Lista 1 - MAE0514

Guilherme NºUSP: 8943160 e Leonardo NºUSP: 9793436

Exercício 1

Seja T uma variável aleatória contínua não negativa com função densidade de probabilidade $f_T(t)$, função de Sobrevivência $S_T(t)$ e função de risco $\alpha_T(t)$. Mostre que

$$S_T(t) = exp\Big\{ - \int_0^t \alpha_T(s)ds \Big\}$$

Resolução

Como sabemos que

$$S_T(t) = \mathbb{P}(T \geq t) = 1 - F_T(t)$$

Utilizando o resultado dado em aula, temos:

$$\alpha_T(s) = \frac{f_T(t)}{S_T(t)}$$

E que:

$$f_T(t) = \frac{d(1 - F_T(t))}{dt} = -\frac{dS_T(t)}{dt} = -S'_T(t)$$

Seja $A_T(t) = \int_0^t \alpha_T(s)ds$, logo

$$A_T(t) = \int_0^t \frac{f_T(s)}{S_T(s)}ds = \int_0^t \frac{-S'_T(s)}{S_T(s)}ds = -[ln(S_T(s))|_0^t] = -[ln(S_T(t)) - ln(S_T(0))] = -ln(S_T(t))$$

Podemos notar que

$$exp\{-A_T(t)\} = exp\Big\{ - \int_0^t \alpha_T(s)ds \Big\} =$$

E pela propriedade dos logritmos:

$$exp\Big\{ - (-ln(S_T(t)))\Big\} = S_T(t) \Rightarrow S_T(t) = \Big\{ - \int_0^t \alpha_T(s)ds \Big\}$$

Exercício 2

Prove os seguintes resultados:

(a) Se $\alpha(t) = \alpha_1(t) + \alpha_2(t)$, em que $\alpha_1(t)$ e $\alpha_2(t)$ são funções de taxa de falha de tempos de falha independentes T_1 e T_2, então T tem a mesma distribuição de $min(T_1, T_2)$.

Resolução

Utilizando o resultado do exercício anterior, temos:

$$S(t) = exp\Big\{-\int_0^t \alpha(u)\ du\Big\}$$

Logo, se $\alpha(t) = \alpha_1(t) + \alpha_2(t)$ então

$$S(t) = exp\Big\{-\int_0^t \alpha_1(u) + \alpha_2(u)\ du\Big\}$$

E sabendo que $\alpha(t) = \frac{f(t)}{S(t)}$

$$\Rightarrow exp\Big\{-\int_0^t \frac{f_1(u)}{S_1(u)} + \frac{f_2(u)}{S_2(u)}\ du\Big\} = exp\Big\{-\Big[\int_0^t \frac{f_1(u)}{S_1(u)}\ du + \int_0^t \frac{f_2(u)}{S_2(u)}\ du\Big]\Big\} =$$

$$exp\Big\{-\int_0^t \frac{f_1(u)}{S_1(u)}\ du\Big\}exp\Big\{\int_0^t \frac{f_2(u)}{S_2(u)}\ du\Big\} = exp\Big\{-\int_0^t \alpha_1(t)\ du\Big\}exp\Big\{\int_0^t \alpha_2(t)\ du\Big\} = S_1(t) * S_2(t)$$

$$\Rightarrow S(t) = S_1(t) * S_2(t) = (1 - F_1(t))(1 - F_2(t))$$

Agora vamos verificar a distribuição de $min(T_1, T_2)$

Seja $T = min(T_1, T_2)$ temos que:

$$\mathbb{P}(T = min(T_1, T_2) \le t) \Rightarrow \mathbb{P}(T > t) = \mathbb{P}(T_1 > t, T_2 > t) \overset{ind}{=} \mathbb{P}(T_1 > t)\mathbb{P}(T_2 > t) = (1 - \mathbb{P}(T_1 \le t))(1 - \mathbb{P}(T_2 \le t))$$

$$\Rightarrow \mathbb{P}(T = min(T_1, T_2) \le t) = (1 - F_1(t))(1 - F_2(t))$$

Como podemos ver os resultados coincidem logo T tem a mesma distribuição de $min(T_1, T_2)$.

(b) Se T é variável aleatória contínua com função de sobrevivência $S(t)$, então vale

$$\mathbb{E}(T-t|T>t) = \frac{\int_0^\infty S(u)du}{S(t)}$$

.

Resolução

Podemos escrever $\mathbb{E}(T-t|T>t) = \int_t^\infty \frac{(u-t)f(u)}{S(t)}du = \frac{1}{S(t)}\int_t^\infty (u-t)f(u)\ du$

Do exercício 1, temos o resultado:

$$f(t) = -S'(t)$$

Logo,

$$\frac{1}{S(t)}\int_t^\infty -(u-t)S'(u)du$$

Agora integrando por partes, com:

$$\begin{cases} g' = S'(u)\ e\ f = (u-t) \\ g = S(u)\ e\ f' = -1 \end{cases}$$

Temos:

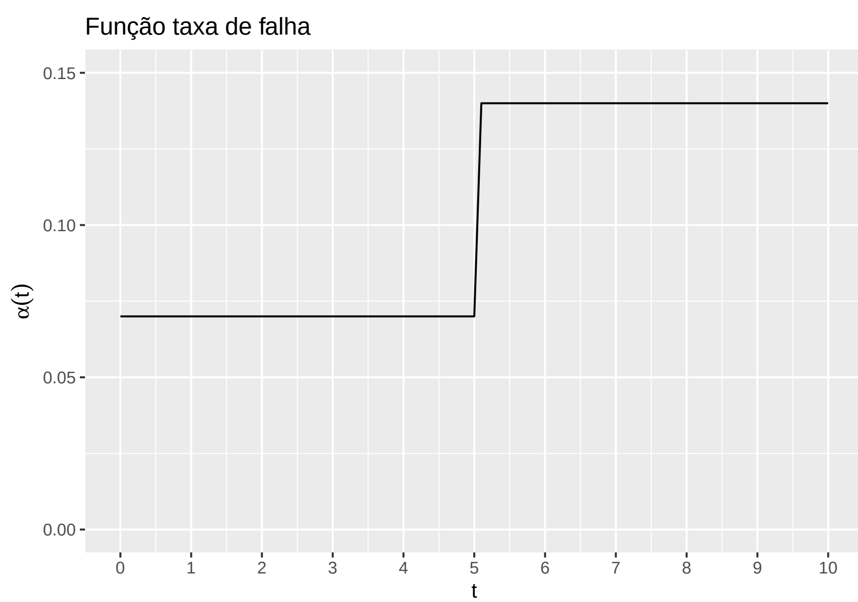
$$\frac{1}{S(t)}\left\{\mathbb{E}(u-t)S(u)|_0^\infty + \int_t^\infty S(u)du\right\} = \frac{1}{S(t)}\left\{0 + \int_t^\infty S(u)du\right\} \Rightarrow \mathbb{E}(T-t|T>t) = \frac{\int_0^\infty S(u)du}{S(t)}$$

Exercício 3

Considere a distribução com função de taxa de falha constante $\lambda = 0.07$ no intervalo $t = 0$ até $t = 5$ e constante igual a $\lambda = 0.14$ para $t > 5$.

(a) Faça o gráfico da função de taxa de falha.

Resolução



(b) Obtenha a função de sobrevivência e faça o gráfico da função obtida.

