

MAE0217 - Estatística Descritiva - Lista 2

Guilherme Marthe - NUSP:8661962

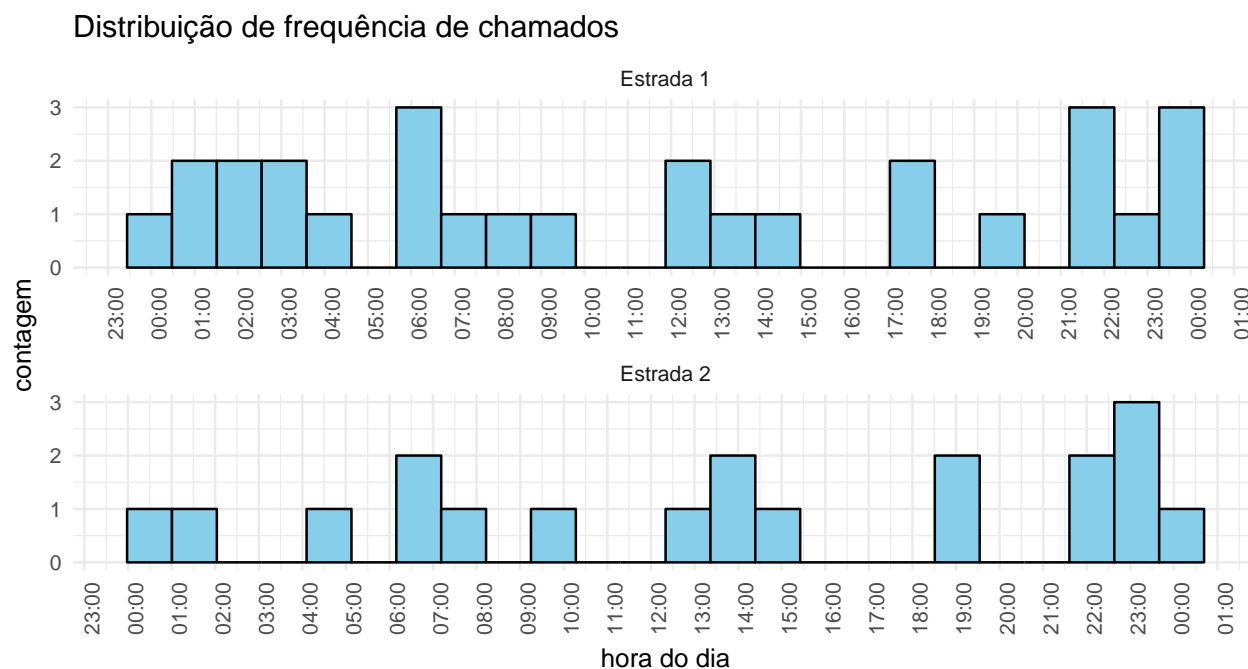
April 18, 2018

Exercício 1

Os dados apresentados na tabela abaixo referem-se a instantes nos quais o centro de controle operacional de estradas rodoviárias recebeu chamados solicitando algum tipo de auxílio em duas estradas, num determinado dia.

```
## # A tibble: 28 x 2
##   `Estrada 1`      `Estrada 2`
##   <dtm>          <dtm>
## 1 2018-04-18 00:07:00 2018-04-18 00:03:00
## 2 2018-04-18 00:58:00 2018-04-18 01:18:00
## 3 2018-04-18 01:24:00 2018-04-18 04:35:00
## 4 2018-04-18 01:35:00 2018-04-18 06:13:00
## 5 2018-04-18 02:05:00 2018-04-18 06:59:00
## 6 2018-04-18 03:14:00 2018-04-18 08:03:00
## 7 2018-04-18 03:25:00 2018-04-18 10:07:00
## 8 2018-04-18 03:46:00 2018-04-18 12:24:00
## 9 2018-04-18 05:44:00 2018-04-18 13:45:00
## 10 2018-04-18 05:56:00 2018-04-18 14:07:00
## # ... with 18 more rows
```

