

## ANÁLISE DO CONSUMO DE FRUTAS POR CRIANÇAS DE UM MUNICÍPIO DO INTERIOR DO PARANÁ USANDO MODELOS MULTINOMIAIS COM EFEITOS ALEATÓRIOS

Letícia de Viveiros INÁCIO<sup>1</sup>  
Andressa Avendaño FORBELLONE<sup>2</sup>  
Claudia Choma Bettega ALMEIDA<sup>3</sup>  
Cesar Augusto TACONELI<sup>4</sup>

- **RESUMO:** A introdução adequada da alimentação complementar ao leite materno, para crianças na faixa de 6 a 24 meses de idade, é de fundamental importância no sentido de garantir um crescimento saudável, com desenvolvimento físico, motor e intelectual adequados. Neste contexto, a inclusão de frutas em quantidade e variedade adequadas na alimentação complementar acrescenta um grande número de nutrientes essenciais à dieta da criança. O presente trabalho teve por objetivo identificar fatores demográficos, socioeconômicos e comportamentais relacionados à variedade de frutas consumidas por crianças entre 6 e 24 meses de idade, residentes e atendidas pelas Unidades de Saúde do Sistema Único de Saúde (SUS) do município de Guarapuava – Paraná. Por meio da aplicação de modelos multinomiais mistos foi possível identificar que a variedade de frutas consumidas está relacionada com a escolaridade das mães, à idade das mesmas quando do nascimento de seus primeiros filhos e ao número de refeições oferecidas diariamente às crianças. Um efeito aleatório foi incorporado ao modelo de forma a refletir o agrupamento das crianças quanto às Unidades de Saúde em que eram atendidas. Embora não tenha mostrado efeito estatisticamente significativo, sua inclusão permitiu identificar grupos de crianças com consumo menos variado de frutas.
- **PALAVRAS-CHAVE:** Dados categorizados; modelos multinomiais; efeitos aleatórios; consumo alimentar; frutas.

### 1 Introdução

Os primeiros anos de vida da criança (especialmente os dois primeiros) são caracterizados por um crescimento acelerado e enormes aquisições no processo de desenvolvimento, incluindo habilidades para receber, mastigar e digerir outros alimentos, além do leite materno, e pelo autocontrole do processo de ingestão de alimentos, para atingir o padrão alimentar de um adulto (BRASIL. Ministério da Saúde, 2010). Dessa forma, a alimentação que é ofertada à criança nos seus primeiros anos de vida tem grande

<sup>1</sup>Universidade Federal do Paraná. E-mail: [leticiviveiros@gmail.com](mailto:leticiviveiros@gmail.com)

<sup>2</sup>Universidade Federal do Paraná. E-mail: [andressa\\_forbellone@yahoo.com.br](mailto:andressa_forbellone@yahoo.com.br)

<sup>3</sup>Universidade Federal do Paraná, Departamento de Nutrição. E-mail: [chomac@ufpr.br](mailto:chomac@ufpr.br)

<sup>4</sup>Universidade Federal do Paraná, Departamento de Estatística, Caixa Postal: 190811, CEP: 81531-990, Curitiba, Paraná, Brasil. E-mail: [taconeli@ufpr.br](mailto:taconeli@ufpr.br)

impacto em sua saúde. Segundo o Ministério da Saúde (2005), uma alimentação infantil adequada compreende a prática do aleitamento materno e a introdução, em tempo oportuno, de alimentos apropriados que o complementem.

A Organização Mundial da Saúde (OMS) e o Ministério da Saúde recomendam a prática do aleitamento materno exclusivo nos seis primeiros meses da criança e do aleitamento materno complementado a partir dessa idade. A recomendação se baseia no fato de que o trato gastrointestinal e as funções renais da criança não estão completamente desenvolvidos antes dos seis meses de idade (AKRÉ, 1997). Assim, segundo Monte e Giugliani (2004), a introdução precoce (antes dos seis meses de idade) de outros alimentos, que não o leite humano, pode se tornar prejudicial à criança, causando, dentre outros problemas: maior número de episódios de diarreia; aumento no número de hospitalizações por doenças respiratórias; risco de desnutrição, se os alimentos introduzidos forem nutricionalmente inferiores ao leite materno; menor absorção de nutrientes importantes, como o ferro e o zinco.

A partir dos seis meses de idade, os alimentos complementares ao leite materno devem começar a ser incluídos na dieta da criança, pois nesse estágio do crescimento, o leite materno já não supre todas as necessidades energéticas da criança, especialmente a de ferro. Alimentos complementares são aqueles oferecidos às crianças enquanto elas ainda estão em período de aleitamento materno e esses podem ser considerados alimentos de transição, se forem especialmente preparados para as crianças até que elas passem a receber a mesma alimentação de sua família (BRASIL. Ministério da Saúde, 2005).

A alimentação complementar deve prover quantidades suficientes de água, energia, proteínas, gorduras, vitaminas e minerais, por meio de alimentos seguros, culturalmente aceitos, economicamente acessíveis e que sejam agradáveis à criança (BRASIL. Ministério da Saúde, 2009). Para que a alimentação complementar consiga cumprir adequadamente o seu papel nutricional é necessário que as crianças recebam uma dieta equilibrada, de forma que todos os grupos alimentares estejam presentes nas quantidades adequadas.

Dado o que foi exposto, o levantamento e análise estatística de dados referentes à alimentação complementar de crianças menores de 2 anos mostra-se essencial para se obter conhecimento sobre o estado nutricional das crianças e auxiliar os profissionais da saúde na implementação de estratégias que possam atingir as famílias e as crianças de forma a melhorar o estado nutricional das mesmas. Nesse contexto, o presente trabalho visa contribuir com o estudo da dieta de crianças no estágio de alimentação complementar, fornecendo subsídios para uma compreensão adequada do problema, baseado em dados de Guarapuava, Paraná (Brasil). Guarapuava é um município da região centro-oeste do Paraná, distante a 252,7 km de Curitiba, capital do estado. Sua população, estimada pelo IBGE para 2014, é de 176.973 habitantes, sendo 7306 crianças menores de dois anos, segundo o Censo/2010 do IBGE. É o segundo município do estado do Paraná em extensão territorial e o nono em população.

O principal objetivo deste estudo foi analisar como se dá a inclusão de variedades de frutas na alimentação de crianças de 6 a 24 meses que possuem informações registradas no banco de dados citado. Neste trabalho, são utilizados dados provenientes de um estudo transversal apresentado em Brecailo (2007), realizado em Guarapuava com crianças de 0 a 2 anos de idade que faziam parte do Programa Saúde da Família. A expressão da grande quantidade de variáveis disponíveis, referentes aos consumos de frutas e sucos pelas crianças, em termos de um indicador de consumo variado de frutas, foi possível mediante

criação de uma variável de contagem posteriormente categorizada. Desta forma, fez-se necessário o ajuste e comparação de diferentes modelos para dados multinomiais. Adicionalmente, o agrupamento das crianças, segundo as Unidades de Saúde em que eram atendidas, foi incorporado à análise por meio da inclusão de efeitos aleatórios, configurando os chamados modelos mistos.