Resolução

A função de sobrevivência pode ser obtida através do resultado do exercício 1 $S(t) = exp\Big\{ - \int_0^t \alpha(u)\ du \Big\}$

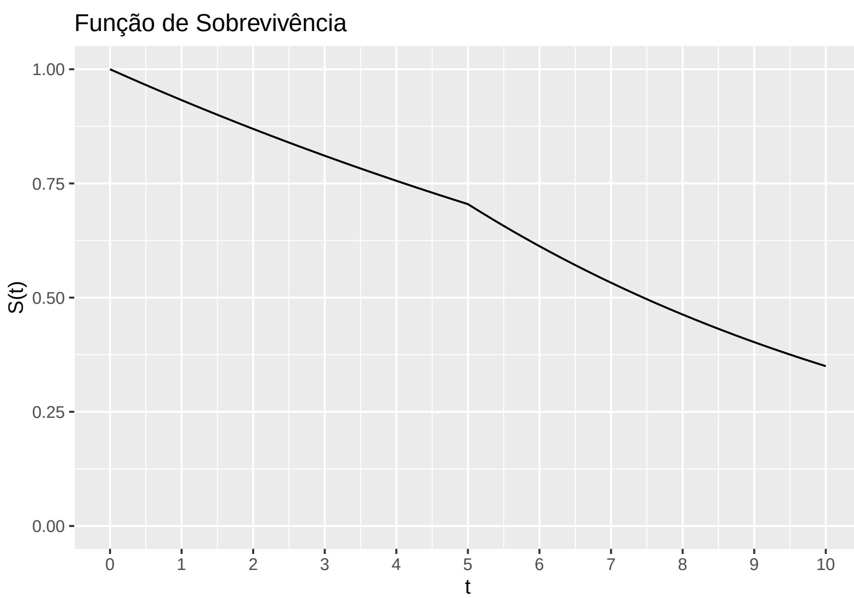
Como a função de taxa de falha é dada por:

$$\alpha(t) = \begin{cases} 0.07\ se\ 0 \leq t \leq 5 \\ 0.14\ se\ t > 5 \end{cases}$$

$$\Rightarrow S(t) = exp\Big\{ - \Big[\int_0^5 0.07\ du + \int_5^t 0.14\ du \Big] \Big\}$$

Logo a função de sobrevivência é dada por:

$$S(t) = \begin{cases} exp(-0.07t)\ se\ 0 \leq t \leq 5 \\ exp(-0.14t + 0.35)\ se\ t > 5 \end{cases}$$



(c) Qual é o tempo mediano de sobrevivência?

Resolução

Sabemos que o tempo mediano de sobrevivência é dado por

$$S(t_{0.5}) = 0.5$$

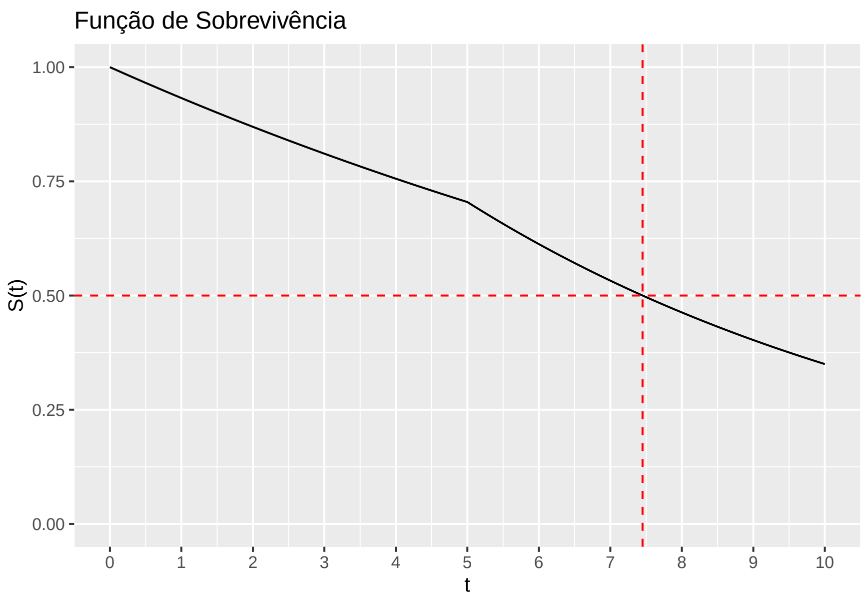
Assim temos que:

$$exp\{- \int_0^t \alpha(u) du\} = 0.5 \Rightarrow \int_0^t \alpha(u) du = -ln(0.5) \Rightarrow A(t) - A(0) = -ln(0.5) \Rightarrow A(t) = -ln(0.5) + 0$$

$$\Rightarrow A(t) = \int_0^5 0.07\ du + \int_5^t 0.14\ du = -ln(0.5) \Rightarrow 0.35 + 0.14u|_5^t = -ln(0.5) = 0.35 + 0.14t - 0.14 * 5 = -ln(0.5)$$

$$\Rightarrow 0.14t = 0.35 - ln(0.5) \Rightarrow t = \frac{1.043}{0.14} = 7.45$$

Como podemos verificar no gráfico:



Exercício 4

Suponha que o tempo até o óbito de pacientes submetidos a transplante de rim (em dias) segue uma distribuição log-logística, com função de sobreviência dada por

$$S(t) = \frac{1}{1 + \lambda t^{\alpha}} \; t \geq 0$$

Se $\alpha = 1.5$ e $\lambda = 0.001$,

(a) Encontre a probabilidade de sobrevivência nos dias 50, 100 e 150.

Resolução

Como sabemos que a probabilidade de sobrevivência é dada por

$$S(t_0)$$

Logo a probabilidade de sobrevivência em 50 dias

$$S(50) =$$

```
## [1] 0.7387961
```

Logo a probabilidade de sobrevivência em 100 dias

$$S(100) =$$

```
## [1] 0.5
```

Logo a probabilidade de sobrevivência em 150 dias

$$S(150) =$$

```
## [1] 0.3524704
```

(b) Encontre o tempo mediano de vida dos pacientes após o transplante.

Resolução

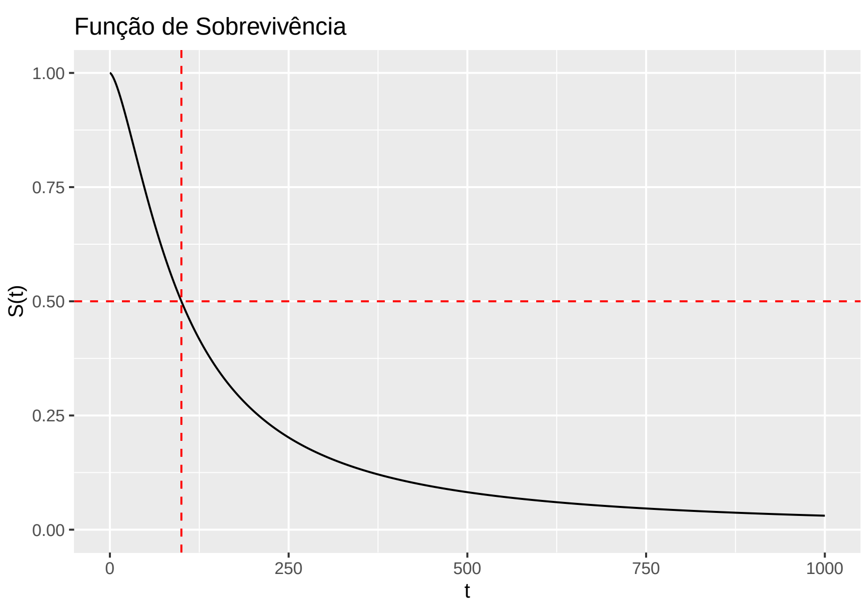
Sabemos que o tempo mediano de sobrevivência é dado por

$$S(t_{0.5}) = 0.5$$

Assim temos que:

$$0.5 = \frac{1}{1 + 0.001t^{1.5}} \Rightarrow 1 = 0.5 + 0.0005t^{1.5} \Rightarrow 0.5 = 0.0005t^{1.5} \Rightarrow 1000 = t^{\frac{3}{2}} \Rightarrow t = 1000^{\frac{2}{3}} = 100$$

