

UNIVERSIDADE DE LISBOA  
FACULDADE DE CIÊNCIAS  
DEPARTAMENTO DE ESTATÍSTICA E INVESTIGAÇÃO OPERACIONAL



**Ciências**  
**ULisboa**

**Cálculo da Dimensão Amostral em Estudos  
Seroepidemiológicos: da Teoria ao Desenvolvimento de uma  
Ferramenta Computacional**

Márcia Maria Gonçalves de Brito Morato da Graça

**Mestrado em Bioestatística**

Dissertação orientada por:  
Professora Doutora Maria Fernanda Nunes Diamantino  
Professor Doutor Nuno Henriques dos Santos de Sepúlveda



# Agradecimentos

Gostaria de agradecer aos meus orientadores pela ajuda e pelas sugestões, que foram fundamentais para aprimorar o meu trabalho e a sua qualidade final.

À minha família, especialmente aos meus avós, Augusta e Joaquim.

A todas as pessoas que, de alguma forma, fizeram parte da minha jornada e me inspiraram nas mais diversas áreas da minha vida.



# Resumo

Este trabalho explora o cálculo da dimensão amostral, com foco em estudos transversais seroepidemiológicos sobre a malária, uma doença infecciosa transmitida pela picada do mosquito *Anopheles*.

O trabalho inicia com a apresentação de métodos para a construção de intervalos de confiança para proporções e o respetivo cálculo da dimensão amostral. Em seguida, são explorados os modelos catalíticos reversíveis (RCM). Estes modelos são baseados numa cadeia de Markov com dois estados (seronegativo e seropositivo), nos quais os indivíduos podem transitar entre os estados segundo um processo de seroconversão e serorreversão, respetivamente.

Com base na teoria dos RCM, desenvolveu-se uma ferramenta computacional com recurso ao *software* R, projetada para calcular a dimensão amostral necessária, considerando a estrutura etária e as características seroepidemiológicas da população. A ferramenta demonstrou um bom desempenho nos diferentes cenários testados.

Os resultados indicam que a dimensão amostral necessária varia com a intensidade de transmissão e a correspondente seroprevalência (SP), sendo também influenciada pela distribuição etária, que afeta a inclinação da curva que relaciona a taxa de seroconversão (SCR) com a SP. Nomeadamente, valores de SP próximos de 0 ou 1 necessitam de maiores dimensões amostrais, em comparação com valores de SP intermédios.

Para  $SCR \leq 0.058 \text{ ano}^{-1}$ , estudos sobre populações africanas requerem maiores dimensões amostrais do que as do Sudeste Asiático e América do Sul para obter o mesmo nível de precisão.

Observou-se também que a aplicação da correção de continuidade ao método de Wald melhora a sua cobertura. O método de Wilson demonstrou ser o melhor em diversos cenários, apresentando uma boa probabilidade de cobertura (perto ou acima do nível de confiança nominal) e uma das menores dimensões amostrais necessárias. O método de Agresti-Coull apresentou melhores resultados no Sudeste Asiático e América do Sul em comparação com África.

**Palavras chave:** Dimensão amostral, Estudos seroepidemiológicos, Malária, Modelos Catalíticos Reversíveis, Intervalos de confiança



# Abstract

This work explores sample size calculation, with a specific focus on cross-sectional seroepidemiological studies of malaria, an infectious disease transmitted by the bite of the *Anopheles* mosquito.

The study begins by presenting methods for constructing confidence intervals for proportions and the corresponding sample size calculation. It then delves into the application of Reversible Catalytic Models (RCM). These models are based on a Markov chain with two states (seronegative and seropositive), in which individuals can transition between states through the processes of seroconversion and seroreversion, respectively.

Building upon RCM theory, a computational tool was developed using R software to calculate the required sample size, accounting for the age structure and seroepidemiological characteristics of the target population. The tool demonstrated good performance in the tested scenarios.

The results indicate that the required sample size varies with the transmission intensity and the corresponding seroprevalence (SP), and is also influenced by the age structure of the population, which affects the slope of the seroconversion rate (SCR) vs. SP curve. In particular, SP values close to 0 or 1 require larger sample sizes compared to intermediate SP values.

It was also observed that for  $SCR \leq 0.058 \text{ year}^{-1}$ , studies on African populations require larger sample sizes than those in Southeast Asia and South America to achieve the same level of precision.

Furthermore, applying a continuity correction to the Wald method resulted in improved coverage compared to the uncorrected approach. The Wilson method was found to be the best in several scenarios, providing good coverage (close to or above the nominal confidence level) and one of the smallest required sample sizes. On the other hand, the Agresti-Coull method performed better in Southeast Asian and South American populations compared to African populations.

**Keywords:** Sample size, Seroepidemiological studies, Malaria, Reversible Catalytic Models, Confidence intervals



# Lista de Abreviaturas

- CP *Coverage probability*, Probabilidade de cobertura  
CRAN *The Comprehensive R Archive Network*  
EIR *Entomological inoculation rate*, Taxa de inoculação entomológica  
EL *Expected length*, Amplitude esperada  
RCM *Reversible Catalytic Models*, Modelos Catalíticos Reversíveis  
RL *Relative length*, Amplitude relativa  
SCR *Seroconversion rate*, Taxa de seroconversão  
SP *Seroprevalence*, Seroprevalência  
SRR *Seroreversion rate*, Taxa de serorreversão



# Índice

<b>Lista de Figuras</b>	<b>xiii</b>
<b>Lista de Tabelas</b>	<b>xv</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Objetivo . . . . .	3
1.2 Estrutura do trabalho . . . . .	3
<b>2 Cálculo da dimensão amostral para medidas epidemiológicas baseadas em proporções</b>	<b>5</b>
2.1 Algoritmo geral . . . . .	6
2.2 Métodos de construção de intervalos de confiança para proporções . . . . .	7
2.2.1 Métodos exatos . . . . .	8
2.2.1.1 Método de Clopper-Pearson . . . . .	8
2.2.2 Métodos assintóticos . . . . .	9
2.2.2.1 Método de Wald . . . . .	10
2.2.2.2 Método de Wilson . . . . .	11
2.2.2.3 Método de Agresti-Coull . . . . .	12
2.2.2.4 Métodos baseados em transformações . . . . .	13
2.2.3 Sugestões práticas para seleção do método . . . . .	15
2.3 Medidas de desfecho principal . . . . .	15
2.3.1 Prevalência de infecção . . . . .	16
2.3.2 Seroprevalência . . . . .	16
<b>3 Cálculo da dimensão amostral para a taxa de seroconversão</b>	<b>17</b>
3.1 Modelos catalíticos reversíveis . . . . .	17
3.2 Algoritmo para o cálculo da dimensão amostral para a taxa de seroconversão . . . . .	18
<b>4 Ferramenta computacional</b>	<b>21</b>
4.1 Tutorial . . . . .	22
4.1.1 Cenário 1: Um único valor da SCR . . . . .	23
4.1.2 Cenário 2: Um conjunto de valores da SCR . . . . .	23
4.1.3 Cenário 3: Comparação de distribuições etárias . . . . .	25
4.1.4 Cenário 4: Comparação de diferentes métodos de construção de intervalos de confiança . . . . .	27
4.1.5 Informações . . . . .	29

## ÍNDICE

<b>5 Resultados</b>	<b>33</b>
5.1 Reprodução dos resultados publicados . . . . .	33
5.2 Desempenho dos métodos de construção de intervalos de confiança para a SCR . . . . .	37
<b>6 Discussão e conclusão</b>	<b>41</b>
<b>A Cálculo da dimensão amostral no método de Wald com correção de continuidade</b>	<b>43</b>
<b>B Método de Wilson</b>	<b>44</b>
<b>C Transmissão estável</b>	<b>46</b>
<b>D Funções disponíveis no pacote <code>RCMsize</code></b>	<b>47</b>
<b>E Funções para o cálculo da probabilidade de cobertura no método de Wald com correção de continuidade</b>	<b>49</b>
<b>Bibliografia</b>	<b>51</b>

# Lista de Figuras

1.1	Ciclo da transmissão de malária. Fonte: [Centers for Disease Control and Prevention, 2024]	2
3.1	Representação da cadeia de Markov associado a um RCM onde $S^+$ e $S^-$ correspondem aos estados de seropositividade e seronegatividade, respetivamente. As transições entre os estados são representadas por $\lambda$ (SCR) e $\rho$ (SRR).	17
3.2	Relação entre a SCR e a SP, para duas populações distintas, considerando $\rho = 0.017$ .	19
4.1	Distribuição etária para as duas populações.	23
4.2	Cálculo da dimensão amostral com uma taxa de seroconversão de 0.0324, taxa de serorreversão de 0.017 e nível de confiança de 95%, para o conjunto de dados referentes à população africana.	24
4.3	Intervalos de confiança para uma taxa de seroconversão de 0.0324, taxa de serorreversão de 0.017 e nível de confiança de 95%, para o conjunto de dados referentes à população africana.	24
4.4	SCR vs. SP para uma taxa de seroconversão de 0.0324, taxa de serorreversão de 0.017 e nível de confiança de 95%, para o conjunto de dados referentes à população africana.	25
4.5	Distribuição etária da população africana.	25
4.6	Resumo da teoria para o cálculo da dimensão amostral necessária.	26
4.7	Dimensão amostral calculada para diferentes valores de SCR.	26
4.8	Função da seroprevalência para a população africana com uma SRR de 0.017.	27
4.9	Gráfico da dimensão amostral calculada para diferentes valores de SCR.	27
4.10	Cálculo da dimensão amostral para duas populações com distribuições etárias distintas.	28
4.11	Intervalos de confiança para a SCR e a SP nas duas populações com distribuições etárias distintas.	28
4.12	Relação entre a SCR e a SP para duas distribuições etárias distintas.	29
4.13	Distribuição etária para duas populações distintas.	29
4.14	Tabela comparativa dos métodos de construção de intervalos de confiança, para uma amplitude relativa de 1.	30
4.15	Comparação gráfica dos intervalos de confiança calculados.	30
4.16	Menu de informações.	31
5.1	Derivada da função de seroprevalência.	34



# Lista de Tabelas

5.1	Resultados do cálculo da dimensão considerando $SRR = 0.017$ . . . . .	36
5.2	Resultados do cálculo da probabilidade de cobertura para diferentes métodos de construção de intervalos de confiança, considerando uma $SRR$ de 0.017, um nível de confiança de 0.95 e uma amplitude relativa de 1. . . . .	39



# Capítulo 1

## Introdução

O cálculo da dimensão amostral é uma etapa fundamental no planeamento de estudos científicos, nomeadamente em investigações epidemiológicas. Em primeiro lugar, este cálculo permite garantir que os parâmetros epidemiológicos sejam estimados com uma precisão adequada. Assim como, contribui para a otimização de recursos financeiros e humanos ao determinar a dimensão mínima da amostra necessária para atingir os objetivos do estudo. Desta forma, evita-se a subamostragem, que pode comprometer a representatividade da amostra, embora a aleatorização desempenhe um papel importante nesta questão. Da mesma forma, evita-se a recolha excessiva de dados, poupando recursos.

Neste trabalho, vamos explorar a questão do cálculo da dimensão amostral com foco em estudos transversais, aplicados no contexto da malária. Os estudos transversais são um tipo de estudos observacionais frequentemente utilizados em epidemiologia e seroepidemiologia. Nestes estudos a amostragem é realizada num único momento temporal, permitindo calcular a prevalência de determinadas características ou condições na população-alvo.

A malária é uma doença infecciosa transmitida aos humanos através da picada de mosquitos *Anopheles* infetados por um parasita unicelular protozoário do género *Plasmodium*. Existem cerca de 156 espécies de *Plasmodium*, sendo que destas já foram reportadas infeções em humanos por cinco espécies diferentes, com o *Plasmodium falciparum* predominante na zona da África Subsariana e o *Plasmodium vivax* na zona da América do Sul e do Sudeste Asiático.

As fêmeas do mosquito *Anopheles* necessitam alimentar-se de sangue para obter os nutrientes necessários à produção de ovos. Quando se alimentam, secretam saliva para a corrente sanguínea, impedindo a coagulação do sangue. Se o mosquito estiver infetado, é neste momento que o parasita tem a chance de entrar na corrente sanguínea, na forma de esporozoítos, migrando até ao fígado. No fígado, os esporozoítos multiplicam-se assexuadamente. Esta multiplicação resulta em milhares de merozoítos, que são então libertados de volta para a corrente sanguínea, onde infetam os glóbulos vermelhos. Dentro dos glóbulos vermelhos, cada merozoíto multiplica-se, e é libertado na corrente sanguínea, onde alguns diferenciam-se em gametócitos. São estes gametócitos que, quando ingeridos por um mosquito durante a picada, infetam-no, sendo responsáveis pela propagação da malária [Sato, 2021]. Este ciclo é apresentado na Fig 1.1.

Os sintomas mais frequentes são febre, calafrios, dores de cabeça, fraqueza, vômitos e diarreia. Podendo evoluir para situações mais graves, como esplenomegalia, anemia, trombocitopenia, hipoglicemia, disfunção pulmonar ou renal e alterações neurológicas. Os sintomas variam consoante o protozoário responsável pela infeção, sendo que as infeções causadas pelo *Plasmodium falciparum* são as mais propensas a evoluir para formas mais graves e letais [Centers for Disease Control and Prevention, 2024].



diferenças nas condições favoráveis à transmissão.

Como forma de contornar as limitações das medidas acima apresentadas, surgiram medidas baseadas na concentração de anticorpos em análises serológicas.

Quando indivíduos são expostos a agentes infecciosos, como vírus ou parasitas, as células B do sistema imunitário produzem uma resposta biológica na forma de anticorpos. Numa fase inicial, esses anticorpos não são detetáveis, e o indivíduo é considerado seronegativo. Quando os anticorpos passam a ser detetáveis na corrente sanguínea através de uma análise serológica, o indivíduo é considerado seropositivo.

A seroprevalência (SP) representa a proporção de indivíduos de uma população que têm anticorpos detetáveis na sua amostra serológica, sugerindo exposição à doença ou agente infeccioso, ou seja, representa a proporção de seropositivos numa população. É útil para verificar a exposição de uma população a um agente infeccioso e pode fornecer informações importantes sobre a imunidade de grupo.

O processo segundo o qual os indivíduos de uma população passam de seronegativos a seropositivos é denominado seroconversão, e a taxa associada a este processo é designada de taxa de seroconversão (*seroconversion rate*, SCR). O processo inverso, segundo o qual os indivíduos seropositivos passam a seronegativos, chama-se serorreversão, e a respetiva taxa é designada por taxa de serorreversão (*seroreversion rate*, SRR) [Sepúlveda et al., 2015b].

Em cenários de baixa intensidade de transmissão, a SP é utilizada como alternativa à prevalência de infeção, pois os anticorpos funcionam como biomarcadores de exposição ao agente infeccioso, permitindo identificar indivíduos que foram expostos ao parasita, mesmo que não estejam atualmente infetados. E ao refletir exposições passadas, não é tão influenciada por variações sazonais.

A principal desvantagem da seroepidemiologia é a dificuldade em lidar com a variabilidade entre diferentes testes de diagnóstico. Questões como a sensibilidade e especificidade dos testes serológicos podem levar a erros nas estimativas de exposição e imunidade, assim como, saber quais os antigénios clinicamente relevantes [Sepúlveda et al., 2015b].

## 1.1 Objetivo

O principal objetivo deste trabalho é o desenvolvimento de uma ferramenta computacional que permita capacitar profissionais sem formação em estatística ou programação a aplicarem o método apresentado no artigo [Sepúlveda and Drakeley, 2015], permitindo-lhes assim calcular a dimensão amostral necessária para os seus estudos de forma autónoma.

Como objetivo secundário, procurou-se retirar conclusões sobre o cálculo da dimensão amostral e avaliar o desempenho de diferentes métodos de construção de intervalos de confiança para proporções.

## 1.2 Estrutura do trabalho

O Capítulo 2 tem como tema o cálculo da dimensão amostral para medidas epidemiológicas baseadas em proporções, onde são introduzidos diversos métodos de construção de intervalos de confiança. Embora não se trate de uma lista exaustiva de todos os métodos existentes, são apresentados os mais comuns, e o respetivo cálculo da dimensão amostral necessária.

O Capítulo 3 foca-se no cálculo da dimensão amostral para a SCR, medida esta que está relacionada com a SP através de modelos catalíticos reversíveis.

No Capítulo 4 é apresentada a ferramenta desenvolvida no âmbito deste trabalho, acompanhada do respetivo tutorial.

## 1. INTRODUÇÃO

O Capítulo 5 apresenta os resultados obtidos, divididos em duas categorias. A primeira reproduz os resultados publicados no artigo [Sepúlveda and Drakeley, 2015], com o intuito de avaliar o desempenho da ferramenta e o cálculo da dimensão amostral. A segunda compara o desempenho dos diferentes métodos de construção de intervalos de confiança, utilizando a probabilidade de cobertura e a dimensão amostral como critérios de comparação.

No último capítulo, são discutidos os resultados obtidos e resumidas as principais conclusões do estudo. São também apresentadas sugestões para o trabalho futuro, destacando possíveis melhorias à ferramenta.

## Capítulo 2

# Cálculo da dimensão amostral para medidas epidemiológicas baseadas em proporções

Num estudo seroepidemiológico, uma das questões essenciais é determinar a dimensão amostral adequada para conseguir extrair conclusões confiáveis desse mesmo estudo. Este cálculo previne tanto a recolha insuficiente, quanto a recolha excessiva de dados. De notar que existem diversos fatores relacionados com outras componentes da amostra, para além da dimensão amostral, que podem influenciar a qualidade dos dados e, conseqüentemente, do estudo. A aleatorização é um desses fatores importantes, pois mesmo uma amostra de grande dimensão pode não ser representativa da população-alvo se a seleção dos indivíduos não for feita de forma aleatória. Outros fatores que também podem afetar a qualidade do estudo incluem dificuldades de acesso à população-alvo, que podem tornar mais fácil ou difícil recrutar determinados subgrupos da população. Assim como, dificuldades logísticas e financeiras, e limitações nos métodos de diagnóstico. Contudo, neste trabalho, estes fatores não são considerados diretamente no cálculo da dimensão amostral, assumindo-se que todos os outros aspetos do planeamento do estudo foram devidamente controlados.

Os intervalos de confiança fornecem-nos uma gama de valores plausíveis para o parâmetro de interesse. Os intervalos contêm todos os valores hipotéticos de um parâmetro que não podem ser rejeitados com base nos dados observados e num nível de confiança especificado [Smithson, 2003].

De notar que não existe apenas uma forma de construir intervalos de confiança e diferentes abordagens devem ser utilizadas consoante as características de cada estudo [Pires and Amado, 2008] [Gonçalves et al., 2012].

A escolha do(s) método(s) adequado(s) depende das características do estudo. Para avaliar os métodos, utilizam-se dois conceitos fundamentais, o conceito de probabilidade de cobertura (*coverage probability*, CP) e o de amplitude esperada (*expected length*, EL).

Como descrito no artigo [Gonçalves et al., 2012], com base numa amostra de dimensão  $n$ , a CP quantifica a frequência com que o intervalo de confiança inclui o verdadeiro valor do parâmetro  $p$  em amostragens repetidas, e é dada por:

$$CP(n, p) = \sum_{j=0}^n \binom{n}{j} p^j (1-p)^{n-j} I_{[L(j); U(j)]}(p).$$

Aqui,  $I_{[L(j); U(j)]}(p)$  é uma função indicatriz que assume o valor 1 se o verdadeiro valor do parâmetro

## 2. CÁLCULO DA DIMENSÃO AMOSTRAL PARA MEDIDAS EPIDEMIOLÓGICAS BASEADAS EM PROPORÇÕES

estiver dentro do intervalo  $[L(j), U(j)]$ , onde  $L(j)$  e  $U(j)$  representam, respetivamente, o limite inferior e superior do intervalo de confiança e toma o valor 0, caso contrário.

