Questão 2.

a) **Solução:**

Algoritmo:

Passo 1) Gere uma observação do processo de poisson homogêneo, $N(t)$, com o tempo variando até 1 hora e com taxa 5 chegadas por hora. No caso, gere um valor de $N(t) \sim poisson(5t)$ com t = 1.

Passo 2) Gere uma amostra de tamanho $N(t)$ de uma uniforme discreta, X, no espaço $\{20, 21, \cdots, 40\}$ de modo a obter um vetor $\mathbf{X} = \{X_1, X_2, \cdots, X_{N(t)}\}$

Passo 3) Faça o número de torcedores que chegam ao evento,$N_{\text{Torcedores}}$, igual a soma dos elementos de $\mathbf{X}$, ou seja, $N_{\text{Torcedores}} = \sum_{k=1}^{N(t)} X_k$.

b) **Solução** O número de torcedores que chega ao evento é dado por,

$$T = \sum_{j=1}^{N(t)} X_j$$

em que $X_j \sim U\{20, 21, \cdots, 40\}$, cujo a média é $E[X_j] = \mu = \frac{40+20}{2} = 30$, e $N(t) \sim PPH$ com taxa de 5 ônibus por hora com o tempo variando até 1 hora. Ou seja, $N(1) \sim Poisson(5)$. Assim,

$$\begin{aligned} E(T|N(1) = n) &= E\Big(\sum_{j=1}^{N(t)} X_j | N(1) = n\Big) \\ &= E\Big(\sum_{j=1}^{n} X_j\Big) = \sum_{j=1}^{n} E\big(X_j\big) \\ &= \sum_{j=1}^{n} \mu = n\mu. \end{aligned}$$

Portanto $E(T|N(1) = n) = N(1)\mu$ e o valor esperado do número de torcedores é, $E(T) = E(E(T|N(1) = n)) = E(N(1)\mu) = \mu E(N(1)) = 30 * 5 = 150$.

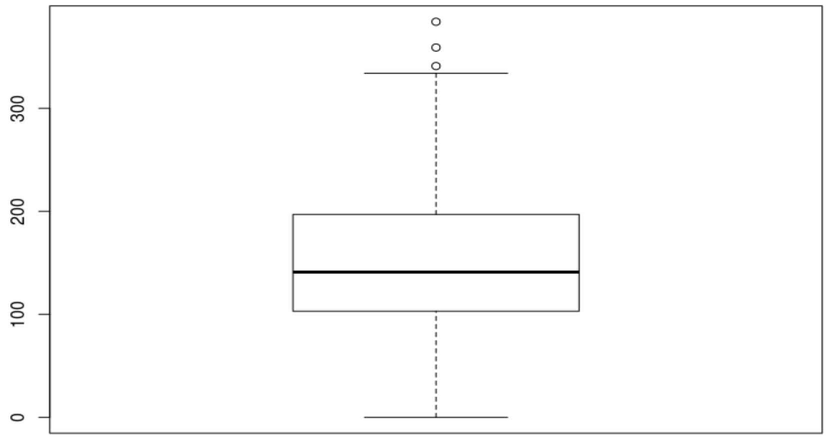
Na tabela temos as estatísticas resumo para uma amostra de tamanho 500.

Min	1º Q	Mediana	Média	3º Q	Max
0.0	103.0	141.0	149.5	197.0	384.0

Box-plot dos dados:

Código:

```
#Maneira 1
 t <- 1
 lambda <- 5
 N_Torced1 <- numeric(500)
 for( i in 1:500){
           Nt <- rpois(1,lambda*t)
           X <- sample(20:40,Nt,replace = T)
           N_Torced1[i] <- sum(X)
 }
```




```
mean(N_Torcedores);  var(N_Torcedores)


# Maneira 2
N_Torced2 <-  replicate(500,sum(sample(20:40,rpois(1,lambda*t),replace = T)))
mean(N_Torced2);  var(N_Torced2)
```

c) **Solução**

Sabendo que a variância de X é $\frac{(40-20+1)^2-1}{12} = \frac{21^2-1}{12}$ e que a variância de $T|N(1)$ é

$$
\begin{aligned}
Var(T|N(1)=n) &= Var(\sum_{j=1}^{N(1)} X_j|N(1)=n) \\
&= Var(\sum_{j=1}^{n} X_j), \text{ por independência} \\
&= \sum_{j=1}^{n} Var(X_j) \\
&= n\frac{21^2-1}{12}
\end{aligned}
$$

A variância do número de torcedores que chegam ao evento, T, é dada por,

$$
\begin{aligned}
V(T) &= Var(E(T|N(1)=n)) + E(Var(T|N(1)=n)) \\
&= Var(30N(1)) + E(N(1)\frac{21^2-1}{12}) \\
&= 30^2 * 5 + \frac{21^2-1}{12} * 5 = 4683.333
\end{aligned}
$$

O valor obtido com a simulação foi 4688.326 , que é um valor próximo do valor real.

Questão 3. **Solução:**

O algoritmo para gerar o PPNH é

i) t = 0, I = 0

ii) Gere um número aleatório U.

iii) Faça $t = t - \frac{1}{\lambda} log(U)$.

iv) Gere um número aleatório U.

v) Se $U \leq \frac{\lambda(t)}{\lambda}$, faça $I = I + 1, S(I) = t$.

vi) Retorne ao passo ii) até que I = 10.

em que $\lambda(t)$ é a função de intensidade, λ é um valor tal que $\lambda(t) \leq \lambda$, I é o número de eventos e $S(1), \cdots, S(I)$ são os tempos dos eventos. Para esta questão temos que, $\lambda(t) = 3 + \frac{4}{t+1}$ e $\lambda = 7$.

