Estrutura Espacial

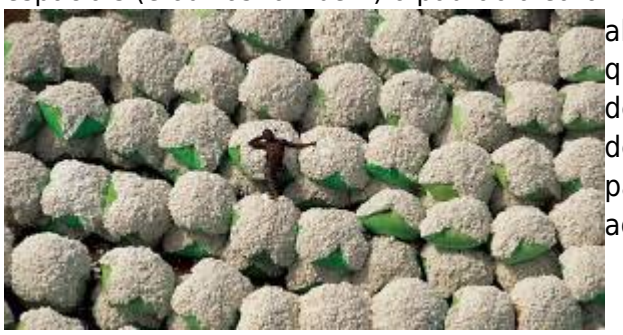
Nesse tutorial vamos tratar do reconhecimento de um dos padrões mais básicos de uma população de plantas: se os indivíduos estão espacialmente mais próximos ou mais afastados do que seria esperado se simplesmente fossem distribuídos ao acaso [1)].

Objetivo

Investigar o padrão espacial em populações de plantas e discutir quais processos subjacentes poderiam gerar os padrões observados. Antes de tudo, porém, precisamos definir alguns conceitos.

Contexto

Um padrão espacial é uma estrutura previsível que pode ser detectada e quantificada. Em geral, considera-se que um padrão é uma estrutura diferente do aleatório, entretanto, no caso dos padrões espaciais (e outros também) o padrão aleatório também pode ser considerado um padrão, afinal tem alguma previsibilidade [2)] e pode ser detectado e quantificado. Existem diversas métricas utilizadas para descrever a distribuição de indivíduos que são capazes de diferenciar, com maior ou menor eficiência, os três padrões espaciais básicos: aleatório, homogêneo e agregado.



Padrões Espaciais

- `aleatório`: a distribuição dos indivíduos não é diferente do que seria esperado por uma distribuição ao acaso;
- `regular` ou `homogêneo`: os indivíduos estão regularmente espaçados. É chamado também de padrão disperso, pois está relacionado ao maior distanciamento possível entre indivíduos;
- `agregado`: os indivíduos estão mais próximos do que esperado por um padrão aleatório, formando agrupamentos.

Detectar um padrão espacial pode ser importante tanto para entender os mecanismos que geram o padrão, como para decidir o método e a escala de amostragem e planejar o manejo de uma população. Algumas propriedades desejáveis de uma medida do padrão espacial são:

- diferenciar claramente o padrão;
- não ser afetada por: tamanho da amostra, densidade populacional ou pela variação no tamanho e na forma da amostra;
- ser estatisticamente tratável: passível de calcular a incerteza do valor e testar a diferenças entre amostras.

Para essa prática usaremos uma estimativa de aleatoriedade de pontos chamada K-Ripley (e sua padronização chamada L-Ripley). Primeiro iremos utilizar dados de distribuição simulados com diferentes padrões e em seguida utilizar a mesma técnica para detectar o padrão espacial em uma população natural.

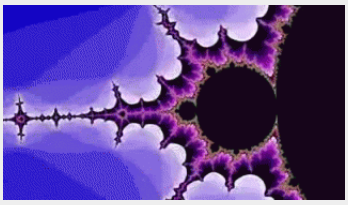
Roteiro

Padrões multiescala

Nesta prática vamos quantificar o padrão espacial usando um método multiescala. Os métodos de multiescala permitem, com uma única métrica, avaliar como o padrão espacial varia com a escala. Iremos descrever o padrão espacial para o conjunto total de indivíduos em uma população em uma área delimitada e iremos avaliar o padrão desde a escala da vizinhança dos indivíduos até a escala mais ampla da população.

Para a prática vamos utilizar um programinha chamado Programita, feito pelo pesquisador Thorsten Wiegand para quantificar o padrões espaciais usando medidas multiescala baseadas em distância entre pontos. Para baixar o manual do **Programita** clique

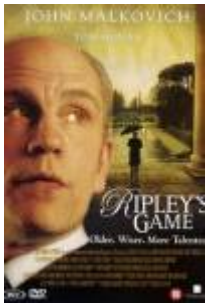
aqui.



