

Respostas

Capítulo 1

Exercício 1.1.

a.

```
> total <- c(3,5,4,5)
> maduros <- c(0,1,3,5)
> x <- c(12.5,22.5,32.5,62.5)
```

b.

```
> pobs <- maduros/total
> pobs
[1] 0.00 0.20 0.75 1.00
```

c.

```
> beta0 <- 1.1630
> beta1 <- 0.2669
> p <- exp(beta0+beta1*(x-32.5)) / (1+exp(beta0+beta1*(x-32.5)))
> p
[1] 0.01514237 0.18153235 0.76187740 0.99989589
```

d.

```
> plot(x,pobs)
```

e. Figura 1

```
> plot(x,pobs)
> lines(x,p)
```

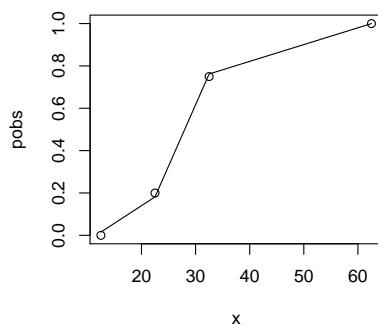


Figura 1:

Exercício 1.2.

```
> alfa <- 0.00001; beta <- 3
```

a.

```
> comp <- 40:80
```

b. Figura 2

```
> p <- alfa*comp^beta
> plot(comp,p)
```

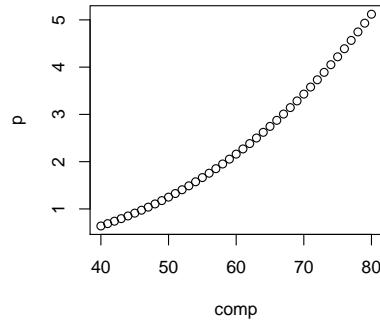


Figura 2:

c. Figura 3

```
> erro <- rnorm(length(comp), 0, 0.25)
> pr <- p+erro
> plot(comp,pr)
```

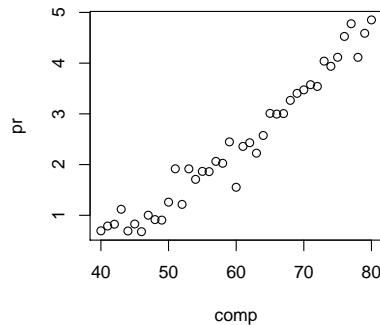


Figura 3:

Exercício 1.3.

a.

```
> modelo <- lm(log(pr)~log(comp))
```

b.

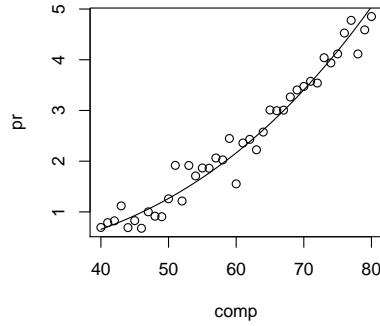
```
> alfa.est <- exp(coef(modelo)[1])
> beta.est <- coef(modelo)[2]
> c(alfa, alfa.est, beta, beta.est)
   (Intercept)          log(comp)
1.000000e-05 4.306399e-05 3.000000e+00 2.653661e+00
```

c.

```
> ppred <- alfa.est * comp^beta.est
```

d. Figura 4

```
> plot(comp, pr)
> lines(comp, ppred)
```



```
> AplusA <- c(A, A[A%in%A==FALSE])
> AplusA
[1] 1 15 29 43
```

A função `A%in%A` utilizada acima indica com valores lógicos (TRUE ou FALSE) os elementos do termo da esquerda que também aparecem no termo da direita.

```
> A%in%A
[1] TRUE TRUE TRUE TRUE
```

Vemos então que todos os termos de A tem correspondentes em A . Então, como pode ser visto abaixo, não há valores de A que não constam em A

```
> A[A%in%A==FALSE]
integer(0)
```

o que é também óbvio. Assim a concatenação dos vetores A e $A[A%in%A==FALSE]$ acaba resultando em um vetor equivalente a A .

Vamos mostrar agora que $AA = A$. O produto lógico AA deve conter somente os elementos que constam nas duas proposições que, neste caso são equivalentes. Assim:

```
> A[A%in%A==TRUE]
[1] 1 15 29 43
```

Vemos portanto que a listagem dos elementos de A que também constam em A , é obviamente equivalente a A .

Propriedade 2: $A + B = B + A$ e $AB = BA$ com A (idade = 1) e B (ipê amarelo).

As linhas que correspondem a A já foram selecionadas no exercício anterior. Vamos fazer a seleção das linhas que correspondem a B .

```
> B <- linhas[tab$sp=="ipa"]
```

Para obter a união lógica $A + B$ precisamos unir os valores de A e os valores de B que não constam em A . Para isto fazemos novamente uso da função `%in%`.

```
> AplusB <- c(A, B[B%in%A==FALSE])
```

De maneira similar obtemos a solução para a união lógica $B + A$.

```
> BplusA <- c(B, A[A%in%B==FALSE])
```

Ao listar $A + B$ e $B + A$ vemos então que estas uniões lógicas são equivalentes.

```
> AplusB
[1] 1 15 29 43 2 3 4 5 6 7 8 9 10 11 12 13 14
> BplusA
[1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 29 43
```

O produto lógico AB corresponde aos valores de A que também constam em B e, é fácil ver que ele é equivalente a BA que, da mesma forma, corresponde aos valores de B que constam em A :

```
> AB <- A[A%in%B]
> BA <- B[B%in%A]
> AB
[1] 1
> BA
[1] 1
```

Propriedade 3: $(A + B) + C = A + (B + C)$ e $(AB)C = A(BC)$ com A (idade = 1), B (ipê amarelo) e C (ipê qualquie

As seleções das linhas que atendem a A e a B já foram realizadas nos exercícios anteriores. Resta a seleção das linhas que correspondem a C .

```
> C <- linhas[(tab$sp=="ipa" | tab$sp=="ipb") & tab$idade >= 12]
```

Já obtivemos anteriormente a união lógica $A + B$ (`AplusB`) e portanto, podemos obter a união lógica $(A + B) + C$ com:

```
> AplusBplusC <- c(AplusB, C[C%in%AplusB==FALSE])
> AplusBplusC
[1] 1 15 29 43 2 3 4 5 6 7 8 9 10 11 12 13 14 26 27 28
```

Para o termo da direita da igualdade, primeiro obtemos a união lógica $B + C$:

```
> BplusC <- c(B, C[C%in%B==FALSE])
```

e, finalmente a união lógica $A = (B + C)$:

```
> c(A, BplusC[BplusC%in%A==FALSE])
[1] 1 15 29 43 2 3 4 5 6 7 8 9 10 11 12 13 14 26 27 28
```

Verificamos portanto que $(A + B) + C = A + (B + C)$.

Vamos mostrar agora que $(AB)C = A(BC)$, no entanto vamos considerar a proposição E (a árvore é um ipê) no lugar da C , o que resulta em uma solução com pelo menos uma árvore, o que é mais didático. Assim queremos verificar $(AB)E = A(BE)$. De todos os termos envolvidos resta selecionar E

```
> E <- linhas[tab$sp=="ipa" | tab$sp=="ipb"]
```

e obter BE

```
> BE <- B[B%in%E]
```

Agora podemos obter facilmente $(AB)E$

```
> AB[AB%in%E]
[1] 1
```

e $A(BE)$

```
> A[A%in%BE]
[1] 1
```

Vemos portanto que se confirma a igualdade $(AB)E = A(BE)$.

Propriedade 4: $A(B + C) = AB + AC$ com A (idade = 1) e B (ipê amarelo).

O único termo de que não dispomos é AC , o qual é obtido com:

```
> AC <- A[A%in%C==TRUE]
```

Seguimos agora com o cálculo de $A(B + C)$

```
> A[A%in%BplusC==TRUE]
[1] 1
```

e de $AB + AC$

```
> c(AB, AC)
[1] 1
```

Verifica-se portanto a igualdade.

Propriedade 5: $A + BC = (A + B)(A + C)$ com A (idade = 1), B (ipê amarelo) e C (ipê qualquer e idade maior ou igual a 12)

Ainda não dispomos de $A + C$ que é:

```
> AplusC <- c(A, C[C%in%A==FALSE])
```

Vamos agora verificar a igualdade $A + BC = (A + B)(A + C)$. O termo da direita é

```
> c(A, BC[BC%in%A==FALSE])
[1] 1 15 29 43 12 13 14
```

que é igual ao termo da esquerda

```
> AplusB[AplusB%in%AplusC]
[1] 1 15 29 43 12 13 14
```

Propriedade 6: $A'B' = (A + B)'$ com A (idade = 1) e B (ipê amarelo).

Começamos por selecionar A' e B' :

```
> Alinha <- linhas[linhas%in%A==FALSE]
> Blinha <- linhas[linhas%in%B==FALSE]
```

Agora vemos que $(A + B)'$ é:

```
> AplusBlinha <- linhas[linhas%in%AplusB==FALSE]
> AplusBlinha
[1] 16 17 18 19 20 21 22 23 24 25 26 27 28 30 31 32 33 34 35 36 37 38 39 40 41
[26] 42 44 45 46 47 48 49 50 51 52 53 54 55 56
```

o que equivale a $A'B'$ que é:

```
> Alinha[Alinha%in%Blinha]
[1] 16 17 18 19 20 21 22 23 24 25 26 27 28 30 31 32 33 34 35 36 37 38 39 40 41
[26] 42 44 45 46 47 48 49 50 51 52 53 54 55 56
```

Propriedade 7: $A' + B' = (AB)'$ com A (idade = 1) e B (ipê amarelo).

Já dispomos de todos os elementos necessários. O termo da esquerda da igualdade $A' + B'$ é:

```
> c(Alinha,Blinha[Blinha%in%Alinha==FALSE])
[1] 2 3 4 5 6 7 8 9 10 11 12 13 14 16 17 18 19 20 21 22 23 24 25 26 27
[26] 28 30 31 32 33 34 35 36 37 38 39 40 41 42 44 45 46 47 48 49 50 51 52 53 54
[51] 55 56 15 29 43
```

o que equivale a $(AB)'$ que é:

```
> linhas[linhas%in%AB==FALSE]
[1] 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
[26] 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51
[51] 52 53 54 55 56
```

Propriedade 8: Se $B \Rightarrow E$ então $BE = B$ com B (a árvore é um ipê amarelo)

A proposição a árvore é um ipê amarelo (B) implica obviamente que a árvore é um ipê (E). Verificamos então que BE

```
> B[B%in%E==TRUE]
[1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14
```

é igual a B

```
> B
[1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14
```

Propriedade 9: Se $A \Rightarrow F$ e $F \Rightarrow A$, então $A = F$.

Se $A \Rightarrow F$ isto significa que $AF = A$. Se $F \Rightarrow A$ isto significa que $FA = F$. Mas, pela propriedade 2 temos que $AF = FA$ e portanto $A = F$.

b.

Se A (idade = 1), C (ipê qualquer e idade maior ou igual a 12) e D (idade maior ou igual a 4 e menor que 12) são exclusivos, eles não podem ser simultaneamente verdadeiros. De fato, é fácil verificar que não há elementos que atendem simultaneamente a A , C e D . Primeiro selecionamos D que é a única proposição para a qual ainda não temos valores.

```
> D <- linhas[tab$idade>=4&tab$idade<12]
```

Por fim vemos que não há qualquer componente de A que também satisfaça as proposições C e D :

```
> A%in%C%in%D
[1] FALSE FALSE FALSE FALSE
```

Vamos agora verificar que a proposição $F = (A + C + D)'$ torna o grupo de proposições exaustivo e que F de fato define o restante das árvores. Primeiro selecionamos $A + C + D$ e fazemos uso de $A + C$ de que já dispomos.

```
> AplusCplusD <- c(AplusC,D[D%in%AplusC==FALSE])
> AplusCplusD
[1] 1 15 29 43 12 13 14 26 27 28 4 5 6 7 8 9 10 11 18 19 20 21 22 23 24
[26] 25 32 33 34 35 36 37 38 39 46 47 48 49 50 51 52 53
```

Agora selecionamos F

```
> F <- linhas[linhas%in%AplusCplusD==FALSE]
> F
[1] 2 3 16 17 30 31 40 41 42 44 45 54 55 56
```

Verifica-se portanto que F define o restante das árvores. Para mostrar que $F = (A + C + D)' = A'C'D'$ primeiro vamos selecionar os termos C' e D' que ainda não temos.

```
> Clinha <- linhas[linhas%in%C==FALSE]
> Dlinha <- linhas[linhas%in%D==FALSE]
```

Selecionamos a proposição $A'C'D'$ em duas etapas. Primeiro obtemos $A'C'$

```
> AlinhaCinha <- Alinha[Alinha%in%Cinha]
```

e finalmente $A'C'D'$

```
> AlinhaCinhaDlinha <- AlinhaCinha[AlinhaCinha%in%D==FALSE]
> AlinhaCinhaDlinha
[1] 2 3 16 17 30 31 40 41 42 44 45 54 55 56
```

que é equivalente à proposição F que já foi mostrada mais acima.

Exercício 2.2.

Nesta parte fazemos uso de alguns dos elementos construídos na resposta para o exercício anterior.

a.

Cálculo de $P(A)$:

```
> pa <- length(A) / length(linhas)
> pa
[1] 0.07142857
```

Cálculo de $P(B)$:

```
> pb <- length(B) / length(linhas)
> pb
[1] 0.25
```

Cálculo de $P(C)$:

```
> pc <- length(C) / length(linhas)
> pc
[1] 0.1071429
```

Cálculo de $P(D)$:

```
> pd <- length(D) / length(linhas)
> pd
[1] 0.5714286
```

b.

A probabilidade $P(B|C)$ corresponde a idéia de se calcular a probabilidade de B dado que C é fato. Pela Lei do Produto temos que $P(BC) = P(C)P(B|C)$. Portanto $P(B|C) = P(BC)/P(C)$. Fazemos então uso dessa relação para resolver o problema proposto. Primeiro calculamos $P(BC)$:

```
> pbc <- length(BC) / length(linhas)
> pbc
[1] 0.05357143
```

Agora é só calcular a razão:

```
> pbc / pc
[1] 0.5
```

Verificamos portanto que $P(B) \neq P(B|C)$, o que indica que as proposições (ou eventos) B e C não são independentes.

c.

Pela Lei da Adição para proposições exclusivas temos que $P(A + D) = P(A) + P(D)$. Portanto a solução é simplesmente:

```
> pa + pd
[1] 0.6428571
```

Exercício 2.3.

$$\begin{aligned}
P(E_1 + E_2|H) &= 1 - P[(E_1 + E_2)'|H] \\
&= 1 - P(E'_1 E'_2|H) \\
&= 1 - P(E'_1|H)P(E'_2|E'_1 H) \\
&= 1 - P(E'_1|H)[1 - P(E_2|E'_1 H)] \\
&= 1 - P(E'_1|H) + P(E'_1|H)P(E_2|E'_1 H) \\
&= P(E_1|H) + P(E_2|H)P(E'_1|E_2 H) \\
&= P(E_1|H) + P(E_2|H)[1 - P(E_1|E_2 H)] \\
&= P(E_1|H) + P(E_2|H) - P(E_2|H)P(E_1|E_2 H) \\
&= P(E_1|H) + P(E_2|H) - P(E_1 E_2|H)
\end{aligned}$$

Exercício 2.4

Para mostrar isso recorremos à definição de eventos exclusivos para concluir que $E'_1 + E'_2$ é dedutível de H e portanto $P(E'_1 + E'_2|H) = 1$. Portanto

$$P(E'_1 + E'_2|H) = P(E'_1 E'_2|H) = 1 - P(E_1 E_2|H)$$

o que leva ao resultado esperado

$$P(E_1 E_2|H) = 1 - P(E'_1 + E'_2|H) = 1 - 1 = 0$$

Exercício 2.5.

A partir da solução mostrada no exercício prévio temos que $P(B + C) = P(B) + P(C) - P(BC)$. Todos estes termos já haviam sido calculados anteriormente com o R e portanto, fazendo uso do que já dispomos, a resposta é:

```
> pb+pc-pbc
[1] 0.3035714
```

Exercício 2.6.

Denominamos de F a proposição “ganha o oscar pela fotografia” e D a proposição “ganha o oscar pela direção”. Portanto temos que $P(F) = 0.3$, $P(D) = 0.2$ e $P(DF) = 0.05$.

a. $P(D + F) = P(D) + P(F) - P(DF) = 0.45$

b.

A probabilidade de que não ganhe prêmio algum é complementar da probabilidade de que ganhe pelo menos um prêmio. Portanto: $P((D + F)') = 1 - P(D + F) = 0.55$.

Exercício 2.7.

Vamos denotar o par de intensidade dos dois fenômenos com “intensidade do primeiro – intensidade do segundo”. Por exemplo o termo 2 – 4 indica que o primeiro fenômeno tem intensidade 2, enquanto que o segundo tem intensidade 4. Como há seis possibilidades de intensidade (1a6) para cada fenômeno, há $6 \cdot 6 = 36$ pares possíveis. Podemos construir uma tabela com todas essas combinações da seguinte forma:

```
> pares <- data.frame(expand.grid(c(1:6), c(1:6)))
```

A título de ilustração segue abaixo as sete primeiras linhas dessa tabela:

```
> pares[1:7,]
  Var1 Var2
1     1     1
2     2     1
3     3     1
4     4     1
5     5     1
6     6     1
7     1     2
```

Pelo princípio da permutabilidade temos que a ocorrência de cada um dos pares tem probabilidade de 1/36.

a.

```
> sum((pares$Var1+pares$Var2<=8)&pares$Var1<4)/nrow(pares)
[1] 0.4166667
```

b.

```
> sum(pares$Var1<=3|pares$Var2<=3)/nrow(pares)
[1] 0.75
```

Exercício 2.8.

A proposição “a bactéria se mostra suscetível a pelo menos uma droga” é $(S_1 + S_2 + S_3)$. A probabilidade que queremos é então:

$$P(S_1 + S_2 + S_3) = 1 - P[(S_1 + S_2 + S_3)']$$

Pela propriedade 5 mostrada no capítulo 2 temos que o complementar da ocorrência de pelo menos uma proposição é o produto lógico do complementar de cada uma das proposições individuais. Assim:

$$P(S_1 + S_2 + S_3) = 1 - P[(S_1 + S_2 + S_3)'] = 1 - P(S'_1 S'_2 S'_3)$$

Como o efeito de uma das drogas não afeta os efeitos das demais, S_1 , S_2 e S_3 são proposições independentes. Portanto:

$$\begin{aligned} P(S_1 + S_2 + S_3) &= 1 - P(S'_1 S'_2 S'_3) \\ &= 1 - [P(S'_1)P(S'_2)P(S'_3)] \\ &= 1 - 0.1^3 = 0.999 \end{aligned}$$

Exercício 2.9.

A probabilidade de sucesso (S) na i -ésima tentativa é $P(S_i) = 0.16$. A probabilidade de sucesso em pelo menos $n = 12$ tentativas é:

$$\begin{aligned} P(S_1 + \dots + S_n) &= 1 - P[(S_1 + \dots + S_n)'] \\ &= 1 - \prod_{i=1}^n P(S'_i) = 1 - 0.84^{12} \cong 0.877 \end{aligned}$$

Exercício 2.10.

Denominamos as proposições A “o indivíduo sob júdice é realmente o progenitor” e B “o teste de paternidade resulta positivo”. Pelas informações disponíveis temos que $P(B|A) = 0.98$ e que $P(B|A') = 0.05$. Os autos indicam que $P(A) = 0.15$. Portanto:

$$\begin{aligned} P(A|B) &= \frac{P(A)P(B|A)}{P(B)} \\ &= \frac{P(A)P(B|A)}{P(A)P(B|A) + P(A')P(B|A')} \\ &= \frac{0.15 \cdot 0.98}{0.15 \cdot 0.98 + 0.85 \cdot 0.05} \cong 0.776 \end{aligned}$$

Exercício 2.11.

Vamos denotar por A a proposição “o candidato acerta a questão” e por E a proposição “o candidato estudou o tema da questão”. A partir das informações disponíveis no enunciado temos que $P(E) = p$, $P(A|E) = 1$ e que $P(A|E') = 1/m$. Portanto a resposta procurada é:

$$\begin{aligned} P(E|A) &= \frac{P(E)P(A|E)}{P(A)} \\ &= \frac{P(E)P(A|E)}{P(E)P(A|E) + P(E')P(A|E')} \\ &= \frac{p \cdot 1}{p \cdot 1 + (1-p) \cdot \frac{1}{m}} = \frac{p \cdot m}{1 + p(m-1)} \end{aligned}$$

Exercício 2.12.

Vamos denominar de A a proposição “o primeiro técnico identifica corretamente a bactéria” e de B a proposição “o segundo técnico identifica corretamente a bactéria”. A partir das informações disponíveis temos que $P(A) = 2/3$,

$P(B|A) = 3/4$ e $P(B|A') = 1/5$.

a.

Pela “extensão da conversa” temos:

$$P(B) = P(A)P(B|A) + P(A')P(B|A') = \frac{2}{3} \cdot \frac{3}{4} + \frac{1}{3} \cdot \frac{1}{5} = \frac{17}{30} \cong 0.567$$

b.

$$\begin{aligned} P(A+B) &= P(A) + P(B) - P(AB) \\ &= P(A) + P(B) - P(A)P(B|A) \\ &= \frac{2}{3} + \frac{17}{30} - \frac{2}{3} \cdot \frac{3}{4} = \frac{11}{15} \cong 0.733 \end{aligned}$$

Exercício 2.13.

Denotaremos como E a proposição “o empreiteiro ganha a concorrência para a parte elétrica” e como H a proposição “o empreiteiro ganha a concorrência para a parte hidráulica. A partir das informações disponíveis temos $P(E) = 1/2$, $P(H|E) = 3/4$ e $P(H|E') = 1/3$.

a.

$$P(EH) = P(E)P(H|E) = \frac{1}{2} \cdot \frac{3}{4} = \frac{3}{8} = 0.375$$

b.

Há mais de uma maneira de resolver o problema. Uma delas começa com a consideração direta das maneiras de ganhar apenas um contrato. Pode-se ganhar a elétrica e perder a hidráulica, e vice-versa. Partindo disto temos que a probabilidade de ganhar apenas um contrato é:

$$P(EH' + E'H) = P(EH') + P(E'H) = P(E)P(H'|E) + P(E')P(H|E')$$

No enunciado não é fornecida as probabilidades $P(E')$ e $P(H'|E)$, mas estes dois valores são facilmente obtidos com as soluções $P(E') = 1 - P(E) = 1/2$ e $P(H'|E) = 1 - P(H|E) = 1/4$. Agora é só fazer as substituições devidas:

$$P(EH' + E'H) = \frac{1}{2} \cdot \frac{1}{4} + \frac{1}{2} \cdot \frac{1}{3} = \frac{7}{24} \cong 0.292$$

c.

Pela “extensão da conversa” temos:

$$P(H) = P(E)P(H|E) + P(E')P(H|E') = \frac{1}{2} \cdot \frac{3}{4} + \frac{1}{2} \cdot \frac{1}{3} = \frac{13}{24} \cong 0.542$$

d.

$$P(E|H) = \frac{P(E)P(H|E)}{P(H)} = \frac{\frac{1}{2} \cdot \frac{3}{4}}{\frac{13}{24}} = \frac{9}{13} \cong 0.692$$

Exercício 2.14.

A partir das informações disponíveis temos $P(A'|M') = 0.2$, $P(A|M) = 0.3$ e $P(M') = 0.78$.

a.

A probabilidade de que o freguês seja mulher é $P(M) = 1 - P(M') = 1 - 0.78 = 0.22$.

b.

A probabilidade de seja escolhido um prato de salada dado que o freguês é homem é $P(A'|M') = 0.2$ como consta no enunciado.

c.