## 2 Materiais e métodos

### 2.1 Material

Os dados utilizados no desenvolvimento deste trabalho referem-se a um estudo transversal, realizado na cidade de Guarapuava, interior do Estado do Paraná (Brasil), com crianças e respectivas mães biológicas que integravam o Programa Saúde da Família (PSF) do Sistema Único de Saúde (SUS) (BRECAILO, 2007). A população sob estudo era composta por 2454 crianças que residiam na área urbana da cidade e eram beneficiárias do PSF. A amostra inicial contava com 519 crianças, considerando que foram excluídas do estudo as crianças gêmeas, as que não estavam sob os cuidados das mães biológicas, as que completaram dois anos antes da realização da entrevista e as que foram visitadas por três vezes sem que se conseguisse contato com a mãe. Como o foco deste trabalho foi analisar a complementação alimentar, na análise foi considerado apenas as 370 crianças que estavam em idade recomendada para receberem alimentos além do leite materno, ou seja, maiores de 6 meses.

As informações foram obtidas por meio de entrevistas realizadas com as mães biológicas das crianças selecionadas. As entrevistas foram realizadas nas casas das participantes por meio da aplicação de questionários. O principal instrumento deste levantamento de dados foi o Questionário de Frequência Alimentar (QFA), em que as mães informavam as frequências com que as crianças consumiam diferentes tipos de alimentos, compreendendo carnes, verduras, legumes, frutas e alimentos industrializados, dentre outros. As frequências foram registradas numa escala ordinal, categorizadas como: (a) todos os dias; (b) 4 a 6 vezes por semana; (c) 1 a 3 vezes por semana; (d) 2 a 3 vezes por mês; (e) 1 vez por mês; (f) nunca ou quase nunca. Em particular, compõem o QFA as seguintes 15 variedades de frutas (e respectivos sucos), que foram mencionados numa amostragem piloto: abacaxi, banana, kiwi, laranja, limão, maçã, mamão, manga, maracujá, melão, melancia, mexerica, pera, pêssego e uva. O consumo de frutas é o objeto de estudo do presente trabalho.

A variável resposta foi construída de forma que englobasse todas as frutas e sucos presentes no QFA em uma única variável refletindo o grau de variedade da dieta da criança, quanto às frutas consumidas com alguma regularidade. Assim, definiu-se inicialmente como variável resposta o número de variedades de frutas consumidas pela criança ao menos uma vez por semana. Para que fique mais claro, essa variável corresponde ao número de variedades de frutas para as quais a frequência de consumo relatada pela mãe se enquadre em uma das seguintes categorias: 'todos os dias'; '4 a 6 vezes por semana' e '1 a 3 vezes por semana'. Nos casos em que a fruta e o respectivo suco fossem consumidos com alguma frequência, atribuiu-se, para essa fruta, a frequência mais elevada com relação às duas formas de consumo. A título de ilustração, considerando que o consumo de abacaxi de determinada criança fosse '2 a 3 vezes por

mês' e o de suco de abacaxi '1 a 3 vezes por semana', a frequência de consumo de abacaxi atribuída a essa criança foi '1 a 3 vezes na semana'.

Dessa forma, a variável resposta, originalmente, correspondia a uma variável de contagem, podendo assumir valores entre zero (caso nenhuma das variedades de fruta fosse consumida ao menos uma vez por semana pela criança) a 15 (caso todas as variedades presentes no QFA fossem consumidas pela criança ao menos uma vez por semana). No entanto, foi decidida, em consenso com profissionais da área da saúde, a categorização dessa variável em três faixas que configurassem diferentes níveis de consumo variado: '0 ou 1 variedade de fruta consumida ao menos uma vez na semana', o que configuraria ausência de variedade no consumo regular de frutas; '2 ou 3 variedades de frutas consumidas ao menos uma vez na semana', indicando uma variedade intermediária no consumo regular de frutas, muitas vezes compreendendo banana, maçã e laranja, que são frutas de consumo mais frequente ou '4 a 15 variedades de frutas consumidas ao menos uma vez na semana', correspondente à maior variedade no consumo regular de frutas. É importante esclarecer que a categorização da variável resposta em três faixas de consumo teve como objetivo facilitar a interpretação e aplicação dos resultados para os nutricionistas, o que não impede que estudos similares possam utilizar alguma outra forma de construção para uma variável correspondente à variedade do consumo alimentar.

O questionário também contemplava características socioeconômicas e demográficas referentes às famílias, além de informações referentes aos hábitos alimentares das crianças. Entre as variáveis complementares, têm-se as 17 variáveis explicativas que foram incorporadas à análise. As variáveis relativas à mãe da criança são: idade, ocupação, estado civil, número de filhos nascidos vivos, idade no nascimento do primeiro filho, anos de estudo e se ela recebeu orientação quanto ao aleitamento materno; variáveis relativas ao pai e à família: anos de estudo do pai, se o pai reside com a criança, renda familiar per capita, número de residentes no domicílio; e as variáveis relativas à criança: sexo, idade, peso ao nascer, tipo de parto, número de refeições diárias e tempo de aleitamento materno. Todas as variáveis numéricas foram convenientemente categorizadas no próprio questionário.

A análise estatística foi realizada utilizando o software R (R CORE TEAM, 2013). Em particular, para o ajuste dos modelos de regressão foi utilizado o pacote ordinal (CHRISTENSEN, 2012).

## **2.2 Métodos**

A análise estatística dos dados fundamentou-se no ajuste de diferentes tipos de modelos de regressão para dados categorizados, contemplando o modelo de regressão multinomial, o modelo de regressão multinomial de chances proporcionais, o modelo de regressão multinomial de chances proporcionais parcial e as extensões dos modelos citados com a inclusão de efeitos aleatórios.

### **2.2.1 Modelo de regressão multinomial**

Sejam  $Y_i$ ,  $i = 1, 2, \dots, n$ , variáveis aleatórias politômicas, denotando seus resultados possíveis por  $j$ ,  $j = 1, 2, \dots, J$ . Adicionalmente, num contexto de análise de regressão,

sejam  $\underline{x}_i' = (x_{i1}, x_{i2}, \dots, x_{ip})'$  os correspondentes vetores de covariáveis, sendo  $p + 1$  o número de parâmetros do modelo. O modelo de regressão multinomial fica definido por:

$$g_j(\underline{x}_i) = \ln \left[ \frac{P(Y_i = j | \underline{x}_i)}{P(Y_i = k | \underline{x}_i)} \right] = \beta_{0j} + \beta_{1j}x_{i1} + \beta_{2j}x_{i2} + \dots + \beta_{pj}x_{ip}, \quad (1)$$

Sendok algum dos  $J$  resultados possíveis de  $Y$  e  $j = 1, 2, \dots, J$ , com a restrição  $j \neq k$ . Assim, o modelo de regressão multinomial configura uma extensão do modelo de regressão logística binária, em que são ajustados  $J - 1$  logitos distintos. Embora seja um modelo bastante geral, peca pela complexidade inerente ao grande número de parâmetros a serem estimados e interpretados  $((j - 1)(p + 1))$ .

