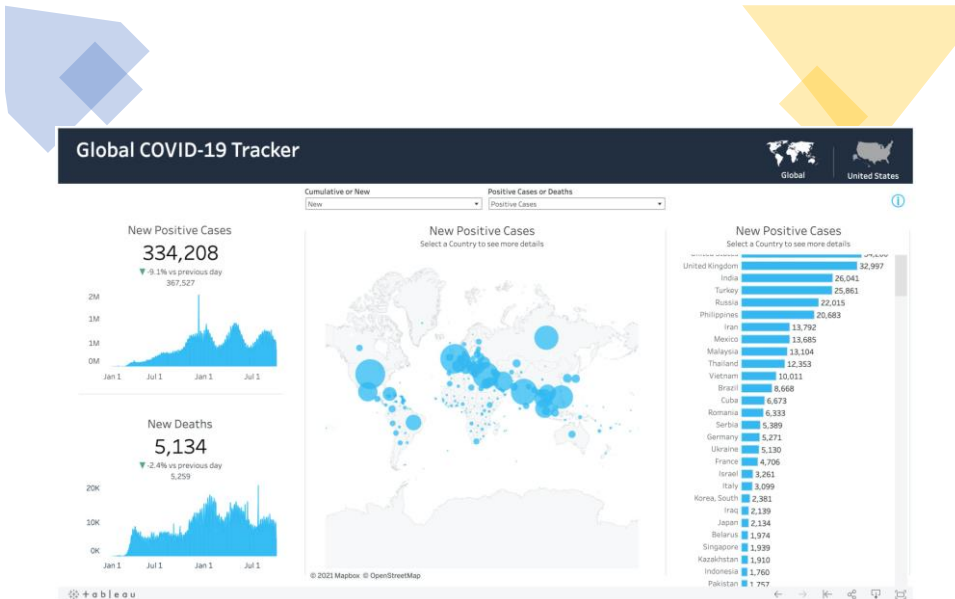


Στατιστικά Μοντέλα Παρακολούθησης της Δημόσιας Υγείας

1

1



2

2

Επιτήρηση της Δημόσιας Υγείας με Χρήση Μεθόδων Στατιστικού Ελέγχου Διεργασιών

3

3

Επιτήρηση της Δημόσιας Υγείας με Χρήση Μεθόδων Στατιστικού Ελέγχου Διεργασιών

- Έλεγχος Διεργασιών (Process Control)
 - Τι ορίζουμε διεργασία;
 - Τι εννοούμε με τον όρο έλεγχος διεργασίας;
- Διαγράμματα Ελέγχου (Control Charts)
 - Τι είναι τα διαγράμματα ελέγχου;
 - Τι σκοπό έχει η χρήση τους;
- Εφαρμογή των Διαγραμμάτων Ελέγχου στον Έλεγχο Μεταβλητών σχετικών με τη Φροντίδα Υγείας
 - Επίπεδα εφαρμογής
 - Παραδείγματα εφαρμογής
- Εφαρμογή των Διαγραμμάτων Ελέγχου στον Έλεγχο Μεταβλητών σχετικών με τη Δημόσια Υγεία
 - Πεδία εφαρμογής
 - Παραδείγματα εφαρμογής

4



Μέρος I: Έλεγχος Διεργασιών

5

Η έννοια της Διεργασίας και του Ελέγχου αυτής

- Διεργασία είναι μια συστηματική σειρά ενεργειών που στοχεύουν στην επίτευξη ενός στόχου.
- Έλεγχος διεργασίας είναι μια διαδικασία ανάδρασης μέσω της οποίας
 - Μετρούμε το αποτέλεσμα της διεργασίας
 - Το συγκρίνουμε με κάποιο πρότυπο – προδιαγραφή
 - Ενεργούμε βάσει της διαφοράς.

Παράδειγμα η διεργασία εξυπηρέτησης ασθενών στα επείγοντα ενός νοσοκομείου...

6

Στατιστικός Έλεγχος Διεργασιών

- Ο σκοπός του στατιστικού ελέγχου διεργασιών είναι να διακρίνει έγκαιρα μια εκτός ελέγχου διεργασία.
 - **Εντός στατιστικού ελέγχου Διεργασία:**
 - Μια διεργασία που παρουσιάζει μόνο τη φυσική της τυχαία μεταβλητότητα (είναι σε σταθερή κατάσταση).
 - **Εκτός στατιστικού ελέγχου Διεργασία:**
 - Μια διεργασία στην οποία εμφανίζονται άλλες συστηματικές - ειδικές μορφές μεταβλητότητας.
 - Ο στατιστικός έλεγχος διεργασιών είναι ένα σύνολο από στατιστικές τεχνικές που αναπτύχθηκαν αρχικά για τον έλεγχο παραγωγικών διεργασιών.
- **Ο μέσος χρόνος εξυπηρέτησης των ασθενών στα επείγοντα ενός νοσοκομείου είναι ίσος με 10 λεπτά...**
- **Ο μέσος χρόνος εξυπηρέτησης των ασθενών στα επείγοντα ενός νοσοκομείου είναι ίσος με 25 λεπτά...**

7

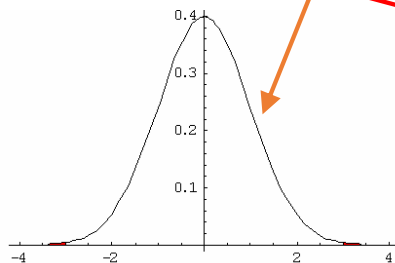
Διαγράμματα Ελέγχου

- Το κυριότερο εργαλείο του στατιστικού ελέγχου διεργασιών είναι το **διάγραμμα ελέγχου**.
- Το **διάγραμμα ελέγχου** ως στόχο έχει να αναγνωρίσει την ύπαρξη ειδικών λόγων μεταβλητότητας σε μια μη σταθερή διεργασία.
- Τα **διαγράμματα ελέγχου** που έχουν αναπτυχθεί μέχρι σήμερα είναι:
 - Shewhart type Control Charts
 - CUmulative SUM type Control Charts (CUSUM)
 - Exponential Weighted Moving Average type Control Charts (EWMA).
- Τα πιο συχνά χρησιμοποιούμενα διαγράμματα ελέγχου είναι τα διαγράμματα τύπου Shewhart.
- Στα διαγράμματα ελέγχου τύπου Shewhart απεικονίζουμε κατάλληλα μια παράμετρο της διεργασίας.

8

Διαγράμματα Τύπου Shewhart

- Θεωρώντας ότι το υπό παρακολούθηση χαρακτηριστικό X ακολουθεί την Κανονική Κατανομή με μέση τιμή μ_0 και διακύμανση σ^2 τότε η μέση τιμή ενός δείγματος n τιμών της μεταβλητής στο χρόνο $t=1,2,\dots$ θα ανήκει με πιθανότητα 0.9973 στο διάστημα



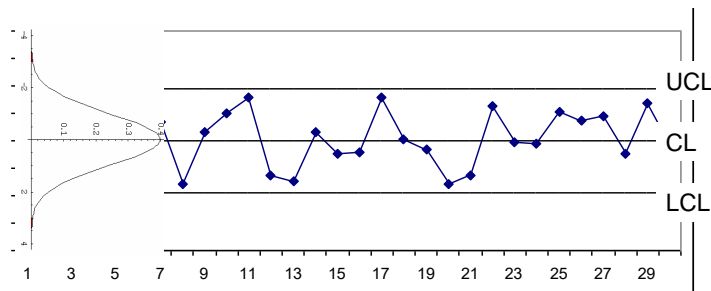
$$\left[\mu_0 - 3\sqrt{\frac{\sigma^2}{n}}, \mu_0 + 3\sqrt{\frac{\sigma^2}{n}} \right]$$

Ο χρόνος εξυπηρέτησης των ασθενών στα επείγοντα ενός νοσοκομείου, κατάλληλα μετασχηματισμένος, ακολουθεί κανονική κατανομή...

9

Διαγράμματα Τύπου Shewhart

Κάθε χρονική στιγμή ελέγχουμε εάν η μέση τιμή είναι ίση με 10 λεπτά...



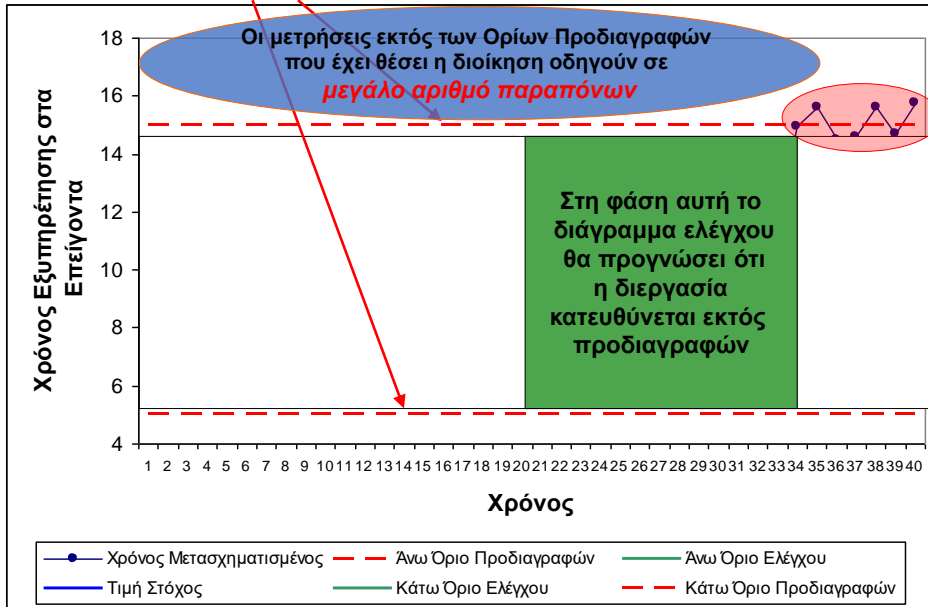
Ο έλεγχος για το εάν ο δειγματικός μέσος ανήκει στο διάστημα αυτό είναι αντίστοιχος με διαδοχικούς ελέγχους υποθέσεων της μορφής

$$\begin{aligned} LCL &= \mu_0 - 3\sqrt{\frac{\sigma^2}{n}} \\ CL &= \mu_0 \\ UCL &= \mu_0 + 3\sqrt{\frac{\sigma^2}{n}} \end{aligned}$$

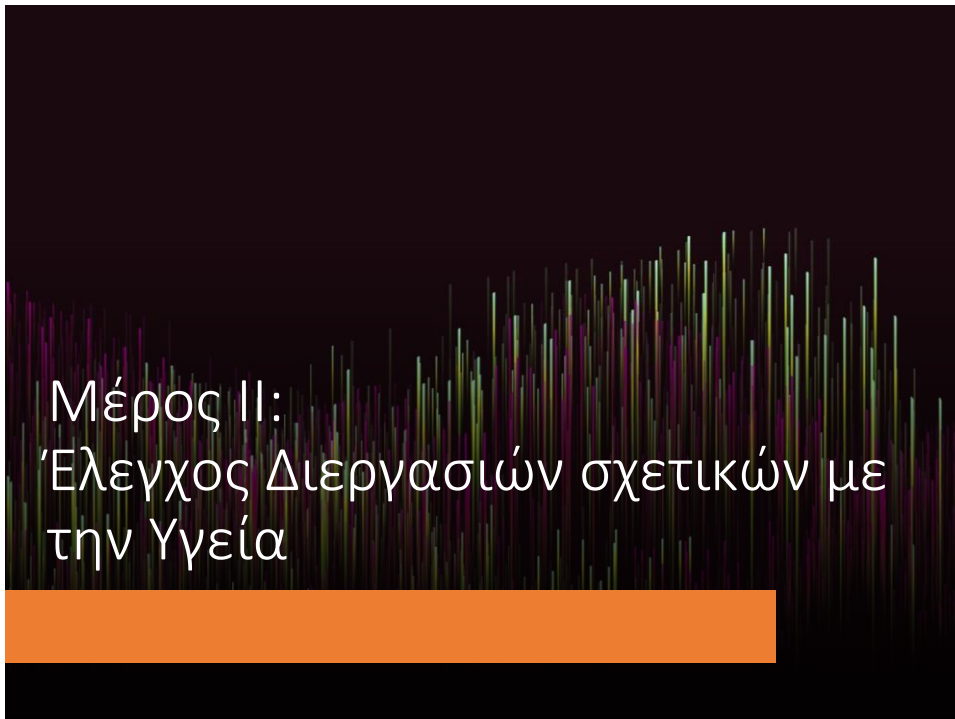
10

Παράδειγμα ροής μίας διεργασίας εξυπηρέτησης στα Επείγοντα....

Η διοίκηση θέτει προδιαγραφές εξυπηρέτησης....



11



12

Έλεγχος Διεργασιών σχετικών με την Υγεία

- Το **διάγραμμα ελέγχου** ως σκοπό έχει να αναγνωρίσει έγκαιρα την ύπαρξη ειδικών λόγων μεταβλητότητας σε μια μη σταθερή διεργασία.
- Το **διάγραμμα ελέγχου** δηλαδή έχει ένα και μοναδικό στόχο, την έγκαιρη διάγνωση μιας μη σταθερής διεργασίας, η οποία παρουσιάζει συστηματική μεταβλητότητα (**αυξητική τάση** κτλ).
- Τα **διαγράμματα ελέγχου** είναι αποτελεσματικότερα όταν χρησιμοποιούνται σε ένα γενικότερο πλαίσιο ελέγχου της ποιότητας των παρεχόμενων υπηρεσιών (ISO, 6-Sigma).
- Σε αντίθεση με τις μεταβλητές που συναντώνται στη βιομηχανία, οι μεταβλητές που ελέγχονται στην υγεία είναι συνήθως είτε ποιοτικές, είτε ποσοτικές, οι οποίες δεν ακολουθούν την κανονική κατανομή.