No **Programita** existem várias medidas que podem ser usadas para calcular o padrão espacial, vamos usar duas delas: o **K de Ripley (na verdade, vamos usar sua padronização L-Ripley)** e o **O-ring**.

Ambas são abordagens baseadas em pontos, que utilizam o cálculo de distâncias ponto a ponto dentro de uma área delimitada. Essas medidas podem ser usadas para análises univariadas, ou seja, identificando o padrão para uma única classe de pontos, ou para análises bivariadas, que identifica o padrão entre dois tipos de pontos. As análises bivariadas podem ser usadas no contexto de populações para verificar se indivíduos de um dado estágio estão espacialmente associados a outro, ou no contexto de estruturação de comunidades para analisar se há atração ou repulsão na ocorrência de uma espécie em relação a outra.

K de Ripley



O K de Ripley é uma medida da densidade média ao redor de cada ponto. Para cada ponto na área de estudo é calculada a densidade no interior de um círculo de raio r centrado no ponto (área cinza da figura). Em seguida, calcula-se uma média desses valores obtidos para todos os pontos.

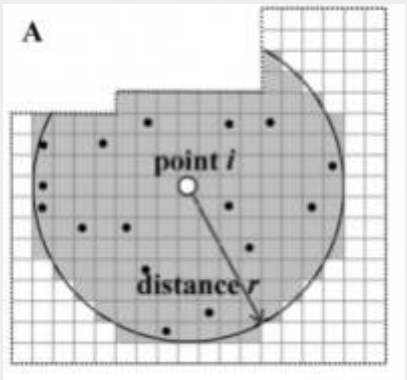


Figura: Implementação da estatística L de Ripley: contagem do número de pontos distantes de `i` no interior do círculo de raio `r`. Extraído de Wiegand & Moloney (2004).

A operação é repetida para diferentes valores de `r`, que permite avaliar de maneira contínua o valor de K para diferentes escalas.

$$ K_{(r)} = \frac{\sum_{i\neq{j}}^{i}I({d_{ij}<r})}{n}\frac{1}{\lambda}$$

Onde:

- d_{ij} é a distância do ponto i ao ponto j;
- $I({d_{ij}<r})$ função indicadora, sendo 1 se o ponto j está a uma distância menor que r do ponto i e 0 se o ponto j está fora desse raio r ao redor de i;
- n é o número de pontos total;
- λ é a densidade dos pontos.

A interpretação visual do K não é muito intuitiva por ser uma função cumulativa associada à área do círculo relativo a `r`. O L de Ripley, por sua vez, é a padronização deste valor:

$$ L_{(r)}= (\sqrt{\frac{K_{(r)}}{\pi}}-r) $$

Esta transformação faz com que o valor de L para **uma distribuição completamente aleatória seja sempre `0`**, para **uma distribuição agregada `L > 0`** e para **uma distribuição homogênea `L < 0`**.

O-ring (O(r))

A estatística **O-ring** é similar ao L de Ripley, mas baseada em um **anel**, ao invés de um círculo. É medida pela contagem do número de pontos em um anel de raio *r* e largura fixa. Da mesma forma que o L-Ripley,também são calculadas as intensidades para diferentes tamanhos de anel, mantendo a largura fixa.

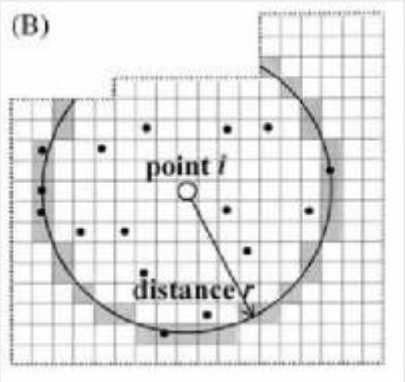


Figura: Implementação da estatística *O-ring*: contagem do número de pontos

distantes de *i* ao longo do raio *r*. Extraído de Wiegand & Moloney (2004).

