



Trabalho de Conclusão de Curso

**Tamanho de amostra e poder com os pacotes
longpower e *HMISC* do Software R**

Carolina Crestani

17 de janeiro de 2020

Carolina Crestani

**Tamanho de amostra e poder com os pacotes *longpower* e
HMISC do Software R**

Trabalho de Conclusão apresentado à comissão de Graduação do Departamento de Estatística da Universidade Federal do Rio Grande do Sul, como parte dos requisitos para obtenção do título de Bacharel em Estatística.

Orientadora: Profa. Dra. Stela Maris de Jesus Castro

Porto Alegre
Dezembro de 2019

Carolina Crestani

**Tamanho de amostra e poder com os pacotes *longpower* e
HMISC do Software R**

Este Trabalho foi julgado adequado para obtenção dos créditos da disciplina Trabalho de Conclusão de Curso em Estatística e aprovado em sua forma final pela Orientadora e pela Banca Examinadora.

Orientadora: _____
Prof. Dra. Stela Maris de Jezus Castro, UFRGS
Doutora pela Universidade Federal do Rio Grande do Sul, Porto Alegre, RS

Banca Examinadora:

Prof. Dr. Álvaro Vigo, UFRGS
Doutor pela Universidade Federal do Rio Grande do Sul, Porto Alegre, RS

Porto Alegre
Dezembro de 2019

Resumo

Esta monografia tem como propósito a construção de um material de apoio prático e objetivo sobre cálculos de tamanho de amostra e poder através do software R para pesquisadores e para a comunidade interessada.

São abordados dois pacotes, *longpower* e *Hmisc*, em que, para cada uma de suas respectivas funções selecionadas, é explicada a sua aplicação, seus parâmetros e são apresentados exemplos de utilização.

No pacote *longpower*, são abordadas funções para cálculo de tamanho de amostra no ambiente de estudos envolvendo dados longitudinais, já na no pacote *Hmisc*, as funções tratam de tamanho de amostra e poder para cenários mais gerais, além de abrangerem a área de análise de sobrevivência.

Palavras-Chave: Amostragem, Tamanho de amostra, Poder, Dados Longitudinais, Análise de Sobrevivência, Software R, Sample Size, Power.

Abstract

This paper aims to build practical support material and objective about sample size calculations and power through the R software for researchers and to the community interested.

Two packages are addressed, *longpower* and *Hmisc*, where, for each of their respective selected functions, their application, their parameters and examples of use are presented.

In the *longpower* package, functions for sample size calculation in the study environment involving longitudinal data are discussed, while in the *Hmisc* package, functions not only deal with more general cases, but approach the area of survival analysis.

Keywords: Sampling, Longitudinal Data, Survival Analysis, Software R.

Sumário

1	Introdução	8
2	Dados Longitudinais	10
3	Pacote <i>longpower</i>	12
3.1	Função <i>diggle.linear.power</i>	12
3.1.1	Aplicação	12
3.1.2	Argumentos	13
3.1.3	Exemplo	14
3.2	Função <i>edland.linear.power</i>	15
3.2.1	Aplicação	15
3.2.2	Argumentos	15
3.2.3	Exemplo	16
3.3	Função <i>liu.liang.linear.power</i>	17
3.3.1	Aplicação	17
3.3.2	Argumentos	17
3.3.3	Exemplo	18
3.4	Função <i>lmmpower</i>	20
3.4.1	Aplicação	20
3.4.2	Argumentos	20
3.4.3	Exemplo	21
3.5	Função <i>power.mmr.ar1</i>	22
3.5.1	Aplicação	22
3.5.2	Argumentos	23
3.5.3	Exemplo	23
3.6	Função <i>power.mmr</i>	25
3.6.1	Aplicação	25
3.6.2	Argumentos	26
3.6.3	Exemplo	26
4	Análise de Sobrevida	28
5	Pacote <i>HMISC</i>	31
5.1	Função <i>samplesize.bin</i>	31
5.1.1	Aplicação	31
5.1.2	Argumentos	32
5.1.3	Exemplo	33

5.2	Função <i>posamsize</i>	33
5.2.1	Aplicação	33
5.2.2	Argumentos	34
5.2.3	Exemplo	34
5.3	Funções <i>popower</i>	35
5.3.1	Aplicação	35
5.3.2	Argumentos	36
5.3.3	Exemplo	36
5.4	Funções <i>bpower</i> e <i>bpower.sim</i>	37
5.4.1	Aplicação	37
5.4.2	Argumentos	37
5.4.3	Exemplo	37
5.5	Funções <i>bsamsize</i> e <i>ballocation</i>	38
5.5.1	Aplicação	38
5.5.2	Argumentos	38
5.5.3	Exemplos	39
5.6	Função <i>cpower</i>	40
5.6.1	Aplicação	40
5.6.2	Argumentos	40
5.6.3	Exemplo	41
5.7	Função <i>ciapower</i>	42
5.7.1	Aplicação	42
5.7.2	Argumentos	42
5.7.3	Exemplo	42
5.8	Função <i>spower</i>	44
5.8.1	Aplicação	44
5.8.2	Argumentos	44
5.8.3	Exemplo	45
6	Considerações Finais	47
9	Referências Bibliográficas	47

1 Introdução

A amostragem é cotidianamente usada na tomada de decisões. Na cozinha, por exemplo, verifica-se se o prato está temperado retirando apenas uma pequena parte dele, ou ainda, durante um exame médico, em que se coleta uma pequena quantidade de sangue a fim de detectar as condições do paciente. Essas são decisões baseadas em procedimentos amostrais (Bolfarine e Bussab, 2005).

Dentro do planejamento de um estudo científico, a amostragem tem um papel fundamental para a validade e precisão dos resultados. Pesquisas investigativas objetivam descrever fenômenos ou comparar o comportamento de certas características de subgrupos de uma população. Usualmente, não é acessível ou viável realizar um estudo de todo o universo populacional (censo), porém, este também não se faz necessário quando se dispõe de uma amostra representativa para a realização de inferências à população-alvo (Miot, 2011).

Obter essa amostra que produz resultados confiáveis e livres de vieses é o objetivo de qualquer pesquisador, podendo ser definida, segundo Bolfarine e Bussab (2005), como aquela que é uma micro representação do universo. Entretanto, esse conceito não é trivial e precisa ser estabelecido para o uso científico dos processos amostrais. Entre outros fatores como tipo de coleta e escopo do estudo, o tamanho de amostra (n) e o poder envolvido são duas características importantes a serem consideradas no momento de avaliar a representatividade da amostra.

Apesar da grande relevância, as dificuldades em se obter um cálculo adequado são muitas e raramente encontra-se uma resposta rápida e precisa (Paes, 2008), fazendo com que essa etapa do planejamento amostral seja deixada de lado na hora do estudo. O uso inadequado de um procedimento de cálculo amostral pode levar a um viés de interpretação do resultado, enquanto uma amostra muito grande pode acarretar custos desnecessários ou implicações éticas, uma amostra pequena pode tornar a pesquisa inconclusiva (Bolfarine e Bussab, 2005).

Estudos mais complexos como os que envolvem dados longitudinais e análise de sobrevivência, abordados nesta monografia, necessitam de um cuidado especial com o planejamento amostral, e conseqüentemente, com o cálculo do tamanho da amostra e poder. Esse elemento a mais é dado pelo fato do tempo ser uma variável de interesse em ambas as metodologias. Enquanto a análise de sobrevivência designa a análise estatística de dados quando a variável em estudo representa o tempo desde um instante inicial bem definido até à ocorrência de determinado acontecimento de interesse (Bastos e Rocha, 2006) ou censura, os estudos longitudinais envolvem duas ou mais observações, realizadas em instantes diferentes, da variável resposta em cada unidade amostral monitorada (Singer et al., 2018).

Existe uma infinidade de 'calculadoras amostrais' de acesso livre na internet que auxiliam no momento de decidir o tamanho de amostra, dois exemplos são as hospedadas pela UCSF ([Insitute, 2019](#)) e pela HyLown Consulting ([LLC, 2019](#)) que podem ser utilizadas de forma gratuita. Vários softwares, como o PASS ([NCSS e LLC, 2019](#)) e o GLIMMPSE ([Kreidler et al., 2013](#)) desempenham cálculos para medidas repetidas e não requerem experiência em programação, enquanto o MEDCALC (<https://www.medcalc.org/index.php>) e o XLSTAT (<https://www.xlstat.com/en/>), suportam o planejamento amostral para dados de sobrevivência, porém, apenas o GLIMMPSE é código aberto e não necessita pagamento. Um software que comporta ambas as técnicas e cobre uma gama ainda maior de cálculos estatísticos, incluindo os amostrais, é o software livre R ([Team, 2019](#)).

Conforme [Aquino \(2014\)](#), o R possibilita uma grande produtividade no trabalho e, o que é mais importante, eficácia. É aconselhável que o leitor sempre atente ao documento de ajuda de cada função, visto que a maioria delas possui argumentos adicionais que são úteis em situações diversas.

Assim, este trabalho apresentará os procedimentos para calcular tamanho de amostra e poder dos pacotes *longpower* ([Donohue, 2019](#)) e *Hmisc* ([Jr., 2019](#)) do software R. Serão abordadas, respectivamente, *diggle.linear.power*, *edland.linear.power*, *liu.liang.powewr*, *lmpower*, *power.mmrn.ar1* e *power.mmrn* pertencentes ao *longpower* e *samplesize.bin*, *posamsize*, *popower*, *bpower*, *bpower.sim*, *bsamsize*, *ballocation*, *cpower*, *ciapower* e *spower* do *Hmisc*. A metodologia abordada será elucidar a técnica estatística e exemplificar cada uma destas funções.

Para noções de uso e entendimento mais gerais do R, recomenda-se [Landeiro \(2011\)](#), enquanto para pacotes focados em cálculos de amostra e poder mais elementares, indica-se ([Pedrotti, 2017](#)), para um maior aprofundamento em outras funções de tamanho de amostra de poder na área de análise de sobrevivência, sugere-se ([Fontoura, 2018](#)).

Com intuito de facilitar o desenvolvimento dos conteúdos abordados nesta monografia, foi utilizado o ambiente de desenvolvimento integrado (IDE) RStudio, sendo uma ferramenta que contribui na melhora da comunicação entre usuário e software. O RStudio foi cuidadosamente pensado para antecipar as necessidades dos usuários de R, simplificando sua abordagem para a criação e transferência de conhecimento [Racine \(2012\)](#). Seu IDE incorpora um editor (neste caso, adaptado à linguagem R), uma interface de vários documentos para que se possa ter uma variedade de documentos abertos em diferentes guias, uma janela para interagir com o R e uma janela que serve como um arquivo navegador e referência de idioma.

É importante salientar que esta monografia não é um trabalho introdutório, que objetiva o ensino de conceitos básicos. Por isso, espera-se que o leitor tenha algum conhecimento de conceitos e técnicas para as quais as funções destes pacotes disponibilizam o cálculo do tamanho amostral ou poder estatístico, já que o texto é destinado àqueles que sabem o que estão buscando em termos de análise de dados, e precisam realizar estes cálculos em estudos específicos. Importante também destacar que os exemplos aqui presentes foram adaptados de outros trabalhos somente para fins didáticos, os valores usados são puramente ilustrativos.

2 Dados Longitudinais

Em muitas situações práticas, conforme [Singer et al. \(2018\)](#), existe o interesse em se modelar o comportamento de uma ou mais variáveis respostas medidas nas unidades de uma ou mais populações ao longo de alguma dimensão ordenada. Geralmente, essa escala ordenada ao longo da qual se fazem as medidas repetidas é o tempo.

Assim, as pesquisas ditas 'longitudinais' são aquelas em que as mesmas (ou diferentes) variáveis são medidas para os mesmos indivíduos em, pelo menos, dois pontos distintos do tempo. Dessa forma, consistem em uma importante fonte para estudos sobre mudanças demográficas e socioeconômicas, estudos epidemiológicos e ambientais, entre outros [Carneiro \(2012\)](#).

Segundo [Singer et al. \(2018\)](#):

Estudos longitudinais têm interesse especial quando o objetivo é avaliar tanto mudanças globais quanto individuais ao longo do tempo. Primeiramente, eles permitem que a variável resposta seja observada em unidades amostrais sob níveis constantes de outras covariáveis que podem influenciá-la. Esta característica é de especial interesse nos casos em que a variabilidade interunidades amostrais é maior do que a variabilidade intraunidades amostrais. Em segundo lugar, planejamentos longitudinais proveem informações sobre variações individuais nos níveis da variável resposta. Finalmente, é importante notar que alguns parâmetros dos modelos estatísticos subjacentes podem ser estimados de forma mais eficiente sob planejamentos longitudinais do que sob planejamentos transversais com o mesmo número de observações.

As metodologias utilizadas na análise deste tipo de dados são várias, e dependem muito da natureza das variáveis do estudo. Entre elas estão a ANOVA para medidas repetidas, os modelos lineares generalizados mistos, para dados com distribuição na família exponencial, os modelos baseados em Equações de Estimção Generalizada (GEE - Generalized Estimating Equation), para casos em que apenas a relação entre a variância e a média é especificada, sem menção à forma da distribuição, entre outros.

A maior desvantagem dos estudos longitudinais está relacionada com seu custo, pois em muitas situações exige-se um grande esforço para garantir a observação das unidades amostrais nos instantes pré-determinados, e em outras, o período de observação pode ser muito longo ([Singer et al., 2018](#)). Por isso, o planejamento amostral

adequado, com uma taxa de alocação, um poder, uma significância e demais parâmetros desenhados da forma mais vantajosa para o pesquisador, é tão importante na condução deste tipo de estudo.