### 2.2.2 Modelo de regressão multinomial de chances proporcionais

O modelo de regressão multinomial de chances proporcionais é uma opção adequada quando os dados são politômicos e as categorias possuem alguma conotação de ordem, ou seja, a variável resposta é de natureza ordinal. Nesse modelo, probabilidades acumuladas são modeladas de forma que as razões de chances para qualquer covariável sejam as mesmas, independente da categoria da resposta que se está avaliando, ou seja, as chances são proporcionais em todas as categorias da resposta (ABREU; SIQUEIRA; CAIAFFA, 2009). Modelada-se a razão da probabilidade de resposta “menor ou igual” a uma categoria  $j$  ( $P(Y \leq j)$ ) com relação à probabilidade de resposta “maior que  $j$ ” ( $P(Y > j)$ ). O modelo de regressão multinomial de chances proporcionais, em termos dos logitos acumulados, fica definido por:

$$g_j(\underline{x}_i) = \ln \left[ \frac{P(Y \leq j | \underline{x}_i)}{P(Y > j | \underline{x}_i)} \right] = \beta_{0j} - (\beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}), \quad (2)$$

para  $j = 1, 2, \dots, J - 1$ . Repare que os coeficientes correspondentes às covariáveis são constantes para os  $J - 1$  logitos, sendo que apenas o intercepto muda conforme o logito. A definição do modelo supõe, portanto, igual efeito de uma covariável para qualquer um dos logitos, o que implicará na suposição de chances proporcionais. Embora conveniente, a suposição de chances proporcionais é bastante forte e, não sendo verificada com base no diagnóstico do ajuste, inviabiliza a aplicação do modelo. Nesse caso é possível adotar um modelo multinomial, não considerando a natureza ordinal dos dados (HOSMER e LEMESHOW, 2000) ou, ainda, um modelo intermediário, assumindo chances proporcionais apenas para um subconjunto de covariáveis, produzindo o modelo de regressão multinomial de chances proporcionais parcial.

### 2.2.3 Modelo de regressão multinomial de chances proporcionais parcial

Abreu, Siqueira e Caiaffa (2009) sugerem, para o caso de violação da suposição de chances proporcionais, que seja adotado um modelo de chances proporcionais parcial, tal que, para as variáveis em que a suposição é válida, se tem um parâmetro comum para todas as categorias da resposta, enquanto para as variáveis em que a suposição não é

válida adota-se um parâmetro específico para cada logito ajustado. Suponha que a suposição de chances proporcionais não seja verificada para  $q$  variáveis presentes no modelo, sendo  $q < p$ . Dessa forma, seja  $\gamma_{jr}$  o parâmetro que incrementa o efeito da  $r$ -ésima covariável ( $r = 1, 2, \dots, q$ ) que não atende a suposição de chances proporcionais. O modelo multinomial de chances proporcionais parcial fica definido por:

$$g_j(\underline{x}_i) = \ln \left[ \frac{P(Y \leq j | \underline{x}_i)}{P(Y > j | \underline{x}_i)} \right] \\ = \beta_{0j} - [(\beta_1 + \gamma_{j1})x_{i1} + \dots + (\beta_q + \gamma_{jq})x_{iq} + \beta_{q+1}x_{i(q+1)} + \dots + (\beta_px_{ip})], \\ \text{para } j = 1, 2, \dots, J - 1 \quad (3)$$

#### 2.2.4 Modelo de regressão multinomial de chances proporcionais com efeitos aleatórios

Modelos com inclusão de efeitos aleatórios podem ser usados quando os dados, de alguma forma, são agrupados, agrupamento esse que é modelado através da adição de um termo aleatório ao preditor linear. O agrupamento mencionado pode ser decorrente, a título de exemplo, de observações coletadas num mesmo indivíduo ao longo do tempo (configurando dados longitudinais) ou de observações “regionalmente” agrupadas, como funcionários de uma mesma empresa, estudantes de uma mesma escola ou crianças atendidas por uma mesma unidade de saúde. A incorporação do efeito dos grupos ao modelo tem como principais objetivos acomodar a variabilidade entre grupos e possível correlação intra-grupo (HOSMER e LEMESHOW, 2000).

Embora a representação da inclusão de efeitos aleatórios seja apresentada aqui apenas para o modelo de regressão de chances proporcionais, a extensão é imediata para os demais modelos considerados. Suponha que os dados originais estejam agrupados em  $M$  grupos e seja  $\alpha_i$  o efeito do grupo ao qual pertence a  $i$ -ésima observação, o qual é denotado por  $\alpha_{mi}$ , apenas para reforçar o grupo correspondente ( $m = 1, 2, \dots, M$ ). O modelo multinomial de chances proporcionais de efeitos mistos, neste caso, fica definido por:

$$g_j(\underline{x}_i) = \ln \left[ \frac{P(Y_i \leq j | \underline{x}_i)}{P(Y_i > j | \underline{x}_i)} \right] = \beta_{0j} - (\beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}) - \alpha_{mi},$$

com  $j = 1, 2, \dots, J - 1$ . Para o efeito aleatório, assume-se  $\alpha_{mi} \sim N(0, \sigma_\alpha^2)$ , sendo  $\sigma_\alpha^2$  a variância da variável considerada como termo aleatório. Adicionalmente, dois ou mais efeitos aleatórios podem ser incluídos num modelo, refletindo, por exemplo, diferentes níveis de agrupamento. No presente trabalho, no entanto, será considerada apenas a inclusão de um termo aleatório, correspondente ao agrupamento das crianças segundo as unidades de saúde em que eram atendidas.

### 2.2.5 Descrição do processo de modelagem

O primeiro passo da análise envolveu uma fase de análise descritiva e exploratória dos dados, através de análise gráfica, medidas descritivas e verificação da existência de associação entre a variável resposta e as variáveis explicativas por meio do teste qui-quadrado de Pearson.

O ajuste de modelos de regressão teve por objetivo avaliar, conjuntamente, os efeitos das 17 variáveis explicativas na variedade de frutas consumidas pelas crianças. A seleção das variáveis a serem incluídas ao modelo foi realizada por meio do método de seleção “forward”, no qual se tem um modelo inicial sem nenhuma covariável e as covariáveis são adicionadas sucessivamente até que não se tenha alguma variável fora do modelo que produza melhora na explicação da variável resposta (CHARNET *et al.*, 2008). O critério da informação de Akaike (AIC) foi utilizado como medida de qualidade de ajuste para avaliação da entrada das variáveis no modelo, de forma que se a inclusão da variável reduzisse o valor do AIC ela permaneceria no modelo (PAULA, 2013).

A estimação dos parâmetros dos modelos de regressão ajustados foi realizada pelo método da Máxima Verossimilhança (CASELLA e BERGER, 2010). A necessidade de inclusão de interações ao modelo, dentre as variáveis selecionadas anteriormente, foi verificada novamente por meio de possível redução no AIC. Testes para as hipóteses de interesse foram executados com base na razão de verossimilhanças (TRV) e na estatística de Wald (CASELLA e BERGER, 2010). Estimativas dos efeitos aleatórios referentes às Unidades de Saúde, baseadas na moda de suas distribuições condicionais, foram obtidas através algoritmo EM (TUTZ e HENNEVOGL, 1996).