O tempo mediano de vida dos pacientes após o transplante é de 100 dias, como podemos ver no gráfico a seguir:



(c) Mostre que a função de taxa de falha é inicialmente crescente e depois descrescente com o tempo. Encontre o ponto em que a taxa de falha muda de crescente para decrescente.

Resolução

Do resultado que diz que a função de taxa de falha é dada por

$$\alpha(t) = -\frac{d}{dt}(ln(S(t)))$$

Assim,

$$ln(S(t)) = ln\Big(\frac{1}{1+\lambda t^{\alpha}}\Big) = -ln(1+\lambda t^{\alpha})$$

Calculando a derivada, temos:

$$\alpha(t) = \frac{\alpha\lambda t^{\alpha-1}}{1+\lambda t^{\alpha}}$$

Para encontar o ponto em que a taxa de falha muda de crescente para decrescente, basta encontrar o ponto de máximo da função e verificar se é ponto de inflexão da função $\alpha(t)$

Para isso deve-se calcular a derivada de $\alpha(t)$ e igualar a zero, assim:

$$\frac{d}{dt}\alpha(t) = \frac{(\alpha-1)\alpha\lambda t^{\alpha-2}}{\lambda t^{\alpha}+1} - \frac{\alpha^2\lambda^2 t^{\alpha-2}}{(1+\lambda t^{\alpha})^2} = 0 \Rightarrow t = \alpha^{-\frac{1}{\alpha}}\lambda^{-\frac{2}{\alpha}}(\alpha^2\lambda - \alpha\lambda)^{1/\alpha}$$

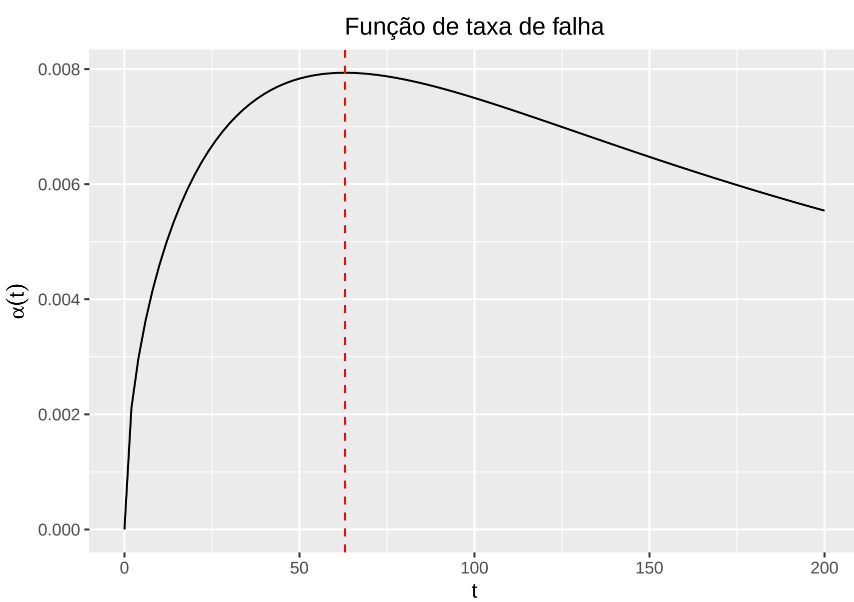
Substituindo os valores de $\lambda = 0.001$ e $\alpha = 1.5$, temos que o ponto de máximo da função de taxa de falha é $t = 63$.

Para verificar se é um ponto de inflexão a segunda derivada de $\alpha(63)$, assim

$$\frac{d^2}{dt^2}\alpha(t) = \alpha\lambda\Big(t^{\alpha-1}\Big(\frac{2\alpha^2\lambda^2 t^{2\alpha-2}}{(\lambda t^{\alpha}+1)^3} - \frac{(\alpha-1)\alpha\lambda t^{\alpha-2}}{(\lambda t^{\alpha}+1)^2}\Big) + \frac{(\alpha-2)(\alpha-1)t^{\alpha-3}}{\lambda t^{\alpha}+1)} - \frac{2(\alpha-1)\alpha\lambda t^{2\alpha-3}}{(\lambda t^{\alpha}+1)^2}\Big)$$

Calculando o valor de $\frac{d^2}{dt^2}\alpha(63)$ obtemos um valor > 0, ou seja, temos um ponto de máximo e além disso um ponto de inflexão.

E também podemos observar o resultado acima no gráfico:



(d) Encontre o tempo médio de vida dos pacientes após o transplante (você pode consultar o livro do Klein e Moeschberger, pg. 38).

Resolução

Para encontrar o tempo médio de vida dos pacientes após o transplante, basta calular a seguinte integral:

$$\int_0^\infty S(t)\ dt = \int_0^\infty \frac{1}{1+\lambda t^\alpha}\ dt$$

Calculando a integral via R, obtemos o seguinte resultado:

```
## 241.8401 with absolute error < 0.0062
```

Logo temos que o tempo médio de de vida dos pacientes após o transplante é de aproximadamente 242 dias.

Exercício 5

Considere a distribuição Weibull, com função de sobrevivência dada por

$$S(t) = exp\{-\lambda t^\rho\},\ t \geq 0$$

Utilizando qualquer software ou pacote estatístico de sua preferência, construa gráficos da função de taxa de falha da distribuição Weibull, variando-se os valores dos parâmetros λ e ρ. Considere 6 combinações diferentes de valores de λ e ρ: utilize dois valores diferentes de λ e três valores diferentes de ρ, sendo um deles necessariamente igual a 1 (ou seja, $\rho = 1$). Construa também gráficos das respectivas funções de sobrevivência.

