Πανεπιστήμιο Αιγαίου

Χωρική Ανάλυση

Ενότητα 5β: Έλεγχος τυχαίας χωρικής κατανομής σημειακού προτύπου

Κυριακίδης Φαίδων

Τμήμα Γεωγραφίας

Χρηματοδότηση

Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.

Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Αιγαίου**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.

Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Έλεγχος Υποθέσεως
Τυχαίου Χωρικού Σημειακού Προτύπου

Φαίδων Κυριακίδης

Καθηγητής

phkyriakidis@geo.aegean.gr

ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ

Λόφος Πανεπιστημίου, 81100 Μυτιλήνη

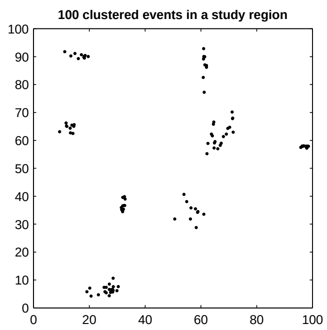
Χωρική Ανάλυση

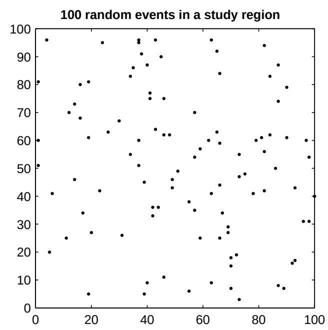
ΤΜΗΜΑ ΓΕΩΓΡΑΦΙΑΣ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

ΓΕΩΓΡΑΦΙΑ ΚΑΙ ΕΦΑΡΜΟΣΜΕΝΗ ΓΕΩΠΛΗΡΟΦΟΡΙΚΗ

Χωρικό Σημειακό Πρότυπο

Σύνολο σημειακών θέσεων με καταγεγραμμένα συμβάντα, π.χ., θέσεις δέντρων, θέσεις εκδήλωσης νόσου ή εγκληματικής ενέργειας





Για το μάθημα αυτό:

- βασική παραδοχή: όλες οι θέσεις συμβάντων έχουν καταγραφεί
- αντικείμενο ανάλυσης: χωρική κατανομή θέσεων συμβάντων κι όχι άλλων ιδιοτήτων που έχουν τυχόν καταμετρηθεί στις θέσεις αυτές

Στόχος μαθήματος

Έλεγχος υποθέσεως τυχαίας χωρικής κατανομής συμβάντων σημειακού προτύπου: Επισκόπηση σχετικών εννοιών και επι μέρους βημάτων που ακολουθούνται στη διαδικασία ελέγχου

Έλεγχος Υποθέσεως Τυχαίου Σημειακού Προτύπου

Προεπισκόπηση

1. μηδενική υπόθεση: Επιλογή μιας τυχαίας (στοχαστικής) διεργασίας ως το γενεσιουργό μηχανισμό για ένα παρατηρούμενο σημειακό πρότυπο.
   Μια στοχαστική διεργασία είναι μια μαθηματική εξίσωση, ή ένα σύνολο εξισώσεων ή κανόνων, γενικά ένας μηχανισμός, που "γεννά" σημειακά πρότυπα στο χώρο.
   Στην περίπτωση του μαθήματος αυτού, η μηδενική υπόθεση είναι η παραδοχή ότι τα συμβάντα κατανείμονται εντελώς τυχαία στην περιοχή μελέτης, ή αλλιώς ότι προέρχονται από μια εντελώς τυχαία χωρική διεργασία
2. στατιστική δειγματοληψία: κατασκευή (προσομοίωση) πολλαπλών εναλλακτικών σημειακών προτύπων (με τυχαίο τρόπο) από τη στοχαστική διεργασία. Τα σημειακά πρότυπα αυτά ονομάζονται δείγματα ή πραγματοποιήσεις της στοχαστικής διεργασίας
3. επιλογή στατιστικού, π.χ., της εμπειρικής συνάρτησης $\hat{G}(d)$: μέσω του οποίου το παρατηρούμενο σημειακό πρότυπο θα συγκρίθει με τα πολλαπλά πρότυπα που προσομοιώθηκαν. Το ιστόγραμμα των τιμών του στατιστικού που υπολογίζονται από τα προσομοιωμένα σημειακά πρότυπα είναι μια προσέγγιση της **δειγματοληπτικής κατανομής** του στατιστικού με βάση τη μηδενική υπόθεση
4. έλεγχος υποθέσεως: Απόρριψη ή όχι της μηδενικής υπόθεσης με βάση το βαθμό στον οποίο διαφέρει το παρατηρούμενο στατιστικό από τη δειγματοληπτική κατανομή του εν λόγω στατιστικού

Βασικές Έννοιες

Σημειακά συμβάντα

Σύνολο n θέσεων συμβάντων που καταγράφηκαν σε μια περιοχή μελέτης:

$$\{\mathbf{u}_i, i = 1, \ldots, n\}, \quad \mathbf{u}_i \in D$$

$\mathbf{u}_i$ = διάνυσμα συντεταγμένων της θέσης του συμβάντος i, π.χ., σε 2 διαστάσεις $\mathbf{u}_i = (x_i \ y_i)$,
n = αριθμός συμβάντων, $\in$ = ανήκει, D = περιοχή μελέτης

Γενικευμένη ή μέση ένταση λ

Εκτιμάται ως: $\hat{\lambda} = \dfrac{n}{|D|}$, όπου $|D|$ = είναι το μέτρο (εμβαδόν) της υπο μελέτης περιοχής D.

Με άλλα λόγια, στις δύο διαστάσεις, $\hat{\lambda}$ είναι η ένταση στη μοναδιαία επιφάνεια

Τοπική ένταση $\lambda(s)$

Αριθμός συμβάντων (συχνότητα) σε μια οποιοδήποτε υποδιαίρεση του χώρου (περιοχή/ζώνη/φατνίο) s δια το εμβαδόν $|s|$ του υπο-χώρου αυτού

Εντελώς Τυχαία Χωρική Σημειακή Διεργασία (ΕΤΧΣΔ)

Μη μαθηματικός ορισμός

Μια στοχαστική χωρική διεργασία που θεωρείται ως ο μηχανισμός δημιουργίας πολλαπλών χωρικών σημειακών προτύπων, με τα εξής χαρακτηριστικά:

- η ένταση λ (αναμενόμενος αριθμός συμβάντων στη μοναδιαία επιφάνεια) είναι σταθερή στην περιοχή μελέτης $D \Rightarrow$ απουσία εξωγενών ή περιβαλλοντικών επιδράσεων
- η θέση εμφάνισης ενός συμβάντος είναι ανεξάρτητη από την εμφάνιση ενός άλλου συμβάντος $\Rightarrow$ απουσία αλληλεπιδράσεων μεταξύ συμβάντων