A probabilidade de que seja escolhido um prato a base de carne dado que o freguês é uma mulher é $P(A|M) = 0.3$, o que também já havia sido fornecido no enunciado.

d.

A probabilidade de que um freguês selecionado aleatoriamente seja homem e que o prato seja a base de carne é $P(AM') = P(M')P(A|M') = P(M')[1 - P(A'|M')] = 0.78(1 - 0.2) = 0.624$.

e.

A probabilidade de que o freguês seja homem ou que escolha um prato a base de carne é $P(A + M') = P(A) + P(M') - P(AM')$. Ainda não dispomos da probabilidade de que seja escolhido um prato a base de carne $P(A)$ mas, podemos obter uma solução pela “extensão da conversa”:

$$\begin{aligned} P(A) &= P(M)P(A|M) + P(M')P(A|M') \\ &= [1 - P(M')][P(A|M) + P(M')[1 - P(A'|M')]] \\ &= (1 - 0.78) \cdot 0.3 + 0.78 \cdot (1 - 0.2) = 0.69 \end{aligned}$$

Agora dispomos de todos os elementos para resolver o problema e portanto $P(A + M') = P(A) + P(M') - P(AM') = 0.690 + 0.780 - 0.624 = 0.846$.

f.

A probabilidade de que o freguês seja uma mulher dado que tenha sido escolhido um prato a base de salada é:

$$\begin{aligned} P(M|A') &= \frac{P(M)P(A'|M)}{P(A')} \\ &= \frac{[1 - P(M')][1 - P(A|M)]}{1 - P(A)} \\ &= \frac{(1 - 0.78)(1 - 0.3)}{1 - 0.69} \cong 0.497 \end{aligned}$$

Exercício 2.15.

Denominamos inicialmente as seguintes proposições: A (a cor do táxi criminoso é azul); A' (a cor do táxi criminoso é verde); a (a testemunha identifica a cor do táxi como sendo azul); e a' (a testemunha identifica a cor do táxi como sendo verde).

A partir da informação fornecida no item “a” temos que $P(A') = 0.85$, o que implica que $P(A) = 1 - P(A') = 0.15$. Já a informação do item “c” indica que $P(a|A) = P(a'|A') = 0.80$. Agora podemos responder a questão. A probabilidade de que o táxi criminoso seja de fato azul dado que a testemunha o identificou como azul é:

$$\begin{aligned} P(A|a) &= \frac{P(A)P(a|A)}{P(a)} \\ &= \frac{P(A)P(a|A)}{P(A)P(a|A) + P(A')P(a|A')} \\ &= \frac{P(A)P(a|A)}{P(A)P(a|A) + P(A')[1 - P(a'|A')]} \\ &= \frac{0.15 \cdot 0.80}{0.15 \cdot 0.80 + 0.85 \cdot (1 - 0.80)} \cong 0.414 \end{aligned}$$

Ou seja, é mais provável que o táxi seja verde $P(A'|a) = 1 - P(A|a) \cong 0.586$, apesar da testemunha o ter identificado como azul. Isto se deve à probabilidade de falsa identificação por parte da testemunha e à dominância da cor verde nos táxis da cidade.

Exercício 2.16.

a.

```
> 100*((1-0.1)^2)*(1/4) + 100*(0.1^2)*(3/4)
[1] 21
```

b.

```
> 100*((1-0.25)^2)*(1/4) + 100*(0.25^2)*(3/4)
[1] 18.75
```

c. Figura 5

```

> p <- seq(0.01, 0.99, by = 0.01)
> S.medio <- 100*((1-p)^2)*(1/4) + 100*(p^2)*(3/4)
> c(min(S.medio),p[S.medio==S.min])
[1] 18.75 0.25
> plot(p,S.medio,type="l")
> points(p.min,S.min, pch=19)
> segments(p.min,0,p.min,S.min)

```

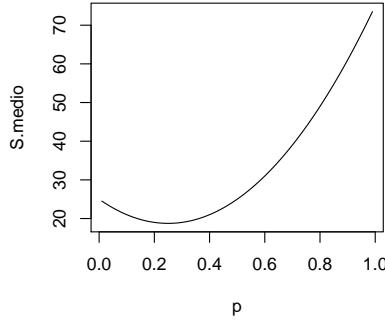


Figura 5:

Exercício 2.17.

a.

Começamos por solucionar $P(B|D'L_1)$

$$P(B|D'L_1) = \frac{BD'L_1}{P(D'L_1)} = \frac{P(BL_1)P(D'|BL_1)}{P(D'L_1)}$$

O numerador da equação indicada acima é:

$$P(BL_1)P(D'|BL_1) = P(L_1)P(B|L_1)[1 - P(D|BL_1)] = 0.33 \cdot 0.80 \cdot (1 - 0.6) = 0.1056$$

O denominador é:

$$\begin{aligned}
P(D'L_1) &= P(D'(B + B')L_1) \\
&= P(D'BL_1) + P(D'B'L_1) \\
&= P(D'BL_1) + P(B'L_1)P(D'|B'L_1) \\
&= P(D'BL_1) + P(L_1)P(B'|L_1)P(D'|B'L_1) \\
&= P(D'BL_1) + P(L_1)[1 - P(B|L_1)]P(D'|B'L_1) \\
&= 0.1056 + 0.33 \cdot (1 - 0.8) \cdot 1 = 0.1716
\end{aligned}$$

Portanto $P(B|D'L_1) = 0.1056/0.1716 \cong 0.615$.

O desenvolvimento da solução para $P(B|D'L_3)$ é similar à mostrada acima, a única diferença é que estamos tratando do lago L_3 ao invés do L_1 . Substituindo L_1 por L_3 nas equações acima e, efetuando as operações algébricas obtemos $P(B|D'L_3) = 0.0017/0.1547 \cong 0.0110$.

b.

Não deveriam porque as condicionantes $D'L_1$, $D'L_2$ e $D'L_3$ não são exaustivas. Falta a consideração das probabilidades de B condicionadas a DL_1 , DL_2 e DL_3 .

c.

A probabilidade desejada para um lago j qualquer é:

$$P(L_j|D') = \frac{P(L_j)P(D'|L_j)}{P(D')}$$

Na seção que trata do “Teorema de Bayes” constam as probabilidades $P(L_1) = 0.33$, $P(L_2) = 0.50$, $P(L_3) = 0.17$ e também o desenvolvimento da solução $P(D') = 0.651$. Resta portanto o cálculo de $P(D'|L_j)$, que é:

$$\begin{aligned} P(D'|L_j) &= P(D'(B + B')|L_j) \\ &= P(D'B|L_j) + P(D'B'|L_j) \\ &= P(D'|BL_j)P(B|L_j) + P(D'|B'L_j)P(B'|L_j) \\ &= [1 - P(D|BL_j)]P(B|L_j) + 1 \cdot [1 - P(B|L_j)] \end{aligned}$$

Substituindo os termos conhecidos e efetuando as operações necessárias obtemos $P(D'|L_1) = (1 - 0.60) \cdot 0.80 + (1 - 0.80) = 0.52$, $P(D'|L_2) = (1 - 0.70) \cdot 0.50 + (1 - 0.50) = 0.65$ e $P(D'|L_3) = (1 - 0.90) \cdot 0.10 + (1 - 0.10) = 0.91$. Portanto, as probabilidades desejadas são:

$$P(L_1|D') = \frac{0.33 \cdot 0.52}{0.651} \cong 0.264 \quad P(L_2|D') = \frac{0.50 \cdot 0.65}{0.651} \cong 0.499 \quad P(L_3|D') = \frac{0.17 \cdot 0.91}{0.651} \cong 0.238$$

d.

Precisamos obter os três termos à direita da marginalização

$$P(B|D') = P(B(L_1 + L_2 + L_3)|D') = P(BL_1|D') + P(BL_2|D') + P(BL_3|D')$$

O desenvolvimento para um lago genérico j é:

$$P(BL_j|D') = \frac{P(BL_j)P(D'|BL_j)}{P(D')} = \frac{P(L_j)P(B|L_j)[1 - P(D|BL_j)]}{P(D')}$$

Dessa forma temos que

$$\begin{aligned} P(BL_1|D') &= \frac{0.33 \cdot 0.80 \cdot 0.40}{0.651} \cong 0.162 \\ P(BL_2|D') &= \frac{0.50 \cdot 0.50 \cdot 0.30}{0.651} \cong 0.115 \\ P(BL_3|D') &= \frac{0.17 \cdot 0.10 \cdot 0.10}{0.651} \cong 0.00261 \end{aligned}$$

Finalmente obtemos $P(B|D') = 0.162 + 0.115 + 0.00261 \cong 0.280$.

Capítulo 3

Exercício 3.1.

a.

Identifica-se que é um problema que pode ser resolvido com a distribuição geométrica, que é um caso particular da binomial negativa. Sendo U o número de experimentos e X o número de fracassos antes de se obter a sucessos, temos que $U = X + a$. Sabendo que $a = 1$, a probabilidade pretendida é:

$$P(U > 8) = P(X+a > 8) = P(X > 8-a) = P(X > 7) = 1 - P(X \leq 7) = 1 - [P(X = 0) + \dots + P(X = 7)]$$

Com o auxílio do R temos que a solução é:

```
> x <- 0:7
> 1-sum(dnbinom(x, 1, 1/6))
[1] 0.2325680
```

b. A distribuição geométrica (equivalente à binomial negativa com $a = 1$) definida em termos da variável U para este problema é:

$$P(u) = \theta(1 - \theta)^{u-1} = (1/6)(1 - 1/6)^{u-1} = (1/5)(5/6)^u$$

Precisamos de um valor de u^* tal que

$$\begin{aligned} P(U \leq u^*) &= [P(u = 1) + \dots + P(u = u^*)] \geq 0.95 \\ &= (1/5) \cdot [(5/6)^1 + \dots + (5/6)^{u^*}] \geq 0.95 \end{aligned}$$

Na prática facilita pensar somente na igualdade e encontrar o valor u^* que corresponde a

$$\begin{aligned} P(U \leq u^*) &= (1/5) \cdot [(5/6)^1 + \cdots + (5/6)^{u^*}] = 0.95 \\ &= [(5/6)^1 + \cdots + (5/6)^{u^*}] = 4.75 \end{aligned}$$

Se u^* for um inteiro teremos encontrado valor que procuramos, caso contrário a resposta será o próximo inteiro maior que u^* , que certamente resulta em uma probabilidade ≥ 0.95 . A expressão $[(5/6)^1 + \cdots + (5/6)^{u^*}]$ corresponde a uma progressão geométrica. O primeiro termo é $t_1 = 5/6$, a razão entre termos é $q = 5/6$. A soma dos $n = u^*$ termos dessa progressão deve ser $S_n = 4.75$. A soma para uma progressão geométrica finita é:

$$S_n = \frac{t_1(q^n - 1)}{q - 1}$$

Resolvendo para n e substituindo os termos conhecidos temos:

$$n = \frac{\log[S_n(q - 1)/t_1 + 1]}{\log q} = \frac{\log \left[\frac{4.75(5/6 - 1)}{5/6} + 1 \right]}{\log(5/6)} \cong 16.431$$

Portanto, precisamos de um valor equivalente ao inteiro $u^* = \lceil n \rceil = 17$ como número de experimentos para garantir uma probabilidade de sucesso de no mínimo 0.95.

Utilizamos as propriedades da progressão geométrica e mostramos acima uma maneira analítica de se obter o valor de u^* que satisfaz a expressão $[(5/6)^1 + \cdots + (5/6)^{u^*}] = 4.75$. No entanto, poderíamos ter obtido este valor na “força bruta” com auxílio do R. Um código para isto seria:

```
> n <- 1
> sol <- (5/6)^n
> while(sol<4.75) {
+   n <- n+1
+   sol <- sol+(5/6)^n
+ }
> c(n,sol)
[1] 17.000000 4.774634
```

Neste cálculo vamos fazendo sequencialmente cálculos, com incrementos gradativos em n até que atingimos a meta, ou seja uma solução (sol no código) de no mínimo 4.75. Vemos então que com n exatamente igual a 17 obtemos aproximadamente 4.775.

Exercício 3.2.

Denotamos inicialmente as proposições: C_1 (o vendedor faz um contato no dia); C_2 (o vendedor faz dois contatos no dia) e v (ocorre alguma venda). A partir das informações disponíveis temos então que: $P(C_1) = 1/3$, $P(C_2) = 2/3$, $P(v|C_1) = 1/20$ e $P(v|C_2) = 1/10$. As entradas respectivas para estes elementos no R são:

```
> pc1 <- 1/3; pc2 <- 2/3; pv.c1 <- 1/20; pv.c2 <- 1/10
```

A probabilidade de que em um dia qualquer (assumindo independência) ocorra alguma venda pode ser obtida com a extensão da conversa $P(v) = P(C_1)P(v|C_1) + P(C_2)P(v|C_2)$. O valor calculado corresponde ao objeto pv no código abaixo:

```
> pv <- pc1*pv.c1+pc2*pv.c2
> pv
[1] 0.08333333
```

Se assumimos que o fato de ocorrer ou não venda em um dia não afeta a probabilidade de que ocorra uma venda em qualquer dos demais dias, temos então que o problema pode ser resolvido com a distribuição binomial. Temos $n = 5$ dias em que a probabilidade de venda em cada um deles é $\theta = P(V)$. A variável Y (número de dias com venda) pode assumir então os valores de 0 a 5. Portanto a distribuição de probabilidade para Y é:

```
> y <- 0:5
> n <- 5
> prob <- dbinom(y,n,pv)
> prob
[1] 6.472278e-01 2.941945e-01 5.348990e-02 4.862719e-03 2.210327e-04
[6] 4.018776e-06
```

e, pelas propriedades da binomial $[E(Y) = n \cdot \theta]$ a esperança é:

```
> n*pv
[1] 0.4166667
```

Exercício 3.3.

Identifica-se que o problema pode ser resolvido com a distribuição binomial.

a.

Queremos $P(X > 7)$ mas o número de operações é menor se resolvemos usando o complementar $P(X > 7) = 1 - P(X \leq 7)$. Até por que há funções prontas para o cálculo de distribuições de probabilidade $F(x) = P(X \leq x)$. Portanto:

```
> teta <- 0.18; n <- 30; x <- 7
> 1-pbinom(x,n,teta)
[1] 0.1582117
```

b.

Aqui precisamos da esperança e da raiz quadrada da variância da binomial, que são, respectivamente:

```
> c(n*teta,sqrt(n*teta*(1-teta)))
[1] 5.400000 2.104281
```

Exercício 3.4.

Identifica-se que o problema pode ser resolvido com uma distribuição binomial se é assumido independência entre os óbitos. A probabilidade de que um óbito qualquer tenha como causa o infarto do miocárdio é $\theta = 0.149$.

a.

Aqui precisamos da esperança e da raiz quadrada da variância de uma variável que tem distribuição binomial:

```
> teta <- 0.149; n <- 5000
> media <- n*teta; dp <- sqrt(n*teta*(1-teta))
> c(media,dp)
[1] 745.00000 25.17926
```

b.

Verificamos a quantos desvios padrões de distância se encontra 846 em relação à média:

```
> mortes <- 836
> (mortes-media)/dp
[1] 3.614086
```

Portanto, pelo critério utilizado, esta é uma cidade incomum no que diz respeito à proporção de óbitos em virtude de infarto do miocárdio.

c.

A suposição do patologista é $\theta = 0.149$. Procuramos pela probabilidade $P(X \geq 836)$ sob essa suposição em um cenário com $n = 5000$. A quantidade de cálculos necessários é menor se optarmos por resolver o complementar $P(X \geq 836) = 1 - P(X < 836) = 1 - P(X \leq 835)$. Ainda assim o número de operações é elevado, mas não é impraticável:

```
> x <- 835
> 1-pbinom(x,n,teta)
[1] 0.0001987085
```

Se optássemos por resolver o problema com a aproximação pela normal com $\mu = n \cdot \theta$ e $\sigma = \sqrt{n \cdot \theta \cdot (1 - \theta)}$, como é o sugerido tradicionalmente, teríamos:

```
> 1-pnorm(x,media,dp)
[1] 0.0001755356
```

Que é de fato uma boa aproximação neste caso. Independente do mecanismo de cálculo vemos que a probabilidade de que ocorram 836 ou mais óbitos por infarto do miocárdio é realmente pequena. Portanto, em sendo verdade que para as cidades da região é esperada uma proporção de aproximadamente 0.15 ($\theta = 0.149$), a cidade avaliada realmente apresenta um resultado realmente preocupante, com um número de casos elevado para o padrão da região.

Exercício 3.5.

Assumindo independência podemos utilizar a distribuição binomial para resolver o problema, com probabilidade $\theta = 0.4$.

a.

```
> teta <- 0.4; n <- 8
> dbinom(0,n,teta)
[1] 0.01679616
```

b.

```
> pbinom(2,n,teta)
[1] 0.3153946
```

c.

O fato de que mais da metade é quitado até o vencimento, implica que menos da metade tem atraso no pagamento $P(X \leq 3)$.

```
> pbinom(3,n,teta)
[1] 0.5940864
```

Exercício 3.6.

A probabilidade de que um passageiro que comprou o bilhete falte é $\theta = 0.05$ e foram vendidos $n = 84$ bilhetes para 80 assentos. Assumindo independência entre os comparecimentos, podemos usar a distribuição binomial.

a.

Para que todos os que comparecem consigam assentos devem comparecer 80 ou menos, ou seja, devem faltar 4 ou mais passageiros. Se X é a quantidade de passageiros que falta, procuramos $P(X \geq 4) = 1 - P(X < 4) = 1 - P(X \leq 3)$:

```
> teta <- 0.05; n <- 84
> pbinom(3,n,teta)
[1] 0.3896925
```

b.

Aqui estamos interessados na esperança:

```
> n*teta
[1] 4.2
```

c.

Se n é o número de bilhetes vendidos, $n - 80$ é o número de bilhetes em excesso. A intenção é de que a probabilidade seja alta de que falte um quantidade de passageiros pelo menos igual à quantidade de bilhetes em excesso ($X \geq n - 80$). Assim garante-se que os que comparecerem terão assento. Queremos especificamente calcular um n tal que $P(X \geq n - 80) > 0.80$. Na prática é mais fácil resolver pelo complementar: $P(X \geq n - 80) = 1 - P(X < n - 80) = 1 - P(X \leq n - 81) > 0.8$. Podemos calcular essa probabilidade com cálculos subsequentes com o incremento gradual de n enquanto a meta (probabilidade > 0.80) for atendida. Quando a meta for ultrapassada (probabilidade ≤ 0.80) teremos exagerado no valor de n , e o número de bilhetes procurado será equivalente a $n - 1$. Segue abaixo um código com a solução:

```
> n <- 81
> prob <- 1-pbinom(n-81,n,teta)
> while(prob>0.80) {
+   n <- n+1
+   prob <- 1-pbinom(n-81,n,teta)
+ }
> c(n,prob)
[1] 83.0000000 0.7905045
```

Vemos então que com 83 bilhetes a probabilidade de que todos consigam assento cai para 0.79. Portanto para garantir que essa probabilidade seja maior que 0.80 é preciso vender 82 ou menos bilhetes.

Exercício 3.7.

Estamos procurando $P(X > 3) = 1 - P(X \leq 3)$ para uma variável que tem distribuição binomial:

```
> 1-pbinom(3,6,0.3)
[1] 0.07047
```

Exercício 3.8.

Identifica-se que a variável Y tem uma distribuição hipergeométrica. Portanto:

a.

```
> N <- 38; M <- 14; n <- 6
> dhyper(3,M,N-M,n)
[1] 0.2668675
```

b.

A esperança e desvio padrão (raiz quadrada da variância) são respectivamente:

```
> c(n*M/N,sqrt(n*M/N*(1-M/N)*(1-(n-1)/(N-1))))
[1] 2.210526 1.098843
```

Exercício 3.9.

Primeiramente definimos os objetos “número de golfinhos”, “número de golfinhos identificáveis” e “número de golfinhos fotografados”, respectivamente.

```
> N <- 85; M <- 35; n <- 10
```

Se a amostragem é realizada com reposição, estamos diante de uma distribuição binomial em que a probabilidade de que seja fotografado um golfinho com marca natural é sempre $\theta = M/N$:

```
> teta <- M/N
```

a.

Queremos a probabilidade $P(X > 5) = 1 - P(X \leq 5)$ que é:

```
> 1-pbinom(5,n,teta)
[1] 0.1865817
```

b.

A esperança e o desvio padrão são, respectivamente:

```
> c(n*teta,sqrt(n*teta*(1-teta)))
[1] 4.117647 1.556324
```

c.

No cenário em que a amostragem é sem reposição temos uma distribuição hipergeométrica ao invés de uma binomial. Portanto, os novos cálculos para os itens anteriores são, respetivamente:

```
> 1-phyper(5,M,N-M,n)
[1] 0.171982
> c(n*M/N,sqrt(n*M/N*(1-M/N)*(1-(n-1)/(N-1))))
[1] 4.117647 1.470588
```

Exercício 3.10.

Identifica-se que temos um problema que pode ser resolvido com a ditribuição binomial negativa, em que o sucesso é determinado pela captura de um peixe grande, enquanto o fracasso é a captura de um peixe pequeno. Definimos inicialmente a probabilidade de capturar um peixe grande, o número de peixes grandes necessários para encerrar a pescaria, e os custos de cada peixe grande e pequeno, respectivamente:

```
> teta <- 0.4; a <- 2; cg <- 6; cp <- 3
```

a.

A esperança (média), a variância e o desvio padrão para a variável U “número de peixes capturados” até que se obtenha a sucessos são, respectivamente:

```
> c(a/teta,a*(1-teta)/teta^2,sqrt(a*(1-teta)/teta^2))
[1] 5.000000 7.500000 2.738613
```

Já para X (fracassos) a variância e o desvio padrão são iguais aos de U , o que muda é a esperança:

```
> a*(1-teta)/teta
[1] 3
```

b.

No cenário médio esperamos que o número de fracassos seja 3 antes que se consiga obter 2 peixes grandes. Portanto o custo seria de R\$ 21.00:

```
> a*cg+a*(1-teta)/teta*cp
[1] 21
```

Exercício 3.11.

Temos um processo de poisson com média 0.005 que o número de intervalos (metros quadrados) é 212. Procuramos pela probabilidade $P(X \geq 3) = 1 - P(X < 3) = 1 - P(X \leq 2)$, que é:

```
> mu <- 0.005; k <- 212
> 1-ppois(2, mu*k)
[1] 0.09166216
```

Exercício 3.12.

a.

Precisamos da probabilidade $P(X \geq 4) = 1 - P(X < 4) = 1 - P(X \leq 3)$, que é:

```
> 1-ppois(3, 1.5)
[1] 0.06564245
```

b.

Precisamos encontrar k tal que $P(X \leq k) > 0.95$. Precisamos então do percentil 95% que pode ser obtido diretamente com:

```
> qpois(0.95, 1.5)
[1] 4
```

Exercício 3.13.

a.

Temos um processo de poisson com média 1.8 por 60 minutos e estamos obervando 20 minutos. Procuramos pela probabilidade $P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0)$ que é:

```
> mu <- 1.8; k <- 20/60
> 1-dpois(0, mu*k)
[1] 0.4511884
```

b.

Aqui precisamos da esperança em um cenário com $k = 75/60$:

```
> k <- 75/60
> mu*k
[1] 2.25
```

c.

Precisamos de um valor k tal que $P(X \geq 1) = 1 - P(X = 0) > 0.80$, ou seja,

$$\begin{aligned} 1 - \frac{\exp(-\mu k) \cdot (\mu k)^0}{0!} &> 0.80 \\ \exp(-\mu k) &< 0.2 \\ -\mu k &< \log_e(0.2) \\ k &> -\log_e(0.2)/\mu \end{aligned}$$

Portanto precisamos de um k (“número de intervalos de 60 minutos) qualquer maior que:

```
> -log(0.2)/mu
[1] 0.8941322
```

Em outras palavras precisamos fazer a amostragem por pelo menos

```
> -log(0.2)/mu*60
[1] 53.64793
```

minutos.