Resolução

Sabendo do resultado:

$$\alpha(t) = -\frac{d}{dt}ln(S(t))$$

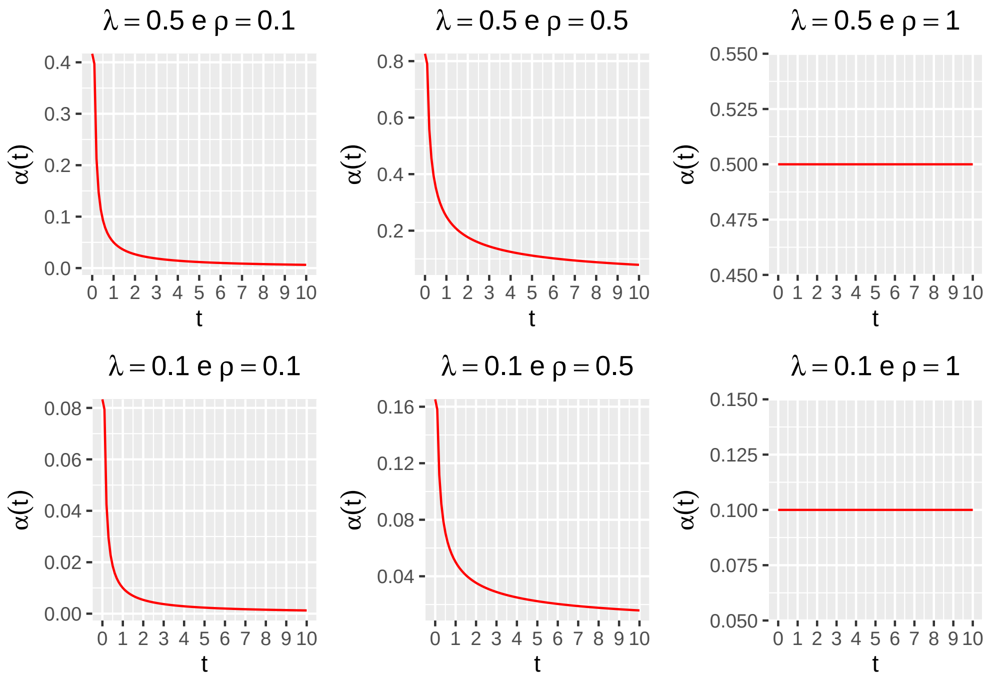
Assim, aplicando o $ln(.)$ em $S(t)$, temos:

$$ln(S(t)) = ln(exp\{-\lambda t^\rho\}) = -\lambda t^\rho$$

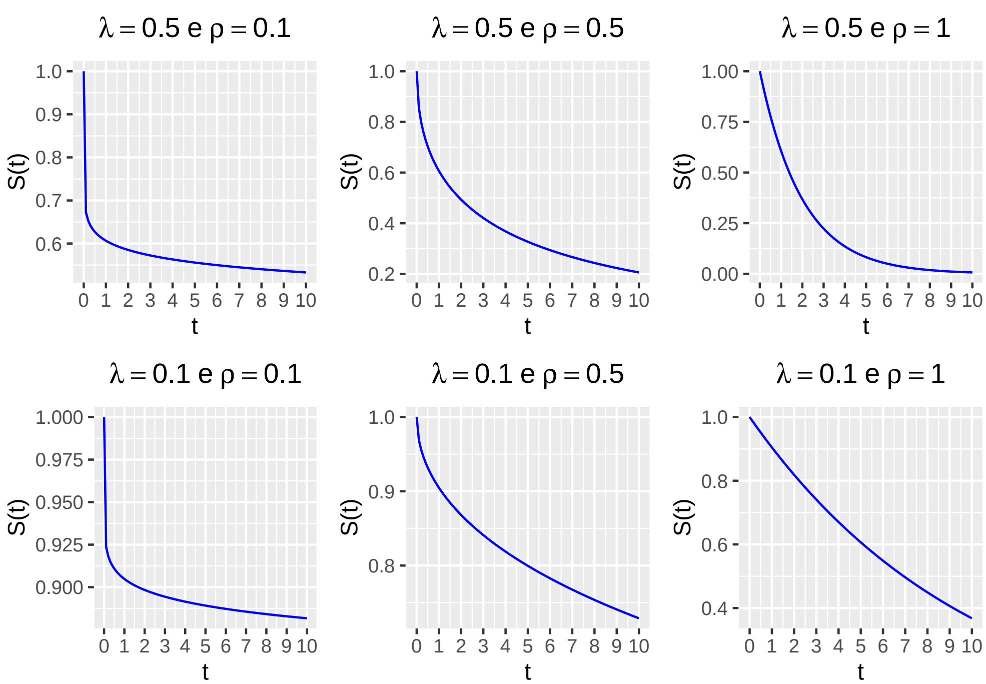
E derivando em relação a t:

$$-\frac{d}{dt}ln(S(t)) = -\frac{d}{dt} - \lambda t^\rho = \lambda \rho t^{\rho-1} \Rightarrow \alpha(t) = \lambda \rho t^{\rho-1}$$

Assim sendo constuindo os gráficos utilizando o software R para a função de taxa de falha ($\alpha(t)$), temos que:



E os gráficos para a função de sobrevivênia $S(t)$, temos que:



Exercício 6

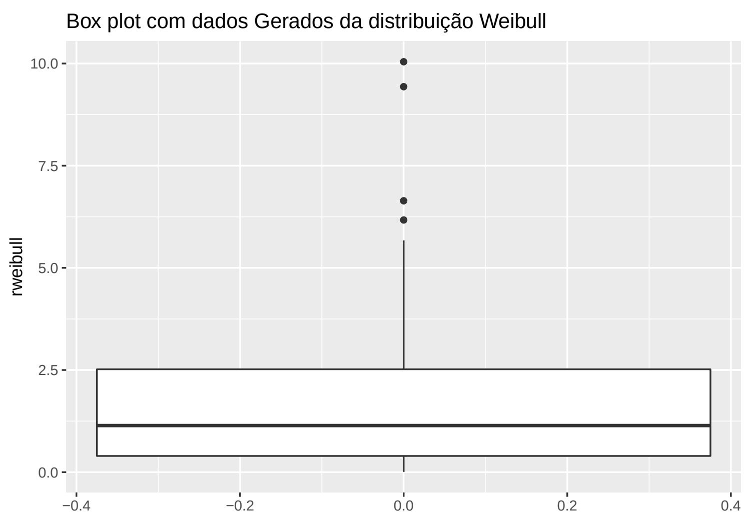
Considere a distribuição Weibull do exercício anterior e escolha uma das combinaçõoes de λ e ρ utilizadas. Utilizando qualquer *software* ou pacote estatístico de sua preferência, gere dados com a distribuição Weibull escolhida, com tamanho amostral $n = 100$.

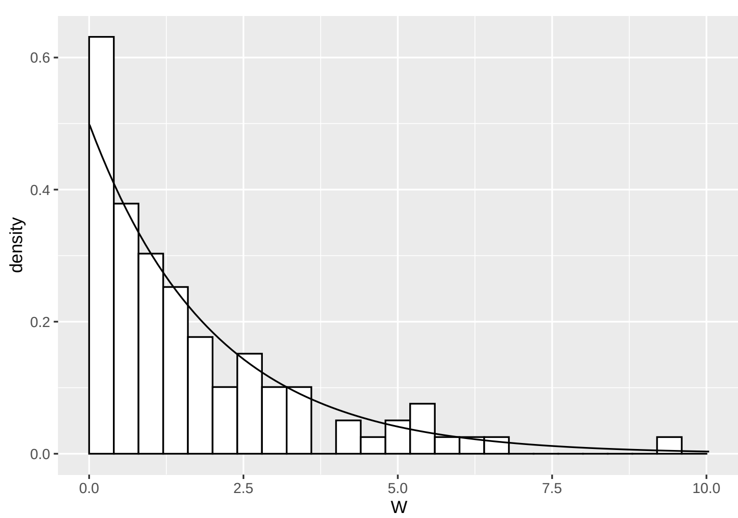
IMPORTANTE: Escreva claramente qual foi o software ou pacote estatístico utilizado e inclua necessariamente os códigos utilizados para a resolução do exercício.