Δύο μορφές εντελώς τυχαίας ΧΣΔ

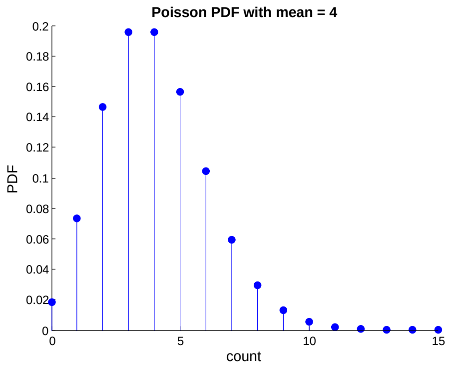
- Διωνυμική (Binomial) σημειακή διεργασία: Αν προσομοιώσει κανείς M τυχαία σημειακά πρότυπα στην περιοχή D, κάθε πρότυπο θα περιέχει τον ίδιο αριθμό n τυχαίων συμβάντων
- Poisson σημειακή διεργασία: Ο αριθμός των συμβάντων n_m που εμφανίζονται τυχαία στο πρότυπο m είναι πραγματοποίηση μιας τυχαίας μεταβλητής N με κατανομή Poisson. Στην περίπτωση M προσομοιωμένων σημειακών προτύπων, υπάρχουν M αριθμοί συμβάντων $\{n_m, m = 1, \ldots, M\}$

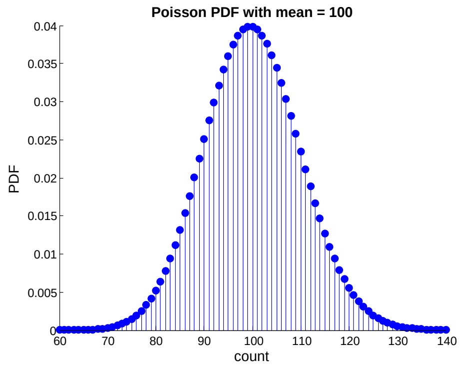
Κατανομή Poisson

Συνάρτηση πυκνότητας πιθανότητας

Μια διακριτή τυχαία μεταβλητή (ΤΜ) N ακολουθεί την κατανομή Poisson, με αναμενόμενη τιμή $\mu = \mathbb{E}\{N\}$, όταν η συνάρτηση πυκνότητας πιθανότητας της ΤΜ N δίνεται ως:

$$f_N(n) = \frac{\mu^n}{n!} \exp(-\mu)$$





- πιθανότητα εμφάνισης αριθμού συμβάντων n σε ένα χώρο με μέτρο (μήκος/εμβαδόν/όγκο) $|D|$, όπου $\mu = \lambda|D|$ και λ είναι η μέση ένταση στο μοναδιαίο εμβαδόν
- η αναμενόμενη τιμή μ και η διακύμανση σ^2 της ΤΜ N είναι ίσες: $\mu = \sigma^2$

Προσομοίωση Από μια Σημειακή Διεργασία Poisson

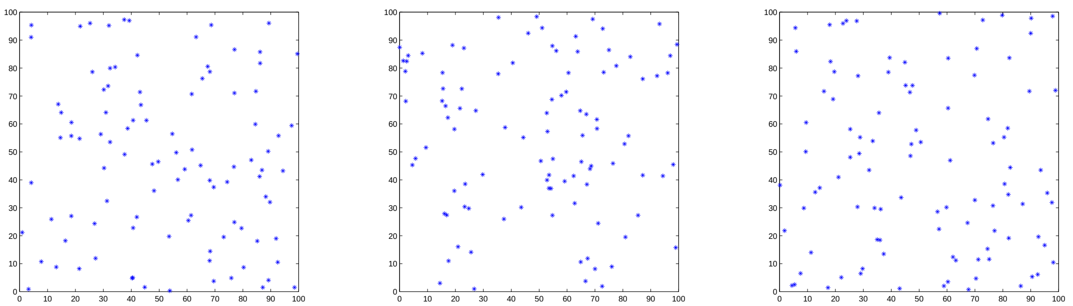
Διαδικασία προσομοίωσης

Για την κατασκευή M προσομοιωμένων σημειακών προτύπων, όπου n_m είναι ο αριθμός των συμβάντων του προτύπου m:

1. προσομοίωση M αριθμών $\{n_m, m = 1, \ldots, M\}$ από μια ΤΜ με κατανομή Poisson με αναμενόμενη τιμή $\lambda|D|$. Οι M αριθμοί αντιστοιχούν στους αριθμούς συμβάντων των M σημειακών προτύπων που θα προσομοιωθούν
2. προσομοίωση n_m τυχαίων θέσεων συμβάντων σε ένα παραλληλόγραμμο που περικλείει (περιγράφει) την περιοχή μελέτης D. Αυτό επιτυγχάνεται με την προσομοίωση n_m ζευγών τιμών συντεταγμένων από δύο ανεξάρτητες ΤΜ με ομοιόμορφη κατανομή στα διαστήματα $[x_{min}, x_{max}]$ και $[y_{min}, y_{max}]$ που αντιστοιχούν στις δύο πλευρές του παραλληλογράμμου
3. απόρριψη προσομοιωμένων θέσεων που δεν εμπίπτουν στην περιοχή D, και επανάληψη των βημάτων 1 και 2 μέχρι να προσομοιωθούν n_m θέσεις μέσα στην περιοχή D
4. επανάληψη των βημάτων 2 και 3 με ένα άλλο προσομοιωμένο αριθμό συμβάντων, $n_{m'}$, για την προσομοίωση ενός άλλου, m', σημειακού προτύπου

Παραδείγματα Εντελώς Τυχαίων Σημειακών Προτύπων

3 πραγματοποιήσεις (προσομοιωμένα χωρικά σημειακά πρότυπα) από μια εντελώς τυχαία διεργασία με $n = 100$ συμβάντα σε μια τετραγωνική περιοχή 100×100:

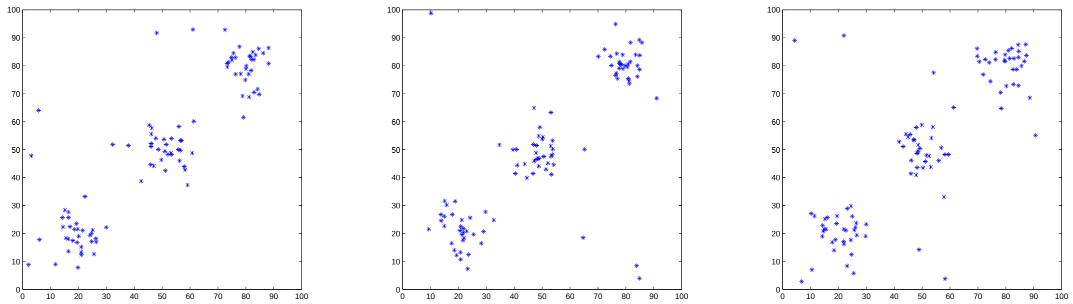


Χαρακτηριστικά

- δεν αναμένονται συγκεντρώσεις συμβάντων στο χώρο, αλλά είναι πιθανή η κατά τόπους τυχαία συγκέντρωσή τους
- όταν εμφανίζονται θύλακες με συγκεντρωμένα συμβάντα, ο αριθμός, η θέση και το σχήμα/έκταση των θυλάκων αυτών είναι τυχαία και αλλάζουν από πρότυπο σε πρότυπο

Παραδείγματα Μη Εντελώς Τυχαίων Σημ. Προτύπων (1)

3 πραγματοποιήσεις (προσομοιωμένα χωρικά σημειακά πρότυπα) από μια μη εντελώς τυχαία διεργασία με $n = 100$ συμβάντα σε μια τετραγωνική περιοχή 100×100:

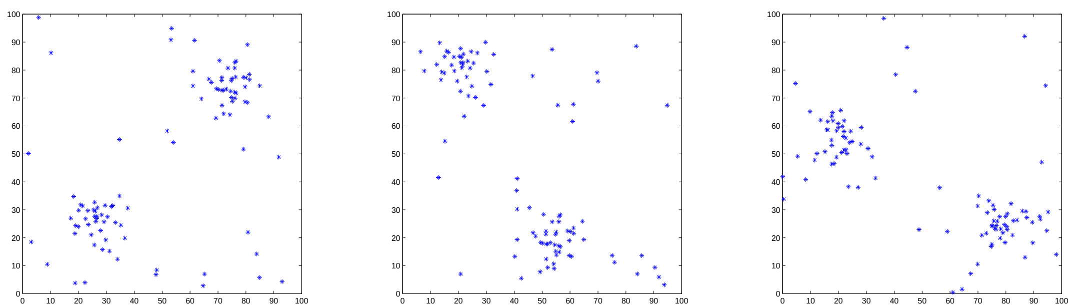


Χαρακτηριστικά

- στην περίπτωση αυτή, αναμένονται συγκεντρώσεις συμβάντων στο χώρο λόγω της επίδρασης εξωγενών (περιβαλλοντικών) παραγόντων που αυξάνουν την πιθανότητα εμφάνισης συμβάντων σε διάφορες υπο-περιοχές
- ο αριθμός, η θέση και το σχήμα/έκταση των θυλάκων αυτών είναι δεν αλλάζουν σημαντικά από πρότυπο σε πρότυπο

Παραδείγματα Μη Εντελώς Τυχαίων Σημ. Προτύπων (2)

3 πραγματοποιήσεις (προσομοιωμένα χωρικά σημειακά πρότυπα) από μια μη εντελώς τυχαία διεργασία με $n = 100$ συμβάντα σε μια τετραγωνική περιοχή 100×100:



Χαρακτηριστικά

- στην περίπτωση αυτή, εμφανίζονται συγκεντρώσεις συμβάντων στο χώρο λόγω της αλληλεπίδρασης μεταξύ τους κι όχι λόγω της επίδρασης εξωγενών (περιβαλλοντικών) παραγόντων
- η θέση των θυλάκων συγκέντρωσης συμβάντων αλλάζει σημαντικά από πρότυπο σε πρότυπο κι όχι τόσο το σχήμα και ο αριθμός των θυλάκων

Περιγραφή Προσομοιωμένου ή Παρατηρούμενου Προτύπου

Ανακεφαλαίωση

- μέχρι στιγμής, είδαμε με ποιόν τρόπο προσομοιώνονται χωρικά σημειακά πρότυπα από μια εντελώς τυχαία σημειακή διεργασία, είτε αυτή λέγεται Διωνυμική διεργασία είτε διεργασία Poisson
- μια στοχαστική χωρική διεργασία οδηγεί σε πολλαπλά σημειακά πρότυπα, κι ο στόχος μας τώρα είναι τα συγκρίνουμε (με βάση κάποια ποσοτικά στατιστικά) το παρατηρούμενο σημειακό πρότυπο με τα M προσομοιωμένα σημειακά πρότυπα

Τρόποι περιγραφής χωρικής κατανομής ενός σημειακού προτύπου

- ανάλυση έντασης: καταμέτρηση έντασης συμβάντων σε φατνία, ή εκτίμηση επιφάνειας έντασης με χρήση πυρήνα
- ανάλυση αλληλεπίδρασης: μελέτη της κατανομής των αποστάσεων μεταξύ πλησιεστέρων συμβάντων, ή των αποστάσεων μεταξύ συμβάντων γενικότερα (εξαιρουμένων των μηδενικών αποστάσεων)

Δειγματοληπτική Κατανομή του Στατιστικού $\overline{d_{min}}$ (1)

Μεταβλητή περιγραφής σημειακού προτύπου

Απόσταση μεταξύ πλησιέστερων συμβάντων. Για ένα πρότυπο (δείγμα) με n συμβάντα, υπάρχουν n τέτοιες αποστάσεις $\{d_i^{min}, i = 1, \ldots, n\}$, όπου d_i^{min} είναι η απόσταση από το συμβάν με διάνυσμα συντεταγμένων $\mathbf{u}_i$ στο πλησιέστερο συμβάν

Στατιστικό δείγματος

Μέση απόσταση μεταξύ πλησιέστερων συμβάντων: $\overline{d_{min}} = \frac{1}{n}\sum_{i=1}^{n} d_i^{min}$

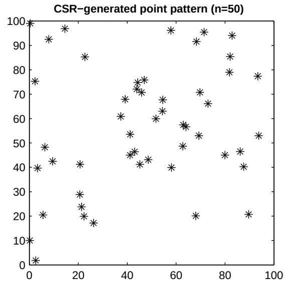
Κατασκευή της δειγματοληπτικής κατανομής του στατιστικού

μέσω προσομοίωσης από τη μηδενική υπόθεση:

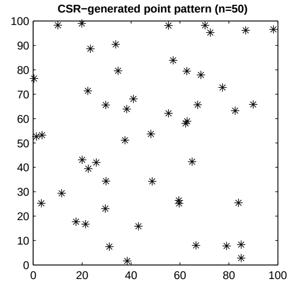
1. προσομοίωση ενός δείγματος, δηλαδή ενός χωρικού σημειακού προτύπου, από την εντελώς τυχαία χωρική σημειακή διεργασία με μέση ένταση λ
2. υπολογισμός της τιμής του στατιστικού $\overline{d_{min}}$ από το προσομοιωμένο πρότυπο
3. επανάληψη των βημάτων (1) και (2), πολλές, π.χ. $M = 1000$, φορές για τον υπολογισμό M προσομοιωμένων τιμών στατιστικού $\overline{d_{min}}$
4. το ιστόγραμμα των M προσομοιωμένων τιμών $\overline{d_{min}}$ αποτελεί (κατά προσέγγιση) τη δειγματοληπτική κατανομή του εν λόγω στατιστικού **με βάση τη μηδενική υπόθεση**

Δειγματοληπτική Κατανομή του Στατιστικού $\overline{d_{min}}$ (2)

Δύο πραγματοποιήσεις (δείγματα) από μια Διωνυμική τυχαία διαδικασία με $n = 50$

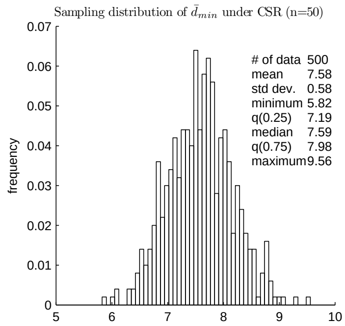


$\overline{d_{min}} = 7.13$



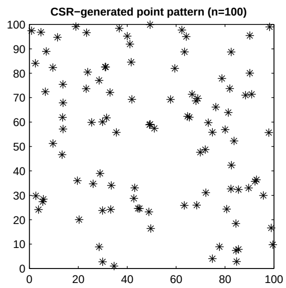
$\overline{d_{min}} = 8.13$

Δειγματοληπτική κατανομή ή ιστόγραμμα τιμών $\overline{d_{min}}$ που υπολογίστηκαν από $M = 500$ προσομοιωμένα εντελώς τυχαία σημειακά πρότυπα, με $n = 50$ συμβάντα στο καθένα

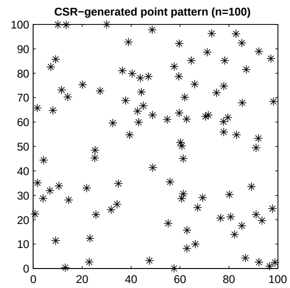


Δειγματοληπτική Κατανομή του Στατιστικού $\overline{d_{min}}$ (3)

Δύο πραγματοποιήσεις (δείγματα) από μια Διωνυμική τυχαία διαδικασία με $n = 100$

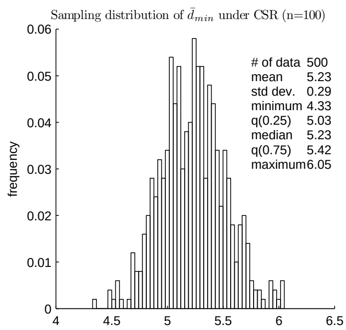


$\overline{d_{min}} = 4.88$



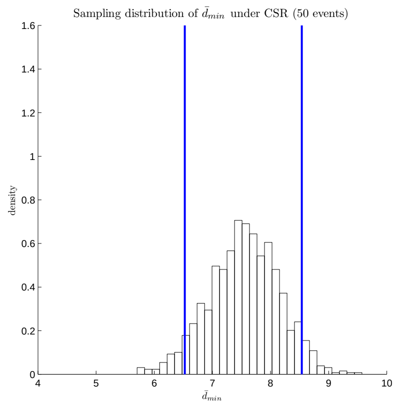
$\overline{d_{min}} = 5.67$

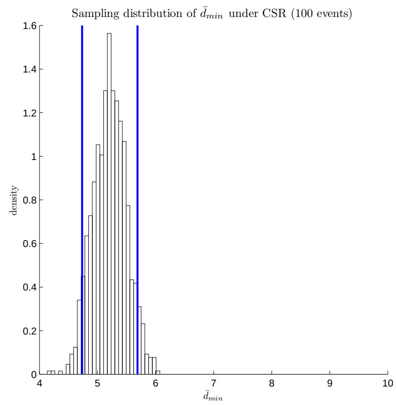
Δειγματοληπτική κατανομή ή ιστόγραμμα τιμών $\overline{d_{min}}$ που υπολογίστηκαν από 500 προσομοιωμένα εντελώς τυχαία σημειακά πρότυπα, με $n = 100$ συμβάντα στο καθένα



Διάστημα Εμπιστοσύνης για το Στατιστικό $\overline{d_{min}}$

Δειγματοληπτικές κατανομές του στατιστικού $\overline{d_{min}}$ με βάση 1000 εντελώς τυχαία προσομοιωμένα σημειακά πρότυπα με αναμενόμενο (μέσο) αριθμό συμβάντων $n = 50$ (αριστερά) και $n = 100$ (δεξιά) σε μια περιοχή 100×100



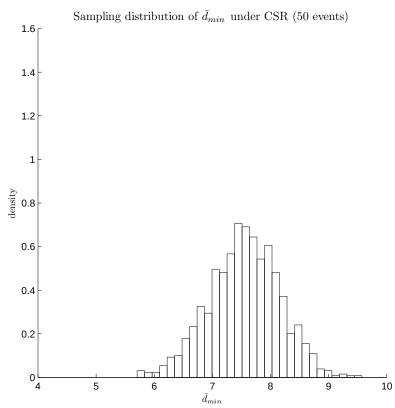


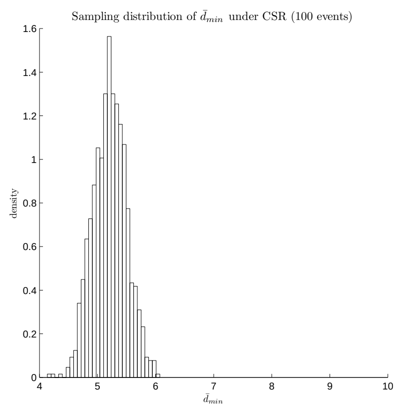
Διάστημα εμπιστοσύνης (ΔΕ)

Για επίπεδο σημαντικότητας α, εδώ $\alpha = 0.1$, τα ποσοστημόρια που αντιστοιχούν στις αθροιστικές πιθανότητες $\alpha/2$ και $1 - (\alpha/2)$, οριοθετούν το $1 - \alpha$ ΔΕ για το επίπεδο αυτό. Στο παράδειγμα αριστερά, οι τιμές 6.52 και 8.54 είναι τα 0.05 και 0.95 ποσοστημόρια, και οριοθετούν το ΔΕ μέσα στο οποίο βρίσκεται η τιμή της μεταβλητής $\overline{d_{min}}$ για το 90% των προσομοιωμένων σημειακών προτύπων με βάση τη μηδενική υπόθεση