Exercício 3.14.

Temos um processo de poisson com média

```
> mu <- 6
```

a.

Queremos que a probabilidade $P(X > 8) = 1 - P(X \leq 7)$ que é:

```
> 1-ppois(7, 6)
[1] 0.2560202
```

b.

Em 98% dos dias chegam

```
> qpois(0.98, 6)
[1] 12
```

ou menos navios. Portanto, se podem ser atendidos pelo menos 12 navios por dia, há a probabilidade de em 98% dos casos todos os navios que aportam são inspecionados.

Exercício 3.15.

Temos uma distribuição binomial negativa em que a probabilidade de sucesso (θ) e o número de sucessos necessários (a) são, respectivamente:

```
> teta <- 0.3; a <- 3
```

a.

Denotamos como U o número de encontros necessários para se atingir a sucessos e X como o número de fracassos antes que essa meta seja atingida. Estamos então procurando pela probabilidade $P(U \leq 8) = P(X + a \leq 8) = P(X \leq 8 - a) = P(X \leq 5)$, que é:

```
> pnbinom(5,a,teta)
[1] 0.4482262
```

b.

A esperança para U é:

```
> a/teta
[1] 10
```

Exercício 3.16.

a.

Vamos considerar inicialmente que a variável Y (“o número de enchentes em um determinado ano”) segue uma distribuição de poisson com parâmetro μ . Este parâmetro pode ser calculado a partir da relação

$$P(Y \geq 1) = 1 - P(Y < 1) = 1 - \frac{e^{-\mu} \mu^0}{0!} = 0.01$$

Temos portanto que $\mu = -\log(0.99)$.

```
> mu <- -log(.99)
> mu
[1] 0.01005034
```

O tempo que transcorre entre ocorrências sucessivas de uma variável que segue a poisson, segue por sua vez uma distribuição exponencial. Portanto temos que $X \sim Exp(\mu)$, o consequentemente implica que $E(X) = 1/\mu$ e que $V(X) = 1/\mu^2$. Assim, a esperança e a variância para o tempo de ocorrência entre duas enchentes é:

```
> c(E.x=1/mu,V.x=1/mu^2)
      E.x          V.x
 99.49916 9900.08333
```

b.

Se definimos W como “tempo até a ocorrência de duas enchentes” vemos que $W = X_1 + X_2$ com X_1 sendo o “tempo até a ocorrência da primeira enchente” e X_2 sendo o “tempo entre a primeira e a segunda enchente”. Ambas as variáveis são exponenciais independentes. Logo $W \sim Gama(2, \mu)$.

Exercício 3.17.

Para simplificar a notação denominamos de A a proposição $X < 2$ ou $X > 4$, e B $0.5 < X < 5.5$. Estamos procurando pela probabilidade:

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}(B) = \frac{P(AB)}{P(B)}$$

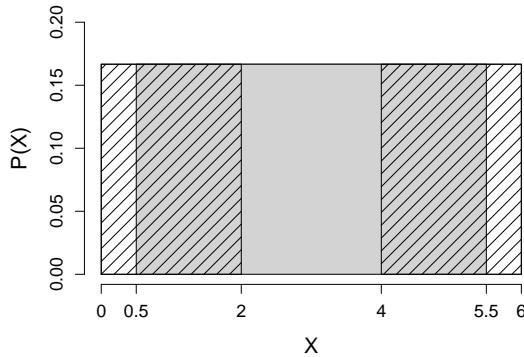


Figura 6:

O termo do denominador $P(B)$ equivale a área em cinza na figura 6.

A probabilidade $P(A)$ equivale a área marcada com linhas em paralelo, e a probabilidade $P(AB)$ (ocorrência conjunta de A e B) corresponde a área em que se sobrepõe o cinza e as linhas. Precisamos da razão entre as áreas que definem $P(AB)$ e $P(B)$. Lembramos que o cálculo de probabilidades para intervalos podem ser facilmente obtido ao considerar que $P(x_1 \leq X \geq x_2) = P(X \leq x_2) - P(X \leq x_1) = F(x_2) - F(x_1)$. Portanto a resposta é:

```
> ((punif(2,0,6)-punif(0.5,0,6))+
+ (punif(5.5,0,6)-punif(4,0,6)))/(punif(5.5,0,6)-punif(0.5,0,6))
[1] 0.6
```

Exercício 3.18.

a.

```
> beta <- 0.2
> pexp(2,beta)
[1] 0.3296800
```

b.

Procuramos por

$$P(X > 7 | X > 5) = \frac{P(X > 7)P(X > 5 | X > 7)}{P(X > 5)}$$

Como $P(X > 5 | X > 7) = 1$ temos que

$$P(X > 7 | X > 5) = \frac{P(X > 7)}{P(X > 5)} = \frac{1 - P(X \leq 7)}{1 - P(X \leq 5)}$$

Portanto a resposta é:

```
> (1-pexp(7,beta))/(1-pexp(5,beta))
[1] 0.67032
```

Exercício 3.19.

a.

Procuramos por $P(Z > 2.33) = 1 - P(Z \leq 2.33)$ e por $P(-1.46 < Z < 0.87) = P(X \leq 0.87) - P(X \leq -1.46)$, cujas respostas são respectivamente:

```
> 1-pnorm(2.33)
[1] 0.009903076
> pnorm(0.87)-pnorm(-1.46)
[1] 0.7357048
```

b.

```
> qnorm(c(0.25,0.75,0.8))
[1] -0.6744898 0.6744898 0.8416212
```

c.

```
> qnorm(c(0.08, 0.95))
[1] -1.405072  1.644854
```

Exercício 3.20.

Temos uma distribuição normal com média (μ) e desvio padrão (σ)

```
> mu <- 170; sigma <- 10
```

a.

Procuramos pelo produto $P(X > 160) \cdot 1000 = [1 - P(X \leq 160)] \cdot 1000$ que é:

```
> (1-pnorm(160, mu, sigma)) * 1000
[1] 841.3447
```

b.

Procuramos por um valor x^* tal que $P(\mu - x^* < X < \mu + x^*) = 0.75$. Isto implica que $P(X \leq \mu - x^*) = 0.25/2 = 0.125$. Portanto, o valor de $\mu - x^*$ é:

```
> qnorm(0.125, mu, sigma)
[1] 158.4965
```

Finalmente temos que $x^* = \mu - 158.4965 = 11.50349$.

Exercício 3.21.

Temos os seguintes valores de média (μ) e desvio padrão (σ):

```
> mu <- 178.1; sigma <- 40.7
```

a.

As distâncias que os valores

```
> x <- c(279, 150, 310, 57)
```

se encontram da média em termos de número de desvios padrões são:

```
> (x-mu)/sigma
[1] 2.4791155 -0.6904177 3.2407862 -2.9754300
```

Portanto podemos classificar os quatro valores, respectivamente, em: *Levemente Alterado, Satisfatório, Fortemente Alterado e Fortemente Alterado*.

b.

Queremos $P[(\mu - 2.5 \cdot \sigma < X < \mu - 1.5 \cdot \sigma) + (\mu + 1.5 \cdot \sigma < X < \mu + 2.5 \cdot \sigma)] = P(\mu - 2.5 \cdot \sigma < X < \mu - 1.5 \cdot \sigma) + P(\mu + 1.5 \cdot \sigma < X < \mu + 2.5 \cdot \sigma) = 2 \cdot [P(X \leq \mu - 1.5 \cdot \sigma) - P(X \leq \mu - 2.5 \cdot \sigma)]$, que é:

```
> 2 * (pnorm(mu-1.5*sigma, mu, sigma) - pnorm(mu-2.5*sigma, mu, sigma))
[1] 0.1211951
```

Exercício 3.22.

Temos uma população com média (μ) e desvio padrão (σ) de:

```
> mu <- 62; sigma <- 8
```

Se definimos X como o peso de uma pessoa, estamos interessados na probabilidade de que o somatório de k pessoas ultrapasse o limite de 450 kg, em que k pode assumir os valores:

```
> k <- c(5, 7, 9)
```

Queremos então calcular:

$$P\left(\sum_{i=1}^k x_i > 450\right) = P\left(\frac{\sum_{i=1}^k x_i}{k} > \frac{450}{k}\right) = P(\bar{X} > 450/k) = 1 - P(\bar{X} \leq 450/k)$$

Como a distribuição amostral para a média é normal $\bar{X} \sim N(\mu, \sigma/\sqrt{n})$, temos que para este caso em que $n = k$, a resposta é:

```
> 1-pnorm(450/k, mu, sigma/sqrt(k))
[1] 2.553513e-15 2.248459e-01 9.999966e-01
```

Vemos então que o elevador está mal dimensionado pois a probabilidade de que o limite de peso seja ultrapassado com sete passageiros (lotação máxima recomendada) é relativamente elevada (0.225).

Exercício 3.23.

A variável “tempo entre ocorrências sucessivas” de eventos que se distribuem segundo uma poisson com média (μ), segue uma distribuição exponencial com parâmetro $\beta = \mu$, que neste caso é:

```
> beta <- 4.16
```

a.

Estamos interessados na probabilidade $P(X > 2) = 1 - P(X \leq 2)$:

```
> 1-pexp(2,beta)
[1] 0.0002435959
```

b.

A esperança e a raiz quadrada da variância de uma variável com distribuição exponencial são, respectivamente:

```
> c(1/beta,sqrt(1/beta^2))
[1] 0.2403846 0.2403846
```

c.

O valor de (α) da distribuição gama é:

```
> alfa <- 20
```

Já a esperança e o desvio padrão para essa distribuição são, respectivamente:

```
> c(alfa/beta,sqrt(alfa/beta^2))
[1] 4.807692 1.075033
```

d.

Finalmente, queremos o percentil de 95% para a variável “tempo decorrido entre duas paradas”:

```
> qgamma(0.95,alfa,beta)
[1] 6.70174
```

Vemos que em média deverá passar no máximo cerca de 7 minutos entre uma e outra parada.

Exercício 3.24.

Primeiramente definimos os objetos necessários para o desenvolvimento dos cálculos.

```
> n <- 11; k <- 6
> alfa <- k; beta <- n-k+1
```

a.

A densidade de probabilidade teórica para a mediana é $\tilde{X} \sim Beta(\alpha = 6, \beta = 11 - 6 + 1 = 6)$. Utilizando as propriedades de distribuições beta, temos que a esperança e a variância são, respectivamente:

```
> c(alfa/(alfa+beta), alfa*beta/((alfa+beta)^2*(alfa+beta+1)))
[1] 0.50000000 0.01923077
```

b.

Segue abaixo uma simulação de 1000 amostras de tamanho $n = 11$ de uma distribuição uniforme $U(0, 1)$:

```
> amostras <- numeric()
> for(i in 1:1000) amostras <- rbind(amostras, runif(n, 0, 1))
```

A título de ilustração mostramos as três primeiras amostras obtidas:

```
> amostras[1:3,]
     [,1]      [,2]      [,3]      [,4]      [,5]      [,6]      [,7]
[1,] 0.2937354 0.5644759 0.8244650 0.1205136 0.4580158 0.58175173 0.2931685
[2,] 0.6503456 0.4909045 0.1641856 0.4259940 0.6141929 0.72793253 0.5832343
[3,] 0.6907260 0.5391746 0.5712495 0.8080558 0.4163163 0.09242827 0.6500682
     [,8]      [,9]      [,10]     [,11]
[1,] 0.2110506 0.04333835 0.81864370 0.8314228
[2,] 0.1970378 0.56673374 0.17833821 0.5530820
[3,] 0.7291020 0.76764775 0.07977433 0.2732723
```

As medianas para essas amostras podem ser obtidas com:

```
> medianas <- apply(amostras, 1, median)
```

As três primeiras medianas são:

```
> medianas[1:3]
[1] 0.4580158 0.5530820 0.5712495
```

Finalmente temos que a esperança (média) e a variância destas medianas são, respectivamente:

```
> c(mean(medianas), var(medianas))
[1] 0.50453325 0.01878659
```

Note que estes valores são similares aos anteriores obtidos com as propriedades da distribuição $Beta(\alpha = 6, \beta = 6)$.

Capítulo 4

Exercício 4.1.

Definimos inicialmente os objetos referentes à “primeira aferição”, “segunda aferição”, “tamanho da amostra” e “incidência de grávidas no grupo RU 486” como:

```
> r <- 0.45; rm <- 0.55; n <- 4; x <- 2
```

Os parâmetros da priori, α e β , são, respectivamente:

```
> alfa <- r*(1-rm)/(rm-r)
> beta <- (1-r)*(1-rm)/(rm-r)
> c(alfa,beta)
[1] 2.025 2.475
```

Como temos uma distribuição $Beta(\alpha = 2.025, \beta = 2.475)$ como priori, a distribuição posterior será $Beta(\alpha^* = \alpha + x = 4.025, \beta^* = \beta + n - x = 4.475)$.

Exercício 4.2.

Neste novo cenário temos que o tamanho amostral, a média e desvio padrão dos pesos amostrais são, respectivamente:

```
> n <- 10; x.bar <- 30; s <- 10
```

Como $n \leq 30$ utilizaremos o desvio padrão “corrigido”:

```
> Sc <- s*(1+20/n^2)
> Sc
[1] 12
```

Com a priori uniforme $p(\mu) \sim U(0, 100)$ e a verossimilhança normal $p(\bar{X}|\mu) \sim N(\mu, Sc/\sqrt{n})$, a distribuição posterior é a normal $p(\mu|\bar{X}) \sim N(\bar{X} = 30, Sc/\sqrt{n} = 12/\sqrt{10} \cong 3.795)$. Já com a priori normal $p(\mu) \sim N(\mu_0, \sigma_0)$ e verossimilhança normal $p(\bar{X}|\mu) \sim N(\mu, Sc/\sqrt{n})$, a distribuição posterior $p(\mu|\bar{X}) \sim N(\mu_1, \sigma_1)$. Para obter os parâmetros μ_1 e σ_1 da posterior, começamos pela definição dos parâmetros da priori, μ_0 e σ_0 , que são:

```
> mu0 <- 45; sigma0 <- 11.7
```

Posteriormente calculamos respectivamente as precisões priori, amostral e posterior, da seguinte forma:

```
> C0 <- 1/sigma0^2; C <- n/Sc^2; C1 <- C0+C
```

Finalmente, obtemos os valores de μ_1 e σ_1 , que são respectivamente:

```
> mu1 <- C0/C1*mu0+C/C1*x.bar
> sigma1 <- 1/sqrt(C1)
> c(mu1,sigma1)
[1] 31.427722 3.609624
```

Portanto, a distribuição posterior é $P(\mu|\bar{X}) \sim N(\mu_1 \cong 31.428, \sigma_1 \cong 3.610)$.

Exercício 4.3.

A título de exemplo vamos considerar inicialmente $\mu_0 = 1$ ano e $\sigma_0 = 0.08$ anos. Para aferir podemos utilizar a idéia de que os percentis de 10% e 90% devem ser de cerca de 0.80 e 1.20 anos. Agora verificamos se a primeira aferição de σ_0 satisfaz a essa demanda.

```
> qnorm(c(0.1, 0.9), 1, 0.08)
[1] 0.8974759 1.1025241
```

Vemos que o desvio padrão deve ser um pouco maior e podemos fazer mais algumas tentativas, até que se obtenha valores satisfatórios. Por exemplo:

```
> qnorm(c(0.1, 0.9), 1, 0.1)
[1] 0.8718448 1.1281552
> qnorm(c(0.1, 0.9), 1, 0.15)
[1] 0.8077673 1.1922327
```

Poderíamos nos dar então por satisfeitos já que a aferição $\sigma_0 = 0.15$ anos propicia que sejam obtidos percentis muito próximos dos valores pretendidos, 0.80 e 1.20 anos.

Exercício 4.4.

Começamos inicialmente definindo os objetos que correspondem, para o Alto Vale, ao tamanho da amostra (n_a), à média (\bar{X}_a) e ao desvio padrão amostral (S_a):

```
> na <- 5; Sa <- 3.32; Xa.bar <- 17.32
```

Para o Baixo Vale o tamanho da amostra (n_b), a média (\bar{X}_b) e o desvio padrão amostral (S_b) são, respectivamente:

```
> nb <- 14; Sb <- 4.96; Xb.bar <- 12.56
```

a.

Vamos utilizar como prioris para a média no Alto (μ_a) e no Baixo Vale (μ_b) distribuições uniforme definida no intervalo $[0, 30]$, ou seja, $p(\mu_a) = p(\mu_b) = U(0, 30)$. Este intervalo é amplo o suficiente para que os truncamentos tenham efeito irrelevante.

Se μ tem priori uniforme e se a verossimilhança é normal $p(\bar{X}|\mu) \sim N(\mu, S)$, então a distribuição posterior é normal $p(\mu|\bar{X}) \sim N(\bar{X}, S/\sqrt{n})$. Antes de aplicar essa solução para os casos do Alto e do Baixo Vale, é importante notar que os tamanhos amostrais são relativamente pequenos (≤ 30). Portanto os desvios padrões amostrais (S) da solução acima devem ser “corrigido” para $S_c = S(1 + 20/n^2)$. Os desvios padrões corrigidos para o Alto (S_{ca}) e o Baixo Vale (S_{cb}) são, respectivamente:

```
> Sca <- Sa * (1 + 20/na^2)
> Scb <- Sb * (1 + 20/nb^2)
> c(Sca, Scb)
[1] 5.976000 5.466122
```

Agora podemos aplicar a solução mencionada anteriormente e obtemos que as distribuições posteriores para o Alto e o Baixo Vale são $p(\mu_a|\bar{X}_a) \sim N(\mu_a^* = \bar{X}_a, \sigma_a^* = S_{ca}/\sqrt{n_a})$ e $p(\mu_b|\bar{X}_b) \sim N(\mu_b^* = \bar{X}_b, \sigma_b^* = S_{cb}/\sqrt{n_b})$, respectivamente. Os parâmetros dessas distribuições posteriores, $\mu_a^*, \sigma_a^*, \mu_b^*$ e σ_b^* são, respectivamente:

```
> mua.p <- Xa.bar
> sigmaa.p <- Sca/sqrt(na)
> mub.p <- Xb.bar
> sigmab.p <- Scb/sqrt(nb)
> c(mua.p, sigmaa.p, mub.p, sigmab.p)
[1] 17.320000 2.672548 12.560000 1.460883
```

Portanto as posteriores são $p(\mu_a|\bar{X}_a) \sim N(\mu_a^* = 17.32, \sigma_a^* \cong 2.67)$ e $p(\mu_b|\bar{X}_b) \sim N(\mu_b^* = 12.56, \sigma_b^* \cong 1.46)$. Um gráfico com essas distribuições é mostrado na figura 7.

b.

Se $\delta = \mu_a - \mu_b$, $\mu_a \sim N(\mu_a^*, \sigma_a^*)$ e $\mu_b \sim N(\mu_b^*, \sigma_b^*)$ então a diferença segue a uma normal $\delta \sim N(\mu_\delta, \sigma_\delta)$, em que os parâmetros são $\mu_\delta = \mu_a^* - \mu_b^*$ e $\sigma_\delta = \sqrt{(\sigma_a^*)^2 + (\sigma_b^*)^2}$. Os valores de μ_δ e σ_δ são, respectivamente:

```
> mu.delta <- mua.p - mub.p
> sigma.delta <- sqrt(sigmaa.p^2 + sigmab.p^2)
> c(mu.delta, sigma.delta)
[1] 4.760000 3.045766
```

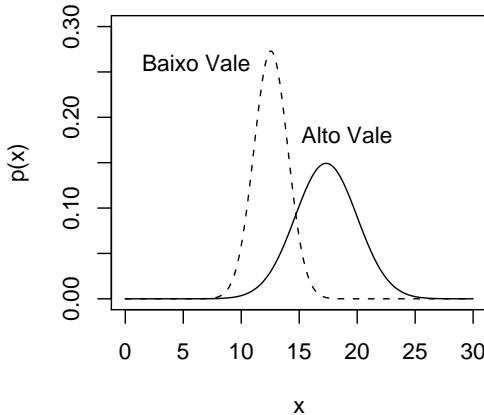


Figura 7:

Procuramos agora pela probabilide $P(\delta > 0) = 1 - P(\delta \leq 0)$ que é:

```
> 1-pnorm(0,mu.delta,sigma.delta)
[1] 0.9409531
```

Fica então claro que é muito provável que a diferença entre as concentrações médias seja maior que zero. Portanto há fortes evidências em favor da concepção de que a concentração média no Alto Vale é maior.

Exercício 4.5.

A projeção inicial do especialista para θ é $r = 1/(3\text{min.})$. No experimento mental, após ser informado de que em 15 minutos ($n = 5$ períodos de 3 min.) houve $x = 20$ atividades de pouso ou decolagem, o especialista atualizou a sua projeção de θ para $rp = 2/(3\text{min.})$. Definimos então, respectivamente, os objetos para r , rp , x e n :

```
> r <- 1; rp <- 2; x <- 20; n <- 5
```

a.

Podemos agora calcular facilmente os parâmetros de uma priori $p(\theta) \sim Gama(\alpha, \beta)$, que são, respectivamente:

```
> alfa <- (x-n*rp)*r/(rp-r)
> beta <- (x-n*rp)/(rp-r)
> c(alfa,beta)
[1] 10 10
```

b.

Temos agora informações de que em 30 minutos ($n = 10$ períodos de 3 min.) foram obvervadas $x = 20$ atividades de decolagem ou aterrissagem. Se a priori é $p(\theta) \sim Gama(\alpha, \beta)$ e a verossimilhança é $p(X|\theta) \sim Poi(\mu = n\theta)$, resulta que a distribuição posterior é $p(\theta|X) \sim Gama(\alpha^* = \alpha + x = 10 + 21 = 31, \beta^* = \beta + n = 10 + 10 = 20)$.

Consta na figura 8 um gráfico com os parâmetros das distribuições priori ($\alpha_0 = 2$ e $\beta_0 = 100$) e posterior ($\alpha_1 = 12$ e $\beta_1 = 1254.54$) para o problema do ganho de peso dos leitões. Segue abaixo o código para gerar a figura.

```
tau <- seq(0.001,0.08,by=0.0003)
sigma <- 1/sqrt(tau)
alfa1 <- 12; beta1 <- 1254.54
densidade1 <- dgamma(tau,alfa1,beta1)*(2/sigma^3)
alfa0 <- 2; beta0 <- 100
densidade0 <- dgamma(tau,alfa0,beta0)*(2/sigma^3)
plot(sigma,densidade1,xlab=expression(sigma),
      ylab="Densidade de Probabilidade",type="l")
lines(sigma,densidade0,lty=2)
text(c(5,15),c(0.18,0.23),c("priori","posterior"))
```

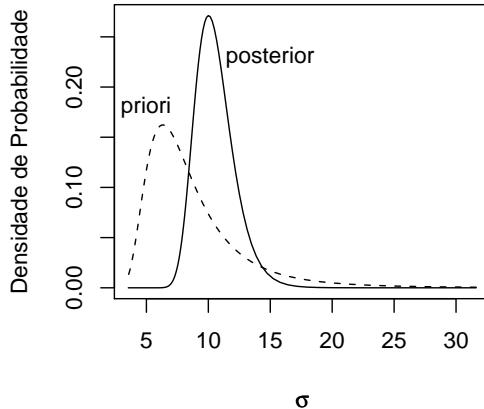


Figura 8:

Exercício 4.7.

Primeiramente definimos os elementos g_1, μ_1, σ_1

```
> g1 <- 13; mu1 <- 25; sigma1 <- 5
```

e uma seqüência de valores T

```
> T <- seq(-4, 4, 0.1)
```

A seguir definimos a função inversa que transforma T em μ :

```
> mu <- T*sigma1+mu1
```

Agora resta simplesmente aplicar a solução dada no enunciado. Segue um código com aplicação e o gráfico resultante (Figura 9).

```
> densidade <- dt(T,g1)/sigma1
> plot(mu,densidade,xlab=expression(mu),
+       ylab="Densidade de Probabilidade",type="l")
```

Exercício 4.8.

a.