(a) Obtenha um boxplot dos dados e um histograma (com a curva da densidade teórica também no gráfico). Obtenha também a curva de sobrevivência empírica dos dados e coloque num mesmo gráfico a curva empírica e a curva teórica utilizada para gerar os gráficos.

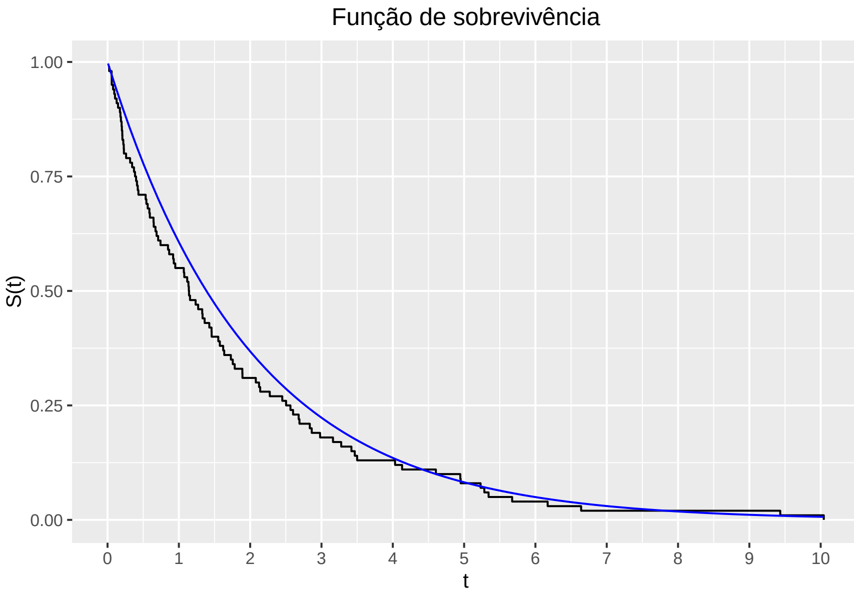
Resolução

Utilizando o R como software estatístico e os parâmetros $\lambda = 0.5$ e $\rho = 1$ obtemos o boxplot e o histograma:





Agora, a curva de sobrevivência empirica:

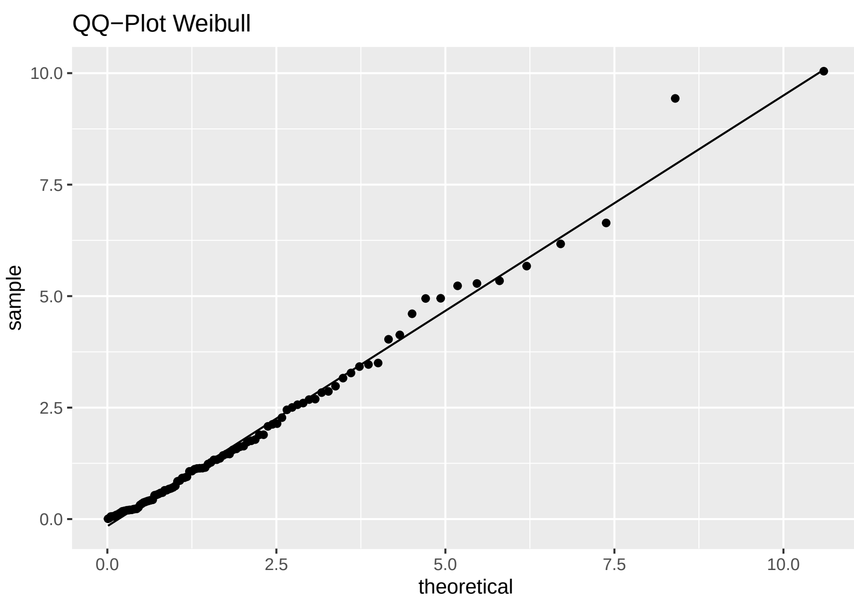


Em que podemos notar que está se adequando muito bem.

(b) Faça um gráfico de quantis (QQ Plot) comparando os quantis empíricos com os quantis teóricos da distribuição Weibull utilizada para gerar os dados.

Resolução

Gráfico QQ Plot da distribuição Weibull com os parâmetros $\lambda = 0.5$ e $\rho = 1$:

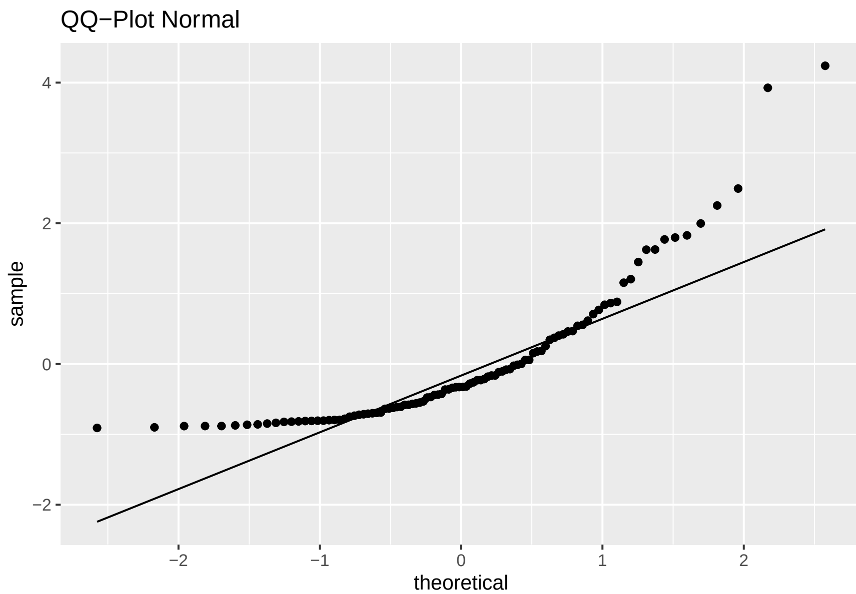


Como se ve no gráfico acima os pontos estão muito próximos da reta o que indica que existe uma boa adequação dos dados simulados em relação a distibuição utilizadas como base para a simulação no caso a distribuição Weibull.

(c) Padronize os dados (ou seja, subtraia a média amostral e divida pelo desvio padrão) e faça um gráfico de quantis (QQ Plot) comparando os quantis empíricos da variável padronizada com os da normal padrão. Discuta a adequalbilidade da distribuição normal aos dados.

Resolução

Subtraindo da média e dividindo pelo desvio padão,temos o seguinte QQ PLot:



Comparando com a distribuição normal padrão, assim como esperado os pontos mais extremos não se adequam aos quantis da distribuição normal.

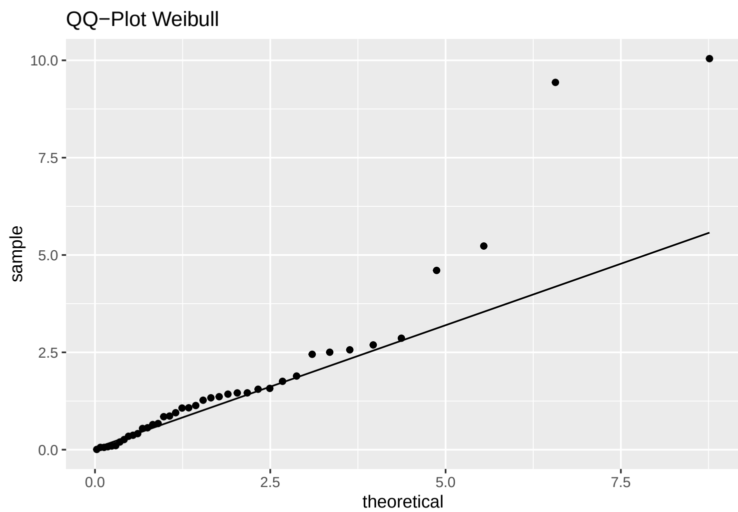
(d) Repita os itens (b) e (c) para $n = 40$, $n = 300$ e $n = 1200$.