3 Pacote *longpower*

O pacote *longpower* contém funções para calcular poder e tamanho da amostra para modelos lineares de dados longitudinais com base nos métodos descritos por Diggle et al. (2002) e Liu e Liang (1997), sendo expressas em termos de modelos marginais ou parâmetros das Equações de Estimação Generalizadas (GEE). Algumas das funções convertem parâmetros piloto de modelos de efeito misto (por exemplo, intercepto e / ou inclinação aleatória) em parâmetros do modelo marginal, de modo que as fórmulas supracitadas possam ser aplicadas para produzir cálculos para duas amostras com delineamentos longitudinais supondo variância conhecida.

Além disso, ainda existem funções que lidam com o tempo categórico dos Modelos Mistos de Medidas Repetidas (MMRM - Mixed Model for Repeated Measurements), usando o método de Lu et al. (2008).

As funções abordadas deste pacote serão seis, *diggle.linear.power*, *edland.linear.power*, *liu.liang.linear.power*, *lmpower*, *power.mmr.ar1* e *power.mmr*. Michael C. Donohue, professor da Universidade do Sul da Califórnia, é o principal autor e mantenedor do pacote, tendo contado com o apoio de Anthony C. Gamst e Steven D. Edland na implementação de algumas das funções.

Os códigos implementados no R das funções descritas a seguir podem ser encontrados em <https://github.com/cran/longpower/tree/master/R>, repositório do GitHub criado pelo autor do pacote e sustentado pela comunidade interessada.

3.1 Função *diggle.linear.power*

3.1.1 Aplicação

Esta função calcula do tamanho da amostra necessário para se diferenciar o comportamento de certa variável de interesse em dois grupos independentes através do tempo, tratado como quantitativo.

A relação entre a variável de interesse e o tempo é linear, por isso esta função retorna o número de sujeitos necessários por grupo para detectar possíveis diferenças nas inclinações das duas retas (uma de cada grupo), atingindo certo poder especificado. As hipóteses podem ser unilaterais ou bilaterais, dadas por:

$$\begin{aligned}
 H_0 &: \beta_{1B} = 0 \\
 H_1 &: \beta_{1B} \neq 0 \text{ ou} \\
 H_1 &: \beta_{1B} < 0 \text{ ou} \\
 H_1 &: \beta_{1B} > 0
 \end{aligned}
 \tag{3.1}$$

Considerando β_{1B} como a diferença na inclinação da reta média do grupo B (controle) em relação à reta média do grupo A (intervenção). A *diggle.linear.power* pode ser aplicada considerando tanto um modelo GEE com estrutura de correlação permutável (exchangeable), quanto um modelo de efeitos aleatórios, que inclui um intercepto aleatório para cada grupo de observações correlacionadas.

A função avalia a seguinte fórmula de [Diggle et al. \(2002\)](#):

$$m = \frac{2(Z_{1-\alpha/2} + Z_{1-\beta})^2(\sigma_b^2 + \sigma_e^2)^2(1 - \sigma_b^2/\sigma^2)}{n s_x^2 d^2} \quad (3.2)$$

Onde:

m = número de sujeitos por grupo para alcançar os valores especificados para os erros Tipo I e Tipo II;

α = erro Tipo I, probabilidade do estudo rejeitar a hipótese nula quando esta for correta;

β = erro Tipo II, probabilidade de o estudo não rejeitar a hipótese nula quando falsa;

d = menor diferença significativa entre as inclinações das duas retas a ser detectada. É desejado, normalmente, que estudo rejeite a hipótese nula com alta probabilidade quando o parâmetro de interesse se desvia do valor sob a hipótese nula por uma quantidade d ou mais, de modo que o valor de d é escolhido como tendo uma significância prática;

$\sigma^2 = \sigma_b^2 + \sigma_e^2$, variância entre as medições, para uma variável de resposta contínua Y_{ij} , a quantidade $\sigma^2 = Var(Y_{ij})$ é a variância total. O valor de σ^2 às vezes está disponível para uma aproximação razoável por meio de estudos piloto ou estudos relatados anteriormente na literatura. Caso contrário, o estatístico deve obter do investigador uma estimativa mais plausível do seu valor;

σ_b^2 = variância do intercepto aleatório;

σ_e^2 = variância do erro aleatório;

n = número de observações por indivíduo no estudo, pode ser restringido por considerações práticas, ou pode precisar ser equilibrado com o tamanho da amostra;

s_x^2 = variância dentro de cada sujeito observado.

3.1.2 Argumentos

Para a função *diggle.linear.power* são necessários os seguintes argumentos:

- **n**, número de medidas repetidas no tempo.
- **delta**, diferença dos grupos nas inclinações.
- **t**, tempos observados.
- **sigma2**, variância residual.
- **R**, a matriz de correlações (ou matriz de variâncias-covariâncias se $\text{sigma2} = 1$). Se **R** for escalar, será assumida uma matriz de correlação do tipo permutável.
- **sig.level**, erro Tipo I.

- **power**, poder.
- **alternative**, teste uni ou bicaudal.
- **tol**, tolerância numérica usada na localização de raízes, por padrão `.Machine$double.eps2`, usualmente, sem necessidade de modificações.

3.1.3 Exemplo

Retirado de (Ndow, 2018). Deseja-se realizar um ensaio clínico sobre o efeito de um novo medicamento na redução da pressão arterial. São planejadas quatro visitas para medições dos pacientes, incluindo a primeira, nos anos de estudo 0, 2, 5 e 8, portanto, $n = 4$.

Considerando um α de 0,05 e um poder de 0,8, assume-se a menor diferença significativa como 0,5 mmHg / ano (d) para que a função retorne o número de sujeitos necessários nos grupos tratamento (A) e controle (B) em doze cenários. Baseado em estudos prévios, encontrou-se quatro valores para a correlação (ρ) = 0,2; 0,3; 0,5; 0,8 e para a variância estimada do modelo os valores (σ^2) = 69, 110, 183; que são o resultado da soma da variância do intercepto (σ_b^2) = 55 com os valores das estimativas da variância do erro (σ_e^2) = 14, 55, 128.

O modelo é expresso por:

$$Y_{ij} = (\beta_{0A} + \beta_{0B}TRAT_i + b_{0i}) + (\beta_{1A} + \beta_{1B}TRAT_i)x_{ij} + \epsilon_{ij} \quad (3.3)$$

Onde:

Y_{ij} = pressão observada do indivíduo i no ano j ;

$TRAT_i$ = tratamento a que pertence o indivíduo i (A se do tratamento intervenção e B se do tratamento controle);

β_{0A} = intercepto da reta média da intervenção;

β_{0B} = diferença no intercepto da reta média do controle em relação à intervenção;

b_{0i} = efeito aleatório de intercepto $\rightarrow b_{0i} \sim N(0, \sigma_b^2)$;

β_{1A} = inclinação da reta média da intervenção;

β_{1B} = diferença na inclinação da reta média do controle em relação à intervenção;

x_{ij} = ano j do indivíduo i ;

ϵ_{ij} = efeito aleatório de indivíduo $\rightarrow \epsilon_{ij} \sim N(0, \sigma_e^2)$.

A estrutura de correlação utilizada:

$$\mathbf{R} = \begin{bmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix}$$

As hipóteses testadas são:

$$\begin{aligned} H_0 &: \beta_{1B} = 0 \\ H_1 &: \beta_{1B} < 0 \end{aligned} \quad (3.4)$$

No RStudio:

```
> n = 4
> t = c(0, 2, 5, 8)
> rho = c(0.2, 0.3, 0.5, 0.8)
> sigma2 = c(69, 110, 183)
> tab = outer(rho, sigma2,
```

```

+           Vectorize(function(rho, sigma2){
+               ceiling(diggle.linear.power(
+                   delta=0.5,
+                   t=t,
+                   sigma2=sigma2,
+                   R=rho,
+                   alternative="one.sided",
+                   power = 0.80)$n)}}))
> colnames(tab) = paste("sigma2=", sigma2)
> rownames(tab) = paste("rho=", rho)
> tab
      sigma2= 69 sigma2= 110 sigma2= 183
rho= 0.2      75      119      198
rho= 0.3      66      104      173
rho= 0.5      47       75      124
rho= 0.8      19       30       50

```

Dentro de cada valor de variância (σ^2), o tamanho da amostra necessário diminui à medida que a correlação entre as observações (ρ) aumenta. Assim, com uma variância menor e maior correlação, seriam necessários apenas 19 indivíduos em cada grupo para testar os efeitos da medicação.

3.2 Função *edland.linear.power*

3.2.1 Aplicação

Esta função desempenha o cálculo do tamanho da amostra para testar diferenças de um modelo misto linear com inclinação aleatória, dadas as hipóteses:

$$\begin{aligned}
 H_0 &: \beta_{1B} = 0 \\
 H_1 &: \beta_{1B} \neq 0 \text{ ou} \\
 H_1 &: \beta_{1B} < 0 \text{ ou} \\
 H_1 &: \beta_{1B} > 0
 \end{aligned}
 \tag{3.5}$$

β_{1B} é a diferença na inclinação da reta média do grupo controle em relação ao grupo tratamento. Caso a variância da inclinação aleatória seja zero ($\text{sig2.s} = 0$), a função fornecerá estimativas de tamanho de amostra para modelos mistos lineares com intercepto aleatório.

3.2.2 Argumentos

Para a função *edland.linear.power* são necessários os seguintes argumentos:

- **n**, número de medidas repetidas por grupo.
- **delta**, diferença mínima entre os grupos a ser considerada significativamente diferente de zero.
- **t**, tempos observados.
- **sig2.s**, variância da inclinação aleatória.
- **sig2.e**, variância residual.
- **sig.level**, erro Tipo I.
- **power**, poder.
- **alternative**, teste uni ou bicaudal.
- **tol**, tolerância numérica usada na localização de raízes, por padrão `.Machine$double.eps^2`, usualmente, sem necessidade de modificações.

3.2.3 Exemplo

Retirado de [Donohue et al. \(2019\)](#). Deseja-se estimar o tamanho da amostra necessário para conduzir um estudo sobre a doença de Alzheimer, para isso, o modelo assumido é com inclinação aleatória, já que os pacientes tendem a ter uma grande variância dentro dos grupos (controle e tratamento). São previstas avaliações realizadas a cada três meses, durante 18 meses ($n = 7$ repetições).

Assume-se um menor tamanho de efeito detectável de 1,5 pontos na porção cognitiva da Escala de Avaliação de Doenças de Alzheimer (ADAS-Cog - Alzheimer's Disease Assessment Scale - Cognitive Subscale), sendo esta uma escala de 70 pontos com grande variabilidade entre indivíduos doentes. Define-se, através de estudos pilotos, que a inclinação aleatória tem uma variância de 24, o modelo tem variância residual de 10, e deseja-se alcançar um poder de 0,8 em um teste bicaudal de 0,05 de significância.

O modelo é expresso por:

$$Y_{ij} = (\beta_{0A} + \beta_{0B}Grupo_i) + (\beta_{1A} + \beta_{1B}Grupo_i + b_{1i})x_{ij} + \epsilon_{ij} \quad (3.6)$$

Onde:

Y_{ij} = pontos observados do indivíduo i na avaliação j ;

$Grupo_i$ = grupo a que pertence o indivíduo i (A se do grupo tratamento e B se do grupo controle);

β_{0A} = intercepto da reta média do grupo tratamento;

β_{0B} = diferença no intercepto da reta média do grupo controle em relação ao grupo tratamento;

β_{1A} = inclinação da reta média do grupo tratamento;

β_{1B} = diferença na inclinação da reta média do grupo controle em relação ao grupo tratamento;

b_{1i} = efeito aleatório de inclinação $\rightarrow b_{1i} \sim N(0, \sigma_b^2)$;

x_{ij} = avaliação j do indivíduo i ;

ϵ_{ij} = efeito aleatório de indivíduo $\rightarrow \epsilon_{ij} \sim N(0, \sigma^2)$.

As hipóteses testadas são:

$$\begin{aligned} H_0 &: \beta_{1B} = 0 \\ H_1 &: \beta_{1B} \neq 0 \end{aligned} \quad (3.7)$$

No RStudio:

```
> t = c(0.00, 0.25, 0.50, 0.75, 1.00, 1.25, 1.50)
> n = length(t)
> edland.linear.power(delta=1.5, t=t, sig2.s = 24, sig2.e = 10, sig.level=0.05,
power = 0.80)
```

```
Power for longitudinal linear model with random slope (Edland, 2009)
```

```
      n = 207.3101
      delta = 1.5
      sig2.s = 24
      sig2.e = 10
      sig.level = 0.05
      t = 0.00, 0.25, 0.50, 0.75, 1.00, 1.25, 1.50
      power = 0.8
      alternative = two.sided
```

```
NOTE: n is number in *each* group
```

Portanto, o estudo exigiria cerca de 208 indivíduos por grupo.

3.3 Função *liu.liang.linear.power*

3.3.1 Aplicação

Estima o tamanho de amostra necessário para detectar diferenças em um parâmetro de interesse entre dois grupos, para modelos com intercepto aleatório ou modelos com intercepto e inclinação aleatórias. Além do número total da amostra, e os valores por grupo, caso o estudo seja desbalanceado, a função também retorna, por padrão, a matriz de correlações do estudo.

A função *liu.liang.linear.power* performa a seguinte fórmula de [Liu e Liang \(1997\)](#):

$$m = \frac{2(Z_{1-\alpha/2} + Z_{1-\beta})^2 \sigma^2}{\pi_0 \pi_1 \psi_1^2 (\mathbf{1}' R^{-1} \mathbf{1})} \quad (3.8)$$

Onde:

m = número de sujeitos por grupo para alcançar valores específicos de erro Tipo I e poder;

ψ_1 = parâmetro de interesse sob a hipótese alternativa;

π_0 = proporção de sujeitos no grupo 0 (controle);

π_1 = proporção de sujeitos no grupo 1 (tratamento);

σ^2 = variância do modelo;

R = matriz de variâncias-covariâncias;

$\mathbf{1}'$ = vetor $n \times 1$ de uns;

$1 - \beta$ = poder;

α = nível de significância.

