

Modelos de distribuição espacial

A função L de Ripley

Exercício prático da disciplina Ecologia de Populações de Comunidades Vegetais, IB-USP

Contents

Introdução	1
O modelo Poisson homogêneo é uma boa aproximação?	1
A função L de Ripley	4

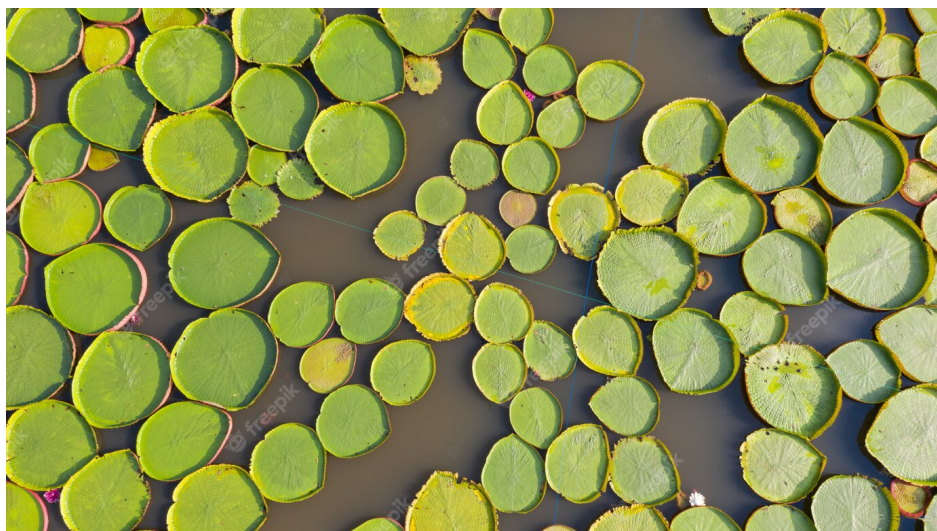
Introdução

No roteiro anterior, apresentamos quatro modelos de distribuição de pontos num plano: Poisson homogêneo, Poisson não-homogêneo, Thompson homogêneo e Thompson não-homogêneo. Estes modelos são úteis para descrever, entre outros padrões, a estrutura espacial de plantas. Por exemplo, se descobrimos que o modelo Thompson homogêneo aproxima bem a distribuição de plântulas de uma espécie, já podemos inferir que a limitação à dispersão de propágulos das plantas-mãe é importante. Se o modelo Poisson não homogêneo aproxima bem a distribuição dos adultos de uma espécie de árvore, podemos suspeitar de que a sobrevivência até a fase adulta é afetada por alguma característica do ambiente, que varia na área.

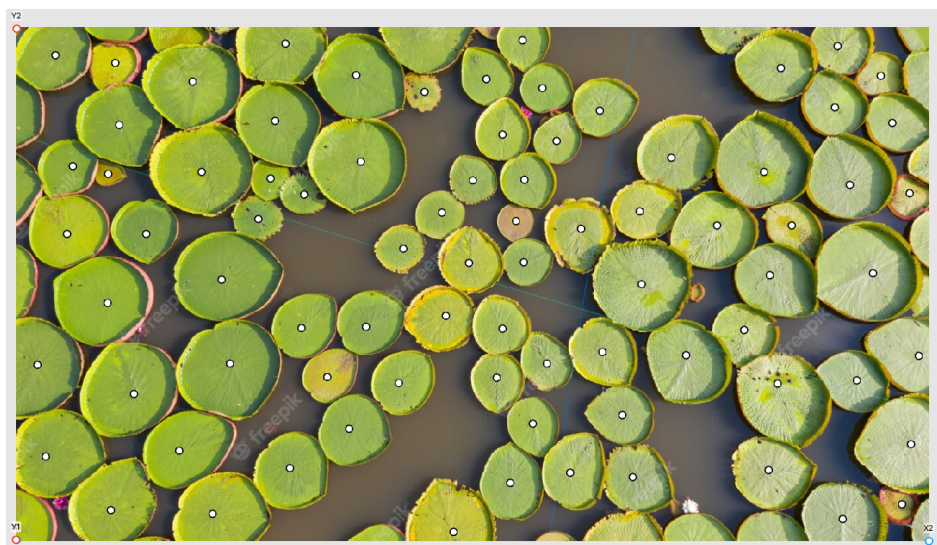
Mas como avaliar se um destes modelos aproxima bem um padrão observado? Neste roteiro, vamos apresentar uma das maneiras de fazer isso, com o modelo mais simples, que é o Poisson homogêneo. A ideia geral que vamos mostrar aqui vale para os demais modelos, mas deixaremos para mostrar isso em outro roteiro.

