Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

Caroline Rodrigues Lopes

Consumo de álcool entre adolescentes: uma aplicação de regressão logística

Brasília - DF
2017

Caroline Rodrigues Lopes

Consumo de álcool entre adolescentes: uma aplicação de regressão logística

Projeto apresentado para obtenção do título de Bacharel em Estatística

Orientador: Prof. Dr. Luís Gustavo do Amaral Vinha

Brasília - DF
2017

Resumo

O consumo de álcool tem se mostrado cada vez mais presente entre adolescentes. Esse uso precoce é motivo de preocupação por trazer consequências a curto e longo prazo, como queda no rendimento escolar e dependência química. Assim, é importante caracterizar quais são os comportamentos de risco e proteção relacionados a ingestão de bebidas alcoólicas. Para tanto, aplicou-se a técnica de regressão logística em uma amostra selecionada no Distrito Federal e na cidade de Goiânia para identificar os principais fatores associados ao consumo de álcool entre crianças e adolescentes. Os modelos foram ajustados para o uso de álcool na vida, no último ano, no último mês e para ocorrência de *binge* alcoólico no último ano. As variáveis significativas incluídas nos modelos referem-se a idade, recebimento do programa Bolsa Família, escolaridade da mãe, práticas parentais e autoeficácia. Pela modelagem observou-se o papel que a prática parental intrusividade e a autoeficácia para aprendizagem autorregulada exercem enquanto fatores de risco e proteção, respectivamente, ao consumo de álcool. As técnicas usadas no diagnóstico dos modelos indicaram que esses se adequam aos dados e possuem poder de discriminação aceitável, podendo auxiliar estudos futuros que busquem reduzir fatores de risco associados ao uso de álcool.

Palavras-chave: Consumo de álcool. Práticas parentais. Autoeficácia. Regressão logística.

Lista de ilustrações

Figura 1 – Distribuição dos participantes segundo o ano escolar 20
Figura 2 – Distribuição dos participantes segundo a escolaridade da mãe 21
Figura 3 – Distribuição dos participantes segundo os planos para o futuro 21
Figura 4 – Curva ROC para os quatro modelos 35
Figura 5 – Resíduos *deviance versus* preditor linear para os quatro modelos 36
Figura 6 – Envelope simulado dos resíduos *deviance* para os quatro modelos 37
Figura 7 – Gráficos para detecção de pontos influentes no modelo 1 38
Figura 8 – Gráficos para detecção de pontos influentes no modelo 2 39
Figura 9 – Gráficos para detecção de pontos influentes no modelo 3 40
Figura 10 – Gráficos para detecção de pontos influentes no modelo 4 41

Lista de tabelas

Tabela 1 – Distribuição dos participantes segundo o consumo de álcool 22
Tabela 2 – Distribuição dos participantes segundo o uso de cigarro 22
Tabela 3 – Distribuição dos participantes segundo o uso de inalantes 23
Tabela 4 – Distribuição dos participantes segundo o uso de maconha 23
Tabela 5 – Média e desvio padrão das escalas de autoeficácia 24
Tabela 6 – Média e desvio padrão das escalas de práticas parentais 24
Tabela 7 – Associação entre consumo de álcool e sexo 25
Tabela 8 – Associação entre consumo de álcool e escolaridade da mãe 25
Tabela 9 – Associação entre consumo de álcool e Bolsa Família 25
Tabela 10 – Associação entre consumo de álcool e planos para o futuro 26
Tabela 11 – Associação entre consumo de álcool e outras drogas na vida 26
Tabela 12 – Associação entre consumo de álcool e outras drogas no último ano 27
Tabela 13 – Associação entre consumo de álcool e outras drogas no último mês 27
Tabela 14 – Associação entre consumo de álcool e práticas parentais 28
Tabela 15 – Associação entre consumo de álcool e autoeficácia 29
Tabela 16 – Variáveis explicativas 30
Tabela 17 – Modelo 1 - Uso de álcool na vida 30
Tabela 18 – Modelo 2 - Uso de álcool no último ano 31
Tabela 19 – Modelo 3 - Uso de álcool no último mês 32
Tabela 20 – Modelo 4 - *Binge* alcoólico no último ano 32
Tabela 21 – Associação das variáveis com o uso de álcool 33
Tabela 22 – Teste de Hosmer-Lemeshow para os quatro modelos 33
Tabela 23 – Tabela de classificação 34
Tabela A.1 – Intervalo de confiança - Modelo 1 46
Tabela A.2 – Intervalo de confiança - Modelo 2 46
Tabela A.3 – Intervalo de confiança - Modelo 3 46
Tabela A.4 – Intervalo de confiança - Modelo 4 46

Sumário

1 INTRODUÇÃO 6

2 OBJETIVOS 8
2.1 Objetivo geral 8
2.2 Objetivos específicos 8

3 METODOLOGIA 9
3.1 Regressão logística simples 9
3.1.1 Estimação dos parâmetros 10
3.1.2 Interpretação dos parâmetros 11
3.1.3 Inferências sobre o modelo 12
3.1.4 Intervalos de confiança 13
3.2 Regressão logística múltipla 13
3.2.1 Inferências sobre o modelo 14
3.2.2 Intervalos de confiança 14
3.3 Seleção de variáveis 15
3.4 Diagnóstico do modelo 15
3.4.1 Análise de resíduos 16
3.4.2 Testes de adequabilidade de ajustamento 17
3.4.3 Tabela de classificação e curva ROC 18
3.4.4 Análise de influência 18

4 RESULTADOS 20
4.1 Análise descritiva 20
4.2 Associações 24
4.3 Modelagem 29
4.3.1 Qualidade dos modelos 33
4.3.2 Análise dos resíduos 35
4.3.3 Observações Influentes 37

5 CONCLUSÃO 42

REFERÊNCIAS 44

APÊNDICE A – INTERVALO DE CONFIANÇA PARA RAZÃO DE CHANCES 46

1 Introdução

Na adolescência o indivíduo encontra-se em um período de desenvolvimento biopsicossocial, sendo, portanto, mais vulnerável aos efeitos das drogas, incluindo o álcool (BOUZAS, 2007). A construção da identidade, a inserção em grupos sociais e os conflitos familiares são alguns dos motivos que instigam os jovens ao uso das drogas lícitas e/ou ilícitas. No Brasil, a venda ou entrega de bebidas alcoólicas e outros produtos que possam causar dependência física ou psíquica em menores de idade é proibida por lei, mas as falhas na fiscalização contribuem para o consumo dessas substâncias.

Os dados de um levantamento realizado em 2010 no Brasil, com estudantes do ensino fundamental e médio, indicaram que 5,7% já usaram maconha, 8,7% solventes/inalantes, 16,9% cigarro e 60,5% álcool (CARLINI et al., 2010). Em um estudo semelhante nos Estados Unidos os resultados mostraram que entre os estudantes da 8ª série 7,7% já usaram inalantes, 12,8% maconha, 9,8% cigarro e 22,8% álcool (JOHNSTON et al., 2017). A experimentação dessas drogas não implica necessariamente em um uso contínuo, mas o contato precoce é um alerta para este comportamento. O cigarro e a maconha são comumente vistos como menos ofensivos pelos adolescentes, e os inalantes são substâncias mais acessíveis, por isso essas drogas são usadas nessa fase da vida por parte dos adolescentes. Nesse mesmo levantamento feito no Brasil, 42,4% afirmaram ter bebido no último ano e 21,1% no último mês anterior à pesquisa, esses resultados apontam para uma maior ocorrência do uso de álcool.

Os fatores sociodemográficos e familiares estão associados ao consumo de bebidas alcoólicas, dessa forma eles assumem papel importante como preditor para o uso frequente. Em uma pesquisa desenvolvida no estado do Paraná, observou-se que 66,39% dos estudantes entre 8 e 14 anos já haviam experimentado bebidas alcoólicas, o que sugere uma iniciação de consumo cada vez mais cedo (ALAVARSE e CARVALHO, 2006).

O uso regular de álcool entre os adolescentes é caracterizado pela frequência e a quantidade consumida habitualmente. O Primeiro Levantamento Nacional sobre os Padrões de Consumo de Álcool na População Brasileira mostrou que 13,0% do total dos adolescentes pesquisados, com idade entre 14 e 17 anos, apresentou padrão intenso de consumo de álcool (LARANJEIRA et al., 2007). A ocorrência de episódio de bebedeira, ou seja, o consumo excessivo de álcool em uma única ocasião, agrava a exposição do adolescente a situações de risco, como acidentes, brigas e violência sexual. Ainda nesse levantamento verificou-se que 30,0% dos jovens tiveram episódios de bebedeira 2 vezes por mês ou mais, apresentando regularidade no consumo de grandes quantidades. Assim, o comportamento abusivo de bebidas alcoólicas de forma precoce e contínua aumenta a chance de dependência. Consequentemente, o adolescente que apresenta quadro de dependência alcoólica está mais propenso a ter problemas de saúde, conflitos interpessoais, baixo desempenho escolar, entre outros prejuízos.

Diante deste contexto que estão inseridas políticas públicas que visam prevenir o uso de álcool e outras drogas fornecendo informações a fim de modificar o comportamento dos jovens em relação à essas substâncias. O programa Famílias Fortes [1] é um exemplo desta ação que tem uma proposta de intervenção com famílias com objetivo de reduzir os fatores de risco ao uso e abuso de substâncias por adolescentes, construir e fortalecer os vínculos familiares.

Os fatores de risco são aqueles que atuam como motivadores ou facilitadores do uso/abuso de drogas, enquanto os fatores de proteção são os que protegem o adolescente contra o abuso de drogas (KHOURI, 2016). Dois constructos relacionados a esses fatores são as práticas parentais, que se refere às ações dos pais para educar e socializar os filhos, e a autoeficácia que pode ser entendida como a percepção do adolescente sobre sua capacidade de realizar determinadas tarefas e tomar certas decisões.

Dessa forma, o presente estudo visa analisar o padrão de uso de álcool entre jovens de 10 a 15 anos e avaliar o perfil dos jovens quanto ao consumo de álcool através dos elementos que estruturam suas vidas, como características sociodemográficas, relações familiares e comportamentos sociais que agem como fatores de risco ou proteção.

[1] http://prevencaoepesquisa.icict.fiocruz.br

2 Objetivos

2.1 Objetivo geral

Identificar os principais fatores associados ao consumo de álcool entre crianças e adolescentes de uma amostra selecionada no Distrito Federal e na cidade de Goiânia.

2.2 Objetivos específicos

- Avaliar o padrão de consumo de álcool por meio de uma análise descritiva dos dados;
- Analisar as associações entre o uso de álcool e práticas parentais e autoeficácia;
- Aplicar o modelo logístico para identificar os fatores associados ao consumo de álcool;
- Avaliar a qualidade do modelo logístico proposto.

3 Metodologia

A regressão logística é uma técnica usada para estudar a associação entre uma variável resposta categórica e uma ou mais variáveis independentes categóricas ou numéricas. Nesse estudo a variável resposta de interesse é dicotômica, por isso será abordado a modelagem para esse caso específico.

3.1 Regressão logística simples

Seja Y uma variável resposta binária que assume valores 1 para sucesso e 0 para fracasso. A distribuição de Y é dada pela probabilidade de sucesso $P(Y = 1) = \pi$ e probabilidade de fracasso $P(Y = 0) = 1 - \pi$.

Em um modelo de regressão linear assumimos que o valor esperado de Y dado o valor de x pode ser expresso com a seguinte equação linear

$$E(Y|x) = \beta_0 + \beta_1 x. \tag{3.1}$$

Considerando que a variável resposta Y tem distribuição Bernoulli (π), então

$$E(Y) = \pi. \tag{3.2}$$

Igualando (2.2) e (2.1), tem-se que

$$E(Y) = \pi = \beta_0 + \beta_1 x.$$

Essa igualdade viola as suposições do modelo linear, pois

(i) Os erros não são normais;

(ii) A variância dos erros é heterogênea;

(iii) A resposta média é obtida em termos de probabilidade, então $0 \leq \beta_0 + \beta_1 x \leq 1$. Entretanto, esta restrição é inapropriada para resposta em um modelo linear, que assume valores no intervalo $(-\infty, +\infty)$.

Dessa forma, um modelo linear não é adequado quando se tem uma variável resposta binária, por isso o uso da regressão logística. Definimos por $\pi(x)$ a probabilidade de sucesso para um dado valor de x. A função resposta para um modelo de regressão logística simples é

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}. \tag{3.3}$$

A transformação logito definida na equação (2.4), em termos de $\pi(x)$, é a função de ligação que relaciona a média da variável resposta às suas variáveis preditoras. Essa transformação

garante que $g(x)$ não está limitada no intervalo $[0, 1]$, podendo ser qualquer número real e possui linearidade em seus parâmetros

$$g(x) = ln\left[\frac{\pi(x)}{1-\pi(x)}\right] = \beta_0 + \beta_1 x. \tag{3.4}$$

Uma característica dos modelos com variável resposta binária é que os erros têm distribuição binomial. Quando $y = 1$ então $\varepsilon = 1 - \pi(x)$ com probabilidade $\pi(x)$, e quando $y = 0$ então $\varepsilon = -\pi(x)$ com probabilidade $1 - \pi(x)$. Assim, o termo ε tem distribuição com média 0 e variância igual a $\pi(x)[1 - \pi(x)]$.

