

MAE0116 – Noções de Estatística

Lista de exercícios 2 – Estatística Descritiva II – Casa

Nesta lista, serão analisados dados coletados pelos autores Cheng et. al (2019) e publicados no artigo referenciado a seguir.

Cheng et al.(2019). T. L. Cheng, A. Gerson, M. S. Moore, J. D. Reichard, J. DeSimone, C. K. Willis, W. F. Frick e A. M. Kilpatrick. *Higher fat stores contribute to persistence of little brown bat populations with white-nose syndrome.* ***Journal of Animal Ecology***, 88:591–600.

Os dados são referentes à gordura corporal de morcegos coletados em seis locais do nordeste dos Estados Unidos durante o inverno de 2009 (que abrange a estação de hibernação de outubro de 2008 a abril de 2009) e de 2016 (outubro de 2015 a abril 2016). Os locais de coleta foram *Aeolus Cave*, *Hibernia Mine*, *Graphite Mine*, *Williams Preserve Mine*, *Barton Hill Mine*, *Williams Hotel Mine*. Os autores investigam se a persistência da síndrome do nariz branco (white-nose syndrome) em uma espécie de pequenos morcegos marrons (*M. lucifigus*) pode ser explicada pelo aumento da gordura corporal no início do inverno, o que permitiria aos morcegos tolerar o aumento dos custos energéticos associados à síndrome. O banco de dados (adaptado) é composto por 555 morcegos. Nesta análise, as variáveis presentes no conjunto de dados serão estudadas de forma descritiva. A variável de interesse é a proporção de gordura corporal (y). As demais variáveis são:

- ▷ `local`: local em que o morcego foi amostrado (*Aeolus*, *Hibernia*, *Graphite*, *Williams Preserve*, *Barton Hill*, *Williams Hotel*);
- ▷ `sexo`: sexo do morcego (M para macho e F para fêmea);
- ▷ `ano`: ano em que o morcego foi amostrado;
- ▷ `dias`: tempo de hibernação (definido como o número de dias desde o equinócio de outono).

O conjunto de dados está disponível no arquivo `Morcegos_Cheng`. Responda os itens a seguir.

Exercício 1

A tabela a seguir apresenta algumas medidas descritivas da variável y por local.

Tabela 1: Medidas resumo para a variável y por local						
	n	média	desvio padrão	mediana	mínimo	máximo
Aeolus	159	0.122	0.071	0.096	0.025	0.323
Barton Hill	57	0.168	0.070	0.142	0.087	0.335
Graphite	117	0.176	0.074	0.165	0.048	0.393
Hibernia	140	0.161	0.069	0.144	0.037	0.324
Williams Hotel	62	0.179	0.035	0.179	0.102	0.250
Williams Preserve	20	0.259	0.042	0.255	0.201	0.364

- (a) Classifique as variáveis y e local. Obtenha a média e os valores mínimo e máximo de y .
- (b) Utilizando o coeficiente de variação, compare os grupos de morcegos dos diferentes locais quanto à variabilidade da variável y .
- (c) Os valores do primeiro e terceiro quartil de y para os morcegos da região Aeolus são, respectivamente, 0.067 e 0.164. Com essas informações adicionais é possível construir o boxplot para a proporção de gordura corporal dos morcegos amostrados nessa região? Se sim, obtenha o boxplot. Se não, justifique.
- (d) Obtenha os boxplots da variável y para os diferentes locais (utilize o R e o conjunto de dados disponível no arquivo `Morcegos_Cheng`). Comente.

Exercício 2

A Figura 1 apresenta o histograma da variável y .

- (a) Obtenha os valores aproximados para a mediana e o terceiro quartil de y com base no histograma. Compare esses valores com os exatos, obtidos através do software R.
- (b) Apresente a tabela de frequências que originou o histograma da Figura 1. Inclua as frequências relativas e absolutas para cada classe.
- (c) Utilizando o software R, obtenha um histograma com classes iguais e o boxplot da variável y .

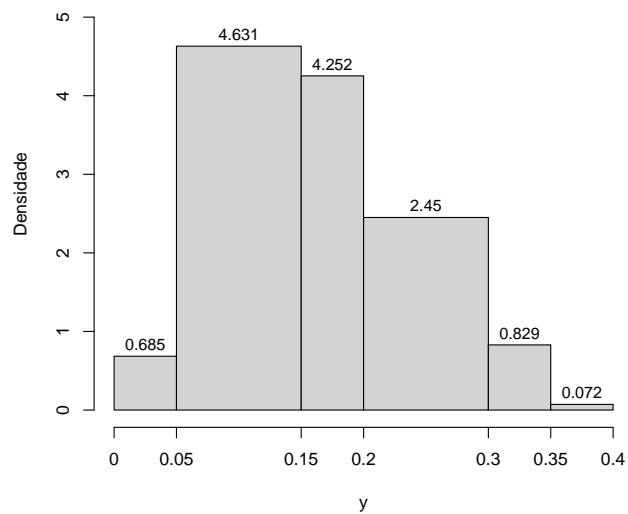


Figura 1: Histograma da variável y .

Exercício 3

O gráfico de setores para a variável local é apresentado na Figura 2.

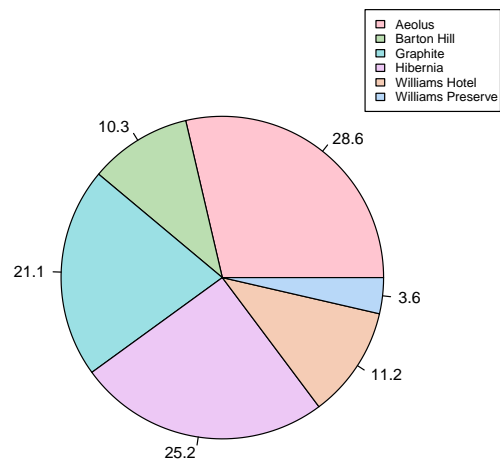


Figura 2: Gráfico de setores da variável local.

(a) Observando o gráfico da Figura 2, o que podemos dizer a respeito da variável local?

- (b) Qual outro gráfico poderia ser utilizado para essa variável? Desenhe, manualmente, esse gráfico.
- (c) Utilizando o R, obtenha um gráfico de setores para a variável ano.

Exercício 4

Nesta questão, vamos estudar a relação entre as variáveis y e dias, considerando os dados dos morcegos amostrados na região Barton Hill. A tabela a seguir apresenta algumas medidas descritivas da variável y por dias.

Tabela 2: Medidas resumo para a variável y para cada valor da variável dias.

dias	n	média	desvio padrão	mediana	mínimo	máximo	Q1	Q3
37	19	0.253		0.245				0.285
155							0.102	
175	18		0.025	0.124		0.172		

- (a) Utilizando o R, complete a Tabela 2 com os valores faltantes.
- (b) Sem utilizar ferramentas computacionais, desenhe, em um mesmo gráfico, os boxplots da variável y para cada valor da variável dias. Comente sobre a relação entre as variáveis y e dias. Considerando o contexto do problema, essa relação é esperada?