

Lista de Exercícios 3

Análise de Sobrevida

Márcio Roger Piaggio
RA: 67384

Universidade Estadual de Maringá
Prof. Dr. Brian Melo

Maringá-Pr
08/03/23

1º Questão

O tempo em dias para o desenvolvimento de tumor em ratos expostos a uma substância cancerígena segue uma distribuição de Weibull tal que

$$S(t) = \exp \left\{ - \left(\frac{t}{\alpha} \right)^\gamma \right\}$$

Considerando $\alpha = 100$ e $\gamma = 2$, responda

a) Qual é a probabilidade de um rato sobreviver sem tumor aos primeiros 30 dias? E aos primeiros 45 dias?

```
S <- function(t){ exp(-(t/100)**2)}  
S(c(30,45))  
## [1] 0.9139312 0.8166865
```

A probabilidade de 0,914 de um rato sobreviver sem tumor aos primeiros 30 dias e probabilidade de 0,817 aos primeiros 45 dias

b) Qual é o tempo médio até o aparecimento do tumor?

```
tm <- 100*gamma(1+(1/2)); tm  
## [1] 88.62269
```

∴ o tempo médio até o aparecimento do tumor é de 88,62 dias

c) Qual é o tempo mediano até o aparecimento do tumor?

```
t0.5 <- 100*(-log(1-.5))**(1/2); t0.5  
## [1] 83.25546
```

∴ o tempo mediano até o aparecimento do tumor é de 83,26 dias.

d) Encontre a taxa de falha de aparecimento de tumor aos 30, 45 e 60 dias. Interprete estes valores.

```
lambda <- function(t){(2/100**2)*t**(2-1)}  
lambda(c(30,45,60))  
## [1] 0.006 0.009 0.012
```

A taxa de falha aumenta com o tempo. Isso pode ser um indicativo de aumento de ocorrência de tumores nos ratos com maior tempo de vida.

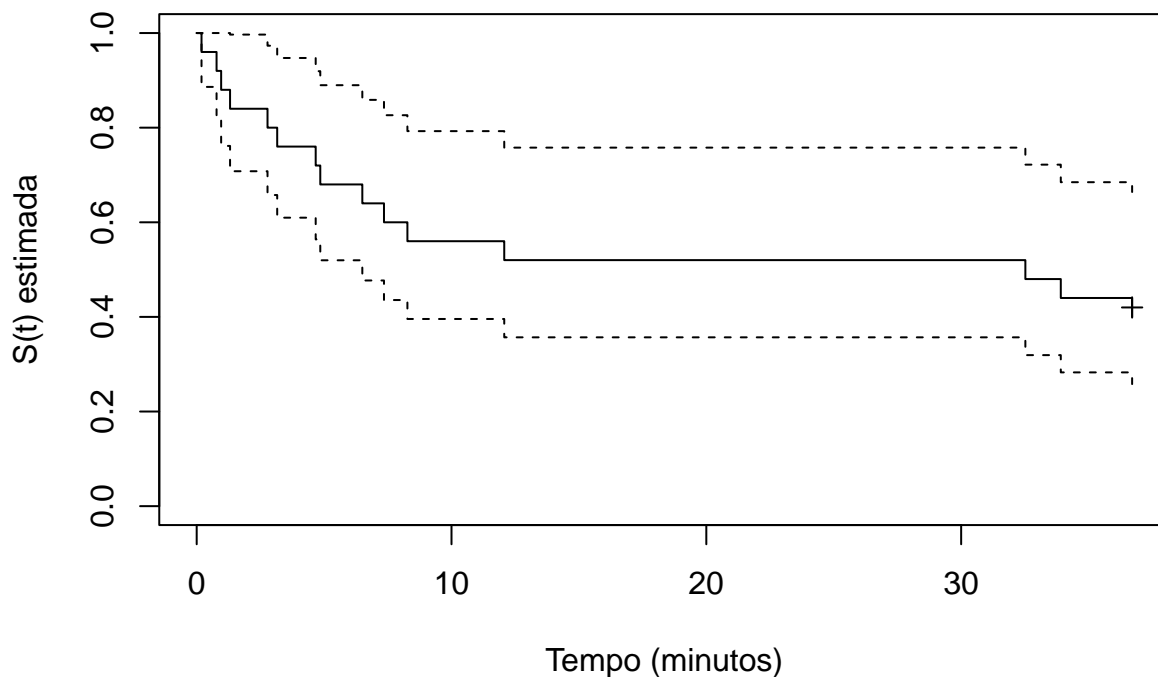
2º Questão

Os dados mostrados a seguir representam o tempo até a ruptura de um tipo de isolante elétrico sujeito a uma tensão de estresse de 35 Kvolts. O teste consistiu em deixar 25 destes isolantes funcionando até que 15 deles falhassem (censura do tipo II), obtendo-se os seguintes resultados (em minutos):

```
tempo <- c(0.19, 0.78, 0.96, 1.31, 2.78, 3.16, 4.67, 4.85, 6.5,  
          7.35, 8.27, 12.07, 32.52, 33.91, rep(36.71,11))  
cens <- c(rep(1,15), rep(0,10))  
ruptura <- data.frame(tempo, cens)
```

