

**Universidade de São Paulo
Instituto de Matemática e Estatística**

Centro de Estatística Aplicada

Relatório de Análise Estatística

RAE-CEA-24P26

RELATÓRIO DE ANÁLISE ESTATÍSTICA SOBRE O PROJETO:

“Desigualdade no acesso à saúde materna em Moçambique”

Bruno Groper Morbin

Carlos Alberto Silva de Souza

Francisco Felipe de Queiroz

São Paulo, novembro de 2024

CENTRO DE ESTATÍSTICA APLICADA - CEA – USP

TÍTULO: Relatório de Análise Estatística sobre o Projeto: “Desigualdade no acesso à saúde materna em Moçambique”.

PESQUISADOR: Manuel Jorge Mahoche

ORIENTADOR: Prof. Dr. Eliseu Alves Waldmann

INSTITUIÇÃO: Faculdade de Saúde Pública da USP

FINALIDADE DO PROJETO: Outra - Epidemiologia

RESPONSÁVEIS PELA ANÁLISE: Bruno Groper Morbin
Carlos Alberto Silva de Souza
Francisco Felipe de Queiroz

REFERÊNCIA DESTE TRABALHO: MORBIN, B.G.; SOUZA, C.A.S.; QUEIROZ, F.F. **Relatório de análise estatística sobre o projeto: “Desigualdade no acesso à saúde materna em Moçambique”.** São Paulo, IME-USP, 2024. (RAE-CEA-24P26)

FICHA TÉCNICA

REFERÊNCIAS BIBLIOGRÁFICAS:

DEPARTAMENTO DE ANÁLISE EPIDEMIOLÓGICA E VIGILÂNCIA DE DOENÇAS NÃO TRANSMISSÍVEIS. **Indicadores da Saúde do Ministério da Saúde e Determinantes Sociais: Objetivo 03 - Saúde e bem-estar**. Disponível em: <https://svs.aids.gov.br/daent/acesso-a-informacao/acoes-e-programas/ods/indicadores-da-saude/objetivo-03/>. Acesso em: 07 out. 2024.

INSTITUTO NACIONAL DE ESTATÍSTICA (INE) – **Projeção da População 2017-2050**. 2017. Disponível em: <https://www.ine.gov.mz/censo-2017>. Acesso em 29 de setembro de 2024. [não citado no texto... verificar se vamos citar ou retirar]

INSTITUTO NACIONAL DE ESTATÍSTICA (INE) e ICF. **Inquérito Demográfico e de Saúde em Moçambique 2022-23**. Maputo, Moçambique e Rockville, Maryland, EUA: INE e ICF, 2024. Disponível em: <https://dhsprogram.com/pubs/pdf/FR389/FR389.pdf>. Acesso em: 07 out. 2024.

INSTITUTO NACIONAL DE ESTATÍSTICA (INE) e MACRO INTERNATIONAL INC. **Inquérito Demográfico e de Saúde – 1997**. Maputo, Moçambique e Calverton, Maryland, EUA: INE e MACRO, 1998. Disponível em: <https://dhsprogram.com/pubs/pdf/FR92/FR92.pdf>. Acesso em: 07 out. 2024.

LOHR, S. L. **Sampling: Design and Analysis**. 2nd ed. Boca Raton: Chapman & Hall/CRC, 2009. (Chapman & Hall/CRC Texts in Statistical Science).

LUMLEY, T.. **Weights in statistics**, 2020. Disponível em: <https://notstatschat.rbind.io/2020/08/04/weights-in-statistics/>. Acesso em: 17 nov. 2024.

LUMLEY, T. **Survey: analysis of complex survey samples**. R package version 4.4. 2024.

MULEVA, B. R. **Qualidade da assistência pré-natal em Nampula, Moçambique**. Dissertação (Mestrado em Enfermagem em Saúde Coletiva) - Escola de Enfermagem, Universidade de São Paulo, São Paulo, 2020. doi:10.11606/D.7.2020.tde-25022021-132650. Acesso em: 7 out. 2024.

PAULA, G. A. **Modelos de Regressão com apoio computacional**. São Paulo: Instituto de Matemática e Estatística, Universidade de São Paulo, 2024. Disponível em: https://www.ime.usp.br/~giapaula/texto_2024.pdf. Acesso em: 19 nov. 2024.

PESSOA, D. G. C.; SILVA, P. L. N.. **Análise de dados amostrais complexos**, 2018. Disponível em: <https://djalmapessoa.github.io/adac/>. Acesso em: 16 nov. 2024.

PROGRAMAS COMPUTACIONAIS UTILIZADOS:

Google Colab – Jupyter Notebook – (versão 2024-02-21)

Microsoft Word versão 2403

Python versão 3.10.12

Rstudio (versão 4.4.0)

TÉCNICAS ESTATÍSTICAS UTILIZADAS

Análise Descritiva Unidimensional (03:010)

Análise Descritiva Multidimensional (03:020)

Análise de Dados Categorizados (06:030)

Regressão Logística (07:090)

Amostragem Complexa (10:990 – outros tipos de amostragem)

ÁREA DE APLICAÇÃO

Medicina/Epidemiologia (14:040)

RESUMO

A taxa de mortalidade materna é um indicador crucial do desenvolvimento social e econômico de um país, refletindo a eficácia dos cuidados de saúde. Este estudo investiga como a qualidade dos cuidados pré-natais (CPN) em Moçambique é influenciada por fatores sociais, econômicos e demográficos, utilizando dados do Inquérito Demográfico de Saúde 2022-2023. Através da aplicação de um modelo de regressão logística, foram identificados os principais determinantes que contribuem para uma melhor qualidade dos CPN (com destaque para o nível educacional e cobertura de seguro-saúde da mãe), subsidiando uma possível discussão de políticas públicas voltadas à saúde materna.

Sumário

1. Introdução.....	10
2. Objetivos.....	11
3. Descrição do estudo.....	11
4. Descrição das variáveis.....	12
4.1 Variável resposta.....	12
4.2 Variáveis explicativas.....	13
5. Análise descritiva.....	14
5.1 Variável resposta.....	15
5.2 Variáveis explicativas.....	16
5.3 Associação entre as variáveis explicativas e os componentes pré-natais.....	17
6. Ajuste do modelo.....	17
6.1 Variáveis explicativas utilizadas nos ajustes.....	18
6.2 Regressão logística com ajuste incorporando o plano amostral.....	18
6.2.1 Por que incorporar o plano amostral?.....	18
6.2.2 Por que regressão logística?.....	19
6.2.3 Especificação do modelo.....	20
6.2.4 Estimação dos parâmetros.....	21
6.2.5 Seleção de modelos.....	21
6.2.6 Análise de diagnóstico.....	24
6.2.7 Resultados e interpretação dos coeficientes.....	27
7. Conclusão.....	32
APÊNDICE A.....	33
Tabelas da seleção do modelo.....	34
Tabelas da análise de diagnóstico do modelo.....	37
APÊNDICE B.....	45
Características gerais do conjunto de dados original: Variável resposta.....	46
Características gerais do conjunto de dados original: Variáveis explicativas.....	51
Relação entre a quantidade de CPN e as variáveis explicativas.....	53
Proporção de respondentes em cada nível das variáveis explicativas.....	63
Análise de diagnóstico do modelo.....	66

1. Introdução

A saúde materna é um indicador crucial do desenvolvimento social e econômico de um país, refletindo não apenas a qualidade dos serviços de saúde, mas também as condições sociais, econômicas e culturais que afetam as mulheres. Em Moçambique, a mortalidade materna continua sendo uma preocupação significativa, com taxas que superam 200 mortes por 100.000 nascidos vivos (INE e ICF, 2024, p.381), muito acima da meta global de menos de 70 mortes por 100.000 até 2030 estabelecida pelos Objetivos de Desenvolvimento Sustentável (ODS) da ONU (Departamento de Análise Epidemiológica e Vigilância de Doenças Não Transmissíveis). A qualidade do atendimento pré-natal é um fator determinante para a redução dessa mortalidade, pois está diretamente relacionada à prevenção de complicações durante a gravidez e ao parto (Muleva, 2020).

Neste contexto, o presente estudo utiliza indicadores de qualidade do atendimento pré-natal em Moçambique, considerando aspectos como aferição da pressão arterial, administração de medicamentos contra parasitas intestinais, malária, realização de testes laboratoriais e administração de medicamentos anti-HIV. Esses indicadores são fundamentais para avaliar a adequação do cuidado recebido pelas gestantes e como isso se relaciona com os resultados maternos e neonatais (Muleva, 2020, p.33-34).

Além disso, fatores sociais e econômicos como cobertura de um seguro de saúde, nível educacional e província de ocupação aparentam influenciar a qualidade do atendimento pré-natal recebido. Segundo o Inquérito Demográfico de Saúde (IDS) relativo ao ano de 1997, em áreas urbanas de Moçambique, a cobertura pré-natal já era quase universal (96%), enquanto nas áreas rurais essa taxa cai para 65% (INE e MACRO, 1998, p.131). Este estudo busca relacionar esses indicadores sociais com a qualidade dos tratamentos pré-natais, contribuindo para uma compreensão mais profunda das barreiras enfrentadas pelas mulheres moçambicanas na busca por

cuidados adequados durante a gestação. Para tanto, foram utilizados os dados amostrais coletados pelo IDS do ano de 2022-23.

2. Objetivos

O estudo tem como objetivo identificar diferentes características individuais, comunitárias ou provinciais que influenciam na qualidade dos cuidados pré-natais às mulheres de Moçambique.

3. Descrição do estudo

A amostra é constituída por mulheres de 15 a 49 anos que, no momento da entrevista do IDS 2022/23, tiveram ao menos um filho (quer estivesse vivo ou não no momento da entrevista). O tamanho amostral desse público-alvo é de 9.956 mulheres do total de 13.183 entrevistadas pelo IDS.

O esquema amostral realizado no IDS foi dividido em dois estágios. No primeiro estágio foram selecionados 619 conglomerados (*clusters*), chamados de Áreas de Enumeração (AE), com probabilidades proporcionais ao número de Agregados Familiares (AF) de acordo com o Censo de 2017. No segundo estágio, para cada AE, foram selecionados 26 AF de forma aleatória. Dessa forma, foram selecionados 16.045 AF (inferior aos 16.094 ideais por problemas de segurança no recenseamento). Todas as mulheres de 15 a 49 anos que eram residentes habituais ou visitantes na noite anterior à entrevista no AF foram elegíveis a serem entrevistadas, enquanto numa subamostra de 50% dos AF todos os homens de 15 a 54 anos foram elegíveis a serem entrevistados. Como informado anteriormente, entretanto, apenas os questionários das mulheres foram utilizados neste estudo.

4. Descrição das variáveis

Os dados foram obtidos por meio de questionários-padrão do IDS (INE e ICF, 2024, p. 605-808). Em especial, os dados filtrados para o público-alvo desejado foram fornecidos pelo pesquisador e eles contemplaram diversas variáveis demográficas, sociais, geográficas, econômicas, clínicas e indicadores de exames laboratoriais, descritas a seguir.

4.1 Variável resposta

Todas as variáveis que compõem a variável resposta (quantidade de componentes pré-natais) referem-se a todas as gestações anteriores: isto é, se recebeu o componente em qualquer uma das gestações, a resposta é “sim”; se não recebeu o componente em nenhuma gestão, a resposta é “não”. Os componentes considerados são:

- **Suplementação de ferro:** sim (1); não (0)
- **Medição da pressão arterial:** sim (1); não (0)
- **Realização de exame de sangue:** sim (1); não (0)
- **Realização de exame de urina:** sim (1); não (0)
- **Ingestão de medicamentos contra malária:** sim (1); não (0)
- **Uso de medicamento para prevenção de transmissão de HIV ao bebê:** sim (1); não (0)
- **Ingestão de medicamentos contra parasitas intestinais:** sim (1); não (0)

4.2 Variáveis explicativas

Todas as variáveis explicativas referem-se ao momento da entrevista. São elas:

- **Faixa etária (em anos):** 15 a 19 anos, 20 a 29 anos, 30 a 39 anos ou 40 a 49 anos
- **Nível educacional:** Sem Educação, Primário, Secundário ou Superior
- **Educação do cônjuge:** Sem Escolaridade, Primário, Secundário ou Superior
- **Exposição à mídia:** número de mídias que a mulher possui exposição pelo menos uma vez por semana, considerando jornal/revista, rádio e televisão. Os valores possíveis são inteiros e vão de zero (nenhuma mídia) a três (três mídias)
- **Categoria de renda:** Baixa renda (dois menores quintis de renda), Média renda (terceiro quintil de renda) ou Alta renda (dois maiores quintis de renda)
- **Decisão de saúde:** quem toma as decisões sobre a saúde da mulher. Possui os valores: Sozinha, Respondente e cônjuge, Cônjuge sozinho ou outra pessoa
- **Intervalo médio entre nascimentos das gestações precedentes (em meses)**
- **Número de filhos:** 1 a 2, 3 a 4 ou mais de 4
- **Estado civil:** Solteira, Casada ou Viúva/Divorciada
- **Religião:** Católica, Cristã, Islâmica ou Sem religião/outras
- **Mobilidade familiar:** número de anos desde a última mudança de casa. Assume os valores: menos de 5 anos, entre 5 e 14 anos ou 15 anos ou mais, sempre na mesma residência e visitante.
- **Desejo de ter filhos:** desejo de ter filhos no presente, desejo de ter filhos mais tarde, desejo de não ter mais filhos
- **Sexo de chefe de agregado:** Masculino ou Feminino

- **Cobertura do seguro de saúde:** Segurada ou Não segurada
- **Local dos cuidados pré-natais:** Institucional ou Não institucional
- **Qualificação do cuidador do CPN:** Qualificado ou Não qualificado
- **Local de residência:** Urbano ou Rural
- **Província:** Cidade de Maputo, Maputo, Gaza, Inhambane, Manica, Sofala, Tete, Zambézia, Nampula, Niassa e Cabo Delgado.
- **Escore de desvantagem socioeconômica:** primeiro componente principal de três variáveis que são agregações referentes à área de enumeração. São elas: taxa de pobreza (proporção de mulheres com a variável “Categoria de renda” possuindo o valor “Baixa renda”), taxa de desemprego (proporção de mulheres com a variável “Situação do emprego” possuindo o valor “desempregada”) e taxa de falta de exposição à mídia (proporção de mulheres com a variável “Exposição à mídia” possuindo o valor zero)

5. Análise descritiva

Nesta seção, é apresentada a análise descritiva dos dados. Primeiramente são analisados os componentes formadores da variável resposta do estudo e em seguida as variáveis explicativas, além da associação entre a resposta e as outras variáveis.

Iniciamos com a análise de respostas faltantes, seja por omissão ou pela declaração explícita de desconhecimento, e observamos que 48,4% das mulheres entrevistadas não responderam a respeito de ter possuído suplementação de ferro, 50,8% não responderam a respeito de uso do medicamentos contra parasitas intestinais e 51,9% para medicamentos contra malária nos cuidados pré-natais. Além disso, 52,9% das mesmas não responderam se houve medição de pressão arterial, enquanto que 52,6% não responderam a respeito de exame de urina e exame de

sangue, conforme Figura B.1. Tais observações nos levaram a desconsiderar as entrevistadas com ao menos uma dessas variáveis-resposta não preenchidas, resultando em uma base final para análise de 3.813 respondentes.

Vale destacar o fato de que nem todas as mulheres seriam elegíveis para o uso de certos medicamentos como contra HIV, malária e contra parasitas intestinais por não possuírem a respectiva doença. Portanto, a ausência desses componentes não implica necessariamente que não houve assistência pré-natal completa a depender da situação da mulher perante cada uma dessas doenças. Entretanto, por incompletude do próprio questionário do IDS, essa questão não foi tratada. Em um futuro estudo, com o intuito de realizar conclusões mais confiáveis, o público-alvo deveria ser restrito apenas a mulheres elegíveis a tal ou qual componente de cuidado pré-natal.

5.1 Variável resposta

Aplicando os pesos populacionais do desenho amostral do conjunto de dados, analisamos a proporção de respondentes da população de mulheres que já tiveram algum filho e que não deixaram de responder nenhuma das componentes pré-natais que compõem a variável resposta.

Conforme os gráficos das Figuras B.2-B.8, observamos a proporção estimada de mulheres e seu respectivo erro padrão¹ para cada componente. Destaca-se o fato de que as proporções estimadas de mulheres que fizeram suplementação de ferro e exame de sangue foram, respectivamente, 90,3% e 88%. Já a proporção estimada de entrevistadas que não realizaram ingestão de medicamentos contra parasitas intestinais foi de 48,9%, apresentando uma ligeira predominância na falta do uso deste medicamento.

¹ Os erros-padrão são calculados considerando os métodos de estimação para médias e proporções (para amostragem complexa) descritos em Pessoa e Silva (2018, Capítulo 4). Esses métodos estão implementados nas funções *svyratio* e *svymean*, do pacote *survey* do *software R* (LUMLEY, 2024).

Observamos pela Figura B.9 a distribuição da soma das componentes pré-natais, possuindo um pico na contagem de 7 componentes e com uma forte assimetria, com a massa de probabilidade concentrada nos maiores números de componentes. Em especial, o grupo com menos de 4 componentes pré-natais representa menos de 20% das entrevistadas do total.

5.2 Variáveis explicativas

Observamos que há alguns dados faltantes para as variáveis explicativas entre as 3.813 respondentes pela Figura B.10. Contudo, percebe-se que a variável Intervalo médio entre gestações possui valores faltantes para o caso em que as entrevistadas possuíam apenas um filho. Logo, para essa variável, os valores faltantes foram preenchidos com o valor de zero. As demais variáveis foram preenchidas como categoria “*Missing*” para análise do comportamento desses dados.

As variáveis Qualificação do cuidador do CPN, Local do CPN e Desejo de ter filhos aparecem com alto índice de concordância² de respostas faltantes, enquanto Decisor de cuidados de saúde e Educação do Cônjuge apresentam concordância total de respostas faltantes, conforme Figura B.11.

Verificamos também a distribuição do número de respondentes na base de análise por província, bem como a estimativa da proporção populacional e respectivo erro padrão apresentado na Figura B.12. Nela é possível entender os pesos populacionais e sua influência nas proporções populacionais estimadas. Por exemplo, Cabo Delgado possui o maior número absoluto de respondentes, sendo elas 575,

² A taxa de concordância de valores faltantes é calculada via $(\text{concordantes-discordantes})/\text{total}$, em que “concordantes” é o número de linhas em que há valores faltantes nas duas colunas ou valores preenchidos nas duas colunas e “discordantes” é o número de linhas em que há um valor faltante em uma coluna e um preenchido na outra.

apesar de Nampula ser a província com maior peso populacional, correspondendo a 29,2% da população estimada da base de análise.

5.3 Associação entre as variáveis explicativas e os componentes pré-natais

Analisando os *box plots* relacionando a contagem de componentes pré-natais com cada variável explicativa, podemos observar nas Figuras B.13-B.17 que algumas delas não apresentam diferença de mediana de categoria para categoria, sendo elas: Número de filhos, Desejo de mais filhos, Mobilidade familiar, Qualificação do cuidador, Intervalo Médio entre gestações.