Primeiro definimos o hiper-parâmetro μ_0 e os objetos y_1 e y_2 a partir da informação disponível sobre a média esperada e das observações sobre valores extremos:

```
> mu0 <- 32; y1 <- 25; y2 <- 40
```

A seguir calculamos a variância global (V):

```
> V <- ((y2-y1)/6)^2
> V
[1] 6.25
```

Tomaremos a título de exemplo $p = 0.4$ que implica que a maioria da variância global decorre da variância intrínseca da variável V_X , e uma parte menor está associada à incerteza sobre a média μ , denotada pela variância V_μ . Agora calculamos os valores de V_X e V_μ que são, respectivamente:

```
> p <- 0.4
> VX <- (1-p)*V
> Vmu <- p*V
> c(VX,Vmu)
[1] 3.75 2.50
```

A seguir fixamos o hiper-parâmetro α_0 em um valor pequeno para denotar uma priori “mente aberta”.

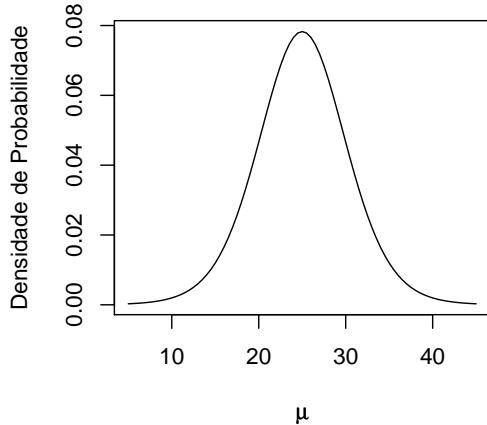


Figura 9:

```
> alfa0 <- 1.5
```

Agora os valores dos hiper-parâmetros β_0 e λ_0 podem ser obtido com:

```
> beta0 <- alfa0*VX
> lambda0 <- beta0/((alfa0-1)*Vm)
```

As soluções para os quatro hiper-parâmetros são:

```
> c(mu0,lambda0,alfa0,beta0)
[1] 32.000 4.500 1.500 5.625
```

Podemos então escrever a prioris $p(\tau) \sim Gama(\alpha_0 = 1.5, \beta_0 = 5.625)$ e $p(\mu|\tau) \sim N^*(\mu_0 = 32, \lambda_0\tau = 4.5 \cdot \tau)$.

É importante lembrar que $N^*(\cdot)$ indica uma distribuição normal parametrizada em função da precisão. b.

O tamanho da amostral (n), a média (\bar{X}) e o desvio padrão amostral (S) são:

```
> n <- 18; x.bar <- 31.5; S <- 1.8
```

Com base nestes dados e nos hiper-parâmetros da priori normal-gama especificada acima, podemos obter os parâmetros da posterior para a precisão $p(\tau|\bar{X}, S^2) \sim Gama(\alpha_1, \beta_1)$ e da posterior marginal para a média $p(\mu|\tau, \bar{X}, S^2) \sim N^*(\mu_1, \lambda_1\tau)$. São eles:

```
> mu1 <- (lambda0*mu0+n*x.bar)/(lambda0+n)
> lambda1 <- lambda0+n
> alfa1 <- alfa0+n/2
> beta1 <- beta0+(n-1)*S^2/2+n*lambda0*(x.bar-mu0)^2/(2*(lambda0+n))
> c(mu1,lambda1,alfa1,beta1)
[1] 31.600 22.500 10.500 33.615
```

Portanto temos que $p(\tau|\bar{X}, S^2) \sim Gama(\alpha_1 = 10.5, \beta_1 = 33.615)$ e que $p(\mu|\tau, \bar{X}, S^2) \sim N^*(\mu_1 = 31.6, \lambda_1\tau = 22.5 \cdot \tau)$.

c. Os percentis para a precisão τ são:

```
> qgamma(c(0.025,0.5,0.975),alfa1,beta1)
[1] 0.1529510 0.3025023 0.5277239
```

A posterior para a média μ é condicionada a τ , portanto, vamos obter os valores dos percentis por simulação. Primeiro amostramos uma grande quantidade (ex: 5000) de valores da posterior para τ e, condicionado a estes valores, obtemos também uma grande amostra a partir da posterior marginal de μ . Finalmente, calculamos os percentis procurados a partir dessa amostra. Ao operacionalizar a solução é importante lembrar que a posterior $p(\mu|\tau, \bar{X}, S^2) \sim N^*(\mu_1, \lambda_1\tau)$ está parametrizada em função da precisão que é $1/\sigma^2$. As funções do R relativas à distribuição normal estão parametrizadas em relação a $\sigma = 1/\sqrt{\tau}$. Portanto precisaremos fazer os cálculos da posterior com a seguinte parametrização $p(\mu|\tau, \bar{X}, S^2) \sim N(\mu_1, \sigma_1 = 1/\sqrt{\lambda_1\tau})$. Segue abaixo o código com a solução:

```

> amostra.tau <- rgamma(5000, alfa1, betal)
> sigma1 <- 1/sqrt(lambdal*amostra.tau)
> amostra.mu <- rnorm(5000, mu1, sigma1)
> quantile(amostra.mu, c(0.025, 0.5, 0.975))
  2.5%      50%    97.5%
30.83436 31.59781 32.37874

```

Uma outra opção seria amostrar da posterior marginal para μ condicionado unicamente ao valor da esperança de τ que é:

```

> esp.tau <- alfa1/betal
> esp.tau
[1] 0.3123606

```

Assim os percentis calculados para μ são:

```

> qnorm(c(0.025, 0.5, 0.975), mu1, 1/sqrt(lambdal*esp.tau))
[1] 30.86069 31.60000 32.33931

```

Mas, neste caso, parte da variabilidade de μ se perde em virtude de que a variabilidade de τ foi de certa forma desconsiderada. Isto pode ser notado a partir da comparação dos percentis de 2.5% e 97.5% obtidos com a simulação considerando a variabilidade de τ e fixando τ ao valor da sua esperança, que são, neste último caso, denotam uma menor amplitude.

Segue abaixo os gráficos das posteriores de μ e τ (Figura 10) e o código utilizado para a construção dos mesmos. No caso de μ , a título de ilustração, mostramos as posteriores obtidas com simulação (linha tracejada) e com a condicional fixada ao valor da esperança de τ (linha contínua).

```

mu <- seq(29, 34, 0.1)
tau <- seq(0, 0.7, 0.01)
par(mfrow=c(1, 2))
plot(tau,dgamma(tau, alfa1, betal),type="l",xlab=expression(tau) )
plot(mu,dnorm(mu,mu1,1/sqrt(lambdal*esp.tau)),type="l",
      ylab="Densidade",xlab=expression(mu),ylim=c(0,1.2))
lines(density(amostra.mu),lty=2)
legend(29,1,c("simulação","esperança"),lty=c(2,1))

```

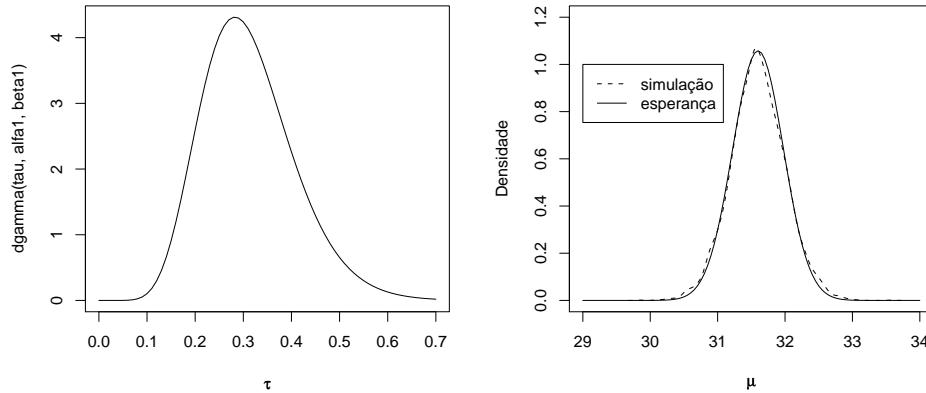


Figura 10:

d.

Aqui vamos considerar que não é necessário modelar a incerteza sobre a variância, que é $\sigma^2 = 1/\tau$. Neste cenário se a priori $p(\mu)$ é uniforme e a verossimilhança é dada pela normal $p(\bar{X}|\mu) \sim N(\mu, S/\sqrt{n})$, a posterior é simplesmente $p(\mu|\bar{X}) \sim N(\bar{X}, S/\sqrt{n})$. A única ressalva é que como o tamanho amostral é relativamente reduzido ($n = 18$), devemos usar $S_c = S(1 + 20/n^2)$ no lugar de S . Segue abaixo as soluções para S_c e para o desvio padrão da posterior S_c/\sqrt{n} .

```

> Sc <- S*(1-20/n^2)
> Sc/sqrt(n)
[1] 0.3980749

```

Portanto temos que $p(\mu|\bar{X}) \sim N(\bar{X} = 31.5, Sc/\sqrt{n} \cong 0.398)$.

e.

Os percentis são:

```
> qnorm(c(0.025,0.5,0.975),x.bar,Sc/sqrt(n))
[1] 30.71979 31.50000 32.28021
```

Segue abaixo o gráfico da distribuição posterior (Figura 11) e o código utilizado para a sua construção.

```
> par(mfrow=c(1,1))
> mu <- seq(29,34,0.1)
> plot(mu,dnorm(mu,x.bar,Sc/sqrt(n)),type="l",xlab=expression(mu))
```

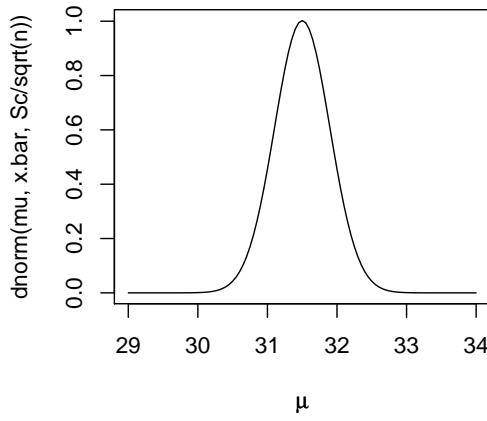


Figura 11:

Exercício 4.9.

Temos um processo de poisson com taxa θ de observações por intervalo de 10 milhas náuticas (mn). Observamos $x = 10$ ocorrências em 150 mn ($n = 15$ intervalos de 10 mn).

```
> x <- 10; n <- 15
```

Queremos obter a distribuição posterior $p(\theta|X)$. Resolveremos o problema utilizando uma priori gama $p(\theta) \sim Gama(\alpha, \beta)$ que é conjugada da poisson $p(X|\theta) \sim Poi(\mu = n \cdot \theta)$. Assim a distribuição posterior será $p(\theta|X) \sim Gama(\alpha^* = \alpha + x, \beta^* = \beta + n)$.

a.

A título de exemplo vamos utilizar os seguintes valores para α e β , respectivamente:

```
> alfa <- 1; beta <- 0.1
```

Como a variância em uma gama é α/β^2 os valores acima implicam em uma variância bastante elevada $1/(0.1)^2 = 100$, o que denota uma priori não-informativa. Assim os parâmetros α^* e β^* da posterior são, respectivamente:

```
> alfa.p <- alfa+x; beta.p <- beta+n
> c(alfa.p,beta.p)
[1] 11.0 15.1
```

Segue abaixo um gráfico com as distribuições priori $p(\theta) \sim Gama(\alpha = 1, \beta = 0.1)$ e posterior $p(\theta|X) \sim Gama(\alpha^* = 11, \beta^* = 15.1)$ (Figura 12), e o código usado para a sua construção.

```
> teta <- seq(0,2,0.01)
> plot(teta,dgamma(teta,alfa.p,beta.p),type="l",xlab=expression(theta))
> lines(teta,dgamma(teta,1,0.1),lty=2)
> text(c(1.7,1.1),c(0.2,1.6),c("priori","posterior"))
```

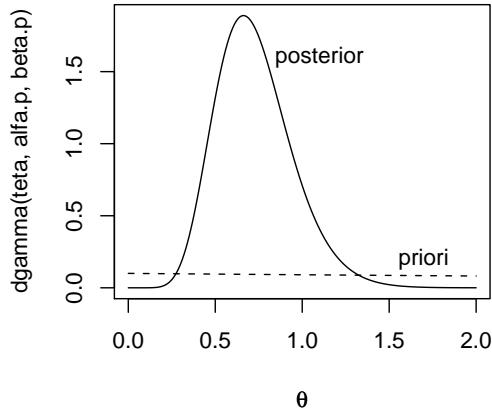


Figura 12:

b.

Pretendemos usar uma priori baseada na concepção de que a esperança para θ é 0.45 avistagens/1 intervalo de 10 mn. Se igualarmos esse valor à esperança de uma distribuição $Gama(\alpha, \beta)$ temos que $\alpha/\beta = 0.45$. Como pretendemos que a priori seja informativa, o valor de α deve ser relativamente elevado, por exemplo $\alpha = 4.5$. Portanto temos que $\beta = 4.5/0.45 = 10$. Como no item anterior, a distribuição posterior será $p(\theta) \sim Gama(\alpha^* = \alpha + x, \beta^* = \beta + n)$. Segue abaixo o cálculo dos parâmetros desta distribuição posterior

```
> alfa <- 4.5; beta <- 10
> alfa.p <- alfa+x; beta.p <- beta+n
> c(alfa.p,beta.p)
[1] 14.5 25.0
```

e também um gráfico com as distribuições priori $p(\theta) \sim Gama(\alpha = 4.5, \beta = 10)$ e posterior $p(\theta|X) \sim Gama(\alpha^* = 14.5, \beta^* = 25)$ (Figura 13), e o código usado para a sua construção.

```
> teta <- seq(0,2,0.01)
> plot(teta,dgamma(teta,alfa.p,beta.p),type="l",xlab=expression(theta))
> lines(teta,dgamma(teta,alfa,beta),lty=2)
> text(c(0.1,1),c(2.2,2.5),c("priori","posterior"))
```

Capítulo 5

Exercício 5.1.

Os valores de α e β das prioris $p(\theta_A) = p(\theta_B) = Beta(\alpha, \beta)$ são, respectivamente:

```
> alfa <- 1.5; beta <- 1.5
```

Os tamanhos amostrais para as áreas A e B (n_A e n_B), e o número de fêmeas obtido nestas áreas (x_A e x_B) são, respectivamente:

```
> nA <- 396; nB <- 460; xA <- 195; xB <- 198
```

a.

Se a priori é $p(\theta) \sim Beta(\alpha, \beta)$ e a verossimilhança é $p(X|\theta) \sim Bin(n, \theta)$ então a distribuição posterior é $p(\theta|X) \sim Beta(\alpha^* = \alpha + x, \beta^* = \beta + n - x)$. Portanto os parâmetros α_A^* e β_A^* da distribuição posterior para a área A são:

```
> alfa.A <- alfa+xA; beta.A <- beta+nA-xA
> c(alfa.A,beta.A)
[1] 196.5 202.5
```

Similarmente, os parâmetros α_B^* e β_B^* da distribuição posterior para a área B são:

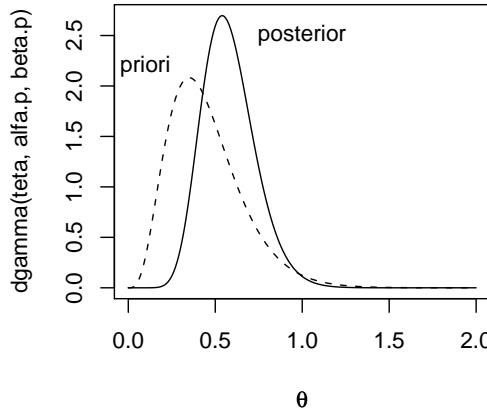


Figura 13:

```
> alfa.B <- alfa+xB; beta.B <- beta+nB-xB
> c(alfa.B,beta.B)
[1] 199.5 263.5
```

Queremos agora a probabilidade $p(\theta > 0.5) = 1 - p(\theta \leq 0.5)$ para as áreas A e B , que são, respectivamente:

```
> c(1-pbeta(0.5,alfa.A,beta.A),1-pbeta(0.5,alfa.B,beta.B))
[1] 0.381823419 0.001432782
```

Vemos então que a probabilidade de que a maioria (mais de 50%) dos espécimes seja fêmea é bem maior para a área A $P(\theta_A > 0.5) \cong 0.38$. Portanto, seguindo o critério proposto, essa deve ser a área escolhida para se proibir a captura no período reprodutivo.

b.

Começamos com a simulação de amostras das posteriores $p(\theta_A|X)$ e $p(\theta_B|X)$ (etapa i):

```
> amostra.A <- rbeta(2500,alfa.A,beta.A)
> amostra.B <- rbeta(2500,alfa.B,beta.B)
```

Calculamos então as diferenças $\delta = \theta_A - \theta_B$ a partir desta amostras (etapa ii):

```
> delta <- amostra.A-amostra.B
```

Segue o histograma (Figura 14) (etapa iii) construído com o código:

```
> hist(delta,xlab=expression(delta))
```

Finalmente, como a amostra da distribuição posterior é relativamente grande, pelo princípio da integração de Monte Carlo, pode-se obter uma aproximação satisfatória de $P(\delta > 0)$ a partir da amostra. Portanto, essa probabilidade é aproximadamente igual à razão (etapa iv):

```
> sum(delta>0)/length(delta)
[1] 0.9532
```

Vemos então que é bastante provável que a proporção de fêmeas na área A seja maior do que na área B .

Exercício 5.2.

O número de fêmeas obtidas nas áreas C e D (x_C e x_D), e os tamanhos amostrais para estas áreas (n_C e n_D) são, respectivamente:

```
> xC <- 200; xD <- 206; nC <- xc+186; nD <- xd+219
```

Os valores de α e β das prioris $p(\theta_C) = p(\theta_D) = Beta(\alpha, \beta)$ são, respectivamente:

```
> alfa <- 1.5; beta <- 1.5
```

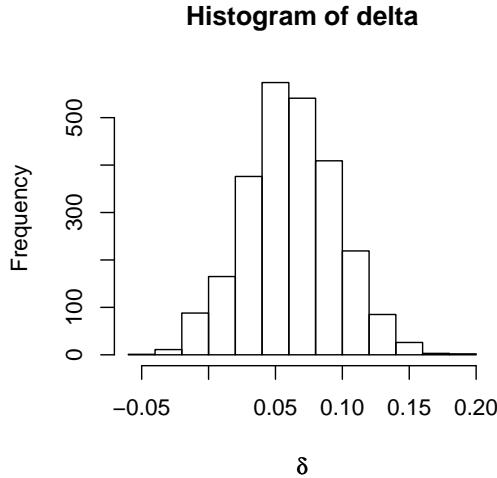


Figura 14:

a.

Da mesma forma que no primeiro item do exercício anterior, temos que se a priori é $p(\theta) \sim Beta(\alpha, \beta)$ e a verossimilhança é $p(X|\theta) \sim Bin(n, \theta)$ então a distribuição posterior é $p(\theta|X) \sim Beta(\alpha^* = \alpha + x, \beta^* = \beta + n - x)$. Portanto os parâmetros α_C^* e β_D^* da distribuição posterior para a área C são:

```
> alfa.C <- alfa+xC; beta.C <- beta+nC-xC
> c(alfa.C,beta.C)
[1] 201.5 187.5
```

Similarmente, os parâmetros α_D^* e β_D^* da distribuição posterior para a área D são:

```
> alfa.D <- alfa+xD; beta.D <- beta+nD-xD
> c(alfa.D,beta.D)
[1] 207.5 220.5
```

b.

Segue o gráfico com as distribuições posteriores sobrepostas (Figura 15) e o código usado para a sua construção.

```
> teta <- seq(0.3,0.7,0.001)
> plot(teta,dbeta(teta,alfa.A,beta.A),type="l",xlab=expression(theta),
+       ylab=expression(p(paste(theta,"|",X))),ylim=c(0,20))
> lines(teta,dbeta(teta,alfa.B,beta.B),lty=2)
> lines(teta,dbeta(teta,alfa.C,beta.C),lty=3)
> lines(teta,dbeta(teta,alfa.D,beta.D),lty=4)
> legend(0.58,19,lty=c(1:4),c("A","B","C","D"))
```

c.

As amostras das posteriores para as áreas A e B (`amostra.A` e `amostra.B`) já foram obtidas na resolução do exercício anterior. Resta obter as amostras de $p(\theta_C|X)$ e $p(\theta_D|X)$ (etapa i), que são:

```
> amostra.C <- rbeta(2500,alfa.C,beta.C)
> amostra.D <- rbeta(2500,alfa.D,beta.D)
```

Em cada uma das 2500 amostras pareadas temos que identificar aquela que corresponde ao valor máximo. Para isto primeiro montamos uma tabela com quatro colunas, cada uma delas contendo uma amostra.

```
> tabela <- cbind(amostra.A,amostra.B,amostra.C,amostra.D)
```

A título de exemplo, vejamos as três primeiras linhas dessa tabela:

```
> tabela[1:3,]
  amostra.A amostra.B amostra.C amostra.D
[1,] 0.4793591 0.4122973 0.5428833 0.5023401
[2,] 0.5045010 0.4422018 0.4744282 0.5449193
[3,] 0.4965926 0.4483586 0.4835623 0.5142590
```

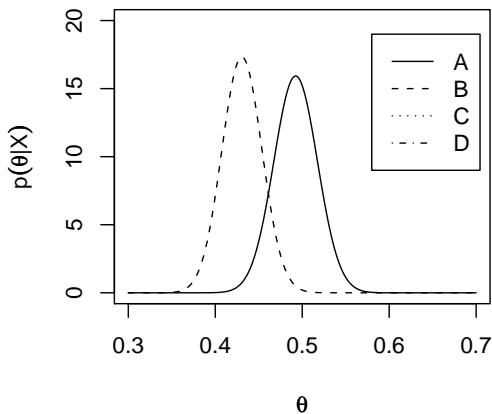


Figura 15:

Vemos então que a primeira linha da tabela, que corresponde à primeira amostra pareada, indica que a área C é a “vencedora” pois ela corresponde ao maior valor de θ .

Agora definimos um vetor que contém a identificação da área para cada uma das quatro colunas da tabela.

```
> areas <- c("A", "B", "C", "D")
```

Para uma linha particular da tabela (por exemplo a linha 1 podemos facilmente identificar a área que corresponde ao valor máximo com a seguinte função:

```
> areas[tabela[1,]==max(tabela[1,])]  
[1] "C"
```

No entanto queremos aplicar essa função em toda a tabela, linha a linha. Para temos que reescrever a função acima porém de maneira genérica, pronta para receber qualquer linha com quatro valores (vetor x) e não somente a linha 1 do caso específico mostrado acima. Criamos então a função `fn`:

```
> fn <- function(x) areas[x==max(x)]
```

Agora resta aplicar esta função ao longo de toda a tabela, linha a linha. Vamos fazer isto e armazenar o resultado em um novo vetor com:

```
> max.teta <- apply(tabela, 1, fn)
```

O argumento 1 na função `apply()` mostrada acima indica que a aplicação será linha a linha. Se estivéssemos em face de um problema que precisaria ser resolvido coluna a coluna, o argumento deveria ser 2.

O vetor `max.teta` contém para cada uma das 2500 linhas da tabela com as amostras pareadas, a indicação de qual é a área que corresponde ao valor máximo de θ . Por exemplo, vemos abaixo que as áreas C e D foram vencedoras, cada uma, em duas das primeiras quatro simulações:

```
> max.teta[1:4]  
[1] "C" "D" "D" "C"
```

Para verificar em quantas simulações cada uma das quatro áreas foi a vencedora (etapa ii) podemos realizar uma contagem com o auxílio da função `table`:

```
> table(max.teta)  
max.teta  
A     B     C     D  
510    5 1673  312
```

Finalmente, podemos então avaliar rapidamente as freqüências relativas (etapa iii) com:

```
> table(max.teta)/length(max.teta)  
max.teta  
A     B     C     D  
0.2040 0.0020 0.6692 0.1248
```

Há portanto uma probabilidade posterior de 0.669 de que a proporção de fêmeas seja maior na área C .