Teoricamente, o nível de confiança  $1 - \alpha$  corresponde à probabilidade de cobertura esperada do intervalo. No entanto, na prática, a probabilidade de cobertura observada pode ser inferior ao nível de confiança nominal. Isto ocorre porque, na construção de um intervalo de confiança para proporções, utilizam-se geralmente duas aproximações: uma baseada no Teorema Limite Central e outra para estimar a variância (e o erro padrão). O que faz com que a diferença entre o nível de confiança e a cobertura real possa aumentar conforme mais aproximações forem utilizadas.

Em termos práticos, a probabilidade de cobertura é estimada como o número de vezes que o intervalo de confiança contém o verdadeiro valor do parâmetro dividido pelo número total de amostragens realizadas.

A amplitude de um intervalo de confiança pode ser definida como a diferença entre o limite superior e o limite inferior do intervalo. Se o intervalo for simétrico em relação à estimativa do parâmetro, i.e, possa ser escrito da forma estimativa do parâmetro  $\pm$  margem de erro, podemos também considerar amplitude =  $2 \times$  margem de erro.

A amplitude esperada refere-se ao valor médio da amplitude dos intervalos de confiança quando se realizam múltiplas amostragens. Em outras palavras, é uma estimativa de quão largo, em média, será o intervalo de confiança, considerando a dimensão amostral e a proporção observada, e é definida como a média ponderada das amplitudes dos intervalos para diferentes valores de sucessos  $j$  numa amostra de dimensão  $n$ .

$$EL(n, p) = \sum_{j=0}^n \binom{n}{j} p^j (1-p)^{n-j} [U(j) - L(j)]$$

Se dois métodos tiverem coberturas similares é dada preferência àquele que tiver menor amplitude esperada.

Estes são os critérios utilizados nesta dissertação, assim como no artigo [Gonçalves et al., 2012], para avaliar o desempenho dos métodos. No entanto, na literatura existem outros critérios como *p-confidence* e *p-bias* apresentados em [Vos and Hudson, 2005]. A *p-confidence* indica o grau de confiança de que apenas os valores do parâmetro dentro do intervalo de confiança não são rejeitados, e garante que os valores do parâmetro fora do intervalo de confiança não têm valores- $p$  associados maiores que  $\alpha$ , sendo assim rejeitados. O *p-bias* mede a probabilidade de o intervalo de confiança incluir, indevidamente, valores do parâmetro para os quais os valores- $p$  indicariam rejeição. Quanto maior for a *p-confidence* e menor o *p-bias*, melhor o método. Contudo, optou-se por não utilizar estes critérios por serem mais complexos e menos intuitivos.

### 2.1 Algoritmo geral

Queremos encontrar a dimensão amostral ( $n$ ) tal que a amplitude esperada do intervalo de confiança para a nossa medida tenha uma determinada amplitude ( $a$ ).

$$EL(n, p) = a$$

No entanto, não é garantido que esta igualdade possa ser satisfeita, uma vez que a função da amplitude esperada do intervalo de confiança, não possui, em geral, uma fórmula fechada. Assim como, dificilmente a equação  $EL(n, p) = a$  possui uma solução inteira.

## 2.2 Métodos de construção de intervalos de confiança para proporções

Para contornar esta limitação, procuramos um valor de  $n$  tal que  $EL(n, p)$  seja aproximadamente  $a$ , definindo um valor de tolerância ( $\xi$ ). De seguida, encontramos todas as soluções que satisfaçam

$$|EL(n, p) - a| \leq \xi.$$

A partir deste conjunto de soluções, podemos seleccionar o valor de  $n$  com base em diferentes critérios. Podemos optar pela menor dimensão amostral que satisfaça a condição, o que pode ser ideal quando temos dificuldades de recrutamento. Alternativamente, podemos escolher o valor  $n$  que maximize a probabilidade de cobertura, ou seleccionar o menor  $n$  que assegure que o intervalo atinja o nível de confiança nominal desejado.

Nos métodos mais simples (simétricos), a amplitude esperada do intervalo de confiança pode ser expressa como o dobro da margem de erro, ou seja,  $EL(n, p) = 2 \times z_{1-\frac{\alpha}{2}} \sqrt{\text{Var}[\hat{p}]}$ .

Ao lidar com métodos mais complexos, especialmente aqueles que não geram intervalos por natureza simétricos, encontrar uma solução explícita é impossível. Diante desta dificuldade, uma abordagem prática e frequentemente adotada é utilizar a diferença entre o limite superior e o limite inferior do intervalo como uma aproximação da amplitude esperada.

Como

$$E[U(X) - L(X)] \approx U(E[X]) - L(E[X]),$$

e

$$U(E[X]) - L(E[X]) = U(np) - L(np),$$

pois  $X \sim \text{Binomial}(n, p)$ .

Queremos encontrar o  $n$  tal que:

$$|U(np) - L(np) - a| \leq \xi.$$

No entanto encontrar todos os valores de  $n$  que satisfaçam  $|U(np) - L(np) - a| \leq \xi$ , pode ser computacionalmente exigente. Por essa razão, diversos investigadores, como [Piegorisch, 2004] e [Krishnamoorthy and Peng, 2007] optam por uma abordagem mais prática: seleccionar o primeiro valor inteiro que satisfaça  $U(np) - L(np) \leq a$ .

## 2.2 Métodos de construção de intervalos de confiança para proporções

Existem diferentes métodos de construção de intervalos de confiança, podendo os métodos dividir-se em duas principais categorias (exatos e assintóticos).

Os métodos assintóticos assumem que, à medida que a dimensão amostral aumenta, a distribuição da estatística de teste tende a aproximar-se duma distribuição Normal. Por sua vez, os métodos exatos baseiam-se em distribuições exatas. Uma característica importante dos métodos exatos é que a probabilidade de cobertura do intervalo de confiança é no mínimo  $1 - \alpha$ , para todos os valores do parâmetro, enquanto nos métodos assintóticos a cobertura pode ser menor que  $1 - \alpha$ . Embora os problemas de cobertura nos métodos assintóticos sejam mais comuns em dimensões amostrais menores, também podem ocorrer em amostras grandes, especialmente para certas combinações de dimensão amostral ( $n$ ) e proporção ( $p$ ), conhecidas como *unlucky pairs* [Brown et al., 2001]. Em tais casos, a probabilidade de cobertura do intervalo pode ser significativamente diferente do nível de confiança nominal.

Os intervalos de confiança apresentados em seguida são intervalos bilaterais.

## 2. CÁLCULO DA DIMENSÃO AMOSTRAL PARA MEDIDAS EPIDEMIOLÓGICAS BASEADAS EM PROPORÇÕES

### 2.2.1 Métodos exatos

No contexto da construção de intervalos de confiança para proporções, os métodos exatos são valorizados pela sua capacidade de garantir uma probabilidade de cobertura mínima de  $1 - \alpha$ , independentemente da dimensão amostral ou da distribuição subjacente aos dados. Entre estes métodos, o método de Clopper-Pearson é frequentemente considerado o “clássico” exato. Este método calcula os intervalos de confiança com base na distribuição binomial exata. Uma das principais críticas ao método de Clopper-Pearson é que este pode resultar em intervalos de confiança mais amplos do que o estritamente necessário.

Para mitigar essa limitação, foram desenvolvidas variantes do método anterior. Entre essas variantes está o método Mid-p, que permite diminuir a amplitude dos intervalos sem sacrificar a cobertura [Berry and Armitage, 1995].

Também existem outros métodos que podemos caracterizar como exatos, apesar de não terem por base uma distribuição exata. Um exemplo é o método de Blaker, que tem por base a utilização de *acceptability functions*, e tem como objetivo calcular intervalos de confiança mais precisos (com menor amplitude) que o de Clopper-Pearson, mantendo uma probabilidade de cobertura mínima de  $1 - \alpha$  [Blaker, 2000]. No entanto, neste trabalho exploramos apenas o método de Clopper-Pearson.

#### 2.2.1.1 Método de Clopper-Pearson

A base do método de Clopper-Pearson é a “inversão” do teste binomial. Ou seja, encontrar os valores possíveis para o parâmetro  $p$ , que não são rejeitados pelo teste ao nível de significância  $\alpha$  [Clopper and Pearson, 1934].

Os limites do intervalo de confiança ( $p_l$  e  $p_u$ ) satisfazem as seguintes equações:

$$\sum_{x=n_1}^n \binom{n}{x} p_l^x (1-p_l)^{n-x} = \frac{\alpha}{2}$$
$$\sum_{x=0}^{n_1} \binom{n}{x} p_u^x (1-p_u)^{n-x} = \frac{\alpha}{2}.$$

Estas equações também podem ser utilizadas para o cálculo da dimensão amostral, neste caso fixamos os limites do intervalo ( $p_l; p_u$ ) e resolvemos as equações em ordem a  $n$ . O que resulta em cálculos da dimensão separados para o limite inferior e superior. O que se considera a dimensão amostral necessária é o maior valor dos calculados anteriormente. Como resultado, a dimensão amostral pode ser maior do que o estritamente necessário para alcançar o nível de confiança nominal desejado.

Na distribuição Binomial, a probabilidade de sucesso  $p$  é um parâmetro fixo, a partir do qual inferimos a probabilidade de  $x$  sucessos. Por outro lado, a distribuição Beta é usada para inferir sobre a probabilidade dum evento quando temos alguma informação sobre a quantidade de sucessos e insucessos, por isso, é muito usada como distribuição *a priori* no contexto da estatística bayesiana.

Devido à relação entre a distribuição binomial e a distribuição Beta, o intervalo de confiança de Clopper-Pearson pode ser representado num formato que utiliza os quantis da distribuição Beta [Thulin, 2014]:

$$(p_l, p_u) = \left( \text{Beta} \left( \frac{\alpha}{2}; x, n-x+1 \right), \text{Beta} \left( 1 - \frac{\alpha}{2}; x+1, n-x \right) \right).$$

Estes limites têm de ser calculados numericamente, com exceção dos casos  $x = 0$  ou  $x = n$ .

## 2.2 Métodos de construção de intervalos de confiança para proporções

Para  $x = 0$ :

$$IC_{(1-\alpha)100\%} = \left[ 0, 1 - \left( \frac{\alpha}{2} \right)^{\frac{1}{n}} \right].$$

Para  $x = n$ :

$$IC_{(1-\alpha)100\%} = \left[ \left( \frac{\alpha}{2} \right)^{\frac{1}{n}}, 1 \right].$$

Para calcular a dimensão amostral, fixa-se a amplitude desejada e o nível de confiança, e de forma iterativa ajusta-se o valor de  $n$  até obtermos uma solução que satisfaça as condições.

### 2.2.2 Métodos assintóticos

A teoria assintótica permite determinar o comportamento de estimadores, como a proporção amostral, em populações teoricamente infinitas.

Os métodos assintóticos baseiam-se nesta teoria para extrair conclusões sobre amostras finitas, mas de dimensão grande. Eles partem da premissa de que, para amostras suficientemente grandes, a distribuição de certas estatísticas de interesse aproximam-se de uma distribuição Normal, mesmo que a distribuição original das variáveis não o seja. Por isso, estes métodos não são recomendados para amostras de dimensão pequena. Por outro lado, são especialmente úteis quando temos grandes amostras, onde os métodos exatos, por não possuírem uma forma fechada, necessitam do uso de métodos numéricos, os quais podem ser computacionalmente exigentes. Neste caso, a aproximação assintótica oferece uma solução prática. Embora uma regra prática comum seja considerar uma dimensão amostral superior a 30 como suficientemente grande para que as aproximações assintóticas sejam válidas, é importante realçar que isso pode variar dependendo da distribuição dos dados, da precisão desejada e da complexidade do problema em questão.

Para fundamentar este raciocínio, consideramos  $X$  variável aleatória que representa o número de sucessos em  $n$  tentativas.

$$X \sim Binomial(n, p)$$

$$E[X] = np \text{ e } Var[X] = np(1-p).$$

Sendo uma Binomial a soma de  $n$  *Bernoullis* independentes e identicamente distribuídas, para grandes valores de  $n$ , pelo Teorema Limite Central, a variável aleatória  $X$  segue aproximadamente uma distribuição Normal.

E o estimador de máxima verosimilhança da proporção de sucessos ( $\hat{p}$ ), é dado por  $\hat{p} = \frac{X}{n}$  e também segue uma distribuição Normal aproximada.

$$E \left[ \frac{X}{n} \right] = \frac{1}{n} E[X] = \frac{1}{n} np = p$$
$$Var \left[ \frac{X}{n} \right] = \frac{1}{n^2} Var[X] = \frac{1}{n^2} np(1-p) = \frac{p(1-p)}{n}.$$

Donde surge que

$$\frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0, 1). \quad (2.1)$$

## 2. CÁLCULO DA DIMENSÃO AMOSTRAL PARA MEDIDAS EPIDEMIOLÓGICAS BASEADAS EM PROPORÇÕES

A partir da equação 2.1 surge que os limites do intervalo de confiança  $(1 - \alpha)100\%$  são dados por

$$\hat{p} \pm z_{1-\frac{\alpha}{2}} \sqrt{\text{Var}[\hat{p}]}.$$

Equivalentemente,

$$\hat{p} \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}. \quad (2.2)$$

É a partir daqui que os métodos assintóticos começam a diferir entre si.

### 2.2.2.1 Método de Wald

O método de Wald é o mais simples e intuitivo, e portanto, aquele que é apresentado em diversos livros de introdução à estatística. Neste método, o parâmetro populacional ( $p$ ) é substituído na equação 2.2 pela sua estimativa ( $\hat{p}$ ).

O intervalo de confiança é dado por

$$\hat{p} \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}.$$

De forma a determinar a dimensão amostral necessária, queremos garantir que a amplitude esperada do intervalo de confiança seja igual a um valor pré-definido ( $a$ ). Considerando que a amplitude esperada do intervalo de confiança pode ser expressa como  $EL(n, p) = 2 \times z_{1-\frac{\alpha}{2}} \sqrt{\text{Var}[\hat{p}]}$ , e da equação 2.2, surge

$$EL_{Wald}(n, p) = 2 \times z_{1-\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}.$$

Queremos encontrar a dimensão  $n$ , que satisfaça a seguinte condição:

$$EL_{Wald}(n, p) = a.$$

Equivalentemente

$$2 \times z_{1-\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}} = a.$$

Resolvendo a equação anterior em ordem a  $n$ , que designaremos  $n_{Wald}$ , obtém-se a seguinte fórmula para o cálculo da dimensão amostral:

$$n_{Wald} = \frac{p(1-p)}{\left(\frac{a}{2}\right)^2} (z_{1-\frac{\alpha}{2}})^2. \quad (2.3)$$

Na prática, para aplicar a fórmula 2.3, é necessário dispor de uma estimativa preliminar de  $p$  ( $p_0$ ). Esta estimativa pode ser obtida através de um estudo anterior ou de um estudo piloto. Caso não se disponha de nenhum valor prévio, utiliza-se  $p_0 = 0.5$ , escolha conservadora, que maximiza a dimensão amostral necessária [Thulin, 2014].

No entanto, é um método que apresenta diversos problemas, nomeadamente, quando a proporção estimada se encontra perto de 0 ou 1 (estamos perante poucos ou muitos sucessos na amostra, respetivamente). Para proporções baixas, o limite inferior do intervalo de confiança pode estar abaixo de 0. Equivalentemente, para proporções próximas de 1, o limite superior do intervalo, pode estar acima de 1. A isto chama-se *overshooting*. Nestes casos, é comum truncar o intervalo de confiança para garantir

## 2.2 Métodos de construção de intervalos de confiança para proporções

que ele permaneça dentro do intervalo  $[0, 1]$ . Quando a proporção é exatamente 0 ou 1, obtém-se um intervalo degenerado, i.e, um intervalo com apenas um ponto (0 ou 1, respetivamente).

De forma a melhorar a aproximação da distribuição binomial pela distribuição normal, surge a ideia de adicionar uma correção de continuidade ao método, e apesar dessa correção melhorar a probabilidade de cobertura do intervalo de confiança, ainda assim pode não ser aconselhado, pois aumenta a ocorrência de *overshooting* [Newcombe, 1998].

Existem diversas correções, a mais usual é a correção de continuidade de Yates, que adiciona  $\frac{1}{2n}$  ao limite superior e subtrai o mesmo valor ao limite inferior. Donde se obtém os seguintes limites para o intervalo de confiança:

$$\begin{aligned} \text{limite inferior: } & \hat{p} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n} - \frac{1}{2n}} \\ \text{limite superior: } & \hat{p} + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{1}{2n}}, \end{aligned}$$

sendo a amplitude esperada  $EL_{Waldcc}(n, p) = 2 \times z_{1-\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n} + \frac{1}{n}}$ .

Considerando  $a$  a amplitude desejada, obtém-se que a dimensão amostral, que designaremos  $n_{Waldcc}$  é dada pela solução positiva da seguinte equação:

$$n_{Waldcc} = \frac{2a + 4z_{1-\frac{\alpha}{2}}^2 p(1-p) \pm \sqrt{(2a + 4z_{1-\frac{\alpha}{2}}^2 p(1-p))^2 - 4a^2}}{2a^2}. \quad (2.4)$$

A derivação deste resultado pode ser consultada no apêndice A. Assim como no exemplo anterior, para aplicar a fórmula 2.4, é necessário dispor de uma estimativa preliminar de  $p$  ( $p_0$ ).

### 2.2.2.2 Método de Wilson

Ao contrário do método de Wald, que utiliza  $\hat{p}$  na estimativa do erro padrão, o método de Wilson adota uma abordagem diferente ao resolver analiticamente a fórmula em função de  $p$ .

Da equação 2.1 e dos respetivos limites dados por 2.2, surge

$$(\hat{p} - p)^2 = z_{1-\frac{\alpha}{2}}^2 \frac{p(1-p)}{n} \Leftrightarrow \hat{p}^2 - 2p\hat{p} + p^2 = \frac{z_{1-\frac{\alpha}{2}}^2}{n} p - \frac{z_{1-\frac{\alpha}{2}}^2}{n} p^2.$$

Donde se obtém que, segundo o método de Wilson, o intervalo de confiança  $(1-\alpha)100\%$  para  $p$  é dado por

$$\frac{\hat{p} + \frac{z_{1-\frac{\alpha}{2}}^2}{2n} \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z_{1-\frac{\alpha}{2}}^2}{4n^2}}}{1 + \frac{z_{1-\frac{\alpha}{2}}^2}{n}}. \quad (2.5)$$

A derivação deste resultado pode ser consultada no apêndice B.

De notar que

$$\lim_{n \rightarrow \infty} \frac{z_{1-\frac{\alpha}{2}}^2}{2n} \rightarrow 0; \quad \lim_{n \rightarrow \infty} \frac{z_{1-\frac{\alpha}{2}}^2}{4n^2} \rightarrow 0; \quad \lim_{n \rightarrow \infty} \left( 1 + \frac{z_{1-\frac{\alpha}{2}}^2}{n} \right) \rightarrow 1,$$

e portanto quando  $n \rightarrow \infty$

$$\hat{p} \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}},$$

## 2. CÁLCULO DA DIMENSÃO AMOSTRAL PARA MEDIDAS EPIDEMIOLÓGICAS BASEADAS EM PROPORÇÕES

que é a fórmula dum intervalo de confiança de Wald. Ou seja, quanto maior a dimensão amostral, maior vai ser a semelhança entre os resultados obtidos através de ambos os métodos.

Se considerarmos  $\tilde{n} = n + z^2$  e  $\tilde{p} = \frac{n\hat{p} + 0.5z^2}{n + z^2}$ , como sugerido em [Piegorisch, 2004], a fórmula do intervalo de confiança simplifica para:

$$\tilde{p} \pm \frac{z}{\tilde{n}} \sqrt{np(1-p) + \frac{z^2}{4}}.$$

Considerando  $\varepsilon = \frac{a}{2}$ , com  $a$  amplitude desejada, obtém-se a seguinte equação quadrática:

$$\varepsilon^2 n^2 + z^2 (2\varepsilon^2 - p(1-p)) n - z^4 \left( \varepsilon^2 - \frac{1}{4} \right) = 0,$$

que resolvendo em ordem a  $n$ , prova-se, pela regra dos sinais de *Descartes*, que a única solução positiva é dada por:

$$n = z^2 \frac{p(1-p) - 2\varepsilon^2 + \sqrt{p^2(1-p)^2 + 4\varepsilon^2 \left( p - \frac{1}{2} \right)^2}}{2\varepsilon^2}. \quad (2.6)$$

Para aplicar a fórmula 2.6 é necessário dispor de uma estimativa preliminar de  $p$  ( $p_0$ ).