Θεωρητικά Αναμενόμενη Τιμή για το Στατιστικό $\overline{d_{min}}$

Δειγματοληπτικές κατανομές του στατιστικού $\overline{d_{min}}$ που υπολογίστηκαν από 1000 εντελώς τυχαία προσομοιωμένα σημειακά πρότυπα με αναμενόμενο (μέσο) αριθμό συμβάντων $n = 50$ (αριστερά) και $n = 100$ (δεξιά) σε μια περιοχή 100×100





Θεωρητικά αναμενόμενη τιμή

Έχει αποδειχθεί μαθηματικά ότι, για μια εντελώς τυχαία σημειακή διεργασία Poisson και για μια περιοχή με σχετικά απλή γεωμετρία, η αναμενόμενη τιμή του στατιστικού $\overline{d_{min}}$ δίνεται από τη σχέση: $\mathbb{E}\{\overline{D_{min}}\} = 0.5\sqrt{|D|/n} + 0.051(P/n) + 0.041(P/n^{3/2})$, όπου $|D|$ και P είναι, αντίστοιχα, το εμβαδόν και η περίμετρος της περιοχής D. Στο παράδειγμα αριστερά, ο μέσος όρος των 1000 προσομοιωμένων τιμών $\overline{d_{min}}$ είναι 7.548, και προσεγγίζει την αντίστοιχη θεωρητικά αναμενόμενη τιμή $\mathbb{E}\{\overline{D_{min}}\} = 7.525$

Δειγματοληπτική Κατανομή του Στατιστικού $\hat{G}(d)$ (1)

Μεταβλητή περιγραφής σημειακού προτύπου

Απόσταση μεταξύ πλησιέστερων συμβάντων. Για ένα πρότυπο (δείγμα) με n συμβάντα, υπάρχουν n τέτοιες αποστάσεις $\{d_i^{min}, i = 1, \ldots, n\}$

Στατιστικό δείγματος

Συνάρτηση αθροιστικής κατανομής n τιμών απόστασεων μεταξύ πλησιέστερων συμβάντων: $\hat{G}(d)$ = ποσοστό των n αποστάσεων $d_i^{min} \leq d$

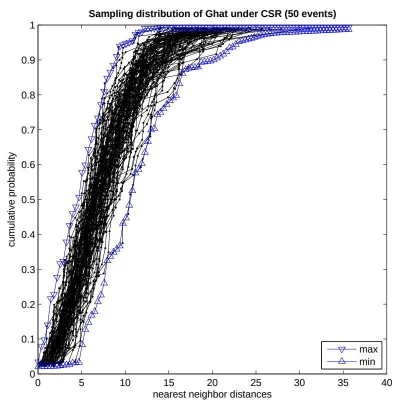
Κατασκευή της δειγματοληπτικής κατανομής του στατιστικού

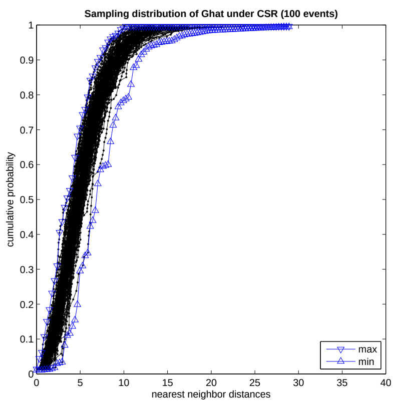
μέσω προσομοίωσης από τη μηδενική υπόθεση:

1. προσομοίωση ενός δείγματος, δηλαδή ενός χωρικού σημειακού προτύπου, από την εντελώς τυχαία χωρική σημειακή διεργασία με μέση ένταση λ
2. υπολογισμός της τιμής του στατιστικού $\hat{G}(d)$ από το προσομοιωμένο πρότυπο
3. επανάληψη των βημάτων (1) και (2), πολλές, π.χ. $M = 1000$, φορές για τον υπολογισμό M προσομοιωμένων τιμών στατιστικού $\hat{G}(d)$
4. η δέσμη των M προσομοιωμένων συναρτήσεων $\hat{G}(d)$ αποτελεί (κατά προσέγγιση) τη δειγματοληπτική κατανομή του εν λόγω στατιστικού **με βάση τη μηδενική υπόθεση**

Δειγματοληπτική Κατανομή του Στατιστικού $\hat{G}(d)$ (2)

Δειγματοληπτικές κατανομές του στατιστικού $\hat{G}(d)$, που κατασκευάστηκαν από 500 προσομειωμένα σημειακά πρότυπα (από τη Poisson διεργασία) με αναμενόμενο αριθμό συμβάντων $E\{N\} = 50$ (αριστερά) και $E\{N\} = 100$ (δεξιά) σε μια τετραγωνική περιοχή 100×100



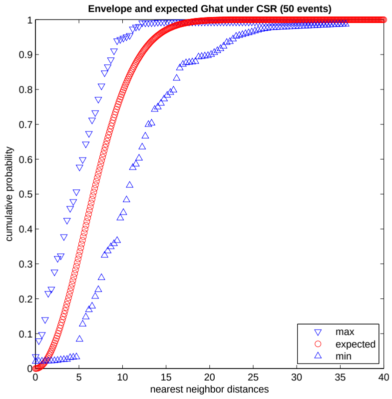


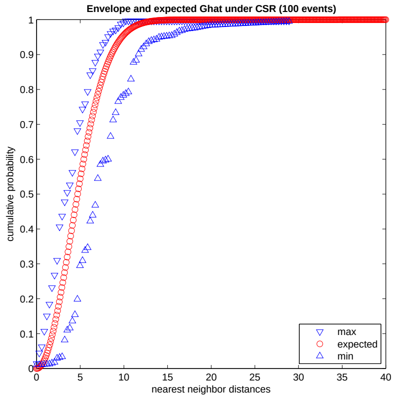
Όσο περισσότερα τυχαία συμβάντα εμφανίζονται σε μια περιοχή, τόσο πιο γρήγορα (σε μικρότερες αποστάσεις d) ανέρχεται και τόσο πιο συγκεντρωμένη είναι η δέσμη των προσομοιωμένων συναρτήσεων $\hat{G}(d)$

Η ελάχιστη $\hat{G}_{min}(d)$ και η μέγιστη $\hat{G}_{max}(d)$ τιμή των εμπειρικών συναρτήσεων $\hat{G}(d)$ για την απόσταση d, που υπολογίστηκαν από M προσομοιωμένα σημειακά πρότυπα, ορίζουν ένα διάστημα εμπιστοσύνης (ΔΕ) για επίπεδο σημανιτκότητας $\alpha = 1/(M+1)$

