

## ΚΕΦΑΛΑΙΟ 4

### Ανάλυση Χωρικών Σημειακών προτύπων

#### Στόχοι κεφαλαίου

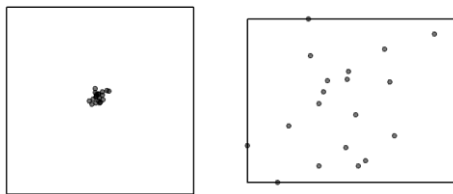
1. Μέθοδοι ανάλυσης χωρικού σημειακού προτύπου

#### 4.1. Βασικοί ορισμοί

##### 4.1.1. Χωρικό σημειακό πρότυπο

Το **χωρικό σημειακό πρότυπο** (*spatial point pattern*) είναι μια ομάδα από τοποθεσίες  $L$  σε ένα προσδιορισμένο χώρο  $A$ , όπου  $x$  γεγονότα λαμβάνουν μέρος. Το **Χωροχρονικό πρότυπο** είναι ένα χωρικό πρότυπο γεγονότων που συμβαίνουν στο χρόνο. Ως **συμβάν** (*event*), ορίζεται η εμφάνιση ενός φαινομένου, μιας κατάστασης ή μιας παρατήρησης σε μια συγκεκριμένη τοποθεσία.

Ο στόχος του κεφαλαίου είναι η ανάλυση των εργαλείων για την ποσοτικοποίηση μεταξύ των συμβάντων ενός χωρικού σημειακού προτύπου. Χρησιμοποιούνται για να περιγράψουν τη μεταβλητότητα των συμβάντων στο χώρο καθώς και να προσδιορίσουν σχέσεις ανάμεσα στα γεγονότα, με την αναγνώριση των χωρικών διαδικασιών που καθοδηγούν τη διάταξη τους.

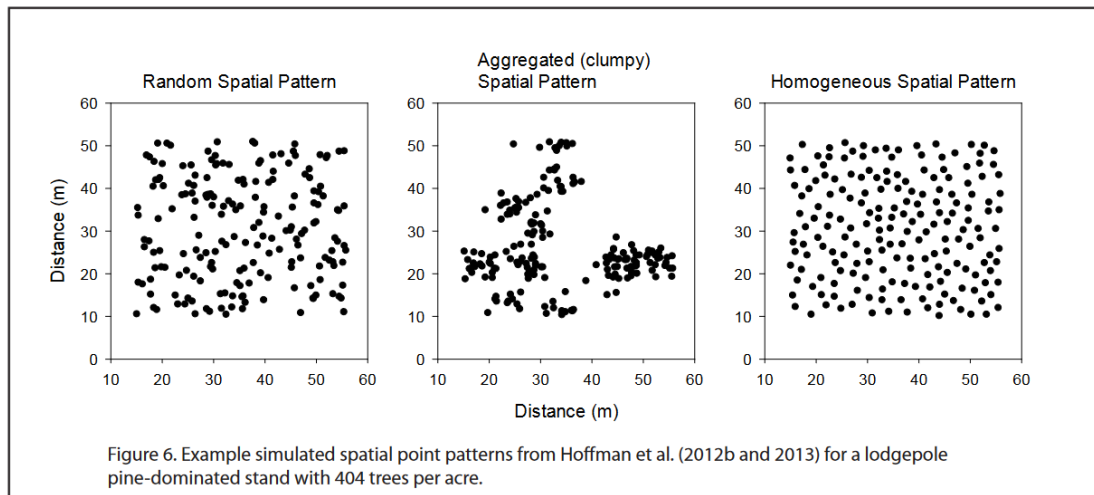


##### 4.1.2. Χωρική διαδικασία (*spatial process*)

Η χωρική διαδικασία είναι η περιγραφή πως μπορεί να δημιουργηθεί ένα χωρικό πρότυπο. Υπάρχουν τρεις κύριοι τύποι του χωρικού προτύπου.

- **Διαδικασία τυχαίου σημειακού προτύπου:** Τα χωρικά αντικείμενα κατανέμονται στο χώρο με βάση δύο αρχές. 1) Υπάρχει όμοια πιθανότητα για την εμφάνιση των γεγονότων σε κάθε τοποθεσία στην περιοχή (*first-order stationary*). 2) Η θέση κάθε γεγονότος είναι ανεξάρτητη από τις θέσεις των άλλων γεγονότων (*second-order stationary*). Αυτό το πρότυπο οδηγεί σε τυχαία διάταξη των γεγονότων σε όλη την περιοχή.
- **Χωρική ομαδοποίηση (clustering):** Είναι η διαδικασία που οδηγεί τα γεγονότα να διατάσσονται σε συστάδες (*clusters*) ως αποτέλεσμα μιας δύναμης ώθησης. Δυνάμεις σύγκλισης που κάνουν τις θέσεις να «έλκονται» μεταξύ τους. Π.χ. τα περισσότερα ξενοδοχεία τείνουν να είναι συγκεντρωμένα γύρω από ένα ιστορικό κέντρο, ένα είδος φυτού παράγει σπόρια που δεν έχουν την ικανότητα να μεταφερθούν πολύ μακριά.

- **Χωρική απώθηση (dispersion):** Με αυτή τη διαδικασία, τα γεγονότα οδηγούνται να διατάσσονται όσο το δυνατό πιο μακριά το ένα από το άλλο, πέραν του αναμενόμενου (σύμφωνα με το Τυχαίο Πρότυπο) . Αυτό οδηγεί σε μια ομοιόμορφη κατανομή των συμβάντων. Π.χ. τα μεγάλα αρπακτικά εξαιτίας του ενδο-ειδικού ανταγωνισμού έχουν ίσες αποστάσεις ανάμεσα στις επικράτειες τους, τα φωτόφιλα είδη δέντρων (που κάθε δέντρο χρειάζεται αυτόνομη επιφάνεια εδάφους).



#### 4.1.3. Χωρική παραλλαγή πρώτης τάξης – δεύτερης τάξης (First-order spatial variation – Second-order spatial variation)

**Η χωρική παραλλαγή πρώτης τάξης** συμβαίνει όταν οι τιμές ή οι θέσεις των χωρικών αντικειμένων διαφέρουν από περιοχή σε περιοχή εξαιτίας μιας τοπικής επίδρασης στο χώρο (π.χ. τα προβλήματα υγείας διαφέρουν από θέση σε θέση σε μια πόλη, αλλά αναμένεται να συγκεντρώνονται σε περιοχές όπου οι άνθρωποι είναι συγκεντρωμένοι κοντά σε βιομηχανικές ζώνες). Τα γεγονότα δεν συμβαίνουν με την ίδια πιθανότητα σε όλες τις θέσεις. Άλλο παράδειγμα είναι η κατανομή των βελανιδιών, η οποία θα ποικίλλει σε ένα τοπίο με βάση τα υποκείμενα χαρακτηριστικά του εδάφους (με αποτέλεσμα κάποιες περιοχές να έχουν πυκνές συστάδες βελανιδιών και άλλες περιοχές όχι). Οι επιδράσεις πρώτης τάξης σχετίζονται με μετρήσεις πυκνότητας /έντασης (*density / intensity measures*).

**Η χωρική παραλλαγή δεύτερης τάξης** συμβαίνει όταν υπάρχει μια αλληλεπίδραση ανάμεσα σε γειτονικές θέσεις. Για παράδειγμα οι μετανάστες σε μια νέα πόλη είναι πιθανότερο να κατοικήσουν σε μια γειτονιά με ανθρώπους ίδιας εθνικότητας. Υπάρχει μια ισχυρή τοπική έλξη που προσελκύει νεοεισερχόμενους. Οι επιδράσεις δεύτερης τάξης σχετίζονται κυρίως με μετρήσεις απόστασης (*distance measures*).

