Biologia Quantitativa Dispersão de Organismos na Paisagem: Poisson, Uniforme, Binomial

Depto de Zoologia
16 de março de 2021

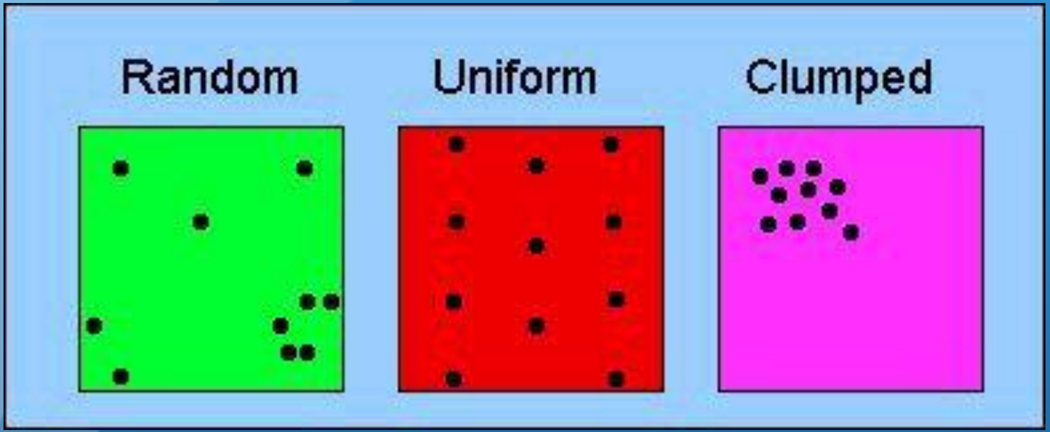
Roteiro da Aula

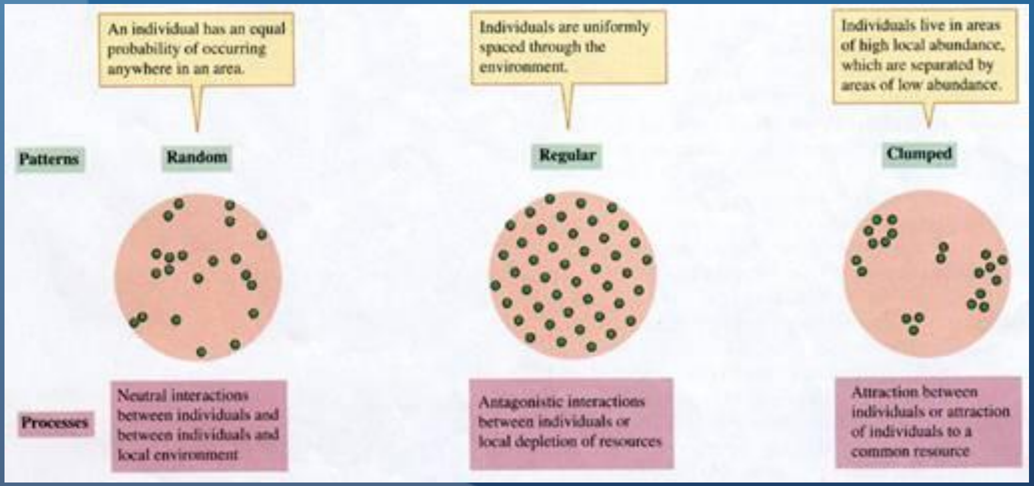
- Andrade e Ogliari sec 4.4.3 e 4.5 Poisson
- Andrade e Ogliari 4.4.2 Binomial
- Vieira Bioestatística cap 9 Binomial
- Sokal e Rohlf cap 5.2 (binomial) 5.3 (poisson)
- Ludwig & Reynolds Statistical Ecology cap 3

Tipos de dispersão

- Aleatória – a presença do organismo é independente da presença dos demais
- Uniforme – o espaçamento entre organismos é maior do que seria esperado ao acaso – interação negativa
- Agregada – os organismos tem distribuição mais agregada do que ao acaso – interações positivas

Distribuição Espacial





Distribuições e tipo de dispersão

- Poisson – aleatória – média igual à variância
- Binomial negativa – média menor que a variância – agregada
- Binomial positiva – média maior que a variância – uniforme

- Um teste prático para estimar agregação é simplesmente calcular a razão média/variância
- Entretanto isto não quer dizer que a distribuição seja uma das acima
- O teste mais aceito hoje é o índice de morisita, que representa um estimador não tendencioso da distribuição estatística