Inicialmente, deve-se identificar um modelo paramétrico adequado para explicar estes dados e, em seguida, responder às perguntas.

```
require(survival)  
ekm<-survfit(Surv(tempo,cens)~1)  
plot(ekm, xlab="Tempo (minutos)", ylab="S(t) estimada", conf.int=T, mark.time=T)
```



Ajuste dos modelos paramétricos

Modelo Exponencial

```
fit_exp <- survreg(Surv(tempo,cens)~1,dist='exponential')
# summary(fit_exp)
alpha_exp <- 1/exp(fit_exp$coefficients[1])
```

Modelo Weibull

```
fit_weib <- survreg(Surv(tempo,cens)~1,dist='weibull')
# summary(fit_weib)
alpha_weib <- exp(fit_weib$coefficients[1])
gama_weib <- 1/fit_weib$scale
```

Modelo Log-normal

```
fit_logn <- survreg(Surv(tempo,cens)~1,dist='lognorm')
# summary(fit_logn)
sigma_logn <- exp(fit_logn$icoef[2])
mi_logn <- fit_logn$icoef[1]
```

Modelo Gama generalizada para comparar TRV

```
require(flexsurv)
fit_gamm <- flexsurvreg(Surv(tempo,cens)~1,dist='gengamma')
```

Criterio de AIC

```
data.frame(Modelo = c('Exponencial','Weibull','Log-normal'),
           AIC = c(AIC(fit_exp),AIC(fit_weib),AIC(fit_logn)))
```

Modelo	AIC
Exponencial	138.5534
Weibull	134.1893
Log-normal	132.6558

Gráficos de adequação dos modelos

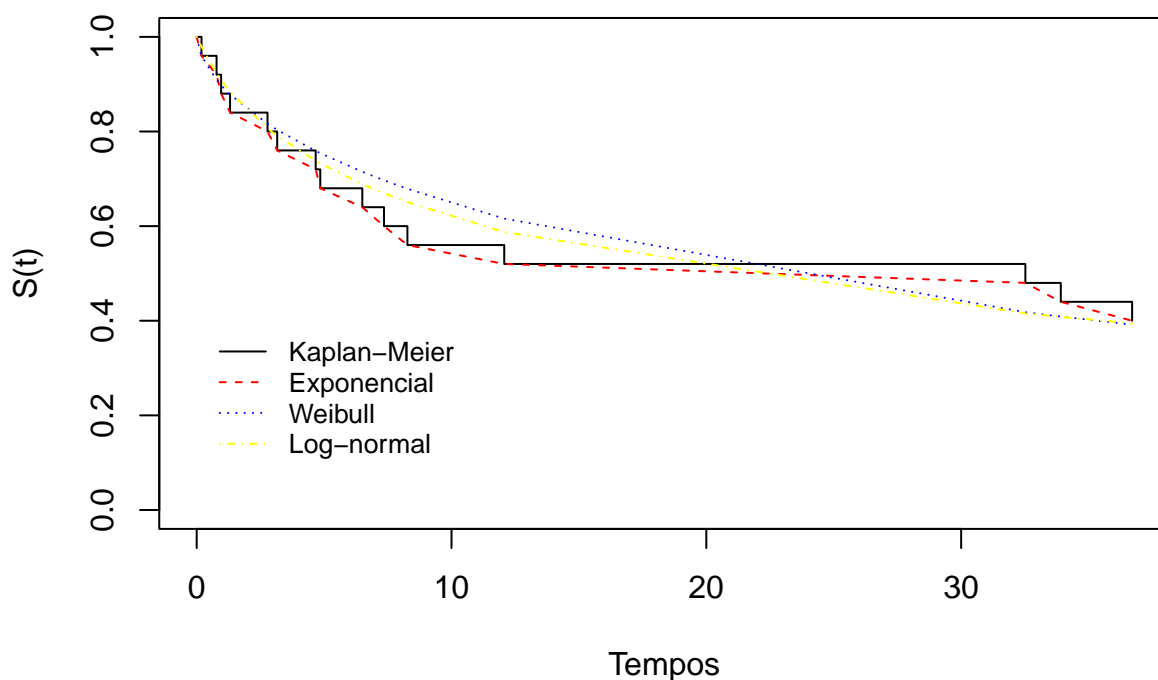
```
time <- ekm$time
st <- ekm$surv
ste <- exp(-time/alpha_exp)
```

```

stw <- exp(-(time/alpha_weib)^gama_weib)
stln <- pnorm((-log(time)+ mi_logn)/sigma_logn)

plot(ekm, conf.int=F, xlab="Tempos", ylab="S(t)")
lines(c(0, time), c(1, st), lty=2, col = "red")
lines(c(0, time), c(1, stw), lty=3, col = "blue")
lines(c(0, time), c(1, stln), lty=4, col = "yellow")
legend(0,0.4,lty=1:4,c("Kaplan-Meier", 'Exponencial','Weibull','Log-normal'),
      col = c("black", "red", "blue", "yellow"), bty="n",cex=0.8)