Já as variáveis explicativas que se destacaram na associação com a contagem de componentes pré-natais por apresentarem mediana diferente entre suas categorias, vide Figuras B.18-B.31, são listadas a seguir: Faixa etária, Estado Civil, Decisor de Cuidados de Saúde, Chefe do agregado, Exposição a mídias, Religião, Província, Local do CPN, Local de Residência, Categoria de renda, Educação do cônjuge, Nível educacional, Cobertura de seguro, Escore de desvantagem socioeconômica.

6. Ajuste do modelo

Neste estudo, será adotado o modelo de regressão logística considerando um plano amostral complexo, tendo a quantidade de componentes de CPN (Cuidados Pré-Natais) como uma variável binária. Essa metodologia foi escolhida para compreender e estimar a influência das variáveis explicativas no acesso aos cuidados pré-natais.

6.1 Variáveis explicativas utilizadas nos ajustes

Conforme a análise descritiva indicou, ficaram fora dos ajustes de modelos as variáveis que apresentaram pouca diferença de mediana (da quantidade de componentes de CPN) de categoria para categoria. A análise univariada dessas variáveis, com a proporção populacional estimada e quantidade de respostas, é apresentada nas Figuras B.32-B.36.

Das demais variáveis, a variável Educação do cônjuge (cuja análise univariada encontra-se na Figura B.37) não fez parte dos ajustes pois é fortemente associada com a variável Nível educacional.

Por fim, para simplicidade da análise, algumas das variáveis finais tiveram as suas categorias unificadas:

- Faixa etária: “entre 20 e 39 anos” e “abaixo de 20 ou acima de 39 anos”
- Decisão de saúde: “respondente e cônjuge ou sozinha” e “outra pessoa ou resposta omitida”.
- Exposição à mídia: “acesso a 1 ou mais mídias” e “sem acesso às mídias”
- Província: “mora na Cidade de Maputo ou em Maputo” e “mora em outra província do país”
- Nível educacional: “educação até nível secundário, podendo ser primário ou sem ensino também” e “educação superior”

6.2 Regressão logística com ajuste incorporando o plano amostral

6.2.1 Por que incorporar o plano amostral?

Embora as estimativas pontuais obtidas com o uso da regressão logística para amostragem com pesos amostrais sejam numericamente equivalentes às estimadas por

métodos que incorporam a amostragem complexa (Lumley, 2020), a variância desses estimadores – e, conseqüentemente, a significância estatística aferida por meio de testes – apresenta diferenças.

Uma maneira de quantificar essas discrepâncias é por meio do Efeito do Plano Amostral (EPA), conforme definido por Kish (Pessoa e Silva, 2018, Capítulo 4). Levando isso em consideração, optou-se por realizar os ajustes necessários para incorporar a estrutura de amostragem complexa ao modelo, garantindo maior precisão e representatividade nas análises estatísticas.

6.2.2 Por que regressão logística?

A aplicação de modelos lineares generalizados em que os procedimentos de estimação incorporam o plano amostral é um campo ainda em estudo. Por conta disso, muitas metodologias ainda estão em desenvolvimento e, conseqüentemente, não há muitas possibilidades em termos de implementação computacional. A ferramenta mais abrangente e completa para análise de amostragem complexa disponível no momento da redação deste estudo é a biblioteca *survey* do *software R*.

Para a abordagem original do pesquisador (Quantidade de componentes de CPN), seria necessária uma família de distribuições com suporte inteiro e finito (pois o número de componentes vai de 0 a 7), para a qual, com base no nosso conhecimento, ainda não existe uma implementação para amostragem complexa. Alternativamente, poderia se tomar a proporção de componentes de CPN, porém, ainda assim, seria necessária uma distribuição discreta e finita – que também não está implementada.

Portanto, a discretização da variável resposta em nível alto (igual ou acima de 4 componentes) ou baixo (igual ou menor que 3 componentes) serviu como adaptação entre os objetivos do pesquisador e a disponibilidade de ferramentas para obter resultados mais confiáveis (i.e. utilização de amostragem complexa). Iremos nos referir, daqui em diante, a nível (alto ou baixo) de CPN para essa variável binarizada.

Nota-se, ainda, que a definição do ponto de corte que separa esses dois grupos baseia-se no fato de que apenas 4 dos 7 componentes são obrigatórios para todas as gestantes. Conforme destacado anteriormente na análise descritiva, os outros 3 componentes – administração de medicamentos contra o HIV, contra a malária e contra parasitas intestinais – são recomendados apenas em situações específicas, quando há indicação clínica devido à presença de determinadas condições ou doenças.

6.2.3 Especificação do modelo

A especificação do modelo de regressão logística para amostragem complexa é a mesma que para amostragem aleatória simples (Pessoa e Silva, 2018, Capítulo 6). Tal modelo é representado por uma parte aleatória e uma parte sistemática, conforme equações (1) e (2), respectivamente, a seguir:

A equação (1) descreve a distribuição de probabilidade do nível de CPN para a i -ésima respondente, dados os valores das covariáveis (variáveis explicativas):

$$Y_i \sim \text{Bernoulli}(p_i), \quad (1)$$

em que $Y_i = 1$ se a i -ésima respondente tem nível alto de CPN e 0 caso contrário, para $i=1, \dots, n$.

$$\log \left(\frac{p_i}{1-p_i} \right) = \beta_0 + \sum_{k=1}^r \beta_k x_{ik} \quad (2)$$

Logo, temos que:

- $i \in \{1, \dots, n\}$ é o índice da respondente e n é o tamanho da amostra;
- Y_i é a indicadora de nível alto de CPN para a i -ésima respondente;
- p_i é a probabilidade de nível alto de CPN para a i -ésima respondente;

- $x_{i1}, x_{i2}, \dots, x_{ip}$ é o vetor de valores das variáveis para a i -ésima respondente. Por exemplo, x_{i1} pode representar o Escore de desvantagem socioeconômica da i -ésima respondente, x_{i2} representar a indicadora de Local de residência urbano (1 se reside em local urbano, 0 se reside em local rural), e x_{i3} representar o nível de Exposição a mídias (1 se alta exposição, 0 se baixa exposição);
- β_k , $k = 0, \dots, r$ são os coeficientes associados às covariáveis para as r covariáveis consideradas, sendo β_0 o intercepto.

6.2.4 Estimação dos parâmetros

O estudo em questão objetiva identificar o peso (tecnicamente conhecido como estimativa) dos parâmetros que representam as covariáveis para inferir sobre o nível de CPN.

As estimativas dos parâmetros coincidem com os pontos que maximizam a função de verossimilhança conforme desenvolvido em Lohr (2009, Capítulo 11), bastando apenas considerar os pesos amostrais de cada observação, ou seja, ponderar os fatores da verossimilhança com o respectivo peso na população. Já para a estimação da variância dos estimadores dos coeficientes (erros-padrão), o método utilizado em amostragem complexa é baseado em Linearização de Taylor (Lohr, 2009, Capítulo 9) e modelos de superpopulação (Pessoa e Silva, 2018, Capítulo 3).

6.2.5 Seleção de modelos

Nota-se de início que o conjunto de dados usado para o ajuste apresenta ao todo 3.813 mulheres e cada nível da variável resposta binarizada é composto por 534 para

as mulheres que não possuíram 4 ou mais cuidados pré-natais, e 3.279 para as que possuíram.

De acordo com Paula (2024), “[...] a seleção de um modelo logístico deve ser um processo conjugado de seleção estatística de modelos e bom senso”. Dessa forma, foi pensado em considerar um primeiro ajuste com todas as covariáveis previamente selecionadas pela análise descritiva envolvidas, chamado de modelo completo.

Tabela 1: Modelo completo.

Coeficiente	Estimativa	Erro Padrão	Valor-p
(Intercepto)	0,8701	0,2940	0,003
Faixa Etária _ 20-39	0,1583	0,1404	0,260
Nível Educacional _ Superior	13,2064	0,2450	<0,001
Exposição à Mídia _ 1+	0,2897	0,1425	0,043
Categoria de Renda _ Alta Renda	0,6116	0,2067	0,003
Chefe do Agregado _ Masculino	-0,1719	0,1487	0,248
Cobertura de Saúde _ Segurado	13,3685	0,3281	<0,001
Local do CPN _ Não Institucional	-0,5768	0,1894	0,002
Local de Residência _ Urbano	0,1803	0,2485	0,468
Província _ Maputo	2,8139	0,6286	<0,001
Decisor de Saúde _ Respondente	0,3175	0,1533	0,039
Escore de Desvantagem Socioeconômica	0,8437	0,6165	0,172

Nota-se, na Tabela 1, que algumas covariáveis não são significativas. São elas (junto ao valor-p):

- Local de Residência (valor-p de 0,468)
- Faixa Etária (valor-p de 0,260)
- Chefe do Agregado (valor-p de 0,248)
- Escore de Desvantagem Socioeconômica (valor-p de 0,172)

Portanto, considerou-se utilizar o método *backward* (Paula, 2024) para seleção do modelo adequado, considerando um nível de significância de 5%. Os passos do método utilizado e os dados de cada modelo estão apresentados nas Tabelas A.1 a A.3 além das Tabelas 1 e 3 com o passo inicial (modelo completo) e final (modelo selecionado) do algoritmo. Sendo assim, tem-se, na Tabela 2, as métricas apresentadas para o modelo de cada passo do método.

Tabela 2: Métricas dos modelos do método *backward*.

Modelo	AIC	<i>Deviance</i>	<i>Null Deviance</i>	AUC
Completo	3654.471	3294.461	3584.68	0.6755
Passo 1	3653.378	3296.271	3584.68	0.6761
Passo 2	3653.194	3298.497	3584.68	0.6750
Passo 3	3652.622	3300.977	3584.68	0.6733
Passo 4	3645.768	3305.311	3584.68	0.6685

A métrica AIC (*Akaike's Information Criterion*) contribui para observar a verossimilhança do modelo com penalização para a quantidade de parâmetros utilizados; Portanto quanto menor, melhor o modelo com menos parâmetros. Já para a métrica AUC (*Area Under the Curve*), o objetivo é analisar o inverso: quanto maior, melhor o modelo de classificação consegue discriminar os dados e trazer valores de especificidade e sensibilidade.

Logo, foi considerado o modelo ajustado do passo 4 com os resultados apresentados na Tabela 3. Portanto, com a remoção da variável Escore de Desvantagem Socioeconômica do modelo do Passo 3.

Tabela 3: Modelo de regressão logística ajustado selecionado

Coeficiente	Estimativa	Erro Padrão	Valor-p
(Intercepto)	1,1537	0,1789	<0,001
Nível Educacional _ Superior	13,2347	0,2332	<0,001
Exposição à Mídia _ 1+	0,2790	0,1413	0,049
Categoria de Renda _ Alta Renda	0,6008	0,1631	<0,001
Cobertura de Saúde _ Segurado	13,3503	0,3142	<0,001
Local do CPN _ Não Institucional	-0,5582	0,1857	0,003
Província _ Maputo	2,7803	0,6266	<0,001
Decisor de Saúde _ Respondente	0,2961	0,1482	0,046

Observa-se que esse último modelo ajustado possui todos coeficientes significativos a nível de 5% observando os respectivos valores-p. Além disso, comparando as métricas dos modelos de cada passo do algoritmo, este último apresentou menor AIC. Dessa forma, segue para a análise de diagnóstico e interpretação dos resultados.

6.2.6 Análise de diagnóstico

Para análise de diagnóstico em modelos generalizados lineares com amostragem complexa, encontramos uma escassez de referências na literatura

(Pessoa e Silva, 2018), contudo é possível realizar análises com o teste de Wald (que traduzem no valor-p apresentado aos coeficientes), além de distância de Cook para identificar possíveis observações de destaque e outras formas avaliadas no pacote *svydiags* do *software R*.

Inicialmente nota-se pelo gráfico da curva ROC, Figura B.38, que considerando um ponto de corte no valor predito de 0,81, poderia-se ter uma relação sensibilidade-especificidade melhor. Ou seja, considerando todos valores abaixo de 0,81 com classificação de “menos de 4 CPN” e acima desse número com classificação “mais de 4 CPN”.

Observa-se também pela Tabela A.4 os valores de Fator de Inflação da Variância (VIF) para cada componente do modelo selecionado. Esses valores representam uma derivação do R^2 -ajustado da regressão linear de um componente com os demais. A análise disso sugere que para valores de VIF acima de 10, pode-se considerar haver multicolinearidade de um componente com os demais. Nota-se, portanto, que para os componentes analisados, não há nenhuma indicação de presença de multicolinearidade.

Para análise dos resíduos, inicialmente visualiza-se na Figura B.39 a disposição do desenho amostral complexo que indica em qual estrato³ e conglomerado (área de enumeração) a observação foi encaixada. Logo, verifica-se na Figura B.40 a correlação entre as observações, o que parece seguir o desenho amostral.

Dessa forma, a Figura B.41 mostra os resíduos padronizados com destaque para os outliers expostos na Tabela A.5. Nota-se, portanto, que esses *outliers* apresentam perfis em comum:

- Possuem menos de 4 CPN
- Nível educacional até o secundário

³ Os estratos possuem uma relação de quase dois estratos por província. Para a relação exata, verificar a Tabela A.14.

- Sem cobertura de seguro de saúde

Porém,

- Moradoras de Maputo ou Cidade de Maputo
- Com realização dos cuidados pré-natais em locais institucionais
- Expostas a uma ou mais mídias
- *Status* financeiro de alta renda

Observe que são as próprias responsáveis pelas decisões dos cuidados de saúde, com exceção da observação de índice 3489. Logo, esses *outliers* demonstram uma distorção nas estimativas dessas últimas variáveis listadas com a variável resposta.

Com a remoção desses 3 pontos *outliers*, e usando os mesmos parâmetros do modelo selecionado, chega-se em um novo ajuste, apresentado na Tabela 4. As métricas desse ajuste (AUC, *Deviance*, *Null Deviance* e AIC) encontram-se na Tabela A.15.

Tabela 4: Modelo de regressão logística selecionado e ajustado com a remoção dos *outliers* da Tabela A.5.

Coeficiente	Estimativa	Erro Padrão	Valor-p
(Intercepto)	1,1532	0,1793	<0,001
Nível Educacional _ Superior	15,0955	0,2437	<0,001
Exposição à Mídia _ 1+	0,2837	0,1419	0,046
Categoria de Renda _ Alta Renda	0,6033	0,1639	<0,001
Cobertura de Saúde _ Segurado	15,2854	0,3207	<0,001
Local do CPN _ Não Institucional	-0,5585	0,1859	0,003

	Coeficiente	Estimativa	Erro Padrão	Valor-p
Província _ Maputo		16,2811	0,1538	<0,001
Decisor de Saúde _ Respondente		0,2935	0,1487	0,049

Portanto, nota-se uma mudança significativa na estimativa do parâmetro relacionado à variável explicativa Província que a mulher vive.

Para análise dos pontos de alavanca, o objetivo é identificar a influência do ponto sobre o valor preditor dele mesmo, de acordo com Paula (2024). Portanto, identifica-se esses pontos de alavanca na Figura B.42 com detalhamento do componente das observações na Tabela A.6.

No mesmo sentido, verifica-se os pontos a partir da métrica DFFITS (Tabela A.7) que indica a influência sobre os parâmetros ajustados no modelo. Já para identificar a relação direta de cada observação diante cada parâmetro, observa-se a medida DFBETAS pela Figura B.43 e Tabelas A.8 a A.12.

Por fim, conforme ilustrado na Figura B.44, observa-se que as maiores distâncias de Cook estão associadas aos perfis descritos na Tabela A.13. Esses perfis incluem mulheres com nível educacional até o secundário, sem cobertura de saúde e que residem fora de Maputo ou da Cidade de Maputo, mas que apresentam alto nível de CPN.

6.2.7 Resultados e interpretação dos coeficientes

Com o modelo ajustado apresentado na Tabela 3, pode-se medir a razão de chances para cada covariável conforme a Tabela 5.

Tabela 5: Razão de chances entre as categorias dentro de cada covariável.

Covariável	Razão de chances	95% IC ⁴
Nível Educacional		
Até secundário	—	—
Superior	559.426	353.818; 884.515
Exposição à Mídias		
0	—	—
1+	1,32	1,00; 1,74
Categoria de Renda		
Até Média Renda	—	—
Alta Renda	1,82	1,32; 2,51
Cobertura de Saúde		
Não segurado	—	—
Segurado	627.979	338.791; 1.164.014
Local do CPN		
Institucional	—	—
Não Institucional	0,57	0,40; 0,82
Província		
Outro	—	—
Maputo	16,1	4,71; 55,2
Decisor de Saúde		
Outro	—	—
Respondente	1,34	1,01; 1,80

Destacam-se os componentes Cobertura de saúde e Nível educacional, com a estimativa indicando que uma mulher com educação superior tem 559.426 vezes a chance de realizar quatro ou mais cuidados pré-natais em comparação com uma mulher com o mesmo perfil, mas sem nível de escolaridade superior.

⁴ Intervalo de Confiança.

Na mesma lógica, de acordo com esse modelo ajustado, a mulher que possui seguro de saúde apresenta uma estimativa de 627.979 vezes a chance de ter mais de 4 componentes de cuidados pré-natais que uma não segurada.

Outra variável de destaque é a de moradoras em Maputo ou Cidade de Maputo, que possuem 16 vezes a chance estimada de ter os devidos cuidados pré-parto do que as que moram nas demais províncias de Moçambique. Nota-se pela Tabela 6 que considerando o modelo ajustado com a remoção dos *outliers* e, cujas estimativas dos parâmetros já foram apresentadas na Tabela 4, chega-se em estimativas da razão de chances ainda maiores para essas três principais variáveis discutidas: Nível educacional, Cobertura de seguro de saúde e Província em que vive.

Tabela 6: Razão de chances entre as categorias dentro de cada covariável para o modelo ajustado com a remoção dos *outliers*.

Covariável	Razão de chances	95% IC
Nível Educacional		
Até secundário	—	—
Superior	3.596.688	2.228.666; 5.804.443
Exposição à Mídias		
0	—	—
1+	1,33	1,00; 1,75
Categoria de Renda		
Até Média Renda	—	—
Alta Renda	1,83	1,32; 2,52
Cobertura de Saúde		
Não segurado	—	—
Segurado	4.348.743	2.316.364; 8.164.332
Local do CPN		
Institucional	—	—
Não Institucional	0,57	0,40; 0,82

Covariável	Razão de chances	95% IC
Província		
Outro	—	—
Maputo	11.770.236	8.701.714; 15.920.823
Decisor de Saúde		
Outro	—	—
Respondente	1,34	1,00; 1,80

Vale destacar, no entanto, que após o último ajuste com a remoção dos *outliers*, a variável com maior razão de chances estimada para possuir um alto nível de CPN passou a ser o terceiro item: moradoras da Cidade de Maputo ou da província de Maputo. Sendo assim, as mulheres residentes nessas áreas apresentam uma chance estimada de 11.770.236 vezes de alcançar um alto nível de CPN em comparação com aquelas que residem em outras províncias.