A qualidade do ajuste do modelo foi verificada através da estatística C de Hosmer e Lemeshow (HOSMER e LEMESHOW, 2000). O TRV também foi utilizado para verificar a suposição de chances proporcionais para as variáveis presentes no modelo, além de um gráfico de logitos, calculados separadamente para cada categoria das covariáveis (HARREL, 2001).

A interpretação dos parâmetros estimados dos modelos foi realizada por meio de razões de chances (AGRESTI, 1996). Também se fez uso intervalos de confiança baseados nas razões de chances estimadas e para os efeitos de cada Unidade de Saúde. Tais intervalos foram baseados nas distribuições assintóticas dos estimadores (CASELLA e BERGER, 2010).

## 3 Resultados

Na Figura 1 têm-se as distribuições de frequências dos consumos das 15 variedades de frutas estudadas, e na Figura 2 para o número de variedades de frutas consumidas ao menos uma vez por semana.

Pela Figura 1 é possível observar que banana, maçã e laranja apresentaram maiores frequências de consumo diário em relação às demais variedades de frutas, registrando frequências superiores a 50 crianças. Essas três frutas também apresentaram frequências elevadas de consumo quando consideradas as crianças que as consumiam ao menos uma vez por semana, sendo essas frequências superiores a 200 crianças. Com uma frequência um pouco menor, mas ainda elevada, pode-se destacar os consumos de mamão, mexerica, manga e melancia. As demais variedades de frutas apresentaram frequências

relativamente baixas, sendo que kiwi, limão e melão apresentaram os menores valores, com menos de 25 crianças que as consumiam pelo menos uma vez por semana.

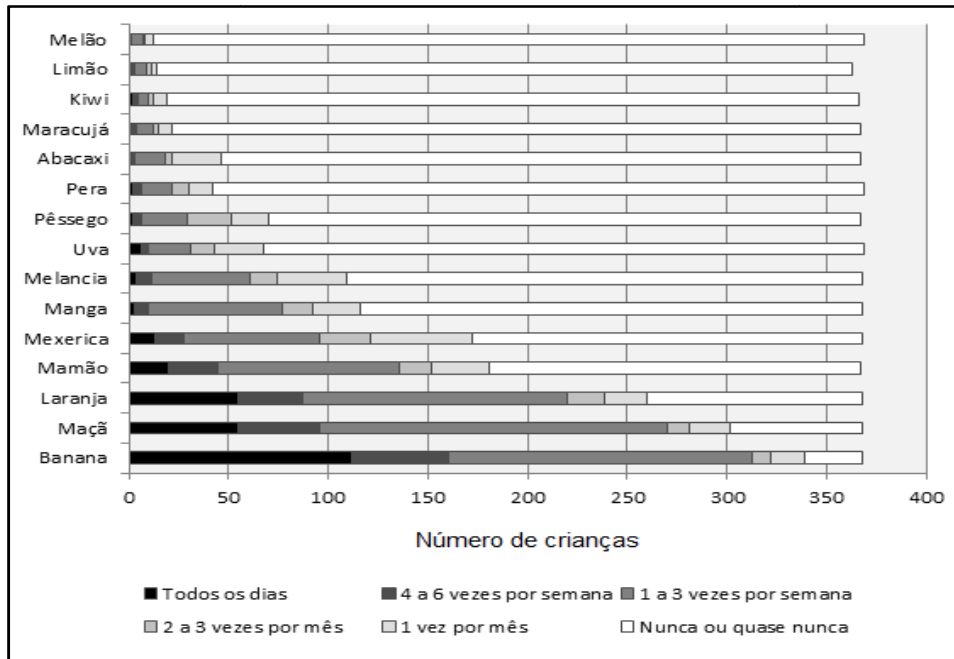


Figura 1 - Frequências de consumo das 15 variedades de frutas.

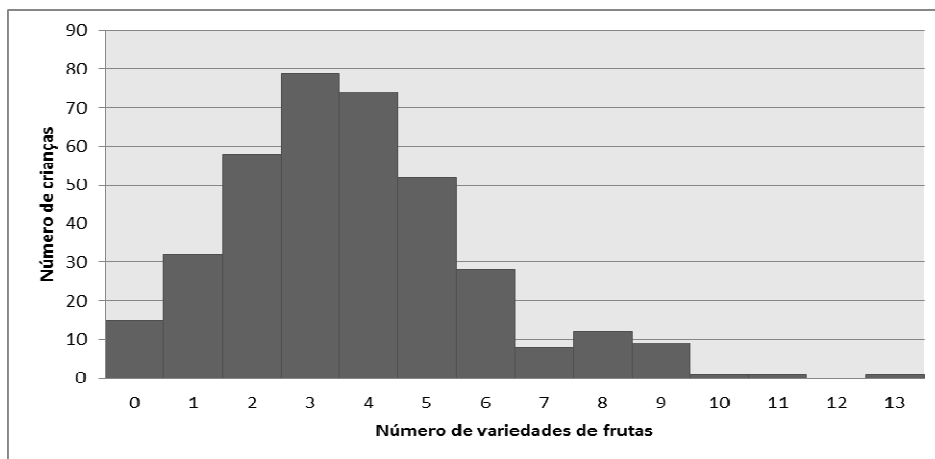


Figura 2 - Distribuição de frequências para o número de variedades de frutas incluídas na dieta da criança com uma frequência mínima semanal.



Na Figura 2 são verificadas maiores frequências para crianças que consumiam três ou quatro variedades de frutas com frequência mínima semanal. Quanto aos quartis, tem-se que 25% das crianças consumiam duas variedades de frutas ou menos com frequência mínima semanal e 25% delas consumiam cinco ou mais variedades de frutas, sendo o consumo mediano igual a quatro variedades de frutas. O número máximo de variedades consumidas com frequência mínima semanal foi igual a treze variedades e o mínimo igual a zero.

Na Tabela 1 são apresentadas as variáveis explicativas do estudo com as correspondentes distribuições de frequências, bem como os resultados dos testes qui-quadrado para analisar as associações dessas variáveis com a variedade de frutas consumidas.

Com base nos resultados apresentados na Tabela 1 é possível verificar, de forma preliminar (sem ajustar pelo efeito das demais variáveis), que algumas variáveis apresentaram associação significativa com a com o número de variedades de frutas incluídas na dieta da criança. Ao nível de significância de 5%, destacam-se as seguintes variáveis: anos de estudo da mãe, anos de estudo do pai, ocupação da mãe, renda familiar *per-capita*, o fato de o pai biológico viver com a criança, a idade da mãe no nascimento de seu primeiro filho e o número de refeições diárias recebidas pela criança.

