

O PROBLEMA DO ANIVERSÁRIO, O PACOTE IPSUR E SEU PLUGIN PARA O R COMMANDER

Uma possibilidade para sala de aula

Felipe Rafael Ribeiro Melo

Em uma sala com n pessoas, qual a probabilidade de pelo menos um par destas pessoas fazer aniversário no mesmo dia? Esse é o conhecido Problema do Aniversário. Sua solução engloba conceitos da Probabilidade, como o conceito clássico da Probabilidade e a relação entre as probabilidades de um evento e do seu evento complementar, além de técnicas básicas em Análise Combinatória, como arranjos e combinações. Há pacotes no *software* R com funções que calculam probabilidades associadas ao Problema do Aniversário. Este capítulo aborda inicialmente a solução do Problema do Aniversário, explicitando seus pressupostos, e apresenta os pacotes `IPSUR` e `RcmdrPlugin.IPSUR`. `IPSUR`, que trazem consigo funções associadas ao Problema do Aniversário. Este último pacote, em particular, é um *plugin* a ser utilizado na interface R Commander, que permite o uso de várias funcionalidades do R por meio de menus, sem a necessidade de digitação e/ou execução de linhas de comando. Isto pode motivar docentes de Matemática do Ensino Médio e docentes do Ensino Superior em disciplinas de Probabilidade a realizar atividades em sala de aula que abordem o Problema do Aniversário, inclusive associando o valor n ao número de alunos em sala de aula e realizando dinâmicas em sala de aula.

Palavras Chave: Problema do Aniversário; Linguagem R; `IPSUR`; R Commander; Sala de Aula.

1 Introdução

O chamado Problema do Aniversário (*Birthday Problem*) versa sobre a probabilidade de pelo menos um par de pessoas aniversariar em um mesmo dia em uma sala com n pessoas. Em pesquisas sobre este assunto, também é comum se deparar com o termo Paradoxo do Aniversário (*Birthday Paradox*). Esta expressão deriva do fato contra-intuitivo de que bastam 23 pessoas para que a probabilidade de ao menos um par de pessoas fazer aniversário no mesmo

dia exceda 50%, ou seja, a ocorrência deste evento é mais provável que a sua não ocorrência em um grupo de apenas 23 pessoas - quantitativo pequeno face aos 365 dias possíveis que uma pessoa, aleatoriamente escolhida, possa aniversariar.

O Problema do Aniversário geralmente é atribuído ao matemático inglês Harold Davenport, por volta de 1927. Embora Davenport não o tenha publicado na época, ele não reivindicou ser seu descobridor porque “ele não podia acreditar que não havia sido estabelecido anteriormente” (Singmaster 2004). A primeira publicação de uma versão do Problema do Aniversário se deve a von Mises (1939), que descobre que o número esperado de pares com o mesmo aniversário é de cerca de 1 quando o grupo tem 29 pessoas, porém não resolve o problema usual (Singmaster 2004). Feller (1950) estabelece a probabilidade de todas as n pessoas aniversariarem em dias diferentes, obtém uma aproximação para esta probabilidade e descobre que, para $n = 23$, tal probabilidade é inferior a 50% e, conseqüentemente, a probabilidade de ao menos um par de pessoas aniversariar no mesmo dia num grupo de 23 pessoas excede 50%.

O cálculo da probabilidade de interesse do Problema do Aniversário diretamente pelo evento “pelo menos um par de pessoas na sala aniversaria no mesmo dia” é trabalhoso no sentido de computar a quantidade de casos favoráveis a este evento, sobretudo em salas com muitas pessoas (ou seja, com n grande). Contudo, uma estratégia que facilita substancialmente este cálculo, tornando-o computacionalmente viável e muito mais rápido, é calcular a probabilidade do evento complementar de “pelo menos um par de pessoas na sala aniversaria no mesmo dia”, ou seja, calcular a probabilidade do evento “todas as pessoas na sala aniversariam em dias diferentes”. Subtraindo esta última probabilidade de 1, chega-se na probabilidade de interesse do Problema do Aniversário.

Uma vez desenvolvida a solução do Problema do Aniversário, o interesse volta-se ao uso de uma ou mais ferramentas computacionais para a realização dos cálculos de interesse. As ferramentas utilizadas neste capítulo são o *software*/linguagem R (R Core Team 2023) e os seus pacotes IPSUR (Kerns 2023a) e RcmdrPlugin.IPSUR (Kerns 2019). Tais pacotes possuem funções específicas voltadas ao Problema do Aniversário. Enquanto o primeiro deles é conveniente para quem tem um domínio mínimo da linguagem de programação R, o segundo é um *plugin* para a interface “point & click” R Commander, contornando a dificuldade de quem tem pouca ou nenhuma experiência com a linguagem R. Em particular, uma aplicação de atividade em sala de aula abordando o Problema do Aniversário com o uso da interface R Commander pode se tornar uma experiência mais agradável tanto para docentes como para discentes em relação a aplicação desta mesma atividade usando apenas linhas de comando.

