

AULA PRÁTICA

R e Python

Naiara Caroline dos Santos

naicaroline2@usp.br

ESTATÍSTICA 1 - ICMC/USP

Professor Jorge Luis Bazán

27 Abril, 2022

1. Instalação do RStudio e Python

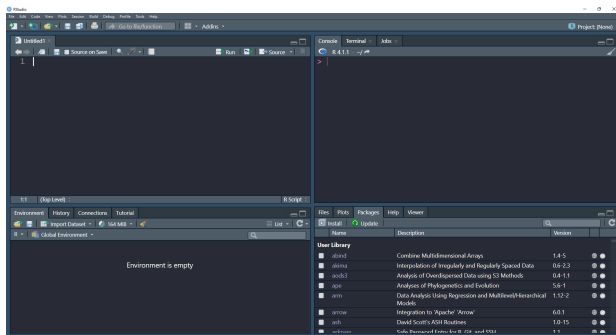
• RStudio

- <https://cran.r-project.org/mirrors.html>;
- <https://www.rstudio.com/products/rstudio/download/>;

• Python

- Vídeo tutorial Khennedy;
- <https://www.python.org/>;
- <https://pandas.pydata.org/>;
- https://colab.research.google.com/?utm_source=scs-index

2. Introdução ao RStudio

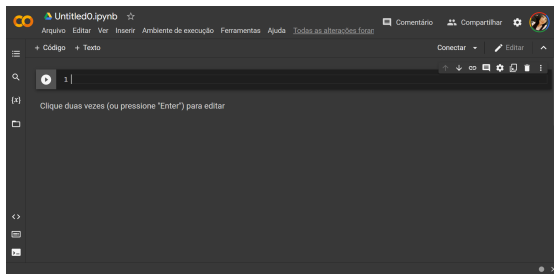


- i **Source:** visualiza e edita *scripts*;
- ii **Environment e History:** carrega objetos na memória do R e guarda linhas de comando que já foram executados;
- iii **Console:** executa funções e *scripts*;
- iv **Files, Plots, Packages, Help e Viewer:** para arquivos, gráficos, pacotes, ajuda e visualizador.

2.1. Distribuições de Probabilidade

- O R inclui funcionalidades para operações com distribuições de probabilidades, em que para cada distribuição há quatro operações básicas indicadas pelas letras:
 - **d** calcula a densidade de probabilidade $f(x)$ no ponto;
 - **p** calcula a função de probabilidade acumulada $F(x)$ no ponto;
 - **q** calcula o quantil correspondente a uma dada probabilidade;
 - **r** retira uma amostra aleatória da distribuição.
- Para usar os funções deve-se combinar uma das letras acima com a abreviatura do nome da distribuição.
Exemplo: `pnorm()` para distribuição Normal, `pexp()` para exponencial, `ppois()` para Poisson, e assim por diante.

3. Introdução ao Google Colab

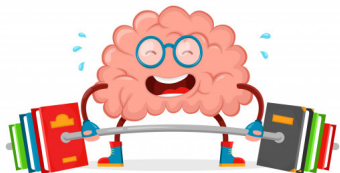


- Linguagem Python
 - Anaconda; Spider; Jupyter; Colab.
- Criando um notebook:
 - Acesse o link
https://colab.research.google.com/?utm_source=scs-index
 - Você deve estar autenticado com sua **conta do Google**.



Vamos praticar?

- 1 Gerar uma amostra aleatória da distribuição Normal Padrão.
 - 1 Alterar valores da média e desvio padrão.
- 2 Construir um histograma.
- 3 Medidas descritivas.



EXERCÍCIO 1. (Distribuição Normal)

Jorge viaja diariamente de sua casa até o escritório onde trabalha. O tempo médio de uma viagem só de ida é de 24 minutos, com desvio padrão de 3,8 minutos. Suponha que o tempo de viagem seja normalmente distribuído.

$$X = \{ \text{Tempo de viagem} \}$$

$$X \sim N(24, 3.8)$$

$\mu \quad \sigma$



- Se o escritório abre às 9h e Jorge sai de casa às 8h45m diariamente, em que porcentagem do tempo ele se atrasa para o trabalho?

$$\frac{8h45}{\text{CASA}} \xrightarrow[15m]{\quad} \frac{9h}{\text{Escritório}}$$

Atraso
+15m

$$P(X > 15) = 1 - P(X \leq 15)$$

$$= 1 - P\left(Z \leq \frac{15 - 24}{3.8}\right)$$

$$= 1 - P(Z \leq -2.37)$$

$$= 1 - 0.0089 = 0.9911$$



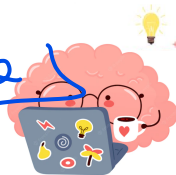
- b) Qual desvio padrão dos tempos é necessário para que a probabilidade de uma viagem gastar não mais do que 26 minutos seja 0,85?

$$P(X < 26) = 0,85$$

$$P\left(Z < \frac{26-24}{\sigma_x}\right) = 0,85$$

$$\frac{26-24}{\sigma_x} = \Phi^{-1}(0,85)$$

$$\sigma_x = \frac{26-24}{\Phi^{-1}(0,85)} = \frac{2}{1,04} \approx 1,92$$



- Encontre o período de tempo acima do qual encontramos os 20% mais lentos das viagens.

$$P(x > \pi) = 0,20 \Rightarrow 1 - P(x \leq \pi) = 0,20$$

$$\Rightarrow P\left(z \leq \frac{\pi - 24}{3,8}\right) = 0,80$$

$$\Rightarrow \frac{\pi - 24}{3,8} = \Phi^{-1}(0,80)$$

$$\Rightarrow \pi = 3,8 \times \Phi^{-1}(0,80) + 24$$

$$\Rightarrow \pi \approx 27,19$$



- Qual a probabilidade de que 2 das próximas 3 viagens levem pelo menos 0,5 hora?

$$Y \sim \text{Binomial}(n=3, p=?)$$

$$p = P(X \geq 30) \approx 0,0572$$

$N(24, 3 \cdot 8)$

$$P(Y=2) = \binom{3}{2} (0,0572)^2 (1 - 0,0572)^{3-2}$$

$$\approx \underline{0,0093}$$





MUITO OBRIGADA!

Naiara Caroline