Dando sequência à análise, foi utilizado o método *forward* para a seleção de um conjunto de covariáveis para o modelo de regressão multinomial de chances proporcionais com efeitos aleatórios de unidade de saúde, considerando como variável resposta a variável categorizada correspondente ao número de variedades de frutas consumidas ao menos uma vez na semana. O modelo resultante contemplava as seguintes covariáveis: idade da mãe no nascimento do primeiro filho; se a criança ainda estava em aleitamento materno; se a mãe havia recebido orientações sobre o aleitamento materno; anos de estudo da mãe; renda familiar *per capita*; anos de estudo do pai e número de refeições diárias da criança (exceto o leite materno).

Uma vez ajustado o modelo indicado no processo de seleção verificou-se, via teste da razão de verossimilhanças (TRV), que as covariáveis relativas aos anos de estudo do pai e à renda familiar *per capita*, conjuntamente, não tinham efeito significativo na explicação da resposta ( $p = 0,73$ ). Aliado a isso, as duas covariáveis apresentaram elevada associação com a variável relativa aos anos de estudo da mãe ( $p < 0,001$  em ambos os casos). Como, além disso, essas duas variáveis apresentaram um grande número de dados faltantes, optou-se, então, por retirá-las do modelo, permanecendo as outras cinco anteriormente relacionadas.

Para verificar se existia algum efeito de interação entre as cinco covariáveis selecionadas para compor o modelo, foram calculados os valores dos AIC's dos modelos resultantes da inclusão das interações duplas, uma de cada vez. Nenhuma das interações proporcionou redução no valor do AIC do modelo sem interações. O mesmo foi verificado mediante inclusão conjunta de todas elas. Optou-se, portanto, pela não inclusão de interações ao modelo.

Tabela 1 - Frequências observadas para o número de variedades de frutas inseridas na dieta das crianças ao menos uma vez na semana

Variável	Categorias	n(%)	Número de variedade de frutas consumidas					
			0 ou 1		2 ou 3		4 ou mais	
Geral	-	370	-	47 (12,7%)	137 (37,0%)	186 (50,3%)		
Idade da mãe	≤ 20 anos	97 (25,4%)	12 (12,8%)	37 (39,4%)	45 (47,9%)			
	> 20 anos	276 (74,6%)	35 (12,7%)	100 (36,2%)	141 (51,1%)			
Anos de estudo da mãe ***	< 4 anos	55 (14,9%)	13 (23,6%)	26 (47,3%)	16 (29,1%)			
	4 a 8 anos	195 (52,7%)	27 (13,8%)	75 (38,5%)	93 (47,7%)			
	> 8 anos	119 (32,2%)	6 (5,0%)	36 (30,3%)	77 (64,7%)			
Anos de estudo do pai **	< 4 anos	45 (12,2%)	6 (13,3%)	21 (46,7%)	18 (40,0%)			
	4 a 8 anos	186 (50,3%)	30 (16,1%)	70 (37,6%)	86 (46,2%)			
	> 8 anos	112 (30,3%)	7 (6,3%)	33 (29,5%)	72 (64,3%)			
Ocupação da mãe *	Trabalha	104 (28,1%)	7 (6,7%)	33 (31,7%)	64 (61,5%)			
	Não trabalha	266 (71,9%)	40 (15,0%)	104 (39,1%)	122 (45,9%)			
Renda familiar per capita **	< 0,5 SM	283 (76,5%)	33 (11,7%)	118 (41,7%)	132 (46,6%)			
	0,5 a 1 SM	57 (15,4%)	5 (8,8%)	11 (19,3%)	41 (71,9%)			
Pai biológico mora com a criança *	Sim	311 (84,1%)	38 (12,2%)	107 (34,4%)	166 (53,4%)			
	Não	59 (15,9%)	9 (15,3%)	30 (50,8%)	20 (33,9%)			
Gênero da criança	Masculino	199 (53,8%)	27 (13,6%)	72 (36,2%)	100 (50,3%)			
	Feminino	171 (46,2%)	20 (11,7%)	65 (38,0%)	86 (50,3%)			
Faixa etária da criança	6 a 8 meses	48 (13,0%)	9 (18,8%)	21 (43,8%)	18 (37,5%)			
	9 a 11 meses	65 (17,6%)	10 (15,4%)	28 (43,1%)	27 (41,5%)			
	12 a 17 meses	148 (40,0%)	17 (11,5%)	51 (34,5%)	80 (54,1%)			
	18 a 23 meses	109 (29,5%)	11 (10,1%)	37 (33,9%)	61 (56,0%)			
Número de filhos nascidos vivos	1 filho	149 (40,3%)	15 (10,1%)	52 (34,9%)	82 (55,0%)			
	2 a 3 filhos	157 (42,4%)	21 (13,4%)	58 (36,9%)	78 (49,7%)			
	> 3 filhos	64 (17,3%)	11 (17,2%)	27 (42,2%)	26 (40,6%)			
Tipo de parto	Normal	284 (76,8%)	37 (13,0%)	125 (44,0%)	122 (43,0%)			
	Cesariana	116 (31,4%)	10 (8,6%)	42 (36,2%)	64 (55,2%)			
Idade da mãe no nasc. do 1º filho **	≤ 20 anos	267 (72,2%)	37 (13,9%)	109 (40,8%)	121 (45,3%)			
	> 20 anos	103 (27,8%)	10 (9,7%)	28 (27,2%)	65 (63,1%)			
Orientação sobre o aleitamento materno	Sim	309 (83,5%)	39 (12,6%)	110 (35,6%)	160 (51,8%)			
	Não	60 (16,2%)	8 (13,3%)	26 (43,3%)	26 (43,3%)			
Peso da criança ao nascer	Inadequado	58 (15,7%)	7 (12,1%)	25 (43,1%)	26 (44,8%)			
	Adequado	312 (84,3%)	40 (12,8%)	112 (35,9%)	160 (51,3%)			
Criança mama no peito	Sim	171 (46,2%)	24 (14,0%)	69 (40,4%)	78 (45,6%)			
	Não	199 (53,8%)	23 (11,6%)	68 (34,2%)	108 (54,3%)			
Duração do aleitamento materno exclusivo	< 120 dias	240 (64,9%)	29 (12,1%)	86 (35,8%)	125 (52,1%)			
	≥ 120 dias	130 (35,1%)	18 (13,8%)	51 (39,2%)	61 (46,9%)			
Nº de refeições diárias da criança ***	1 a 3 refeições	98 (26,5%)	16 (16,3%)	44 (44,9%)	38 (38,8%)			
	4 a 5 refeições	204 (55,1%)	28 (13,7%)	80 (39,2%)	96 (47,1%)			
	> 5 refeições	64 (17,3%)	2 (3,1%)	12 (18,8%)	50 (78,1%)			

Códigos referentes à probabilidade de significância do teste qui-quadrado de associação: \*\*\*: 0<p<0,001; \*\*: 0,001≤p<0,01; \*: 0,01≤p<0,05; · : 0,05≤p<0,10.