3.1.1 Estimação dos parâmetros

Para ajustar um modelo de regressão deve-se estimar os parâmetros β do modelo. O método da máxima verossimilhança é usado para estimação desses parâmetros e consiste na seleção de valores que maximizam a probabilidade de se obter a amostra observada. Para aplicar esse método deve-se construir primeiro a função verossimilhança.

Sabendo que Y assume os valores 0 ou 1, então a expressão para $\pi(x)$ dada na equação (2.3) fornece a probabilidade condicional de Y para esses valores, $P(Y = 1|x)$ e $P(Y = 0|x)$. Dessa forma, quando $y_i = 1$ a contribuição para a função verossimilhança é $\pi(x_i)$ e quando $y_i = 0$ a contribuição para a função verossimilhança é $1 - \pi(x_i)$, sendo que $\pi(x_i)$ denota o valor de $\pi(x)$ calculado em x_i. Através da expressão abaixo podemos mostrar a contribuição para a função verossimilhança para o par (x_i, y_i)

$$\pi(x_i)^{y_i}[1 - \pi(x_i)]^{1-y_i}. \tag{3.5}$$

Assumindo que as observações são independentes, a função de verossimilhança é obtida pelo produto dos termos da expressão (2.5)

$$L(\beta) = \prod_{i=1}^{n} \pi(x_i)^{y_i}[1 - \pi(x_i)]^{1-y_i}. \tag{3.6}$$

Pelo princípio da máxima verossimilhança, a estimativa de β para o modelo de regressão logística simples é o valor que maximiza a expressão dada na equação (2.6). Entretanto, matematicamente é preferível que se trabalhe com o logaritmo dessa equação. Assim, a log-verossimilhança é definida como

$$l(\beta) = ln[L(\beta)] = \sum_{i=1}^{n} \{y_i ln[\pi(x_i)] + (1 - y_i) ln[1 - \pi(x_i)]\}. \tag{3.7}$$

Derivando $l(\beta)$ em relação a β_0 e β_1 e igualando a zero obtemos o valor de β que maximiza $l(\beta)$. As equações de verossimilhança encontradas são

$$\sum_{i=1}^{n} [y_i - \pi(x_i)] = 0 \tag{3.8}$$

e

$$\sum_{i=1}^{n} x_i \left[y_i - \pi(x_i)\right] = 0. \tag{3.9}$$

Na regressão logística as equações de verossimilhança definidas acima são não-lineares em β_0 e β_1, por isso requerem métodos iterativos para a sua solução que podem ser vistos no texto de McCullagh e Nelder (1989).

Com a solução dessas equações tem-se a estimativa da máxima verossimilhança denotada por $\hat{\beta}$, a partir dessas estimativas pode-se encontrar os valores ajustados do modelo dados por

$$\hat{\pi}(x) = \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 x}}{1 + e^{\hat{\beta}_0 + \hat{\beta}_1 x}}$$

e o valor estimado do logito $\hat{g}(x)$ fazendo as devidas substituições.

3.1.2 Interpretação dos parâmetros

No modelo de regressão logística a chance (*odds*) de sucesso de determinado evento pode ser calculada através da razão entre a probabilidade de sucesso e a probabilidade de fracasso e é definida como

$$\frac{\pi(x)}{1 - \pi(x)} = e^{\beta_0 + \beta_1 x}.$$

Para comparar a chance de sucesso de um evento para diferentes níveis de x podemos calcular a razão de chances (*odds ratio*) dada por

$$\frac{e^{\beta_0 + \beta_1 (x+1)}}{e^{\beta_0 + \beta_1 x}} = e^{\beta_1}.$$

A função *odds ratio* acima fornece os seguintes resultados:

(i) se $\beta_1 = 0$, então a chance de ocorrer sucesso nos níveis $x + 1$ e x é a mesma;

(ii) se $\beta_1 < 0$, então a *odds ratio* < 1 o que significa que a chance de sucesso no nível $x + 1$ é menor que a chance de sucesso no nível x;

(iii) se $\beta_1 > 0$, então a *odds ratio* > 1 o que significa que a chance de sucesso no nível $x + 1$ é maior que a chance de sucesso no nível x.

Quando x é uma variável contínua, a razão de chances fornece a variação da chance para cada aumento de uma unidade em x. Por exemplo, considerando a variável resposta consumo de álcool e a idade (em anos) como variável x, se $e^{\beta_1} = 1,65$, então a chance de uso de álcool aumenta 65% a cada aumento de um ano de vida. A variável independente x também pode ser qualitativa dicotômica ou politômica. No caso em que ela é dicotômica a razão de chances demonstra quantas vezes é mais provável ocorrer sucesso comparando as duas categorias de x. Assim, se x for a variável sexo, tendo o sexo masculino como referência, e $e^{\beta_1} = 0,56$, então a chance de uma menina consumir bebida alcoólica é 44% menor quando comparado com a chance

de um menino consumir bebida alcoólica. Para uma variável independente com k categorias é necessário o uso de variáveis *dummy* que representem essas categorias. A interpretação da razão de chances dependerá do método escolhido para criação das variáveis *dummy*, por exemplo, se x for a variável escolaridade com as categorias ensino médio incompleto, ensino médio completo e nunca estudou, então escolhe-se uma categoria de referência e a compara com as demais categorias.

3.1.3 Inferências sobre o modelo

A significância da variável explicativa no modelo pode ser verificada pelos testes de Wald ou da Razão de Verossimilhança.

- Teste de Wald

O teste de Wald é obtido pela razão da estimativa do parâmetro β_1 com a estimativa de seu erro padrão. As hipóteses testadas são

$$H_0 : \beta_1 = 0$$

$$H_A : \beta_1 \neq 0.$$

A estatística do teste é dada por

$$W^2 = \left(\frac{\hat{\beta}_1}{EP(\hat{\beta}_1)} \right)^2 \sim \chi^2_{(1)} \text{ para amostras grandes.}$$

A hipótese nula será rejeitada quando $W^2 > \chi^2_{(1-\alpha,1)}$.

- Teste da Razão de Verossimilhança

Esse teste consiste em comparar o modelo completo dado por $\pi(x) = \frac{e^{\beta_0+\beta_1 x}}{1+e^{\beta_0+\beta_1 x}}$ com o modelo reduzido dado por $\pi(x) = \frac{e^{\beta_0}}{1+e^{\beta_0}}$. Dessa forma, as hipóteses são

$$H_0 : \beta_1 = 0$$

$$H_A : \beta_1 \neq 0.$$

A estatística do teste é

$$G^2 = -2\left[lnL_0 - lnL_A\right] \sim \chi^2_{(1)} \text{ para amostras grandes,}$$

em que L_0 é a função de verossimilhança do modelo reduzido e L_A é a função de verossimilhança do modelo completo. Valores elevados da estatística do teste indicam rejeição da hipótese nula.

3.1.4 Intervalos de confiança

Os intervalos são calculados para um nível de confiança $(1-\alpha)\%$ com $\alpha \in (0,1)$, complementando a interpretação dos resultados de um modelo logístico.

O intervalo de confiança para β_0 e β_1 é, respectivamente,

$$\hat{\beta}_0 \pm Z_{1-\frac{\alpha}{2}} \hat{EP}(\hat{\beta}_0) \quad \text{e} \quad \hat{\beta}_1 \pm Z_{1-\frac{\alpha}{2}} \hat{EP}(\hat{\beta}_1).$$

Também é possível calcular um intervalo de confiança para a estimativa do logito, $g(x) = \beta_0 + \beta_1 x$, dado por

$$\hat{g}(x) \pm Z_{1-\frac{\alpha}{2}} \hat{EP}(\hat{g}(x)),$$

em que a variância estimada de $\hat{g}(x)$ é $\hat{Var}[\hat{g}(x)] = \hat{Var}(\hat{\beta}_0) + x^2 \hat{Var}(\hat{\beta}_1) + 2x\hat{Cov}(\hat{\beta}_0, \hat{\beta}_1)$.

O valor estimado do logito e seu intervalo de confiança fornecem o intervalo de confiança para a probabilidade $\pi(x)$

$$\frac{e^{\hat{g}(x) \pm Z_{1-\frac{\alpha}{2}} \hat{EP}(\hat{g}(x))}}{1 + e^{\hat{g}(x) \pm Z_{1-\frac{\alpha}{2}} \hat{EP}(\hat{g}(x))}}.$$

Outro importante intervalo de confiança é o da razão de chances dado da seguinte forma

$$\left[e^{\hat{\beta}_1 - Z_{1-\frac{\alpha}{2}} \hat{EP}(\hat{\beta}_1)}; e^{\hat{\beta}_1 + Z_{1-\frac{\alpha}{2}} \hat{EP}(\hat{\beta}_1)} \right].$$

3.2 Regressão logística múltipla

Na regressão logística múltipla a variável resposta de interesse é explicada por mais de uma variável independente. Alguns procedimentos adotados para o modelo simples são generalizados para o caso em que há mais de uma variável explicativa no modelo.

O modelo regressivo, considerando p variáveis explicativas, é dado por

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p}}.$$

A transformação logito obtida na regressão logística simples estende-se para o modelo múltiplo, assim tem-se que

$$ln\left[\frac{\pi(x)}{1-\pi(x)}\right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p.$$

Os parâmetros na regressão múltipla também são estimados pelo método da máxima verossimilhança. A função log-verossimilhança é definida como

$$ln[L(\beta_0, \beta_1, \beta_2, ..., \beta_p)] = \sum_{i=1}^{n} y_i(\beta_0 + \beta_1 + \beta_2 + ... + \beta_p) - \sum_{i=1}^{n} ln(1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p}).$$

3.2.1 Inferências sobre o modelo

Após o ajuste do modelo é necessário verificar se as variáveis explicativas são significativas. Assim como na regressão simples, o teste da Razão de Verossimilhança permite essa avaliação do modelo.

O teste baseia-se na comparação de modelos completos com modelos reduzidos, ou seja, é possível realizar um teste global onde todos os parâmetros são incluídos no modelo e também testar se determinados β_j são iguais a zero. Dessa maneira as hipóteses testadas são

$$H_0 : \beta_1 = \beta_2 = ... = \beta_p = 0$$

$$H_A : \text{Existe pelo menos um } \beta_j \text{ diferente de 0, } j = 1, ..., p.$$

A estatística do teste definida abaixo segue uma distribuição Qui-Quadrado com graus de liberdade iguais à diferença entre o número de parâmetros no modelo reduzido e no modelo completo para grandes amostras

$$G^2 = -2\left[lnL_0 - lnL_A\right].$$

O teste de Wald também pode ser aplicado em um modelo múltiplo para avaliar quais variáveis são estatisticamente significativas no modelo. Tem-se como hipóteses

$$H_0 : \beta_j = 0$$

$$H_A : \beta_j \neq 0.$$

A estatística de Wald é obtida da seguinte forma

$$W^2 = \left(\frac{\hat{\beta}_j}{EP(\hat{\beta}_j)}\right)^2 \sim \chi^2_{(1)} \text{ para grandes amostras.}$$

Em ambos os testes, valores elevados da estatística do teste indicam rejeição da hipótese nula.

3.2.2 Intervalos de confiança

Os métodos utilizados na construção dos intervalos de confiança na regressão simples estendem-se para o caso múltiplo.

O intervalo de confiança para cada β_j baseado no teste de Wald é definido como

$$\hat{\beta}_j \pm z_{1-\frac{\alpha}{2}} EP(\hat{\beta}_j).$$

A variância estimada de $\hat{g}(x)$ no modelo múltiplo é dada por

$$\hat{Var}[\hat{g}(x)] = \sum_{j=0}^{p} x_j^2 \hat{Var}(\hat{\beta}_j) + \sum_{j=0}^{p} \sum_{k=j+1}^{p} 2x_j x_k \hat{Cov}(\hat{\beta}_j, \hat{\beta}_k).$$

Assim, o intervalo de confiança é estabelecido da seguinte forma

$$\hat{g}(x) \pm z_{1-\frac{\alpha}{2}} EP(\hat{g}(x)).$$

Com base no valor estimado do logito tem-se o intervalo de confiança para os valores ajustados

$$\frac{e^{\hat{g}(x) \pm Z_{1-\frac{\alpha}{2}} \hat{EP}(\hat{g}(x))}}{1 + e^{\hat{g}(x) \pm Z_{1-\frac{\alpha}{2}} \hat{EP}(\hat{g}(x))}}.$$

O intervalo de confiança para a razão de chances é obtido considerando os seguintes limites

$$\left[e^{\hat{\beta}_j - Z_{1-\frac{\alpha}{2}} EP(\hat{\beta}_j)}; e^{\hat{\beta}_j + Z_{1-\frac{\alpha}{2}} EP(\hat{\beta}_j)} \right].$$

3.3 Seleção de variáveis

Nos casos em que há p variáveis explicativas para um modelo é possível construir subconjuntos de modelos com as variáveis regressoras que produzem os melhores resultados. A construção desses subconjuntos baseia-se na adição ou eliminação das variáveis preditoras do modelo.

