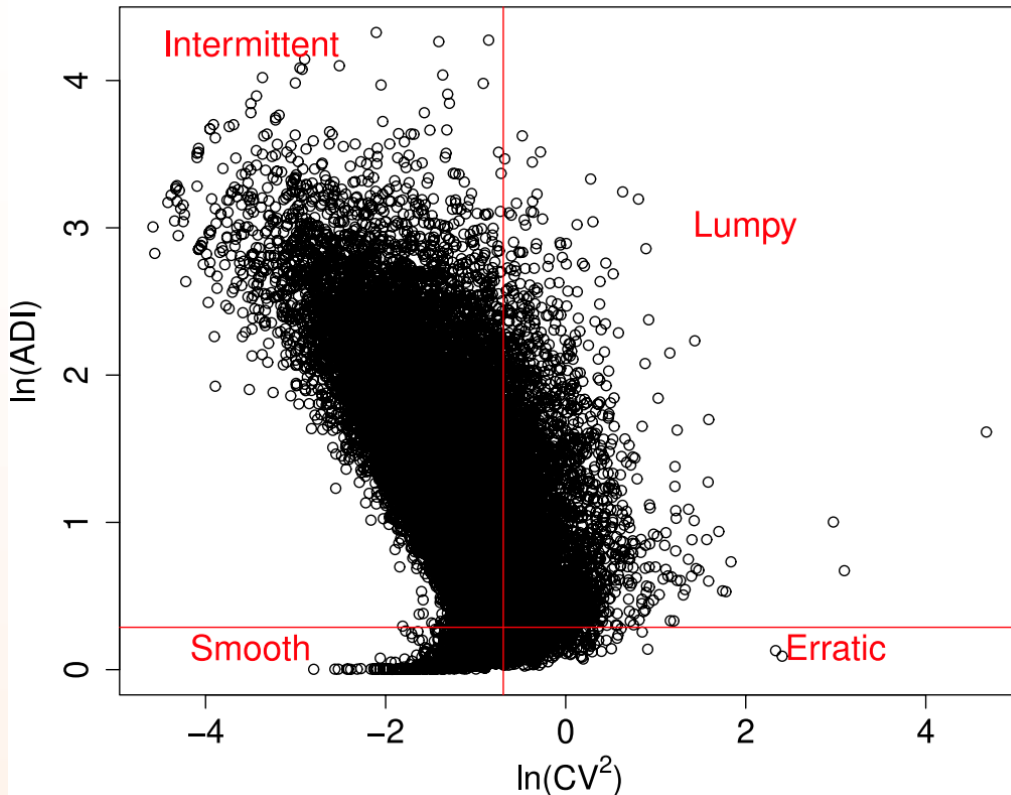


- Οι lumpy χρονοσειρές είναι δύσκολο να προβλεφθούν καθώς εμφανίζουν αβεβαιότητα τόσο ως προς το **πότε** θα εμφανιστεί η ζήτηση, όσο και ως προς το **πόση** θα είναι αυτή
- Αντίθετα, οι smooth χρονοσειρές είναι αρκετά πιο εύκολες να προβλεφθούν

Κατηγοριοποίηση Χρονοσειρών



Κατηγοριοποίηση Χρονοσειρών



Στο σύνολο 30,490 χρονοσειρών αναγνωρίζονται

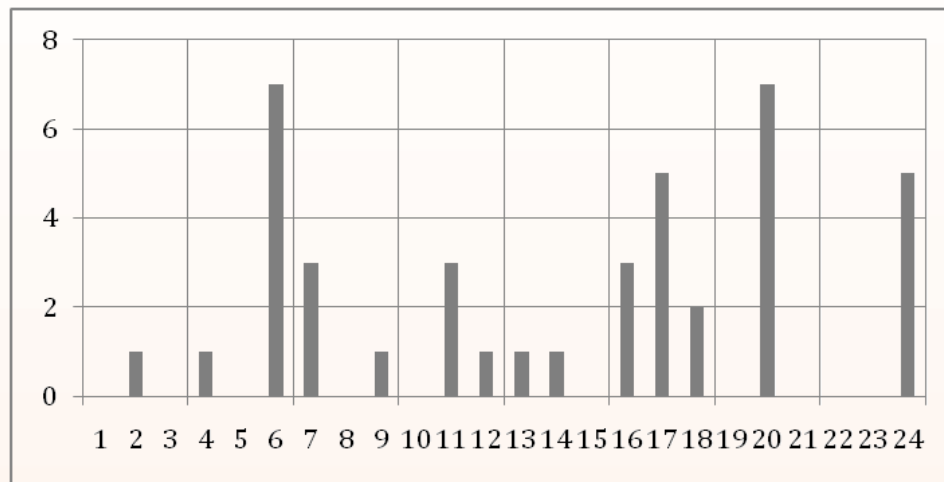
- Smooth: 7%
- Intermittent: 73%
- Erratic: 3%
- Lumpy: 17%

Μέθοδοι Πρόβλεψης Διακοπτόμενης Ζήτησης

- Οι μέθοδοι της εκθετικής εξομάλυνσης χρησιμοποιούνταν συχνά όταν πρέπει να γίνει χειρισμός και πρόβλεψη δεδομένων διακοπτόμενης φύσης. Καθώς, όμως, οι μέθοδοι αυτές αποδίδουν μεγαλύτερο βάρος στα πιο πρόσφατα δεδομένα, καταλήγουν σε ένα μοντέλο πρόβλεψης όπου οι εκτιμήσεις είναι μέγιστες έπειτα από μια εμφάνιση ζήτησης και ελάχιστες ακριβώς πριν από περίοδο μη μηδενικής ζήτησης.
- Αυτή η αδυναμία των μοντέλων εξομάλυνσης οδήγησε τους ερευνητές στην αναζήτηση νέων μεθόδων και τεχνικών για την αποτελεσματικότερη πρόβλεψη χρονοσειρών με έντονη παρουσία μηδενικών τιμών.
 - ✓ *Croston Method*
 - ✓ *Syntetos & Boylan Approximation*
 - ✓ *Aggregate-Disaggregate Intermittent Demand Approach*

Διακοπτόμενη Ζήτηση

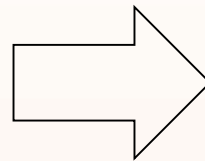
t	Y_t	t	Y_t
1	0	13	1
2	1	14	1
3	0	15	0
4	1	16	3
5	0	17	5
6	7	18	2
7	3	19	0
8	0	20	7
9	1	21	0
10	0	22	0
11	3	23	0
12	1	24	5



Croston Decomposition

t	Y_t
1	0
2	1
3	0
4	1
5	0
6	7
7	3
8	0
9	1
10	0
11	3
12	1

t	Y_t
13	1
14	1
15	0
16	3
17	5
18	2
19	0
20	7
21	0
22	0
23	0
24	5



Demands	Intervals
1	2
1	2
7	2
3	1
1	2
3	2
1	1
1	1
1	1
3	2
5	1
2	1
7	2
5	4

Μέθοδος Croston

Demands	Intervals
1	2
1	2
7	2
3	1
1	2
3	2
1	1
1	1
1	1
3	2
5	1
2	1
7	2
5	4

Average Demand Interval

$$\bar{Y}_{\text{demands}} = \frac{1 + 1 + \dots + 5}{14} = 2,93$$

$$\bar{Y}_{\text{intervals}} = \frac{2 + 2 + \dots + 4}{14} = 1,71$$

*Συντελεστής μεταβλητότητας
μη μηδενικών ζητήσεων*

$$c_v = \frac{\sigma_{\text{demands}}}{\bar{Y}_{\text{demands}}} \cdot 100 \Rightarrow$$

$$c_v = \frac{\sqrt{\frac{(1 - 2,93)^2 + (1 - 2,93)^2 + \dots + (5 - 2,93)^2}{14}}}{2,93} \cdot 100 = 73,54\%$$