O modelo Poisson homogêneo é uma boa aproximação?

Será que as folhas flutuantes de Vitoria Régia (*Victoria amazonica*) da figura abaixo estão distribuídas ao acaso pela lâmina d'água?



Um dos jeitos de responder essa pergunta é representar a distribuição das folhas como um padrão de pontos no plano, para então comparar com o processo Poisson homogêneo. Aqui escolhemos como ponto de referência a inserção do pecíolo, que está marcada com um ponto branco em cada folha, na figura abaixo. Note que as folhas que têm este ponto fora dos limites da figura não são marcadas.

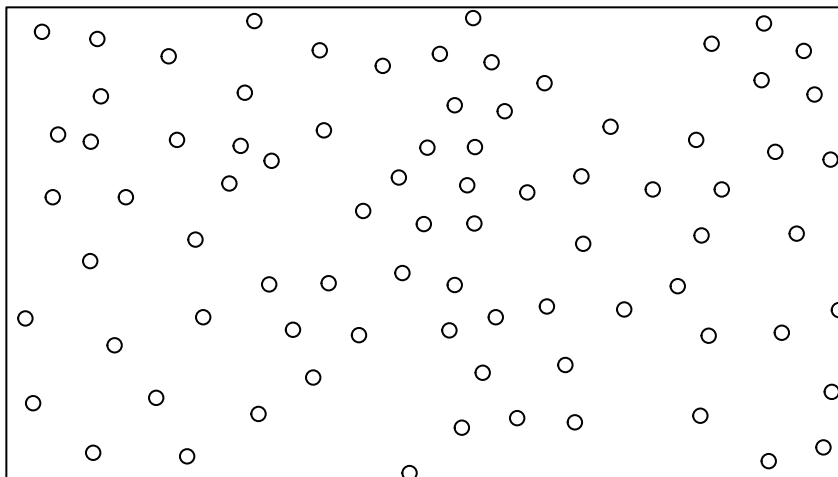


Comparando com padrões que o modelo simula

Use os comandos abaixo para carregar um arquivo com as coordenadas destes pontos, e gerar um gráfico do padrão de pontos:

```
library(spatstat) ## caso vc não tenha carregado o pacote
## carrega um objeto com as coordenadas dos pontos de inserção dos pecíolos de cada folha
vit <- read.csv("https://github.com/piLaboratory/BIE-0320/raw/main/data/vit_regia_coord_m.csv")
## Cria objeto com o padrão de pontos para analisar
## Define a área de observação
vit.w <- owin(xrange = c(0, 13.70323), yrange = c(0, 7.695469))
## Cria o padrão de pontos nesta área
vit.p <- ppp(x = vit$xm, y = vit$ym, window = vit.w)
## O objeto se chama vit.p para plotar
plot(vit.p, main = paste0("Vitoria regia, n = ", vit.p$n))
```

Vitoria regia, n = 81



Se a distribuição desses pontos nessa área for aleatória, uma simulação de um processo **Poisson homogêneo** deveria produzir padrões similares. Vamos então gerar algumas réplicas do processo Poisson homogêneo, com a mesma intensidade. Nosso melhor palpite da intensidade deste processo é a densidade de pontos que observamos na área. Obtemos este valor dividindo o número de pontos (81), pela área (105.45 m^2). Este cálculo já está feito no objeto do padrão de pontos que criamos anteriormente. Podemos ver isso quando pedimos um resumo numérico do padrão:

```
## Resumo do padrão de pontos  
summary(vit.p)
```

```
## Planar point pattern: 81 points  
## Average intensity 0.7681163 points per square unit  
##  
## Coordinates are given to 7 decimal places  
##  
## Window: rectangle = [0, 13.70323] x [0, 7.695469] units  
## Window area = 105.453 square units
```

