

FFL5158 - Estudos Experimentais: Metodologia e Análise

Lista 2

prazo: 21/10/2021

1. O Censo Americano (The American Community Survey) de 2010 divulgou que entre americanos, (i) 14.6% vivem abaixo da linha da pobreza; (ii) 20.7% falam uma língua estrangeira em casa e (iii) 4.2% pertencem a ambas categorias. Considerando esses dados, responda:
 - a. Viver abaixo da linha da pobreza e falar uma língua estrangeira em casa são eventos disjuntos ou não disjuntos? Justifique.
 - b. Desenhe um diagrama de Venn no R indicando a relação entre as variáveis. Ver Como fazer um diagrama de Venn no R
 - c. Qual é a porcentagem de americanos que vivem abaixo da linha da pobreza e falam apenas inglês em casa?
 - d. Qual é a porcentagem de americanos que vivem abaixo da linha de pobreza ou falam uma língua estrangeira em casa?
 - e. Qual é a porcentagem de americanos que vivem acima da linha de pobreza e falam apenas inglês em casa?
 - f. O evento de alguém viver abaixo da linha de pobreza é independente do evento em que uma pessoa fala um língua estrangeira em casa?
2. Em uma tarefa de julgamento de valor de verdade sobre implicatura escalar, uma criança forneceu julgamentos para 50 sentenças no total. O teste foi dividido em 3 dias, sendo que no dia 1 ela julgou 18 sentenças e nos demais ela julgou 16 sentenças. A criança forneceu um porcentagem de 50% respostas adultas no primeiro dia, 68.7% no segundo dia e 56.2% no terceiro dia. Calcule a probabilidade de que a criança forneceu, em cada um dos dias, respostas ao acaso nos itens do teste. Em seguida, use o teste de hipótese da binomial para verificar se a hipótese de que a criança não forneceu respostas ao acaso é verdadeira (use o código abaixo e substitua os valores). Supondo que o experimento não apresenta nenhuma falha, o que você concluiria sobre a gramática dessa criança?

```
binom.test(0, n = 1, p = 0.3, alternative = "two.sided", conf.level = 0.95)
```
3. Para mulheres, falantes nativas de inglês americano, a distribuição da frequência do primeiro formante da vogal [E] possui aproximadamente uma distribuição normal com média 608Hz e desvio padrão 77.5Hz. Faça um gráfico ilustrando essa distribuição. Qual é a probabilidade de que a frequência do primeiro formante do [E] de uma mulher, aleatoriamente selecionada, seja entre 555Hz e 697Hz? Mostre o código do R utilizado para calcular a sua resposta.
4. Um doutorando propôs testar 250 imigrantes holandeses em seu projeto sobre perguntas-qu em holandês. Suponha que holandeses aceitam participar de um experimento de linguística 35% das vezes. Considerando que há 500 holandeses no país em que o doutorando está conduzindo sua pesquisa, e que na melhor das hipóteses ele consiga contatar todos os sujeitos, qual é a probabilidade de que o doutorando consiga recrutar pelo menos 250 participantes? Considerando os seus resultados, você acha que essa marca é realista? Por quê?
5. Faça um gráfico da probabilidade da frequência de participantes, considerando os dados do exercício anterior. Para isso:

- a) crie um vetor `x` com a função `seq()` para indicar que o doutorando pode selecionar no máximo 500 participantes (população = 500): `seq(1, 500, by = 1)`.
 - b) faça um gráfico de frequência utilizando a função `plot()`, tendo o vetor `x` e `dbinom(x, n, p)` como argumentos. Selecione `type = h` (histograma) em `plot()`.
 - c) baseando-se no gráfico, escolha um tamanho de amostra realista para o doutorando, e em seguida, repita o cálculo realizado no exercício anterior.
6. Dryer (2013) conduziu uma pesquisa sobre a ordem de palavras em 1228 línguas, encontrando a seguinte distribuição. Se fecho os meus olhos e coloco 40 vezes o meu indicador no mapa “Languages of the World”, qual é a probabilidade que em pelo menos 5 vezes meu dedo caia em um local onde a língua falada não tem ordem de palavras fixa?

Order	Number	Percentage
SOV	565	41.00%
SVO	488	35.00%
VSO	95	6.90%
VOS	25	1.80%
OVS	11	0.80%
OSV	4	0.30%
Unfixed	189	13.00%