Observa-se, pelas Figuras B.29 e B.30, que as variáveis explicativas com estimativas de parâmetros elevadas, e consequentemente com razões de chances altas, estão relacionadas ao fato de o nível aparentar associação direta a um elevado número de consultas pré-natais (CPN). Em outras palavras, todas as mulheres com Seguro de saúde presentes na base de dados analisada possuem 4 ou mais CPN (conforme indicado nos *box plots* das figuras mencionadas). O mesmo padrão é observado em relação ao Nível educacional superior.

Quanto às moradoras da província de Maputo ou da cidade de Maputo, ao remover os *outliers* dos *box plots*, vide Figura B.24 (o que coincide com a remoção dos *outliers* do modelo ajustado selecionado), verifica-se o mesmo cenário: todas as mulheres analisadas apresentam 4 ou mais CPN.

Conclui-se, então, que a importância de cada variável sobre o nível adequado de cuidados pré-natais segue a ordem (iniciando pela maior):

1. **Cobertura do seguro de saúde**, demonstrando que uma mulher com seguro possui uma estimativa de 627.979 vezes a chance de possuir 4 ou mais cuidados pré-natais que as mulheres não seguradas;
2. **Nível educacional superior**, em que mulheres com educação além da secundária possuem 559.426 vezes a chance estimada de alto nível de CPN daquelas que não possuem esse nível educacional;
3. **Moradoras da Cidade de Maputo ou na província de Maputo**, representando uma chance estimada de ter os devidos cuidados 16,1 vezes a chance que as que moram em outras localidades do país;
4. **Possuir *status* financeiro de alta renda**, estimando ser 1,82 vezes a chance de CPN dos demais níveis de renda;
5. **Realizar os cuidados pré-natais em locais institucionalizados** apresenta uma vantagem significativa, indicando que mulheres que não utilizam esses locais têm apenas 57% (0,57) da chance estimada de alcançar um alto nível de CPN, em comparação com aquelas que utilizam locais institucionalizados. Em outras palavras, realizar os cuidados pré-natais em locais institucionalizados aumenta a chance estimada de um alto nível de CPN em 1,75 vezes (inverso de 0,57) em relação aos cuidados realizados em locais não institucionalizados.
6. **A decisão pelos cuidados de saúde partindo da mulher** é um fator que contribui em 1,34 vezes (com intervalo de confiança de 95% sendo de 1,01 a 1,80 vezes) a chance estimada de realização de uma quantidade alta de cuidados pré-natais em relação a mulheres com outros tomadores de decisão sobre a sua saúde. Fator importante que conversa com a busca de empoderamento feminino discutida no país.
7. Por fim, **a exposição a uma ou mais mídias** contribui para o aumento de 32% da chance estimada de realizar os 4 ou mais cuidados pré-natais,

sugerindo políticas de conscientização e aderência aos meios de informação.

7. Conclusão

Apesar das variáveis Faixa etária, Chefe do agregado, Local de residência (Urbano ou Rural) e Escore de desvantagem socioeconômica se apresentarem como fatores inicialmente importantes para explicar a quantidade de cuidados pré-natais para as mulheres grávidas durante a análise descritiva, o mesmo não se confirmou ao ajustar o modelo de regressão logística para amostragem complexa.

Ainda de acordo com o mesmo modelo, as variáveis Cobertura de seguro de saúde, Nível educacional superior e residência em Maputo ou na Cidade de Maputo demonstraram ser fatores de extrema importância para o aumento das chances estimadas de acesso adequado aos cuidados pré-natais pelas mulheres em Moçambique. Esses resultados ressaltam a necessidade de implementar políticas públicas que fortaleçam e ampliem a viabilidade desses componentes para as mulheres do país. Tais medidas podem gerar um impacto significativo na redução dos índices de mortalidade pós-parto, além de contribuir para a melhoria geral da saúde materna e neonatal em Moçambique.

Uma continuação do trabalho poderia envolver: o ajuste de modelos com diferentes tipos de variáveis-resposta (ordinal, *multi-output*); a própria redefinição da variável-resposta, com a inclusão dos critérios de elegibilidade e, por fim, a inclusão da dinâmica temporal, a partir dos IDS de anos anteriores.

APÊNDICE A

Tabelas

Tabelas da seleção do modelo

Tabela A.1: Modelo ajustado no algoritmo backward para a regressão logística no Passo 1, com a remoção da variável Local de Residência do modelo completo.

Componente	Estimativa	Erro Padrão	Valor-p
(Intercepto)	0.9278	0.2715	<0.001
Faixa Etária _ 20-39	0.1574	0.1407	0.264
Nível Educacional _ Superior	13.2396	0.2408	<0.001
Exposição à Mídia _ 1+	0.2948	0.1419	0.038
Categoria de Renda _ Alta Renda	0.6948	0.1903	<0.001
Chefe do Agregado _ Masculino	-0.1695	0.1481	0.253
Cobertura de Saúde _ Segurado	13.3883	0.3235	<0.01
Local do CPN _ Não Institucional	-0.5838	0.1907	0.002
Província _ Maputo	2.8204	0.6296	<0.001
Decisor de Saúde _ Respondente	0.3099	0.1525	0.043
Escore de Desvantagem Socioeconômica	0.7310	0.5623	0.194

Tabela A.2: Modelo ajustado no algoritmo backward para a regressão logística no Passo 2, com a remoção da variável Faixa etária do modelo do Passo 1.

Componente	Estimativa	Erro Padrão	Valor-p
(Intercepto)	1.0477	0.2555	<0.001
Nível Educacional _ Superior	13.2523	0.2393	<0.001
Exposição à Mídia _	0.2979	0.1419	0.036

Componente	Estimativa	Erro Padrão	Valor-p
1+			
Categoria de Renda _ Alta Renda	0.6954	0.1905	<0.001
Chefe do Agregado _ Masculino	-0.1621	0.1491	0.277
Cobertura de Saúde _ Segurado	13.4009	0.3224	<0.001
Local do CPN _ Não Institucional	-0.5821	0.1914	0.002
Província _ Maputo	2.8205	0.6299	<0.001
Decisor de Saúde _ Respondente	0.3160	0.1515	0.037
Escore de Desvantagem Socioeconômica	0.7032	0.5615	0.211

Tabela A.3: Modelo ajustado no algoritmo backward para a regressão logística no Passo 3, com a remoção da variável Chefe do agregado do modelo do Passo 2.

Componente	Estimativa	Erro Padrão	Valor-p
(Intercepto)	0.9367	0.2653	<0.001
Nível Educacional _ Superior	13.2563	0.2380	<0.001
Exposição à Mídia _ 1+	0.2956	0.1420	0.038
Categoria de Renda _ Alta Renda	0.7020	0.1917	<0.001
Cobertura de Saúde _ Segurado	13.3885	0.3200	<0.001
Local do CPN _ Não Institucional	-0.5867	0.1916	0.02
Província _ Maputo	2.8387	0.6299	<0.001
Decisor de Saúde _ Respondente	0.2888	0.1482	0.052

Componente	Estimativa	Erro Padrão	Valor-p
Escore de Desvantagem Socioeconômica	0.7228	0.5662	0.202

Tabelas da análise de diagnóstico do modelo

Tabela A.4: VIF (*Variance Inflation Factor*) das componentes utilizadas no modelo selecionado ajustado para regressão logística de amostragem complexa.

Componente	VIF
Nível Educacional _ Superior	1.0000
Exposição à Mídia _ 1+	1.3296
Categoria de Renda _ Alta Renda	2.1896
Cobertura de Saúde _ Segurado	1.0000
Local do CPN _ Não Institucional	1.1734
Província _ Maputo	1.1053
Decisor de Saúde _ Respondente	0.9874

Tabela A.5: Pontos *outliers* com resíduos padronizados do ajuste do modelo.

Índice	Resíduo Padronizado	Mais de 3 CPN	Nível Educacional	Exposição à Mídia	Categoria de Renda	Cobertura de Saúde	Local do CPN	Província	Decisor de Saúde
3761	-32.1774	0	Até secundário	1+	Alta Renda	Não segurado	Institucional	Maputo	Respondente
3789	-32.1774	0	Até secundário	1+	Alta Renda	Não segurado	Institucional	Maputo	Respondente
3489	-23.9796	0	Até secundário	1+	Alta Renda	Não segurado	Institucional	Maputo	Outro

Tabela A.6: Pontos de alavanca destacados.

Índice	Hat	Mais de 3 CPN	Nível Educacional	Exposição à Mídia	Categoria de Renda	Cobertura de Saúde	Local do CPN	Província	Decisor de Saúde
1547	0.1587	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente

Índice	Hat	Mais de 3 CPN	Nível Educacional	Exposição à Mídia	Categoria de Renda	Cobertura de Saúde	Local do CPN	Província	Decisor de Saúde
1655	0.1277	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
1189	0.1124	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
1571	0.0846	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
1559	0.0758	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
1187	0.0705	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente
2447	0.0624	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
28	0.0621	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
1531	0.0563	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Outro
1537	0.0537	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Outro
1568	0.0531	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente
1049	0.0525	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente
473	0.0524	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Outro
1979	0.0495	1	Até secundário	1+	Até Média Renda	Segurado	Institucional	Outro	Respondente
436	0.0456	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
2482	0.0438	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
773	0.0398	1	Superior	1+	Até Média Renda	Não segurado	Institucional	Outro	Respondente
2660	0.0392	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
1976	0.0379	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente
134	0.0378	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
2394	0.0322	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Outro
2406	0.0310	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Outro
1723	0.0307	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente
1717	0.0303	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente
1779	0.0283	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente
3054	0.0268	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente
535	0.0260	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
5	0.0259	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Outro
2395	0.0248	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Outro
2858	0.0248	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Outro
2404	0.0230	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente
2861	0.0226	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Outro
1561	0.0218	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente
1564	0.0218	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente

Tabela A.7: Pontos de influência com DFFITS absolutos destacados.

Índice	DFFITS	Mais de 3 CPN	Nível Educacional	Exposição à Mídia	Categoria de Renda	Cobertura de Saúde	Local do CPN	Província	Decisor de Saúde
1547	0.6609	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
1655	0.5129	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
1189	0.4436	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
1187	0.3945	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente
1571	0.3240	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
1568	0.2917	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente

Índice	DFFITS	Mais de 3 CPN	Nível Educacional	Exposição à Mídia	Categoria de Renda	Cobertura de Saúde	Local do CPN	Provincia	Decisor de Saúde
1049	0.2880	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente
1559	0.2874	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
1531	0.2747	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Outro
1537	0.2612	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Outro
2447	0.2333	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
28	0.2322	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
1976	0.2047	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente
473	0.1721	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Outro
436	0.1674	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
773	0.1655	1	Superior	1+	Até Média Renda	Não segurado	Institucional	Outro	Respondente
1723	0.1647	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente
1979	0.1630	1	Até secundário	1+	Até Média Renda	Segurado	Institucional	Outro	Respondente
1717	0.1623	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente
2482	0.1606	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
2394	0.1530	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Outro
1779	0.1514	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente
2406	0.1471	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Outro
2660	0.1431	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
3054	0.1431	1	Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente
134	0.1377	1	Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente
1038	-0.1254	0	Até secundário	1+	Alta Renda	Não segurado	Institucional	Outro	Outro
1484	-0.1310	0	Até secundário	0	Até Média Renda	Não segurado	Institucional	Outro	Outro
1416	-0.1332	0	Até secundário	0	Até Média Renda	Não segurado	Não Institucional	Outro	Respondente
1420	-0.1332	0	Até secundário	0	Até Média Renda	Não segurado	Não Institucional	Outro	Respondente
1422	-0.1332	0	Até secundário	0	Até Média Renda	Não segurado	Não Institucional	Outro	Respondente
1186	-0.1345	0	Até secundário	1+	Até Média Renda	Não segurado	Institucional	Outro	Respondente
1399	-0.1347	0	Até secundário	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente
1040	-0.1355	0	Até secundário	1+	Alta Renda	Não segurado	Não Institucional	Outro	Respondente
1413	-0.1415	0	Até secundário	0	Até Média Renda	Não segurado	Não Institucional	Outro	Outro
1415	-0.1415	0	Até secundário	0	Até Média Renda	Não segurado	Não Institucional	Outro	Outro
1492	-0.1472	0	Até secundário	0	Até Média Renda	Não segurado	Institucional	Outro	Respondente
1185	-0.1578	0	Até secundário	0	Alta Renda	Não segurado	Institucional	Outro	Outro
1384	-0.1624	0	Até secundário	1+	Até Média Renda	Não segurado	Institucional	Outro	Outro
1473	-0.1638	0	Até secundário	1+	Alta Renda	Não segurado	Institucional	Outro	Outro
1009	-0.1652	0	Até secundário	0	Alta Renda	Não segurado	Não Institucional	Outro	Outro
1608	-0.1709	0	Até secundário	0	Até Média Renda	Não segurado	Não Institucional	Outro	Respondente
1154	-0.1740	0	Até secundário	1+	Alta Renda	Não segurado	Institucional	Outro	Outro
1158	-0.1740	0	Até secundário	1+	Alta Renda	Não segurado	Institucional	Outro	Outro
1159	-0.1740	0	Até secundário	1+	Alta Renda	Não segurado	Institucional	Outro	Outro
1286	-0.1752	0	Até secundário	1+	Até Média Renda	Não segurado	Institucional	Outro	Respondente
1287	-0.1752	0	Até secundário	1+	Até Média Renda	Não segurado	Institucional	Outro	Respondente
1398	-0.1765	0	Até secundário	0	Alta Renda	Não segurado	Institucional	Outro	Respondente
1609	-0.1816	0	Até secundário	0	Até Média Renda	Não segurado	Não Institucional	Outro	Outro
1107	-0.1825	0	Até secundário	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente
1075	-0.1876	0	Até secundário	0	Alta Renda	Não segurado	Institucional	Outro	Outro
1289	-0.1939	0	Até secundário	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente
1554	-0.1977	0	Até secundário	1+	Até Média Renda	Não segurado	Não Institucional	Outro	Outro
1176	-0.2076	0	Até secundário	0	Alta Renda	Não segurado	Institucional	Outro	Outro

Índice	DFFITS	Mais de 3 CPN	Nível Educacional	Exposição à Mídia	Categoria de Renda	Cobertura de Saúde	Local do CPN	Provincia	Decisor de Saúde
1155	-0.2099	0	Até secundário	0	Alta Renda	Não segurado	Institucional	Outro	Outro
1662	-0.2120	0	Até secundário	1+	Até Média Renda	Não segurado	Institucional	Outro	Outro
1419	-0.2284	0	Até secundário	1+	Até Média Renda	Não segurado	Não Institucional	Outro	Respondente
1421	-0.2284	0	Até secundário	1+	Até Média Renda	Não segurado	Não Institucional	Outro	Respondente
1607	-0.2488	0	Até secundário	1+	Até Média Renda	Não segurado	Institucional	Outro	Outro
1483	-0.2850	0	Até secundário	1+	Alta Renda	Não segurado	Institucional	Outro	Outro
3761	-0.3404	0	Até secundário	1+	Alta Renda	Não segurado	Institucional	Maputo	Respondente
3789	-0.5155	0	Até secundário	1+	Alta Renda	Não segurado	Institucional	Maputo	Respondente
3489	-0.7592	0	Até secundário	1+	Alta Renda	Não segurado	Institucional	Maputo	Outro

Tabela A.8: Pontos com DFBETAS absolutos destacados para o valor de intercepto.

Índice	DFBETAS	Mais de 3 CPN	Nível Educacional	Exposição à Mídia	Categoria de Renda	Cobertura de Saúde	Local do CPN	Provincia	Decisor de Saúde
1587	-0.1133	0	Até secundário	0	Até Média Renda	Não segurado	Institucional	Outro	Outro
1662	-0.1139	0	Até secundário	1+	Até Média Renda	Não segurado	Institucional	Outro	Outro
1380	-0.1239	0	Até secundário	0	Até Média Renda	Não segurado	Institucional	Outro	Outro
1385	-0.1239	0	Até secundário	0	Até Média Renda	Não segurado	Institucional	Outro	Outro
1607	-0.1336	0	Até secundário	1+	Até Média Renda	Não segurado	Institucional	Outro	Outro
1676	-0.1511	0	Até secundário	0	Até Média Renda	Não segurado	Institucional	Outro	Outro
1677	-0.1511	0	Até secundário	0	Até Média Renda	Não segurado	Institucional	Outro	Outro
1678	-0.1511	0	Até secundário	0	Até Média Renda	Não segurado	Institucional	Outro	Outro
1659	-0.1617	0	Até secundário	0	Até Média Renda	Não segurado	Institucional	Outro	Outro
1660	-0.1617	0	Até secundário	0	Até Média Renda	Não segurado	Institucional	Outro	Outro
1661	-0.1617	0	Até secundário	0	Até Média Renda	Não segurado	Institucional	Outro	Outro
1484	-0.1846	0	Até secundário	0	Até Média Renda	Não segurado	Institucional	Outro	Outro

Tabela A.9: Pontos com DFBETAS absolutos destacados para a variável Categoria de renda.

Índice	DFBETAS	Mais de 3 CPN	Nível Educacional	Exposição à Mídia	Categoria de Renda	Cobertura de Saúde	Local do CPN	Província	Decisor de Saúde
1607	0.1765	0	Até secundário	1+	Até Média Renda	Não segurado	Institucional	Outro	Outro
1662	0.1504	0	Até secundário	1+	Até Média Renda	Não segurado	Institucional	Outro	Outro
1384	0.1152	0	Até secundário	1+	Até Média Renda	Não segurado	Institucional	Outro	Outro
1176	-0.1216	0	Até secundário	0	Alta Renda	Não segurado	Institucional	Outro	Outro
1155	-0.1230	0	Até secundário	0	Alta Renda	Não segurado	Institucional	Outro	Outro

Tabela A.10: Pontos com DFBETAS absolutos destacados para a variável Decisor de saúde.

Índice	DFBETAS	Mais de 3 CPN	Nível Educacional	Exposição à Mídia	Categoria de Renda	Cobertura de Saúde	Local do CPN	Província	Decisor de Saúde
1483	0.1541	0	Até secundário	1+	Alta Renda	Não segurado	Institucional	Outro	Outro
1607	0.1395	0	Até secundário	1+	Até Média Renda	Não segurado	Institucional	Outro	Outro
1609	0.1318	0	Até secundário	0	Até Média Renda	Não segurado	Não Institucional	Outro	Outro
1484	0.1216	0	Até secundário	0	Até Média Renda	Não segurado	Institucional	Outro	Outro
1662	0.1189	0	Até secundário	1+	Até Média Renda	Não segurado	Institucional	Outro	Outro

Tabela A.11: Pontos com DFBETAS absolutos destacados para a variável Exposição à mídia.