```



```

# Comparando Gama generalizada com Exponencial, Weibull e Log-normal (TRV)
gamm <- fit_gamm$loglik
exp <- fit_exp$loglik
weib <- fit_weib$loglik
logn <- fit_logn$loglik

kable(data.frame(Modelo = c('Gama generalizada','Exponencial','Weibull','Log-normal'),
  loglik = c(gamm, exp[1], weib[1], logn[1]),
  lambda = c(" ", 2*(gamm-exp[1]), 2*(gamm-weib[1]), 2*(gamm-logn[1])),
  pvalor = c(" ", 1-pchisq(2*(gamm-exp[1]), df=2), 1-pchisq(2*(gamm-weib[1]), df=1),
    1-pchisq(2*(gamm-logn[1]), df=1))),align = "c")

```

Modelo	loglik	lambda	pvalor
Gama generalizada	-64.15942		
Exponencial	-68.27670	8.2345512704446	0.0162888307644301
Weibull	-65.09467	1.87050231715043	0.171417559594494
Log-normal	-64.32788	0.336923404680164	0.561610497300125

Pelo critério de AIC o modelo Log-normal é o menor, visualmente todos os modelos têm um bom ajuste e o TRV não rejeitou a hipóteses de um bom ajuste dos modelos weibull e log-normal, assim selecionamos o modelo log-normal para descrever os dados.

a) Uma estimativa para o tempo mediano de vida deste tipo de isolante elétrico funcionando a 35 Kvolts.

```

t.05 <- exp(qnorm(0.5)*sigma_logn+mi_logn); t.05 # minutos
## Log(scale)
## 19.9797

```

O tempo mediano é de 19,9797 minutos.

b) Uma estimativa (por ponto e por intervalo) para a fração de defeituosos esperada nos dois primeiros minutos de funcionamento.

```
mu <- fit_logn$coefficients
sigma <- fit_logn$scale

# variâncias e covariância dos parâmetros estimados
Mvar <- fit_logn$var

# fração de defeituosos
fd <- 1 - pnorm((-log(2)+ mu)/sigma); fd
## (Intercept)
## 0.1563221
```

Método delta para dois parâmetros:

$$Var(\hat{\theta}) = Var(\hat{\mu}) \left(\frac{d\theta}{d\mu} \right)^2 + 2Cov(\hat{\mu}, \hat{\sigma}) \left(\frac{d\theta}{d\mu} \right) \left(\frac{d\theta}{d\sigma} \right) + Var(\hat{\sigma}) \left(\frac{d\theta}{d\sigma} \right)^2$$

```
varFd <- 0.26423124*((dnorm((-log(2)+ mu)/sigma)/sigma)**2) +
  2*0.03286996*(dnorm((-log(2)+ mu)/sigma)/sigma)*dnorm((-log(2)+ mu)/sigma)*(log(2)+mu) +
  0.03994483*(dnorm((-log(2)+ mu)/sigma)*(log(2)+mu)**2)

# Intervalo de confiança para a fração de defeituosos
fd + c(- 1.96, 1.96) * sqrt(varFd)
## [1] -0.5749635 0.8876076
```

c) Uma estimativa (por ponto e por intervalo) para o tempo médio de vida destes isoladores funcionando a 35 Kvolts.

```
# Tempo médio
tm <- exp(mu+sigma**2/2); tm # minutos
## (Intercept)
## 268.4694

# Método delta para dois parâmetros:
varMu <- 0.26423124*(exp(mu+sigma**2/2)**2) +
  2*0.03286996*exp(mu+sigma**2/2)*sigma*exp(mu+sigma**2/2) +
  0.03994483*(sigma*exp(mu+sigma**2/2)**2)

# Intervalo de confiança para o tempo médio
tm + c(- 1.96, 1.96) * sqrt(varMu)
## [1] -105.5170 642.4559
```

d) O tempo necessário para 20% dos isolantes estarem fora de operação.

```
t.02 <- exp(qnorm(0.2)*sigma+mu); t.02 # minutos
## (Intercept)
## 2.933674
```