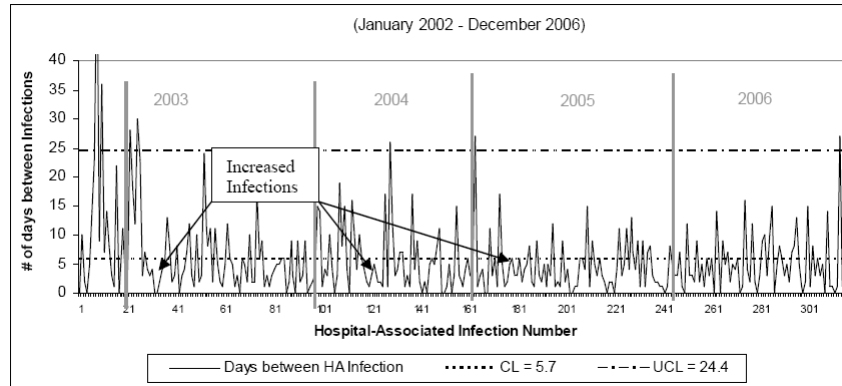
13

Έλεγχος Διεργασιών σχετικών με την Υγεία

- Ο χρόνος
 - διενέργειας εργαστηριακών εξετάσεων
 - από την ταχυβιοψία στην τελική βιοψία
 - από τη στιγμή της τηλεφωνικής κλήσης έως και την παραλαβή από ασθενοφόρο και τη μεταφορά στα επείγοντα
 - παραμονής στο νοσοκομείο μετά την επέμβαση
 - Ο αριθμός
 - των εσφαλμένων αγωγών / θεραπειών
 - των μολύνσεων (ενδοноσοκομειακών)
 - των επιμολύνσεων στις ΜΕΘ
 - των θανάτων σε μια κλινική / νοσοκομείων
 - Η ικανοποίηση των ασθενών
 - Παράμετροι Ιατρικού φακέλου Ασθενούς
-
- The diagram consists of three blue boxes on the right side, each connected to a list of variables on the left by a blue bracket. The top box, labeled 'Εκθετική ή γεωμετρική κατανομή', is connected to the 'Ο χρόνος' category. The middle box, labeled 'Κατανομή Poisson', is connected to the 'Ο αριθμός' category. The bottom box, labeled 'Διάφορες διακριτές ή μη κανονικές μεταβλητές', is connected to the 'Η ικανοποίηση των ασθενών' and 'Παράμετροι Ιατρικού φακέλου Ασθενούς' categories.

14

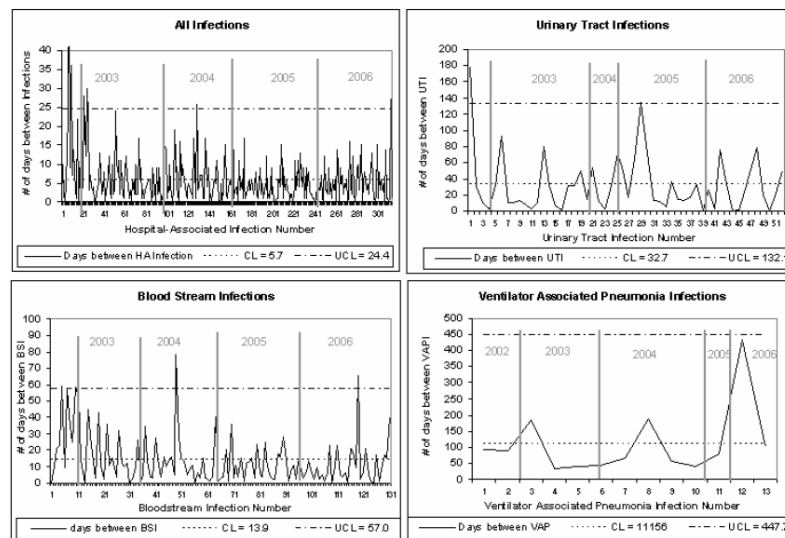
Στατιστικός Έλεγχος Διεργασιών Νοσηλευτικών Ιδρυμάτων



- Hospital-associated infections are a major concern in the medical community due to the potential loss of life and high costs. Monitoring the incidences of infections is an established part of quality maintenance programs in hospitals. However, traditional methods of analysis are often inadequate since the incidences of infections are infrequent. In order to address this issue, techniques such as the cumulative sum (CUSUM) chart for counted data, the exponentially weighted moving average (EWMA) chart, and the g-type control chart have been suggested. This paper demonstrates how these charts may be applied to infection control surveillance data Children's Hospital, compares their performance, and makes recommendations for a control chart most suitable for monitoring hospital-associated infections.

15

Στατιστικός Έλεγχος Διεργασιών Νοσηλευτικών Ιδρυμάτων



16

Ιατρικά Λάθη

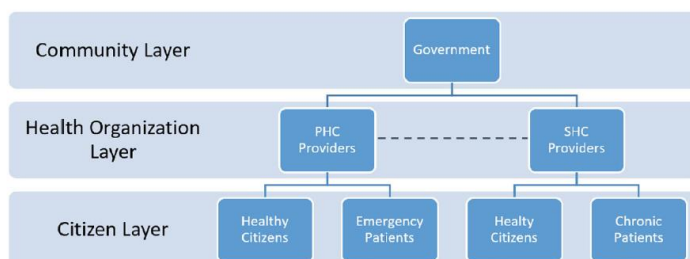
- Στις ΗΠΑ με βάση στοιχεία που δημοσιεύθηκαν σε μελέτη του 2000 οι θάνατοι που είναι πιθανόν να σχετίζονται με ιατρικά λάθη υπολογίζονται σε 44.000-98.000 ετησίως.
- Συγκεκριμένα οι θάνατοι
 - από μη απαιτούμενες εγχειρήσεις ήταν 12.000
 - από ενδονοσοκομειακές μολύνσεις ήταν 80.000
 - από εσφαλμένη αγωγή ήταν 7.000.
- Στην Αυστραλία σε μία μελέτη 14.000 ιατρικών φακέλων αποκαλύφθηκε ότι περίπου το 16,6 % των εισαγωγών σε νοσοκομεία οφείλονταν σε πιθανά ιατρικά λάθη.
- Οι θάνατοι αυτοί αναφέρονται διεθνώς ως “preventable adverse events” και στην Ελλάδα ως «Ιατρικά Λάθη».
- <http://www.mednet.gr/app/webroot/happenings/iatrika-lathi-praktika.pdf>

17

17

Επίπεδα Εφαρμογής στη φροντίδα Υγείας

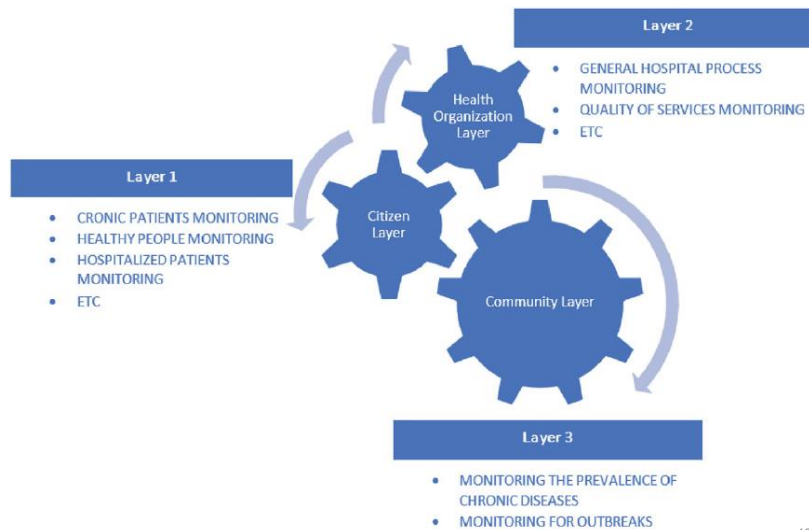
- Οι μεθοδολογίες του ελέγχου διεργασιών και ειδικότερα των διαγραμμάτων ελέγχου είναι δυνατό να εφαρμοσθούν σε 3 επίπεδα
 - 1^ο επίπεδο: το επίπεδο του ατόμου
 - 2^ο επίπεδο: το επίπεδο του Οργανισμού Υγείας
 - 3^ο επίπεδο: το επίπεδο της υγείας της Κοινότητας (εκτός νοσοκομείου)



18

18

Επίπεδα Εφαρμογής στη φροντίδα Υγείας



19

1ο Επίπεδο Εφαρμογής: Επίπεδο του Ατόμου / Πλαίσιο

- Τα διαγράμματα ελέγχου συνδυάζονται επιτυχώς με τον ηλεκτρονικό φάκελο του ασθενή προκειμένου να παρακολουθούνται και να ελέγχονται κρίσιμοι ατομικοί δείκτες αυτού μέσα από πληθώρα διαγνωστικών εξετάσεων.
- Συνήθως παρακολουθούνται άτομα που εμπίπτουν στις ακόλουθες ομάδες:
 - Πάσχοντες από Υπέρταση,
 - Πάσχοντες από Σακχαρώδη Διαβήτη τύπου II,
 - Πάσχοντες από καρδιαγγειακά νοσήματα.

20

20

1ο Επίπεδο Εφαρμογής: Επίπεδο του Ατόμου / Πλαίσιο

- Δείκτες που παρακολουθούνται σε ατομικό επίπεδο, είναι:
 - Η αρτηριακή πίεση, το βάρος, το σάκχαρο, η LDH και η HDL σε υπερτασικούς ασθενείς αλλά και σε ασθενείς με καρδιαγγειακά νοσήματα.
 - Διάφοροι δείκτες που σχετίζονται με την εκτίμηση του διαβητικού ασθενή (HbA1c, BMI, κτλ).
 - Δεδομένα από την κάρτα του μαθητή (εμβολιαστική κάλυψη, BMI, χοληστερόλη, κτλ).
 - Το όφελος της εφαρμογής ελέγχου διεργασιών στους ατομικούς δείκτες ασθενών, αλλά και υγιών πολιτών της κοινότητας, είναι η έγκαιρη ανίχνευση πιθανών κινδύνων και η ενημέρωση του ιατρού που έχει την ευθύνη.
-
- Ενδεικτικές αναφορές:
 - a. Henderson et al (2008). Use of statistical process control charts in stroke medicine to determine if clinical evidence and changes in service delivery were associated with improvements in the quality of care, *Quality and Safety in Health Care*;17:301-306.
 - b. Haroon et al (2006). Continuous monitoring of the performance of hip prostheses. *The Journal of Bone and Joint Surgery*, 88-B(6), 716 – 720.

21

1ο Επίπεδο Εφαρμογής: Επίπεδο του Ατόμου / Πλαίσιο

Διάγραμμα ελέγχου Φάσης II

3σ Όρια Ελέγχου Φάσης II

$$UCL = \mu + 3\sigma$$

$$CL = \mu$$

$$LCL = \mu - 3\sigma$$

Διάγραμμα ελέγχου Φάσης I

3σ Όρια Ελέγχου Φάσης I

$$UCL = \bar{X} + 3\frac{\overline{MR}}{d_2}$$

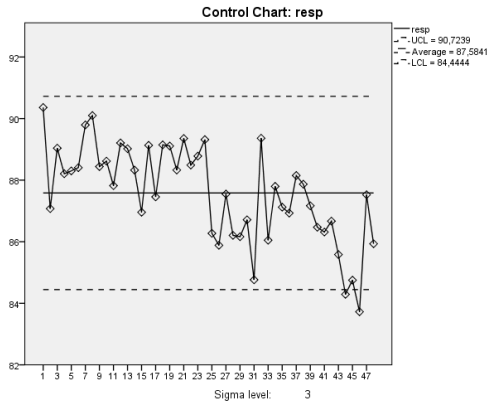
$$CL = \bar{X}$$

$$LCL = \bar{X} - 3\frac{\overline{MR}}{d_2}$$

22

22

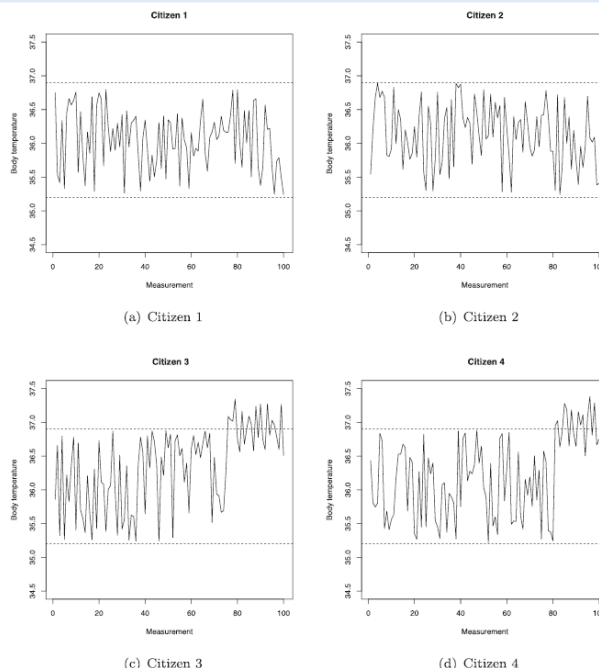
Εφαρμογή: Οξυγόνο ασθενούς στην ΜΕΘ



Ωρα	ημέρα 1η	Ημέρα 2η
1	90,36078	86,27219
2	87,0744	85,88259
3	89,03916	87,54861
4	88,20893	86,2059
5	88,30101	86,16197
6	88,40544	86,71219
7	89,7939	84,76305
8	90,09844	89,36186
9	88,43993	86,0486
10	88,61756	87,79824
11	87,82154	87,12223
12	89,20836	86,92951
13	89,0209	88,15063
14	88,3258	87,86828
15	86,95731	87,16642
16	89,1311	86,47007
17	87,4571	86,31393
18	89,14529	86,6687
19	89,10857	85,58326
20	88,32774	84,2903
21	89,35456	84,75099
22	88,48936	83,72402
23	88,78062	87,52656
24	89,31633	85,93324

23

Εφαρμογή: Παρακολούθηση θερμοκρασίας σώματος



24

24

2ο Επίπεδο Εφαρμογής: Επίπεδο του Οργανισμού Υγείας / Πλαίσιο

- Τα διαγράμματα ελέγχου εφαρμόζονται στο σύνολο των διεργασιών του οργανισμού υγείας και στοχεύουν στον έλεγχο της καλής λειτουργίας κάθε μίας από αυτές.
- Η παρακολούθηση και ο έλεγχος όλων των διακριτών διεργασιών του οργανισμού υγείας εγγυάται τη συνολική καλή λειτουργία του.
- Η παρακολούθηση και ο έλεγχος συνήθως στοχεύει
 - στην ποιότητα και την ασφάλεια των παρεχόμενων υπηρεσιών κυρίως μέσα από κατάλληλα επιλεγμένους υγειονομικούς δείκτες όσο και
 - στη σωστή διαχείριση του οργανισμού και των πόρων αυτού με χρήση κατάλληλα επιλεγμένων ή κατασκευασμένων ποσοτικών δεικτών.