Exercício 5.3.

As informações sobre o número de períodos de 12 horas amostrados em cada área (n_A e n_B) e as respectivas quantidades de peixes capturados (x_A e x_B) são:

```
> nA <- 10; nB <- 10; xA <- 8; xB <- 5
```

Temos que a priori não-informativa (ni) é comum às duas áreas $p(\theta_A) = p(\theta_B) \sim Gama(\alpha_{ni} = 1, \beta_{ni} = 0.3)$. Os objetos no R com os parâmetros destas prioris são:

```
> alfa.ni <- 1; beta.ni <- 0.3
```

Já as prioris informativas (i) são $p(\theta_A) \sim Gama(\alpha_i^{(A)}, \beta_i^{(A)})$ e $p(\theta_B) \sim Gama(\alpha_i^{(B)}, \beta_i^{(B)})$. Os quatro parâmetros ($\alpha_i^{(A)}, \beta_i^{(A)}, \alpha_i^{(B)}$ e $\beta_i^{(B)}$) são, respectivamente:

```
> alfaA.i <- 3; betaA.i <- 3; alfaB.i <- 4; betaB.i <- 5
```

a.

A esperança para uma distribuição gama é α/β , enquanto que a variância é α/β^2 . Portanto, as esperanças para as prioris não-informativa, informativa para A e informativa para B são, respectivamente:

```
> c(alfa.ni/beta.ni, alfaA.i/betaA.i, alfaB.i/betaB.i)
[1] 3.333333 1.000000 0.800000
```

Já as variâncias para estes mesmos três casos, são:

```
> c(alfa.ni/beta.ni^2, alfaA.i/betaA.i^2, alfaB.i/betaB.i^2)
[1] 11.111111 0.3333333 0.1600000
```

b.

Segue o gráfico com as prioris (Figura 16) e o código usado para a sua construção.

```
> teta <- seq(0,3.4,0.02)
> plot(teta,dgamma(teta,alfa.ni,beta.ni),type="l",xlab=expression(theta),
+       ylab=expression(p(theta)),ylim=c(0,1.2))
> lines(teta,dgamma(teta,alfaA.i,betaA.i),lty=2)
> lines(teta,dgamma(teta,alfaB.i,betaB.i),lty=3)
> legend(1.3,1.1,c("não-informativa","informativa A","informativa B"),
+         lty=c(1,2,3),cex=0.8)
```

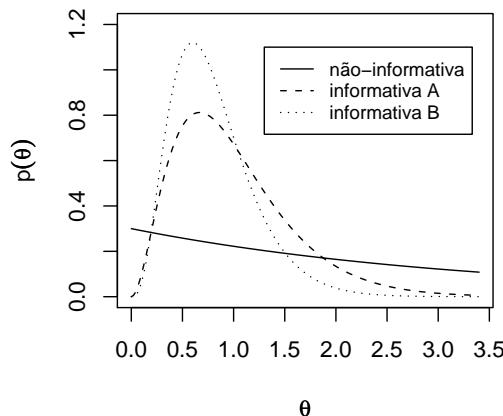


Figura 16:

c.

As funções de verossimilhanças são definidas por processos de poisson em que temos $p(X|\theta) \sim Poi(\mu = n \cdot \theta)$. Segue o gráfico com as verossimilhanças para as áreas A e B (Figura 17) e o código usado para a sua construção.

```

> teta <- seq(0,3.4,0.02)
> plot(teta,dpois(xA,nA*teta),type="l",xlab=expression(theta),
+       ylab=expression(p(paste(X,"|",theta))),ylim=c(0,0.2))
> lines(teta,dpois(xB,nB*teta),lty=2)
> legend(2,0.2,c("A","B"),lty=c(1,2))

```

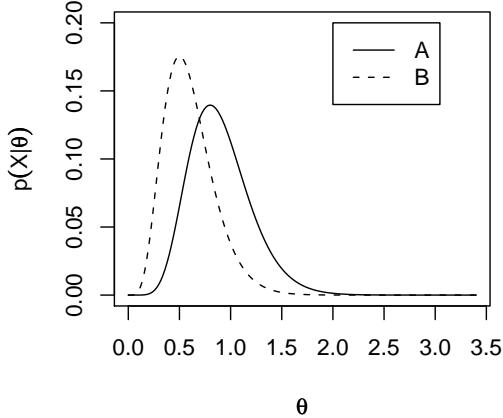


Figura 17:

d.

Se a priori é $p(\theta) \sim Gama(\alpha, \beta)$ e a verossimilhança é $p(X|\theta) \sim Poi(\mu = n \cdot \theta)$ então a distribuição posterior é $p(\theta|X) \sim Gama(\alpha^* = \alpha + x, \beta^* = \beta + n)$. Portanto os parâmetros α^* e β^* das distribuições posteriores para a área A com as prioris não-informativa (marcadas com n_i) e informativa (i) são, respectivamente:

```

> alfaAp.ni <- alfa.ni+xA; betaAp.ni <- beta.ni+nA
> c(alfaAp.ni,betaAp.ni)
[1] 9.0 10.3
> alfaAp.i <- alfaA.i+xA; betaAp.i <- betaA.i+nA
> c(alfaAp.i,betaAp.i)
[1] 11 13

```

Já os parâmetros α^* e β^* das distribuições posteriores para a área B com as prioris não-informativa e informativa são, respectivamente:

```

> alfaBp.ni <- alfa.ni+xB; betaBp.ni <- beta.ni+nB
> c(alfaBp.ni,betaBp.ni)
[1] 6.0 10.3
> alfaBp.i <- alfaB.i+xB; betaBp.i <- betaB.i+nB
> c(alfaBp.i,betaBp.i)
[1] 9 15

```

Finalmente, segue o gráfico com as distribuições posteriores (Figura 18) e o código usado para a sua construção.

```

> teta <- seq(0,3.4,0.02)
> par(mfrow=c(1,2))
> plot(teta,dgamma(teta,alfaAp.i,betaAp.i),type="l",xlab=expression(theta),
+       ylab=expression(p(paste(theta,"|",X))),ylim=c(0,2.1))
> lines(teta,dgamma(teta,alfaAp.ni,betaAp.ni),lty=2)
> legend(1.5,2,lty=c(1,2),c("informativa A","não-informativa A"))
> plot(teta,dgamma(teta,alfaBp.i,betaBp.i),type="l",xlab=expression(theta),
+       ylab=expression(p(paste(theta,"|",X))),ylim=c(0,2.1))
> lines(teta,dgamma(teta,alfaBp.ni,betaBp.ni),lty=2)
> legend(1.5,2,lty=c(1,2),c("informativa B","não-informativa B"))

```

e.

As amostras com 5000 valores, retiradas das posteriores para a área A construídas com as prioris não-informativa e informativa são, respectivamente:

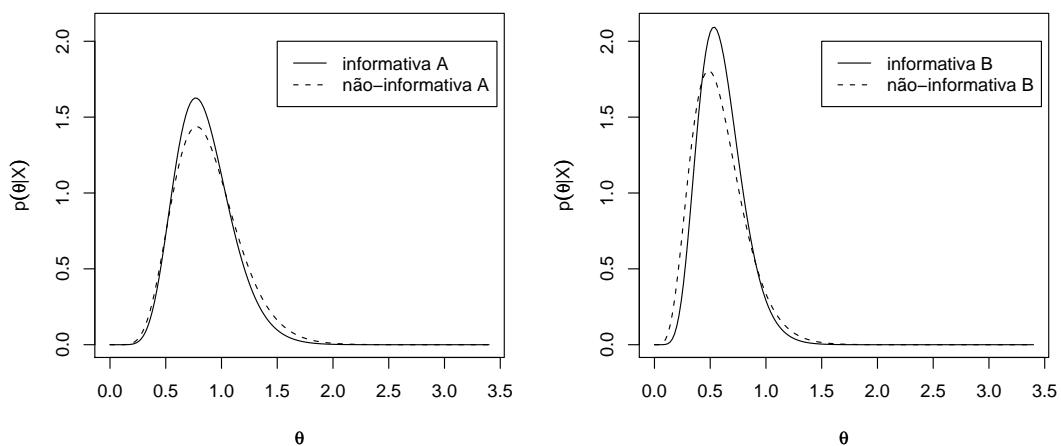


Figura 18:

```
> amostraA.ni <- rgamma(5000, alfaAp.ni, betaAp.ni)
> amostraA.i <- rgamma(5000, alfaAp.i, betaAp.i)
```

Similarmente, temos que para a área B as amostras das posteriores obtidas com as prioris não-informativa e informativa são, respectivamente:

```
> amostraB.ni <- rgamma(5000, alfaBp.ni, betaBp.ni)
> amostraB.i <- rgamma(5000, alfaBp.i, betaBp.i)
```

Os valores de δ para cada caso são:

```
> delta.ni <- amostraA.ni+amostraB.ni
> delta.i <- amostraB.i+amostraA.i
```

Os histogramas das distribuições empíricas obtidas com as duas prioris, construídos com o código abaixo, são mostrados na figura 19.

```
> par(mfrow=c(1,2))
> hist(delta.ni)
> hist(delta.i)
```

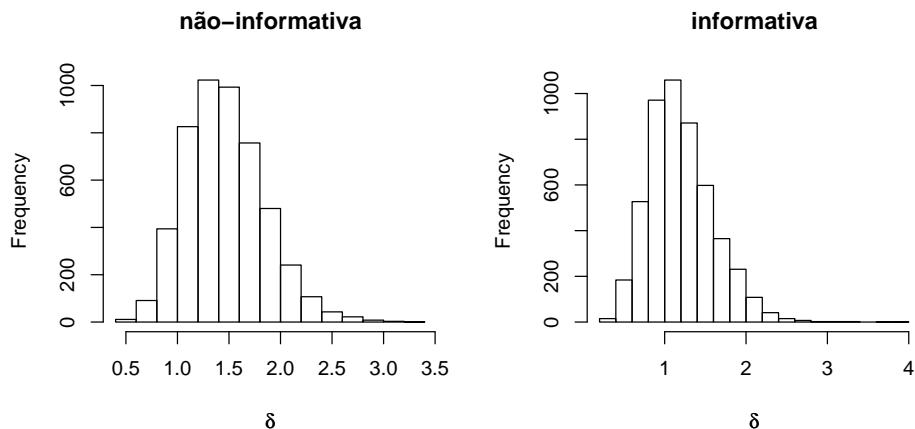


Figura 19:

f.

As probabilidades $p(\delta > 1.5)$ para as abordagens “não-informativa” e “informativa” são, respectivamente:

```
> sum(delta.ni>1.5)/length(delta.ni)
[1] 0.4176
> sum(delta.i>1.5)/length(delta.i)
[1] 0.2222
```

Exercício 5.4.

Vamos denotar com b e m os objetos referentes à *boa* e à *má* prática de amostragem, respectivamente.

a.

Primeiramente definimos o tamanho da amostra do *SIR*:

```
> n <- 5000
```

Segue o código com a retirada das amostras de θ e o cálculo dos pesos normalizados para a *boa*

```
> amostra.b <- rt(n, 3)
> peso.b <- dnorm(amostra.b) / dt(amostra.b, 3)
> pesonorm.b <- peso.b / sum(peso.b)
```

e a *má* prática

```
> amostra.m <- rnorm(n)
> peso.m <- dt(amostra.m, 3) / dnorm(amostra.m)
> pesonorm.m <- peso.m / sum(peso.m)
```

b.

As Entropias Relativas à Uniformidade (*ERU*) para a boa e a má prática são, respectivamente:

```
> -sum(pesonorm.b * log(pesonorm.b)) / log(n)
[1] 0.9916753
> -sum(pesonorm.m * log(pesonorm.m)) / log(n)
[1] 0.9644569
```

Vemos que os valores de entropia (*ERU*) são similares mas que o obtido com a boa prática é um pouco maior que o outro. c. Figura 20

```
> par(mfrow=c(1, 2))
> plot(amostra.b, pesonorm.b)
> abline(h=0.01, lty=2)
> plot(amostra.m, pesonorm.m)
> abline(h=0.01, lty=2)
```

A linha tracejada horizontal no segundo painel marca o valor 0.01.

d.

Vemos que na má prática há uma amostra que tem peso superior a 0.01. Neste caso é então indicada uma amostragem sem reposição. Ver comentários na seção “Re-Amostragem por Importância”.

e.

Um exemplo de código que poderia ser utilizado para extrair a amostra *SIR* seria:

```
> n.sir <- 500
> teta.bc <- sample(amostra.b, size=n.sir, replace=TRUE, prob=pesonorm.b)
> teta.bs <- sample(amostra.b, size=n.sir, replace=FALSE, prob=pesonorm.b)
> teta.mc <- sample(amostra.m, size=n.sir, replace=TRUE, prob=pesonorm.m)
> teta.ms <- sample(amostra.m, size=n.sir, replace=FALSE, prob=pesonorm.m)
```

No código acima as terminações *mc*, *ms*, *bc* e *bs* indicam, respectivamente, *má* prática com e sem reposição e *boa* prática com e sem reposição.

f.

Segue os cálculos da média, dos desvio padrão e do percentil de 90% para a boa prática com reposição

```
> c(media=mean(teta.bc), desvio=sd(teta.bc), quantile(teta.bc, prob=0.9))
      media    desvio    90%
0.05842466 0.98940362 1.29676844
```

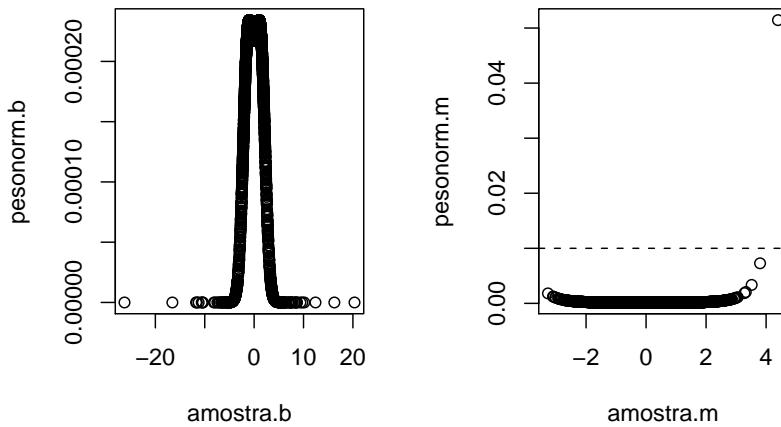


Figura 20:

para a boa prática sem reposição

```
> c(media=mean(teta.bs), desvio=sd(teta.bs), quantile(teta.bs, prob=0.9))
  media      desvio      90%
-0.02042704 1.03536363 1.24955034
```

para a má prática com reposição

```
> c(media=mean(teta.mc), desvio=sd(teta.mc), quantile(teta.mc, prob=0.9))
  media      desvio      90%
0.3087424 1.5473291 2.2671195
```

e, finalmente, para a má prática sem reposição

```
> c(media=mean(teta.ms), desvio=sd(teta.ms), quantile(teta.ms, prob=0.9))
  media      desvio      90%
0.02335291 1.17774915 1.43673371
```

Vemos que na boa prática os valores calculados a partir da amostra se aproximam bastante dos valores verdadeiros. Isto é particularmente evidente no caso da amostragem realizada sem reposição. Já no caso da má prática somente o valor de média obtido com a amostragem sem reposição se aproxima do valor verdadeiro, o que indica que o procedimento é falho. Isto é especialmente evidente no caso da amostragem com reposição. Toda essa análise deve ser vista com cautela pois aqui foi realizada uma única simulação de cada uma das abordagens (boa e má prática, com e sem reposição). No próximo exercício há resultados obtidos com uma maior quantidade de simulações.

Exercício 5.5.

Há pouca diferença quanto ao viés para a função μ em qualquer um dos casos, boa e má prática, com ou sem reposição. Vemos no entanto que o viés para σ e $P_{90\%}$ no caso da boa prática (com ou sem reposição) é pequeno enquanto que no caso da má prática o viés é mais acentuado, particularmente para a estimativa de σ . Também fica evidente que no caso da má prática, o viés é um pouco maior se utilizada uma amostragem sem reposição.

Ao avaliarmos DP como índice das precisões das estimativas vemos que no caso da boa prática há pouca diferença. Já no caso da má prática as amostragens com reposição são menos precisas.

De maneira geral vemos que se utilizada uma boa prática não há grandes distinções entre os resultados obtidos em amostragens com e sem reposição. Já no caso da má prática é aparente que o uso de amostragens com reposição resultam em estimativas menos viesadas mas, menos precisas.

Capítulo 6

Exercício 6.1.

a.

Os estimadores para as três perdas são, respectivamente, a média, a mediana e a moda, que em uma distribuição normal são iguais. No problema temos então que os três estimadores de Bayes são equivalentes a 30.109.

b.

Os limites do intervalo de credibilidade $IC_{95\%}$ correspondem simplesmente os percentis de 2.5% e 97.5% que são, respectivamente:

```
> qnorm(c(0.025, 0.975), 30.109, 0.996)
[1] 28.15688 32.06112
```

c.

Iremos comparar a hipótese $H_0: \mu > 35$ versus $H_1: \mu \leq 35$. Rejeitamos H_0 se a sua probabilidade for inferior a $w_1/(w_0 + w_1)$. Como $w_0 = w_1$ temos que esse limite crítico é $w_1/(w_0 + w_1) = 0.5$. A probabilidade para H_0 é $p(\mu > 35) = 1 - p(\mu \leq 35)$ que é:

```
> 1-pnorm(35, 30.109, 0.996)
[1] 4.538921e-07
```

Como essa probabilidade é menor que 0.5 rejeitamos H_0 .

Exercício 6.2.

a.

Como os $w_0 = w_1$ a hipótese $H_0: \theta \geq 0.5$ será rejeitada se a probabilidade posterior $p(\theta \geq 0.5|X)$ for menor que $w_1/(w_0 + w_1) = 0.5$.

b.

O valor de $p(\theta \geq 0.5|X)$ é

```
> 1-pbeta(0.5, 1, 5)
[1] 0.03125
```

e é menor que $w_1/(w_0 + w_1) = 0.5$. Conclui-se portanto que rejeitamos H_0 . É mais provável que o novo tratamento seja mais eficiente.

c.

Para rejeitar H_0 deve ser satisfeita a condição

$$\begin{aligned} p(\theta \geq 0.5|X) &< \frac{w_1}{w_0 + w_1} \\ w_0 \cdot p(\theta \geq 0.5|X) + w_1 \cdot p(\theta \geq 0.5|X) &> w_1 \\ w_0 \cdot p(\theta \geq 0.5|X) &> w_1[1 - p(\theta \geq 0.5|X)] \\ \frac{w_0}{w_1} &> \frac{[1 - p(\theta \geq 0.5|X)]}{p(\theta \geq 0.5|X)} \end{aligned}$$

Portanto a razão w_0/w_1 deve ser de no mínimo

```
> pbeta(0.5, 1, 5) / (1-pbeta(0.5, 1, 5))
[1] 31
```

d.

Com $h = 0.15$ estaríamos considerando que valores no intervalo $0.35 \leq \theta \leq 0.65$ constituem indícios favoráveis a $H_0: \theta = 0.5$ em um teste bilateral. Temos que a probabilidade $p_0^h(x) = P(0.35 \leq \theta \leq 0.65|X) = P(\theta \leq 0.65|X) - P(\theta \leq 0.35|X)$ é

```
> pbeta(0.65, 1, 5) - pbeta(0.35, 1, 5)
[1] 0.1107769
```

Se considerarmos o proposto inicialmente, $w_0 = w_1$, o limite de decisão continua sendo $w_1/(w_0 + w_1) = 0.5$. Como $p_0^{h=0.15}(x) \cong 0.110$ é menor que 0.5, optamos pela rejeição de H_0 .

Exercício 6.3.

Ao resolver o problema 4.4 havíamos obtido a posterior $p(\delta|X) \sim N(\mu_\delta = 4.76, \sigma_\delta \cong 3.05)$. Se fixarmos $h = 1.5$ temos que valores no intervalo $0 - 1.5 \leq \delta \leq 0 + 1.5$ são favoráveis à hipótese $H_0: \delta = 0$ em um teste bilateral. A probabilidade posterior $P(-1.5 \leq \delta \leq 1.5|X) = P(\delta \leq 1.5|X) - P(\delta \leq -1.5|X)$ é:

```
> pnorm(1.5, 4.76, 3.05) - pnorm(-1.5, 4.76, 3.05)
[1] 0.1225055
```

Como $w_0 = w_1 = 1$, temos que o limite crítico é $w_1/(w_0 + w_1) = 0.5$. Visto que $P(-1.5 \leq \delta \leq 1.5|X) < 0.5$, optamos então por rejeitar H_0 . É mais provável que a diferença entre as concentrações médias de óxidos não seja igual a zero.

Exercício 6.4.

a.

A tabela com as consequências é:

```
> perdas <- matrix(c(1200, 500, -1000, 600, 200, -100, -800, -100, 1000), ncol=3)
> perdas
 [,1] [,2] [,3]
[1,] 1200  600 -800
[2,]  500  200 -100
[3,] -1000 -100 1000
```

É importante notar que nesse caso as consequências correspondem a reais e que portanto estaremos procurando pela decisão que maximiza a *utilidade esperada*, e não pela que minimiza a *perda esperada*.

O vetor linha com as probabilidades posteriores para os estados da natureza, $p(\theta|X)$, é:

```
> post <- c(0.2, 0.5, 0.3)
```

As *utilidades esperadas* são:

```
> perdas %*% post
 [,1]
[1,] 300
[2,] 170
[3,] 50
```

A operação `%*%` acima indica um produto

```
> 0.2*1200+0.5*600+0.3*(-800)
[1] 300
```

A decisão de Bayes é a primeira (“Plantio em Larga Escala”) visto que essa é a alternativa que corresponde à maior *utilidade esperada* (300).

b.

Neste caso as probabilidades posteriores e as *utilidades esperadas* são:

```
> postb <- c(0.1, 0.5, 0.4)
> perdas %*% postb
 [,1]
[1,] 100
[2,] 110
[3,] 250
```

Isto implica que a melhor alternativa é não plantar nada.

Exercício 6.5.

Primeiro retiramos as constantes para fora da integral e adicionamos os expoentes dos termos potenciais que tem base comum

$$\int_0^1 \theta^{\tilde{x}} (1-\theta)^{1-\tilde{x}} \frac{\Gamma(6)}{\Gamma(1)\Gamma(5)} \theta^0 (1-\theta)^4 d\theta$$

$$\frac{\Gamma(6)}{\Gamma(1)\Gamma(5)} \int_0^1 \theta^{\tilde{x}} (1-\theta)^{5-\tilde{x}} d\theta$$

Se fizermos $\tilde{x} = \alpha - 1$ isto implica que $\alpha = \tilde{x} + 1$. Da mesma forma, se fizermos $5 - \tilde{x} = \beta - 1$ isto implica que $\beta = 6 - \tilde{x}$. Agora podemos pensar no integrando como sendo o núcleo de uma distribuição $Beta(\alpha = \tilde{x} + 1, \beta = 6 - \tilde{x})$. Incluímos agora na função termos que tornam o integrando uma distribuição beta completa.

$$\frac{\Gamma(6)}{\Gamma(1)\Gamma(5)} \cdot \frac{\Gamma(\alpha = \tilde{x} + 1)\Gamma(\beta = 6 - \tilde{x})}{\Gamma(\alpha + \beta = 7)} \int_0^1 \frac{\Gamma(\alpha + \beta = 7)}{\Gamma(\alpha = \tilde{x} + 1)\Gamma(\beta = 6 - \tilde{x})} \theta^{\alpha-1=\tilde{x}}(1-\theta)^{\beta-1=5-\tilde{x}} d\theta$$

A integral da distribuição beta completa a direita equivale a 1 e, obtemos então a solução procurada:

$$\begin{aligned} & \frac{\Gamma(6)}{\Gamma(1)\Gamma(5)} \cdot \frac{\Gamma(\tilde{x} + 1)\Gamma(6 - \tilde{x})}{\Gamma(7)} \\ & \frac{5!}{0!4!} \cdot \frac{\Gamma(\tilde{x} + 1)\Gamma(6 - \tilde{x})}{6!} \\ & \frac{\Gamma(\tilde{x} + 1)\Gamma(6 - \tilde{x})}{6 \cdot \Gamma(5)} \end{aligned}$$

Exercício 6.6.