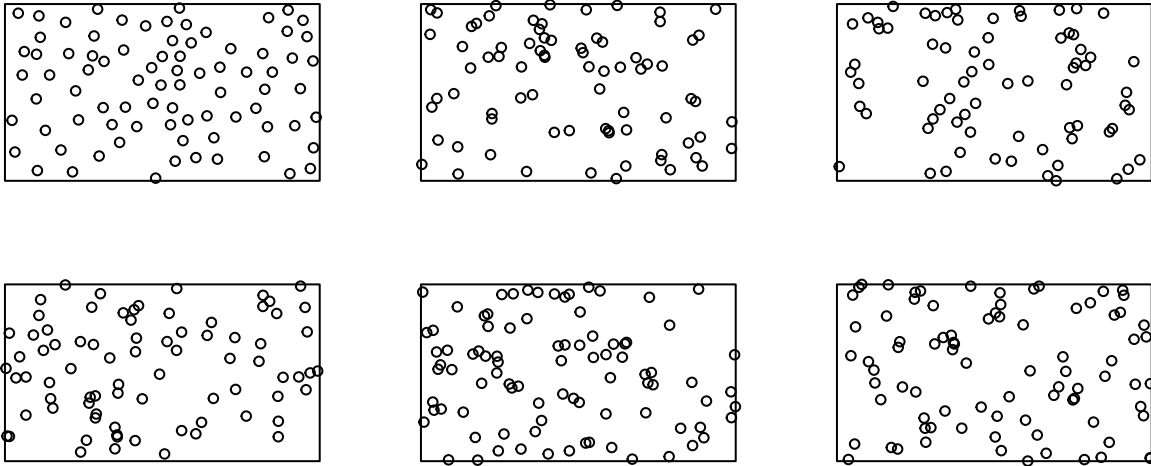
```
## E aqui guardamos o valor da intensidade em um objeto  
(vit.i <- summary(vit.p)$intensity)
```

```
## [1] 0.7681163
```

Vamos então simular alguns padrões Poisson homogêneos, de mesma intensidade do padrão observado. O comando abaixo faz isso, e plota estes padrões, junto com o padrão observado (primeiro painel):

```
## 5 simulações de Poisson Homogêneo  
vit.PH <- rpoispp(lambda = vit.i, win = vit.w, nsim = 5)  
## Junta o padrão observado
```

```
vit.PH$Observado <- vit.p
## Plota todos (o primeiro é o observado)
plot(rev(vit.PH), multiplot=TRUE, main = "")
```



Os padrões gerados pelo modelo são parecidos com o padrão observado? Se você está em dúvida, não é a única pessoa nessa situação. De fato é bastante difícil fazer esta avaliação apenas com a inspeção visual destes padrões. Precisamos de um critério mais rigoroso.

A função L de Ripley

Existem várias maneiras de caracterizar padrões de pontos por meio de números. Uma das maneiras mais informativas é usar o **número médio de vizinhos** de cada ponto. Este número pode ser calculado para diferentes distâncias. Assim, podemos nos perguntar sobre o número de vizinhos num raio de meio metro, um metro, dois metros, e assim por diante.

