

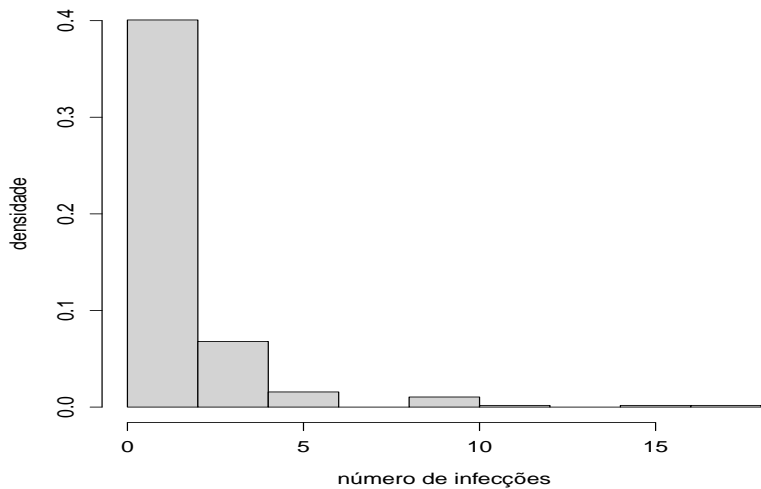
Análise de dados: recrutas

Prof. Caio Azevedo

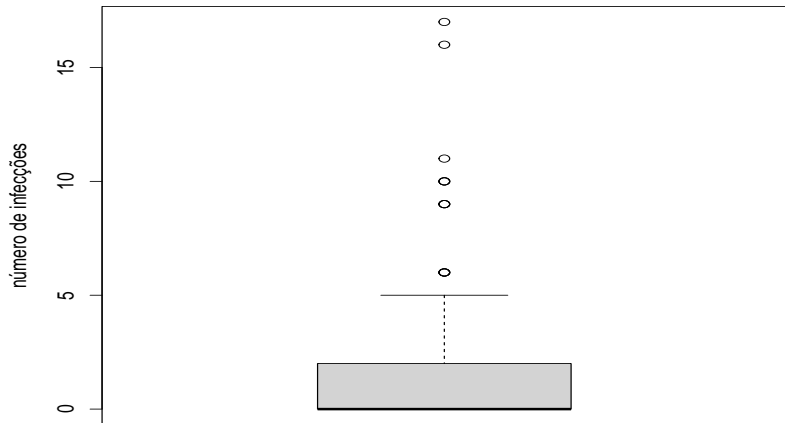
Exemplo 4

- No arquivo `recrutas.dat` são descritos os resultados de um estudo desenvolvido em 1990 com recrutas americanos referente a associação entre o número de infecções de ouvido e alguns fatores. Os dados são apresentados na seguinte ordem: hábito de nadar (ocasional ou frequente), local onde costuma nadar (piscina ou praia), faixa etária (15-19, 20-25 ou 25-29), sexo (masculino ou feminino) e número de infecções de ouvido diagnosticadas pelo próprio recruta (`ninf`). Veja mais detalhes [aqui](#).
- Vamos considerar como resposta a variável “`ninf`”. Por se tratar de uma contagem, uma escolha (inicial) natural é a distribuição de Poisson

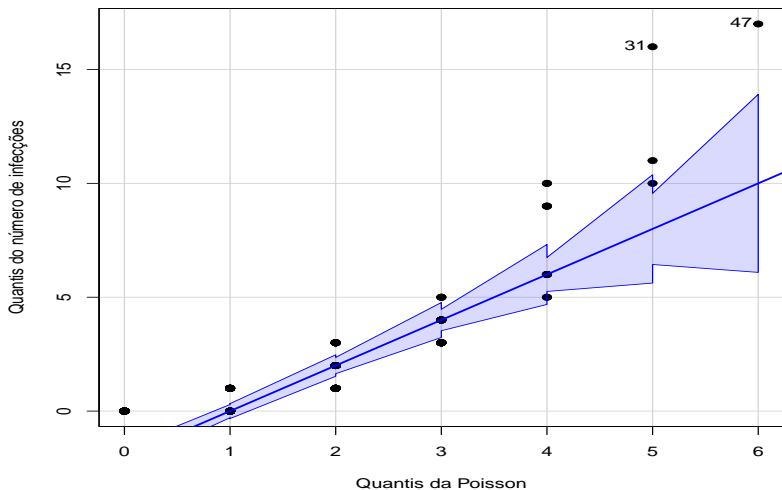
Histograma



Box-plot



QQ-plot



Medidas resumo

MR	Valor
media	1,39
dp	2,34
variância	5,47
cv (%)	168,63
minimo	0,00
mediana	0,00
maximo	17,00
ca	3,19
curtose	16,91

Priori

- Vamos usar a família conjugada de prioris ($\text{gama}(a, b^{-1})$) ([aqui](#)).
- Vamos considerar diferentes pares de hiperparâmetros:

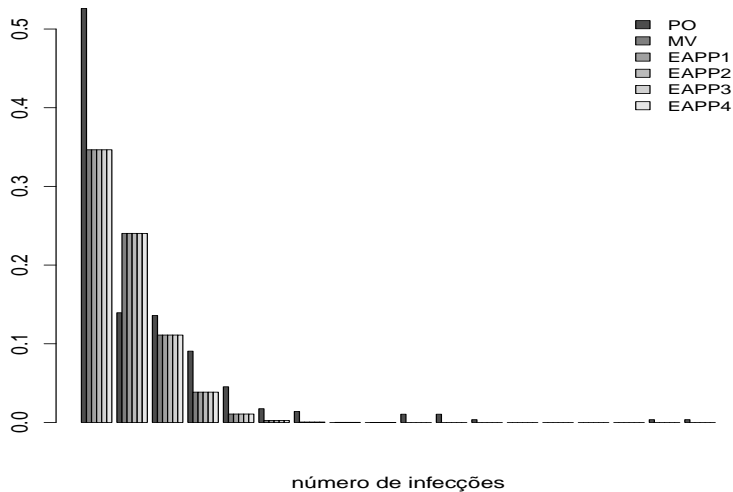
priori	média	variância
(1) $\text{gama}(0,0013;1/0,001)$	1,30	1300,00
(2) $\text{gama}(0,013;1/0,01)$	1,30	130,00
(3) $\text{gama}(0,13;1/0,1)$	1,30	13,00
(4) $\text{gamma}(1,3;1/1)$	1,30	1,30

Posteriori e Inferências

- Resultados numéricos:

Método	Est.	EP/DPA
EMV	1,387	0,070
EAP(priori 1)	1,387	0,070
EAP(priori 2)	1,387	0,070
EAP(priori 3)	1,387	0,069
EAP(priori 4)	1,386	0,069

Gráficos de barras



Comparação entre as distribuições

- Desvio médio absoluto entre as frequências relativas observadas e esperadas.

MV	EAPP1	EAPP2	EAPP3	EAPP4
0,0251	0,0251	0,0251	0,0251	0,0251

- Percebemos que as representações, em termos de uma distribuição de Poisson, sob cada uma das estimativas de θ , são bem próximas.