#### 4.2. Χωρικές διεργασίες

Κάθε χωρικό πρότυπο σε μια συγκεκριμένη θέση και σε μια συγκεκριμένη στιγμή είναι το αποτέλεσμα μιας διεργασίας που συμβαίνει μέσα σε ένα ευρύτερο χώρο και χρόνο. Για ένα

συγκεκριμένο σετ συμβάντων, μέσα από τη χωρική ανάλυση ενδιαφερόμαστε να αναλύσουμε α) το παρατηρούμενο χωρικό πρότυπο και β) τις διεργασίες που έχουν δημιουργήσει αυτή τη διάταξη.

Μια χωρική διεργασία μπορεί να είναι:

1. **Ντετερμινιστική**, όπου οι εισροές και εκροές είναι συγκεκριμένες σε μια χρονική στιγμή
2. **Στοχαστική**, όταν το αποτέλεσμα υπόκειται σε διαφοροποίηση που δεν μπορεί να περιγραφεί με ακρίβεια (π.χ. τα τυχαίοποιημένα πρότυπα).

#### 4.2.1. Έλεγχος υποθέσεως Τυχαίου Σημειακού Προτύπου

Ο έλεγχος υποθέσεων γενικά στη στατιστική επιστήμη είναι μια διαδικασία προσδιορισμού αν μια δεδομένη υπόθεση ισχύει ή όχι. Στη χωρική ανάλυση, υπάρχει η ανάγκη να επιβεβαιώσουμε / απορρίψουμε την υπόθεση ότι κάποιο χωρικό σημειακό πρότυπο είναι αποτέλεσμα τυχαίας χωρικής διεργασίας.

**Η Εντελώς Τυχαία Χωρική Σημειακή Διεργασία** (*Complete Spatial Randomness*) είναι μια στοχαστική χωρική διεργασία με τα εξής δύο χαρακτηριστικά:

- Η μέση ένταση (αναμενόμενος αριθμός συμβάντων στη μοναδιαία επιφάνεια) είναι σταθερή στην περιοχή μελέτης. *Απουσία εξωγενών ή περιβαλλοντικών επιδράσεων οι οποίοι θα μπορούσαν να διαφοροποιήσουν τη μέση ένταση.*
- Η θέση εμφάνισης ενός συμβάντος είναι ανεξάρτητη από την εμφάνιση ενός άλλου συμβάντος. *Δεν υπάρχουν αλληλεπιδράσεις μεταξύ των συμβάντων.*

#### Τυχειότητα συμβάντων

Με βάση την τυχειότητα των συμβάντων, στη χωρική διάσταση η **μηδενική υπόθεση** είναι η πλήρης χωρική τυχειότητα, όπου η πιθανότητα να βρούμε συναθροίσεις (συστάδες) είναι ελάχιστη και δεν υπάρχει χωρική αυτοσυσχέτιση.

Για να αποδεχτούμε ή να απορρίψουμε τη μηδενική υπόθεση χρησιμοποιούμε δύο μετρικές: **το z-score και το p-value.**

- **To z-score** είναι η κρίσιμη τιμή που χρησιμοποιείται για να υπολογιστεί το p-value με βάση την κανονική κατανομή. Αντανακλά τον αριθμό των τυπικών αποκλίσεων στην οποία μια τιμή βρίσκεται από τη μέση τιμή της.
- **Η p-value** είναι η πιθανότητα που προσδιορίζει το επίπεδο εμπιστοσύνης. Όταν η p-value είναι πολύ μικρή, η πιθανότητα το παρατηρούμενο πρότυπο να έχει δημιουργηθεί με τυχειότητα είναι ελάχιστη και μπορούμε να απορρίψουμε τη μηδενική υπόθεση (της τυχειότητας). Επομένως υπάρχει έντονη ένδειξη για την ύπαρξη μιας υποκείμενης χωρικής διαδικασίας.

Αν υπάρχει χωρική τυχειότητα, δεν υπάρχει κάποιο πρότυπο ή υποκείμενη διεργασία και δεν χρειάζεται περαιτέρω ανάλυση. Αν τα δεδομένα δεν έχουν τυχειότητα και η χωρική ανάλυση είναι πολύ σημαντική, καθώς ο χώρος έχει μια επίδραση στις εμφανίσεις των συμβάντων και στις σχετικές τιμές τους.

Όταν συμβαίνουν πρώτης τάξης επιδράσεις (*first-order effects*), η θέση της παρατήρησης είναι ο βασικός καθοριστικός παράγοντας για την εμφάνιση του συμβάντος. Αν είναι δεύτερης τάξης αποτελέσματα (*second-order effects*), οι αλληλεπιδράσεις ανάμεσα στα συμβάντα επηρεάζονται κυρίως από την απόστασή τους.

### 4.3. Μέθοδοι ανάλυσης σημειακού χωρικού προτύπου

Υπάρχουν δύο κύριες μέθοδοι για την ανάλυση των σημειακών προτύπων

1. **Μέθοδοι με βάση την Απόσταση** (*Distance-based methods*)
2. **Μέθοδοι με βάση την Πυκνότητα** (*Density-based methods*)

Οι μέθοδοι απόστασης χρησιμοποιούν τις αποστάσεις ανάμεσα στα συμβάντα και περιγράφουν **επιδράσεις δευτέρας τάξης**. Περιλαμβάνουν μεθόδους όπως

- Πλησιέστερου γείτονα (nearest neighbor)
- Εμπειρικές συναρτήσεις G και F
- Εμπειρικές συναρτήσεις Ripley's K και ο μετασχηματισμός της L

### Πλήρης χωρική τυχαιότητα (CSR)

#### Η απλούστερη μηδενική υπόθεση της χωρικής κατανομής

Τα σημεία κατανέμονται τυχαία με καμία αλληλεπίδραση

- Χωρίς ομαδοποίηση (έλξη, προνομιακές συνθήκες . . . )
- Χωρίς διασπορά (απώθηση, περιορισμένοι πόροι στην περιοχή . . . )

Παράγεται με μια (χωρικά) ομοιογενή διαδικασία Poisson με σταθερή ένταση  $\lambda(s) = \lambda > 0, \forall s \in A$

Η πιθανότητα οποιουδήποτε αριθμού σημείων στην ίδια μεγέθους περιοχή είναι η ίδια, σε όλο το εύρος κατανομής

### Κατανομή της συνάρτησης G

Μεταβλητή που περιγράφει την απόσταση μεταξύ των πλησιέστερων σημείων. Το στατιστικό είναι η συνάρτηση αθροιστικής κατανομής η τιμών αποστάσεων μεταξύ πλησιέστερων συμβάντων. Είναι μια συνεχής συνάρτηση του  $r$  (απόσταση) και δεν χρειάζεται διακριτοποίηση.