3º Questão

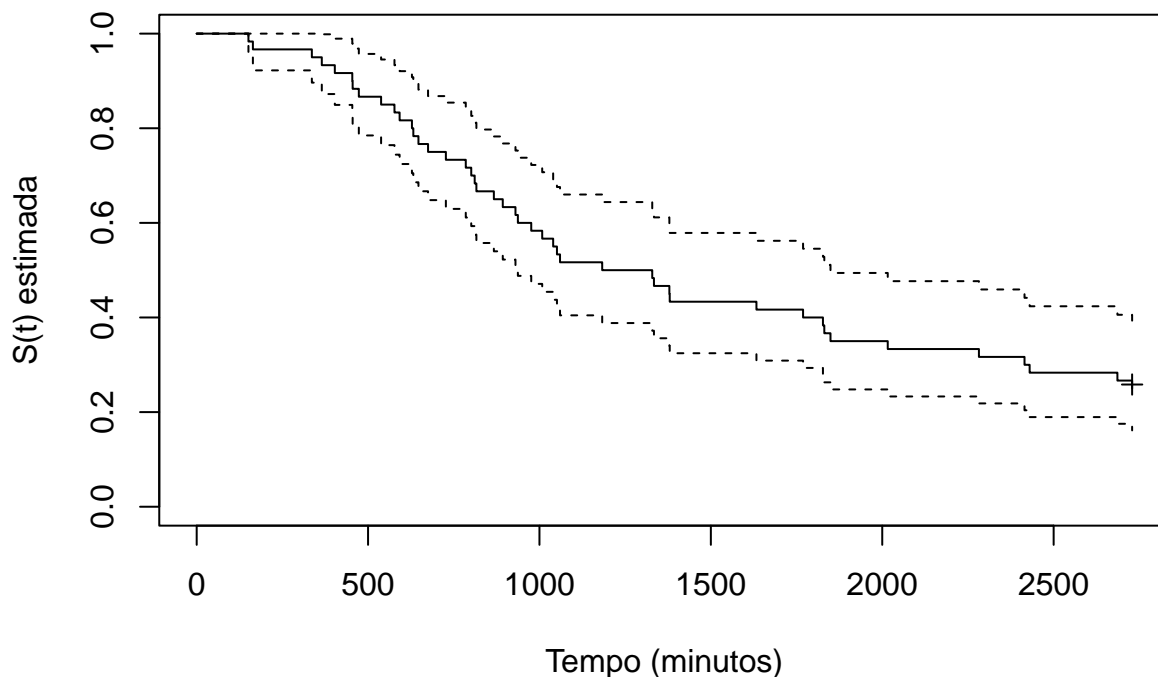
O fabricante de um tipo de isolador elétrico quer conhecer o comportamento de seu produto funcionando a uma temperatura de 200°C. Um teste de vida foi realizado nestas condições usando-se 60 isoladores elétricos. O teste terminou quando 45 deles havia falhado (censura do tipo II). As 15 unidades que não haviam falhado ao final do teste foram, desta forma, censuradas no tempo $t = 2729$ horas. O fabricante tem interesse em estimar o tempo médio e mediano de vida do isolador e o percentual de falhas após 500 horas de uso. Os tempos (em horas) obtidos são apresentados abaixo. Responda às questões de interesse do fabricante fazendo uso do modelo paramétrico que se apresentar mais apropriado para descrever os dados.

```
tempo <- c(151, 164, 336, 365, 403, 454, 455, 473, 538, 577,
  592, 628, 632, 647, 675, 727, 785, 801, 811, 816,
  867, 893, 930, 937, 976, 1008, 1040, 1051, 1060, 1183,
  1329, 1334, 1379, 1380, 1633, 1769, 1827, 1831, 1849, 2016,
  2282, 2415, 2430, 2686, 2729, rep(2729, 15))

cens <- c(rep(1,45), rep(0,15))

isolador <- data.frame(tempo, cens)
```

```
ekm<-survfit(Surv(tempo,cens)~1)
plot(ekm, xlab="Tempo (minutos)",ylab="S(t) estimada",conf.int=T,mark.time=T)
```



```
## Ajuste dos modelos paramétricos
```

```
# Modelo Exponencial
fit_exp <- survreg(Surv(tempo,cens)~1,dist='exponential')
# summary(fit_exp)
alpha_exp <- 1/exp(fit_exp$coefficients[1])
```

```
# Modelo Weibull
fit_weib <- survreg(Surv(tempo,cens)~1,dist='weibull')
# summary(fit_weib)
alpha_weib <- exp(fit_weib$coefficients[1])
gama_weib <- 1/fit_weib$scale
```

```
# Modelo Log-normal
fit_logn <- survreg(Surv(tempo,cens)~1,dist='lognorm')
# summary(fit_logn)
sigma_logn <- exp(fit_logn$icoef[2])
mi_logn <- fit_logn$icoef[1]
```

