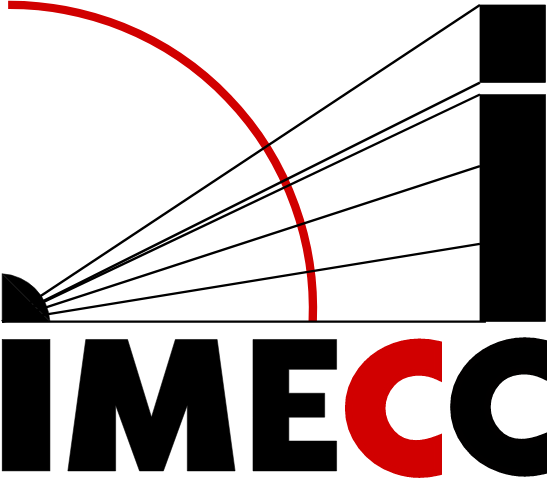
****

**Universidade Estadual de Campinas – UNICAMP**

**Instituto de Matemática, Estatística e Computação Científica- IMECC**

**Atividade Final de ME110**

**Thales Crivilari 236312**

**Parte 1: Mega Sena**

“As chances de você ser atingido por um raio são em torno de uma em 1 milhão, ou seja, você tem 50 vezes mais chance de ser atingido por um raio do que de ganhar na Mega-Sena", frases como essa são repetidas em todos os lugares inclusive em filmes, apesar de ser senso comum ainda existem alguns espíritos céticos que jogam na esperança de uma fortuna escolhendo números aleatórios de 1 a 60 .

**Proposta**

Calcular qual a probabilidade de cada bola ao ser sorteada, para assim , descobrir se todas possuem a mesma probabilidade de cair ou se algumas se sobressaem sobre outras.

Para isso, utilizei o software R, para realização de um teste, sendo que com uma hipótese nula, todos os resultados têm a mesma probabilidade. Enquanto, uma hipótese alternativa, pelo menos um resultado tem a probabilidade diferente das demais.

**TABELA DA ESTATÍSTICA DO TESTE**

Nessa etapa, nosso principal objetivo é achar a Estatística do Teste, através da Tabela da Mega Sena( no final do arquivo).

* Para começar, se conta quantas vezes cada bolinha saiu.
* Como nossa hipótese nula é de que todos os resultados têm a mesma probabilidade, considerando que o número esperado, ou Oj, que cada bolinha deve ter saído é de 248,8 (6 x 2488 / 60).
* Calculamos o módulo da subtração entre o Valor Observado e o Valor Esperado de cada bolinha. Depois, somamos todos os módulos obtidos, resultando na estatística do nosso teste, que chamaremos de W (W = ⅀ |Oj - Ej|).

| Valor observado da estatística do teste |
| --- |
| 847,2 |

|  | observados | Esperado | | Oj -Ej| |
| --- | --- | --- | --- |
| 1 | 241 | 248,8 | 7,8 |
| 2 | 255 | 248,8 | 6,2 |
| 3 | 228 | 248,8 | 20,8 |
| 4 | 265 | 248,8 | 16,2 |
| 5 | 272 | 248,8 | 23,2 |
| 6 | 250 | 248,8 | 1,2 |
| 7 | 236 | 248,8 | 12,8 |
| 8 | 253 | 248,8 | 4,2 |
| 9 | 234 | 248,8 | 14,8 |
| 10 | 282 | 248,8 | 33,2 |
| 11 | 259 | 248,8 | 10,2 |
| 12 | 240 | 248,8 | 8,8 |
| 13 | 253 | 248,8 | 4,2 |
| 14 | 233 | 248,8 | 15,8 |
| 15 | 219 | 248,8 | 29,8 |
| 16 | 258 | 248,8 | 9,2 |
| 17 | 261 | 248,8 | 12,2 |
| 18 | 244 | 248,8 | 4,8 |
| 19 | 236 | 248,8 | 12,8 |
| 20 | 241 | 248,8 | 7,8 |
| 21 | 214 | 248,8 | 34,8 |
| 22 | 218 | 248,8 | 30,8 |
| 23 | 267 | 248,8 | 18,2 |
| 24 | 256 | 248,8 | 7,2 |
| 25 | 236 | 248,8 | 12,8 |
| 26 | 205 | 248,8 | 43,8 |
| 27 | 262 | 248,8 | 13,2 |
| 28 | 263 | 248,8 | 14,2 |
| 29 | 259 | 248,8 | 10,2 |
| 30 | 264 | 248,8 | 15,2 |
| 31 | 235 | 248,8 | 13,8 |
| 32 | 257 | 248,8 | 8,2 |
| 33 | 268 | 248,8 | 19,2 |
| 34 | 262 | 248,8 | 13,2 |
| 35 | 263 | 248,8 | 14,2 |
| 36 | 257 | 248,8 | 8,2 |
| 37 | 270 | 248,8 | 21,2 |
| 38 | 258 | 248,8 | 9,2 |
| 39 | 235 | 248,8 | 13,8 |
| 40 | 240 | 248,8 | 8,8 |
| 41 | 262 | 248,8 | 13,2 |
| 42 | 273 | 248,8 | 24,2 |
| 43 | 257 | 248,8 | 8,2 |
| 44 | 260 | 248,8 | 11,2 |
| 45 | 240 | 248,8 | 8,8 |
| 46 | 248 | 248,8 | 0,8 |
| 47 | 238 | 248,8 | 10,8 |
| 48 | 229 | 248,8 | 19,8 |
| 49 | 258 | 248,8 | 9,2 |
| 50 | 244 | 248,8 | 4,8 |
| 51 | 258 | 248,8 | 9,2 |
| 52 | 249 | 248,8 | 0,2 |
| 53 | 284 | 248,8 | 35,2 |
| 54 | 264 | 248,8 | 15,2 |
| 55 | 211 | 248,8 | 37,8 |
| 56 | 255 | 248,8 | 6,2 |
| 57 | 237 | 248,8 | 11,8 |
| 58 | 242 | 248,8 | 6,8 |
| 59 | 234 | 248,8 | 14,8 |
| 60 | 236 | 248,8 | 12,8 |

