

Índice de Dispersão

Vamos testar se existe agregação na população de palmito usando a razão variância/média, que é uma medida baseada em contagens em parcelas. Esta razão é também conhecida como **índice de dispersão**, e é definida como:

$$I_d = \frac{s^2}{\bar{x}}$$

Esse índice está baseado em uma propriedade da distribuição probabilística¹⁾ de contagens em unidades amostrais: quando a média e a variância são iguais, ou seja, quando o índice é igual a 1, a distribuição é aleatória. No caso dos indivíduos estarem distribuídos de maneira completamente uniforme, a variância tenderia a zero e o índice também. Quando estão agregados, a variância seria maior que a média e o índice de dispersão maior que um.

Nesses exercícios, vamos usar dados de abundância de palmitos-juçara na floresta de restinga. Avaliando o trabalho antes de ir a campo, tomamos algumas decisões que definem muito como os resultados serão interpretados:

- a nossa amostra será feita aleatoriamente em uma área de 10 hectares de restinga
- o total de amostra que podemos fazer, dadas limitações de tempo e dinheiro, é de 16.000 m²
- cada unidade amostral será uma parcela de 40 x 40 m (1.600 m²)
- as parcelas foram distribuídas ao acaso pela área e o número de palmitos em cada uma foi registrado.

Atividade 1:



A população de palmito é agregada?

- Abra a pasta **DadosAlunos** na área de trabalho da máquina virtual em uso e abra o arquivo com seu nome. Cada arquivo contém os dados da contagem de indivíduos de palmito nas subparcelas. Na primeira aba da planilha você terá uma amostra de 10 subparcelas de 40 x 40 m, tomadas ao acaso na área da parcela.

ATENÇÃO: Você precisa abrir o arquivo na máquina virtual pois lá o excel já está com o Resampling instalado.

Ao abrir o Excel você verá os seguintes botões: 

Caso não visualize, entre no menu *Ferramentas* e depois em *Personalizar*. Na aba *Barras de Ferramentas* selecione a

caixa *Resampling Stats in Excel.*

Calculando índice de dispersão para o palmito

1. calcule a média utilizando a função `MÉDIA()` no Excell
2. calcule a variância utilizando a função `VAR()`
3. calcule o I_d ;



Interprete o resultado e responda à pergunta formulada nesta seção.

Uma questão que pode ser colocada nesse momento é: o valor observado é de fato diferente do que seria esperado caso o Palmito estivesse distribuído aleatoriamente na parcela?

Lembre-se que qualquer configuração é possível de ser gerada pelo acaso, só que umas são mais comuns e outras são mais raras. A configuração da nossa amostra é possível de ser gerada pelo acaso, resta saber se a probabilidade dela acontecer pelo acaso é pequena ou grande. Isso nos conduz a uma segunda pergunta, relativa à nossa metodologia:

Pergunta:

Qual a probabilidade do acaso gerar um valor como esse ou maior?

Testando o índice de dispersão

Para responder a essa pergunta nós vamos gerar amostras em que as unidades amostrais (parcelas fictícias) tenham a mesma média da nossa amostra real, mas que palmeiras se distribuam ao acaso pelas parcelas. O arranjo de números, ou configuração obtida com esse procedimento é chamado de **modelo nulo**.

Em nosso caso, queremos um modelo nulo de *completa aleatoriedade espacial*. Para cada amostra gerada, vamos calcular um índice de dispersão. Se o modelo nulo gerar valores dos índices de dispersão iguais ou maiores que o observado com muita frequência, não teremos segurança para dizer que existe algum padrão diferente do aleatório nos dados reais. Mas qual é esse limiar de frequência de valores que nos deixa desconfiados? É usual na área de biologia utilizar 5% para esse limiar, ou seja, se menos de 5% dos valores gerados são maiores ou iguais ao índice real observado, dizemos que o valor é significativamente diferente do aleatório.

Vamos fazer o teste usando um *plugin* do Excel chamado Resampling.