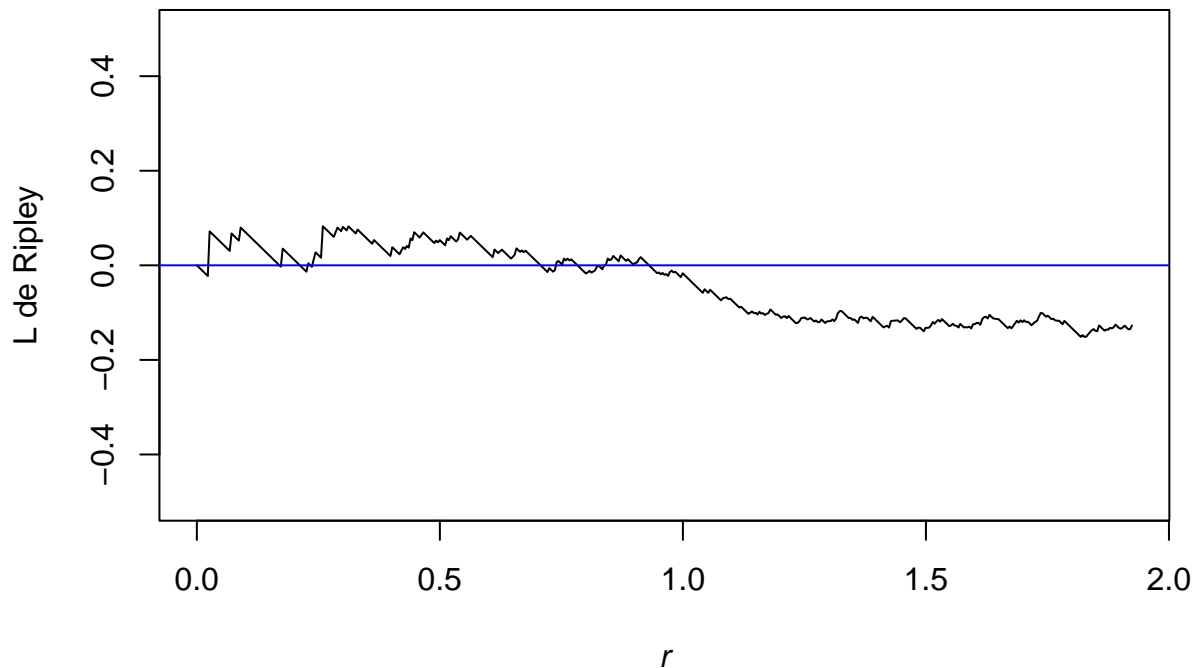
A função **L de Ripley** expressa a diferença entre este número de vizinhos a uma certa distância, e o número esperado por um processo Poisson homogêneo de mesma intensidade ¹. Se há mais vizinhos do que o esperado pelo processo Poisson Homogêneo, a função L terá valor positivo. Se há menos vizinhos que o esperado, a função L terá valor negativo.

Os comandos abaixo calculam a função L de Ripley para o primeiro padrão de pontos que simulamos com o processo Poisson Homogêneo, acima:

```
## calcula L de Ripley de um padrão Poisson Homogêneo
L.ph1 <- Lest(vit.PH[[1]])
```

¹Os detalhes do cálculo desta função estão explicados no Portal EcoVirtual. Para os propósitos deste roteiro, basta lembrar que o L de Ripley aqui expressa a congruência com um processo Poisson Homogêneo, em diferentes escalas.