Existem diferentes métodos para selecionar as variáveis explicativas mais adequadas para um modelo regressivo, entre eles tem-se os critérios de seleção e a seleção automática. O primeiro considera todos os possíveis subconjuntos do modelo e escolhe o melhor de acordo com algum critério. Os critérios mais utilizados são Critério de Informação de Akaike (AIC) e o Critério Bayesiano de Schwarz (SBC). Para a escolha do melhor modelo considera-se os menores valores de AIC ou SBC, dados por

$$AIC = -2lnL(b) + 2p$$

$$SBC = -2lnL(b) + pln(n).$$

O segundo método faz uma seleção automática das melhores variáveis explicativas sem considerar todos os subconjuntos possíveis. Para realizar a seleção automática um dos procedimentos mais utilizados em modelos logísticos é o *stepwise*, que seleciona as variáveis que mais produzem mudança na log-verossimilhança, uma vez que a significância dos parâmetros é assegurada pelo teste da Razão de Verossimilhança. Esse método inicia-se com um modelo apenas com o intercepto, nos passos seguintes as variáveis vão sendo adicionadas ou excluídas ao comparar a log-verossimilhança do modelo atual com a do passo anterior. O modelo final gerado pelo método será considerado o mais adequado, porém esse método pode ocultar outros modelos apropriados dado que ele não considera todos os subconjuntos possíveis.

3.4 Diagnóstico do modelo

Após o ajuste do modelo logístico proposto é importante avaliar a qualidade do ajustamento, identificar observações discrepantes, influentes e verificar o seu poder de discriminação.

3.4.1 Análise de resíduos

O i-ésimo resíduo calculado pela diferença entre os valores observados e os valores preditos no modelo logístico dicotômico pode ser definido como

$$e_i = \begin{cases} 1 - \hat{\pi}_i, & se \;\; y_i = 1 \\ -\hat{\pi}_i, & se \;\; y_i = 0 \end{cases}$$

Os resíduos nessa forma não permitem detectar *outliers*, por isso utiliza-se os resíduos de Pearson ou *deviance* para tal análise.

- Resíduos de Pearson

  Os resíduos de Pearson são definidos como

$$r_{p_j} = \frac{y_j - m_j\hat{\pi}_j}{\sqrt{m_j\hat{\pi}_j(1 - \hat{\pi}_j)}},$$

sendo que $m_j\hat{\pi}_j$ é igual a $\hat{y}_j$ e representa os valores ajustados calculados para cada combinação distinta das variáveis regressoras, e m_j é o número de observações com os mesmos valores das variáveis regressoras.

Os resíduos de Pearson compõem a estatística Qui-Quadrado de Pearson do teste de adequabilidade de ajustamento, pois a soma de quadrados dos resíduos de Pearson é igual a estatística do teste.

- Resíduos de Pearson studentizados

  Os resíduos de Pearson studentizados são definidos como

$$r_{sp_j} = \frac{r_{p_j}}{\sqrt{1 - h_j}},$$

em que h_j é o j-ésimo elemento da diagonal da matriz $\hat{H}$.

- Resíduos *deviance*

  Os resíduos *deviance* são definidos como

$$dev_j = \pm\sqrt{2\left[y_j ln\left(\frac{y_j}{m_j\hat{\pi}_j}\right) + (m_j - y_j)ln\left(\frac{m_j - y_j}{m_j(1 - \hat{\pi}_j)}\right)\right]}$$

que terá o mesmo sinal de $(y_j - m_j\hat{\pi}_j)$. A soma de quadrados desses resíduos é igual a estatística do teste *Deviance* de adequabilidade.

Valores extremos dos resíduos sugerem que o modelo não está bem ajustado, portanto as inferências sobre o modelo podem estar comprometidas.

- Gráficos para os resíduos

Alguns gráficos podem contribuir para a validação da adequabilidade do modelo logístico proposto. No caso do gráfico de resíduos *versus* as probabilidades estimadas com *lowess smooth* espera-se uma linha aproximadamente horizontal com intercepto zero quando o modelo é adequado, pois se o modelo está bem ajustado então $E(y_i - \hat{\pi}_i) = E(e_i) = 0$. Portanto, se o comportamento da linha for diferente o modelo pode não ser adequado.

O gráfico semi-normal de probabilidades com envelope simulado é uma ferramenta utilizada para detectar resíduos *deviance* que sejam discrepantes e analisar se o logito do modelo é adequado. Caso o modelo esteja bem ajustado os resíduos estarão dentro de uma faixa criada a partir do envelope simulado.

3.4.2 Testes de adequabilidade de ajustamento

Os testes de Pearson, *Deviance* e Hosmer-Lemeshow são métodos utilizados para verificação da qualidade de ajuste do modelo aos dados.

- Teste Qui-Quadrado de Pearson

Para a realização desse teste supõe-se a presença de observações com os mesmos valores para as variáveis preditoras. As hipóteses testadas são

$$H_0 : E(Y) = [1 + exp(-X'\beta)]^{-1}$$

$$H_A : E(Y) \neq [1 + exp(-X'\beta)]^{-1}$$

e a estatística do teste é dada por

$$\chi^2 = \sum_{j=1}^{J} r_{p_j}^2 \sim \chi^2_{J-(p+1)},$$

em que J indica o número de valores distintos para cada variável regressora. A hipótese nula será rejeitada quando $\chi^2 > \chi^2_{(1-\alpha, J-(p+1))}$.

- Teste *Deviance*

Nesse teste também é necessário identificar repetições entre as observações para todas as variáveis preditoras. As hipóteses testadas são as mesmas do teste citado anteriormente e a estatística do teste é dada por

$$G^2 = \sum_{j=1}^{J} dev_j^2 \sim \chi^2_{J-(p+1)}.$$

Valores elevados da estatística do teste indicam a rejeição da hipótese nula.

- Teste de adequabilidade de ajustamento de Hosmer-Lemeshow

Esse teste propõe o agrupamento das probabilidades estimadas em classes de tamanho aproximadamente iguais para observar se há diferença estatisticamente significativa entre os valores observados e esperados. Duas estratégias podem ser seguidas para agrupar os dados, a primeira consiste em usar os percentis das probabilidades estimadas e a outra opção é fixar valores para a probabilidade estimada. A estatística de teste é mesma do teste Qui-Quadrado de Pearson para tabelas de frequências esperadas e observadas com $c - 2$ graus de liberdade. Um modelo adequado apresenta valores baixos para a estatística de teste.

3.4.3 Tabela de classificação e curva ROC

Os resultados de uma regressão logística podem ser resumidos em uma tabela de classificação que mostra quantas observações foram classificadas corretamente de acordo com o modelo proposto a partir de um ponto de corte. Uma forma de definir esse ponto de corte é através da curva ROC, que permite identificar a variação da sensibilidade e especificidade do modelo.

A sensibilidade é definida como $P(\hat{Y} = 1|Y = 1)$, ou seja, é a probabilidade de um indivíduo ser classificado como evento dado que ele é evento. Enquanto a especificidade é definida como $P(\hat{Y} = 0|Y = 0)$, que é a probabilidade de um indivíduo ser classificado como não evento dado que ele é não evento. Portanto, pode-se escolher o ponto de corte pela combinação ótima dessas medidas.

O gráfico da curva ROC produz a área abaixo da curva (AUC - *Area Under the Curve*) que é uma medida de discriminação do modelo. Quanto maior a AUC maior é o poder de classificação do modelo. Uma forma de definir a discriminação de um modelo (HOSMER e LEMESHOW, 2000) é

$$\begin{cases} \text{se AUC = 0,5: não há discriminação} \\ \text{se 0,7 < AUC < 0,8: discriminação aceitável} \\ \text{se 0,8 < AUC < 0,9: discriminação excelente} \\ \text{se AUC > 0,9: discriminação excepcional} \end{cases}$$

3.4.4 Análise de influência

As medidas apresentadas a seguir são usadas para detectar observações mal ajustadas e aquelas que influenciam os valores ajustados do modelo logístico.

- DFCHISQ

Essa medida é calculada a partir da diferença entre a estatística χ^2 dos dados completos e a estatística χ_j^2 quando a observação j é excluída dos dados

$$\Delta\chi_j^2 = r_{sp_j}^2.$$

- DFDEV

A medida DFDEV é calculada considerando a mudança na estatística *deviance* quando a observação j é excluída do conjunto de dados

$$\Delta dev_j = \frac{dev_j^2}{1 - h_j}.$$

- DFBETA

Essa medida é obtida pela diferença entre o parâmetro estimado do modelo e o parâmetro quando a observação j é excluída dos dados

$$\Delta \hat{\beta}_j = \frac{r_{sp_j}^2 h_j}{1 - h_j}.$$

- Diagonal da matriz $\hat{H}$

A matriz $\hat{H}$ é dada por

$$H = \hat{W}^{\frac{1}{2}} X (X' \hat{W} X)^{-1} X' \hat{W}^{\frac{1}{2}},$$

em que $\hat{W}$ é a matriz diagonal com os elementos $m_j \hat{\pi}_j (1 - \hat{\pi}_j)$ e X é a matriz de delineamento das variáveis explicativas. O j-ésimo elemento da diagonal da matriz $\hat{H}$ é definido como o ponto de alavanca da observação j

$$h_j = m_j \hat{\pi}_j (1 - \hat{\pi}_j) x_j' (X' W X)^{-1} x_j.$$

Essas medidas calculam o impacto de cada observação na regressão logística, portanto valores elevados sugerem influência no modelo. A identificação de pontos influentes é importante, pois eles podem viesar as estimativas dos coeficientes e consequentemente superestimar ou subestimar as inferências sobre a razão de chances.

4 Resultados

4.1 Análise descritiva

A amostra desse estudo foi composta por estudantes da rede pública de ensino que responderam um questionário utilizado como instrumento de avaliação da efetividade de um programa social (MURTA et al., 2015). Esse questionário avalia características sociodemográficas, padrão de uso e problemas decorrentes do uso de drogas, práticas parentais, habilidades de autocontrole, habilidades sociais assertivas e de autoeficácia, e perspectiva de tempo futuro.

Os participantes tinham de 10 a 15 anos de idade, sendo 381 meninas (54,5%), 307 meninos (43,9%) e 11 (1,6%) que não informaram o sexo. A Figura 1 mostra que a maioria dos alunos cursavam o 6º e 7º ano do ensino fundamental, 28,5% e 34,8% respectivamente.

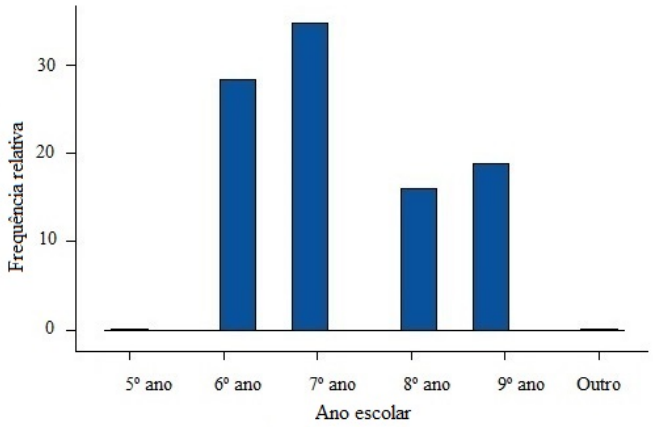

Figura 1: Distribuição dos participantes segundo o ano escolar

Uma outra característica da amostra refere-se à composição familiar. Dentre as pessoas que residem com os adolescentes observou-se a presença da mãe (88,7%) e de irmãos (75,2%) na maioria dos casos, enquanto a presença do pai acontecia em 58,6% dos lares. A Figura 2 mostra que cerca de 27% das mães ou responsáveis concluíram o ensino médio, mas não terminaram a faculdade e 24,3% não souberam informar a escolaridade dos responsáveis.

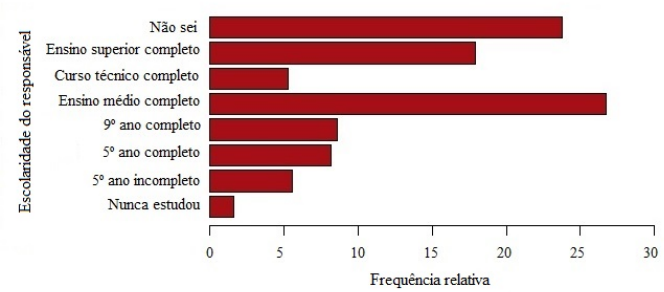

Figura 2: Distribuição dos participantes segundo a escolaridade da mãe

Observou-se ainda que 58,9% das famílias não participavam do programa Bolsa Família, 25,0% recebiam o auxílio e 16,1% dos adolescentes não sabiam informar se a família participava ou não do programa. Em relação aos planos para o futuro verifica-se, na Figura 3, que 75,4% dos adolescentes mostraram interesse em fazer faculdade, 14,8% não sabiam ainda o que fazer após concluir o ensino médio e 9,8% tinham interesses diversos, que incluíam fazer curso técnico, trabalhar e outros planos.