Resolução

Repetindo os itens (b) e (c) para $n = 40$:

(b__40)

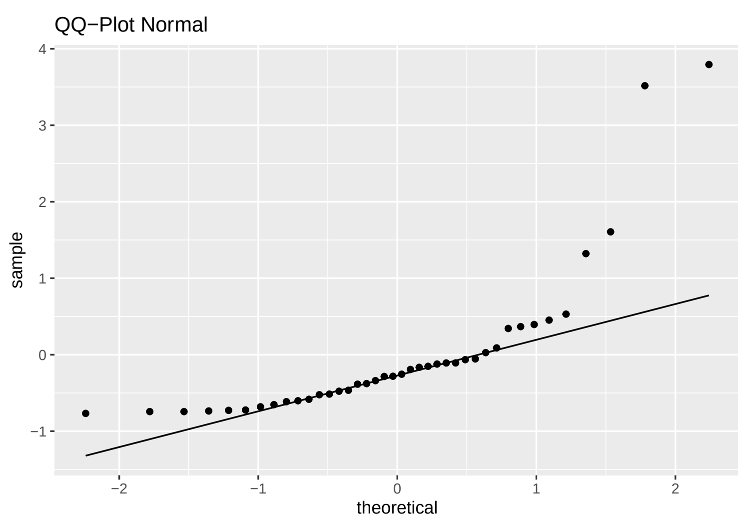
Gráfico QQ Plot da distribuição Weibull com os parâmetros $\lambda = 0.5$ e $\rho = 1$:



Apesar de ter um n "pequeno" como se ve no gráfico acima os pontos estão muito próximos da reta o que indica que existe uma boa adequação dos dados simulados.

(c__40)

Subtraindo da média e dividindo pelo desvio padão,temos o seguinte QQ PLot:

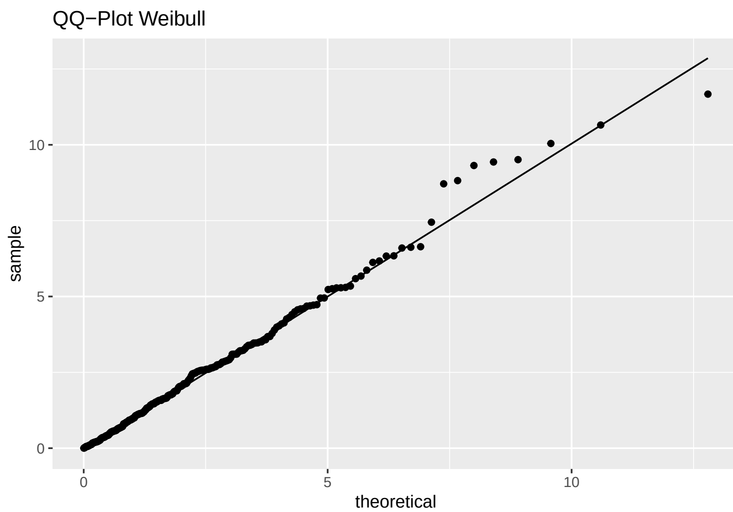


Apesar de ter um n "pequeno" comparando com a distribuição normal padrão, assim como esperado os pontos mais extremos não se adequam aos quantis da distribuição normal.

Repetindo os itens (b) e (c) para $n = 300$:

(b__300)

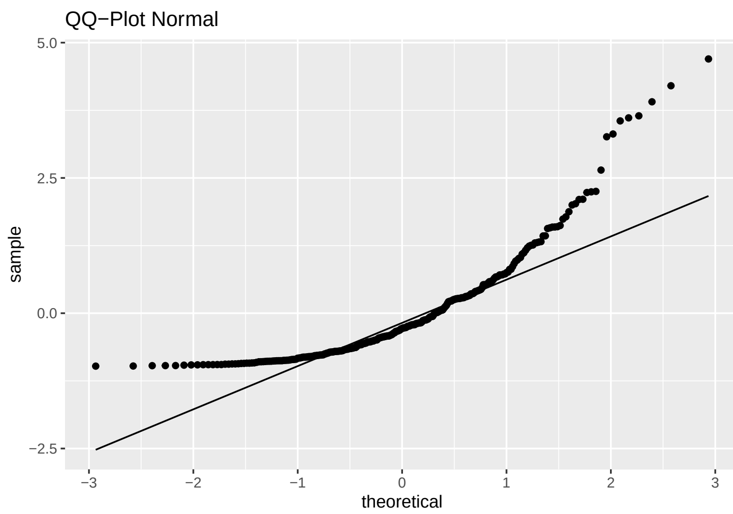
Gráfico QQ Plot da distribuição Weibull com os parâmetros $\lambda = 0.5$ e $\rho = 1$:



Com um $n = 300$ como podemos ver no gráfico acima os pontos estão muito próximos da reta o que indica que existe uma boa adequação dos dados simulados.

(c__300)

Subtraindo da média e dividindo pelo desvio padão,temos o seguinte QQ PLot:

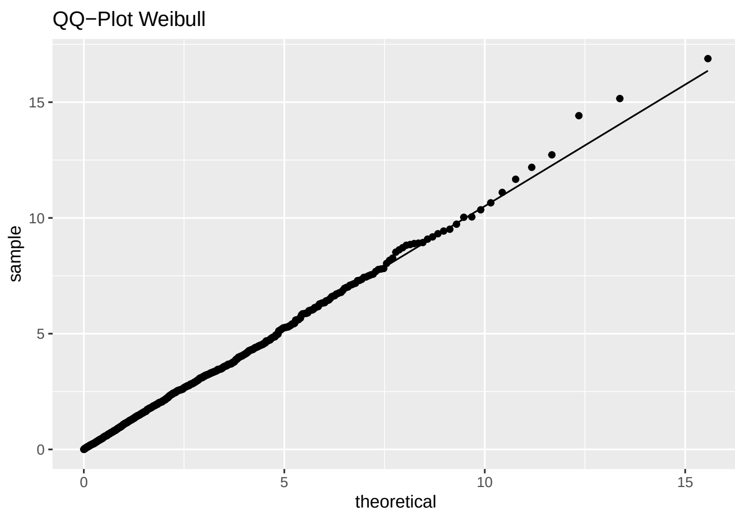


Assim como no gráfico com $n = 100$ comparando com a distribuição normal padrão, assim como esperado os pontos mais extremos não se adequam aos quantis da distribuição normal.

