01. Seja X uma variável aleatória com função densidade de probabilidade dada por:

$$f(x|\theta) = \theta^2 x \, e^{-\theta x} \, I_A(x), A = (0, \infty), \theta > 0.$$

Queremos testar:

$$H_0 : 0 = 1 \quad vs \quad H_1 : 0 = 2.$$

Seja $X_1, X_2, \ldots, X_n$ uma amostra aleatória de X.

(i) Qual é a região crítica se $n = 5$ e $\alpha = 0,05$.

(ii) Se $n = 1$, qual o teste que minimiza $\alpha + \beta$? E qual o valor de $\alpha + \beta$?

Solução: Notemos que

$$X \sim Gama(\theta, 2).$$

Note que:

$$L(\theta; \mathbf{x}) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} \theta^2 x_i \, e^{-\theta x_i}.$$

Seja $s = \sum_{i=1}^{n} x_i$.

$$L(\theta; \mathbf{x}) = \theta^{2n} \prod_{i=1}^{n} x_i \, e^{-\theta s}$$

Se a hipótese nula é verdadeira temos $\theta = 1$:

$$L_0(\mathbf{x}) = \prod_{i=1}^{n} x_i \, e^{-s}$$

Se a hipótese alternativa é verdadeira temos $\theta = 2$:

$$L_1(\mathbf{x}) = 2^{2n} \prod_{i=1}^{n} x_i \, e^{-2s} = 4^n \prod_{i=1}^{n} x_i \, e^{-2s}.$$

Pelo Lema de Neyman-Pearson , utilizando a razão de verossimilhança simples, temos que o teste mais poderoso será aquele com região crítica dada por

$$A_1^* = \{\mathbf{x}; \frac{L_1(\mathbf{x})}{L_0(\mathbf{x})} \geq k\}.$$

Vamos com calma:

$$\frac{L_1(\mathbf{x})}{L_0(\mathbf{x})} = \frac{4^n \prod_{i=1}^{n} x_i \, e^{-2s}}{\prod_{i=1}^{n} x_i \, e^{-s}} = 4^n \, e^{-s} \geq c.$$

Assim

$$e^{-s} \geq c\, 4^{-n}$$

$$\log(e^{-s}) \geq \log(c4^{-n})$$

$$-s \geq \log(c4^{-n})$$

$$s = x_1 + x_2 + \ldots + x_n \leq k$$

Devemos achar a distribuição amostral de

$$S = \sum_{i=1}^{n} X_i \sim Gama(2n, \theta).$$

Sabemos que $S \sim Gama(r, \theta)$

$$V = 2\theta S \sim \chi^(2r).$$

No nosso caso temos $n = 5$

$$V = 2\theta S \sim \chi^2(4n) = \chi^2(20).$$

Olhando a tabela IV do Bussab& Morettin com $p = 0,95$ e $\ni= 20$ temos:

$$P(V \leq 10,851) = 0,05.$$

Se H_0 é verdade temos:

$$V = 2\theta S = 2S \sim \chi^2(20).$$

$$P(2S \leq 10,851) = 0,05.$$

$$P(S \leq 5,4255) = 0,05.$$

Assim nossa regra de decisão fica:

Se $S \leq 5,4255$ rejeitar H_0. caso contrário não rejeitar.

```
> alfa=0.05
> n=5
>
> teta_0=1;teta_2=2
>
> r=4*n;r
[1] 20
>
```

```
> P_5=qchisq(0.05,r);P_5
[1] 10.85081
>
> k=P_5/(2*teta_0);k; round(k,4)
[1] 5.425406
[1] 5.4254
>
> ##Podemos olhar direto na Gama(r=2n=10,teta_0=1)
>
> k=qgamma(0.05,2*n,teta_0);k
[1] 5.425406
>
```

Agora vamos fazer uma simulação para olhar com carinho a nossa região crítica:

Passo 1: Vamos usar o número de matrícula de cada aluno para gerar $N = 100000$ amostras de tamanho $n = 5$ de uma gama $(r = 2, \theta = 1)$. O meu número de matrícula será 33.

Passo 2: Vamos definir uma função para indicar de dada uma amostra de tamanho 5 rejeitamos ou não H_0.

Passo 3: Agora vamos contar em quantas amostras de tamanho 5 das 100000 amostras rejeitamos H_0 e calcular sua frequência relativa. Isto ns dará um estimativa de α.