### 2.2.2.3 Método de Agresti-Coull

Na construção de intervalos de confiança para uma proporção  $p$ , o método de Wald é conhecido por apresentar problemas de cobertura. Por sua vez, o método de Wilson apresenta uma fórmula complexa, isto não é um problema do ponto de vista computacional, mas pode ser do ponto de vista do ensino.

O método de Agresti-Coull surge como um compromisso entre estes dois métodos, utilizando uma aproximação do ponto médio do intervalo de Wilson, numa fórmula semelhante à do intervalo de Wald.

Através da equação 2.5, verifica-se que o ponto médio do intervalo de Wilson é

$$\frac{\hat{p} + \frac{z^2}{2n}}{1 + \frac{z^2}{n}} = \frac{\frac{x}{n} + \frac{z^2}{2n}}{\frac{n+z^2}{n}} = \frac{x + \frac{z^2}{2}}{n + z^2}.$$

Agora utiliza-se na fórmula do método de Wald,  $\tilde{p} = \frac{x + \frac{z^2}{2}}{\tilde{n}}$ , considerando  $\tilde{n} = n + z^2$ .

Donde surge que o intervalo de confiança é dado por

$$\tilde{p} \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}}.$$

Note-se que quando  $\alpha = 0.05$ , temos  $z^2 \approx 4$ , e portanto

$$\tilde{p} = \frac{x + 2}{n + 4}.$$

Isto é equivalente a adicionar 4 pseudo observações, duas aos sucessos e duas aos insucessos, o que ajuda com a problemática de valores de  $x$ , realização da variável aleatória  $X$ , perto de 0 ou de  $n$ . Este simples ajustamento apresenta um desempenho melhor que o método simples (de Wald), até para amostras mais pequenas [Agresti and Caffo, 2000].

Numa perspetiva bayesiana este ajuste pode ser interpretado como resultado de uma distribuição *a priori*  $Beta(2, 2)$ . A distribuição *a posteriori*, considerando  $x$  o número de sucessos observados, é  $p|x \sim Beta(x + 2, n - x + 2)$ .

## 2.2 Métodos de construção de intervalos de confiança para proporções

$$E[p | x] = \frac{x + 2}{n + 4}$$

Embora o método Agresti-Coull possa ser interpretado sob uma perspectiva bayesiana, neste trabalho não iremos discutir este tipo de métodos.

Utilizando a formulação inicial, queremos encontrar a dimensão amostral. Seja  $a$  a amplitude máxima desejada

$$\begin{aligned} a &= 2 \times z_{1-\frac{\alpha}{2}} \sqrt{\text{Var}[\tilde{p}]} \Leftrightarrow \left(\frac{a}{2}\right)^2 = z_{1-\frac{\alpha}{2}}^2 \frac{\tilde{p}(1-\tilde{p})}{\tilde{n}} \Leftrightarrow \\ \Leftrightarrow \tilde{n} &= \frac{z_{1-\frac{\alpha}{2}}^2 \tilde{p}(1-\tilde{p})}{\left(\frac{a}{2}\right)^2} \Leftrightarrow n + z_{1-\frac{\alpha}{2}}^2 = \frac{z_{1-\frac{\alpha}{2}}^2 \tilde{p}(1-\tilde{p})}{\left(\frac{a}{2}\right)^2} \Leftrightarrow \\ n &= \frac{z_{1-\frac{\alpha}{2}}^2 \tilde{p}(1-\tilde{p})}{\left(\frac{a}{2}\right)^2} - z_{1-\frac{\alpha}{2}}^2. \end{aligned}$$

Designando a dimensão amostral por  $n_{AC}$ , de forma simplificada, obtém-se:

$$n_{AC} \approx n_{Wald} - z_{1-\frac{\alpha}{2}}^2.$$

De notar que quando  $\alpha = 0.05$ , o valor crítico  $z_{1-\frac{\alpha}{2}}^2$  é aproximadamente 4. Neste contexto, o método de Agresti-Coull, requer aproximadamente menos 4 participantes por grupo em comparação com o método de Wald. Este resultado é expectável, visto que tínhamos adicionado 4 pseudo observações.

### 2.2.2.4 Métodos baseados em transformações

Seja  $X \sim \text{Binomial}(n, p)$ , como nos casos anteriores. A variância de  $X$  depende do parâmetro  $p$ , pois  $\text{Var}[X] = np(1-p)$ . O que faz com que para amostras diferentes retiradas da mesma população não é só a estimativa do parâmetro que vai ser diferente, como também a estimativa do erro padrão. De forma a combater isto, surge a utilização de transformações estabilizadoras da variância, que como o nome indica, vão estabilizar (padronizar) o valor da variância e respetivo erro padrão. Existem diferentes transformações que podem ser utilizadas, entre elas a arcsine e a logit.

A transformação logit, usada principalmente no contexto de regressão logística, é mais intuitiva, pois baseia-se no logaritmo das *odds*, ao contrário da arcsine que não tem uma relação direta com um conceito probabilístico. No entanto, quando a proporção está próxima de 0 ou 1, a transformação logit torna-se problemática, pois  $\log(0)$  e  $\log(\infty)$  não estão definidos.

Uma desvantagem destes métodos é que os intervalos de confiança calculados com recurso a transformações podem ser mais difíceis de interpretar, dado que os limites precisam ser revertidos para a escala original.

No método arcsine, utiliza-se a transformação  $T(\hat{p}) = \arcsin(\sqrt{\hat{p}})$ . Pode-se facilmente demonstrar que a  $\text{Var}[\arcsin(\sqrt{\hat{p}})]$  não depende de  $p$ .

Seja  $X \sim \text{Binomial}(n, p)$  e  $\hat{p} = \frac{x}{n}$ .

$$T(x, n) = \arcsin\left(\sqrt{\frac{x}{n}}\right).$$

Seja  $U = \frac{X}{n}$ ,

## 2. CÁLCULO DA DIMENSÃO AMOSTRAL PARA MEDIDAS EPIDEMIOLÓGICAS BASEADAS EM PROPORÇÕES

$$E[U] = \frac{1}{n} E[X] = \frac{1}{n} np = p$$

$$Var[U] = \frac{1}{n^2} Var[X] = \frac{p(1-p)}{n}$$

$$U \sim N\left(p, \sqrt{\frac{p(1-p)}{n}}\right).$$

Seja  $g(U) = \arcsin \sqrt{U}$ , pelo método Delta,  $Var(g(U)) \approx (g'(p))^2 Var[U]$

$$g'(u) = \arcsin \sqrt{u} \frac{d}{du} = \frac{1}{\sqrt{1-\sqrt{u}^2}} \times \frac{1}{2\sqrt{u}} = \frac{1}{2\sqrt{u(1-u)}}.$$

Para  $u = p$ ,  $g'(p) = \frac{1}{2\sqrt{p(1-p)}}$ .

Então  $Var[\arcsin \sqrt{U}] \approx (g'(p))^2 Var[U] \approx \left(\frac{1}{2\sqrt{p(1-p)}}\right)^2 \times \frac{p(1-p)}{n} \approx \frac{1}{4n}$ , valor este que não depende do parâmetro  $p$ .

O intervalo de confiança para a transformada é dado por:

$$\arcsin\left(\sqrt{\frac{x}{n}}\right) \pm z_{1-\frac{\alpha}{2}} \times \frac{1}{2\sqrt{n}}.$$

Para voltarmos à escala original, utilizamos a inversa da transformação, ou seja, a função seno. Como,

$$T(\hat{p}) = \arcsin(\sqrt{\hat{p}}) \Leftrightarrow \sin(T(\hat{p})) = \sqrt{\hat{p}} \Leftrightarrow \hat{p} = \sin^2(T(\hat{p})).$$

O intervalo de confiança para  $p$  é então dado por:

$$\sin^2\left(\arcsin\left(\sqrt{\frac{x}{n}}\right) \pm z_{1-\frac{\alpha}{2}} \times \frac{1}{2\sqrt{n}}\right).$$

Baseado no anterior e no trabalho de Anscombe [Anscombe, 1948], foi proposto, utilizar-se a seguinte transformação,  $T(x, n) = \arcsin\left(\sqrt{\frac{x+\frac{3}{8}}{n+\frac{3}{4}}}\right)$ , neste caso,  $Var[T(x, n)] = \frac{1}{4(n+0.5)}$ . Segundo [Brown et al., 2001], esta modificação oferece melhor estabilização da variância. Esta alteração ao método de arcsine é frequentemente referida na literatura como método de Anscombe.

Neste caso o intervalo de confiança é dado por

$$\sin^2\left(\arcsin\left(\sqrt{\frac{x+\frac{3}{8}}{n+\frac{3}{4}}}\right) \pm z_{1-\frac{\alpha}{2}} \times \frac{1}{2\sqrt{n+0.5}}\right).$$

A função logit é definida como o logaritmo das *odds* dum determinado evento.

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right)$$

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = \ln\left(\frac{\frac{X}{n}}{1-\frac{X}{n}}\right) = \ln\left(\frac{X}{n-X}\right).$$

Pode-se demonstrar que  $Var\left[\ln\left(\frac{X}{n-X}\right)\right] = \frac{n}{X(n-X)}$  [Brown et al., 2001].

O intervalo de confiança  $(1-\alpha)100\%$  para o logaritmo das *odds* é dado por

$$\ln\left(\frac{X}{n-X}\right) \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{n}{X(n-X)}}. \quad (2.7)$$

E os limites do intervalo de confiança para o parâmetro  $p$  são dados por

$$\begin{aligned} \text{limite inferior} &= \frac{e^{ll}}{1 + e^{ll}} \\ \text{limite superior} &= \frac{e^{ul}}{1 + e^{ul}}, \end{aligned}$$

com  $ll$  o valor do limite inferior do intervalo para o logaritmo das *odds* (2.7), e  $ul$  o valor do limite superior.

Neste método não ocorre *overshooting*, mas tem a desvantagem de não poder ser calculado para casos extremos ( $x = 0$  ou  $x = n$ ); para estes casos é sugerido utilizar  $\hat{p} = \frac{x+0.5}{n+1}$  [Newcombe, 2001].

Assim, como tem uma amplitude excessiva, a sua amplitude esperada é maior que a do método de Clopper-Pearson [Brown et al., 2001].

### 2.2.3 Sugestões práticas para seleção do método

As sugestões práticas para a seleção do método descritas a seguir são baseadas no trabalho de Gonçalves et al. [Gonçalves et al., 2012]. De acordo com esse estudo, o critério de maior CP geralmente resulta em menores dimensões amostrais, enquanto garantir uma CP de pelo menos o nível de confiança nominal  $(1 - \alpha)$  exige amostras maiores.

Para valores de  $p < 0.03$ , o método de Agresti-Coull requer maior dimensão que os outros métodos. No entanto, a partir deste ponto, os métodos exatos como o de Clopper-Pearson passam a exigir as maiores dimensões amostrais. As discrepâncias entre o método de Wald e os restantes métodos é notável para valores  $p < 0.05$  e  $p > 0.95$ , sendo que para os restantes valores de  $p$  as diferenças não são tão evidentes.

Em termos de recomendação, para valores  $p < 0.01$ , são recomendados os métodos bayesianos (embora não discutidos neste trabalho), assim como o método de Anscombe que minimiza a dimensão amostral e o método de Agresti-Coull que, neste caso, garante o nível de confiança nominal. Para  $0.01 < p < 0.05$ , os métodos bayesianos e o método de Clopper-Pearson são recomendados. Para  $0.05 < p < 0.1$ , os métodos de Clopper-Pearson e Anscombe apresentam bons resultados. Para  $0.1 < p < 0.5$ , todos os métodos produzem resultados semelhantes. De notar que o método de Wald deve ser utilizado apenas para valores de  $p$  entre 0.2 e 0.8.

Não existe um método universalmente superior que seja o melhor em todos os cenários. A escolha do método ideal deve sempre considerar o contexto específico de cada estudo. Contudo, se o objetivo for escolher um método que seja adequado a uma variedade de situações, o método de Wilson é preferido por diversos investigadores, como Newcombe em [Newcombe, 1998]. Também, o método de Agresti-Coull, conhecido pela sua simplicidade e bom desempenho em grandes dimensões amostrais, é preferido por autores como Brown et al. em [Brown et al., 2001].

## 2.3 Medidas de desfecho principal

As duas principais medidas epidemiológicas baseadas em proporções são a prevalência de infeção e a SP. O uso de cada uma delas como medida de desfecho principal depende da intensidade de transmissão observada na população.

## **2. CÁLCULO DA DIMENSÃO AMOSTRAL PARA MEDIDAS EPIDEMIOLÓGICAS BASEADAS EM PROPORÇÕES**

### **2.3.1 Prevalência de infecção**

A maior parte dos estudos sobre a malária é projetada considerando a prevalência de infecção como medida de desfecho principal [Sepúlveda and Drakeley, 2015]. No entanto, esta abordagem apresenta limitações em cenários de baixa intensidade de transmissão, onde há poucos casos de infecção, ou em cenários com muitos indivíduos assintomáticos, e portanto, não diagnosticados. No entanto, é uma medida útil em cenários de moderada ou elevada intensidade de transmissão, pois nestes cenários é mais fácil identificar indivíduos infetados numa amostra aleatória, não sendo necessário recorrer à serologia.

Sendo esta medida uma proporção, pode-se aplicar qualquer um dos métodos apresentados na secção 2.2. A prevalência de infecção é, neste caso, o parâmetro de controlo, ou seja, fixamos a amplitude do intervalo de confiança para esta medida de forma a determinar a dimensão amostral necessária, conforme o algoritmo apresentado em 2.1.

### **2.3.2 Seroprevalência**

É em cenários de baixa intensidade de transmissão ou com muitos indivíduos assintomáticos, que a serologia têm a sua principal vantagem. Pois a presença de anticorpos, detetados por meio de análises serológicas, indica exposição prévia ao agente infeccioso, mesmo que essa exposição tenha ocorrido no passado.

A SP, por ser uma proporção, possui métodos bem estabelecidos para a construção de intervalos de confiança e para o cálculo da dimensão amostral necessária, nomeadamente, qualquer um dos métodos apresentados na secção 2.2. Neste caso, fixamos a amplitude do intervalo de confiança para a SP de forma a determinar a dimensão amostral necessária.

## Capítulo 3

# Cálculo da dimensão amostral para a taxa de seroconversão

Nesta secção, vamos abordar o cálculo da dimensão amostral para a SCR. Para tal, introduzimos os modelos catalíticos reversíveis (*reversible catalytic models*, RCM), que permitem estimar a SCR com base em dados de seroprevalência. Consideramos apenas situações de transmissão estável, ou seja, onde a SCR é constante ao longo do tempo. Não levaremos em conta cenários de heterogeneidade na transmissão e abordamos apenas o caso em que a SRR é conhecida e constante. Caso considerássemos a SRR desconhecida, a dimensão amostral necessária seria maior, uma vez que estaríamos a estimar um parâmetro adicional [Sepúlveda and Drakeley, 2015].

### 3.1 Modelos catalíticos reversíveis

Os RCM baseiam-se numa cadeia de Markov com dois estados (seropositivo e seronegativo), onde cada indivíduo pode transitar entre os estados segundo a SCR ( $\lambda$ ) e a SRR ( $\rho$ ), como podemos ver na Fig 3.1.

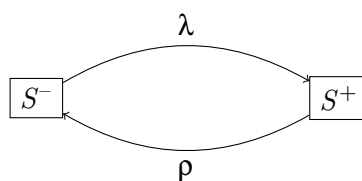


Figura 3.1: Representação da cadeia de Markov associado a um RCM onde  $S^+$  e  $S^-$  correspondem aos estados de seropositividade e seronegatividade, respetivamente. As transições entre os estados são representadas por  $\lambda$  (SCR) e  $\rho$  (SRR).

Seja  $P$  a proporção de indivíduos seropositivos numa amostra. A dinâmica da população é descrita pela seguinte equação diferencial:

$$\frac{dP}{dt} = \lambda(1-P) - \rho P.$$

O termo  $\lambda(1-P)$  descreve o processo de seroconversão, por sua vez o termo  $\rho P$  descreve o processo de serorreversão.

Obtém-se que a probabilidade de um indivíduo de idade  $t$  ser seropositivo ( $p_{S^+}$ ), partindo do estado inicial de seronegatividade, é dada por:

### 3. CÁLCULO DA DIMENSÃO AMOSTRAL PARA A TAXA DE SEROCONVERSÃO

$$p_{S+} = \frac{\lambda}{\lambda + \rho} \left( 1 - e^{-(\lambda + \rho)t} \right), \quad \lambda \in \mathbb{R}_0^+, \rho \in \mathbb{R}_0^+. \quad (3.1)$$

A derivação deste resultado pode ser consultada no apêndice C.

Os parâmetros dos RCM são geralmente estimados pelo método de máxima verosimilhança. A função de verosimilhança é definida com base no número de indivíduos seropositivos observados na amostra estratificada por idades. Nomeadamente, consideramos uma amostra de  $n$  indivíduos independentes entre si, com  $n = \sum_{t=1}^{A_{\max}} n_t$ , onde  $n_t$  é o número total de indivíduos de idade  $t$  na amostra e  $A_{\max}$  é a idade máxima considerada. Seja  $m_t$  o número de indivíduos seropositivos de idade  $t$  observados. A função de verosimilhança é dada por:

$$L(\lambda, \rho) = \prod_{t=1}^{A_{\max}} \binom{n_t}{m_t} (p_{S+})^{m_t} (1 - p_{S+})^{n_t - m_t}. \quad (3.2)$$

As estimativas de máxima verosimilhança são os valores dos parâmetros que maximizam a equação 3.2.

A SP refere-se à proporção de indivíduos numa população que apresenta anticorpos contra um determinado agente infeccioso num determinado momento temporal, e também pode ser expressa como um somatório, de 1 até a  $A_{\max}$ , das probabilidades de seropositividade ( $p_{S+}$ ) para cada faixa etária multiplicadas pela proporção ou peso ( $\alpha_t$ ) daquela faixa etária na população total.

$$SP = \sum_{t=1}^{A_{\max}} \alpha_t \times p_{S+}$$

Mais especificamente, e assumindo que SRR é conhecida e constante,  $\rho_t = \rho_0, \forall t \in [1, A_{\max}]$ , a SP pode ser representada pela seguinte equação:

$$SP = \sum_{t=1}^{A_{\max}} \left[ \alpha_t \times \frac{\lambda}{\lambda + \rho_0} \left( 1 - e^{-(\lambda + \rho_0)t} \right) \right], \quad (3.3)$$

onde  $\alpha_t$  representa a proporção de indivíduos com idade  $t$  na população geral,  $\lambda$  representa a SCR e  $\rho_0$  a SRR.

#### 3.2 Algoritmo para o cálculo da dimensão amostral para a taxa de seroconversão

A SCR é uma taxa, e não existem métodos diretos para o cálculo da dimensão amostral necessária. No entanto, a SP, por ser uma proporção, possui métodos bem estabelecidos para a construção de intervalos de confiança e para o cálculo da dimensão amostral necessária, como apresentado na secção 2.3.2. E sabemos que estas duas medidas estão relacionadas entre si através de RCM, nomeadamente, através da equação 3.3, e como ilustrado na Fig 3.2. É então possível calcular a SP a partir da SCR, assim como, utilizar o intervalo de confiança para a SP como base para construir o intervalo de confiança para SCR. A partir deste intervalo, consegue-se então determinar a dimensão amostral necessária para o estudo.

Começamos por calcular a SP a partir da SCR, utilizando a equação 3.3, e de seguida construímos um intervalo de confiança para a SP. Para construir o intervalo de confiança para a SP, utiliza-se um dos métodos apresentados na secção 2.2, embora outros métodos também possam ser utilizados.