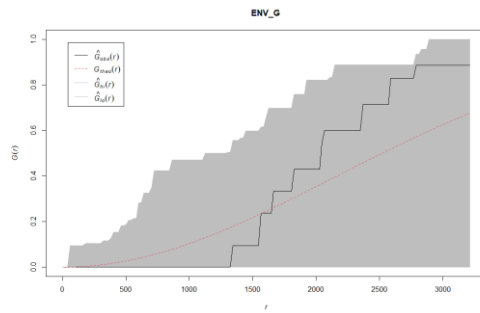
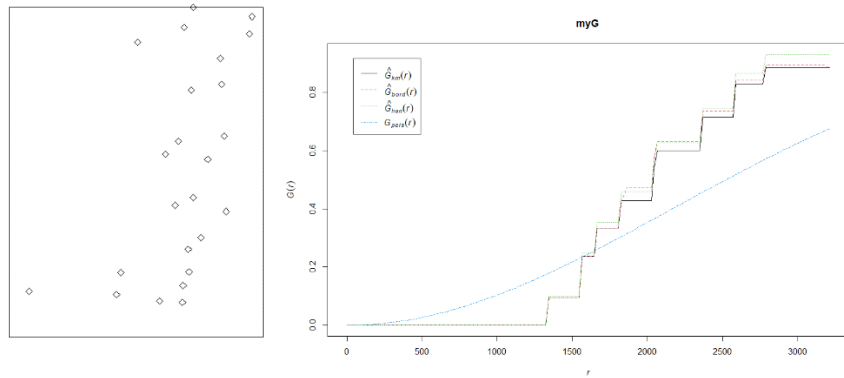
Το αποτέλεσμα της ομοιογενούς διαδικασίας Poisson είναι η θεωρητική συνάρτηση:

$G(r) = 1 - e^{-\lambda r^2}$  όπου  $\lambda$  είναι η ένταση διεργασίας, δηλ., ο μέσος αριθμός σημείων ανά μονάδα επιφάνειας.

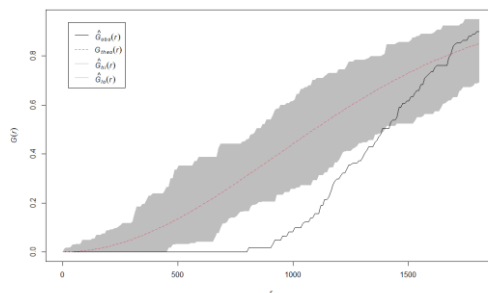
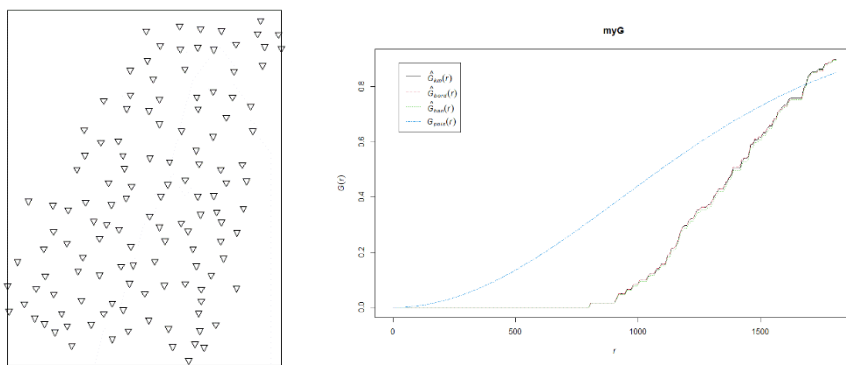
**Ομαδοποιημένα μοτίβα:**  $\hat{G}(r) > G(r)$  (περισσότερα κοντινά σημεία από τα αναμενόμενα κάτω από CSR)

**Διάσπαρτα (κανονικά) μοτίβα** :  $\hat{G}(r) < G(r)$  (λιγότερα κοντινά σημεία από τα αναμενόμενα κάτω από CSR)

Όλα αυτά αξιολογούνται σε οποιοδήποτε όριο (ακτίνα)  $r$ , που μπορεί να συμπεράνουμε ακτίνα ομαδοποίησης/διασποράς



clustered:  $\hat{G}(r) > G(r)$ , αλλά όχι σημαντικό

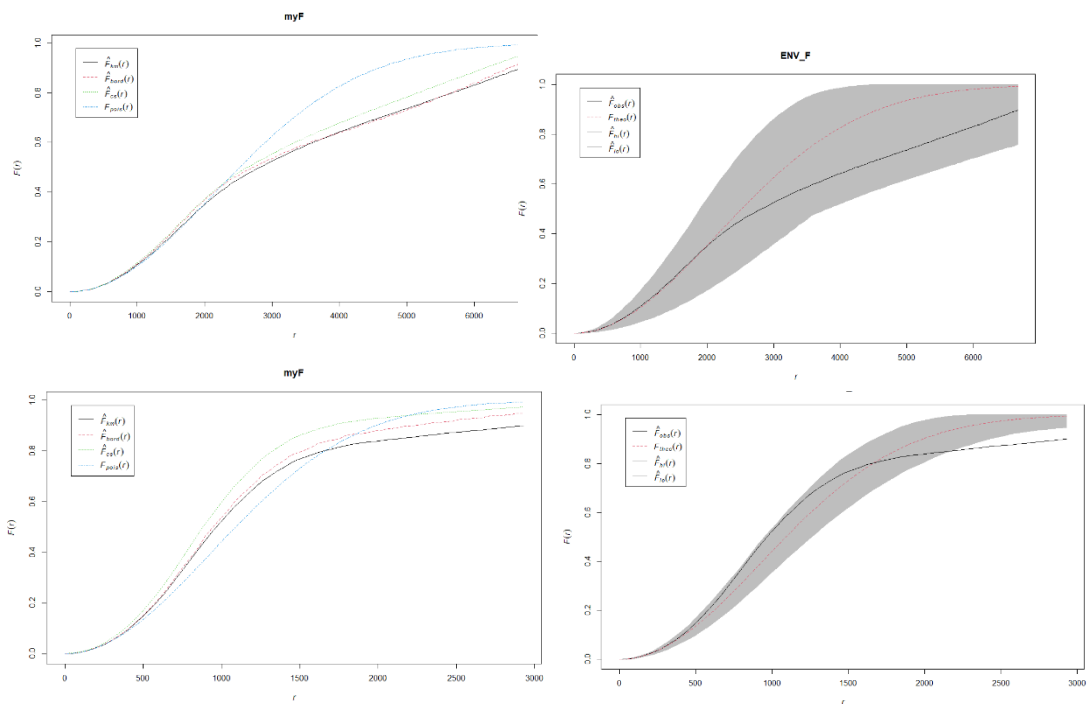


dispersed:  $\hat{G}(r) < G(r)$

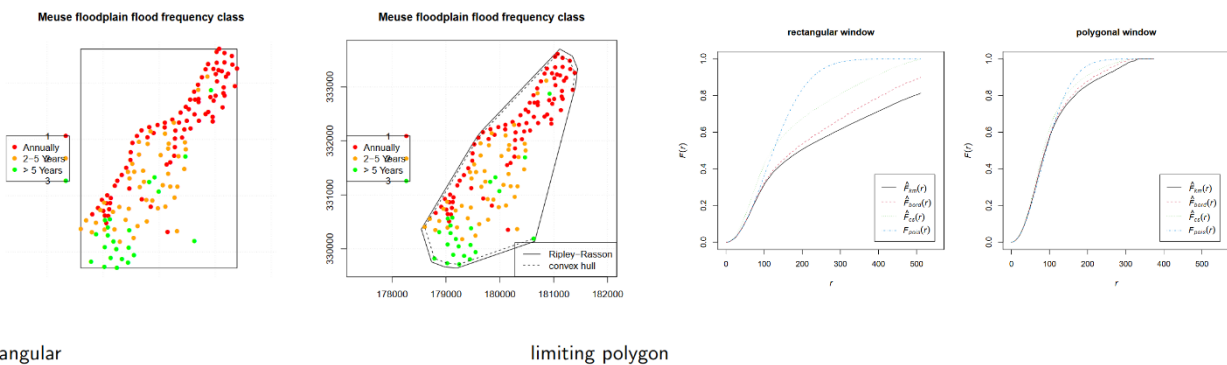
## Κατανομή της συνάρτησης F

Η συνάρτηση αυτή εξετάζει τον κενό χώρο σε ένα χωρικό σημειακό πρότυπο. Η συγκεκριμένη μέθοδος κάνει τυχαία επιλογή κάποιου σημείου εντός της περιοχής μελέτης (όχι απαραίτητα από τις παρατηρήσεις μας) και στη συνέχεια υπολογίζει τη μικρότερη απόσταση προς τη κοντινότερη παρατήρηση.