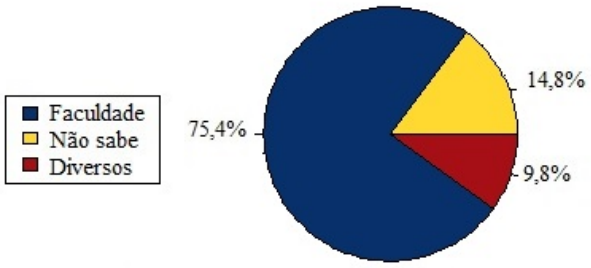

Figura 3: Distribuição dos participantes segundo os planos para o futuro

Sobre o uso de álcool tem-se que 47,6% dos jovens já experimentaram bebida alcoólica (Tabela 1), entre esses 25% bebeu a primeira vez aos 12 anos. A idade média para o primeiro consumo foi de 11,58 anos, sendo que a idade mínima identificada foi de 6 anos e a idade máxima de 15 anos. Após ingerir álcool pela primeira vez 34,2% responderam que nunca mais beberam, os demais afirmaram beber algumas vezes, seja em datas especiais (Natal, fim de ano, carnaval), festas ou saídas com amigos. Nos últimos 12 meses que precederam a pesquisa, 32,0% dos participantes disseram ter ingerido algum tipo de bebida alcoólica e 13,9% bebeu nos últimos 30 dias (Tabela 1).

Também foi estudado o uso pesado de álcool, em que observou-se a ingestão de cinco doses ou mais de álcool em uma única ocasião. Os resultados mostram a ocorrência de *binge* alcoólico na vida para 14,2% dos jovens, enquanto para a ocorrência no último ano observou-se 13,3% de casos e 15,2% para a ocorrência no último mês. O registro de maior ocorrência de *binge* alcoólico no último mês indica inconsistência nos dados, pois não se espera que o uso no mês seja maior que o uso no ano (Tabela 1).

Tabela 1: Distribuição dos participantes segundo o consumo de álcool

Uso de álcool	Resposta	Total
Na vida	Não	356 (52,4%)
	Sim	324 (47,6%)
No último ano	Não	463 (68,0%)
	Sim	218 (32,0%)
No último mês	Não	583 (85,9%)
	Sim, 1 a 5 dias	70 (10,3%)
	Sim, 6 a 19 dias	11 (1,6%)
	Sim, 20 dias ou mais	14 (2,0%)
Binge alcoólico na vida	Não	589 (85,8%)
	Sim	97 (14,2%)
Binge alcoólico no último ano	Não	587 (86,7%)
	Sim	90 (13,3%)
Binge alcoólico no último mês	Nenhuma vez	577 (84,8%)
	1 vez	40 (5,9%)
	2 vezes	30 (4,4%)
	3 a 5 vezes	16 (2,4%)
	Mais de 5 vezes	17 (2,5%)

A Tabela 2 apresenta os resultados para o uso de cigarro e aponta que 11,0% dos jovens já fumaram alguma vez na vida e 5,3% afirmaram ter fumado no último ano. Nos últimos 30 dias observou-se que 3,1% fez uso de cigarro, sendo que não há ocorrência de uso no intervalo de 6 a 19 dias. Além disso, constatou-se que a quantidade de cigarros fumados por dia varia de 1 a 10 entre os fumantes.

Tabela 2: Distribuição dos participantes segundo o uso de cigarro

Uso de cigarro	Resposta	Total
Na vida	Não	610 (89,0%)
	Sim	76 (11,0%)
No último ano	Não	646 (94,7%)
	Sim	36 (5,3%)
No último mês	Não	665 (96,9%)
	Sim, 1 a 5 dias	17 (2,5%)
	Sim, 6 a 19 dias	(0%)
	Sim, 20 dias ou mais	4 (0,6%)

As Tabelas 3 e 4 apresentam os resultados para o uso das drogas ilícitas inalantes e maconha. Quanto ao uso na vida, o maior índice apresentado está no tipo inalante (19,5%), que abrange o consumo de loló, lança, cola, éter, removedor de tinta, gasolina, entre outros. Esse é um tipo de droga mais acessível, o que pode explicar o consumo entre estes jovens. O uso de cocaína, crack

e merla também foram avaliados e apresentaram baixos percentuais, por isso não fazem parte do estudo. Assim, a segunda maior incidência de uso está na maconha, 6,6% dos participantes já experimentaram. No último ano e no último mês a ocorrência do uso de inalantes também foi maior do que os casos de uso de maconha.

Tabela 3: Distribuição dos participantes segundo o uso de inalantes

Uso de inalantes	Resposta	Total
Na vida	Não	551 (80,5%)
	Sim	133 (19,5%)
No último ano	Não	637 (92,8%)
	Sim	49 (7,2%)
No último mês	Não	660 (96,1%)
	Sim, 1 a 5 dias	21 (3,1%)
	Sim, 6 a 19 dias	3 (0,4%)
	Sim, 20 dias ou mais	3 (0,4%)

Tabela 4: Distribuição dos participantes segundo o uso de maconha

Uso de maconha	Resposta	Total
Na vida	Não	638 (93,4%)
	Sim	45 (6,6%)
No último ano	Não	645 (95,5%)
	Sim	30 (4,5%)
No último mês	Não	667 (97%)
	Sim, 1 a 5 dias	13 (1,9%)
	Sim, 6 a 19 dias	3 (0,4%)
	Sim, 20 dias ou mais	5 (0,7%)

A capacidade dos participantes em expressar opiniões, organizar e executar tarefas escolares e de pedir auxílio está relacionada à autoeficácia e foi medida através da escala Likert de 5 pontos, variando de “com certeza eu não consigo fazer isso” (1) até “com certeza eu consigo fazer isso” (5). Observa-se na Tabela 5 que a habilidade dos jovens em conseguir se expressar e defender seu próprio ponto de vista, chamada eficácia autoassertiva, apresenta maior média (3,90). Para aqueles que acreditam serem capazes de cumprir suas tarefas escolares, autoeficácia para aprendizagem autorregulada, a média foi 3,73. E por último a autoeficácia para conseguir suporte parental, que indica a capacidade para pedir ajuda aos familiares, obteve média 3,49. Esses valores acima de 3, considerado o ponto central, sugerem que os participantes tendem a ter mais certeza quanto a sua capacidade de realizar determinadas atividades.

Tabela 5: Média e desvio padrão das escalas de autoeficácia

Autoeficácia	Média	Desvio Padrão
Autoeficácia para aprendizagem autorregulada	3,73	0,69
Autoeficácia para conseguir suporte parental	3,49	1,01
Eficácia autoassertiva	3,90	0,81

As práticas parentais apresentadas na Tabela 6 também foram medidas pela escala Likert de 5 pontos, variando de "nunca" (1) até "sempre" (5). As dimensões estudadas e suas respectivas médias foram apoio emocional (3,92), cobrança de responsabilidade (4,33), controle punitivo (3,32), incentivo à autonomia (3,24), intrusividade (2,64) e supervisão do comportamento (4,05). As médias mais próximas de 5 indicam maior ocorrência das práticas, dessa forma nota-se que cobrança de responsabilidade e supervisão de comportamento são muito frequentes nas famílias desses jovens, enquanto intrusividade aparece como a prática parental com menor média.

Tabela 6: Média e desvio padrão das escalas de práticas parentais

Prática parental	Média	Desvio Padrão
Apoio emocional	3,92	0,84
Cobrança de responsabilidade	4,33	0,84
Controle punitivo	3,32	1,09
Incentivo à autonomia	3,24	1,11
Intrusividade	2,46	1,07
Supervisão do comportamento	4,05	0,98

4.2 Associações

Para verificar quais variáveis, dentre sexo, escolaridade da mãe, participação no programa Bolsa Família, planos para o futuro, consumo de cigarro, maconha e inalantes, práticas parentais e autoeficácia estão relacionadas significativamente com o uso de álcool, foram realizados testes de associação e comparação de médias considerando 5% de significância para rejeição da hipótese nula.

Na Tabela 7 os resultados do teste Qui-Quadrado para os três tipos de consumo de álcool analisados indicam que não há associação significativa entre as variáveis. Assim, não se pode afirmar que existe diferença no consumo de álcool entre meninas e meninos.

Tabela 7: Associação entre consumo de álcool e sexo

Uso de álcool		Sexo		χ^2	p-valor
		Feminino	Masculino		
Na vida	Não	209	148	2,80	0,094
	Sim	169	155		
No último ano	Não	262	201	0,88	0,347
	Sim	115	103		
No último mês	Não	327	256	0,67	0,412
	Sim	49	46		

A escolaridade da mãe foi organizada em três categorias: não sabe, ensino médio incompleto e ensino médio completo. A categoria "não sabe"refere-se aos adolescentes que por algum motivo não souberam informar a escolaridade da mãe. Os resultados apresentados na Tabela 8 indicam que existe associação significativa entre a escolaridade da mãe e o consumo de álcool na vida e no último ano, enquanto para o uso de álcool no último mês a associação não foi significativa.

Tabela 8: Associação entre consumo de álcool e escolaridade da mãe

Uso de álcool		Escolaridade da mãe			χ^2	p-valor
		Não sabe	E.M. incompleto	E.M. completo		
Na vida	Não	101	95	156	15,16	< 0,001
	Sim	61	72	190		
No último ano	Não	129	114	215	14,07	< 0,001
	Sim	34	53	129		
No último mês	Não	141	139	297	2,01	0,365
	Sim	18	28	48		

Observa-se na Tabela 9 que há associação significativa entre o uso de álcool e a participação no programa Bolsa Família para o consumo na vida e no último ano. Entretanto, não houve associação significativa para o uso de álcool no último mês.

Tabela 9: Associação entre consumo de álcool e Bolsa Família

Uso de álcool		Bolsa Família			χ^2	p-valor
		Sim	Não	Não sabe		
Na vida	Não	186	105	68	15,77	< 0,001
	Sim	217	66	42		
No último ano	Não	248	134	84	20,71	< 0,001
	Sim	156	37	25		
No último mês	Não	340	154	91	4,10	0,128
	Sim	63	16	17		

Na Tabela 10 verifica-se que não existe associação significativa entre planos para o futuro e consumo de álcool nos três padrões de uso analisados.

Tabela 10: Associação entre consumo de álcool e planos para o futuro

Uso de álcool		Planos para o futuro			χ^2	p-valor
		Não sabe	Faculdade	Diversos		
Na vida	Não	37	286	30	3,77	0,151
	Sim	35	234	40		
No último ano	Não	48	360	46	0,53	0,765
	Sim	24	159	24		
No último mês	Não	61	450	58	0,94	0,623
	Sim	10	67	12		

Os resultados do teste para medir as associações entre o uso de álcool e as outras drogas aqui estudadas (cigarro, inalante e maconha) são apresentados nas Tabelas 11, 12 e 13. Entre os participantes que ingeriram bebida alcoólica na vida observou-se maior ocorrência para o uso de inalantes, sendo essa associação significativa. O uso de cigarro e maconha na vida também mostraram-se associados ao consumo de álcool.

Para o uso nos últimos 12 meses há evidências de associação significativa entre o uso de álcool e as três drogas analisadas. Mais uma vez os inalantes apresentaram maior tendência de uso entre aqueles que consumiram bebidas alcoólicas. Entre os jovens que fizeram uso de álcool nos últimos 30 dias verificou-se uma tendência maior ao uso de maconha e cigarro, sendo que as associações com as três drogas mostraram-se estatisticamente significativas.

Tabela 11: Associação entre consumo de álcool e outras drogas na vida

Droga		Uso de álcool		χ^2	p-valor
		Não	Sim		
Cigarro	Não	354	250	79,8	$< 0{,}001$
	Sim	3	72		
Inalantes	Não	322	223	47,3	$< 0{,}001$
	Sim	34	98		
Maconha	Não	351	280	36,7	$< 0{,}001$
	Sim	4	41		

Tabela 12: Associação entre consumo de álcool e outras drogas no último ano

Droga		Uso de álcool Não	Uso de álcool Sim	χ^2	p-valor
Cigarro	Não	454	185	46,4	$< 0{,}001$
	Sim	6	30		
Inalantes	Não	446	184	33,6	$< 0{,}001$
	Sim	15	34		
Maconha	Não	451	187	38,6	$< 0{,}001$
	Sim	5	25		

Tabela 13: Associação entre consumo de álcool e outras drogas no último mês

Droga		Uso de álcool Não	Uso de álcool Sim	χ^2	p-valor
Cigarro	Não	577	78	70,2	$< 0{,}001$
	Sim	5	16		
Inalantes	Não	568	82	27,1	$< 0{,}001$
	Sim	14	13		
Maconha	Não	579	78	80,6	$< 0{,}001$
	Sim	4	17		

Utilizou-se o teste t de Student para avaliar se há diferença significativa nas médias das práticas parentais e autoeficácia entre adolescentes que ingeriram bebida alcoólica e aqueles que não ingeriram.

Na Tabela 14 observa-se que os participantes que fizeram uso de álcool na vida apresentaram média de apoio emocional significativamente menor ao comparar com aqueles que não ingeriram bebida alcoólica. As médias das práticas parentais incentivo à autonomia e intrusividade mostraram-se estatisticamente diferentes, sendo que os participantes que beberam apresentaram as maiores médias. As demais práticas parentais não indicaram diferença significativa nas médias para o uso de álcool na vida.

