

Nome: Deivid da Silva Galvão

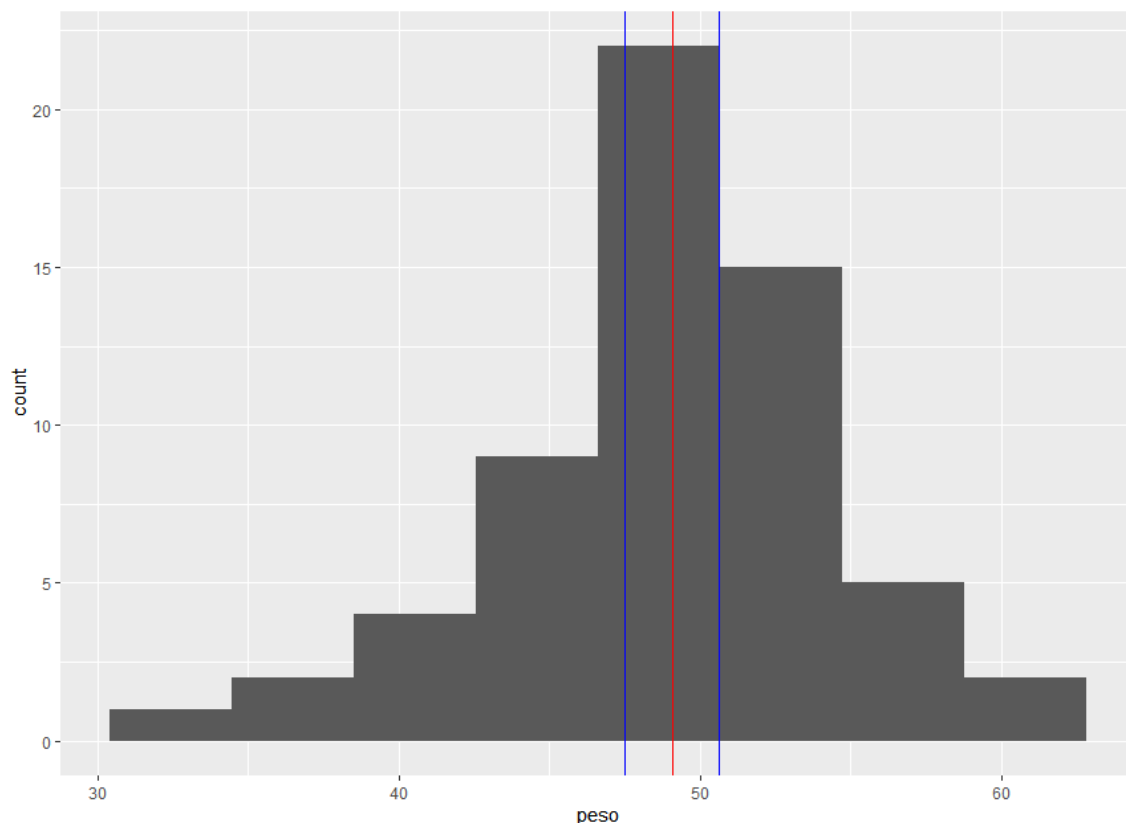
RA: 2408740

Para todas as questões abaixo, **interprete os resultados e apresente os códigos**. As bases de dados estão em anexo do google Classroom, já salvas em CSV, com o separador decimal em Inglês, ou seja, as decimais estão separadas por ponto. **Cada questão vale 2 pontos.**

ex1) Considerando os dados **p3ex1.csv**, construa um intervalo de confiança para a variável peso, considerando 4% de significância. Apresente um histograma da variável, apresentando a média e os limites do intervalo encontrados. (dica: utilize `geom_vline()` no ggplot para adicionar as linhas)

Resposta: A Partir do output gerado pelo código abaixo podemos concluir que o intervalo de confiança, com 4% de significância, estará dentro do intervalo de 47,51532 a 50,63513 e a média da amostra é estimada em 49,07522.

```
alternative hypothesis: true mean is not equal to 0
96 percent confidence interval:
 47.51532 50.63513
sample estimates:
mean of x
49.07522
```



Códigos:

`dados=read.csv('p3ex1.csv', sep = ',', dec='.', header = T)`

`install.packages("ggplot2")`

```
# Carregando a biblioteca ggplot2

library(ggplot2)

# IC para média

#considerando um nível de 4% de significância, ou seja, 96% de confiança.

t.test(dados$peso, conf.level = 0.96)
```

ex2) Em uma amostra de 100 peças, observou-se quais estavam com o tamanho correto ou não (**p3ex2.csv**). Determine a estimativa de peças com o tamanho incorreto, bem como uma estimativa confiável para a verdadeira proporção de equipamentos com tamanho incorreto, considerando um nível de significância de 4%. (Dica: Table())