2 Objetivo

O Objetivo principal deste capítulo é apresentar o Problema do Aniversário, incluindo sua solução sob certos pressupostos, e ferramentas do *software* R voltadas a este problema. De forma mais específica, deseja-se:

- apresentar, de maneira clara e didática, a solução do Problema do Aniversário, de forma que estudantes de ensino médio e de ensino superior com conhecimentos basilares em Probabilidade e Análise Combinatória possam compreendê-la;
- mostrar como instalar e carregar o pacote `IPSUR` do *software* R, o qual carrega funções voltadas ao Problema do Aniversário.
- explorar as funções do pacote `IPSUR` voltadas ao Problema do Aniversário.
- por meio de representação gráfica, verificar como a probabilidade de interesse do Problema do Aniversário evolui conforme o número de pessoas na sala aumenta;
- apresentar a interface R Commander, proveniente do pacote `Rcmdr`, e o *plugin* do pacote `IPSUR` nesta interface, denominado `RcmdrPlugin.IPSUR`;
- sugerir dinâmicas em sala de aula envolvendo o Problema do Aniversário, com suporte da interface R Commander e do pacote `RcmdrPlugin.IPSUR` para a realização de cálculos de probabilidades relacionadas a este problema.

3 Aplicação

Esta seção aborda inicialmente a solução do Problema do Aniversário. A solução clássica considera um conjunto de três pressupostos, que seguem abaixo:

1. Desconsiderar 29 de fevereiro como uma data possível de aniversário.
2. Não ter irmãos gêmeos na sala.
3. Assumir a mesma probabilidade de aniversário para cada um dos 365 dias possíveis, ou seja: $P(\text{"01/jan"}) = P(\text{"02/jan"}) = \dots = P(\text{"30/dez"}) = P(\text{"31/dez"}) = 1/365$.

Em particular, a última destas suposições permite o uso da interpretação clássica da Probabilidade para a obtenção da probabilidade de interesse. Considerando uma sala com n pessoas, seja o evento

$$B = \{\text{"pelo menos um par de pessoas aniversariar no mesmo dia"}\}.$$

A probabilidade de interesse é $P(B)$, onde B é um evento num espaço amostral finito uniforme Ω composto por 365^n elementos. Cada elemento de Ω é uma n -upla, onde a ordem é relevante. Por exemplo, para $n = 3$, $(18/mai, 16/out, 05/nov)$ e $(18/mai, 05/nov, 16/out)$ são dois elementos diferentes em Ω . O j -ésimo elemento de cada um destes ternos ordenados pode ser pensado como o aniversário da j -ésima pessoa arguída sobre sua data de aniversário.

O fato de Ω ser um espaço amostral finito uniforme faz com que a probabilidade do evento B possa ser calculada por meio da simples razão entre o número de elementos de B e o número de elementos de Ω , ou seja,

$$P(B) = \frac{\#B}{\#\Omega} = \frac{\#B}{365^n}.$$

Contudo, o numerador da fração acima é de difícil obtenção até mesmo para valores de n relativamente pequenos. Voltando ao caso de uma sala com apenas 3 pessoas (ou seja, $n = 3$), o evento B pode ser escrito como a união dos seguintes eventos mutuamente exclusivos:

- $E_1 = \{\text{“aniversário da 1ª e da 2ª pessoa no mesmo dia e da 3ª pessoa em outro dia”}\},$
- $E_2 = \{\text{“aniversário da 1ª e da 3ª pessoa no mesmo dia e da 2ª pessoa em outro dia”}\},$
- $E_3 = \{\text{“aniversário da 2ª e da 3ª pessoa no mesmo dia e da 1ª pessoa em outro dia”}\},$
- $E_4 = \{\text{“todas as 3 pessoas fazem aniversário no mesmo dia”}\}.$

O número de elementos do evento E_1 é o número de arranjos simples de 365 elementos tomados 2 a 2: $A_{365,2} = 365 \times 364 = 132860$. O mesmo vale para os eventos E_2 e E_3 , ao passo que o evento E_4 possui 365 elementos. Como $B = E_1 \cup E_2 \cup E_3 \cup E_4$ e os eventos $\{E_j : j = 1, 2, 3, 4\}$ são mutuamente exclusivos,

$$\#B = \#E_1 + \#E_2 + \#E_3 + \#E_4 = 3 \times 132860 + 365 = 398945.$$

Portanto, em uma sala com $n = 3$ pessoas, a probabilidade de ao menos um par de pessoas aniversariarem no mesmo dia é

$$P(B) = \frac{398945}{365^3} = \frac{398945}{48627125} \approx 0,0082.$$

Na tentativa de estender tal raciocínio para valores de n maiores que 3, é possível que se chegue, de maneira equivocada, em

$$\sum_{j=2}^n \frac{n!}{j!(n-j)!} \times \frac{365!}{(365 - (n-j+1))!} = \sum_{j=2}^n C_{n,j} \times A_{365,n-j+1} \quad (1)$$

como forma geral para $\#B$. Tal equívoco se dá pelo fato da Equação (1) subestimar $\#B$. Por exemplo, para $n = 4$, a Equação (1) retorna 289.900.885 elementos, quando a cardinalidade de B para $n = 4$ é, na verdade, 290.299.465. Os 398.580 elementos não computados pela Equação (1) são os elementos do evento (contido em B) que contempla duas das quatro pessoas aniversariando no mesmo dia e as outras duas pessoas também fazendo aniversário num mesmo dia, porém, numa data diferente (por exemplo, o elemento (02/ago, 19/nov, 19/nov, 02/ago)).