Μέθοδος Croston

- Το 1972 ο Croston πρότεινε μια εναλλακτική μέθοδο, η οποία λαμβάνει υπόψη τόσο το μέγεθος της ζήτησης όσο και το χρόνο μεταξύ των ζητήσεων.
- Στην πράξη, ο Croston εξήγαγε προβλέψεις εφαρμόζοντας ανεξάρτητα απλή εκθετική εξομάλυνση τόσο στις μη μηδενικές τιμές των χρονοσειρών όσο και στα χρονικά διαστήματα μεταξύ των μη μηδενικών τιμών των χρονοσειρών.
- Πιο συγκεκριμένα, το σκεπτικό της μεθόδου *Croston*, είναι να διαχωρίζεται η εκάστοτε χρονοσειρά σε **δύο επιμέρους**, όπου η μία αποτελείται από τα χρονικά διαστήματα μεταξύ των μη μηδενικών ζητήσεων (*intervals*) και η άλλη από το πλήθος των ανεξάρτητων μη μηδενικών ζητήσεων (*demands*).
- Οι δύο χρονοσειρές προεκτείνονται **ανεξάρτητα**, με χρήση της μεθόδου εκθετικής εξομάλυνσης σταθερού επιπέδου. Στη βιβλιογραφία ως παράμετρος εξομάλυνσης στην εφαρμογή της μεθόδου είναι συνήθης η χρήση της τιμής $a=0,05$.
- Η προέκταση των χρονοσειρών ακολουθείται από υπολογισμό της πρόβλεψης *Croston* μέσω της εύρεσης του πηλίκου των δύο ανεξάρτητων προβλέψεων, ως εξής:

$$F_{\text{Croston}} = \frac{F_{\text{demands}}}{F_{\text{intervals}}}$$

Μέθοδος Croston

- Γενικά, η διακοπτόμενη ζήτηση μπορεί να οδηγήσει σε αύξηση του επιπέδου προμήθειας αποθεμάτων επιφέροντας προκατάληψη στις εκτιμήσεις της μέσης ζήτησης.
- Ο Croston, όμως, υποστηρίζει πως η προσέγγιση των ξεχωριστών εκτιμήσεων για ζήτηση και χρόνο μεταξύ ζητήσεων, οδηγεί σε μείωση της προκατάληψης. Επίσης, η πρόσθετη συνιστώσα της συχνότητας μεταξύ των ζητήσεων επιτρέπει στο διαχειριστή των αποθεμάτων να καθορίσει τις παραγγελίες και το κόστος με μεγαλύτερη ακρίβεια, αποφεύγοντας την περίσσεια αποθεμάτων.
- Η μέθοδος έχει το πλεονέκτημα ότι μπορεί να χρησιμοποιηθεί με τον ίδιο ακριβώς τρόπο και για δεδομένα συνεχούς ζήτησης, αφού πλέον συμπίπτει με απλή εφαρμογή της εκθετικής εξομάλυνσης σταθερού επιπέδου.
- Στην περίπτωση, όμως, δεδομένων διακοπτόμενης ζήτησης ουσιαστικά αποτελεί μια παραλλαγή της απλής εκθετικής εξομάλυνσης, με τις προβλέψεις να παραμένουν σταθερές και αμετάβλητες μεταξύ μη μηδενικών ζητήσεων και να ανανεώνονται μόνο μετά την παρατήρηση νέας μη μηδενικής τιμής.

Μέθοδος Croston

- Έχουν προταθεί κατά καιρούς διάφορες παραλλαγές αυτής, όσον αφορά τη μέθοδο που θα χρησιμοποιηθεί για πρόβλεψη. Έχει προταθεί, για παράδειγμα, η χρήση της μεθόδου *Theta* έναντι της εκθετικής εξομάλυνσης σταθερού επιπέδου (Nikoloroulos et al., 2007). Επίσης, έχει ελεγχθεί σε μεγάλο πλήθος χρονοσειρών η χρήση βελτιστοποιημένης παραμέτρου εκθετικής εξομάλυνσης έναντι σταθερής, ίσης με 0,05, δίνοντας ικανοποιητικά αποτελέσματα (Petropoulos et al., 2008).
- Σε κάθε περίπτωση, η διαδικασία που πρότεινε ο Croston θα μπορούσε να αναχθεί σε ένα γενικό μεθοδολογικό πλαίσιο, αποτελούμενο από τα ακόλουθα βήματα:
 1. Αποσύνθεση της αρχικής χρονοσειράς διακοπτόμενης ζήτησης σε δύο σειρές, από τις οποίες η μία περιλαμβάνει τις μη μηδενικές ζητήσεις και η άλλη τα μεσοδιαστήματα μεταξύ αυτών.
 2. Προέκταση καθεμίας από τις χρονοσειρές που προέκυψαν από την αποσύνθεση με κατάλληλη μέθοδο πρόβλεψης.
 3. Συνδυασμός των παραχθέντων προβλέψεων μέσω διαίρεσης της σειράς πρόβλεψης ζήτησης και της σειράς πρόβλεψης μεσοδιαστημάτων, ώστε να υπολογισθεί η τελική σειρά προβλέψεων της μεθόδου.

Μέθοδος Croston

Παράδειγμα

$$a=0,05$$

$$S_0^{\text{demands}} = \bar{Y}_{\text{demands}} = \frac{1 + 1 + \dots + 5}{14} = 2,93$$

Demands	F_t	e_t	S_t
			2,93
1	2,93	-1,93	2,83
1	2,83	-1,83	2,74
7	2,74	4,26	2,95
3	2,95	0,05	2,96
1	2,96	-1,96	2,86
3	2,86	0,14	2,87
1	2,87	-1,87	2,77
1	2,77	-1,77	2,68
1	2,68	-1,68	2,60
3	2,60	0,40	2,62
5	2,62	2,38	2,74
2	2,74	-0,74	2,70
7	2,70	4,30	2,92
5	2,92	2,08	3,02
			3,02

$$S_0^{\text{intervals}} = \bar{Y}_{\text{intervals}} = \frac{2 + 2 + \dots + 4}{14} = 1,71$$

Intervals	F_t	e_t	S_t
			1,71
2	1,71	0,29	1,72
2	1,72	0,28	1,74
2	1,74	0,26	1,75
1	1,75	-0,75	1,71
2	1,71	0,29	1,73
2	1,73	0,27	1,74
1	1,74	-0,74	1,70
1	1,70	-0,70	1,67
1	1,67	-0,67	1,64
2	1,64	0,36	1,65
1	1,65	-0,65	1,62
1	1,62	-0,62	1,59
2	1,59	0,41	1,61
4	1,61	2,39	1,73
			1,73

Μέθοδος Croston

Παράδειγμα

Demands Forecast	Intervals Forecast	Croston's Forecast
2,93	1,71	1,71
2,83	1,72	1,64
2,74	1,74	1,58
2,95	1,75	1,69
2,96	1,71	1,73
2,86	1,73	1,65
2,87	1,74	1,65
2,77	1,70	1,63
2,68	1,67	1,61
2,60	1,64	1,59
2,62	1,65	1,58
2,74	1,62	1,69
2,70	1,59	1,70
2,92	1,61	1,81
3,02	1,73	1,75

t	Y_t	F_{Croston}
1	0	1,71
2	1	1,71
3	0	1,64
4	1	1,64
5	0	1,58
6	7	1,58
7	3	1,69
8	0	1,73
9	1	1,73
10	0	1,65
11	3	1,65
12	1	1,65

t	Y_t	F_{Croston}
13	1	1,63
14	1	1,61
15	0	1,59
16	3	1,59
17	5	1,58
18	2	1,69
19	0	1,70
20	7	1,70
21	0	1,81
22	0	1,81
23	0	1,81
24	5	1,81
25		1,75

Μέθοδος SBA

- Η μέθοδος *Croston* έχει αποδειχθεί μεγίστης αξίας για τους κατασκευαστές που αντιμετωπίζουν δεδομένα διακοπτόμενης φύσης.
- Παρά όμως τη σπουδαία θεωρητική ανωτερότητα αυτής της μεθόδου πρόβλεψης, οι εμπειρικές ενδείξεις έχουν δείξει πως τα κέρδη από την εφαρμογή της είναι χειρότερα του αναμενόμενου, συγκρινόμενη με απλούστερες τεχνικές πρόβλεψης. Σε μερικές περιπτώσεις, παρατηρείται ακόμα και χειρότερη επίδοση.
- Οι Συντετός και Boylan (2001), προσπαθώντας να εντοπίσουν το αίτιο αυτής της απροσδόκητης συμπεριφοράς, διαπίστωσαν ότι η μέθοδος *Croston* είναι θετικά προκατειλημμένη (*positively biased*), δηλαδή παρουσιάζει μια αισιόδοξη τάση στα αποτελέσματα των προβλέψεών της.
- Κατάφεραν, μάλιστα, να συνδέσουν το επίπεδο της αισιόδοξης τάσης της μεθόδου με την τιμή της παραμέτρου εξομάλυνσης a που χρησιμοποιείται για την προέκταση των δύο αποσυντεθμένων χρονοσειρών.
- Η μέγιστη προκατάληψη παρατηρείται όταν το a λάβει τη μέγιστη τιμή, δηλαδή $a=1$. Γενικά παρατηρήθηκε μια αναλογία, με μεγάλες τιμές του a να παρουσιάζουν μεγάλη προκατάληψη και η μέθοδος *Croston* να ενδείκνυται μόνο όταν το a είναι μικρότερο του 0,15.

