

## POWER AND SAMPLE SIZE FOR HEALTH RESEARCHERS: UMA FERRAMENTA PARA CÁLCULO DE TAMANHO AMOSTRAL E PODER DO TESTE VOLTADO A PESQUISADORES DA ÁREA DA SAÚDE

### *POWER AND SAMPLE SIZE FOR HEALTH RESEARCHERS: A TOOL FOR CALCULATING SAMPLE SIZE AND STATISTICAL POWER DESIGNED FOR HEALTH RESEARCHERS*

Rogério Boff Borges<sup>1</sup>, Aline Castello Branco Mancuso<sup>1</sup>,  
Suzi Alves Camey<sup>1,2</sup>, Vanessa Bielefeldt Leotti<sup>1,2</sup>, Vânia Naomi Hirakata<sup>1</sup>,  
Guilherme Serpa Azambuja<sup>2</sup>, Stela Maris de Jezus Castro<sup>1,2</sup>

#### RESUMO

Nas próximas edições da seção de Bioestatística da revista *Clinical & Biomedical Research* uma nova série de artigos será publicada abordando um assunto de grande importância ao planejar uma pesquisa: o tamanho de amostra mínimo necessário para atingir os objetivos do estudo. Nessa série será apresentado como calcular o tamanho de uma amostra usando a ferramenta PSS Health (*Power and Sample Size for Health Researchers*), construído em linguagem R por meio do pacote Shiny, para diferentes tipos e objetivos de estudo, direcionado à pesquisadores da área da saúde, utilizando termos e conceitos comumente utilizados nesta área. Além disso, o pacote fornece uma sugestão de texto com as informações consideradas no cálculo, e como devem ser descritas, com a finalidade de minimizar problemas de interpretação por parte dos pesquisadores. Neste primeiro artigo será apresentada essa ferramenta desenvolvida pela Unidade de Bioestatística do Grupo de Pesquisa e Pós-Graduação do Hospital de Clínicas de Porto Alegre, que permite calcular não apenas o tamanho de amostra, mas também o poder de um teste de hipóteses.

**Palavras-chave:** *Tamanho de amostra; poder do teste; estimação de parâmetros; comparação de grupos; R*

#### ABSTRACT

In the next issues of *Clinical and Biomedical Research*, the Biostatistics section will introduce a new series of articles addressing a very important subject for research planning: the minimum sample size to achieve the aim of a study. This series will show how to calculate sample size using PSS Health (Power and Sample Size for Health Researchers). This tool was built using R language through the Shiny package. It can be used for different types of study and is designed for health researchers by using terms and concepts commonly used in this area. PSS Health also suggests a text with information considered in the calculation to minimize problems of interpretation by the researchers. In this first article, a general overview of PSS Health will be presented. This tool, which was developed by the Research and Graduate Group Biostatistics Unit of the Hospital de Clínicas de Porto Alegre, is useful not only to calculate sample size but also to determine power of a hypothesis test.

**Keywords:** *Sample size; statistical power; parameter estimation; group comparison; R*

*Clin Biomed Res.* 2020;40(4):247-253

1 Grupo de Pesquisa e Pós-graduação (GPPG), Hospital de Clínicas de Porto Alegre (HCPA). Porto Alegre, RS, Brasil.

2 Departamento de Estatística, Instituto de Matemática e Estatística (IME), Universidade Federal do Rio Grande do Sul (UFRGS). Porto Alegre, RS, Brasil.

#### Autor correspondente:

Rogério Boff Borges  
l-bioestatistica@hcpa.edu.br  
Grupo de Pesquisa e Pós-graduação (GPPG), Hospital de Clínicas de Porto Alegre (HCPA)  
Rua Ramiro Barcelos, 2350.  
90035-007, Porto Alegre, RS, Brasil.

## INTRODUÇÃO

A maioria dos métodos estatísticos utilizados no processo inferencial pressupõe que os dados foram obtidos por meio de um processo de amostragem probabilística. Consequentemente, estudos que utilizam este processo de amostragem precisam de um tamanho amostral mínimo para ter resultados estatisticamente relevantes. Uma pesquisa com um número muito pequeno de indivíduos pode ter resultados de diferenças clinicamente relevantes que não sejam estatisticamente significativas. Por outro lado, uma pesquisa com tamanho amostral muito grande pode levar a um gasto desnecessário de fontes e recursos<sup>1</sup>.