Médias e Variâncias

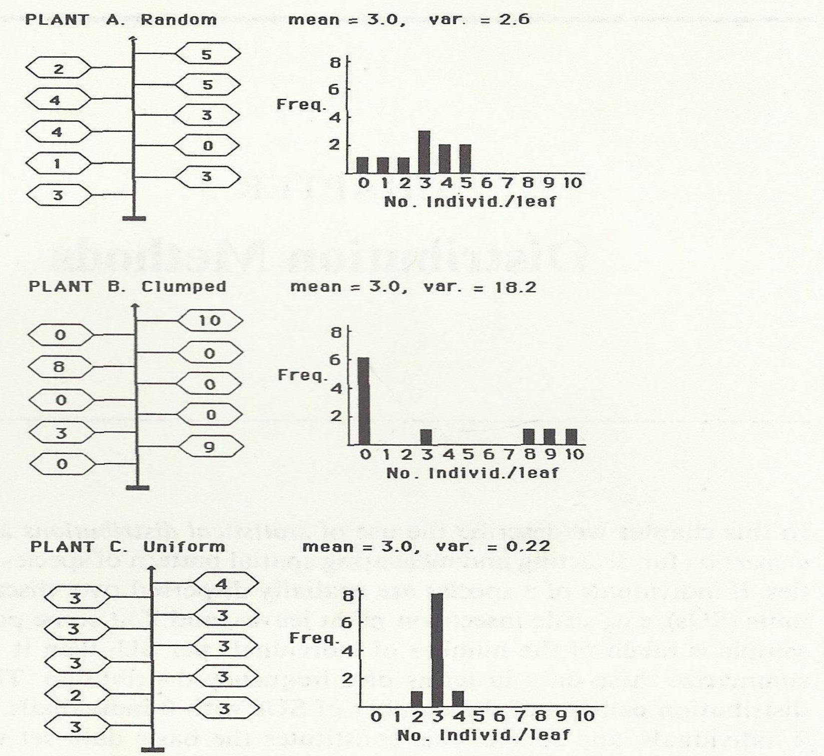


Figure 3.1 *Possible dispersals for 30 scale insects on 10 leaves (natural SUs) showing number of insects found on each leaf. The mean and variance for the number of insects per leaf for plants A (random), B (clumped), and C (uniform) is given.*

Distribuição Poisson

- Usada para contagem de indivíduos por intervalo de tempo, comprimento, área ou volume.
- Valores por unidade tem de ser baixos para poder usar (escala)
- Cada unidade tem igual probabilidade de ocorrencia de individuos
- A ocorrência de indivíduos não afeta a probabilidade de presença de outros
- Cada unidade está igualmente acessível
- Hipótese: número de indivíduos por amostra é de distribuição poisson. Se Ho for rejeitada, sabemos apenas que não é aleatória, mas não a qual pertence

Distribuição Poisson

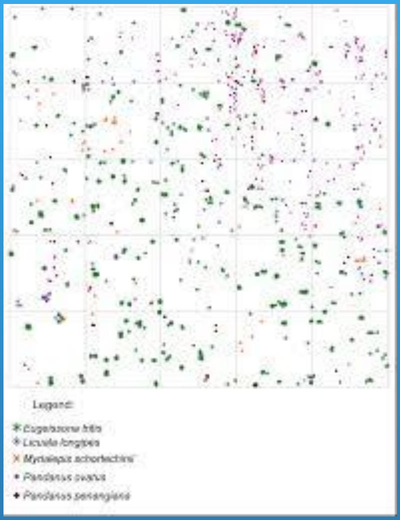
Pode esse padrão de dispersão aleatória ser descrito matematicamente? A resposta é sim, e a descrição é feita através de um modelo, cuja função de probabilidade, ou seja, a probabilidade de encontrar x indivíduos por quadrante, é dada por:

$$\boxed{Prob(X = x) = \frac{e^{-\lambda} \times \lambda^{x}}{x!}} \quad (4.7)$$

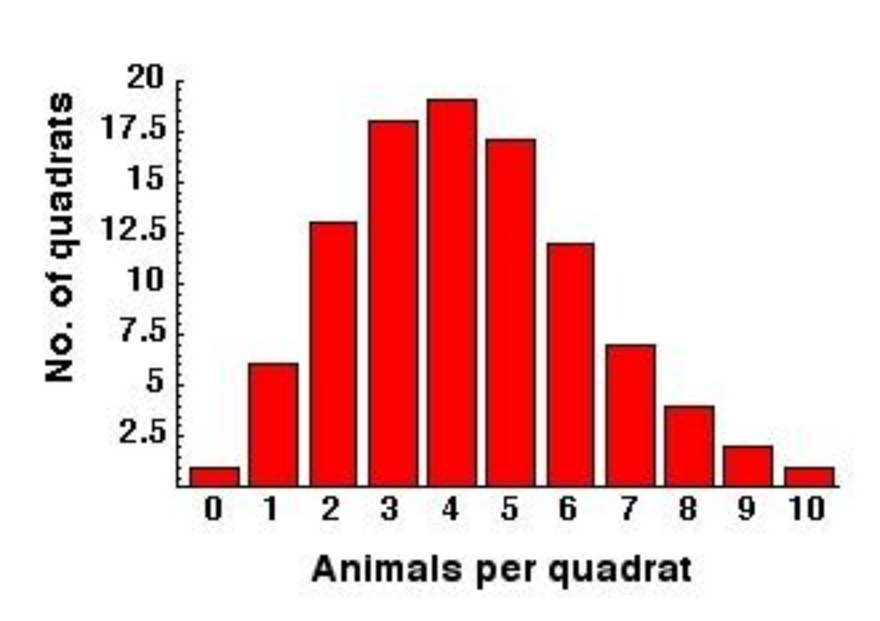
com $x = 0,1,2,3,\ldots$, onde e é o número de Euler e vale 2,718282 e $\lambda > 0$ é o parâmetro da distribuição e corresponde ao número médio ou esperado de

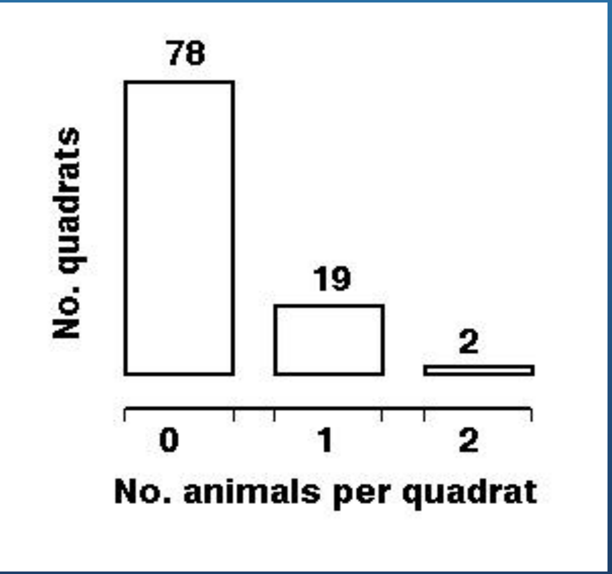
[2] O nome desta distribuição está associado ao matemático francês Siméon Denis Poisson (1781-1840).

Distribuição Poisson



Abaixo esq: poisson média 4.3
Abaixo dir: poisson média 0,25 n=25 100 quadrats
Dir: distribuição palmeiras em Penang, Malásia





Distribuição Binomial

4.5 Aproximação da distribuição binomial pela distribuição de Poisson

O modelo de Poisson pode ser considerado como limite da distribuição binomial, isto é, para valores de n **grande** (fazendo-se n cada vez maior) e π **pequeno** (fazendo-se π cada vez menor), verifica-se a seguinte aproximação:

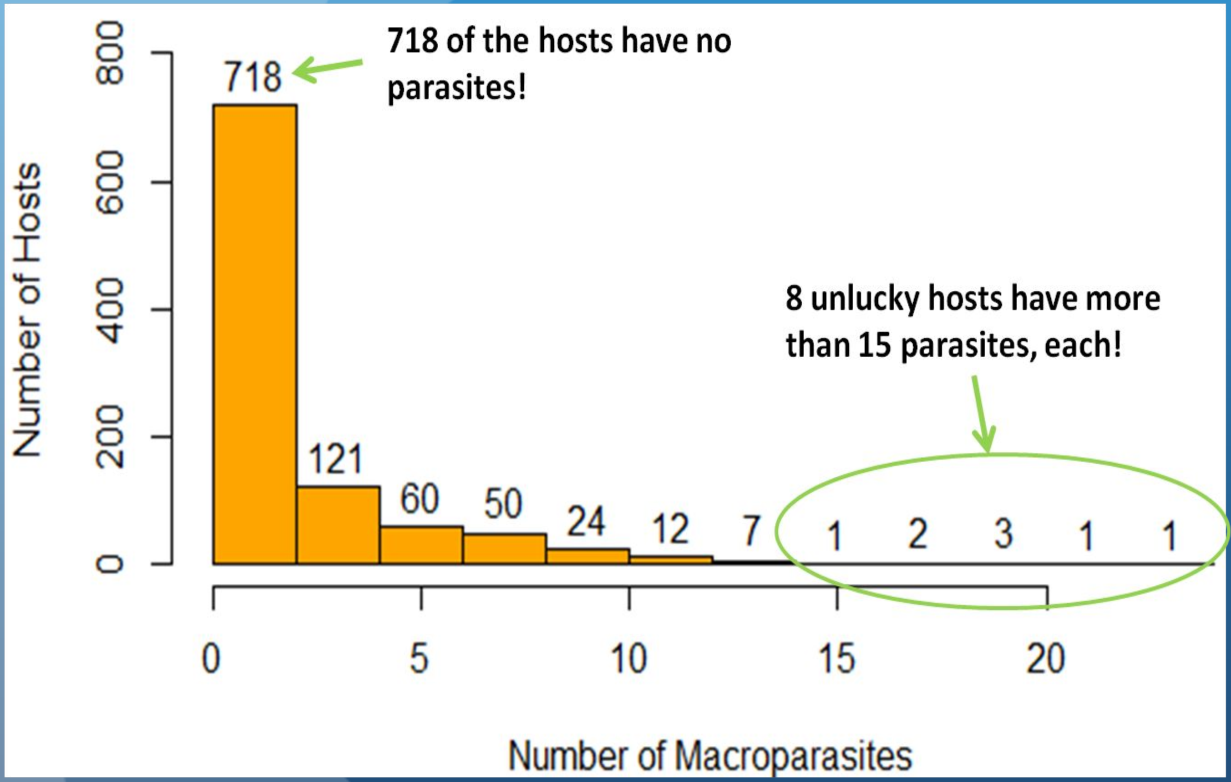
$$\binom{n}{x}\pi^x(1-\pi)^{n-x} \cong \frac{e^{-\lambda}\times\lambda^x}{x!}, \text{com} \quad x = 0, 1, 2, ..., n, \tag{4.8}$$

com parâmetro $\lambda = n\pi$, a média da distribuição binomial. Observemos que, no lado esquerdo de (4.8), temos a função de probabilidade da binomial; e, no lado direito, a função de probabilidade da Poisson, com parâmetro dado pela média da binomial. Para saber se a aproximação é boa, uma recomendação prática é verificar se a desigualdade $n\pi \leq 10$ é válida. No Apêndice H estão apresentadas algumas situações do cálculo das probabilidades usando os dois modelos. Podemos notar que para $n \geq 500$ e $\lambda \leq 10$ a aproximação é boa.

Distribuição Binomial / Negativa

- Mais usada de todas para modelar dados agregados.
- Aplica-se no caso de duas premissas da poisson não serem válidas: a) cada amostra ter probabilidade igual de hospedar organismo e b) presença de organismo não afeta a ocorrência dos demais.
- Indicada por alta variância em relação à média
- Dois parâmetros: média e k= parâmetro indicando grau de agregação

