

Modelos Lineares I

*Daniel dos Santos
Lyncoln Sousa de Oliveira*

25 de Novembro de 2019

a) Especificando o modelo teórico. (Modelo A ajustado veja as tabelas tabela 1 e tabela 5)

$$Y_i = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \beta_3 X_{i,3} + \beta_4 X_{i,4} + \varepsilon_i; \quad i = 1, 2, \dots, 54$$

Onde:

Y_i : Tempo de vida em escala logarítmica do i -ésimo paciente;

$X_{i,1}$: Idade do i -ésimo paciente;

$$X_{i,2} = \begin{cases} 1, & \text{se o } i\text{-ésimo paciente é do sexo feminino;} \\ 0, & \text{c.c.} \end{cases}$$

$$X_{i,3} = \begin{cases} 1, & \text{se o } i\text{-ésimo paciente tem histórico de uso de álcool severo;} \\ 0, & \text{c.c.} \end{cases}$$

$$X_{i,4} = \begin{cases} 1, & \text{se o } i\text{-ésimo paciente tem histórico de uso de álcool moderado;} \\ 0, & \text{c.c.} \end{cases}$$

β_0 : É intercepto do modelo, neste caso, sem interpretação prática;

β_1 : Variação do tempo de vida em escala logarítmica para cada unidade de idade;

β_2 : Efeito do tempo de vida em escala logarítmica quando o indivíduo é do sexo feminino;

β_3 : Efeito do tempo de vida em escala logarítmica quando o indivíduo possui histórico de álcool severo;

β_4 : Efeito do tempo de vida em escala logarítmica quando o indivíduo possui histórico de álcool moderado;

ε_i : Erro aleatório do i -ésimo paciente;

Onde as hipóteses básicas são:

$$E[\varepsilon_i] = 0; \quad i = 1, 2, \dots, 54;$$

$$Var(\varepsilon_i) = \sigma^2; \quad i = 1, 2, \dots, 54;$$

$$Cor(\varepsilon_i, \varepsilon_j) = 0; \quad i = 1, 2, \dots, 54; \quad i \neq j;$$

$$\varepsilon_i \sim N(0, \sigma^2); \quad i = 1, 2, \dots, 54;$$

Testes de hipóteses para avaliar significância entre as variáveis do estudo. Para $\alpha = 5\%$ e $n = 54$.

Teste de Hipótese para β_0

Passo 1: Definição das Hipóteses.

$$\begin{cases} H_0 : \beta_0 = 0; \\ H_1 : \beta_0 \neq 0; \end{cases}$$

Passo 2: Cálculo da Estatística de teste sob H_0 .

$$T = \frac{\hat{\beta}_0}{\sqrt{\hat{Var}(\hat{\beta}_0)}} \sim T_{49}$$

desta forma,

$$t_{obs} = 22.092;$$

Veja a [tabela 1](#)

Passo 3: Região Crítica.

$$RC = \{t \in R : t > t_{49;0,025} = 2,009575 \text{ ou } t < -t_{49;0,025} = -2,009575\}$$

Passo 4: Tomada de decisão.

Como obteve-se um $t_{obs} = 22.092$, então $t_{obs} \in RC$, desta forma, rejeitamos a hipótese nula a um nível de significância $\alpha = 5\%$. Conclui-se. com base no teste T que o intercepto é significativo.

Teste de Hipótese para β_1

Passo 1: Definição das Hipóteses.

$$\begin{cases} H_0 : \beta_1 = 0; \\ H_1 : \beta_1 \neq 0; \end{cases}$$

Passo 2: Calculo da Estatística de teste sob H_0 .

$$T = \frac{\hat{\beta}_1}{\sqrt{\hat{Var}(\hat{\beta}_1)}} \sim T_{49}$$

onde,

$$\hat{Var}(\hat{\beta}_1) = \frac{\hat{\sigma}^2}{\sum_{i=1}^{54} (X_{i,1} - \bar{X}_1)^2};$$

desta forma,

$$t_{obs} = -0,886;$$

Veja a [tabela 1](#)

Passo 3: Região Crítica.

$$RC = \{t \in R : t > t_{49;0,025} = 2,009575 \text{ ou } t < -t_{49;0,025} = -2,009575\}$$

Passo 4: Tomada de decisão.