Uma questão importante a ser definida previamente à realização do cálculo do tamanho da amostra é o objetivo principal da pesquisa a ser realizada. A realização deste cálculo está diretamente relacionada ao objetivo do estudo, que será a questão norteadora de como o cálculo deverá ser realizado e quais parâmetros serão considerados. No caso de haver mais do que um objetivo, é necessário calcular o tamanho da amostra para cada um deles e, então, utilizar o maior tamanho de amostra calculado. Outra opção é definir um dos objetivos como sendo o principal, ou primário, tendo claro que o tamanho amostral calculado pode ser insuficiente para os demais objetivos.

Atualmente, para determinar o tamanho amostral mínimo ou o poder de um teste de hipóteses, podem ser utilizadas diversas calculadoras on-line e *softwares* estatísticos. No entanto, grande parte destes acabam utilizando termos técnicos que nem sempre são acompanhados de uma explicação clara e objetiva sobre o significado de cada informação que deve ser inserida, o que pode acarretar erros graves no cálculo. Posteriormente, ao justificar o tamanho amostral do estudo, é necessário que todas estas informações utilizadas para o cálculo estejam claras. Inconsistências ou informações incompletas tornam o cálculo irreproduzível e questionável.

Com o objetivo de tornar acessível os cálculos do tamanho da amostra e do poder do teste dos projetos de pesquisa, foi desenvolvido o *Power and Sample Size for Health Researchers* (PSS Health), uma ferramenta que permite calcular o tamanho mínimo amostral necessário e o poder de um teste de hipóteses de forma simples, gratuita e interativa. Além disso, é voltada à pesquisadores da área da saúde, apresentando termos e conceitos comumente utilizados nesta área.

## PSS HEALTH

Inspirado no trabalho de conclusão de curso de Azambuja (2019)<sup>2</sup>, o PSS Health, desenvolvido pela Unidade de Bioestatística do Grupo de Pesquisa e

Pós-Graduação do Hospital de Clínicas de Porto Alegre, foi construído em linguagem R<sup>3</sup> por meio do pacote Shiny<sup>4</sup>, que possibilita uma interface gráfica amigável e interativa. O usuário pode acessar o programa de forma on-line ou via pacote no software R e navegar entre as abas para encontrar o cenário que melhor atenda o objetivo principal de sua pesquisa.

Dentre as funcionalidades disponíveis é possível calcular o poder ou o tamanho da amostra para estimar ou testar hipóteses sobre média e proporção (incluindo testes de não inferioridade, equivalência e superioridade), associação, coeficiente de correlação, coeficiente de regressão (linear, logística e Cox), modelos de medidas repetidas, alfa de Cronbach, concordância interobservador via Kappa, limites de concordância do gráfico de Bland-Altman, coeficiente de correlação intraclasse, área sob a curva, sensibilidade e especificidade incorporando a prevalência da doença (novas funcionalidades serão adicionadas nas próximas versões). Cabe salientar que todos os cálculos realizados na ferramenta pressupõem que os dados foram ou serão obtidos utilizando um processo de amostragem probabilística, caso não esteja explicitamente especificado, pressupõe amostragem aleatória simples. Silva et al.<sup>5</sup> apontam, que em um processo de amostragem probabilística, três condições devem ser satisfeitas:

1. todas as unidades da população de interesse devem ter uma probabilidade positiva de serem incluídas na amostra;
2. a escolha dessas unidades deve ser realizada utilizando algum mecanismo aleatório que permita saber a probabilidade de cada unidade ser selecionada;
3. essa probabilidade de ser selecionada deve ser levada em consideração no momento da análise de dados.