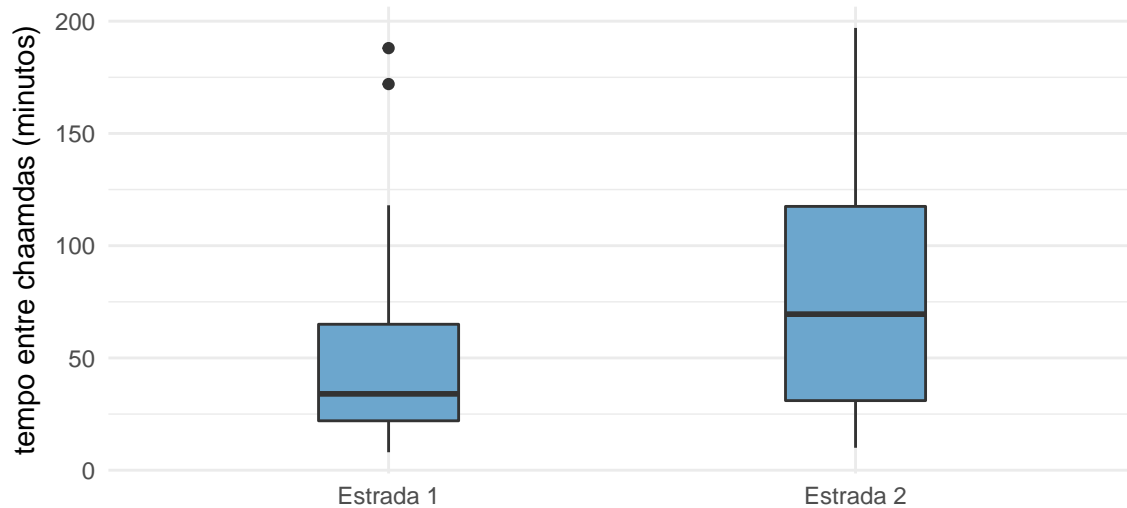
a) Construa um histograma para a distribuição de frequências de chamados em cada uma das estradas:



b) Calcule os intervalos de tempo entre as sucessivas chamadas e descreva-os, para cada uma das estradas, utilizando medidas resumo gráficos do tipo boxplot. Existe alguma relação entre o tipo de estrada e o intervalo de tempo entre as chamadas?

Observando o gráfico a seguir, existem indicativas de que o tempo entre chamadas da estrada 1 é menor que na estrada 2. Isso se dá ao fato de que a mediana do tempo na estrada 1 é menor que na estrada 2.

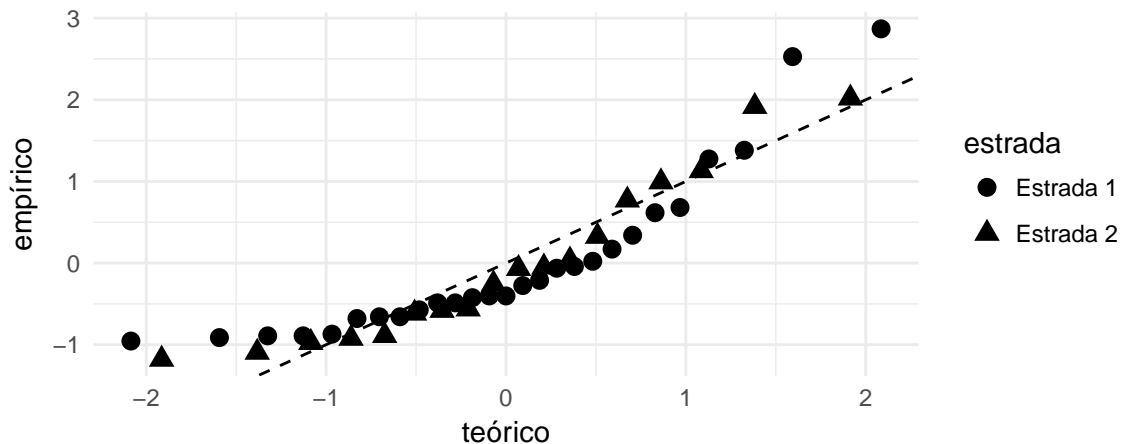
Comparação da distribuição de tempos entre chamadas