Como obteve-se um $t_{obs} = -0,886$, então $t_{obs} \notin RC$ (verifique a [figura 1](#)), desta forma, não rejeitamos a hipótese nula a um nível de significância $\alpha = 5\%$. Conclui-se. com base no teste T que não existe uma relação estatisticamente significativa entre a idade (X_1) e o tempo de vida do paciente em escala logarítmica (Y).

Teste de Hipótese para β_2

Passo 1: Definição das Hipóteses.

$$\begin{cases} H_0 : \beta_2 = 0; \\ H_1 : \beta_2 \neq 0; \end{cases}$$

Passo 2: Calculo da Estatística de teste sob H_0 .

$$T = \frac{\hat{\beta}_2}{\sqrt{\hat{Var}(\hat{\beta}_2)}} \sim T_{49}$$

onde,

$$\hat{Var}(\hat{\beta}_2) = \frac{\hat{\sigma}^2}{\sum_{i=1}^{54} (X_{i,2} - \bar{X}_2)^2};$$

desta forma,

$$t_{obs} = -2,012;$$

Veja a [tabela 1](#)

Passo 3: Região Crítica.

$$RC = \{t \in R : t > t_{49;0,025} = 2,009575 \text{ ou } t < -t_{49;0,025} = -2,009575\}$$

Passo 4: Tomada de decisão.

Como obteve-se um $t_{obs} = -2,012$, então $t_{obs} \in RC$ (verifique a [figura 1](#)), desta forma, rejeitamos a hipótese nula a um nível de significância $\alpha = 5\%$. Conclui-se. com base no teste T que existe uma relação estatisticamente significativa entre o sexo do paciente (X_2) e o tempo de vida em escala logarítmica (Y).

Teste de Hipótese para β_3

Passo 1: Definição das Hipóteses.

$$\begin{cases} H_0 : \beta_3 = 0; \\ H_1 : \beta_3 \neq 0; \end{cases}$$

Passo 2: Calculo da Estatística de teste sob H_0 .

$$T = \frac{\hat{\beta}_3}{\sqrt{\hat{Var}(\hat{\beta}_3)}} \sim T_{49}$$

onde,

$$\hat{Var}(\hat{\beta}_3) = \frac{\hat{\sigma}^2}{\sum_{i=1}^{54} (X_{i,3} - \bar{X}_3)^2};$$

desta forma,

$$t_{obs} = -2,1891;$$

Veja a [tabela 1](#)

Passo 3: Região Crítica.

$$RC = \{t \in R : t > t_{49;0,025} = 2,009575 \text{ ou } t < -t_{49;0,025} = -2,009575\}$$

Passo 4: Tomada de decisão.

Como obteve-se um $t_{obs} = -2,1891$, então $t_{obs} \in RC$ (verifique a [figura 1](#)), desta forma, rejeitamos a hipótese nula a um nível de significância $\alpha = 5\%$. Conclui-se. com base no teste T que existe uma relação estatisticamente significativa entre o consumo moderado de álcool (X_3) e o tempo de vida em escala logarítmica (Y).

Teste de Hipótese para β_4

Passo 1: Definição das Hipóteses.

$$\begin{cases} H_0 : \beta_4 = 0; \\ H_1 : \beta_4 \neq 0; \end{cases}$$

Passo 2: Cálculo da Estatística de teste sob H_0 .

$$T = \frac{\hat{\beta}_4}{\sqrt{\hat{Var}(\hat{\beta}_4)}} \sim T_{49}$$

onde,

$$\hat{Var}(\hat{\beta}_4) = \frac{\hat{\sigma}^2}{\sum_{i=1}^{54} (X_{i,4} - \bar{X}_4)^2};$$

desta forma,

$$t_{obs} = -2,012;$$

Veja a [tabela 1](#)

Passo 3: Região Crítica.

$$RC = \{t \in R : t > t_{49;0,025} = 2,009575 \text{ ou } t < -t_{49;0,025} = -2,009575\}$$

Passo 4: Tomada de decisão.