Μέθοδος SBA

Η μέθοδος *SBA* (*Syntetos and Boylan Approximation*) αποτελεί μια τροποποίηση της μεθόδου *Croston*, στην οποία η πρόβλεψη υπολογίζεται από τον ακόλουθο μαθηματικό τύπο:

$$F_{SBA} = \left(1 - \frac{a}{2}\right) \cdot \frac{F_{\text{demands}}}{F_{\text{intervals}}}$$

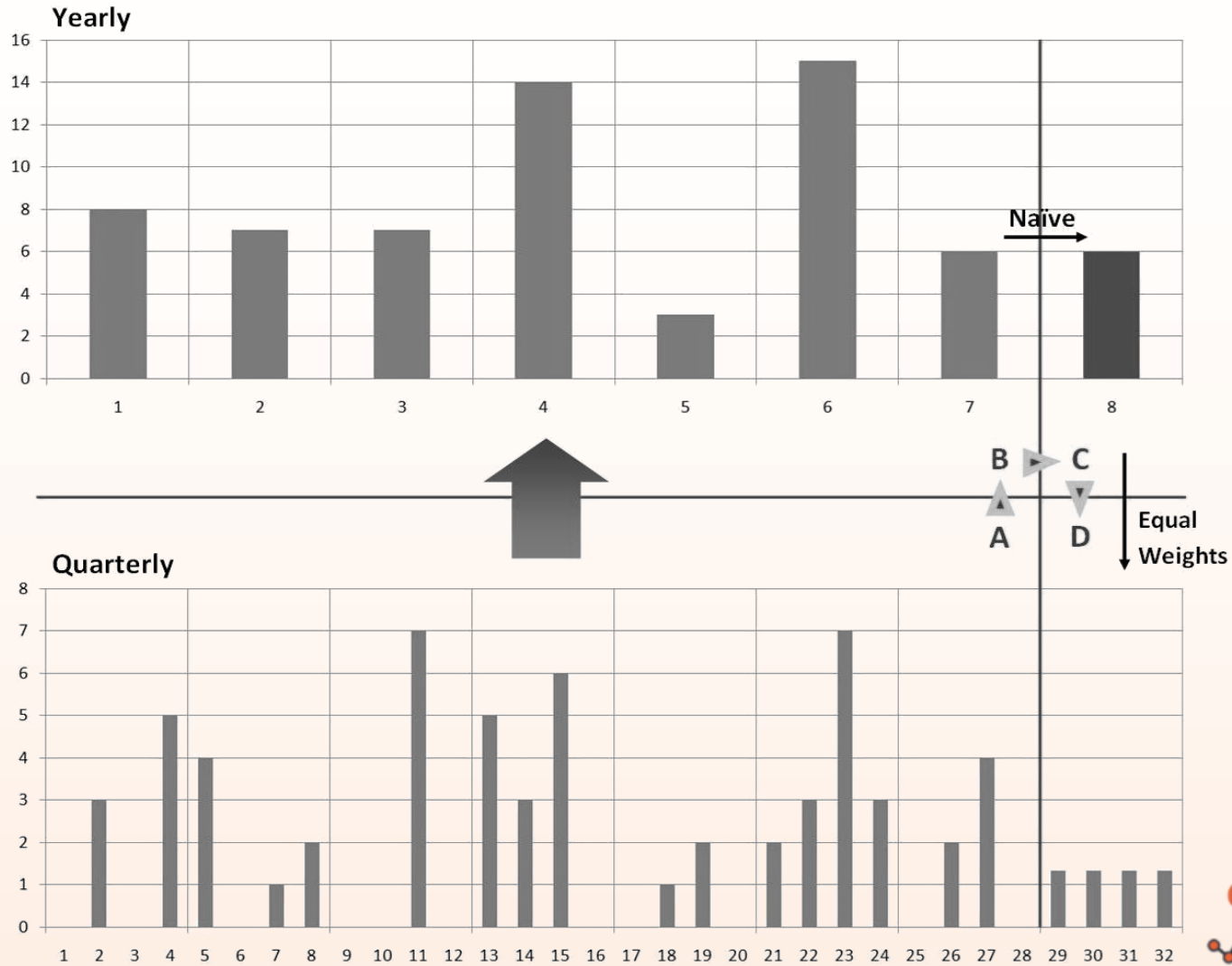
Μέθοδος ADIDA

- Μια διαδικασία που μπορεί να ακολουθηθεί προκειμένου να αποφεύγεται η ασυνέχεια των δεδομένων διακοπτόμενης ζήτησης, όσων αφορά τις μηδενικές τιμές, είναι η μη επικαλυπτόμενη συνάθροιση (*aggregation*) των δεδομένων σε περιόδους μικρότερης συχνότητας.
- Για παράδειγμα, στην περίπτωση των μηνιαίων δεδομένων, θα μπορούσε να εφαρμοσθεί συνάθροιση σε τριμηνιαίο επίπεδο, θέτοντας το επίπεδο συνάθροισης ίσο με τρεις περιόδους.
- Με το σκεπτικό αυτό ενδέχεται να μειωθεί, αν όχι απαλειφθεί, η ασυνέχεια λόγω ύπαρξης μηδενικών παρατηρήσεων. Επίσης, η διακύμανση της προκύπτουσας νέας χρονοσειράς αναμένεται να είναι μικρότερη, έχοντας χρησιμοποιήσει ουσιαστικά μη επικαλυπτόμενο κινητό μέσο όρο για εξομάλυνση.
- Ο κατάλληλος καθορισμός του επιπέδου συνάθροισης (*aggregation level*) θα οδηγήσει σε χρονοσειρά συνεχούς ζήτησης, χωρίς μηδενικές τιμές, στην οποία θα μπορεί πλέον να εφαρμοσθεί οιαδήποτε τεχνική πρόβλεψης δεδομένων κατάλληλη για συνεχή ζήτηση για την παραγωγή προβλέψεων στο επίπεδο συνάθροισης.
- Ακολούθως, εφαρμόζεται η διάσπαση (*disaggregation*) των υπολογισμένων προβλέψεων, ώστε να υπολογισθούν οι τελικές προβλέψεις χρονικής συχνότητας ίσης με του αρχικού επιπέδου.

Μέθοδος ADIDA

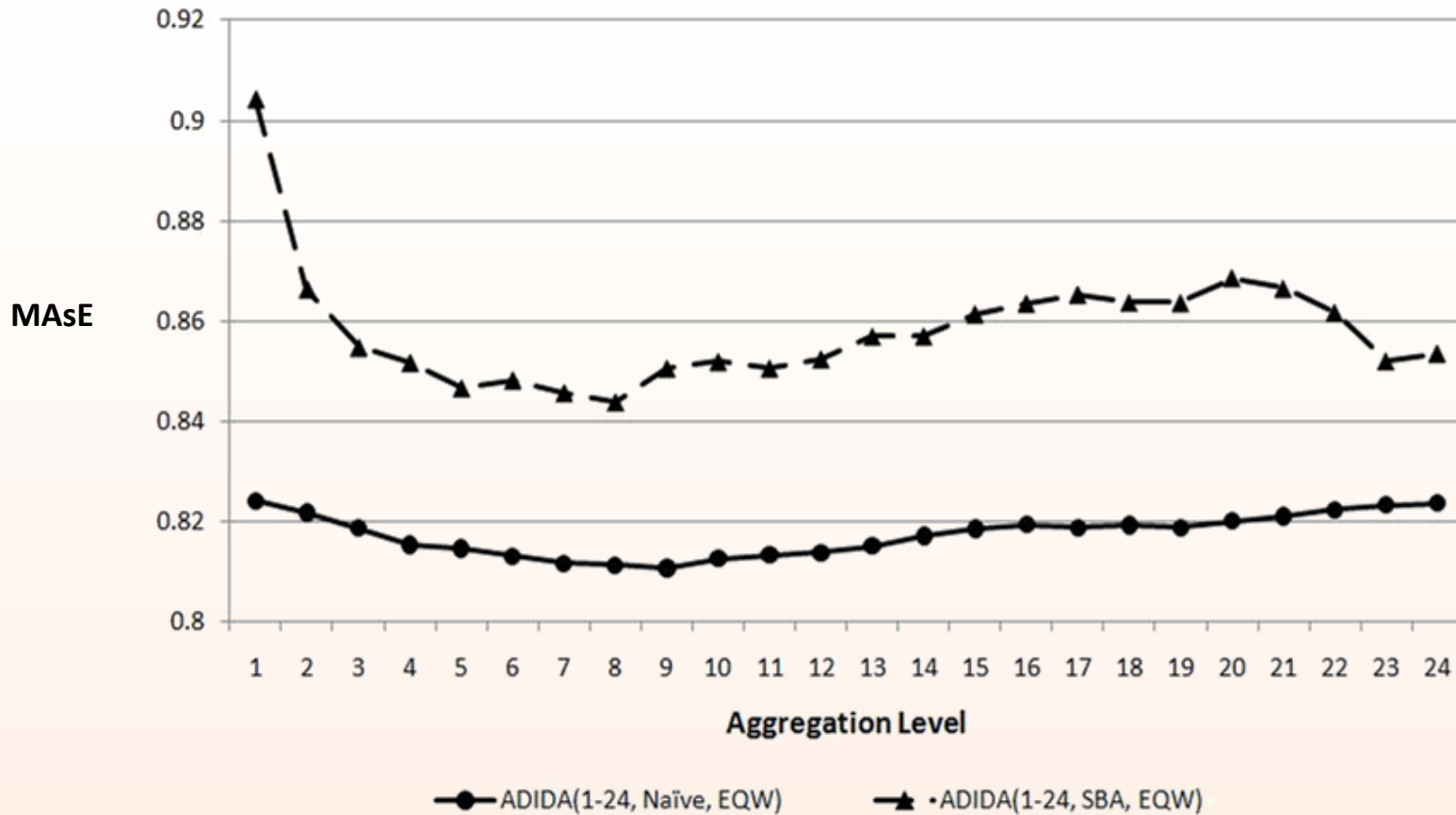
- Η παραπάνω φιλοσοφία οδηγεί σε μια νέα μέθοδο, η οποία ονομάζεται *ADIDA* (Nikolopoulos et al., 2010) και προτείνει τη συνάθροιση δεδομένων σε ένα υψηλότερο χρονικά επίπεδο, όπου έχει αποφευχθεί η ασυνέχεια των δεδομένων.
- Ανάλογα με το επίπεδο της συνάθροισης των δεδομένων, δημιουργείται μια καινούρια χρονοσειρά με σαφώς ελαττωμένη διακοπτόμενη συμπεριφορά.
- Σε αυτήν μπορούν πλέον να εφαρμοστούν μέθοδοι πρόβλεψης για δεδομένα συνεχούς ζήτησης, πέραν των μεθόδων που εφαρμόζονται ως επί το πλείστον σε δεδομένα διακοπτόμενης ζήτησης.
- Ύστερα από το στάδιο της παραγωγής των προβλέψεων, απαιτείται ο διαχωρισμός των προβλέψεων στα αρχικά επίπεδα των χρονοσειρών, χρησιμοποιώντας εμπειρικές τεχνικές.
- Το μεθοδολογικό πλαίσιο ολοκληρώνεται σε τρία βήματα:
 1. Συνάθροιση των δεδομένων με κατάλληλο επίπεδο συνάθροισης
 2. Πρόβλεψη της συναθροισμένης χρονοσειράς
 3. Διαχωρισμός των προβλέψεων που παρήχθησαν στο προηγούμενο βήμα