Θεωρητικά Αναμενόμενη Τιμή του Στατιστικού $\hat{G}(d)$

Ελάχιστες και μέγιστες τιμές του στατιστικού $\hat{G}(d)$ για κάθε απόσταση d, που υπολογίστηκαν από 500 προσομειωμένα σημειακά πρότυπα (από τη Poisson διεργασία) με αναμενόμενο αριθμό συμβάντων $E\{N\} = 50$ (αριστερά) και $E\{N\} = 100$ (δεξιά) σε μια τετραγωνική περιοχή 100×100





Θεωρητικά αναμενόμενη τιμή

Έχει αποδειχθεί μαθηματικά ότι, για μια εντελώς τυχαία σημειακή διεργασία Poisson και για μια περιοχή με σχετικά απλή γεωμετρία, η αναμενόμενη τιμή του στατιστικού $\hat{G}(d)$ δίνεται από τη σχέση: $\mathbb{E}\{\hat{G}(d)\} = 1 - \exp(-\lambda\pi d^2)$. Στα παραπάνω παραδείγματα, οι θεωρητικά αναμενόμενες τιμές της εμπειρικής συνάρτησης $\hat{G}(d)$ απεικονίζονται με *κόκκινους κύκλους* για τις αποστάσεις d

Δειγματοληπτική Κατανομή του Στατιστικού $\hat{L}(d)$ (1)

Μεταβλητή περιγραφής σημειακού προτύπου

Απόσταση μεταξύ συμβάντων, εξαιρουμένων των μηδενικών αποστάσεων. Για ένα πρότυπο (δείγμα) με n συμβάντα, υπάρχουν $n^2 - n$ τέτοιες αποστάσεις

Στατιστικό δείγματος

Συνάρτηση αθροιστικής κατανομής $n^2 - n$ απόστασεων μεταξύ συμβάντων (επί το εμβαδόν $|D|$ της περιοχής): $\hat{K}(d)$. Το στατιστικό αυτό έχει θεωρητικά αναμενόμενη τιμή (για μια εντελώς τυχαία σημειακή διεργασία) $\mathbb{E}\{\hat{K}(d)\} = \pi d^2$, και συνίσταται η μετατροπή του στο στατιστικό $\hat{L}(d) = \sqrt{\hat{K}(d)/\pi} - d$ με θεωρητικά αναμενόμενη τιμή 0

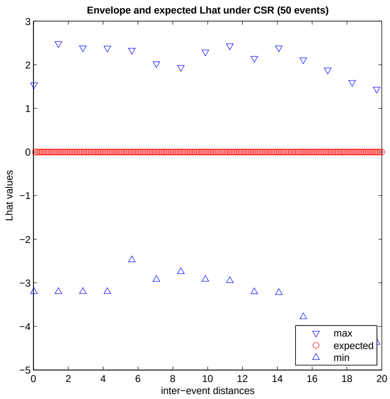
Κατασκευή της δειγματοληπτικής κατανομής του στατιστικού

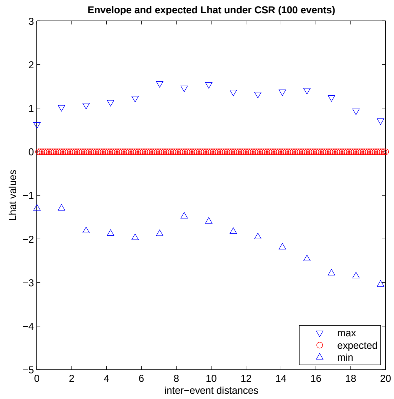
μέσω προσομοίωσης από τη μηδενική υπόθεση:

1. προσομοίωση ενός δείγματος, δηλαδή ενός χωρικού σημειακού προτύπου, από την εντελώς τυχαία χωρική σημειακή διεργασία με μέση ένταση λ
2. υπολογισμός της τιμής του στατιστικού $\hat{L}(d)$ από το προσομοιωμένο πρότυπο
3. επανάληψη των βημάτων (1) και (2), πολλές, π.χ. $M = 1000$, φορές για τον υπολογισμό M προσομοιωμένων τιμών στατιστικού $\hat{L}(d)$
4. η δέσμη των M προσομοιωμένων συναρτήσεων $\hat{L}(d)$ αποτελεί (κατά προσέγγιση) τη δειγματοληπτική κατανομή του εν λόγω στατιστικού **με βάση τη μηδενική υπόθεση**

Δειγματοληπτική Κατανομή του Στατιστικού $\hat{L}(d)$ (2)

Δειγματοληπτικές κατανομές του στατιστικού $\hat{L}(d)$ (ελάχιστη, μέση και μέγιστη τιμή για κάθε απόσταση d), που κατασκευάστηκαν από 500 εντελώς τυχαία προσομειωμένα σημειακά πρότυπα (από μια Poisson διεργασία) με αναμενόμενο αριθμό συμβάντων $E\{N\} = 50$ (αριστερά) και $E\{N\} = 100$ (δεξιά) σε μια τετραγωνική περιοχή 100×100.



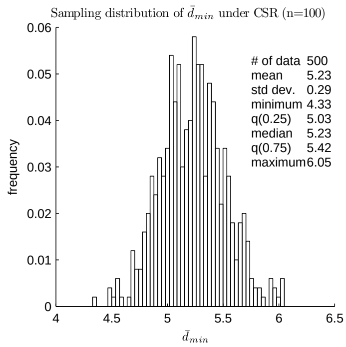


Όσο περισσότερα τυχαία συμβάντα εμφανίζονται σε μια περιοχή, τόσο πιο συγκεντρωμένη είναι η δέσμη των προσομοιωμένων εμπειρικών συναρτήσεων $\hat{L}(d)$

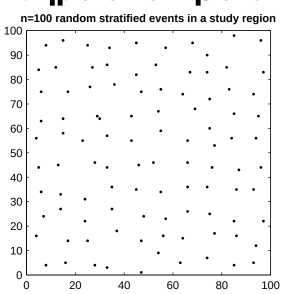
Η ελάχιστη $\hat{L}_{min}(d)$ και η μέγιστη $\hat{L}_{max}(d)$ τιμή των εμπειρικών συναρτήσεων $\hat{L}(d)$ για την απόσταση d, που υπολογίστηκαν από M προσομοιωμένα σημειακά πρότυπα, ορίζουν ένα διάστημα εμπιστοσύνης για επίπεδο σημαντικότητας $\alpha = 1/(M+1)$

Έλεγχος Υποθέσεως με Βάση το Στατιστικό $\bar{d}_{min}$ (I)