25

25

2ο Επίπεδο Εφαρμογής: Επίπεδο του Οργανισμού Υγείας / Πλαίσιο

- Ενδεικτικά, μεταξύ άλλων παρακολουθούνται:
 - στο επίπεδο της ποιότητας των παρεχόμενων υπηρεσιών και σε υγειονομικούς δείκτες: (i) οι χρόνοι εξυπηρέτησης ασθενών, (ii) ο αριθμός παραπόνων ασθενών, (iii) ο βαθμός ικανοποίησης (iv) ο χρόνος διενέργειας εργαστηριακών εξετάσεων, (v) ο αριθμός των ιατρικών σφαλμάτων, (vi) ποσοστά επιμολύνσεων, (vii) χρονική περίοδος νοσηλείας, κ.α.
 - στο επίπεδο της διαχείρισης του οργανισμού και των πόρων αυτού: (i) η κατανάλωση υλικών, (ii) το απόθεμα πρώτων υλών, (iii) η κατανάλωση συγκεκριμένων αναλώσιμων ειδών, (iv) η κατανάλωση φαρμάκων, (v) τα αποθέματα σε είδη και τρόφιμα, (vi) το κόστος λειτουργίας κτλ.

26

26

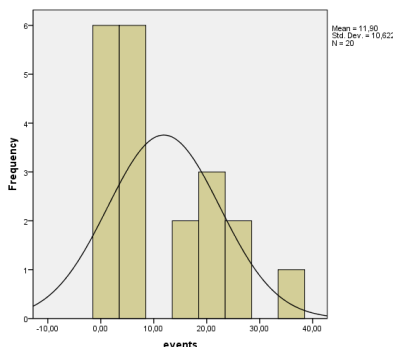
2ο Επίπεδο Εφαρμογής: Επίπεδο του Οργανισμού Υγείας / Πλαίσιο

- Το όφελος της εφαρμογής μεθόδων ελέγχου στο σύνολο των διεργασιών του οργανισμού είναι:
 - η έγκαιρη ανίχνευση πιθανών κινδύνων στους χώρους του οργανισμού υγείας (επιμολύνσεις κτλ),
 - η έγκαιρη ανίχνευση της υποβάθμισης των παρεχόμενων υπηρεσιών προς τους ασθενείς (αύξηση χρόνων διενέργειας εργαστηριακών εξετάσεων, αύξηση χρόνων αναμονής ασθενών κτλ),
 - η βελτίωση της διοίκησης μέσω της διαρκούς παρακολούθησης οικονομικών ή σχετικών στοιχείων (κατανάλωση συγκεκριμένων αναλωσίμων ειδών, κτλ).
- Steiner S et al. (1999) Monitoring Paired Binary Surgical Outcomes Using Cumulative Sum Charts. *Statistics in Medicine*, 18.
 - Steiner et al. (2000). Monitoring surgical performance using risk-adjusted cumulative sum charts, *Biostatistics*, 1, 441-452.
 - Flaatten H. (2005). Effects of a major structural change to the intensive care unit on the quality and outcome after intensive care, *Quality and Safety in Health Care*, 14(4), 270 - 272.
 - M. Goddard, et al. (2002). Clinical performance measurement: part 1--getting the best out of it, *Journal of the Royal Society of Medicine*, 95 (10), 508 - 510.
 - R.G. Carey and L.V. Stake (2003). *Improving Healthcare with Control Charts: Basic and Advanced SPC Methods and Case Studies*, American Society for Quality Publications.

27

Εφαρμογή: Χρόνοι μεταξύ συμβάντων

- Σε νοσοκομείο της επαρχίας καταγράφηκε ο αριθμός των ημερών ανάμεσα σε κρούσματα ενδονοσοκομειακών λοιμώξεων.
- Χρόνοι μεταξύ εμφανίσεων σπάνιων ενδεχομένων προσεγγίζονται από την εκθετική κατανομή.



28

Ημέρες Μεταξύ Λοιμώξεων
4
8
8
1
19
28
23
36
7
8
3
22
15
1
15
8
3
27
1
1

Εφαρμογή: Χρόνοι μεταξύ συμβάντων

- Εφαρμόζουμε το μετασχηματισμό του Nelson $x = \gamma^{1/3.6}$ (περνάμε από μια εκθετική γ σε μια Weibull x η οποία προσεγγίζεται από την κανονική).

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
transf	,123	20	,200 [*]	,927	20	,136

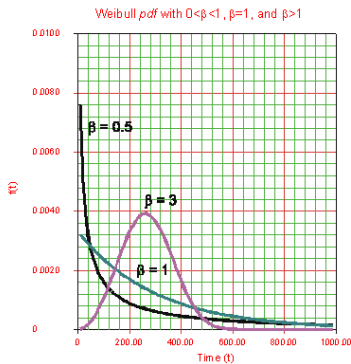


Figure 6-1: The effect of the Weibull shape parameter on the pdf.

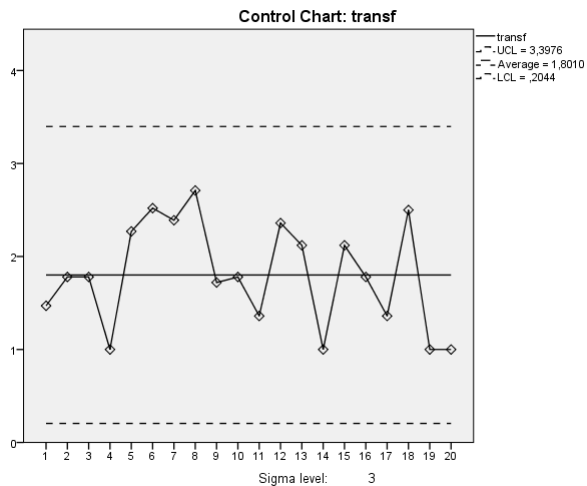
29

Ημέρες Μεταξύ Λοιμώξεων

1,47
1,78
1,78
1,00
2,27
2,52
2,39
2,71
1,72
1,78
1,36
2,36
2,12
1,00
2,12
1,78
1,36
2,50
1,00
1,00

Εφαρμογή: Χρόνοι μεταξύ συμβάντων

- Παρατηρούμε ότι η διεργασία είναι εντός ελέγχου



30

30

Εφαρμογή: Αριθμός συμβάντων

Το c διάγραμμα ελέγχου με 3σ όρια ελέγχου δίνεται ως

3σ Όρια Ελέγχου Φάσης II
$UCL = c + 3\sqrt{c}$
$CL = c$
$LCL = c - 3\sqrt{c}$

3σ Όρια Ελέγχου Φάσης I
$UCL = \bar{C} + 3\sqrt{\bar{C}}$
$CL = \bar{C}$
$LCL = \bar{C} - 3\sqrt{\bar{C}}$

Αν το LCL είναι αρνητικό, το θέτουμε ίσο με το 0.

31

31

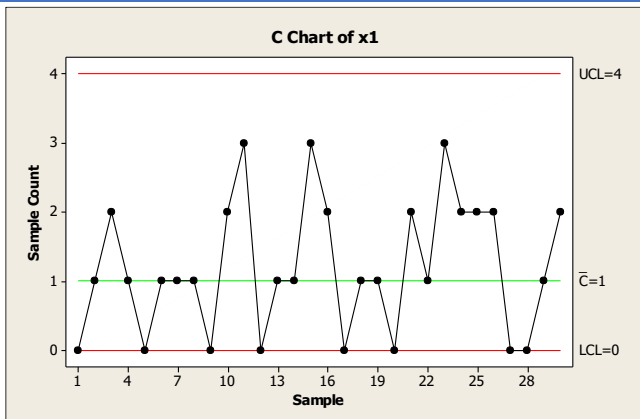
Εφαρμογή: Αριθμός συμβάντων

Ημέρα	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Κρούσματα λοιμώξεων του αναπνευστικού	0	1	2	1	0	1	1	1	0	2	3	0	1	1	3
Κρούσματα λοιμώξεων του ουροποιητικού	0	2	3	0	1	1	3	1	0	1	1	0	1	1	1
Κρούσματα λοιμώξεων του κυκλοφορικού	2	2	1	0	2	1	0	1	2	2	1	0	1	1	1
Ημέρα	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Κρούσματα λοιμώξεων του αναπνευστικού	2	0	1	1	0	2	1	3	2	2	2	0	0	1	2
Κρούσματα λοιμώξεων του ουροποιητικού	0	1	1	0	2	1	2	4	4	4	0	0	1	1	2
Κρούσματα λοιμώξεων του κυκλοφορικού	1	1	3	1	1	3	1	4	4	4	2	1	1	3	2

32

α. Θεωρώντας ότι τα κρούσματα λοιμώξεων του αναπνευστικού είναι τυχαία μεταβλητή (έστω X_1) η οποία ακολουθεί την κατανομή Poisson με παράμετρο λ_1 ίση με 1.0 (μελέτη Φάσης II).

Το κατάλληλο διάγραμμα ελέγχου είναι το c-chart.

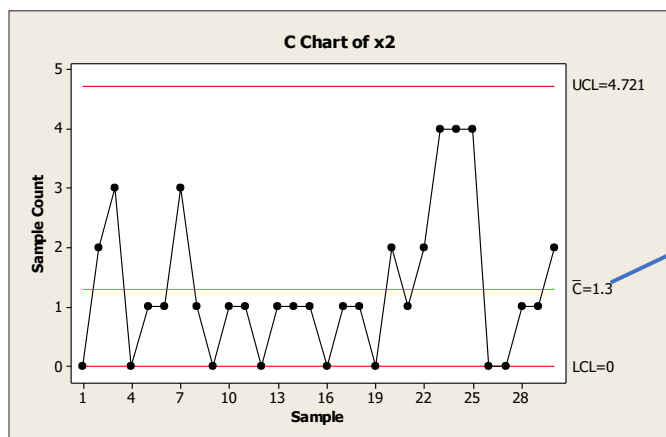


Όπως προκύπτει από το σχήμα ο αριθμός των κρουσμάτων στην περίπτωση αυτή είναι εντός ελέγχου. Δεν παρατηρούμε κάποιο ιδιαίτερο μοτίβο στο διάγραμμα ελέγχου.

33

β. Θεωρώντας αυτή τη φορά ότι βρισκόμαστε σε μελέτη Φάσης I και ότι τα κρούσματα λοιμώξεων του Ουροποιητικού είναι τυχαία μεταβλητή (έστω X_2) η οποία ακολουθεί την κατανομή Poisson με κάποια άγνωστη παράμετρο λ_2 ,

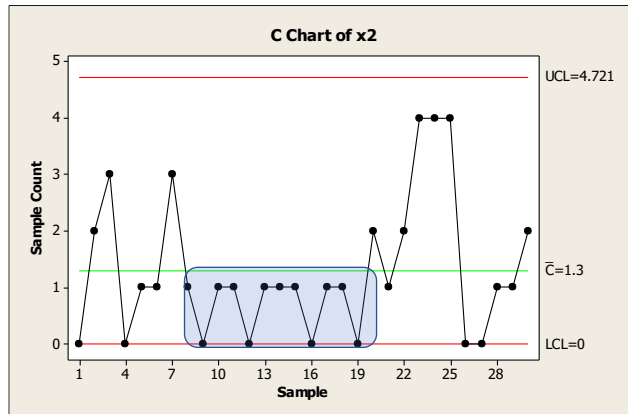
Το κατάλληλο διάγραμμα ελέγχου είναι το c-chart.



$$\lambda_2 = 1.3$$

34

Όπως προκύπτει ο αριθμός των κρουσμάτων και στην περίπτωση αυτή είναι εντός ελέγχου εάν θεωρήσουμε τον απλό κανόνα του ενός σημείου εκτός των ορίων ελέγχου. Σε αντίθετη περίπτωση θεωρώντας και έναν σύνθετο κανόνα όπως «9 σημεία κάτω της κεντρικής γραμμής» μπορούμε να συμπεράνουμε ότι η διεργασία είναι εκτός ελέγχου.

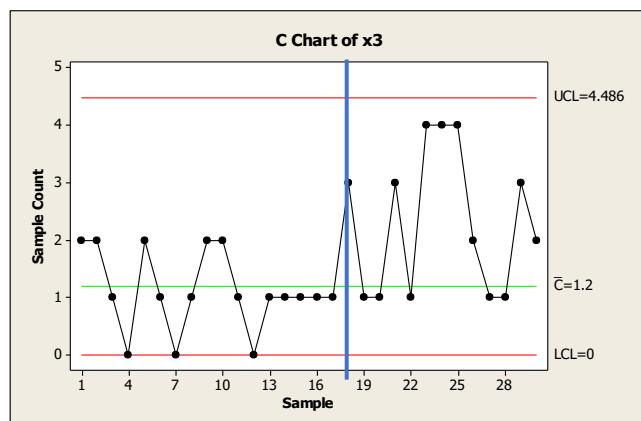


35

γ. Θεωρώντας ότι τα κρούσματα λοιμώξεων του κυκλοφορικού είναι τυχαία μεταβλητή (έστω X_3) η οποία ακολουθεί την κατανομή Poisson με παράμετρο λ_3 ίση με 1.2 (μελέτη Φάσης II),

Το κατάλληλο διάγραμμα ελέγχου είναι το c-chart.