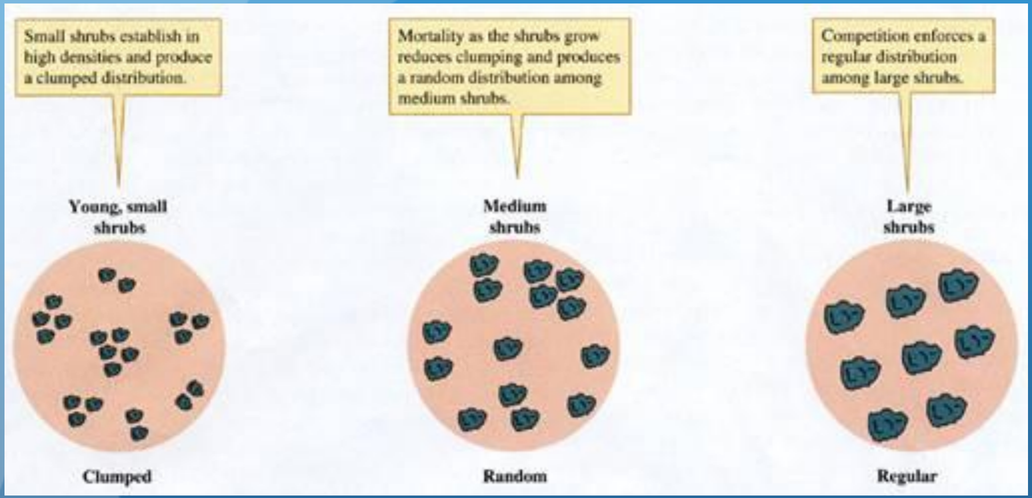
Agregação de Parasitas

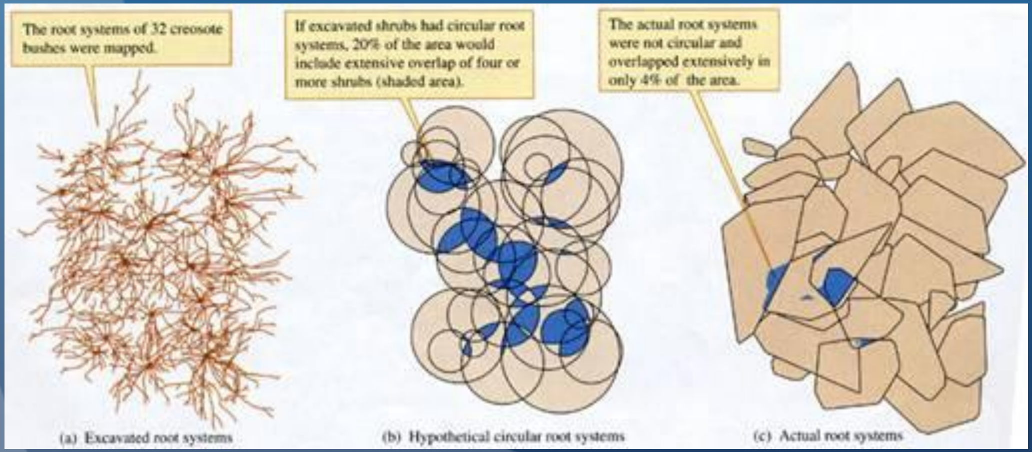


Distribuição Binomial / Positiva

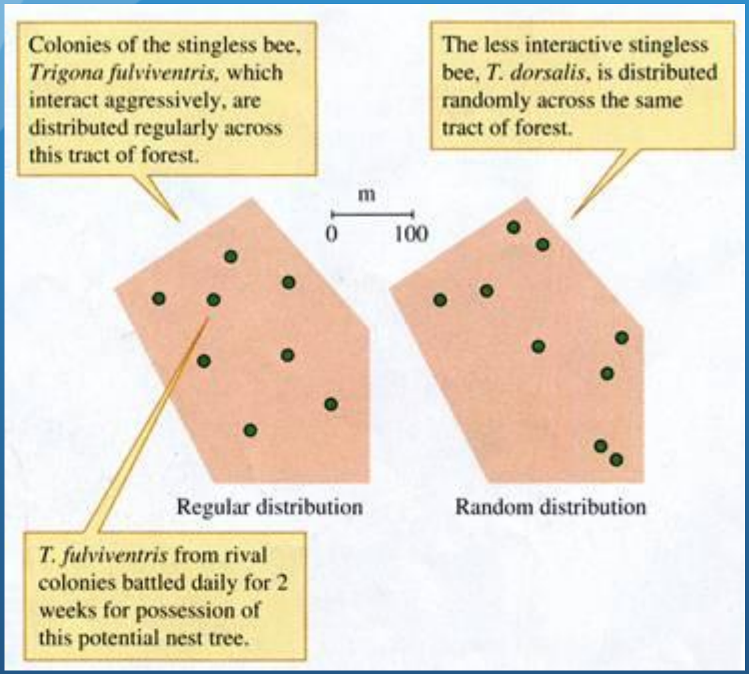
- Assim como a binomial negativa, usada quando as premissas da poisson não se aplicam
- Entretanto está associada a baixa variância em relação à média
- Interações negativas levam a super-dispersão

Interações Intra-Específicas





Abelhas – Hubbell e Johnson



Como medir

- Métodos diversos
- Quadrat
- Transectos
- Contagens instantâneas

Quadrats – Curso Biomar 2011

Ajustar Distribuição

- Trabalho de Elida Cunha (2011) distribuição cupins
- Índice de morisita para gerar hipótese
- Agregada ou poisson?
- Gerou distribuição esperada
- Comparou com observada qui-quadrado