Μέθοδος ADIDA



Μέθοδος ADIDA

Επίπεδο Συνάθροισης



5000 SKUs RAF

Μέθοδος ADIDA

Μέθοδοι Διαχωρισμού

Equal Weights. Αναφέρεται στον απλό ισοβαρή διαχωρισμό. Ενδείκνυται για χρονοσειρές με μεγάλη τυχαιότητα και χωρίς εποχιακή συμπεριφορά.

Previous Weights. Αναφέρεται στην εφαρμογή των βαρών που έχουν οι m προηγούμενες παρατηρήσεις, όπου m ισούται με το επίπεδο συνάθροισης.

Average Weights. Αναφέρεται στον υπολογισμό των μέσων βαρών που υπολογίζονται αν χωρίσουμε τις παρατηρήσεις σε k ομάδες m παρατηρήσεων η καθεμία, όπου $k \times m$ ισούται με το σύνολο των διαθέσιμων παρατηρήσεων και m ισούται με το επίπεδο συνάθροισης. Ενδείκνυται σε περιπτώσεις όπου η συνιστώσα της εποχιακότητας είναι έντονη.

Μέθοδος ADIDA

Μέθοδοι Διαχωρισμού

- Συνοψίζοντας, η μη επικαλυπτόμενη συνάθροιση δεδομένων είναι μια πολλά υποσχόμενη προσέγγιση για χρονοσειρές διακοπτόμενης ζήτησης, καθώς οι προβλέψεις σε υψηλότερα επίπεδα συνάθροισης περιόδων είναι γενικά πιο ακριβείς και έχουν μικρότερες διακυμάνσεις σε σχέση με εκείνες μικρότερων επιπέδων συνάθροισης.
- Ως επίπεδο συνάθροισης μπορεί κάλλιστα να επιλεγεί ο ορίζοντας πρόβλεψης, ή ακόμα και να τον υπερβεί, όπου εν συνεχεία απαιτούνται αποσυνθετικοί μηχανισμοί.
- Στα πλεονεκτήματα της μεθόδου *ADIDA* πρέπει να προστεθεί η δυνατότητα που παρέχεται από τα περισσότερα πακέτα λογισμικού προβλέψεων όσον αφορά τη συνάθροιση των δεδομένων σε υψηλότερα χρονικά επίπεδα.
- Η μέθοδος *ADIDA* μπορεί να οδηγήσει σε σημαντική βελτίωση της ακρίβειας πρόβλεψης σε σχέση με μεμονωμένες μεθόδους, γεγονός που την καθιστά μηχανισμό «αυτοβελτίωσης» μιας μεθόδου.
- Τα εμπειρικά αποτελέσματα αποδεικνύουν ότι είναι πιθανό να υπάρχει βέλτιστο επίπεδο συνάθροισης, ενώ ο ορισμός του επιπέδου συνάθροισης ίσου με το *lead time* συν μία ακόμα περίοδο οδηγεί σε πολύ καλά αποτελέσματα, αναφορικά με την ακρίβεια πρόβλεψης.

Μέθοδος ADIDA

Παράδειγμα

Υ	Y _{aggregated}
0	
1	1
0	
1	
0	8
7	
3	
0	4
1	
0	
3	4
1	

Υ	Y _{aggregated}
1	
1	2
0	
3	
5	10
2	
0	
7	7
0	
0	
0	5
5	

Y _{aggregated}	F _{aggregated}
1	
8	
4	
4	4,33
2	5,33
10	3,33
7	5,33
5	6,33
	7,33

$$Y_1^{\text{aggregated}} = Y_1 + Y_2 + Y_3 = 0 + 1 + 0 = 1$$

$$Y_2^{\text{aggregated}} = Y_4 + Y_5 + Y_6 = 1 + 0 + 7 = 8$$

$$F_4^{\text{aggregated}} = \frac{Y_1^{\text{aggregated}} + Y_2^{\text{aggregated}} + Y_3^{\text{aggregated}}}{3} \Rightarrow$$

$$F_4^{\text{aggregated}} = \frac{1 + 8 + 4}{3} = 4,33$$

Μέθοδος ADIDA

Παράδειγμα

Y_t	$F_{\text{aggregated}}$	F_t
0		
1		
0		
1		
0		
7		
3		
0		
1		
0	4,33	1,44
3		1,44
1		1,44

$$F_{10} = F_{11} = F_{12} = \frac{F_4^{\text{aggregated}}}{3} = \frac{4,33}{3} = 1,44$$

$$F_{25} = F_{26} = F_{27} = \frac{F_9^{\text{aggregated}}}{3} = \frac{7,33}{3} = 2,44$$

Y_t	$F_{\text{aggregated}}$	F_t
1	5,33	1,78
1		1,78
0		1,78
3		1,11
5	3,33	1,11
2		1,11
0	5,33	1,78
7		1,78
0		1,78
0		2,11
0	6,33	2,11
5		2,11
???	7,33	2,44
???		2,44
???		2,44

Μέθοδος ADIDA

Παράδειγμα

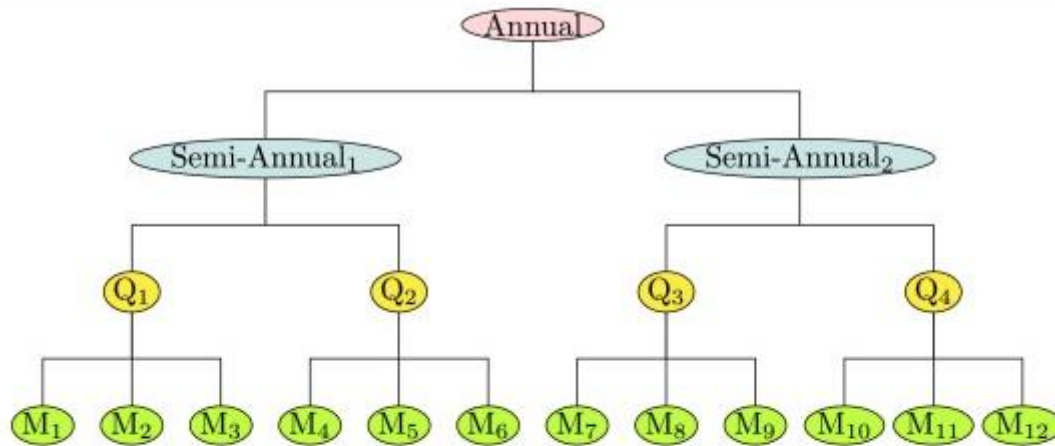
$$\text{MASe} = \frac{\frac{1}{n} \sum_{i=1}^n |Y_i - F_i|}{\frac{1}{n-1} \sum_{i=2}^n |Y_i - Y_{i-1}|} \Rightarrow$$

$$\text{MASe} = \frac{\frac{1,44 + 1,56 + \dots + 2,89}{15}}{\frac{1 + 1 + \dots + 5}{23}} \Rightarrow$$