Logo, definimos $O(r)$ como: $$ O_{(r)} = L_{(r)} - L_{(r-l)}$$

Onde:

- $r -l$: é o raio menos a largura do anel [3)]

Na completa aleatoriedade espacial (CAE) $O(r) = \lambda$ (intensidade do padrão), quando o padrão é agregado $O(r) > \lambda$ e quando o padrão é homogêneo $O(r) < \lambda$.

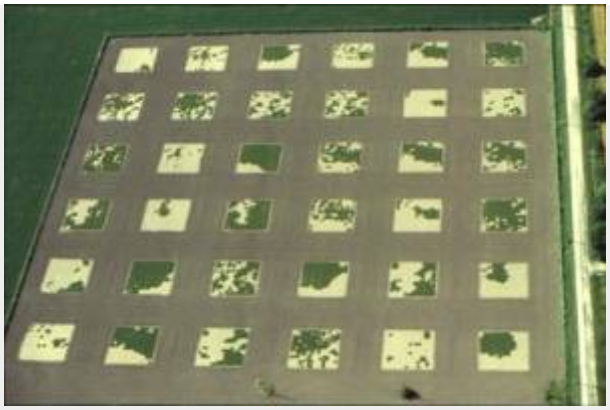
As medidas $K_{(r)}$, $L_{(r)}$ ou $O_{(r)}$ apresentam soluções analíticas teóricas para o padrão definido como processo Poisson ou Completa Aleatoriedade Espacial (CAE). Ou seja, quando a distribuições dos pontos no espaço estudado não é diferente do esperado pelo acaso. Para uma dada densidade de pontos conseguimos calcular esses valores teóricos para qualquer raio. Dessa forma, para interpretar o padrão espacial dos pontos observados precisamos:

- calcular os valores observados e o teóricos para CAE;
- comparar esses valores;
- definir quando uma diferença é ou não aceitável para afirmar que o padrão é diferente do aleatório;

Para os dois primeiros tópicos acima, usamos as fórmulas e calculamos os valores. Para tirar a subjetividade do terceiro, podemos calcular intervalos de confiança ou gerar envelopes[4)] de confiança por simulações computacionais, para definir objetivamente o que é algo diferente do esperado para a CAE.

Padrões de Pontos Simulados

Atividade 1:



Qual processo gerou o padrão de pontos?

Instruções gerais

- 1. baixe os arquivos relacionados ao padrão espacial 01 **OU** 02 (você escolhe) na mesma pasta em que o Programita esteja instalado. Caso abra uma página mostrando os dados, clique no link com o botão direito do mouse para salvar o arquivo. Salve no formato ".dat":

Dados para Análise Espacial

Padrão 1

- População (todos os indivíduos, sem distinguir adultos e jovens)
- Parentais (somente os adultos)
- Prole (somente os jovens)
- Bivariado (todos os indivíduos, distinguindo adultos e jovens)

Padrão 2

- População (todos os indivíduos, sem distinguir adultos e jovens)
- Parentais (somente os adultos)
- Prole (somente os jovens)
- Bivariado (todos os indivíduos, distinguindo adultos e jovens)

- caso não tenha o `programita` instalado, baixe o programita aqui na mesma pasta do arquivo de dados;
- descompacte o arquivo *programita.zip*;
- clique 2x para abrir o arquivo executável `ProgramitaJulio2006.exe`.

Bem vindo(a) ao **Programita**! Agora vamos abrir os dados que iremos trabalhar.

O **Programita** aceita arquivos de texto das extensões .dat e .asc. São arquivos em formato de texto, separados por tabulação (ou espaço). Os arquivos de dados possuem a seguinte estrutura:

A primeira linha contém informações gerais sobre o arquivo de dados:

- valor mínimo da coordenada x;
- valor máximo da coordenada x;
- valor mínimo da coordenada y;
- valor máximo da coordenada y; e
- número total de indivíduos

A partir da segunda linha, estão os dados dos pontos que serão analisados:

- primeira coluna com as coordenadas x dos indivíduos;
- segunda coluna com as coordenadas y dos indivíduos;
- no caso de dados univariados, a terceira coluna será sempre 1 e a quarta coluna sempre 0.
- no caso de dados bivariados a terceira coluna tem os pontos dos indivíduos tipo A (adultos, por exemplo) identificados por 1 e do tipo B (jovens, por exemplo) identificados por 0 ;
- ainda no caso de dados bivariados, a quarta coluna tem os pontos dos indivíduos do tipo A identificados por 0 e do tipo B identificados por 1 .

