**autotestR: Um pacote R para facilitar testes estatísticos básicos com visualização automatizada**

**Resumo**

O pacote autotestR foi desenvolvido para auxiliar estudantes, pesquisadores e profissionais de ciências biológicas e áreas afins na realização de testes estatísticos básicos. Ele automatiza a seleção e aplicação de testes como t de Student, Mann-Whitney, ANOVA, Kruskal-Wallis e correlação de Pearson, oferecendo interpretação e visualização gráfica dos resultados. Com comandos simples e foco na usabilidade, autotestR representa uma ferramenta acessível para a análise exploratória de dados.

**1. Introdução**

A análise de dados é um componente fundamental no desenvolvimento de pesquisas em diversas áreas do conhecimento. Com os avanços tecnológicos e a ampla disponibilidade de dados, essa etapa tornou-se cada vez mais presente na rotina científica. No entanto, dois obstáculos ainda se destacam como barreiras recorrentes para o avanço científico e tecnológico: (i) a correta aplicação de testes estatísticos e (ii) o domínio de linguagens de programação voltadas à análise de dados.

A estatística é essencial para a interpretação de resultados experimentais. A escolha apropriada do teste estatístico, aliada à verificação de pressupostos (como normalidade e homogeneidade de variância), é indispensável para garantir conclusões válidas e reprodutíveis. Contudo, estudantes e profissionais com pouca formação na área frequentemente enfrentam dificuldades nessa etapa. Além disso, a apresentação visual dos resultados — etapa igualmente importante para a comunicação científica — é muitas vezes negligenciada ou executada de forma limitada.

Embora existam diversos softwares e ferramentas que auxiliam na análise estatística, o uso de linguagens de programação continua sendo uma exigência cada vez mais presente, especialmente para tarefas que envolvem reprodutibilidade, automação ou customização. A linguagem R, amplamente adotada na comunidade científica, destaca-se por sua flexibilidade, código aberto e extenso ecossistema de pacotes. No entanto, seu uso ainda representa um desafio para iniciantes, que frequentemente se deparam com blocos complexos de código para realizar tarefas simples, como a execução de testes básicos ou a geração de gráficos interpretáveis.

Diante desse cenário, desenvolvemos o pacote autotestR, com o objetivo de democratizar o acesso à estatística básica aplicada por meio da linguagem R. O pacote foca em simplificar a execução de testes estatísticos comuns, automatizando verificações de pressupostos e oferecendo saídas interpretáveis acompanhadas de visualizações informativas. Dessa forma, o autotestR visa ser uma ponte entre a análise estatística rigorosa e uma experiência de usuário mais acessível, especialmente para estudantes, pesquisadores iniciantes ou profissionais de áreas correlatas.

**2. Desenvolvimento e estrutura do pacote**

O desenvolvimento do pacote autotestR foi realizado com foco em facilitar a aplicação de testes estatísticos básicos por usuários com pouca familiaridade com a linguagem R ou com estatística inferencial. O pacote foi inteiramente desenvolvido na linguagem R (versão 4.4.2), utilizando um conjunto de bibliotecas amplamente empregadas na análise de dados, como ggplot2, dplyr, stats, car, nortest e multcompView.

A estrutura do pacote é composta por um conjunto de funções que seguem uma lógica comum: identificar automaticamente o teste estatístico adequado a partir das características dos dados, realizar o teste estatístico correspondente, gerar uma saída interpretável e apresentar uma visualização gráfica intuitiva. As funções foram projetadas para receber múltiplos vetores numéricos como entrada (via ...) e verificar, internamente, os pressupostos de normalidade (por meio do teste de Shapiro-Wilk ou Lilliefors) e homogeneidade de variância (teste de Levene).

Cada função principal do pacote — como teste.t(), teste.u(), teste.anova() e teste.kruskal() — incorpora tanto a execução do teste estatístico quanto a visualização dos dados por meio de gráficos tipo boxplot, enriquecidos com informações adicionais, como letras de significância (no caso de ANOVA) ou rótulos indicando valores de p.

Além disso, cada função conta com um argumento ajuda = TRUE que, quando ativado, exibe no console um resumo textual explicativo sobre o teste, incluindo sua aplicação, pressupostos, diferenças em relação a outros testes e exemplos de uso. Essa funcionalidade visa promover o aprendizado dos usuários durante o uso do pacote.