O PSS Health também apresenta um texto contendo todas as informações utilizadas no cálculo, que pode ser incluído no artigo, projeto de pesquisa, tese, dissertação, etc, sendo esta uma das vantagens da ferramenta. Assim, fica garantida a compreensão e reprodução do cálculo realizado e, consequentemente, reduz a divulgação inconsistente ou incompleta das informações utilizadas no cálculo do tamanho da amostra ou do poder.

## UTILIZANDO PELO COMPUTADOR

O PSS Health pode ser utilizado pelo computador por meio do pacote PSS.Health<sup>6</sup>. Neste caso é necessário ter instalado previamente o *software* R<sup>3</sup>, preferencialmente a última versão, que pode ser baixada em <https://cran.r-project.org/>. Com o R já instalado, é necessário instalar o pacote PSS.Health

utilizando o comando: `install.packages("PSS.Health")`. O processo descrito até aqui é necessário realizar somente uma vez ou quando houver atualizações tanto do programa R quanto do pacote PSS.Health. Recomendamos fortemente que as atualizações sejam realizadas periodicamente.

Com o programa R e o pacote PSS.Health instalados, basta executar o PSS Health utilizando os comandos: `library(PSS.Health)`, para carregar a ferramenta, e `PSS_Health()`, para executar. Executados os comandos, especificamente nesta ordem, uma aba abrirá em seu navegador iniciando a ferramenta.

## UTILIZANDO DE FORMA ON-LINE

A ferramenta também pode ser utilizada de maneira on-line no site da Unidade de Bioestatística (<https://sites.google.com/hcpa.edu.br/bioestatistica>) ou diretamente no endereço [\[bioestatistica.shinyapps.io/PSS\\\_Health\]\(https://bioestatistica.shinyapps.io/PSS\_Health\). Este modo de utilização tem a vantagem de não ser necessário se fazer instalações, bastando acessar utilizando o navegador em um dispositivo com acesso à internet.](https://hcpa-unidade-</a></p>
</div>
<div data-bbox=)

Podem haver diferenças nas funcionalidades presentes entre a versão on-line e a do pacote. Além disso, a versão on-line pode estar indisponível devido a manutenções ou indisponibilidade do servidor.

## UTILIZANDO A FERRAMENTA

Ao acessar o PSS Health, via site ou pacote, será apresentada a tela da página inicial da ferramenta, conforme ilustrado na Figura 1, que contém instruções gerais para seu uso, leituras recomendadas e uma seção sobre a concepção do aplicativo. Nas abas apresentadas na parte superior são listados os grupos de parâmetros ou técnicas estatísticas disponíveis na ferramenta.

Boas vindas! Médias ▾ Proporções ▾ Qui-quadrado Correlação Regressão ▾ Classificação ▾ Concordância ▾ Cronbach Outras ferramentas ▾

## PSS Health

Power and Sample Size for Health Researchers

Navegue entre as abas para encontrar o cenário correspondente ao objetivo do estudo. Altere os parâmetros conforme necessidade e utilize o texto como auxílio para entender o cálculo. **Sempre** procure um profissional de estatística para orientações no planejamento do estudo.

O PSS Health está disponível no CRAN e pode ser utilizado pelo computador por meio do pacote `PSS.Health`.

```
install.packages("PSS.Health")
PSS.Health::PSS_Health()
```

### Leituras recomendadas

Frequentemente a Unidade de Bioestatística do Grupo de Pesquisa e Pós-Graduação do Hospital de Clínicas de Porto Alegre publica artigos na seção de bioestatística da revista *Clinical & Biomedical Research*. Nessas publicações são abordados temas que podem te auxiliar na definição do tamanho amostral e do poder do teste.