3.3.2 Argumentos

Para a função *liu.liang.linear.power* são necessários os seguintes argumentos:

- **N**, tamanho total da amostra, pode acomodar alocação entre grupos desequilibrada com o parâmetro π_i .
- **delta**, diferença entre os interceptos ou entre as inclinações dos grupos (possivelmente um vetor de diferenças).
- **u**, uma lista de covariáveis (vetores/matrizes) associadas ao parâmetro de interesse.
- **v**, uma lista respectiva de covariáveis (vetores/matrizes) associadas ao parâmetro de incômodo (aquele em que não se está interessado).
- **sigma2**, variância residual.
- **R**, matriz variância-covariância para as medidas repetidas.
- **R.list**, lista de matrizes de variância-covariância para as medidas repetidas, se consideradas diferentes nos dois grupos.
- **sig.level**, erro Tipo I.
- **power**, poder.
- **Pi**, a proporção de alocação em cada grupo.
- **alternative**, teste uni ou bicaudal.
- **tol**, tolerância numérica usada na localização de raízes, por padrão `.Machine$double.eps2`, usualmente, sem necessidade de modificações.

3.3.3 Exemplo

Adaptado de [Liu e Liang \(1997\)](#). Uma pesquisa deseja testar a eficácia de um novo fármaco (tratamento = A) que auxilia na diminuição da perda de cabelo frente à droga já existente no mercado (controle = B), durante um certo período de tratamento. Para isso, deve-se calcular o número de pacientes necessários para validar os resultados com um poder de 0,9 e uma significância de 0,05 em um teste bicaudal.

A diferença mínima esperada entre os grupos é uma taxa média de diminuição de 0,5% na queda de cabelo, tendo sido planejadas 4 medições, ao longo de um ano de uso dos produtos. Serão testados dois tipos de alocação, a uniforme e a com 80% dos pacientes recebendo a nova droga. Foi assumida uma estrutura de correlação permutável, indexada pelo coeficiente de correlação intra-sujeito ($\rho = 0,7$) e com a diagonal principal composta de uns.

Um modelo GEE com estrutura de correlação permutável (exchangeable) equivale a um modelo misto de intercepto aleatório ([Donohue et al., 2019](#)).

O modelo misto deste exemplo seria expresso por:

$$Y_{ij} = (\beta_{0A} + \beta_{0B}Grupo_i + b_{0i}) + (\beta_{1A} + \beta_{1B}Grupo_i)x_{ij} + \epsilon_{ij} \quad (3.9)$$

Onde:

Y_{ij} = diminuição observada na queda de cabelo do indivíduo i na medição j ;

$Grupo_i$ = grupo a que pertence o indivíduo i (A se do grupo tratamento e B se do grupo controle);

β_{0A} = intercepto da reta média do tratamento;

β_{0B} = diferença no intercepto da reta média do controle em relação ao tratamento;

b_{0i} = efeito aleatório de intercepto $\rightarrow b_{0i} \sim N(0, \sigma_b^2)$;

β_{1A} = inclinação da reta média do tratamento;

β_{1B} = diferença na inclinação da reta média do controle em relação ao tratamento;

x_{ij} = medição j do indivíduo i ;

ϵ_{ij} = efeito aleatório de indivíduo $\rightarrow \epsilon_{ij} \sim N(0, (\sigma)^2)$.

A estrutura de correlação utilizada:

$$\mathbf{R} = \begin{bmatrix} 1 & 0,7 & 0,7 & 0,7 \\ 0,7 & 1 & 0,7 & 0,7 \\ 0,7 & 0,7 & 1 & 0,7 \\ 0,7 & 0,7 & 0,7 & 1 \end{bmatrix}$$

As hipóteses testadas são:

$$\begin{aligned} H_0 : \beta_{1B} &= 0 \\ H_1 : \beta_{1B} &\neq 0 \end{aligned} \quad (3.10)$$

Ao reescrever o modelo na forma dos modelos GEE para melhor entendimento dos argumentos desta função, tem-se na notação de [Diggle et al. \(1994\)](#) que este modelo assume que a resposta para cada indivíduo i no tempo j , (Y_{ij}) é a média do grupo, dependente do tempo (t_{ij}) e tratamento (A ou B), além do termo de erro (ϵ_{ij}). Ou seja, respectivamente para os grupos A e B:

$$\begin{aligned} Y_{Aij} &= \beta_{0A} + \beta_{1A}t_{ij} + \epsilon_{ij} \\ Y_{Bij} &= \beta_{0B} + \beta_{1B}t_{ij} + \epsilon_{ij} \end{aligned} \quad (3.11)$$

Onde:

β_{0A} = parâmetro do intercepto da reta do grupo A;

β_{1A} = parâmetro da inclinação da reta do grupo A;

β_{0B} = parâmetro do intercepto da reta do grupo B;

β_{1B} = parâmetro da inclinação da reta do grupo B.

A hipótese a ser testada a respeito da igualdade das inclinações das duas retas é:

$$H_0 : d = \beta_{1A} - \beta_{1B} = 0 \quad (3.12)$$

Neste exemplo, os parâmetros de interesse (aqueles que serão testados) são β_{1A} e β_{1B} , logo o argumento 'u' (vetor u) da função deverá ser $u_1 = 1$, (referindo-se ao grupo A) e $u_2 = 0$, (referindo-se ao grupo B). Os parâmetros de distúrbio (aqueles que não serão testados) são β_{0A} e β_{0B} , logo o argumento 'v' (vetor v) da função deverá receber valores iguais para os dois grupos. Por fim, os argumentos u, v e Pi devem ter o mesmo comprimento e ser chamados de forma ordenada entre si. Informações mais detalhadas são encontradas em [Donohue \(2019\)](#).

No RStudio:

```
> ##Alocacao desbalanceada (0.8/0.2)
> R <- matrix(0.7, nrow = 4, ncol = 4)
> diag(R) <- 1
> u = list(u1 = rep(1, 4), # tratamento
+         u2 = rep(0, 4)) # controle
>
> v = list(v1 = rep(1, 4), v2 = rep(1, 4)) # intercepto
> Pi = c(0.8, 0.2)
>
> liu.liang.linear.power(delta=0.5, u=u, v=v, R=R, sig.level=0.05, power=0.90, Pi=Pi)

      Longitudinal linear model power calculation (Liu & Liang, 1997)
              N = 203.5813
              n = 162.86506, 40.71626
              delta = 0.5
              sigma2 = 1
              sig.level = 0.05
              power = 0.9
              alternative = two.sided

NOTE: N is total sample size and n is sample size in each group.
R:
      [,1] [,2] [,3] [,4]
[1,]  1.0  0.7  0.7  0.7
[2,]  0.7  1.0  0.7  0.7
[3,]  0.7  0.7  1.0  0.7
[4,]  0.7  0.7  0.7  1.0

> ##Alocacao balanceada (0.5/0.5)
> R <- matrix(0.7, nrow = 4, ncol = 4)
> diag(R) <- 1
>
> u = list(u1 = rep(1, 4), # tratamento
+         u2 = rep(0, 4)) # controle
>
> v = list(v1 = rep(1, 4), v2 = rep(1, 4)) # intercepto
>
> liu.liang.linear.power(delta=0.5, u=u, v=v, R=R, sig.level=0.05, power=0.90)
```

```

Longitudinal linear model power calculation (Liu & Liang, 1997)
  N = 130.292
  n = 65.14602, 65.14602
  delta = 0.5
  sigma2 = 1
  sig.level = 0.05
  power = 0.9
  alternative = two.sided

NOTE: N is total sample size and n is sample size in each group.
R:
  [,1] [,2] [,3] [,4]
[1,]  1.0  0.7  0.7  0.7
[2,]  0.7  1.0  0.7  0.7
[3,]  0.7  0.7  1.0  0.7
[4,]  0.7  0.7  0.7  1.0
>

```

Portanto, o estudo exigiria cerca de 204 indivíduos no delineamento com alocação 0,8|0,2, sendo que destes 163 receberiam a nova droga, enquanto 41 receberiam a já existente no mercado. Em comparação, no delineamento 0,5|0,5 menos pacientes seriam necessários, apenas 130 divididos entre os dois grupos. A função ainda retornou as duas matrizes de correlações, que são iguais, já que essa definição não mudou para os dois exemplos.

3.4 Função *lmpower*

3.4.1 Aplicação

A função *lmpower* calcula o tamanho da amostra para modelos mistos lineares com base nas fórmulas de [Diggle et al. \(2002\)](#) ou [Liu e Liang \(1997\)](#), que são expressas em termos de modelo marginal ou parâmetros de Equações de Estimação Generalizadas (GEE).

Essas funções convertem parâmetros do modelo piloto de efeito misto (por exemplo, intercepto aleatório e / ou inclinação, efeitos fixos, etc.) em parâmetros do modelo marginal, para que qualquer fórmula possa ser aplicada a um efeito equivalente.

Presume-se que as estimativas piloto (das funções *lmer*, *lme* ou *gee*) sejam de um grupo "placebo" apropriado e o parâmetro de interesse seja a taxa de mudança ao longo do tempo observado da variável resposta. Nem todos os parâmetros são utilizados sempre, seu uso está condicionado ao delineamento experimental escolhido.

3.4.2 Argumentos

Para a função *lmpower* são necessários os seguintes argumentos:

- **object**, objeto retornado pelo pacote *lme4*.
- **n**, tamanho da amostra por grupo de um objeto de modelo de efeitos mistos para dados de placebo.
- **parameter**, o nome ou a posição do parâmetro de taxa de variação de interesse, por exemplo ("hora", "t" ou 2 se for o segundo efeito fixo especificado).
- **pct.change**, a variação percentual na estimativa piloto do parâmetro de interesse (beta, efeito placebo / nulo).
- **delta**, a alteração na estimativa piloto do parâmetro de interesse, calculada a partir de **pct.change**, se deixada em falta.

- **t**, vetor de tempos observados.
- **sig.level**, erro Tipo I.
- **power**, poder.
- **alternative**, teste uni ou bicaudal.
- **beta**, estimativa piloto do efeito placebo (inclinação ou taxa de variação na variável resposta).
- **beta.CI**, limites de confiança de 95% da estimativa piloto de beta.
- **delta.CI**, limites de confiança de 95% do tamanho do efeito.
- **sig2.i**, estimativa da variância do efeito aleatório do intercepto.
- **sig2.s**, estimativa da variância do efeito aleatório de inclinação.
- **sig2.e**, estimativa da variância residual.
- **cov.s.i**, estimativa da covariância entre inclinação e intercepto.
- **R**, estimativa da matriz de correlação de trabalho modelo marginal.
- **method**, a fórmula a ser usada. O padrão é "diggle" para [Diggle et al. \(2002\)](#). Alternativamente, "liuliang" e "edland" podem ser setados para [Liu e Liang \(1997\)](#) e [Edland \(2009\)](#).
- **tol**, tolerância numérica usada na localização de raízes, por padrão $.Machine\$double.eps^2$, usualmente, sem necessidade de modificações.
- ..., outros argumentos.

3.4.3 Exemplo

Considerando o mesmo cenário do exemplo de *edland.linear.power*, em que deseja-se estimar o tamanho da amostra necessário para conduzir um estudo sobre a doença de Alzheimer. Embora agora o modelo assumido tenha inclinação e intercepto aleatórios, ainda são realizadas sete medições durante 18 meses, buscando encontrar diferenças mínimas de 1,5 pontos na escala de avaliação.

Pelos estudos pilotos foram tomadas como 24, 55 e 10 as variâncias da inclinação aleatória, do intercepto aleatório e a residual, respectivamente. Dado um coeficiente de correlação de 0,8, o valor da covariância entre o intercepto e a inclinação é aproximadamente 29. Deseja-se alcançar um poder de 0,8 em um teste bicaudal de 0,05 de significância.

O modelo é expresso por:

$$Y_{ij} = (\beta_{0A} + \beta_{0B}Grupo_i + b_{0i}) + (\beta_{1A} + \beta_{1B}Grupo_i + b_{1i})x_{ij} + \epsilon_{ij} \quad (3.13)$$

Onde:

Y_{ij} = pontos observados do indivíduo i na avaliação j ;

$Grupo_i$ = grupo a que pertence o indivíduo i (A se do grupo tratamento e B se do grupo controle);

β_{0A} = intercepto da reta média do grupo tratamento;

β_{0B} = diferença no intercepto da reta média do grupo controle em relação ao grupo tratamento;

b_{0i} = efeito aleatório de intercepto $\rightarrow b_{0j} \sim N(0, (\sigma_{b0})^2)$;

β_{1A} = inclinação da reta média do grupo tratamento;

β_{1B} = diferença na inclinação da reta média do grupo controle em relação ao grupo tratamento;

b_{1i} = efeito aleatório de inclinação $\rightarrow b_{1j} \sim N(0, (\sigma_{b1})^2)$;

x_{ij} = avaliação j do indivíduo i;
 ϵ_{ij} = efeito aleatório de indivíduo $\rightarrow \epsilon_{ij} \sim N(0, (\sigma)^2)$.