Índice	DFBETAS	Mais de 3 CPN	Nível Educacional	Exposição à Mídia	Categoria de Renda	Cobertura de Saúde	Local do CPN	Província	Decisor de Saúde
1155	0.1297	0	Até secundário	0	Alta Renda	Não segurado	Institucional	Outro	Outro
1176	0.1282	0	Até secundário	0	Alta Renda	Não segurado	Institucional	Outro	Outro

Índice	DFBETAS	Mais de 3 CPN	Nível Educacional	Exposição à Mídia	Categoria de Renda	Cobertura de Saúde	Local do CPN	Província	Decisor de Saúde
1075	0.1159	0	Até secundário	0	Alta Renda	Não segurado	Institucional	Outro	Outro
1384	-0.1222	0	Até secundário	1+	Até Média Renda	Não segurado	Institucional	Outro	Outro
1554	-0.1408	0	Até secundário	1+	Até Média Renda	Não segurado	Não Institucional	Outro	Outro
1662	-0.1595	0	Até secundário	1+	Até Média Renda	Não segurado	Institucional	Outro	Outro
1607	-0.1872	0	Até secundário	1+	Até Média Renda	Não segurado	Institucional	Outro	Outro
1419	-0.1892	0	Até secundário	1+	Até Média Renda	Não segurado	Não Institucional	Outro	Respondente
1421	-0.1892	0	Até secundário	1+	Até Média Renda	Não segurado	Não Institucional	Outro	Respondente

Tabela A.12: Pontos com DFBETAS absolutos destacados para a variável Local do CPN.

Índice	DFBETAS	Mais de 3 CPN	Nível Educacional	Exposição à Mídia	Categoria de Renda	Cobertura de Saúde	Local do CPN	Província	Decisor de Saúde
1413	-0.1177	0	Até secundário	0	Até Média Renda	Não segurado	Não Institucional	Outro	Outro
1415	-0.1177	0	Até secundário	0	Até Média Renda	Não segurado	Não Institucional	Outro	Outro
1416	-0.1326	0	Até secundário	0	Até Média Renda	Não segurado	Não Institucional	Outro	Respondente
1420	-0.1326	0	Até secundário	0	Até Média Renda	Não segurado	Não Institucional	Outro	Respondente
1422	-0.1326	0	Até secundário	0	Até Média Renda	Não segurado	Não Institucional	Outro	Respondente
1609	-0.1511	0	Até secundário	0	Até Média Renda	Não segurado	Não Institucional	Outro	Outro
1419	-0.1557	0	Até secundário	1+	Até Média Renda	Não segurado	Não Institucional	Outro	Respondente
1421	-0.1557	0	Até secundário	1+	Até Média Renda	Não segurado	Não Institucional	Outro	Respondente
1608	-0.1702	0	Até secundário	0	Até Média Renda	Não segurado	Não Institucional	Outro	Respondente

Tabela A.13: Observações com distância de Cook maior ou igual a 4 do Modelo ajustado selecionado.

Nível Educacional	Exposição à Mídia	Categoria de Renda	Cobertura de Saúde	Local do CPN	Província	Decisor de Saúde	Mais de 3 CPN	Índice	Distância de Cook
Superior	1+	Alta Renda	Segurado	Institucional	Outro	Respondente	1	2065	10.245845
Superior	1+	Alta Renda	Não segurado	Institucional	Maputo	Respondente	1	3538	9.720871
Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente	1	3068	9.304335
Até secundário	1+	Alta Renda	Segurado	Institucional	Maputo	Respondente	1	3499	7.736348
Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente	1	134	7.293154
Superior	1+	Alta Renda	Não segurado	Institucional	Maputo	Respondente	1	3500	6.243689
Superior	0	Alta Renda	Não segurado	Institucional	Maputo	Respondente	1	3548	5.473826
Superior	1+	Alta Renda	Não segurado	Institucional	Maputo	Respondente	1	3471	5.375913
Superior	1+	Alta Renda	Não segurado	Institucional	Maputo	Respondente	1	3502	4.537890
Superior	0	Alta Renda	Não segurado	Institucional	Maputo	Respondente	1	3531	4.466792
Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente	1	3054	4.156633
Até secundário	1+	Alta Renda	Segurado	Institucional	Maputo	Respondente	1	3469	4.139223
Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente	1	1559	3.998323
Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Outro	1	2406	3.845373
Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente	1	1564	3.630032
Superior	1+	Alta Renda	Não segurado	Institucional	Outro	Respondente	1	3388	3.582722
Superior	1+	Alta Renda	Segurado	Institucional	Maputo	Respondente	1	3530	3.524926
Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente	1	1655	3.127088
Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente	1	1547	3.013674
Até secundário	1+	Alta Renda	Segurado	Institucional	Outro	Respondente	1	2660	3.004148

Tabela A.14: Código do estrato e respectiva região de província.

Estrato	Província
1	Niassa
2	Niassa
3	Cabo Delgado
4	Cabo Delgado
5	Nampula
6	Nampula

Estrato	Província
7	Zambézia
8	Zambézia
9	Tete
10	Tete
11	Manica
12	Manica
13	Sofala
14	Sofala
15	Inhambane
16	Inhambane
17	Gaza
18	Gaza
19	Maputo
20	Maputo
21	Cidade De Maputo

Tabela A.15: Métricas do modelo selecionado ajustado (Passo 4 do método *backward* realizado), removendo os 3 *outliers* da Tabela A.5.

Modelo	AIC	<i>Deviance</i>	<i>Null Deviance</i>	AUC
Passo 4 (com remoção dos <i>outliers</i>)	3621,3644	3277,7136	3576,2175	0,6718

APÊNDICE B

Figuras

Características gerais do conjunto de dados original: Variável resposta

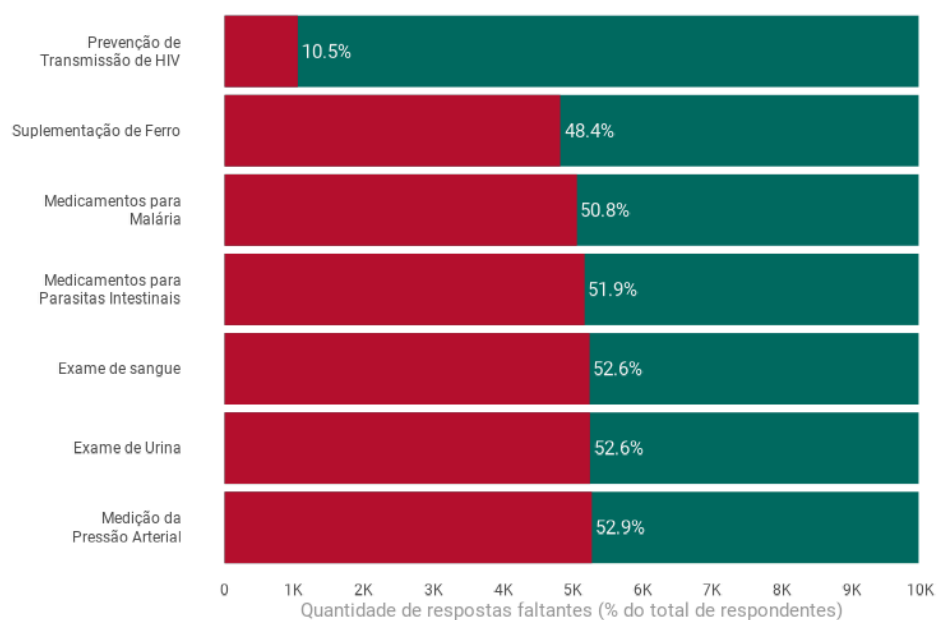


Figura B.1: Quantidade de entrevistadas da base inicial com respostas faltantes por cada componente de cuidado pré-natal que compõem a variável resposta.

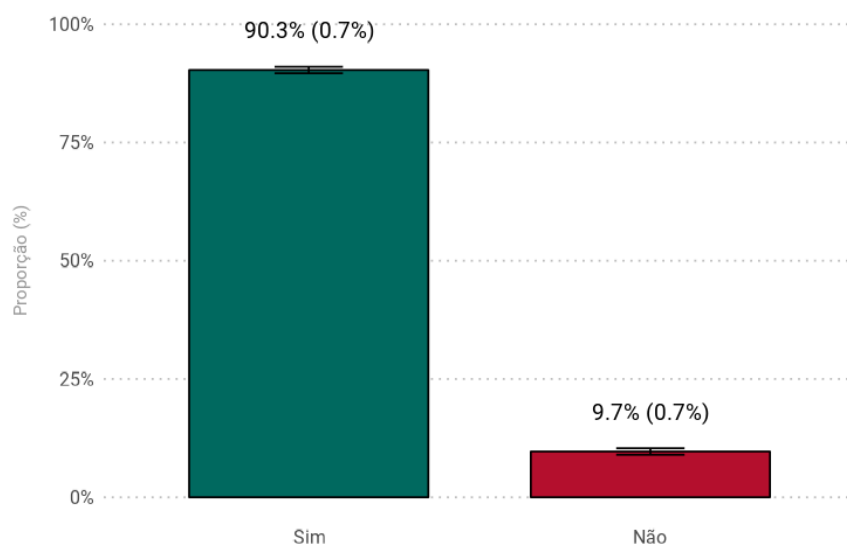


Figura B.2: Proporção de respondentes e respectivo erro padrão para a componente Suplementação de ferro.

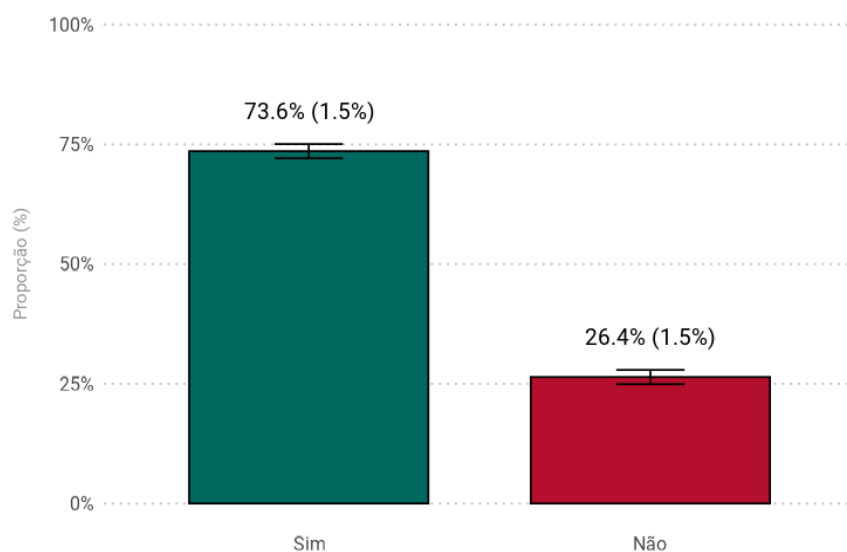


Figura B.3: Proporção de respondentes e respectivo erro padrão para a componente Medição da pressão arterial.

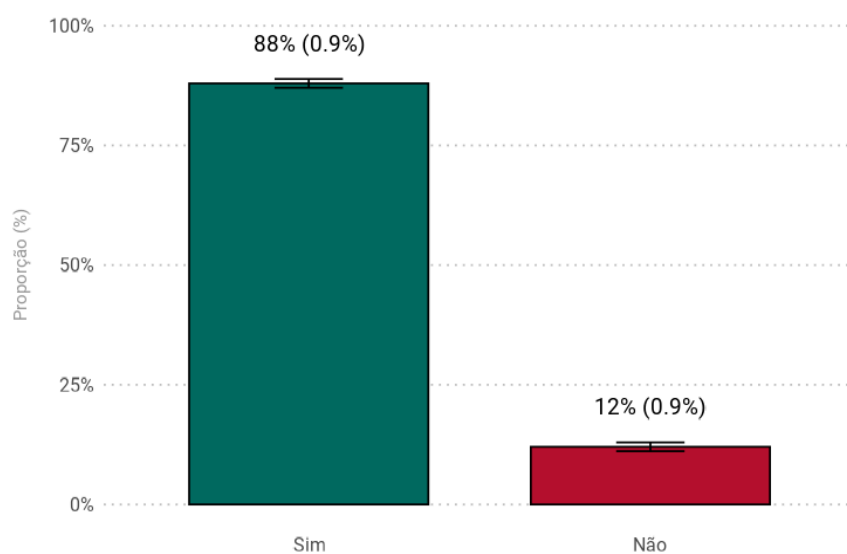


Figura B.4: Proporção de respondentes e respectivo erro padrão para a componente Realização de exame de sangue.

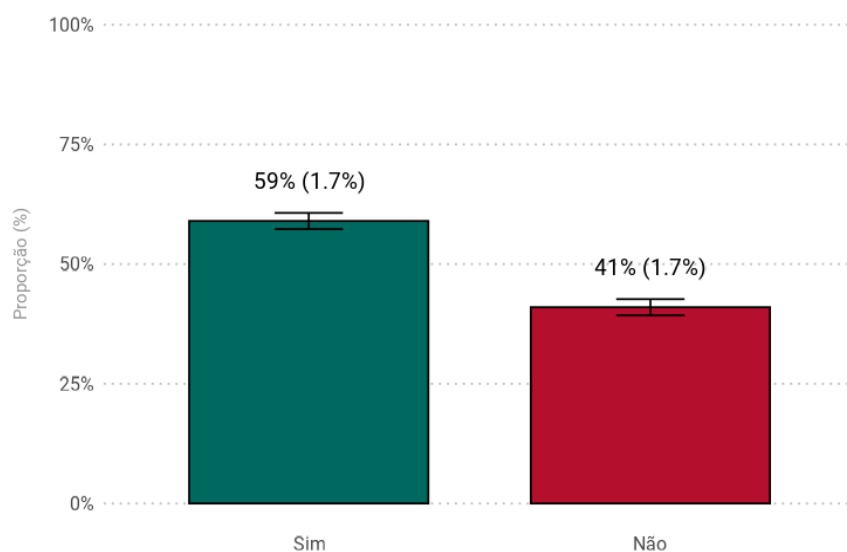


Figura B.5: Proporção de respondentes e respectivo erro padrão para a componente Realização de exame de urina.

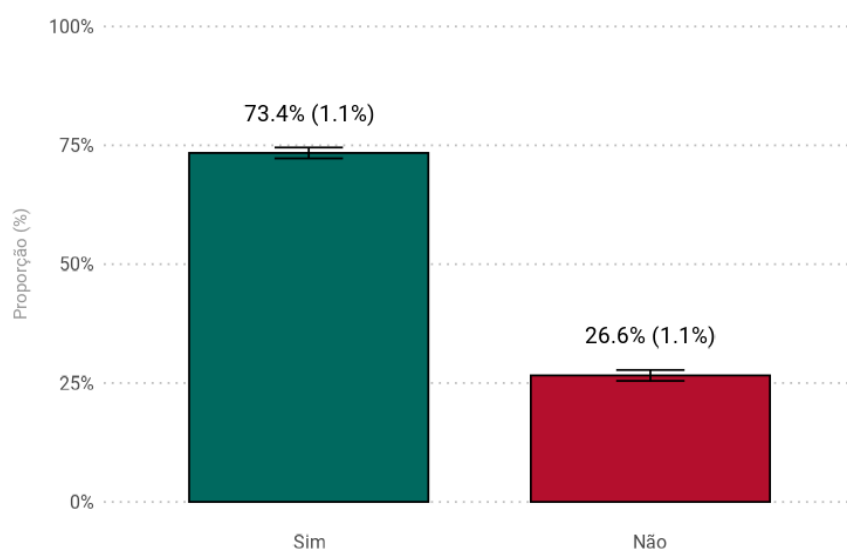


Figura B.6: Proporção de respondentes e respectivo erro padrão para a componente Ingestão de medicamentos contra malária.

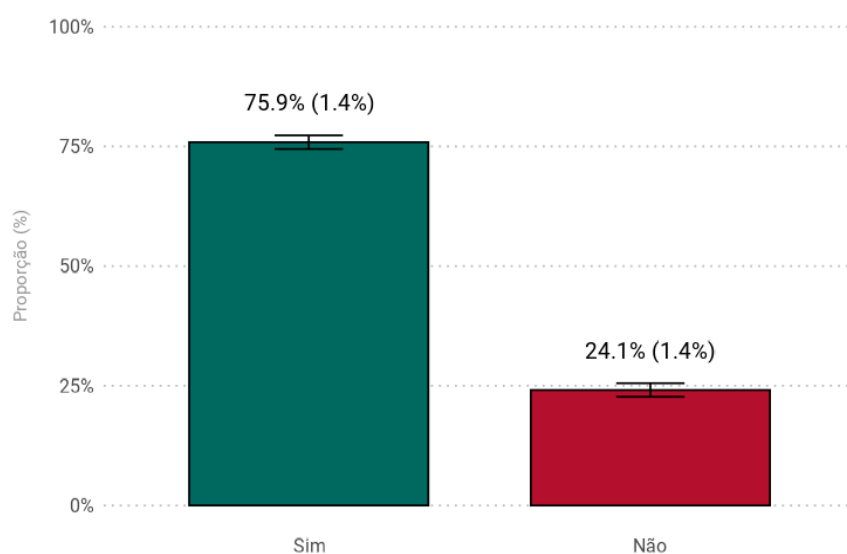


Figura B.7: Proporção de respondentes e respectivo erro padrão para a componente Uso de medicamento para prevenção de transmissão de HIV ao bebê.

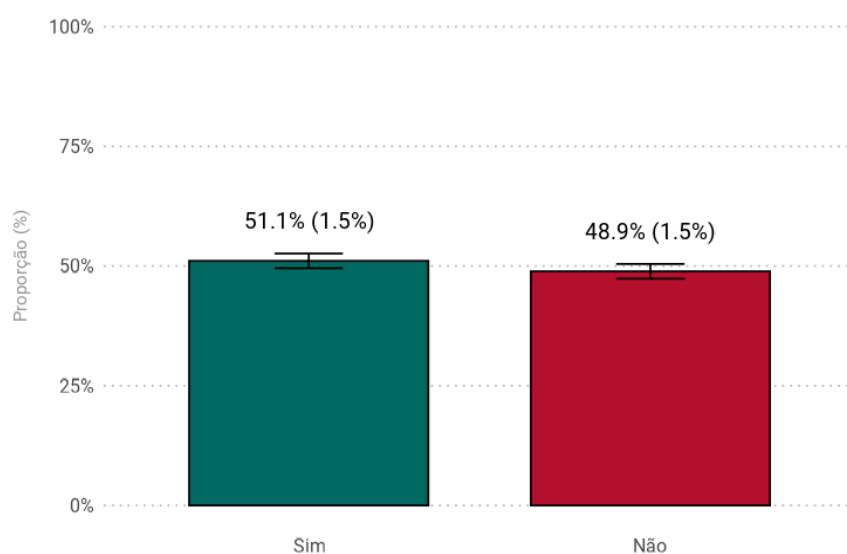


Figura B.8: Proporção de respondentes e respectivo erro padrão para a componente Ingestão de medicamentos contra parasitas intestinais.