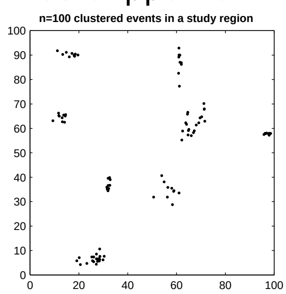
Δειγματοληπτική κατανομή
του στατιστικού $\overline{d_{min}}$ για μια ΕΤΧΣΔ



Δύο παρατηρούμενα σημειακά πρότυπα με $n = 100$ συμβάντα



$\overline{d_{min}} = 7.86$

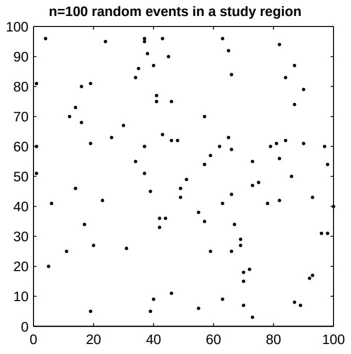


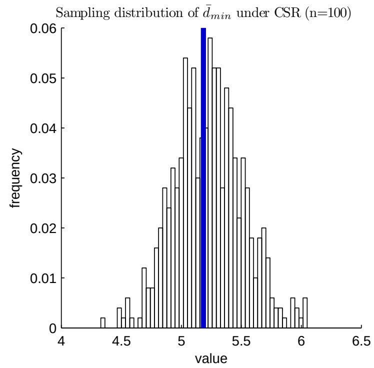
$\overline{d_{min}} = 1.17$

Ερώτηση: Μπορεί να θεωρηθεί ότι τα 2 αυτά πρότυπα προέρχονται από μια εντελώς τυχαία ΧΣΔ;
Η απάντηση είναι **όχι**, και μπορεί να δωθεί με μεγάλη εμπιστοσύνη, γιατί το πρότυπο στα αριστερά (δεξιά) έχει πολύ μεγαλύτερη (μικρότερη) τιμή $\overline{d_{min}}$ απ' ότι αναμένεται με βάση τη μηδενική υπόθεση της ΕΤΧΣΔ.

Έλεγχος Υποθέσεως με Βάση το Στατιστικό $\bar{d}_{min}$ (II)

Παρατηρούμενο πρότυπο με $n = 100$ συμβάντα,
και δειγματοληπτική κατανομή του στατιστικού $\overline{d_{min}}$ για μια ΕΤΧΣΔ



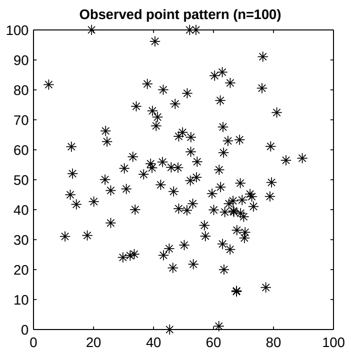


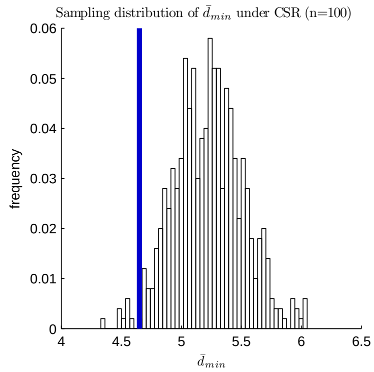
$\overline{d_{min}} = 5.18$

Ερώτηση: Είναι το παρατηρούμενο πρότυπο πιο συγκεντρωμένο απ' ότι αναμένεται για μια ΕΤΧΣΔ; Η απάντηση είναι πιθανότατα **όχι**, γιατί η παρατηρούμενη τιμή $\overline{d_{min}} = 5.18$ (κατακόρυφη γραμμή) βρίσκεται στο κέντρο της δειγματοληπτικής κατανομής του στατιστικού $\overline{d_{min}}$ για τη μηδενική υπόθεση μιας ΕΤΧΣΔ

Έλεγχος Υποθέσεως με Βάση το Στατιστικό $\bar{d}_{min}$ (III)

Παρατηρούμενο σημειακό πρότυπο με $n = 100$ συμβάντα, και δειγματοληπτική κατανομή του στατιστικού $\overline{d_{min}}$ για μια ΕΤΧΣΔ





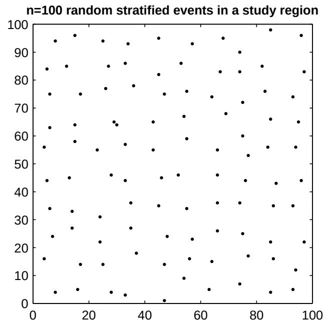
$\overline{d_{min}} = 4.65$

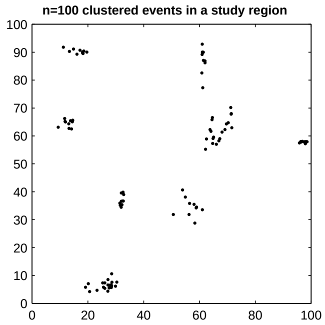
Ερώτηση: Είναι το παρατηρούμενο πρότυπο πιο συγκεντρωμένο απ' ότι αναμένεται για μια ΕΤΧΣΔ; Η ερώτηση αυτή υποδηλώνει ότι ενδιαφερόμαστε για μικρότερες τιμές $\overline{d_{min}}$ από αυτές που αντιστοιχούν σε μια ΕΤΧΣΔ, δηλαδή μας ενδιαφέρει πόσο στα αριστερά της δειγματοληπτικής κατανομής βρίσκεται η παρατηρούμενη τιμή

Απάντηση: Το εμβαδό των ραβδών στα αριστερά της παρατηρούμενης τιμής $\overline{d_{min}} = 4.65$ ποσοτικοποιεί την πιθανότητα p να **ανήκει** η παρατηρούμενη τιμή στη δειγματοληπτική κατανομή του στατιστικού. Στη συγκεκριμένη περίπτωση: $p = 0.017$

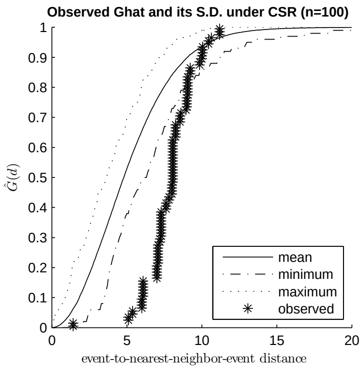
Έλεγχος Υποθέσεως με Βάση το Στατιστικό $\hat{G}(d)$ (I)