As hipóteses testadas são:

$$\begin{aligned} H_0 : \beta_{1B} &= 0 \\ H_1 : \beta_{1B} &\neq 0 \end{aligned} \tag{3.14}$$

No RStudio:

```
> lmpower(delta=1.5, t = c(0.00, 0.25, 0.50, 0.75, 1.00, 1.25, 1.50), sig2.i = 55,
sig2.s = 24, sig2.e = 10, cov.s.i=0.8*sqrt(55)*sqrt(24), power = 0.80, sig.level = 0.05)
```

```
Power for longitudinal linear model with random slope (Edland, 2009)
```

```
      n = 207.3101
    delta = 1.5
  sig2.s = 24
  sig2.e = 10
sig.level = 0.05
      t = 0.00, 0.25, 0.50, 0.75, 1.00, 1.25, 1.50
    power = 0.8
alternative = two.sided
```

```
NOTE: n is number in *each* group
```

A função concorda com o exemplo anterior, alocando 208 pessoas em cada um dos grupos.

3.5 Função *power.mmrn.ar1*

3.5.1 Aplicação

Esta função executa o cálculo do tamanho de amostra/poder para testar diferenças médias entre dois grupos em um modelo misto de medidas repetidas com a estrutura de correlação autorregressiva de primeira ordem, $\mathbf{R} = \text{AR}(1)$, ao final do período de acompanhamento. Quando é explicitado o N, a função devolve o poder a ser alcançado, caso contrário, o número de amostra estimado por grupo é retornado. O tempo é categórico e são aceitas proporções de alocação de sujeitos diferentes entre os grupos.

O modelo é expresso como:

$$Y_i = X_i\beta + Z_iu_i + \epsilon_i \tag{3.15}$$

Onde:

Y_i = vetor $nx1$ de dados observados;

X_i = matriz $n \times p$ de efeitos fixos;

β = vetor $px1$ de efeitos fixos, geralmente desconhecidos;

Z_i = matriz $n_i \times q$ de efeitos aleatórios;

u_i = vetor $qx1$ de efeitos aleatórios, geralmente desconhecidos;

ϵ_i = vetor $nx1$ de erros aleatórios, não observáveis.

As hipóteses testadas são:

$$\begin{aligned} H_0 : \mu_{A.t} - \mu_{B.t} &= 0 \\ H_1 : \mu_{A.t} - \mu_{B.t} &\neq 0 \end{aligned} \tag{3.16}$$

Onde:

$\mu_{A.t}$ = resposta média do grupo A no tempo final t;

$\mu_{B.t}$ = resposta média do grupo B no tempo final t.

A matriz com estrutura de correção AR(1) é, considerando quatro medidas de tempo:

$$\mathbf{AR}(1) = \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}$$

3.5.2 Argumentos

Para a função *power.mmrn.ar1* são necessários os seguintes argumentos:

- **N**, tamanho total da amostra.
- **rho**, parâmetro de correlação AR(1).
- **ra**, retenção no grupo a (% de sujeitos que ainda estão no estudo em dado ponto no tempo).
- **sigmaa**, desvio padrão da observação de interesse no grupo a.
- **rb**, retenção no grupo b, se deixada em branco é assumido igual à ra (% de sujeitos que ainda estão no estudo em dado ponto no tempo).
- **sigmab**, desvio padrão da observação de interesse no grupo b. Caso NULL, sigmab é considerado igual à sigmaa. Caso não NULL, sigmaa e sigmab são calculados como média.
- **lambda**, taxa de alocação.
- **times**, tempos observados.
- **delta**, tamanho do efeito.
- **sig.level**, erro Tipo I.
- **power**, poder.
- **alternative**, teste uni ou bicaudal.
- **tol**, tolerância numérica usada na localização de raízes, por padrão `.Machine$double.eps2`, usualmente, sem necessidade de modificações.

3.5.3 Exemplo

Adaptado de [Lu et al. \(2008\)](#). Supõe-se um ensaio clínico com dois grupos (tratamento e placebo) e quatro momentos de medições, em que o objetivo é avaliar o efeito de tratamento ao final do período de acompanhamento. A última medição possui uma estrutura de correlação AR(1) com $\rho = 0,6$, e o vetor de desvios padrões é de (0,7; 0,8; 0,9 e 1). O vetor de médias do grupo placebo foi assumido (0; 0; 0; 0), enquanto o do grupo tratamento é (0,3; 0,5; 0,8; 0,9). Foram realizados quatro ensaios, para estimar tanto os tamanhos de amostra quanto o poder dos testes.

$$\mathbf{R} = \begin{bmatrix} 1 & 0,6 & 0,6^2 & 0,6^3 \\ 0,6 & 1 & 0,6 & 0,6^2 \\ 0,6^2 & 0,6 & 1 & 0,6 \\ 0,6^3 & 0,6^2 & 0,6 & 1 \end{bmatrix}$$

As taxas de retenção durante o estudo são estimadas em (100%, 76%, 63%, 52%) e (100%, 87%, 81%, 78%) nos grupos placebo e tratamento, respectivamente. Essas taxas são provenientes de um estudo prévio, em que se deu a retirada de pacientes do estudo em uma taxa de 90%, visto que suas medições ficavam abaixo do limiar do protocolo pré-estabelecido (MAR - missing at random).

Deseja-se detectar uma diferença de 0,9 entre as médias dos grupos no último tempo de medição, com nível de significância 0,05, assumindo proporções tanto iguais, quanto diferentes de alocação de amostra. Os dois primeiros testes devolvem o número de sujeitos por grupo para um poder de 0,9, enquanto os últimos retornam o poder para tamanhos de amostras de 80 e 84, respectivamente.

No RStudio:

```
> ra <- c(1.00, 0.76, 0.63, 0.52)
> rb <- c(1.00, 0.87, 0.81, 0.78)
> rho <- 0.6
> power.mmr.ar1(rho=rho, ra=ra, sigmaa=1, rb = rb,lambda = 1, power = 0.9, delta = 0.9)

Power for Mixed Model of Repeated Measures (Lu, Luo, & Chen, 2008)

      n1 = 38.90636
      n2 = 38.90636
      rho = 0.6
retention1 = 1.00, 0.76, 0.63, 0.52
retention2 = 1.00, 0.87, 0.81, 0.78
      phi1 = 1.752276
      phi2 = 1.246951
      delta = 0.9
      times = 1, 2, 3, 4
sig.level = 0.05
      power = 0.9
alternative = two.sided

> power.mmr.ar1(rho=rho , ra=ra, sigmaa=1, rb = rb,lambda = 2, power = 0.9, delta = 0.9)

Power for Mixed Model of Repeated Measures (Lu, Luo, & Chen, 2008)

      n1 = 55.08197
      n2 = 27.54099
      rho = 0.6
retention1 = 1.00, 0.76, 0.63, 0.52
retention2 = 1.00, 0.87, 0.81, 0.78
      phi1 = 1.752276
      phi2 = 1.246951
      delta = 0.9
      times = 1, 2, 3, 4
sig.level = 0.05
      power = 0.9
alternative = two.sided

> power.mmr.ar1(N=80, rho=rho, ra=ra, sigmaa=1, rb = rb,lambda = 1, delta = 0.9)

Power for Mixed Model of Repeated Measures (Lu, Luo, & Chen, 2008)

      n1 = 40
      n2 = 40
      rho = 0.6
retention1 = 1.00, 0.76, 0.63, 0.52
retention2 = 1.00, 0.87, 0.81, 0.78
      phi1 = 1.752276
      phi2 = 1.246951
      delta = 0.9
      times = 1, 2, 3, 4
sig.level = 0.05
      power = 0.9077117
alternative = two.sided

> power.mmr.ar1(N=84, rho=rho, ra=ra, sigmaa=1, rb = rb,lambda = 2, delta = 0.9)
```



```

Power for Mixed Model of Repeated Measures (Lu, Luo, & Chen, 2008)

      n1 = 56
      n2 = 28
      rho = 0.6
retention1 = 1.00, 0.76, 0.63, 0.52
retention2 = 1.00, 0.87, 0.81, 0.78
      phi1 = 1.752276
      phi2 = 1.246951
      delta = 0.9
      times = 1, 2, 3, 4
sig.level = 0.05
      power = 0.9046401
alternative = two.sided

```

No terceiro cenário, apresenta-se o delineamento com maior poder, quase 0,91, e menor tamanho de amostra por grupo, 40 indivíduos, o que seria o cenário ideal quando se pensa sob a óptica dos custos de pesquisa.

3.6 Função *power.mmrn*

3.6.1 Aplicação

Esta função executa uma fórmula de cálculo de tamanho de amostra/poder baseada em uma família de análises de medidas repetidas, frequentemente referidas como modelo misto para medidas repetidas (MMRM), advindo de (Lu et al., 2008), com foco em comparar diferenças médias entre tratamentos no último momento da coleta de dados, quando a variável de tempo é considerada categórica.

O método MMRM inclui uma modelagem não estruturada de tempo e a estrutura de correlação de erros dentro do sujeito. Especificamente, o modelo inclui apenas os efeitos principais de tratamento e tempo e a sua interação, sendo usada quaisquer matriz de covariâncias para modelar a correlação entre as medidas repetidas (AR1, não-estruturada, permutável, etc.).

O modelo é expresso como:

$$Y_i = X_i\beta + Z_iu_i + \epsilon_i \quad (3.17)$$

Onde:

Y_i = vetor $nx1$ de dados observados;

X_i = matriz $n \times p$ de efeitos fixos;

β = vetor $px1$ de efeitos fixos, geralmente desconhecidos;

Z_i = matriz $n_i \times q$ de efeitos aleatórios ;

u_i = vetor $qx1$ de efeitos aleatórios, geralmente desconhecidos;

ϵ_i = vetor $nx1$ de erros aleatórios, não observáveis.

As hipóteses testadas são:

$$\begin{aligned} H_0 : \mu_{A.t} - \mu_{B.t} &= 0 \\ H_1 : \mu_{A.t} - \mu_{B.t} &\neq 0 \end{aligned} \quad (3.18)$$

Onde:

$\mu_{A.t}$ = resposta média do grupo A no tempo final t;

$\mu_{B.t}$ = resposta média do grupo B no tempo final t.

3.6.2 Argumentos

Para a função `power.mmrn` são necessários os seguintes argumentos:

- **N**, tamanho total da amostra.
- **Ra**, matriz de correlação para o grupo a.
- **ra**, retenção no grupo a (% de sujeitos que ainda estão no estudo em dado ponto no tempo).
- **sigmaa**, desvio padrão da observação de interesse no grupo a.
- **Rb**, matriz de correlação para o grupo b.
- **rb**, retenção no grupo b (% de sujeitos que ainda estão no estudo em dado ponto no tempo).
- **sigmab**, desvio padrão da observação de interesse no grupo b. Caso NULL, sigmab é considerado igual à sigmaa. Caso não NULL, sigmaa e sigmab são calculados como média.
- **lambda**, taxa de alocação.
- **delta**, tamanho do efeito.
- **sig.level**, erro Tipo I.
- **power**, poder.
- **alternative**, teste uni ou bicaudal.
- **tol**, tolerância numérica usada na localização de raízes, por padrão `.Machine$double.eps2`, usualmente, sem necessidade de modificações.

3.6.3 Exemplo

Neste exemplo será utilizado o mesmo cenário do exemplo anterior (Lu et al., 2008), porém, agora o modelo conta com uma estrutura de correlação permutável, em que $\rho = 0,25$ e taxas de retenção iguais entre os grupos controle (A) e tratamento (B). As medições ainda serão feitas em quatro momentos, em que o objetivo é avaliar o efeito do tratamento ao final do período de observação, quando são comparadas as médias dos grupos. As taxas de retenção em cada momento, para o grupo controle, são dadas por $ra = (1; 0,9; 0,8; 0,7)$.

$$\mathbf{R} = \begin{bmatrix} 1 & 0,25 & 0,25 & 0,25 \\ 0,25 & 1 & 0,25 & 0,25 \\ 0,25 & 0,25 & 1 & 0,25 \\ 0,25 & 0,25 & 0,25 & 1 \end{bmatrix}$$

Deseja-se detectar uma diferença de 0,5 entre as médias dos grupos em relação ao momento final de medição, tendo um nível de significância de 0,05 e assumindo proporções tanto iguais, quanto diferentes de alocação de amostra.

Novamente foram realizados quatro testes, onde os dois primeiros ensaios devolvem o poder e tamanho de amostra para alocação igual de sujeitos nos grupos, enquanto os últimos também o fazem para uma taxa de alocação desbalanceada. Os testes que devolvem o poder receberam 174 indivíduos, e aqueles que devolvem a amostra para os grupos receberam um poder de 0,8.

No RStudio:

```

> Ra <- matrix(0.25, nrow = 4, ncol = 4)
> diag(Ra) <- 1
> ra <- c(1, 0.90, 0.80, 0.70)
> sigmaa <- 1
> power.mmrn(Ra = Ra, ra = ra, sigmaa = sigmaa, delta = 0.5, power = 0.80)

Power for Mixed Model of Repeated Measures (Lu, Luo, & Chen, 2008)

      n1 = 86.99175
      n2 = 86.99175
retention1 = 1.0, 0.9, 0.8, 0.7
retention2 = 1.0, 0.9, 0.8, 0.7
      delta = 0.5
      sig.level = 0.05
      power = 0.8
alternative = two.sided

> power.mmrn(N = 174, Ra = Ra, ra = ra, sigmaa = sigmaa, delta = 0.5)

Power for Mixed Model of Repeated Measures (Lu, Luo, & Chen, 2008)

      n1 = 87
      n2 = 87
retention1 = 1.0, 0.9, 0.8, 0.7
retention2 = 1.0, 0.9, 0.8, 0.7
      delta = 0.5
      sig.level = 0.05
      power = 0.8000372
alternative = two.sided

> power.mmrn(Ra = Ra, ra = ra, sigmaa = sigmaa, delta = 0.5, power = 0.80, lambda = 2)

Power for Mixed Model of Repeated Measures (Lu, Luo, & Chen, 2008)

      n1 = 130.4876
      n2 = 65.24381
retention1 = 1.0, 0.9, 0.8, 0.7
retention2 = 1.0, 0.9, 0.8, 0.7
      delta = 0.5
      sig.level = 0.05
      power = 0.8
alternative = two.sided

> power.mmrn(N = 174, Ra = Ra, ra = ra, sigmaa = sigmaa, delta = 0.5, lambda = 2)

Power for Mixed Model of Repeated Measures (Lu, Luo, & Chen, 2008)

      n1 = 116
      n2 = 58
retention1 = 1.0, 0.9, 0.8, 0.7
retention2 = 1.0, 0.9, 0.8, 0.7
      delta = 0.5
      sig.level = 0.05
      power = 0.7522291
alternative = two.sided

```

Comparando os resultados dos cenários que envolvem o número de sujeitos por grupo, percebe-se que alocação desbalanceada faz requer um número total muito maior de indivíduos (195) do que a balanceada (174), para o mesmo nível de poder.