### **SIMULANDO NO R**

Nessa etapa, nosso principal objetivo é simular os resultados da Mega Sena, obter os dados através da Estatística do Teste (essa obtida na etapa anterior) e, assim, chegar a uma conclusão segundo as hipóteses já feitas.

* Primeiro, com o RGui aberto, se coloca t = seq(1:60).
* Depois para simular para 10.000, usamos essas funções:

wstar = c()

esperado = rep(248.8,60)

for( j in 1:10000){

resultados = matrix(0,2488,6)

for( i in 1:2488){

resultados[i,] = sample(x,6)}

observado = table(resultados)

esperado = rep(248.8,60)

* E para achar a estatística do teste eu usei a função:

wstar[j] = sum(abs( observado - esperado))

* Lembrando que “wstar” é a estatística do teste, obtivemos as estatísticas do teste de 10.000 simulações.
* Depois, calculamos o p-valor e escrevemos no R:

pvalor= sum(wstar > 847.2)/10000

* também podemos obter o histograma das estatísticas do teste “Westrela” e a linha da estatística do teste observada (W=847,2).

png("histograma.png")

hist(wstar,col=”a cor desejada”, main= "", ylab ="Frequência")

abline(v = 847.2, col=2)

dev.off()

* Por último, podemos construir um Intervalo de Confiança, considerando 𝛂 = 0,05. Para isso, colocamos os seguintes comandos:

ME = sqrt(pvalor\*(1-pvalor)/10000)

LI = pvalor - 1.96\*ME

LS = pvalor + 1.96\*ME

IC = c(LI,LS)

## 

## 

## 

## 

## 

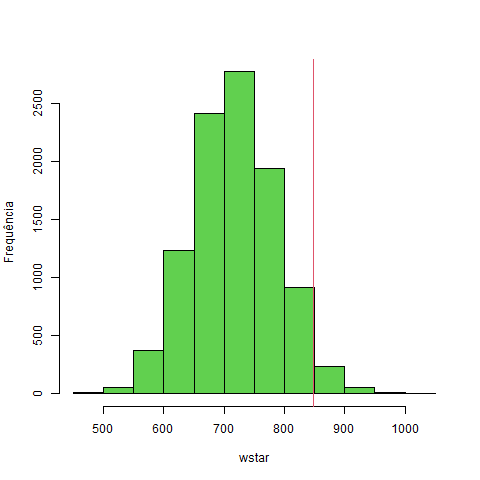
## 

## 

## 

## 

## **Dados chegados no R:**

**** O p-valor obtido:0.0334

intervalo de confiança: 0.02841956 0.03838044

## **CONCLUSÃO**

Usando o histograma acima, podemos concluir que a probabilidade de obter um valor acima de 847,2, depois da linha vermelha, é muito baixa, o que pode ser mostrado no p-valor, que é menor que 5%.Portanto, rejeitamos a hipótese nula de que todos os resultados têm a mesma probabilidade, ficando a favor da hipótese alternativa.

# 

# **Parte 2: Regressão**

Uma tradição muito utilizada em lojas de roupas é a tradição de medir as calças no pescoço, essa crença popular diz que se a largura da calça conseguir envolver o colarinho do pescoço então a calça serve. Mas será que essa crença é verdadeira?