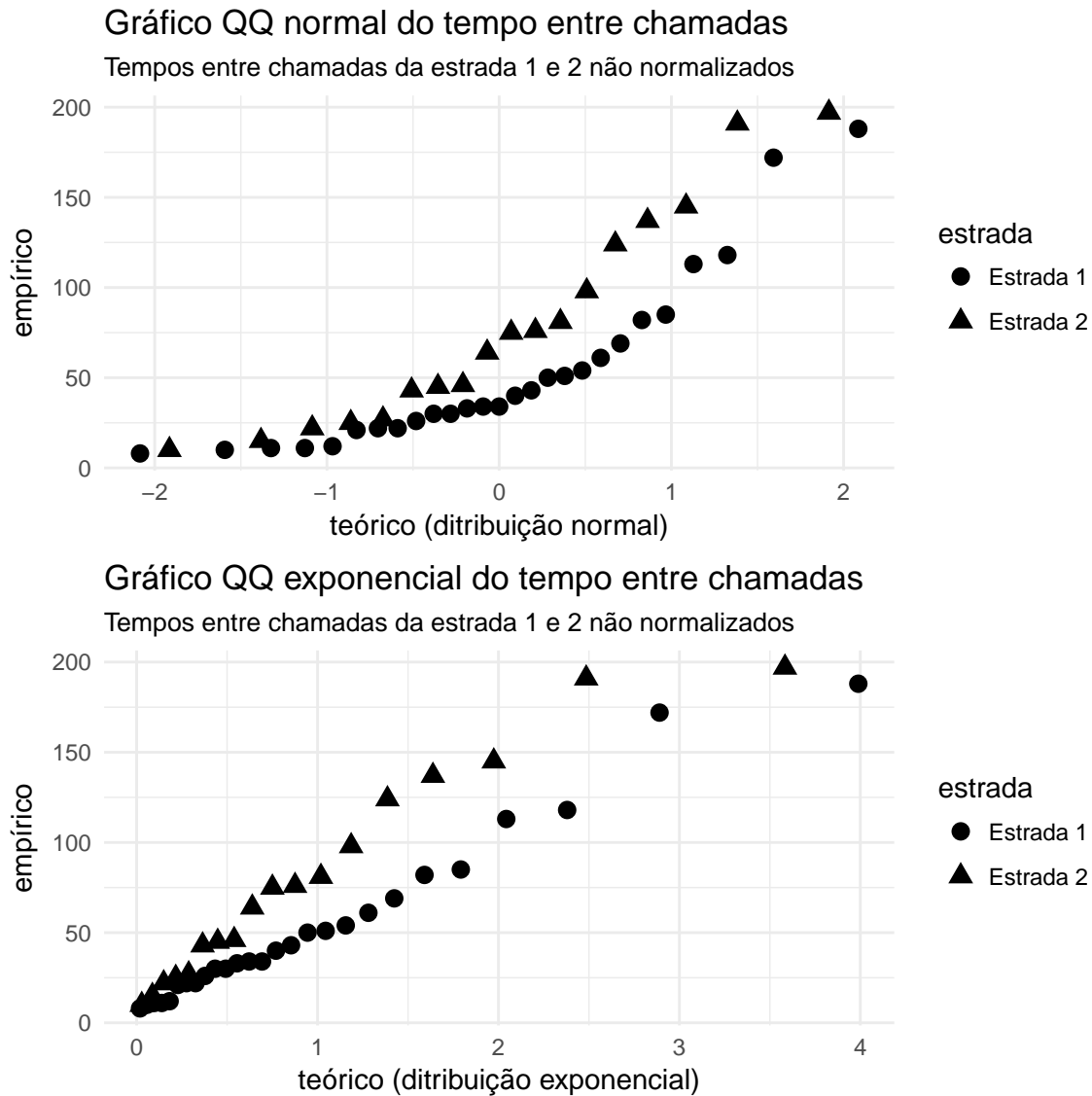


c) Por intermédio de um gráfico QQ, verifique se a distribuição da variável intervalo de tempo entre as chamadas em cada estrada é compatível com um modelo normal. Faça o mesmo para um modelo exponencial. Compare as distribuições de frequências correspondentes às duas estradas.

Gráfico QQ normal do tempo entre chamadas

Tempos entre chamadas da estrada 1 e 2 normalizados.
Linha pontilhada indica a linha $y = x$.