4 Análise de Sobrevivência

A análise de sobrevivência, conforme [Colosimo e Giolo \(2006\)](#), é uma das áreas da estatística que mais cresceu nas últimas duas décadas do século passado. [Bastos e Rocha \(2006\)](#) definem Análise de Sobrevivência como a análise estatística de dados quando a variável em estudo representa o tempo desde um instante inicial bem definido até à ocorrência de determinado acontecimento de interesse (tempo de falha) ou censura. O delineamento epidemiológico pode ser observacional (coorte) ou experimental (ensaio clínico randomizado), outros tipos de delineamentos experimentais também podem ser utilizados principalmente no contexto de experimentos industriais que visam a melhoria de processos ou produtos, todos mantêm o acompanhamento dos indivíduos por um determinado período de tempo.

A variável aleatória em estudo, então, é não negativa e pode representar, por exemplo, o tempo até à falha de determinada componente elétrica no âmbito da Fiabilidade Industrial, a duração de uma greve ou período de desemprego no contexto da Economia, o tempo de resposta a um inquérito no âmbito da Psicologia ou o tempo até à morte de um indivíduo com determinada doença em Medicina.

A principal característica dos dados de sobrevivência é a presença de censura, que é a observação parcial da resposta, sendo situações em que, por alguma razão, o acompanhamento do paciente foi interrompido ([Colosimo e Giolo, 2006](#)). Não havendo a presença de censura, técnicas clássicas como análise de regressão e planeamento de experimentos poderiam ser aplicadas para a análise desse tipo de dados.

Assim, o tempo de falha e a censura são suas principais características:

O tempo de falha é composto pelo *tempo inicial*, data precisa do início do estudo; pela *escala de medida*, geralmente, o tempo real ou "do relógio"; e pelo *evento de interesse*, usualmente são casos indesejados, definidos de forma clara previamente.

As censuras podem ocorrer por várias razões, como a perda de acompanhamento do paciente no decorrer do estudo ou a não ocorrência do evento de interesse até o término do experimento, podendo ser classificadas em três tipos: censura à direita, à esquerda e intervalar. **À direita** é a censura em que o tempo entre o início do estudo e a ocorrência do evento é maior do que o tempo registrado, já **à esquerda** ocorre quando o momento da ocorrência do evento não é conhecido, mas é sabido que este ocorreu antes do tempo registrado. Assim, a **intervalar** é um tipo mais geral de censura ocorrendo, por exemplo, em estudos nos quais os pacientes são acompanhados em visitas periódicas e é sabido somente que o evento de interesse ocorreu em um certo intervalo de tempo.

Alguns outros mecanismos de censura são diferenciados em estudos clínicos. **Censura do tipo I** é aquela em que o estudo será terminado após um período

pré-estabelecido de tempo. **Censura do tipo II** é aquela em que o estudo será terminado após ter ocorrido o evento de interesse em um número pré-estabelecido de indivíduos. Um terceiro mecanismo de censura, o do **tipo aleatório**, é o que mais ocorre na prática médica, mas não se restringe à ela, e se dá quando um paciente é retirado no decorrer do estudo sem ter ocorrido a falha (por exemplo, morte por outra razão que não a doença observada) [Colosimo e Giolo \(2006\)](#).

Função de Sobrevivência

Dado tempo de sobrevivência (T) uma variável aleatória contínua e não negativa, a função de sobrevivência é definida como a probabilidade de uma observação não falhar até certo tempo t , ou seja, a probabilidade de uma observação sobreviver ao tempo t , expressa por:

$$S(t) = P(T \geq t) \quad (4.1)$$

Relacionada à função $S(t)$, a função taxa de falha ou taxa de risco $\lambda(t)$ mede a probabilidade do evento de interesse ocorrer dentro de um intervalo $[t, t + \delta t)$, dado que não ocorreu antes de t , expressa por:

$$\lambda(t) = \frac{S(t) - S(t + \delta t)}{\delta t S(t)} \quad (4.2)$$

Em que $S(t)$ e $S(t + \delta t)$ representam os valores das funções de sobrevivência nos tempos t e δt , respectivamente, e δt representa o tamanho do intervalo de tempo. Quando assumido δt bem pequeno, $\lambda(t)$ representa a taxa de falha instantânea no tempo t dado que o evento não ocorreu até t . As funções $S(t)$ e $\lambda(t)$ são inversamente proporcionais, isto é, quando o risco aumenta, a probabilidade de sobrevivência diminui.

Regressão de Cox

Dentre todas as técnicas de análise de sobrevivência, o modelo de Cox ([Cox, 1972](#)) está entre as mais utilizadas ([Bessegato, 2014](#)). O interesse é modelar o efeito de covariáveis sobre o tempo de sobrevivência (risco), considerando os riscos proporcionais, o que significa que as razões entre as taxas de risco para indivíduos com diferentes valores ou em diferentes categorias das variáveis explicativas x são assumidas como independentes do tempo (variável contínua).

Sua forma mais simples se dá quando a variável resposta é explicada apenas pela covariável indicadora de grupo, tomando por exemplo, tratamento padrão (grupo 0) e novo tratamento (grupo 1). As funções de risco são, para os grupos respectivamente, $\lambda_0(t)$ e $\lambda_1(t)$, então, pela propriedade de proporcionalidade, satisfaz:

$$\frac{\lambda_0(t)}{\lambda_1(t)} = K, \forall t \quad (4.3)$$

A forma do modelo considerando p covariáveis, usa um risco basal $\lambda_0(t)$ e o vetor de covariáveis x :

$$\lambda(t) = \lambda_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p) \quad (4.4)$$

Os parâmetros β_i indicam o efeito de cada variável x_i sobre a função de risco, isto é, mede o quanto uma covariável pode acelerar ou desacelerar a função de risco (Fontoura, 2018).

O modelo de Cox é dito semi-paramétrico, já que supõe que as covariáveis agem multiplicativamente (componente paramétrico) e não assume distribuição para o tempo de sobrevivência na estimação dos efeitos das covariáveis nem na função de risco basal (componente não paramétrico).

5 Pacote *HMISC*

O pacote *HMISC*, é uma miscelânea de funções úteis para análise de dados, gráficos de alto nível, operações utilitárias, funções para cálculo de tamanho de amostra e poder, importação e criação datasets, imputação de valores ausentes, criação avançada de tabelas, clusterização de variáveis, manipulação de cadeias de caracteres, recodificação de variáveis, conversão de objetos em R para códigos LaTeX e html, entre outros. Frank Harrell, da Universidade de Vanderbilt, é seu principal autor e mantenedor, contando ainda com outros contribuintes em sua implementação. As funções aqui sobre esse pacote serão dez, *samplesize.bin*, *posamsize*, *popower*, *bpower*, *bpower.sim*, *bsamsize*, *ballocation*, *cpower*, *ciapower* e *spower*, das quais apenas a primeira não foi implementada por Harrell, mas sim por Rick Chappell, da Universidade de Wisconsin em Madison, tratando-se quatro destas de tamanhos de amostra, enquanto as demais abordam a questão do poder da amostra.

As funções *cpower*, *ciapower* e *spower* tratam de amostragem dentro da área de dados de sobrevivência, fazendo estimativas para modelos de Cox e log-rank. Enquanto isso, as demais funções executam procedimentos para comparação entre duas amostras com respostas binomiais ou ordinais.

Os códigos implementados no R das funções descritas a seguir podem ser encontrados em <https://github.com/harrelfe/Hmisc/tree/master/R>, repositório do GitHub criado pelo autor do pacote e sustentado com comentários, dúvidas e exemplos da comunidade interessada.

5.1 Função *samplesize.bin*

5.1.1 Aplicação

Essa função utiliza uma aproximação do teste exato de Fisher para calcular o tamanho da amostra (ou tamanhos de amostras) necessário para testar a diferença entre duas proporções de um evento de interesse, assumindo observações independentes. Essa aproximação é a transformação seno-inverso, usada para estabilizar a variância, mais detalhes em (Bell et al., 2014).

A função performa as seguintes fórmulas de Meinert e Tonascia (1986):

Quando $\lambda = 1$, alocação uniforme:

$$\begin{aligned} n_c &= \frac{(Z_\alpha + Z_\beta)^2}{2(\sin^{-1}\sqrt{P_c} - \sin^{-1}\sqrt{P_t})^2} \\ n_t &= n_c \\ N &= (r + 1)n_c \end{aligned} \tag{5.1}$$

Quando $\lambda \neq 1$, alocação não-uniforme:

$$\begin{aligned} n_c &= \frac{(Z_\alpha + Z_\beta)^2 + (\lambda + 1)/\lambda}{4(\sin^{-1}\sqrt{P_c} - \sin^{-1}\sqrt{P_t})^2} \\ n_t &= \lambda n_c \\ N &= r n_t + n_c \end{aligned} \tag{5.2}$$

Onde:

$\lambda = n_t/n_c$, proporção entre o número de pacientes do grupo tratamento e o número de pacientes do grupo controle;

n_c , tamanho de amostra necessário para o grupo controle;

n_t , tamanho de amostra necessário para o grupo tratamento;

N , tamanho total da amostra necessário em todos os grupos combinados;

α , probabilidade de erro Tipo I;

β , probabilidade de erro Tipo II;

Z_α , ponto na abcissa de uma curva N(0,1) até a direita, onde é encontrada 100(α)% da área total sob a curva;

Z_β , ponto na abcissa de uma curva N(0,1) até a direita, onde é encontrada 100(β)% da área total sob a curva;

P_c , taxa de eventos assumidos (expressa em proporção) para o resultado de interesse no grupo controle;

P_t , taxa de eventos assumidos (expressa em proporção) para o resultado de interesse no grupo tratamento;

$r =$ número de tratamentos.

As hipóteses testadas são:

$$\begin{aligned} H_0 &: P_c = P_t \\ H_1 &: P_c \neq P_t \end{aligned} \tag{5.3}$$

Em que P_c e P_t representam as probabilidades de sucesso (ocorrência do evento de interesse) nos grupos controle e tratamento.

5.1.2 Argumentos

Para a função *samplesize.bin* são necessários os seguintes argumentos:

- **alpha**, teste uni ou bicaudal.
- **beta**, poder.
- **pit**, probabilidade de sucesso (ocorrência do evento de interesse) hipotética no grupo tratamento.

- **pic**, probabilidade de sucesso (ocorrência do evento de interesse) hipotética no grupo controle.
- **rho**, proporção da amostra dedicada ao grupo tratamento, caso não declarado $\rho = 0,5$ ($0 < \rho < 1$).

5.1.3 Exemplo

Um laboratório deseja testar a efetividade de um novo fármaco para pacientes com problemas de acne severa. A probabilidade de cura no grupo controle (fármaco já existente) foi estimada em 50% por estudos prévios, e espera-se que o novo tratamento eleve essa chance para 65%. Assume-se um nível de significância de 5% (unicaudal), fixa-se o poder do teste em 0,8, e considera-se o número de pacientes igual em ambos os grupos ($\rho = 0,5$).

No RStudio:

```
> samplesize.bin(alpha = 0.05, beta = 0.8, pit = 0.65, pic = 0.5, rho = 0.5)
[1] 267
```

Estima-se que serão necessários 267 pacientes para esse estudo, divididos igualmente entre controle e tratamento.

5.2 Função *posamsize*

5.2.1 Aplicação

A função *posamsize* calcula o tamanho da amostra necessário para se realizar um teste de hipótese para resultados ordinais, comparando dois grupos independentes, com diferentes tamanhos. Foi implementada por Harrell e usa o método de [Whitehead \(1993\)](#), que considera as razões de chances como tamanho de efeito, podendo ser considerada como uma abordagem 'não paramétrica', pois a derivação das fórmulas de tamanho da amostra e a análise dos dados são baseadas no teste U de Mann-Whitney com permissão para empates ([Campbell e Walters, 2014](#)). A *posamsize* ainda retorna eficiência de seu delineamento em comparação com um modelo no qual a variável resposta é contínua.

São implementados dois cálculos nessa função, $OR_{Ordinal_i}$, que calcula a razão de chances e $n_{IndividualOrdinal}$ que fornece o número de sujeitos por grupo, para um nível de significância α bicaudal e poder $1 - \beta$, dados pelas equações:

$$OR_{Ordinal_i} = \frac{\frac{CP_{iT}}{(1-CP_{iT})}}{\frac{CP_{iC}}{(1-CP_{iC})}} \quad (5.4)$$

$$n_{IndividualOrdinal} = \frac{6 \left[\frac{(z_{1-\alpha/2} + z_{1-\beta})^2}{(\log OR_{Ordinal})^2} \right]}{[1 - \sum_{i=1}^k \bar{p}_i^3]} \quad (5.5)$$

Onde:

i , categorias ordenadas da variável de interesse, $i = 1, 2, \dots, co$, sendo co as categorias ordenadas de y_i ;

$OR_{Ordinal_i}$, ou odds ratio, são as chances de um sujeito estar em até um determinada categoria (i) em um grupo, em comparação com as chances de um sujeito estar em até determinada categoria no outro grupo, assumindo o mesmo valor de odds ratio para todas as categorias $i = 1, 2, \dots, co$;

C , grupo controle;

T , grupo tratamento;

P_{iC} , a probabilidade de pertencer à categoria i do grupo C;

P_{iT} , a probabilidade de pertencer à categoria i do grupo T;

CP_{iC} , a probabilidade acumulada esperada de estar inclusa até a categoria i no grupo C;

CP_{iT} , a probabilidade acumulada esperada de estar inclusa até a categoria i no grupo T.