Resposta: Em resumo, com base na amostra analisada, estima-se que cerca de 53% das peças estão com o tamanho incorreto na população, com um intervalo de confiança de 42.35% e 63.39% considerando um nível de confiança de 96%.

```
X-squared = 0.25, df = 1, p-value = 0.6171
alternative hypothesis: true p is not equal to 0.5
96 percent confidence interval:
 0.4235209 0.6339348
sample estimates:
      p
0.53
```

Códigos:

```
dados2=read.csv('p3ex2.csv', sep = ',', dec='.', header = T)

# Contando o número de peças com tamanho correto e tamanho incorreto

aux=table(dados2$tamanho)

# Estimativa de peças com tamanho incorreto

incorretas=aux["incorreto"]

# Intervalo de confiança com 4% de significância

prop.test(53, 100, conf.level = 0.96)
```

ex3) O tempo médio ideal **máximo** para a produção de determinado componente é de **50h ou menos**. Para verificar se uma nova máquina atende os requisitos mínimos, foi realizado o teste em 40 produtos (**p3ex3.csv**). Verifique, considerando um nível de significância de 5%, se a máquina nova atende o tempo médio ideal, ou seja, se ela será comprada ou não. (Dica: H0: Compra x H1: não compra)

Resposta:

=> H0: media = 50

=> H1: media > 50

Como o p valor associado de 0.0001766 é menor do que o nível de significância de 0.05, pode-se rejeitar a hipótese nula de que a média de tempo para a produção dos componentes é igual a 50 horas. A hipótese alternativa afirma que a média para a produção é maior que 50 horas. Com um nível de significância de 5% a estimativa da média de tempo é de 53,6949 horas com limite superior do intervalo sendo infinito.

Logo, pode-se concluir que a nova maquina não atende ao requisito de tempo médio ideal de no máximo de 50 horas, então a decisão recomendada é de não comprar.

```
t = 3.9146, df = 39, p-value = 0.0001766
alternative hypothesis: true mean is greater than 50
95 percent confidence interval:
 52.1046      Inf
sample estimates:
mean of x
 53.6949
```

Códigos:

```
dados3=read.csv('p3ex3.csv', sep = ',', dec='.', header = T)
t.test(dados3$tempo, mu = 50, alternative = "greater", conf.level = 0.95)
```

ex4) Em um estudo, 30 ligas de metal foram testadas para verificar a resistência **das mesmas, antes e após** a aplicação de um produto. As forças necessárias para o rompimento das ligas foram anotadas (**p3ex4.csv**). Verifique, considerando 6% de significância, se o produto utilizado teve efeito nas resistências das ligas.

Resposta:

=> H0: media antes = media depois

=> H1: media antes != media depois

Como o p valor(0.0005982) associado é menor do que o nível de significância 0.06, pode-se rejeitar a hipótese nula de que o produto não teve efeito nas resistências das ligas metálicas, logo pode-se concluir que houve uma alteração na média das resistências após a aplicação do produto.

```
t = -3.8513, df = 29, p-value = 0.0005982
alternative hypothesis: true mean difference is not equal to 0
94 percent confidence interval:
-15.766467 -5.140979
sample estimates:
mean difference
-10.45372
```

Códigos:

```
dados4=read.csv('p3ex4.csv', sep = ',', dec='.', header = T)
t.test(dados4$santes, dados4$depois, alternative = "two.sided", paired = T, conf.level = 0.94)
```

ex5) Uma empresa pretende comprar uma nova máquina, porém está preocupada com a durabilidade dos itens produzidos pelas máquinas existentes no mercado. Para escolher entre as duas máquinas mais baratas, um teste de produção de 60 peças foi realizado em **cada uma das máquinas** (**p3ex5.csv**), ou seja, totalizando 120 testes. Determine, ao nível de 8% de significância, se a durabilidade média das máquinas testadas são as mesmas ou não. Caso não, qual a melhor?

Resposta:

=> H0: média durabilidade maquina A = média durabilidade maquina B

=> H1: média durabilidade maquina A != média durabilidade maquina B

Como o p valor é maior que o nível de significância de 0.08, não há como rejeitar com certeza a hipótese nula, que afirma que a durabilidade média das máquinas é a mesma.

```
t = -0.94123, df = 110.07, p-value = 0.3487
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -52.92830  18.84149
sample estimates:
mean of x mean of y
 595.2724  612.3158
```

Códigos:

`dados5=read.csv('p3ex5.csv', sep = ',', dec='.', header = T)`

`t.test(dados5$A, dados5$B, alternative = "two.sided", conf.level=0.92)`