Repetindo os itens (b) e (c) para $n = 1200$:

(b__1200)

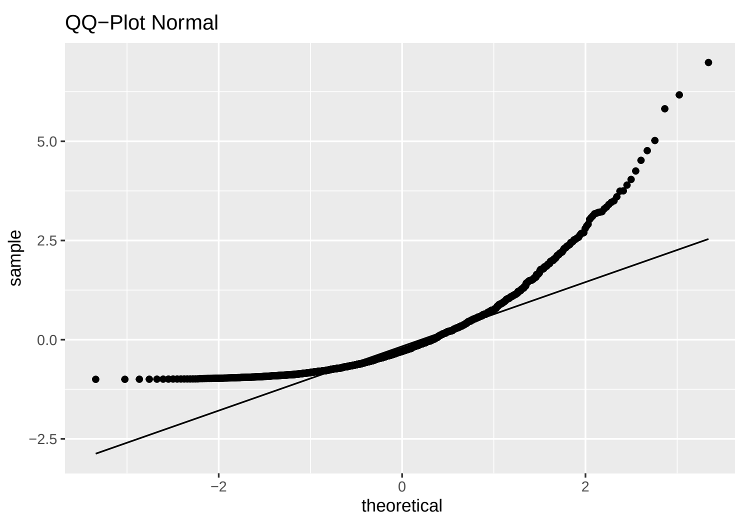
Gráfico QQ Plot da distribuição Weibull com os parâmetros $\lambda = 0.5$ e $\rho = 1$:



Apesar de ter um n "grande" como podemos ver no gráfico acima os pontos estão muito próximos da reta o que indica que existe uma adequação muito dos dados simulados.

(c__1200)

Subtraindo da média e dividindo pelo desvio padão,temos o seguinte QQ PLot:



Apesar de ter um n "grande" comparando com a distribuição normal padrão, assim como esperado os pontos mais extremos não se adequam aos quantis da distribuição normal.