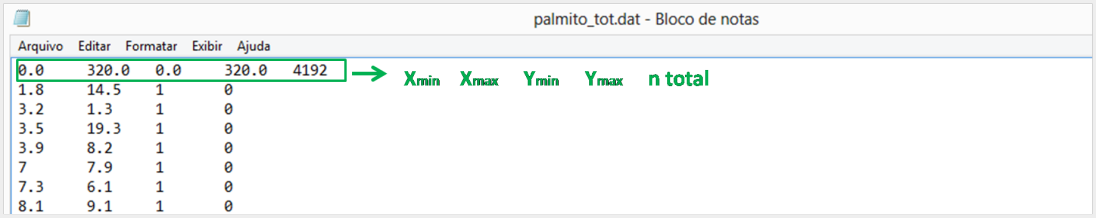


Fig. Exemplo de arquivo .dat no formato de uso no *Programita*.

Padrão Univariado: todos os pontos

- 1. Verifique se na janela *Input data file* estão aparecendo os arquivos .dat. Caso não esteja, **verifique se o arquivo executável do programita está na mesma pasta dos arquivos *.dat***.

Dependendo da configuração do seu navegador o arquivo salvo pode aparecer com uma extensão diferente (p.ex. ".bin"). Nesse caso é preciso mudar a extensão do arquivo para ".dat".

- 2. no menu à esquerda selecione o arquivo **padrao"0X"all.dat**. No caso **X** vai ser 1 ou 2 dependendo da sua escolha;

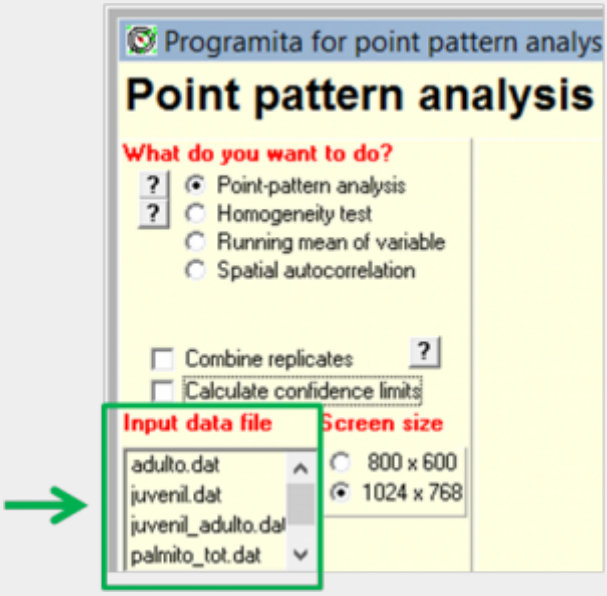
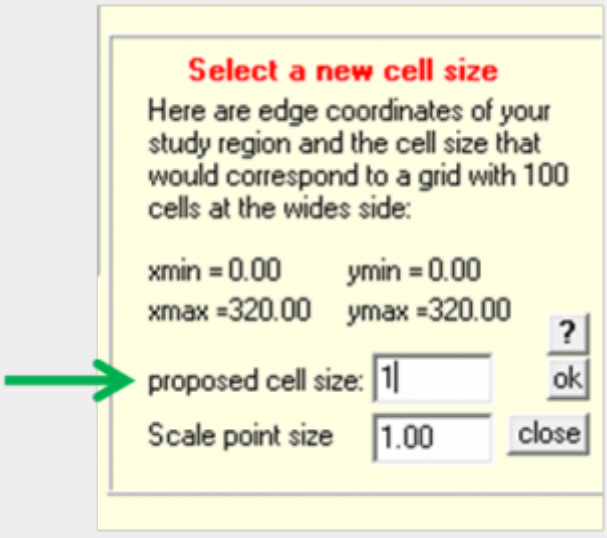