### 3.2 Algoritmo para o cálculo da dimensão amostral para a taxa de seroconversão

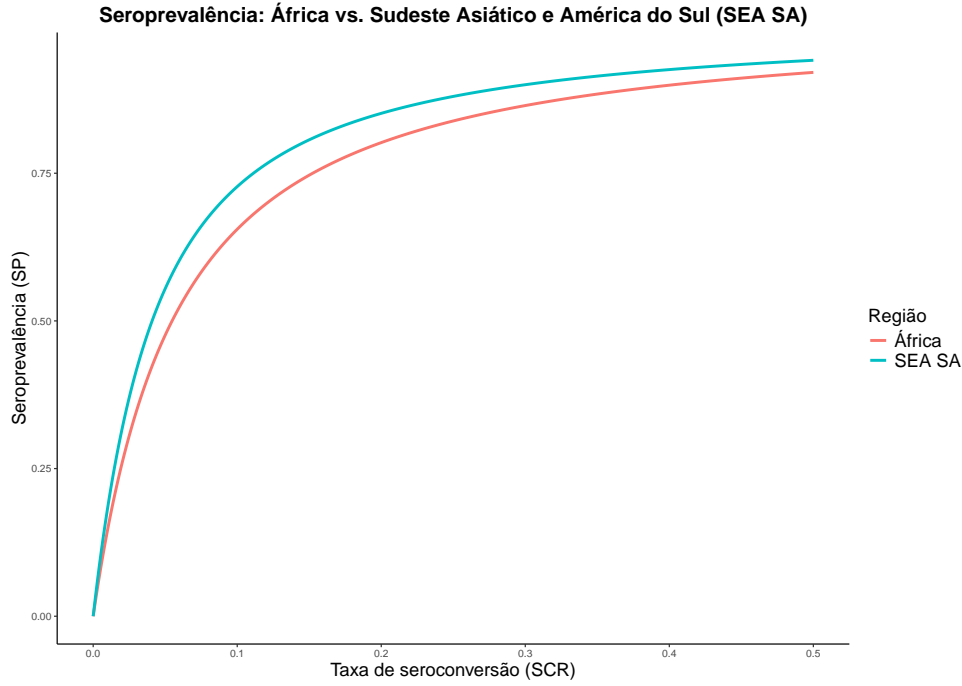


Figura 3.2: Relação entre a SCR e a SP, para duas populações distintas, considerando  $\rho = 0.017$ .

Uma vez obtido o intervalo de confiança para a SP, podemos utilizá-lo para derivar o intervalo de confiança para a SCR. Os limites inferior e superior do intervalo de confiança para a SCR, designados por  $\lambda_l$  e  $\lambda_u$ , respetivamente, são os valores que satisfazem:

$$SP_l = \sum_{t=1}^{A_{\max}} \left[ \alpha_t \times \frac{\lambda_l}{\lambda_l + \rho_0} \left( 1 - e^{-(\lambda_l + \rho_0)t} \right) \right], \quad (3.4)$$

$$SP_u = \sum_{t=1}^{A_{\max}} \left[ \alpha_t \times \frac{\lambda_u}{\lambda_u + \rho_0} \left( 1 - e^{-(\lambda_u + \rho_0)t} \right) \right], \quad (3.5)$$

onde  $SP_l$  representa o limite inferior do intervalo de confiança para a SP, e  $SP_u$  o limite superior. No entanto, não é possível encontrar uma solução analítica para os valores de  $\lambda$ , sendo necessário recorrer a métodos numéricos [Sepúlveda and Drakeley, 2015].

Define-se a amplitude relativa (*relative length*,  $RL$ ) como a razão entre a amplitude do intervalo de confiança e o valor do parâmetro que originou esse intervalo. Neste caso, como queremos controlar a amplitude do intervalo de confiança para a  $SCR(\lambda)$ , a amplitude relativa é dada por  $(\lambda_u - \lambda_l)/\lambda$ .

O objetivo é determinar a dimensão amostral ( $n$ ) para a qual a amplitude relativa do intervalo de confiança para a SCR seja menor ou igual à definida pelo utilizador. Para isso, a função *sample\_s* disponível no pacote *RCMsize* [Graça et al., 2025] vai iterando sobre sucessivos valores de  $n$  até encontrar um valor que satisfaça a condição, ou até ser atingido o número máximo de iterações (parâmetro definido pelo utilizador, que representa a dimensão máxima permitida). Este critério de paragem funciona porque, à medida que  $n$  aumenta, a amplitude absoluta do intervalo de confiança diminui, e, conseqüentemente, a amplitude relativa também diminui. Portanto, o primeiro valor de  $n$  que satisfizer a condição será o menor valor necessário, sendo assim, a solução ideal. Caso o número máximo de iterações seja atingido, a função retorna uma mensagem indicativa.

Embora existam outros critérios discutidos no subcapítulo 2.1, que poderiam ser aplicados neste contexto, este é o critério utilizado na ferramenta computacional apresentada no capítulo seguinte.



## Capítulo 4

# Ferramenta computacional

Como parte deste trabalho, desenvolveu-se uma ferramenta computacional e um pacote R, com o objetivo de tornar acessível o cálculo da dimensão amostral necessária para estudos seroepidemiológicos que utilizem modelos catalíticos reversíveis, modelos aplicados ao caso da malária, conforme proposto por [Sepúlveda and Drakeley, 2015]. A ferramenta, desenvolvida com recurso aos pacotes `shiny` (v1.9.1) [Chang et al., 2024] e `flexdashboard` (v0.6.2) [Aden-Buie et al., 2023] do *software* R (v4.3.1), apresenta uma interface interativa que dispensa conhecimentos prévios em programação, tornando o cálculo da dimensão amostral acessível a profissionais de saúde ou investigadores sem formação em estatística. Em paralelo, foi criado um pacote R, de nome `RCMsize`, disponível no CRAN [Graça et al., 2025], destinado a utilizadores com conhecimentos de R e que queiram realizar o cálculo da dimensão amostral diretamente no *software*.

Esta iniciativa justifica-se pelo facto de que, até ao momento, a aplicação prática dos métodos descritos em [Sepúlveda and Drakeley, 2015] não havia sido disponibilizada de forma pública. O que limitava o cálculo a investigadores com competências em programação e estatística. A ferramenta e o pacote desenvolvidos neste trabalho eliminam essa barreira, permitindo a qualquer utilizador calcular a dimensão amostral necessária para o seu estudo, ao mesmo tempo que oferece funcionalidades extra, como a comparação de diferentes distribuições etárias e da dimensão amostral obtida com recurso a diferentes métodos de construção de intervalos de confiança. A implementação faz também uso dos pacotes `binom` (v1.1.1.1) [Dorai-Raj, 2022], `ggplot2` (v3.5.1) [Wickham, 2016] e `plotly` (v4.10.4) [Sievert, 2020].

Algumas premissas são assumidas para o correto funcionamento da ferramenta. Em primeiro lugar, considera-se que a SRR é conhecida e constante. Isto porque a SRR não pode ser calculada a partir de dados de estudos transversais, pois a informação pontual sobre o estado de seronegatividade dum indivíduo não permite afirmar se esse mesmo indivíduo esteve sempre seronegativo, ou passou por um ou mais processos de serorreversão. Também é pressuposto que a SCR mantém-se constante ao longo do tempo, ou seja, não ocorrem mudanças na intensidade de transmissão durante o período de estudo. Assim como, pressupõe-se que estamos perante um cenário de diagnósticos perfeitos, i.e, sem erros de classificação serológica.

A interface está organizada em menus que orientam o utilizador a trabalhar com diferentes cenários. O primeiro cenário refere-se ao cálculo da dimensão amostral para um único valor da SCR. O segundo permite ao utilizador inserir um conjunto de valores da SCR e observar como estas variações têm impacto na SP e na dimensão amostral. O terceiro oferece a possibilidade de comparar duas distribuições etárias distintas e analisar a influência que elas exercem sobre a SP e a dimensão amostral. Por último, a

## 4. FERRAMENTA COMPUTACIONAL

ferramenta permite também a comparação das dimensões amostrais obtidas com recurso a diferentes métodos de construção de intervalos de confiança para a SP.

Para isso, o utilizador deve fornecer um conjunto de parâmetros, a começar pela seleção do método de construção dos intervalos de confiança para a SP. De seguida, é necessário carregar um ficheiro de dados no formato ".txt", contendo duas colunas: uma com as idades, dadas por valores inteiros, e outra com a frequência absoluta de cada idade na população em estudo. Para garantir o correto funcionamento, a coluna correspondente às idades deve ser a primeira, seguida pela coluna da frequência. Outros parâmetros são a SRR e a SCR, que podem ser obtidas a partir de um estudo piloto ou de um estudo anterior, assim como, os limites inferior e superior para a SCR, a dimensão amostral máxima, o nível de confiança desejado e a amplitude relativa desejada. É importante notar que, embora a SCR geralmente tome valores entre 0 e 1, em cenários de alta intensidade de transmissão pode ser necessário ajustar os limites definidos pelo utilizador para permitir valores superiores a 1.

Relativamente aos cálculos, a ferramenta utiliza 6 funções principais, as quais podem ser consultadas no apêndice D. A primeira calcula a probabilidade de um indivíduo ser seropositivo, considerando a sua idade  $t$ , a SCR e a SRR. A segunda calcula a SP com base na distribuição etária e na probabilidade de seropositividade. A partir da SP calculada, duas funções são responsáveis por construir o intervalo de confiança associado: uma para o método de Wald com correção de continuidade (desenvolvida neste trabalho) e outra para os restantes métodos (com recurso ao pacote `binom`). Uma quinta função calcula o intervalo de confiança para a SCR. Por fim, uma sexta função utiliza um algoritmo iterativo para determinar a dimensão amostral necessária, tendo em conta os requisitos de nível de confiança e da amplitude relativa do intervalo de confiança para a SCR especificados pelo utilizador. Estas funções estão compiladas no pacote `RCMsize`, disponível no CRAN [Graça et al., 2025].

A ferramenta apresenta os resultados tanto em formato textual quanto gráfico. O texto exibido inclui informações sobre a dimensão amostral recomendada, o método utilizado e o nível de confiança. Paralelamente, são gerados gráficos que mostram a distribuição etária da população, a relação entre a SP e a dimensão amostral e a relação entre a SCR e a SP.

Os métodos disponíveis são parte dos que estão implementados na função `binom.confint`, do pacote `binom` do R. O método exato refere-se ao método de Clopper-Pearson, também estão disponíveis os métodos de Wald, Wald com correção de continuidade de Yates, Wilson, Agresti-Coull, bem como os métodos `logit` e `cloglog`.

### 4.1 Tutorial

Esta secção explica como usar cada uma das funcionalidades da ferramenta para calcular a dimensão amostral e os intervalos de confiança em diferentes contextos. O utilizador pode ajustar os parâmetros de cada cenário conforme necessário e visualizar os resultados em gráficos e tabelas. Após a inserção dos parâmetros, basta clicar em “Calcular” para que a ferramenta determine automaticamente o número de participantes necessários para o estudo.

De seguida, são apresentados quatro cenários distintos, cada um com um exemplo prático que demonstra como a ferramenta pode ser utilizada. Utiliza-se dois conjuntos de dados, os mesmos que foram utilizados no artigo [Sepúlveda and Drakeley, 2015], sendo um deles referente a África e outro ao Sudeste Asiático e América do Sul. A distribuição etária em África é caracterizada por uma maior concentração de indivíduos nas faixas etárias mais jovens, e uma diminuição gradual à medida que se

avança para as idades mais elevadas. Já nas populações não africanas, a distribuição etária tende a ser mais homogênea, apresentando uma menor discrepância entre os diferentes grupos etários (Fig 4.1).

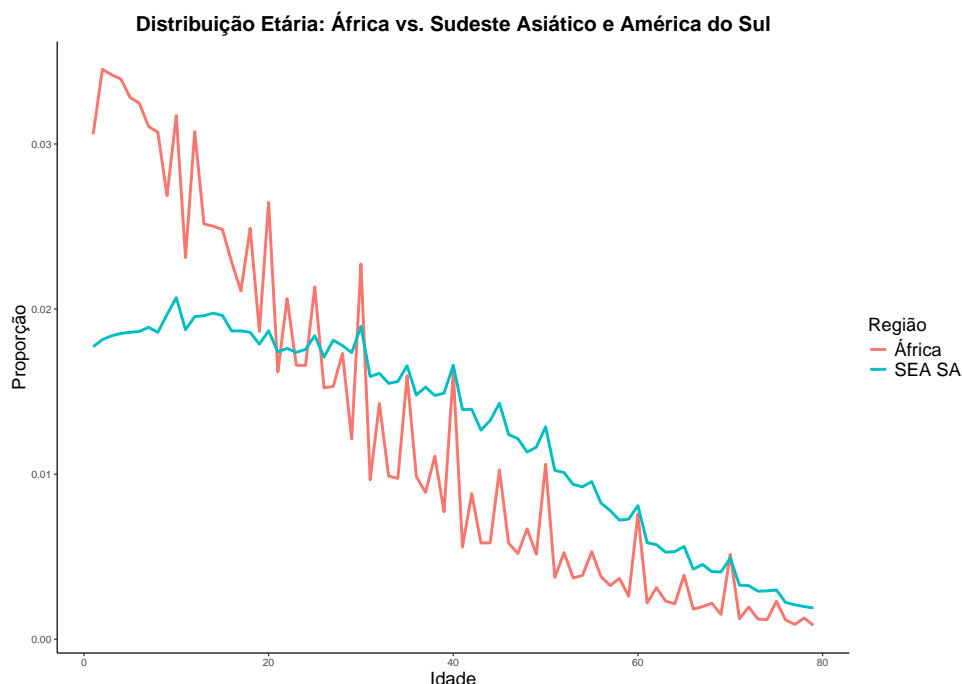


Figura 4.1: Distribuição etária para as duas populações.

#### 4.1.1 Cenário 1: Um único valor da SCR

Neste cenário, o objetivo é calcular a dimensão amostral necessária com base em um único valor da SCR e nos demais parâmetros fornecidos pelo utilizador.

A Fig 4.2 mostra o resultado da dimensão amostral calculada para um cenário hipotético. Nela, é possível ver que, com uma SCR de 0.0324, uma SRR de 0.017, um nível de confiança de 95%, e uma amplitude relativa desejada de 1, a ferramenta determina que é necessária uma amostra de 73 participantes, utilizando o método de Wald com correção de continuidade no conjunto de dados referentes à população africana.

Além disso, a ferramenta também apresenta o intervalo de confiança associado (Fig 4.3), assim como, gera um gráfico da relação entre a SCR e SP (Fig 4.4), um gráfico da distribuição etária da população (Fig 4.5), conforme os dados carregados pelo utilizador e um pequeno resumo da teoria utilizada (Fig 4.6).

#### 4.1.2 Cenário 2: Um conjunto de valores da SCR

Neste cenário, expandiu-se a análise para incluir um conjunto de valores da SCR, que é introduzido como um vetor separado por vírgulas. Os demais parâmetros permanecem inalterados, o que permite observar como a dimensão amostral necessária varia com os diferentes níveis de SCR/SP.

A Fig 4.7 apresenta os resultados obtidos para um cenário similar ao anterior, em forma de tabela, que inclui os intervalos de confiança para a SCR e SP. Foram considerados os seguintes valores de SCR (0.0036, 0.0108, 0.0324, 0.0969, 0.29) e obteve-se as seguintes dimensões amostrais (316, 128, 73, 79, 151). Paralelamente, são gerados dois gráficos: um que mostra a relação entre a SCR e a SP (Fig 4.8),

## 4. FERRAMENTA COMPUTACIONAL

Método: Wald com correção de continuidade

Upload ficheiro: data\_africa.txt

Taxa de serorreversão (SRR): 0.017

Taxa de seroconversão (SCR): 0.0324

Dimensão amostral máxima: 1000

Limites para a SCR [vetor; separado por vírgula]: 0,1

Nível de confiança: 0,95

Amplitude relativa (relative length): 1

Calcular

Dimensão Intervalos de confiança Gráfico SCR vs. SP Gráfico distribuição de idades Teoria

Tendo por base a distribuição de idades e os parâmetros selecionadas, a dimensão necessária é aproximadamente 73. Foi utilizado o método Wald com correção de continuidade e um nível de confiança de 0.95.

Figura 4.2: Cálculo da dimensão amostral com uma taxa de seroconversão de 0.0324, taxa de serorreversão de 0.017 e nível de confiança de 95%, para o conjunto de dados referentes à população africana.

Método: Wald com correção de continuidade

Upload ficheiro: data\_africa.txt

Taxa de serorreversão (SRR): 0.017

Taxa de seroconversão (SCR): 0.0324

Dimensão amostral máxima: 1000

Limites para a SCR [vetor; separado por vírgula]: 0,1

Nível de confiança: 0,95

Amplitude relativa (relative length): 1

Calcular

Dimensão	Intervalos de confiança	Gráfico SCR vs. SP	Gráfico distribuição de idades	Teoria
SCR	SP	Intervalo.SCR	Intervalo.SP	n
0.0324	0.3654132	0.0189 ; 0.0513	0.2481 ; 0.4827	73

Figura 4.3: Intervalos de confiança para uma taxa de seroconversão de 0.0324, taxa de serorreversão de 0.017 e nível de confiança de 95%, para o conjunto de dados referentes à população africana.

e outro que ilustra a relação entre a SP e a dimensão amostral (Fig 4.9). Assim, é possível visualizar a relação entre os diferentes valores de SP e o número recomendado de participantes para cada um.

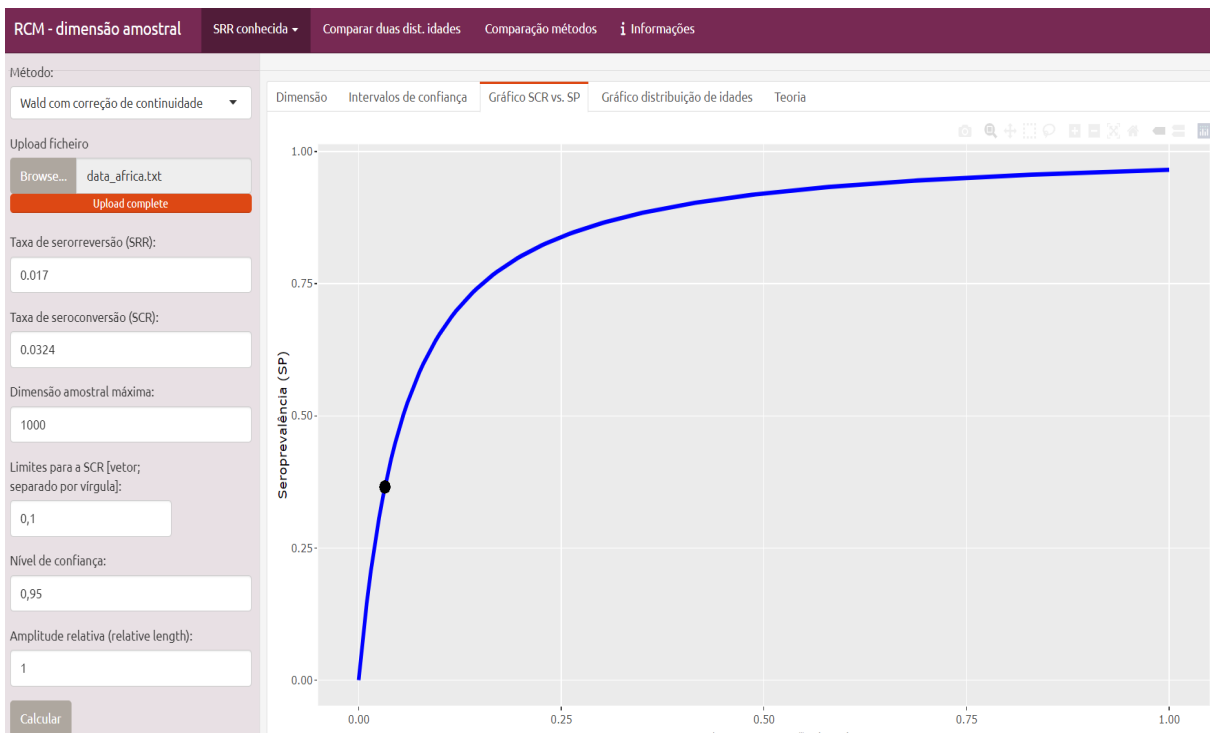


Figura 4.4: SCR vs. SP para uma taxa de seroconversão de 0.0324, taxa de serorreversão de 0.017 e nível de confiança de 95%, para o conjunto de dados referentes à população africana.