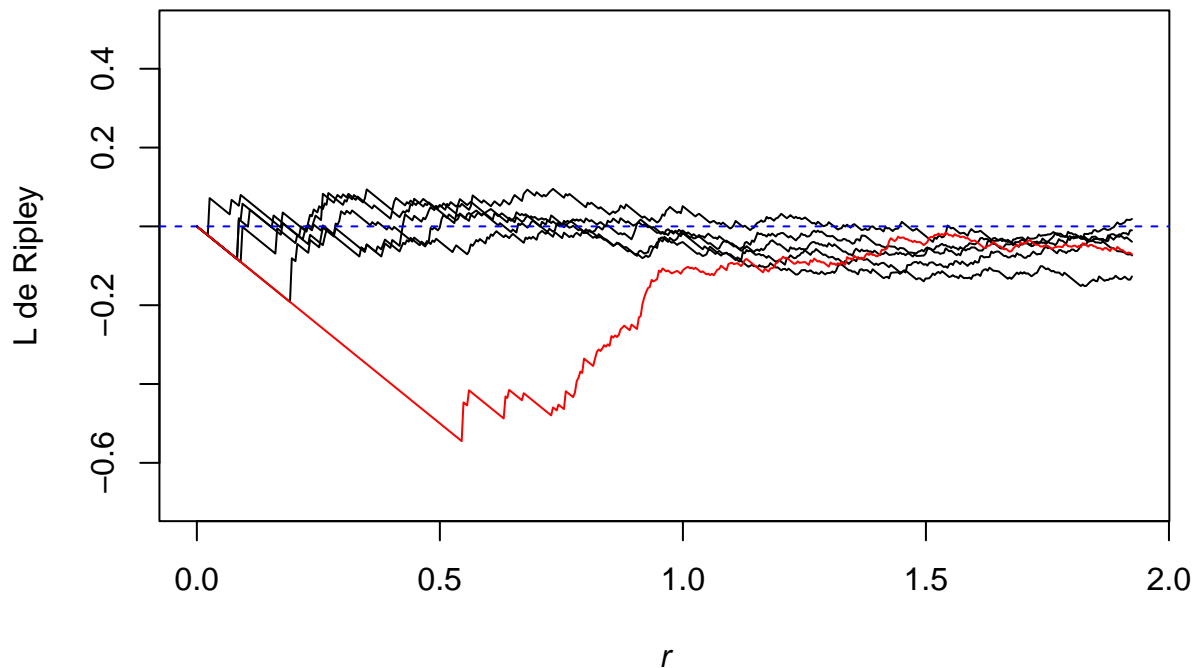
```
## Plota a função L em função do raio
plot(L.ph1, iso=r ~ r, ylim = c(-0.5,0.5), main="", ylab = "L de Ripley")
abline(h=0, col="blue")
```



O gráfico mostra o valor da função L no eixo y , e a distância no eixo x . Esta distância é o raio de um círculo cujo centro é cada indivíduo. Se o valor de L é zero para um certo raio, a média de vizinhos neste círculo de vizinhança é o esperado por um processo Poisson Homogêneo. Isso não impede que o valor de L seja maior ou menor que zero para outros raios. Portanto, dizemos que o L de Ripley expressa a distribuição espacial em diferentes escalas espaciais. Ou seja, é um **índice multi-escala**.

No gráfico acima, o valor de L parece estar bem próximo de zero para todos os raios, mas há alguma variação. Será que esta variação leva a algum desvio importante em alguma escala? Podemos ter uma ideia do quão comum (ou incomum) é esta variação plotando a função L para cada uma das 5 simulações do processo Poisson Homogêneo que fizemos, e também para o padrão de pontos observado para as vitórias-régias.

```
## Calcula o L de Ripley para os 5 padrões simulados
## E o observado
L.PH <- lapply(vit.PH, Lest)
## Plota as funções L em função do raio
plot(L.PH[[1]], iso=r ~ r, ylim = c(-0.7,0.5), main="", ylab = "L de Ripley")
for(i in 2:6)
  plot(L.PH[[i]], iso=r ~ r, add=TRUE, col=ifelse(i==6,"red","black"))
abline(h=0, col="blue", lty =2)
```



A linha vermelha é a função L de Ripley para o padrão de pontos das folhas de vitória régia. Os valores de L negativos para raios pequenos sugerem que há em média menos vizinhos em escalas pequenas do que o esperado por uma distribuição aleatória. Para ter mais certeza deste resultado, podemos fazer muitas mais simulações do processo Poisson homogêneo, e calcular as funções L para cada uma destas simulações. Se a linha para o padrão das vitórias-régias se mantiver discrepante de tantas simulações, teremos mais certeza de que há uma diferença em relação ao esperado, e para quais escalas.

Uma maneira simples de fazer isso é definir a faixa dentro da qual está a maioria dos valores de L gerados pelas simulações do processo Poisson homogêneo. Chamamos esta faixa de **envelope de confiança**. Em geral, usamos um envelope que contém os valores de L de 95% das simulações.

