

Affonso Amendola
NUSP 9301753

Análise de Dados em Astronomia
Exercício 1

Estatística descritiva e análise de variância do arquivo LMC_distance.dat
(https://astrostatistics.psu.edu/datasets/LMC_distance.dat)

1. e 2. Calcule estatísticas descritivas em cada caso, compare as medias e medianas em cada caso:

Usando a função summary() do R, foram obtidos os seguintes valores para os dois casos:

Population I:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
18.40	18.47	18.55	18.54	18.58	18.70

Population II:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
18.30	18.43	18.45	18.47	18.53	18.69

No entanto, o uso dessa função não leva em conta os valores de erro para o cálculo da média, no entanto, ao usar a função weighted.mean() é possível obter um resultado melhor:

Population I:

weighted mean: 18.55231

Population II:

weighted mean: 18.48804

Temos que o resultado da média nos dois casos é bem próximo da mediana, o caso da População I sendo ~0.1 maior que a População II, com o resultado da média com peso sendo muito similar a mediana no caso da Pop I e muito similar a da média na Pop II.

3. Compare os desvios padrão com a dispersão obtida da distancia intraquartil:

A dispersão obtida da distância intraquartil é, para cada caso:

Population I:

dispersion: 0.081543

Population II:

dispersion: 0.07413

Os erros dos dados tem valor médio de:

Population I:
mean_err: 0.1107692

Population II:
mean_err: 0.1533333

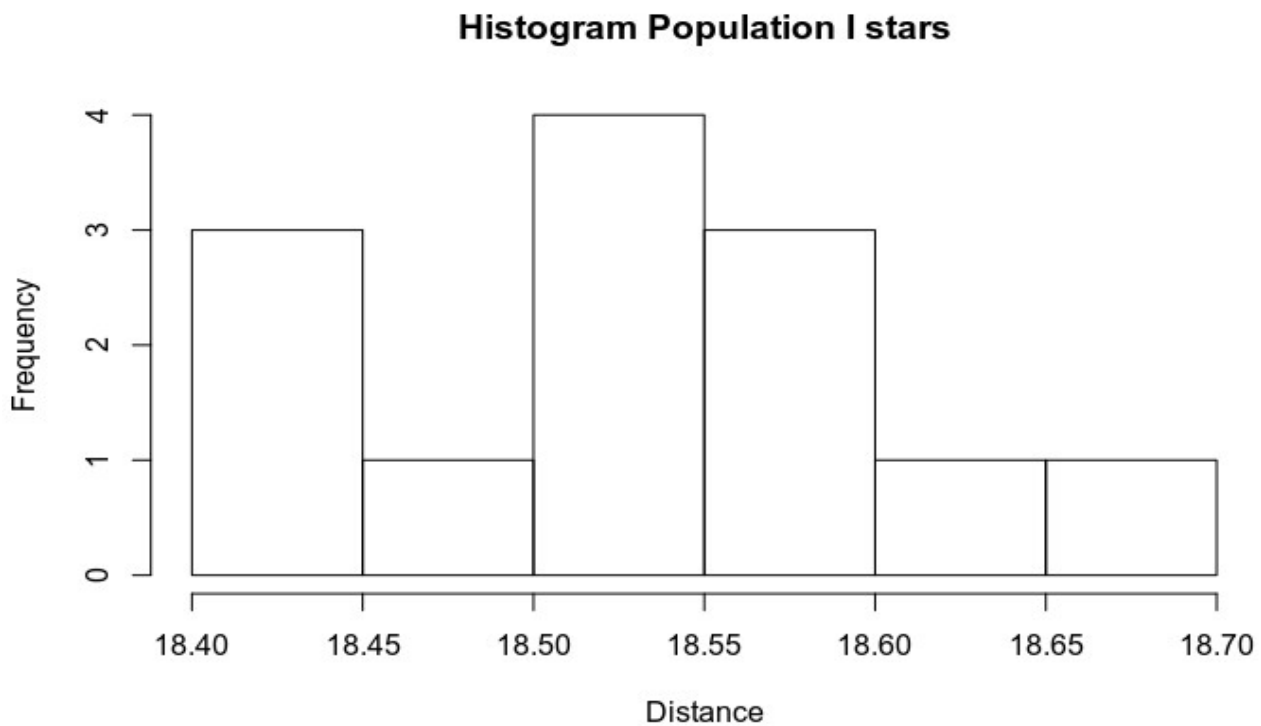
E o desvio padrão dos dados tem valor de:

Population I:
std_dev: 0.08562073

Population II:
std_dev: 0.09930058

Podemos ver que os valores de dispersão obtidos da distância intraquartil é comparável ao valor de desvio padrão, especialmente no caso da Pop I, no caso da Pop II o valor de desvio padrão é ~ 0.03 maior, mas os erros obtidos nas medidas tem um valor significativamente maior (em média)

4.



Histogram Population II stars