Como obteve-se um $t_{obs} = -2,619$, então $t_{obs} \in RC$ (verifique a figura 1), desta forma, rejeitamos a hipótese nula a um nível de significância $\alpha = 5\%$. Conclui-se. com base no teste T que existe uma relação estatisticamente significativa entre o consumo severo de álcool (X_4) e o tempo de vida em escala logarítmica (Y).

Escolha do modelo Pelo teste de hipótese T de significância individual que a variável “Idade” não possui efeito estatisticamente significativo, logo prefere-se a escolha de um modelo que não leva em conta a variável Idade. Escolhe-se o modelo que leva em conta as variáveis dummy: “Sexo”, “Histórico de álcool”, para um nível alfa de significância de 5%, todas essas variáveis possuem efeito estatisticamente significativo.

O modelo adotado é: (Modelo B ajustado veja as tabelas [tabela 2](#) e [tabela 6](#))

$$Y_i = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \beta_3 X_{i,3} + \varepsilon_i; \quad i = 1, 2, \dots, 54$$

Onde:

Y_i : Tempo de vida em escala logarítmica do i -ésimo paciente;

$$X_{i,1} = \begin{cases} 1, & \text{se o } i\text{-ésimo paciente é do sexo feminino;} \\ 0, & \text{c.c.} \end{cases}$$

$$X_{i,2} = \begin{cases} 1, & \text{se o } i\text{-ésimo paciente tem histórico de uso de álcool severo;} \\ 0, & \text{c.c.} \end{cases}$$

$$X_{i,3} = \begin{cases} 1, & \text{se o } i\text{-ésimo paciente tem histórico de uso de álcool moderado;} \\ 0, & \text{c.c.} \end{cases}$$

β_0 : É intercepto do modelo, é o efeito causado no tempo de vida em escala logarítmica, quando o paciente é do sexo feminino e não possui histórico de álcool;

β_1 : Efeito do tempo de vida em escala logarítmica quando o indivíduo é do sexo feminino;

β_2 : Efeito do tempo de vida em escala logarítmica quando o indivíduo possui histórico de álcool severo;
 β_3 : Efeito do tempo de vida em escala logarítmica quando o indivíduo possui histórico de álcool moderado;
 ε_i : Erro aleatório do i -ésimo paciente;

Onde as hipóteses básicas são:

$$\begin{aligned} E[\varepsilon_i] &= 0; \quad i = 1, 2, \dots, 54; \\ Var(\varepsilon_i) &= \sigma^2; \quad i = 1, 2, \dots, 54; \\ Cor(\varepsilon_i, \varepsilon_j) &= 0; \quad i = 1, 2, \dots, 54; \quad i \neq j; \\ \varepsilon_i &\sim N(0, \sigma^2); \quad i = 1, 2, \dots, 54; \end{aligned}$$

b) Avaliar a significância do efeito global das variáveis “Histórico de álcool” e “Sexo” usando o teste F de comparabilidade de modelos.

Teste de Hipótese para comparação do modelo reduzido que leva em conta apenas a variável “Sexo” (Modelo C ajustado veja as tabelas [tabela 3](#) e [tabela 7](#)) contra o modelo completo, que leva em conta as variáveis “Sexo” e “Histórico de álcool”

Passo 1: Definição das Hipóteses.

$$\begin{cases} H_0 : Y = \beta_0 + \beta_1 X_{i,1} + \varepsilon_i; \\ H_1 : Y = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \beta_3 X_{i,3} + \varepsilon_i \end{cases} \quad i = 1, 2, \dots, 54$$

Passo 2: Calculo da Estatística de teste sob H_0 .

$$F = \frac{\frac{SQres_0 - SQres_1}{4-2}}{\frac{SQres_1}{54-4}} \sim F_{2,50}$$

desta forma,

$$f_{obs} = 4,9108$$

Passo 3: Região Crítica.

$$RC = \{f \in R : f > f_{2,50;0.05} = 3,18261\}$$

Passo 4: Tomada de decisão.

Como obteve-se um $f_{obs} = 4,9108$, então $f_{obs} \in RC$ (verifique a [figura 2](#)), desta forma, rejeitamos a hipótese nula a um nível de significância $\alpha = 5\%$. Ou seja, o modelo completo que leva em conta as variáveis “Sexo” e “Histórico de álcool” é mais adequado que o modelo reduzido que leva em conta somente a variável “Sexo”. Conclui-se que a variável “Histórico de Álcool” possui efeito estatisticamente significante.