Cupins – Iporá - GO

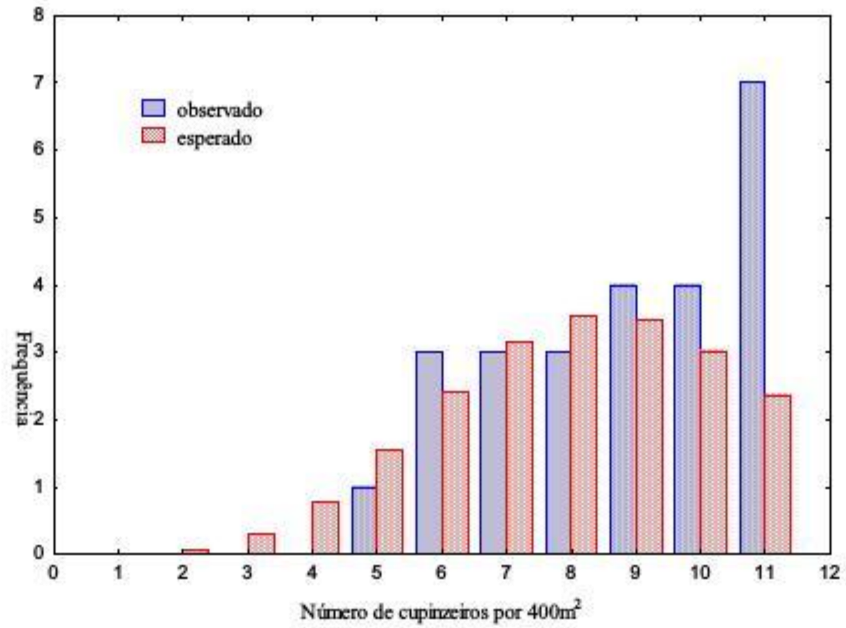


Figura 1. Distribuição de cupinzeiros observados em 25 parcelas da pastagem e a frequência esperada pela Distribuição Poisson, Iporá-GO, 2005.

Cupins – Iporá - GO

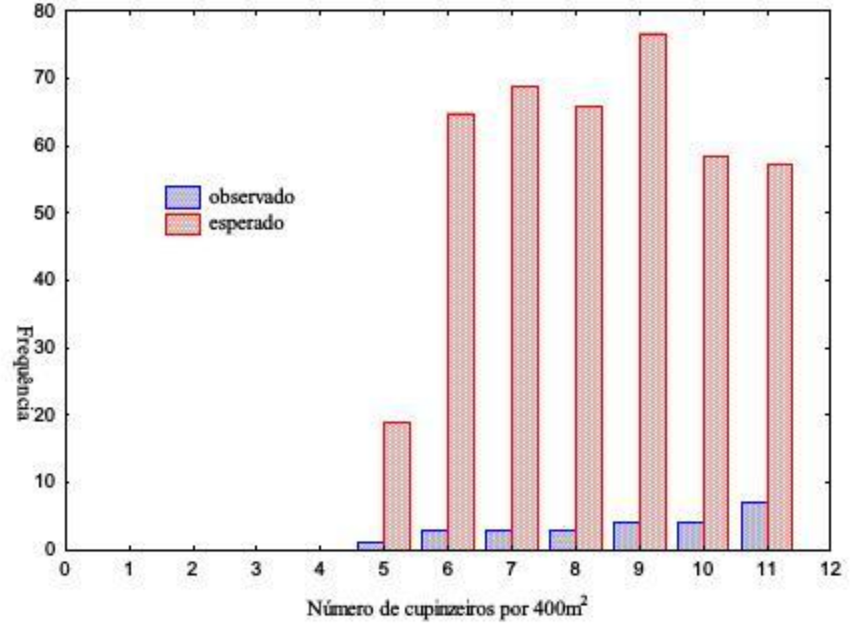


Figura 2. Distribuição de cupinzeiros observados em 25 parcelas da pastagem e a frequência esperada pela Distribuição Binomial Positiva, Iporá-GO, 2005

Ajustar Distribuição

- Lima-Ribeiro e Prado 2007 arvore Vernonia aurea cerradão Caiapônia, GO
- Distribuição Poisson (aleatória) e binomial negativa (agregada)
- Teste G (esperado vs observado)
- Indice de dispersao

Ajustar Distribuição

Tabela 2. Distribuição de freqüências esperadas de acordo com a Distribuição Binomial Negativa e cálculos do teste G para parcelas de tamanho 0,09 m^2 em um fragmento de Cerradão no extremo oeste do município de Caiapônia, Goiás, Brasil.