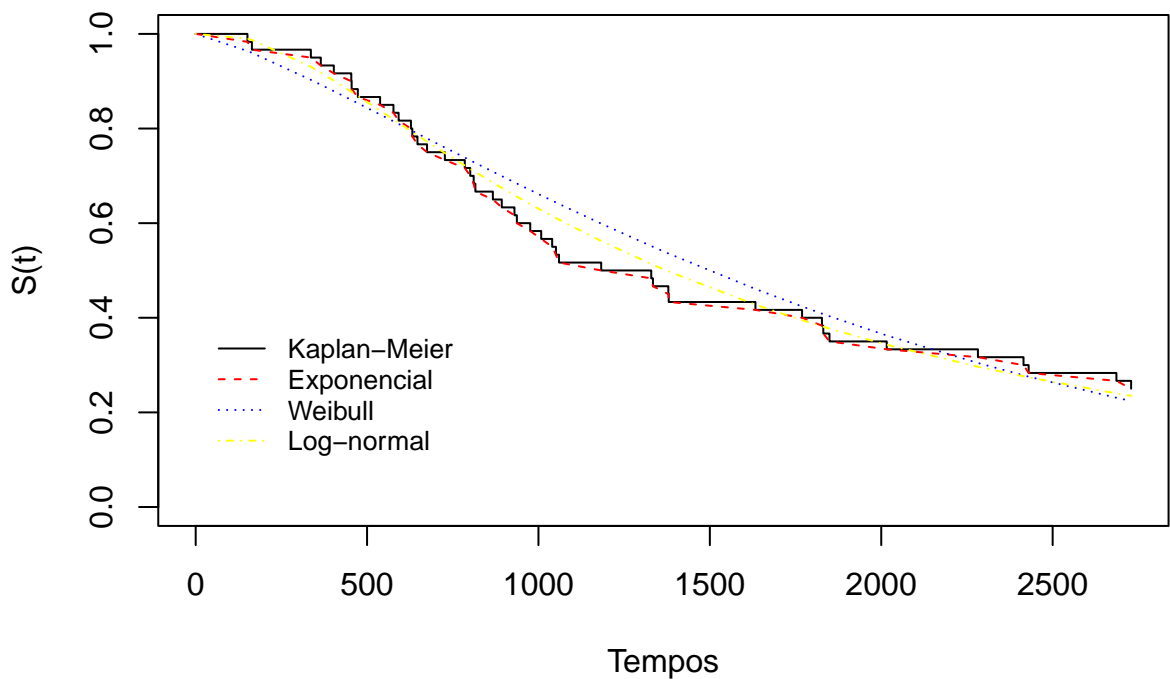
```
# Modelo Gama generalizada para comparar TRV
require(flexsurv)
fit_gamm <- flexsurvreg(Surv(tempo,cens)~1,dist='gengamma')
```

```
## Criterio de AIC
data.frame(Modelo = c('Exponencial','Weibull','Log-normal'),
           AIC = c(AIC(fit_exp),AIC(fit_weib),AIC(fit_logn)))
```

Modelo	AIC
Exponencial	776.8767
Weibull	775.3940
Log-normal	769.3506

```
# Gráficos de adequação dos modelos
time <- ekm$time
st <- ekm$surv
ste <- exp(-time/alpha_exp)
stw <- exp(-(time/alpha_weib)^gama_weib)
stln <- pnorm((-log(time)+ mi_logn)/sigma_logn)

plot(ekm, conf.int=F, xlab="Tempos", ylab="S(t)")
lines(c(0, time), c(1, st), lty=2, col = "red")
lines(c(0, time), c(1, stw), lty=3, col = "blue")
lines(c(0, time), c(1, stln), lty=4, col = "yellow")
legend(0,0.4,lty=1:4,c("Kaplan-Meier", 'Exponencial','Weibull','Log-normal'),
      col = c("black", "red", "blue", "yellow"), bty="n",cex=0.8)
```



```
# Comparando Gama generalizada com Exponencial, Weibull e Log-normal (TRV)
gamm <- fit_gamm$loglik
exp <- fit_exp$loglik
weib <- fit_weib$loglik
logn <- fit_logn$loglik

kable(data.frame(Modelo = c('Gama generalizada','Exponencial','Weibull','Log-normal'),
  loglik = c(gamm, exp[1], weib[1], logn[1]),
  lambda = c(" ", 2*(gamm-exp[1]), 2*(gamm-weib[1]), 2*(gamm-logn[1])),
  pvalor = c(" ",1-pchisq(2*(gamm-exp[1]), df=2), 1-pchisq(2*(gamm-weib[1]), df=1),
    1-pchisq(2*(gamm-logn[1]), df=1))),align = "c")
```