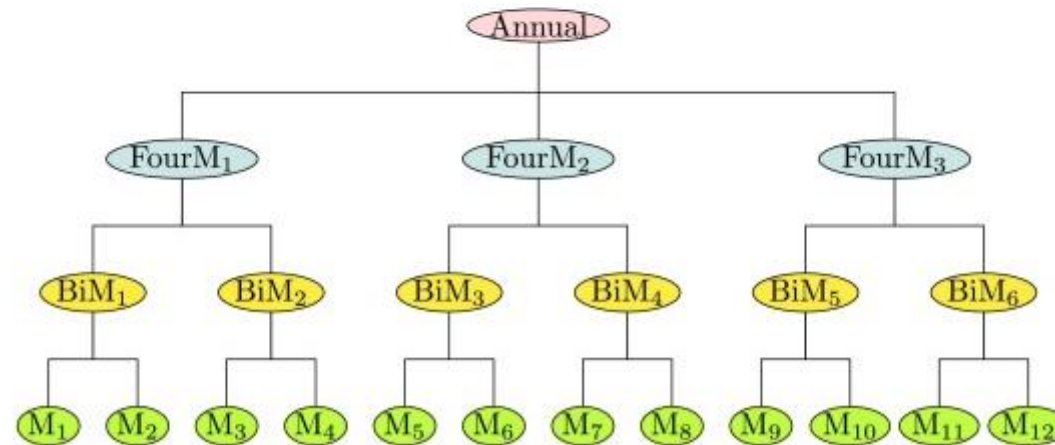
$$\text{MASe} = \frac{1,96}{2,39} = 0,82$$

Y_t	F_t	$ e_t $	Naïve	$ e_t $
0				
1			0	1
0			1	1
1			0	1
0			1	1
7			0	7
3			7	4
0			3	3
1			0	1
0	1,44	1,44	1	1
3	1,44	1,56	0	3
1	1,44	0,44	3	2
1	1,78	0,78	1	0
1	1,78	0,78	1	0
0	1,78	1,78	1	1
3	1,11	1,89	0	3
5	1,11	3,89	3	2
2	1,11	0,89	5	3
0	1,78	1,78	2	2
7	1,78	5,22	0	7
0	1,78	1,78	7	7
0	2,11	2,11	0	0
0	2,11	2,11	0	0
5	2,11	2,89	0	5
???	2,44		5	
???	2,44			
???	2,44			

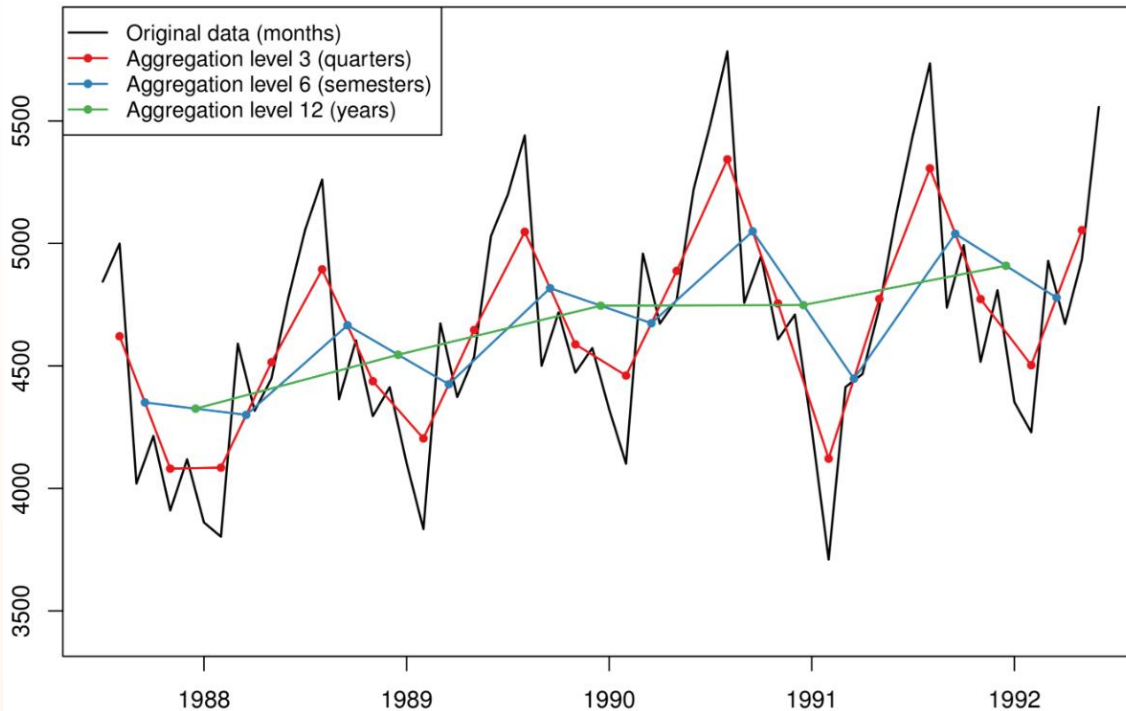
Συνάθροιση Δεδομένων



- Η διαδικασία αυτή μπορεί να γενικευτεί και για δεδομένα συνεχούς ζήτησης
- Σε αυτήν την περίπτωση, και βάσει της συχνότητας των δεδομένων, η συνάθροιση μπορεί να πραγματοποιηθεί σε διαστήματα που δεν είναι τυχαία επιλεγμένα, αλλά ταυτίζονται με πραγματικές ημερολογιακές περιόδους (π.χ. έτος, τρίμηνο, δίμηνο, μήνας)



Συνάθροιση Δεδομένων



- Συναθροίσεις μεγαλύτερων περιόδων αναδεικνύουν τα **μακροπρόθεσμα** χαρακτηριστικά των χρονοσειρών, όπως η **τάση**
- Συναθροίσεις μικρότερων περιόδων αναδεικνύουν τα **βραχυπρόθεσμα** χαρακτηριστικά των χρονοσειρών, όπως η **εποχιακότητα** και το **επίπεδο**
- **Ενδεικτικές μέθοδοι:** MAPA (Multiple Aggregation Prediction Algorithm) & Thief (Temporal Hierarchies)

1. **MAPA:** Kourentzes, N., Petropoulos, F., Trapero, J.R. (2014). Improving forecasting by estimating time series structural components across multiple frequencies, *International Journal of Forecasting*, 30(2), 291-302
2. **Thief:** Athanasopoulos, G., Hyndman, R.J., Kourentzes, N., Petropoulos, F.(2017). Forecasting with temporal hierarchies, *European Journal of Operational Research*, 262(1), 60-74