Teste de Hipótese para comparação do modelo reduzido que leva em conta apenas a variável “Histórico de álcool” (Modelo D ajustado veja as tabelas [tabela 4](#) e [tabela 8](#)) contra o modelo completo, que leva em conta as variáveis “Sexo” e “Histórico de álcool”

Passo 1: Definição das Hipóteses.

$$\begin{cases} H_0 : Y = \beta_0 + \beta_2 X_{i,2} + \beta_3 X_{i,3} + \varepsilon_i; \\ H_1 : Y = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \beta_3 X_{i,3} + \varepsilon_i \end{cases} \quad i = 1, 2, \dots, 54$$

Passo 2: Calculo da Estatística de teste sob H_0 .

$$F = \frac{\frac{SQres_0 - SQres_1}{4-3}}{\frac{SQres_1}{54-4}} \sim F_{1,50}$$

desta forma,

$$f_{obs} = 4,0743$$

Passo 3: Região Crítica.

$$RC = \{f \in R : f > f_{1,50;0.05} = 4,03431\}$$

Passo 4: Tomada de decisão.

Como obteve-se um $f_{obs} = 4,0743$, então $f_{obs} \in RC$ (verifique a [figura 3](#)), desta forma, rejeitamos a hipótese nula a um nível de significância $\alpha = 5\%$. Ou seja, o modelo completo que leva em conta as variáveis “Sexo” e “Histórico de álcool” é mais adequado que o modelo reduzido que leva em conta somente a variável “Histórico de álcool”. Conclui-se que a variável “Sexo” possui efeito estatisticamente significante.

c) Verificação das hipóteses básicas do modelo.

Serão utilizadas as [figuras 4 e 5](#) para observar se há alguma violação algumas das hipóteses básicas dos modelos, que são: Homocedasticidade e normalidade. Será suposto a independência dos erros aleatórios do modelo.

- Homocedasticidade:
Pela [figura 4](#) é possível notar uma nuvem de pontos aleatórios em torno de 0(zero), o que indica que os erros aleatórios possuem variâncias constante. Também é possível notar que não há presença de um padrão sistemático dos pontos, o que indica que a hipótese de linearidade não foi violada. O que foi corroborado pelo teste de Breusch-Pagan, obteve-se um P-valor = 0.795
- Normalidade:
Pela [figura 5](#), pode-se notar que os quantis dos resíduos studentizados se aproximam dos quantis teóricos de uma distribuição normal, o que é um bom indicativo de normalidade dos erros. Para corroborar o resultado foi realizado um teste de Shapiro-Wilk nos resíduos studentizados e obteve-se um P-valor = 0.8599

d) Modelo selecionado e coeficiente de determinação do modelo

O modelo ajustado:

$$\hat{Y}_i = 6,7121 + 0,2484X_{i,1} - 0,5404X_{i,2} - 0,4589X_{i,3}$$

Onde:

\hat{Y}_i : Tempo de vida estimado em escala logarítmica do i-ésimo paciente;

$$X_{i,1} = \begin{cases} 1, & \text{se o i-ésimo paciente é do sexo feminino;} \\ 0, & \text{c.c.} \end{cases}$$

$$X_{i,2} = \begin{cases} 1, & \text{se o i-ésimo paciente tem histórico de uso de álcool severo;} \\ 0, & \text{c.c.} \end{cases}$$

$$X_{i,3} = \begin{cases} 1, & \text{se o i-ésimo paciente tem histórico de uso de álcool moderado;} \\ 0, & \text{c.c.} \end{cases}$$

$\hat{\beta}_0 = 6,7121$: É o valor estimado do tempo de vida em escala logarítmica quando o paciente é do sexo masculino e nunca ingeriu bebida alcólica;

$\hat{\beta}_1 = 0,2484$: O valor estimado do tempo de vida em escala logarítmica para o i-ésimo indivíduo é maior quando ele é do sexo feminino comparado com o sexo masculino e esse aumento é de 0.2484;