A principal suposição do modelo adotado é a de chances proporcionais. Nesse sentido, torna-se necessário verificar se essa suposição do modelo está realmente sendo atendida. A suposição de chances proporcionais foi avaliada, inicialmente, por meio de um gráfico, apresentado na Figura 3, em que, para cada categoria de cada covariável são apresentados os valores dos dois logitos, calculados com base na amostra:  $logito_1 = \ln(P(y \leq 1)/P(y > 1)|x)$  e  $logito_2 = \ln(P(y \leq 2)/P(y > 2)|x)$ , sendo  $y$  a variável resposta categorizada, 1,2,3 números que simbolizam suas categorias, em ordem crescente de variedade de frutas consumidas e  $x$  o valor (categoria) fixado de uma das covariáveis. Sob a suposição de chances proporcionais, a diferença dos dois logitos deve ser aproximadamente constante para as diferentes categorias de uma covariável (HARRELL, 2001). Fica evidente, pela Figura 3, que a distância entre os dois diferentes símbolos, correspondentes aos dois logitos, é aproximadamente constante ao longo das categorias de uma mesma covariável, o que fornece indícios favoráveis à suposição de chances proporcionais.

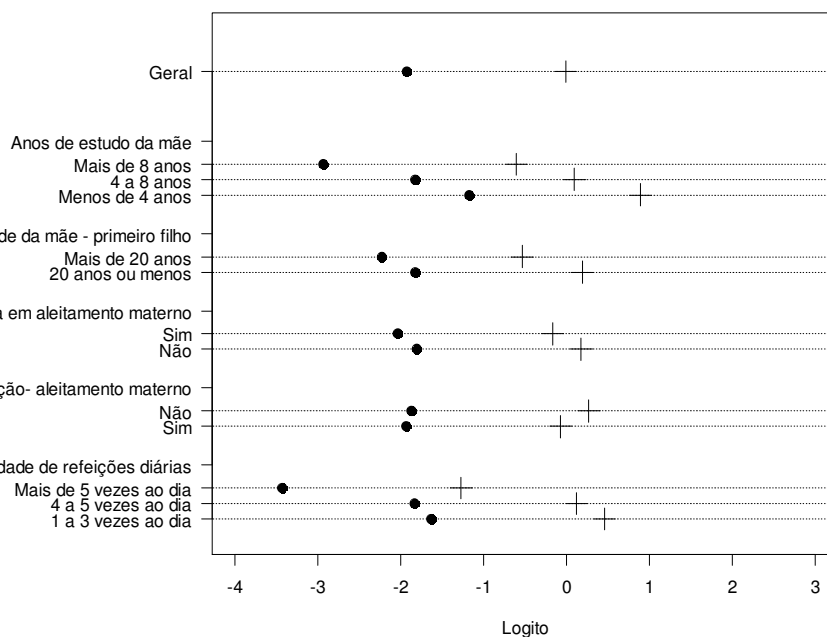


Figura 3 - Gráfico de logitos para as categorias de variedades de frutas consumidas, calculados para cada categoria das covariáveis. Os círculos correspondem aos logitos para  $y \leq 1$  e as cruzes aos logitos de  $y \leq 2$ .

Adicionalmente, utilizou-se o teste da razão de verossimilhanças, baseado em modelos encaixados, a fim de comparar o modelo em que se assumem chances proporcionais com os modelos em que uma das variáveis tem efeitos nominais, ou seja,

efeitos distintos para cada categoria da resposta (configurando modelos de chances proporcionais parciais). Também foi aplicado o teste da razão de verossimilhança para comparação de modelos encaixados a fim de confrontar o modelo de chances proporcionais ao modelo multinomial em que não se assume chances proporcionais para nenhuma das covariáveis. Os resultados são apresentados na Tabela 2, verificando-se que em nenhum dos casos têm-se diferença significativa entre os ajustes dos modelos com e sem a suposição de chances proporcionais. Assim, a suposição de chances proporcionais está validada e a opção pelo modelo de regressão multinomial de chances proporcionais está feita.

Tabela 2 - Testes da razão de verossimilhanças para a hipótese de chances proporcionais

Variável considerada com efeito nominal	Deviance	Diferença de deviance	Valor $p$
Nenhuma	560,94	-----	-----
Quantidade de refeições diárias recebidas pela criança	559,33	1,61	0,45
Idade da mãe no nascimento do 1º filho	559,71	1,22	0,27
Se a criança está em aleitamento materno	560,65	0,29	0,59
Se a mãe recebeu orientação sobre aleitamento materno	559,29	1,64	0,20
Anos de estudo da mãe	560,51	0,42	0,81
Todas	555,17	5,77	0,58

Ainda com o objetivo de diagnosticar o ajuste do modelo fez-se uso da estatística  $C$  de Hosmer e Lemeshow (HOSMER e LEMESHOW, 2000), que permite comparar as frequências amostrais com aquelas ajustadas pelo modelo. Como o modelo é composto por duas equações distintas (uma para cada logito) foram calculados valores separados para cada uma delas, obtendo-se, respectivamente  $C_1 = 5,61$  ( $p = 0,69$ ) e  $C_2 = 8,58$  ( $p = 0,38$ ), não havendo, em nenhum dos casos, evidências significativas contrárias à hipótese nula de que o ajuste é adequado.

Tendo verificado que o modelo multinomial de chances proporcionais tem bom ajuste, pôde-se, finalmente, analisar seus resultados. A Tabela 3 apresenta um resumo dos resultados produzidos pelo ajuste do modelo de chances proporcionais com efeito aleatório.

Com base no modelo ajustado, percebe-se que a chance de uma dieta mais variada de frutas aumenta conforme a escolaridade da mãe, tendo-se, em relação às mães com menos de quatro anos de estudo, chance 2,05 vezes maior para mães com 4 a 8 anos de estudo e 3,18 vezes maior para mães com mais de oito anos de estudo. Quanto à idade da mãe no nascimento do primeiro filho, tem-se uma chance 66% maior da criança receber maior variedade de frutas entre as mães que tiveram seu primeiro filho após completar 20 anos de idade. Já em relação ao número de refeições diárias recebidas pela criança, a chance de dieta mais variada de frutas é 4,73 vezes do grupo de crianças que recebem mais de 5 refeições diárias em relação às que recebem no máximo três. Todas as conclusões apresentadas são válidas controlando-se o efeito das demais variáveis incluídas no modelo, bem como o efeito aleatório correspondente às unidades de saúde.