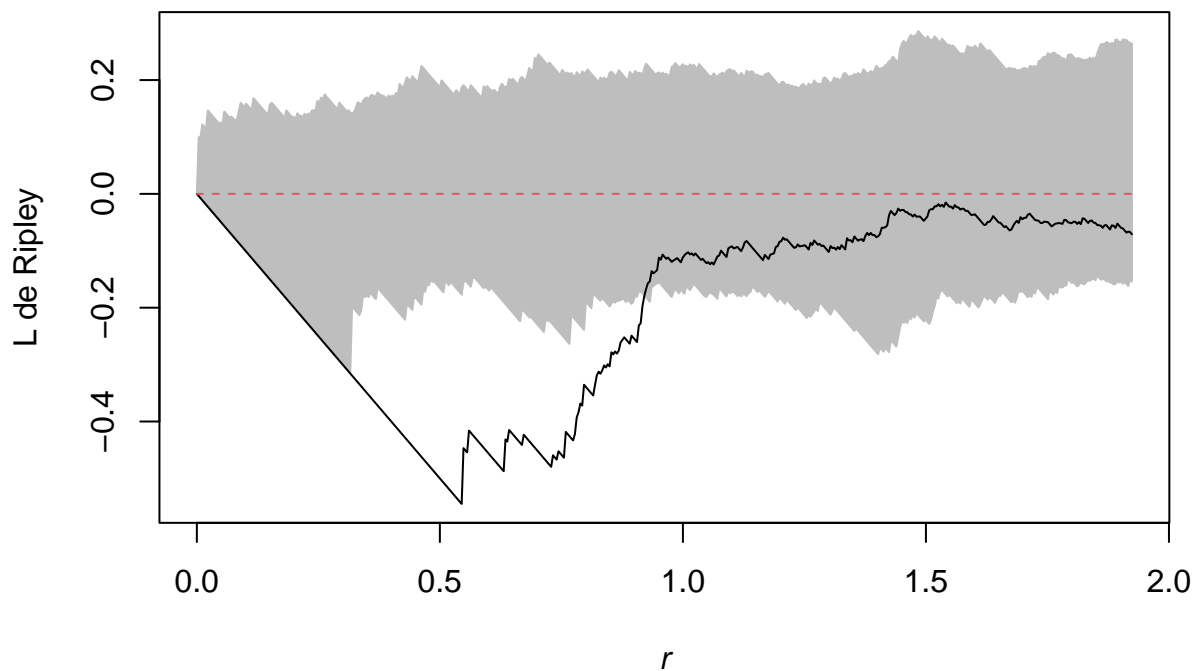
Os comandos abaixo calculam este envelope, usando mil simulações do processo Poisson homogêneo. O gráfico mostra a função L em função do raio de vizinhança para o padrão observado para as vitórias-régias em preto, e o envelope de 95% de confiança em cinza.

```
vit.env <- envelope(vit.p, fun = Lest, nsim = 1000)
```

```
## Generating 1000 simulations of CSR ...
## 1, 2, 3, .....10.....20.....30.....40.....50.....60.....
## .70.....80.....90.....100.....110.....120.....130.....
## ...140.....150.....160.....170.....180.....190.....200...
## ....210.....220.....230.....240.....250.....260.....270..
## .....280.....290.....300.....310.....320.....330.....340
## .....350.....360.....370.....380.....390.....400.....
## .410.....420.....430.....440.....450.....460.....470.....
## ...480.....490.....500.....510.....520.....530.....540...
## ....550.....560.....570.....580.....590.....600.....610..
## .....620.....630.....640.....650.....660.....670.....680
```

```
## .....690.....700.....710.....720.....730.....740.....
## .750.....760.....770.....780.....790.....800.....810.....
## ...820.....830.....840.....850.....860.....870.....880....
## .....890.....900.....910.....920.....930.....940.....950..
## .....960.....970.....980.....990..... 1000.
##
## Done.
```

```
plot(vit.env, . -r ~ r, ylab = "L de Ripley", legend=FALSE, main = "")
```



A função L para as vitórias-régias está fora do envelope, do raio zero até aproximadamente o raio de 1 metro em torno da inserção do pecíolo de cada folha. Portanto, podemos dizer, com 95% de chance de acerto, que há menos folhas vizinhas que o esperado por uma distribuição aleatória **nas menores escalas espaciais**. Já a partir da escala de um metro de raio, o número de folhas vizinhas é compatível com o esperado por uma distribuição aleatória.

Você consegue imaginar uma razão para este resultado?