Portanto, não é difícil concluir que, conforme n aumenta, torna-se cada vez mais difícil o cálculo de $\#B$. Por este motivo, a solução do Problema do Aniversário volta-se ao cálculo do número de elementos do evento complementar de B , isto é,

$$B^C = \{\text{“todas as } n \text{ pessoas na sala aniversariam em dias diferentes”}\},$$

e a probabilidade do evento B pode ser obtida por:

$$P(B) = 1 - P(B^C) = 1 - \frac{\#B^C}{365^n}.$$

Para qualquer $n \in \{2, 3, \dots, 365\}$, é direto concluir que

$$\#B^C = 365 \times \dots \times (365 - (n - 1)) = \frac{365!}{(365 - n)!} = A_{365, n}.$$

Para salas com mais de 365 pessoas, o Princípio de Dirichlet (Morgado et al. 1991) garante que ao menos 2 pessoas nesta sala fazem aniversário no mesmo dia, uma vez que há mais pessoas na sala que datas de aniversário possíveis. Dito isto, $P(B) = 1$ (ou ainda, $P(B^C) = 0$) para $n > 365$. Portanto, em uma sala com n pessoas, a probabilidade de ao menos um par de pessoas aniversariar no mesmo dia é dada por

$$P(B) = \begin{cases} 1 - \frac{365!}{365^n \times (365 - n)!} & , \text{ se } n \leq 365, \\ 1 & , \text{ se } n > 365, \end{cases}$$

ou ainda, utilizando notação de número de arranjos simples,

$$P(B) = \begin{cases} 1 - \frac{A_{365, n}}{365^n} & , \text{ se } n \leq 365, \\ 1 & , \text{ se } n > 365. \end{cases}$$

Note que a solução acima vale para qualquer n inteiro positivo. Em particular, para $n = 1$, tem-se $P(B) = 0$, o que faz todo o sentido, pois é impossível um par de pessoas aniversariar no mesmo dia se não há sequer um par de pessoas nesta sala.

Uma vez desenvolvida e finalizada a solução do Problema do Aniversário, cabe o uso de alguma ferramenta computacional para calcular a expressão

$$1 - \frac{A_{365,n}}{365^n} \quad (2)$$

para o(s) valor(es) desejado(s) de n (ou ainda, para uma probabilidade pré-definida, calcular qual o menor n que faz a expressão acima ser maior ou igual a esta probabilidade). A ferramenta utilizada aqui é o *software* R, que consiste em um programa gratuito e de código aberto, comumente usado para tratamento de dados e análises estatísticas. De forma mais genérica, ele também pode ser pensado como uma linguagem de programação, em particular uma linguagem orientada a objetos. Seu *download* pode ser feito em «The Comprehensive R Archive Network» (2023).

Uma forte característica do *software*/linguagem R é sua abundante quantidade de pacotes disponíveis. Em particular, dois pacotes que trazem consigo funções associadas ao Problema do Aniversário são objeto de estudo neste capítulo: os pacotes `IPSUR` e `RcmdrPlugin.IPSUR`. Pacotes no *software* R podem ser facilmente instalados por meio do menu *Pacotes > Instalar pacote(s)* (ou *Tools > Install packages*, para usuários do ambiente de desenvolvimento integrado Rstudio), ou ainda via linha de comando, na forma

```
install.packages("nome_do_pacote")
```

para apenas um pacote ou

```
install.packages(c("nome_do_pacote_1", ..., "nome_do_pacote_n"))
```

para dois ou mais pacotes, desde que o(s) pacote(s) em questão esteja(m) no repositório CRAN (*Comprehensive R Archive Network*). Entretanto, nem todo pacote atualmente disponível para o R está listado neste repositório, uma vez que alguns pacotes que outrora pertenceram a ele foram, em algum momento, arquivados. Este é o contexto no qual se encontram os pacotes `IPSUR` e `RcmdrPlugin.IPSUR`. Felizmente, há como instalar pacotes ausentes do repositório CRAN, porém presentes no repositório *Github*, por meio da função `install_github` do pacote `devtools` (Wickham et al. 2022), o qual está disponível no repositório CRAN. Ainda em relação ao uso do pacote `devtools`, é necessário que usuários do sistema operacional Windows instalem previamente um programa auxiliar denominado Rtools, conforme explicitado no Passo 2 em «devtools» (2014), que também esclarece a necessidade da instalação do Xcode para usuários de Mac e de um compilador e várias bibliotecas de desenvolvimento para usuários de distribuições Linux. Os parágrafos seguintes desta seção direcionam-se aos procedimentos para instalação dos pacotes `IPSUR` e `RcmdrPlugin.IPSUR`, considerando que as instalações solicitadas acima tenham sido realizadas conforme o sistema operacional em uso.

O pacote `IPSUR`, cujo nome é uma sigla para *Introduction to Probability and Statistics Using R*, não está mais disponível no repositório CRAN desde 30 de maio de 2019, “pois os problemas de verificação não foram corrigidos a tempo, apesar dos lembretes” («Package ‘IPSUR’ was removed from the CRAN repository» 2019). Logo, não é possível instalá-lo pelos métodos