$\hat{\beta}_2 = -0,5404$: O valor estimado do tempo de vida em escala logarítmica para o i-ésimo indivíduo é menor quando ele possui histórico de álcool severo comparado com quem não possui histórico de álcool, e essa redução é de -0,5404;

$\hat{\beta}_3 = -0,4589$: O valor estimado do tempo de vida em escala logarítmica para o i-ésimo indivíduo é menor quando ele possui histórico de álcool moderado comparado com quem não possui histórico de álcool, e essa redução é de -0,4589;

Conclusões

Por não apresentar nenhuma violação nas hipóteses básicas, o modelo é adequado para representar os dados observados, porém seu coeficiente de determinação (R^2) é de apenas 20.92%, ou seja, 20.92 da variação é explicada pelo modelo ajustado que leva em consideração as variáveis sexo e histórico de álcool.

É importante observar que para o “Modelo A” a variável “Idade” para o teste T de significância individual apresentou P-valor > 0.05 , ou seja, a um nível de significância de 5% a variável não apresentou uma relação estatisticamente significativa. Pelo princípio da parcimônia, opta-se pelo modelo reduzido B, que leva em conta as variáveis “Sexo” e “Histórico de álcool”. Foi utilizado este modelo (B) como modelo completo para realização dos teste do exercício b).

Apêndice 1 - Tabelas

Tabela 1: Tabela Modelo A

Estimadores	Estimativa	Erro padrão	Valor da estatística de teste T	P-valor
$\hat{\beta}_0$	6,956899	0,314901	22,092	< 0,001
$\hat{\beta}_1$	-0,004994	0,005638	-0,886	0,38011
$\hat{\beta}_2$	0,248113	0,123322	2,012	0,04974
$\hat{\beta}_3$	-0,533359	0,184475	-2,891	0,00571
$\hat{\beta}_4$	-0,438230	0,167344	-2,619	0,01171

Tabela 2: Tabela Modelo B

Estimadores	Estimativa	Erro padrão	Valor da estatística de teste T	P-valor
$\hat{\beta}_0$	6,7121	0,1506	44,574	< 0,001
$\hat{\beta}_1$	0,2484	0,1231	2,018	0,04892
$\hat{\beta}_2$	-0,5404	0,1839	-2,2938	0,00498
$\hat{\beta}_3$	-0,4589	0,1653	-2,775	0,00773

Tabela 3: Tabela Modelo C

Estimadores	Estimativa	Erro padrão	Valor da estatística de teste T	P-valor
$\hat{\beta}_0$	6,32567	0,08963	70,57	< 0,001
$\hat{\beta}_1$	0,22653	0,13173	1,72	0,0915

Tabela 4: Tabela Modelo D

Estimadores	Estimativa	Erro padrão	Valor da estatística de teste T	P-valor
$\hat{\beta}_0$	6,8115	0,1465	46,482	< 0,001
$\hat{\beta}_1$	-0,5238	0,1892	-2,769	0,00782
$\hat{\beta}_2$	-0,4384	0,1699	-2,580	0,01282

Tabela 5: Tabela ANOVA Modelo A

Fontes de variação	Soma dos quadrados	gl	Quadrado médio	Valor da estatística de teste F	P-valor
Idade	0,2694	1	0,26941	1,3245	0.25538
Sexo	0,6940	1	0,69398	3,4118	0.07077
Álcool severo	0,4794	1	0,47937	2,3567	0.13118
Álcool moderado	1,3949	1	1,39491	6,8578	0.01171
Resíduos	9,9668	49	0,20340		

Tabela 6: Tabela ANOVA Modelo B

Fontes de variação	Soma dos quadrados	gl	Quadrado médio	Valor da estatística de teste F	P-valor
Sexo	0,6889	1	0,68895	3,4017	0.071056
Álcool severo	0,4291	1	0,42907	2,1186	0.151774
Álcool moderado	1,5601	1	1,56009	7,7031	0.007735
Resíduos	10,1264	50	0,20253		

Tabela 7: Tabela ANOVA Modelo C

Fontes de variação	Soma dos quadrados	gl	Quadrado médio	Valor da estatística de teste F	P-valor
Sexo	0,6889	1	0,68895	2,957	0.09146
Resíduos	12,1156	52	0,23299		