Para o consumo de álcool nos últimos 12 meses, as médias de apoio emocional e controle punitivo dos jovens que não ingeriram bebida alcoólica revelaram-se significativamente maiores. Para aqueles que consumiram álcool, as médias de incentivo à autonomia e intrusividade são estatisticamente maiores. Cobrança de responsabilidade e supervisão do comportamento não apresentaram diferença significativa.

No que se refere ao uso de álcool nos últimos 30 dias verificou-se que apoio emocional e intrusividade apontaram diferença significativa entre os grupos analisados, sendo que para a primeira prática parental citada a média é maior para o grupo de jovens que não ingeriram álcool e no segundo caso a relação é inversa, ou seja, os jovens que consumiram álcool apresentaram maior média. Para as outras escalas de práticas parentais não foram identificadas diferenças significativas entre as médias.

Tabela 14: Associação entre consumo de álcool e práticas parentais

Uso de álcool	Práticas parentais	Resposta	Média	t	p-valor
Na vida	Apoio emocional	Não	4,05	3,97	$< 0{,}001$
		Sim	3,80		
	Cobrança de responsabilidade	Não	4,39	1,59	0,110
		Sim	4,28		
	Controle punitivo	Não	3,30	-0,44	0,655
		Sim	3,34		
	Incentivo à autonomia	Não	3,16	-2,12	0,033
		Sim	3,33		
	Intrusividade	Não	2,25	-5,41	$< 0{,}001$
		Sim	2,69		
	Supervisão do comportamento	Não	4,08	0,46	0,642
		Sim	4,04		
No último ano	Apoio emocional	Não	4,02	3,78	$< 0{,}001$
		Sim	3,75		
	Cobrança de responsabilidade	Não	4,37	1,66	0,097
		Sim	4,25		
	Controle punitivo	Não	3,38	2,29	0,022
		Sim	3,17		
	Incentivo à autonomia	Não	3,18	-2,22	0,026
		Sim	3,37		
	Intrusividade	Não	2,36	-3,22	0,001
		Sim	2,65		
	Supervisão do comportamento	Não	4,09	1,11	0,260
		Sim	4,00		
No último mês	Apoio emocional	Não	3,98	3,79	$< 0{,}001$
		Sim	3,61		
	Cobrança de responsabilidade	Não	4,35	0,94	0,340
		Sim	4,26		
	Controle punitivo	Não	3,31	-0,18	0,850
		Sim	3,33		
	Incentivo à autonomia	Não	3,24	-0,06	0,940
		Sim	3,25		
	Intrusividade	Não	2,39	-3,18	0,001
		Sim	2,82		
	Supervisão do comportamento	Não	4,06	-0,31	0,750
		Sim	4,09		

Conforme observado na Tabela 15, para os três padrões de consumo de álcool não foram identificadas diferenças significativas para eficácia autoassertiva entre os participantes que fizeram ou não uso de bebidas alcoólicas. As médias para aprendizagem autorregulada e autoeficácia para conseguir suporte parental mostraram-se significativamente maiores para aqueles que não ingeriram álcool em todos os padrões de consumo analisados.

Tabela 15: Associação entre consumo de álcool e autoeficácia

Uso de álcool	Autoeficácia	Resposta	Média	t	p-valor
Na vida	Autoeficácia para aprendizagem autorregulada	Não Sim	3,90 3,53	7,06	$< 0{,}001$
	Autoeficácia para conseguir suporte parental	Não Sim	3,73 3,27	6,08	$< 0{,}001$
	Eficácia autoassertiva	Não Sim	3,95 3,87	1,27	0,200
No último ano	Autoeficácia para aprendizagem autorregulada	Não Sim	3,82 3,51	5,76	$< 0{,}001$
	Autoeficácia para conseguir suporte parental	Não Sim	3,63 3,24	5,01	$< 0{,}001$
	Eficácia autoassertiva	Não Sim	3,88 3,96	-1,27	0,201
No último mês	Autoeficácia para aprendizagem autorregulada	Não Sim	3,77 3,45	4,78	$< 0{,}001$
	Autoeficácia para conseguir suporte parental	Não Sim	3,56 3,19	3,89	$< 0{,}001$
	Eficácia autoassertiva	Não Sim	3,90 3,96	-0,83	0,400

4.3 Modelagem

Foram construídos quatro modelos logísticos referentes aos padrões de consumo de álcool. No primeiro a variável resposta é o uso de álcool na vida, no segundo o uso de álcool no último ano, no terceiro o uso de álcool no último mês e por fim, ocorrência de *binge* alcoólico no último ano. As variáveis explicativas consideradas nos modelos são apresentadas na Tabela 16 e apesar de o consumo de cigarro, inalantes e maconha ter se mostrado associado ao uso de álcool, essas variáveis não foram analisadas na regressão logística. A categoria "não sabe"das variáveis Bolsa Família e escolaridade da mãe refere-se aos adolescentes que não sabiam informar sobre a participação da família no programa e o grau de estudo da mãe. Os modelos foram gerados a partir do método *stepwise* considerando 5% de significância para entrada e saída das variáveis.

Tabela 16: Variáveis explicativas

Variável	Descrição
Idade	Anos
Sexo	1 - Feminino (referência) 2 - Masculino
Bolsa Família	1 - Não (referência) 2 - Sim 3 - Não sabe
Escolaridade da mãe	1 - Não sabe 2 - Ensino médio incompleto 3 - Ensino médio completo (referência)
Apoio emocional Intrusividade Incentivo à autonomia Supervisão do comportamento Cobrança de responsabilidade Controle punitivo	Fatores relacionados com o constructo práticas parentais obtidos por meio de Análise Fatorial (COLARES, 2016)
Autoeficácia para aprendizagem autorregulada Eficácia autoassertiva Autoeficácia para suporte parental	Fatores relacionados com o constructo autoeficácia obtidos por meio de Análise Fatorial (VASCONCELOS, 2016)

- Modelo para consumo de álcool na vida

Observa-se na Tabela 17 que a chance de um adolescente ter experimentado bebida alcoólica na vida aumenta 91,1% a cada aumento de um ano de vida. Para os adolescentes que recebem *Bolsa Família* a chance de ter ingerido álcool é 49,2% menor que a chance de um adolescente que não recebe o *Bolsa Família* e para aqueles que não sabem se participam do programa a chance de ter ingerido álcool é 46,6% menor quando comparado com quem não recebe. A cada mudança em um desvio padrão na prática parental *Intrusividade* a chance de ter ingerido bebida alcoólica aumenta 43,8%, enquanto a cada mudança em um desvio padrão em autoeficácia para *Aprendizagem Autorregulada* a chance é 48,8% menor.

Tabela 17: Modelo 1 - Uso de álcool na vida

Variável	Estimativa	Erro Padrão	Wald	p-valor	$\hat{RC}$
Intercepto	-6,437	1,221	27,756	< 0,001	-
Idade	0,647	0,081	62,563	< 0,001	1,911
Bolsa Família					
Sim	-0,676	0,216	9,792	0,001	0,508
Não sabe	-0,625	0,249	6,286	0,012	0,534
Intrusividade	0,363	0,085	18,188	< 0,001	1,438
Aprendizagem autorregulada	-0,667	0,135	24,191	< 0,001	0,512

- Modelo para consumo de álcool no último ano

Para o uso de álcool no último ano (Tabela 18) tem-se que a chance de um adolescente ter ingerido bebida alcoólica aumenta 86,2% a cada aumento de um ano de vida. Ter o auxílio do programa *Bolsa Família* diminui 55,6% a chance de usar álcool quando comparado com quem não recebe esse auxílio, assim como para aqueles que não sabem se recebem o *Bolsa Família* a chance é 48,5% menor do que a chance de quem não recebe. Entre os adolescentes que não sabem a *Escolaridade da Mãe* a chance de ter ingerido bebida alcoólica é 43,2% menor que a chance de adolescentes que têm mãe com ensino médio completo.

A chance de ter ingerido bebida alcoólica no último ano aumenta 28,2% a cada mudança em um desvio padrão na prática parental *Intrusividade*. Para as três medidas de autoeficácia tem-se uma redução de 49,4% na chance de uso de álcool a cada mudança em um desvio padrão em autoeficácia para *Aprendizagem Autorregulada* e uma redução de 21,1% para cada mudança em um desvio padrão em autoeficácia para *Suporte Parental*, por outro lado a chance de uso de álcool aumenta 50,4% a cada mudança em um desvio padrão em *Eficácia Autoassertiva*.

Tabela 18: Modelo 2 - Uso de álcool no último ano

Variável	Estimativa	Erro Padrão	Wald	p-valor	$\hat{RC}$
Intercepto	-7,381	1,391	28,151	< 0,001	-
Idade	0,621	0,090	46,908	< 0,001	1,862
Bolsa Família					
Sim	-0,810	0,257	9,885	0,001	0,444
Não sabe	-0,662	0,289	5,242	0,022	0,515
Escolaridade da mãe					
Não sabe	-0,565	0,260	4,716	0,029	0,568
E.M. incompleto	-0,055	0,240	0,052	0,819	0,946
Intrusividade	0,248	0,091	7,350	0,006	1,282
Aprendizagem autorregulada	-0,680	0,170	15,986	< 0,001	0,506
Eficácia autoassertiva	0,408	0,139	8,589	0,003	1,504
Suporte parental	-0,235	0,105	4,946	0,026	0,789

- Modelo para consumo de álcool no último mês

Na Tabela 19, referente ao uso de álcool no último mês, pode-se observar que a chance de um adolescente ter ingerido bebida alcoólica aumenta 95% a cada aumento de um ano de vida. A cada mudança em um desvio padrão na prática parental *Intrusividade* a chance de uso de álcool aumenta 33,3%. Enquanto a mudança em um desvio padrão em autoeficácia para *Aprendizagem Autorregulada* reduz 39,5% a chance de um adolescente ingerir bebida alcoólica.

Tabela 19: Modelo 3 - Uso de álcool no último mês

Variável	Estimativa	Erro Padrão	Wald	p-valor	$\hat{RC}$
Intercepto	-9,493	1,712	30,743	< 0,001	-
Idade	0,668	0,113	34,411	< 0,001	1,950
Intrusividade	0,288	0,106	7,336	0,006	1,333
Aprendizagem autorregulada	-0,501	0,173	8,363	0,003	0,605

- Modelo para *binge* alcoólico no último ano

Para a ocorrência de *binge* alcoólico no último ano (Tabela 20) tem-se que a chance de um adolescente ter ingerido álcool excessivamente em uma única ocasião aumenta 75,6% a cada aumento de um ano de vida. A cada mudança em um desvio padrão em autoeficácia para *Aprendizagem Autorregulada* a chance de um adolescente ter episódio de bebedeira diminui 35,9%.

Tabela 20: Modelo 4 - *Binge* alcoólico no último ano

Variável	Estimativa	Erro Padrão	Wald	p-valor	$\hat{RC}$
Intercepto	-7,605	1,633	21,681	< 0,001	-
Idade	0,563	0,109	26,310	< 0,001	1,756
Aprendizagem autorregulada	-0,444	0,168	6,940	0,008	0,641

Na Tabela 21 são apresentadas as variáveis explicativas dos modelos finais de acordo com a direção da razão de chances. Deste modo, as variáveis com sinal positivo estão relacionadas com o aumento da chance de uso de álcool e as variáveis com sinal negativo com a redução da chance de uso de álcool. É importante ressaltar que para a variável escolaridade da mãe, a chance de consumo de álcool diminui para níveis mais baixos de escolaridade. Os intervalos de confiança para a razão de chances dos quatro modelos estão no Apêndice A.

Tabela 21: Associação das variáveis com o uso de álcool

Variáveis	Padrão de uso de álcool			
	Uso na vida	Uso no ano	Uso no mês	*Binge* no ano
Idade	+	+	+	+
Bolsa Família	-	-		
Escolaridade da mãe		-		
Intrusividade	+	+	+	
Aprendizagem autorregulada	-	-	-	-
Eficácia autoassertiva		+		
Suporte parental		-		

(+) variável com razão de chances maior que 1
(-) variável com razão de chances menor que 1

4.3.1 Qualidade dos modelos

A Tabela 22 apresenta os resultados do teste de Hosmer-Lemeshow indicando que não existe diferença significativa entre os valores observados e esperados dos modelos. Portanto, pela estatística de teste e seu respectivo p-valor há evidências de que os modelos propostos adequam-se bem aos dados.

Tabela 22: Teste de Hosmer-Lemeshow para os quatro modelos

Modelo	χ^2	GL	p-valor
Uso de álcool na vida	4,52	8	0,80
Uso de álcool no último ano	7,37	8	0,49
Uso de álcool no último mês	7,49	8	0,48
Binge alcoólico no último ano	6,84	8	0,55

O ponto de corte utilizado na construção da tabela de classificação dos modelos (Tabela 23) foi baseado na combinação ótima da sensibilidade e especificidade. No modelo 1, o ponto de corte usado foi 0,49, e tem-se que 70,9% dos indivíduos estão bem classificados, sendo a sensibilidade 71,3% e especificidade 70,5%. No modelo 2, com ponto de corte 0,34, a proporção de classificações corretas é de 72,7%, com 72,4% de sensibilidade e 73,3% de especificidade. Para o modelo 3, com ponto de corte 0,14, observa-se 66,5% de indivíduos classificados corretamente, sendo a sensibilidade e especificidade 65,6% e 72,0%, respectivamente. No modelo 4, para o ponto de corte 0,14, tem-se 63,8% de classificações corretas, com 63,8% de sensibilidade e 64,0% de especificidade.