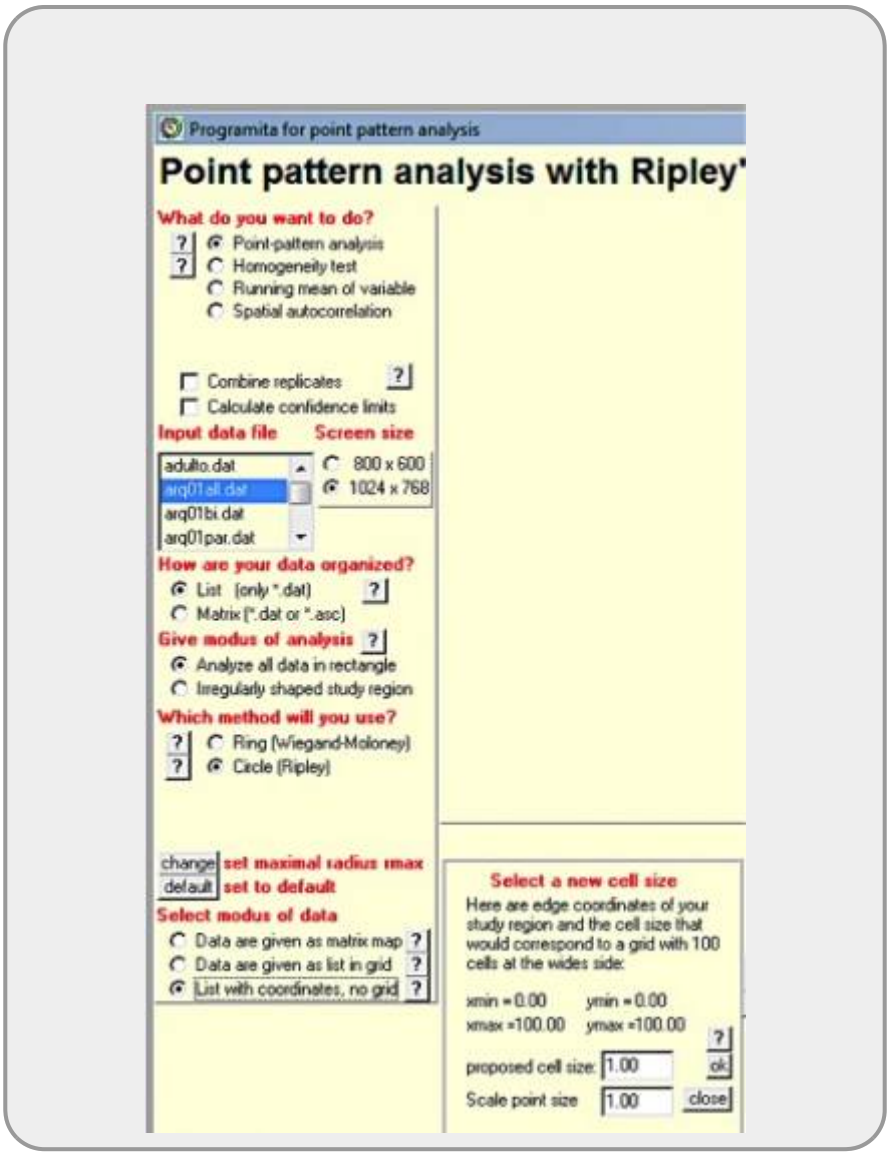
Figura. Janela de entrada de dados do *Programita*.

- 3. Em ***How your data are organized*** selecione ***List***
- 4. Vamos começar usando o L de Ripley então em ***Which method to use*** selecione ***Circle ou Ripley*** *(a depender da versão que foi baixada)*
- 5. Em ***Select modus of data*** selecione ***List with coordinates no grid***. Ao selecionar esta opção aparecerá uma janela com a opção ***Select a new cell size***:



.