Tabela 3 - Resumo do modelo de chances proporcionais ajustado

Parâmetro do modelo	Estimativa	Razão de Chance (IC:95%)
Intercepto (logito 1)	-0,73	--
Intercepto (logito 2)	1,45	--
Anos de estudo da mãe:		
Menos de quatro anos	0,00	--
4 a 8 anos	0,72	2,05 (1,14;3,70)
Mais de 8 anos	1,16	3,18 (2,50;3,86)
Idade da mãe ao nascimento do primeiro filho:		
Menos de 20 anos	0,00	--
20 anos ou mais	0,51	1,66 (1,14;2,18)
Presença do aleitamento materno:		
Não	0,00	--
Sim	0,30	1,35 (0,93;1,78)
Orientação sobre aleitamento materno:		
Sim	0,00	--
Não	-0,44	0,64 (0,36;1,14)
Quantidade de refeições diárias da criança (exceto o leite materno)		
0 a 3	0,00	--
4 ou 5	0,35	1,42 (0,94;1,89)
6 ou mais	1,80	4,29 (2,16;8,53)

De acordo com o modelo ajustado, a maior probabilidade de consumo mais variado de frutas é verificada para crianças que não recebem mais o leite materno, têm mais de 5 refeições diárias, cujas mães têm mais de 8 anos de estudo, tendo sido orientadas sobre o aleitamento materno e tendo seu primeiro filho com mais de 20 anos. A probabilidade estimada de uma criança com este perfil pertencer à categoria referente ao consumo de 4 ou mais variedades de frutas com uma frequência mínima semanal é 0,9308 enquanto a probabilidade dela estar na categoria referente ao consumo de duas ou três variedades é 0,0609 e na categoria de 0 ou uma variedade de frutas é 0,0083. Por outro lado, as crianças que possuem menor probabilidade de ter um consumo variado de frutas são aquelas que ainda são amamentadas, recebem de 1 a 3 refeições diárias, cujas mães têm menos de 4 anos de estudo, não tendo sido orientadas sobre o aleitamento materno e tendo seus primeiros filhos com menos de 20 anos. A probabilidade estimada de uma criança com estas características registrar um consumo de 4 ou mais variedades de frutas com frequência mínima semanal é igual a 0,0967, para duas ou três frutas é de 0,7646 e para 0 ou uma variedade de frutas é igual a 0,1387. Tais probabilidades foram estimadas

considerando-se o efeito de unidade de saúde igual à média (zero). Na Figura 4 são apresentadas outras probabilidades estimadas pelo modelo ajustado, permitindo avaliar melhor o efeito das covariáveis.

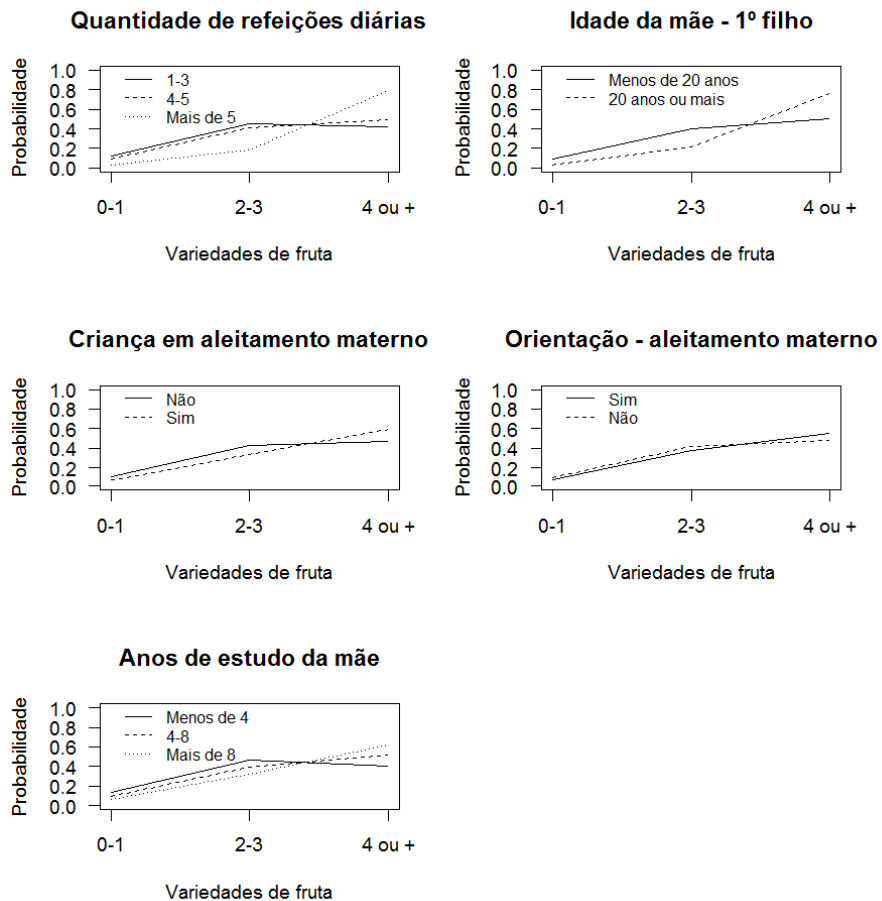


Figura 4 - Probabilidades ajustadas pelo modelo multinomial de chances proporcionais com efeito aleatório para as categorias de consumo de frutas segundo cada covariável. Em cada gráfico as probabilidades são calculadas fixando as demais covariáveis (inclusive o efeito aleatório) em seus valores médios.

Com relação ao efeito aleatório associado às unidades de saúde, teve-se  $\hat{\sigma}_\alpha^2 = 0,1$ , ( $p = 0,25$ , pelo TRV). Embora o efeito relativo às unidades de saúde seja não significativo, optou-se por mantê-lo no modelo, por refletir um agrupamento regional das crianças e para explorar as estimativas produzidas para cada unidade, buscando identificar unidades de saúde (e regiões do município) em situações mais críticas quanto ao consumo menos variado de frutas por parte das crianças atendidas.

Com as estimativas desses efeitos seria possível criar um plano de ação para as Unidades de Saúde que apresentassem estimativas mais baixas, possivelmente tomando como referência de atuação as Unidades de Saúde que apresentaram maiores estimativas. Os efeitos aleatórios estimados para as Unidades de Saúde, acompanhados de intervalos de 95% de confiança, são apresentados na Figura 5, com as unidades de saúde codificadas por números de 1 a 22. Embora todas as unidades apresentem os correspondentes intervalos incluindo o valor zero (efeito nulo), nota-se que existe variação no valor das estimativas deste efeito, sendo o menor valor verificado para a unidade 2, seguida pelas unidades 5, 1 e 19, o que fornece indício de dieta com menor variedade de frutas dentre as crianças atendidas por tais unidades.

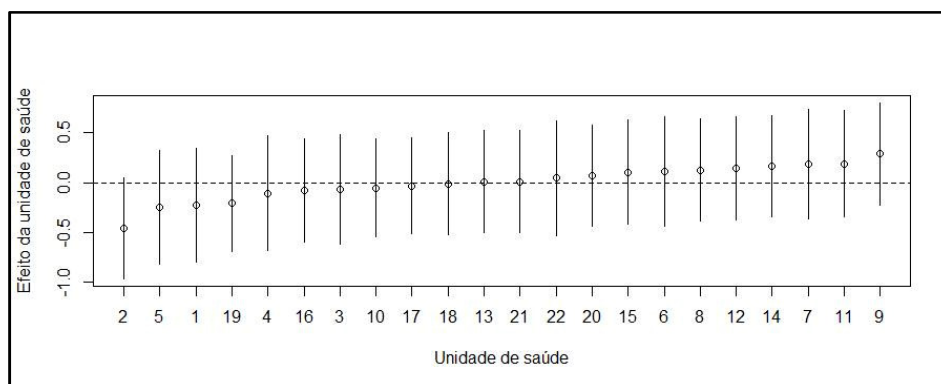


Figura 5 - Efeitos estimados e intervalos de confiança (95%) para as unidades de saúde.