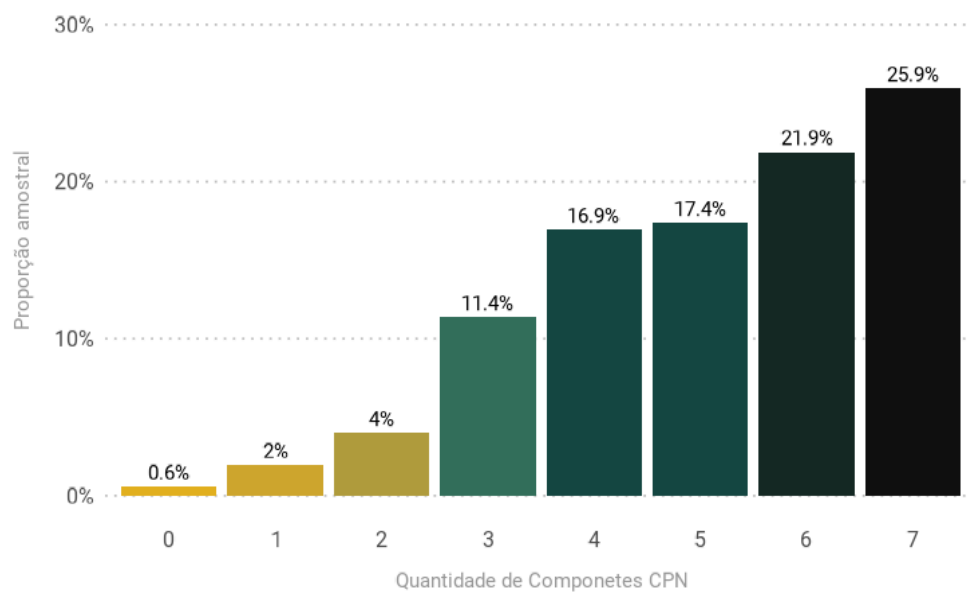


Figura B.9: Distribuição da variável resposta - soma das componentes pré-natais.

Características gerais do conjunto de dados original: Variáveis explicativas

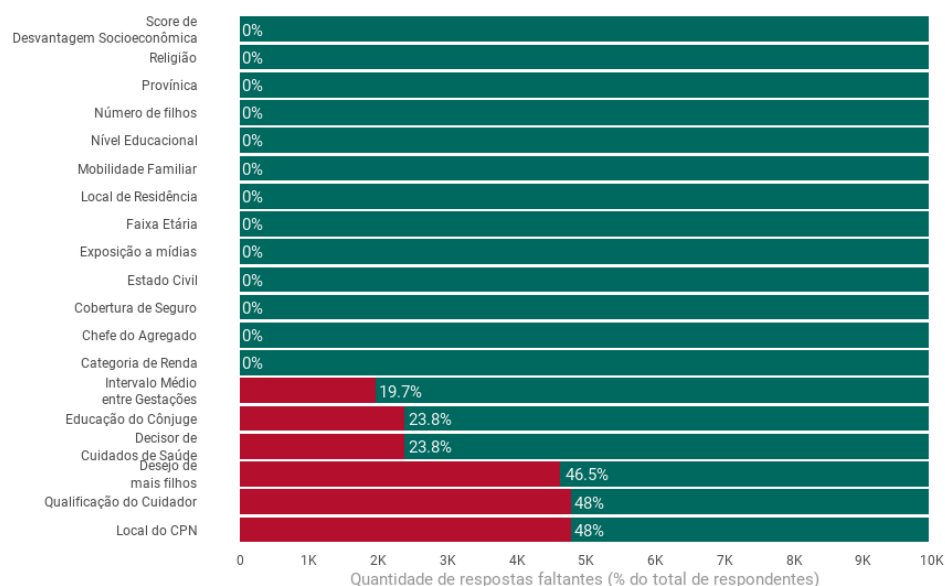


Figura B.10: Quantidade de entrevistadas com respostas faltantes para as variáveis explicativas.

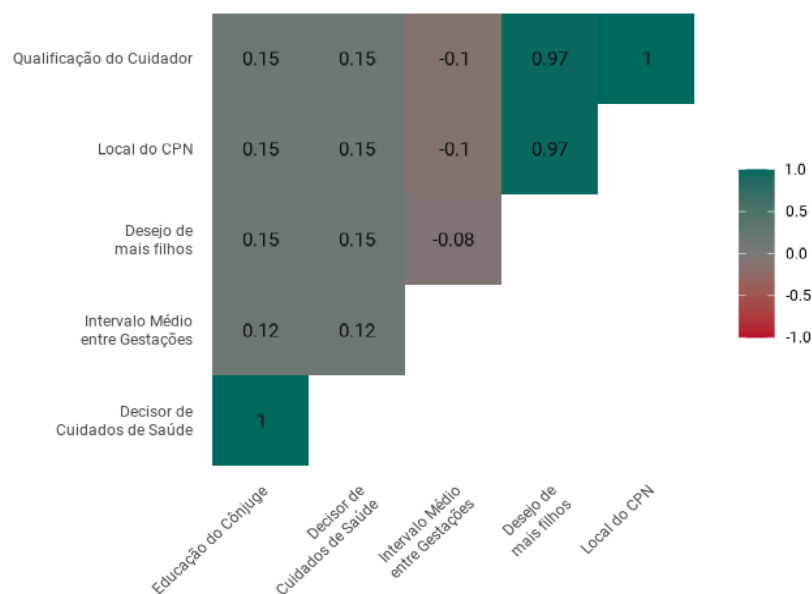


Figura B.11: Taxa de concordância total (1) e discordância total (-1) entre as variáveis explicativas com algum valor faltante.

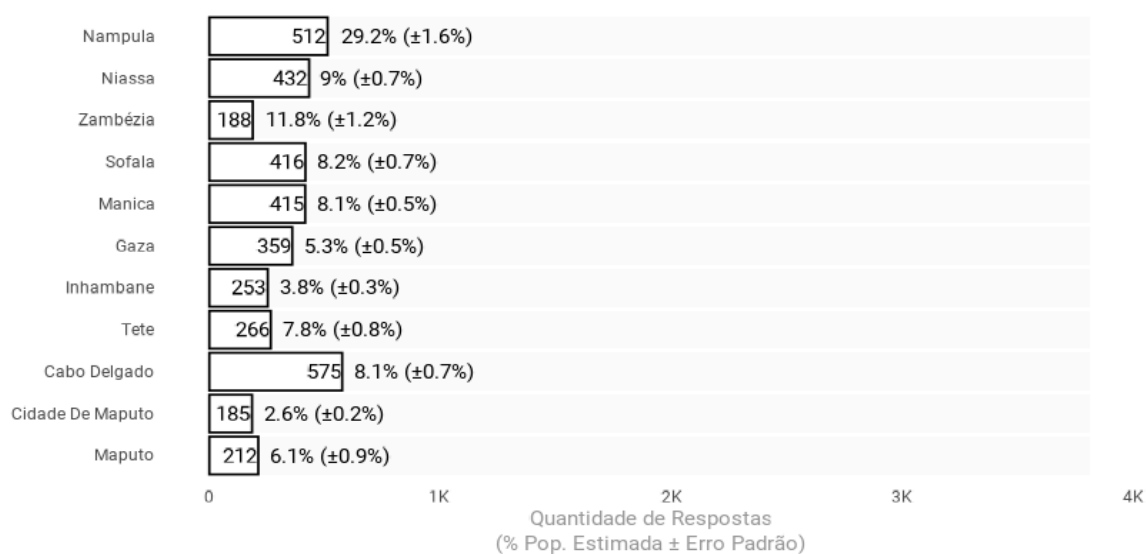


Figura B.12: Número absoluto de respondentes e respectiva proporção populacional na base de análise, contemplando o erro padrão.

Relação entre a quantidade de CPN e as variáveis explicativas

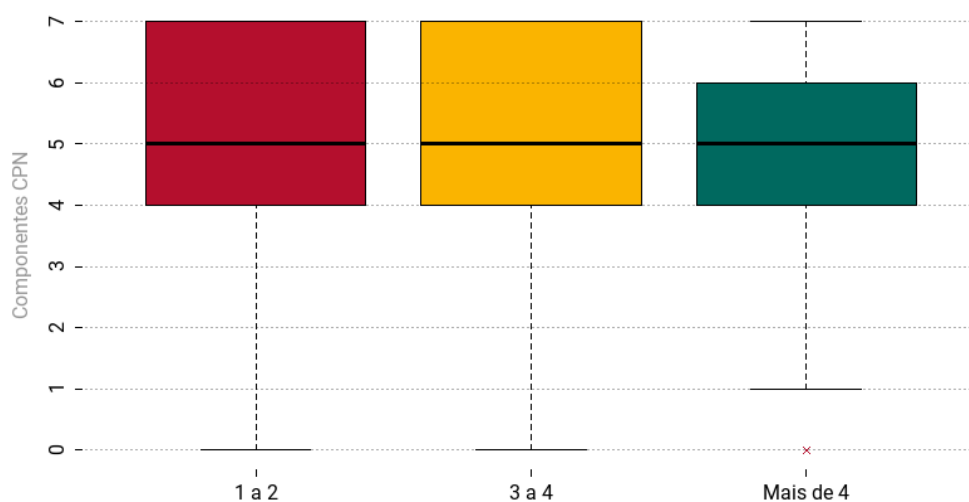


Figura B.13: *Box plot* da quantidade de componentes CPN em relação ao Número de filhos

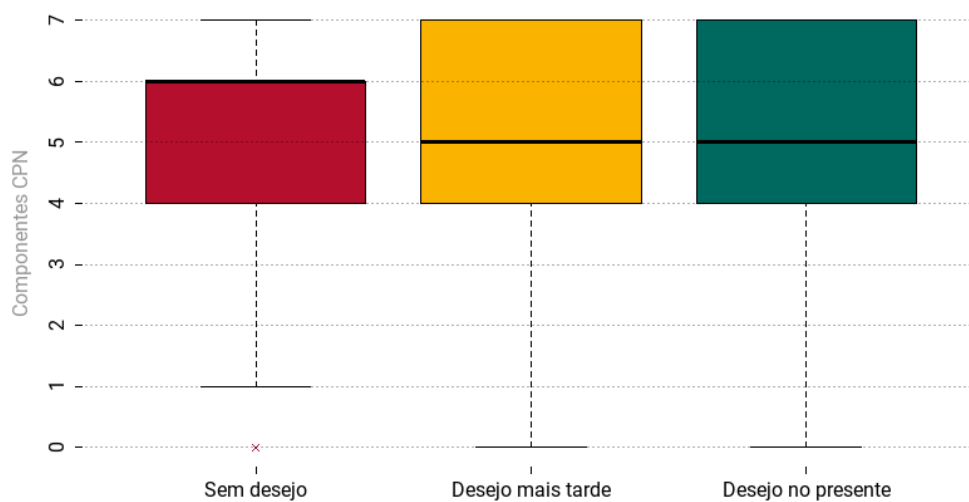


Figura B.14: *Box plot* da quantidade de componentes CPN em relação ao Desejo de ter mais filhos

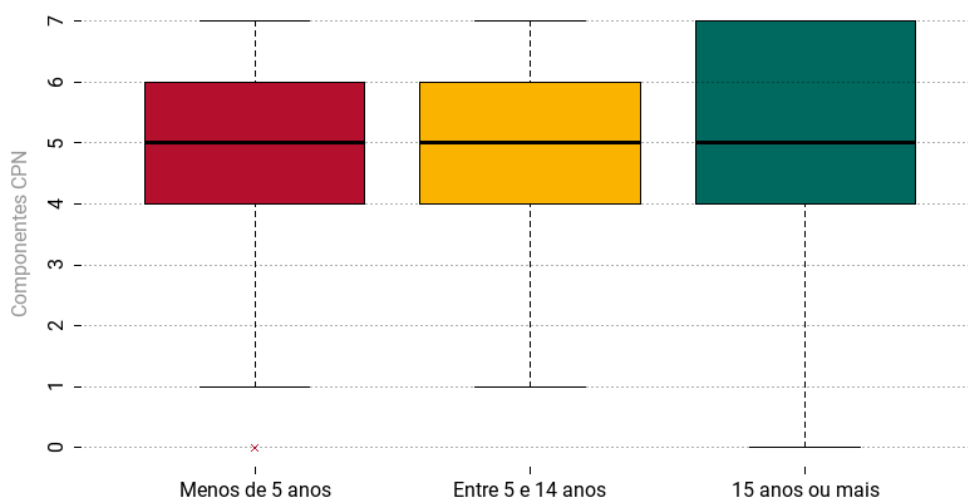


Figura B.15: *Box plot* da quantidade de componentes CPN em relação à Mobilidade familiar

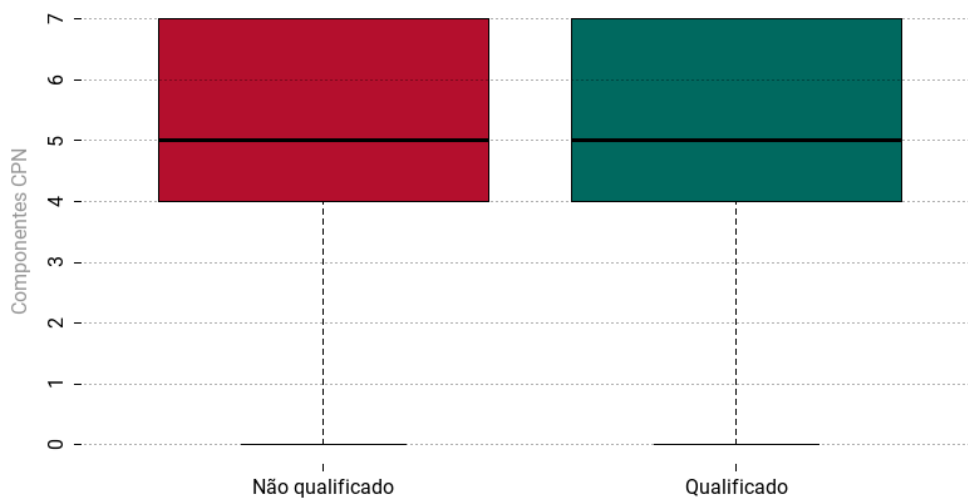


Figura B.16: *Box plot* da quantidade de componentes CPN em relação à Qualificação do cuidador

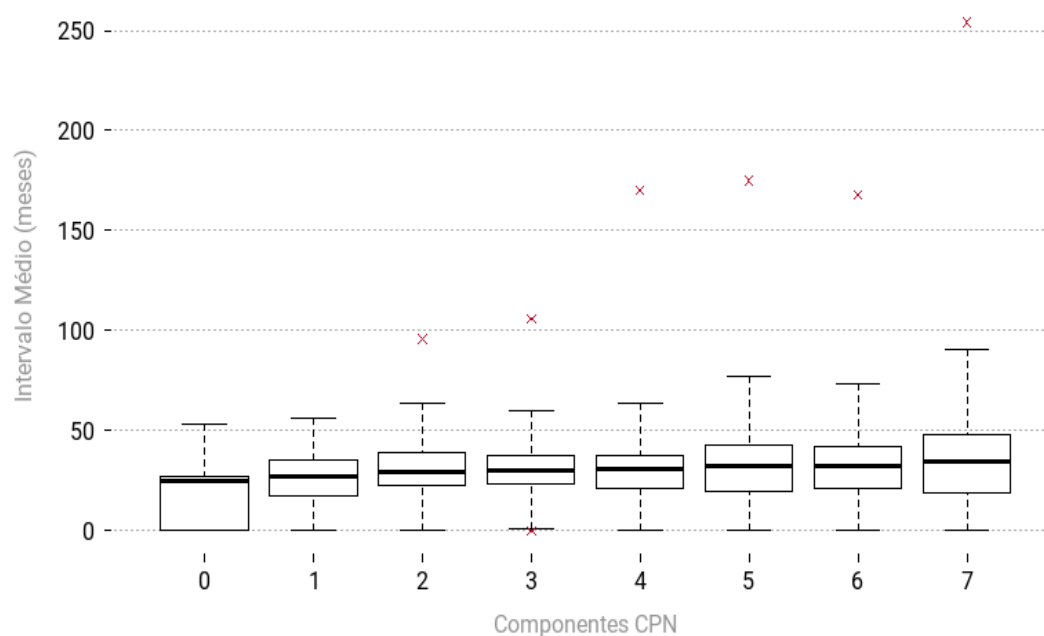


Figura B.17: *Box plot* do intervalo médio entre gestações (em meses) em relação à quantidade de componentes CPN

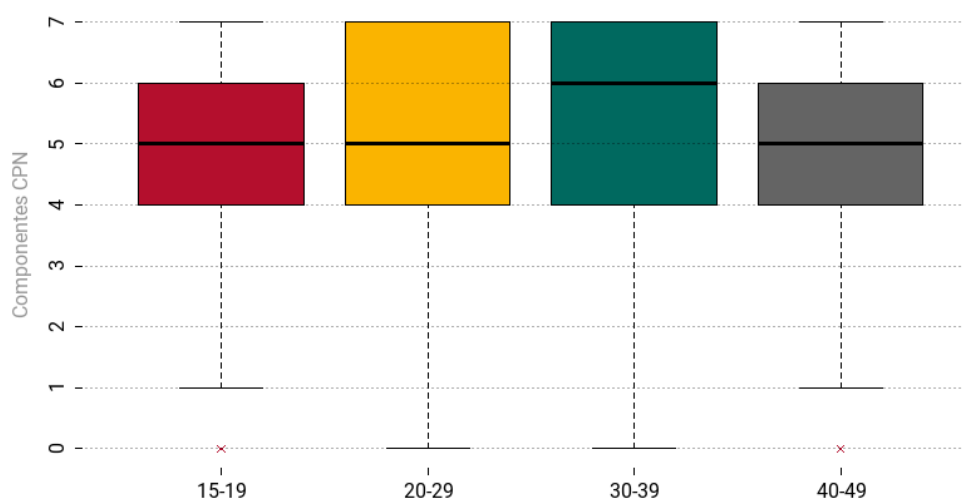


Figura B.18: *Box plot* da quantidade de componentes CPN em relação à Faixa etária

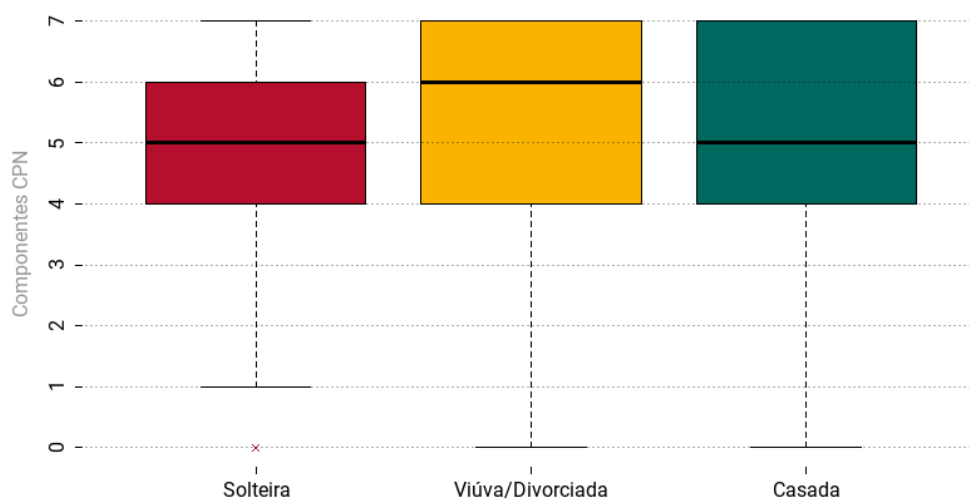


Figura B.19: *Box plot* da quantidade de componentes CPN em relação ao Estado civil