Códigos

```r
# Lista 1 - MAE 514

library(ggplot2)
library(gridExtra)

# Exercício 3
# item a

# função taxa de falha
alpha <- function(t){
  l <- vector()
  for(i in 1:length(t)){
    if(t[i]>=0 & t[i]<=5){
      l[i]=0.07
    }
    else {
      l[i]=0.14
    }
  }
  return(l)
}

t <- data.frame(t=seq(0,10,0.1))

ggplot(t, aes(x=t)) +
  stat_function(fun = alpha) +
  scale_y_continuous(limits = c(0,0.15)) +
  scale_x_continuous(breaks = seq(0,10)) +
  labs(y = expression(alpha(t)),
       x = "t",
       title = "Função taxa de falha"
       )

# item b

Alpha <- vector()
for(i in 1:length(t$t)){
  Alpha[i]=integrate(alpha,0,t$t[i])$value
}

S <- data.frame(S=exp(-Alpha),t=t$t)

ggplot(S, aes(x=t,y=S)) +
  geom_line() +
  scale_y_continuous(limits = c(0,1)) +
  scale_x_continuous(limits = c(0,10),breaks = seq(0,10)) +
  labs(y = "S(t)",
       x = "t",
       title = "Função de Sobrevivência")

ggplot(S, aes(x=t,y=S)) +
```

```
  geom_line() +
  scale_y_continuous(limits = c(0,1)) +
  scale_x_continuous(limits = c(0,10),breaks = seq(0,10)) +
  labs(y = "S(t)",
       x = "t",
       title = "Fun??o de Sobreviv?ncia")+
  geom_abline(intercept = 0.5, slope = 0, color = "red", lty=2)+
  geom_vline(xintercept = 7.45, color = "red",lty=2)

# Exercício 4
# item a

# função taxa de sobrevivência
S_4 <- function(t,l,a){
  S <- vector()
  for(i in 1:length(t)){
    S[i] <- 1/(1+l*t[i]^a)
  }
  return(S)
}

S_4(50,0.001,1.5)
S_4(100,0.001,1.5)
S_4(150,0.001,1.5)

# item b

S_4(100,0.001,1.5)

t <- seq(0,1000,0.1)
S <-S_4(t,0.001,1.5)

dad <- data.frame(t,S)
ggplot(dad, aes(x=t,y=S)) +
  geom_line() +
  scale_y_continuous(limits = c(0,1)) +
  scale_x_continuous(limits = c(0,1000)) +
  labs(y = "S(t)",
       x = "t",
       title = "Fun??o de Sobreviv?ncia")+
  geom_abline(intercept = 0.5, slope = 0, color = "red", lty=2)+
  geom_vline(xintercept = 100, color = "red",lty=2)

# item c

# função taxa de taxa de falha
alpha_4 <- function(t,l,a){
  alfa <- vector()
  for(i in 1:length(t)){
    alfa[i] <- (a*l*t[i]^(a-1))/(1+l*t[i]^a)
  }
  return(alfa)
}
```

```
t1 <- data.frame(t1=seq(0,200,0.1))
ggplot(t1, aes(x=t1)) +
  stat_function(fun = function(y) alpha_4(y,0.001,1.5)) +
  geom_vline(xintercept = 63,colour='red') +
  theme(plot.title = element_text(hjust = 0.5)) +
  labs(y = bquote(alpha(t)),
       x = "t",
       title = "Função de taxa de falha")

# item d

R <- function(t){
  R <- vector()
  for(i in 1:length(t)){
    R[i] <- 1/(1+0.001*t[i]^1.5)
  }
  return(R)
}
integrate(R,0,Inf)$value

# Exercício 5

# função taxa de sobrevivência
S_5 <- function(t,l,p){
  S <- vector()
  for(i in 1:length(t)){
    S[i] <- exp(-l*t[i]^p)
  }
  return(S)
}

plot1 <- ggplot(t, aes(x=t)) +
  stat_function(fun = function(y) S_5(y,0.5,0.1),colour='blue') +
  scale_x_continuous(breaks = seq(0,10)) +
  theme(plot.title = element_text(hjust = 0.5)) +
  labs(y = "S(t)",
       x = "t",
       title = bquote(lambda == 0.5 ~ 'e' ~ rho == 0.1))

plot2 <- ggplot(t, aes(x=t)) +
  stat_function(fun = function(y) S_5(y,0.5,0.5),colour='blue') +
  scale_x_continuous(breaks = seq(0,10)) +
  theme(plot.title = element_text(hjust = 0.5)) +
  labs(y = "S(t)",
       x = "t",
       title = bquote(lambda == 0.5 ~ 'e' ~ rho == 0.5))

plot3 <- ggplot(t, aes(x=t)) +
  stat_function(fun = function(y) S_5(y,0.5,1),colour='blue') +
  scale_x_continuous(breaks = seq(0,10)) +
  theme(plot.title = element_text(hjust = 0.5)) +
  labs(y = "S(t)",
       x = "t",
```

```
       title = bquote(lambda == 0.5 ~ 'e' ~ rho == 1))

plot4 <- ggplot(t, aes(x=t)) +
  stat_function(fun = function(y) S_5(y,0.1,0.1),colour='blue') +
  scale_x_continuous(breaks = seq(0,10)) +
  theme(plot.title = element_text(hjust = 0.5)) +
  labs(y = "S(t)",
       x = "t",
       title = bquote(lambda == 0.1 ~ 'e' ~ rho == 0.1))

plot5 <- ggplot(t, aes(x=t)) +
  stat_function(fun = function(y) S_5(y,0.1,0.5),colour='blue') +
  scale_x_continuous(breaks = seq(0,10)) +
  theme(plot.title = element_text(hjust = 0.5)) +
  labs(y = "S(t)",
       x = "t",
       title = bquote(lambda == 0.1 ~ 'e' ~ rho == 0.5))

plot6 <- ggplot(t, aes(x=t)) +
  stat_function(fun = function(y) S_5(y,0.1,1),colour='blue') +
  scale_x_continuous(breaks = seq(0,10)) +
  theme(plot.title = element_text(hjust = 0.5)) +
  labs(y = "S(t)",
       x = "t",
       title = bquote(lambda == 0.1 ~ 'e' ~ rho == 1))

grid.arrange(plot1, plot2, plot3,
             plot4, plot5, plot6, ncol=3,nrow=2)

# função taxa de falha
alpha_5 <- function(t,l,p){
  alpha <- vector()
  for(i in 1:length(t)){
    alpha[i] <- (exp(-l*t[i]^p)*l*p*t[i]^(-1+p))/exp(-l*t[i]^p)
  }
  return(alpha)
}

plot1 <- ggplot(t, aes(x=t)) +
  stat_function(fun = function(y) alpha_5(y,0.5,0.1),colour='red') +
  scale_x_continuous(breaks = seq(0,10)) +
  theme(plot.title = element_text(hjust = 0.5)) +
  labs(y = bquote(alpha(t)),
       x = "t",
       title = bquote(lambda == 0.5 ~ 'e' ~ rho == 0.1))

plot2 <- ggplot(t, aes(x=t)) +
  stat_function(fun = function(y) alpha_5(y,0.5,0.5),colour='red') +
  scale_x_continuous(breaks = seq(0,10)) +
  theme(plot.title = element_text(hjust = 0.5)) +
  labs(y = bquote(alpha(t)),
       x = "t",
       title = bquote(lambda == 0.5 ~ 'e' ~ rho == 0.5))
```

```r
plot3 <- ggplot(t, aes(x=t)) +
  stat_function(fun = function(y) alpha_5(y,0.5,1),colour='red') +
  scale_x_continuous(breaks = seq(0,10)) +
  theme(plot.title = element_text(hjust = 0.5)) +
  labs(y = bquote(alpha(t)),
       x = "t",
       title = bquote(lambda == 0.5 ~ 'e' ~ rho == 1))

plot4 <- ggplot(t, aes(x=t)) +
  stat_function(fun = function(y) alpha_5(y,0.1,0.1),colour='red') +
  scale_x_continuous(breaks = seq(0,10)) +
  theme(plot.title = element_text(hjust = 0.5)) +
  labs(y = bquote(alpha(t)),
       x = "t",
       title = bquote(lambda == 0.1 ~ 'e' ~ rho == 0.1))

plot5 <- ggplot(t, aes(x=t)) +
  stat_function(fun = function(y) alpha_5(y,0.1,0.5),colour='red') +
  scale_x_continuous(breaks = seq(0,10)) +
  theme(plot.title = element_text(hjust = 0.5)) +
  labs(y = bquote(alpha(t)),
       x = "t",
       title = bquote(lambda == 0.1 ~ 'e' ~ rho == 0.5))

plot6 <- ggplot(t, aes(x=t)) +
  stat_function(fun = function(y) alpha_5(y,0.1,1),colour='red') +
  scale_x_continuous(breaks = seq(0,10)) +
  theme(plot.title = element_text(hjust = 0.5)) +
  labs(y = bquote(alpha(t)),
       x = "t",
       title = bquote(lambda == 0.1 ~ 'e' ~ rho == 1))

grid.arrange(plot1, plot2, plot3,
             plot4, plot5, plot6, ncol=3,nrow=2)

# Exercício 6
# item a
#                                         rho         1/lambda
set.seed(4)
rweibull <- data.frame(W = rweibull(100,shape = 1, scale = 2))
ggplot(rweibull, aes(y=W)) +
  geom_boxplot()+
  labs(y = "rweibull",
       title = "Box plot rweibull")

ggplot(rweibull, aes(x = W)) +
  geom_histogram(aes(y =..density..),
                 breaks = seq(0, 10, by = 0.4),
                 colour = "black",
                 fill = "white")+
  stat_function(fun = dweibull,
                args = list(shape = 1, scale = 2))
```

```r
dat <- data.frame(t=knots(ecdf(rweibull$W)),
                  S=1-ecdf(rweibull$W)(knots(ecdf(rweibull$W))))

ggplot(dat,aes(x=t, y=S)) +
  geom_step() +
  stat_function(fun = function(y) S_5(y,0.5,1),colour='blue') +
  scale_x_continuous(breaks = seq(0,10)) +
  theme(plot.title = element_text(hjust = 0.5)) +
  labs(y = "S(t)",
       x = "t",
       title = 'Função de sobrevivência')

# item b

# QQplot
ggplot(rweibull, aes(sample = W)) +
  stat_qq(distribution = qweibull,dparams = list(shape = 1, scale = 2)) +
  stat_qq_line(distribution = qweibull, dparams = list(shape = 1, scale = 2))+
  ggtitle("QQ-Plot Weibull")

# item c

# substraindo da média e dividindo pelo desvio padrão
rweibull2 <- data.frame(W_pad=(rweibull$W-mean(rweibull$W))/sd(rweibull$W))

ggplot(rweibull2, aes(sample = W_pad)) +
  stat_qq(distribution = qnorm) +
  stat_qq_line(distribution = qnorm) +
  ggtitle("QQ-Plot Normal")

# item d

#n=40
set.seed(4)
rweibull <- data.frame(W = rweibull(40,shape = 1, scale = 2))

ggplot(rweibull, aes(sample = W)) +
  stat_qq(distribution = qweibull,dparams = list(shape = 1, scale = 2)) +
  stat_qq_line(distribution = qweibull, dparams = list(shape = 1, scale = 2))+
  ggtitle("QQ-Plot Weibull")

# substraindo da média e dividindo pelo desvio padrão
rweibull2 <- data.frame(W_pad=(rweibull$W-mean(rweibull$W))/sd(rweibull$W))

ggplot(rweibull2, aes(sample = W_pad)) +
  stat_qq(distribution = qnorm) +
  stat_qq_line(distribution = qnorm) +
  ggtitle("QQ-Plot Normal")

#n=300
set.seed(4)
rweibull <- data.frame(W = rweibull(300,shape = 1, scale = 2))
```

```
ggplot(rweibull, aes(sample = W)) +
  stat_qq(distribution = qweibull,dparams = list(shape = 1, scale = 2)) +
  stat_qq_line(distribution = qweibull, dparams = list(shape = 1, scale = 2))+
  ggtitle("QQ-Plot Weibull")

# substraindo da média e dividindo pelo desvio padrão
rweibull2 <- data.frame(W_pad=(rweibull$W-mean(rweibull$W))/sd(rweibull$W))

ggplot(rweibull2, aes(sample = W_pad)) +
  stat_qq(distribution = qnorm) +
  stat_qq_line(distribution = qnorm) +
  ggtitle("QQ-Plot Normal")

#n=1200
set.seed(4)
rweibull <- data.frame(W = rweibull(1200,shape = 1, scale = 2))

ggplot(rweibull, aes(sample = W)) +
  stat_qq(distribution = qweibull,dparams = list(shape = 1, scale = 2)) +
  stat_qq_line(distribution = qweibull, dparams = list(shape = 1, scale = 2))+
  ggtitle("QQ-Plot Weibull")

# substraindo da média e dividindo pelo desvio padrão
rweibull2 <- data.frame(W_pad=(rweibull$W-mean(rweibull$W))/sd(rweibull$W))

ggplot(rweibull2, aes(sample = W_pad)) +
  stat_qq(distribution = qnorm) +
  stat_qq_line(distribution = qnorm) +
  ggtitle("QQ-Plot Normal")
```