Tabela 8: Tabela ANOVA Modelo D

Fontes de variação	Soma dos quadrados	gl	Quadrado médio	Valor da estatística de teste F	P-valor
Álcool severo	0,4241	1	0,42406	1,9748	0.16600
Álcool moderado	1,4289	1	1,42888	6,6541	0.01282
Resíduos	10,9516	51	0,21474		

Apêndice 2 - Figuras

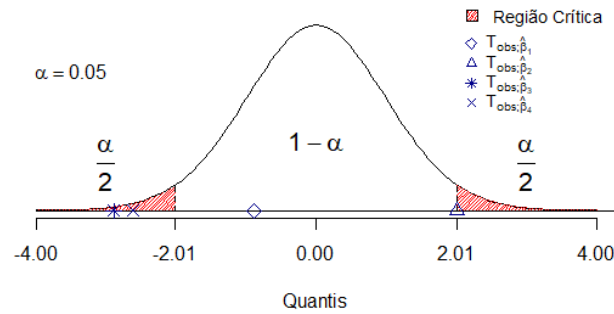


Figura 1: Densidade e região crítica de T_{49} .

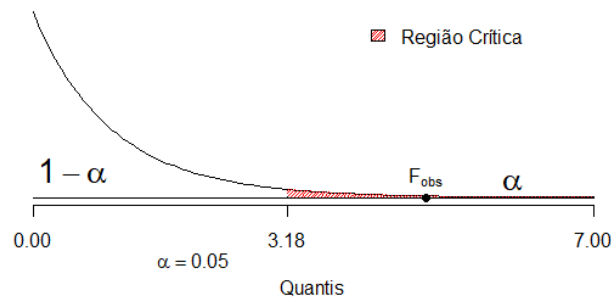


Figura 2: Densidade e região crítica de $F_{2,50}$.

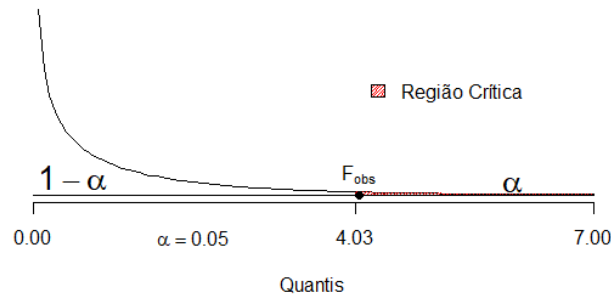


Figura 3: Densidade e região crítica de $F_{1,50}$.

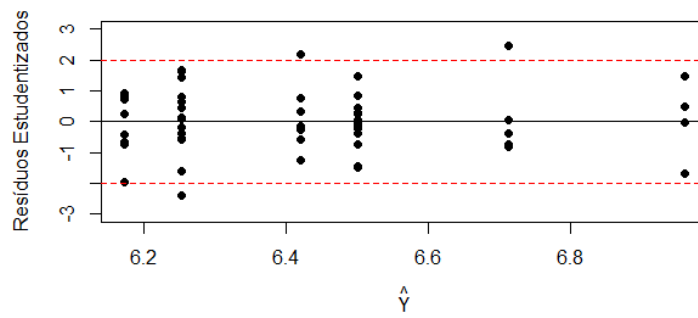


Figura 4: Dispersão dos resíduos estudentizados e valores estimados.