Η γραφική παράσταση της F μας βοηθάει να κατανοήσουμε κάποια χαρακτηριστικά του χωρικού σημειακού συνόλου μας. Για ομαδοποιημένα δεδομένα (clustered) ανέρχεται αργά αρχικά σε μικρές αποστάσεις και ανέρχεται αργότερα γρήγορα σε μεγάλες αποστάσεις. Για κανονικοποιημένα πρότυπα (απωθητικά), η F ανέρχεται γρήγορα για μικρές αποστάσεις και στη συνέχεια ανέρχεται αργά σε μεγάλες αποστάσεις.



Ευαίσθητη στο μέγεθος του παραθύρου εάν υπάρχει «κενός» χώρος στις άκρες της περιοχής.

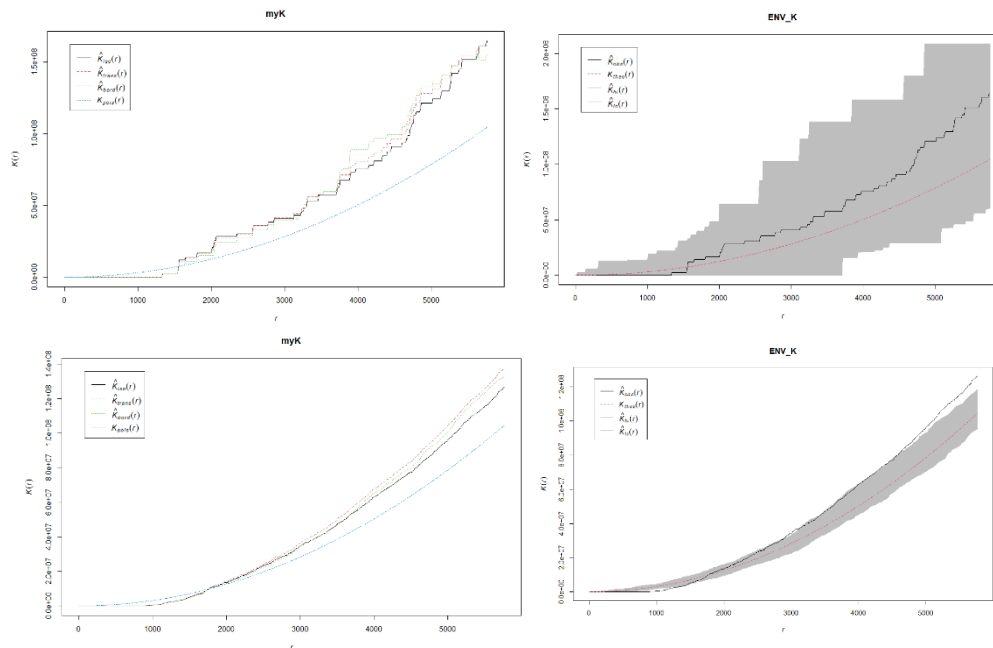


Rectangular

limiting polygon

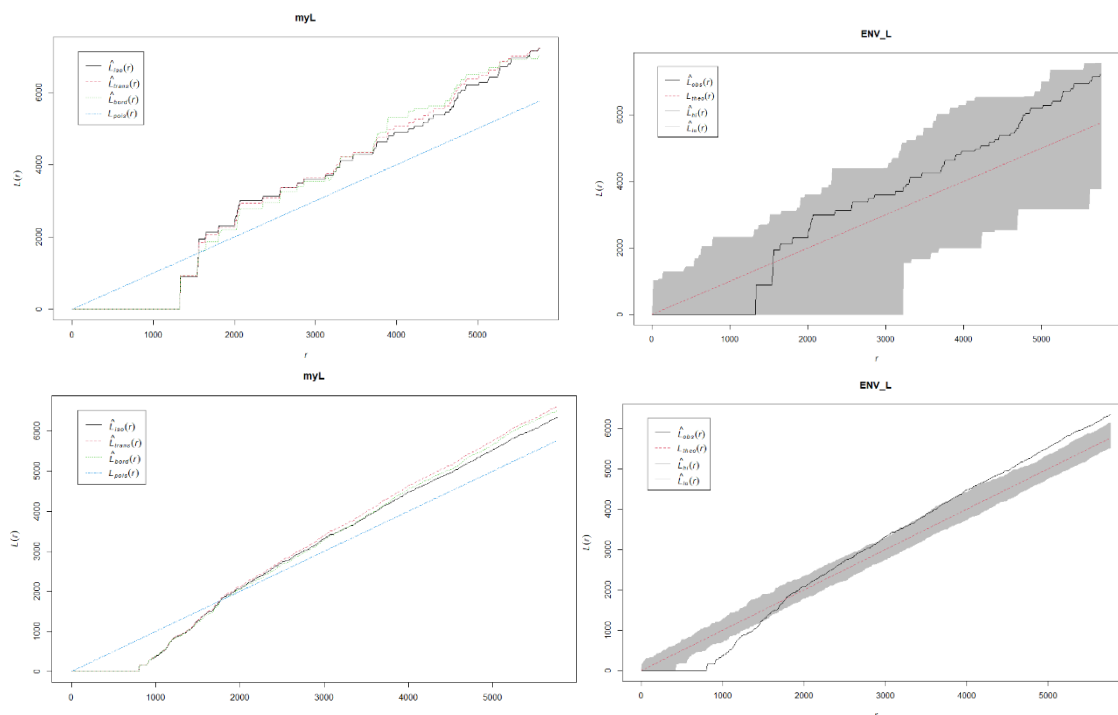
## Κατανομή της συνάρτησης K (Ripley's K)

Η συνάρτηση αυτή εξετάζει ομόκεντρους κύκλους από κάθε συμβάν. Αποδίδει το βαθμό χωρικής συγκέντρωσης του χωρικού προτύπου. Για κάθε  $i$  συμβάν, τοποθετείται ένας κύκλος  $d$  ακτίνας και υπολογίζεται πόσα άλλα συμβάντα υπάρχουν μέσα σε αυτό τον κύκλο. Ο συνολικός αριθμός συμβάντων είναι ο indicator factor  $i_d$ . Η διαδικασία εκτελείται για όλα τα συμβάντα σε ένα εύρος αποστάσεων. Όταν η παρατηρούμενη τιμή  $K$  είναι μεγαλύτερη από την αναμενόμενη, η κατανομή των συμβάντων τείνουν σε ομαδοποίηση (clustered pattern).