1. calcule o número total de indivíduos da nossa amostra (use a função SOMA())
2. atribua para cada indivíduo a posição de uma parcela ao acaso
 - clique no símbolo **R** ²⁾
 - na janela que abre, complete:
 - Input Range: as células com o nome das parcelas
 - Top Cell of Output Range: a posição onde quer retornar o resultado
 - Number of Cells in Output Range: o valor do total de indivíduos da amostra



Se tudo correu bem, teremos uma nova coluna na planilha com o rótulo da parcela sorteada para cada indivíduo. Agora precisamos contar quantas vezes cada parcela ocorre nessa coluna, o que representa o número de indivíduos no nosso cenário nulo de “Completa Aleatoriedade Espacial”:

1. copie os nomes das parcelas da coluna A e coloque ao lado da coluna onde acabou de gerar as parcelas para cada indivíduo
2. use a função CONT.SE() para contar quantas vezes aparece a primeira parcela na nossa simulação, veja a figura abaixo para ver como fazer
3. puxe a formula para outras linhas para contar as outras parcelas
4. calcule o índice de dispersão conforme fizemos anteriormente



Dica: Coloque o símbolo \$ entre os valores da coluna e linha no intervalo do CONT.SE(), isso fixa o intervalo para usar a mesma função nas outras linhas.

Pronto! Calculamos o índice para o cenário de “Completa Aleatoriedade Espacial”. Agora só falta repetir isso muitas vezes e guardar o valor do índice resultante a cada vez. Que tal mil vezes? Parece bastante entediante ficar repetindo o mesmo procedimento tantas vezes... Vamos fazer o Resampling fazer isso por nós!

Só precisamos repetir a re-amostragem e depois pedir para ele repetir o procedimento. Refaça os passos:

- clique no símbolo **R** ³⁾
- na janela que abre, complete:
 - Input Range: as células com o nome das parcelas

Dica: Para selecionar todas as células com o nome das parcelas de uma só vez aperte:

“Ctrl + Shift + ↓”

- Top Cell of Output Range: a posição onde quer retornar o resultado
- Number of Cells in Output Range: o valor do total de indivíduos da amostra

Veja como o valor do índice mudou! Agora selecione a célula com o valor do Id e clique no botão **RS**⁴⁾. Na janela que se abre modifique apenas o campo **Iterations** para 1000. Se tudo correr bem, você verá o Resampling trabalhando, ao final responda sim para ver os valores que foram calculados nas 1000 simulações. Coloque esses resultados em ordem decrescente.



**Dica:** Altere o nome da planilha **Results** para guardar o resultado da simulação para a planilha de parcelas de 40 x 40m.

Responda:

- Quantas vezes apareceram valores simulados iguais ou maiores do que o Id observado?
- Qual a probabilidade do valor observado ter vindo de uma população com distribuição aleatória?
- Responda a pergunta inicial, agora incorporando a sua simulação de “Completa Aleatoriedade Espacial”

A escala de amostragem

Na década de 70 houve um debate muito acalorado e importante para a conservação chamado **SLOSS** (Single Large or Several Small). Em função das limitações de recursos para estabelecer Unidades de Conservação, a grande questão em jogo era decidir entre ter poucas reservas grandes ou muitas pequenas.

Aqui temos uma questão parecida. Quando iniciamos o trabalho, tomamos⁵⁾ a decisão de fazer parcelas de 40 x 40 m. O que aconteceria se fizéssemos parcelas menores? O resultado poderia mudar? Se mudar, o que isso significa?

Vamos responder essas perguntas agora. Refaça os procedimentos anteriores para parcelas de 20 x 20 m e 10 x 10 m, mantendo o mesmo esforço amostral ($\$16.000 \text{ m}^2$). Já fizemos essas amostras para você! Estão nas outras abas da planilha que baixou.

 **Dica:** Altere o novamente o nome da planilha **Results** para guardar o resultado da simulação para a planilha de parcelas de 20 x 20 m e 10 x 10 m.

Próximos passos

Siga para a segunda parte do roteiro:

[Padrões multiescala](#)

1)

http://en.wikipedia.org/wiki/Probability_distribution

2) 3)

Resample WITH Replacement

4)

Repeat and Score

5)

os professores!

From:

<http://labtrop.ib.usp.br/> - **Laboratório de Ecologia de Florestas Tropicais**

Permanent link:

<http://labtrop.ib.usp.br/doku.php?id=cursos:popcom:2015:roteiros:ep1>



Last update: **2021/07/20 12:43**