Código:

```
t <- 0 ; I <- 0
lambda <- 7
S <- numeric(10)

while(I < 10) {
  U <- runif(2)
  t = t - log(U[1])/lambda
  if(U[2] <= (3 + 4/(t + 1))/lambda){
    I = I + 1 ; S[I] = t
  }
}
```

Evento(I)	1	2	3	4	5	6	7	8	9	10
Tempo de ocorrência (S)	0.05	0.10	0.22	0.31	0.91	0.93	1.10	1.77	1.90	2.23

Questão 6.

a) Explique como usar o método de bootstrap para estimar p.

Solução: Primeiro encontraremos a distribuição de $\delta = \bar{X} - \mu$ usando a aproximação bootstrap e para isso procedemos da seguinte forma:

i) Calcule a média dos dados, $\bar{x}$, e a considere como uma estimativa para μ.

ii) Calcule o vetor de médias $\bar{\mathbf{x}}^*$ que contém a média de cada uma das B amostras bootstrap.

iii) Calcule a diferença $\boldsymbol{\delta}^* = \bar{\mathbf{x}}^* - \bar{x}$.

A partir da distribuição bootstrap, representada pela amostra $\boldsymbol{\delta}^*$, podemos calcular a estimativa de p como sendo: $\hat{p} = \frac{\text{Cardinalidade } \left(a < \boldsymbol{\delta}^* < b\right)}{B}$.

b) Considere $n = 10$,$a = -5$ e $b = 5$.Estime p considerando que os valores observados são : 56,101, 78, 67, 93, 87,64,72,80 e 69.

Solução: A estimativa para $P(-5 < \bar{X} - \mu < 5) \approx P(-5 < \bar{X}^* - \bar{X} < 5) = 0.7638$ para $B = 10^4$.

Código:

```
dados <-  c(56,101, 78, 67, 93, 87,64,72,80,69)
B <- 10^4 # quantidade de amostra bootstrap
xbar <- mean(dados)
xbar_ast <- replicate(B, mean(sample(dados,10,replace = T)))
delta_ast <- xbar_ast - xbar
P <- length(delta_ast[delta_ast > -5 & delta_ast < 5])/B
```

Questão 7. Usando a propriedade da variância da soma obtemos,

$$Var[\alpha X + (1-\alpha)Y] = \alpha^2 Var[X] + (1-\alpha)^2 Var[Y] + 2\alpha(1-\alpha)Cov[X,Y].$$

Derivando com relação a α,

$$\frac{\partial}{\partial\alpha}Var[\alpha X + (1-\alpha)Y] = 2\alpha Var[X] - 2(1-\alpha)Var[Y] + 2(1-2\alpha)Cov[X,Y].$$

Igualando a zero e isolando α,

$$\begin{aligned}
2\alpha Var[X] - 2(1-\alpha)Var[Y] + 2(1-2\alpha)Cov[X,Y] &= 0 \\
\alpha(Var[X] + Var[Y] - 2Cov[XY]) - Var[Y] + Cov[X,Y] &= 0 \\
\alpha(Var[X] + Var[Y] - 2Cov[X,Y]) &= Var[Y] - Cov[X,Y] \\
\alpha &= \frac{Var[Y] - Cov[X,Y]}{Var[X] + Var[Y] - 2Cov[X,Y]}.
\end{aligned}$$

A segunda derivada é igual a $Var[X] + Var[Y] - 2Cov(X,Y) = Var[X-Y] > 0$. Portanto, trata-se de um ponto de mínimo.

Questão 8.

a) Na tabela abaixo apresentamos as estimativas dos parâmetros da regressão linear simples.

	Estimativa	Erro padrão	Valor t	Pr(>|t|)
Intercepto	0.3794	0.6127	0.62	0.5465
LSTA	0.0045	0.0010	4.44	0.0007

b) Sabendo que $\hat{\sigma}^2 = QMRes = \frac{\sum_{i=1}^{n}(y_i - (\hat{\beta}_0 + \hat{\beta}_1 \times LSTA_i))^2}{n-2}$, com $n = 15$. O erro padrão obtido via bootstrap é 0.00085 e o erro padrão obtido pela expressão $ep(\hat{\beta}) = \sqrt{\frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}}$ é 0.0010. Portanto, o erro-padrão obtido via bootstrap é menor.

```
 LSTA<-c(576,635,558,578,666,580,555,661,651,605,653,575,545,572,594)

GPA <- c(3.39,3.30,2.81,3.03,3.44,3.07,3.00,3.43,3.36,3.13,3.12,2.74,2.76,2.88,2.96)

Dados <- data.frame(GPA,LSTA)

ajuste <- lm(GPA ~ LSTA)

summary(ajuste)

B <- 100000


boot <- function(dados){
                  ib <- sample(1:15,15, replace = T)
                  y <- dados[ib,1]
                  x <- dados[ib,2]
                  beta1_b <- cov(y,x)/var(x) # sxy/sxx
                  return(beta1_b)
}


beta1_B <- replicate(B,boot(Dados))

sd(beta1_B) # estimatica ep
```