Όπως προκύπτει ο αριθμός των κρουσμάτων είναι εντός ελέγχου, αξιοποιώντας τον κανόνα που κάνει χρήση αποκλειστικά των ορίων ελέγχου. Με μια πιο προσεκτική ματιά παρατηρούμε ότι μετά την 18^η παρατήρηση έχουμε γενικά μεγαλύτερες τιμές, δηλαδή, παρουσιάζεται κάποια έξαρση κρουσμάτων.

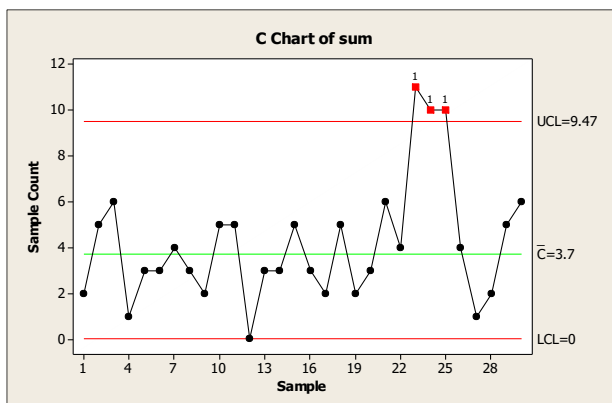


36

δ. Θεωρώντας ότι τα κρούσματα λοιμώξεων του αναπνευστικού (X_1), τα κρούσματα λοιμώξεων του ουροποιητικού (X_2) και τα κρούσματα λοιμώξεων του κυκλοφορικού (X_3) είναι ανεξάρτητες τυχαίες μεταβλητές οι οποίες ακολουθούν την κατανομή Poisson με παράμετρο $\lambda_1=1.0$, $\lambda_2=1.5$, και $\lambda_3=1.2$, αντίστοιχα

Αθροίζουμε τα δεδομένα από τα 3 διαφορετικά είδη λοιμώξεων. Υποθέτοντας την ανεξαρτησία, καταλήγουμε σε μια μεταβλητή που ακολουθεί την Poisson με παράμετρο το άθροισμα των παραμέτρων. Επιλέγουμε c-chart και διενεργούμε ανάλυση Φάσης II / ανάλυση με δεδομένο πρότυπο, και το αποτέλεσμα είναι το ακόλουθο διάγραμμα ελέγχου

Όπως προκύπτει εξετάζοντας τον συνολικό αριθμό των κρουσμάτων μια έξαρση εμφανίζεται την περίοδο ανάμεσα στην 23 και την 25 χρονική περίοδο.



37

Εφαρμογή: Αριθμός συμβάντων

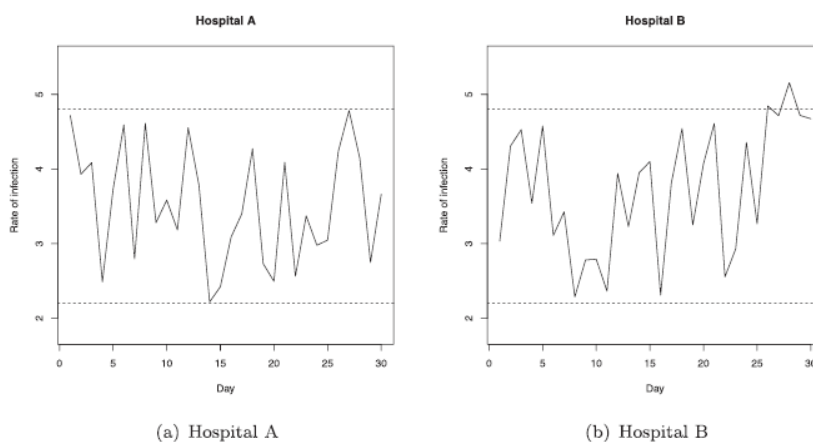


Figure 5. p -charts for the monitoring of the rate of patients experiencing a hospital infection into two hospitals.

38

38

3ο Επίπεδο Εφαρμογής: Επίπεδο της κοινότητας / Πλαίσιο

- Τα διαγράμματα ελέγχου εφαρμόζονται με πολλαπλούς τρόπους στον τομέα της παρακολούθησης της υγείας της κοινότητας (γενικά μιας στενής γεωγραφικής περιοχής).
- Με την καταγραφή των περιστατικών που εμφανίζονται στην κοινότητα εφαρμόζεται συνεχής και αδιάλειπτος έλεγχος της συχνότητας εμφάνισης συγκεκριμένων περιστατικών.
- Υποθέτουμε ότι υπό κανονικές συνθήκες ο αριθμός των καθημερινών (μηνιαίων κτλ) συμβάντων σε μια περιοχή είναι ανεξάρτητος της προηγούμενης.
- Παρακολουθούνται συνεχώς δείκτες όπως
 - Ημερήσιος αριθμός κρουσμάτων γρίπης και άλλων λοιμωδών νόσων (AIDS κτλ)
 - Μηνιαίος αριθμός νέων κρουσμάτων μη λοιμωδών νόσων (καρκίνος του τραχήλου της μήτρας, του μαστού, λευχαιμίες κτλ),
 - Μηνιαίος αριθμός νέων διαγνώσεων Alzheimer, κτλ.

39

39

3ο Επίπεδο Εφαρμογής: Επίπεδο της κοινότητας / Πλαίσιο

- Ο στόχος της εφαρμογής των παραπάνω μεθόδων σε περιστατικά συγκεκριμένων ασθενειών είναι η έγκαιρη ανίχνευση της αύξησης του επιπολασμού μη μεταδιδόμενων (χρόνιων) νοσημάτων, εξάρσεων γρίπης ή η έγκαιρη ανίχνευση παραγόντων κινδύνου, αλλά και η αποτίμηση της αποτελεσματικότητας των προγραμμάτων παρέμβασης.

- Heffernan et al. (2004). Syndromic surveillance in public health practice, New York City. *Emerging Infectious Diseases*; 10(5), 858–864.
- Carey and Stake (2003). *Improving Healthcare with Control Charts: Basic and Advanced SPC Methods and Case Studies*, American Society for Quality Publications.
- Sonesson and Bock (2003). A review and discussion of prospective statistical surveillance in public health. *Journal of the Royal Statistical Society A*; 166(1), 5–21.

40

40

3ο Επίπεδο Εφαρμογής: Επίπεδο της κοινότητας / Πλαίσιο

Το p διάγραμμα ελέγχου με 3σ όρια ελέγχου δίνεται ως

3σ Όρια Ελέγχου Φάσης II

$$UCL = p + 3\sqrt{\frac{p(1-p)}{n}}$$

$$CL = p$$

$$LCL = p - 3\sqrt{\frac{p(1-p)}{n}}$$

Παρακολουθούμε τον επιπολασμό μιας νόσου, ποσοστά ανεπιτυχών εγχειρήσεων κτλ

3σ Όρια Ελέγχου Φάσης I

$$UCL = \bar{P} + 3\sqrt{\frac{\bar{P}(1-\bar{P})}{n}}$$

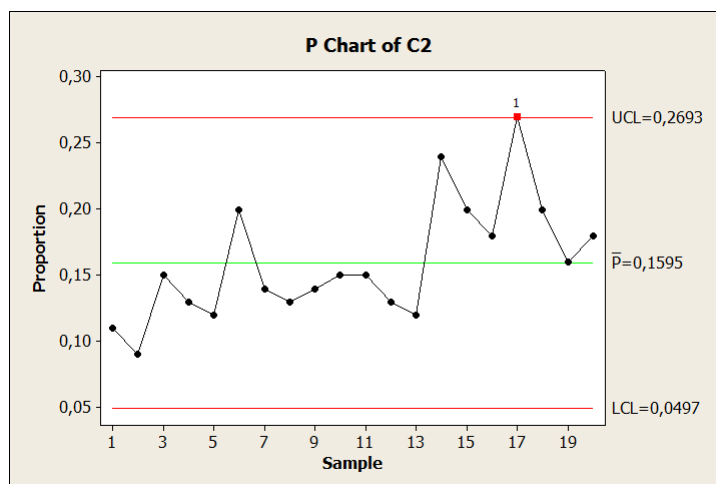
$$CL = \bar{P}$$

$$LCL = \bar{P} - 3\sqrt{\frac{\bar{P}(1-\bar{P})}{n}}$$

41

41

Εφαρμογή: Ποσοστά rotavirus στο σύνολο των παιδιών που εξετάστηκαν ανά μήνα



42

42

3ο Επίπεδο Εφαρμογής: Επίπεδο της κοινότητας / Πλαίσιο Risk Adjustment

- Πολλές φορές στην πράξη ανακύπτει η ανάγκη να υπολογίσουμε την πιθανότητα ενός συμβάντος (π.χ. ο θάνατος ενός ασθενούς μετά από κάποιο δύσκολο χειρουργείο, ή ο επιπολασμός μιας ασθένειας σε μια περιοχή / ηλικιακή ομάδα / ομάδα γενικά) λαμβάνοντας υπόψη μια σειρά από παράγοντες κινδύνου:
- Για παράδειγμα σε χειρουργεία καρδιακών νόσων:
 - Η ηλικία, το φύλο, η υπέρταση, ο διαβήτης, η κατάσταση των νεφρών, η κατάσταση των αρτηριών, η παρουσία ενδοκαρδίτιδας, η κρεατίνη, κτλ
- Τα διαγράμματα ελέγχου προσαρμοσμένα ως προς τον κίνδυνο (risk adjusted control charts) λαμβάνουν υπόψη τις διάφορες συνθήκες υγείας των ασθενών.
- Τέτοιου είδους διαγράμματα χρησιμοποιούνται για την παρακολούθηση ορισμένων ιατρικών διαδικασιών, όπως χειρουργικές επεμβάσεις, θνησιμότητα και εμπειρία των γιατρών.

43

43

Εφαρμογή: Ποσοστά rotavirus στο σύνολο των παιδιών που εξετάστηκαν ανά μήνα

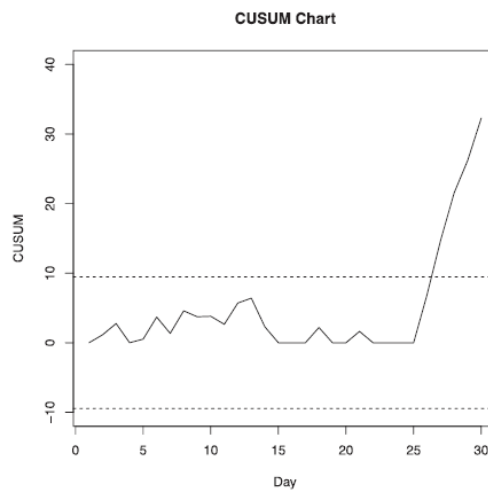


Figure 6. A CUSUM control chart for monitoring the number of citizens appearing symptoms of high body temperature.