## **OBJETIVO**

Vamos tomar como objetivo verificar se a crença popular apresentada na introdução é verdadeira, utilizando regressão linear, com uma amostra da medida do pescoço e da cintura de 20 pessoas, calcularemos se existe uma associação entre ambas.

### **ETAPA 1 - MEDIR 20 PESSOAS**

1. Os métodos que iremos usaremos para medir são :

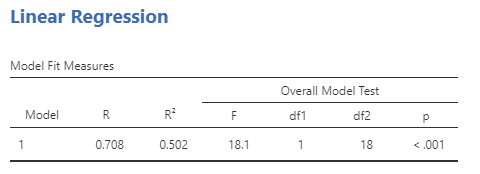
Para medir a circunferência do pescoço, olhe para a frente e relaxe os ombros. Enrole a fita métrica em volta do pescoço logo abaixo do pomo de Adão (osso hióide), paralela ao chão. Para medir sua cintura, fique relaxado com os braços ao lado do corpo. Enrole a fita métrica ao redor da menor parte do seu abdômen, paralela ao chão. Isso geralmente é o meio caminho entre o umbigo e a base do esterno, alinhado com a dobra do cotovelo. Colocar a fita métrica contra a pele e tirar as medidas no final de sua expiração normal e relaxada.

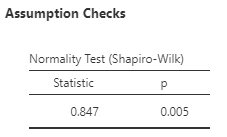
**Dados obtidos**

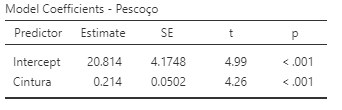
|  | pescoço | cintura |
| --- | --- | --- |
| pessoa1 | 37 | 73 |
| pessoa2 | 40 | 90 |
| pessoa3 | 35 | 67 |
| pessoa4 | 38 | 76 |
| pessoa5 | 37 | 75 |
| pessoa6 | 38,5 | 90 |
| pessoa7 | 40 | 95 |
| pessoa8 | 38 | 82 |
| pessoa9 | 43 | 98 |
| pessoa10 | 37 | 97 |
| pessoa11 | 39 | 86 |
| pessoa12 | 42 | 100 |
| pessoa13 | 37 | 78 |
| pessoa14 | 35 | 70 |
| pessoa15 | 38 | 68 |
| pessoa16 | 37 | 74 |
| pessoa17 | 49 | 97 |
| pessoa18 | 36 | 82 |
| pessoa19 | 38 | 78 |
| pessoa20 | 35 | 74 |

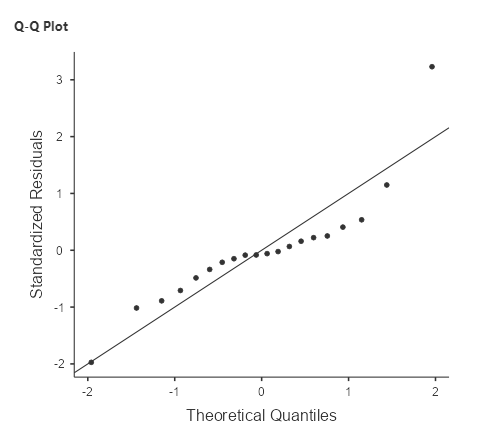
## **DADOS CHEGADOS NO JAMOVI**

Dados obtidos na regressão linear no software Jamovi:







Gráfico da Dispersão

## 

## 

## 

## 

## 

## 

## **CONCLUSÃO**

Analisando, concluímos que o “coeficiente de relação” é 0.708, indicando uma forte correlação entre a medida do pescoço e da cintura. Outro fator é o R², que nos indica qual a porcentagem da variável dependente é afetada pela variável independente, que nesse caso é de 0,502. Esse número nos diz que 50% da variável independente é afetada pela circunferência do pescoço e a outra parte pode ser por diversos fatores como um acidente durante a vida das pessoas. A distribuição normal dos resíduos tem um p-valor de 0.005, isso significa que os valores se distanciam muito um do outro da função do gráfico. Utilizando o teste de variância f, concluímos também que o nosso modelo é significativo por ter o p menos que 0,05.

Assim, podemos concluir com confiança que a crença popular realmente tem uma associação.

**Arquivos e Documentos usados no trabalho:**

PARTE 1: Tabela contendo 2488 jogos da Mega Sena ([MegaSena2488.xlsx](https://moodle.ggte.unicamp.br/pluginfile.php/3066464/mod_assign/introattachment/0/MegaSena2488.xlsx?forcedownload=1)), Software RGui

PARTE 2: https://www.youtube.com/watch?v=DdlDTR0EaNM