convencionais abordados no parágrafo anterior. Uma alternativa para a sua instalação se dá por executar as linhas de comando abaixo, conforme disponibilizado em «IPSUR: Introduction to Probability and Statistics Using R» (2022). Note que o pacote `prob` (Kerns 2023b), uma das dependências do pacote IPSUR, também não está disponível no repositório CRAN, uma vez que é necessário instalá-lo por meio da função `install_github`. Tal pacote foi arquivado pelo CRAN em 29 de abril de 2022, “pois requer o pacote `fAsiaOptions`” («Package ‘prob’ was removed from the CRAN repository» 2022), sendo que este último foi removido, na mesma data, por “deturpação da autoria e propriedade de direitos autorais” («Package ‘fAsianOptions’ was removed from the CRAN repository» 2022). A dependência do pacote `fAsiaOptions` para o pacote `prob` foi removida pelo autor do último, sem perdas significativas.

```
## INSTALANDO PACOTES DO REPOSITÓRIO CRAN REFERENCIADOS NO IPSUR:
install.packages(c("actuar", "aplpack", "binom", "boot", "coin", "diagram",
                  "distrEx", "e1071", "emdbook", "ggplot2", "HH", "Hmisc",
                  "lmtree", "mvtnorm", "qcc", "reshape", "RcmdrMisc",
                  "scatterplot3d", "TeachingDemos", "vcd"))

## INSTALAÇÃO DO PACOTE prob, QUE NÃO ESTÁ MAIS NO REPOSITÓRIO CRAN:
install.packages("combinat") # combinat é uma dependência do pacote prob
install.packages("devtools") # caso devtools não esteja instalado ainda
devtools::install_github("gjkerns/prob")

## FINALIZANDO INSTALAÇÃO DO PACOTE IPSUR:
devtools::install_github("gjkerns/IPSUR")
```

O pacote `RcmdrPlugin.IPSUR` é um *plugin* do pacote IPSUR para a interface R Commander, fornecida pelo pacote `Rcmdr` (Fox e Bouchet-Valat 2022). Mais detalhes sobre esta interface, incluindo instalação e carregamento do pacote `Rcmdr`, podem ser consultadas em Melo (2019). Em particular, há vários pacotes que são *plugins* para a interface R Commander no repositório CRAN, todos iniciados em “`RcmdrPlugin`”. Todavia, o pacote `RcmdrPlugin.IPSUR` não é mais um destes pacotes disponibilizados neste repositório desde 10 de agosto de 2022 por conta de “problemas que não foram corrigidos a tempo” («Package ‘RcmdrPlugin.IPSUR’ was removed from the CRAN repository» 2022). Assumindo que o pacote `devtools` já está instalado, conforme parte do passo a passo acima, o pacote `RcmdrPlugin.IPSUR` pode ser instalado simplesmente pela execução da linha de comando

```
devtools::install_github("gjkerns/RcmdrPlugin.IPSUR")
```

Caso algum procedimento de instalação acima falhe por conta de versões antigas de outros pacotes, uma solução é acessar o menu *Pacotes > Atualizar pacotes* (ou, equivalentemente, o menu *Tools > Check for Packages Updates* do Rstudio) e atualizar todos os pacotes para suas versões mais recentes. Feito isto, refaça o procedimento de instalação de pacote que falhou.

4 Resultados e Discussão

Uma vez instalado o pacote IPSUR, carregue-o por meio dos comandos

```
library(IPSUR)
```

ou

```
require(IPSUR)
```

Para saber quais funções o pacote IPSUR fornece, basta executar a linha de comando

```
ls("package:IPSUR")
```

```
[1] "pbirthday.ipsur" "qbirthday.ipsur" "read"
```

Como pode ser visto acima, são apenas três funções. Quanto a função `read`, executar a linha de comando

```
read(IPSUR)
```

abre o arquivo *IPSUR.pdf*, que é a edição mais recente, tomando por base o momento de instalação do pacote, do livro *Introduction to Probability and Statistics Using R* (Kerns 2021). No sistema operacional Windows, em particular, o arquivo *IPSUR.pdf* encontra-se na pasta de usuário em `AppData\Local\R\win-library\4.3\IPSUR\doc` (para usuários de versão 4.3.x do R e que optaram pela instalação padrão deste). Com 15 capítulos e 5 anexos distribuídos em cerca de 350 páginas, a quarta edição do livro *Introduction to Probability and Statistics Using R* traz o Problema do Aniversário em seu Exemplo 4.7, o qual é finalizado com um curto *script* para a geração de um gráfico com o número de pessoas na sala no eixo horizontal e a probabilidade de ao menos um par destas pessoas aniversariar num mesmo dia no eixo vertical. Tal *script* utiliza a função `pbirthday.ipsur`, uma das três funções fornecidas pelo pacote IPSUR e detalhada a seguir com seus valores *default*, tal como a função `qbirthday.ipsur`.

- `pbirthday.ipsur(n, classes=365, coincident=2)`: calcula a probabilidade de, numa sala com n pessoas, ao menos 2 fazerem aniversário no mesmo dia, considerando 365 possibilidades de dias.
- `qbirthday.ipsur(prob=0.5, classes=365, coincident=2)`: retorna o menor valor n cuja probabilidade de, numa sala com n pessoas, ao menos 2 fazerem aniversário no mesmo dia, considerando 365 possibilidades de dias, seja maior ou igual ao argumento `prob`.

Os valores *default* dos argumentos `classes` e `coincident` consistem, de fato, no Problema do Aniversário. As linhas de comando abaixo ilustram que a probabilidade de, em uma sala com n pessoas, pelo menos um par delas fazer aniversário no mesmo dia é (aproximadamente): 11,69% para $n = 10$, 50,73% para $n = 23$, 70,63% para $n = 30$, 89,12% para $n = 40$ e 97,04% para $n = 50$. Para $n = 100$, note que é necessário uma precisão de pelo menos 7 casas decimais para que a probabilidade solicitada não retorne 1. Para $n = 120$, note que a saída foi 1, porém isto é um valor aproximado, por conta do padrão de precisão do R ser 7 casas decimais. Alterando a precisão de 7 para 22 casas decimais, a linha de comando que anteriormente retornou 1 passa a retornar 0,9999999997560852227352. De fato, o evento “pelo menos duas de 120 pessoas aniversariarem em um mesmo dia” não é um evento certo para ter probabilidade 1, por mais improvável que seja a sua não ocorrência.

```
pbirthday.ipsur(10)
```

```
[1] 0.1169482
```

```
pbirthday.ipsur(23)
```

```
[1] 0.5072972
```

```
pbirthday.ipsur(30)
```

```
[1] 0.7063162
```

```
pbirthday.ipsur(40)
```

```
[1] 0.8912318
```

```
pbirthday.ipsur(50)
```

```
[1] 0.9703736
```

```
pbirthday.ipsur(100)
```

```
[1] 0.9999997
```

```
pbirthday.ipsur(120)
```

```
[1] 1
```

```
options(digits=22)  
pbirthday.ipsur(120)
```

```
[1] 0.9999999997560852227352
```