5.2.2 Argumentos

Para a função *posamsize* são necessários os seguintes argumentos:

- **p**, um vetor de probabilidades marginais, deve somar um. O i -ésimo elemento especifica a probabilidade de um paciente estar no nível de resposta i , é necessário calcular a média entre os dois grupos.
- **odds.ratio**, a razão de chances para poder detectar diferença, não importa qual grupo está no numerador, é constante entre as categorias.
- **alpha**, erro Tipo I.
- **x**, (auxiliar) um objeto criado por *posamsize*, ou um vetor de valores de dados à *pomodm* que corresponde ao vetor p de probabilidades. Se x for omitido para *pomodm*, a *odds.ratio* será aplicada e o novo vetor de probabilidades individuais será retornado. Caso contrário, se x é dado a *pomodm*, é retornado um duplo vetor com a média e a mediana x após a aplicação da *odds.ratio*.
- **fraction**, a fração de sujeitos que serão alocados ao grupo controle.
- **power**, o poder desejado, o padrão é 0,8.

5.2.3 Exemplo

Retirado de (Campbell e Walters, 2014). Supondo que um estudo deseja comparar dois grupos quanto à qualidade de vida, obtida por um determinado instrumento o qual avalia este traço latente. Os dados de avaliação foram recodificados em uma variável ordinal de quatro categorias, sendo essas 0–24, 25–49, 50–74 e 75–100. As proporções esperadas (CP) em cada categoria, suas proporções acumuladas (π_i) e proporção média de sujeitos previstos na categoria ($\bar{\pi}_i$) são mostradas na tabela a seguir.

Acredita-se que a diferença média nos escores da qualidade de vida entre os dois grupos não é uma medida comparativa apropriada. Porém, a comparação da probabilidade do paciente do grupo tratamento ter uma pontuação de qualidade de vida em até determinada categoria com a probabilidade do paciente do grupo controle é considerada uma medida resumida adequada.

Dado que, supostamente, 50% dos pacientes do grupo controle tenham uma boa qualidade de vida (uma pontuação de 75 ou mais), espera-se que o número melhore para 60% no grupo de tratamento após um regime diário de exercícios por um ano.

Categoria i		Controle		Tratamento		$(\bar{\pi}_i)$
		Proporção na categoria	Proporção acumulada	Proporção na categoria	Proporção acumulada	
		(π_{iC})	(CP_{iC})	(π_{iT})	(CP_{iT})	
1	0-24	0,122	0,122	0,085	0,085	0,1034
2	25-49	0,144	0,266	0,110	0,194	0,1269
3	50-74	0,236	0,502	0,207	0,402	0,2216
4	75-100	0,498	1,0	0,598	1,0	0,5482
		1,0		1,0		

Assim, calcula-se a razão de chances através de (5.4) e definem-se demais parâmetros, aqui, toma-se o poder de 0,8, significância de 5% e uma taxa taxa de alocação de 0,5, ou seja, metade da amostra entra no grupo controle, e outra no grupo tratamento.

$$OR_{Ordinal_i} = \frac{\frac{0.5}{(1-0.5)}}{\frac{0.6}{(1-0.6)}} = \frac{1.0}{1.5} = 0.67$$

As hipóteses testadas são:

$$\begin{aligned} H_0 : \theta &= 0 \\ H_1 : \theta &\neq 0 \end{aligned} \quad (5.6)$$

Em que θ representa o log da razão de chances da ocorrência de um desfecho na categoria C, ou melhor que C, para um sujeito do grupo tratamento em relação a um sujeito de controle. H_0 então é a hipótese em que os grupos tratamento e controle tem desempenho igual, e H_1 afirma que os grupos tem níveis de qualidade de vida diferentes.

No RStudio:

```
> pic<-c(0.122, 0.144, 0.236, 0.498)
> pit<-c(0.085, 0.11, 0.207, 0.598)
> p<-(pic+pit)/2
> posamsize(p, 0.67, 0.5, 0.05, 0.8)
Total sample size: 714.9
Efficiency of design compared with continuous response: 0.821
```

Assim, para este estudo, estima-se que serão necessários 715 pacientes, divididos igualmente entre controle e tratamento. A eficiência deste modelo em comparação com um delineamento no qual a variável de resposta é contínua, não categórica, é de 0,82.

5.3 Funções *popower*

5.3.1 Aplicação

A função *popower* pode ser vista como complementar da *posamsize*, já que suas aplicabilidades e princípios são os mesmas, estando ambas documentadas conjuntamente no R. O objetivo desta, no entanto, é calcular o poder necessário para a realização do teste, não a amostra, por isso os parâmetros apresentados diferem em alguns pontos da 'irmã'. Esta função também retorna a sua eficiência em comparação com um delineamento no qual a variável resposta é contínua.

5.3.2 Argumentos

Para a função *popower* são necessários os seguintes argumentos:

- **p**, um vetor de probabilidades marginais, deve somar um. O *i*-ésimo elemento especifica a probabilidade de um paciente estar no nível de resposta *i*, é necessário calcular a média entre os dois grupos.
- **odds.ratio**, a razão de chances para poder detectar diferença, não importa qual grupo está no numerador, é constante entre as categorias.
- **n**, tamanho total da amostra, é necessário especificar *n* ou *n1* e *n2*. No caso de *n* ser declarado, é assumido que $n1 = n2$ ($n/2$).
- **n1**, o número de sujeitos no grupo 1.
- **n2**, o número de sujeitos no grupo 2.
- **alpha**, erro Tipo I.
- **x**, (auxiliar) um objeto criado por *popower*, ou um vetor de valores de dados à *pomodm* que corresponde ao vetor *p* de probabilidades. Se *x* for omitido para *pomodm*, a *odds.ratio* será aplicada e o novo vetor de probabilidades individuais será retornado. Caso contrário, se *x* é dado a *pomodm*, é retornado um duplo vetor com a média e a mediana *x* após a aplicação da *odds.ratio*.

5.3.3 Exemplo

Agora o objetivo é o cálculo do poder necessário para testar a existência de diferenças entre a proporção das pessoas que possuem qualidade de vida considerada boa no grupo controle e no tratamento, utilizando os mesmos dados de 5.2.3. O cenário é de uma amostra total de 715 pacientes, dividido igualmente entre os dois grupos, considerando um nível de significância de 5%.

As hipóteses testadas são:

$$\begin{aligned} H_0 : \theta &= 0 \\ H_1 : \theta &\neq 0 \end{aligned} \tag{5.7}$$

Em que θ representa o log da razão de chances da ocorrência de um desfecho na categoria *C*, ou melhor que *C*, para um sujeito do grupo tratamento em relação a um sujeito de controle. H_0 então é a hipótese em que os grupos tratamento e controle tem desempenho igual, e H_1 afirma que os grupos tem níveis de qualidade de vida diferentes.

No RStudio:

```
> pic<-c(0.122, 0.144, 0.236, 0.498)
> pit<-c(0.085, 0.11, 0.207, 0.598)
> p<-(pic+pit)/2
> popower(p, 0.67, 715, alpha=0.05)
Power: 0.834
Efficiency of design compared with continuous response: 0.9
```

Portanto, para este estudo, a probabilidade de detectar diferença entre os grupos, se elas existirem, é de 0,83. A eficiência de optar por este modelo com respostas categóricas em comparação com um delineamento com variável contínua, é de 0,9.

5.4 Funções *bpower* e *bpower.sim*

5.4.1 Aplicação

As funções *bpower* e *bpower.sim* usam o método de [Fleiss et al. \(1980\)](#) (sem a correção da continuidade) para estimar o poder para testar diferenças entre duas proporções de um evento de interesse. Os tamanhos das amostras podem ser diferentes nos dois grupos. Enquanto *bpower* calcula um poder aproximado, *bpower.sim* devolve o poder exato, advindo de uma simulação, e o intervalo de confiança de 95% com seus limites superior e inferior.

5.4.2 Argumentos

Para as funções *bpower* e *bpower.sim* são necessários os seguintes argumentos:

- **p1**, probabilidade de ocorrência do evento de interesse no grupo 1.
- **p2**, probabilidade de ocorrência do evento de interesse no grupo 2.
- **odds.ratio**, razão de chances.
- **percent.reduction**, % de redução.
- **n**, tamanho total da amostra nos dois grupos.
- **n1**, o tamanho da amostra do grupo 1. Se **n** for escolhido, **n1** e **n2** são definidos como $n/2$ ($n1=n2$).
- **n2**, o tamanho da amostra do grupo 2. Se **n** for escolhido, **n1** e **n2** são definidos como $n/2$ ($n1=n2$).
- **alpha**, erro Tipo I.
- **nsim**, número de simulações de respostas binomiais.

As hipóteses testadas são:

$$\begin{aligned} H_0 : P_1 &= P_2 \\ H_1 : P_1 &\neq P_2 \end{aligned} \tag{5.8}$$

Em que P_1 e P_2 representam as probabilidades de sucesso (ocorrência do evento de interesse) nos grupos com e sem câncer de mama.

5.4.3 Exemplo

Retirado de [Breheny \(2014\)](#). Neste exemplo, é desejado verificar o poder de um estudo prospectivo (veja [Hochman et al. \(2005\)](#)) de câncer de mama conduzido por um Centro de Controle e Prevenção de Doenças (CDC). São dados os seguintes resultados sobre a relação entre câncer de mama e a idade em que uma mulher deu à luz seu primeiro filho:

Câncer	Não	Sim
Antes do 25 anos	4475	65
Aos 25 anos ou depois	1597	31

Para o cálculo de poder, os tamanhos de ambas as amostras devem ser especificados, $n1 = 4540$ e $n2 = 1628$. Em seguida, é especificada a frequência com

que o evento (câncer de mama) ocorre, aqui, este evento é relativamente incomum, ocorrendo em cerca de 1% das mulheres do estudo (para ambos os grupos).

Finalmente, deve-se definir um tamanho de efeito (*odds.ratio*) que, devido à estudos prévios, será assumido 1,5, enquanto o valor de α , para este exemplo é 0,05.

No RStudio:

```
> bpower (p1 = .01, p2 = .01, odds.ratio = 1.5, n1 = 4540, n2 = 1628, alpha = 0.05)
      Power
0.37643
```

O poder estimado é de 37%.

Agora, em um delineamento alternativo com grupo controle, são incluídos 300 casos, 150 mulheres com câncer de mama e 150 sem câncer de mama, com um foco retrospectivo. O evento medido é a exposição ao fator de risco - independente da mulher ter tido ou não filhos após os 25 anos, e ocorre cerca de 35% das vezes em ambos os grupos. Considera-se o mesmo tamanho de efeito (1,5) no cálculo do poder para comparação dos estudos.

No RStudio:

```
> bpower(p1 = .35, p2 = 0.35, odds.ratio = 1.5, n = 300, alpha = 0.05)
      Power
0.4019139
```

Para este exemplo, o poder é de 40%. Foi obtido maior poder com um estudo retrospectivo de 300 pessoas, do que com um estudo prospectivo de 6.000 pessoas (20x menos pessoas).

5.5 Funções *bsamsize* e *ballocation*

5.5.1 Aplicação

A função *bsamsize* também foi implementada usando o método de [Fleiss et al. \(1980\)](#) (sem a correção da continuidade) para estimar o poder para testar diferenças entre proporções para duas variáveis binomiais, considerando um teste bicaudal. Novamente, é permitido um tamanho de amostra desigual entre os dois grupos, devendo ser especificada a fração de observações no grupo 1. A função devolve um vetor contendo os tamanhos de amostra nos dois grupos.

Quanto à *ballocation*, para dados valores de proporção nos dois grupos (p_1 e p_2), essa função usa o método de [Brittain e Schlesselman \(1982\)](#) para procurar a taxa de alocação ótima das amostras para comparações de proporções com respostas binárias. Retorna um vetor com os quatro valores ótimos de observações a serem alocadas no grupo 1, otimizando os seguintes critérios, respectivamente:

(1) minimizar a variância da diferença entre duas proporções; (2) minimizar a variância da razão entre as duas proporções; (3) minimizar a variância do log da razão de chances, esta fração é a complementar da (1); (4) maximizar o poder do teste bicaudal para diferenças (o tamanho total da amostra deve ser fornecido ou a fração que otimiza o poder não é retornada).

5.5.2 Argumentos

Para as funções *bsamsize* e *ballocation* são necessários os seguintes argumentos:

- **p1**, probabilidade no grupo 1.
- **p2**, probabilidade no grupo 2.
- **odds.ratio**, razão de chances.
- **n**, tamanho total da amostra nos dois grupos, se omitida, a fração que otimiza o poder não será devolvida.
- **alpha**, erro Tipo I.
- **fraction**, fração de observações no grupo 1.
- **power**, a probabilidade desejada de detectar uma diferença.

5.5.3 Exemplos

Adaptado de [Cohen e Cohen \(2008\)](#). Dois grupos de 40 pacientes cada foram selecionados para um estudo da eficácia das vacinas contra a gripe. Os membros do grupo de tratamento receberam uma vacina contra a gripe, enquanto membros do grupo de controle receberam uma injeção salina. A história médica de ambos os grupos foi seguida durante toda a temporada de gripe. Do grupo controle, 15 apresentaram sintomas de gripe pelo menos uma vez, no grupo de tratamento, 9 o fizeram.