Modelo	loglik	lambda	pvalor
Gama generalizada	-382.3885		
Exponencial	-387.4383	10.099636562313	0.00641049824874795
Weibull	-385.6970	6.61692954144689	0.0101013835432968
Log-normal	-382.6753	0.573530490730263	0.448859394187131

Pelo critério de AIC o modelo log-normal é o menor, visualmente todos os modelos têm um bom ajuste e o TRV não rejeitou a hipóteses de um bom ajuste somente para o modelo log-normal, assim selecionamos o modelo log-normal para descrever os dados.

```

# Tempo médio
mu      <- fit_logn$coefficients
sigma   <- fit_logn$scale
tm      <- exp(mu+sigma**2/2); tm # horas
## (Intercept)
##      2157.131

# Método delta para dois parâmetros:
varTheta <- 0.016454219*(exp(mu+sigma**2/2)**2) +
  2*0.002395354*exp(mu+sigma**2/2)*sigma*exp(mu+sigma**2/2) +
  0.012570933*(sigma*exp(mu+sigma**2/2)**2)

# Intervalo de confiança para o tempo médio
tm + c(- 1.96, 1.96) * sqrt(varTheta)
## [1] 1389.579 2924.684

# tempo mediano
t.05 <- exp(qnorm(0.5)*sigma+mu); t.05 # horas
## (Intercept)
##      1373.018

# percentual de falhas após 500 horas
1 - pnorm((-log(500)+ mu)/sigma)
## (Intercept)
##      0.1439557

```

4° Questão

Seja $T \sim \text{Pareto}(\theta)$, tal que

$$f(t|\theta) = \frac{\theta}{t^{\theta+1}}, \quad t > 1, \quad \theta > 0$$

.

Determine

a) A função de sobrevivência de T.

$$\begin{aligned}
 S(t) &= \int_t^{\infty} \frac{\theta}{u^{\theta+1}} du = \theta \int_t^{\infty} u^{-(\theta+1)} du \\
 &= \frac{\theta u^{-\theta-1+1}}{-\theta-1+1} \Bigg|_t^{\infty} = -u^{-\theta} \Bigg|_t^{\infty} \\
 &= -0 - (-t^{-\theta}) = t^{-\theta}
 \end{aligned}$$

b) As funções de risco, $h(t)$ e de risco acumulado $H(t)$.

$$\begin{aligned}
 h(t) &= \frac{f(t)}{S(t)} = \frac{\frac{\theta}{t^{\theta+1}}}{t^{-\theta}} \\
 &= \frac{\theta \cdot t^{\theta}}{t \cdot t^{\theta}} = \frac{\theta}{t}
 \end{aligned}$$

lembrando que $t > 1$, temos,

$$\begin{aligned}
 H(t) &= \int_1^t \frac{\theta}{u} du = \theta \ln(u) \Bigg|_1^t \\
 &= \theta \ln(t) - \theta \times 0 \\
 &= \theta \ln(t)
 \end{aligned}$$

c) A função de verossimilhança de θ baseada em uma amostra $(t_1, \delta_1), \dots, (t_n, \delta_n)$

$$\begin{aligned} L(\theta) &\propto \prod_{i=1}^n \left[f(t_i; \theta) \right]^{\delta_i} \left[S(t_i; \theta) \right]^{1-\delta_i} \\ &\propto \prod_{i=1}^n \left[\lambda(t_i; \theta) \right]^{\delta_i} S(t_i; \theta) \\ &\propto \prod_{i=1}^n \left[\frac{\theta}{t_i} \right]^{\delta_i} \times t_i^{-\theta} \end{aligned}$$

d) O estimador de máxima verossimilhança de θ .

$$\begin{aligned} l(\theta) &\propto \sum_{i=1}^n \ln \left[\left(\frac{\theta}{t_i} \right)^{\delta_i} \times t_i^{-\theta} \right] \\ &\propto \sum_{i=1}^n \left[\delta_i \left(\ln(\theta) - \ln(t_i) \right) - \theta \ln(t_i) \right] \\ &\propto \sum_{i=1}^n \left[\delta_i \left(\ln(\theta) - \ln(t_i) \right) - \theta \ln(t_i) \right] \\ &\propto \ln(\theta) \sum_{i=1}^n \delta_i - \sum_{i=1}^n \ln(t_i) \delta_i - \theta \sum_{i=1}^n \ln(t_i) \end{aligned}$$

$$\frac{\partial l(\hat{\theta})}{\partial \hat{\theta}} \propto \frac{\sum_{i=1}^n \delta_i}{\hat{\theta}} - 0 - \sum_{i=1}^n \ln(t_i) = 0$$

$$\hat{\theta} \propto \frac{\sum_{i=1}^n \delta_i}{\sum_{i=1}^n \ln(t_i)} = \frac{r}{\sum_{i=1}^n \ln(t_i)}$$

e) Um intervalo de confiança para θ .