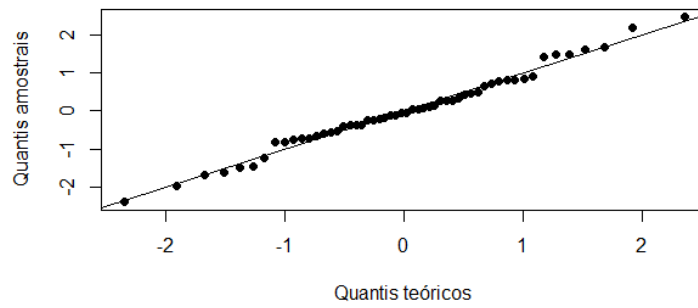


Figura 5: Gráfico quantil-quantil

Apêndice 3 - Códigos

Resolução das questões

```
require(dplyr)
data = foreign::read.spss("tvida.sav", to.data.frame = TRUE)
head(data)

colnames(data) = tolower(colnames(data))

data = data %>% mutate(sexo = ifelse(sexo == "Mulher", 1, 0),
                             alcool_severo = ifelse(historia_alcool == "Uso severo", 1, 0),
                             alcool_moderado = ifelse(historia_alcool == "Uso moderado", 1, 0)) %>%
  select(-historia_alcool) %>%
  select(ln_tempo_vida, everything())
head(data)
# a) -----
modelo_1 = lm(ln_tempo_vida ~ idade + sexo + alcool_severo + alcool_moderado, data); summary(modelo_1)
qt(0.05/2, 54-5, lower.tail = F)
# Variável idade não se mostrou significativa a um nível de 5%.
modelo_2 = lm(ln_tempo_vida ~ sexo + alcool_severo + alcool_moderado, data)
# b) -----
modelo_3 = lm(ln_tempo_vida ~ sexo, data); summary(modelo_3)
anova(modelo_3, modelo_2)
modelo_4 = lm(ln_tempo_vida ~ alcool_severo + alcool_moderado, data); summary(modelo_4)
anova(modelo_4, modelo_2)
# c) -----
ris = rstandard(modelo_2)
ychapeu = fitted(modelo_2)
plot(ris~ychapeu,
     pch = 19,
     ylim = c(-3,3),
     ylab = "Resíduos Estudentizados",
     xlab = latex2exp::TeX("$\\hat{Y}$"))
abline(h = c(-2,0,2),
       lty = c(2,1,2),
       col = c("red", "black", "red"))
lmtest::bptest(modelo_2)
qqnorm(ris,
       pch = 19,
       main = " ",
       ylab = "Quantis amostrais",
       xlab = 'Quantis teóricos')
abline(0,1)
shapiro.test(ris)
# d) -----
summary(modelo_2)
summary(modelo_2)$r.squared
```

Gerando o gráfico da densidade da F.

```
degree_1 = 2
degree_2 = 50
quantile = qf(0.05, df1 = degree_1, df2 = degree_2, lower.tail = F)
test_stat = 4.9108
rc_values = seq(quantile, 7, length = 54)
denisty_rc_values = df(rc_values, df1 = degree_1, df2 = degree_2)
ic_values = seq(0, quantile, length = 54)
denisty_ic_values = df(ic_values, df1 = degree_1, df2 = degree_2)
plot(
  function(x)
    df(x,
      df1 = degree_1,
      df2 = degree_2,
      xlim = c(0, 7),
      ylab = '',
      xlab = 'Quantis',
      bty="n",
      yaxt='n',
      xaxt='n'
    )
  axis(side=1, at=round(c(0, quantile, 7), 2))
  polygon(
    x = c(quantile, rc_values, 7),
    y = c(0, denisty_rc_values, 0),
    border = FALSE,
    col = 'red',
    density = 50
  )
  lines(
    x = c(quantile, quantile),
    y = c(0, denisty_rc_values[length(denisty_rc_values)]),
    lty = 2
  )
  lines(x=c(0, 7), y=c(0,0))
  text(6, 0.09, expression(alpha), cex=1.7)
  text(0.5, 0.15, expression(1 - alpha), cex=1.7)
  par(xpd=TRUE)
  points(test_stat, 0, pch = 19)
  text(test_stat, 0, latex2exp::TeX("$F_{obs}$"), pos = 3)
  text(2, -0.33, expression(alpha == 0.05))
  legend(4, 1, legend = c('Região Crítica'), box.col = "white",
    fill = c('red'),
    density = 50)
  par(xpd=FALSE)
  degree_1 = 1
  degree_2 = 50
  quantile = qf(0.05, df1 = degree_1, df2 = degree_2, lower.tail = F)
  test_stat = 4.0743
  rc_values = seq(quantile, 7, length = 54)
  denisty_rc_values = df(rc_values, df1 = degree_1, df2 = degree_2)
  ic_values = seq(0, quantile, length = 54)
  denisty_ic_values = df(ic_values, df1 = degree_1, df2 = degree_2)
```

```

plot(
  function(x)
    df(x,
      df1 = degree_1,
      df2 = degree_2),
  xlim = c(0, 7),
  ylab = '',
  xlab = 'Quantis',
  bty="n",
  yaxt='n',
  xaxt='n'
)
axis(side=1, at=round(c(0,quantile, 7), 2))
polygon(
  x = c(quantile, rc_values, 7),
  y = c(0, denisty_rc_values, 0),
  border = FALSE,
  col = 'red',
  density = 50
)
lines(
  x = c(quantile, quantile),
  y = c(0, denisty_rc_values[length(denisty_rc_values)]),
  lty = 2
)
lines(x=c(0, 7), y=c(0,0))
text(6, 0.09, expression(alpha), cex=1.7)
text(0.5, 0.15, expression(1 - alpha), cex=1.7)
par(xpd=TRUE)
points(test_stat, 0, pch = 19)
text(test_stat, 0, latex2exp::TeX("$F_{obs}$"), pos = 3)
text(2, -0.33, expression(alpha == 0.05))
legend(4, 1, legend = c('Região Crítica'), box.col = "white",
  fill = c('red'),
  density = 50)
par(xpd=FALSE)