- **Principais conceitos em Epidemiologia:** Tem dúvidas sobre os tipos de delineamento e métodos de amostragem?
  - Os principais delineamentos na Epidemiologia
  - Os principais delineamentos na Epidemiologia: Ensaios Clínicos I
  - Os principais delineamentos na Epidemiologia: Ensaios Clínicos II
- **Série "Perguntas que você sempre quis fazer, mas nunca teve coragem":** Tem dúvidas sobre conceitos comumente utilizados em estatística e epidemiologia?
  - Estatística Descritiva: Perguntas que você sempre quis fazer, mas nunca teve coragem
  - Bioestatística e epidemiologia: perguntas que você sempre quis fazer, mas nunca teve coragem
  - Teste de hipóteses: perguntas que você sempre quis fazer, mas nunca teve coragem
  - Modelagem estatística: Perguntas que você sempre quis fazer, mas nunca teve coragem
  - Equívocos Estatísticos: Perguntas que você sempre quis fazer, mas nunca teve coragem
- **Outras publicações:**
  - Um alerta sobre o uso de amostras pequenas na regressão logística
  - Cálculo de tamanho de amostra: proporções
  - Estudos transversais e longitudinais com desfechos binários: qual a melhor medida de efeito a ser utilizada?
  - Calculando o tamanho de efeito no SPSS
  - Beanplot uma nova ferramenta gráfica
  - Análise de concordância entre métodos de Bland-Altman
  - Uso do Modelo de Equações de Estimáveis Generalizadas na análise de dados longitudinais
  - Normalidade de variáveis: métodos de verificação e comparação de alguns testes não-paramétricos por simulação

### Sobre

Esse aplicativo foi concebido no trabalho de conclusão do curso de Bacharelado em estatística do aluno Guilherme Azambuja, sob a orientação da professora Stela Castro. Recebeu incrementos, novas funcionalidades e é mantido pela equipe da Unidade de Bioestatística Grupo de Pesquisa e Pós-Graduação do Hospital de Clínicas de Porto Alegre.

- Aline Castello Branco Mancuso (Orcid ID) (Lattes ID)
- Rogério Boff Borges (Orcid ID) (Lattes ID)
- Stela Maris de Jezus Castro (Orcid ID) (Lattes ID)
- Suzi Alves Carney (Orcid ID) (Lattes ID)
- Vanessa Bielefeldt Leotti (Orcid ID) (Lattes ID)
- Vânia Naomi Hirakata (Orcid ID) (Lattes ID)

**Sugestão, críticas ou bugs?** Mande um e-mail para [bioestatistica@hcpa.edu.br](mailto:bioestatistica@hcpa.edu.br).

Figura 1: Página inicial do PSS Health.

Para exemplificar, vamos calcular o tamanho amostral mínimo de um estudo hipotético cujo objetivo é detectar diferenças nas proporções de cura entre os indivíduos que receberão uma nova droga (A) ou o tratamento padrão (B). Neste caso o desfecho é a cura (sim/não) e há dois grupos independentes (nova droga e tratamento padrão). As hipóteses para comparação dessas duas proporções são:

$$H_0: \pi_A = \pi_B$$

$$H_1: \pi_A \neq \pi_B$$

onde  $\pi_A$  e  $\pi_B$  são as proporções de cura entre os indivíduos que utilizaram a nova droga e que utilizaram o tratamento padrão, respectivamente. Vamos supor que o pesquisador defina um poder de 80%, um nível de significância de 5% e que os grupos tenham o mesmo tamanho.

Definidas as hipóteses a serem testadas, no menu superior da ferramenta clique na aba “Proporções” e em seguida em “Dois grupos independentes” (Figura 2).



Figura 2: Aba proporções.

Uma nova página será aberta, onde três opções de cálculo serão apresentadas: “Testar”, “Estimar” e “Poder”. Seguindo o exemplo apresentado, na opção “Testar”, o usuário deve alterar as informações dos campos do painel lateral, à esquerda, conforme o planejamento do estudo. Os resultados, então, serão exibidos no painel à direita e são atualizados automaticamente conforme a interação do usuário.

Para auxiliar na compreensão dos resultados, podem ser alterados o nome do desfecho e dos grupos para o contexto do nosso exemplo clicando em “Mudar nomes” conforme seta indicada na Figura 3. Ao selecionar o item “Mudar nomes”, uma modal será aberta onde podem ser definidos os nomes do desfecho e dos grupos conforme indicado pela Figura 4.