Abaixo, seguem resultados obtidos pela função `qbirthday.ipsur`. Note que bastam 15 pessoas na sala para que a probabilidade de ao menos um par destas pessoas aniversariem no mesmo dia seja igual ou ultrapasse 25% (com 14 pessoas, tal probabilidade é menor que 25%). De maneira análoga, a justificativa da expressão “paradoxo do aniversário” se dá ao aplicar esta função em 50%, resultando em apenas 23 pessoas. Aplicando esta função em probabilidades cada vez maiores (75%, 90%, 95% e 99%), claramente os resultados são números cada vez maiores, porém ainda distantes de 365. Basta uma sala ter 57 pessoas, por exemplo, para que a probabilidade de ocorrer ao menos uma coincidência de aniversários atinja ou ultrapasse 99%. Por fim, note que `qbirthday.ipsur(1)` retorna 366, uma vez que este é o menor valor para n que faz com que o evento “Ao menos um par destas n pessoas faz aniversário no mesmo dia” seja um evento certo. Para $n = 365$, este evento não é um evento certo, em que pese ser extremamente improvável não ser satisfeito.

```
qbirthday.ipsur(0.25)
```

```
[1] 15
```

```
qbirthday.ipsur(0.50)
```

```
[1] 23
```

```
qbirthday.ipsur(0.75)
```

```
[1] 32
```

```
qbirthday.ipsur(0.90)
```

```
[1] 41
```

```
qbirthday.ipsur(0.95)
```

[1] 47

```
qbirthday.ipsur(0.99)
```

[1] 57

```
qbirthday.ipsur(1)
```

[1] 366

Voltando ao Exemplo 4.27 em Kerns (2021), mais especificamente ao *script* que encerra este exemplo, é possível verificar graficamente a evolução da probabilidade de interesse do Problema do Aniversário conforme n cresce. A Figura 1, baseada neste *script*, porém com algumas adaptações, ilustra tal evolução de $n = 1$ até $n = 50$, com destaque para $n = 23$, menor valor de n cuja probabilidade associada ultrapassa 50%.

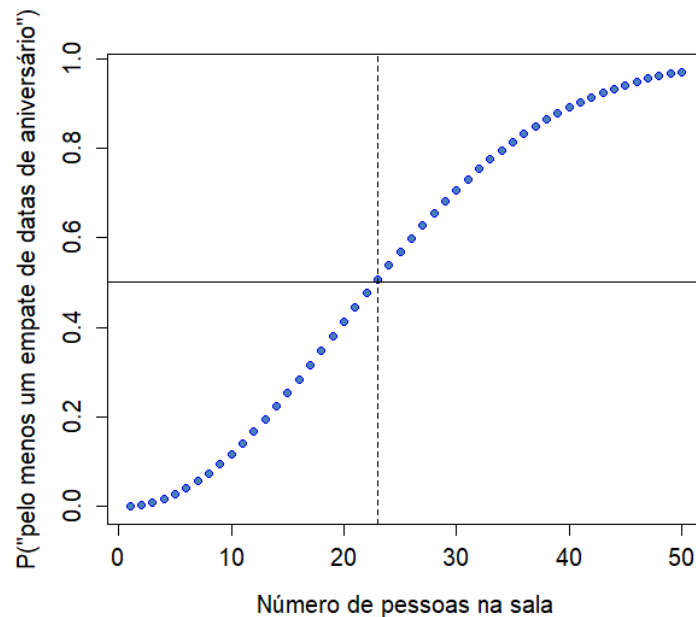


Figura 1: Evolução da probabilidade de interesse do Problema do Aniversário conforme n cresce. Fonte: Os autores.

O *script* que gera a Figura 1 segue abaixo:

```

g <- Vectorize(pbirthday.ipsur)
plot(1:50, g(1:50), xlab = 'Número de pessoas na sala',
     ylab = 'P("pelo menos um empate de datas de aniversário")',
     pch=21, bg='steelblue', col='blue', cex.lab=1.25, cex.axis=1.25)
abline(h = 0.5)
abline(v = 23, lty = 2)
remove(g)

```

Exploradas todas as funções do pacote IPSUR, o interesse volta-se ao pacote `RcmdrPlugin.IPSUR`. Conforme já mencionado, este pacote é um *plugin* para a interface R Commander, proveniente de carregamento do pacote `Rcmdr`. Como o pacote `Rcmdr` está disponível no repositório CRAN, sua instalação pode ser feita por um dos métodos convencionais mencionados na Seção 3, como por exemplo, executar a linha de comando

```
install.packages("Rcmdr")
```

Feita esta instalação, carregue o pacote `Rcmdr` por meio dos comandos

```
library(Rcmdr)
```

ou

```
require(Rcmdr)
```

para que seja aberta a janela R Commander (Figura 2), que fornece ao usuário do R uma interface mais “amigável”, com menus que propiciam várias funcionalidades do R sem a necessidade de escrever linhas de comando. Para dinâmicas em sala de aula envolvendo o Problema do Aniversário, seja em turmas de ensino básico ou superior, é recomendável o uso desta interface com menus, de forma a tornar a atividade mais prática, fácil e motivadora, inclusive para os estudantes mais interessados a reproduzirem em outros ambientes fora da sala de aula.