- 6. Caso tenha menos de 500 pontos, altere o ***proposed cell size*** para 1. Caso contrário deixe no padrão do programa.
- 7. Feito tudo isso, você deve estar assim:



- 8. Você pode agora respirar fundo e apertar o botão ***Calculate index***;

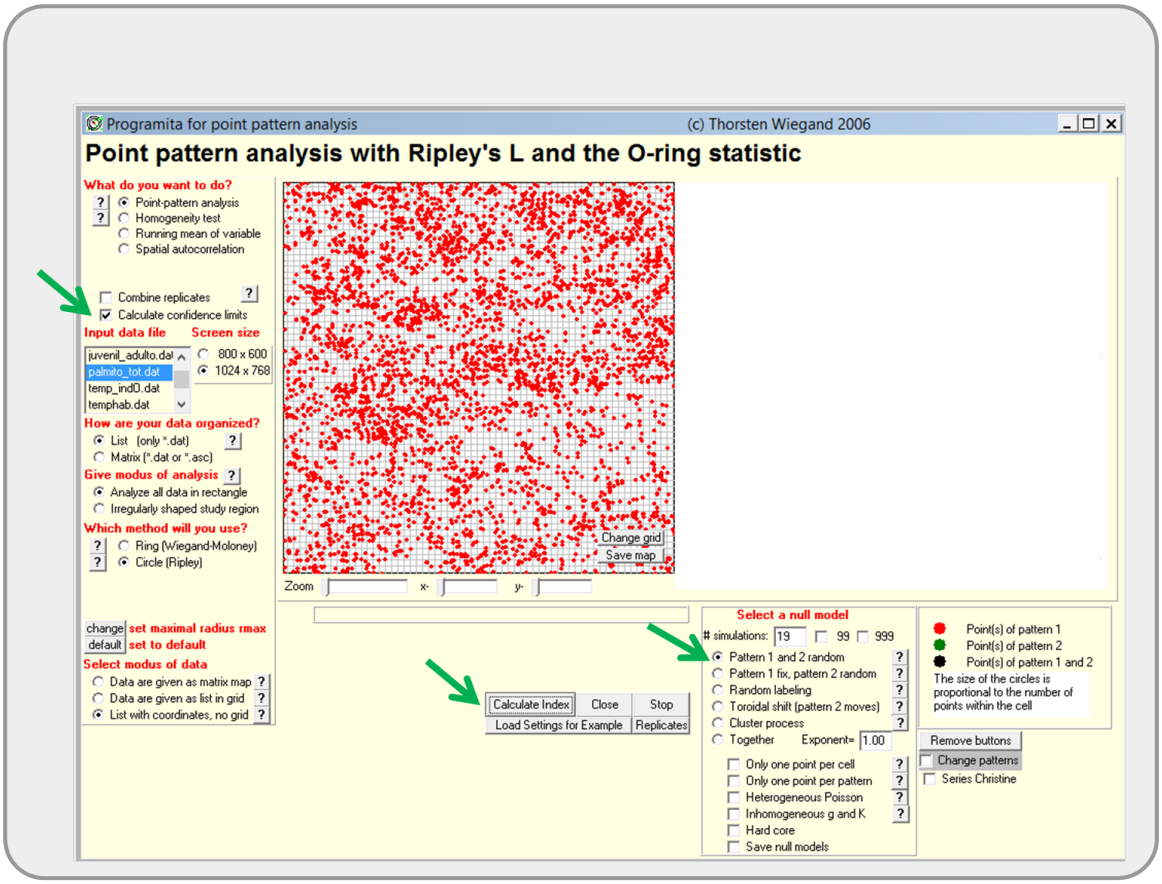
A saída visual do programa é um mapa onde os indivíduos aparecem em pontos vermelhos, seguindo as coordenadas do arquivo de dados. O gráfico no canto superior direito corresponde ao valor do L-Ripley para diferentes raios. Nessa saída gráfica é possível analisar como o padrão espacial varia de acordo com a escala. Para dados univariados, ignore o gráfico inferior.