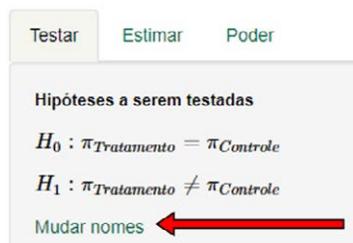


Figura 3: Local para mudar os nomes do desfecho e dos grupos.

#### Ajustes

Preencha os campos abaixo de acordo com seu estudo para que sirvam de guia no preenchimento dos demais campos.

Descreva o nome do desfecho

Descreva o nome do desfecho para que sirva de guia no preenchimento dos demais valores. Essa informação completará o texto do sugerido do tamanho amostral calculado para relatar nos projetos de pesquisa ou nos trabalhos científicos. O desfecho é a variável mais relevante do estudo, que servirá para testar a hipótese em questão (Castro et al. 2019).

Descreva um nome para o grupo Tratamento

Em alguns estudos o grupo Tratamento também pode ser chamado de grupo Intervenção ou grupo Exposto.

Descreva um nome para o grupo Controle

Em alguns estudos o grupo Controle também pode ser chamado de grupo Placebo/ Sham ou grupo Não exposto.

Figura 4: Modal para mudar os nomes do desfecho e dos grupos.

Após a edição dos nomes, ainda no painel esquerdo, entramos com o valor esperado para o percentual de cura no grupo B e no grupo A, no caso 15% (destacado com o círculo 1 na Figura 5) e 45% (destacado com o círculo 2 na Figura 5), respectivamente. Por padrão, a ferramenta já considera poder de 80%, nível de significância de 5% e grupos com o mesmo tamanho amostral, mas o usuário pode alterá-los no mesmo painel.

Conforme ilustrado na Figura 5, no painel do lado direito da tela, a ferramenta retornará:

- O tamanho amostral calculado (destacado com o círculo 3), neste caso 84 pacientes, sendo 42 receberão a droga A e 42 a droga B;
- Uma sugestão de texto (destacado com o círculo 4) contendo todas as informações utilizadas para o cálculo. Incluindo a incorporação de perdas e recusas;
- Sugestão para citação da ferramenta (destacado com o círculo 5). Cabe salientar a importância de descrever a ferramenta utilizada no cálculo para fins de transparência e reprodutibilidade;
- O comando do R utilizado para calcular o tamanho amostral (destacado com o círculo 6), permitindo que o usuário identifique a função utilizada no cálculo;
- E informações complementares, presente em algumas abas, como neste caso em que aparecerá um valor aproximado do intervalo de confiança da diferença entre as proporções (destacado com o círculo 7), dado o tamanho amostral calculado e demais informações do painel lateral.

Testar Estimar Poder

Hipóteses a serem testadas

$H_0 : \pi_A = \pi_B$

$H_1 : \pi_A \neq \pi_B$

Mudar nomes

% de cura esperado no grupo B

15

Medida do grupo A

% esperado no grupo A

% de cura esperado no grupo A

45

Balanço da amostra (B:A)

1

Poder (%)

80

**Tamanho amostral calculado: 84 (42 no A e 42 no grupo B)**

Sugestão de texto:

Foi calculado o tamanho de amostra para detectar diferenças entre as proporções de cura entre os A e B, por meio da ferramenta PSS Health versão 0.1.5 (citação abaixo). Considerando poder de 80%, nível de significância de 5% e proporção de cura no A de 45% e no B de 15% como é referida em Fulano (1900), chegou-se ao tamanho de amostra total de 84 sujeitos, sendo 42 em cada grupo. Acrescentando 10% para possíveis perdas e recusas o tamanho de amostra deverá ser 94

Obs.: Com esse tamanho de amostra o intervalo de confiança da diferença entre as proporções será aproximadamente [9.1% ; 50.9%].

Sugestões de citação:

Borges RB, Azambuja GS, Mancuso ACB, Leotti VB, Hirakata VN, Camey SA, Castro SMJ (2020). PSS Health: Power and Sample Size for Health Researchers via Shiny. R package version 0.1.5. Available from: <https://CRAN.R-project.org/package=PSS.Health>

Comando R utilizado:

```
EnvStats::propTestN(p.or.p1 = 0.45, p0.or.p2 = 0.15, alpha = 0.05, power = 0.8, sample.type = 'two.sample', alternative = 'two.sided', ratio = 1, correct = TRUE, warn = FALSE)
```

Figura 5: Tela para testar duas proporções de grupos independentes.