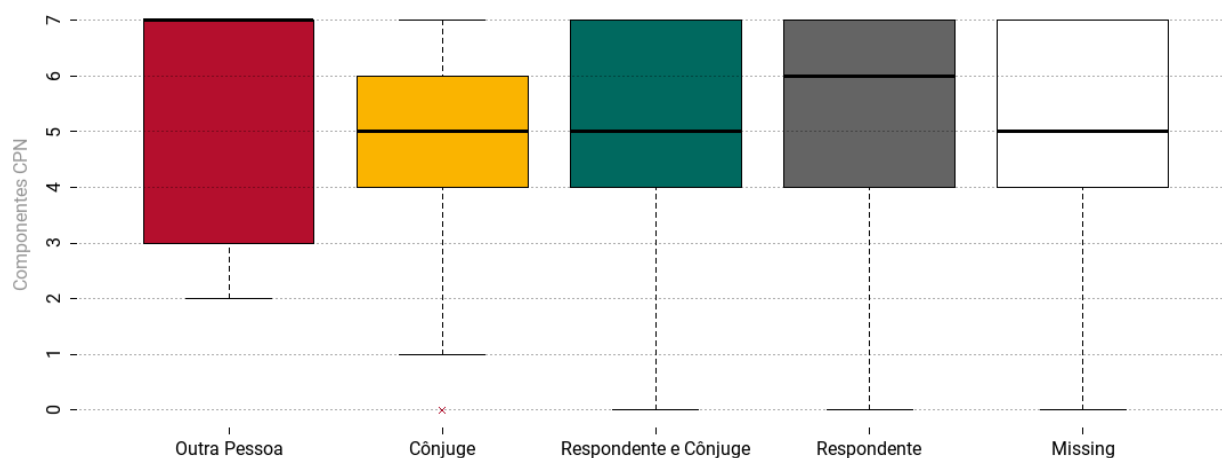


Figura B.20: *Box plot* da quantidade de componentes CPN em relação ao Decisor de cuidados de saúde

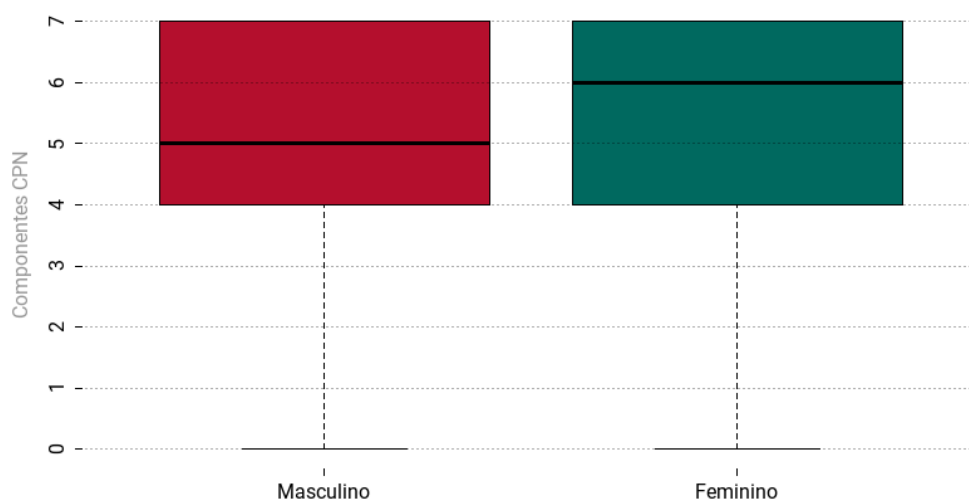


Figura B.21: *Box plot* da quantidade de componentes CPN em relação a Chefe do agregado

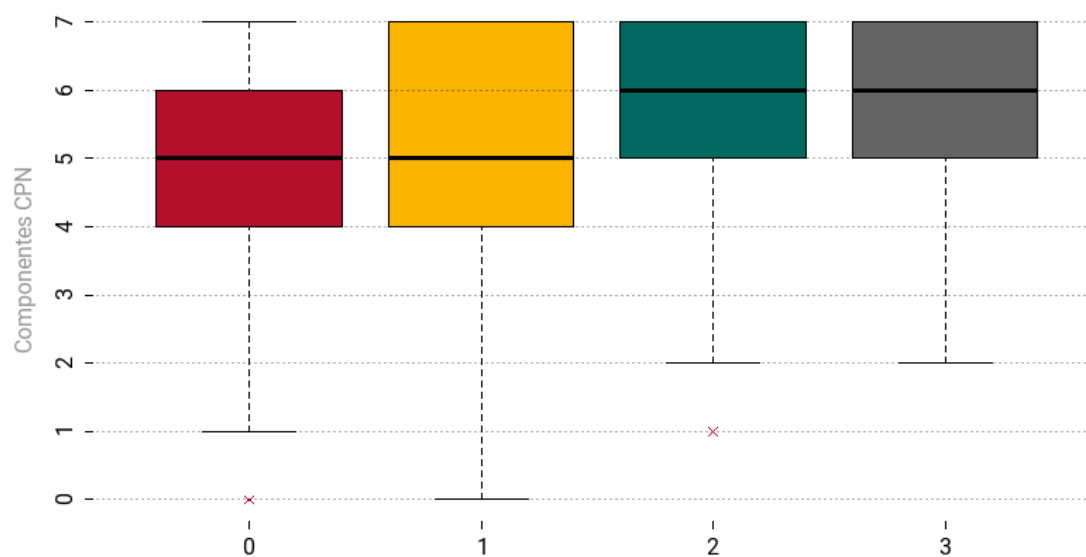


Figura B.22: *Box plot* da quantidade de componentes CPN em relação à Exposição à mídias

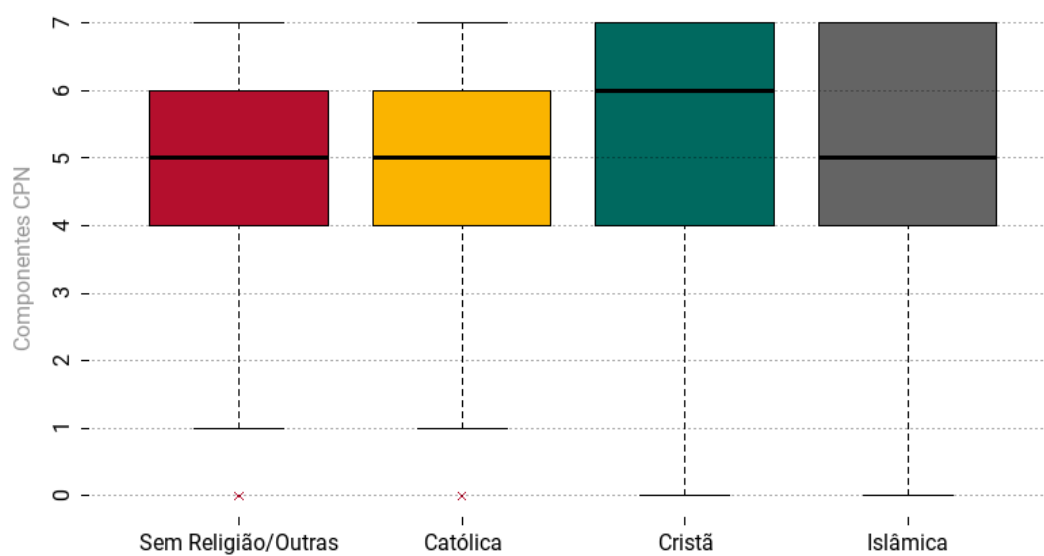


Figura B.23: *Box plot* da quantidade de componentes CPN em relação à Religião

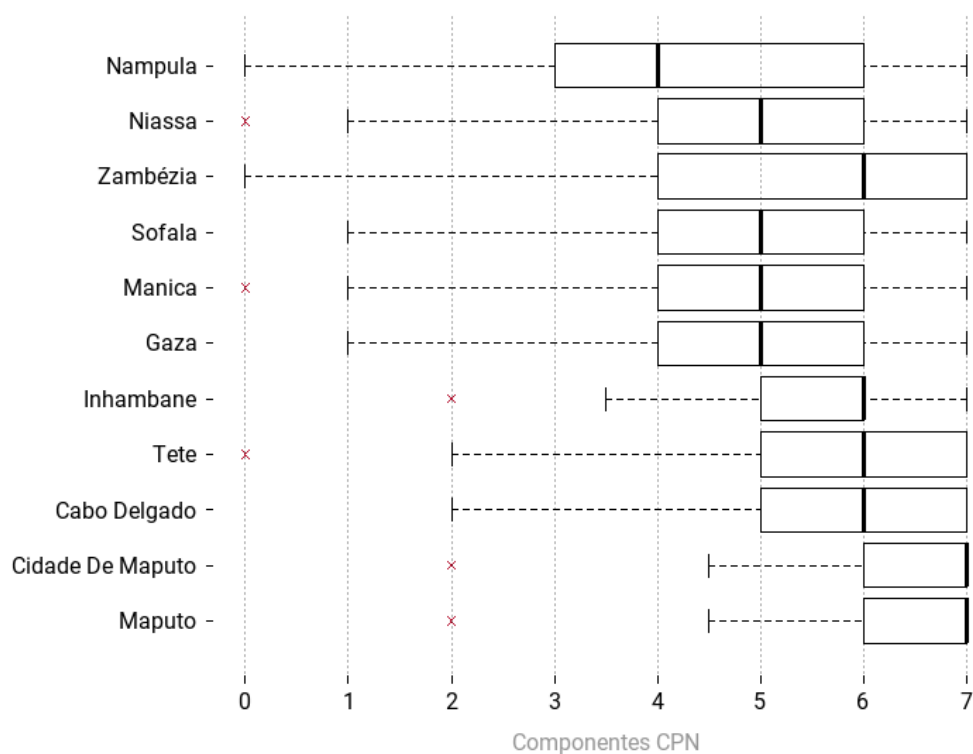


Figura B.24: *Box plot* da quantidade de componentes CPN em relação à Província

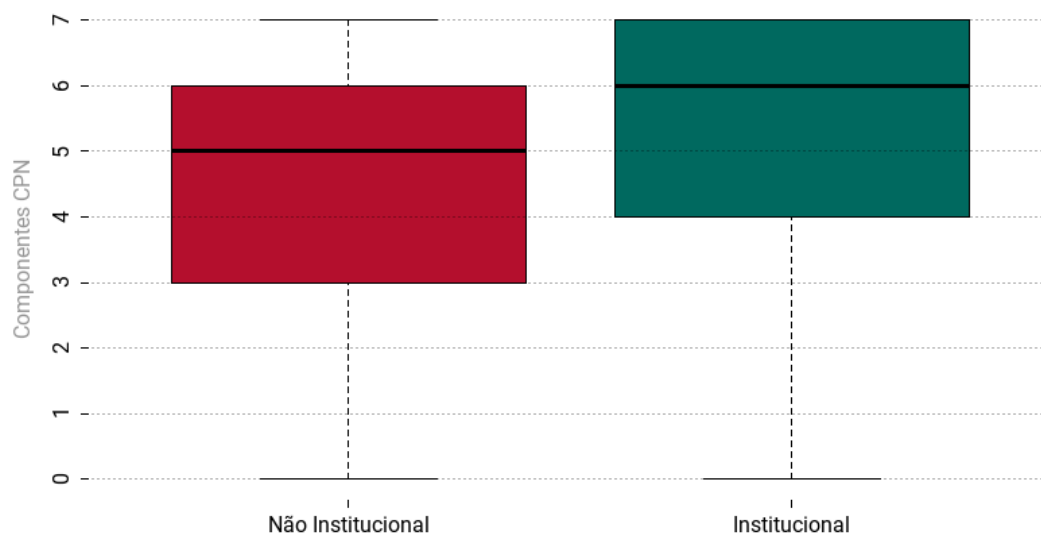


Figura B.25: *Box plot* da quantidade de componentes CPN em relação ao Local do CPN

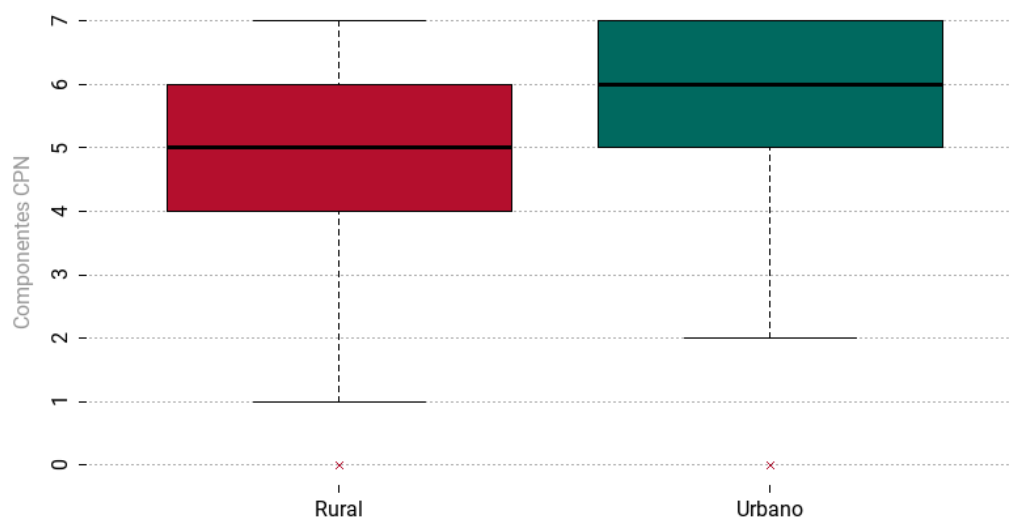


Figura B.26: *Box plot* da quantidade de componentes CPN em relação ao Local de residência

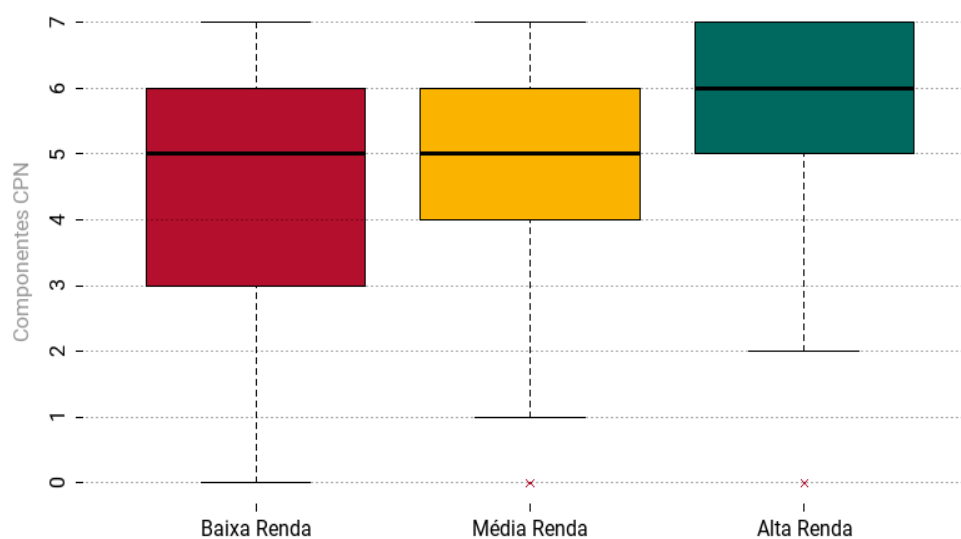


Figura B.27: *Box plot* da quantidade de componentes CPN em relação à Categoria de renda

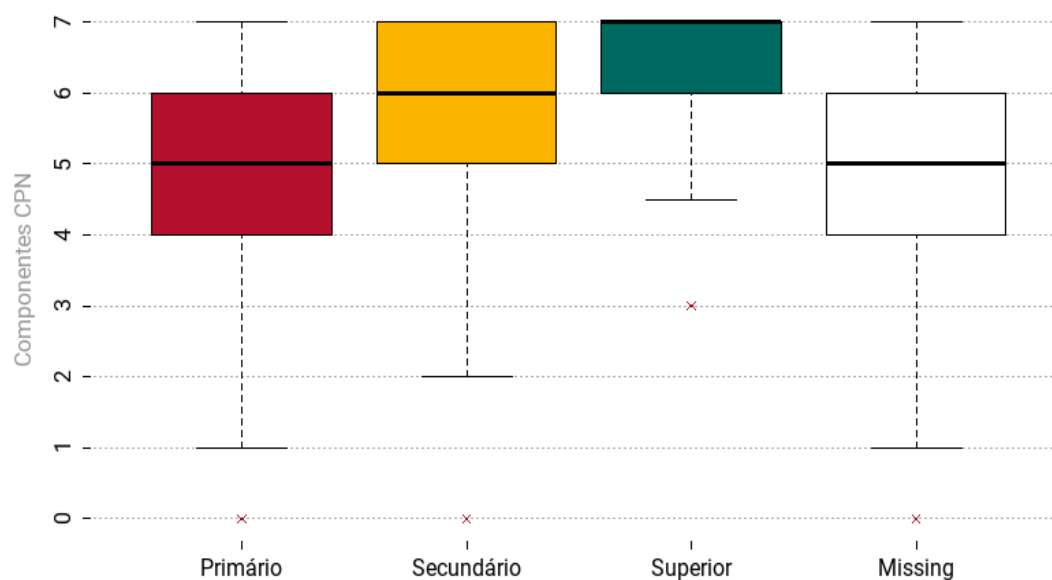


Figura B.28: *Box plot* da quantidade de componentes CPN em relação à Educação do cônjuge

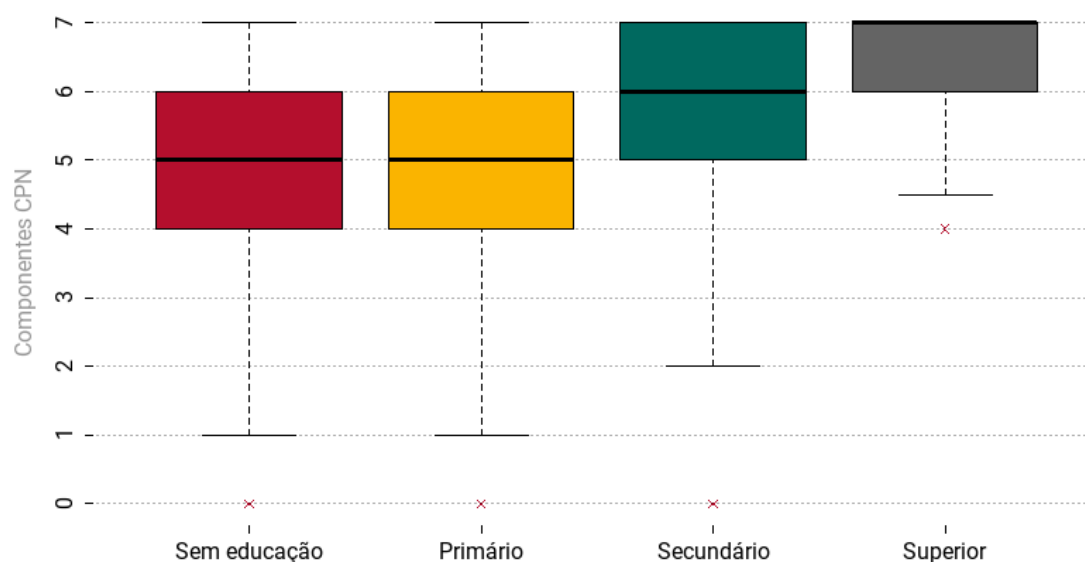


Figura B.29: *Box plot* da quantidade de componentes CPN em relação ao Nível educacional