44

44

Εφαρμογή: RA CC HbA1C

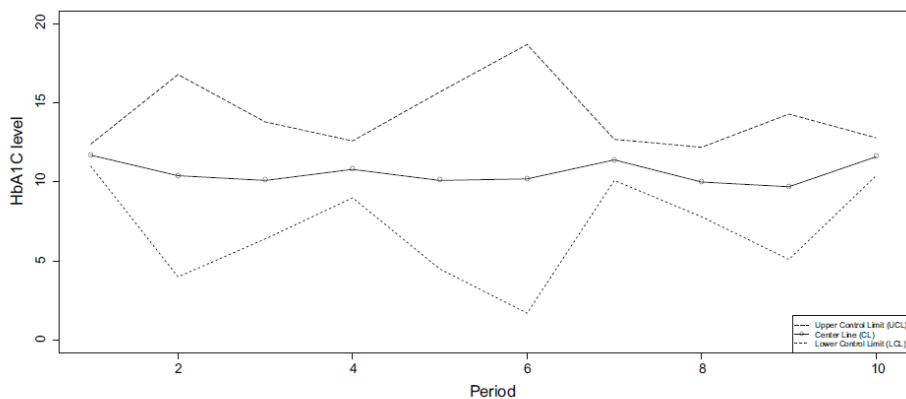


Fig. 2 The RA CC for monitoring the HbA1C level

45

45

Εφαρμογή: RA CC fall rate

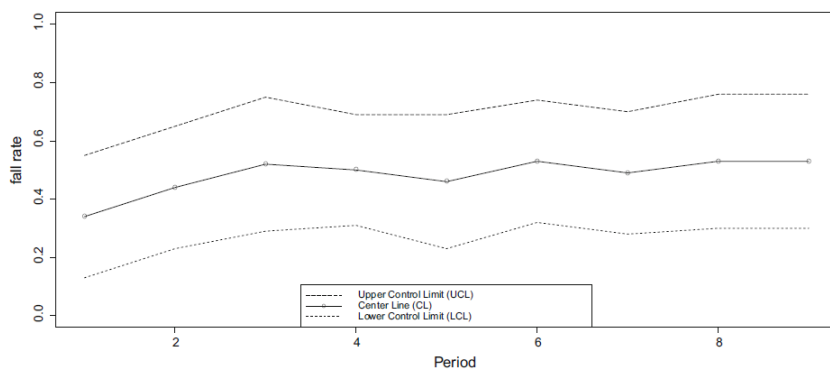


Fig. 3 The RA CC for monitoring the fall rate

46

46

Εφαρμογή: TW RA CC performance

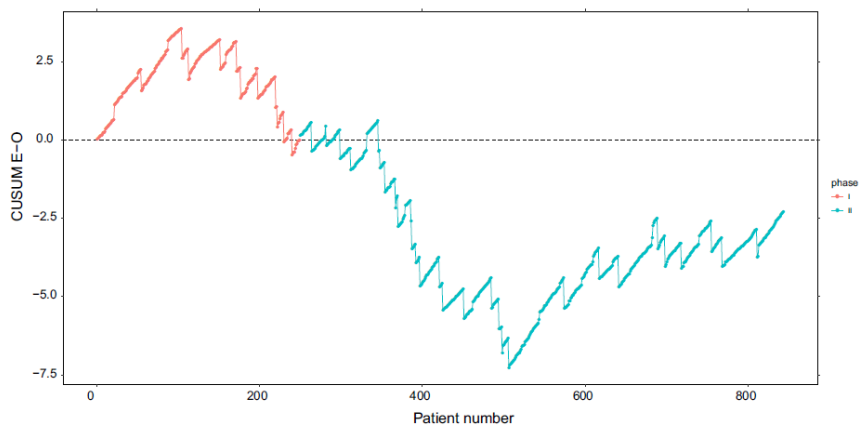


Fig. 4 The RA CC for monitoring a surgeon's performance

47

47

Table 5 Applications of RA control charts

Method	Reference	Application
RA p chart, RA-CUSUM	Cook et al. [22]	Monitoring intensive care unit outcomes
RA sequential probability ratio test	Spiegelhalter et al. [77]	Monitoring adverse outcomes
$E - O$ chart, RA p chart	Cockings et al. [21]	Monitoring intensive care unit outcomes
Shewhart \bar{X} and s charts	Hart et al. [38]	Comparison of several healthcare organizations
RA-CUSUM	Novick et al. [61]	Analysis of surgery outcomes
RA p chart	Marshall and Mohammed [52]	Monitoring mortality rates
RA sequential probability ratio test	Matheny et al. [54]	Monitoring mortality rates
RA sequential control charts	Baghurst et al. [12]	Monitoring pediatric intensive care performance
CUSUM, CRAM, VLAD and cumulative excess mortality charts	Coory et al. [24]	Monitoring the quality of hospital care
RA model	Sousa et al. [75]	Monitoring adverse cardiac and cerebrovascular events
RA-CUSUM	Chen et al. [19]	Monitoring out-of-hospital cardiac arrest patient mortality
RA control charts	Morris et al. [58]	Monitoring postoperative mortality
RA \bar{X} and R chart	Fry et al. [26]	Measuring the quality in surgical care
RA-CUSUM	Chiu et al. [20]	Monitoring the medical information in shoulder surgery study
RA-EWMA	Moran et al. [57]	Monitoring mortality
Simple control charts	Norton et al. [60]	Monitoring risk-adjusted quality indicators
Shewhart-type control charts	Maruthappu et al. [53]	Monitoring surgical procedures
RA-CUSUM	Kim et al. [45]	Analysis of the learning curve for single-incision laparoscopic anterior resection for sigmoid colon cancer
RA \bar{X} control chart	Sasikumar and Devi [68]	Monitoring HemoglobinA1C level of Type 2 diabetic patients
VLAD control chart	Patella et al. [64]	Prediction of morbidity
RA-CUSUM	Tomassini et al. [85]	Evaluation of single-surgeon learning curve
RA-CUSUM	Schrem et al. [69]	Monitoring clinical practice

Statistics in Biosciences
<https://doi.org/10.1007/s12561-019-09257-z>

48

Risk-Adjusted Control Charts για μεταβλητές

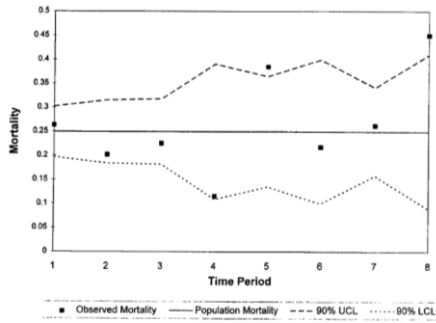


Figure 1. Unadjusted mortality rate control chart.

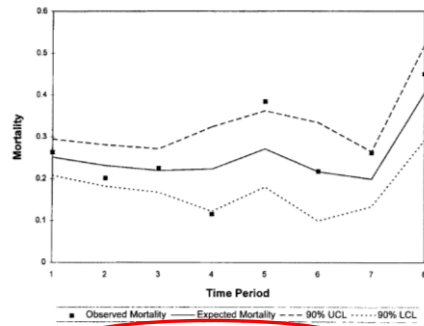


Figure 2. Risk-adjusted mortality rate control chart.

49

Στατιστικές συναρτήσεις
σάρωσης

50

50

ΣΤΑΤΙΣΤΙΚΕΣ ΣΥΝΑΡΤΗΣΕΙΣ ΣΑΡΩΣΗΣ

- Οι στατιστικές συναρτήσεις σάρωσης μελετήθηκαν με λεπτομέρεια για πρώτη φορά από τον Naus το 1965.
- Μια σημαντική πρόκληση για τις στατιστικές συναρτήσεις σάρωσης είναι να βρεθούν αναλυτικά αποτελέσματα για τις πιθανότητες να παρατηρηθεί μια συστάδα συγκεκριμένου μεγέθους (Glaz and Balakrishnan 1999; Glaz et al., 2001).
- Τα χαρακτηριστικά των αρχικών εργασιών σχετικά με τις συναρτήσεις σάρωσης ήταν δύο.
- Το πρώτο ήταν ότι έκαναν χρήση ενός σταθερού μεγέθους παραθύρου σάρωσης.
- Το δεύτερο ήταν ότι χρησιμοποιούσαν δεδομένα καταμετρήσεων (count data), όπου υπό την μηδενική υπόθεση, ο παρατηρημένος αριθμός περιπτώσεων ακολουθεί την ομοιόμορφη κατανομή.
- Στην παρακολούθηση των ασθενειών, όμως δεν ισχύει κανένα από αυτά τα χαρακτηριστικά, αφού δεν γνωρίζουμε εκ των προτέρων το μέγεθος της συστάδας και επιπλέον ο πληθυσμός σε κίνδυνο είναι γεωγραφικά ανομοιογενής.

51

ΣΤΑΤΙΣΤΙΚΕΣ ΣΥΝΑΡΤΗΣΕΙΣ ΣΑΡΩΣΗΣ

- Όταν ισχύει η μηδενική υπόθεση της ισότητας του κινδύνου εκδήλωσης της ασθένειας περιμένουμε να παρατηρήσουμε περισσότερες περιπτώσεις ασθενών σε μια πόλη σε σύγκριση με μια παρόμοιου μεγέθους περιοχή στην ύπαιθρο, ακριβώς λόγω της υψηλότερης πυκνότητας του πληθυσμού στην πόλη.
- Οι **στατιστικές συναρτήσεις σάρωσης** (scan statistics) χρησιμοποιούνται για την ανίχνευση και την αξιολόγηση συστάδων από περιπτώσεις είτε σε αμιγώς χρονικό, αμιγώς χωρικό ή χωρο-χρονικό πλαίσιο.
- Αυτό γίνεται με τη σταδιακή σάρωση ενός παραθύρου στο χρόνο ή / και το χώρο, σημειώνοντας τον αριθμό των παρατηρημένων και των αναμενόμενων περιπτώσεων μέσα στο παράθυρο σε κάθε τοποθεσία.
- Οι στατιστικές συναρτήσεις σάρωσης χρησιμοποιούν διαφορετικό μοντέλο πιθανότητας ανάλογα με τη φύση των δεδομένων.
- Στην περίπτωση που μας ενδιαφέρουν καταγραφές, όπως ο αριθμός των ατόμων που πάσχουν από άσθμα, συνήθως χρησιμοποιείται ένα μοντέλο που βασίζεται στην κατανομή Poisson.

52

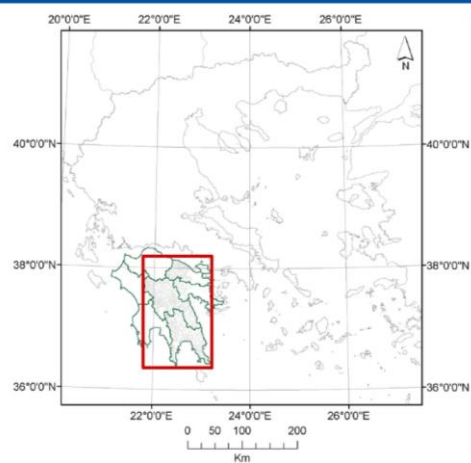
Διδιάστατες συναρτήσεις σάρωσης

53

53

ΔΙΔΙΑΣΤΑΤΕΣ ΣΥΝΑΡΤΗΣΕΙΣ ΣΑΡΩΣΗΣ

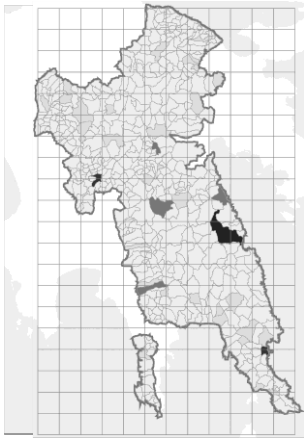
- Υποθέτουμε ότι διερευνούμε μια γεωγραφική περιοχή $R = [0, L_1] \times [0, L_2]$ για την εμφάνιση συγκεκριμένων γεγονότων (με $L_1, L_2 > 0$)..



54

ΔΙΔΙΑΣΤΑΤΕΣ ΣΥΝΑΡΤΗΣΕΙΣ ΣΑΡΩΣΗΣ

- Έχουμε $h_1=L_1/n_1$ και $h_2=L_2/n_2$ (με n_1 και n_2 θετικούς ακέραιους αριθμούς).
- Η τυχαία μεταβλητή Y_{ij} , για $1 \leq i \leq n_1$ και $1 \leq j \leq n_2$, αναπαριστά τον αριθμό των υπό μελέτη γεγονότων, τα οποία εμφανίστηκαν στην υπό-περιοχή $[(i-1)h_1, i h_1] \times [(j-1)h_2, j h_2]$.



	1	2	3	4	5	6	7	8	9	10	11	12
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	1	0	0	0	0
3	0	0	0	0	0	1	0	3	0	0	0	0
4	0	1	0	0	0	0	0	0	3	0	0	0
5	2	1	1	1	2	0	0	1	0	0	0	0
6	0	0	0	2	2	1	0	0	0	0	0	0
7	0	0	1	1	0	3	0	0	0	0	0	0
8	0	0	4	0	0	1	0	0	0	0	0	0
9	0	0	2	0	0	1	0	2	3	0	0	0
10	0	0	0	0	0	3	1	0	0	0	0	0
11	0	0	0	0	0	0	0	0	5	0	0	0
12	0	0	0	0	1	0	0	0	0	0	0	0
13	0	0	0	0	0	1	2	0	0	0	0	0
14	0	0	0	0	0	4	0	0	0	1	0	0
15	0	0	0	0	0	0	0	0	1	0	0	0
16	0	0	0	0	0	0	0	0	0	1	1	0
17	0	0	0	0	0	0	0	0	0	0	4	0
18	0	0	0	0	0	0	0	0	0	0	3	1
19	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0

55

	1	2	3	4	5	6	7	8	9	10	11	12
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	1	0	0	0	0
3	0	0	0	0	0	1	0	3	0	0	0	0
4	0	1	0	0	0	0	0	0	3	0	0	0
5	2	1	1	1	2	0	0	1	0	0	0	0
6	0	0	0	2	2	1	0	0	0	0	0	0
7	0	0	1	1	0	3	0	0	0	0	0	0
8	0	0	4	0	0	1	0	0	0	0	0	0
9	0	0	2	0	0	1	0	2	3	0	0	0
10	0	0	0	0	0	3	1	0	0	0	0	0
11	0	0	0	0	0	0	0	0	5	0	0	0
12	0	0	0	0	1	0	0	0	0	0	0	0
13	0	0	0	0	0	1	2	0	0	0	0	0
14	0	0	0	0	0	4	0	0	0	1	0	0
15	0	0	0	0	0	0	0	0	1	0	0	0
16	0	0	0	0	0	0	0	0	0	1	1	0
17	0	0	0	0	0	0	0	0	0	0	4	0
18	0	0	0	0	0	0	0	0	0	0	3	1
19	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0

56

ΔΙΔΙΑΣΤΑΤΕΣ ΣΥΝΑΡΤΗΣΕΙΣ ΣΑΡΩΣΗΣ

- Υποθέτοντας ότι τα $Y_{i,j}$, για $1 \leq i \leq n_1$ και $1 \leq j \leq n_2$, είναι ανεξάρτητες και ισόνομες τυχαίες μεταβλητές, οι οποίες ακολουθούν την Poisson κατανομή με παράμετρο λ_0 , οι Chen and Glaz (1996) πρότειναν έναν έλεγχο της μορφής

- H_0 : Τυχαία διασπορά των υπό μελέτη Γεγονότων στο Χώρο
- H_1 : Υπαρξη συστάδων στο χώρο

ο οποίος στηρίζεται στην κατανομή της στατιστικής

$$S_m = \max\{S(i_1, i_2); 1 \leq i_1 \leq n_1 - m_1 + 1, 1 \leq i_2 \leq n_2 - m_2 + 1\}$$

με

$$S(i_1, i_2) = \sum_{j=i_2}^{i_2+m_2-1} \sum_{i=i_1}^{i_1+m_1-1} Y_{i,j}$$

57

ΔΙΔΙΑΣΤΑΤΕΣ ΣΥΝΑΡΤΗΣΕΙΣ ΣΑΡΩΣΗΣ

- Ο έλεγχος μπορεί να γραφτεί ως εξής
 - H_0 : $\lambda = \lambda_0$ σε όλες τις υποπεριοχές
 - H_1 : $\lambda = \lambda_1$ σε μια τουλάχιστον υπό περιοχή
- Υπό την εναλλακτική υπόθεση ουσιαστικά υποθέτουμε ότι υπάρχει μια υπό-περιοχή, μεγέθους $[(i_1-1)h_1, (i_1 + m_1-1) h_1] \times [(i_2-1)h_2, (i_2 + m_2-1) h_2]$, στην οποία οι τυχαίες μεταβλητές $Y_{i,j}$, για $1 \leq i \leq n_1$ και $1 \leq j \leq n_2$, ακολουθούν την Poisson κατανομή με παράμετρο λ_1 , με $\lambda_1 > \lambda_0$.
- Η κατανομή της τυχαίας μεταβλητής S_m έχει μελετηθεί μόνο προσεγγιστικά ενώ έχουν δοθεί και πληθώρα φραγμάτων.
- Στις προσεγγίσεις αυτές το πλέγμα $n_1 \times n_2$ μελετάται στην ειδική περίπτωση $n_1 = n_2 = n$ κάτι το οποίο στην πράξη σπάνια ισχύει.