## Conclusões

Uma alimentação adequada, de acordo com as recomendações dos órgãos responsáveis pela promoção da saúde, é necessária para que a criança cresça e se desenvolva adequadamente, particularmente quanto às crianças menores de dois anos, pois é nessa fase da vida que ocorrem as maiores aquisições em termos de crescimento, tanto físico e motor quanto intelectual. Tendo isso em vista, o presente trabalho procurou analisar o consumo de frutas por crianças na faixa etária mencionada e identificar fatores associados a um número mais variado de frutas na alimentação de crianças do município de Guarapuava-PR, participantes do Programa Saúde da Família.

Inicialmente, foi possível verificar que banana, maçã e laranja eram as frutas consumidas com maior frequência. Os consumos das 15 frutas relacionadas no Questionário de Frequência Alimentar foram resumidos por meio de uma nova variável, referente ao número de variedades de frutas consumidas ao menos uma vez por semana. Procedeu-se então com a categorização dessa variável e ajuste de diferentes modelos de regressão para dados multinomiais, considerando ainda a inclusão de um efeito aleatório para acomodar o agrupamento das crianças segundo as unidades de saúde em que eram atendidas.

Após avaliação e comparação de diferentes modelos, optou-se pelo modelo de regressão multinomial de chances proporcionais (com efeito aleatório), segundo o qual tinha-se uma dieta mais variada de frutas para crianças que se alimentavam mais de cinco vezes ao dia e cujas mães tinham mais de oito anos de escolaridade e mais de 20 anos no nascimento do primeiro filho. Outras duas variáveis compuseram o modelo final: se a mãe tinha recebido orientação quanto ao aleitamento materno durante a gravidez e se a criança ainda era amamentada. Embora não se tenha constatado variância significativa associada às unidades de saúde, o efeito aleatório foi mantido no modelo e permitiu identificar unidades de saúde com resultados ligeiramente mais críticos.

Embora o presente trabalho configure um estudo local, boa parte de seus resultados são compatíveis aos observados em pesquisas mais amplas. Segundo dados da Pesquisa de Orçamentos Familiares – POF, realizada no biênio 2008-2009 (IBGE, 2011), as três frutas mais consumidas foram banana, laranja e maçã, coincidindo com as três frutas mais consumidas pelas crianças em Guarapuava. Além disso, outro resultado revelado pela POF é a associação positiva entre as classes de renda e frequências de consumo e valores per capita de consumo de verduras, saladas cruas e frutas. Vale ressaltar que, no presente estudo, verificou-se associação positiva entre a escolaridade da mãe e a variedade de frutas consumidas pela criança, destacando-se que a escolaridade da mãe estava fortemente associada à escolaridade do pai e à renda da família.

INÁCIO, L. V., FORBELLONE, A. A., TACONELI, C. A., ALMEIDA, C. C. B. Analysis of fruit intake by children from a county in the State of Paraná-Brazil using multinomial models with random effects. *Rev. Bras. Biom.* São Paulo, v.33, n.4, p.517-533, 2015.

- *ABSTRACT: The introduction of complementary foods to breast milk in children aged 6-24 months has fundamental importance to ensure a healthy, physical, suitable motor and intellectual growth. In this sense, the inclusion of fruits, in quantity and variety, adds a large number of nutrients, essential to the child's diet. This study aimed to identify demographic, socioeconomic and behavioral factors related to the variety of fruits consumed by children aged 6-24 months, attended by health units of Sistema Único de Saúde (SUS) in Guarapuava, Paraná (Brasil). Through application of mixed multinomial models, we found that the consumed variety of fruits is related to the mothers' education, their ages at the birth of first child and the number of meals provided to children. A random effect was incorporated referent to health units. Although not statistically significant, their inclusion identified groups under less varied fruit consumption.*
- *KEYWORDS: Categorical data; multinomial data; random effects; food intake; fruits.*

## Referências

- ABREU, M. N. S.; SIQUEIRA, A. L.; CAIAFFA, W. T. Ordinal logistic regression in epidemiological studies. *Revista de Saúde Pública*, v.43, n.1, p.183-194, 2009.
- AGRESTI, A. *An Introduction to Categorical Data Analysis*. New York: John Wiley & Sons, 1996.



- AKRÉ, J. *Alimentação infantil: bases fisiológicas*. Genebra: Organização Mundial da Saúde, 1997.
- BRASIL. Ministério da Saúde. *Guia alimentar para crianças menores de 2 anos*. Brasília: Editora do Ministério da Saúde, 2005. (Serie A. Normas e Manuais Técnicos).
- BRASIL. Ministério da Saúde. *Saúde da Criança: nutrição infantil, aleitamento materno e alimentação complementar*. Brasília: Editora do Ministério da Saúde, 2009. (Cadernos de Atenção Básica, n. 23) (Série A. Normas e Manuais Técnicos).
- BRASIL. Ministério da Saúde. *Dez passos para uma alimentação saudável: guia alimentar para crianças menores de dois anos*. Brasília: Editora do Ministério da Saúde, 2010. (Serie A. Normas e Manuais Técnicos).
- BRECAILO, M. K. *Aleitamento materno: estudo com crianças menores de dois anos atendidas pelo Programa Saúde da Família no município de Guarapuava-PR*. 113 f. Tese (Mestrado em Nutrição) - Centro de Ciências da Saúde, Universidade Federal de Santa Catarina, Florianópolis, 2007.
- CASELLA, G.; BERGER, R. L. *Inferência Estatística*. São Paulo: Cengage Learning, 2010.
- CHARNET, R. *et al. Análise de modelos de regressão linear: com aplicações*. 2. ed. Campinas: Editora da UNICAMP, 2008.
- CHRISTENSEN, R. H. B. *ordinal – Regression Models for Ordinal Data R package version 2012.09-11*. 2012. <<http://www.cran.r-project.org/package=ordinal>>.
- HARRELL, Frank E. *Regression modeling strategies*. Springer Science & Business Media, 2001.
- HOSMER, D. W.; LEMESHOW, S. *Applied Logistic Regression*. 2. ed. New York: John Wiley & Sons, 2000.
- IBGE. *Pesquisa de Orçamentos Familiares 2008-2009: análise do consumo alimentar pessoal no Brasil*. Rio de Janeiro: IBGE, 2011.
- MONTE, C. M. G.; GIUGLIANI, E. R. J. Recomendações para alimentação complementar da criança em aleitamento materno. *J Pediatr.*, v. 80, n. 5, p. 131-41, 2004.
- PAULA, G. A. *Modelos de Regressão: com apoio computacional*. São Paulo: IME-USP, 2013. Disponível em: <[http://www.ime.usp.br/~giapaula/texto\\_2013.pdf](http://www.ime.usp.br/~giapaula/texto_2013.pdf)>. Acesso em: 16/09/2013.
- R CORE TEAM (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: <<http://www.R-project.org>>.
- TUTZ, G.; HENNEVOGL, W. Random effects in ordinal regression models. *Computational Statistics & Data Analysis*, v.22, n.5, p.537-557, 1996.

Recebido em 26.11.2014

Aprovado após revisão em 31.07.2015