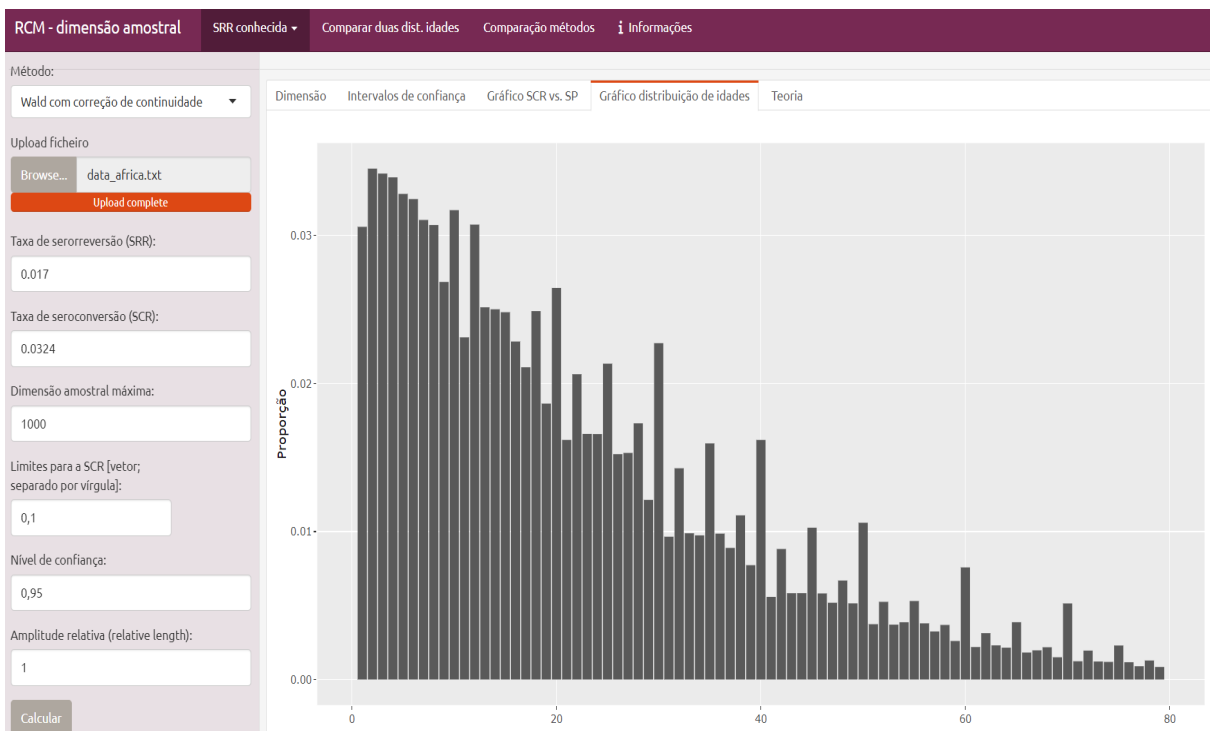


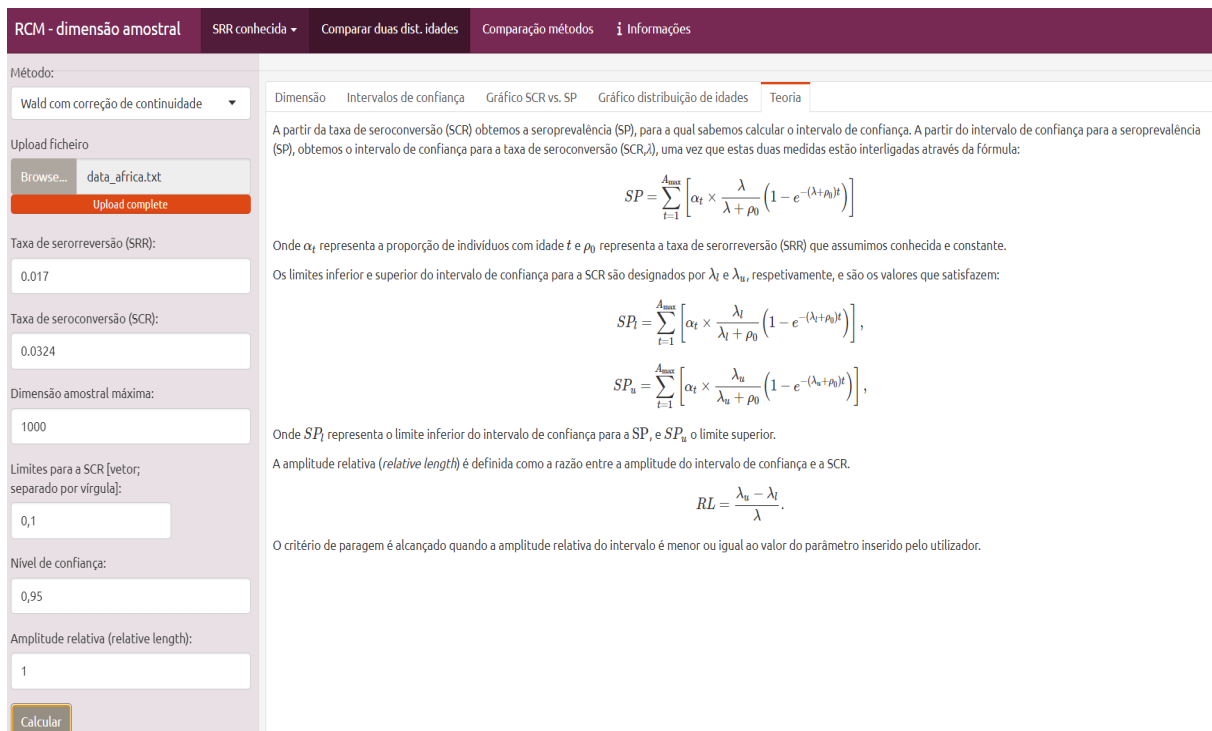
Figura 4.5: Distribuição etária da população africana.

### 4.1.3 Cenário 3: Comparação de distribuições etárias

Neste cenário, o objetivo é observar como a estrutura etária duma população pode influenciar a dimensão amostral necessária.

A Fig 4.10 apresenta o cálculo da dimensão amostral para as duas distribuições etárias, e para as quais

## 4. FERRAMENTA COMPUTACIONAL



Método: Wald com correção de continuidade

Upload ficheiro: data\_africa.txt

Taxa de serorreversão (SRR): 0.017

Taxa de seroconversão (SCR): 0.0324

Dimensão amostral máxima: 1000

Límites para a SCR (vetor; separado por vírgula): 0,1

Nível de confiança: 0,95

Amplitude relativa (relative length): 1

Calcular

Dimensão Intervalos de confiança Gráfico SCR vs. SP Gráfico distribuição de idades **Teoria**

A partir da taxa de seroconversão (SCR) obtemos a seroprevalência (SP), para a qual sabemos calcular o intervalo de confiança. A partir do intervalo de confiança para a seroprevalência (SP), obtemos o intervalo de confiança para a taxa de seroconversão (SCR,  $\lambda$ ), uma vez que estas duas medidas estão interligadas através da fórmula:

$$SP = \sum_{t=1}^{A_{max}} \left[ \alpha_t \times \frac{\lambda}{\lambda + \rho_0} (1 - e^{-(\lambda + \rho_0)t}) \right]$$

Onde  $\alpha_t$  representa a proporção de indivíduos com idade  $t$  e  $\rho_0$  representa a taxa de serorreversão (SRR) que assumimos conhecida e constante.

Os limites inferior e superior do intervalo de confiança para a SCR são designados por  $\lambda_l$  e  $\lambda_u$ , respetivamente, e são os valores que satisfazem:

$$SP_l = \sum_{t=1}^{A_{max}} \left[ \alpha_t \times \frac{\lambda_l}{\lambda_l + \rho_0} (1 - e^{-(\lambda_l + \rho_0)t}) \right],$$

$$SP_u = \sum_{t=1}^{A_{max}} \left[ \alpha_t \times \frac{\lambda_u}{\lambda_u + \rho_0} (1 - e^{-(\lambda_u + \rho_0)t}) \right],$$

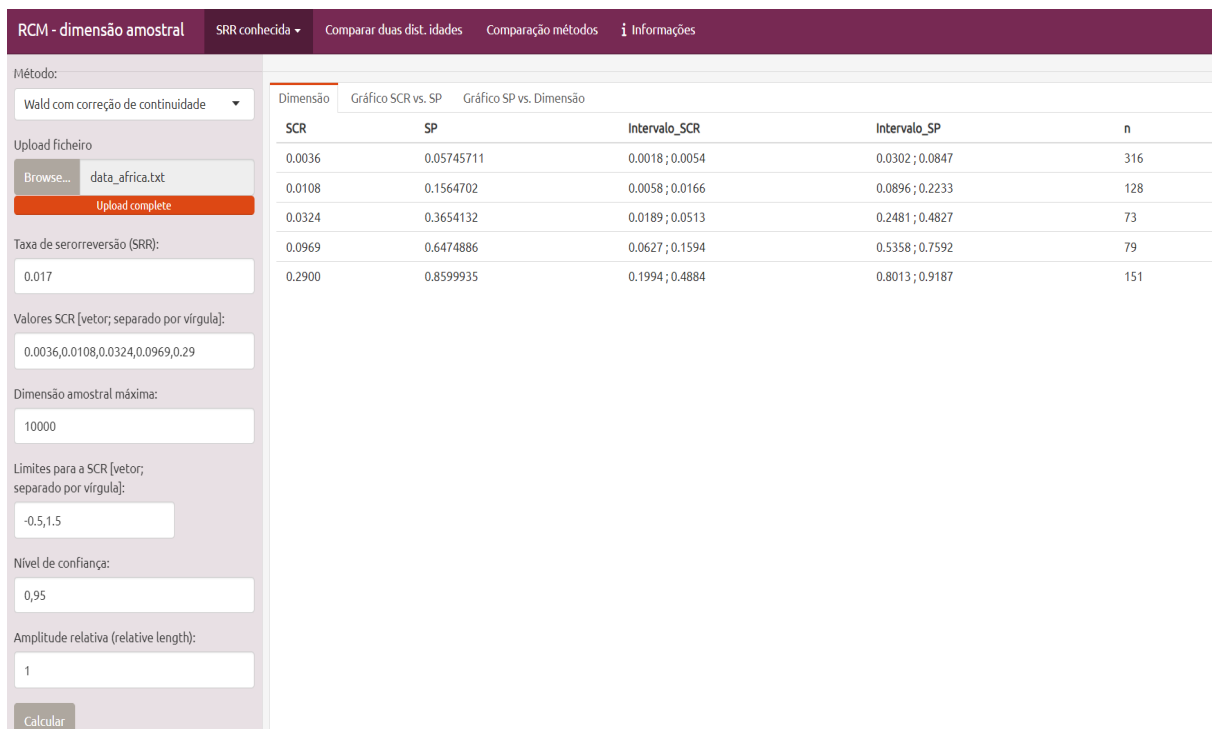
Onde  $SP_l$  representa o limite inferior do intervalo de confiança para a SP, e  $SP_u$  o limite superior.

A amplitude relativa (*relative length*) é definida como a razão entre a amplitude do intervalo de confiança e a SCR.

$$RL = \frac{\lambda_u - \lambda_l}{\lambda}$$

O critério de paragem é alcançado quando a amplitude relativa do intervalo é menor ou igual ao valor do parâmetro inserido pelo utilizador.

Figura 4.6: Resumo da teoria para o cálculo da dimensão amostral necessária.



Método: Wald com correção de continuidade

Upload ficheiro: data\_africa.txt

Taxa de serorreversão (SRR): 0.017

Valores SCR (vetor; separado por vírgula): 0.0036,0.0108,0.0324,0.0969,0.29

Dimensão amostral máxima: 10000

Límites para a SCR (vetor; separado por vírgula): -0.5,1.5

Nível de confiança: 0,95

Amplitude relativa (relative length): 1

Calcular

Dimensão Gráfico SCR vs. SP Gráfico SP vs. Dimensão

SCR	SP	Intervalo_SCR	Intervalo_SP	n
0.0036	0.05745711	0.0018; 0.0054	0.0302; 0.0847	316
0.0108	0.1564702	0.0058; 0.0166	0.0896; 0.2233	128
0.0324	0.3654132	0.0189; 0.0513	0.2481; 0.4827	73
0.0969	0.6474886	0.0627; 0.1594	0.5358; 0.7592	79
0.2900	0.8599935	0.1994; 0.4884	0.8013; 0.9187	151

Figura 4.7: Dimensão amostral calculada para diferentes valores de SCR.

foram obtidas dimensões amostrais de 73 e 67, respetivamente, considerando os mesmos parâmetros do exemplo apresentado no cenário 1 (4.1.1).

Assim como, também é apresentada uma tabela com os intervalos de confiança (Fig 4.11), um gráfico da relação entre a SCR e a SP (Fig 4.12) e um gráfico com as distribuições etárias das populações em estudo (Fig 4.13).

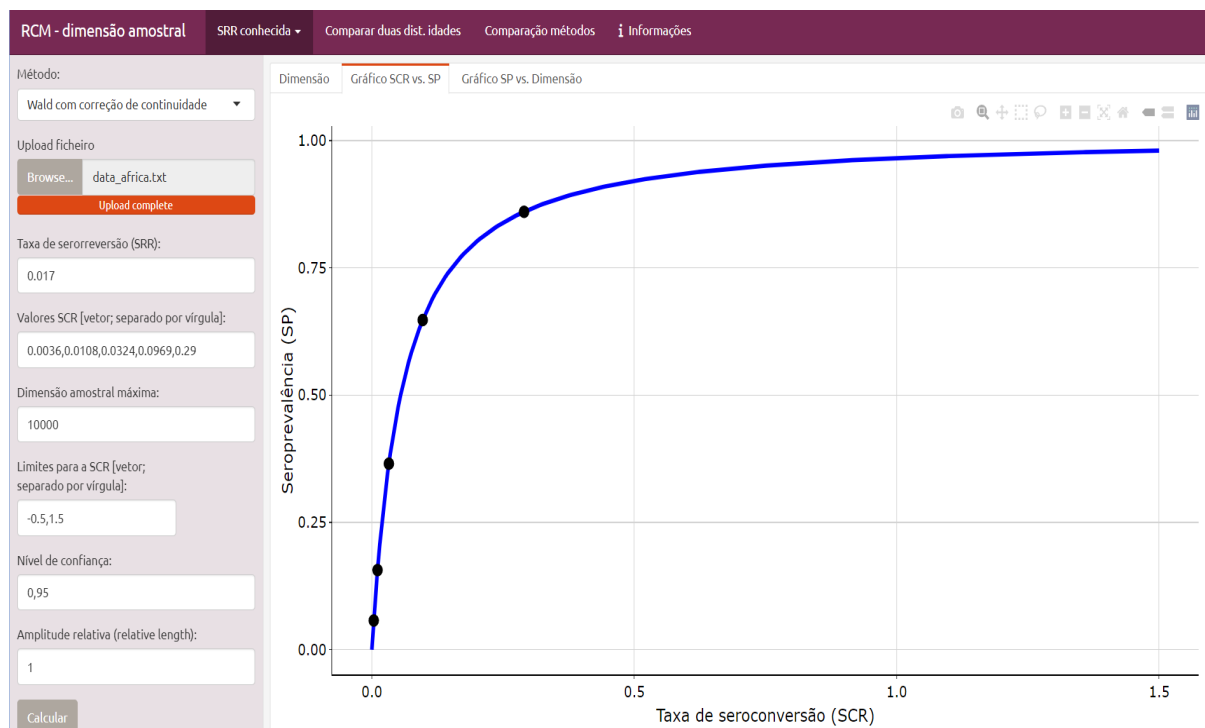


Figura 4.8: Função da seroprevalência para a população africana com uma SRR de 0.017.

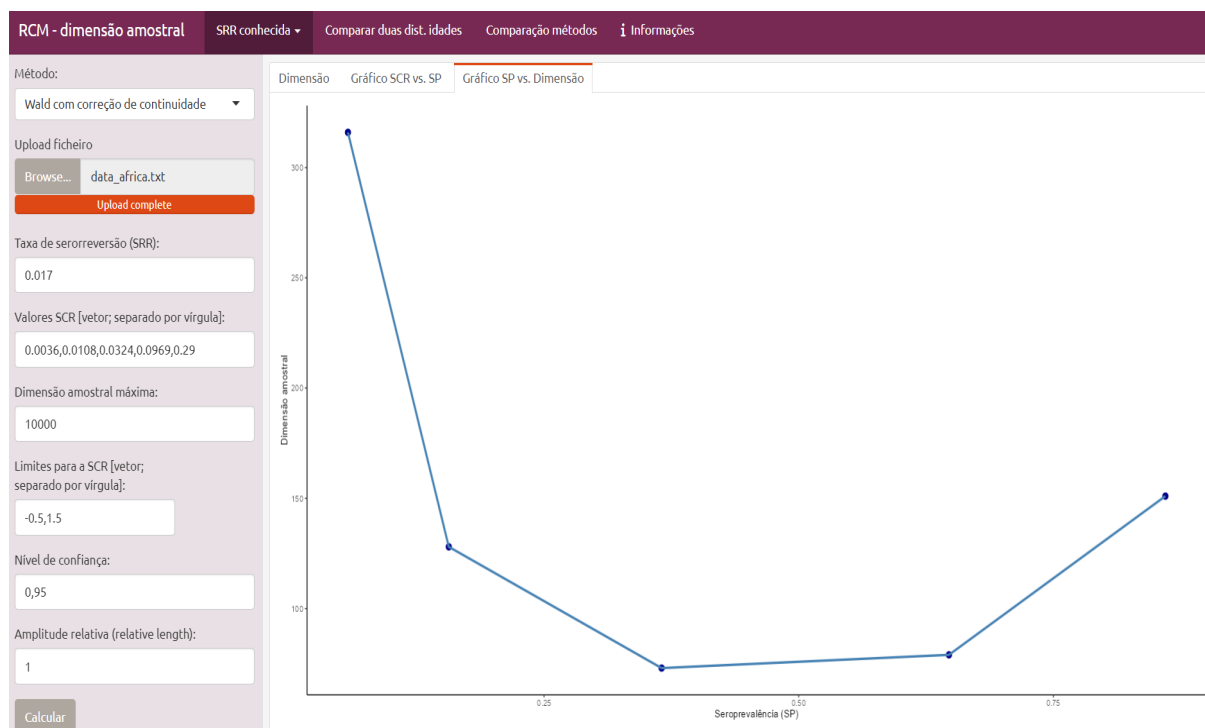


Figura 4.9: Gráfico da dimensão amostral calculada para diferentes valores de SCR.