$$\begin{aligned} \frac{\partial^2 l(\theta)}{\partial \theta^2} &\propto \sum_{i=1}^n \delta_i \left(-\frac{1}{\theta^2} \right) \\ &\propto -\frac{r}{\theta^2} \end{aligned}$$

Calculando informação de fisher

$$\begin{aligned} I(\theta) &= \mathbb{E} \left(-\frac{\partial^2 l(\theta)}{\partial \theta^2} \right) \\ &= \mathbb{E} \left[-\left(-\frac{r}{\theta^2} \right) \right] \\ &= \frac{r}{\theta^2} \end{aligned}$$

$$\begin{aligned} Var(\hat{\theta}) &= I(\theta)^{-1} \\ &= \left(\frac{r}{\theta^2} \right)^{-1} \\ &= \frac{\theta^2}{r} \end{aligned}$$

assim o intervalo de de confiança para θ é:

$$\begin{aligned} IC(\theta; (1 - \alpha) \times 100\%) &= \hat{\theta} \pm z_{(1-\alpha)} \sqrt{Var(\hat{\theta})} \\ &= \frac{r}{\sum_{i=1}^n \ln(t_i)} \pm z_{(1-\alpha)} \frac{\theta}{\sqrt{r}} \end{aligned}$$

f) Um intervalo de confiança para $E(T) = \frac{\theta}{\theta-1}$ (assumindo $\theta > 1$).

$$\begin{aligned} Var(g(\hat{\theta})) &= \left(\frac{\partial g(\theta)}{\partial \theta} \right)^2 \times Var(\hat{\theta}) \\ &= \left(-\frac{1}{(\theta-1)^2} \right)^2 \times \frac{\theta^2}{r} \\ &= \frac{\theta^2}{(\theta-1)^4 r} \end{aligned}$$

$$\begin{aligned} IC(g(\theta); (1 - \alpha) \times 100\%) &= g(\hat{\theta}) \pm z_{(1-\alpha)} \sqrt{Var(g(\hat{\theta}))} \\ &= \frac{\theta}{\theta-1} \pm z_{(1-\alpha)} \left(\frac{\theta}{(\theta-1)^2} \frac{1}{\sqrt{r}} \right) \end{aligned}$$

5° Questão

Deseja-se comparar duas populações de tempos de vida. Uma amostra de tamanho n ($r \leq n$ falhas) foi obtida da população 1 que tem distribuição Exponencial com média α . Uma amostra de tamanho m ($s \leq m$ falhas) foi obtida da população 2 que tem distribuição Exponencial com média $\alpha + \gamma$.

a) Estabeleça as hipóteses que se deseja testar.

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta \neq \theta_0 \end{cases} \Rightarrow \begin{cases} H_0 : (\alpha; \gamma) = (\alpha; 0) \\ H_1 : (\alpha; \gamma) \neq (\alpha; 0) \end{cases}$$

b) Apresente a função de verossimilhança para $\theta = (\alpha, \gamma)$.

Baseada nas duas amostras $\left((t_1, \delta_1), \dots, (t_n, \delta_n); (u_1, v_1), \dots, (u_m, v_m) \right)$

$$\begin{aligned} L(\theta) &= L(\alpha, \gamma) \propto \\ &\propto \prod_{i=1}^n \left[\lambda(t_i; \theta) \right]^{\delta_i} S(t_i; \theta) \times \prod_{j=1}^m \left[\lambda(u_j; \theta) \right]^{v_j} S(u_j; \theta) \\ &\propto \prod_{i=1}^n \left[\frac{1}{\alpha} \right]^{\delta_i} \exp \left\{ -\frac{t_i}{\alpha} \right\} \times \prod_{j=1}^m \left[\frac{1}{\alpha + \gamma} \right]^{v_j} \exp \left\{ -\frac{u_j}{\alpha + \gamma} \right\} \\ &\propto (\alpha)^{-\sum_{i=1}^n \delta_i} \exp \left\{ -\frac{\sum_{i=1}^n t_i}{\alpha} \right\} \times (\alpha + \gamma)^{-\sum_{j=1}^m v_j} \exp \left\{ -\frac{\sum_{j=1}^m u_j}{\alpha + \gamma} \right\} \\ &\propto (\alpha)^{-r} (\alpha + \gamma)^{-s} \exp \left\{ -\frac{\sum_{i=1}^n t_i}{\alpha} - \frac{\sum_{j=1}^m u_j}{\alpha + \gamma} \right\} \end{aligned}$$

c) Apresente o vetor escore $(U(\theta))$ e a matriz de informação $I(\theta)$.

