

Material Auxiliar de Estudo para o Laboratório 2

Roteiro do Laboratório 2

Tema: Probabilidade básica

Referência: Capítulo 1 do livro-texto.

- 1) Revisão de atribuição de probabilidade em espaços amostrais equiprováveis
- 2) Exemplo 1. O método de captura-recaptura para estimação de populações de animais.
 - i) Descrição do método de estimação de populações desconhecidas
 - ii) A probabilidade de x indivíduos recapturados
 - iii) Estimativa da população por maximização da probabilidade
- 3) Revisão do teorema de Bayes
- 4) Exemplo 2. Sensibilidade e especificidade de testes clínicos
 - i) Definições de sensibilidade e especificidade (Jacob Yerushalmy, 1947)
 - ii) Falsos positivos, falsos negativos, e por que testes são refeitos
 - iii) A probabilidade a posteriori como função da probabilidade a priori

A seguir é apresentado o material didático completo para auxiliar a compreensão dos dois exemplos desse laboratório.

Material Auxiliar de Estudo para o Laboratório 2 – Exemplo 1

Para melhor compreensão dos exercícios (exemplos) do Laboratório 2 da disciplina de Probabilidade e Estatística, recomenda-se a leitura (e exercitação, quando for o caso) deste material didático auxiliar.

Exemplo 1. O método de captura-recaptura para estimação do tamanho de população de animais

Este exemplo baseia-se em um problema que envolve o cálculo de probabilidades em espaços amostrais finitos equiprováveis e constitui uma interessante aplicação do cálculo combinatório na solução de um problema prático – a estimação do tamanho de população de animais silvestres, sob certas condições. Uma apresentação detalhada desse método é encontrada no material didático complementar do curso, do tipo aplicação, sobre a distribuição hipergeométrica – O método de captura, marca e recaptura – já antes enviado. Recomenda-se inicialmente uma leitura desse documento, no qual é mostrado o modelo probabilístico empregado.

Em seguida, pode ser útil a leitura do texto a seguir que mostra o resultado da execução do programa do exemplo 1, caso 1, conforme o script (em R) do mesmo.

Para auxiliar a compreensão do problema são a seguir apresentados em detalhe os resultados do primeiro caso do exemplo 1, com a especificação dos seguintes parâmetros envolvidos:

- i) Tamanho da população de tubarões (N), fixado por hipótese na primeira parte do exercício, com o valor igual a 100.
- ii) Número de tubarões capturados inicialmente (ou seja, os animais conseguidos na primeira operação de captura), os quais são todos marcados (K), com valor igual 15.
- iii) Número de tubarões obtidos na segunda captura (n), com valor 15.

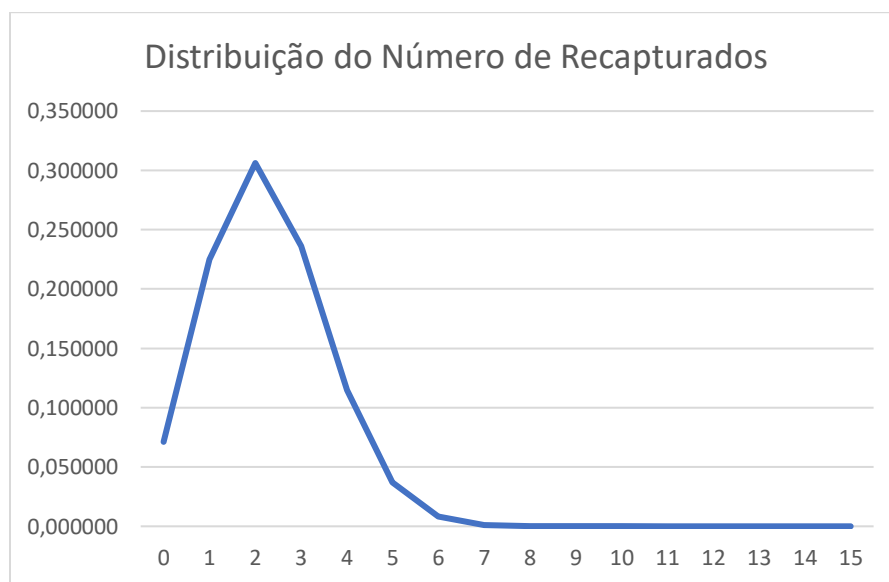
Nessas condições, empregando o modelo probabilístico especificado é possível determinar a distribuição de probabilidades do número eventual de tubarões marcados que são recapturados. O resultado obtido com a utilização do programa (script) em R é aqui representado e comentado em detalhe.