#### 4.1.4 Cenário 4: Comparação de diferentes métodos de construção de intervalos de confiança

Neste cenário, a ferramenta permite ao utilizador comparar diferentes métodos de construção de intervalos de confiança para a SCR. O objetivo é analisar como diferentes abordagens afetam a amplitude dos intervalos, a probabilidade de cobertura e a dimensão amostral necessária, e assim ajudar o utilizador

## 4. FERRAMENTA COMPUTACIONAL

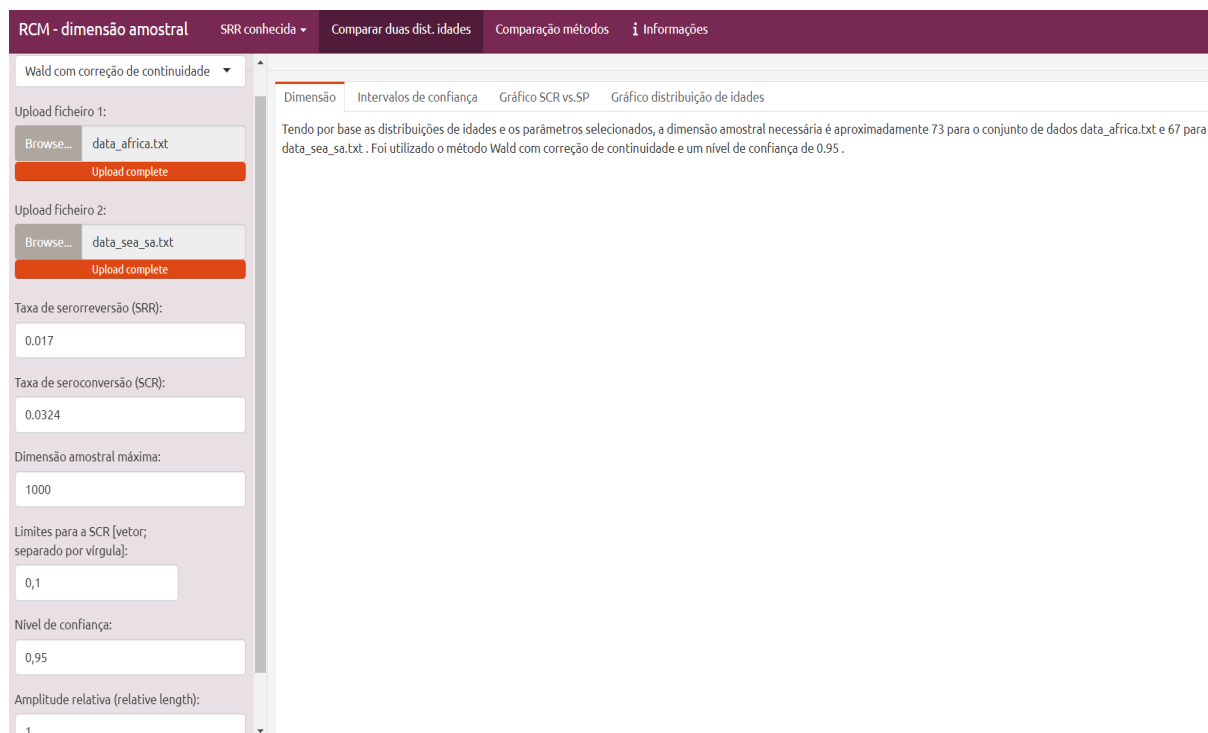
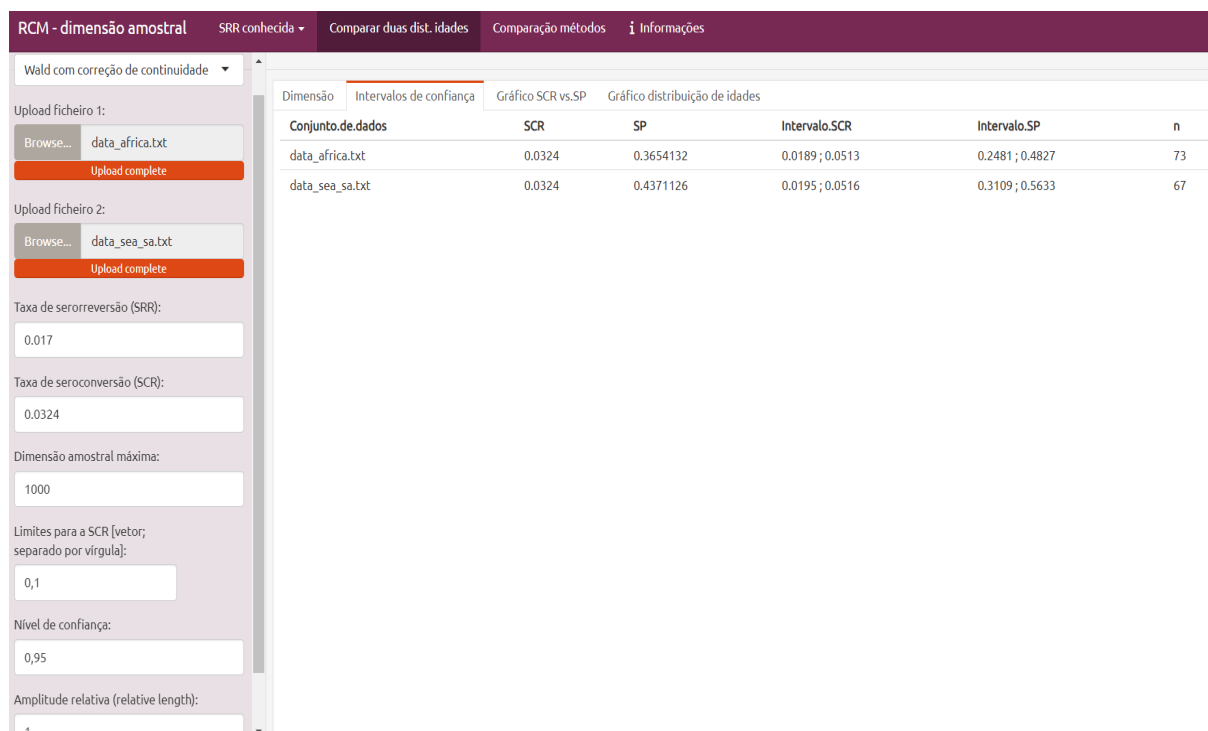


Figura 4.10: Cálculo da dimensão amostral para duas populações com distribuições etárias distintas.



Conjunto.de.dados	SCR	SP	Intervalo.SCR	Intervalo.SP	n
data_africa.txt	0.0324	0.3654132	0.0189 ; 0.0513	0.2481 ; 0.4827	73
data_sea_sa.txt	0.0324	0.4371126	0.0195 ; 0.0516	0.3109 ; 0.5633	67

Figura 4.11: Intervalos de confiança para a SCR e a SP nas duas populações com distribuições etárias distintas.

a escolher o método mais adequado para o seu estudo.

Depois de selecionar os métodos a comparar, a ferramenta gera uma tabela, como podemos ver na Fig 4.14 , que resume os intervalos de confiança calculados para cada método, considerando uma amplitude relativa de 1. A Fig 4.15 ilustra graficamente os intervalos de confiança calculados através dos métodos selecionados.

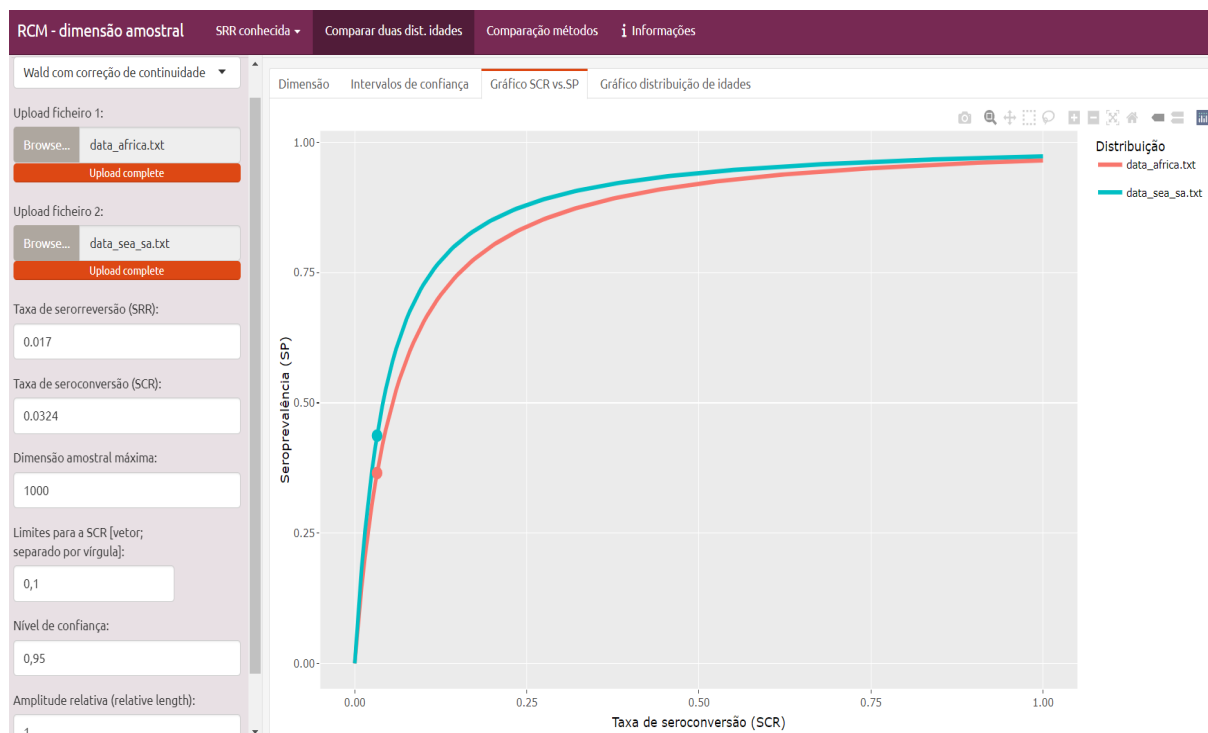


Figura 4.12: Relação entre a SCR e a SP para duas distribuições etárias distintas.



Figura 4.13: Distribuição etária para duas populações distintas.

#### 4.1.5 Informações

A ferramenta inclui ainda um menu de informações (Fig 4.16), no qual são apresentadas as suas funcionalidades. O utilizador pode consultar as instruções sobre o processo de importação do ficheiro de dados, assim como os métodos disponíveis para a construção dos intervalos de confiança.

## 4. FERRAMENTA COMPUTACIONAL

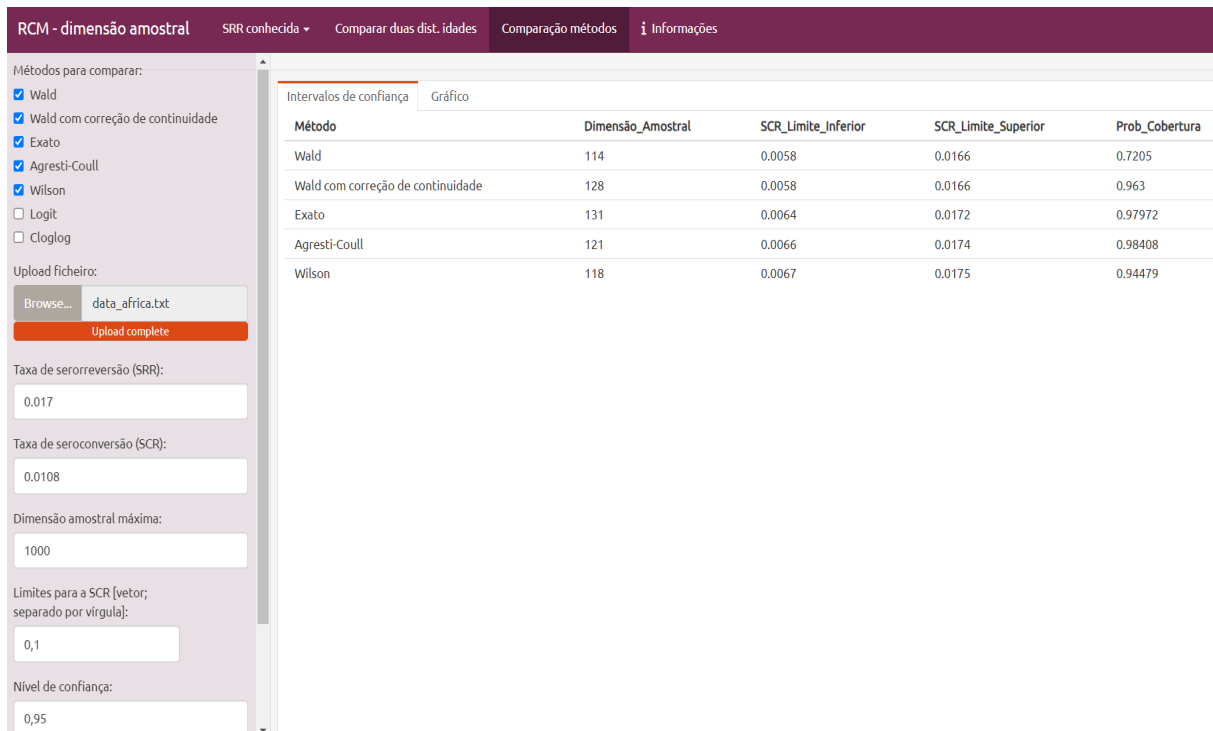


Figura 4.14: Tabela comparativa dos métodos de construção de intervalos de confiança, para uma amplitude relativa de 1.

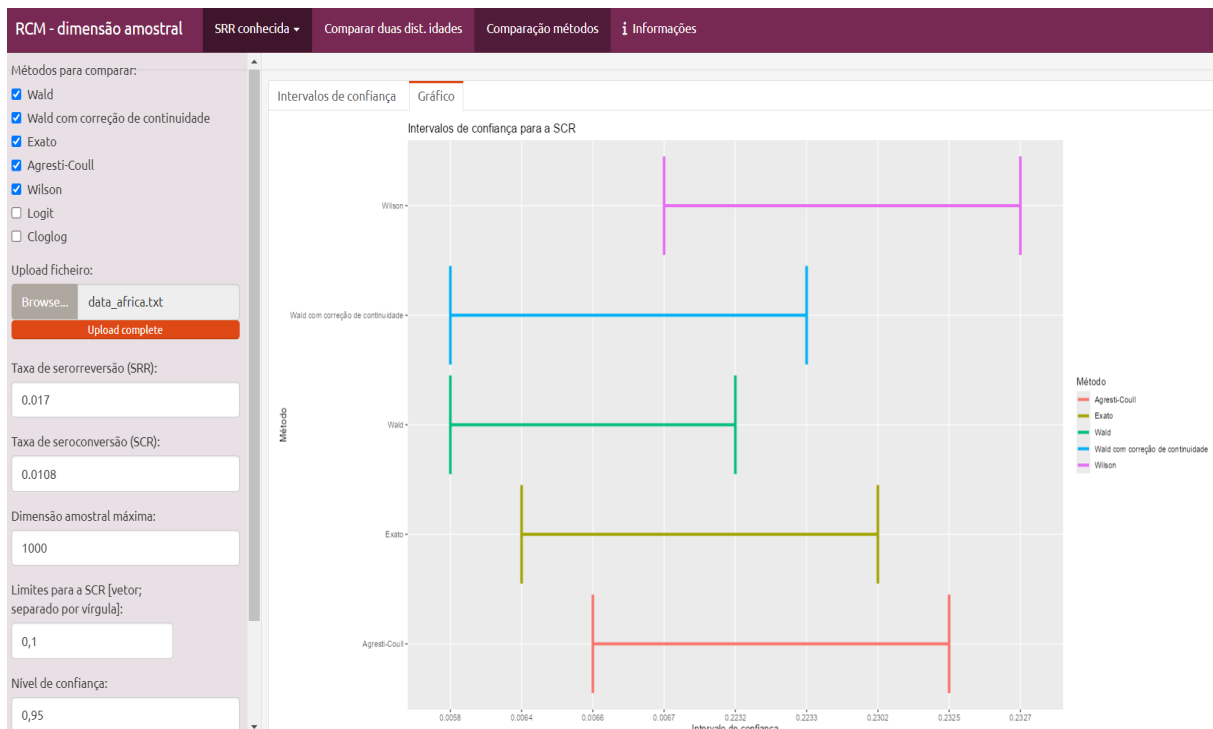


Figura 4.15: Comparação gráfica dos intervalos de confiança calculados.

RCM - dimensão amostral   SRR conhecida ▾   Comparar duas dist. idades   Comparação métodos   **i** Informações

Permite calcular a dimensão amostral necessária para uma população (uma distribuição de idades) ou comparar duas distribuições de idades.  
Também permite comparar as dimensões amostrais necessárias utilizando os diferentes métodos de construção de intervalos de confiança disponíveis.  
Presume-se que a taxa de serorreversão (SRR) é conhecida e constante.

### Como Importar os Dados

O ficheiro a carregar deve estar no formato ".txt". Este ficheiro deve conter duas colunas, a primeira coluna deve conter a variável **idade** e a segunda coluna a **frequência**.

### Métodos Utilizados

Os métodos utilizados são parte dos disponíveis na função `binom.confint` do pacote `binom`, ao qual se acrescentou o método de *Wald com correção de continuidade* ( $\pm \frac{1}{2n}$ ).  
O método de *Agresti-Coull* utiliza a fórmula proposta por [Brown et al.](#); o método exato utilizado é o de *Clopper-Pearson*.

### Lista de pacotes utilizados

- `shiny`
- `flexdashboard`
- `binom`
- `ggplot2`
- `plotly`

Figura 4.16: Menu de informações.



## Capítulo 5

# Resultados

Nesta secção, discutiremos os resultados obtidos no cálculo da dimensão amostral para estudos seroepidemiológicos sobre a malária, utilizando dois conjuntos de dados, os mesmos utilizados no artigo [Sepúlveda and Drakeley, 2015]. Esta análise permitirá comparar os resultados com os apresentados no referido estudo e evidenciar como as diferentes distribuições etárias têm impacto na SP e na necessidade de dimensões amostrais diferentes. Assim como discutiremos o desempenho dos diferentes métodos de construção de intervalos de confiança.

### 5.1 Reprodução dos resultados publicados

De seguida, apresentamos os resultados do cálculo da dimensão amostral necessária (Tabela 5.1) para duas populações com perfis demográficos diferentes (Fig 4.1). Fixamos a SRR em 0.017 e variámos a SCR entre os seguintes valores (0.0036, 0.0108, 0.0324, 0.0969 e 0.29), a amplitude relativa de (1, 0.75, 0.5 e 0.25), e comparamos os resultados com os apresentados no artigo [Sepúlveda and Drakeley, 2015].

Os valores para a SCR utilizados são baseados em dados de estudos epidemiológicos anteriores realizados no nordeste da Tanzânia, uma região caracterizada por uma variação natural na endemicidade da malária, com variações na SCR consoante a altitude [Drakeley et al., 2005]. Nomeadamente, verificou-se que a prevalência de *Plasmodium falciparum* é inversamente proporcional à altitude. Isto pode ser explicado pelo facto de a altitudes mais elevadas estarem associadas temperaturas mais baixas, e portanto, menos favoráveis à sobrevivência e reprodução dos mosquitos *Anopheles*. A partir desta relação entre a altitude e as medidas de intensidade de transmissão como a EIR foi possível determinar os valores da SCR a serem utilizados. Observou-se também que a estimativa da SRR parecia ser constante em todas as aldeias, com um valor de 0.017 [Sepúlveda and Drakeley, 2015].

Os resultados obtidos foram, em geral, semelhantes aos apresentados no artigo, com a maior discrepância observada para uma SCR de 0.0036. Esta discrepância aumenta à medida que a precisão desejada cresce, ou seja, à medida que a amplitude relativa diminui.

Todas as dimensões amostrais mencionadas nesta secção estão compiladas na Tabela 5.1 e todos os intervalos de confiança foram obtidos com recurso à ferramenta.

Um resultado apresentado no artigo [Sepúlveda and Drakeley, 2015] indica que, para valores de SCR  $\leq 0.058 \text{ ano}^{-1}$ , estudos sobre populações africanas requerem maiores dimensões amostrais em comparação com estudos sobre populações não africanas, a fim de obter o mesmo nível de precisão. O oposto acontece nos restantes níveis de intensidade de transmissão. Este resultado foi obtido utilizando o método de Wald com correção de continuidade conforme descrito em 2.2.2.1.