A análise dos gráficos QQ normal e exponencial nos indicam algumas observações sobre o conjunto de dado:

- analisando o gráfico QQ normal com as variáveis normalizadas, podemos ver que ambos os conjuntos de pontos não ficam em linha reta da maneira, indicando que o modelo normal não é mais indicado;
- analisando o QQ normal com as variáveis originais existe um certo descasamento entre os dados da estrada 1 e da estrada 2, indicando que talvez haja diferença na distribuição delas.
- analisando o gráfico QQ exponencial, podemos observar que ambos os conjuntos de dados (estrada 1 e 2) ficam em linha reta, indicando a adequação do modelo exponencial. Além disso, a inclinação dessas retas são possivelmente diferentes, indicando distribuições com parâmetros distintos.

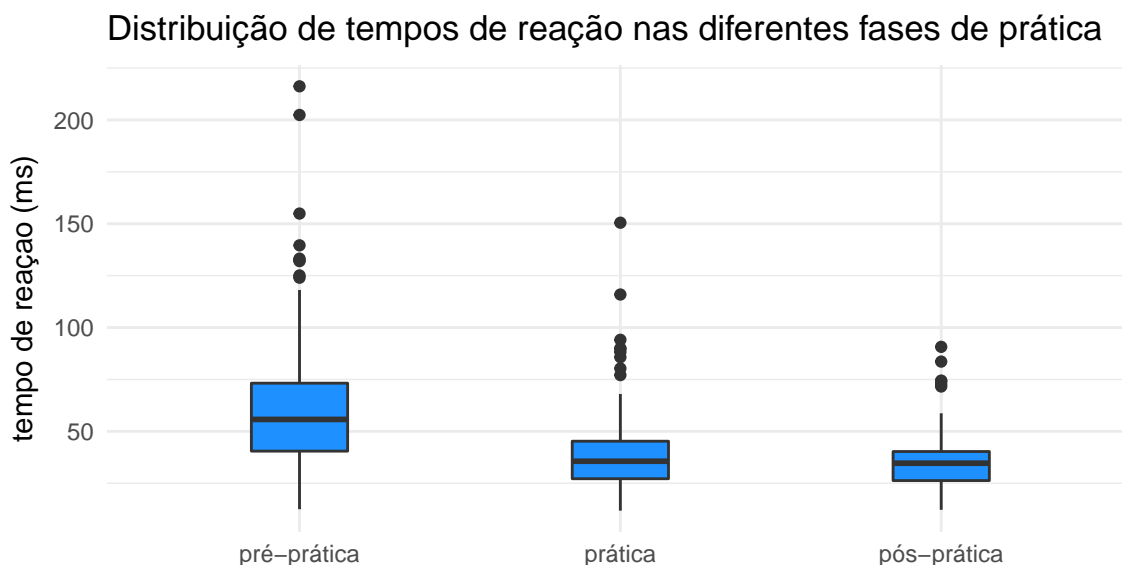
Exercício 4

Este exercício é inspirado em uma análise estatística realizada no Centro de Estatística Aplicada (CEA) do IME/USP. Para resolvê-lo, você precisa baixar o arquivo MAE217_2018-reacao.xls. O principal objetivo do estudo é avaliar o efeito do treinamento de certo tipo de movimento, a ser realizado pela mão dominante dos participantes (mão direita ou esquerda - ambidestros foram descartados). Os 20 participantes, sem prévio treino com relação ao movimento, foram submetidos a certo estímulo e seu tempo de resposta foi anotado (em ms). A fim de obter valores basais (fase pré- prática), inicialmente cada participante repetiu o procedimento 5 vezes. Após um certo período de tempo pré-fixado, o participante voltou e entrou em fase de prática, quando pode repetir o movimento por 8 vezes. Nessa fase, ele recebia informações luminosas cujo objetivo era melhorar o seu reflexo. Após um novo período de tempo pré-fixado, o participante voltou e realizou mais 5 repetições do movimento, sem o auxílio luminoso (fase pós-prática).

Após as transformações necessárias, os dados que foram utilizados ficaram com o seguinte formato:

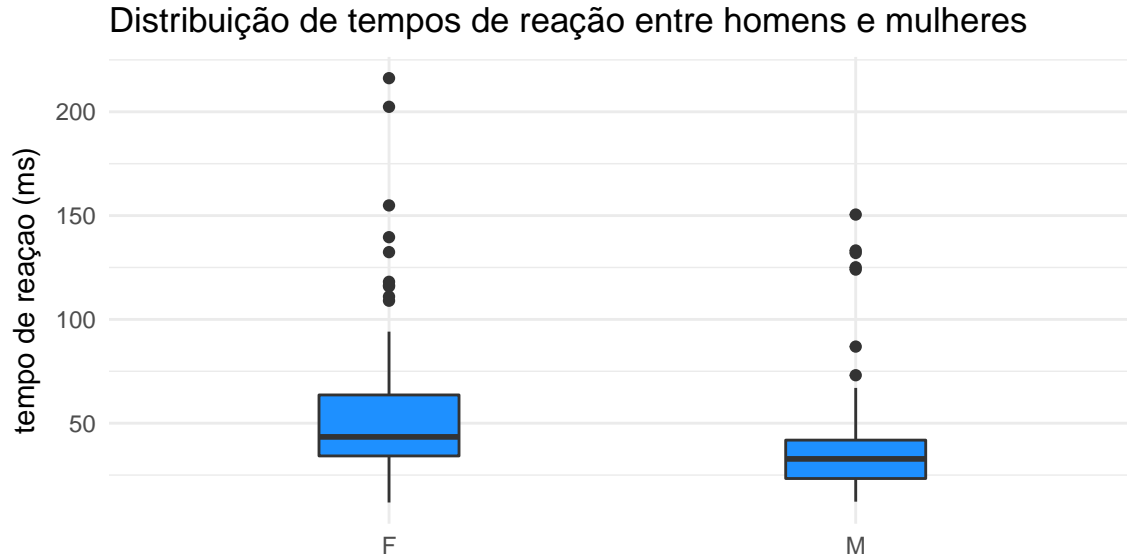
```
## # A tibble: 360 x 7
##   Id Sexo Idade Dom. fase tempo_de_reacao fase_treinamento
##   <dbl> <chr> <dbl> <chr> <chr>         <dbl> <chr>
## 1 1. M 32. D Bloco1      125. pré-prática
## 2 2. M 36. D Bloco1      26.4 pré-prática
## 3 3. M 19. D Bloco1      37.9 pré-prática
## 4 4. F 30. E Bloco1      65.5 pré-prática
## 5 5. F 19. E Bloco1      61.8 pré-prática
## 6 6. F 36. E Bloco1     116. pré-prática
## 7 7. F 31. D Bloco1      31.0 pré-prática
## 8 8. F 26. D Bloco1     118. pré-prática
## 9 9. M 27. E Bloco1      31.8 pré-prática
## 10 10. M 21. D Bloco1      48.5 pré-prática
## # ... with 350 more rows
```

a) construa gráficos do tipo box-plot (um ao lado do outro) para cada uma das 3 fases. O que pode ser dito?



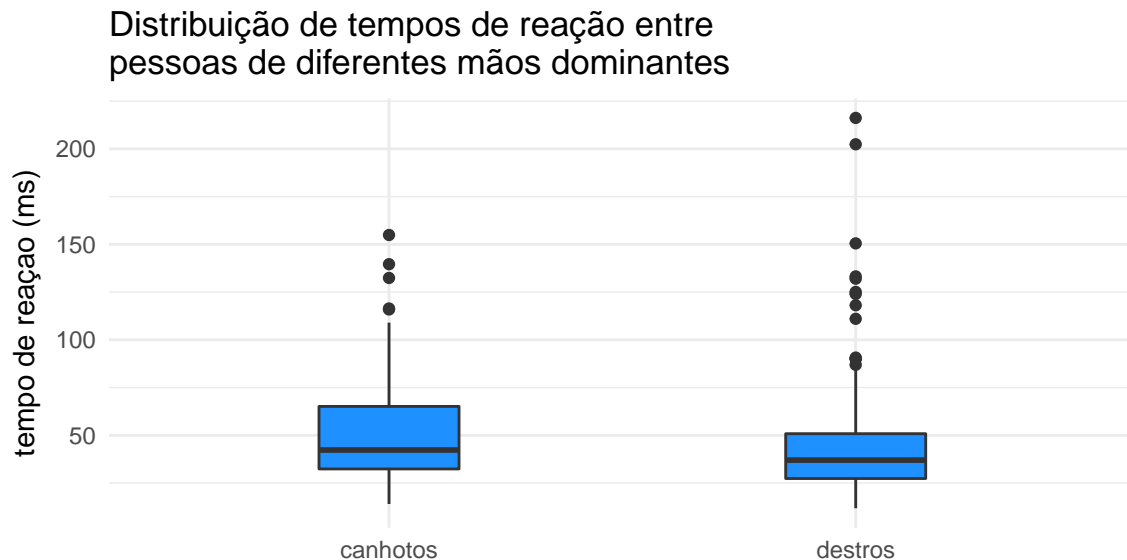
Como evidenciado pela inspeção do boxplot do tempo de reação nas fases de prática, parece haver um ganho considerável no tempo de reação da pré-prática para a prática. Essa diminuição parece não ser tão evidente na comparação do tempo de reação da prática para pós-prática.

b) Repita o item anterior para cada sexo. Existem diferenças nos padrões observados, para cada sexo?