## CONSTRUÇÃO DE CENÁRIOS

Ao planejar um estudo, o pesquisador pode estar com dúvidas sobre os valores a serem inseridos no cálculo. Para auxiliar na escolha, algumas abas também possuem a funcionalidade de construir e comparar diferentes cenários. No exemplo, é possível criar cenários de tamanho amostral utilizando uma sequência de valores para a proporção esperada no grupo que usou a nova droga (A) (destacado com o círculo 1 na Figura 6) e diferentes valores para o

poder (destacado com o círculo 2 na Figura 6). No exemplo foi definido “Mínimo = 0”, “Máximo = 60” e “Intervalo = 5”, ou seja, as proporções esperadas no grupo que usou a nova droga são 0%, 5%, 10%, 15%, ..., 55% e 60% e um poder de 80%, 90% e 95%. A Figura 7 e a tabela da Figura 8, ambas apresentadas de forma interativa na página em questão, apresentam os diferentes tamanhos de amostra para todas as possíveis combinações de proporções e poder definidas pelo pesquisador. Aqui é possível avaliar o impacto dessas escolhas no tamanho da amostra.

## Construção de cenários

Utilize os argumentos abaixo para construir diferentes cenários. Você pode definir um intervalo de % para o grupo Tratamento e especificar valores do poder (%). Demais informações serão recuperadas do painel lateral.

Defina a sequência de valores (%) para o grupo Tratamento:

Mínimo	Máximo	Intervalo
5	60	5

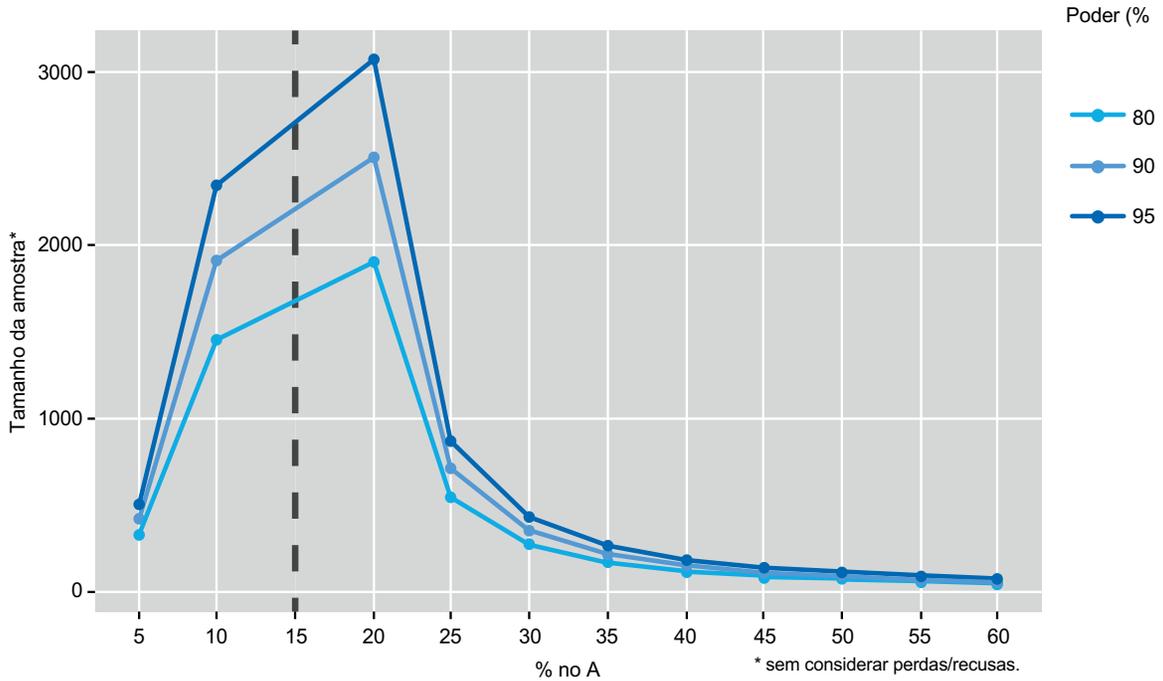
1

Digite valores de poder (%) para fazer o gráfico

80, 90, 95

2

Figura 6: Bloco para entrada de valores na construção de cenários.