O processo de desenvolvimento foi realizado em ambiente RStudio, e a documentação das funções foi gerada utilizando o sistema roxygen2. O pacote foi estruturado de acordo com as boas práticas recomendadas para publicação no CRAN, com a inclusão de documentação (.Rd), exemplos reprodutíveis, licença de uso (MIT), e verificação de dependências.

*Comparação com Pacotes Existentes*

Diversos pacotes do R oferecem suporte à realização de testes estatísticos e à geração de visualizações associadas. Entre os mais utilizados estão rstatix, ggpubr e summarytools.

O pacote rstatix, por exemplo, disponibiliza funções encadeáveis com o pipe (%>%) para análise estatística, oferecendo uma sintaxe amigável para usuários acostumados com o estilo tidyverse. No entanto, ele exige que o usuário tenha familiaridade com manipulação de data frames e estruturas do tipo tibble, o que pode ser um obstáculo para iniciantes.

Já o pacote ggpubr facilita a adição de resultados estatísticos em gráficos do ggplot2, como valores de p e comparações múltiplas. Contudo, sua utilização depende da construção prévia de objetos ggplot e da combinação com outros pacotes, como rstatix ou car, para execução dos testes. Assim, exige um nível intermediário de conhecimento em R.

O summarytools é voltado para sumarizações descritivas e apresenta tabelas amigáveis para relatórios. Embora seja útil para descrever dados, não contempla testes inferenciais de forma integrada com verificação de pressupostos e visualização gráfica dos resultados.

O autotestR diferencia-se por oferecer uma abordagem integrada: a partir de uma única linha de código, o usuário pode realizar testes estatísticos, verificar pressupostos, obter uma interpretação textual dos resultados e visualizar os dados de maneira intuitiva. Além disso, as funções do autotestR não exigem conhecimentos prévios sobre a estrutura de data frames ou o uso do pipe, sendo acessíveis a iniciantes e facilitando a curva de aprendizado no R. Uma tabela resumindo o comparativo entre o autotestR com diferentes pacotes pode ser observada abaixo (Tabela 1).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tabela 1. Comparação do autotestR com diferentes pacotes para análise estatística.** | | | | |
| **Recurso** | **autotestR** | **rstatix** | **ggpubr** | **summarytools** |
| Verificação de pressupostos | Sim | Manual | Manual | Não |
| Testes estatísticos básicos | Sim | Sim | Não | Não |
| Visualização integrada | Sim | Não | Não | Não |
| Uso direto sem data frame | Sim | Não | Não | Sim |
| Saída interpretável | Sim | Não | Não | Sim |
| Interface em português e linguagem acessível | Sim | Não | Não | Não |

**3. Funcionalidades principais**

As principais funções são:

pre.teste(): avaliação inicial dos dados, sugerindo o teste estatístico mais adequado para os tipos de dados do usuário.

* teste.t(): Teste t de Student para dois grupos.
* teste.t.pareado(): Teste t pareado.
* teste.u(): Teste de Mann-Whitney.
* teste.anova(): ANOVA com pos-teste de Tukey e letras de significância.
* teste.kruskal(): Teste de Kruskal-Wallis com pos-teste de Dunn.
* teste.correlacao(): Correlação de Pearson ou Spearman com visualização.

Cada uma delas exibe gráficos automaticamente e é capaz de exibir uma "ajuda embutida" quando chamado ajuda = TRUE ou quando é chamada a função sem a inclusão de parâmetros.

Exemplo:

x <- rnorm(30, 10)

y <- rnorm(30, 12)

teste.t(x, y)

**4. Aplicabilidade e público-alvo**

O pacote se destina a:

* Estudantes de graduação e pós-graduação;
* Professores em aulas de estatística aplicada;
* Pesquisadores da área biológica, biomédica e veterinária.

Seu diferencial está em oferecer um fluxo simples para executar testes com feedback interpretado e visualização incorporada.

**5. Considerações finais**

O autotestR está disponível no GitHub () e em breve será submetido ao CRAN. Futuros desenvolvimentos incluem suporte a testes para variáveis categóricas, integração com shiny e tradução para inglês.

**Disponibilidade**

GitHub: [inserir URL do repositório]

Licença: MIT

**Agradecimentos**

Ao apoio da UFPR e FIOCRUZ pela formação e infraestrutura que possibilitaram o desenvolvimento deste pacote.

R Core Team (2024). \_R: A Language and Environment for Statistical Computing\_. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>