$$U(\theta) = U(\alpha, \gamma) = \left(\frac{\partial l(\alpha, \gamma)}{\partial \alpha}, \frac{\partial l(\alpha, \gamma)}{\partial \gamma} \right)$$

$$\frac{\partial l(\alpha, \gamma)}{\partial \alpha} = -\frac{r}{\alpha} - \frac{s}{\alpha + \gamma} + \frac{\sum_{i=1}^n t_i}{\alpha^2} + \frac{\sum_{j=1}^m u_j}{(\alpha + \gamma)^2}$$

$$\frac{\partial l(\alpha, \gamma)}{\partial \gamma} = -\frac{s}{\alpha + \gamma} + \frac{\sum_{j=1}^m u_j}{(\alpha + \gamma)^2}$$

$$\therefore U(\theta) = \left(-\frac{r}{\alpha} - \frac{s}{\alpha + \gamma} + \frac{\sum_{i=1}^n t_i}{\alpha^2} + \frac{\sum_{j=1}^m u_j}{(\alpha + \gamma)^2} ; -\frac{s}{\alpha + \gamma} + \frac{\sum_{j=1}^m u_j}{(\alpha + \gamma)^2} \right)$$

A matriz de informação de Fisher, é dada por:

$$I(\theta) = \begin{vmatrix} \frac{\partial^2 l(\theta)}{\partial \alpha^2} & \frac{\partial^2 l(\theta)}{\partial \alpha \partial \gamma} \\ \frac{\partial^2 l(\theta)}{\partial \alpha \partial \gamma} & \frac{\partial^2 l(\theta)}{\partial \gamma^2} \end{vmatrix}$$

$$\frac{\partial^2 l(\theta)}{\partial \alpha^2} = \frac{r}{\alpha^2} - \frac{s}{(\alpha + \gamma)^2} - 2 \frac{\sum_{i=1}^n t_i}{\alpha^3} - 2 \frac{\sum_{j=1}^m u_j}{(\alpha + \gamma)^3}$$

$$\frac{\partial^2 l(\theta)}{\partial \gamma^2} = \frac{s}{(\alpha + \gamma)^2} - 2 \frac{\sum_{j=1}^m u_j}{(\alpha + \gamma)^3}$$

$$\frac{\partial^2 l(\theta)}{\partial \alpha \partial \gamma} = \frac{s}{(\alpha + \gamma)^2} - 2 \frac{\sum_{j=1}^m u_j}{(\alpha + \gamma)^3}$$

$$\therefore I(\theta) = \begin{pmatrix} \frac{r}{\alpha^2} + \frac{s}{(\alpha + \gamma)^2} - 2 \frac{\sum_{i=1}^n t_i}{\alpha^3} - 2 \frac{\sum_{j=1}^m u_j}{(\alpha + \gamma)^3} & \frac{s}{(\alpha + \gamma)^2} - 2 \frac{\sum_{j=1}^m u_j}{(\alpha + \gamma)^3} \\ \frac{s}{(\alpha + \gamma)^2} - 2 \frac{\sum_{j=1}^m u_j}{(\alpha + \gamma)^3} & \frac{s}{(\alpha + \gamma)^2} - 2 \frac{\sum_{j=1}^m u_j}{(\alpha + \gamma)^3} \end{pmatrix}$$

d) Obtenha a expressão da estatística do teste de Wald para as hipóteses em (a).

$$W = (\hat{\theta} - \theta_0) I(\theta_0) (\hat{\theta} - \theta_0)^T$$

$$W = (\hat{\alpha} - \alpha_0 \quad \hat{\gamma} - 0) \begin{pmatrix} \frac{r}{\alpha_0^2} + \frac{s}{(\alpha_0)^2} - 2 \frac{\sum_{i=1}^n t_i}{\alpha_0^3} - 2 \frac{\sum_{j=1}^m u_j}{(\alpha_0)^3} & \frac{s}{(\alpha_0)^2} - 2 \frac{\sum_{j=1}^m u_j}{(\alpha_0)^3} \\ \frac{s}{(\alpha_0)^2} - 2 \frac{\sum_{j=1}^m u_j}{(\alpha_0)^3} & \frac{s}{(\alpha_0)^2} - 2 \frac{\sum_{j=1}^m u_j}{(\alpha_0)^3} \end{pmatrix} \begin{pmatrix} \hat{\alpha} - \alpha_0 \\ \hat{\gamma} - 0 \end{pmatrix}$$

Sob H_0 , tem aproximadamente uma distribuição qui-quadrado com 2 graus de liberdade χ_2^2 . A um nível $100\alpha\%$ de significância, valores de W superiores ao valor tabelado da distribuição $\chi_{2,1-\alpha}^2$ indicam a rejeição de H_0