Tabela 23: Tabela de classificação

			Observado	
Modelo 1			Y = 0	Y = 1
	Estimado	Y = 0	241	95
		Y = 1	97	228
Modelo 2			Observado	
			Y = 0	Y = 1
	Estimado	Y = 0	324	53
		Y = 1	123	146
Modelo 3			Observado	
			Y = 0	Y = 1
	Estimado	Y = 0	376	26
		Y = 1	197	67
Modelo 4			Observado	
			Y = 0	Y = 1
	Estimado	Y = 0	371	32
		Y = 1	210	57

Em um modelo com alto poder de discriminação a curva ROC se aproxima do canto superior esquerdo do gráfico. Na Figura 4 estão dispostos os gráficos da curva ROC para os modelos propostos. A área abaixo da curva no modelo 1 (Figura 4a) é de 77,4%, no modelo 2 (Figura 4b) é de 78,8%, no modelo 3 (Figura 4c) é de 75,4% e no modelo 4 (Figura 4d) é de 70,8%. Pode-se afirmar que os quatro modelos possuem poder de discriminação aceitável, sendo que o modelo para uso de álcool no último ano é o que melhor classifica os indivíduos, enquanto o modelo para *binge* alcoólico no último ano tem o menor poder de classificação.

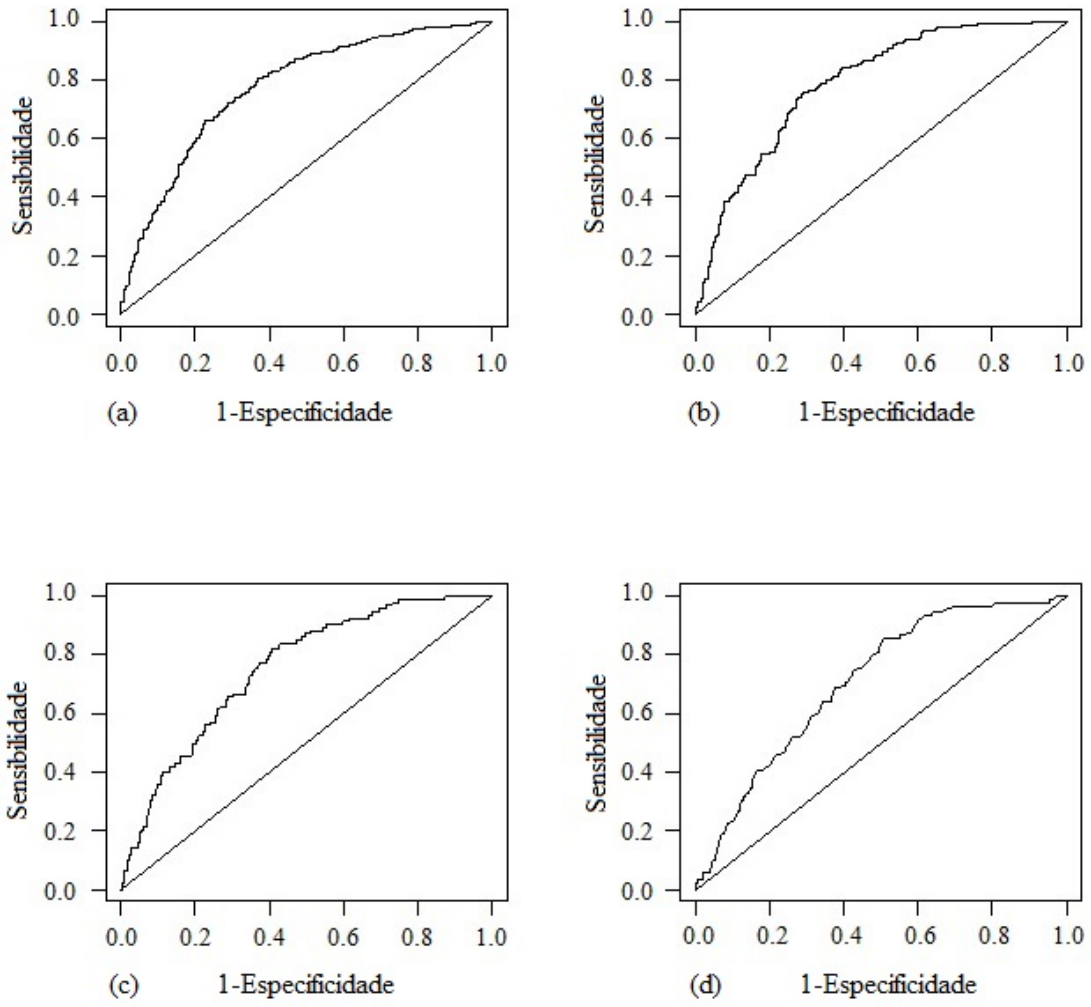

Figura 4: Curva ROC para os quatro modelos

4.3.2 Análise dos resíduos

Nos gráficos dos resíduos *deviance versus* o preditor linear com *lowess smooth* espera-se uma linha suavizada aproximadamente horizontal com intercepto zero quando há adequação do modelo aos dados. Tal fenômeno é observado nos modelos 1 e 2 (Figura 5a e 5b), enquanto nos modelos 3 e 4 (Figura 5c e 5d) a linha apresenta uma leve queda, o que sugere presença de *outliers*, porém essas observações discrepantes não influenciam fortemente a adequação de ambos modelos.

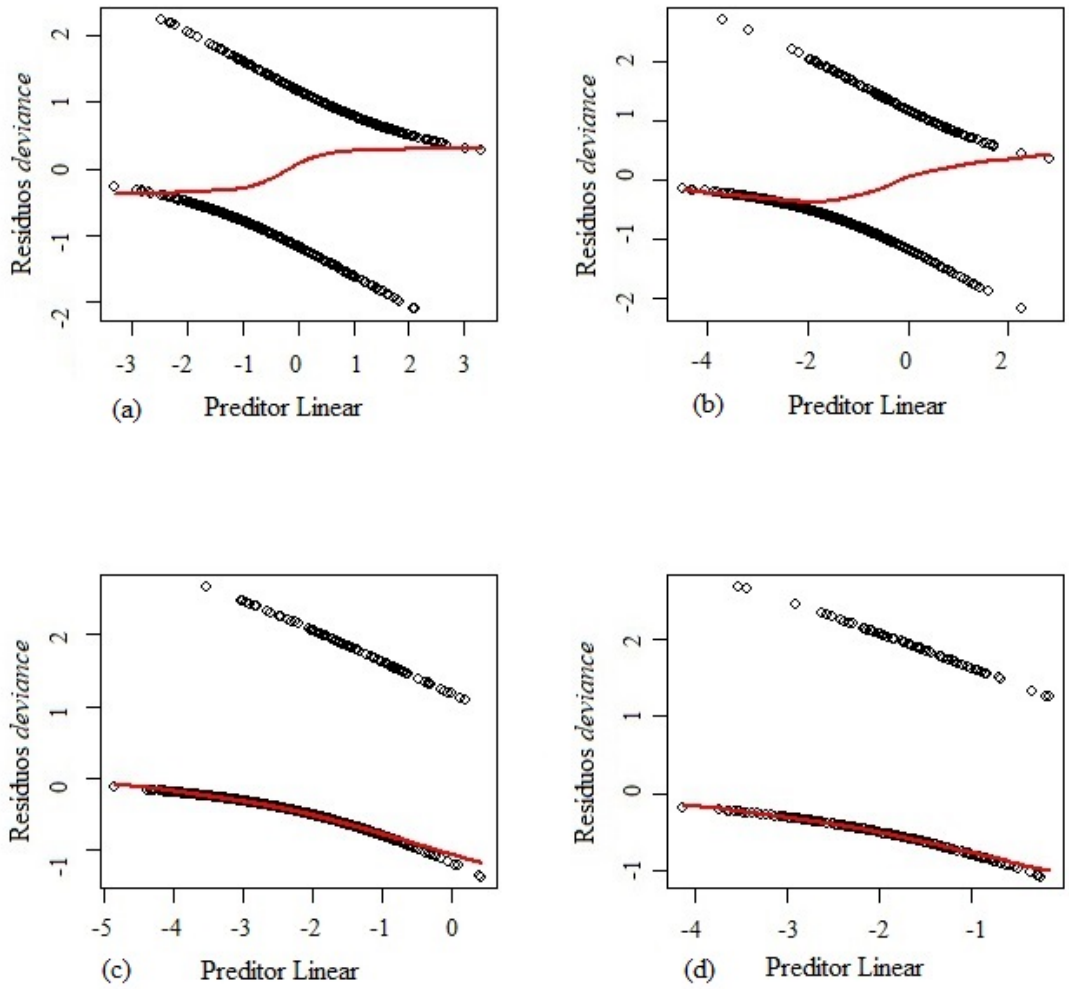

Figura 5: Resíduos *deviance versus* preditor linear para os quatro modelos

Na Figura 6 estão plotados os gráficos semi-normal de probabilidade com envelope simulado que também permitem identificar resíduos *deviance* discrepantes e consequentemente avaliar a adequação do modelo proposto. Nos gráficos (a) e (d) não são detectados *outliers*, pois não há pontos fora da faixa. No gráfico (b) quatro pontos estão fora da faixa, equivalente a cerca de 0,62% da amostra, porém tal proporção não é suficiente para invalidar o modelo. No gráfico (c) apenas um ponto encontra-se fora da faixa, dessa forma também não se pode afirmar que há inadequação no modelo.

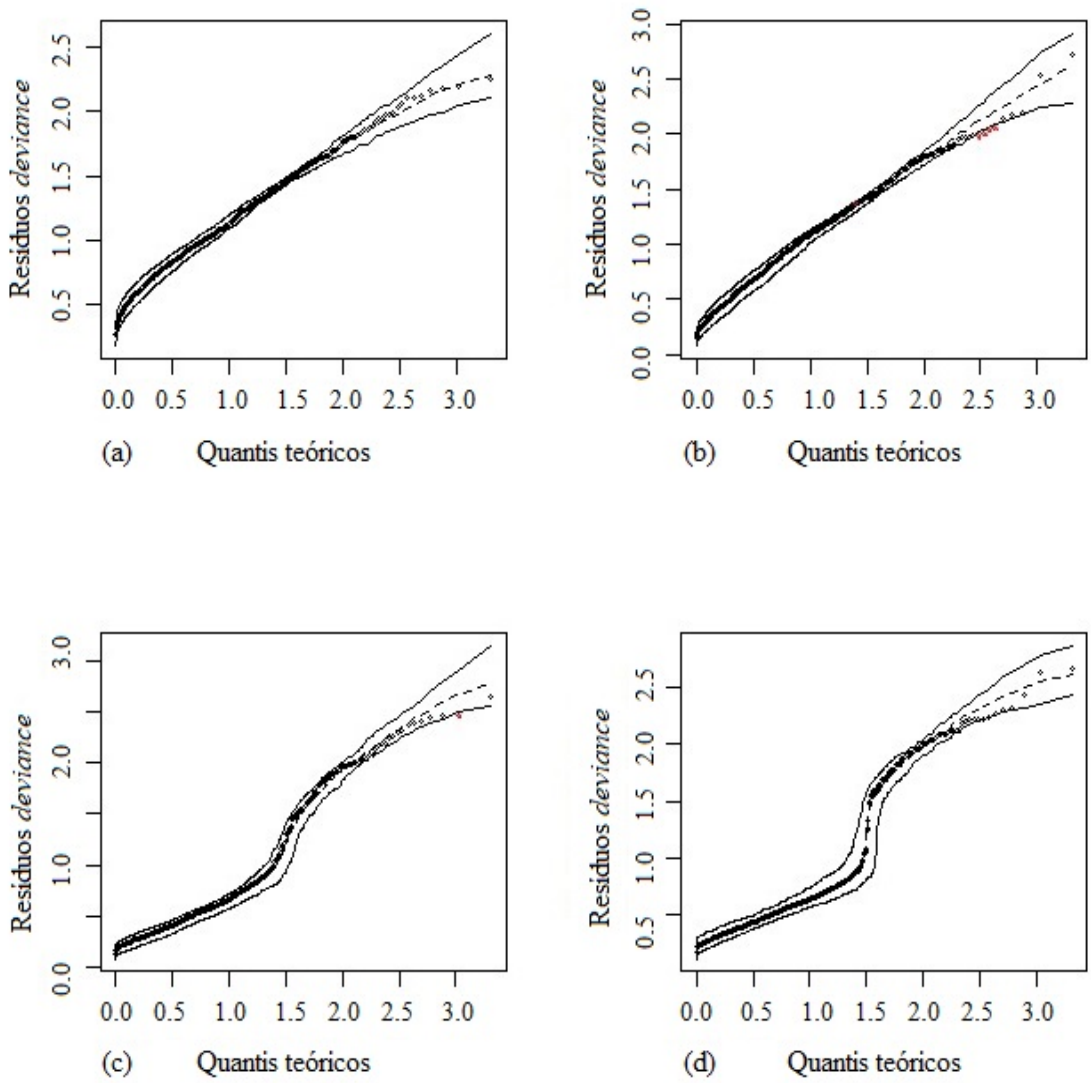

Figura 6: Envelope simulado dos resíduos *deviance* para os quatro modelos

4.3.3 Observações Influentes

A análise foi realizada por meio de gráficos pela dificuldade em se definir os valores usados para classificar uma observação como influente ou discrepante. As medidas DFCHISQ, DFDEV, DFBETA e *leverage* foram analisadas conjuntamente considerando o intervalo da probabilidade estimada em que a observação se encontra.