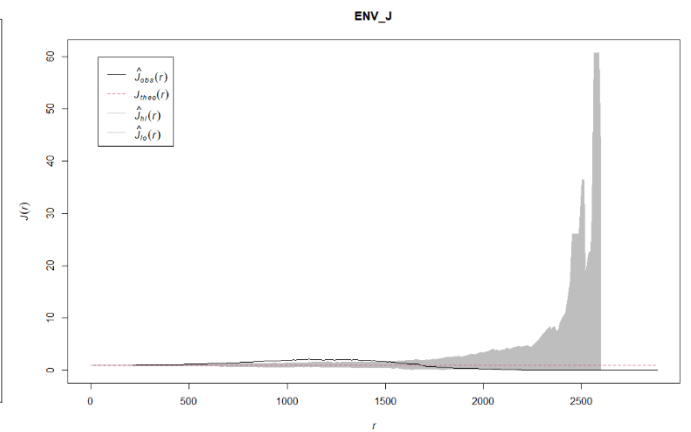
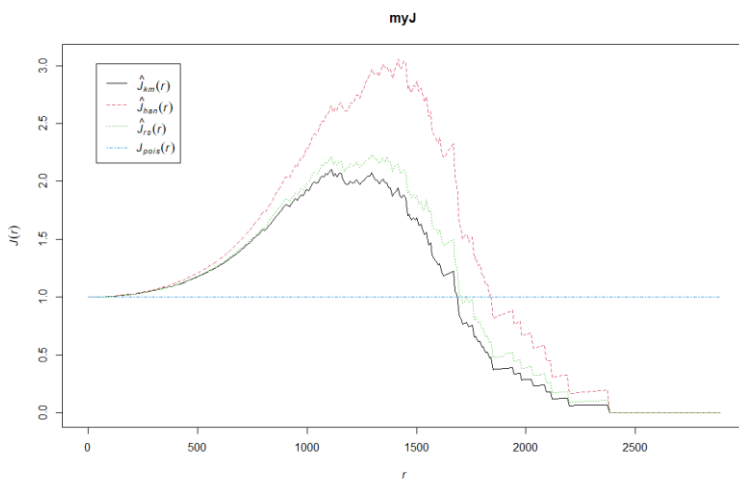
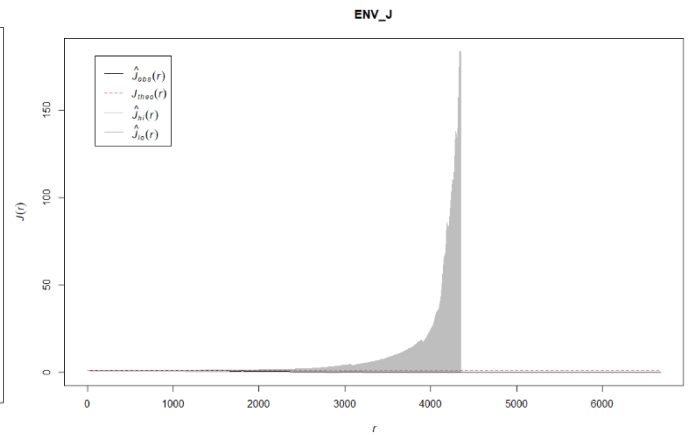
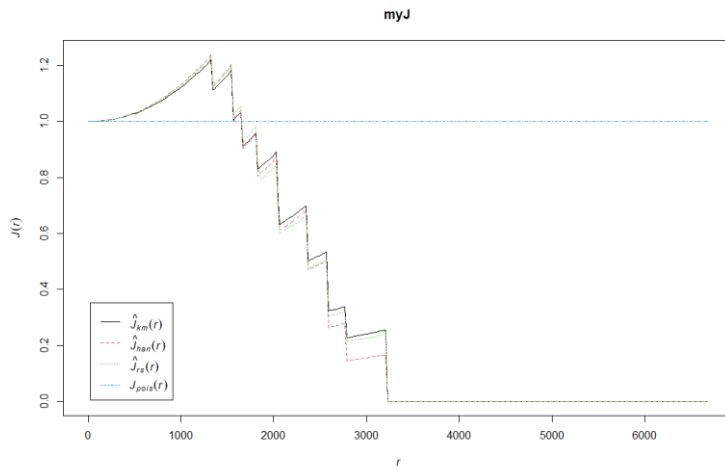
## Κατανομή της συνάρτησης L

Η συνάρτηση  $K$  είναι κάπως προβληματική, καθώς ο δείκτης  $K$  παίρνει μεγαλύτερες τιμές όταν οι αποστάσεις μεγαλώνουν και δεν αποδίδει πλήρως την απόκλιση από τη θεωρητική κατανομή. Ένας ευρέως μετασχηματισμός της  $K$  συνάρτησης, είναι το στατιστικό  $L$ , το οποίο έχει θεωρητικά αναμενόμενη τιμή 0 και έχει γραμμική σχέση με την κατανομή Poisson, κάνοντας ευκολότερη την ερμηνεία της



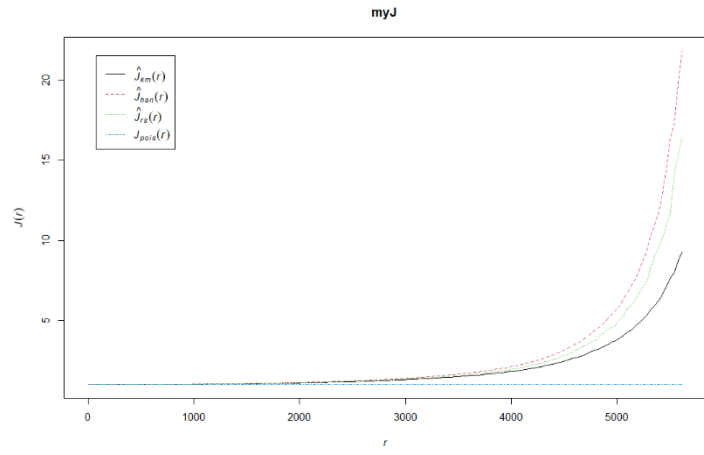
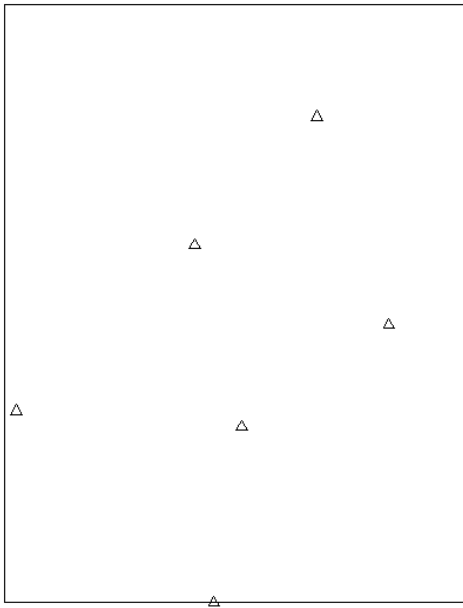
## Κατανομή της συνάρτησης J

Η συνάρτηση J είναι μια μετατροπή των G(d) και F(d) συναρτήσεων:  $J(r) = (1 - G(r)) / (1 - F(r))$ . Συνδυάζει την G (από σημείο σε σημείο) και την F (χώρος σε σημείο) και η αναμενόμενη τιμή για CSR είναι 1. Τιμή  $J(r) < 1$  υποδηλώνει ομαδοποίηση και  $J(r) > 1$  υποδηλώνει διασπορά (χωρική κανονικότητα). Το σημαντικό πλεονέκτημα της είναι ότι δεν είναι ευαίσθητη στην επίδραση των ορίων (edge effects).