58

	1	2	3	4	5	6	7	8	9	10	11	12
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	1	0	0	0	0
3	0	0	0	0	0	1	0	3	0	0	0	0
4	0	1	0	0	0	0	0	0	3	0	0	0
5	2	1	1	1	2	0	0	1	0	0	0	0
6	0	0	0	2	2	1	0	0	0	0	0	0
7	0	0	1	1	0	3	0	0	0	0	0	0
8	0	0	4	0	0	1	0	0	0	0	0	0
9	0	0	2	0	0	1	0	2	3	0	0	0
10	0	0	0	0	0	3	1	0	0	0	0	0
11	0	0	0	0	0	0	0	0	5	0	0	0
12	0	0	0	0	1	0	0	0	0	0	0	0
13	0	0	0	0	0	1	2	0	0	0	0	0
14	0	0	0	0	0	4	0	0	0	1	0	0
15	0	0	0	0	0	0	0	0	1	0	0	0
16	0	0	0	0	0	0	0	0	0	1	1	0
17	0	0	0	0	0	0	0	0	0	0	4	0
18	0	0	0	0	0	0	0	0	0	0	3	1
19	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0

59

ΔΙΔΙΑΣΤΑΤΕΣ ΣΥΝΑΡΤΗΣΕΙΣ ΣΑΡΩΣΗΣ

	1	2	3	4	5	6	7	8	9	10	11	12
1						0	0	0				
2					0	0	0	1	0			
3					0	1	0	3				
4	0	1	0	0	0	0	0	2	3			
5	1	1	1	1	1	0	0	1				
6	0	0	0	1	2	1	0	0				
7	0	2	2	0	3	0						
8	0	6	0	0	1	0	0	0				
9	0	1	0	0	1	0	2	3				
10	0	0	0	0	3	1	0	0				
11				0	0	0	0	4	0			
12					1	0	0	1	0			
13					0	1	2	0	0	0		
14					0	4	0	1	0	1	0	
15						0	0	0	2	0	0	
16					0	0			0	2	1	0
17						1				0	4	
18						0				0	3	1
19					0	0						
20						0						

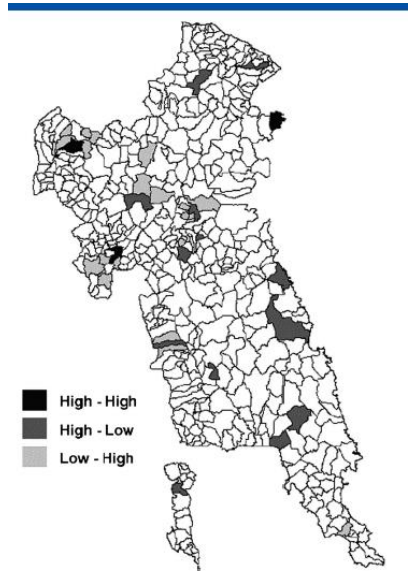
Table I. The distribution of S_m .

s	$Pr[S_m = s]$	$Pr[S_m \leq s]$
≤ 4	0.002	0.002
5	0.070	0.072
6	0.367	0.439
7	0.359	0.798
8	0.142	0.940
9	0.042	0.983
10	0.015	0.997
11	0.002	0.999
≥ 12	0.001	1.000

60

60

LOCAL INDICATORS of SPATIAL ASSOCIATION (LISA)



Global Moran's I

$$I = \frac{N}{W} \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2}$$

Local Moran's I

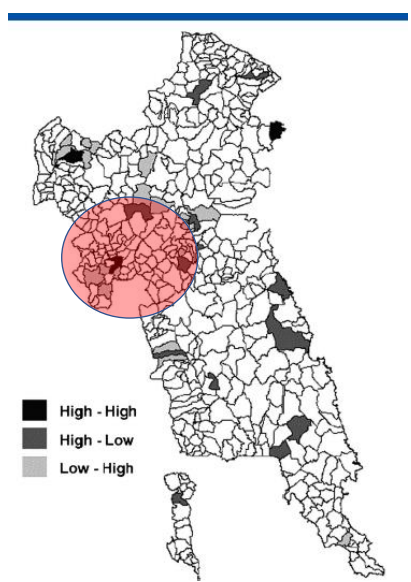
$$I_i = \frac{x_i - \bar{x}}{m_2} \sum_{j=1}^N w_{ij} (x_j - \bar{x})$$

$$m_2 = \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N}$$

Anselin, 1995 ⁶¹

61

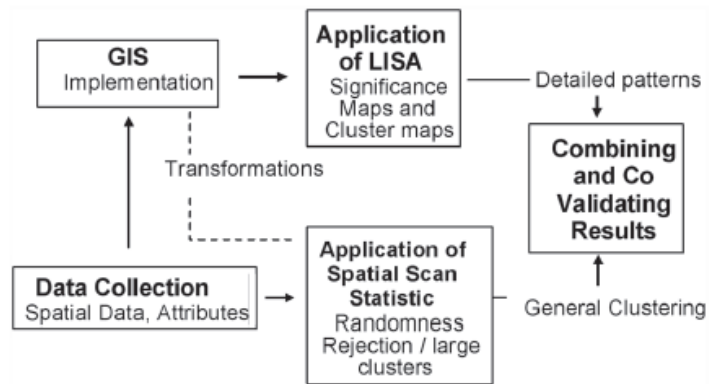
LOCAL INDICATORS of SPATIAL ASSOCIATION (LISA)



	1	2	3	4	5	6	7	8	9	10	11	12
1						0	0	0				
2					0	0	0	1	0			
3					0	1	0	3				
4	0	1	0	0	0	0	0	2	3			
5	1	1	1	1	1	0	0	1				
6	0	0	0	1	2	1	0	0				
7	0	0	2	2	0	3	0					
8	0	0	6	0	0	1	0	0	0			
9	0	0	1	0	0	1	0	2	3			
10	0	0	0	0	0	3	1	0	0			
11					0	0	0	0	0	4	0	
12					1	0	0	0	1	0		
13					0	1	2	0	0	0	0	
14					0	4	0	1	0	1	0	
15					0	0	0	2	0	0		
16					0	0		0	2	1	0	
17						1			0	4		
18					0	0			0	3	1	
19					0	0						
20						0						

62

LOCAL INDICATORS of SPATIAL ASSOCIATION (LISA) & SPATIAL SCAN STATISTIC



63

63

KULLDORF SCAN STATISTIC

Είδη στατιστικών συναρτήσεων σάρωσης του Kulldorf

- Χωρικές στατιστικές συναρτήσεις σάρωσης
- Χωρο-χρονικές στατιστικές συναρτήσεις σάρωσης
- Χρονικές στατιστικές συναρτήσεις σάρωσης

64

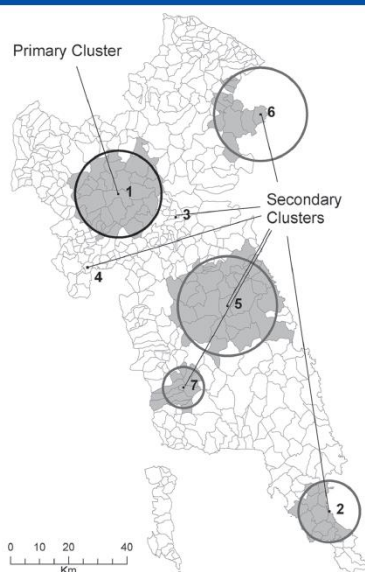
64

ΧΩΡΙΚΕΣ ΣΤΑΤΙΣΤΙΚΕΣ ΣΥΝΑΡΤΗΣΕΙΣ ΣΑΡΩΣΗΣ ΤΟΥ KULLDORF

- Η τυπική **χωρική στατιστική συνάρτηση σάρωσης** (spatial scan statistic) του Kulldorf δημιουργεί ένα κυκλικό παράθυρο στο χάρτη.
- Το παράθυρο με τη σειρά του επικεντρώνεται σε κάθε ένα από τα πολλά πιθανά σημεία πλέγματος που υπάρχουν σε όλη την υπό μελέτη περιοχή.
- Για κάθε σημείο του πλέγματος, η ακτίνα του παραθύρου διαρκώς μεταβάλλεται σε μέγεθος από το μηδέν σε κάποιο ανώτατο όριο που καθορίζεται από τον χρήστη.
- Με τον τρόπο αυτό, το κυκλικό παράθυρο είναι ευέλικτο τόσο ως προς τη θέση του, όσο και προς το μέγεθός του.
- Συνολικά, η μέθοδος αυτή δημιουργεί έναν «άπειρο» αριθμό διαφορετικών γεωγραφικών κύκλων με διαφορετικές ομάδες των γειτονικών περιοχών στο εσωτερικό τους. Κάθε κύκλος είναι μια πιθανή υποψήφια συστάδα.
- Για την εφαρμογή της ανάλυσης ο Kulldorf δημιούργησε το λογισμικό SatScan.
- Ο χρήστης με το λογισμικό SatScan ορίζει το σύνολο των σημείων πλέγματος που χρησιμοποιούνται σε ένα αρχείο πλέγματος.
- Εάν δεν έχει οριστεί αρχείο πλέγματος, τα σημεία του πλέγματος ορίζονται να είναι ίδια με τις συντεταγμένες των τοποθεσιών που ορίζονται στο αρχείο συντεταγμένων.
- Η τελευταία επιλογή διασφαλίζει ότι κάθε τοποθεσία δεδομένων είναι μια πιθανή συστάδα από μόνη της και είναι η συνιστώμενη επιλογή για τους περισσότερους τύπους αναλύσεων.

65

ΧΩΡΙΚΕΣ ΣΤΑΤΙΣΤΙΚΕΣ ΣΥΝΑΡΤΗΣΕΙΣ ΣΑΡΩΣΗΣ ΤΟΥ KULLDORF



Η τυπική **χωρική στατιστική συνάρτηση σάρωσης** (spatial scan statistic) δημιουργεί ένα κυκλικό παράθυρο στο χάρτη.

Εναλλακτικά μπορώ να κάνω χρήση ελλειπτικής συνάρτησης σάρωσης.

66

66

ΧΩΡΙΚΕΣ ΣΤΑΤΙΣΤΙΚΕΣ ΣΥΝΑΡΤΗΣΕΙΣ ΣΑΡΩΣΗΣ ΤΟΥ KULLDORF

- Ο Kulldorff (1997) όρισε τη στατιστική σάρωσης (scan statistic) ως το maximum likelihood ratio επάνω σε όλους τους πιθανούς κύκλους Z

$$S = \frac{\max_Z \{L(Z)\}}{L_0} = \max_Z \left\{ \frac{L(Z)}{L_0} \right\}$$

- Εξέτασε αρχικά το μοντέλο Bernoulli και το μοντέλο Poisson

67

ΧΩΡΙΚΕΣ ΣΤΑΤΙΣΤΙΚΕΣ ΣΥΝΑΡΤΗΣΕΙΣ ΣΑΡΩΣΗΣ ΤΟΥ KULLDORF

- Για το διακριτό μοντέλο Poisson, ο αριθμός των περιπτώσεων σε κάθε τοποθεσία ακολουθεί την κατανομή Poisson.
- Υπό την μηδενική υπόθεση, και όταν δεν υπάρχουν συμμεταβλητές, ο αναμενόμενος αριθμός των περιπτώσεων σε κάθε περιοχή είναι ανάλογη με το μέγεθος του πληθυσμού της ή τα ανθρωπο-έτη στην περιοχή αυτή.
- Τα δεδομένα που περιγράφονται από το διακριτό μοντέλο Poisson μπορούν να αναλυθούν είτε με την χρονική, είτε με την χωρική είτε την χωρο-χρονική στατιστική συνάρτηση σάρωσης.

68

ΧΩΡΙΚΕΣ ΣΤΑΤΙΣΤΙΚΕΣ ΣΥΝΑΡΤΗΣΕΙΣ ΣΑΡΩΣΗΣ ΤΟΥ KULLDORF

- Για κάθε θέση και μέγεθος του παραθύρου σάρωσης, η εναλλακτική υπόθεση είναι ότι υπάρχει αυξημένος κίνδυνος μέσα στο παράθυρο σε σχέση με αυτό που συμβαίνει εκτός του παραθύρου.
- Σύμφωνα με την υπόθεση Poisson, η συνάρτηση πιθανοφάνειας για ένα συγκεκριμένο παράθυρο είναι ανάλογη με την ποσότητα

$$\frac{L(Z)}{L_0} = \left(\frac{n_Z}{E(n_Z)} \right)^{n_Z} \left(\frac{N-n_Z}{N-E(n_Z)} \right)^{N-n_Z} I(\cdot)$$

- όπου N είναι ο συνολικός αριθμός των περιπτώσεων, n_Z είναι ο παρατηρημένος αριθμός των περιπτώσεων μέσα στο παράθυρο και $E(n_Z)$ είναι ο αναμενόμενος αριθμός περιπτώσεων εντός του παραθύρου υπό την μηδενική υπόθεση.