```

Gerando o gráfico da densidade da T.

```
degree = 49
quantile = qt(0.975, df = degree)
b0_test_stat = 22.092
b1_test_stat = -0.886
b2_test_stat = 2.012
b3_test_stat = -2.891
b4_test_stat = -2.619
rc_values = seq(-4, -quantile, length = 54)
denisty_rc_values = dt(rc_values, df = 49)
ic_values = seq(-quantile, quantile, length = 54)
denisty_ic_values = dt(ic_values, df = 49)
plot(
  function(x)
    dt(x, df = 46),
  xlim = c(-4, 4),
  ylab = '',
  xlab = 'Quantis',
  bty="n",
  yaxt='n',
  xaxt='n'
)
axis(side=1, at=round(c(-4, -quantile, 0, quantile, 4), 2))
polygon(
  x = c(-4, rc_values, -quantile),
  y = c(0, denisty_rc_values, 0),
  border = FALSE,
  col = 'red',
  density = 50
)
polygon(
  x = c(quantile, sort(-1 * rc_values), 7),
  y = c(0, sort(denisty_rc_values, decreasing = TRUE), 0),
  border = FALSE,
  col = 'red',
  density = 50
)
lines(
  x = c(-quantile, -quantile),
  y = c(0, denisty_rc_values[length(denisty_rc_values)]),
  lty = 2
)
lines(
  x = c(quantile, quantile),
  y = c(0, denisty_rc_values[length(denisty_rc_values)]),
  lty = 2
)
lines(x=c(-4, 7), y=c(0,0))
points(b1_test_stat, 0, pch=5, cex = 1.2, col = "darkblue")
points(b2_test_stat, 0, pch=2, cex = 1.2, col = "darkblue")
points(b3_test_stat, 0, pch=8, cex = 1.2, col = "darkblue")
points(b4_test_stat, 0, pch=4, cex = 1.2, col = "darkblue")
text(-3, 0.1, expression(frac(alpha, 2)), cex=1.3)
```

```

text(3, 0.1, expression(frac(alpha,2)), cex=1.3)
text(0, 0.15, expression(1 - alpha), cex=1.3)
par(xpd=TRUE)
text(-3.5, 0.3, expression(alpha == 0.05))
legend(2, 0.4, legend = c(latex2exp::TeX("$T_{obs; \\hat{\\beta}_1}$"),
                          latex2exp::TeX("$T_{obs; \\hat{\\beta}_2}$"),
                          latex2exp::TeX("$T_{obs; \\hat{\\beta}_3}$"),
                          latex2exp::TeX("$T_{obs; \\hat{\\beta}_4}$")),
      pch = c(5,2,8,4),
      col = 'darkblue',
      bty = 'n',
      cex = 0.85)
legend(1.93, 0.4597, legend = 'Região Crítica',
      fill = 'red',
      density = 50,
      bty = 'n',
      cex = 0.85)
par(xpd=FALSE)

```