Nº indivíduos por parcela (r)	Interior do fragmento			Borda do fragmento		
	Freqüências		Teste G	Freqüências		Teste G
	Fo(r)	Fe(r)	cumulativo		Fe(r)	cumulativo
0	72	71,20	0,80	52	52,75	-0,75
1	59	59,19	0,61	54	50,15	3,24
2	31	35,34	-3,46	34	36,69	0,65
3	24	18,35	2,99	21	24,16	-2,29
4	5	8,81	0,16	16	15,03	-1,29
5	7	4,02	4,04	10	9,02	-0,26
6	1	1,77	-	6	5,28	0,51
7	1	1,32	-	2	3,04	-0,33
8	-	-	-	4	1,72	-
9	-	-	-	1	2,16	-
-	*2	*3,09	3,17	*5	*3,88	0,94
Total	200	200,0	G = 6,34	200	200,0	G = 1,88

Fo(r): distribuição de freqüências observadas (número de parcelas com 0, 1, 2, ..., r indivíduos); Fe(r): distribuição de freqüências esperadas (modelo teórico); *Somatória das frequências esperadas menores que três. Suas respectivas frequências observadas também foram somadas para o calculo do teste G.

Ajustar Distribuição

Tabela 1. Distribuição de freqüências esperadas de acordo com o modelo de Poisson e cálculos do test parcelas de tamanho 0,09 m² em um fragmento de Cerradão no extremo oeste do muni Caiapônia, Goiás, Brasil.

Nº indivíduos por parcela (r)	Interior do fragmento			Borda do fragmento		
	Freqüências		Teste G	Freqüências		Teste G
	Fo(r)	Fe(r)	cumulativo	Fo(r)	Fe(r)	cumulativo
0	72	54,24	20,39	52	28,48	31,31
1	59	70,78	9,65	54	55,48	29,85
2	31	46,18	-2,71	34	54,10	14,06
3	24	20,08	1,57	21	35,16	3,24
4	5	6,56	0,21	16	17,14	2,14
5	7	1,72	-	10	6,68	6,17
6	1	0,38	-	6	2,18	-
7	1	0,06	-	2	0,60	-
8	-	-	-	4	0,14	-
9	-	-	-	1	0,04	-
-	*9	*2,16	13,06	*13	*2,96	25,40
Total	200	200,0	G = 26,13	200	200,0	G = 50,80

Fo(r): distribuição de freqüências observadas (número de parcelas com 0, 1, 2, ..., r indivíduos); Fe(r): distribuição de f esperadas (modelo teórico); *Somatória das frequências esperadas menores que três. Suas respectivas frequências observada foram somadas para o calculo do teste G.

Ajustar Distribuição

- Medeiros et al 2015 distribuição moluscos em estuários
- Diversas distâncias a partir do litoral
- Densidades
- Índice de morisita