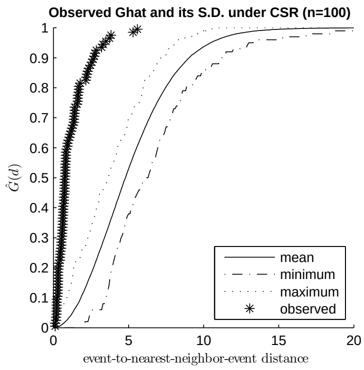
2 παρατηρούμενα σημειακά πρότυπα με $n = 100$ συμβάντα





Ερώτηση: Είναι τα πρότυπα αυτά τυχαία ή όχι;

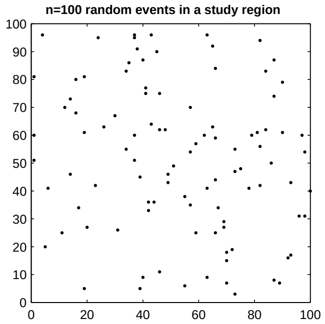




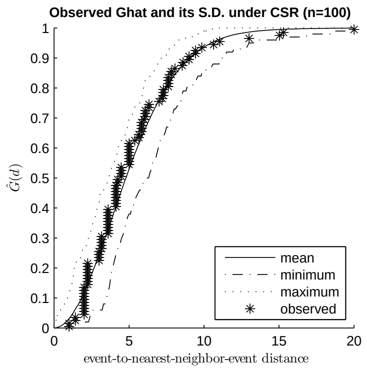
Πιθανότατα **όχι**, γιατί η παρατηρούμενη εμπειρική συνάρτηση $\hat{G}(d)$ δεν εμπίπτει στα όρια της δειγματοληπτικής κατανομής του στατιστικού $\hat{G}(d)$ με βάση μια ΕΤΧΣΔ

Έλεγχος Υποθέσεως με Βάση το Στατιστικό $\hat{G}(d)$ (II)

Παρατηρούμενο πρότυπο με $n = 100$ συμβάντα



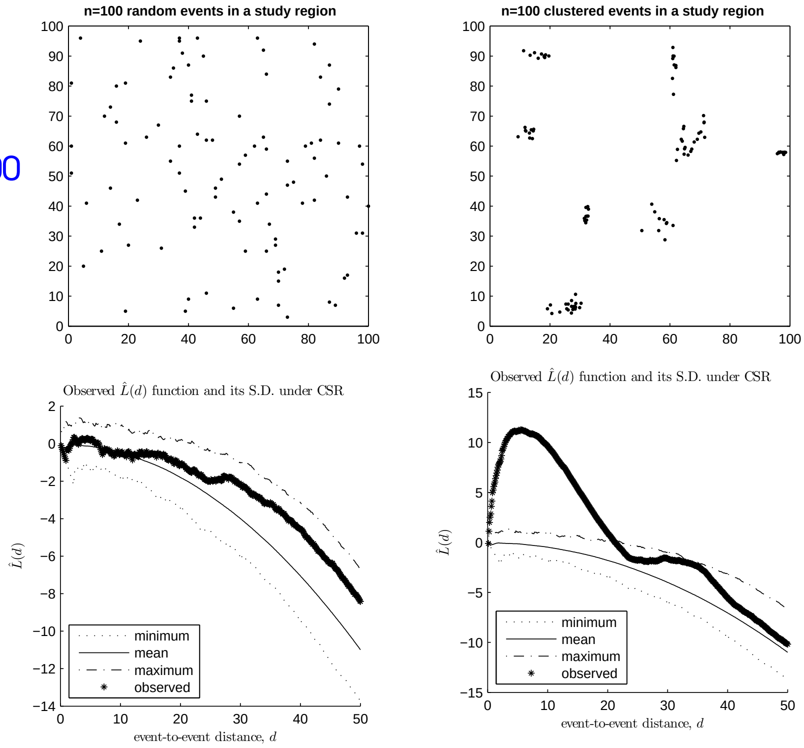
Ερώτηση: Είναι το παρατηρούμενο σημειακό πρότυπο τυχαίο ή όχι;



Πιθανότατα **ναι**, γιατί η παρατηρούμενη συνάρτηση $\hat{G}(d)$ βρίσκεται στο κέντρο της δειγματοληπτικής κατανομής του στατιστικού $\hat{G}(d)$ με βάση μια ΕΤΧΣΔ

Έλεγχος Υποθέσεως με Βάση το Στατιστικό $\hat{L}(d)$

Παρατηρούμενα πρότυπα, $n = 100$



Η παρατηρούμενη εμπειρική συνάρτηση $\hat{L}(d)$ βρίσκεται υψηλότερα (για μικρές αποστάσεις d) από ότι αναμένεται με βάση μια ΕΤΧΣΔ $\Rightarrow$ συγκεντρωτικότητα

Βασικά Σημεία Διάλεξης

Έλεγχος υποθέσεως για χωρικά σημειακά πρότυπα

- ποσοτικοποίηση απόκλισης μεταξύ της τιμής του στατιστικού, π.χ., $\overline{d_{min}}$ ή $\hat{G}(d)$, που υπολογίστηκε από ένα σημειακό πρότυπο από τις αντίστοιχες αναμενόμενες τιμές του στατιστικού με βάση μια μηδενική υπόθεση (εδώ μια ΕΤΣΧΔ)
- απόρριψη της μηδενικής υπόθεσης αν η παραπάνω απόκλιση υπερβαίνει ένα υποκειμενικά ορισμένο όριο (επίπεδο σημαντικότητας)

Μια εντελώς τυχαία σημειακή διεργασία υποδηλώνει: (ι) χωρικά σταθερή αναμενόμενη ένταση συμβάντων, και (ιι) απουσία αλληλεπιδράσεων μεταξύ συμβάντων

Δειγματοληπτική κατανομή στατιστικού

- κατανομή τιμών ενός στατιστικού που υπολογίζονται από προσομοιωμένα δείγματα (εδώ χωρικά σημειακά πρότυπα) με βάση μια μηδενική υπόθεση (εδώ μια ΕΤΧΣΔ). **Σημείωση:** Τα πολλαπλά σημειακά πρότυπα προσομοιώνονται στην ίδια περιοχή μελέτης με τον ίδιο (αναμενόμενο) αριθμό συμβάντων όπως και στο παρατηρούμενο πρότυπο
- μια δειγματοληπτική κατανομή μπορεί να κατασκευαστεί και μαθηματικά χωρίς προσομοίωση, αλλά για περιοχές μελέτης με σχετικά απλό γεωμετρικό σχήμα (υπεισέρχονται επίσης και άλλα προβλήματα λόγω επιδράσεων ορίων)