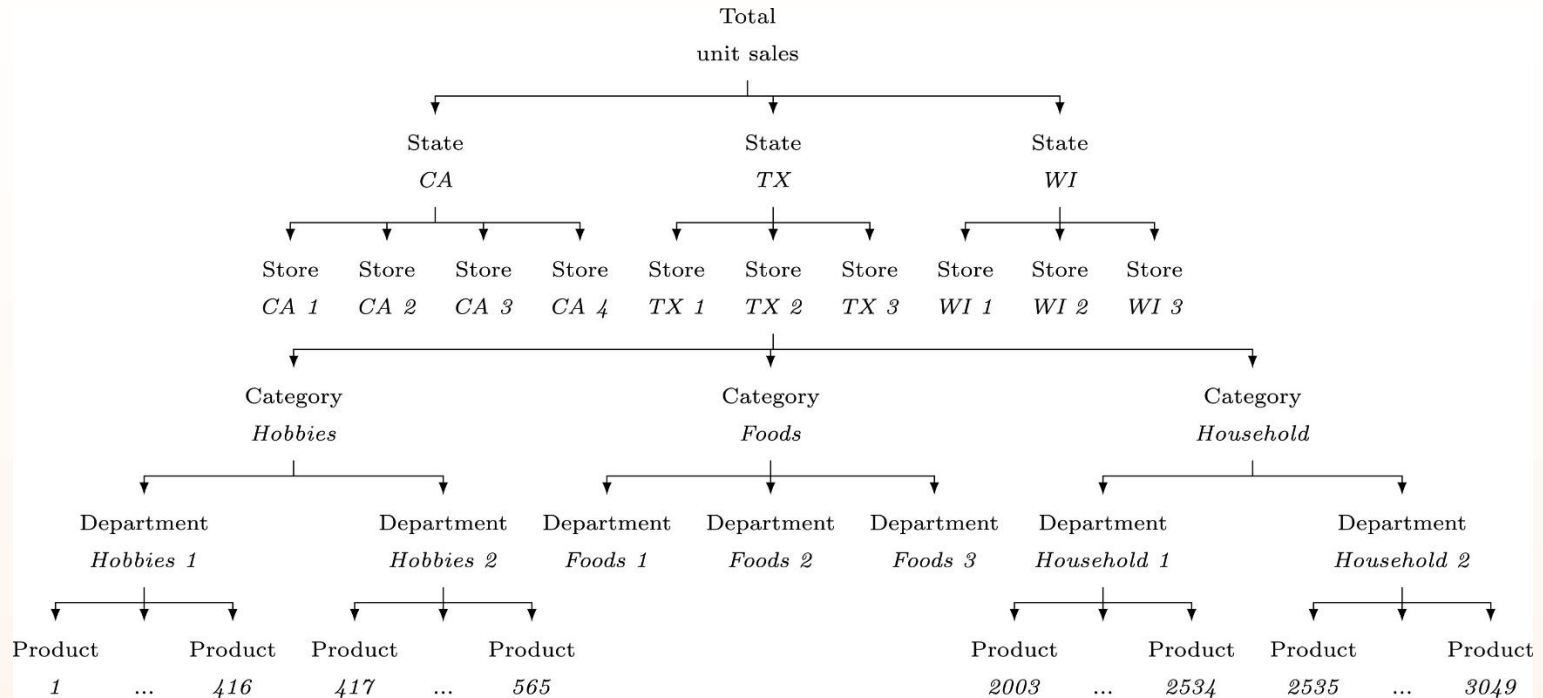
Διαγωνισμός M5

Έτος διεξαγωγής 2020 | 42,840 χρονοσειρές | Περισσότερες από 6000 συμμετοχές

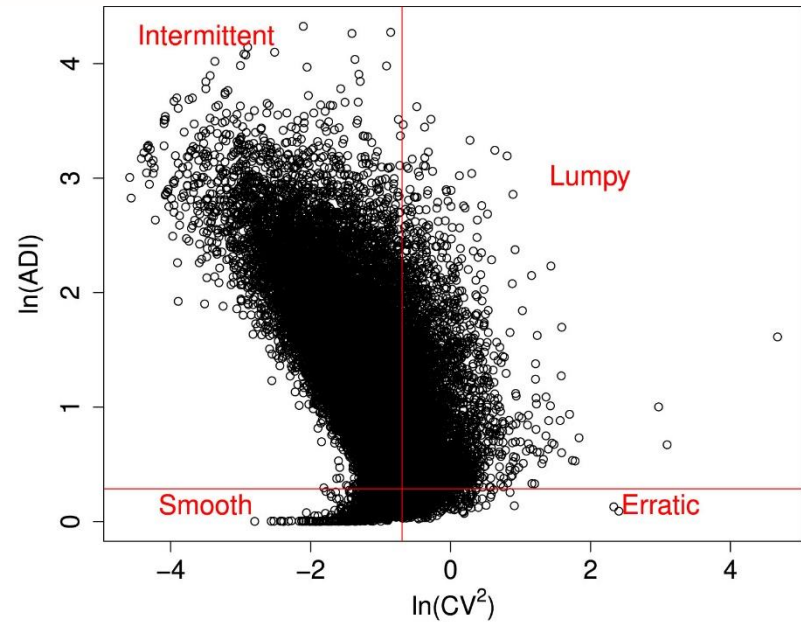
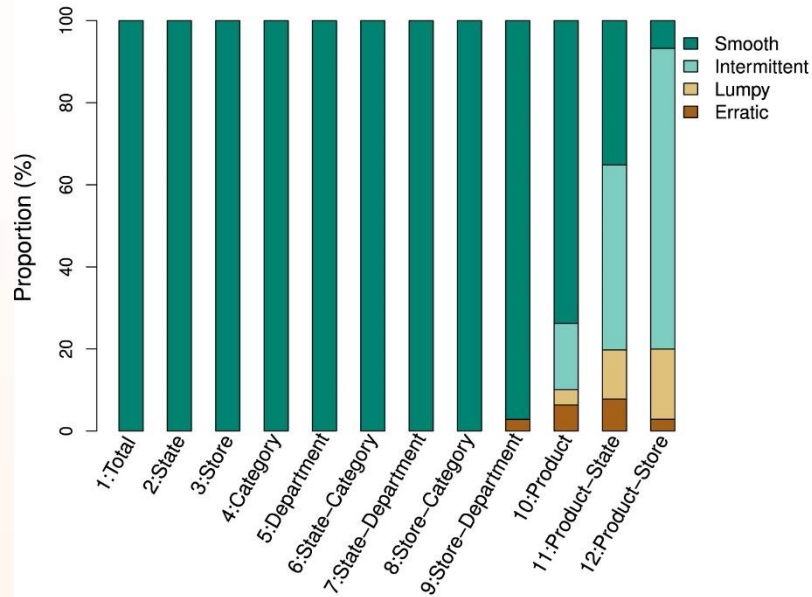
- Ιεραρχικά δεδομένα πωλήσεων που αφορούν επίπεδα προϊόντων, καταστημάτων, κατηγοριών και υποκατηγοριών προϊόντων, γεωγραφικές περιοχές κ.α.
- Εξωγενείς μεταβλητές όπως τιμή προϊόντος, προωθητικές ενέργειες, ημερομηνίες, γιορτές και αργίες.
- Αξιολόγηση τόσο σημειακών προβλέψεων όσο και πιθανοτικών προβλέψεων (50%, 67%, 95%, 99% και median)
- Έμφαση σε δεδομένα διακοπτόμενης ζήτησης
- Τα δεδομένα προέρχονται από τη Walmart και αφορούν 3,049 κωδικούς προϊόντων
- Ξεχωριστοί διαγωνισμοί για σημειακές και πιθανοτικές προβλέψεις

Makridakis, S., Spiliotis, E., Assimakopoulos, V. (2022). The M5 competition: Background, organization, and implementation, International Journal of Forecasting, 38(4), 1325-1336

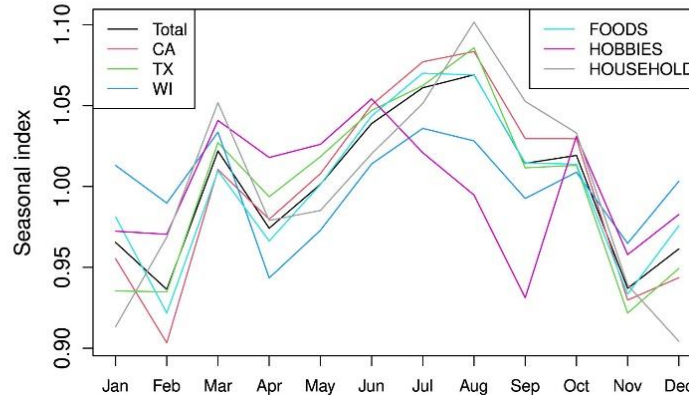
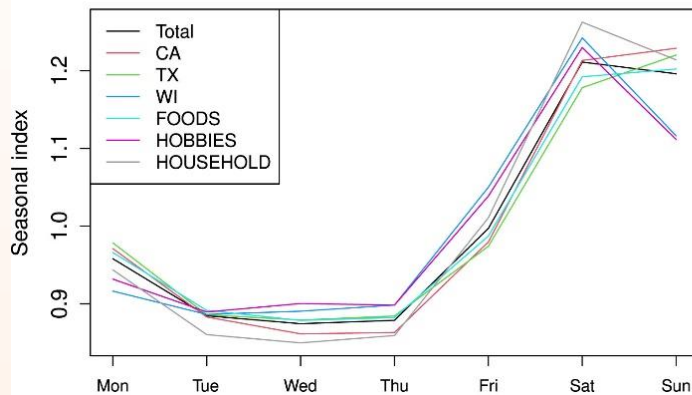
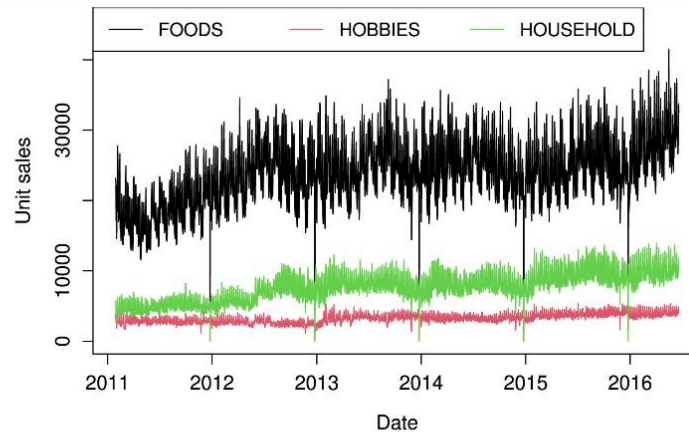
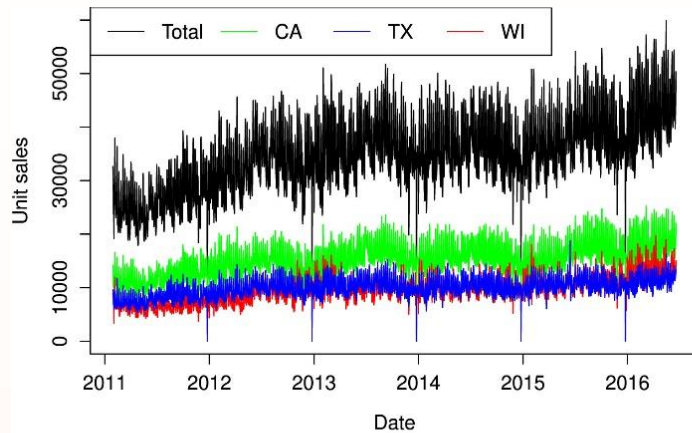
Δεδομένα διαγωνισμού M5