Deseja-se determinar os tamanhos de amostra necessários para estabelecer uma diferença de 0,375 a 0,225 entre a proporção que adoeceu nos grupos controle e tratamento, respectivamente. Utilizando os valores padrão de $\alpha = 0,05$ e poder = 0,8, alocando a mesma fração da amostra total para os dois grupos, aplicam-se os valores na *bsamsize*.

As hipóteses testadas são:

$$\begin{aligned} H_0 &: \pi_{Tratamento} = \pi_{Controle} \\ H_1 &: \pi_{Tratamento} \neq \pi_{Controle} \end{aligned} \quad (5.9)$$

Em que π_1 e π_2 representam as probabilidades de sucesso (ocorrência do evento de interesse) nos grupos tratamento e controle.

No RStudio:

```
> ceiling(bsamsize(p1 = 0.375, p2 = 0.225, fraction = 1/2, alpha = 0.05, power = 0.8))
  n1  n2
146 146
```

Em outras palavras, precisamos de 292 pessoas para alcançar a significância desejada. A função *ceiling* é usada apenas para arredondar.

Pode-se pensar também que é, por exemplo, duas vezes mais caro seguir membros do grupo de tratamento em comparação com o grupo controle. Enquanto acompanhar um membro do grupo de tratamento custa R\$ 100, seguir um membro do grupo de controle custa R\$50. A fração desejada de alocação para o grupo de tratamento seria de 1/3.

No RStudio:

```
> ceiling(bsamsize(p1 = 0.375, p2 = 0.225, fraction = 1/3, alpha = 0.05,
power = 0.8))
  n1  n2
107 214
```

Agora o tamanho das amostras tem distribuição diferente, e seriam necessárias 321 pessoas para atingir a significância (107 no grupo de tratamento e 214 no grupo controle).

Para exemplificar a função *ballocation*, pode-se pensar em investigar a possibilidade de alocar as 321 pessoas para ambos os grupos de forma a maximizar o poder alcançado

.No RStudio:

```
> ballocation(p1 = 0.375, p2 = 0.225, n = 321, alpha = 0.05)

fraction.group1.min.var.diff
                        0.5368971
fraction.group1.min.var.ratio
                        0.4102412
fraction.group1.min.var.logodds
                        0.4631029
fraction.group1.max.power
                        0.4885115
```

Assim, ao invés de 107 pacientes no grupo de tratamento, 157 ($0.4885 \cdot 321$) serão alocados neste grupo, em função da maximização do poder alcançado.

5.6 Função *cpower*

5.6.1 Aplicação

A função *cpower* desempenha o cálculo de poder para o teste log-rank para duas amostras, usando o método de [George e Desu \(1974\)](#) junto com as fórmulas de [Schoenfeld \(1983\)](#), que permitem estimar o número esperado de eventos nos dois grupos comparados. A função ainda suporta especificações para entrada de pacientes no estudo de forma aleatória, considerando certo tempo fixo e um tempo mínimo de acompanhamento.

Outras definições de não cumprimento dos requisitos do estudo dentro do grupo de tratamento (noncompliance) ou migrações do grupo controle para o tratamento (dropin) são aceitas, usando a equação 5.4 de [Lachin e Foulkes \(1986\)](#). O número de eventos em cada grupo é estimado assumindo curvas de sobrevivência exponenciais, e o autor usa o log da razão de chances ao invés da diferença entre riscos.

As hipóteses testadas são:

$$\begin{aligned} H_0 &: S_{Controle(t)} = S_{Tratamento(t)}, \forall t \\ H_1 &: S_{Controle(t)} \neq S_{Tratamento(t)}, \forall t \end{aligned} \quad (5.10)$$

Definindo que $S_{Controle(t)}$ e $S_{Tratamento(t)}$ são as funções (curvas) de sobrevivência em cada tempo t .

5.6.2 Argumentos

Para a função *cpower* são necessários os seguintes argumentos:

- **tref**, período de tempo em que as taxas de mortalidades são estimadas.

- **n**, tamanho total da amostra (ambos os grupos combinados). Se a alocação for desigual, não havendo $n/2$ observações em cada grupo, os tamanhos das amostras n_c e n_i devem ser especificados.
- **mc**, mortalidade em dado período de tempo (tref), grupo controle.
- **r**, % de redução em mc por intervenção.
- **accrual**, duração do período de competência (tempo de recrutamento dos pacientes para o estudo).
- **tmin**, tempo mínimo de acompanhamento.
- **noncomp.c**, % de não conformidade no grupo controle (dropins).
- **noncomp.i**, % de não conformidade no grupo intervenção (não aderentes, pacientes cujo comportamento não corresponde à intervenção necessária protocolada do estudo).
- **alpha**, probabilidade de erro do Tipo I. Um teste bicaudal é assumido.
- **nc**, número de sujeitos no grupo controle.
- **ni**, número de sujeitos no grupo de intervenção. n_c e n_i são especificados exclusivamente para n .
- **pr**, se FALSE, suprime a impressão de detalhes.

5.6.3 Exemplo

Simulando um estudo clínico com 2 grupos (tratamento e controle), sendo que número de eventos em cada grupo é estimado assumindo curvas de sobrevivência exponenciais, deseja-se estimar o poder para a comparação destas curvas. Foi considerado um $n = 1000$, igualmente distribuído, um α de 0,05, e a redução da mortalidade estimada dada pelo grupo intervenção, não havendo dropouts ou dropins) é de 25%.

A taxa de risco é considerada contante e os pacientes entram no estudo uniformemente, em um período de 1 a 3 anos. As probabilidades de não conformidade (dropins) e não aderência também são constantes ao longo do estudo, respectivamente, 10% e 17,5%. A sobrevida em dois anos dos pacientes do grupo controle, na ausência de dropins, é de 0,8 (mortalidade = 0,2).

No RStudio:

```
> cpower(tref = 2, n = 1000, mc = .2, r = 25, accrual = 2, tmin = 1, noncomp.c = 10,
noncomp.i = 17.5, alpha=0.05, pr=TRUE)

Accrual duration: 2 y   Minimum follow-up: 1 y

Total sample size: 1000

Alpha= 0.05

2-year Mortalities
  Control Intervention
    0.20         0.15

Hazard Rates
  Control Intervention
0.11157178  0.08125946

Probabilities of an Event During Study
  Control Intervention
0.1983392  0.1490643
```

```

Expected Number of Events
  Control Intervention
      99.2       74.5

Hazard ratio: 0.7283156

Drop-in rate (controls):10%
Non-adherence rate (intervention):17.5%
Effective hazard ratio with non-compliance: 0.7946607
Standard deviation of log hazard ratio: 0.1532996
Power
0.322786

```

Assim, o poder calculado para a comparação das curvas é igual a 32,3%.

5.7 Função *ciapower*

5.7.1 Aplicação

A função *ciapower* usa o método de [Peterson e George \(1993\)](#) para calcular o poder necessário para testar a interação entre as variáveis em um estudo fatorial 2 x 2, em que os quatro tempos de sobrevivência sigam distribuições exponenciais, sendo o mesmo poder do teste para o modelo Cox se mantidas as suposições. É necessário especificar a duração da competência (assumindo-a constante), assim como o tempo mínimo de acompanhamento. O tempo máximo de acompanhamento do paciente é dado pelo tempo de competência somado ao tempo mínimo ($\text{accrual} + \text{tmin}$). Presume-se que a alocação de tratamento seja 1:1 (mesma quantidade de pacientes nos dois grupos), e não leva em conta especificações de não conformidade.

5.7.2 Argumentos

Para a função *ciapower* são necessários os seguintes argumentos:

- **tref**, período de tempo em que as taxas de mortalidades são estimadas.
- **n1**, tamanho total da amostra, estrato 1.
- **n2**, tamanho total da amostra, estrato 2.
- **m1c**, mortalidade em dado período de tempo (**tref**) para o controle no estrato 1.
- **m2c**, mortalidade em dado período de tempo (**tref**) para o controle no estrato 2.
- **r1**, redução percentual em **m1c** por intervenção, estrato 1.
- **r2**, redução percentual em **m2c** por intervenção, estrato 2.
- **accrual**, duração do período de competência (tempo de recrutamento dos pacientes para o estudo).
- **tmin**, tempo mínimo de acompanhamento.
- **alpha**, probabilidade de erro do Tipo I. Um teste bicaudal é assumido.
- **pr**, se FALSE, suprime a impressão de detalhes.

5.7.3 Exemplo

Um grupo de pesquisadores levantou um estudo sob a hipótese de que dado medicamento para a diabetes possui diferentes efeitos em pacientes de etnia branca e

não-branca. A fim de descobrir qual seria o poder necessário para testar a existência dessa interação entre etnia e tratamento, delineou-se um estudo com amostra hipotética de 15.000 pacientes, sendo 25% não brancos. Baseados em estudos prévios, foi estimado que em 5 anos a mortalidade por diabetes do grupo controle seria de 0,18 para brancos e 0,23 para não brancos. São esperadas reduções nessas taxas de 20% para o primeiro estrato (brancos) e 5% para o segundo (não brancos), durante um período mínimo de acompanhamento de 5 anos, havendo um período inicial de um ano e meio para obter os pacientes (accrual). Por padrão, o cálculo considera uma significância de 5% para um teste bicaudal.

As hipóteses testadas são:

$$\begin{aligned} H_0 : \frac{\Delta_1}{\Delta_2} &= 1 \\ H_1 : \frac{\Delta_1}{\Delta_2} &\neq 1 \end{aligned} \tag{5.11}$$

Em que $\Delta_1 = \frac{\lambda_{\text{controle,nãobranco}}}{\lambda_{\text{tratamento,nãobranco}}}$ é a taxa de risco para o efeito do tratamento no nível 1 do fator etnia (não branco), e $\Delta_2 = \frac{\lambda_{\text{controle,branco}}}{\lambda_{\text{tratamento,branco}}}$ é taxa de risco para o efeito do tratamento no nível 2 do fator etnia (branco).

No RStudio:

```
> n <- 15000
> ciapower(5, .75*n, .25*n, .18, .23, 20, 5, 1.5, 5)

Accrual duration: 1.5 y   Minimum follow-up: 5 y

Sample size Stratum 1: 11250   Stratum 2: 3750

Alpha= 0.05

5-year Mortalities
      Control Intervention
Stratum 1   0.18         0.1440
Stratum 2   0.23         0.2185

Hazard Rates
      Control Intervention
Stratum 1 0.03969019    0.03109698
Stratum 2 0.05227295    0.04930803

Probabilities of an Event During Study
      Control Intervention
Stratum 1 0.2039322     0.1636574
Stratum 2 0.2594139     0.2467011

Expected Number of Events
      Control Intervention
Stratum 1 1147.1        920.6
Stratum 2 486.4         462.6

Ratio of hazard ratios: 0.8306049
Standard deviation of log ratio of ratios: 0.0785863
Power
0.6560863
```

A função retorna um poder de 0,66 para detectar o efeito da interação entre as variáveis etnia e tratamento.

5.8 Função *spower*

5.8.1 Aplicação

Dadas certas funções para gerar variáveis aleatórias para tempos de sobrevivência e tempos de censura, a *spower* simula o poder de um teste log-rank de duas amostras dado pelo usuário para dados censurados. Opcionalmente, um modelo de Cox é ajustado para cada conjunto de dados simulados e os log das taxas de risco são salvos.

Para compor as funções *S* para gerar tempos de sobrevivência aleatórios sob condições complexas, a função *Quantile2* permite que o usuário especifique a taxa de risco de intervenção:controle em função do tempo, a probabilidade de um sujeito do grupo controle receber o tratamento intervenção (dropin) em função de tempo e a probabilidade de um sujeito do grupo intervenção receber apenas o agente placebo em função do tempo (non-compliance, dropout).

Além de fornecer a função *Quantile2*, o pacote *spower* também contém três funções que compõem as funções *S*, que retornam as probabilidades de sobrevivência das distribuições Weibull, log-normal e Gompertz. Essas funções (*Weibull2*, *Lognorm2* e *Gompertz2*) funcionam resolvendo os dois parâmetros de cada uma dessas distribuições que os tornam adequados a dois tempos e probabilidades de sobrevivência especificados pelo pesquisador. Os três tipos de funções criadas são úteis como o primeiro argumento para *Quantile2*.

A função *Quantile2* retorna uma função que gera tempos de sobrevivência sem censura de controle ou intervenção, sujeitos a efeitos de tratamento não constantes, dropin e dropout. Ainda assume que a função de risco para qualquer grupo de tratamento é uma mistura das funções de risco dos grupos controle e intervenção, com proporções de mistura definidas pelas probabilidades de dropin e dropout.

As distribuições de riscos e sobrevivência são calculadas por numérica diferenciação e integração usando uma grade de (por padrão) 7.500 pontos de tempo igualmente espaçados.

5.8.2 Argumentos

Para a função *spower* são necessários os seguintes argumentos:

- **rcontrol**, uma função de *n* que retorna *n* tempos de falha aleatórios sem censura para o grupo de controle. Assume que a não conformidade (dropins) foi levada em consideração por essa função.
- **rinterv**, semelhante ao **rcontrol**, mas para o grupo de intervenção.
- **rcens**, uma função de *n* que retorna *n* tempos de censura aleatórios. Supõe-se que ambos os grupos de tratamento tenham a mesma distribuição de censura.
- **nc**, número de sujeitos no grupo controle.
- **ni**, número de sujeitos no grupo intervenção.
- **scontrol**, uma função de um vetor de tempos que retorna as probabilidades de sobrevivência do grupo controle naqueles momentos, assumindo que todos os pacientes estão em conformidade.
- **hratio**, uma função do tempo que especifica a taxa de risco de intervenção/controlado (efeito do tratamento).
- **x**, um objeto da classe "Quantile2" criado por *Quantile2*.