Note inicialmente que se $H_0 : \theta < 0.5$, então $H_1 : \theta \geq 0.5$. Prosseguimos com o cálculo da razão entre as posteriores $o(1, 0|x) = p_1(x)/p_0(x)$

```
> o10x <- (1-pbeta(0.5, 1, 5))/pbeta(0.5, 1, 5)
> o10x
[1] 0.03225806
```

e, depois com cálculo da razão entre as prioris $o(1, 0) = p_1/p_0$

```
> o10 <- (1-punif(0.5, 0, 1))/punif(0.5, 0, 1)
> o10
[1] 1
```

Agora podemos calcular o Fator de Bayes $FB_{1,0} = o(1, 0|x)/o(1, 0)$

```
> FB <- o10x/o10
> FB
[1] 0.03225806
```

Como regra devemos rejeitar H_0 sempre que for satisfeita a desigualdade

$$FB_{1,0} > \frac{w_0}{w_1} \cdot \frac{1}{o(1, 0)}$$

e portanto, para a rejeição temos que

$$0.03225806 > \frac{w_0}{w_1} \cdot \frac{1}{1} \Rightarrow w_1 > 31 \cdot w_0$$

Isto implica que para rejeitar H_0 , a penalização w_1 deve ser 31 vezes maior que a w_0 . Lembramos que w_1 é a penalização imposta quando equivocadamente escolhemos H_0 em detrimento de H_1 e que, w_0 é a penalização imposta quando equivocadamente escolhemos H_0 em detrimento de H_1 . Como temos fortes evidências a favor de H_0 , nós somente devemos rejeitar essa hipótese se, as consequências de escolher equivocadamente H_0 forem de enorme gravidade.

Capítulo 7

Exercício 7.1.

a.

Temos que $z = x - \bar{x}$ o que implica que $x = z + \bar{x}$. Se substituirmos x por essa solução na regressão original obtemos:

$$\begin{aligned} y &= \beta_0 + \beta_1 x = \beta_0 + \beta_1(z + \bar{x}) \\ y &= \beta_0 + \beta_1 \bar{x} + \beta_1 z \end{aligned}$$

Os dois primeiros termos a direita da igualdade correspondem a uma constante. Então podemos escrever a solução como uma regressão linear simples de y com respeito a z :

$$y = \beta'_0 + \beta'_1 z$$

Vemos então que a inclinação dessa regressão corresponde à inclinação da regressão original $\beta'_1 = \beta_1$ e que o intercepto é $\beta'_0 = \beta_0 + \beta_1 \bar{x}$.

b.

O intercepto β'_0 corresponde então ao valor médio de y para um valor de x equivalente à média \bar{x} . Já a inclinação β'_1 é multiplicativo de $z = (x - \bar{x})$ e corresponde então ao efeito de acréscimos ou decréscimos com respeito à média. Se considerarmos como exemplo um $x^* = \bar{x}$ vemos que $z^* = x^* - \bar{x} = 0$ e $y = \beta'_0 + \beta'_1 z^* = \beta'_0$, que é então o y médio esperado para x igual a \bar{x} . Se tomarmos como exemplo $x^* = \bar{x} + 1$ vemos então que $z = \bar{x} + 1 - \bar{x} = 1$ e $y = \beta'_0 + \beta'_1 z^* = \beta'_0 + \beta'_1 \cdot 1$, que é então a média esperada correspondente ao valor \bar{x} (efeito β'_0) mais o efeito do acréscimo de uma unidade $\beta'_1 \cdot 1$.

Exercício 7.2.

a.

Segue abaixo o código com a entrada de dados, o ajuste do modelo, a retirada da amostra da posterior conjunta dos parâmetros do modelo, como disposto no livro.

```
x <- c(1.9, 3.1, 3.3, 4.8, 5.3, 6.1, 6.4, 7.6, 9.8, 12.4)
y <- c(2, 1, 5, 5, 20, 20, 23, 10, 30, 25)
n <- length(x)
modelo <- lm(y~x)
est <- coef(modelo)
se <- summary(modelo)$sigma
v.beta <- summary(modelo)$cov.unscaled
m <- 5000
prec <- rgamma(m, (n-2)/2, se^2*(n-2)/2)
sig2 <- 1/prec
library(MASS)
beta.conj <- numeric()
for(i in 1:m){
    beta.conj <- rbind(beta.conj,mvrnorm(1,est,sig2[i]*v.beta))
}
```

Agora fixamos o valor pretendido de x_p e calculamos valores de interesse para a estimativa da média e da escala da distribuição preditiva student não central.

```
xp <- 4
xbar <- mean(x)
sxx <- mean(x*x)-xbar^2
```

Segue o cálculo da escala e da média, e o código para obter o gráfico da densidade (Figura 21) para as previsões y_{tilde} e alguns percentis. small

```
escala <- sqrt(se^2*((n+1)/n+(xp-xbar)^2/(n*sxx)))
media <- est[1]+est[2]*xp
ytil <- seq(-20,40,.1)
ttil <- (ytil-media)/escala
dens <- dt(ttil,n-2)/escala
plot(ytil,dens,type="l")
qt(c(0.025,0.25,0.50,0.75,0.975),n-2)*escala+media
[1] -8.049708 3.485792 8.579836 13.673880 25.209380
```

O cálculo da densidade (`dens`) e dos percentis é baseado no uso de soluções para transformadas de variáveis. Em particular a abordagem consiste em transformar a student não-central em uma student padrão. Esse desenvolvimento consta no capítulo sobre as distribuições de probabilidade.

b.

Ao avaliar a figura 21, e também o intervalo de credibilidade ($IC_{95\%}$) indicado pelos percentis mostrados no item anterior, fica evidente há previsões negativas para \hat{y} (depressão em milímetros). Esta é uma variável extritamente positiva e portanto, o modelo é inadequado para previsões.

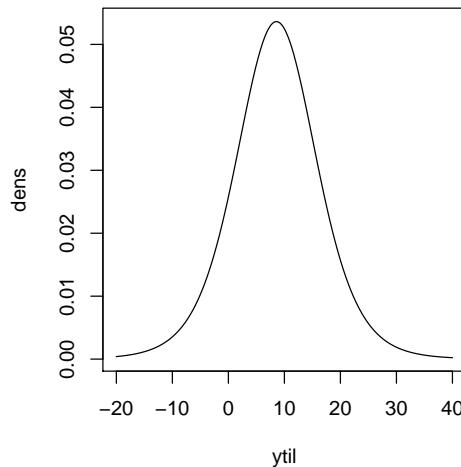


Figura 21:

c.

O gráfico obtido com o código `plot(log(x), log(y))` consta na figura 22. Vemos que um modelo de regressão linear é razoável. Para fins de exercício vamos assumir que tal modelo pode ser usado.

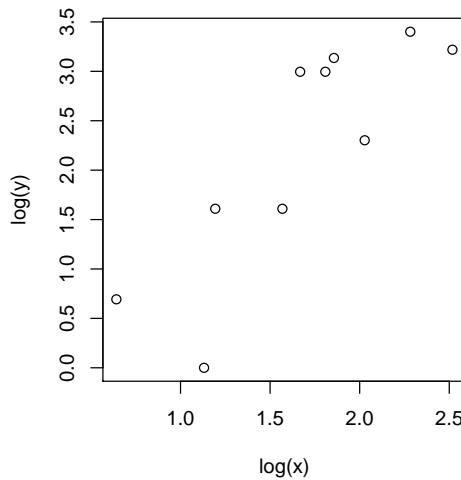


Figura 22:

d.

Depois de converter os dados para a escala logarítmica, repetimos toda a operação realizada no item a.

```

x <- log(x)
y <- log(y)
modelo <- lm(y~x)
est <- coef(modelo)
se <- summary(modelo)$sigma
v.beta <- summary(modelo)$cov.unscaled
m <- 5000
prec <- rgamma(m, (n-2)/2, se^2*(n-2)/2)
sig2 <- 1/prec
library(MASS)
beta.conj <- numeric()
for(i in 1:m){
    beta.conj <- rbind(beta.conj,mvrnorm(1,est,sig2[i]*v.beta))
}

```

```
xbar <- mean(x)
sxx <- mean(x*x)-xbar^2
```

Fixamos agora o valor de x_p e obtemos a escala e a média.

```
> xp <- log(4)
> escala <- sqrt(se^2*((n+1)/n+(xp-xbar)^2/(n*sxx)))
> media <- est[1]+est[2]*xp
```

Note que temos que entrar com o valor de x_p na escala logarítmica, já que o modelo foi ajustado para dados logaritimizados. Da mesma forma é importante lembrar que a média e a escala são para o logaritmo de y . A distribuição preditiva para y é obtida em duas etapas. Primeiro obtemos a preditiva para $\log(y)$ que é uma student não-central e depois obtemos a preditiva para a variável de interesse y .

Para obter a densidade de probabilidade de $\log(y)$ com as funções do R temos que usar a solução indicada na seção que trata de transformadas de variáveis no capítulo sobre distribuições de probabilidade. Podemos obter as densidades de uma student não-central a partir das densidades de uma student padrão se calcularmos essa última e dividirmos pelo parâmetro de escala. Definimos então um vetor com a amplitude de valores de interesse

```
> ytil <- seq(10,50,0.1)
```

e transformamos estes valores para a escala logarítmica a qual sabemos que corresponde a uma student não-central

```
> logytil <- log(ytil)
```

Transformamos agora estes valores para uma variável student padronizada

```
> logytilpad <- (logytil-media)/escala
```

e calculamos agora as densidades preditivas do logaritmo de y que segue uma student não-central

```
> denslogy <- dt(logytilpad,n-2)/escala
```

Vamos agora à segunda etapa que consiste em obter densidades de probabilidade para y a partir da densidade de probabilidade de $\log(y)$ obtida acima. Recorremos novamente à uma das soluções indicadas na seção que trata de probabilidade para transformações de variáveis. As densidades de y correspondem às densidades calculadas para $\log(y)$ divididas por y . Segue então o código para o cálculo dessas densidades e para a construção do gráfico da distribuição preditiva de y (Figura 23).

```
> densytil <- denslogy/ytil
> plot(ytil,densytil,type="l")
```

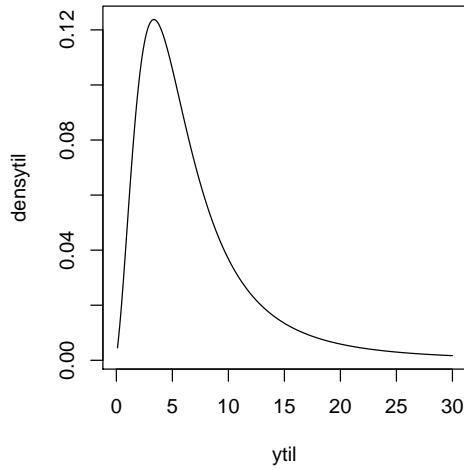


Figura 23:

Percentis para essa distribuição dependem de transformações. Primeiro calculamos os percentis para a distribuição de $\log(y)$ que é uma student não-central

```
> perc <- qt(c(0.025,0.25,0.50,0.75,0.975),n-2)*escala+media
```

e convertemos os valores para a escala da variável de interesse y :

```
> exp(perc)
[1] 1.031495 3.286172 5.481661 9.143955 29.131119
```

e.

A figura 23 e o intervalo de credibilidade ($IC_{95\%}$) que compreende valores de aproximadamente 1 a 29 apontam para previsões positivas para a depressão (mm), o que é bem mais razoável que as obtidas anteriormente no item a. Vemos portanto que o modelo de regressão ajustado aos dados logaritimizados mais útil que o anterior.

Exercício 7.3.

a.

Primeiramente a base de dados é lida e armazenada em um objeto com:

```
> prionotus <- read.csv(file="../dados/prionotus.csv", header=T)
```

É importante notar que o argumento `file` deve corresponder à indicação de onde está o arquivo `prionotus.csv` no seu computador. Segue o código para obter a figura com diagramas de dispersão para as variáveis nas escalas original e logarítmica (Figura 24)

```
plot(prionotus$comprimento, prionotus$peso, cex=0.6)
plot(log(prionotus$comprimento), log(prionotus$peso), cex=0.6)
```

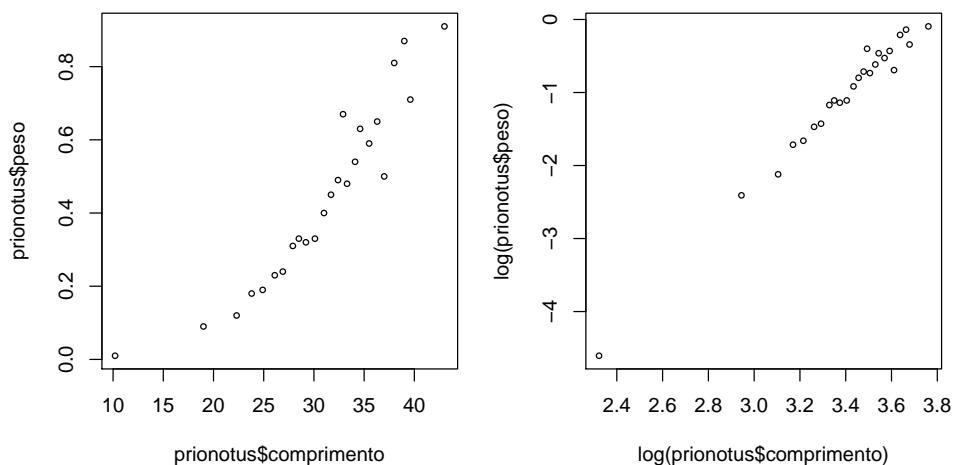


Figura 24:

A seguir armazenamos os logarítmicos das variáveis de interesse em `lx` e `ly`, ajustamos o modelo e obtemos a amostra da posterior conjunta de β'_0 e β'_1 usando as soluções que constam na seção sobre regressão linear simples.

```
> lx <- log(prionotus$comprimento)
> ly <- log(prionotus$peso)
> n <- length(lx)
> modlog <- lm(ly~lx)
> est <- coef(modlog)
> se <- summary(modlog)$sigma
> v.beta <- summary(modlog)$cov.unscaled
> m <- 5000
> prec <- rgamma(m, (n-2)/2, se^2*(n-2)/2)
> sig2 <- 1/prec
> library(MASS)
> beta.conj <- numeric()
> for(i in 1:m){
+   beta.conj <- rbind(beta.conj, mvrnorm(1, est, sig2[i]*v.beta))
+ }
```

A primeira coluna de `beta.conj` corresponde a β'_0 enquanto que a segunda corresponde a β'_1 .

b.

Segue abaixo o código para obter os gráficos de dispersão (Figura 25).

```
> plot(beta.conj[,1],beta.conj[,2],cex=0.6)
> plot(exp(beta.conj[,1]),beta.conj[,2],cex=0.6)
```

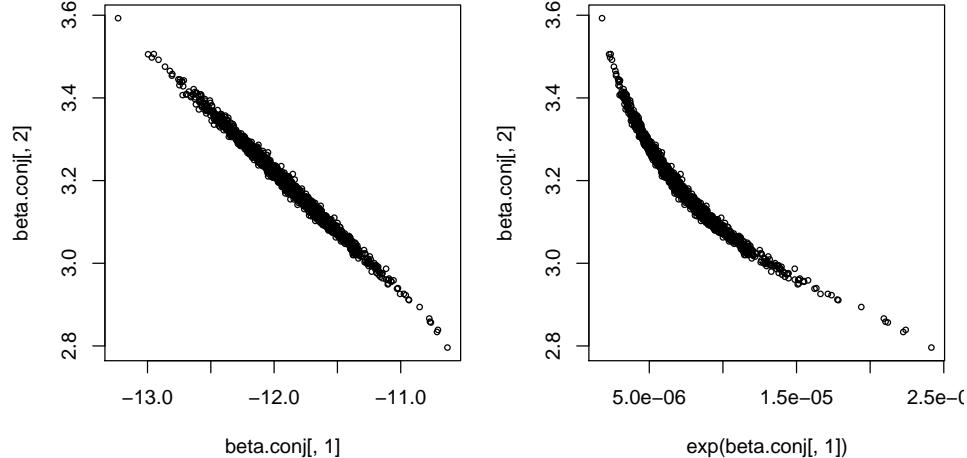


Figura 25:

Como a primeira coluna de `beta.conj` contém a amostra para $\beta'_0 = \log(\beta_0)$ há a aplicação do exponencial para obter a amostra convertida para β_0 , mostrada no painel da direita da figura 25.

c.

Segue o código para a construção dos histogramas (Figura 26) e para o cálculo dos limites dos intervalos de credibilidade.

```
> hist(exp(beta.conj[,1]),main="")
> hist(beta.conj[,2],main="")
> quantile(exp(beta.conj[,1]),c(0.025,0.975))
  2.5%    97.5%
3.594753e-06 1.263133e-05
> quantile(beta.conj[,2],c(0.025,0.975))
  2.5%    97.5%
3.009485 3.386055
```

Vemos que a distribuição posterior de β_1 é simétrica e que a de β_0 tem leve assimetria positiva. O intervalo de credibilidade e a distribuição posterior indicam que é muito provável que β_1 para a espécie avaliada seja superior a 3. Na verdade podemos estimar a probabilidade $P(\beta_1 > 3)$ em:

```
> sum(beta.conj[,2]>3)/m
[1] 0.979
```

d.

Segue o código para o histograma (Figura 27) e o sumário de interesse.

```
> raz <- beta.conj[,2]/(10E5*exp(beta.conj[,1]))
> hist(raz,main="",nclass=50)
> c(quantile(raz,c(0.025,0.975)),mean(raz),median(raz))
  2.5%    97.5%
0.2380576 0.9394876 0.5070293 0.4765877
```

A figura indica que há uma assimetria positiva, o que é também evidenciado na comparação da mediana e da média, visto que esta última maior que a primeira. A probabilidade de interesse $P(\text{raz} > 0.5)$ é:

```
> sum(raz>0.5)/m
[1] 0.4415
```

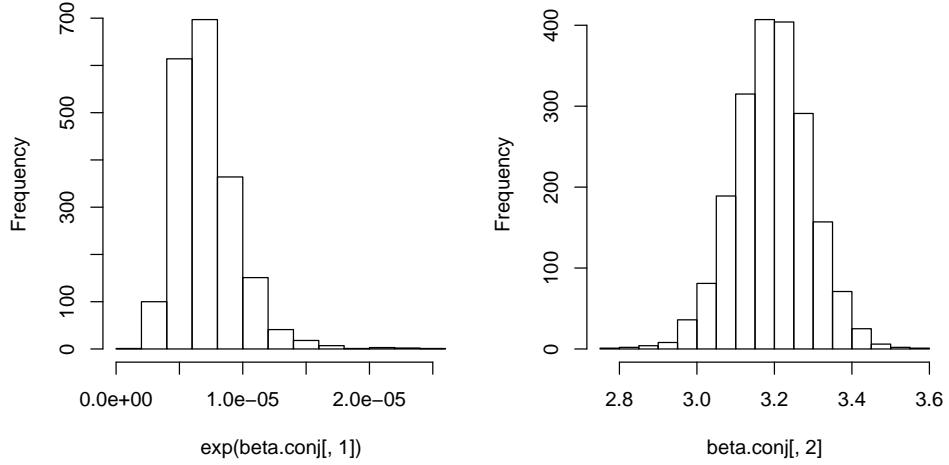


Figura 26:

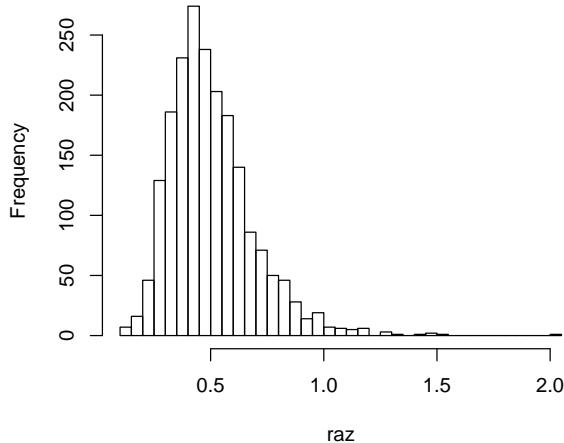


Figura 27:

e.

Queremos a distribuição preditiva para y mas, o ajuste foi realizado para os dados logaritimizados. Assim temos que partir da distribuição preditiva para $\log(y)$ que é uma distribuição student não-central. Para calcular a escala e a média para essa distribuição precisamos primeiro obter a média e a variância de $\log(y)$. No código abaixo calculamos estes valores intermediários, definimos o valor fixo para o qual será realizada a previsão (lxp), calculamos a escala e finalmente, a média da student não-central de interesse.

```
> lxbbar <- mean(lx)
> sxx <- mean(lx*lx)-lxbbar^2
> lxp <- log(20)
> escala <- sqrt(se^2*((n+1)/n+(lxp-lxbbar)^2/(n*sxx)))
> media <- est[1]+est[2]*lxp
```

Para o cálculo e o gráfico das densidades para \tilde{y} , primeiramente definimos intervalo de valores em uma amplitude de interesse e calculamos o logarítmico já que o modelo foi ajustado aos dados logaritimizados

```
> ytil <- seq(0.04,.18,0.001)
> lytil <- log(ytil)
```

Toda a seqüência de cálculo que segue abaixo também foi realizada e devidamente explicada no exercício 7.2, na resolução da letra d. Evitamos aqui repetições e segue somente o código para obter o gráfico de densidades desejado, com comentários mínimos. Primeiro transforma-se o logarítmico de y em uma variável que segue a distribuição student padrão

O código abaixo produz um gráfico de dispersão para as variáveis na escala original e também para as respectivas transformações logarítmicas (Figura 29). Nesta figura as variáveis indicadas nos três primeiros painéis no alto a esquerda referem-se aos dados na escala original, enquanto que as indicações nos três painéis mais abaixos e à direita referem-se à escala logarítmica.

```
> plot(cbind(dados.ph, log(dados.ph)))
```

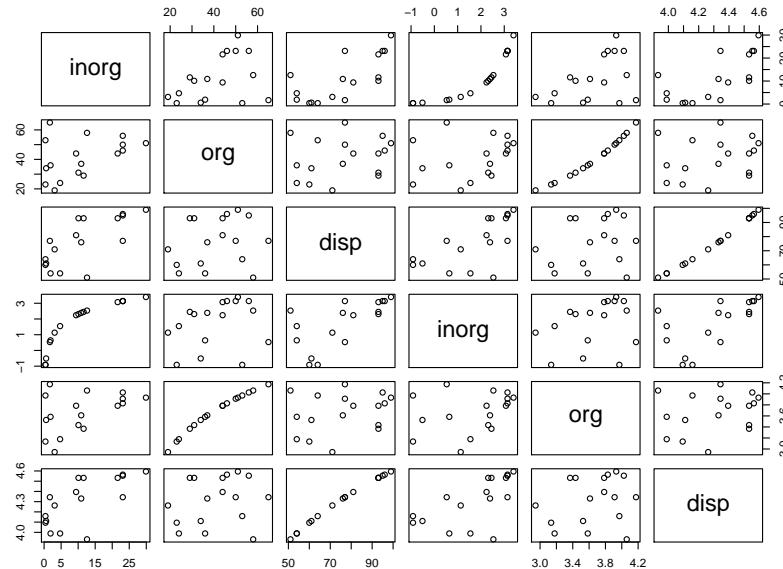


Figura 29:

Verifica-se que as transformações logarítmicas não resultam em linearizações das relações. Portanto, seguimos com a análise sem transformações.

b.