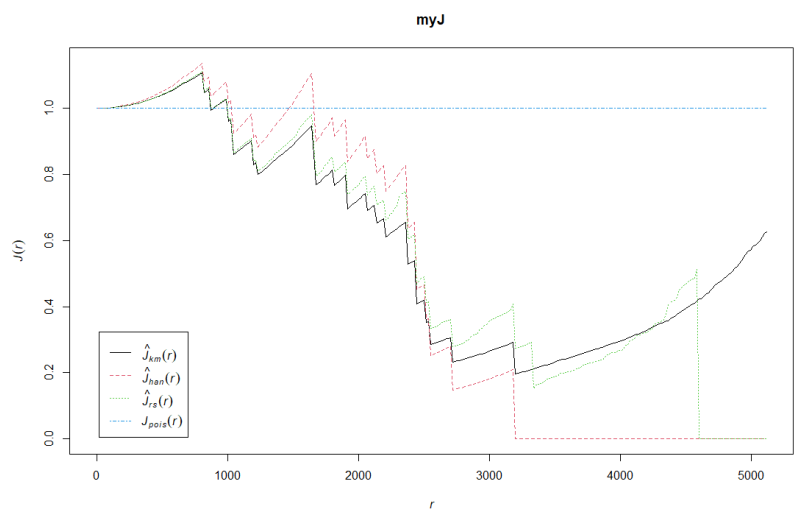
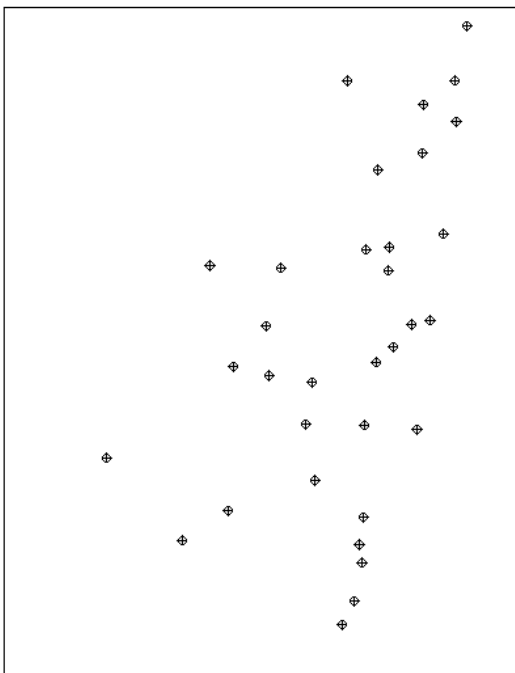




### Achrysaetos



### Cnigra



## Κώδικας στην R

```
#### ΦΟΡΤΩΜΑ ΒΙΒΛΙΟΘΗΚΩΝ
# ΒΙΒΛΙΟΘΗΚΗ ΓΙΑ ΝΑ ΔΙΑΒΑΣΟΥΜΕ SHP ΑΡΧΕΙΑ
library(sf)
DA <- read_sf("D:/DATA/DAD05.shp")

## ΔΟΜΗ ΤΟΥ ΑΡΧΕΙΟΥ
str(DA)

## ΜΕΤΑΤΡΟΠΗ ΚΑΙ ΜΕΤΑΝΟΜΑΣΙΑ ΑΡΧΕΙΩΝ
x<- DA$X
y<- DA$Y
SPEC <- as.factor(da95$SPECIES)

## ΜΕΤΑΤΡΟΠΗ ΤΟΥ ΑΡΧΕΙΟΥ ΣΕ ΠΙΝΑΚΑ
da95 <- as.data.frame(DA)

### ΔΙΑΜΟΡΦΩΣΗ ΤΗΣ ΠΕΡΙΟΧΗΣ ΜΕΛΕΤΗΣ (χ, ψ)
summary(da95$X)
summary(da95$Y)
owin (xrange= c(670680, 693743), yrange = c(4539160, 4569120)) ->w1

## ΔΗΜΙΟΥΡΓΙΑ ΤΟΥ PPP ΑΝΤΙΚΕΙΜΕΝΟΥ
ppp(x,y, window = w1, unitname=c("metre", "metres"),
    marks= factor(da95$SPECIES), check =T) -> x1
is.multitype(x1)

### ΑΠΕΙΚΟΝΙΣΗ ΚΑΤΑΝΟΜΗΣ ΕΙΔΩΝ
plot(x1)
points(x1, col=x1$marks)

## ΕΠΙΛΟΓΗ ΕΙΔΩΝ
myspecies = c("Cnig")
Daselect = subset(x1, x1$marks %in% myspecies , drop=F)
Daselect
plot(Daselect)
```

```
#### KONTINOTEPOΣ ΓΕΙΤΟΝΑΣ
```

```
NNR <- nndist(Daselect, k=1)
NNR
plot(NNR)
scatter.smooth(NNR, pch=20, cex=3, col="red")
```

```
#### ΔΗΜΙΟΥΡΓΙΑ ΣΥΝΑΡΤΗΣΕΩΝ ΑΠΟΣΤΑΣΗΣ
```

```
# G function
```

```
myG <- Gest(Daselect)
ENV_G = envelope(Y=Daselect, fun=Gest, nsim=400)
plot(myG)
plot(ENV_G)
```

```
# F function
```

```
myF <- Fest(Daselect)
ENV_F = envelope(Y=Daselect, fun=Fest, nsim=400)
plot(myF)
plot(ENV_F)
```

```
# K function
```

```
myK <- Kest(Daselect)
ENV_K = envelope(Y=Daselect, fun=Kest, nsim=400)
plot(myK)
plot(ENV_K)
```

```
# L function
```

```
myL <- Lest(Daselect)
ENV_L = envelope(Y=Daselect, fun=Lest, nsim=400)
plot(myL)
plot(ENV_L)
```

```
# J function
```

```
myJ <- Jest(Daselect)
ENV_J = envelope(Y=Daselect, fun=Jest, nsim=400)
plot(myJ)
plot(ENV_J)
```

**Οι μέθοδοι πυκνότητας** χρησιμοποιούν την ένταση (intensity) της εμφάνισης των συμβάντων στο χώρο. Περιγράφουν κυρίως πρώτης τάξης επιδράσεις. Περιλαμβάνουν μεθόδους όπως

- Εκτίμηση έντασης σε Φατνία (Quadrat count method)
- Εκτίμηση πυκνότητας με χρήση πυρήνα (Kernel density estimation)

**Η έννοια της έντασης** έχει να κάνει με το πλήθος των παρατηρήσεων ανά κάποια μονάδα μέτρησης. Είναι η ποσοτικοποίηση της συχνότητας των συμβάντων ανά επιφάνεια περιοχής. Η ανάλυση της χωρικής έντασης, μέσα από την ανάδειξη των αλλαγών στην ένταση, βοηθάει στη μελέτη των συσχετίσεων μεταξύ έντασης συμβάντων και άλλων χωρικών μεταβλητών. Δεν εισέρχεται στη μελέτη της σχετικής θέσης των συμβάντων και δεν εστιάζει στις αιτίες για της αλλαγές στη χωρική ένταση.