Obs.: a linha tracejada representa a % no grupo controle definida no painel lateral.

Figura 7: Gráfico do tamanho de amostra necessário para os cenários construídos.

% A	% B	n total	n A	n B	Nível de significância (%)	Poder (%)	Hipótese alternativa	Balanco da amostra (B/A)	Correção de continuidade
60	15	44	22	22	5	80	two.sided	1	true
55	15	52	26	26	5	80	two.sided	1	true
60	15	54	27	27	5	90	two.sided	1	true
60	15	64	32	32	5	95	two.sided	1	true
50	15	66	33	33	5	80	two.sided	1	true
55	15	66	33	33	5	90	two.sided	1	true
55	15	78	39	39	5	95	two.sided	1	true
45	15	84	42	42	5	80	two.sided	1	true
50	15	84	42	42	5	90	two.sided	1	true
50	15	100	50	50	5	95	two.sided	1	true

Showing 1 to 10 of 33 entries

Figura 8: Tabela com os valores do tamanho de amostra necessário para os cenários construídos.

## CONSIDERAÇÕES FINAIS

O PSS Health é uma ferramenta estatística, voltada à pesquisadores da área da saúde, de fácil acesso e manipulação, com potencial para o ensino e a aprendizagem de quem deseja definir o tamanho amostral de um estudo ou calcular o poder de um teste de hipótese. Suas funcionalidades abrangem diversos parâmetros e técnicas utilizados na área da saúde de forma simples, direta e clara, possibilitando

a compreensão de iniciantes e facilitando a exploração por parte dos usuários mais experientes.

Este é o primeiro artigo de uma série sobre o uso da ferramenta PSS Health no planejamento de uma pesquisa. Os próximos artigos abordarão o seu uso para diferentes parâmetros e técnicas estatísticas.

### Conflitos de interesse

Os autores declaram não ter conflitos de interesse.

## REFERÊNCIAS

1. Paes AT. Qual deve ser o tamanho da minha amostra. *Einstein Educ Contin Saude*. 2008;6(4):153-4.
2. Azambuja GS. *Aplicativo Web Shiny para cálculo do tamanho de amostra e poder* [Internet]. Porto Alegre. Monografia [Bacharelado em Estatística] – Universidade Federal do Rio Grande do Sul, Instituto de Matemática e Estatística; 2019 [citado 30 nov. 2020]. Disponível em: <https://lume.ufrgs.br/handle/10183/212679>
3. R Core Team. *The R Project for Statistical Computing* [Internet]. Vienna: The R Foundation; 2020 [citado 30 nov. 2020]. Disponível em: <https://www.R-project.org/>
4. Chang W, Cheng J, Allaire JJ, Sievert C, Schloerke B, Xie Y, et al. *Shiny: Web Application Framework for R* [Internet]. Boston: Winston Chang RStudio; 2020 [citado 30 nov. 2020]. Disponível em: <https://CRAN.R-project.org/package=shiny>
5. Silva PLN, Bianchini ZM, Dias AJR. *Amostragem: Teoria e Prática Usando R* [Internet]. Rio de Janeiro: ENCE; 2021 [citado 30 nov. 2020]. Disponível em: <https://amostragemcomr.github.io/livro/conceitos.html>
6. Borges RB, Azambuja GS, Mancuso ACB, Leotti VB, Hirakata VN, Camey AS, et al. *PSS.Health: Power and Sample Size for Health Researchers via Shiny* [Internet]. Porto Alegre: HCPA; 2020 [citado 1 dez. 2020]. Disponível em: <https://CRAN.R-project.org/package=PSS.Health>

Recebido: 1 dez, 2020

Aceito: 8 dez, 2020