Porém, olharmos apenas o formato da curva não é suficiente para afirmamos em que escalas a população é agregada. Para isso precisamos comparar o resultado observado com o padrão que seria gerado pela distribuição dos pontos completamente aleatório. Esse modelo nulo é chamado de ***completa aleatoriedade espacial***. Para gerar esse modelo por simulação é necessário recolocar o mesmo número de pontos de forma aleatória na mesma área. Se fizermos isso, muitas e muitas vezes, é possível gerar um envelope de confiança (similar ao intervalo de confiança) no qual o padrão de distribuição aleatória é encontrado. Se os valores observados estão contidos dentro do envelope podemos concluir que nosso padrão não é diferente do aleatório.

Para fazer isso você deve:

- 9. selecionar a opção ***Calculate confidence limits*** e;

- 10. na janela ***Select a null model*** selecionar o modelo nulo ***Pattern 1 and 2 random***;
- 11. verifique se sua tela está como a figura e clique novamente no botão ***Calculate index***.



Caso a simulação esteja demorando muito

- aperte o botão de *stop* ao lado do *Calculate index*;
- selecione outro *"modus of data"* e em seguida selecione novamente *list with coordenate,...*;
- na janela *Select a new cell size*, altere *proposed cell size* para 2;
- na janela *Select a null model* altere *# simulations* para 20;
- aperte novamente o botão *Calculate index*;

Descreva o padrão observado

O *Programita* permite acompanhar graficamente a simulação ao longo do tempo. É possível observar que a cada simulação é gerada uma distribuição aleatória dos indivíduos e recalculado os valores de L-Ripley. Ao final é gerado o gráfico com os valores observados a partir do arquivo de dados, acompanhado do envelope de confiança gerado a partir da simulação de completa aleatoriedade espacial. Valores fora do intervalo de confiança indicam a existência de um

padrão espacial diferente do aleatório.

Dica: Faça um ***Print Screen*** dos seus resultados para salvar o gráfico de cada análise que fizer ao longo da prática.

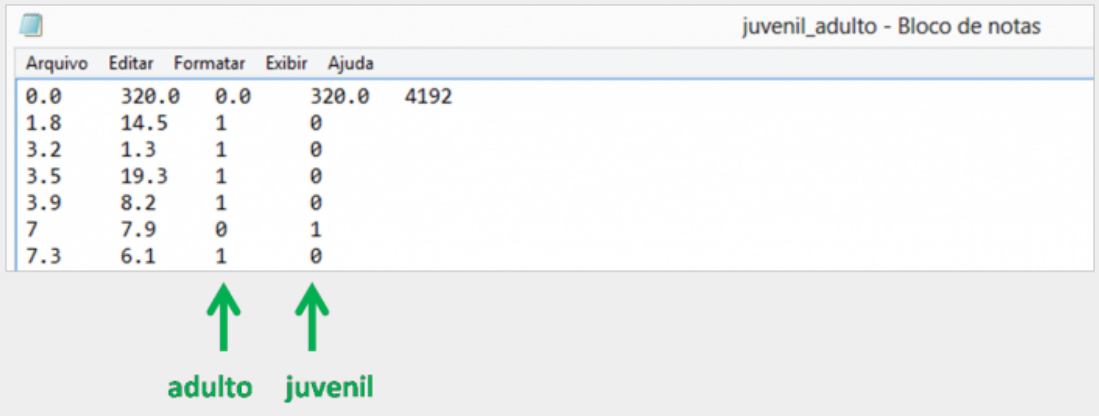
- 12. Faça o mesmo procedimento, porém em ***Which method to use*** selecione ***Ring***
- 13. Compare os resultados entre o L-Ripley e o O-Ring.