No modelo para uso de álcool na vida (Figura 7) é possível identificar algumas observações com DFCHISQ e DFDEV mais distantes de zero sugerindo indivíduos mal ajustados ao modelo. Os pontos com DFBETA e *leverage* destoantes encontram-se no intervalo (0,1;0,9) para a probabilidade estimada, tal comportamento não caracteriza essas observações como influentes.

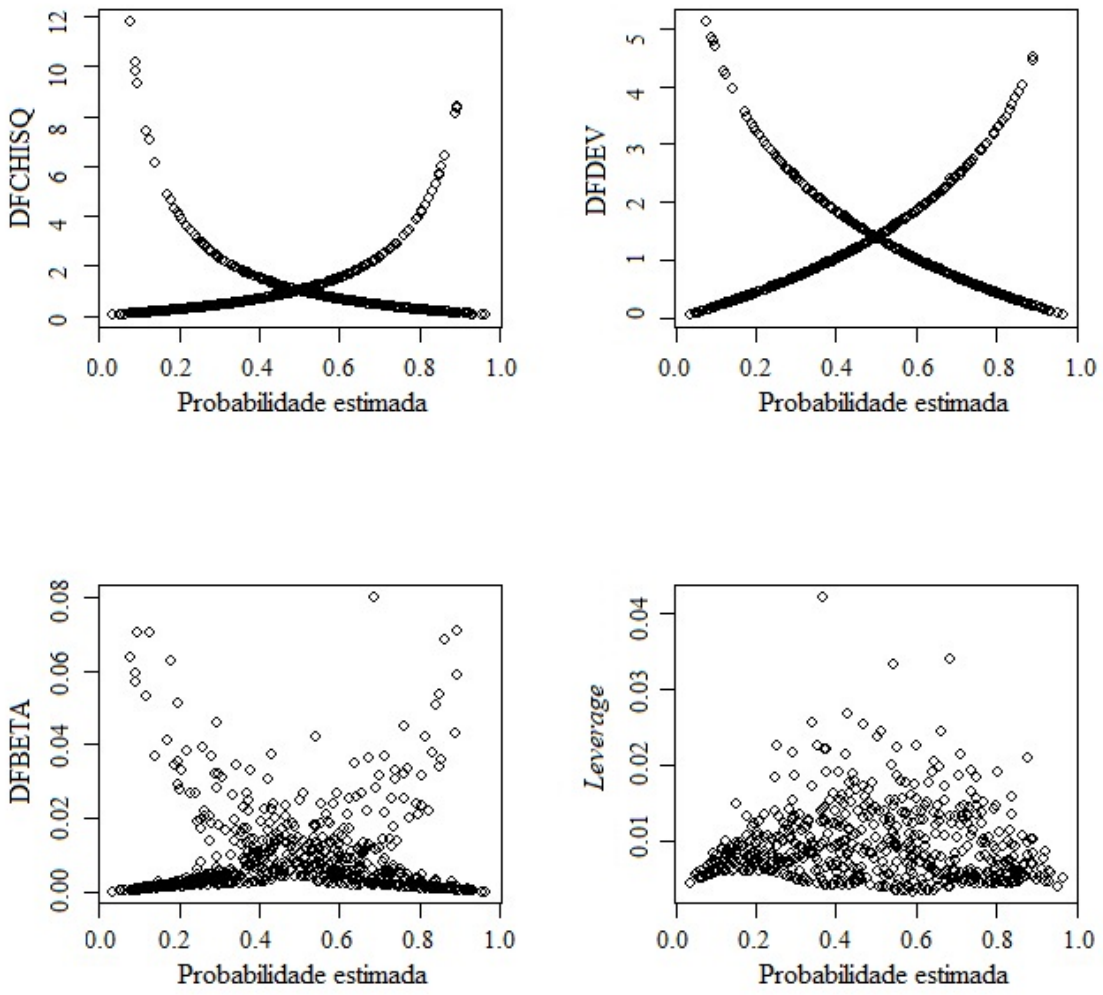

Figura 7: Gráficos para detecção de pontos influentes no modelo 1

Para o modelo de uso de álcool no último ano (Figura 8) duas observações apresentam alto valor para DFCHISQ e DFDEV indicando dados mal ajustados no modelo. A observação com maior valor DFBETA tem alto valor DFCHISQ e DFDEV, porém possui *leverage* pequeno, sendo assim ela pode ser considerada influente devido sua falta de ajuste. As duas observações com maior valor para *leverage* têm probabilidade estimada entre 0,4 e 0,8 e valores pequenos para as demais medidas, portanto não são definidas como pontos influentes no modelo.

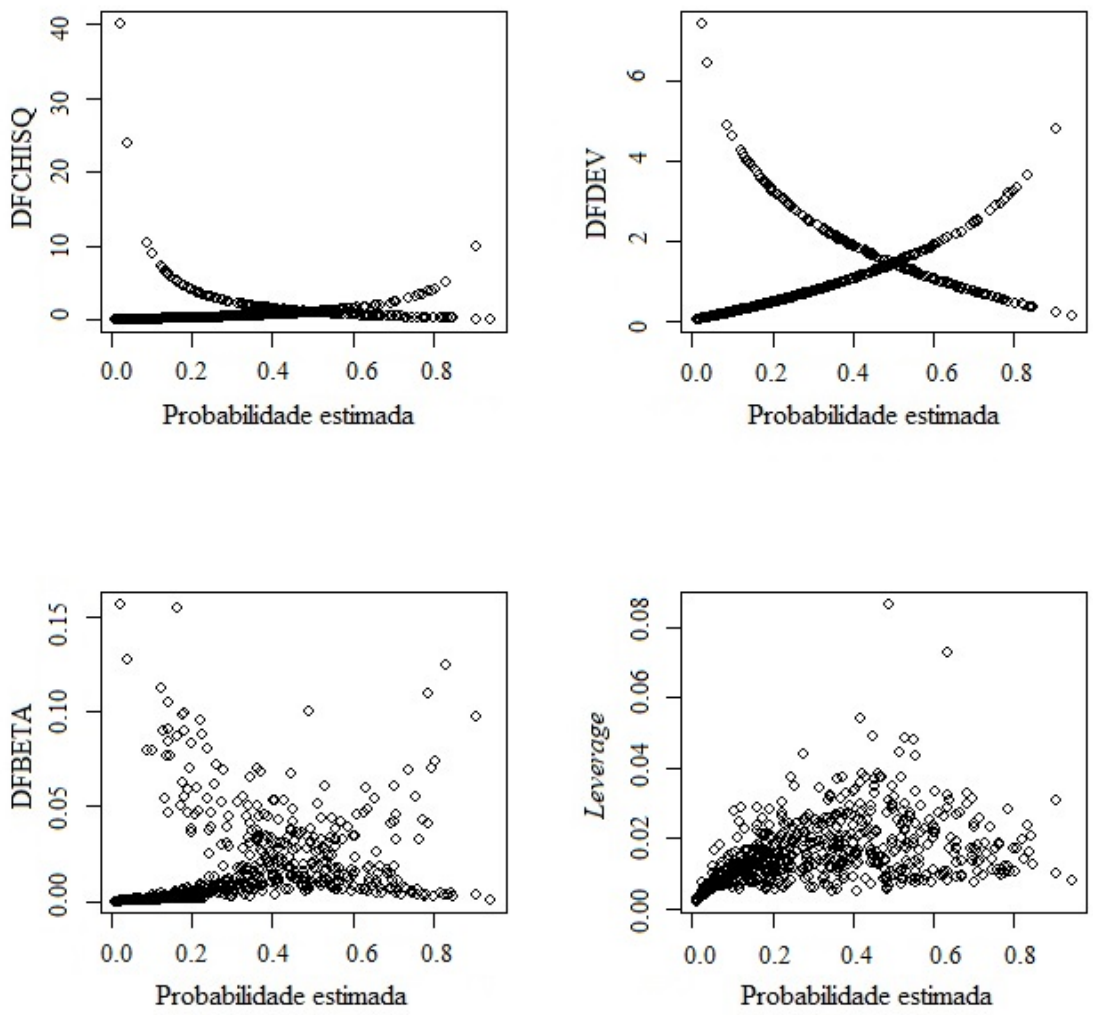

Figura 8: Gráficos para detecção de pontos influentes no modelo 2

No modelo para uso de álcool no último mês (Figura 9) a observação com maior valor para DFCHISQ também apresenta alto valor para DFDEV, enquanto DFBETA e *leverage* são pequenos o que propõe mal ajuste do modelo. A observação com maior valor DFBETA possui valores baixos para as outras medidas e probabilidade estimada menor que 0,1 o que sugere ser uma observação influente. Já a observação com *leverage* discrepante não influencia o modelo, pois esse comportamento é esperado quando a probabilidade estimada não está próxima de 0 e 1.

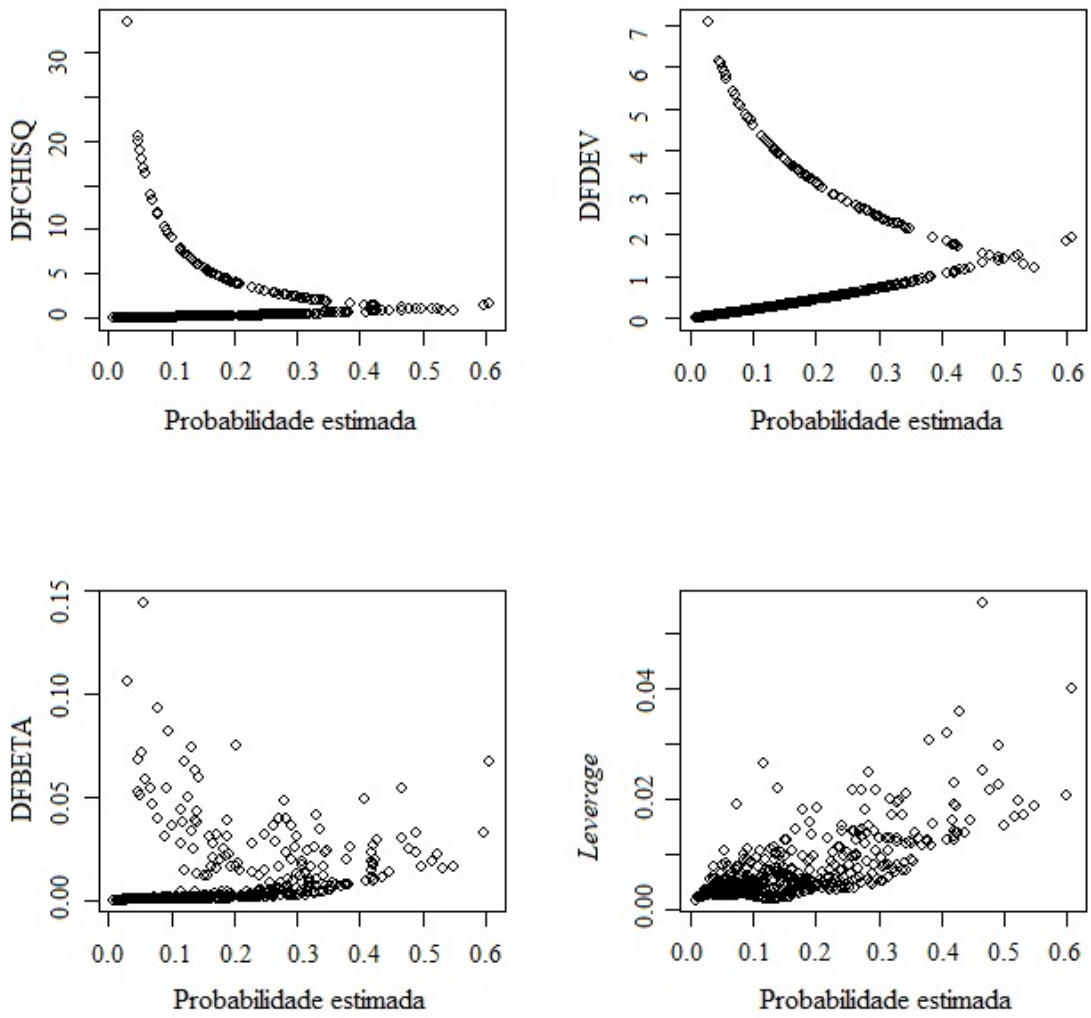

Figura 9: Gráficos para detecção de pontos influentes no modelo 3

No modelo para *binge* alcoólico no último ano (Figura 10) duas observações apresentam alto valor para DFCHISQ e DFDEV o que indica falta de ajuste ao modelo. A observação com maior DFBETA possui baixos valores para as demais medidas e probabilidade estimada próxima a 0,10, sendo considerada ponto influente no modelo. A observação com maior *leverage* apresenta comportamento esperado dado que sua probabilidade estimada está entre 0,1 e 0,9, portanto não é ponto de influência.