Inicialmente, nenhum dos menus da janela R Commander contém funcionalidades do *plugin* `RcmdrPlugin.IPSUR`, pois este ainda não foi carregado. Uma vez que a janela R Commander esteja aberta, qualquer um de seus *plugins* previamente instalados pode ser carregado por meio do menu (da janela R Commander) *Ferramentas > Carregar plug-in(s) do Rcmdr*. Feito este procedimento para carregar o *plugin* `RcmdrPlugin.IPSUR`, a interface R Commander é reiniciada e novas funcionalidades são adicionadas aos menus da janela R Commander, indicadas pelo rótulo “(IPSUR)”, como o menu *Distribuições > Birthday Problem... (IPSUR)*, que possibilita o uso das funções `pbirthday.ipsur` e `qbirthday.ipsur` sem a necessidade de digitar de linhas de comando (Figura 3). Após escolher a opção desejada (Probabilidades ou Quantis) e preencher o campo associado à escolha marcada, o resultado será mostrado na janela *Output* da janela R Commander. Caso esta janela *Output* não esteja em exibição, o resultado será mostrado na janela onde foi carregado o pacote `Rcmdr`.

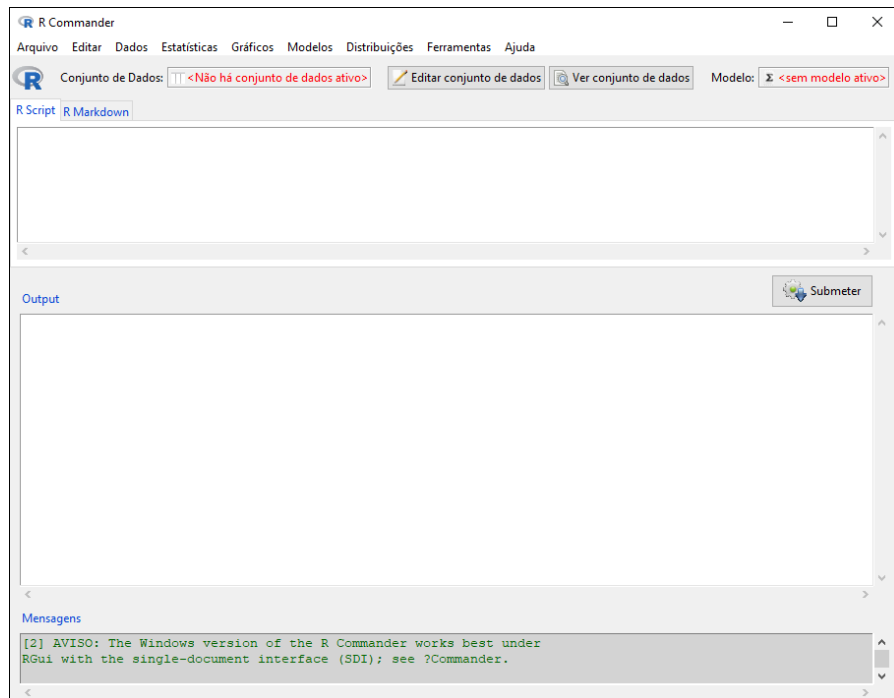


Figura 2: Janela R Commander. Fonte: Os autores.

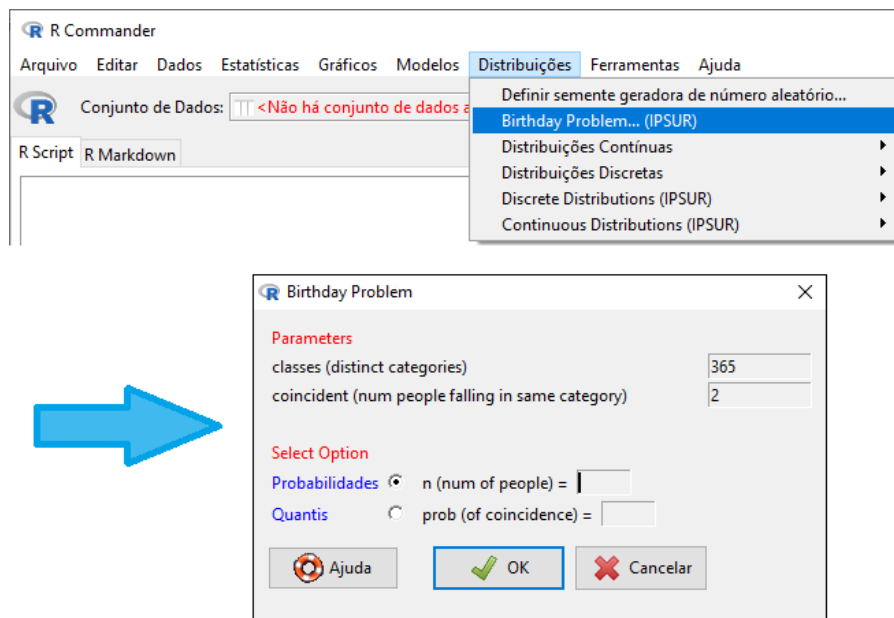


Figura 3: Menu do R Commander para o Problema do Aniversário, adicionado pelo *plugin* RcmdrPlugin.IPSUR. Fonte: Os autores.