Durante a inspeção do boxplot anterior, podemos observar que o tempo de reação entre pessoas do sexo feminino apresentam uma maior variabilidade, evidenciado pela amplitude do intervalo interquartil maior. Além disso, pessoas do gênero masculino parecem ter um tempo de reação menor, uma vez a mediana desse grupo é menor que a das pessoas do grupo feminino.

c) Repita o item anterior para cada mão dominante. O que pode ser dito?



A inspeção do boxplot anterior indica que pessoas destras tem um tempo de reação ligeiramente menor (evidenciado mais uma vez pela menor mediana), e talvez um pouco menos variáveis que o tempo de reação de pessoas canhotas.

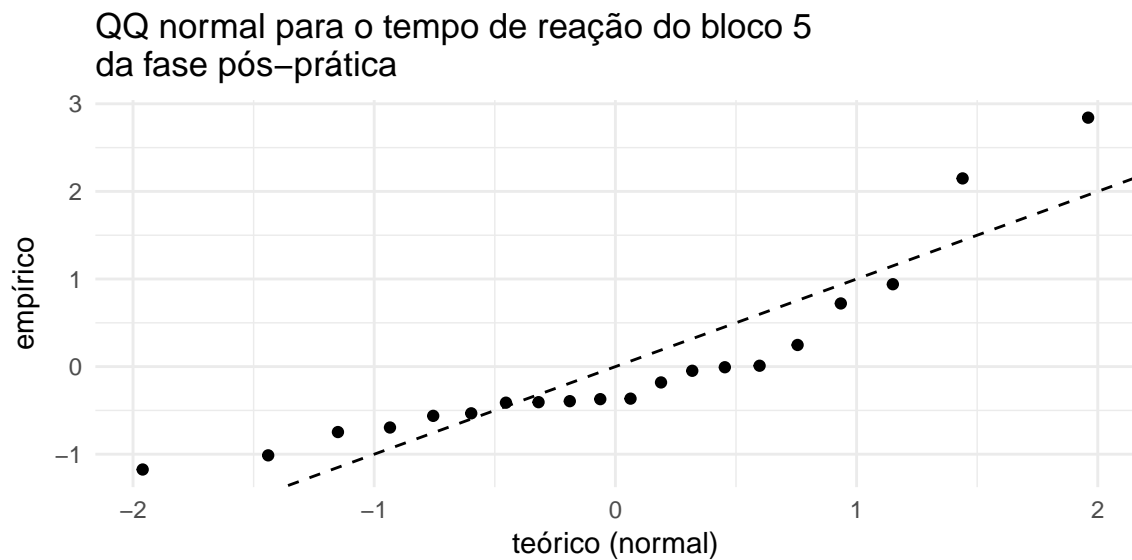
d) Utilizando os coeficientes de assimetria (Bowley e amostral) e curtose, verifique como se comportam as medidas feitas no Bloco 5 da fase pós-prática. Você consideraria o modelo Normal adequado para esses dados?

Conforme podemos ver na tabela abaixo, os valores de estatísticas de assimetria e cortose indicam que os dados:

- tem uma assimetria positiva
- possui um achatamento menor que o esperado para uma normal padrão

```
## # A tibble: 4 x 2
##   estatística      `tempo de reação do bloco 5 do pós-prática`
##   <chr>                <dbl>
## 1 bowey                0.436
## 2 amostral_b          2.63
## 3 amostral_g          1.62
## 4 coef_achatamento    1.57
```

e) Responda à pergunta do item anterior analisando agora um gráfico QQ para o modelo Normal.



Observando o gráfico QQ normal, podemos ver que os pontos não se alinham, indicando que o modelo normal não é o mais indicado para o conjunto de dados de tempos de reação no bloco 5 da fase pós-prática.