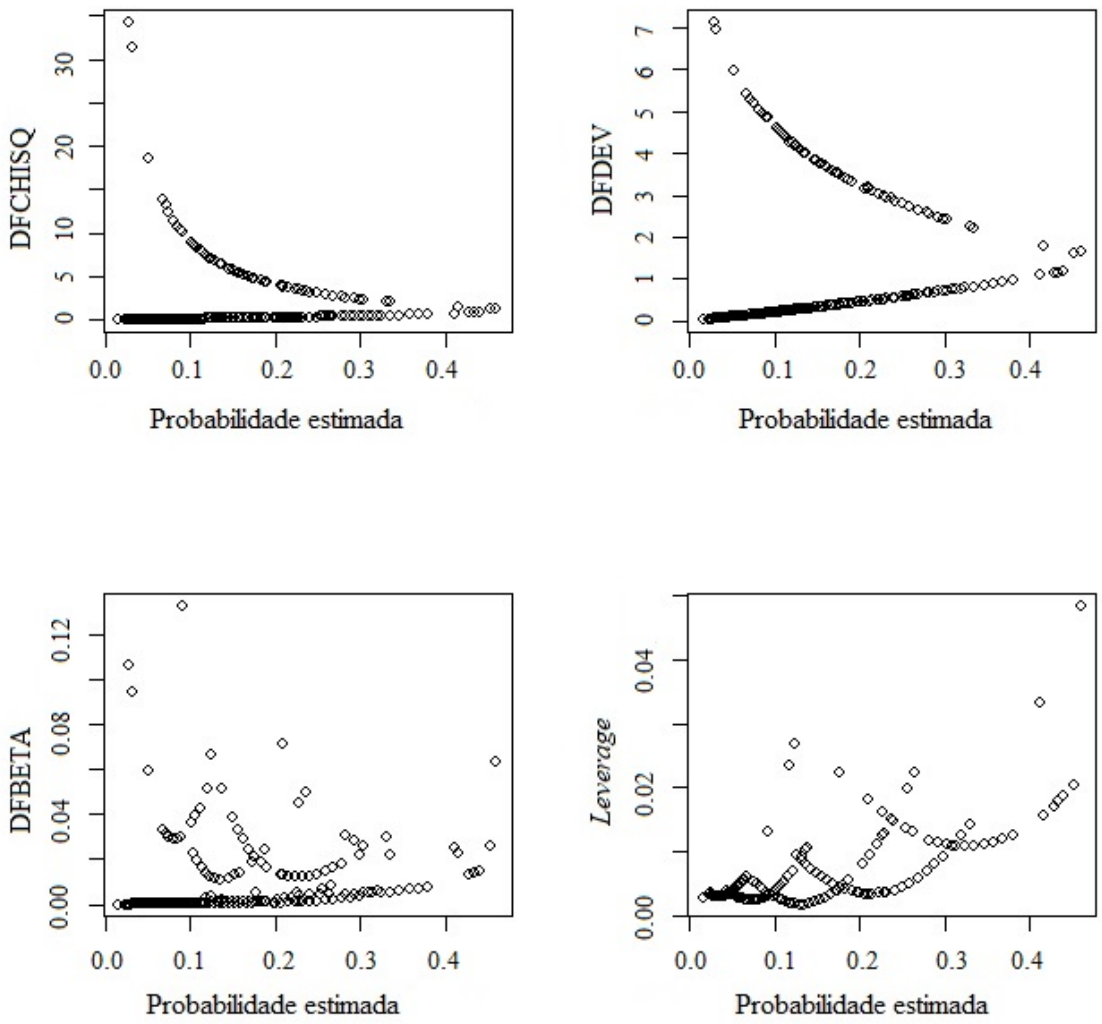

Figura 10: Gráficos para detecção de pontos influentes no modelo 4

5 Conclusão

O objetivo desse estudo consistiu em identificar fatores que estão associados ao uso de álcool em adolescentes da amostra, por se tratar de um assunto importante para a sociedade e ser objeto de estudos mais abrangentes. A experimentação cada vez mais precoce no consumo de álcool é um dado alarmante por aumentar o risco ao uso da droga. Na amostra estudada observou-se média de 11,58 anos para o primeiro contato entre os que afirmaram ter experimentado bebidas alcoólicas, ao comparar com pesquisas anteriores que apresentaram média de 13,0 (CARLINI et al., 2010) e 13,9 anos (LARANJEIRA et al., 2007) para o início do consumo de álcool, percebe-se esse contato prematuro.

Com o intuito de verificar esse comportamento dos jovens, foram gerados modelos logísticos que têm como variável resposta o uso ou não de álcool para quatro padrões de consumo, sendo eles o uso de álcool na vida, no último ano, no último mês e *binge* alcoólico. As variáveis explicativas foram selecionadas pelo método *stepwise* considerando 5% de confiança e observou-se que as variáveis idade e autoeficácia para aprendizagem autorregulada estão presentes nos quatro modelos, enquanto Bolsa Família, escolaridade da mãe, intrusividade, eficácia autoassertiva e suporte parental fazem parte de ao menos um modelo.

No que se refere a razão de chances, verificou-se que a chance de uso de álcool aumenta a cada ano de vida em todos os modelos, portanto quanto mais velho maior a chance de consumo de álcool. A chance de uso de álcool entre os adolescentes que participam do programa Bolsa Família e aqueles que não sabem se participam é menor quando comparados com os adolescentes que não recebem o auxílio. Esse resultado indica que os adolescentes que não recebem o Bolsa Família têm uma chance maior de consumir álcool, o que poderia ser explicado pela condição financeira familiar.

A escolaridade da mãe se mostrou significativa apenas no modelo de uso de álcool no último ano, indicando que a chance de um adolescente que não sabe a escolaridade da mãe ingerir bebida alcoólica é menor quando comparado aqueles que possuem mãe com ensino médio completo. No entanto, não existe diferença entre os que têm mãe com ensino médio incompleto e completo, pois o intervalo de confiança para razão de chances da categoria ensino médio incompleto possui o valor 1 (ver Apêndice A). O fato de não saber a escolaridade da mãe está muitas vezes associado à pouca escolaridade materna, o que dificulta a interpretação dos resultados, visto que alguns adolescentes poderiam pertencer a outra categoria estudada.

A prática parental intrusividade foi significativa em três modelos, mostrando que a chance de um adolescente consumir álcool aumenta com os comportamentos invasivos por parte dos pais. Assim, tal prática atua como fator de risco sobre o comportamento do jovem. A autoeficácia para aprendizagem autorregulada e para conseguir suporte parental diminuem a chance de consumo de álcool. Então, esses são fatores que podem ser definidos como protetivos, pois a confiança

em suas habilidades e capacidade de pedir auxílio aos pais faz com que os jovens busquem uma alternativa que não seja o uso de álcool. Por outro lado, a eficácia autoassertiva aumenta a chance de consumo de álcool. Esse fator está relacionado com a capacidade do indivíduo em defender sua opinião e nesse caso atua como fator de risco.

A adequabilidade dos modelos é assegurada pelo teste de Hosmer-Lemeshow que indicou que os modelos estão bem ajustados aos dados. Pelas tabelas de classificação, os quatro modelos analisados apresentaram boas medidas de classificação e poder de discriminação aceitável. Na análise dos resíduos não foram identificados valores extremos, porém algumas observações mostraram-se pouco ajustadas aos dados. As medidas de influência também detectaram alguns pontos influentes devido a falta de ajuste aos modelos.

Os resultados obtidos nesse estudo permitiram a avaliação dos padrões de consumo de álcool entre os adolescentes, e principalmente a identificação de fatores sociodemográficos, familiares e comportamentais associados a esse consumo. Tais fatores têm efeito positivo ou negativo no comportamento do jovem, podendo aproximá-lo ou afastá-lo das bebidas alcoólicas. É importante destacar uma limitação desse estudo referente à amostra, pois como os dados foram coletados apenas em oito escolas públicas do Distrito Federal e uma de Goiânia não é possível que se faça generalizações das características para uma população mais abrangente. Dessa forma, novos estudos podem ser realizados buscando ampliar o perfil dos jovens a fim de se aprofundar o conhecimento sobre o impacto de fatores associados ao uso de álcool em adolescentes.

Referências

1 ALAVARSE, G. M. A.; CARVALHO, M. D. B. **Álcool e adolescência: o perfil de consumidores de um município do norte do Paraná**. Esc. Anna Nery, Rio de Janeiro, v. 10, n. 3, p. 408-416, Dez., 2006.

2 BOUZAS, I. **O álcool**. [Editorial]. Adolescência & Saúde, v. 4, n. 4, Out./Dez., 2007.

3 CARLINI, E. L.; NOTO, A. R.; SANCHEZ, Z. M.; CARLINI-COTRIM, M.; LOCATELLI, D. P.; ABEID, L. R.; MOURA, Y. G. **VI Levantamento nacional sobre o consumo de drogas psicotrópicas entre estudantes do ensino fundamental e médio das redes pública e privada de ensino nas 27 capitais brasileiras**. São Paulo: CEBRID - Centro Brasileiro de Informações sobre Drogas Psicotrópicas: UNIFESP - Universidade Federal de São Paulo. Brasília: SENAD - Secretaria Nacional de Políticas sobre Drogas, 2010.

4 COLARES, J. M. J. **Avaliação psicométrica de um instrumento para mensuração de práticas parentais**. 2016. 34 f., il. Trabalho de Conclusão de Curso (Bacharelado em Estatística) - Universidade de Brasília, Brasília, 2016.

5 HOSMER, D.W.; LEMESHOW S. **Applied Logistic Regression**. Second Edition, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2000.

6 JOHNSTON, L. D.; O'MALLEY, P. M.; MIECH, R. A.; BACHMAN, J. G.; SCHULENBERG, J. E. **Monitoring the future national survey results on drug use, 1975-2016: Overview, key findings on adolescent drug use**. Ann Arbor: Institute for Social Research, The University of Michigan, 2017.

7 KHOURI, N. D. M. A. A. **Uso de drogas na adolescência: associações com sexo, práticas parentais, autoeficácia e perspectiva de tempo futuro**. 2016. xi, 136 f., il. Dissertação (Mestrado em Psicologia Clínica e Cultura) - Universidade de Brasília, Brasília, 2016.

8 LARANJEIRA, R.; PINSKY, I.; ZALESKI, M.; CAETANO, R. **Primeiro levantamento nacional sobre os padrões de consumo de álcool na população brasileira**. Secretaria Nacional Antidrogas, Brasília, 2007.

9 MCCULLAGH, P.; NELDER, J.A. **Generalized Linear Models**. Second Edition. Chapman & Hall, London, 1989.

10 MURTA, S. G.; NOBRE-SANDOVAL, L. A.; IGLESIAS, F.; VINHA, L. G. A.; ROCHA, V. P.; MIRANDA, A. A. V.; SOUZA, A. S. S. **Avaliação da efetividade, qualidade de implementação e validade social do programa Fortalecendo Famílias**. Brasília, Universidade de Brasília: Relatório Não Publicado, 2015.

11 VASCONCELOS, R. B. B. **Avaliação psicométrica de um instrumento para mensuração de autoeficácia**. 2016. 44 f., il. Trabalho de Conclusão de Curso (Bacharelado em Estatística) - Universidade de Brasília, Brasília, 2016.

APÊNDICE A – Intervalo de confiança para razão de chances

Tabela A.1: Intervalo de confiança - Modelo 1

Variável	Razão de chances	IC (95%)
Idade	1,911	1,627 - 2,244
Bolsa Família		
Sim	0,508	0,332 - 0,776
Não sabe	0,534	0,327 - 0,872
Intrusividade	1,438	1,217 - 1,700
Aprendizagem autorregulada	0,512	0,393 - 0,669

Tabela A.2: Intervalo de confiança - Modelo 2

Variável	Razão de chances	IC (95%)
Idade	1,862	1,558 - 2,225
Bolsa Família		
Sim	0,444	0,268 - 0,736
Não sabe	0,515	0,292 - 0,909
Escolaridade da mãe		
Não sabe	0,568	0,341 - 0,946
E.M. incompleto	0,946	0,590 - 1,517
Intrusividade	1,282	1,071 - 1,534
Aprendizagem autorregulada	0,506	0,362 - 0,706
Eficácia autoassertiva	1,504	1,144 - 1,977
Suporte parental	0,789	0,641 - 0,972

Tabela A.3: Intervalo de confiança - Modelo 3

Variável	Razão de chances	IC (95%)
Idade	1,950	1,560 - 2,438
Intrusividade	1,333	1,082 - 1,643
Aprendizagem autorregulada	0,605	0,431 - 0,850

Tabela A.4: Intervalo de confiança - Modelo 4

Variável	Razão de chances	IC (95%)
Idade	1,756	1,416 - 2,178
Aprendizagem autorregulada	0,641	0,460 - 0,892