Quarta Lista de Exercícios

Exercício 1. A Tabela 1 contém dados sobre a gravidade das lesões sofridas durante acidentes de carro.

		Nível de ferimento			
		nenhum	leve	moderado	grave
uso de sinto de segurança	sim	12813	647	359	42
	não	65963	4000	2642	303

Tabela 1: Dados referentes ao Exercício 1.

(a) A matriz da Figura 1 foi obtida no R a partir dos seguintes comandos:

```
A <- matrix(c(20, 30, 60, 75),ncol = 2,byrow = TRUE)
row.names(A) <- c("um", "dois")
colnames(A) <- c("alpha", "beta")
```

Recicle o código acima para criar uma matriz 2×4 em que suas entradas são os valores numéricos da Tabela 1. Essa matriz deverá ser guardada dentro da variável `dados`. As linhas de `dados`, nessa ordem, deverão se chamar `sim` e `não`. As colunas de `dados` deverão se chamar, nessa ordem, `nenhum`, `leve`, `moderado` e `grave`.

```
> A
     alpha beta
um      20   30
dois    60   75
```

Figura 1: Matriz referente ao código do Exercício 2a.

(b) Realize um teste de hipóteses para verificar se as variáveis **uso de sinto de segurança** e **nível de ferimento** são independentes. Faça o teste considerando que $\alpha = 1\%$. Deixe no código, como comentário, a sua explicação sobre a conclusão do teste.

Exercício 2. O objetivo dessa questão é criar uma função que retornará o p-valor de um teste de hipóteses unilateral à direita para a média de uma população com variância desconhecida. Para isso, vamos relembrar a teoria desse tipo de teste.

Considere a população P_X. Dessa população, extraímos uma amostra aleatória $X_1, X_2, \ldots, X_n$:

$$X_i \;=\; \mathcal{N}(\mu_X, \sigma^2),\ i = 1, 2, \ldots, n.$$

A variância σ^2 é desconhecida. Nosso interesse é testar:

$$H_0 \;:\; \mu_X = \mu_0 \quad \text{versus} \quad H_a \;:\; \mu_X \; > \mu_0$$

A estatística de teste T sob a hipótese nula, nesse caso, será dada por:

$$T = \frac{\bar{X} - \mu_0}{\sqrt{S^2/n}} \sim t_{n-1},$$

em que

$$S^2 \;\; = \;\; \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

Se a hipótese alternativa for da forma unilateral à direita ($H_a \;:\; \mu_X > \mu_0$), então o p-valor será dado por:

$$p\_valor = P(t_{n-1} > t_{obs}),$$

em que t_{n-1} representa uma distribuição t de Student com $n-1$ graus de liberdade e t_{obs} é o valor observado a partir da amostra da estatística de teste T.

(a) Construa uma função chamada `p_valor` cuja entrada é dada pelos argumentos `x, m`. O argumento x é um vetor de tamanho n e o argumento m representa a média μ_0 a ser testada. A saída dessa função deverá ser a frase:

`O p-valor do teste é:` $\star$

em que $\star$ representa o p-valor que a função calculou.

(b) O arquivo `amostraX.txt` apresenta uma amostra aleatória de uma população X. Realize um teste de hipóteses para verificar se esta amostra foi retirada de uma população Normal. Comente o resultado encontrado. Por fim, se a amostra foi retirada de uma população Normal, utilize a função criada em (a) para testar:

$$H_0 : \; \mu_X = 20 \quad \text{versus} \quad H_a : \; \mu_X > 20,$$

considerando que o nível de significância seja de 5%.

Exercício 3. O conjunto de dados `femur.csv` foi retirado do artigo *Height Estimation from Skeletal Remains* publicado em dezembro de 2008. No artigo, tenta-se estimar a altura de um indivíduo do período medieval a partir do tamanho de seu fêmur. Resolva as questões a seguir considerando o nível de significância como 5%.

(a) Realize um teste de hipóteses para verificar se a altura dos homens segue uma distribuição Normal. Em seguida, realize també, um teste de hipóteses para verificar se a altura das mulheres segue uma distribuição Normal.

(b) Se as alturas dos homens e das mulheres segue uma distribuição Normal, realize um teste de hipóteses para verificar se as variâncias dessas duas populações (homens e mulheres) são iguais.

(c) Se as alturas dos homens e das mulheres segue uma distribuição Normal, realize um teste de hipóteses para verificar se as médias dessas duas populações (homens e mulheres) são iguais.

Informações úteis:

Informação 1. Exemplo de uma função com retorno do tipo texto:

```
g <- function(a){
  if(a < 0){
    return(paste("o valor absoluto do número é:", -a))
  }else{
    return(paste("o valor absoluto do número é:", a))
  }
}
g(-2)
g(10)
```

Informação 2. Seja T uma distribuição t-Student com 10 graus de liberdade. Então a probabilidade de T ser menor ou igual a 2 (isto é, $P(T \leq 2)$) pode ser calculada pela seguinte função no R:

```
> pt(2, df = 10)
[1] 0.963306
```