O que foi exposto acima considerou o carregamento do *plugin* do pacote IPSUR para o R Commander com esta interface aberta, via navegação por menus. Caso a janela R Commander não esteja aberta, carregar diretamente o pacote `RcmdrPlugin.IPSUR` abre a janela R Commander com as funcionalidades deste *plugin* carregado. À título de teste, feche a janela R Commander e a sessão atual do R. Abra uma nova sessão do R e carregue o pacote `RcmdrPlugin.IPSUR` por meio de uma das linhas da comando abaixo:

```
library(RcmdrPlugin.IPSUR)
```

ou

```
require(RcmdrPlugin.IPSUR)
```

Cabe ressaltar um ponto para novos usuários (ou usuários pouco experientes) da interface R Commander: quando a janela R Commander é fechada, não é possível reabri-la na atual sessão do R. Carregar novamente o pacote `Rcmdr` ou qualquer um de seus *plugins* na atual sessão não surte efeito. Neste cenário, é necessário abrir uma nova sessão do R e carregar novamente o pacote `Rcmdr` (ou o seu *plugin* de interesse) para acessar novamente a janela R Commander.

Além do menu voltado ao Problema do Aniversário, o pacote `RcmdrPlugin.IPSUR` adiciona ao R Commander outras funcionalidades nos menus *Distribuições*, *Gráficos* e *Estatísticas*, além de três conjuntos de dados que podem ser carregados no menu *Dados > Conjunto de dados em pacotes > Ler dados de pacote “attachado”*. Mais detalhes sobre as funcionalidades deste pacote seguem em Kerns (2014).

5 Conclusão

A solução do Problema do Aniversário pode parecer, a princípio, bastante complicada, sobretudo em salas com um número grande de pessoas, desde que não ultrapasse 365. Contudo, a contagem de n -uplas com ao menos uma coincidência de datas de aniversário é potencialmente facilitada ao se buscar pela diferença entre o total de n -uplas possíveis e o total de n -uplas nas quais todas as datas são diferentes. Ao visualizar a Equação (2), é natural imaginar valores altos para n (superiores à metade de 365) quando esta equação é igualada a probabilidades maiores que 50%. Porém, conforme verificado na Figura 1 e nas saídas ilustradas pela função `qbirthday.ipsur`, tais valores de n estão muito aquém da metade de 365. Por exemplo, a probabilidade de ao menos um par de pessoas fazer aniversário num mesmo dia em uma sala com 41 pessoas ultrapassa 90%, e 41 é inferior a um oitavo de 365.

O fato da solução do Problema do Aniversário ser contra-intuitiva torna este problema ainda mais interessante. Em uma sala de aula, o docente pode fazer o questionamento da probabilidade do evento de interesse do Problema do Aniversário considerando os estudantes presentes

na sala. Supondo que há 30 alunos presentes, é intuitivo que eles atribuam (“chutem”) probabilidades bem baixas, levando em conta que 30 é bem inferior a 365 e também supondo que cada estudante desconheça as datas de aniversário dos colegas. Todavia, para $n = 30$, a solução do Problema do Aniversário estabelece uma probabilidade de 70,63% de que há pelo menos duas pessoas aniversariando num mesmo dia. Ou seja, é bem provável que o professor, ao abordar os alunos, um por um, perguntando cada data de aniversário, receba como resposta, em algum momento, uma data que já foi mencionada, o que pode surpreender os alunos mais incrédulos. Uma dinâmica em sala de aula pode ser iniciada assim para que, na sequência, seja apresentado o Problema do Aniversário, sua solução e aplicações das funções `pbirthday.ipsur` e `qbirthday.ipsur` por meio da interface R Commander, com advento do pacote `RcmdrPlugin.IPSUR`. Ou ainda, para criar um clima de suspense e, desta forma, gerar maior expectativa e interesse da turma, a etapa da arguição das datas de aniversário dos estudantes pode ser deixada para o encerramento da dinâmica.

É importante reforçar, antes e após a solução do Problema do Aniversário, os três pressupostos elencados no começo da Seção 3. O primeiro deles é bem razoável, pois ainda que uma pessoa tenha nascido em um dia 29 de fevereiro, seu registro de aniversário consta como 28 de fevereiro ou 1º de março. Sobre não ter gêmeos na sala, isto pode ser adaptado considerando que os irmãos gêmeos constituem um único elemento. Por exemplo, se há 30 alunos presentes na aula e, dentre eles, há um par de irmãos gêmeos, o problema deve ser considerado com $n = 29$ em vez de $n = 30$. O pressuposto restante (de assumir a mesma probabilidade para todas as 365 datas de aniversário possíveis) é essencial para a solução, pois é por conta disto que é possível obter a solução do problema pela interpretação clássica da probabilidade, isto é, para um evento E em um espaço amostral Ω , a probabilidade do evento E é dada por $P(E) = \#E/\#\Omega$. Contudo, dos três pressupostos, este é, na prática, o menos razoável, uma vez que podem haver períodos tipicamente com maior frequência de aniversários quando comparados a outros períodos com a mesma quantidade de dias, dentro de cada ano.