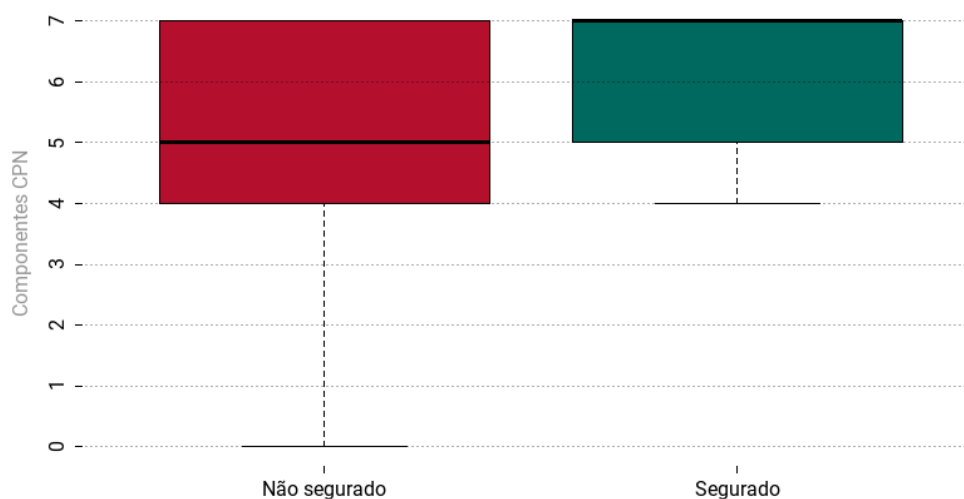


Figura B.30: *Box plot* da quantidade de componentes CPN em relação à Cobertura de seguro

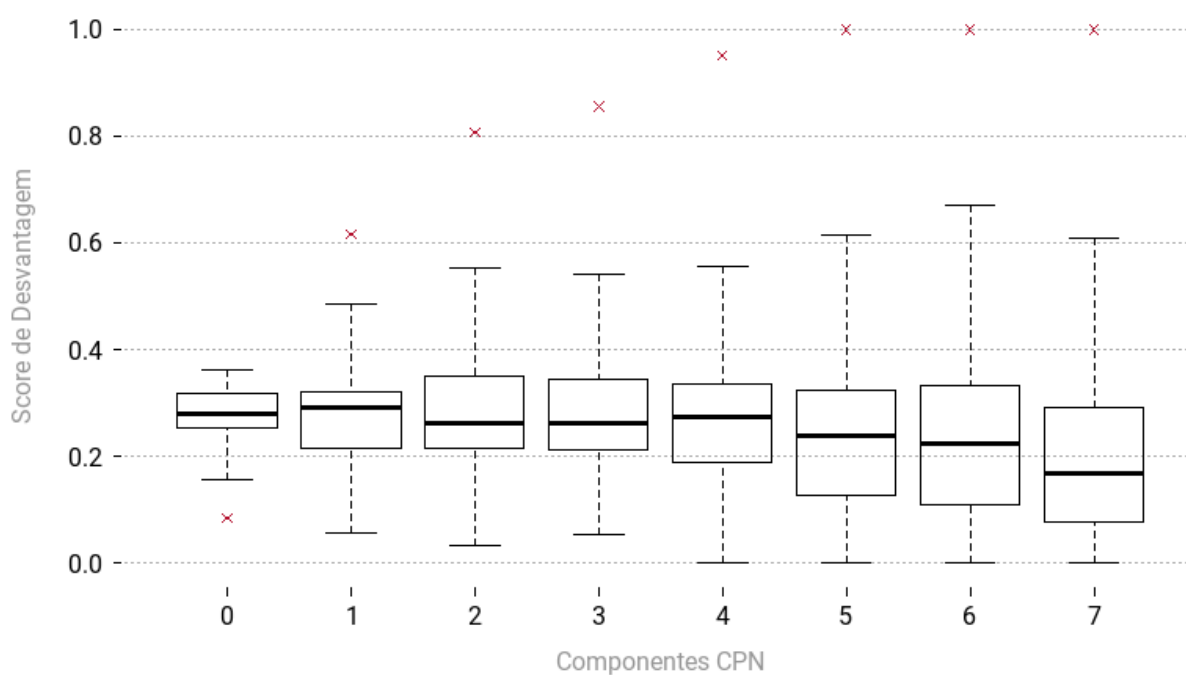


Figura B.31: *Box plot* do Escore de Desvantagem Socioeconômica em relação à quantidade de componentes CPN

Proporção de respondentes em cada nível das variáveis explicativas

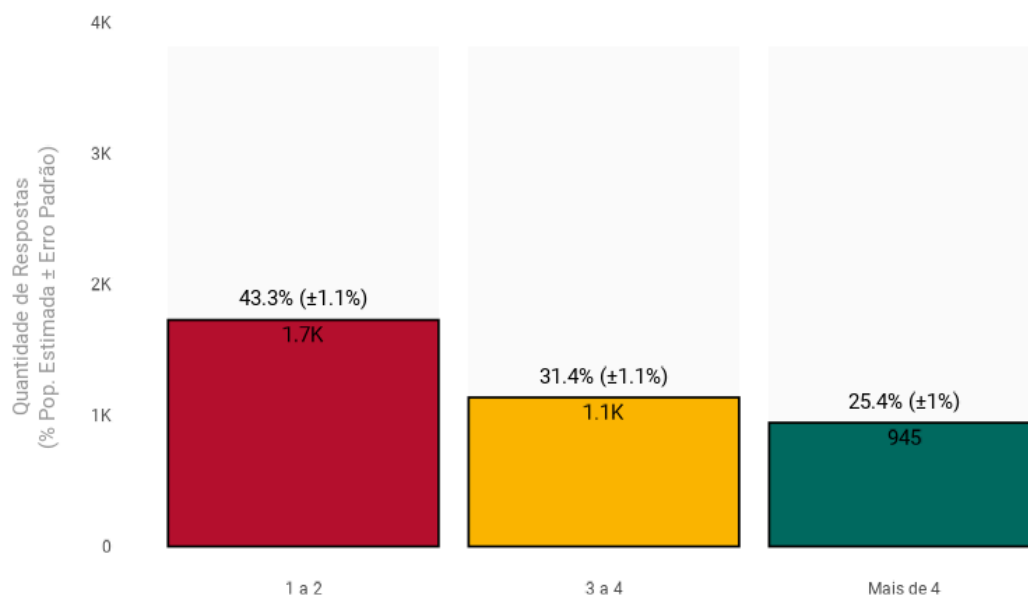


Figura B.32: Quantidade de respostas e proporção populacional estimada para Número de filhos

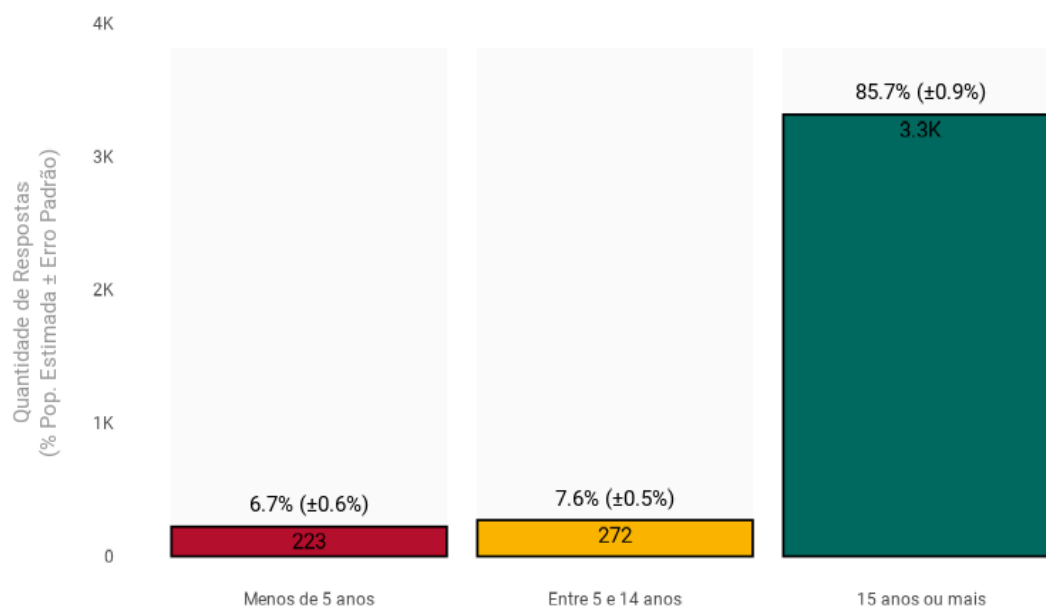


Figura B.33: Quantidade de respostas e proporção populacional estimada para Mobilidade familiar

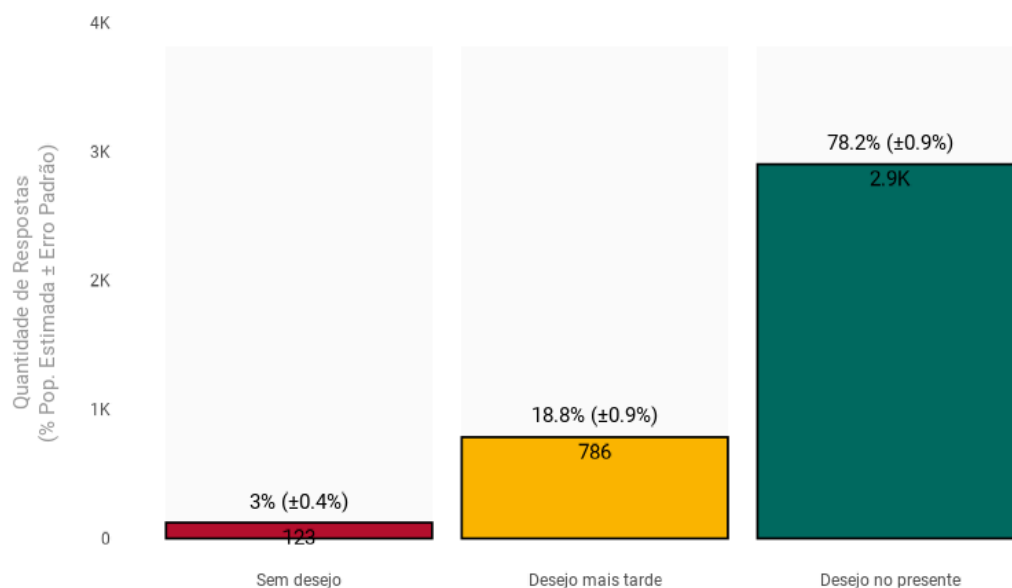


Figura B.34: Quantidade de respostas e proporção populacional estimada para Desejo de ter mais filhos

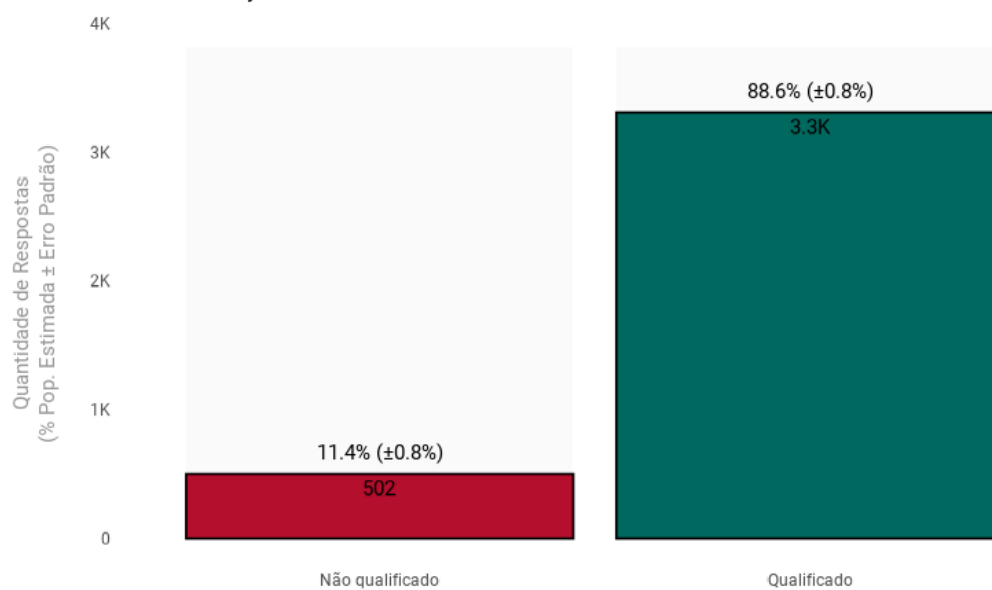


Figura B.35: Quantidade de respostas e proporção populacional estimada para Qualificação do cuidador

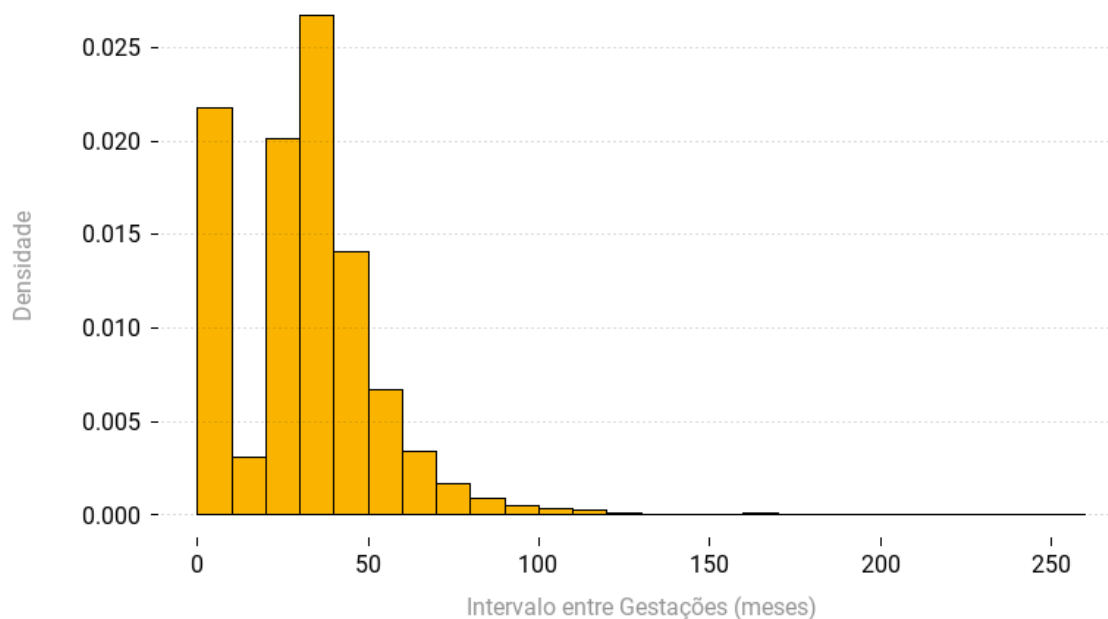


Figura B.36: Histograma do Intervalo entre gestações (em meses)

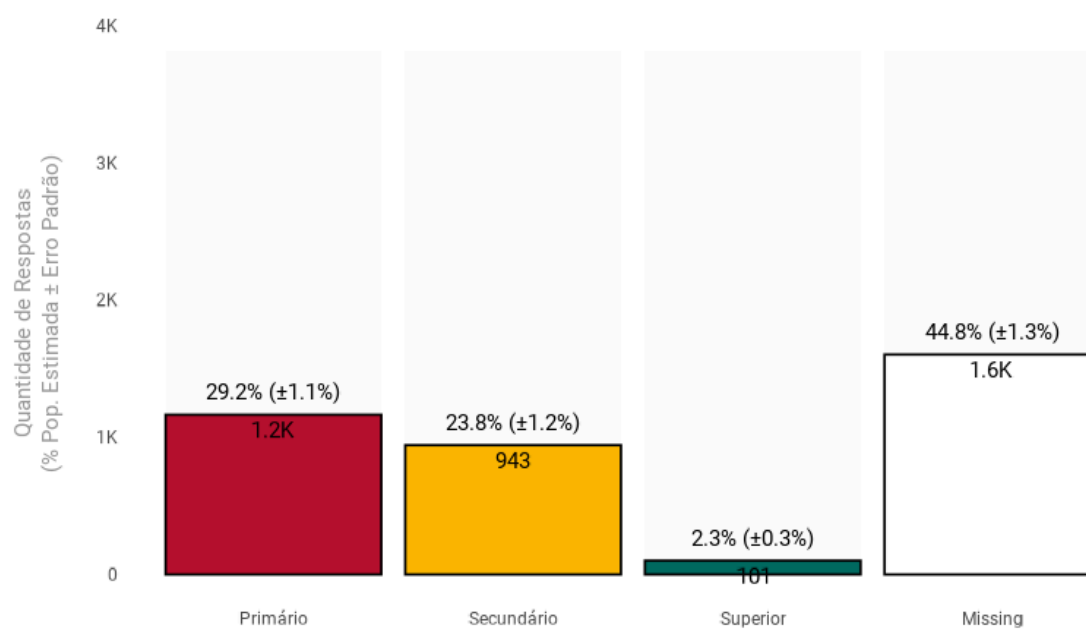


Figura B.37: Quantidade de respostas e proporção populacional estimada para Educação do cônjuge

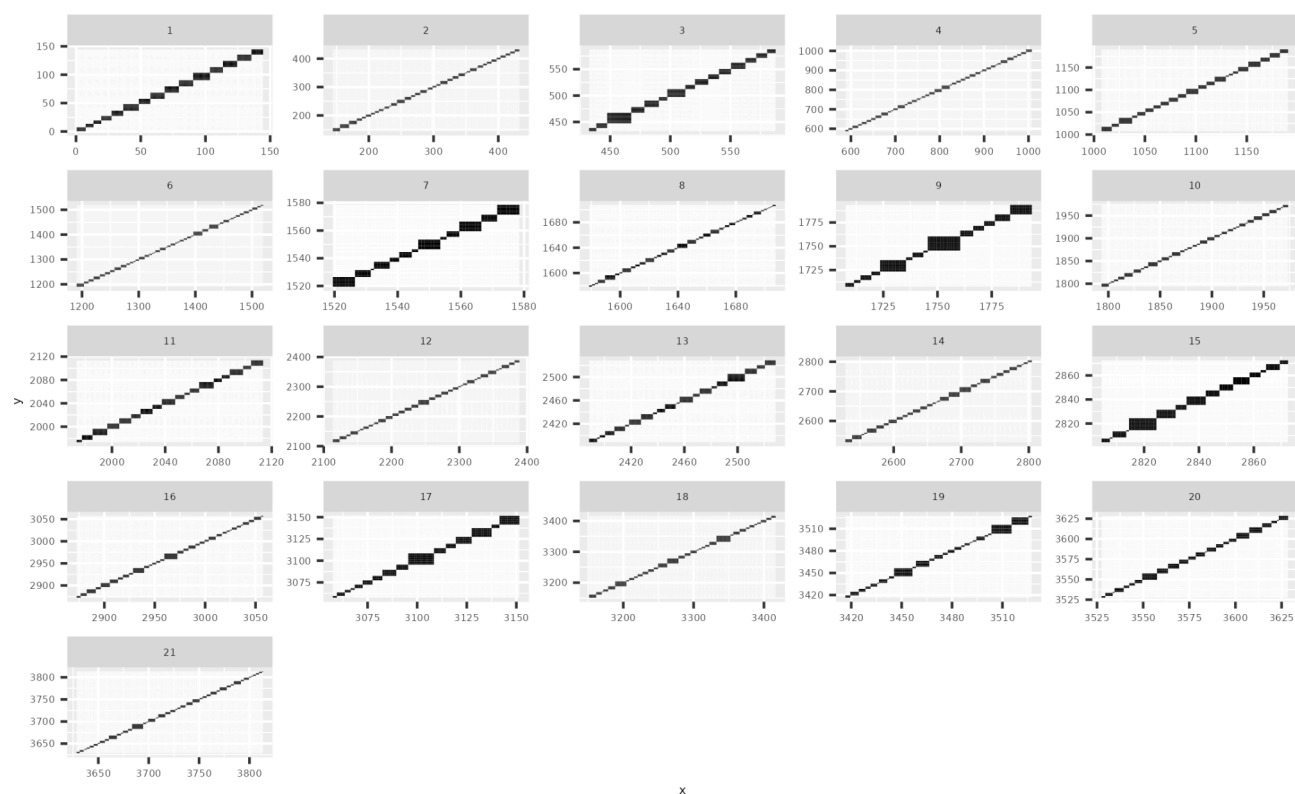


Figura B.39: Observações pertencentes a cada conglomerado dentro de cada estrato pelo desenho amostral complexo.