#### **Μερικές βασικές έννοιες:**

1. **Συχνότητα** (counts) είναι το πλήθος των παρατηρήσεων σε μια περιοχή
2. **Σχετική συχνότητα** (relative frequency) είναι το πλήθος των παρατηρήσεων σε μια περιοχή, προς το συνολικό πλήθος των παρατηρήσεων σε όλη την περιοχή μελέτης
3. **Ένταση**, είναι το πλήθος των παρατηρήσεων προς το εμβαδό μιας έκτασης (υπο-περιοχή του συνόλου)
4. **Πυκνότητα**, είναι η σχετική συχνότητα προς το εμβαδό μιας έκτασης (υπο-περιοχή του συνόλου)

#### **Εκτίμηση έντασης σε Φατνία / Κελιά**

Υπολογισμός και οπτικοποίηση αριθμού παρατηρήσεων σε Φατνία

```
#### QUADRAT COUNT
```

```
Q1 <- quadratcount(x1, nx=4, ny=4)
```

```
Q1
```

```
## Plot (x1) για τις παρατηρήσεις και plot(Q1) για να απεικονιστούν οι τιμές
```

```
## ανά φατνίο
```

```
plot(x1)
```

```
plot(Q1, col="blue", cex=2, add=TRUE)
```

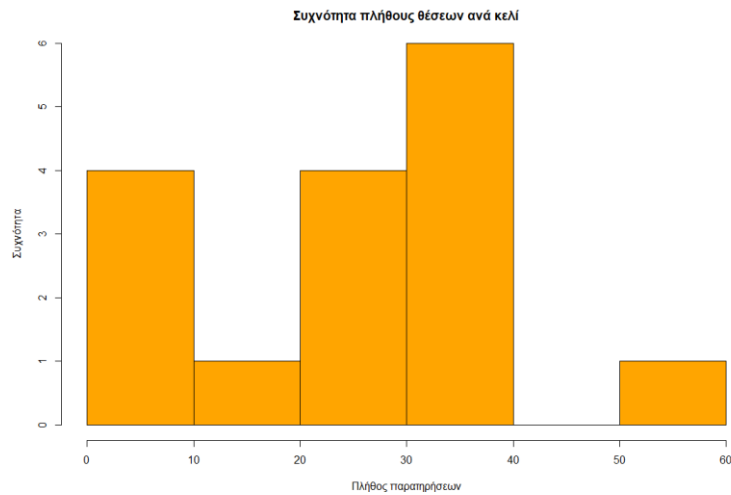
```
## Ιστόγραμμα πλήθους παρατηρήσεων ανά κελί
```

```
# Διαμόρφωση περιθωρίων
```

```
par(mar = c(4,4,2,0))
```

```
hist(Q1, col="orange", main="Συχνότητα πλήθους θέσεων ανά κελί",
```

```
      xlab="Πλήθος παρατηρήσεων", ylab= "Συχνότητα")
```



### Διαμόρφωση έντασης ανά κελί

```
## Ένταση ανά κελί
```

```
# Διαμόρφωση περιθωρίων
```

```
par(mar = c(0,0,2,1))
```

```
# Διαμόρφωση παλέτας διαβάθμισης χρωμάτων
```

```
mypalette = colorRampPalette(c("White", "orange", "red"))
```

```
# Διαμόρφωση plot έντασης ανά κελί
```

```
plot(intensity(Q1, image=TRUE), col=mypalette, lwd=2,  
      main = "'Ένταση θέσεων παρατηρήσεων")
```

```
# Ενσωμάτωση των παρατηρήσεων
```

```
plot(x1, pch=10, add=TRUE)
```

```
# Ιστόγραμμα έντασης θέσεων παρατηρήσεων
```

```
# Διαμόρφωση περιθωρίων
```

```
par(mar = c(4,4,2,0))
```

```
# Ιστόγραμμα
```

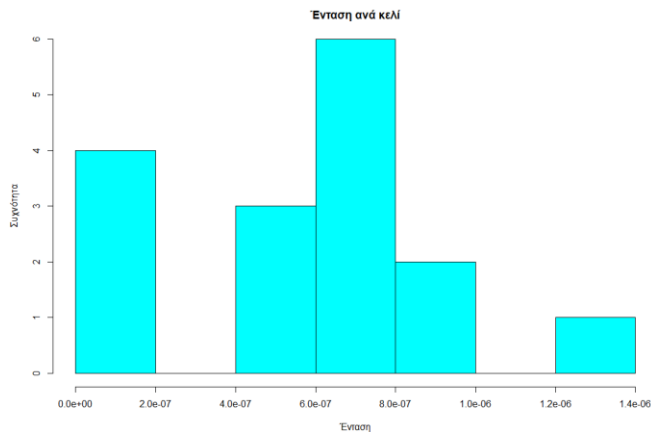
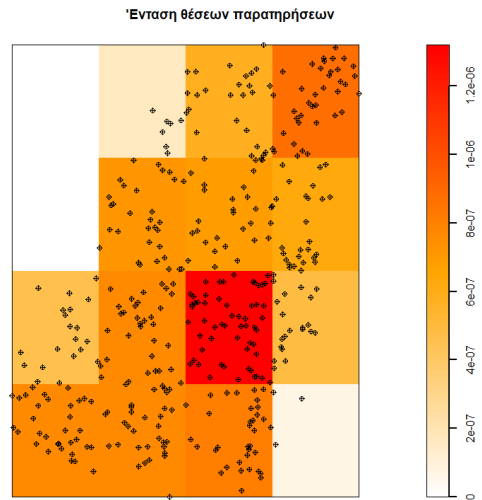
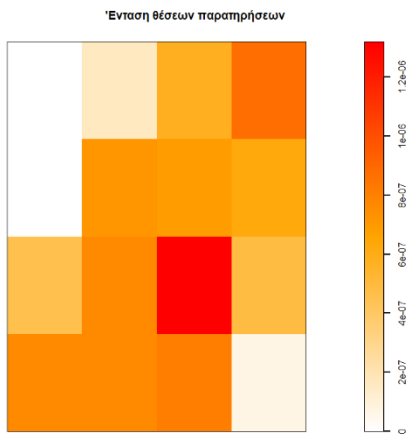
```
hist(intensity(Q1), col="cyan", main="Ένταση ανά κελί", xlab= "Ένταση",  
      ylab="Συχνότητα")
```

```
# Απενεργοποίηση επιστημονικής ονομασίας (2e - 07)
```

```
options(scipen=999)
```

```
# Επαναφορά επιστημονικής ονομασίας
```

```
options(scipen = 0)
```

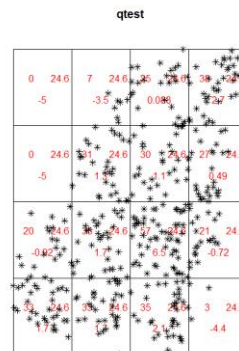


### Έλεγχος Υποθέσεων για CSR με QUADRAT TEST

Πραγματοποίηση ελέγχων για τυχαία διασπορά (μέσω  $\chi^2$  ή Monte Carlos)

qttest

0	24.6	7	24.6	25	24.6	38	24.6
-5		-3.5		0.088		2.7	
0	24.6	31	24.6	30	24.6	27	24.6
-5		1.3		1.1		0.49	
20	24.6	33	24.6	57	24.6	21	24.6
-0.92		1.7		6.5		-0.72	
33	24.6	33	24.6	35	24.6	3	24.6
1.7		1.7		2.1		-4.4	



Υπολογίζονται 3 τιμές α) Παρατηρούμενο πλήθος, β) Αναμενόμενο πλήθος και γ) Κατάλοιπα του Pearson (residuals). Μπορούμε να δούμε αν ισχύει η μηδενική υπόθεση είτε με το  $\chi^2$  είτε με τη μέθοδο Monte Carlo