69

ΧΩΡΙΚΕΣ ΣΤΑΤΙΣΤΙΚΕΣ ΣΥΝΑΡΤΗΣΕΙΣ ΣΑΡΩΣΗΣ ΤΟΥ KULLDORF

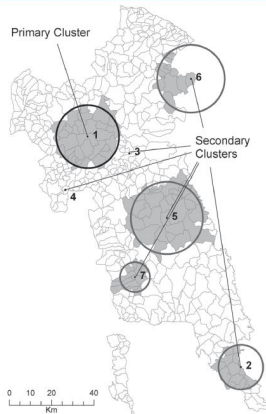
- Εφόσον η ανάλυση εξαρτάται από τον συνολικό αριθμό των περιπτώσεων που παρατηρήθηκε, η ποσότητα $N - E(n_Z)$ είναι ο αναμενόμενος αριθμός των περιπτώσεων εκτός του παραθύρου.
- Η $I(\cdot)$ είναι μια δείκτρια συνάρτηση.
- Όταν θέλουμε να σαρώσουμε μόνο συστάδες με υψηλούς ρυθμούς, η $I(\cdot)$ ισούται με 1 όταν το παράθυρο έχει περισσότερες περιπτώσεις από ότι αναμενόταν υπό την μηδενική υπόθεση και 0 σε διαφορετική περίπτωση.
- Το αντίθετο συμβαίνει όταν θέλουμε να σαρώσουμε μόνο συστάδες με χαμηλούς ρυθμούς.
- Εάν θέλουμε να σαρώσουμε για συστάδες είτε με υψηλούς, είτε με χαμηλούς ρυθμούς, τότε η $I(\cdot)$ ισούται με 1 για όλα τα παράθυρα.
- Το μοντέλο χωρο-χρονικής μετάθεσης χρησιμοποιεί την ίδια συνάρτηση πιθανοφάνειας με το μοντέλο Poisson.

$$\frac{L(Z)}{L_0} = \left(\frac{n_Z}{E(n_Z)} \right)^{n_Z} \left(\frac{N-n_Z}{N-E(n_Z)} \right)^{N-n_Z} I(\cdot)$$

70

ΧΩΡΙΚΕΣ ΣΤΑΤΙΣΤΙΚΕΣ ΣΥΝΑΡΤΗΣΕΙΣ ΣΑΡΩΣΗΣ ΤΟΥ KULLDORF

ID	Centering coordinates	Radius	Number of included neighbours	Population	Cases			Relative risk	Log-likelihood ratio	p-value
					Observed	Expected	Ratio			
1	341818, 4160740	14839.41	27 Subareas	8818	11	2.26	4.873	5.500	9.196856	0.01688
2	414225, 4052370	10541.58	13 Subareas	5856	9	1.50	6.004	6.647	9.006251	0.02036
3	361529, 4152820	0.00	1 Subarea	365	3	0.09	32.110	33.338	7.555090	0.07569

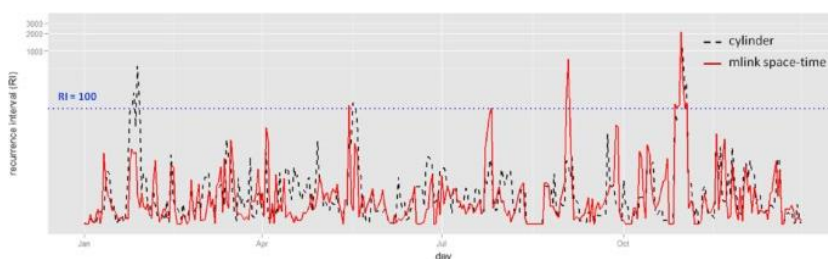


71

71

ΧΡΟΝΙΚΕΣ ΣΤΑΤΙΣΤΙΚΕΣ ΣΥΝΑΡΤΗΣΕΙΣ ΣΑΡΩΣΗΣ

- Η στατιστική συνάρτηση χρονικής σάρωσης (temporal scan statistic) χρησιμοποιεί ένα παράθυρο που κινείται σε μια διάσταση – τον χρόνο – και ορίζεται με τον ίδιο τρόπο όπως το ύψος του κυλίνδρου που χρησιμοποιείται από τη στατιστική συνάρτηση χωρο-χρονικής σάρωσης.
- Αυτό σημαίνει ότι είναι ευέλικτη τόσο στην ημερομηνία έναρξης, όσο και λήξης.
- Η μέγιστη χρονική διάρκεια ορίζεται από τον χρήστη.

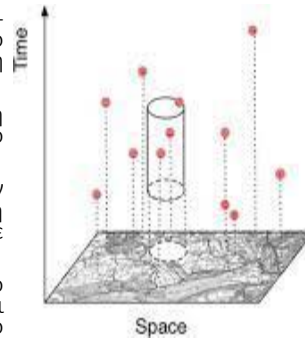


(b) cylindrical versus mlink space-time recurrence intervals

72

ΧΩΡΟ-ΧΡΟΝΙΚΕΣ ΣΤΑΤΙΣΤΙΚΕΣ ΣΥΝΑΡΤΗΣΕΙΣ ΣΑΡΩΣΗΣ

- Η **στατιστική συνάρτηση χωρο-χρονικής σάρωσης** (space-time scan statistic) ορίζεται από ένα κυλινδρικό παράθυρο με μια κυκλική (ή ελλειπτική) γεωγραφική βάση και με ύψος που αντιστοιχεί στο χρόνο.
- Η βάση ορίζεται ακριβώς όπως η στατιστική συνάρτηση χωρικής σάρωσης, ενώ το ύψος αντανακλά το χρονικό διάστημα των πιθανών συστάδων.
- Το κυλινδρικό παράθυρο στη συνέχεια μετακινείται στον χώρο και τον χρόνο έτσι ώστε, για κάθε πιθανή γεωγραφική θέση και μέγεθος, να επισκέπτεται και κάθε πιθανή χρονική περίοδο.
- Στην πραγματικότητα, παίρνουμε έναν άπειρο αριθμό επικαλυπτόμενων κυλίνδρων διαφορετικού μεγέθους και σχήματος, που από κοινού καλύπτουν ολόκληρη την υπό μελέτη περιοχή.
- Κάθε κύλινδρος αντανακλά μια πιθανή συστάδα.
- Η στατιστική συνάρτηση χωρο-χρονικής σάρωσης μπορεί να χρησιμοποιηθεί είτε για μια μόνο αναδρομική (retrospective) ανάλυση, με τη χρήση ιστορικών δεδομένων, είτε για μια χρονο-περιοδική προοπτική (prospective) επιτήρηση, στην οποία η ανάλυση επαναλαμβάνεται για παράδειγμα κάθε μέρα, εβδομάδα, μήνα ή έτος.



73

Εφαρμογές

SaTScan™

Software for the spatial, temporal, and space-time scan statistics

74

74

ΕΦΑΡΜΟΓΗ 1

Spat. Inf. Res. (2019) 27(2):187–204
<https://doi.org/10.1007/s41324-018-0224-9>



The spatial clustering of dengue disease and risk susceptibility mapping: an approach towards sustainable health management in Kharagpur city, India

Subrata Ghosh¹ · Santanu Dinda¹ · Nilanjana Das Chatterjee¹ · Kousik Das¹ · Riya Mahata¹

Abstract Dengue fever becomes one of the serious vector-borne diseases in the world, particularly in tropical countries. Mosquitoes borne disease transmission are mainly transpired by physical, environmental and socio-economic variables. Therefore, disease mapping is essential for monitoring, prediction, and prevention of dengue. Accordingly, the aims of this research are to find out the dengue outbreak, spatial pattern, understand the factors and construct a risk map to accurately predict and control of dengue. For this purpose, the Kharagpur city of West Bengal has been selected which is frequently encountered with the dengue outbreak. Epidemiological data were collected from the district health department. The spatial scan statistics of the Poisson model and Local Moran's I is used for the mapping of spatial patterns and concentration of dengue and Monte Carlo hypothesis testing was applied to test the significance of the result. Therefore, multiple

logistic regression analysis has been applied with selected parameters to produce a risk map for the entire study area. Finally, this study has proposed suitable site-specific management strategies. However, this study provides an example of disease risk estimation which will be remarkable for spatial analysis of epidemiological research as well as health planning and disease surveillance.

Keywords Dengue disease · Poisson model · SaTScan™ · Spatial clustering · Dengue risk mapping · Kharagpur city

1 Introduction

Dengue fever is one of the major vectors borne viral disease in the world, especially tropical and subtropical region [1, 2]. It is a most significant public health issue in globally and approximately, about one-third of the population of the

75

ΕΦΑΡΜΟΓΗ 1

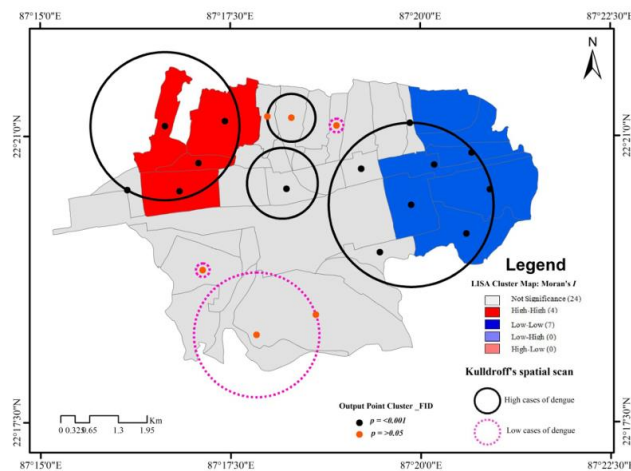


Fig. 8 SaTScan generated circled overlaid on Moran's I cluster map to show the spatial concentration and related risk using Poisson model for entire study area

76

ΕΦΑΡΜΟΓΗ 1

Table 2 SaTScan generated circles showing relative risk, log-likelihood ratio and their significance level

Cluster ID	Radius of cluster (in km)	Ward number	Relative risk	Log-likelihood ratio (LLR)	p-value
1	1.68	10, 11, 12, 13, 14	3.39	36.05	0.000
2	1.86	3, 21, 22, 23, 24, 25, 26, 28	1.14	34.56	0.000
3	0.85	18	2.28	9.90	0.003
4	0.54	8, 9	1.15	7.23	0.034
5	0.00	33	0.12	4.97	0.209
6	1.41	34, 30	0.38	4.30	0.320
7	0.00	6	0.40	1.52	0.990

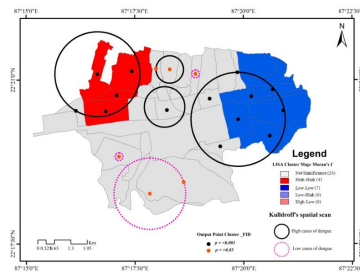


Fig. 4 SaTScan generated circles overlaid on Moran's I cluster map to show the spatial concentration and related risk using Poisson model for entire study area

77

ΕΦΑΡΜΟΓΗ 2

Understanding the Cholera Epidemic, Haiti

Renaud Piarroux, Robert Barrais, Benoît Faucher, Rachel Haus, Martine Piarroux, Jean Gaudart, Roc Magloire, and Didier Raoult

After onset of a cholera epidemic in Haiti in mid-October 2010, a team of researchers from France and Haiti implemented field investigations and built a database of daily cases to facilitate identification of communes most affected. Several models were used to identify spatiotemporal clusters, assess relative risk associated with the epidemic's spread, and investigate causes of its rapid expansion in Artibonite Department. Spatiotemporal analyses highlighted 5 significant clusters ($p < 0.001$): 1 near Mirebalais (October 16–19) next to a United Nations camp with deficient sanitation, 1 along the Artibonite River (October 20–28), and 3 caused by the centrifugal epidemic spread during November. The regression model indicated that cholera more severely affected communes in the coastal plain (risk ratio 4.91) along the Artibonite River downstream of Mirebalais (risk ratio 4.60). Our findings strongly suggest that contamination of the Artibonite and 1 of its tributaries downstream from a military camp triggered the epidemic.

78

ΕΦΑΡΜΟΓΗ 2

Spatiotemporal Modeling

By using SaTScan (Kulldorf), several spatiotemporal clusters were identified (Figure 3): Mirebalais, October 16–19 ($p < 0.001$), and in the Artibonite delta, October 20–28 ($p < 0.001$). Overlapping staggered clusters occurred in the North-West (November 11–29; $p < 0.001$); Port-au-Prince area (November 14–30; $p < 0.001$); and North (November 21–30; $p < 0.001$).

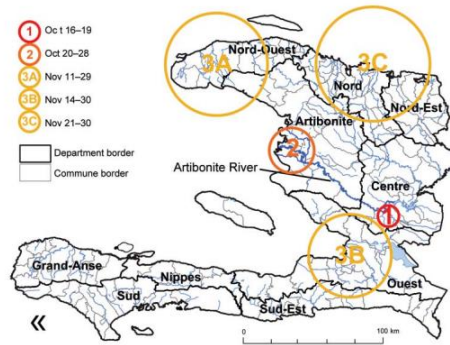


Figure 3. Spatiotemporal clusters of cholera cases, Haiti (results of SaTScan [Kulldorf, Cambridge, UK] analysis). The first cluster covered 1 commune, Mirebalais, October 16–19; the second cluster covered a few communes in or near the Artibonite delta during October 20–28; the next 3 clusters appeared in the North-West Department (A) during November 11–29, in the West Department (B) during November 14–30, and in the North and North-East Departments (C) during November 21–30. Other departments were affected later.