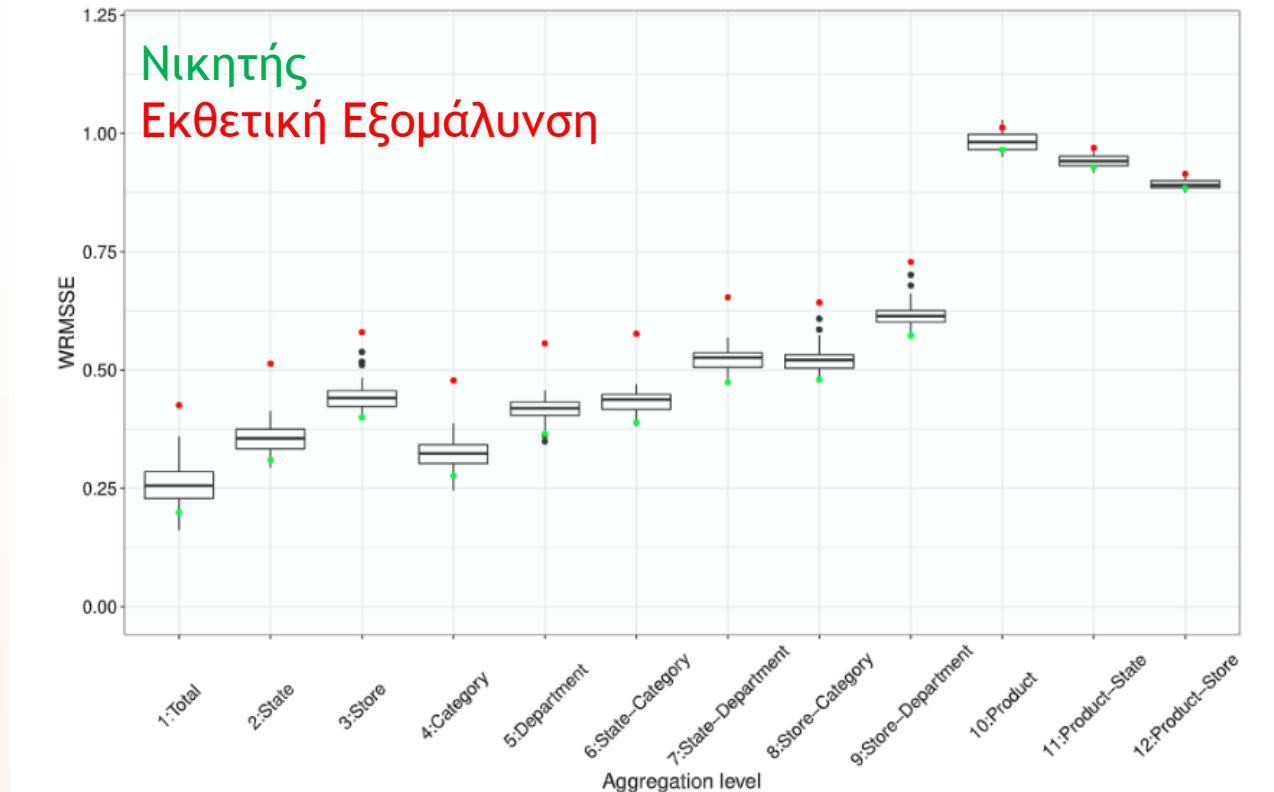
Δεδομένα διαγωνισμού M5



Δεδομένα διαγωνισμού M5



Ευρήματα Διαγωνισμού M5



- Χρονοσειρές διακοπτόμενης ζήτησης μπορούν να εμφανιστούν σε διάφορα ιεραρχικά επίπεδα μιας επιχείρησης. Συνήθως όμως αυτό συμβαίνει σε επίπεδο SKU (κατάστημα-προϊόν ή αποθήκη-προϊόν)
- Σε αυτά τα επίπεδα, η πρόβλεψη γίνεται ιδιαίτερα απαιτητική

Ευρήματα Διαγωνισμού M5

- Οι μέθοδοι μηχανικής μάθησης, όπως τα δέντρα παλινδρόμησης (regression trees) μπορούν να βοηθήσουν σημαντικά στη μείωση του σφάλματος πρόβλεψης
- Μεγάλο πλεονέκτημα των μοντέλων αυτών είναι ότι μπορούν να μάθουν από πολλές χρονοσειρές ταυτόχρονα (**cross-learning**) αντί για κάθε μία χρονοσειρά ξεχωριστά. Ως εκ τούτου, δύναται να ανακαλύψουν και αξιοποιήσουν μοτίβα που παρατηρούνται σε συγκεκριμένες κατηγορίες ή υποκατηγορίες κωδικών, ημέρες και μήνες του έτους, ειδικά γεγονότα κ.ο.κ.
- Σε περιπτώσεις όπου υπάρχει χρήσιμη **πληροφορία** σχετικά το μέλλον (αλλαγή τιμών, αργίες, ειδικά γεγονότα) είναι κρίσιμο αυτή να λαμβάνεται υπόψη μέσω κατάλληλων μεθόδων (π.χ. τεχνικές SEA ή παλινδρόμησης)

Εφαρμογή στην R

```
library(forecast)  
library(Rlab)
```

```
#Δημιουργία χρονοσειράς διακοπόμενης ζήτησης
```

```
n <- 25
```

```
set.seed(102)
```

```
intervals <- rbern(n, 0.4)
```

```
demand <- runif(n,5,20)
```

```
series <- demand*intervals
```

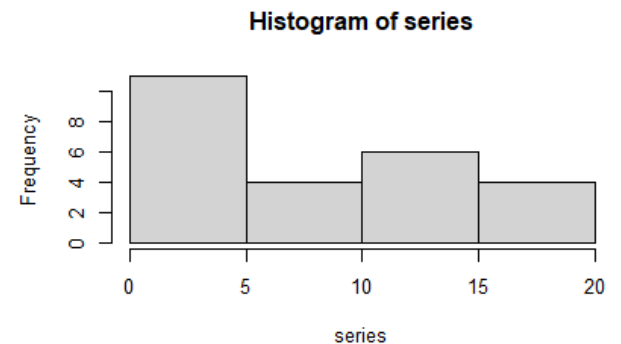
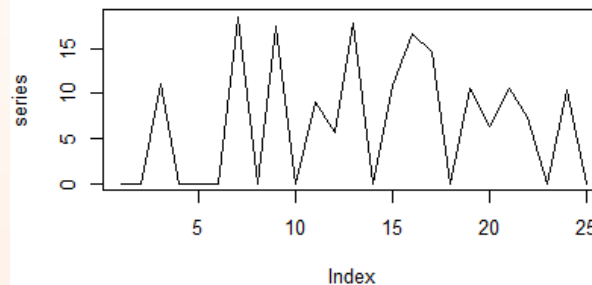
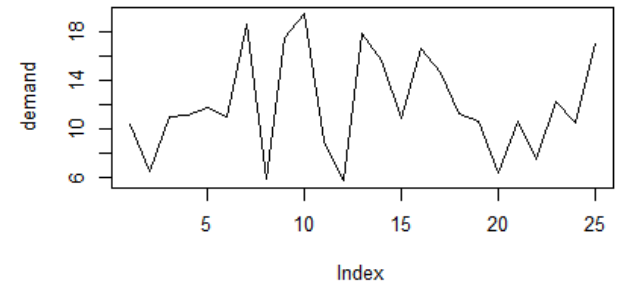
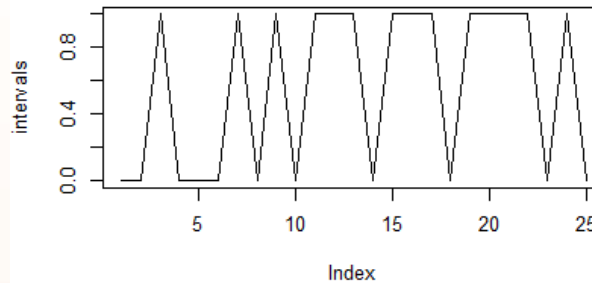
```
par(mfrow=c(2,2))
```

```
plot(intervals, type="l")
```

```
plot(demand, type="l")
```

```
plot(series, type="l")
```

```
hist(series)
```