## 5. RESULTADOS

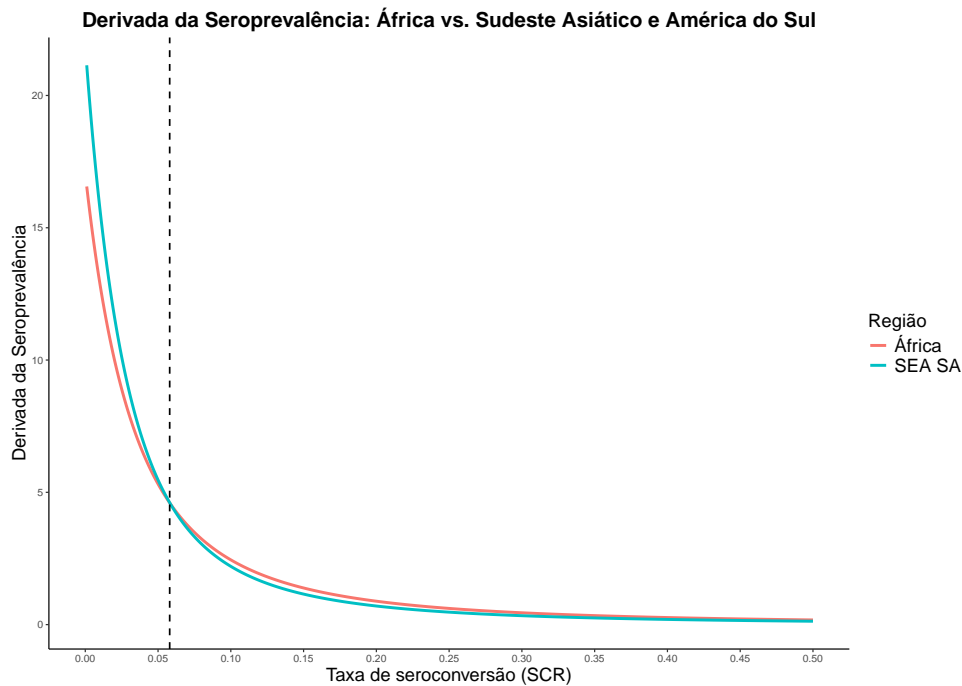


Figura 5.1: Derivada da função de seroprevalência.

A fig 5.1 ilustra a derivada da função de seroprevalência para as duas populações. Observamos que as linhas interseitam-se em  $SCR = 0.058$ , ou seja, para valores de SCR inferiores a 0.058, a derivada para a população africana é menor do que para a população não africana, o que significa que a população africana necessita de maiores dimensões amostrais para atingir o mesmo nível de precisão. Por outro lado, para valores de SCR superiores a 0.058, a situação inverte-se, e a população não africana requer maiores dimensões amostrais para a mesma precisão. Isto porque, uma menor derivada da função de seroprevalência implica uma curva SCR vs. SP menos íngreme, e uma curva menos íngreme significa que a mesma variação na SCR produz uma menor variação na SP. No caso da população africana, para uma SCR de 0.0108 (0.0324), observa-se uma SP de 0.156 (0.365), respetivamente, resultando numa mudança de 0.209. Em contraste, para a população não africana, considerando as mesmas intensidades de transmissão, corresponde uma SP de 0.195 (0.437), o que resulta numa mudança de 0.242. Portanto, para alcançar a mesma precisão na estimativa da SCR, um estudo com uma curva SCR vs. SP menos íngreme necessitará de uma dimensão amostral maior em comparação com um estudo com uma curva mais íngreme.

Por exemplo, para o método de Wald com correção de continuidade, assumindo uma SRR de 0.017, uma SCR de 0.0108, um nível de confiança de 0.95 e uma amplitude relativa de 1, para o conjunto de dados africanos, ao qual corresponde um intervalo de confiança para a SCR de (0.0058;0.0166), a dimensão amostral calculada foi de 128. Para o conjunto de dados referentes a populações do Sudeste Asiático e América do Sul, o intervalo de confiança para a SCR é (0.0059;0.0166) e a dimensão amostral necessária é 105. Se por outro lado a  $SCR = 0.0969$ , mantendo todos os outros parâmetros inalterados, à população africana corresponde um intervalo de confiança para a SCR de (0.0627;0.1594) e uma dimensão amostral de 79, e para a não africana um intervalo de confiança para a SCR de (0.0645;0.1612) e uma dimensão de 90. Isto pode ser explicado pela diferença nas distribuições de idades das duas populações, com a população africana com maior número de crianças e jovens, enquanto a não africana é mais uniforme, como podemos ver na Fig 4.1. A estrutura etária da população influencia a sua função de seroprevalência (equação 3.3) e portanto, diferentes distribuições etárias resultam em curvas com inclinações

## 5.1 Reprodução dos resultados publicados

distintas. Além disso, as dimensões amostrais também são influenciadas por condições socioeconômicas, práticas culturais e comportamentais, assim como por diferenças epidemiológicas. A maior prevalência do *Plasmodium falciparum* em África em comparação com o *Plasmodium vivax* no Sudeste Asiático e na América do Sul [Sepúlveda et al., 2015b], pode contribuir para diferenças na resposta imunológica dos pacientes. Estas variações na resposta imunológica podem afetar o tempo de infecção, a duração dos anticorpos, e a taxa de mortalidade. Todos esses fatores têm um impacto na seroconversão e na serorreversão, o que pode justificar as diferenças na dimensão amostral necessária.

Outro resultado observado é que, com o aumento da precisão, ocorre um aumento da magnitude das diferenças entre as dimensões amostrais em relação às diversas intensidades de transmissão [Sepúlveda and Drakeley, 2015]. De acordo com os resultados apresentados na tabela 5.1, para uma amplitude relativa de 1, as dimensões variam de 73 a 316 para a população africana, e entre 67 e 250 para o Sudeste Asiático e América do Sul (SEA/SA). Quando a amplitude relativa é reduzida para 0.25, a magnitude das diferenças aumenta, variando de 977 a 4618 em África e de 890 a 3649 para o restante conjunto de dados.

Um dos resultados obtidos neste estudo mostra que a dimensão amostral necessária varia conforme a intensidade de transmissão. Nomeadamente, para valores de SCR intermédios a dimensão amostral necessária é inferior à observada para valores altos e baixos da SCR. Como podemos ver na Fig 4.9, para SCR correspondentes a (0.0036; 0.0108; 0.0324; 0.0969; 0.29), a dimensão amostral necessária correspondente é (316; 128; 73; 79; 151). Este comportamento é consistente em todos os casos analisados, conforme podemos confirmar na Tabela 5.1.

Este comportamento pode ser explicado pelo facto da SCR e da SP estarem relacionadas. Para intensidades de transmissão baixa a SP tende a ser baixa, refletindo uma menor exposição da população ao agente infeccioso. De forma análoga, para intensidades de transmissão altas, a SP tende a ser alta, uma vez que uma intensidade de transmissão alta implica maior exposição da população ao agente infeccioso. Quando a intensidade de transmissão toma valores intermédios a relação entre a SCR e SP não é tão linear.

Sendo a SP uma proporção a sua variância é dada por  $\frac{SP(1-SP)}{n}$ . Portanto, quando a SP está próxima de 0 (pouca exposição da população ao agente infeccioso) ou 1 (a maioria da população já foi exposta ao agente infeccioso), o termo  $SP(1-SP)$  tende para 0. Como tal, para garantir uma determinada amplitude relativa a dimensão amostral necessária é maior neste casos, em comparação com valores de SP intermédios.

## 5. RESULTADOS

Tabela 5.1: Resultados do cálculo da dimensão considerando  $SRR = 0.017$ .

<b>População</b>	<b>Amplitude relativa</b>	<b>SCR</b>	<b>Dimensão amostral</b>
África	1	0.0036	316
	1	0.0108	128
	1	0.0324	73
	1	0.0969	79
	1	0.29	151
	0.75	0.0036	546
	0.75	0.0108	220
	0.75	0.0324	123
	0.75	0.0969	127
	0.75	0.29	233
	0.5	0.0036	1191
	0.5	0.0108	478
	0.5	0.0324	264
	0.5	0.0969	262
	0.5	0.29	461
	0.25	0.0036	4618
	0.25	0.0108	1853
	0.25	0.0324	1010
	0.25	0.0969	977
	0.25	0.29	1671
SEA/SA	1	0.0036	250
	1	0.0108	105
	1	0.0324	67
	1	0.0969	90
	1	0.29	209
	0.75	0.0036	431
	0.75	0.0108	179
	0.75	0.0324	111
	0.75	0.0969	142
	0.75	0.29	320
	0.5	0.0036	941
	0.5	0.0108	389
	0.5	0.0324	234
	0.5	0.0969	288
	0.5	0.29	629
	0.25	0.0036	3649
	0.25	0.0108	1502
	0.25	0.0324	890
	0.25	0.0969	1060
	0.25	0.29	2263

### 5.2 Desempenho dos métodos de construção de intervalos de confiança para a SCR

Nesta secção, vamos analisar o desempenho de diferentes métodos de construção de intervalos de confiança, tendo como critério a conjugação entre a maior probabilidade de cobertura e a menor dimensão amostral necessária. Os resultados são comparados para três cenários de intensidade de transmissão: baixa (entre 0.0034 e 0.01), intermédia (entre 0.01 e 0.1) e alta (acima de 0.1), conforme definido em [Sepúlveda and Drakeley, 2015], e para as duas populações distintas.

Na comparação, utilizaremos os métodos de Wald, com e sem correção de continuidade, o método exato (Clopper-Pearson), o de Agresti-Coull e o de Wilson. A escolha destes métodos é motivada pelo facto de terem sido analisados e comparados em diversos estudos anteriores, nomeadamente em [Gonçalves et al., 2012] e [Pires and Amado, 2008]. Os resultados obtidos estão resumidos na Tabela 5.2.

Para o método de Wald com correção de continuidade, o cálculo da probabilidade de cobertura é feito com recurso a simulação ( $n = 1000$ ), através duma função desenvolvida para este trabalho e que pode ser consultada no apêndice E. Para os restantes métodos, é utilizada a função *binom.coverage* do pacote *binom*, que se baseia nas expansões assintóticas propostas em [Brown et al., 2002].

Para o conjunto de dados referentes a África, considerando uma SRR de 0.017, uma SCR de 0.0036 (cenário de muito baixa transmissibilidade), um nível de confiança de 0.95 e uma amplitude relativa de 1, o método de Wald apresenta um fraco desempenho, com a probabilidade de cobertura de 0.639, bastante inferior ao nível de confiança nominal. Por outro lado, tanto o método exato (Clopper-Pearson) quanto o de Agresti-Coull apresentam uma probabilidade de cobertura de cerca 0.99. O método de Wald com correção de continuidade apresenta uma cobertura de cerca de 95%, mas necessita da maior dimensão amostral dentro dos métodos não exatos. O método de Wilson apresenta a menor dimensão amostral (296) entre os métodos com boa cobertura e é, portanto, o mais indicado neste caso.

Para o conjunto de dados não africanos, mantendo todos os outros parâmetros inalterados, o resultado é semelhante: o método de Wald apresenta uma probabilidade de cobertura de 0.556, enquanto o método exato e o de Agresti-Coull têm coberturas em torno de 98%. Assim como no exemplo anterior, o método de Wald com correção de continuidade apresenta uma cobertura de cerca de 95% e a maior dimensão entre os métodos não exatos. O método de Wilson apresenta uma cobertura de cerca de 93%, ligeiramente abaixo do nível de confiança nominal, o que torna a escolha do “melhor” método menos evidente, pois as dimensões amostrais necessárias são relativamente próximas (250, 255, 242 e 233 para o método de Wald com correção de continuidade, exato, Agresti-Coull e Wilson, respetivamente).

Por sua vez, ao considerarmos ainda um cenário de baixa transmissibilidade, mas com uma SCR de 0.009 (superior à anterior), mantendo todos os demais parâmetros inalterados, a probabilidade de cobertura do método de Wald aumenta para aproximadamente 0.7 e 0.625, para as populações africana e não africana, respetivamente. Embora ainda abaixo do nível de confiança nominal, esta cobertura apresenta uma melhoria em relação ao cenário anterior.

Em ambos os casos o método que requer menor dimensão amostral e apresenta uma boa cobertura é o método de Wilson.

## 5. RESULTADOS

Considerando uma SCR de 0.08 (intensidade intermédia), para a população africana, e mantendo os outros parâmetros inalterados, a probabilidade de cobertura do método de Wald é de 0.940, ainda abaixo do nível de confiança nominal, mas superior à do cenário de baixa intensidade, com correção de continuidade a probabilidade de cobertura é de 0.977. O método exato apresenta uma cobertura, em torno de 97%. Por outro lado, os métodos de Wilson e Agresti-Coull ficam empatados, ambos com uma cobertura de 96% e uma dimensão amostral necessária de 59.

Fazendo a mesma análise para a população não africana, o método de Wald apresenta uma probabilidade de cobertura de aproximadamente 0.9 e de 0.975 com correção de continuidade. Tanto o método exato quanto o de Agresti-Coull apresentam uma cobertura em torno de 98%, enquanto o método de Wilson apresenta uma cobertura de 93%. Em relação à dimensão amostral, os métodos de Agresti-Coull e Wilson exigem as menores dimensões amostrais (ambas no valor de 62), mas como o método de Wilson apresenta uma cobertura abaixo do nível de confiança nominal, o método de Agresti-Coull é considerado, neste caso, o melhor.

Considerando agora um cenário de elevada intensidade de transmissão na população não africana, nomeadamente, uma SCR de 0.29, as probabilidades de cobertura situam-se entre 0.94 e 0.97 para os diferentes métodos, com as coberturas dos métodos de Agresti-Coull e Wilson ligeiramente abaixo do nível de confiança nominal (0.948 e 0.944, respetivamente). Mais uma vez, a escolha do melhor método não é imediata: o método de Wilson requer a menor dimensão amostral (139) entre os quatro métodos, mas é o que apresenta a menor cobertura. O método de Agresti-Coull requer a segunda menor dimensão amostral (147) e oferece uma cobertura próxima de 95%, ainda que abaixo do nível de confiança nominal.

Para a população africana e uma SCR de 0.29, o método de Wilson apresenta uma probabilidade de cobertura de 0.956 e a menor dimensão amostral necessária (102).

Resumindo, o método de Wald apresentou um desempenho fraco na maioria dos cenários analisados, especialmente nos casos de baixa transmissibilidade, onde a cobertura é consistentemente inferior ao nível de confiança nominal de 0.95, com valores inadequadamente baixos (por exemplo, 0.639 e 0.556 em cenários de muito baixa transmissibilidade). Apesar da correção de continuidade aumentar a probabilidade de cobertura, para valores de SCR intermédios, onde o método de Wald (simples) apresenta um desempenho razoável, o pequeno aumento na probabilidade de cobertura poderá não compensar o grande aumento na dimensão amostral necessária associado. Por sua vez, tanto o método exato (Clopper-Pearson) como o de Agresti-Coull apresentaram uma probabilidade de cobertura acima do nível de confiança nominal, à exceção do método de Agresti-Coull num dos cenários. O método de Wilson apresenta uma probabilidade de cobertura acima ou próxima do nível de confiança nominal com a menor ou uma das menores dimensões amostrais.

Estes resultados sugerem que o método de Wilson é, geralmente, a melhor opção devido à sua boa cobertura associada à menor ou uma das menores dimensões amostrais. O método de Agresti-Coull destaca-se em regiões não africanas, o que poderá dever-se a uma distribuição etária mais uniforme nestas populações. Estes resultados estão em conformidade com o apresentado em [Gonçalves et al., 2012] e [Pires and Amado, 2008].

## 5.2 Desempenho dos métodos de construção de intervalos de confiança para a SCR

Tabela 5.2: Resultados do cálculo da probabilidade de cobertura para diferentes métodos de construção de intervalos de confiança, considerando uma SRR de 0.017, um nível de confiança de 0.95 e uma amplitude relativa de 1.

Região	SCR	Método	Probabilidade de Cobertura	Dimensão Amostral
África	0.0036	Wald	0.639	280
		Waldcc	0.948	316
		Wilson	0.967	296
		Agresti-Coull	0.991	308
		Clopper-Pearson	0.990	322
	0.009	Wald	0.700	130
		Waldcc	0.966	146
		Wilson	0.951	136
		Agresti-Coull	0.986	140
		Clopper-Pearson	0.983	150
	0.08	Wald	0.940	66
		Waldcc	0.977	74
		Wilson	0.961	59
		Agresti-Coull	0.961	59
		Clopper-Pearson	0.969	69
	0.29	Wald	0.941	134
		Waldcc	0.974	151
		Wilson	0.956	102
		Agresti-Coull	0.949	107
		Clopper-Pearson	0.956	124
SEA/SA	0.0036	Wald	0.556	222
		Waldcc	0.954	250
		Wilson	0.931	233
		Agresti-Coull	0.983	242
		Clopper-Pearson	0.981	255
	0.009	Wald	0.625	105
		Waldcc	0.965	119
		Wilson	0.975	109
		Agresti-Coull	0.973	112
		Clopper-Pearson	0.993	121
	0.08	Wald	0.903	72
		Waldcc	0.976	81
		Wilson	0.931	62
		Agresti-Coull	0.976	62
		Clopper-Pearson	0.978	73
	0.29	Wald	0.953	187
		Waldcc	0.970	209
		Wilson	0.944	139
		Agresti-Coull	0.948	147
		Clopper-Pearson	0.959	170

## 5. RESULTADOS

## Capítulo 6

# Discussão e conclusão

Os resultados mostram que a dimensão necessária varia conforme a intensidade de transmissão. Em particular, para valores de SCR inferiores a 0.058, os estudos sobre a população africana requerem amostras maiores do que os do Sudeste Asiático e América do Sul. No entanto, para valores superiores, verifica-se a tendência oposta, sendo as dimensões amostrais necessárias menores para a população africana em comparação com as outras regiões. Também se observou que com o aumento da precisão desejada, aumenta a magnitude das diferenças entre as diversas intensidades de transmissão.

A ferramenta apresentou um bom desempenho na aplicação de RCM a dados serológicos, no entanto tem algumas limitações a considerar.

Uma das principais limitações é que os dados devem ser fornecidos exclusivamente em formato ".txt", com uma estrutura específica em que a primeira coluna contém as idades e a segunda coluna contém as frequências. Formatos como ".csv" e ".xlsx" não são compatíveis com esta ferramenta, embora sejam frequentemente utilizados. Além disso, para que a ferramenta funcione corretamente, é necessário seguir rigorosamente a estrutura das colunas.

Outra limitação é que, neste momento, a interface está disponível apenas em português. Embora haja planos para traduzir a ferramenta para inglês, até lá, esta restrição vai limitar o número de utilizadores que dela podem usufruir. No entanto, o pacote RCMsize está disponível em inglês.

Adicionalmente, assume-se que a SRR é conhecida e constante. No entanto, na prática, a SRR requer estudos longitudinais para ser determinada e, portanto, terá frequentemente de ser estimada. Num trabalho futuro, seria interessante adicionar à ferramenta um método que permita adaptar o cálculo do tamanho amostral a cenários em que existe incerteza na estimativa da SRR, como discutido em [Sepúlveda et al., 2015a].

Considerar a heterogeneidade na transmissão traz diversas vantagens, nomeadamente, permite avaliar o efeito de programas de controlo ou de erradicação da doença, e considerar o impacto dos movimentos migratórios na exposição e transmissão da malária. A migração pode alterar os padrões de transmissão, por exemplo, introduzindo a doença em regiões anteriormente não afetadas. Além disso, considerar a heterogeneidade na transmissão permite investigar como a exposição à malária varia considerando diferentes faixas etárias, e analisar a heterogeneidade individual, o que permite avaliar o efeito de covariáveis.

Seria também interessante integrar modelos mais complexos, como modelos SIM (*superinfection models*) e os modelos de aquisição de anticorpos. Os modelos SIM permitem considerar mais de um estado de seropositividade (diferentes níveis de exposição), o que os torna especialmente vantajosos em cenários de elevada intensidade de transmissão, onde as segundas infeções são frequentes. Já os modelos de aquisição de anticorpos vão além da simples distinção entre seropositivo e seronegativo, permitindo também medir os níveis de anticorpos presentes.

## **6. DISCUSSÃO E CONCLUSÃO**

Além disso, neste trabalho estamos a assumir diagnósticos serológicos perfeitos, o que na prática não acontece. As ferramentas de deteção disponíveis, por mais avançadas que sejam, possuem limitações inevitáveis em termos de sensibilidade e especificidade.

Espero que este trabalho contribua para o planeamento de estudos sobre a malária e que as melhorias nas limitações da ferramenta, bem como a inclusão de cenários mais complexos possam vir a ser implementadas no futuro, o que contribuirá para melhorar o desempenho da ferramenta e expandir a sua aplicabilidade.