Isto será feito no **R** através dos seguintes comandos:

```
>
> set.seed(33)
> amostras=matrix(rgamma(5*100000,2,1),nrow=5)
>
> round(amostras[,1:25],2)
     [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12] [,13] [,14]
[1,] 1.34 0.72 2.65 0.86 1.80 3.64 3.58 0.79 0.18  2.85  0.66  2.84  0.28  1.87
[2,] 3.70 2.99 1.08 0.47 2.31 1.31 1.97 1.20 3.38  3.00  3.21  0.93  1.07  1.91
[3,] 1.55 1.17 1.48 2.52 4.95 0.96 2.71 6.09 0.91  3.05  0.22  1.62  1.08  1.75
[4,] 1.05 1.71 1.02 2.12 3.85 3.68 2.60 0.40 0.28  1.58  2.11  2.10  1.62  1.52
[5,] 2.17 1.46 1.65 1.74 4.23 1.36 1.60 2.47 1.65  3.77  0.90  0.79  1.93  1.96
     [,15] [,16] [,17] [,18] [,19] [,20] [,21] [,22] [,23] [,24] [,25]
[1,]  3.34  0.25  1.53  2.81  1.42  1.53  0.95  1.88  2.60  1.68  0.57
[2,]  2.91  1.71  1.16  2.72  2.20  1.78  3.75  1.32  0.58  0.87  2.93
[3,]  1.03  0.65  3.06  3.25  1.38  0.45  1.42  1.81  3.06  0.91  1.34
[4,]  0.88  1.84  1.34  6.60  2.93  3.41  1.12  0.52  1.76  2.24  0.50
[5,]  3.20  2.54  1.34  2.17  8.79  0.56  1.93  0.85  0.75  2.02  0.74
>
>
> round(matrix(amostras[1:5,1],ncol=1),2)
     [,1]
[1,] 1.34
[2,] 3.70
[3,] 1.55
```

```
[4,] 1.05
[5,] 2.17
> round(matrix(amostras[1:5,73],ncol=1),2)
     [,1]
[1,] 0.54
[2,] 0.53
[3,] 0.27
[4,] 1.85
[5,] 1.28
>
>
> rejeita=function(x) {
+ ifelse( sum(x) <= qgamma(0.05,10,1),1,0)}
>
>
> rejeita(amostras[1:5,1])
[1] 0
>
> rejeita(amostras[1:5,2])
[1] 0
>
> rejeita(amostras[1:5,73])
[1] 1
>
>
> ####Vamaos aplicar nas 100000 amostras de tamanho 5.
>
> sum(apply(amostras,2,rejeita))/1000000
[1] 0.005085# bem próximo do nosso alfa!!!!
>
```

Vamos resolver o item **b**:

Pelo lema 6.3.1 , páginas 124 e 125, temos que:
O teste que minimiza $a\alpha + b\beta$ tem região crítica dada por:

$$A_1^* = \{\mathbf{x}; \frac{L_1(\mathbf{x})}{L_0(\mathbf{x})} \geq \frac{a}{b}\}.$$

Como queremos minimizar

$$\alpha + \beta = 1 \times \alpha + 1 \times \beta,$$

logo

$$a = b = 1.$$

Assim para $n = 1$ temos:

$$A_1^* = \{x; \frac{L_1(x)}{L_0(x)} \geq 1\}.$$

$$\frac{L_1(x)}{L_0(x)} = \frac{4\ x\ e^{-2x}}{x\ e^{-x}} = 4e^{-x} \geq 1.$$

$$e^{-x} \geq \frac{1}{4}$$

$$-x \geq -\log(4)$$

$$x \leq log(4).$$

Assim a região crítica que minimiza $\alpha + \beta$ é dada por:

$$A_1^* = \{x | x \leq log(4)\}.$$

Vamos calcular o nível de significância :

$$\alpha = P_{H_0}(X \in RC) = P(X < log(4))$$

Sabemos que :

$$X \sim Gama(r, \lambda)$$

$$P(X > x) = \sum_{i=0}^{r-1} \frac{e^{-\lambda x}(\lambda x)^i}{i!}$$

Se H_0 é verdade temos

$$X \sim Gama(2, \lambda = \theta = 1)$$

$$P(X > x) = \sum_{i=0}^{1} \frac{e^{-x}\ x^i}{i!} = e^{-x}\left(1 + x\right).$$

$$P(X > \log(4)) = e^{-log(4)}(1 + log(4)) = e^{log(1/4)}(1 + log(4))$$

$$P(X > \log(4)) = \frac{1 + log(4)}{4} = 0,5966.$$

$$\alpha = 1 - \frac{1 + log(4)}{4} = \frac{3 - log(4)}{4} = 0,4034.$$

```
> alfa=1-(1+log(4))/4
> alfa
[1] 0.4034264
>
>
> pgamma(log(4),2,1)
[1] 0.4034264
>
```

Vamos calcular o tamanho do erro do tipo II :

$$\beta = P_{H_1}(X \in RA) = P(X > log(4))$$

Sabemos que :

$$X \sim Gama(r, \lambda)$$

$$P(X > x) = \sum_{i=0}^{r-1} \frac{e^{-\lambda x}(\lambda x)^i}{i!}$$

Se H_1 é verdade temos

$$X \sim Gama(2, \lambda = \theta = 2)$$

$$P(X > x) = \sum_{i=0}^{1} \frac{e^{-2x}\ (2x)^i}{i!} = e^{-2x}\,(1 + 2x)\,.$$

$$P(X > \log(4)) = e^{-2\log(4)}\,(1 + 2\log(4))\,.$$

$$P(X > \log(4)) = e^{\log(1/16)}\,(1 + \log(16)) = \frac{1 + \log(16)}{16} = 0,2358.$$

```
 beta=(1+log(16))/16;beta
[1] 0.2357868
> pgamma(log(4),2,2,lower.tail=F)
[1] 0.2357868
> alfa+beta
[1] 0.6392132
>
```