Αποτέλεσμα της παραπάνω κατανομής ( $\chi^2 = 148,4$ ,  $p < 0.001$ ) -

### Κώδικας στην R

Αρχικός κώδικας για δημιουργία PPP αρχείου

```
.....  
#### ΦΟΡΤΩΜΑ ΒΙΒΛΙΟΘΗΚΩΝ  
# ΒΙΒΛΙΟΘΗΚΗ ΓΙΑ ΝΑ ΔΙΑΒΑΣΟΥΜΕ SHP ΑΡΧΕΙΑ  
library(sf)  
DA <- read_sf("D:/DATA/DAD05.shp")  
  
## ΔΟΜΗ ΤΟΥ ΑΡΧΕΙΟΥ  
str(DA)  
  
## ΜΕΤΑΤΡΟΠΗ ΚΑΙ ΜΕΤΑΝΟΜΑΣΙΑ ΑΡΧΕΙΩΝ  
x<- DA$X  
y<- DA$Y  
SPEC <- as.factor(da95$SPECIES)  
  
## ΜΕΤΑΤΡΟΠΗ ΤΟΥ ΑΡΧΕΙΟΥ ΣΕ ΠΙΝΑΚΑ  
da95 <- as.data.frame(DA)  
  
### ΔΙΑΜΟΡΦΩΣΗ ΤΗΣ ΠΕΡΙΟΧΗΣ ΜΕΛΕΤΗΣ (χ,ψ)  
summary(da95$X)  
summary(da95$Y)  
owin (xrange= c(670680, 693743), yrange = c(4539160, 4569120)) ->w1  
  
## ΔΗΜΙΟΥΡΓΙΑ ΤΟΥ PPP ΑΝΤΙΚΕΙΜΕΝΟΥ  
ppp(x,y, window = w1, unitname=c("metre","metres"),  
marks= factor(da95$SPECIES), check =T) -> x1  
is.multitype(x1)  
.....  
##### DENSITY ANALYSIS  
#### QUADRAT COUNT  
Q1 <- quadratcount(x1, nx=4, ny=4)  
Q1
```

```

## Plot (x1) για τις παρατηρήσεις και plot(Q1) για να απεικονιστούν οι τιμές
## ανά φατνίο

plot(x1)

plot(Q1, col="blue", cex=2, add=TRUE)

## Ιστόγραμμα πλήθους παρατηρήσεων ανά κελί
# Διαμόρφωση περιθωρίων

par(mar = c(4,4,2,0))

# Ιστόγραμμα

hist(Q1, col="orange", main="Συχνότητα πλήθους θέσεων ανά κελί",
      xlab="Πλήθος παρατηρήσεων", ylab= "Συχνότητα")

## Ένταση ανά κελί
# Διαμόρφωση περιθωρίων

par(mar = c(0,0,2,1))

# Διαμόρφωση παλέτας διαβάθμισης χρωμάτων

mypalette = colorRampPalette(c("White", "orange", "red"))

# Διαμόρφωση plot έντασης ανά κελί

plot(intensity(Q1, image=TRUE), col=mypalette, lwd=2,
      main = "'Ένταση θέσεων παρατηρήσεων")

# Ενσωμάτωση των παρατηρήσεων

plot(x1, pch=10, add=TRUE)

# Ιστόγραμμα έντασης θέσεων παρατηρήσεων
# Διαμόρφωση περιθωρίων

par(mar = c(4,4,2,0))

# Ιστόγραμμα

hist(intensity(Q1), col="cyan", main="Ένταση ανά κελί", xlab= "Ένταση",
      ylab="Συχνότητα")

# Απενεργοποίηση επιστημονικής ονομασίας (2e - 07)

options(scipen=999)

```



```
# Επαναφορά επιστημονικής ονομασίας
```

```
options(scipen = 0)
```

```
#### Έλεγχος Υποθέσεων για CSR με QUADRAT TEST
```

```
# Παρατηρούμενη τιμή
```

```
# Αναμενόμενη τιμή
```

```
# Κατάλοιπα Pearson (residuals)
```

```
# μέθοδος Χ2
```

```
quadrat.test(x1, nx=4, ny=4) -> qtest
```

```
plot(qtest, col="red")
```

```
points(x1, pch = 8)
```

```
qtest$statistic
```

```
qtest$p.value
```

```
# Μέθοδος Monte Carlo
```

```
quadrat.test(x1, nx=4, ny=4, method="MonteCarlo") -> qtest2
```

```
plot(qtest2, col="red")
```

```
points(x1, pch = 8)
```

```
qtest2$statistic
```

```
qtest2$p.value
```

### **Εκτίμηση πυκνότητας με χρήση πυρήνα (Kernel density estimation)**

Τρόπος εκτίμησης της έντασης σε σημειακά χωρικά δεδομένα με χρήση κυκλικών φατνίων, ώστε να έχουμε μια πιο ομαλή επιφάνεια μεταβολής της έντασης των συμβάντων. Η εκτίμηση πυκνότητας με χρήση πυρήνα (*Kernel density estimation*) είναι μια μη παραμετρική μέθοδος που χρησιμοποιεί συνάρτηση Kernel για τη χαρτογράφηση των τιμών πυκνότητας, όπου η πυκνότητα σε κάθε θέση υποδεικνύει τη συγκέντρωση σημείων γειτονικών περιοχών (μεγάλη συγκέντρωση ως «κορυφές») και χαμηλή συγκέντρωση ως «κοιλιάδες».

Η πυκνότητα πυρήνα σε μια θέση υπολογίζεται χρησιμοποιώντας τη συνάρτηση Kernel σε δύο ή παραπάνω διαστάσεις. Ο όρος “Kernel” στη χωρική στατιστική αναφέρεται κυρίως σε μια διεργασία «παραθύρου» κεντραρισμένη πάνω σε μια περιοχή που συστηματικά κινείται σε κάθε τοποθεσία, υπολογίζοντας την αντίστοιχη φόρμουλα.

Η απλούστερη φόρμουλα (κυκλική)

$$\hat{\lambda}(x) = \frac{N(b(x, h))}{\pi h^2}$$

Όπου

- Το  $b(x, h)$  είναι ένας δίσκος ακτίνας  $h$  με κέντρο το  $x$
- Το  $\hat{\lambda}(x)$  είναι η εκτιμώμενη πυκνότητα ανά μονάδα τετραγώνου.
- Το  $h$  είναι το εύρος ζώνης - φάσμα (bandwidth)
  - Μεγαλύτερες τιμές οδηγούν σε πιο ομαλές εκτιμήσεις καθώς ο πυρήνας μετακινείται στον χάρτη
- $N$  είναι ο αριθμός των σημείων στο δίσκο
- Ο παρονομαστής  $\pi h^2$  είναι το εμβαδόν αυτού του κύκλου που κανονικοποιεί την πυκνότητα

Υπάρχουν διάφοροι τύποι Kernel, που περιλαμβάνουν *normal*, *uniform*, *quadratic* και *Gaussian* τύπους. Το ArcPro χρησιμοποιεί τον *quadratic* τύπο.

Εικόνα Kernel density από το ArcPro, με εκτίμηση του bandwidth με βάση αριστερά το Silverman's Rule of Thumb και δεξιά με ακτίνα 5.000 μ. ως εύρος ζώνης). Μικρές τιμές φάσματος πυρήνα δίνουν πιο τραχιά εκτίμηση της έντασης, ενώ μεγάλες τιμές οδηγούν σε υπερ-εξομαλυσμένα αποτελέσματα, αποκρύπτοντας πιθανές λεπτομέρειες της εκτίμησης.