- a) Distribuição do número eventual de animais capturados na segunda operação e que foram marcados, supondo que o tamanho (não conhecido) da população é $N=100$, considerando que na primeira operação foram capturados $K=15$ animais e que na segunda operação foram capturados $n=15$ animais. Essa distribuição é apresentada na forma de uma tabela mostrada a seguir.

k	p(k)
0	0,07099
1	0,22497
2	0,30621
3	0,23630
4	0,11496
5	0,03709
6	0,00813
7	0,00122
8	0,00013
9	0,00001
10	0,00000
11	0,00000
12	0,00000
13	0,00000
14	0,00000
15	0,00000
Total	1,00000

Um exame da tabela acima mostra que o valor mais provável do número de animais marcados recapturados é 2.

O gráfico da função de probabilidade é mostrado a seguir, no qual também se verifica, de forma muito clara, que o número de animais recapturados (marcados) mais provável de ser obtido é igual a 2.



Em seguida, executando também a segunda parte do referido programa (script) em R, é obtida, de forma experimental, a distribuição de probabilidade para os diversos valores de N, tamanho da população de tubarões, admissíveis considerando que (dado que) o número de

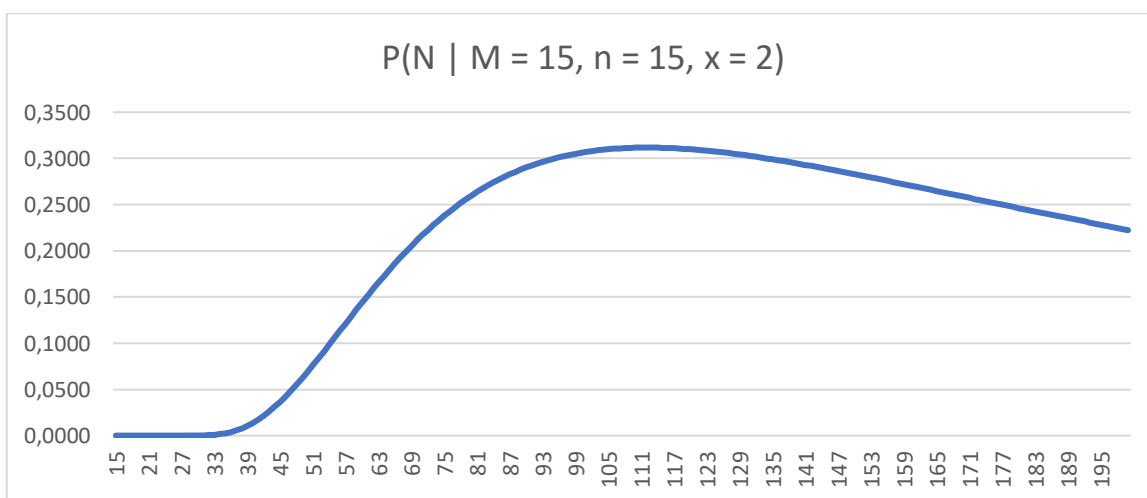
tubarões marcados que foram recapturados é $k = 2$. Os resultados desse cálculo são mostrados a seguir, por meio de uma tabela, e em seguida ilustrados pelo gráfico correspondente.

b) Sabendo que o número de animais recapturados foi 2, a distribuição do tamanho da população dado que $x = 2$ e também que são conhecidos $K = 15$ e $n = 15$ pode ser determinada, como é mostrado a seguir.

Foram calculadas as probabilidades para valores de N variando de 15 a 200. Como essa tabela inteira é muito grande, será apresentada de forma resumida. Examinando a tabela verifica-se que o tamanho da população mais provável é $\hat{N}^* = 112$.

N	p(N)
15	0,0000
...	...
29	0,0001
30	0,0002
...	...
50	0,0689
...	...
100	0,3062
...	...
110	0,3116
111	0,3117
112	0,3118
113	0,3117
114	0,3117
115	0,3115
...	...
150	0,2829
...	...
200	0,2226

O gráfico dessa distribuição é mostrado a seguir.