Ajustar Distribuição

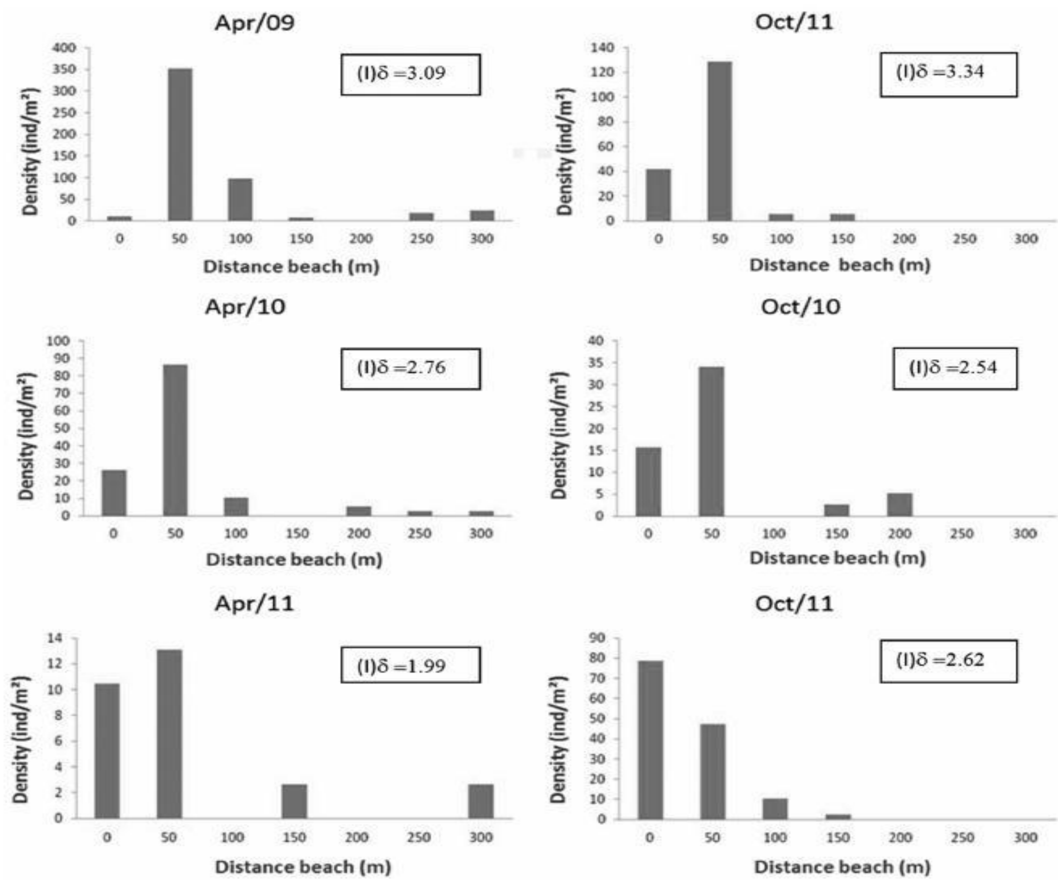


Figure 5. Average density of *Donax striatus*, in the estuary of Apodi/Mossoró River, Rio Grande do Norte State (RN), Northeast Brazil, acoording to different sampling times (Apr/09, Oct/09, Apr/10, Oct/10, Apr/11, and Oct/11) and corresponding Morisita index values (**I**δ).

Índices de Agregação

- Variancia – Média
- Binomial negativa – k
- Coeficiente de Green
- Indice de Morisita
- Distancia – vs regularidade

Abordagem Frequencista

- Desenho experimental / observacional pré-definido Bolker p.10
- Calcular probabilidade de resultado particular, definido como a frequencia média daquele resultado em uma sequencia a longo prazo de experimentos repetidos.
- Calcular o valor p , ou seja a probabilidade do resultado particular ou de resultados mais extremos dada uma hipótese nula
- Caso a probabilidade for baixa, rejeitar a hipótese nula
- Críticas: mesmo com valores de p baixos (< 5%), rejeita-se a hipótese nula em alguns casos – um em 20 para 5%. Se o p> 5% aceitamos a hipótese nula com o risco da mesma ser falsa.

Abordagem Frequencista Slide 2

- Problemas com hipótese nula:
- Pode ser rejeitada mesmo com diferenças pequenas em relação à alternativa, em situações de grandes amostras que dão significancia estatística.
- Hipótese nula pontual (ex: inclinação da reta=0) não é realista. O teste acaba por depender de termos dados suficientes para rejeitar a hipótese nula.
- Recomendação (Bolker). Melhor estimar os valores dos parâmetros biologicamente significativos e obter seus intervalos de confiança do que centrar nos valores de p.

Abordagem Bayesiana (Bolker)

- Os dados observados são a realidade. Os parâmetros ou hipóteses tem distribuições probabilísticas
- Frequencista: existe um conjunto de parâmetros reais. Os dados experimentais fazem parte de uma distribuição de possíveis resultados.
- Vantagens da abordagem bayesiana:respostas são derivadas dos dados e não de uma sequencia hipotética de repetições; permite fazer afirmativas sobre probabilidades de diferentes hipóteses ou valores de parâmetros.
- Dificuldades: na estatística Bayesiana as probabilidades das hipóteses tem de ser definidas a priori.
- Vantagens bayesianas: se há dados anteriores para incorporar à análise; modelos complexos multivariáveis; lacunas de dados; apoio à decisão (eventos raros/catastróficos)

Abordagem por Estimadores- Bolker

- Estimativa por máxima verossimilhança (Andrade e Ogliari sec 7.2). Combinado com análise frequencista.
- Dado um modelo estatístico, este método obtém o conjunto de parâmetros que maximizam a probabilidade de ocorrência dos dados observados.
- Geralmente usa-se o logaritmo da máxima verossimilhança (log-likelihood)
- Aceita-se um intervalo de 2 unidades log $e^{-2} \sim 1/7.4 = 14\%$
- Outra abordagem: em amostragem repetida, distribuição do log negativo segue qui-quadrado. Usa se 95% da distribuição qui-quadrado ou 1.92 unidades log ou razão $e^{1.92} = 6.82$ como fator de decréscimop. Fora do intervalo apenas 5% no tempo