Primeiramente obtemos as variáveis explicativas padronizadas

```
> dados.ph$z1 <- (dados.ph$inorg-mean(dados.ph$inorg))/sd(dados.ph$inorg)
> dados.ph$z2 <- (dados.ph$org-mean(dados.ph$org))/sd(dados.ph$org)
```

ajustamos um modelo linear da maneira tradicional

```
> ph.lm <- lm(disp~z1+z2, data=dados.ph)
```

e extraímos o erro padrão, os coeficientes e a matriz de covariância

```
> se <- summary(ph.lm)$sigma
> est <- coef(ph.lm)
> v.beta <- summary(ph.lm)$cov.unscaled
```

Definimos agora o tamanho da amostra da posterior

```
> m <- 1000
```

retiramos a amostra da precisão e calculamos a variância

```
> prec <- rgamma(m, (nrow(dados.ph)-3)/2, se^2*(nrow(dados.ph)-3)/2)
> sig2 <- 1/prec
```

Carregamos a biblioteca necessária para gerar número aleatórios de uma normal multivariada e, retiramos a amostra dos coeficientes do modelo linear de regressão múltipla

```
> library(MASS)
> beta.ph <- numeric()
> for(i in 1:m){
+   beta.ph <- rbind(beta.ph,mvrnorm(1,est,sig2[i]*v.beta))
+ }
```

Seguem abaixo as três primeiras linhas da amostra de tamanho 3000 retirada da distribuição posterior conjunta.

```
> cbind(beta.ph,sig2)[1:3,]
  (Intercept)      z1       z2      sig2
[1,]    76.40571 12.46775 1.7267641 271.03122
[2,]    76.56951 10.50569 0.9853323 144.56682
[3,]    76.48247  8.86795 3.5565188  99.75168
```

A primeira coluna corresponde ao intercepto da regressão. As duas colunas centrais correspondem às amostras dos coeficientes das variáveis explicativas. A última coluna corresponde à variância do componente que é o “ruído” do modelo.

c.

Diagramas de dispersão para as amostra da posterior constam na figura 30. O código para gerar a figura é:

```
> plot(data.frame(cbind(beta.ph,sig2)),cex=.4)
```

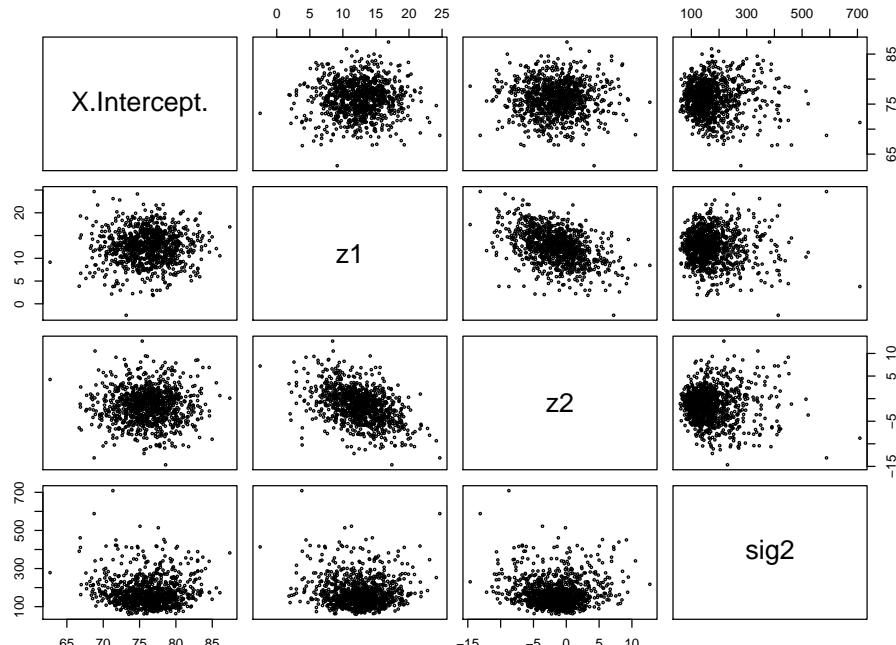


Figura 30:

Note que a forma da dispersão indica que as correlações entre as estimativas não são muito elevadas.

d.

As quantidades de fósforo inorgânico e orgânico para as quais queremos fazer a predição são, respectivamente:

```
> x1p <- 15; x2p <- 40
```

Para fazer os cálculos devemos padronizar estas quantidades já que o modelo foi ajustado a variáveis padronizadas. Armazenamos as quantidades padronizadas nos objetos $z1p$ e $z2p$:

```
> z1p <- (x1p - mean(dados.ph$inorg)) / sd(dados.ph$inorg)
> z2p <- (x2p - mean(dados.ph$org)) / sd(dados.ph$org)
```

Para retirar a amostra da distribuição posterior, aplicamos o procedimento indicado no livro. Calculamos a média da distribuição preditiva para cada uma das 1000 amostras

```
> mup <- beta.ph %*% c(1, z1p, z2p)
```

e prosseguimos com a retirada da amostra da preditiva

```
> ytil <- numeric()
> ytil <- for(i in 1:length(mup)) {
+   ytil <- c(ytil,rnorm(1,mup[i],sqrt(sig2[i])))
+ }
```

Os três primeiros valores da amostra da preditiva são:

```
> ytil[1:3]
[1] 71.67071 94.09193 83.29817
```

e.

Um sumário estatístico é obtido com

```
> quantile(ytil,prob=c(.025,.25,.5,.75,.975))
  2.5%    25%    50%    75%   97.5%
55.52778 72.67019 81.42637 90.29232 107.07403
```

e, o histograma da amostra da preditiva obtida com o código abaixo

```
> hist(ytil,main="")
```

é mostrado na figura 31. O intervalo de confiança (95%) compreende valores entre 56 e 107 ppm. Vemos também que a distribuição é aproximadamente simétrica e centrada em torno de 80.

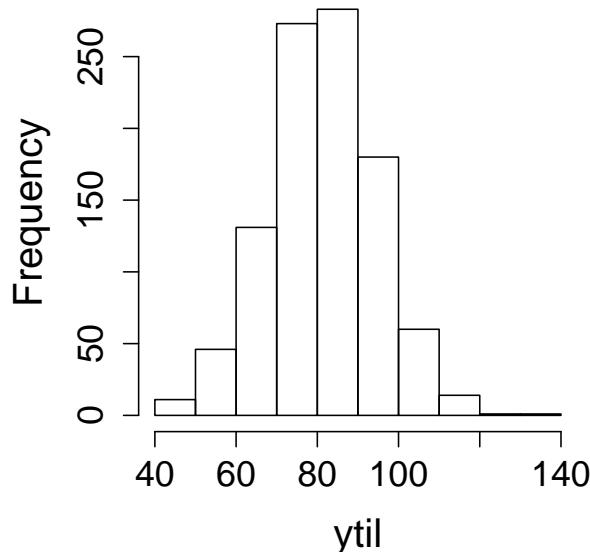


Figura 31:

Exercício 7.5

a.

Começamos com a entrada de dados

```
> praia <- factor(c(1,1,1,1,2,2,2,3,3,3,3,3,4,4,4,4))
> comprimento <- c(100.4,96.7,96.3,96.2,94.6,92.8,89.0,92.2,96.6,
+                   95.3,95.9,93.2,93.2,95.3,97.3,99.3,99.0,96.4)
```

Note que no código acima já há uma indicação de que a variável `praia` é um fator. Seguimos então com os cálculos das médias

```
> ymg <- tapply(comprimento,praia,mean)
> ymg
  1      2      3      4
97.40000 92.13333 94.40000 97.46000
```

e da variância

```
> s2g <- tapply(comprimento, praia, var)
> s2g
  1       2       3       4
4.046667 8.173333 3.124000 2.893000
```

Finalmente, construímos o diagrama de caixa (Figura 32) com o código

```
> boxplot(comprimento~praia)
```

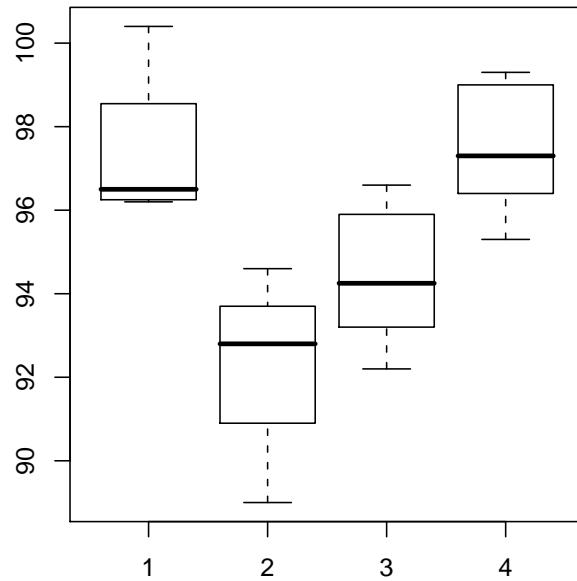


Figura 32:

b.

Antes de retirar a amostra são necessários alguns cálculos preliminares. Especificamos o número de observações em cada praia (n_g), o número total de observações (n) e o número de níveis do fator “praia” (G):

```
> ng <- tapply(comprimento, praia, length)
> n <- sum(ng)
> G <- length(ng)
```

Segue agora o cálculo da variância (S_e^2) e da matriz diagonal (V_μ), representados respectivamente por $s2hat$ e Vmu no código abaixo

```
> s2hat <- sum((ng-1)*s2g) / (n-G)
> Vmu <- diag(1/ng)
```

Para a simular da distribuição posterior, definimos quantidade de amostras desejada

```
> m <- 3000
```

retiramos uma amostra da precisão

```
> tau <- rgamma(m, (n-G)/2, s2hat*(n-G)/2)
```

e obtemos a amostra do desvio padrão (σ)

```
> sigma <- sqrt(1/tau)
```

Para retirar a amostra da posterior para os efeitos, inicialmente chamamos a biblioteca que possibilita a retirada de números aleatórios de uma normal multivariada e, prosseguimos então com a retirada da amostra

```
> library(MASS)
> mupost <- numeric()
> for(i in 1:m) mupost <-
+     rbind(mupost,mvrnorm(1,as.numeric(ymg),Vmu/tau[i]))
```

No código acima a amostra da posterior dos efeitos (μ_g) para o fator “praia” é armazenada no objeto `mupost`.

c.

Um exemplo de sumário estatístico para a amostra da posterior do desvio padrão (σ) seria:

```
> c(quantile(sigma,prob=c(0.025,0.5,0.975)),media=mean(sigma),dp=sd(sigma))
   2.5%      50%    97.5%     media      dp
1.4630281 2.0493367 3.1827688 2.1241082 0.4493866
```

O intervalo de credibilidade ($ICr_{95\%}$) compreende então valores entre 1.46 e 3.18.

Segue abaixo sumários estatísticos para os efeitos dos quatro níveis do fator “praia”

```
> for(i in 1:ncol(mupost)) {
+   print(c(quantile(mupost[,i],prob=c(0.025,0.5,0.975)),
+         media=mean(mupost[,i]),dp=sd(mupost[,i])))
+ }
   2.5%      50%    97.5%     media      dp
95.178637 97.427923 99.549721 97.401565 1.084276
   2.5%      50%    97.5%     media      dp
89.608782 92.128839 94.640067 92.141252 1.267151
   2.5%      50%    97.5%     media      dp
92.6306255 94.3784745 96.2107524 94.3775632 0.8984044
   2.5%      50%    97.5%     media      dp
95.5924960 97.4514180 99.4187497 97.4650962 0.9670546
```

Note que o intervalo de credibilidade ($ICr_{95\%}$) para o segundo nível (praia 2), indicado na segunda linha de resultados no código acima, somente apresenta sobreposição com o $ICr_{95\%}$ do segundo terceiro nível (praia 3). Note também que os intervalos de credibilidade do efeito do primeiro e do quarto nível se sobrepõe quase que completamente.

Segue agora o código que produz o gráfico com as marginais posteriores para os efeitos, mostradas na figura 33.

```
> plot(density(mupost[,1]),xlim=c(88,102),ylim=c(0,0.5),
+       xlab="mu",ylab="p(mu|dados)",main="")
> lines(density(mupost[,2]),col="red")
> lines(density(mupost[,3]),col="blue")
> lines(density(mupost[,4]),col="green")
```

No código acima a repetição de comandos `lines()` poderia ter sido evitada com uma programação menos grosseira. No entanto, fizemos esta opção para maior clareza. Na ordem, fica evidente que as cores preta, vermelha, azul e verde, representam respectivamente as marginais para o primeiro, o segundo, o terceiro e o quarto nível do fator “praia”.

d.

A probabilidade de que μ_3 seja maior que μ_2 equivale a:

```
> sum(mupost[,3]>mupost[,2])/m
[1] 0.929
```

e vemos portanto que há uma probabilidade alta de que o efeito para o nível 3 do fator “praia” seja maior que o efeito do nível 2.

Exercício 7.6

A perda esperada para a decisão de explorar a área 9 é

```
> 4*0.73+0*.27
[1] 2.92
```

enquanto que para a área 15 temos

```
> 2*0.73+6*0.27
[1] 3.08
```

Vemos portanto que a área 9 que apresenta uma menor *CPUE* parece ser a melhor opção por conta de sua maior proximidade do porto.

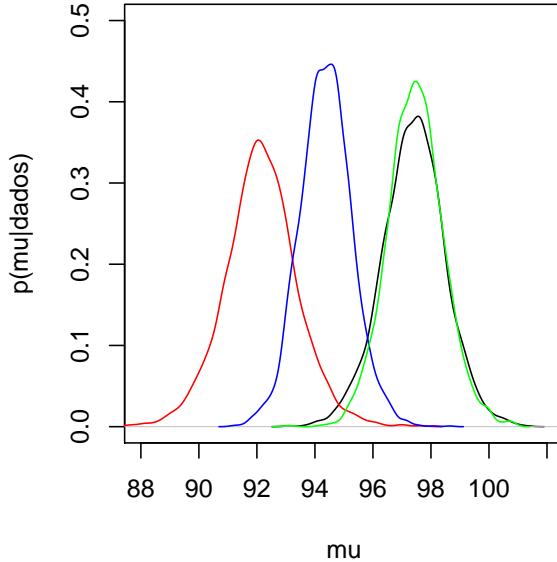


Figura 33:

Capítulo 8

Exercício 8.1.

Utilizaremos o modelo logístico com logit na forma:

$$\log \left(\frac{p_i}{1 - p_i} \right) = \beta_0 + \beta_1 \cdot x_i$$

em que p_i é a proporção de embriões com asas e x_i é a variável “tempo desde a deposição dos ovos”. Pretendemos aqui fazer o ajuste do modelo e retirar uma amostra da distribuição posterior utilizando o procedimento *MCMC*. Inicialmente entramos com as informações sobre o tempo (t), o número de ovos depositados (n) e o número de ovos que já tem asas (y):

```
> t <- c(5, 6, 8, 8, 10, 11, 16, 18)
> n <- c(34, 33, 33, 35, 30, 27, 33, 39)
> y <- c(6, 4, 23, 18, 28, 27, 33, 39)
```

Para a cadeia de Markov é interessante dispor de valores iniciais próximos do que seriam os resultados finais obtidos na convergência. Uma alternativa é ajustar um modelo linear da maneira usual e, utilizar as estimativas como valores iniciais para a cadeia de Markov. Portanto, construímos uma tabela (`dados`) com os “sucessos” (embriões com asas) e “fracassos” (embriões ainda sem asas) e utilizamos a função `glm()` obter estimativas iniciais para β_0 e β_1 .

```
> dados <- cbind(y, n-y)
> modelo <- glm(dados ~ t, family=binomial(link=logit))
> modelo
```

```
Call: glm(formula = dados ~ t, family = binomial(link = logit))
```

Coefficients:

(Intercept)	t
-7.0811	0.9505

```
Degrees of Freedom: 7 Total (i.e. Null); 6 Residual
Null Deviance: 173.5
Residual Deviance: 8.562 AIC: 29.53
```

Note que a variável explicativa da regressão é t “tempo após a deposição dos ovos”, e que obtivemos então -7.08 e 0.95 como estimativas pontuais iniciais para β_0 e β_1 respectivamente.

Seguimos carregando a biblioteca `rjags` e especificando os dados (`data.list`), os parâmetros a serem monitorados (`params`) e os valores iniciais (`inicial`) de β_0 e β_1 para o procedimento *MCMC*.

```
> library(runjags)
> datalist <- dump.format(list(metamorfose=y, total=n, tempo=t))
> params <- c("beta0", "beta1")
> inicial <- dump.format(list(beta0=coef(modelo)[1],
+                                beta1=coef(modelo)[2]))
```

Agora criamos o objeto que contém o modelo

```
> modmcmc <- "model{
+ beta0 ~ dnorm(0, 0.001)
+ beta1 ~ dnorm(0, 0.001)
+ for(i in 1:length(metamorfose)) {
+   logit(p[i]) <- beta0+beta1*tempo[i]
+   metamorfose[i] ~ dbin(p[i],total[i])
+ }}
```

e estamos preparados então para utilizar a função `run.jags()` para obter a amostra da distribuição posterior

```
> ## rodando o modelo
> modfit <- run.jags(model=modmcmc, monitor=params, data=datalist,
+                      inits=inicial, n.chains=1, burnin=50000, thin=3,
+                      sample=9000, check.conv=FALSE)
Calling the simulation... (this may take some time)
Welcome to JAGS 1.0.3 on Tue Jan 5 11:25:00 2010
JAGS is free software and comes with ABSOLUTELY NO WARRANTY
Loading module: basemod
Loading module: bugs
. <>. <Reading data file data.txt
>. <>Compiling model graph
  Resolving undeclared variables
  Allocating nodes
  Graph Size: 49
. <Reading parameter file inits1.txt
>>>. . <>Updating 50000
-----| 50000
*****| 100%
. <>. <>. <>Updating 9000
-----| 9000
*****| 100%
. <>. <><>.
Simulation complete.  Reading coda files...
Coda files loaded successfully
```

Finalmente, extraímos a amostra da posterior conjunta de β_0 e β_1

```
> beta0.post <- as.numeric(modfit$mcmc[[1]][,1])
> beta1.post <- as.numeric(modfit$mcmc[[1]][,2])
```

O código abaixo gera a figura 34 que contém histogramas das amostras das posteriores de β_0 e β_1 , um diagrama de dispersão para elas, e também um histograma para $T-50 = -\beta_0/\beta_1$.

```
par(mfrow=c(2,2))
hist(beta0.post, nclass=50)
hist(beta1.post, nclass=50)
plot(beta0.post, beta1.post)
hist(-beta0.post/beta1.post, nclass=50)
```

Vemos que a distribuição marginal de β_0 tem leve assimetria negativa enquanto que a de β_1 tem assimetria positiva. Vemos também que há uma correlação negativa entre β_0 e β_1 . Já a amostra da distribuição posterior de $T-50$ é relativamente simétrica e compreende valores de 6.5 a 8.5. Segue um sumário dessa amostra:

```
> t50 <- -beta0.post/beta1.post
> c(media=mean(t50), dp=sd(t50), quantile(t50, c(.025, 0.5, 0.975)))
  media      dp    2.5%    50%   97.5%
7.4503467 0.2035256 7.0505140 7.4484012 7.8361016
```

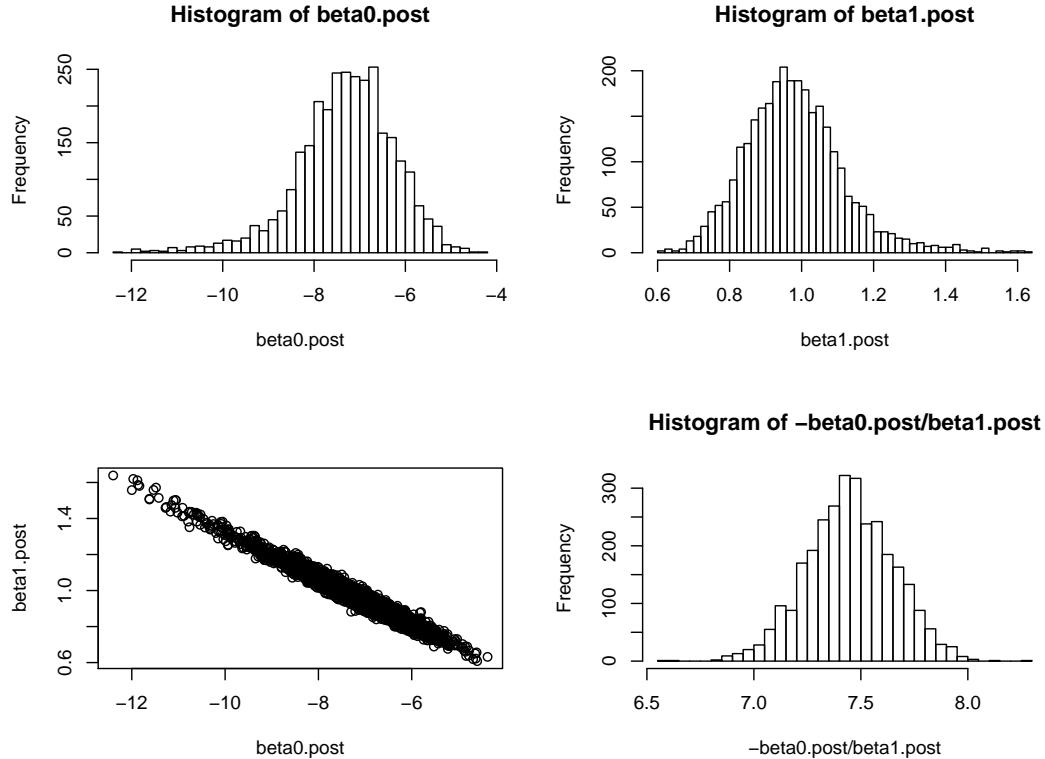


Figura 34:

Os valores da média e mediana são similares e próximos de 7.45 horas. O intervalo de credibilidade (95%) compreende valores de 7.0 a 7.8 aproximadamente. Esta seria uma estimativa do tempo necessário para a formação das asas dos embriões após a deposição dos ovos.

Exercício 8.2.

Vamos assumir então que a i -ésima observação do número de sementes (y_i) provém de uma distribuição de poisson $y_i \sim Pois(\mu_i)$. O logarítmico da média μ_i é função linear da intensidade de vento v_i na forma:

$$\log(\mu_i) = \beta_0 + \beta_1 \cdot v_i$$

Outras parametrizações e modelos poderiam ser também avaliados como por exemplo um em que a função de ligação não é logarítmica. O modelo acima é utilizado a título de ilustração.