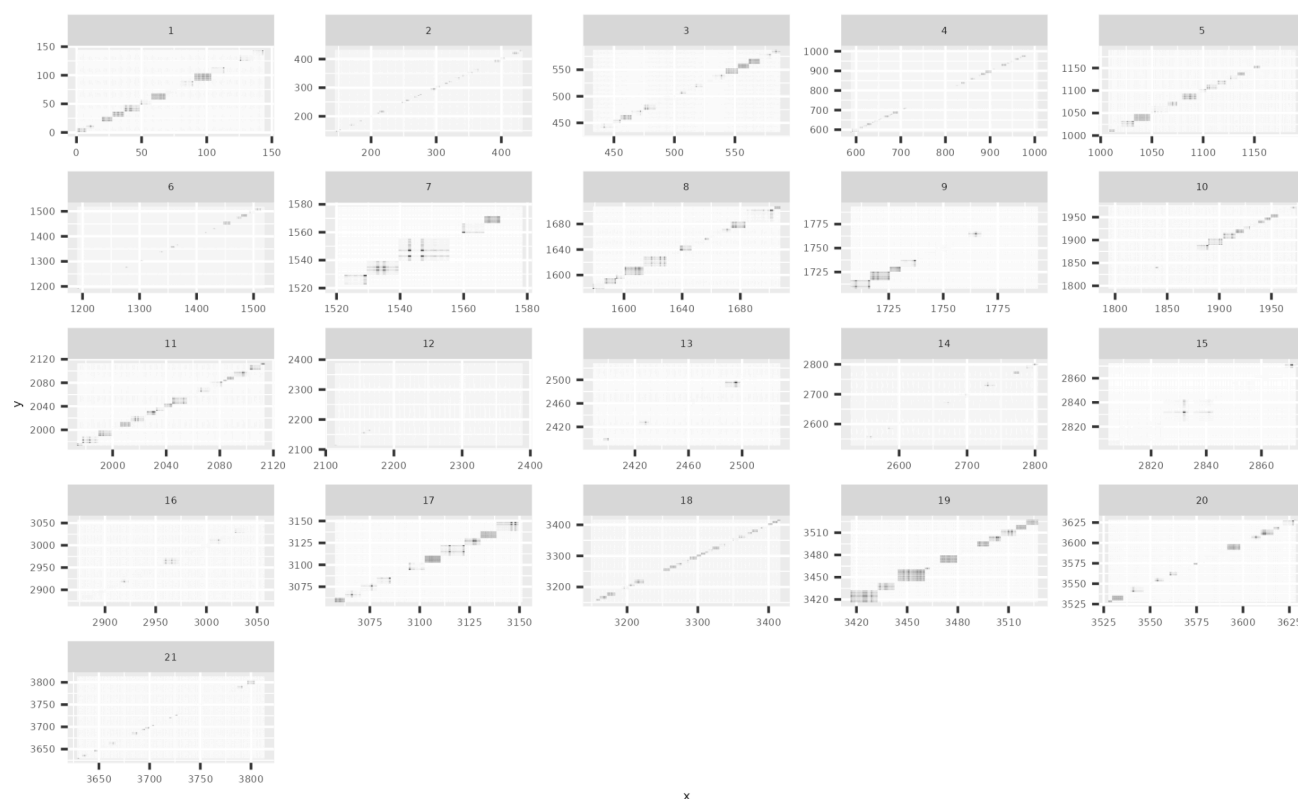


Figura B.40: Covariância entre as observações⁵, representada por tonalidade mais escura para os maiores valores absolutos e tonalidade mais clara para as correlações próximas a zero. Portanto, caselas em branco representam falta de correlação.

⁵ Calculado pela função *Vmat* do pacote *svydiags* do software *R*.

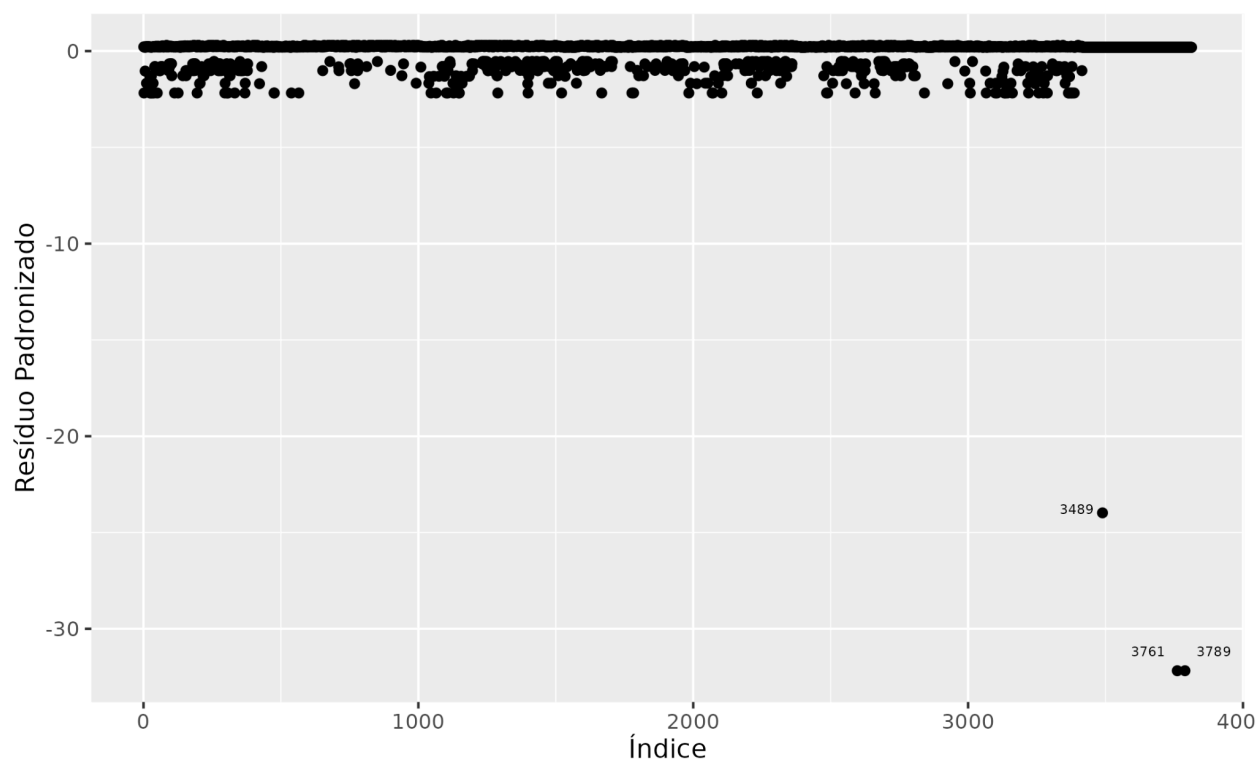


Figura B.41: Resíduos padronizados.

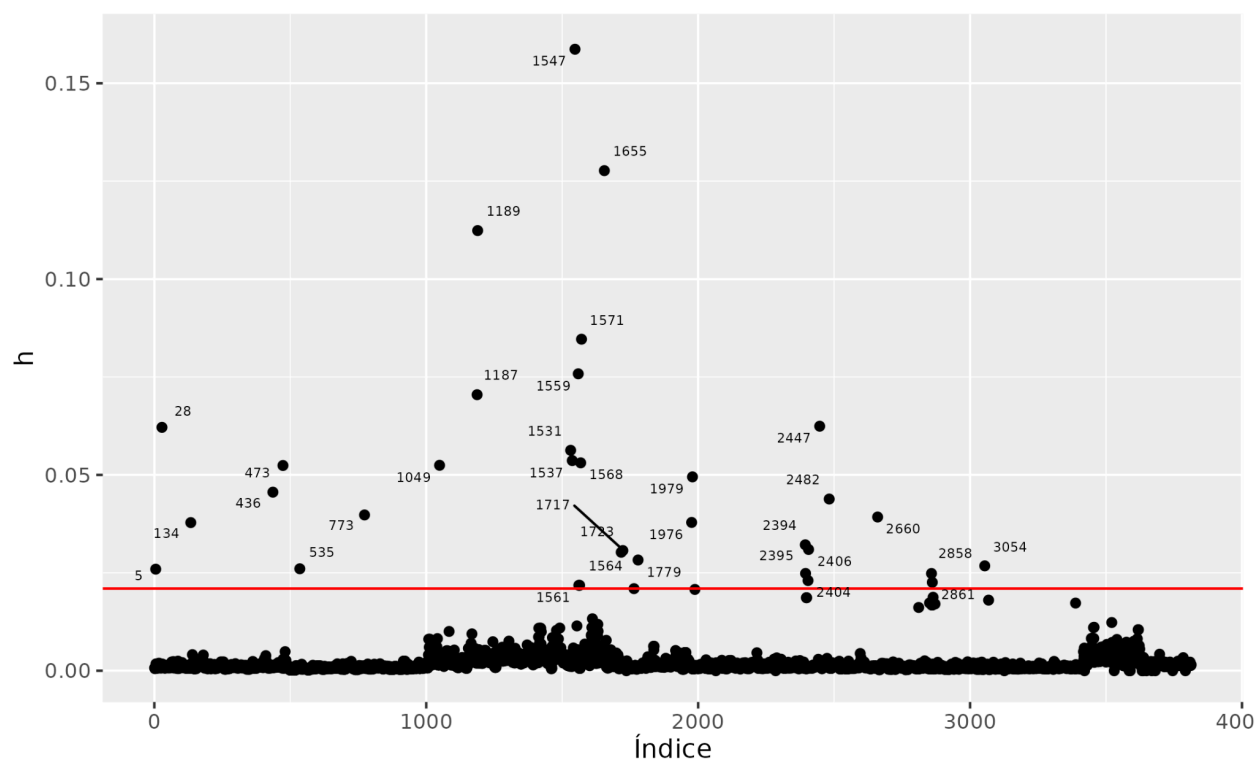


Figura B.42: Pontos de alavanca. Linha vermelha destacando os pontos acima de $10 \times \frac{p}{n}$, onde p é o número de parâmetros do modelo (8) e n é o número de observações no conjunto de dados ajustado (3.813).

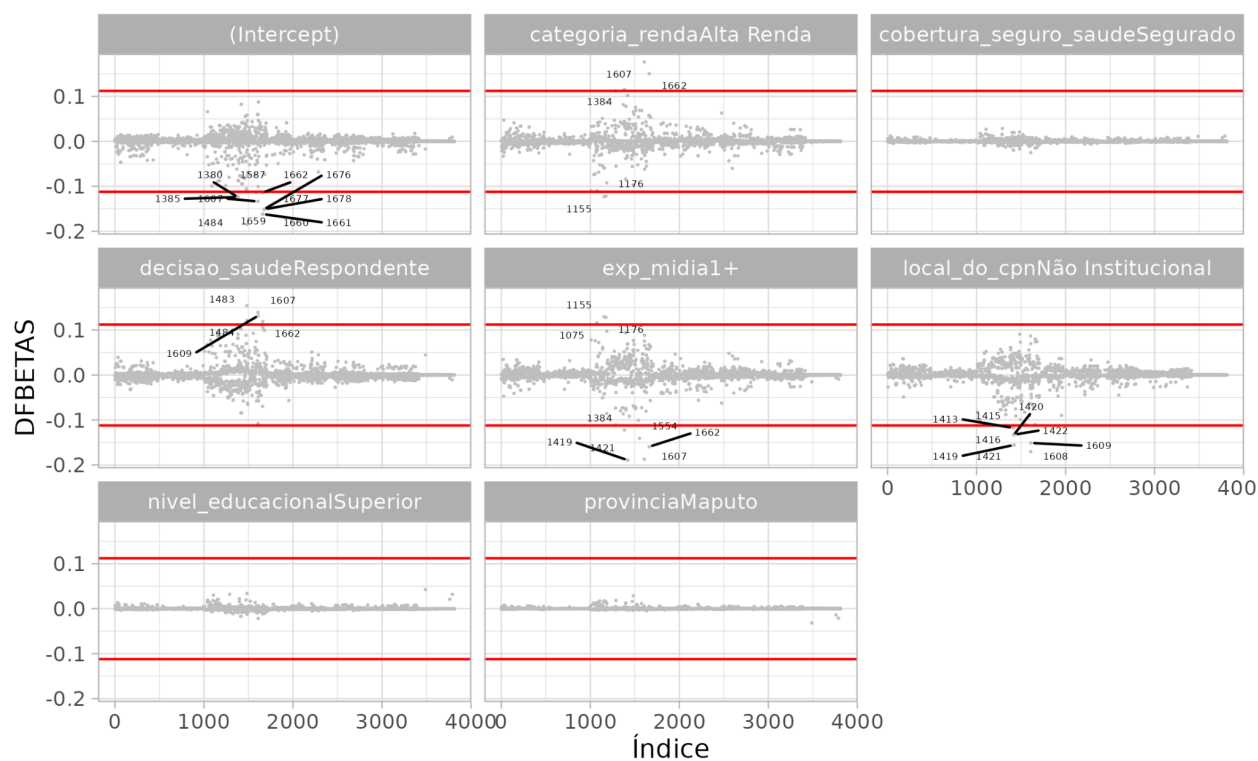


Figura B.43: DFBETAS (normalizado) para as observações.

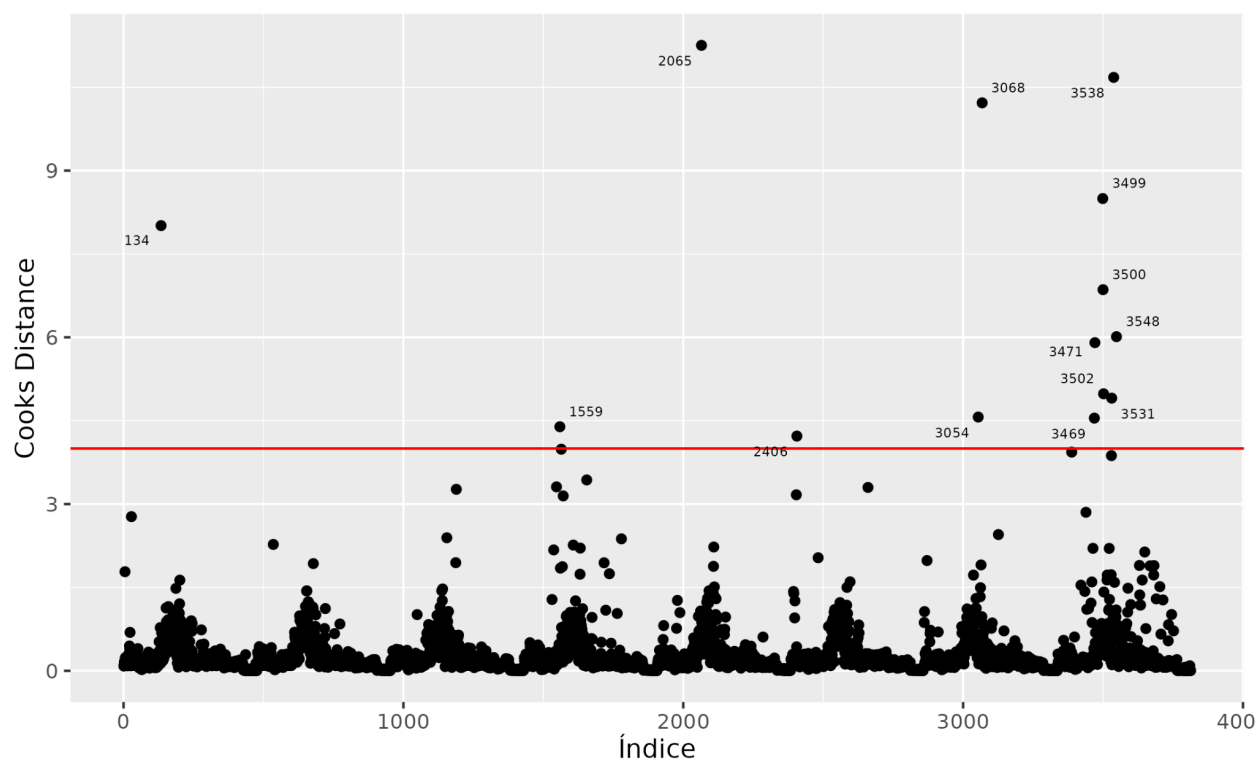


Figura B.44: Distância de Cook para o Modelo Ajustado. Índices com distância maior ou igual a 4 destacados acima da linha vermelha.