Atividade

- repita a análise com L-Ripley e O-Ring para os arquivos com:
  - os pontos dos adultos (parentais): *padrao"0X"par.dat* e;
  - os pontos dos jovens (prole): *padrao"0X"prole.dat*;
- interprete o resultado para cada tipo de ponto;

Padrão Bivariado: dois tipos de pontos

O *Programita* permite a análise de padrão de pontos de uma classe em relação a outra (por exemplo juvenis em relação a adultos). Para isso é necessário diferenciar os pontos no arquivo de dados, utilizando 0 ou 1 nas colunas 3 e 4, como mostra a figura abaixo, em um arquivo que distinguia indivíduos adultos de juvenis:



Vamos agora analisar o padrão dos pontos associados (PROLE) em relação aos parentais (PAR), seguindo o mesmo procedimento anterior.

- 1. selecione o arquivo com a separação de classes de pontos parentais e associados: *padrao“0X”bi.dat*;
- 2. em ***What do you want to do*** selecione a opção ***Point-pattern analysis***
- 3. em ***How your data are organized*** selecione ***List***
- 4. neste caso, estamos interessados na análise do padrão em escala cumulativa para entender até que distância há agregação, por isso, em *Which method to use* selecione **Circle** ou **Ripley** (a depender da versão que estiver usando)
- 5. em ***Select modus of data*** selecione ***List with coordinates no grid***
- 6. para testarmos se existe agregação dos pontos PROLE em relação ao PAR , utilizaremos o envelope de confiança. Selecione a opção ***Calculate confidence limits*** e selecione o modelo nulo ***Pattern 1 fix, 2 random***.
- 7. rode a análise apertando: ***Calculate index***
- 8. interprete os resultados. **Obs.: o gráfico que mostra o padrão de associação é o inferior, denominado “Bivariate L-function(Ripley)”. O gráfico superior é o mesmo que o gráfico do padrão tipo A univariado (no nosso caso, o padrão dos adultos), com pequenas diferenças nos limites do eixo Y.**

Descubra o algoritmo

Algoritmo é uma sequência de passos para executar uma tarefa. Os pontos dos arquivos de dados foram gerados por um algoritmo muito simples em duas fases: primeiro foram gerados os pontos parentais e em seguida os pontos associados (prole). Descreva uma sequencia de tarefas [5)] que seria capaz de gerar a distribuição de pontos (incluindo ambas classes de pontos) que você observou a partir do seu arquivo de dados.

Distribuição Espacial de Palmitos na Restinga

O Palmiteiro (*Euterpe edulis* Mart.) é uma espécie muito característica das florestas atlânticas e costuma ocorrer com densidades altas em áreas mais preservadas. Vamos agora analisar os dados referentes a uma população de palmitos que ocorre em uma parcela de floresta de Restinga na Ilha do Cardoso, Cananéia -SP. Os dados foram coletados nos anos de 2009/2010 em uma área de 10,24ha (320m x 320m).

Preparamos três arquivos no formato lido pelo *Programita*:

1. dados de indivíduos juvenis (diâmetro do tronco entre 1 e 5 cm): juvenil.dat
2. dados de indivíduos adultos (diâmetro do tronco > 5 cm): adulto.dat
3. juvenis e adultos (padrão 1 adulto, padrão 2 juvenil): juvenil_adulto.dat

Utilize as ferramentas disponíveis no *Programita* para descrever os padrões espaciais:

- da população total de palmito;
- apenas dos juvenis e;
- apenas dos adultos.

Investigue se a distribuição dos juvenis está associada a dos adultos.

Padrões & Processos Junte-se em um grupo de 2 a 4 alunos e discuta quais possíveis processos poderiam gerar os padrões descritos.

[1] ou seja, a localização de um indivíduo não melhora a predição de onde outros indivíduos podem estar

[2] por exemplo, em relação ao número médio de indivíduos

3)

igual ao raio interno do anel

4)

equivalente a intervalo de confiança obtido por simulação numérica

5)

p.ex: gerar 10 valores de x a partir de uma distribuição aleatória uniforme de 0 a 100; gerar valores de uma sequência de 10 a 90 a cada intervalo de 5 como o y....

From:
http://ecovirtual.ib.usp.br/ -

Permanent link:
http://ecovirtual.ib.usp.br/doku.php?id=ecovirt:roteiro:pad_spat

Last update: **2022/09/27 12:40**