Εφαρμογή στην R

#Εφαρμογή μεθόδου Croston

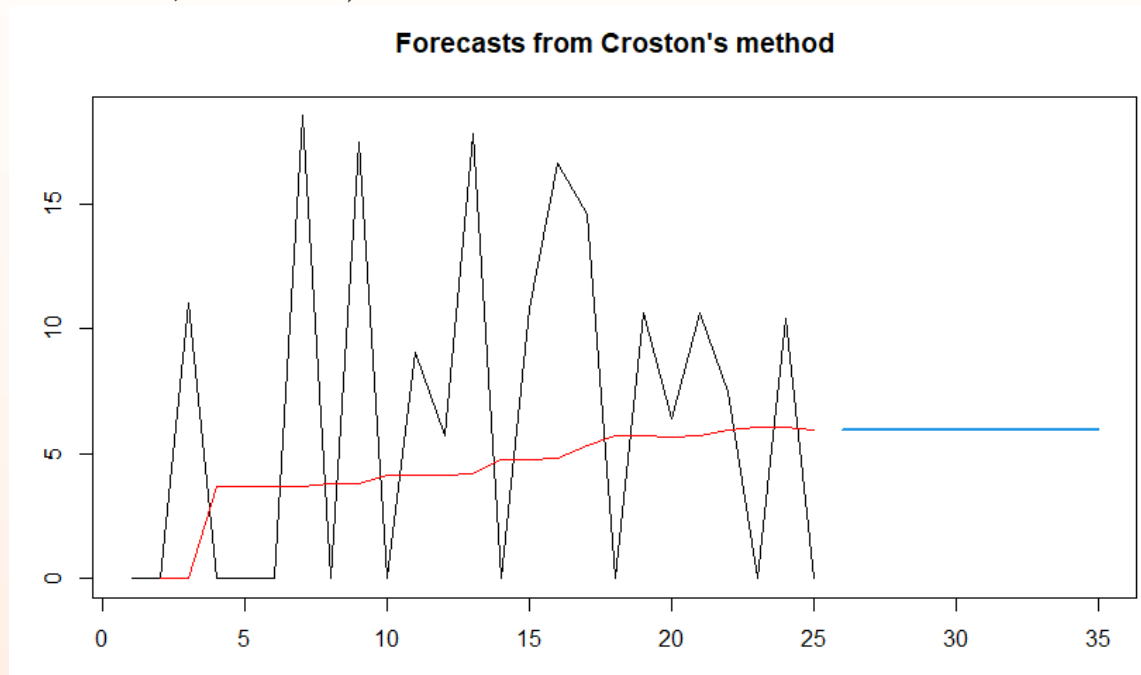
```
model <- croston(series, h = 10, alpha = 0.1)
```

```
model$mean
```

```
par(mfrow=c(1,1))
```

```
plot(model)
```

```
lines(model$fitted, col="red")
```



Εφαρμογή στην R

```
#Εφαρμογή μεθόδου Croston, SBA, TSB και ADIDA
```

```
library(tsintermittent)
```

```
insample <- head(series, 15)
```

```
outsample <- tail(series, 10)
```

```
Croston <- crost(insample, type="croston",h=10)$frc.out
```

```
SBA <- crost(insample, type="sba",h=10)$frc.out
```

```
TSB <- tsb(insample,h=10)$frc.outlibrary(zoo)
```

```
ADIDA <- function(x, h, al){ AS <- as.numeric(na.omit(as.numeric(rollapply(tail(x, (length(x) %/%  
al)*al), al, FUN=sum, by = al))))
```

```
forecast <- rep(as.numeric(ses(AS, h=1)$mean)/al, h)
```

```
return(forecast)}
```

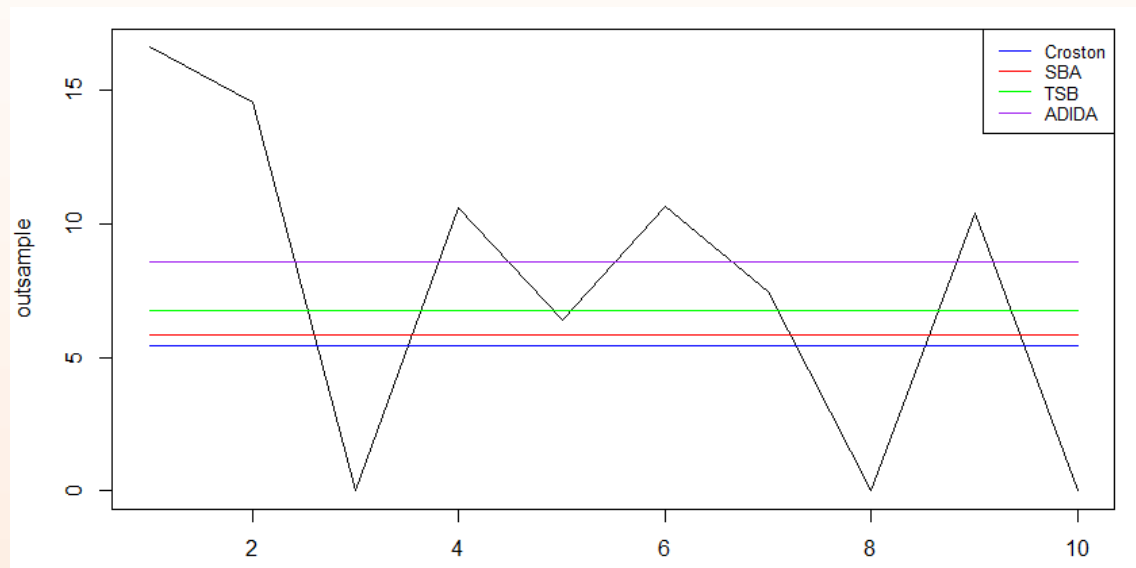
```
ADIDA_frc <- ADIDA(insample, h=10, al=4)
```

Εφαρμογή στην R

#Συγκριτική αξιολόγηση μεθόδων στο out-of-sample με χρήση RMSE

```
plot(outsample, type="l")  
lines(Croston, col="blue") ; lines(SBA, col="red")  
lines(TSB, col="green") ; lines(ADIDA_frc, col="purple")  
legend("topright", legend=c("Croston", "SBA", "TSB", "ADIDA"), col=c("blue", "red", "green", "purple"),  
lty=1, cex=0.8)
```

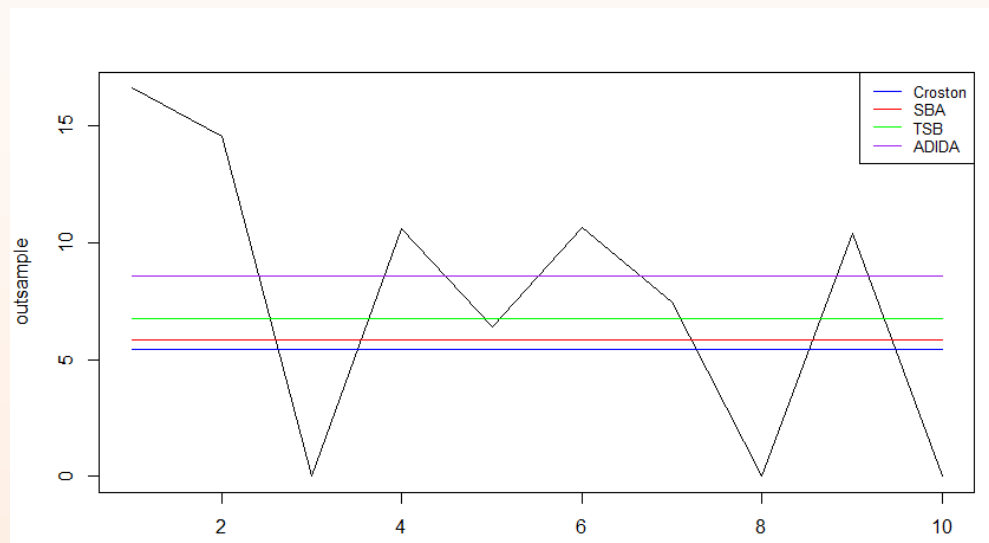
```
rmse_croston <- sqrt(mean((Croston-outsample)^2)) # 6.174  
rmse_sba <- sqrt(mean((SBA-outsample)^2)) # 6.034  
rmse_tsb <- sqrt(mean((TSB-outsample)^2)) # 5.818  
rmse_adida <- sqrt(mean((ADIDA_frc-outsample)^2)) # 5.822
```



Εφαρμογή στην R

```
plot(outsample, type="l")  
lines(Croston, col="blue") ; lines(SBA, col="red")  
lines(TSB, col="green") ; lines(ADIDA_frc, col="purple")  
legend("topright", legend=c("Croston", "SBA", "TSB", "ADIDA"), col=c("blue", "red", "green", "purple"),  
lty=1, cex=0.8)
```

```
rmse_croston <- sqrt(mean((Croston-outsample)^2)) # 6.174  
rmse_sba <- sqrt(mean((SBA-outsample)^2)) # 6.034  
rmse_tsb <- sqrt(mean((TSB-outsample)^2)) # 5.818  
rmse_adida <- sqrt(mean((ADIDA_frc-outsample)^2)) # 5.822  
c(rmse_croston, rmse_sba, rmse_tsb, rmse_adida)
```



Εφαρμογή στην R

#Υπολογισμός επιπέδου παραγγελίας βάσει πιθανοτικής πρόβλεψης

```
n <- 130
```

```
set.seed(105)
```

```
intervals <- rbern(n, 0.4)
```

```
demand <- runif(n,5,20)
```

```
series <- demand*intervals
```

```
insample <- head(series, 100)
```

```
outsample <- tail(series, 30)
```

```
model <- croston(insample, h = 30)
```

```
residuals <- na.omit(insample-as.numeric(model$fitted)) #Compute residual errors
```

```
q <- quantile(residuals, 0.8) #Compute quantile of error
```

```
frc <- as.numeric(model$mean)
```

```
pfrc <- frc + q
```


Εφαρμογή στην R

```
plot(series, type="l", ylim = c(0, max(insample)*1.2))  
lines(c(rep(NA,length(insample)), frc), col="blue")  
lines(c(rep(NA,length(insample)), pfrc), col="red")  
legend("topleft", legend=c("Point", "Quantile"), col=c("blue", "red"), lty=1, cex=0.8)
```

```
#Relative frequency
```

```
1-length(outsample[outsample>pfrc[1]])/length(outsample)
```