79

ΕΦΑΡΜΟΓΗ 3

The Egyptian Journal of Remote Sensing and Space Sciences 22 (2019) 297–304



Geographic information system based spatio-temporal dengue fever cluster analysis and mapping²⁹

Shuchi Mala^{a,1}, Mahesh Kumar Jat

^aDepartment of Civil Engineering, Malaviya National Institute of Technology Jaipur, Jawaharlal Nehru Marg, Malviya Nagar, Jaipur, Rajasthan 302017, India

ABSTRACT

Dengue Fever (DF) is one of the world's most rapidly spreading mosquito-borne viral diseases leading to increased risk of premature mortality and requires critical amount of health care cost. The present work is aimed at investigating the spatial and temporal distribution of DF in Civil Lines Zone of Delhi, capital city of India using Kulldorf's space-time permutation method over a period of 2010–2012. GIS technique such as kernel density and directional distribution have been used to determine the spatial diffusion patterns of DF incidences. Also, Pearson correlation matrix has been used to determine spatial relationship of environmental factor such as wind speed with DF incidences. Highest kernel (intection) density has been found for the year 2010 ranging from 7.62 to 8.89 per km² approximately. The most affected age group was found to be 16–30 years and out of the total confirmed dengue cases 63% were males and 37% were females. The study is successful in detecting the spatio-temporal clusters having higher risk of DF outbreaks in Delhi and negative association of wind speed with DF incidences has been observed. Thus, confirming environmental factors in addition to other socio-economic, demographic and physiographic factors are responsible for the change in the trend of DF incidences over the study period.

© 2019 National Authority for Remote Sensing and Space Sciences. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

80

ΕΦΑΡΜΟΓΗ 3

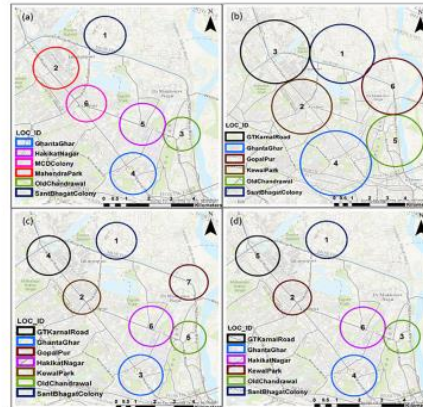


Fig. 5. Potential clusters of DF identified by SaTScan and the numbering of clusters is done on the basis of statistically significance in decreasing order (cluster numbered "1" is the most likely cluster which has the lowest p-value). These clusters are overlaid on the graphical representation of the study area. Numbers indicate clusters identified in Table 2, 4 and 5. (a) A Map Showing Space-Time Clusters; (b) A Map Showing Covariates (age and gender) Adjusted Space-Time Clusters; (c) A Map Showing Covariate (Age) Adjusted Space-Time Clusters; (d) A Map Showing Covariate (Gender) Adjusted Space-Time Clusters.

81

ΕΦΑΡΜΟΓΗ 4

Hallu et al. BMC Infectious Diseases (2020) 20:751
<https://doi.org/10.1186/s12879-020-05456-y>

BMC Infectious Diseases

RESEARCH ARTICLE

Open Access

Spatial patterns and associated factors of HIV Seropositivity among adults in Ethiopia from EDHS 2016: a spatial and multilevel analysis

Bayuh Asmamaw Hallu^{1*}, Fentaw Tadesse¹, Getahun Gebre Bogale², Asressie Molla³, Birhan Asmame Mihretu³ and Joseph Beyene⁴

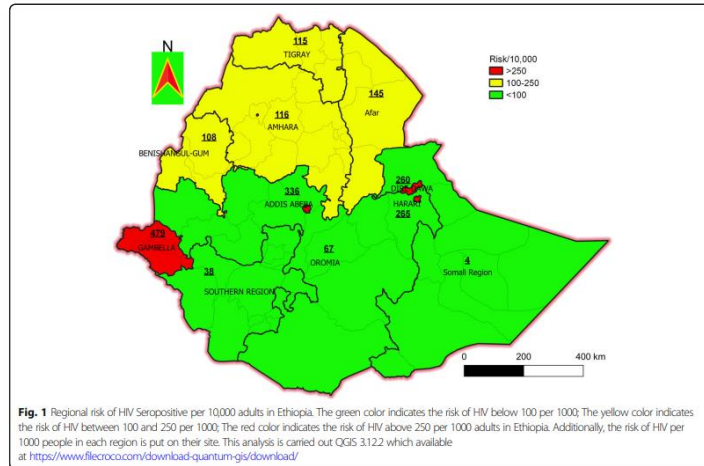


Kulldorff's scan statistic

Kulldorff's scan statistic is a tool that uses spatial scan statistics for the identification and assessment of statistically important spatial cluster risk factors for a particular disease. The final confirmatory spatial analysis was performed using SaTScan with QGIS analysis tools. The SaTScan may distinguish particular locations were higher or lower rate of spatial aggregate. Its output presents the hotspot areas in circular windows, indicating areas of windows are higher than expected distributions compared to the areas outside of the cluster windows [21–24].

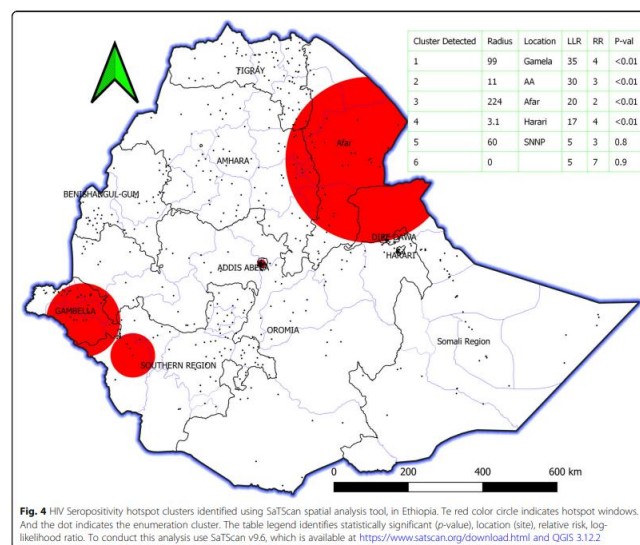
82

ΕΦΑΡΜΟΓΗ 4



83

ΕΦΑΡΜΟΓΗ 4



84

ΕΦΑΡΜΟΓΗ 5



Rapid surveillance of COVID-19 by timely detection of geographically robust, alive and emerging hotspots using Particle Swarm Optimizer

Ankita Wadhwa^{*}, Manish Kumar Thakur

Department of Computer Science Engineering and IT, Joppe Institute of Information Technology, A-10 Sector 62, Noida, UP, 201309, India

ABSTRACT

A novel virus, called severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has rapidly become a pandemic called Coronavirus disease 2019 (COVID-19). According to the World Health Organization, COVID-19 was first detected in Wuhan city in December 2019 and has affected 216 countries with 9473214 confirmed cases and 484249 deaths globally as on June 26th, 2020. Also, this outbreak continues to grow in many countries like the United States of America (U.S.), Brazil, India, and Russia. To ensure rapid surveillance and better decision-making by government authorities in different countries, it is vital to identify alive and emerging hotspots within a country promptly. State-of-the-art methods based on space-time scan statistics (like SaTScan) are not geographically robust. Also, due to the enumeration of many Spatio-temporal cylinders, the computation cost of Spatio-temporal SaTScan (ST-SaTScan) is very high. In the applications like COVID-19 where we need to detect the emerging hotspots daily as soon as the new count of cases gets updated, ST-SaTScan seems inefficient. Therefore, this paper proposes a Particle Swarm Optimizer-based scheme to timely detect geographically robust, alive, and emerging COVID-19 hotspots in a country. Timely detection can help government officials design better control strategies like increasing testing in hotspots, imposing stricter containment rules, or setting up temporary hospital beds. Performance of ST-SaTScan and proposed scheme have been analyzed for four worst-hit U.S. states for the incubation period of 14 days between June 11th, 2020, and June 24th, 2020. Results indicate that the proposed scheme detects hotspots of a higher likelihood ratio (a measure to indicate the significance of hotspot) than ST-SaTScan in significantly less time. We also applied the proposed scheme to detect the emerging COVID-19 hotspots in all states of the U.S. During the study period, the proposed scheme has detected 104 emerging COVID-19 hotspots.

85

ΕΦΑΡΜΟΓΗ 5

Algorithm 1: ST-SaTScan for detecting COVID-19 Hotspots

Input:

- A matrix of cases named 'cases' where, cases(i, j) represents the number of COVID-19 cases in i^{th} county (total C counties in the study area) on j^{th} day (total T days in study duration)
- Loc[C] and Pop[C] containing (latitude, longitude) pairs and population of each county respectively.
- Max_rad representing upper bound on radius of the hotspot.
- The number of Monte Carlo simulations 'm' and p-value threshold α_p .

Output: Non-overlapping alive and emerging COVID-19 hotspots with p-value $\leq \alpha_p$

Algorithm:

Step 1: Enumerate cylinders and calculate their LLR

```
for each county i do
  for each county j except i do
    rad=distance( Loc(i), Loc(j))
    if rad>max_rad
      break;
    end if
    for h =2 : T/2
      calculate LLR using (location(i), rad, h, Loc[], Pop[], cases[[]]) using eqn 1
      cylinders = [cylinders ; location(i) rad h LLR];
    end for
  end for
end for
among overlapping cylinders keep the one with highest LLR
```

Step 2: Generate distribution of LLR under H_0

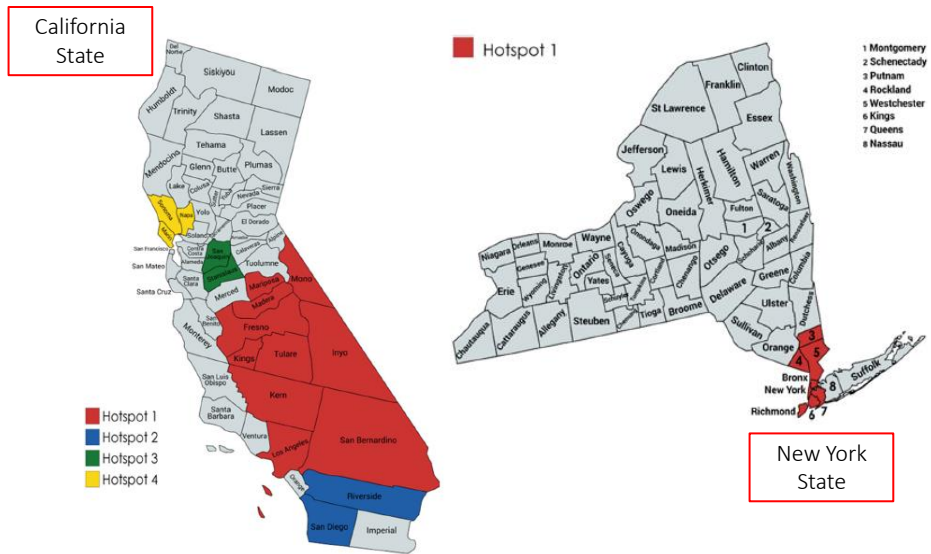
```
maxLLRs=[];
for each Monte Carlo simulation i= 1 to m
  cases' = Poisson(Distribution);
  Compute maximum LLR' cylinder using step 1
  append (maxLLRs, LLR');
end
maxLLRs=sort(maxLLRs, desc);
```

Step 3: Test cylinders generated in step 1 for statistical significance

```
for each cylinder H in cylinders
  pos= the position of LLR of cylinder in the maxLLRs
  p-value0 = pos/m+1;
  if (p-value0) <  $\alpha_p$ 
    append (Hotspots_significant, H);
  end
```

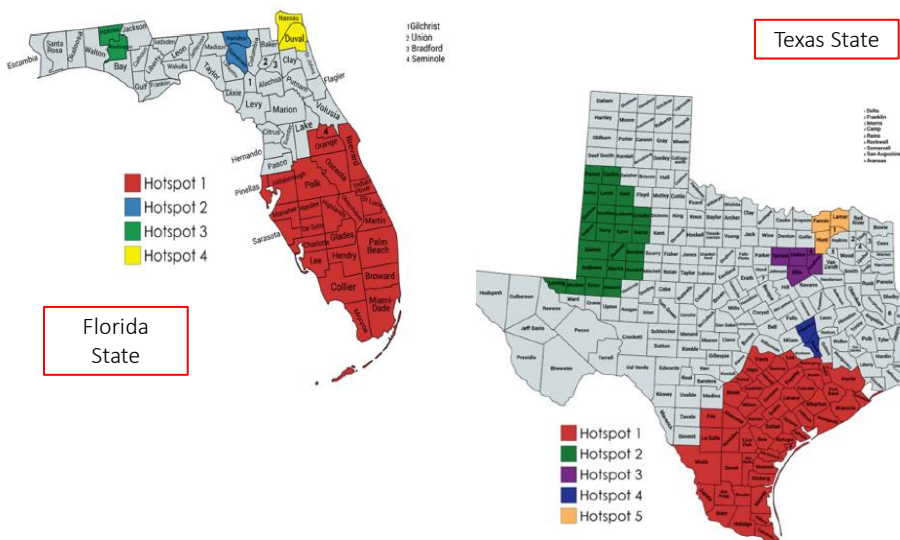
86

ΕΦΑΡΜΟΓΗ 5



87

ΕΦΑΡΜΟΓΗ 5



88

Τέλος μαθήματος

Ιδιαίτερη προσοχή πρέπει να δοθεί στο γεγονός ότι οι διαφάνειες αποτελούν τον οδηγό για τη διεξαγωγή του μαθήματος περιγράφοντας τις γενικές κατευθύνσεις κάθε ενότητας. Για το λόγο αυτό οι διαφάνειες δεν είναι πλήρεις, ενώ η συμπλήρωση, ο εμπλουτισμός τους με επιπλέον ασκήσεις, ερωτήματα ακόμα και διορθώσεις γίνεται κατά τη διάρκεια των μαθημάτων. Συνέπεια αυτού είναι το υλικό αυτό να μην μπορεί να θεωρηθεί, χωρίς την ενεργή συμμετοχή του φοιτητή στη διάρκεια των μαθημάτων, ότι καλύπτει πλήρως την ύλη του μαθήματος.