## Apêndice A

# Cálculo da dimensão amostral no método de Wald com correção de continuidade

$$\begin{aligned}W &= 2z_{1-\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}} + \frac{1}{n} \Leftrightarrow \\ \Leftrightarrow Wn &= 2z_{1-\frac{\alpha}{2}} n \sqrt{\frac{p(1-p)}{n}} + 1 \Leftrightarrow \\ \Leftrightarrow Wn - 1 &= 2z_{1-\frac{\alpha}{2}} \sqrt{np(1-p)} \Leftrightarrow \\ \Leftrightarrow (Wn - 1)^2 &= 4z_{1-\frac{\alpha}{2}}^2 np(1-p) \Leftrightarrow \\ \Leftrightarrow W^2 n^2 - 2Wn + 1 &= 4z_{1-\frac{\alpha}{2}}^2 np(1-p) \Leftrightarrow \\ \Leftrightarrow W^2 n^2 - (2W + 4z_{1-\frac{\alpha}{2}}^2 p(1-p))n + 1 &= 0.\end{aligned}$$

Trata-se de uma equação quadrática ( $an^2 + bn + c = 0$ ), com  $a = W^2$ ,  $b = -(2W + 4z_{1-\frac{\alpha}{2}}^2 p(1-p))$  e  $c = 1$ , sendo as soluções dadas por

$$n = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Obtém-se

$$n = \frac{2W + 4z_{1-\frac{\alpha}{2}}^2 p(1-p) \pm \sqrt{(2W + 4z_{1-\frac{\alpha}{2}}^2 p(1-p))^2 - 4W^2}}{2W^2},$$

sendo apenas considerada a solução positiva.

## Apêndice B

# Método de Wilson

Da equação 2.1 e dos respectivos limites dados por 2.2, surge

$$(\hat{p} - p)^2 = z_{1-\frac{\alpha}{2}}^2 \frac{p(1-p)}{n} \Leftrightarrow \hat{p}^2 - 2p\hat{p} + p^2 = \frac{z_{1-\frac{\alpha}{2}}^2}{n} p - \frac{z_{1-\frac{\alpha}{2}}^2}{n} p^2.$$

Por razões de simplicidade de escrita vamos substituir  $z_{1-\frac{\alpha}{2}}$  por  $z$ .

$$\hat{p}^2 - 2\hat{p}p - \frac{z^2}{n}p + p^2 + \frac{z^2}{n}p^2 = 0 \Leftrightarrow \left(1 + \frac{z^2}{n}\right)p^2 - \left(2\hat{p} + \frac{z^2}{n}\right)p + \hat{p}^2 = 0.$$

Estamos perante uma equação quadrática em função de  $p$ :  $ap^2 + bp + c = 0$ , com  $a = 1 + \frac{z^2}{n}$ ,  $b = -\left(2\hat{p} + \frac{z^2}{n}\right)$ ,  $c = \hat{p}^2$ .

Donde se obtém

$$\begin{aligned} p &= \frac{\left(2\hat{p} + \frac{z^2}{n}\right) \pm \sqrt{\left(2\hat{p} + \frac{z^2}{n}\right)^2 - 4\left(1 + \frac{z^2}{n}\right)\hat{p}^2}}{2\left(1 + \frac{z^2}{n}\right)} \Leftrightarrow \\ \Leftrightarrow p &= \frac{\left(2\hat{p} + \frac{2z^2}{2n}\right) \pm \sqrt{\left(2\hat{p} + \frac{2z^2}{2n}\right)^2 - 4\left(1 + \frac{z^2}{n}\right)\hat{p}^2}}{2\left(1 + \frac{z^2}{n}\right)} \Leftrightarrow \\ \Leftrightarrow p &= \frac{2\left(\hat{p} + \frac{z^2}{2n}\right) \pm \sqrt{4\left(\hat{p} + \frac{z^2}{2n}\right)^2 - 4\left(1 + \frac{z^2}{n}\right)\hat{p}^2}}{2\left(1 + \frac{z^2}{n}\right)} \Leftrightarrow \\ \Leftrightarrow p &= \frac{\hat{p} + \frac{z^2}{2n} \pm \sqrt{\left(\hat{p} + \frac{z^2}{2n}\right)^2 - \left(1 + \frac{z^2}{n}\right)\hat{p}^2}}{1 + \frac{z^2}{n}} \Leftrightarrow \\ \Leftrightarrow p &= \frac{\hat{p} + \frac{z^2}{2n} \pm \sqrt{\hat{p}^2 + \frac{2\hat{p}z^2}{2n} + \left(\frac{z^2}{2n}\right)^2 - \hat{p}^2 - \hat{p}\frac{z^2}{n}}}{1 + \frac{z^2}{n}} \Leftrightarrow \\ \Leftrightarrow p &= \frac{\hat{p} + \frac{z^2}{2n} \pm \sqrt{\frac{z^2}{n}(\hat{p} - \hat{p}^2) + \frac{z^4}{4n^2}}}{1 + \frac{z^2}{n}} \Leftrightarrow \end{aligned}$$

$$\Leftrightarrow p = \frac{\hat{p} + \frac{z^2}{2n} \pm z \sqrt{\frac{1}{n}(\hat{p} - \hat{p}^2) + \frac{z^2}{4n^2}}}{1 + \frac{z^2}{n}} \Leftrightarrow$$

$$\Leftrightarrow p = \frac{\hat{p} + \frac{z^2}{2n} \pm z \sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z^2}{4n^2}}}{1 + \frac{z^2}{n}}.$$

Segundo o método de Wilson, o intervalo de confiança  $(1 - \alpha)100\%$  para  $p$  é dado por

$$\frac{\hat{p} + \frac{z_{1-\frac{\alpha}{2}}^2}{2n} \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z_{1-\frac{\alpha}{2}}^2}{4n^2}}}{1 + \frac{z_{1-\frac{\alpha}{2}}^2}{n}}. \quad (\text{B.1})$$

## Apêndice C

### Transmissão estável

$$\begin{aligned}\frac{dP}{dt} &= \lambda(1-P) - \rho P \\ \int \frac{1}{\lambda - P(\lambda + \rho)} dP &= \int dt \Leftrightarrow \\ \Leftrightarrow -\frac{1}{\lambda + \rho} \int \frac{-(\lambda + \rho)}{\lambda - P(\lambda + \rho)} dP &= \int dt \Leftrightarrow \\ \Leftrightarrow -\frac{1}{\lambda + \rho} \ln(\lambda - P(\lambda + \rho)) &= t + c, c \in \mathbb{R} \Leftrightarrow \\ \Leftrightarrow \ln(\lambda - P(\lambda + \rho)) &= -(\lambda + \rho)(t + c) \Leftrightarrow \\ \Leftrightarrow \lambda - P(\lambda + \rho) &= e^{-(\lambda + \rho)(t + c)} \Leftrightarrow \\ \Leftrightarrow P(\lambda + \rho) &= \lambda - e^{-(\lambda + \rho)(t + c)} \Leftrightarrow \\ \Leftrightarrow P &= \frac{1}{\lambda + \rho} \left( \lambda - e^{-(\lambda + \rho)(t + c)} \right).\end{aligned}$$

Pressupondo que todos os indivíduos nascem seronegativos, i.e,  $P(0) = 0$ , obtém-se

$$\begin{aligned}P(0) &= \frac{1}{\lambda + \rho} \left( \lambda - e^{-(\lambda + \rho)c} \right) = 0 \Leftrightarrow \\ \Leftrightarrow e^{-(\lambda + \rho)c} &= \lambda \Leftrightarrow \\ \Leftrightarrow c &= -\frac{\ln \lambda}{\lambda + \rho}.\end{aligned}$$

Substituindo,

$$\begin{aligned}P &= \frac{1}{\lambda + \rho} \left( \lambda - e^{-(\lambda + \rho)\left(t - \frac{\ln \lambda}{\lambda + \rho}\right)} \right) \Leftrightarrow \\ \Leftrightarrow P &= \frac{1}{\lambda + \rho} e^{-(\lambda + \rho)t + (\lambda + \rho)\frac{\ln \lambda}{(\lambda + \rho)}} \Leftrightarrow \\ \Leftrightarrow P &= \frac{1}{\lambda + \rho} \left( e^{-(\lambda + \rho)t} \lambda \right) \Leftrightarrow \\ \Leftrightarrow P &= \frac{\lambda}{\lambda + \rho} \left( 1 - e^{-(\lambda + \rho)t} \right),\end{aligned}$$

P também pode ser designado por  $p_{S+}$ .

## Apêndice D

# Funções disponíveis no pacote RCMsize

```
1 prob_seropositive <- function(SCR,SRR,t){
2   gamma <- SCR+SRR
3   ps <- (SCR/gamma)*(1-exp(-gamma*t))
4   return (ps)
5 }
6
7 seroprevalence <- function(ages, A_max, SCR,SRR){
8   SP <- 0
9   for (i in 1:A_max){
10    SP = SP + ages[i] * prob_seropositive(SCR,SRR,i)
11  }
12  return (SP)
13 }
14
15 IC_SP <- function(SP, n, conf.level = 0.95, method = "asymptotic"){
16  x <- SP*n
17  info <- binom::binom.confint(x,n,conf.level = conf.level,methods = method)
18  ll <- info$lower
19  ul <- info$upper
20  return(c(ll,ul))
21 }
22
23 IC_SP_Waldcc <- function(SP,n, conf.level=0.95){
24  alpha <- 1 - conf.level
25  z <- stats::qnorm(1 - alpha / 2)
26  ME <- z*sqrt(SP*(1-SP)/n)
27  ll <- SP-ME-1/(2*n)
28  ul <- SP+ME+1/(2*n)
29  return(c(ll,ul))
30 }
31
32 IC_SCR <- function(SP_interval, SRR, ages, A_max, limits=c(0,1)){
33  SCR_ll <- stats::uniroot(function(SCR) seroprevalence(ages, A_max, SCR, SRR) -
34    SP_interval[1], lower = limits[1], upper = limits[2], tol = 1e-5)$root
35  SCR_ul <- stats::uniroot(function(SCR) seroprevalence(ages, A_max, SCR, SRR) -
36    SP_interval[2], lower = limits[1], upper = limits[2], tol = 1e-5)$root
37  return(c(SCR_ll, SCR_ul))
38 }
```

## D. FUNÇÕES DISPONÍVEIS NO PACOTE RCMSIZE

```
39
40 sample_s <- function(SCR, RL, SRR, ages, A_max, limits, max_iter =
    10000, conf.level=0.95, method="asymptotic") {
41   if (RL <= 0) {
42     stop("RL should be a positive value")
43   }
44   ampl <- RL*SCR
45   n <- 50
46   sp <- seroprevalence(ages, A_max, SCR, SRR)
47   if (method == "waldcc") {
48     ic_sp <- IC_SP_Waldcc(sp, n, conf.level)
49   } else {
50     ic_sp <- IC_SP(sp, n, conf.level, method)
51   }
52
53   ic_scr <- IC_SCR(ic_sp, SRR, ages, A_max, limits)
54   amplitude <- ic_scr[2] - ic_scr[1]
55
56   iter <- 0
57
58   while (amplitude > ampl && iter < max_iter) {
59     n <- n + 1
60     if (method == "waldcc") {
61       ic_sp <- IC_SP_Waldcc(sp, n, conf.level)
62     } else {
63       ic_sp <- IC_SP(sp, n, conf.level, method)
64     }
65     ic_scr <- IC_SCR(ic_sp, SRR, ages, A_max, limits)
66     amplitude <- ic_scr[2] - ic_scr[1]
67     iter <- iter + 1
68   }
69
70   if (iter == max_iter) {
71     warning("Maximum number of iteration reached")
72   }
73   return(list(n = n, ci_sp = ic_sp, ci_scr = ic_scr))
74 }
```

## Apêndice E

# Funções para o cálculo da probabilidade de cobertura no método de Wald com correção de continuidade

```
1 IC_SCR_for_Waldcc <- function(SP_interval, SRR, ages, A_max, limits = c(-1, 2))
  {
2   seroprevalence <- function(SCR) {
3     mean(sapply(ages, function(age) prob_seropositivo(SCR, SRR, age)))
4   }
5
6   SCR_ll <- uniroot(function(SCR) seroprevalence(SCR) - SP_interval[1], lower =
  limits[1], upper = limits[2], tol = 1e-5)$root
7   SCR_ul <- uniroot(function(SCR) seroprevalence(SCR) - SP_interval[2], lower =
  limits[1], upper = limits[2], tol = 1e-5)$root
8   return(c(SCR_ll, SCR_ul))
9 }
10
11 calculate_seroprevalence <- function(ages, SCR, SRR) {
12   serologic_states <- sapply(ages, function(age) {
13     prob_seropositivo(SCR, SRR, age)
14   })
15   serologic_states <- rbinom(length(serologic_states), 1, serologic_states)
16   SP <- sum(serologic_states) / length(serologic_states)
17   return(SP)
18 }
19
20 coverage_probability <- function(SCR, SRR, n, A_max, age_distribution,
  num_simulations = 1000) {
21   coverage_count <- 0
22   for (i in 1:num_simulations) {
23     # Randomly create an age vector
24     ages <- sample(1:A_max, n, replace = TRUE, prob = age_distribution)
25     # Calculate seroprevalence
26     SP <- calculate_seroprevalence(ages, SCR, SRR)
27     # Calculate confidence interval for SP
28     SP_interval <- IC_SP_Waldcc(SP, n)
29     # Backtransform the confidence interval for SCR
30     SCR_interval <- IC_SCR_for_Waldcc(SP_interval, SRR, ages, A_max)
31     # Check if the true SCR is within the confidence interval
```

## E. FUNÇÕES PARA O CÁLCULO DA PROBABILIDADE DE COBERTURA NO MÉTODO DE WALD COM CORREÇÃO DE CONTINUIDADE

```
32     if (SCR >= SCR_interval[1] && SCR <= SCR_interval[2]) {
33         coverage_count <- coverage_count + 1
34     }
35 }
36 # Calculate coverage probability
37 coverage_prob <- coverage_count / num_simulations
38 return(coverage_prob)
39 }
```

---

# Bibliografia

- [Aden-Buie et al., 2023] Aden-Buie, G., Sievert, C., Iannone, R., Allaire, J., and Borges, B. (2023). *flexdashboard: R Markdown Format for Flexible Dashboards*. R package version 0.6.2.
- [Agresti and Caffo, 2000] Agresti, A. and Caffo, B. (2000). Simple and Effective Confidence Intervals for Proportions and Differences of Proportions Result from Adding Two Successes and Two Failures. *The American Statistician*, 54(4):280–288.
- [Anscombe, 1948] Anscombe, F. J. (1948). The transformation of poisson, binomial and negative-binomial data. *Biometrika*, 35(3-4):246–254.
- [Berry and Armitage, 1995] Berry, G. and Armitage, P. (1995). Mid-P Confidence Intervals: A Brief Review. *The Statistician*, 44(4):417. Publisher: JSTOR.
- [Blaker, 2000] Blaker, H. (2000). Confidence curves and improved exact confidence intervals for discrete distributions. *Canadian Journal of Statistics*, 28(4):783–798. Publisher: Wiley.
- [Brown et al., 2001] Brown, L. D., Cai, T. T., and DasGupta, A. (2001). Interval Estimation for a Binomial Proportion. *Statistical Science*, 16(2).
- [Brown et al., 2002] Brown, L. D., Cai, T. T., and DasGupta, A. (2002). Confidence Intervals for a binomial proportion and asymptotic expansions. *The Annals of Statistics*, 30(1).
- [Centers for Disease Control and Prevention, 2024] Centers for Disease Control and Prevention (2024). Cdc - dpdx - malaria. Accessed: 2024-06-14.
- [Chang et al., 2024] Chang, W., Cheng, J., Allaire, J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert, A., and Borges, B. (2024). *shiny: Web Application Framework for R*. R package version 1.9.1.
- [Clopper and Pearson, 1934] Clopper, C. J. and Pearson, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 26(4):404–413.
- [Dorai-Raj, 2022] Dorai-Raj, S. (2022). *binom: Binomial Confidence Intervals for Several Parameterizations*. R package version 1.1-1.1.
- [Drakeley et al., 2005] Drakeley, C., Carneiro, I., Reyburn, H., Malima, R., Lusingu, J., Cox, J., Theander, T., Nkya, W., Lemnge, M., and Riley, E. (2005). Altitude-Dependent and -Independent Variations in *Plasmodium falciparum* Prevalence in Northeastern Tanzania. *The Journal of Infectious Diseases*, 191(10):1589–1598.

## BIBLIOGRAFIA

- [Gonçalves et al., 2012] Gonçalves, L., De Oliveira, M. R., Pascoal, C., and Pires, A. (2012). Sample size for estimating a binomial proportion: comparison of different methods. *Journal of Applied Statistics*, 39(11):2453–2473.
- [Graça et al., 2025] Graça, M., Diamantino, F., and Sepúlveda, N. (2025). *RCMsize: Sample Size Calculation in Reversible Catalytic Models*. R package version 1.0.1.
- [Krishnamoorthy and Peng, 2007] Krishnamoorthy, K. and Peng, J. (2007). Some Properties of the Exact and Score Methods for Binomial Proportion and Sample Size Calculation. *Communications in Statistics - Simulation and Computation*, 36(6):1171–1186.
- [Newcombe, 1998] Newcombe, R. G. (1998). Two-sided confidence intervals for the single proportion: comparison of seven methods. *Statistics in Medicine*, 17(8):857–872.
- [Newcombe, 2001] Newcombe, R. G. (2001). Logit Confidence Intervals and the Inverse Sinh Transformation. *The American Statistician*, 55(3):200–202.
- [Piegorisch, 2004] Piegorisch, W. W. (2004). Sample sizes for improved binomial confidence intervals. *Computational Statistics & Data Analysis*, 46(2):309–316.
- [Pires and Amado, 2008] Pires, A. M. and Amado, C. (2008). Interval Estimators for a Binomial Proportion: Comparison of Twenty Methods. *REVSTAT-Statistical Journal*, pages 165–197 Pages. Artwork Size: 165–197 Pages Publisher: REVSTAT-Statistical Journal.
- [Sato, 2021] Sato, S. (2021). Plasmodium—a brief introduction to the parasites causing human malaria and their basic biology. *Journal of Physiological Anthropology*, 40(1):1.
- [Sepúlveda and Drakeley, 2015] Sepúlveda, N. and Drakeley, C. (2015). Sample size determination for estimating antibody seroconversion rate under stable malaria transmission intensity. *Malaria Journal*, 14(1):141.
- [Sepúlveda et al., 2015a] Sepúlveda, N., Paulino, C. D., and Drakeley, C. (2015a). Sample size and power calculations for detecting changes in malaria transmission using antibody seroconversion rate. *Malaria Journal*, 14(1):529.
- [Sepúlveda et al., 2015b] Sepúlveda, N., Stresman, G., White, M. T., and Drakeley, C. J. (2015b). Current Mathematical Models for Analyzing Anti-Malarial Antibody Data with an Eye to Malaria Elimination and Eradication. *Journal of Immunology Research*, 2015:1–21.
- [Sievert, 2020] Sievert, C. (2020). *Interactive Web-Based Data Visualization with R, plotly, and shiny*. Chapman and Hall/CRC.
- [Smithson, 2003] Smithson, M. (2003). *Confidence intervals*. Number no. 07/140 in Sage university papers. Quantitative applications in the social sciences. Sage Publications, Thousand Oaks, Calif.
- [Thulin, 2014] Thulin, M. (2014). The cost of using exact confidence intervals for a binomial proportion. *Electronic Journal of Statistics*, 8(1).
- [Vos and Hudson, 2005] Vos, P. W. and Hudson, S. (2005). Evaluation Criteria for Discrete Confidence Intervals: Beyond Coverage and Length. *The American Statistician*, 59(2):137–142.

## BIBLIOGRAFIA

[Wickham, 2016] Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

[World Health Organization, 2023] World Health Organization (2023). *World Malaria Report 2023*. World Health Organization, Geneva.