O exame desse gráfico também mostra (visualmente) que o valor mais provável para o tamanho da população de animais é aquele já antes apresentado.

c) Empregando a expressão para estimar o tamanho da população, que é mostrada na aplicação sobre o método de captura, marca e recaptura, tem-se

$$\hat{N} = \left[\frac{K \cdot n}{x} \right] = \left[\frac{15 \cdot 15}{2} \right] = \left[\frac{225}{2} \right] = [112,5] = 112 \text{ ou seja, nesse caso, há uma solução: } 112$$

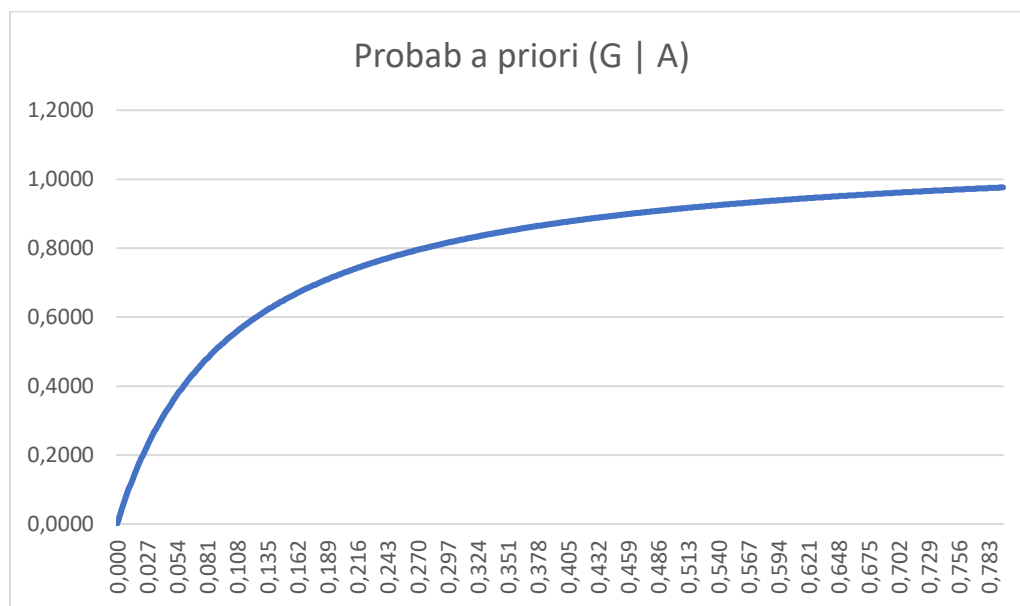
Material Auxiliar de Estudo para o Laboratório 2 – Exemplo 2

Para melhor compreensão do exercício 2 (exemplo 2) do Laboratório 2 da disciplina de Probabilidade e Estatística, recomenda-se a leitura (e exercitação, quando for o caso) deste material didático auxiliar.

Exemplo 2. Sensibilidade e especificidade de testes clínicos.

Este exemplo baseia-se em um problema que envolve o cálculo de probabilidades a posteriori de eventos, ou seja, uma aplicação do teorema de Bayes. Em particular, refere-se a testes para diagnósticos clínicos. Nesse sentido, recomenda-se que antes sejam revistos os exemplos apresentados em aula sobre o tema, em particular o exemplo 27, páginas 39 e 40 das notas de aula do capítulo 1. Além disso, também é importante ler o documento contendo o seguinte material didático complementar do curso, do tipo nota didática: Análise Bayesiana – uma Breve Introdução; este documento já foi anteriormente enviado.

Após essas leituras, pode-se realizar a execução do programa (script em R) do exemplo 2, na sua formulação inicial. O resultado é mostrado a seguir, de forma resumida, no gráfico a seguir.



Para melhor compreensão do cálculo que levou a esse resultado, é recomendável fazer uma árvore de probabilidades, o que torna clara a obtenção da expressão utilizada no programa (script em R) deste exemplo 2. Essa expressão mostra a forma paramétrica da probabilidade a posteriori que é empregada no cálculo