- **S**, um objeto *Surv* ou outra matriz de duas colunas para tempos de sobrevivência censurados à direita.
- **group**, os indicadores de grupo, têm comprimento igual ao número de linhas em S. Valores permitidos somente 1 e 2.
- **times**, um vetor de dois tempos.
- **surv**, um vetor de duas probabilidades de sobrevivência.
- **test**, qualquer função de um objeto *Surv* e uma variável de agrupamento que calcula um qui-quadrado para um teste de dados censurados de duas amostras, o padrão é log-rank.
- **nsim**, número de simulações a serem executadas, padrão é 500.
- **alpha**, erro Tipo I, padrão é 0,05.
- **pr**, FALSE suprime as impressões do progresso das simulações.
- **dropin**, uma função do tempo que especifica a probabilidade de um sujeito de controle se tornar um sujeito de intervenção no momento correspondente.
- **dropout**, uma função do tempo que especifica a probabilidade de um sujeito de intervenção desistir para controlar condições.
- **m**, número de pontos no tempo usados para aproximar funções, o padrão é 7500.
- **tmax**, ponto de tempo máximo a ser usado na grade de m vezes. O padrão é o tempo em que `scontrol (time)` é `qtmax`.
- **qtmax**, probabilidade de sobrevivência correspondente ao último ponto no tempo usado para aproximar as funções de sobrevivência e risco, tendo o padrão 0.001. Para `qtmax` o tempo durante o qual é necessário um tempo simulado que corresponda a uma probabilidade de sobrevivência menor que `qtmax`, o valor simulado será `tmax`.
- **mplot**, número de pontos usados para aproximar funções para uso em plotagem, o padrão é 200 pontos igualmente espaçados.
- ..., argumentos opcionais passados para a função `scontrol` quando validados pelo `Quantile2`.
- **what**, uma constante de caractere único (pode ser abreviada) especificando quais funções devem ser plotadas. O padrão é "both", significando funções de sobrevivência e risco, "drop" para plotar apenas as funções `dropin` e `drop-out`, "hratio" para plotar as funções da taxa de risco ou "all" para criar quatro gráficos separados mostrando todas as funções (6 gráficos se `dropsep = TRUE`).
- **dropsep**, definir TRUE para criar `plot.Quantile2` separa funções puras e contaminadas em plotagens separadas.
- **lty**, vetor de tipos de linha.
- **col**, vetor de cores.
- **xlim**, limites opcionais do eixo x.
- **ylim**, limites opcionais do eixo y.
- **label.curves**, lista opcional que é passada como argumento `opts` para `labcurve`.

5.8.3 Exemplo

Retirado de [Alzola e Harrell \(2006\)](#). Simulando um estudo de acompanhamento de 2 anos com 2 grupos (350 indivíduos por grupo) para o qual a distribuição de

sobrevida do grupo controle é Weibull, com sobrevida em 1 ano de 0,95 e sobrevida em 3 anos de 0,7.

Todos os indivíduos são acompanhados por pelo menos um ano e entram no estudo com probabilidade linearmente crescente, começando com zero.

Supondo (1) que não há chance de dropins durante os primeiros 6 meses, então a probabilidade aumenta linearmente até 0,15 aos 5 anos de estudo; (2) existe uma chance linearmente crescente de dropouts de até 0,3 em 5 anos; e (3) o tratamento não tem efeito nos primeiros 9 meses, depois tem um efeito constante (taxa de risco de 0,75).

No RStudio:

Primeiro, encontra-se a distribuição Weibull correta para os pacientes do grupo controle que estão seguindo corretamente o protocolo, usando a função *Weibull2*.

```
> sc <- Weibull2(c(1,3), c(.95,.7))
> sc
function (times = NULL, alpha = 0.0512932943875506, gamma = 1.76519490623438)
{
  exp(-alpha * (times^gamma))
}
```

Distribuição cumulativa inversa, para o caso em que todos os indivíduos são acompanhados pelo menos um ano, e depois, entre os anos a e b a densidade aumenta, $(time - a)^d$ é $a + (b - a) * u^{1/(d + 1)}$.

```
> rcens <- function(n) 1 + (5-1) * (runif(n) ^ .5)
>
```

Juntando na função *Quantile2*:

```
> f <- Quantile2(sc,
+             hratio=function(x)ifelse(x <= .75, 1, .75),
+             dropin=function(x)ifelse(x <= .5, 0, .15*(x-.5)/(5-.5)),
+             dropout=function(x).3*x/5)
```

```
Interval of time for evaluating functions:[0, 16.07887]
```

A função f criada pelo *Quantile2* tem como principais argumentos n , o número de variáveis aleatórias a serem desenhadas e *what*, informando a função se é necessário retirar amostras dos tempos de sobrevivência sem censura para controle ou intervenção.

Agora, *spower* simula os resultados necessários, baseando a comparação da distribuição de sobrevivência no teste log-rank.

```
> rcontrol<- function(n) f(n, "control")
> rinterv <- function(n) f(n, "intervention")
> set.seed(211)
> spower(rcontrol, rinterv, rcens, nc=350, ni=350, test=logrank, nsim=300)
300
[1] 0.3966667
```

O poder estimado então para comparação de curvas entre os grupos é de 40%.

6 Considerações Finais

A criação desta monografia teve como propósito auxiliar o leitor no manuseio das funções para cálculo de tamanho de amostra e poder do Software R dentro dos pacotes *longpower* e *Hmisc*. Foi também uma continuação de [Pedrotti \(2017\)](#) e [Fountoura \(2018\)](#), abordando tanto funções para técnicas mais recorrentes em estatística, quanto para estudos mais complexos.

Reforçando o que foi ilustrado anteriormente, os delineamentos de dados longitudinais e dados de sobrevivência abrangem as mais variadas áreas de pesquisa. Por isso, é importante lembrar a relevância de um planejamento amostral bem desenvolvido, a fim de evitar desperdício de recursos humanos, temporais e econômicos.

Embora cada pacote tenha sua própria documentação, a complexidade das funções aqui abordadas pode facilmente gerar confusão no leitor. Nesta monografia, foram desenvolvidos ou adaptados exemplos com aplicações práticas para o melhor entendimento das técnicas implementadas em cada cenário.

Futuros trabalhos podem abordar outros pacotes do Software R, como *powerlmm*, sobre cálculos de poder para modelos longitudinais multiníveis, *samplesize*, sobre cálculo de tamanho de amostra para vários testes t e testes de Wilcoxon e *LogrankPower*, que estima o poder do teste log-rank usando conjuntos de dados de simulação.

Não são raras as ocasiões em que o investigador coleta os dados sem ter claro nem mesmo o objetivo principal do estudo e qual a forma correta de se alcançar esta meta. Assim, espera-se que a criação de materiais desta natureza torne o processo inicial de delineamento da pesquisa mais fácil para os pesquisadores, visto que esta é uma das áreas que mais carece de atenção para o bom desenvolvimento do estudo.

A miríade de pacotes implementados no Software R que envolvem tamanho de amostra e poder nas mais diversas áreas pode ser encontrada e explorada facilmente aqui <https://cran.r-project.org/>.

Referências Bibliográficas

- Alzola, C. e Harrell, F. (2006). An introduction to s and the hmisc and design libraries.
- Aquino, J. A. (2014). *R para cientistas sociais*. Editora da UESC, Ilhéus, BR.
- Bastos, J. e Rocha, C. (2006). Análise de sobrevivência: Conceitos básicos. *Arquivos de Medicina*, 20(5/6):185–187.
- Bell, M. L., Pinto, A. T., McKenzie, J. E., e Olivier, J. (2014). A myriad of methods: Calculated sample size for two proportions was dependent on the choice of sample size formula and software. *Journal of Clinical Epidemiology*.
- Bessegato, L. F. (2014). Análise de sobrevivência aplicada á saúde. University Lecture.
- Bolfarine, H. e Bussab, W. O. (2005). *Elementos de Amostragem*. ABE, São Paulo, BR.
- Breheeny, P. (2014). Introduction to biostatistics (171:161). myweb.uiowa.edu/pbreheeny/4120/s14/labs/11.pdf.
- Brittain, E. e Schlesselman, J. J. (1982). Optimal allocation for the comparison of proportions. *Biometrics*, 38(4):1003–1009.
- Campbell, M. J. e Walters, S. J. (2014). *How to Design, Analyse and Report Cluster Randomised Trials in Medicine and Health Related Research*. John Wiley and Sons, Ltd, West Sussex, UK.
- Carneiro, A. C. (2012). Análise de dados longitudinais através de modelos marginais. Trabalho de conclusão de curso em estatística, Universidade Federal de Juiz de Fora, Juiz de Fora, BR.
- Cohen, Y. e Cohen, J. Y. (2008). *Statistics and Data with R: An applied approach through examples*. John Wiley and Sons, Ltd, West Sussex, UK.
- Colosimo, E. e Giolo, S. (2006). *Análise de Sobrevivência Aplicada*. ABE, São Paulo, BR.
- Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society*, 34(2):187–220.

- Diggle, P., Liang, K. Y., e Zeger, S. L. (1994). *Analysis of longitudinal data*. Clarendon Press, first edition.
- Diggle, P. J., Heagerty, P. J., Liang, K. Y., e Zeger, S. L. (2002). *Analysis of longitudinal data*. OUP Oxford, New York, USA, second edition.
- Donohue, M. C. (2019). *longpower: Power and sample size calculations for linear-mixed models*. R package version 1.0-19.
- Donohue, M. C., Edland, S. D., e Gamst, A. C. (2019). Power for linear models of longitudinal data with applications to alzheimer’s disease phase ii study design. Manuscript.
- Edland (2009). *Which MRI measure is best for Alzheimer’s disease prevention trials: Statistical considerations of power and sample size*, Alexandria, USA. American Statistical Association.
- Fleiss, J. L., Tytun, A., e Ury, H. K. (1980). A simple approximation for calculating sample sizes for comparing independent proportions. *Biometrics*, 36(2):343–346.
- Fontoura, R. (2018). Análise de sobrevivência: tamanho de amostra e poder com pacote powersurvepi do software r. *LUME UFRGS*.
- George, S. e Desu, M. (1974). Planning the size and duration of a clinical trial studying the time to some critical event. *J Chron Dis*, 27:15–24.
- Hochman, B., Nahas, F. X., Filho, R. S., e Ferreira, L. M. (2005). Desenhos de pesquisa. *Acta Cirúrgica Brasileira*, 20 (Supl. 2).
- Institute, U. C. . T. S. (2019). Sample size calculators. sample-size.net/.
- Jr., F. E. H. (2019). *Hmisc: Harrell Miscellaneous*. R package version 4.3-0.
- Kreidler, S. M., Muller, K. E., Grunwald, G. K., Ringham, B. M., Coker-Dukowitz, Z. T., Sakhadeo, U. R., Barón, A. E., e Glueck, D. H. (2013). Glimmpse: Online power computation for linear models with and without a baseline covariate. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3882200/>.
- Lachin, J. M. e Foulkes, M. A. (1986). Evaluation of sample size and power for analyses of survival with allowance for nonuniform patient entry, losses to follow-up, noncompliance, and stratification. *Biometrics*, 42(3):507–519.
- Landeiro, V. L. (2011). *Introdução ao uso do programa R*. Programa de Pós Graduação em Ecologia, Instituto Nacional de Pesquisas da Amazônia, Manaus, BR.
- Liu, G. e Liang, K. Y. (1997). Sample size calculations for studies with correlated observations. *Biometrics*, 53(3):937–47.
- LLC, H. C. (2013-2019). Sample size calculators. powerandsamplesize.com/Calculators/.
- Lu, K., Luo, X., e Chen, P.-Y. (2008). Sample size estimation for repeated measures analysis in randomized clinical trials with missing data. *International Journal of Biostatistics*, 4(1) art.9.

- Meinert, C. L. e Tonascia, S. (1986). *Clinical Trials: Design, Conduct, and Analysis*. Oxford University Press, Inc., New York, USA.
- Miot, H. A. (2011). Tamanho da amostra em estudos clínicos e experimentais. *Jornal Vascular Brasileiro*, 10(4):275–278.
- NCSS e LLC (2019). Pass 2019 power analysis and sample size software. [ncss.com/software/pass](https://www.ncss.com/software/pass).
- Ndow, M. (2018). Power calculations for linear mixed effects models. Master's thesis, Maastricht University.
- Paes, A. T. (2008). Qual deve ser o tamanho da minha amostra? *Einstein: Educação Continuada em Saúde*, pages 153–154.
- Pedrotti, L. (2017). Tamanho de amostra e poder no software r. *LUME UFRGS*.
- Peterson, B. e George, S. L. (1993). Sample size requirements and length of study for testing interaction in a 1 x k factorial design when time-to-failure is the outcome. *Controlled Clinical Trials*, 14:511–522.
- Racine, J. S. (2012). Rstudio: A platform-independent ide for r and sweave. *JOURNAL OF APPLIED ECONOMETRICS*, 27:167–172.
- Schoenfeld, D. (1983). Sample-size formula for the proportional-hazards regression model. *Biometrics*, 39(2):499–503.
- Singer, J. M., Nobre, J. S., e Rocha, F. M. M. (2018). *Análise de dados longitudinais*. Departamento de Estatística: Universidade de São Paulo, São Paulo, BR.
- Team, R. C. (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, AT.
- Whitehead, J. (1993). Sample size calculations for ordered categorical data. *Stat in Med*, 12:2257–2271.