Sobre o desenvolvimento da solução do Problema do Aniversário com a turma, é importante que o docente tenha noção da base matemática da sua turma, seja ela de ensino médio ou ensino superior, sobretudo em conceitos básicos de análise combinatória e interpretação clássica da Probabilidade. Em turmas com maior dificuldade de compreensão destes conceitos, uma estratégia é adotar um conjunto de possibilidades menor. Em vez de 365 dias, adotar os 12 meses de ano, ou os 12 signos do zodíaco, ou ainda os 7 dias da semana. Desta forma, o argumento `classes` das funções `pbirthday.ipsur` e `qbirthday.ipsur` deve ser mudado (de 365 para 12 ou 7), e os pressupostos da solução do Problema do Aniversário devem ser devidamente adaptados, ainda que os meses do ano não tenham todos a mesma quantidade de dias, o mesmo ocorrendo para os períodos de cada signo. Entretanto, a identificação de elementos (n -uplas) de um espaço amostral que satisfazem ou não certo evento pode ser mais plausível quando este espaço amostral possui 7^n ou 12^n elementos em vez de 365^n elementos. Uma vez compreendida a forma geral da solução para ao menos uma coincidência de dias da semana, meses ou signos, a extensão do raciocínio para os 365 dias segue de forma natural.

Por fim, cabe ressaltar que o *software* R possui, de forma nativa (pacote `stats`), as funções

`pbirthday` e `qbirthday`, que retornam os mesmos resultados das funções `pbirthday.ipsur` e `qbirthday.ipsur` quando `coincident=2`. Porém, as principais motivações em apresentar e explorar o pacote `IPSUR` neste capítulo se dão pelo livro *Introduction to Probability and Statistics Using R* e como estudo preliminar para apresentar e explorar o pacote `RcmdrPlugin.IPSUR`, de forma a entender quais linhas de comando estão “por trás” do menu dedicado ao Problema do Aniversário propiciado por este pacote à interface R Commander.

6 Referências

- «devtools». 2014. Disponível em: <https://www.r-project.org/nosvn/pandoc/devtools.html>. Acesso em: 31 jul. 2023.
- Feller, William. 1950. *An Introduction to Probability Theory and Its Applications: Vol 1*. New York, USA: John Wiley; Sons, Inc.
- Fox, John, e Milan Bouchet-Valat. 2022. *Rcmdr: R Commander*. <https://socialsciences.mcmaster.ca/jfox/Misc/Rcmdr/>.
- «IPSUR: Introduction to Probability and Statistics Using R». 2022. Disponível em: <https://ipsur.org/>. Acesso em: 31 jul. 2023.
- Kerns, Gary Jay. 2014. «Package 'RcmdrPlugin.IPSUR'». Disponível em: <http://cran.nexr.com/web/packages/RcmdrPlugin.IPSUR/RcmdrPlugin.IPSUR.pdf>. Acesso em: 31 jul. 2023.
- . 2019. *RcmdrPlugin.IPSUR: An IPSUR Plugin for the R Commander*. <https://CRAN.R-project.org/package=RcmdrPlugin.IPSUR>.
- . 2021. *Introduction to Probability and Statistics Using R*. GNU Free Documentation License.
- . 2023a. *IPSUR: Introduction to Probability and Statistics Using R*.
- . 2023b. *prob: Elementary Probability on Finite Sample Spaces*. <https://github.com/gjkerns/prob/>.
- Melo, Felipe Rafael Ribeiro. 2019. *Introdução ao R Commander: Notas de Aula*. Rio de Janeiro, Brasil: Departamento de Métodos Quantitativos da Universidade Federal do Estado do Rio de Janeiro. http://gae.uniriotec.br/7/Notas_de_aula_Rcmdr.pdf.
- Morgado, Augusto César de Oliveira, João Bosco Pitombeira Carvalho, Paulo César Pinto Carvalho, e Pedro Fernandez. 1991. *Análise Combinatória e Probabilidade*. Rio de Janeiro: IMPA.
- «Package 'fAsianOptions' was removed from the CRAN repository». 2022. Disponível em: <https://CRAN.R-project.org/package=fAsianOptions>. Acesso em: 31 jul. 2023.
- «Package 'IPSUR' was removed from the CRAN repository». 2019. Disponível em: <https://CRAN.R-project.org/package=IPSUR>. Acesso em: 31 jul. 2023.
- «Package 'prob' was removed from the CRAN repository». 2022. Disponível em: <https://CRAN.R-project.org/package=prob>. Acesso em: 31 jul. 2023.
- «Package 'RcmdrPlugin.IPSUR' was removed from the CRAN repository». 2022. Disponível em: <https://CRAN.R-project.org/package=RcmdrPlugin.IPSUR>. Acesso em: 31 jul. 2023.

- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Singmaster, David. 2004. «Sources in Recreational Mathematics: An Annotated Bibliography». Disponível em: https://www.puzzlemuseum.com/singma/singma6/SOURCES/singma-sources-edn8-2004-03-19.htm#_Toc69533810. Acesso em: 31 jul. 2023.
- «The Comprehensive R Archive Network». 2023. Disponível em: <https://cloud.r-project.org/>. Acesso em: 31 jul. 2023.
- von Mises, Richard. 1939. «Über Aufteilungs- und Besetzungswahrscheinlichkeiten». *Revue de la faculté des sciences de l'Université d'Istanbul* 4 (1): 145–63.
- Wickham, Hadley, Jim Hester, Winston Chang, e Jennifer Bryan. 2022. *devtools: Tools to Make Developing R Packages Easier*. <https://CRAN.R-project.org/package=devtools>.