Começamos com a entrada dos dados:

```
> v <- c(10,13,5,2,2)
> y <- c(4,11,7,5,6)
```

Vamos obter a amostra da distribuição posterior conjunta de β_0 e β_1 usando o procedimento *MCMC*. Portanto, precisamos de valores iniciais para a cadeia de Markov. Estes valores podem ser estimativas obtidas com o ajuste de um modelo de regressão com o uso da função `glm()`.

```
> sementes.glm <- glm(y~v, family=poisson(link=log))
> sementes.glm
```

```
Call: glm(formula = y ~ v, family = poisson(link = log))
```

```
Coefficients:
(Intercept)          v
1.58814        0.04375
```

```
Degrees of Freedom: 4 Total (i.e. Null); 3 Residual
```

```
Null Deviance: 4.136
Residual Deviance: 2.867 AIC: 25.33
```

Os coeficientes de aproximadamente 1.59 e 0.04 os valores que usaremos como iniciais no *MCMC*. Vamos agora carregar a biblioteca `runjags` e, definir objetos com os dados, os parâmetros a serem monitorados e, os valores iniciais para a cadeia de Markov.

```
> library(runjags)
Loading required package: coda
Loading required package: lattice
> datalist <- dump.format(list(sementes=y, vento=v))
> params <- c("beta0", "beta1")
> inicial <- dump.format(list(beta0=coef(sementes.glm)[1],
+                                beta1=coef(sementes.glm)[2]))
```

O modelo a ser utilizado é:

```
> sem.mcmc <- "model{
+ beta0 ~ dnorm(0, 0.001)
+ beta1 ~ dnorm(0, 0.001)
+ for(i in 1:length(sementes)){
+   lmu[i] <- beta0+beta1*vento[i]
+   mu[i] <- exp(lmu[i])
+   sementes[i] ~ dpois(mu[i])
+ }}"
+ }"
```

Usamos agora a função `run.jags()` para aplicar o procedimento *MCMC*:

```
> semfit <- run.jags(model=sem.mcmc, monitor=params, data=datalist,
+                      inits=inicial, n.chains=1, burnin=50000, thin=3,
+                      sample=9000, check.conv=FALSE)
Calling the simulation... (this may take some time)
Welcome to JAGS 1.0.3 on Wed Jan 6 14:16:10 2010
JAGS is free software and comes with ABSOLUTELY NO WARRANTY
Loading module: basemod
Loading module: bugs
. <>. <Reading data file data.txt
>. <>Compiling model graph
  Resolving undeclared variables
  Allocating nodes
  Graph Size: 26
. <Reading parameter file inits1.txt
>>. . <>Updating 50000
-----| 50000
*****| 100%
. <>. <>. <>Updating 9000
-----| 9000
*****| 100%
. <>. <><>.
Simulation complete. Reading coda files...
Coda files loaded successfully
```

Extraímos a amostra da distribuição posterior conjunta de β_0 e β_1

```
> beta0.sem <- as.numeric(semfit$mcmc[[1]][,1])
> beta1.sem <- as.numeric(semfit$mcmc[[1]][,2])
```

e, com o código abaixo

```
> par(mfrow=c(1, 2))
> hist(beta0.sem, nclass=50)
> hist(beta1.sem, nclass=50)
> plot(beta0.sem, beta1.sem)
```

a mostramos na forma de histogramas e de um diagrama de dispersão (Figura 35). Vemos que as distribuições marginais são aproximadamente simétricas e que há uma correlação negativa entre β_0 e β_1 .

Vamos agora fazer previsões quanto ao número de sementes para um dia em que a intensidade média do vento é de 8 nós. Primeiramente usamos a amostra da posterior ($n = 3000$) para obter valores de $\log(\mu)$ e de μ (`mu.prev` no código abaixo):

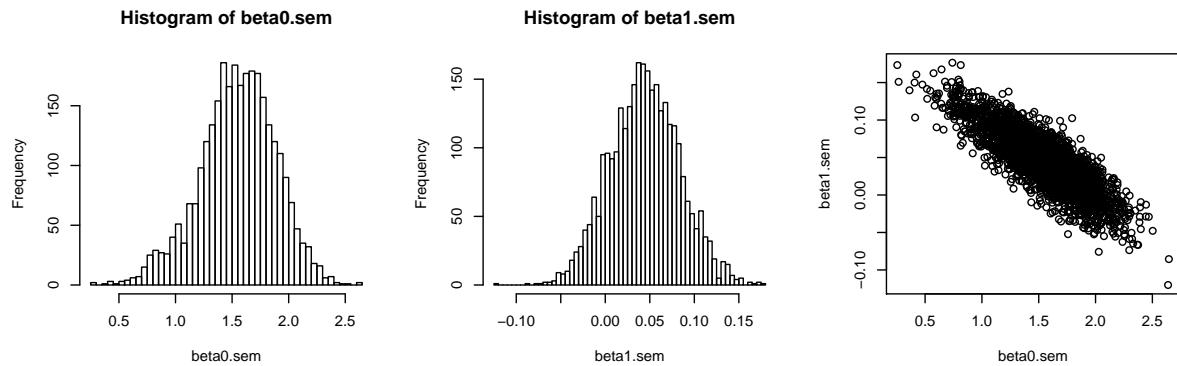


Figura 35:

```
> lmu.prev <- beta0.sem+beta1.sem*8
> mu.prev <- exp(lmu.prev)
```

Agora podemos obter facilmente uma amostra do número de sementes usando uma função para gerar número aleatórios de uma distribuição de poisson com média μ :

```
> sem8.prev <- rpois(rep(1,length(mu.prev)),mu.prev)
```

O código para gerar o histograma da distribuição preditiva (Figura 36) é mostrado abaixo.

```
> par(mfrow=c(1,1))
> hist(sem8.prev)
```

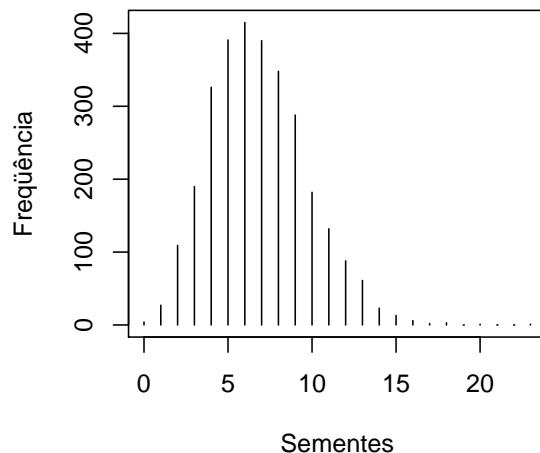


Figura 36:

O sumário para esta distribuição é:

```
> c(media=mean(sem8.prev), dp=sd(sem8.prev),
+ quantile(sem8.prev,c(.025,0.5,0.975)))
media      dp    2.5%      50%    97.5%
6.845333 2.880656 2.000000 7.000000 13.000000
```

Vemos que a distribuição tem assimetria positiva, com média de 6.84 e mediana de 7.00. Já o intervalo de credibilidade (95%) para a previsão do número de sementes para um vento com intensidade média de 8 nós, é de duas a

treze sementes.

Exercício 8.3.

a.

Primeiramente vamos entrar com os dados de comprimento (y) e de idade (t):

```
> x <- c(1,1,1,2,2,3,3,4,6,6,9)
> y <- c(38.9,36.1,39.4,46.2,56.5,61.8,74.0,67.4,96.6,68.2,87.5)
```

Para a construção do modelo, vamos assumir que o comprimento do i -ésimo peixe (y_i) segue uma distribuição normal $y_i \sim N(\mu_i, \sigma)$. A média dessa distribuição é descrita por um modelo de von Bertalanffy tal que:

$$\mu_i = L_\infty [1 - e^{-k(x_i - t_0)}]$$

em que x_i é a idade do i -ésimo peixe e, L_∞ , k e t_0 são parâmetros. Como os parâmetros L_∞ , k e σ são estritamente positivos é conveniente lidar com os mesmos na escala logarítmica. Já o parâmetro t_0 pode assumir valores na escala dos reais. Assim o vetor com os parâmetros de interesse seria $\theta = \{\log L_\infty, \log k, t_0, \log \sigma\}$. Note que somente o terceiro elemento do vetor está expresso na escala original.

No capítulo que trata de modelos não-lineares vimos que neste caso, menos logarítmico da distrição posterior é:

$$-\log p(\theta|y) \propto \log \sigma - \sum_{i=1}^n \log \left\{ \frac{1}{\sqrt{2\pi} \cdot \sigma} \exp \left[-\frac{1}{2} \left(\frac{y_i - \mu_i}{\sigma} \right)^2 \right] \right\}$$

Esta é a função que deve ser minimizada, que podemos escrever como:

```
> neglogpost <- function(tetat) {
+   teta <- c(exp(tetat[1:2]), tetat[3], exp(tetat[4]))
+   mui <- teta[1] * (1-exp(-teta[2] * (x-teta[3])))
+   log(teta[4]) - sum(dnorm(y, mui, teta[4], log=T))
+ }
```

O elemento `tetat` é o vetor com 4 componentes com três deles na escala logarítmica. Já `teta` é o vetor com todos os parâmetros na escala original. O `teta[4]` é o σ da distribuição $y_i \sim N(\mu_i, \sigma)$ enquanto que os outros três elementos são os parâmetros do modelo de von Bertalanffy.

Definida a função a ser otimizada, resta entrar com parâmetros iniciais e proceder com a função `optim()`. Porém, antes de usar a função de otimização, é importante verificar se os valores iniciais são razoáveis e, estão, pelo menos em princípio, próximos da solução. Isto facilita a convergência da iteração na estimativa dos parâmetros. Para verificar se os valores iniciais são razoáveis, basta fazer um gráfico (Figura 37) com os dados observados e as projeções do modelo com os valores iniciais (`teta.ini`):

```
> tetat.ini <- c(log(100), log(1), 0, log(5))
> teta.ini <- c(exp(tetat.ini[1:2]), tetat.ini[3], exp(tetat.ini[4]))
> plot(x, y)
> lines(x, teta.ini[1] * (1-exp(-teta.ini[2] * (x-teta.ini[3]))), type="l")
```

Atente para as transformações utilizadas no código acima. Ao final os valores iniciais para L_∞ , k , t_0 e σ são 100, 1, 0 e 5. Vemos que estes valores iniciais não são muito bons. A assíntota é muito superior e a curvatura parece ser acentuada demais. Vamos readjustar os valores iniciais e refazer o gráfico (Figura 38):

```
> tetat.ini <- c(log(90), log(0.5), 0, log(5))
> teta.ini <- c(exp(tetat.ini[1:2]), tetat.ini[3], exp(tetat.ini[4]))
> plot(x, y)
> lines(x, teta.ini[1] * (1-exp(-teta.ini[2] * (x-teta.ini[3]))), type="l")
```

Vemos que temos valores iniciais bem mais razoáveis no sentido de que devem ser relativamente próximos da solução que deverá ser encontrada por iteração. O entendimento do contexto do problema (crescimento de peixes) e dos componentes matemáticos do modelo facilitam bastante a tarefa de encontrar bons valores iniciais.

Vamos então ao uso da função de otimização para obter estimativas pontuais dos parâmetros e também a matriz cuja inversa é a matriz de covariância, que são, respectivamente, os componentes `tetat.est$par` e `tetat.est$hessian` indicados no código abaixo.

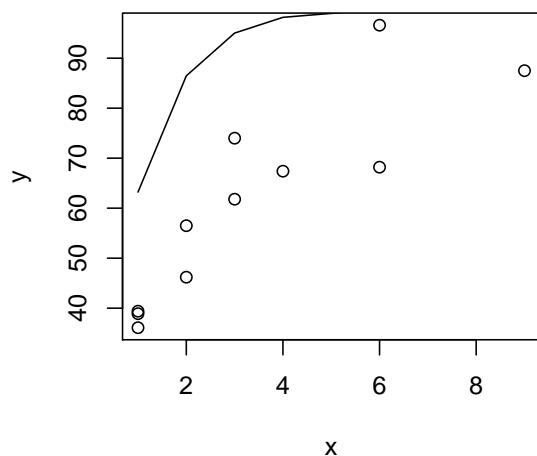


Figura 37:

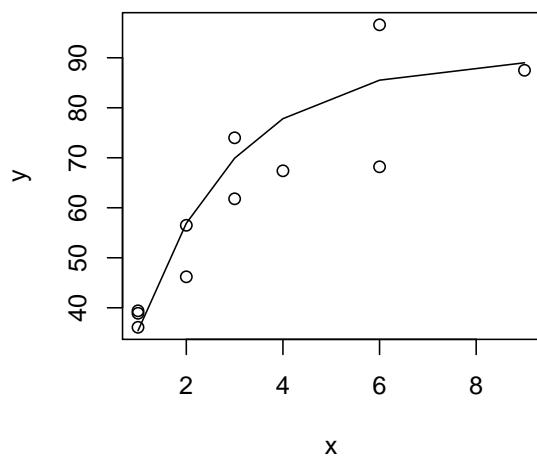


Figura 38:

```

> tetat.est <- optim(par=tetat.ini,fn=neglogpost,hessian=T)
> tetat.est
$par
[1] 4.5131802 -1.0646570 -0.5534917  1.9513263

$value
[1] 39.52411

$counts
function gradient
      241          NA

$convergence

```

```
[1] 0

$message
NULL

$hessian
[,1]          [,2]          [,3]          [,4]
[1,] 8.999655e+02 3.568488e+02 -1.167532e+02 -0.0057962373
[2,] 3.568488e+02 1.769567e+02 -6.869924e+01 -0.0003390506
[3,] -1.167532e+02 -6.869924e+01 3.215952e+01 -0.0010484777
[4,] -5.796237e-03 -3.390506e-04 -1.048478e-03 23.9995055686
```

A visualização do ajuste é muitas vezes desnecessária mas, temos visto que na áreas da biologia, da ecologia e outras do grupo de ciências naturais, isto é bastante comum. Portanto, segue abaixo o código e a figura com o ajuste que obtivemos (Figura 39).

```
> teta.est <- c(exp(tetat.est$par[1:2]), tetat.est$par[3],
+                  exp(tetat.est$par[4]))
> plot(x,y)
> lines(x,teta.est[1]*
+        (1-exp(-teta.est[2] * (x-teta.est[3]))), type="l")
```

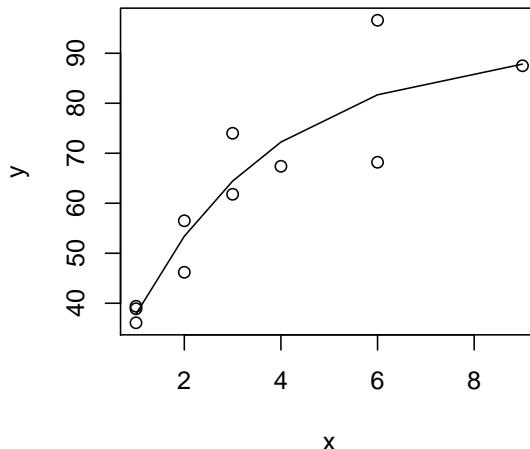


Figura 39:

Agora temos todos os componentes necessários para prosseguir com a análise bayesiana pois podemos facilmente extrair as estimativas pontuais dos parâmetros e a matriz de covariância. Tais componentes são armazenados, respectivamente, nos objetos `est` e `cova` indicados no código abaixo:

```
> est <- tetat.est$par
> cova <- solve(tetat.est$hessian)
```

Retiramos então uma amostra de 3000 vetores da distribuição posterior conjunta, com o auxílio da função `mvtnorm()` da biblioteca MASS

```
> library(MASS)
> tetat.post <- mvtnorm(3000, est, cova)
```

Para simplificar os códigos que seguem obtemos uma versão da mesma com todos os parâmetros na escala original:

```
> teta.post <- cbind(exp(tetat.post[,1:2]), tetat.post[,3],
+                      exp(tetat.post[,4]))
```

Podemos agora explorar essa amostra na forma de gráficos e de sumários estatísticos como por exemplo os mostrados na (Figura 40) e no código abaixo.

```
> pairs(teta.post,
+        labels=c(expression(L[infinity]), "k",
+                  expression(t[0]), expression(sigma)))
> summary(teta.post)
      V1              V2              V3              V4
Min.   : 61.56   Min.   :0.07299   Min.   :-2.6222   Min.   : 3.250
1st Qu.: 85.24   1st Qu.:0.25800   1st Qu.:-0.9588   1st Qu.: 6.143
Median : 91.32   Median :0.34538   Median :-0.5432   Median : 7.040
Mean   : 91.85   Mean   :0.37678   Mean   :-0.5505   Mean   : 7.174
3rd Qu.: 98.39   3rd Qu.:0.45907   3rd Qu.:-0.1454   3rd Qu.: 8.023
Max.   :134.40   Max.   :1.93712   Max.   : 1.7516   Max.   :15.907
```

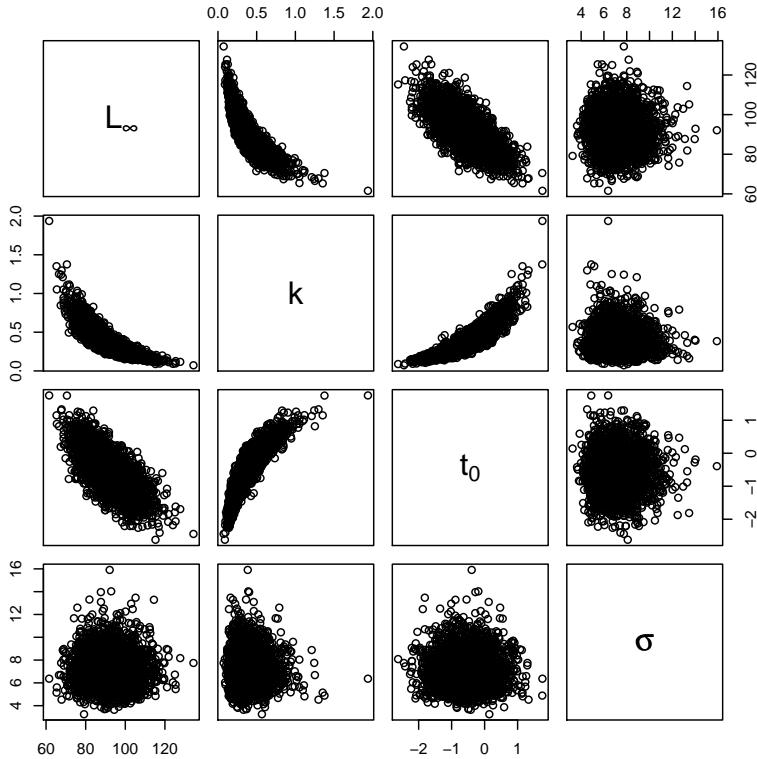


Figura 40:

Vemos que há correlações negativas entre L_∞ e k e entre L_∞ e t_0 . Vemos também que a correlação entre k e t_0 é positiva e, que as correlações que envolvem k e os demais parâmetros de von Bertanlanffy, são não lineares. Se desejado, inferências sobre estes parâmetros e também sobre σ poderiam ser feitas de maneira relativamente simples a partir da exploração da amostra da posterior. Consta no código abaixo um sumário das distribuições marginais posteriores de L_∞ , k , t_0 e σ , nesta ordem.

```
> for(i in 1:4){
+   print(c(media=mean(teta.post[,i]), dp=sd(teta.post[,i]),
+          quantile(teta.post[,i], prob=c(0.025,0.5,0.975))))
+ }
media      dp      2.5%      50%     97.5%
91.851542  9.551371  74.749174  91.317937 111.427999
media      dp      2.5%      50%     97.5%
0.3767813 0.1658680 0.1559206 0.3453829 0.7872439
media      dp      2.5%      50%     97.5%
-0.5505161 0.6044938 -1.7360363 -0.5432193  0.6482596
```

media	dp	2.5%	50%	97.5%
7.174479	1.467332	4.732899	7.040314	10.460330

b.

A distribuição preditiva $\tilde{y} \sim N(\mu^{(t=2.5)}, \sigma)$ em que $\mu^{(t=2.5)} = L_\infty[1 - e^{-k(2.5-t_0)}]$. Portanto, uma amostra da distribuição preditiva seria:

```
> mu.pred <- teta.post[,1] *
+   (1-exp(-teta.post[,2]*(2.5-teta.post[,3])))
> y.til <- rnorm(length(mu.pred), mu.pred, teta.post[,4])
```

O código abaixo gera um sumário estatístico e um histograma (Figura 41) da distribuição preditiva.

```
> c(media=mean(y.til), dp=sd(y.til),
+   quantile(y.til, prob=c(0.025, 0.5, 0.975)))
media      dp    2.5%    50%    97.5%
58.351098 7.993262 42.591246 58.468900 73.742171
> hist(y.til, right=FALSE)
```

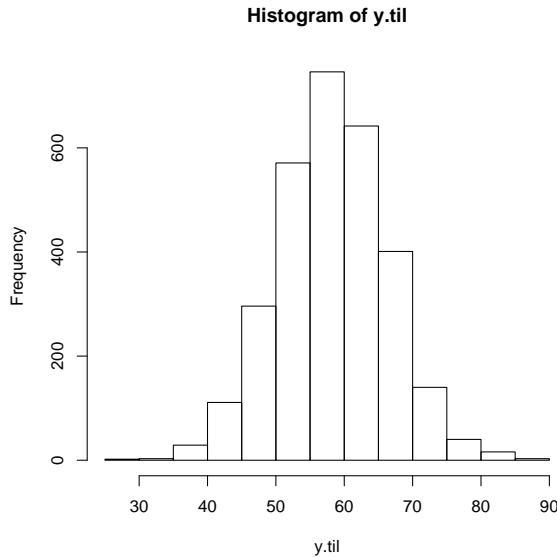


Figura 41:

Vemos que o intervalo de credibilidade (95%) para a predição do comprimento de um peixe com idade de 2.5 anos compreende os valores entre 43 e 74 centímetros aproximadamente.

c.

Como explicado no capítulo que trata de modelos não-lineares queremos os j -ésimos valores $\tilde{x}^{(j)}$ que minimizam a função $-\log p(y_p | \theta^{(j)}, \tilde{x}^{(j)})$ em que $\theta^{(j)}$ corresponde aos j -ésimos valores com $j = 1, \dots, m$ da amostra da distribuição posterior conjunta dos parâmetros do modelo. Os valores de $\theta^{(j)}$ definem o *sinal* e o *ruído* do modelo utilizado.

O *ruído* do modelo utilizado neste exercício é normal e, portanto

$$-\log p(y_p | \theta^{(j)}, \tilde{x}^{(j)}) \propto \frac{1}{2} \left(\frac{y_p - \mu^{(j)}}{\sigma^{(j)}} \right)^2$$

Esta função é minimizada quando $y_p = \mu^{(j)}$, ou seja

$$y_p = \mu^{(j)} = L_\infty^{(j)} [1 - e^{-k^{(j)} (\tilde{x}^{(j)} - t_0^{(j)})}]$$

Para encontrar a j -ésima predição reversa basta então resolver esta equação para $\tilde{x}^{(j)}$. A solução é:

$$\tilde{x}^{(j)} = t_0^{(j)} - \frac{1}{k^{(t)}} \log \left(1 - \frac{y_p}{L_\infty^{(j)}} \right)$$

Portanto, a previsão reversa da idade para $y_p = 45$ cm é:

```
> x.til <- teta.post[,3]-1/teta.post[,2]*log((1-45/teta.post[,1]))
```

Mostramos essa distribuição preditiva na forma de um sumário estatístico e de um histograma (Figura ??).

```
> c(media=mean(x.til),dp=sd(x.til),
+     quantile(x.til,prob=c(0.025,0.5,0.975)))
media      dp    2.5%      50%    97.5%
1.5061303 0.2245751 1.1050054 1.4944673 1.9900609
> hist(x.til,right=FALSE,nclass=50)
```

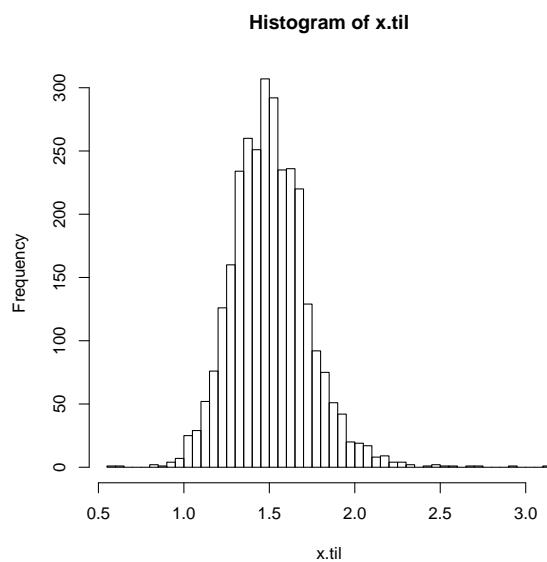


Figura 42: