

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

INSTITUTO DE MATEMÁTICA

DEPARTAMENTO DE ESTATÍSTICA

Os fatores de risco associados aos nascidos vivos com baixo peso em Porto Alegre, estudados a partir do SINASC.

Vânia Naomi Hirakata

Orientadora: Marilene Bandeira

Trabalho apresentado para obtenção do grau de Bacharel em Estatística, Universidade Federal do Rio Grande do Sul.

UFRGS
SISTEMAS DE BIBLIOTECAS
BIBLIOTECA SETORIAL DE MATEMÁTICA

Porto Alegre, janeiro de 1993.

Agradecimentos

A profa. Marilene que, com seu conhecimento e dedicação orientou-me na realização desta monografia.

A epidemiologista Denise que, além do incentivo, com sua competência, me prestou orientação técnica na área da saúde.

A minha família e ao Eduardo, pelo apoio e compreensão.

Aos colegas e amigos, pelo companheirismo e amizade, que se estenderam ao longo de todo este curso.

A Secretaria Municipal da Saúde e Serviço Social, pela concessão das informações utilizadas neste trabalho.

Sumário

1	Introdução	4
1.1	SINASC	6
1.2	Objetivos	8
1.2.1	Geral	8
1.2.2	Específicos	8
2	Metodologia	9
2.1	Fonte	9
2.2	Variáveis	9
2.3	Regressão Logística	12
2.3.1	Univariada	12
2.3.2	Múltipla	20
3	Resultados e Discussão	27
3.1	Análise Univariada	27
3.2	Análise Bivariada	37
3.3	Análise Multivariada	40
4	Conclusão	43
5	Bibliografia	44
6	Anexos	46
6.1	Declaração de Nascido Vivo	46
6.2	Resultados nas saídas de computador	48

1 INTRODUÇÃO

O presente trabalho foi realizado para a conclusão do curso de Bacharelado em Estatística, da Universidade Federal do Rio Grande do Sul.

Este trabalho utiliza as informações disponíveis na Secretaria Municipal da Saúde e Serviço Social do município de Porto Alegre e que fazem parte do Sistema de Informações sobre os Nascidos Vivos (SINASC). O objetivo principal é mostrar como os dados constantes na Declaração de Nascido Vivo podem ajudar a compreender melhor as razões do baixo peso nos recém nascidos deste município, durante os meses de março a dezembro de 1992.

Segundo a Organização Mundial da Saúde (OMS), todos os nascidos vivos que apresentem peso inferior a 2500g devem ser classificados como de baixo peso ao nascer (OMS, 1986).

O interesse no baixo peso ao nascer deve-se à sua relação com a mortalidade infantil (morte de crianças menores de um ano), visto ser ele (baixo peso ao nascer) um dos indicadores de risco para a mesma, juntamente com idade da mãe e idade gestacional, entre outros disponíveis (TESCH, MACHADO & CARVALHO, 1993).

Também o peso ao nascer determina a sobrevivência da criança no primeiro ano de vida e, principalmente, no primeiro mês (TESCH, MACHADO & CARVALHO, 1993). A OMS reconhece o peso ao nascer como fator isolado mais importante para a sobrevivência infantil (OMS, 1986).

Como o evento de interesse deste trabalho é o baixo peso ao nascer, resolveu-se trabalhar apenas com as variáveis do SINASC que fossem fatores de risco para o mesmo. Por este motivo, trabalha-se apenas com as variáveis sexo, idade e instrução maternas, idade gestacional, tipo de gravidez, filhos mortos e vivos de gestações anteriores, e, evidentemente, o peso ao nascer.

Este estudo é feito por meio de estatísticas descritivas (análise univariada) e regressão logística simples (análise bivariada) e múltipla (análise multivariada).

É importante lembrar que na Declaração de Nascido Vivo não se encontram todas as causas que podem determinar o baixo peso ao nascer de um nascido vivo.

Por outro lado, as informações pertencentes ao SINASC poderiam ser, e certamente serão, a base de outros estudos relacionados, principalmente, com a mortalidade infantil e, conseqüentemente, com o baixo peso ao nascer, entre muitos outros fatores de risco para a mesma.

1.1 SINASC

A partir da necessidade de se obter dados sobre nascidos vivos direto da principal fonte, o registro hospitalar, o Grupo de Estatísticas Vitais do Ministério da Saúde (GEVIMS) recomendou a implantação do Sistema de Informação sobre Nascidos Vivos, SINASC, em 1989. Estas informações deveriam percorrer três níveis: o municipal, o estadual e o federal.

A Declaração de Nascido Vivo (DN) é o documento básico do SINASC. Sua emissão é feita em três vias carbonadas de igual teor. A 1ª via, branca, é destinada à Secretaria Municipal da Saúde e Serviço Social (SMSSS), a 2ª, amarela, é destinada à família, sendo indispensável para o registro civil do recém nascido, e a 3ª via, rosa, é enviada à Secretaria de Saúde e Meio Ambiente do Estado do Rio Grande do Sul (SSMA-RS). A DN pode ser preenchida no hospital, se foi prestada assistência hospitalar, ou no cartório caso o parto tenha sido domiciliar.

Em Porto Alegre, a implantação definitiva do SINASC ocorreu em janeiro de 1992 com a coordenação da SSMA-RS, ficando sob sua responsabilidade até março do mesmo ano, quando a SMSSS passou a coordená-lo.

Juntamente ao SINASC foi criado em Porto Alegre um programa pertencente ao Sistema de Vigilância à Saúde da Criança,

conhecido como Prá Nenê.

Atualmente, o programa acompanha todos os recém nascidos de risco que residem nas áreas de abrangência das unidades sanitárias. Com este programa, a SMSSS pretende reduzir a mortalidade neonatal (morte de crianças com menos de um mês de vida), que corresponde a mais de 50% dos óbitos de menores de um ano em Porto Alegre (Aerts, 1992).

O SINASC pode ser considerado, por isso, uma importante fonte de informações, não apenas do ponto de vista epidemiológico, como também para as ações de vigilância da saúde das crianças.

1.2 OBJETIVOS

1.2.1 Geral

Analisar os fatores de risco para o baixo peso ao nascer a partir do SINASC, considerando as suas distribuições e suas relações com a variável de interesse, nos recém nascidos de Porto Alegre, RS, durante os meses de março a dezembro de 1992.

2.2.2 Específicos

- 1 - Descrever e analisar a distribuição dos nascidos vivos de Porto Alegre segundo sexo, peso ao nascer, idade gestacional, tipo de gravidez, idade e instrução materna, filhos tidos vivos e mortos.
- 2 - Estudar a relação de cada variável citada no item anterior com o baixo peso ao nascer por meio de regressão logística.
- 3 - Estudar a relação de todas as variáveis que puderam ser consideradas fatores de risco no item 2 com a variável peso ao nascer.

UFRRS
SISTEMAS DE BIBLIOTECAS
BIBLIOTECA SETORIAL DE MATEMÁTICA

2 METODOLOGIA

2.1 Fonte

O presente trabalho teve como fonte a Declaração de Nascido Vivo (DN), documento primordial do SINASC. As informações encontradas aqui referem-se aos meses de março a dezembro de 1992. Foram consideradas todas as DNs que chegaram à SMSSS até março de 1993. Em anexo, pode ser encontrada uma cópia deste documento.

2.2 Variáveis

As variáveis encontradas na DN que se mostraram mais interessantes para este estudo puderam ser divididas em três grandes grupos de informações:

- sobre o recém nascido;
- sobre a gestação e o parto;
- sobre a mãe.

Relacionadas ao recém nascido estão as seguintes variáveis:

PESO - indica o peso ao nascer da criança, podendo ser identificado até a 5ª hora após o nascimento. Na DN, o peso deve ser informado da maneira mais exata possível.

SEXO - diz respeito ao sexo da criança.

No que diz respeito à gestação e ao parto, podemos encontrar as variáveis:

GESTACAO - informa a duração da gestação, um dos fatores que determinam o crescimento intra-uterino. As categorias em que esta variável encontra-se estratificada na DN seguem a orientação proposta para a 10ª Revisão da Classificação Internacional de Doenças, e são, 0 a 21, 22 a 27, 28 a 36, 37 a 41, 42 e mais semanas de gestação (JORGE et al., 1992).

TIPOGRAV - fornece informações sobre o tipo de gravidez. Existem estudos que apontam um risco maior de morte nas crianças nascidas de gestações múltiplas, já que estas são, também, as que apresentam em média menor peso ao nascer (JORGE et al., 1992).

As variáveis que informam sobre a mãe são:

IDADEMAE - refere-se à informação sobre a idade materna. Esta variável é muito importante, não só para estudos demográficos, mas também como um fator de risco para o

baixo peso ao nascer. Segundo Jorge, estudos têm demonstrado que este risco é maior entre as mães mais jovens e, da mesma forma, entre as mães com 35 anos e mais.

INSTRMAE - indica o grau de instrução da mãe. Constitui-se de uma informação sócio-econômica muito relevante e de fácil obtenção.

FILVIVOS e FILMORTO - fornece informações sobre as gestações anteriores. De acordo com Jorge, há evidências indicando a existência de maior risco de mortalidade perinatal entre as mulheres que estão tendo o primeiro filho e as que têm cinco ou mais filhos anteriores.

2.3 Regressão Logística

A regressão logística é uma técnica estatística de análise de dados que tem sido utilizada com bastante frequência quando se deseja descrever a relação entre uma variável resposta (dependente) e um ou mais fatores de risco (variáveis independentes), principalmente se esta variável resposta for discreta, podendo ter duas ou mais categorias.

O principal objetivo da regressão logística é, como a de muitas outras técnicas de elaboração de modelos, a procura do melhor ajuste que explique a relação entre a variável resposta (dependente) e o conjunto de variáveis independentes (fatores de risco).

2.3.1 O Modelo de Regressão Logística

Seja x a variável independente e Y a variável dependente dicotômica. A forma específica do modelo de regressão logística a ser usada é:

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \quad (1.1)$$

O valor $\pi(x) = E(Y/x)$ representa a média condicional de

Y dado x quando a distribuição logística é utilizada. A transformação logit é uma transformação de $\pi(x)$ que será o centro do estudo de regressão logística. Esta transformação é definida em termos de $\pi(x)$ como:

$$g(x) = \ln \left[\frac{\pi(x)}{1 - \pi(x)} \right] = \beta_0 + \beta_1 x \quad (1.2)$$

O logit $g(x)$ é linear em seus parâmetros, pode ser contínuo, e pode se estender de $-\infty$ a $+\infty$, dependendo apenas da extensão de x .

A distribuição condicional da variável resposta dado por x pode ser expressa como:

$$y = \pi(x) + \varepsilon$$

onde o valor ε pode assumir um de apenas dois valores possíveis. Se $y = 1$, então $\varepsilon = 1 - \pi(x)$ com probabilidade $\pi(x)$ e, se $y = 0$, então $\varepsilon = -\pi(x)$ com probabilidade $1 - \pi(x)$. Logo, ε possui uma distribuição com média zero e variância $\pi(x)[1 - \pi(x)]$. Ou seja, a distribuição condicional da variável resposta segue uma distribuição binomial com probabilidade dada pela média condicional $\pi(x)$.

O Ajuste do Modelo de Regressão Logística

Suponha que se tivesse uma amostra de n observações independentes do par (x_i, y_i) , $i = 1, 2, \dots, n$, onde y_i é o valor de uma variável resposta e x_i é o valor da variável independente

para o i -ésimo sujeito. Além disso, deve-se assumir que a variável resposta tenha sido codificada como zero ou um indicando a ausência ou a presença da característica, respectivamente. Para o ajuste do modelo de um conjunto de dados, deve-se estimar os valores dos parâmetros desconhecidos β_0 e β_1 .

O modelo de estimação utilizado aqui é semelhante ao método de mínimos quadrados utilizado na regressão linear e denomina-se método de máxima verossimilhança. De um modo geral, este método produz valores para os parâmetros desconhecidos, que maximizam a probabilidade de se obter o conjunto observado de dados. Antes que se possa aplicar o método de máxima verossimilhança deve-se construir uma função, chamada **função de verossimilhança**, que expressa a probabilidade dos dados observados como uma função dos parâmetros desconhecidos. Os **estimadores de máxima verossimilhança** destes parâmetros são escolhidos como sendo os valores que maximizem esta função.

Se y estiver codificada como 0 ou 1, então a expressão para $\pi(x)$ dada na equação (1.1) fornece (para um valor arbitrário de $B' = (\beta_0, \beta_1)$, o vetor de parâmetros) a probabilidade condicional de Y ser igual a 1 dado x , e será denotado como $P(Y=1/x)$. Conseqüentemente, o valor $1 - \pi(x)$ dá a probabilidade de Y ser igual a 0 dado x , $P(Y=0/x)$. Então, para aqueles pares (x_i, y_i) , onde $y_i = 1$ a contribuição para a função de verossimilhança é $\pi(x_i)$ e, para aqueles onde $y_i = 0$, a contribuição para a função de verossimilhança é $1 - \pi(x_i)$, onde o valor $\pi(x_i)$ denota o valor de $\pi(x)$ calculado por meio de x_i . Uma

maneira conveniente de expressar a contribuição do par (x_i, y_i) para a função de verossimilhança é através do termo:

$$\zeta(x_i) = \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i} \quad (1.3)$$

Desde que se assuma a independência das observações, a função de verossimilhança é obtida pelo produto dos termos dados na expressão (1.3), como se segue:

$$l(\beta) = \prod_{i=1}^n \zeta(x_i) \quad (1.4)$$

O princípio da máxima verossimilhança especifica que seja utilizado como estimador de β o valor que maximize a expressão na equação (1.4). Porém, é matematicamente mais fácil trabalhar como o logaritmo natural (que, a partir de agora denominaremos apenas de log) da equação (1.4). Esta expressão, o log da verossimilhança, é definida como

$$\begin{aligned} L(\beta) &= \ln[l(\beta)] \\ &= \sum_{i=1}^n \{y_i \ln[\pi(x_i)] + (1 - y_i) \ln[1 - \pi(x_i)]\} \end{aligned} \quad (1.5)$$

Para encontrarmos o valor de β que maximize $L(\beta)$, deve-se calcular a derivada de $L(\beta)$ em relação a β_0 e β_1 e a expressão resultante deve ser igualada a zero. Estas equações são as seguintes

$$\sum_{i=1}^n [y_i - \pi(x_i)] = 0 \quad (1.6)$$

e

$$\sum_{i=1}^n x_i [y_i - \pi(x_i)] = 0 \quad (1.7)$$

e são denominadas **equações de verossimilhança**. Estas equações não são lineares em β_0 e β_1 e, por isso, requerem métodos especiais para sua solução. Estes métodos são iterativos por natureza e já estão programados em softwares estatísticos. Por ora estes métodos não serão estudados.

O valor de β dado pela solução das equações (1.6) e (1.7) é chamado **estimador de máxima verossimilhança** e será denotado como $\hat{\beta}$. Do mesmo modo, $\hat{\pi}(x_i)$ é o **estimador de máxima verossimilhança** de $\pi(x_i)$. Uma consequência interessante da equação (1.6) é que

$$\sum_{i=1}^n y_i = \sum_{i=1}^n \hat{\pi}(x_i)$$

Isto é, a soma dos valores observados de y é igual à soma dos valores estimados (esperados).

Teste de significância dos coeficientes

Depois de estimados os coeficientes, o primeiro olhar para o modelo ajustado concentra-se na avaliação da significância das variáveis no modelo. Isto geralmente envolve a formulação e o teste de uma hipótese estatística para determinar se a variável independente no modelo é relacionada significativamente com a variável resposta. O método para a realização deste teste é

completamente geral e difere do modelo com somente uma variável independente para o modelo multivariado apenas em detalhes específicos. Será discutido agora, somente a aplicação do método no primeiro modelo.

O princípio para a realização do teste de significância na regressão logística é: "Comparar os valores observados da variável resposta com os valores estimados obtidos de modelos com e sem a variável em questão" (HOSMER & LEMESHOW, 1989). Esta comparação é baseada na função log da verossimilhança definida na equação (1.5). Para entender melhor esta comparação, pode-se pensar que um valor observado da variável resposta é também um valor estimado resultante de um modelo saturado.

A comparação dos valores observados para com os estimados a partir da função de verossimilhança baseia-se na seguinte expressão:

$$D = -2 \ln \left[\frac{\text{verossimilhança do presente modelo}}{\text{verossimilhança do modelo saturado}} \right] \quad (1.8)$$

O valor no interior dos colchetes é chamado de razão de verossimilhança. Utiliza-se o seu logaritmo natural e ainda multiplicado por -2 pela necessidade de se obter um valor cuja distribuição fosse conhecida e, assim, pudesse ser usada para testar hipóteses. Este teste é chamado teste da razão de verossimilhança. Das equações (1.5) e (1.8), temos:

$$D = -2 \sum_{i=1}^n \left[y_i \ln \left(\frac{\hat{\pi}_i}{y_i} \right) + (1 - y_i) \ln \left(\frac{1 - \hat{\pi}_i}{1 - y_i} \right) \right] \quad (1.9)$$

onde $\hat{\pi}_i = \hat{\pi}(x_i)$.

Esta estatística D é chamada por alguns autores de "deviance", sendo muito importante em algumas aproximações para avaliar a bondade do ajuste (goodness-of-fit). O "deviance" para a regressão logística é semelhante à soma dos quadrados devido ao erro da regressão linear.

Com o objetivo de avaliar a significância de uma variável independente, deve-se comparar os valores de D com e sem a variável em questão. A alteração em D devido à sua inclusão no modelo é obtida através de:

$$G = D(p/ \text{ o modelo sem a variável}) - D(p/ \text{ o modelo com a variável})$$

Esta estatística exerce o mesmo papel que o numerador do teste F parcial na regressão linear. Pelo motivo de que a verossimilhança do modelo saturado é o mesmo para os dois valores de D , que, por sua vez são diferentes no cálculo de G , este pode ser expresso como:

$$G = -2 \ln \left[\frac{\text{verossimilhança sem a variável}}{\text{verossimilhança com a variável}} \right] \quad (1.10)$$

Para o caso específico da presença de apenas uma variável independente, pode se mostrar que, quando a variável não está no modelo, o estimador de máxima verossimilhança de β_0 é o $\ln(n_1/n_0)$ onde $n_1 = \sum y_i$ e $n_0 = \sum (1 - y_i)$ e que o valor estimado

é constante n_1/n . Neste caso, G pode ser escrito como:

$$G = -2 \ln \left[\frac{\left(\frac{n_1}{n} \right)^{n_1} \left(\frac{n_0}{n} \right)^{n_0}}{\prod_{i=1}^n \hat{\pi}_i^{y_i} (1 - \hat{\pi}_i)^{(1-y_i)}} \right] \quad (1.11)$$

ou

$$G = 2 \left\{ \sum_{i=1}^n [y_i \ln(\hat{\pi}_i) + (1-y_i) \ln(1-\hat{\pi}_i)] - [n_1 \ln(n_1) + n_0 \ln(n_0) - n \ln(n)] \right\} \quad (1.12)$$

Sob a hipótese de que β_1 é igual a zero, a estatística G segue uma distribuição χ^2 com 1 grau de liberdade.

Serão apresentados, também, outros testes estatisticamente equivalentes: o teste de Wald e o teste Score. As suposições necessárias para estes testes são as mesmas necessárias para o teste da razão de verossimilhança na equação (1.11).

O teste de Wald é obtido pela comparação do estimador de máxima verossimilhança do parâmetro $\hat{\beta}_1$, com a estimação do seu Erro Padrão (SE). A razão resultante,

$$W = \frac{\hat{\beta}_1}{\text{SE}(\hat{\beta}_1)}$$

sob a hipótese de que $\beta_1 = 0$, segue uma distribuição normal padrão.

O teste Score univariado é baseado na distribuição

condicional das derivadas, equação (1.7), dadas as derivadas na equação (1.6). Este teste utiliza o valor da equação (1.7), calculada usando $\beta_0 = \ln(n_1/n_0)$ e $\beta_1 = 0$. Sob o valor destes parâmetros, $\hat{\pi} = n_1/n = \bar{y}$. Assim, o lado direito da equação (1.7) fica sendo $\sum x_i (y_i - \bar{y})$. Pode-se demonstrar que sua variância estimada é $\bar{y}(1-\bar{y}) \sum (x_i - \bar{x})^2$. A estatística de teste para o teste Score é

$$ST = \frac{\sum_{i=1}^n x_i (y_i - \bar{y})}{\sqrt{\bar{y}(1-\bar{y}) \sum_{i=1}^n (x_i - \bar{x})^2}}$$

que segue uma distribuição normal padrão.

2.3.2 O Modelo de Regressão Logística Múltipla

Deve-se considerar um conjunto de p variáveis independentes x_1, x_2, \dots, x_p representadas pelo vetor x' . Seja a probabilidade condicional quando a variável independente está presente, denotada por $P(Y = 1 | x) = \pi(x)$. Então, o logit do modelo de regressão logística múltipla é dado pela equação:

$$g(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (2.1)$$

em qualquer caso,

$$\pi(x) = \frac{e^{g(x)}}{1 + e^{g(x)}} \quad (2.2)$$

Se algumas das variáveis independentes estiverem em

escala nominal, recomenda-se não usá-las da maneira em que estiverem. Nesta situação, o método indicado é o uso de variáveis dummy.

Em geral, se uma variável categórica tiver k valores possíveis, então serão necessárias $k - 1$ variáveis dummy. Isto é verdade desde que, exceto se houver declaração diferente, o modelo possua um termo constante.

Suponha que a j -ésima variável independente x_j tenha k_j níveis. A variável dummy $k_j - 1$ será denotada por D_{ju} e os coeficientes para estas variáveis serão denotados por β_{ju} , $u = 1, 2, \dots, k_j - 1$. Então, o logit para um modelo com p variáveis e a j -ésima variável sendo discreta, é expressa por:

$$g(x) = \beta_0 + \beta_1 x_1 + \dots + \sum_{u=1}^{k_j-1} \beta_{ju} D_{ju} + \beta_p x_p \quad (2.3)$$

O Ajuste do Modelo de Regressão Logística Múltipla

É assumida a existência de uma amostra de n observações independentes do par (x_i, y_i) , $i = 1, 2, \dots, n$. O ajuste do modelo exige que se obtenha a estimação do vetor $B' = (\beta_0, \beta_1, \dots, \beta_p)$. O método de estimação a ser usado é o de máxima verossimilhança. Existem $p + 1$ equações de verossimilhança obtidas por meio da diferenciação da função log da verossimilhança que dizem respeito aos $p + 1$ coeficientes. A equação de verossimilhança resultante

pode ser expressa da seguinte maneira:

$$\sum_{i=1}^n [y_i - \pi(x_i)] = 0 \quad (2.4)$$

e

$$\sum_{i=1}^n x_{ij} [y_i - \pi(x_i)] = 0, \text{ para } j = 1, 2, \dots, p \quad (2.5)$$

A solução desta equação requer um software com este propósito, que pode ser encontrado em vários programas estatísticos. Seja $\hat{\beta}$ a solução destas equações. Então, os valores ajustados para o modelo de regressão logística múltipla são $\hat{\pi}(x_i)$, o valor da expressão na equação (2.2) calculada usando $\hat{\beta}$ e x_i .

O método de estimação de variâncias e covariâncias dos coeficientes estimados segue a mesma teoria dos estimadores de máxima verossimilhança. Esta teoria especifica que os estimadores são obtidos da matriz da derivada segunda parcial da função log da verossimilhança. Estas derivadas parciais têm a seguinte forma geral:

$$\frac{\partial^2 L(\beta)}{\partial \beta_j^2} = - \sum_{i=1}^n x_{ij}^2 \pi_i (1 - \pi_i) \quad (2.6)$$

e,

$$\frac{\partial^2 L(\beta)}{\partial \beta_j \partial \beta_u} = - \sum_{i=1}^n x_{ij} x_{iu} \pi_i (1 - \pi_i) \quad (2.7)$$

para $j, u = 0, 1, 2, \dots, p$, onde π_i denota $\pi(x_i)$. Seja a matriz $(p+1) \times (p+1)$ contendo o negativo dos termos dados nas equações (2.6) e (2.7) denotada como $I(\beta)$. Essa matriz é chamada **matriz de**

informação. As variâncias e covariâncias dos coeficientes estimados são obtidos da inversa desta matriz que denota-se por $\Sigma(\beta) = I^{-1}(\beta)$. Exceto em casos muito especiais, não é possível escrever uma expressão explícita para os elementos desta matriz. Por isso, usa-se a notação $\sigma^2(\beta_j)$ para denotar o j-ésimo elemento da diagonal desta matriz, que será a variância de $\hat{\beta}_j$, e $\sigma(\beta_j, \beta_u)$ para denotar um elemento arbitrário que não pertença à diagonal, que será a covariância de $\hat{\beta}_j$ e $\hat{\beta}_u$. Os estimadores das variâncias e covariâncias, que serão denotados por $\hat{\Sigma}(\hat{\beta})$, são obtidos pela estimação do valor de $\Sigma(\beta)$ para cada $\hat{\beta}$. Usa-se $\hat{\sigma}^2(\hat{\beta}_j)$ e $\hat{\sigma}(\hat{\beta}_j, \hat{\beta}_u)$, $j, u = 0, 1, 2, \dots, p$, para denotar os valores nesta matriz. Na maior parte das vezes, utiliza-se somente os erros padrões estimados dos coeficientes estimados, os quais denota-se por:

$$SE(\hat{\beta}_j) = \sqrt{\hat{\sigma}^2(\hat{\beta}_j)}$$

para $j = 0, 1, 2, \dots, p$. Será usada esta notação no desenvolvimento dos métodos para testar os coeficientes e estimação do intervalo de confiança.

Uma formulação possível da matriz de informação, útil na discussão sobre o ajuste do modelo e avaliação do ajuste, pode ser $\hat{I}(\hat{\beta}) = X'VX$ onde X é uma matriz n por $p+1$ contendo os dados de cada sujeito, e V é uma matriz diagonal $n \times n$ com elemento geral $\hat{\pi}_i(1 - \hat{\pi}_i)$. Então, a matriz X será dada por:

$$X = \begin{bmatrix} 1 & x_{11} & \dots & x_{1p} \\ 1 & x_{21} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{np} \end{bmatrix}$$

e a matriz V por:

$$X = \begin{bmatrix} \hat{\pi}_1(1 - \hat{\pi}_1) & 0 & \dots & 0 \\ 0 & \hat{\pi}_2(1 - \hat{\pi}_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \hat{\pi}_n(1 - \hat{\pi}_n) \end{bmatrix}$$

Teste de significância dos coeficientes

Deseja-se ajustar um modelo de regressão logística múltipla particular, por isso deve-se iniciar com o processo de avaliação do modelo. O 1º passo neste processo é, geralmente, o cálculo da significância das variáveis no modelo. O teste da razão de verossimilhança para a significância geral dos p coeficientes para as variáveis no modelo é baseado em:

$$G = D(p/ \text{o modelo sem a variável}) - D(p/ \text{o modelo com a variável})$$

onde

$$D = -2 \sum_{i=1}^n \left[y_i \ln \left(\frac{\hat{\pi}_i}{y_i} \right) + (1 - y_i) \ln \left(\frac{1 - \hat{\pi}_i}{1 - y_i} \right) \right]$$

e, $\hat{\pi}$ será o vetor contendo p+1 parâmetros $\hat{\beta}$.

Compara-se os valores de D com e sem a variável independente na equação para podermos avaliar a significância de uma variável independente. O valor de G indica a variação em D devido à inclusão da variável independente no modelo. A

estatística G representa, na regressão logística, o mesmo que o numerador do teste F parcial, na regressão linear. Pode-se calcular G da seguinte maneira:

$$G = -2 \ln \left[\frac{\text{verossimilhança sem a variável}}{\text{verossimilhança com a variável}} \right]$$

Sob a hipótese nula de que os coeficientes de inclinação das covariáveis no modelo são iguais a 0, a distribuição de G será χ^2 com p graus de liberdade.

Antes de concluir que um ou mais coeficientes não são iguais a 0, ou seja, que rejeita-se H_0 , pode-se olhar a estatística de teste Wald univariada, $W_j = \hat{\beta}_j / \widehat{SE}(\hat{\beta}_j)$. Sob a hipótese de que um coeficiente é zero, esta estatística segue uma distribuição normal. Assim, o valor desta estatística pode nos dar uma indicação de quais das variáveis no modelo podem ou não ser significativas.

Quando uma variável independente categórica for incluída (ou excluída) de um modelo, todas as suas dummies devem ser incluídas (ou excluídas); caso contrário, deve-se recodificar a variável. Em razão dos múltiplos graus de liberdade, é necessário cautela no uso da estatística Wald (W) para avaliar a significância dos coeficientes. Por exemplo, se a estatística W de dois coeficientes excedem o valor 2, então pode-se concluir que as variáveis dummy são significativas. Se, de outro modo, um coeficiente tiver o valor de Wald superior a 2, e outro inferior, então não se tem certeza da contribuição da variável no modelo.

Neste caso, a decisão de incluir ou excluir esta variável deve ser feita em conjunto com pesquisadores que conheçam o assunto em questão.

3 RESULTADOS E DISCUSSÃO

Apresenta-se agora, os resultados deste estudo. Todas as análises foram feitas no software SPSS (Statistical Package for Social Sciences) versão 5.0.2 (de 11/01/93), para Windows. Este pacote estatístico era o único dentro do nosso alcance, que possuía o módulo de regressão logística e que suportava uma quantidade tão grande de informações (21980 registros).

3.1 Análise Univariada

Pode-se observar a distribuição dos recém nascido segundo o peso ao nascer na Figura 1. Os picos que se sobressaem correspondem aos pesos terminados em dígitos 00 e 50, também denominados dígitos de preferência. Eles aparecem devido à facilidade de sua identificação na balança.

Esta variável foi codificada para que pudesse ser utilizada na regressão logística. Esta nova distribuição pode ser encontrada na Tabela 1.

Foram considerados recém nascidos de baixo peso aqueles que apresentaram peso abaixo de 2500g.

BRASILIA, 15 de Novembro de 1993

A prevalência de baixo peso ao nascer nas crianças nascidas em Porto Alegre, 1992, se aproxima à da apresentada para a região Sul (9,5%) pela Pesquisa Nacional sobre Saúde e Nutrição (PNSN) realizada em 1989, pelo IBGE, em todo o território nacional (MONTEIRO et al., 1992).

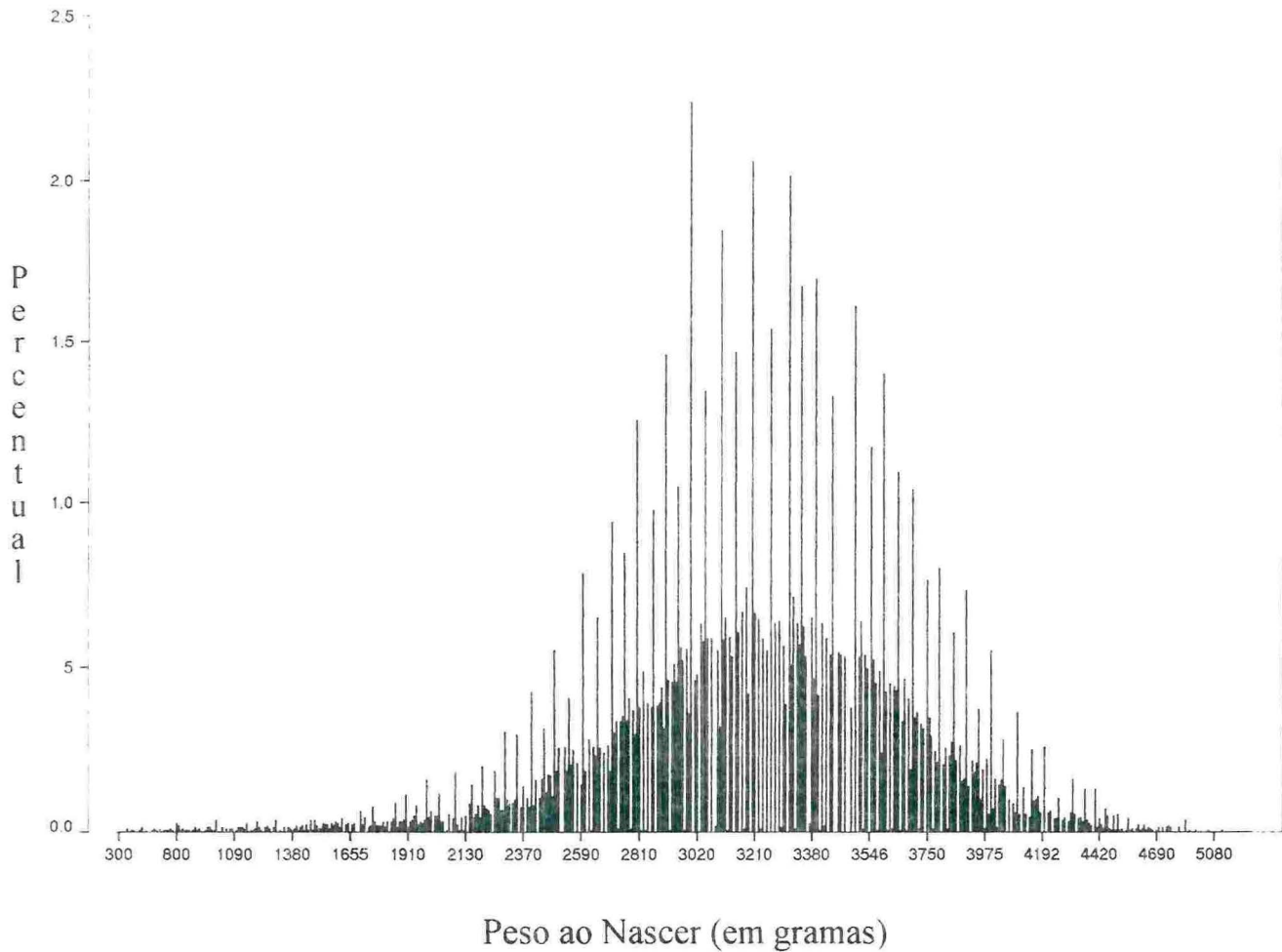


Figura 1: Distribuição dos nascidos vivos segundo a variável peso ao nascer, Porto Alegre, 1992.

TABELA 1 - DISTRIBUIÇÃO DOS NASCIDOS VIVOS SEGUNDO O PESO AO NASCER, 1992, PORTO ALEGRE, RS.

PESO AO NASCER	N	%
Baixo	1952	8,9
Normal	20008	91,0
Ignorado	20	0,1
TOTAL	21980	100,0

A informação sobre sexo confirma os resultados de alguns estudos que têm mostrado haver um diferencial em relação ao sexo, com a ocorrência de mais nascimentos do sexo masculino do que do feminino (JORGE et al., 1992). Esta distribuição pode ser vista na Tabela 2.

TABELA 2 - DISTRIBUIÇÃO DOS NASCIDOS VIVOS SEGUNDO O SEXO, 1992, PORTO ALEGRE, RS.

SEXO	N	%
Masculino	11227	51,1
Feminino	10741	48,9
Ignorado	12	0,1
TOTAL	21980	100,0

A distribuição dos recém nascidos de Porto Alegre, segundo a idade gestacional, da forma em que é encontrada na DN é apresentada na Figura 2. Vale a pena comentar o fato de que 91%

das crianças nascidas em Porto Alegre se encontravam em sua idade gestacional adequada, entre 37 e 41 semanas.

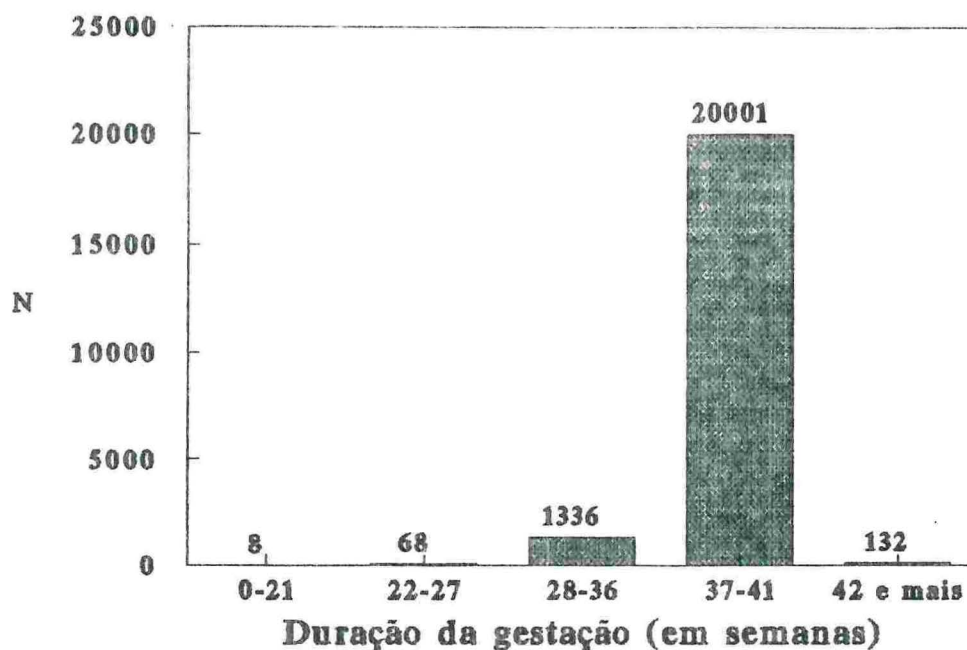


Figura 2: Distribuição dos nascidos vivos segundo a idade gestacional, 1992, Porto Alegre, RS.

Na Tabela 3, esta mesma variável aparece recodificada conforme a maneira em que será utilizada nas análises posteriores.

Os nascidos vivos prematuros são aqueles que nasceram

antes da 37ª semana de gestação, ou seja, correspondentes às 1ª, 2ª e 3ª categorias visualizadas no gráfico.

TABELA 3 - DISTRIBUIÇÃO DOS NASCIDOS VIVOS SEGUNDO A IDADE GESTACIONAL, 1992, PORTO ALEGRE, RS.

IDADE GESTACIONAL	N	%
Prematuro	1412	6,4
A termo	20436	93,0
Ignorado	132	0,6
TOTAL	21980	100,0

A distribuição dos nascidos vivos segundo o tipo de gravidez se encontra na Tabela 4. A categoria de gravidez múltipla inclui as crianças que tiveram 1 ou mais irmãos gêmeos envolvendo, assim, as categorias 2, 3 e 4 da codificação original (Figura 3). Chama a atenção o fato de que as gestações únicas constituíram quase a totalidade (98,3%) dos partos ocorridos no município de Porto Alegre.

TABELA 4 - DISTRIBUIÇÃO DOS NASCIDOS VIVOS SEGUNDO O TIPO DE GRAVIDEZ, 1992, PORTO ALEGRE, RS.

TIPO DE GRAVIDEZ	N	%
Única	21607	98,3
Múltipla	355	1,6
Ignorado	18	0,1
TOTAL	21980	100,0

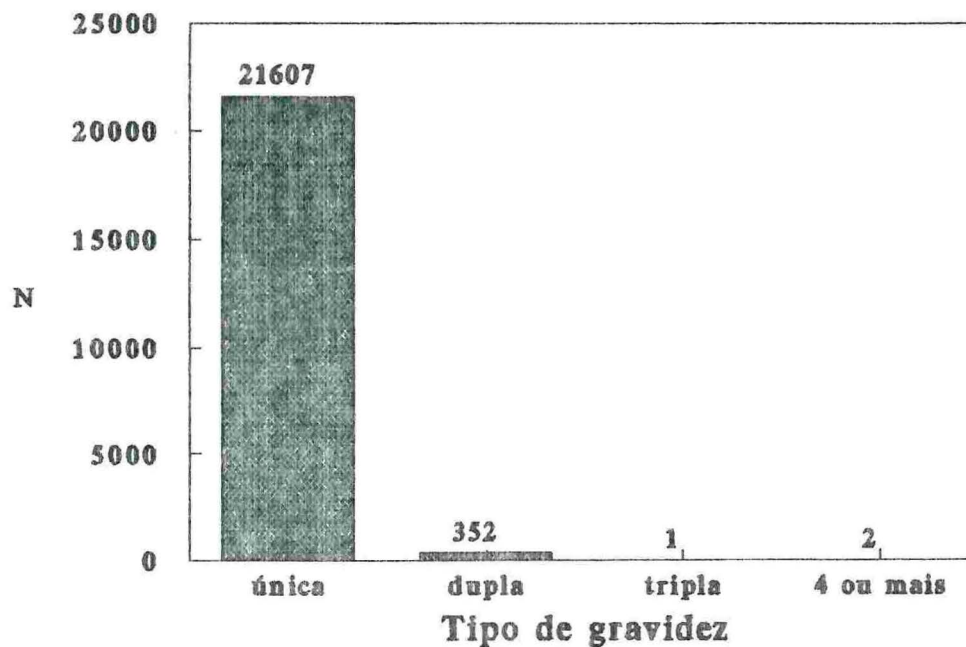


Figura 3: Distribuição dos nascidos vivos segundo o tipo de gravidez, 1992, Porto Alegre, RS.

A distribuição dos nascidos vivos segundo a idade

materna pode ser observada na Figura 4. Para que pudesse ser melhor aproveitada nas análises posteriores, esta variável foi codificada. Esta nova distribuição é apresentada na Tabela 5, onde as mães foram divididas em três faixas etárias.

Cabe salientar o pequeno percentual de mães com menos de 18 anos, quando comparado com os resultados encontrados em estudos semelhantes realizados em outros Estados (aproximadamente 25%) (JORGE et al., 1992).

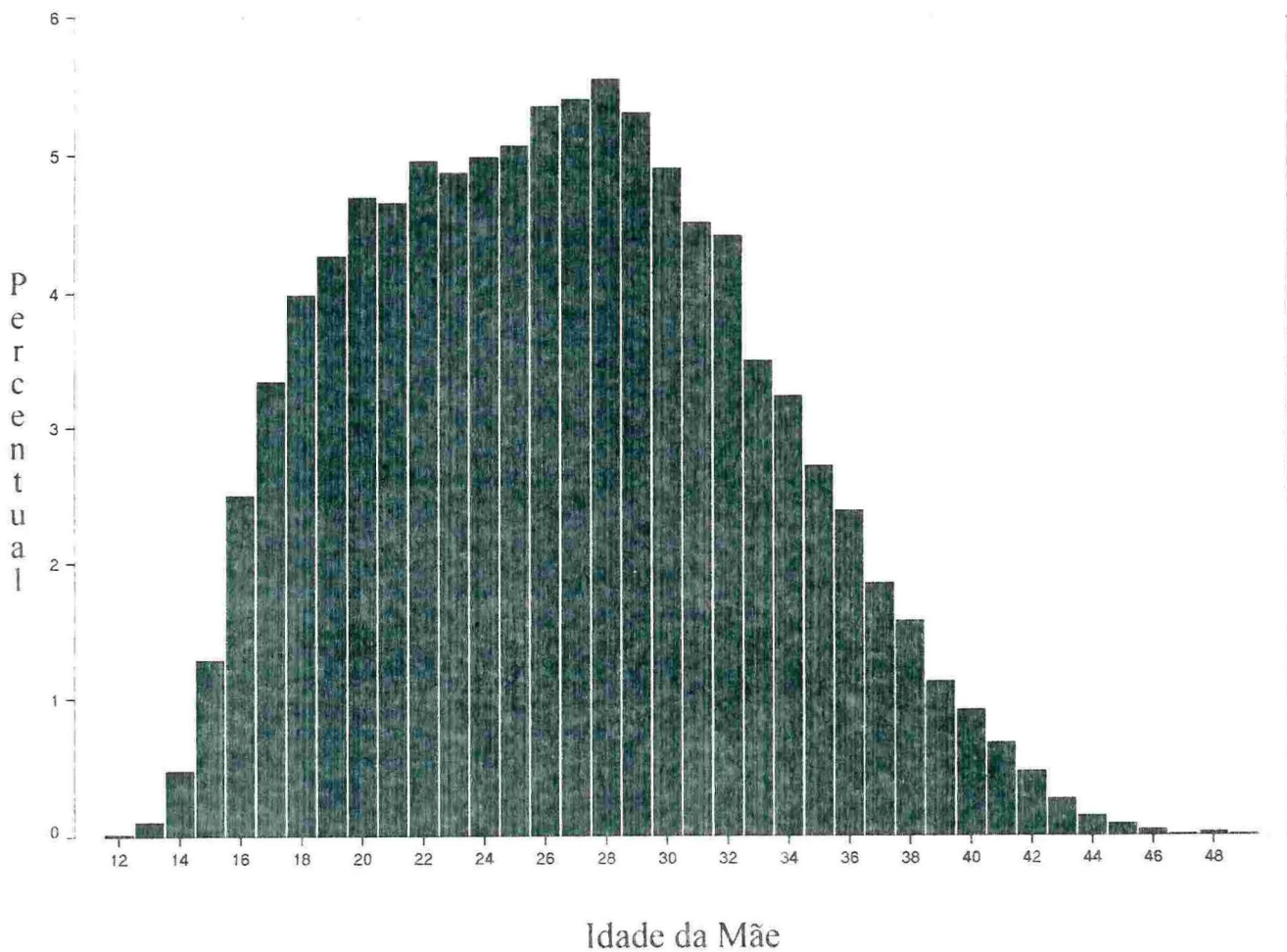


Figura 4: Distribuição dos nascidos vivos segundo a idade materna, Porto Alegre, 1992.

TABELA 5 - DISTRIBUIÇÃO DOS NASCIDOS VIVOS SEGUNDO A IDADE MATERNA, 1992, PORTO ALEGRE, RS.

FAIXA ETARIA	N	%
< 18 anos	1700	7,7
18 a 34 anos	17519	79,7
35 anos ou mais	2714	12,3
Ignorado	47	0,2
TOTAL	21980	100,0

Na Tabela 6, observamos a distribuição dos recém nascidos de acordo com a variável referente ao grau de instrução materna.

TABELA 6 - DISTRIBUIÇÃO DOS NASCIDOS VIVOS SEGUNDO O GRAU DE INSTRUÇÃO DA MÃE, 1992, PORTO ALEGRE, RS.

GRAU DE INSTRUÇÃO	N	%
Nenhum e 1º grau inc.	7839	35,6
1º grau comp. e 2º grau	7516	34,2
Superior	2694	12,3
Ignorado	3931	17,9
Total	21980	100,0

Originalmente, esta variável é apresentada em 6 categorias (incluindo a categoria de ignorados), como mostra a Figura 5. Sua recodificação foi feita mantendo as mães analfabetas e com o 1º grau incompleto num grupo, as com 1º grau

completo e 2º grau em outro e as com nível de instrução superior em um terceiro. Chama a atenção o alto número de DNs com este campo ignorado (17,9%).

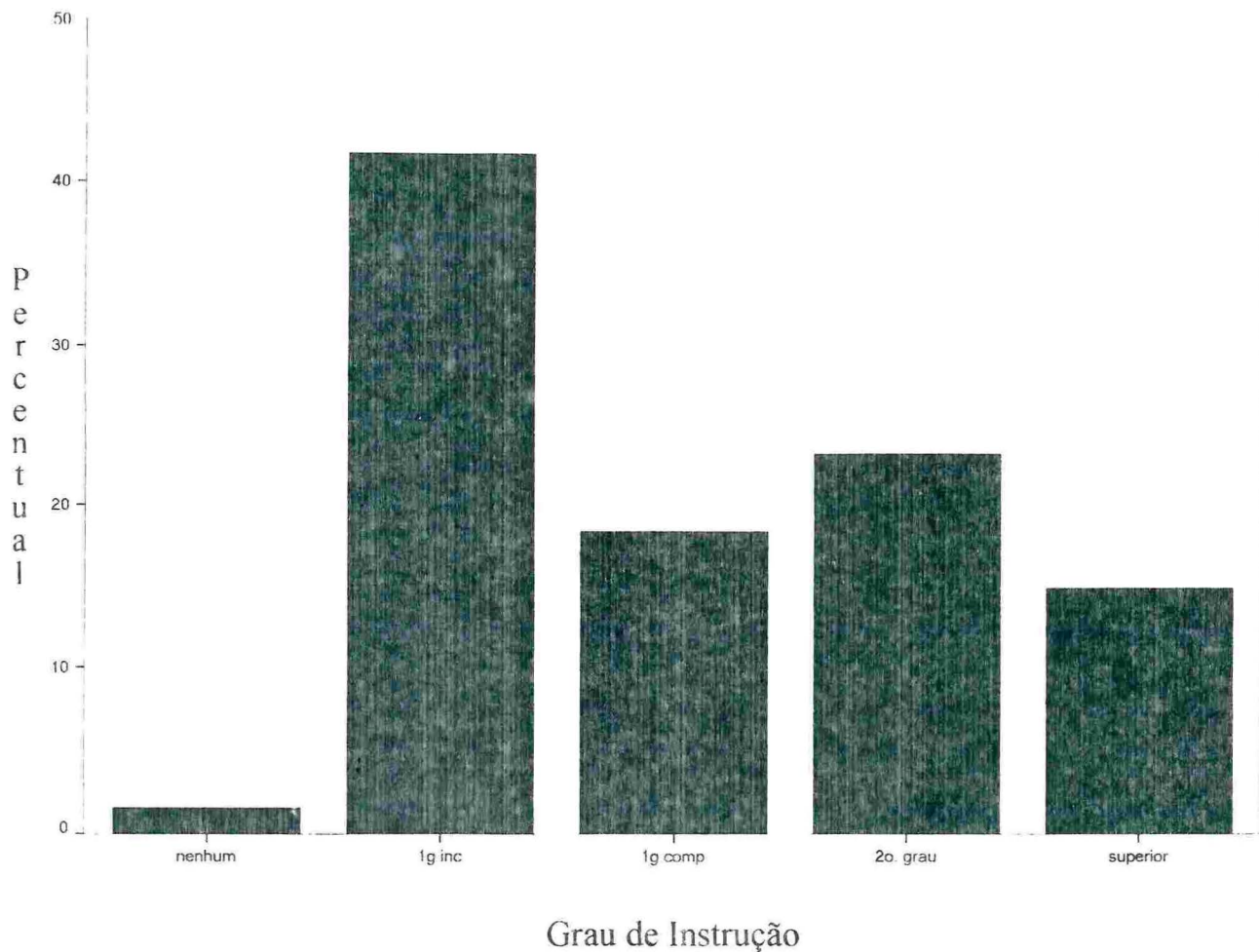


Figura 5: Distribuição dos nascidos vivos segundo o grau de instrução da mãe, Porto Alegre, 1992.

A Tabela 7 apresenta a distribuição dos nascidos vivos de acordo com os filhos vivos de gestações anteriores àquela

constante na DN. A quantidade de filhos tidos vivos superior a 4 foram agrupadas para facilitar as análises posteriores.

Cabe comentar a pequena quantidade de DN's com este campo ignorado, visto que em Rondônia, este valor foi de 38,9% durante todo o ano em que o SINASC foi implantado naquele Estado (JORGE et al., 1992).

TABELA 7 - DISTRIBUIÇÃO DOS NASCIDOS VIVOS SEGUNDO OS FILHOS TIDOS VIVOS, 1992, PORTO ALEGRE, RS.

FILHOS VIVOS	N	%
Nenhum	9110	41,4
Um	6241	28,4
Dois	3212	14,6
Três ou mais	3130	14,2
Ignorado	287	1,3
TOTAL	21980	100,0

A distribuição dos recém nascidos segundo filhos mortos de gestações anteriores é apresentada na Tabela 8. Podemos observar que quase a totalidade (89%) dos nascidos vivos não haviam tido irmãos mortos em gestações anteriores. Nota-se, a exemplo do que ocorreu com filhos tidos vivos, a boa qualidade do preenchimento das DN's, em comparação aos dados de Rondônia, cujos ignorados neste campo foram em torno de 88,8% (JORGE et al., 1992).

TABELA 8 - DISTRIBUIÇÃO DOS NASCIDOS VIVOS SEGUNDO OS FILHOS TIDOS MORTOS, 1992, PORTO ALEGRE, RS.

FILHOS MORTOS	N	%
Nenhum	19584	89,1
Um	1465	6,7
Dois ou mais	500	2,3
Ignorado	431	2,0
TOTAL	21980	100,0

3.2 Análise Bivariada

A relação de cada um dos fatores de risco com a variável dependente, peso ao nascer, analisadas por meio da regressão logística univariada será apresentada agora. Nesta seção, podemos encontrar os principais resultados e, em anexo, as saídas de computador que deram origem a eles.

As regressões logísticas apresentadas neste trabalho que envolviam variáveis com mais de duas categorias foram realizadas utilizando como contraste a opção *indicator* do SPSS, que equivale à utilização de variáveis *dummy*.

Na Tabela 9 observa-se a razão de chances, a significância do Teste de Wald e o intervalo de confiança com 95% de confiança dos fatores de risco em relação ao peso ao nascer.

TABELA 9 - RAZÃO DE CHANCES (RC), SIGNIFICÂNCIA E INTERVALO DE CONFIANÇA (IC) REFERENTES AS REGRESSÕES LOGÍSTICAS UNIVARIADAS DOS FATORES DE RISCO EM RELAÇÃO A VARIÁVEL PESO AO NASCER, 1992, PORTO ALEGRE, RS.

VARIÁVEL	RC	SIG.*	IC (95%)
SEXO			
Masculino	1		
Feminino	1,24	0,0000	(1,13; 1,36)
FILHOS TIDOS MORTOS			
Nenhum	1		
Um	1,28	0,0048	(1,08; 1,52)
Dois ou mais	1,12	0,4767	(0,83; 1,51)
FILHOS TIDOS VIVOS			
Nenhum	1,46	0,0000	(1,29; 1,64)
Um	1		
Dois	1,25	0,0053	(1,07; 1,47)
Três ou mais	1,71	0,0000	(1,47; 1,98)
IDADE GESTACIONAL			
A termo	1		
Prematuro	32,40	0,0000	(28,59;36,71)
IDADE MATERNA			
< 18 anos	1,44	0,0000	(1,23; 1,69)
>= 18 e < 35 anos	1		
>= 35 anos	1,38	0,0000	(1,21; 1,58)
GRAU DE INSTRUÇÃO			
Nenhum e 1º g. inc.	1,68	0,0000	(1,41; 2,00)
1º g. comp. e 2º g.	1,24	0,0197	(1,03; 1,48)
Superior	1		
TIPO DE GRAVIDEZ			
Única	1		
Múltipla	15,72	0,0000	(12,66;19,52)

* Significância do Teste de Wald

Todas as Razões de Chances cujo intervalo de confiança não contenha a unidade devem ser consideradas estatisticamente significativas.

A razão de chances (RC) para sexo, 1,24, é o risco que os recém nascidos do sexo feminino têm de apresentarem baixo peso ao nascer em relação aos do sexo masculino. Em outras palavras, pode-se dizer que a ocorrência de baixo peso ao nascer é 1,24 vezes mais freqüente em recém nascidos do sexo feminino.

A ocorrência de baixo peso ao nascer nos recém nascidos que tinham tido um irmão morto é 1,28 vezes mais freqüente do que naqueles que não o tiveram.

Com relação à variável filhos tidos vivos, todas as outras categorias apresentaram risco em relação à categoria referente aos recém nascidos com um irmão anterior. Isto quer dizer que a presença de apenas um irmão anterior "protege" a criança de ter baixo peso ao nascer.

Para os recém nascidos prematuros a chance de apresentar baixo peso ao nascer é 32,4 vezes maior do que nas crianças nascidas a termo.

Quando se estuda a variável idade materna, pode-se dizer que a ocorrência de baixo peso ao nascer é 1,44 vezes mais freqüente nas mães com idade inferior a 18 anos do que naquelas com idade entre 18 e 34 anos. A chance da mãe com idade superior a 34 anos gerar um recém nascido com baixo peso é 1,38 vezes maior do que aquela com idade entre 18 e 34 anos.

As mães sem grau de instrução ou com apenas o 1º grau

incompleto apresentaram 1,68 vezes mais chance de ter crianças com baixo peso ao nascer do que as mães com nível de instrução superior. Já as mães com 1º grau completo ou 2º grau possuem um risco de 1,24 de apresentarem recém nascidos com baixo peso do que as mães com grau de instrução superior.

Em relação aos nascidos vivos de gravidez múltipla, estes têm 15,72 vezes mais chance de terem baixo peso ao nascer do que os que não têm irmão gêmeo.

3.3 Análise Multivariada

Podemos observar na Tabela 10, os resultados da regressão logística múltipla de uma forma resumida.

O modelo final inclui todas as variáveis, pois não houve nenhum fator de risco que tenha sido completamente não significativo. A única variável que poderia ocasionar alguma dúvida quanto à sua inclusão no modelo é a referente a filhos tidos mortos. Resolvemos incluí-la, pois além de sua importância do ponto de visto epidemiológico, uma de suas Razões de Chances foi estatisticamente significativa.

A interpretação dos coeficientes na análise de regressão logística múltipla é semelhante à da análise univariada. O que deve ser levado em conta é que, neste caso, cada fator de risco está ajustado, estatisticamente, pelas outras variáveis.

TABELA 10 - RAZÃO DE CHANCES (RC), SIGNIFICÂNCIA E INTERVALO DE CONFIANÇA (IC) REFERENTES A REGRESSÃO LOGÍSTICA MÚLTIPLA DOS FATORES DE RISCO EM RELAÇÃO A VARIÁVEL PESO AO NASCER, 1992, PORTO ALEGRE, RS.

VARIÁVEL	RC	SIG.*	IC (95%)
SEXO			
Masculino	1		
Feminino	1,43	0,0000	(1,26; 1,62)
FILHOS TIDOS MORTOS			
Nenhum	1		
Um	1,20	0,1097	(0,96; 1,51)
Dois ou mais	1,22	0,2867	(0,84; 1,77)
FILHOS TIDOS VIVOS			
Nenhum	1,61	0,0000	(1,36; 1,89)
Um	1		
Dois	1,11	0,3459	(0,89; 1,38)
Três ou mais	1,25	0,0456	(1,004; 1,57)
IDADE GESTACIONAL			
A termo	1		
Prematuro	33,00	0,0000	(28,56;38,13)
IDADE MATERNA			
< 18 anos	1,02	0,8839	(0,80; 1,30)
>= 18 e < 35 anos	1		
>= 35 anos	1,33	0,0033	(0,10; 1,62)
GRAU DE INSTRUÇÃO			
Nenhum e 1º g. inc.	1,82	0,0000	(1,45; 2,28)
1º g. comp. e 2º g.	1,49	0,0004	(1,20; 1,85)
Superior	1		
TIPO DE GRAVIDEZ			
Única	1		
Múltipla	12,43	0,0000	(9,16;16,86)

* Significância do Teste de Wald

De um modo geral, os riscos não se alteraram muito da análise anterior para esta. Observa-se apenas, uma mudança maior na Razão de Chances apresentada para as variáveis idade materna, filhos vivos e filhos mortos.

A variável idade da mãe passa a ser não significativa quando há a comparação das mães pertencentes à faixa de menores de 18 anos com a faixa que vai de 18 a 34 anos. Isto é, quando há o ajuste pelas demais variáveis, as mães pertencentes à faixa de menores de 18 anos se igualam àquelas pertencentes à faixa de 18 a 34 anos, em relação ao baixo peso ao nascer. Isto pode ser explicado, talvez, pelo fato de que mães com idade inferior tenham tido, em geral, menos filhos do que as outras mães (principalmente aquelas com 35 anos e mais). Há, também, uma relação de idade materna com grau de instrução, já que mães com menos de 18 anos dificilmente poderiam ter grau de instrução superior.

Com relação aos filhos tidos vivos, pode-se notar que, quando há o ajuste pelas outras variáveis, a categoria que se refere à presença de dois filhos anteriores passa a apresentar Razão de Chance não significativas em relação à categoria de referência (que apresenta menor risco).

Na análise multivariada, a variável filhos tidos mortos, se apresentou não significativa. No anexo, pode ser encontrada um modelo encontrado quando foi realizada uma regressão logística com o método Stepwise Forward (LR) no SPSS, onde esta variável foi excluída.

4 CONCLUSÃO

A regressão logística se mostrou uma técnica de análise estatística bastante adequada para este tipo de estudo.

Os resultados indicam que, de um modo geral, nossas crianças estão nascendo saudáveis. No entanto, deve-se tomar cuidado, principalmente, com aqueles recém nascidos de gravidez múltipla e prematura (fatores que mostraram maiores chances de determinar o baixo peso de um nascido vivo).

Com relação às outras variáveis, apesar de terem apresentado razões de chances relativamente baixas, foram significativas. Por isso, os recém nascidos do sexo feminino, sem nenhum irmão vivo anterior ou mais de 2, cuja mãe possua 35 anos ou mais e grau de instrução inferior (nenhum ou 1º grau incompleto) ou intermediário (1º grau completo e 2º grau), além das características citadas acima, merecem mais cuidados dos órgãos responsáveis pela vigilância à saúde da criança.

Espera-se que os resultados deste tipo de estudo possam auxiliar no combate à mortalidade infantil, melhorando, assim, a qualidade de vida de nossa população.

5 BIBLIOGRAFIA

- AERTS, Denise R.G. de C.. Polígrafo referente ao Sistema de Informação sobre os Nascidos Vivos em Porto Alegre. Secretaria Municipal da Saúde e Serviço Social, 1992.
- DELGADO, H. et al.. Use and Interpretation of Anthropometric indicators of nutritional status. Reviews Analyses, Bulletin of the World Health Organization, 64 (6): 929-941, 1986.
- HOSMER, David W., Jr. & LEMESHOW, Stanley. Applied Logistic Regression. Ed. John Wiley & Sons, Inc., USA, 1989.
- KLEINBAUM, David G., KUPPER, Laurence L. & MULLER, Keith E.. Applied Regression Analysis and Other Multivariable Methods. Ed. PWS-KENT, Boston, 2ª ed., 1987.
- JORGE, Maria Helena P. de M. et al.. O Sistema de Informação sobre nascidos vivos - SINASC. Centro da OMS para a Classificação de Doenças em Português, NEPS/USP, série divulgação nº 7, São Paulo, 1992.
- JORGE, Maria Helena P. de M. et al.. Sistema de Informação sobre nascidos vivos - SINASC. Informe Epidemiológico do SUS, nº 4, Brasília, FNS/CENEPI, 1992.
- MONTEIRO, Carlos A. et al.. Perfil Estatístico de Crianças e Mães

no Brasil: Aspectos de Saúde e Nutrição de Crianças no Brasil, 1989. IBGE, Departamento de Estatísticas e Indicadores Sociais, convênio IBGE/UNICEF/INAN. Rio de Janeiro, 1992

TESCH, Lenira C., MACHADO, Maria Elida e CARVALHO, Rodney F..
Configuração epidemiológica dos nascidos vivos do município de Cachoeira do Sul: um estudo a partir do SINASC. Porto Alegre, Escola de Saúde Pública, Secretaria de Saúde e Meio Ambiente do Rio Grande do Sul/UFRGS, 1993.

6 ANEXOS

6.1 Declaração de Nascido Vivo

6.2 Resultados nas Saídas de Computador

Sexo

Total number of cases: 21980 (Unweighted)
 Number of selected cases: 21980
 Number of unselected cases: 0

Number of selected cases: 21980
 Number rejected because of missing data: 31
 Number of cases included in the analysis: 21949

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Dependent Variable.. PESO

Beginning Block Number 0. Initial Log Likelihood Function

-2 Log Likelihood 13172.175

* Constant is included in the model.

Beginning Block Number 1. Method: Enter

Variable(s) Entered on Step Number

1.. SEXO

Estimation terminated at iteration number 4 because Log Likelihood decreased by less than .01 percent.

-2 Log Likelihood 13152.322
 Goodness of Fit 21948.306

	Chi-Square	df	Significance
Model Chi-Square Improvement	19.853	1	.0000
	19.853	1	.0000

Classification Table for PESO

Observed		Predicted		Percent Correct
		0	1	
0	0	19997	0	100.00%
1	1	1952	0	.00%
		Overall		91.11%

----- Variables in the Equation -----

Variable	B	S.E.	Wald	df	Sig	R	Exp(B)
SEXO	.2115	.0475	19.7860	1	.0000	.0367	1.2355
Constant	-2.4347	.0347	4927.067	1	.0000		

Filhos Tidos Mortos

Total number of cases: 21980 (Unweighted)
 Number of selected cases: 21980
 Number of unselected cases: 0

Number of selected cases: 21980
 Number rejected because of missing data: 444
 Number of cases included in the analysis: 21536

Dependent Variable Encoding:

Original Value	Internal Value
0	0
1	1

	Value	Freq	Parameter Coding	
			(1)	(2)
FILMOPTO	0	19573	.000	.000
	1	1463	1.000	.000
	2	500	.000	1.000

Dependent Variable = PESO

Beginning Block Number 0. Initial Log Likelihood Function

-2 Log Likelihood 12390.46

* Constant is included in the model

Beginning Block Number 1. Method: Enter

Variable(s) Entered on Step Number
 1 FILMOPTO

Estimation terminated at iteration number 4 because
 Log Likelihood decreased by less than .01 percent

-2 Log Likelihood 12882.571
 Goodness of Fit 21535.453

	Chi-Square	df	Significance
Model Chi-Square Improvement	7.889	2	.0194
	7.889	2	.0194

Classification Table for PESO

Observed		Predicted		Percent Correct
		0	1	
0	0	19628	0	100.00%
1	1	1908	0	.00%

Overall 91.14%

Variables in the Equation

Variable	B	S.E.	Wald	df	Sig.	R	Exp(B)
FILMOPTO			8.3920	2	.0157	.0183	
FILMOPTO(1)	.2477	.0877	7.9783	1	.0048	.0215	1.2811
FILMOPTO(2)	1.025	.1539	45.063	1	.0000	.0000	1.1157
Constant	-2.3520	.0254	8592.140	1	.0000		

Filhos Tidos Vivos

Total number of cases: 21980 (Unweighted)
 Number of selected cases: 21980
 Number of unselected cases: 0

Number of selected cases: 21980
 Number rejected because of missing data: 300
 Number of cases included in the analysis: 21680

Dependent Variable Encoding:

Original Value	Internal Value
0	0
1	1

	Value	Freq	Parameter Coding		
			(1)	(2)	(3)
FILVIVOS	0	9108	1,000	,000	,000
	1	6237	,000	,000	,000
	2	3208	,000	1,000	,000
	3	3127	,000	,000	1,000

Dependent Variable: PESO

Beginning Block Number 0 Initial Log Likelihood Function

Log Likelihood = 12987,101

* Constant is included in the model.

Beginning Block Number 1 Method: Enter

Variable(s) Entered on Step Number 1
 1 FILVIVOS

Estimation terminated at iteration number 4 because log likelihood decreased by less than .01 percent

Model Likelihood 12926,554
 Difference of Fit 60530,576

	Chi-Square	df	Significance
Model Chi-Square	60,547	3	,0000
Improvement	60,547	3	,0000

Classification Table for PESO

Observed		Predicted		Percent Correct
		0	1	
0	0	19757	0	100,00%
1	1	1923	0	,00%
Overall				91,13%

Variables in the Equation

Variable	B	S.E.	Wald	df	Sig.	R	Exp(B)
FILVIVOS			58,8019	3	,0000	,9113	
FILVIVOS(1)	,3763	,0615	37,4632	1	,0000	,0133	1,4563
FILVIVOS(2)	,2258	,0810	7,7705	1	,0053	,0311	1,2534
FILVIVOS(3)	,5352	,0758	49,7952	1	,0000	,0607	1,7078
Constant	-2,6128	,0502	2710,166	1	,0000		

Idade Gestacional

Total number of cases 31980 (Unweighted)
 Number of selected cases 31980
 Number of unselected cases 0

Number of selected cases 31960
 Number rejected because of missing data 142
 Number of cases included in the analysis 31838

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Dependent Variable PMSO

Beginning Block Number 0 Initial Log Likelihood Function

-2 Log Likelihood 13063.035

* Constant is included in the model.

Beginning Block Number 1 Method: Enter

Variable(s) Entered on Step Number
 1 GESTACAO

Iteration terminated at iteration number 5 because
 Log Likelihood decreased by less than .01 percent

-2 Log Likelihood 10064.047
 Degrees of Freedom 21837,984

	Chi-Square	df	Significance
Model Chi-Square	3998.988	1	.0000
Improvement	3998.988	1	.0000

Expected	Observed	Predicted		Percent Correct
		0	1	
0	0	19390	515	97.41%
1	1	1039	804	46.25%

Variable	B	S.E.	Wald	df	Sig.	Exp. B	Exp. B1
GESTACAO	1.4781	0.638	2968.907	1	.0000	4.506	32.1161
Constant	-2.9265	.0318	8445.843	1	.0000		

Idade Materna

Total number of cases 21980 (Unweighted)
 Number of selected cases 21980
 Number of unselected cases 0

Number of selected cases 21980
 Number rejected because of missing data 64
 Number of cases included in the analysis 21916

Dependent Variable Encoding

Original Value	Internal Value	Value	Freq	Parameter Coding (1)	Parameter Coding (2)
0	0	1	1649	1.000	0.000
1	1	2	17117	.000	.000
2	2	3	2709	.000	1.000

Dependent Variable: PESO

* Missing Block Number 0. Initial Log Likelihood Function

* Log Likelihood -13114.28

* Constant is included in the model.

Beginning Block Number 1 Method Enter

Variables Entered on Step Number 1 IDADEMAE

Estimation terminated at iteration number 4 because Log Likelihood decreased by less than .01 percent

-2 Log Likelihood 11382.809
 Goodness of Fit 21914.242

	Chi-Square	df	Significance
Model Chi-Square	36.471	2	.0000
Improvement	36.471	2	.0000

Classification Table for PESO

Observed		Predicted		Percent Correct
		0	1	
0	0	19973	0	100.00%
1	1	1942	0	.00%
Overall				91.14%

Variables in the Equation

Variable	B	S.E.	Wald	df	Sig.	R	Exp(B)
IDADEMAE			38.2066	2	.0000	.0511	
IDADEMAE(1)	.3639	.0809	20.2290	1	.0000	.3373	1.4389
IDADEMAE(2)	.3236	.0671	23.2623	1	.0000	.0403	1.3821
Constant	-2.4068	.0274	7689.473	1	.0000		

Grau de Instrução

Total number of cases 31980 (Unweighted)
 Number of selected cases 21980
 Number of unselected cases 0

Number of selected cases: 21980
 Number rejected because of missing data 3943
 Number of cases included in the analysis 18037

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Value	Freq	Parameter Coding	
		(1)	(2)
1	7333	1.000	.000
2	7412	.000	1.000
3	3692	.000	.000

Dependent Variable = PESO

Beginning Block Number 0 Initial Log Likelihood Function

-2 Log Likelihood 10385.243

* Constant is included in the model.

Beginning Block Number 1. Method: Enter

Variable(s) Entered on Step Number

1 INSTRMAE

Estimation terminated at iteration number 4 because Log Likelihood decreased by less than .01 percent

-2 Log Likelihood 10224.943
 Goodness of Fit 18935.184

	Chi-Square	df	Significance
Model Chi-Square	50.299	2	.0000
Improvement	50.299	2	.0000

Contingency Table for PESO

Observed		Predicted		Percent Correct
		0	1	
0	0	18526	0	100.00%
1	1	1511	0	.00%
Overall				91.62%

----- Variables in the Equation -----

Variable	B	S.E.	Wald	df	Sig.	R	Exp(B)
INSTRMAE			49.1975	2	.0000	.0660	
INSTRMAE(1)	.5201	.0886	34.4909	1	.0000	.0559	1.6831
INSTRMAE(2)	.2128	.0913	5.4369	1	.0197	.0132	1.2372
Constant	-2.7220	.0801	1154.514	1	.0000		

Tipo de Gravidez

Total number of cases: 21980 (Unweighted)
 Number of selected cases: 21980
 Number of unselected cases: 0
 Number of selected cases: 21980
 Number rejected because of missing data: 35
 Number of cases included in the analysis: 21947

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Dependent Variable: PESO

Beginning Block Number 0 Initial Log Likelihood Function

Log Likelihood 13167,149

* constant is included in the model.

Beginning Block Number 1. Method: Enter

Variable(s) Entered on Step Number

1 TIPOGRAV

Estimation terminated at iteration number 4 because Log Likelihood decreased by less than .01 percent.

-2 Log Likelihood 12607,290
 Goodness of Fit 21945,380

	Chi-Square	df	Significance
Model Chi-Square	559,859	1	.0000
Improvement	559,859	1	.0000

Classification Table for PESO

Observed		Predicted		Percent Correct
		0	1	
0	0	19817	149	99,25%
1	1	1745	206	10,56%
Overall				91,37%

Variables in the Equation

Variable	B	S.E.	Wald	df	Sig.	F	Exp(B)
TIPOGRAV	2,7552	,1104	622,7597	1	,0000	,2191	15,7237
Constant	-3,4312	,0250	9481,441	1	,0000		

Regressão Logística Múltipla

	Value	Freq	Parameter Coding			
			(1)	(2)	(3)	
FILVIVOS	0	7699	1.000	.000	.000	Total number of cases: 21980 (Unweighted) Number of selected cases: 21980 Number of unselected cases: 0
	1	5169	.000	.000	.000	
	2	2512	.000	1.000	.000	
IDADEMAE	0	2396	.000	.000	1.000	Cases deleted because of missing data: 4082 Number of cases included in the analysis: 2787
	1	7694	1.000	.000	.000	
	2	7386	.000	1.000	.000	
TIPOGRAV	0	2547	.000	.000	.000	
	1	1365	1.000	.000	.000	
	2	8419	.000	.000	.000	
FILMORTO	0	2215	.000	1.000	.000	
	1	15873	.000	.000	.000	
	2	1348	1.000	.000	.000	
	3	458	.000	1.000	.000	

Dependent Variable: PESO

beginning Block Number 0. Initial Log Likelihood function

Log Likelihood = -10130.265

* Constant is included in the model.

beginning Block Number 1. Method: Enter

Variables Entered on Step Number

- 1. SEXO
- FILMORTO
- FILVIVOS
- GESTACAO
- IDADEMAE
- INSTRMAE
- TIPOGRAV

Estimation terminated at iteration number 5 because Log Likelihood decreased by less than .01 percent

Final Log Likelihood = -7365.826
Goodness of Fit = 17183.686

	Chi-Square	df	Significance
Model Chi-Square Improvement	2864.979	12	.0000
	2864.979	12	.0000

Classification Table for PESO

Actual \ Predicted	Predicted		Percent Correct
	0	1	
0	15817	339	97.60%
1	621	650	44.19%
Overall			93.15%

----- Variables in the Equation -----

Variable	B	S.E.	Wald	df	sig.	R	Exp(B)
PESO	.9577	.0656	29.7369	1	.0000	.9521	1.4409
FILMORTO			3.4756	2	.1759	.0023	
FILMORTO(1)	1861	.1163	2.5585	1	1.197	.0074	1.2045
FILMORTO(2)	2013	.1289	1.1350	1	.2867	.0000	1.2220
FILVIVOS			31.2769	2	.0000	.0338	
FILVIVOS(1)	4763	.0628	71.9011	1	.0000	.5431	1.6894
FILVIVOS(2)	1095	.1120	9.8995	1	.0021	.0076	1.1113
FILVIVOS(3)	2089	.1135	8.9958	1	.0028	.0140	1.2547
GESTACAO	4965	.0737	2248.679	1	.0000	.4709	32.9995
IDADEMAE			8.6570	2	.0132	.0014	
IDADEMAE(1)	10181	.1241	6313	1	.0000	.0000	1.0183
IDADEMAE(2)	12885	.0981	8.6544	1	.0031	.0258	1.3344
INSTRMAE			27.9221	2	.0000	.0496	
INSTRMAE(1)	5981	.1148	27.1299	1	.0000	.0498	1.8186
INSTRMAE(2)	13974	.1115	12.7112	1	.0004	.0325	1.4880
TIPOGRAV	2.5198	.1656	262.1430	1	.0000	1.001	12.4258
Constant	-4.0919	.1263	1050.354	1	.0000		

Regressão Logística Múltipla (Método Stepwise Forward - LR)

	Value	Freq	Parameter Coding		
			(1)	(2)	(3)
FILVIVOS					
	0	7699	1,000	0,000	0,000
	1	5169	0,000	1,000	1,000
	2	3513	0,000	1,000	0,000
	3	2296	0,000	0,000	1,000
INSTRMAE					
	1	7654	1,000	0,000	
	2	1386	0,000	1,000	
	3	2677	0,000	0,000	
IDADEMAM					
	1	1265	1,000	0,000	
	2	14199	0,000	0,000	
	3	2313	0,000	1,000	
FILMORTO					
	1	15273	0,000	0,000	
	2	1346	1,000	0,000	
	3	458	0,000	1,000	

Dependent Variable: PESO

Beginning Block Number: 0 Initial Log Likelihood Maximum

Final Log Likelihood: 101,00805

* Constant is included in the model.

Estimation terminated at iteration number 4 because Log Likelihood decreased by less than .01 percent

Classification Table for PESO

Observed \ Predicted	Predicted		Percent Correct
	0	1	
Observed 0	0	0	100,00%
Observed 1	0	0	0,00%
Overall			91,80%

----- Variables in the Equation -----

Variable	B	S.E.	Wald	df	Sig.	F	Exp(B)
Constant	-2,3194	,0372	7764,138	1	,000		

Beginning Block Number: 1 Method: Forward Stepwise (LR)

----- Variables not in the Equation -----
Residual Chi Square: 5,827545 with 12 df Sig.: ,0000

Variable	Score	df	Sig.	F
SEXO	20,1512	1	,0000	,9423
FILMORTO	11,6915	3	,0026	,6279
FILMORTO(1)	10,1309	1	,0015	,0257
FILMORTO(2)	1,3929	1	,2379	,0000
FILVIVOS	44,9472	3	,0000	,0620
FILVIVOS(1)	9,9011	1	,0017	,0379
FILVIVOS(2)	,7532	1	,3855	,0000
FILVIVOS(3)	17,0173	1	,0000	,0325
GESTACAO	4796,3225	1	,0000	,6879
IDADEMAE	32,9429	3	,0000	,0635
IDADEMAE(1)	12,6953	1	,0004	,0325
IDADEMAE(2)	16,8152	1	,0000	,0382
INSTRMAE	46,1412	2	,0000	,0645
INSTRMAE(1)	41,3719	1	,0000	,0623
INSTRMAE(2)	10,8496	1	,0010	,0296
TIPCGPAV	911,0432	1	,0000	2946

Variable(s) Entered on Step Number
1 GESTACAO

Estimation terminated at iteration number 5 because
Log Likelihood decreased by less than .01 percent

2 Log Likelihood 7594.218
Goodness of Fit 2536.592

	Chi-Square	df	Significance
Model Chi-Square	2536.592	1	.0000
Improvement	2536.592	1	.0000

Contingency Table for FREQ

		Predicted		Percent Correct
		0	1	
Observed	0	400	417	97.24%
	1	79	222	40.00%
Overall				93.30%

----- Variables in the Equation -----

Variable	B	S.E.	Wald	df	Sig.	P	Exp(B)
GESTACAO	3.5568	.0709	2518.749	1	.0000	.4984	35.0161
Constant	-3.0612	.0377	6606.434	1	.0000		

----- Model if Term Removed -----

Term Removed	Log Likelihood	-2 Log LP	df	Significance of Log LP
GESTACAO	-5065.402	2536.592	1	.0000

----- Variables not in the Equation -----
Excluded Chi-Square 342.096 with 11 df Sig. .0000

Variable	Score	df	Sig.	P
SEMO	30.7907	1	.0000	.0503
FILMORTO	3.5871	2	.1839	.0000
FILMORTO(1)	3.5084	1	.1132	.0071
FILMORTO(2)	.7340	1	.3916	.0000
FILVIVOS	28.4874	3	.0000	.0471
FILVIVOS(1)	12.2452	1	.0005	.0318
FILVIVOS(2)	1.2833	1	.2573	.0000
FILVIVOS(3)	5.1544	1	.0232	.0176
IDADEMAE	12.5758	2	.0019	.0291
IDADEMAE(1)	3.1789	1	.0746	.0108
IDADEMAE(2)	8.0582	1	.0045	.0345
INSTRMAE	12.3791	2	.0000	.0426
INSTRMAE(1)	15.4418	1	.0001	.0364
INSTRMAE(2)	1.3903	1	.2337	.0000
TIOGPAY	240.4668	1	.0000	1534

Variable(s) Entered on Step Number
2. TIPOGRAV

Estimation terminated at iteration number 5 because
Log Likelihood decreased by less than .01 percent.

-2 Log Likelihood 7370.218
Goodness of Fit 17655.979

	Chi-Square	df	Significance
Model Chi-Square	2760.587	2	.0000
Improvement	223.994	1	.0000

Classification Table for PESO

Observed	Predicted	Percent Correct		
		0	1	
0	0	15759	447	97.24%
1	1	738	733	49.83%
Overall				93.30%

----- Variables in the Equation -----

Variable	B	S.E.	Wald	df	Sig	P	Exp(B)
GESTACAO	3.4908	.0725	2320.938	1	.0000	.4784	32.8109
TIPOGRAV	2.4590	.1538	255.6330	1	.0000	.1582	11.6930
Constant	-3.1362	.0390	6456.655	1	.0000		

----- Model if Term Removed -----

Term Removed	Log Likelihood	-2 Log LR	df	Significance of Log LR
GESTACAO	-4835.812	2301.407	1	.0000
TIPOGRAV	-4897.106	223.994	1	.0000

----- Variables not in the Equation -----

Variable	Score	df	Sig	F
SEXO	30.1845	1	.0000	0427
FILMORTO	4.3460	2	.1138	6058
FILMORTO(1)	3.0345	1	.0815	0101
FILMORTO(2)	1.1161	1	.2938	0000
FILMIVOS	32.8265	3	.0000	0515
FILMIVOS(1)	17.9941	1	.0000	0207
FILMIVOS(2)	1.5922	1	.2070	0000
FILMIVOS(3)	3.2518	1	.0713	0111
IDADENAE	11.0148	2	.0041	0058
IDADENAE(1)	5.0607	1	.0345	0174
IDADENAE(2)	4.6842	1	.0304	0163
INSTRMAE	25.9688	2	.0000	0460
INSTRMAE(1)	18.8281	1	.0000	0383
INSTRMAE(2)	1.1354	1	.2866	.0000

Variable(s) Entered on step Number
SENG

Convergence terminated at iteration number 5 because Log Likelihood decreased by less than .01 percent

Log Likelihood = -1527.976
 Total number of cases = 17547 cases
 Chi-Square = 29.9626
 df = 1
 Sig. = .0000

Classification Table for FREQ
 Observed
 Predicted

Observed \ Predicted	0	1	Total
0	15778	447	16225
1	153	701	854
Total	15931	1148	17079

Overall 93.30%

----- Variables in the Equation -----

Variable	B	S.E.	Wald	df	Sig.	F	Exp(B)
SENG	.0095	.0053	29.9626	1	.0000	09.1	1.4298
GESTACAC	3.5118	.0730	2313.607	1	.0000	47.7	33.5074
TIPOGRAV	3.1621	.1543	254.4709	1	.0000	15.7	11.7293
Constant	-31.250	.0539	3804.759	1	.0000		

Model if Term Removed -----

Term	Log Likelihood	-2 Log LR	df	Significance of Log LR
SENG	-1527.976	30.246	1	.0000
GESTACAC	-462.794	2711.217	1	.0000
TIPOGRAV	-3761.673	223.385	1	.0000

----- Variables not in the Equation -----
 Score on Chi-Square = 73.403 with 11 df, Sig. = .0000

Variable	Score	df	Sig.	F
FILMOFTO	4.7155	2	.0946	.0084
FILMOFTO(1)	3.3897	1	.0656	.0117
FILMOFTO(2)	1.1197	1	.2900	.0000
FILVIVOS	12.7329	3	.0000	.0511
FILVIVOS(1)	18.1253	1	.0000	.0399
FILVIVOS(2)	1.5456	1	.2138	.0000
FILVIVOS(3)	3.0991	1	.0783	.0104
IDADEMAE	10.2906	2	.0060	.0248
IDADEMAE(1)	9.8421	1	.0078	.0187
IDADEMAE(2)	4.1091	1	.0402	.0149
INSTRMAE	36.0842	2	.0000	.0467
INSTRMAE(1)	16.5693	1	.0000	.0379
INSTRMAE(2)	1.0009	1	.3171	.0000

Model entered on Step Number
4 FILVIVOS

Estimation terminated at iteration number 8 because
log likelihood decreased by less than .01 percent

2 Best Log Likelihood	2306.454		
Standard Error of Fit	17396.115		
	Chi-Square	df	Significance
Model Chi-Square	2304.250	8	.0000
Improvement	222.416	2	.0000

Classification Table for PESO

Observed	Predicted		Percent Correct
	0	1	
0	10219	587	94.53%
1	815	655	44.34%
Overall			54.23%

----- Variables in the Equation -----

Variable	B	S.E.	Wald	df	Sig.	F	Exp(B)
SEX	.3576	.0654	29.361	1	.0000	25.4	1.4240
FILVIVOS			32.474	2	.0000	25.11	
FILVIVOS(1)	4362	0817	28.5314	1	.0000	25.11	1.5468
FILVIVOS(2)	1841	1108	2.7582	1	.0968	2.087	1.2021
FILVIVOS(3)	4229	1072	15.5575	1	.0001	12.366	1.5264
GESTACAO	3.5089	.0733	2289.801	1	.0000	1475.2	33.4139
TIPOGRAV	2.4947	1549	259.3383	1	.0000	159.4	12.1184
Constant	-3.6119	.0802	2029.904	1	.0000		

----- Model if Term Removed -----

Term Removed	Log Likelihood	-2 Log LR	df	Significance of Log LR
SEX	-3868.352	30.149	1	.0000
FILVIVOS	-3619.985	33.416	2	.0000
GESTACAO	-4801.436	2296.318	1	.0000
TIPOGRAV	-5167.393	238.232	1	.0000

----- Variables not in the Equation -----
Residual Chi Square 39.883 with 6 df Sig. > .0000

Variable	Score	df	Sig.	F
FILVIVOS	5.3870	2	.076	10.112
FILVIVOS(1)	2.7939	1	.114	1.910
FILVIVOS(2)	1.3138	1	.2517	.8000
IDADEMAE	6.6426	2	.0301	11.62
IDADEMAE(1)	1.7306	1	.1881	.8000
IDADEMAE(2)	4.4908	1	.0341	10.157
INSTRMAE	27.2599	3	.0000	194.79
INSTRMAE(1)	16.6716	1	.0000	338.1
INSTRMAE(2)	7.451	1	.0000	26.96

Após as análises, seis causas gerais se apresentam como fatores decisivos para a evasão do aluno da UFRGS. Elas atuam em grupos diferenciados de estudantes, segundo as características inerentes ao corpo discente evadido.

Concluimos que as causas gerais "Curso Incompatível com o Mercado de Trabalho" e "Falta de Aptidão para o Curso" foram as causas que contribuíram para a evasão de alunos jovens, solteiros e do sexo feminino. Uma diferenciação constatada entre estas duas causas gerais, está em torno da modalidade de evasão, ou seja, os estudantes ligados à causa incompatibilidade com o mercado de trabalho se evadem no final do curso; enquanto os relacionados à segunda, abandonam o curso na metade do período de duração do mesmo. À elas se deve o fato de um mercado de trabalho saturado; do curso não estar voltado para o mercado de trabalho e de não proporcionar um trabalho bem remunerado; por outro lado, o desconhecimento prévio a respeito do curso e a falta de aptidão para o mesmo, além de uma indecisão quanto à escolha da profissão, geraram o descontentamento do estudante em relação ao curso e sua conseqüente evasão.

Conseguimos verificar também que a "Escolha de Outro Curso" é preponderante para a evasão dos estudantes solteiros, do sexo feminino com idade entre 17 e 27 anos, que ingressaram via

vestibular e que abandonaram o curso entre o primeiro e o terceiro semestre. Verificamos que esta causa geral foi constituída por alguns pontos importantísimos apontados pelos alunos evadidos, como: entraram na UFRGS em curso de 2a opção e desejarem experimentar um novo curso superior; além de terem sido aprovados em novo vestibular. Mediante esta situação, tudo nos leva a crer que estas pessoas, pelo simples fato de terem passado no concorrido vestibular da UFRGS, entraram no curso, mesmo sendo de segunda opção, procurando experimentá-lo. Como viram que o curso não era condizente com as suas expectativas, decidiram evadir-se no início do curso.

As "Pressões Familiares" se evidenciaram na decisão de evasão dos estudantes mais velhos, do sexo feminino; não solteiros e que ingressaram na UFRGS através de algum tipo de transferência ou outra modalidade de ingresso. Para estes alunos, as responsabilidades familiares estão intrínsecas no seu perfil e problemas referentes à família se evidenciam através do relacionamento destes com os filhos, com o cônjuge e com a própria casa. Constatamos que, para as pessoas entre 28 e 37 anos e do sexo feminino, estes problemas são preponderantes na decisão de evasão, possivelmente devido ao fato de que as mulheres levam uma vida muito ativa, pois além de trabalhar fora, muitas vezes têm que cuidar de filhos, marido e casa; tornando inviável a

continuidade de um curso superior.

Outra causa ligada ao aluno não solteiro é a "Necessidade de Trabalhar". Ela está intimamente ligada a decisão de evasão do estudante mais velho, cuja forma de ingresso é via diplomado ou algum tipo de transferência ou outra modalidade de ingresso. Para este tipo de aluno a colisão de horários entre o curso e atividade profissional e o não oferecimento de disciplinas à noite são fatores que influenciaram na sua evasão. Estes estudantes já devem ter uma atividade profissional definida, da qual dependam para sustentar a família. Portanto, entre abandonar o curso ou atividade profissional, este tipo de aluno optou pelo abandono de curso.

Quanto à causa geral "Precariedade Material e Docente do Curso" está associada à evasão do aluno que estava frequentando o curso do quarto semestre em diante. Para estes alunos a pouca eficiência de professores, a precariedade de aparelhagem e de material disponível, assim como a dificuldade de relacionamento com professores, geraram a decepção destes com o curso, culminando em sua evasão.

BIBLIOGRAFIA:

AGRESTI, A. (1990) *Categorical Data Analysis*. New York, Wiley.

AGUAYO, N. T. V. (1993) *Análise de Correspondência e Modelos Log-Lineares: Um Enfoque Integrado para a Análise Exploratória de Dados Categóricos*. Dissertação de Mestrado. Departamento de Estatística - Instituto de Matemática, Estatística e Ciência da Computação - Universidade Estadual de Campinas.

CALLEGARI-JACQUES, S. M. (1991) *Análise de Correspondência: Aplicações em Genética*. *Cadernos de Matemática e Estatística*, Série F, nº 2. Porto Alegre, Instituto de Matemática - UFRGS.

DPI-PROPLAN (1991) *Evasão dos Cursos de Graduação da UFRGS em 1985, 1986 e 1987*. Série Estudos e Projetos, nº 15. Porto Alegre, Departamento de Pesquisa Institucional - Pró-Reitoria de Planejamento - UFRGS.

FURTADO, F. J. F. (1992) *Coefficiente de Fidedignidade e Análise de Itens*. Monografia, Bacharelado em Estatística, Instituto de Matemática - UFRGS.

GREENACRE, M. J. (1984) *Theory and applications of correspondence analysis*. Academic Press, New York.

HULL, C. H. & NIE, N. H. (1981) SPSS UPDATE 7 - 9: New Procedures and Facilities for Releases 7 - 9: 248 - 257.

SILVA, N. V. (1990) Introdução à Análise de Dados Qualitativos. São Paulo: Vértice, Editora Revista dos Tribunais, 1990 (Vértice Universitário: 8).

POSSOLI, S. (1992). Análise Multivariada, Cadernos de Matemática e Estatística, Série B: Trabalho de Apoio Didático, Porto Alegre, UFRGS.

SOUZA, N. M. (1990) Análise de Correspondência. Monografia, Bacharelado em Estatística, Instituto de Matemática - UFRGS.

ANEXOS

ANEXO 2:

RESULTADOS FINAIS DA TÉCNICA DE COMPONENTES PRINCIPAIS

Variable	Communality	*	Factor	Character. Eigenvalue	Pct of Var.	Cum Pct
CAU1	.51740	*	1	4.63118	14.5	14.5
CAU2	.61264	*	2	2.92457	9.1	23.6
CAU3	.59193	*	3	2.20696	6.9	30.5
CAU4	.60363	*	4	1.80212	5.6	36.1
CAU5	.58883	*	5	1.43067	4.5	40.6
CAU6	.49914	*	6	1.35169	4.2	44.8
CAU7	.73183	*	7	1.29852	4.1	48.9
CAU8	.53373	*	8	1.22449	3.8	52.7
CAU9	.73742	*	9	1.03460	3.2	56.0
CAU10	.66255	*				
CAU11	.64130	*				
CAU12	.52884	*				
CAU13	.52538	*				
CAU14	.54743	*				
CAU15	.36303	*				
CAU16	.67625	*				
CAU17	.31492	*				
CAU18	.45572	*				
CAU19	.52861	*				
CAU20	.49651	*				
CAU21	.50675	*				
CAU22	.52234	*				
CAU23	.40771	*				
CAU24	.51620	*				
CAU25	.46376	*				
CAU26	.61548	*				
CAU27	.73372	*				
CAU28	.78616	*				
CAU29	.33500	*				
CAU30	.61410	*				
CAU31	.66122	*				
CAU32	.58523	*				

Varimax Rotation 1, Extraction 1, Analysis 1 - Kaiser Normalization

Varimax converged in 14 iterations.

Rotated Factor Matrix:

- - - - FACTOR ANALYSIS - - - -

	PREC	MERCT	FAAPCUR	NETRAB	ESOUCUR
CAU1	.17978	.06583	.63858	-.04480	.09373
CAU2	.00652	.16425	.63176	-.10430	.24085
CAU3	.00542	.10687	.22739	-.09434	.11007
CAU4	.06778	.00101	.71180	.14795	-.08764
CAU5	.08495	.18767	.71702	-.13791	.06580
CAU6	.27796	.61736	.00069	-.04755	.16794
CAU7	-.06562	-.01195	-.18413	.78836	-.19318
CAU8	.02910	.12407	.05540	.53690	-.33093
CAU9	.09477	.82722	.11726	.11429	.05755
CAU10	.04430	.79033	.15031	.04240	-.01208
CAU11	-.07101	-.05357	-.06683	-.02240	-.13096
CAU12	.08075	.12373	.31766	-.08042	.52674
CAU13	.01984	.03554	-.01088	.02209	.09271
CAU14	.38685	.61765	.08405	.05700	-.02724
CAU15	.00730	.25795	.17370	.27417	.11688
CAU16	.71832	.21407	.23908	-.12206	.04629
CAU17	.08772	.04740	.10922	.35572	-.01492
CAU18	.60402	-.02667	.21510	.03450	-.01829
CAU19	.70514	.13895	-.05904	.04687	.06622
CAU20	.08206	.04561	.04268	.51837	.06850
CAU21	-.08782	.03191	.06139	-.17674	-.08307
CAU22	-.06278	-.00523	-.01663	.12384	.02489
CAU23	-.09167	.24744	-.01998	-.04375	.51652
CAU24	.00828	-.06761	-.04519	.02410	.58427
CAU25	-.04250	.11808	-.27777	-.03317	-.29406
CAU26	.08199	.04336	-.06459	.76033	.05934
CAU27	.83231	.17669	.00395	.06509	-.05956
CAU28	.86722	.13685	.07085	.03527	-.07073
CAU29	.37242	-.02447	-.01745	.12010	-.07084
CAU30	.03361	-.00046	-.01894	-.05951	.09837
CAU31	-.01281	.02060	.00645	.10354	-.12162
CAU32	-.03949	.02483	.14861	-.09638	.70617

- - - - FACTOR ANALYSIS - - - -

	PRESFAM	QUINPES	NCOMEP	MUDANC
CAU1	-.02647	.14892	-.12548	-.15347
CAU2	.01081	.33820	-.02739	-.04866
CAU3	.17451	.68330	-.07299	.07121
CAU4	.14091	-.18466	.08309	.04414
CAU5	-.07760	.03193	.02730	.03390
CAU6	.07708	.04988	-.03338	-.02700
CAU7	-.01882	.13477	.12404	-.02732
CAU8	.26186	.12764	.15093	-.09483
CAU9	.06110	.07299	-.06912	-.01396
CAU10	-.00688	.10607	.01034	-.00509
CAU11	.21170	.09051	-.19432	.72146
CAU12	.17487	.28100	-.09989	-.05161
CAU13	.71327	.04145	.01949	.06058
CAU14	-.00545	.01581	.06751	-.01990
CAU15	-.06665	.39198	.12179	-.06718
CAU16	-.01604	.15020	-.03859	-.12621
CAU17	.33143	-.20248	-.03247	.12000
CAU18	.03388	-.20200	.00642	.01950
CAU19	.00626	-.01417	-.00363	.04199
CAU20	.25507	-.17283	-.33972	.04620
CAU21	.37608	-.34453	.37799	-.23044
CAU22	.09427	-.02911	.70118	.02737
CAU23	-.01686	-.22163	.09429	.10335
CAU24	-.15183	.17937	.33321	.03618
CAU25	-.15304	.45820	.21423	.06348
CAU26	-.07232	-.05663	.09569	-.05910
CAU27	-.00285	.03245	-.02998	-.00170
CAU28	-.01979	.01206	-.04673	-.03706
CAU29	.08459	.14873	.36990	.09928
CAU30	-.11316	-.02022	.22004	.73332
CAU31	.78962	.05974	.08807	-.01601
CAU32	.00168	.03640	-.20501	-.09824

ANEXO 3:

COEFICIENTE DE FIDEDIGNIDADE DAS VARIÁVEIS INDICADORAS

RELIABILITY /VARIABLES CAU16 CAU18 CAU19 CAU27 CAU28 /SCALE (PREC) CAU16 CAU18
CAU19 CAU27 CAU28 /MODEL ALPHA /SUMMARY ALL /STATISTICS CORRELATION SCALE
DESCRIPTIVES.

***** METHOD 2 (COVARIANCE MATRIX) WILL BE USED FOR THIS ANALYSIS *****

***** 488 BYTES OF SPACE REQUIRED FOR RELIABILITY *****

RELIABILITY ANALYSIS - SCALE (PREC)

- 1. CAU16
- 2. CAU18
- 3. CAU19
- 4. CAU27
- 5. CAU28

		MEAN	STD DEV	CASES
1.	CAU16	1.0456	1.1744	1205.0
2.	CAU18	.3245	.7438	1205.0
3.	CAU19	.5469	.8848	1205.0
4.	CAU27	.9320	1.0892	1205.0
5.	CAU28	.8946	1.0743	1205.0

CORRELATION MATRIX

	CAU16	CAU18	CAU19	CAU27	CAU28
CAU16	1.0000				
CAU18	.3662	1.0000			
CAU19	.4212	.3056	1.0000		
CAU27	.5654	.3533	.5523	1.0000	
CAU28	.6095	.4970	.5133	.7817	1.0000

RELIABILITY ANALYSIS - SCALE (PREC)

OF CASES = 1205.0

STATISTICS FOR SCALE	MEAN	VARIANCE	STD DEV	VARIABLES		
	3.7436	15.1942	3.8980	5		
ITEM MEANS	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.7487	.3245	1.0456	.7212	3.2225	.0909
ITEM VARIANCES	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	1.0112	.5533	1.3791	.8259	2.4927	.1121
INTER-ITEM COVARIANCES	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.5069	.2011	.9147	.7136	4.5479	.0502
INTER-ITEM CORRELATIONS	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.4965	.3056	.7817	.4761	2.5579	.0191

ITEM-TOTAL STATISTICS

	SCALE MEAN IF ITEM DELETED	SCALE VARIANCE IF ITEM DELETED	CORRECTED ITEM-TOTAL CORRELATION	SQUARED MULTIPLE CORRELATION	ALPHA IF ITEM DELETED
CAU16	2.6979	9.3157	.6276	.4054	.8071
CAU18	3.4191	12.2320	.4630	.2632	.8425
CAU19	3.1967	11.0933	.5629	.3352	.8198
CAU27	2.8116	9.0949	.7478	.6529	.7661
CAU28	2.8490	8.9024	.8014	.6931	.7490

RELIABILITY COEFFICIENTS 5 ITEMS

ALPHA = .8341

STANDARDIZED ITEM ALPHA = .8314

RELIABILITY /VARIABLES CAU6 CAU9 CAU10 CAU14 /SCALE (MERC) CAU6 CAU9
CAU10 CAU14 /MODEL ALPHA /SUMMARY ALL /STATISTICS CORRELATION SCALE
DESCRIPTIVES.

***** METHOD 2 (COVARIANCE MATRIX) WILL BE USED FOR THIS ANALYSIS *****

***** 344 BYTES OF SPACE REQUIRED FOR RELIABILITY *****

RELIABILITY ANALYSIS - SCALE (MERC)

- 1. CAU6
- 2. CAU9
- 3. CAU10
- 4. CAU14

		MEAN	STD DEV	CASES
1.	CAU6	.5395	.8848	1214.0
2.	CAU9	.8509	1.0596	1214.0
3.	CAU10	.6203	.9116	1214.0
4.	CAU14	.6507	.9946	1214.0

CORRELATION MATRIX

	CAU6	CAU9	CAU10	CAU14
CAU6	1.0000			
CAU9	.4358	1.0000		
CAU10	.3309	.6472	1.0000	
CAU14	.3989	.4817	.3992	1.0000

RELIABILITY ANALYSIS - SCALE (MERC)

OF CASES = 1214.0

STATISTICS FOR SCALE	MEAN	VARIANCE	STD DEV	VARIABLES		
	2.6614	8.7682	2.9611	4		
ITEM MEANS	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.6654	.5395	.8509	.3114	1.5771	.0175
ITEM VARIANCES	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.9315	.7829	1.1228	.3400	1.4343	.0241
INTER-ITEM COVARIANCES	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.4202	.2669	.6251	.3582	2.3423	.0148
INTER-ITEM CORRELATIONS	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.4489	.3309	.6472	.3163	1.9558	.0108

ITEM-TOTAL STATISTICS

	SCALE MEAN IF ITEM DELETED	SCALE VARIANCE IF ITEM DELETED	CORRECTED ITEM-TOTAL CORRELATION	SQUARED MULTIPLE CORRELATION	ALPHA IF ITEM DELETED
CAU6	2.1219	5.9324	.4763	.2377	.7559
CAU9	1.8105	4.5626	.6810	.5064	.6442
CAU10	2.0412	5.4295	.5903	.4298	.7002
CAU14	2.0107	5.3379	.5312	.2861	.7310

RELIABILITY COEFFICIENTS 4 ITEMS

ALPHA = .7668

STANDARDIZED ITEM ALPHA = .7652

RELIABILITY /VARIABLES CAU1 CAU2 CAU4 CAU5 /SCALE (FAAPCUR) CAU1 CAU2
 CAU4 CAU5 /MODEL ALPHA /SUMMARY ALL /STATISTICS CORRELATION SCALE
 DESCRIPTIVES.

***** METHOD 2 (COVARIANCE MATRIX) WILL BE USED FOR THIS ANALYSIS *****

***** 344 BYTES OF SPACE REQUIRED FOR RELIABILITY *****

RELIABILITY ANALYSIS - SCALE (FAAPCUR)

- 1. CAU1
- 2. CAU2
- 3. CAU4
- 4. CAU5

		MEAN	STD DEV	CASES
1.	CAU1	.4554	.8192	1212.0
2.	CAU2	.7195	1.0371	1212.0
3.	CAU4	.3515	.7417	1212.0
4.	CAU5	.4117	.8075	1212.0

CORRELATION MATRIX

	CAU1	CAU2	CAU4	CAU5
CAU1	1.0000			
CAU2	.4528	1.0000		
CAU4	.2840	.2356	1.0000	
CAU5	.3854	.4368	.3579	1.0000

RELIABILITY ANALYSIS - SCALE (FAAPCUR)

OF CASES = 1212.0

STATISTICS FOR SCALE	MEAN	VARIANCE	STD DEV	VARIABLES		
	1.9381	6.0961	2.4690	4		
ITEM MEANS	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.4845	.3515	.7195	.3680	2.0469	.0263
ITEM VARIANCES	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.7372	.5502	1.0757	.5255	1.9551	.0537
INTER-ITEM COVARIANCES	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.2623	.1726	.3847	.2121	2.2292	.0078
INTER-ITEM CORRELATIONS	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.3588	.2356	.4528	.2172	1.9217	.0066

ITEM-TOTAL STATISTICS

	SCALE MEAN IF ITEM DELETED	SCALE VARIANCE IF ITEM DELETED	CORRECTED ITEM-TOTAL CORRELATION	SQUARED MULTIPLE CORRELATION	ALPHA IF ITEM DELETED
CAU1	1.4827	3.8007	.5086	.2647	.6010
CAU2	1.2186	3.1569	.5056	.2869	.6100
CAU4	1.5866	4.4095	.3648	.1545	.6840
CAU5	1.5264	3.7739	.5324	.2853	.5871

RELIABILITY COEFFICIENTS 4 ITEMS

ALPHA = .6884

STANDARDIZED ITEM ALPHA = .6912

RELIABILITY /VARIABLES CAU7 CAU8 CAU20 CAU26 /SCALE (NETRAB) CAU7 CAU8
CAU20 CAU26 /MODEL ALPHA /SUMMARY ALL /STATISTICS CORRELATION SCALE
DESCRIPTIVES.

***** METHOD 2 (COVARIANCE MATRIX) WILL BE USED FOR THIS ANALYSIS *****

***** 344 BYTES OF SPACE REQUIRED FOR RELIABILITY *****

RELIABILITY ANALYSIS - SCALE (NETRAB)

- 1. CAU7
- 2. CAU8
- 3. CAU20
- 4. CAU26

		MEAN	STD DEV	CASES
1.	CAU7	1.8470	1.3108	1216.0
2.	CAU8	1.0962	1.1539	1216.0
3.	CAU20	.9038	1.1945	1216.0
4.	CAU26	1.3092	1.3535	1216.0

CORRELATION MATRIX

	CAU7	CAU8	CAU20	CAU26
CAU7	1.0000			
CAU8	.4755	1.0000		
CAU20	.2014	.2038	1.0000	
CAU26	.5514	.2281	.2511	1.0000

RELIABILITY ANALYSIS - SCALE (NETRAB)

OF CASES = 1216.0

STATISTICS FOR SCALE	MEAN	VARIANCE	STD DEV	VARIABLES		
	5.1563	12.4200	3.5242	4		
ITEM MEANS	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	1.2891	.9038	1.8470	.9433	2.0437	.1658
ITEM VARIANCES	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	1.5771	1.3315	1.8319	.5004	1.3758	.0559
INTER-ITEM COVARIANCES	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.5093	.2809	.9782	.6973	3.4827	.0705
INTER-ITEM CORRELATIONS	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.3185	.2014	.5514	.3500	2.7381	.0215

ITEM-TOTAL STATISTICS

	SCALE MEAN IF ITEM DELETED	SCALE VARIANCE IF ITEM DELETED	CORRECTED ITEM-TOTAL CORRELATION	SQUARED MULTIPLE CORRELATION	ALPHA IF ITEM DELETED
CAU7	3.3092	6.6763	.5943	.4332	.4687
CAU8	4.0600	8.3758	.4062	.2420	.6087
CAU20	4.2525	8.9889	.2798	.0858	.6854
CAU26	3.8470	7.1074	.4823	.3277	.5552

RELIABILITY COEFFICIENTS 4 ITEMS

ALPHA = .6561

STANDARDIZED ITEM ALPHA = .6515

RELIABILITY /VARIABLES CAU12 CAU23 CAU24 CAU32 /SCALE (ESOU CUR) CAU12 CAU23
CAU24 CAU32 /MODEL ALPHA /SUMMARY ALL /STATISTICS CORRELATION SCALE
DESCRIPTIVES.

***** METHOD 2 (COVARIANCE MATRIX) WILL BE USED FOR THIS ANALYSIS *****

***** 344 BYTES OF SPACE REQUIRED FOR RELIABILITY *****

RELIABILITY ANALYSIS - SCALE (ESOU CUR)

- 1. CAU12
- 2. CAU23
- 3. CAU24
- 4. CAU32

		MEAN	STD DEV	CASES
1.	CAU12	.6893	1.0923	1178.0
2.	CAU23	.1537	.6232	1178.0
3.	CAU24	.5348	1.0714	1178.0
4.	CAU32	.5730	1.1447	1178.0

CORRELATION MATRIX

	CAU12	CAU23	CAU24	CAU32
CAU12	1.0000			
CAU23	.1638	1.0000		
CAU24	.1203	.1415	1.0000	
CAU32	.3926	.2183	.2279	1.0000

RELIABILITY ANALYSIS - SCALE (ESOU CUR)

OF CASES = 1178.0

STATISTICS FOR SCALE	MEAN 1.9508	VARIANCE 6.5855	STD DEV 2.5662	VARIABLES 4		
ITEM MEANS	MEAN .4877	MINIMUM .1537	MAXIMUM .6893	RANGE .5357	MAX/MIN 4.4862	VARIANCE .0539
ITEM VARIANCES	MEAN 1.0099	MINIMUM .3884	MAXIMUM 1.3103	RANGE .9219	MAX/MIN 3.5733	VARIANCE .1764
INTER-ITEM COVARIANCES	MEAN .2121	MINIMUM .0945	MAXIMUM .4908	RANGE .3964	MAX/MIN 5.1953	VARIANCE .0208
INTER-ITEM CORRELATIONS	MEAN .2107	MINIMUM .1203	MAXIMUM .3926	RANGE .2722	MAX/MIN 3.2626	VARIANCE .0088

ITEM-TOTAL STATISTICS

	SCALE MEAN IF ITEM DELETED	SCALE VARIANCE IF ITEM DELETED	CORRECTED ITEM-TOTAL CORRELATION	SQUARED MULTIPLE CORRELATION	ALPHA IF ITEM DELETED
CAU12	1.2615	3.9061	.3442	.1611	.4068
CAU23	1.7971	5.4737	.2481	.0632	.4994
CAU24	1.4160	4.4080	.2289	.0614	.5159
CAU32	1.3778	3.4230	.4373	.2053	.3039

RELIABILITY COEFFICIENTS 4 ITEMS

ALPHA = .5154 STANDARDIZED ITEM ALPHA = .5164

RELIABILITY /VARIABLES CAU13 CAU31 /SCALE (PRESFAM) CAU13 CAU31
/MODEL ALPHA /SUMMARY ALL /STATISTICS CORRELATION SCALE
DESCRIPTIVES.

***** METHOD 2 (COVARIANCE MATRIX) WILL BE USED FOR THIS ANALYSIS *****

***** 128 BYTES OF SPACE REQUIRED FOR RELIABILITY *****

RELIABILITY ANALYSIS - SCALE (PRESFAM)

1. CAU13
2. CAU31

		MEAN	STD DEV	CASES
1.	CAU13	.1641	.5690	1213.0
2.	CAU31	.4089	.9052	1213.0

CORRELATION MATRIX

	CAU13	CAU31
CAU13	1.0000	
CAU31	.3983	1.0000

RELIABILITY ANALYSIS - SCALE (PRESFAM)

OF CASES = 1213.0

STATISTICS FOR SCALE	MEAN	VARIANCE	STD DEV	VARIABLES		
	.5730	1.5535	1.2464	2		
ITEM MEANS	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.2865	.1641	.4089	.2448	2.4925	.0300
ITEM VARIANCES	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.5716	.3237	.8195	.4957	2.5314	.1229
INTER-ITEM COVARIANCES	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.2051	.2051	.2051	.0000	1.0000	-.0000
INTER-ITEM CORRELATIONS	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.3983	.3983	.3983	.0000	1.0000	.0000

ITEM-TOTAL STATISTICS

	SCALE MEAN IF ITEM DELETED	SCALE VARIANCE IF ITEM DELETED	CORRECTED ITEM-TOTAL CORRELATION	SQUARED MULTIPLE CORRELATION	ALPHA IF ITEM DELETED
CAU13	.4089	.8195	.3983	.1586	.
CAU31	.1641	.3237	.3983	.1586	.

RELIABILITY COEFFICIENTS 2 ITEMS

ALPHA = .5282

STANDARDIZED ITEM ALPHA = .5697

RELIABILITY /VARIABLES CAU3 CAU25 /SCALE (QUINPES) CAU3 CAU25
/MODEL ALPHA /SUMMARY ALL /STATISTICS CORRELATION SCALE DESCRIPTIVES.

***** METHOD 2 (COVARIANCE MATRIX) WILL BE USED FOR THIS ANALYSIS *****

***** 128 BYTES OF SPACE REQUIRED FOR RELIABILITY *****

RELIABILITY ANALYSIS - SCALE (QUINPES)

- 1. CAU3
- 2. CAU25

		MEAN	STD DEV	CASES
1.	CAU3	1.6491	1.2167	1197.0
2.	CAU25	.6717	1.0995	1197.0

CORRELATION MATRIX

	CAU3	CAU25
CAU3	1.0000	
CAU25	.0907	1.0000

RELIABILITY ANALYSIS - SCALE (QUINPES)

OF CASES = 1197.0

STATISTICS FOR SCALE	MEAN	VARIANCE	STD DEV	VARIABLES		
	2.3208	2.9321	1.7123	2		
ITEM MEANS	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	1.1604	.6717	1.6491	.9774	2.4552	.4777
ITEM VARIANCES	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	1.3447	1.2090	1.4805	.2715	1.2245	.0368
INTER-ITEM COVARIANCES	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.1213	.1213	.1213	.0000	1.0000	.0000
INTER-ITEM CORRELATIONS	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.0907	.0907	.0907	.0000	1.0000	.0000

ITEM-TOTAL STATISTICS

	SCALE MEAN IF ITEM DELETED	SCALE VARIANCE IF ITEM DELETED	CORRECTED ITEM-TOTAL CORRELATION	SQUARED MULTIPLE CORRELATION	ALPHA IF ITEM DELETED
CAU3	.6717	1.2090	.0907	.0082	.
CAU25	1.6491	1.4805	.0907	.0082	.

RELIABILITY COEFFICIENTS 2 ITEMS

ALPHA = .1655

STANDARDIZED ITEM ALPHA = .1663

RELIABILITY /VARIABLES CAU22 /SCALE (NCOMEP) CAU22 /MODEL ALPHA /SUMMARY ALL /STATISTICS SCALE DESCRIPTIVES.

***** METHOD 2 (COVARIANCE MATRIX) WILL BE USED FOR THIS ANALYSIS *****

***** 56 BYTES OF SPACE REQUIRED FOR RELIABILITY *****

SCALE (NCOMEP) HAS LESS THAN TWO NON-ZERO VARIANCE ITEMS; CANNOT BE PROCESSED

RELIABILITY /VARIABLES CAU11 CAU30 /SCALE (MUDANC) CAU11 CAU30 /MODEL ALPHA /SUMMARY ALL /STATISTICS CORRELATION SCALE DESCRIPTIVES.

***** METHOD 2 (COVARIANCE MATRIX) WILL BE USED FOR THIS ANALYSIS *****

***** 128 BYTES OF SPACE REQUIRED FOR RELIABILITY *****

RELIABILITY ANALYSIS - SCALE (MUDANC)

- 1. CAU11
- 2. CAU30

		MEAN	STD DEV	CASES
1.	CAU11	.4587	1.0341	1212.0
2.	CAU30	.1601	.6419	1212.0

CORRELATION MATRIX

	CAU11	CAU30
CAU11	1.0000	
CAU30	.2165	1.0000

RELIABILITY ANALYSIS - SCALE (MUDANC)

OF CASES = 1212.0

STATISTICS FOR SCALE	MEAN	VARIANCE	STD DEV	VARIABLES		
	.6188	1.7687	1.3299	2		
ITEM MEANS	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.3094	.1601	.4587	.2987	2.8660	.0446
ITEM VARIANCES	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.7407	.4120	1.0693	.6573	2.5953	.2160
INTER-ITEM COVARIANCES	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.1437	.1437	.1437	.0000	1.0000	-.0000
INTER-ITEM CORRELATIONS	MEAN	MINIMUM	MAXIMUM	RANGE	MAX/MIN	VARIANCE
	.2165	.2165	.2165	.0000	1.0000	.0000

ITEM-TOTAL STATISTICS

	SCALE MEAN IF ITEM DELETED	SCALE VARIANCE IF ITEM DELETED	CORRECTED ITEM-TOTAL CORRELATION	SQUARED MULTIPLE CORRELATION	ALPHA IF ITEM DELETED
CAU11	.1601	.4120	.2165	.0469	.
CAU30	.4587	1.0693	.2165	.0469	.

RELIABILITY COEFFICIENTS 2 ITEMS

ALPHA = .3250

STANDARDIZED ITEM ALPHA = .3559

ANEXO 4:

TABELAS CRUZADAS ENTRE AS VARIÁVEIS INDICADORAS E AS VARIÁVEIS INDEPENDENTES

Crosstabulation: SEXO
By PREC

PREC-> SEXO	Count				Row
	Std Res	0	1	2	Total
1	478 .9	196 -1.0	50 -.7	724 60.1%	
2	285 -1.1	153 1.2	42 .9	480 39.9%	
Column Total	763 63.4%	349 29.0%	92 7.6%	1204 100.0%	

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
5.59402	2	.0610	36.678	None

Crosstabulation: ESCIV
By PREC

PREC-> ESCIV	Count				Row
	Std Res	0	1	2	Total
1	423 -.0	189 -.2	54 .4	666 55.8%	
2	336 .0	154 .2	37 -.5	527 44.2%	
Column Total	759 63.6%	343 28.8%	91 7.6%	1193 100.0%	

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
.53149	2	.7666	40.199	None

Crosstabulation: IDADE
By PREC

PREC->	Count				Row
	Std Res	0	1	2	Total
IDADE	0	375 -.6	192 1.0	44 -.3	611 54.2%
	1	281 .4	121 -.5	32 -.1	434 38.5%
	2	58 .8	16 -1.7	9 1.1	83 7.4%
Column Total		714 63.3%	329 29.2%	85 7.5%	1128 100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
6.47664	4	.1663	6.254	None

Crosstabulation: INGR
By PREC

PREC->	Count				Row
	Std Res	0	1	2	Total
INGR	1	563 -.5	270 .4	74 .6	907 75.5%
	2	104 .6	39 -.9	11 -.2	154 12.8%
	3	95 .6	40 -.1	6 -1.4	141 11.7%
Column Total		762 63.4%	349 29.0%	91 7.6%	1202 100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
4.43693	4	.3501	10.675	None

Crosstabulation: SERAC
By PREC

PREC->	Count				Row
	Std Res	0	1	2	Total
SERAC	1	407	160	39	606
		1.3	-1.3	-1.0	51.1%
	2	258	134	39	431
		-.8	.7	1.1	36.3%
	3	83	54	12	149
		-1.1	1.6	.2	12.6%
Column		748	348	90	1186
Total		63.1%	29.3%	7.6%	100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
10.55976	4	.0320	11.307	None

Crosstabulation: SEXO
By MERCT

MERCT->	Count				Row
	Std Res	0	1	2	Total
SEXO	1	515	187	25	727
		1.3	-.7	-2.8	60.0%
	2	296	142	47	485
		-1.6	.9	3.4	40.0%
Column		811	329	72	1212
Total		66.9%	27.1%	5.9%	100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
24.67911	2	.0000	28.812	None

Crosstabulation: ESCIV
By MERCT

MERCT->	Count				Row
ESCIV	Std Res	0	1	2	Total
1	428	193	54	675	
	-1.1	.7	2.1	56.1%	
2	375	135	18	528	
	1.2	-.7	-2.4	43.9%	
Column Total	803	328	72	1203	
	66.7%	27.3%	6.0%	100.0%	

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
14.00069	2	.0009	31.601	None

Crosstabulation: IDADE
By MERCT

MERCT->	Count				Row
IDADE	Std Res	0	1	2	Total
0	388	185	42	615	
	-1.3	1.5	1.0	54.1%	
1	307	108	21	436	
	.8	-.9	-.9	38.4%	
2	69	13	3	85	
	1.6	-2.1	-.9	7.5%	
Column Total	764	306	66	1136	
	67.3%	26.9%	5.8%	100.0%	

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
14.57853	4	.0057	4.938	1 OF 9 (11.1%)

Crosstabulation: INGR
By MERCT

MERCT->	Count	0	1	2	Row Total
INGR	Std Res				
1		611	252	51	914
		-.0	.2	-.5	75.5%
2		100	44	11	155
		-.4	.3	.6	12.8%
3		99	33	10	142
		.4	-.9	.5	11.7%
Column Total		810	329	72	1211
		66.9%	27.2%	5.9%	100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
2.08514	4	.7201	8.443	None

Crosstabulation: SERAC
By MERCT

MERCT->	Count	0	1	2	Row Total
SERAC	Std Res				
1		418	155	39	612
		.4	-.9	.5	51.3%
2		283	127	19	429
		-.2	.9	-1.2	35.9%
3		97	44	12	153
		-.5	.3	1.0	12.8%
Column Total		798	326	70	1194
		66.8%	27.3%	5.9%	100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
5.14035	4	.2732	8.970	None

Crosstabulation: SEXO
By FAAPCUR

FAAPCUR->	Count				Row Total
		Std Res	0	1	
SEXO					
1	588	.9	-1.5	-1.1	729 60.2%
2	351	-1.2	1.8	1.3	481 39.8%
Column Total	939	77.6%	19.8%	2.6%	1210 100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
10.53780	2	.0051	12.323	None

Crosstabulation: ESCIV
By FAAPCUR

FAAPCUR->	Count				Row Total
		Std Res	0	1	
ESCIV					
1	483	-1.7	2.7	1.8	673 56.1%
2	447	1.9	-3.0	-2.1	527 43.9%
Column Total	930	77.5%	19.9%	2.6%	1200 100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
30.37353	2	.0000	13.614	None

Crosstabulation: IDADE
By FAAPCUR

FAAPCUR->	Count				Row Total	
		Std Res	0	1		2
IDADE	0		436	159	20	615
			-1.9	3.3	1.1	54.2%
	1		366	61	8	435
			1.6	-2.8	-.9	38.3%
	2		78	6	1	85
			1.5	-2.7	-.8	7.5%
	Column Total		880	226	29	1135
			77.5%	19.9%	2.6%	100.0%

<u>Chi-Square</u>	<u>D.F.</u>	<u>Significance</u>	<u>Min E.F.</u>	<u>Cells with E.F.< 5</u>
36.37057	4	.0000	2.172	1 OF 9 (11.1%)

Crosstabulation: INGR
By FAAPCUR

FAAPCUR->	Count				Row Total	
		Std Res	0	1		2
INGR	1		682	205	24	911
			-.9	1.8	.1	75.4%
	2		138	14	3	155
			1.6	-3.0	-.5	12.8%
	3		118	21	4	143
			.7	-1.4	.2	11.8%
	Column Total		938	240	31	1209
			77.6%	19.9%	2.6%	100.0%

<u>Chi-Square</u>	<u>D.F.</u>	<u>Significance</u>	<u>Min E.F.</u>	<u>Cells with E.F.< 5</u>
18.51151	4	.0010	3.667	2 OF 9 (22.2%)

Crosstabulation: SERAC
By FAAPCUR

FAAPCUR->	Count	SERAC			Row Total
		0	1	2	
SERAC	1	452	138	22	612
	Std Res	-1.1	1.5	1.5	51.3%
	2	352	72	5	429
	Std Res	1.0	-1.4	-1.8	36.0%
	3	121	26	4	151
	Std Res	.4	-.7	.0	12.7%
Column Total		925	236	31	1192
Std Res		77.6%	19.8%	2.6%	100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F. < 5
12.86634	4	.0119	3.927	1 OF 9 (11.1%)

Crosstabulation: SEXO
By NETRAB

NETRAB->	Count	SEXO			Row Total
		0	1	2	
SEXO	1	266	328	137	731
	Std Res	-.1	.6	-.8	60.2%
	2	178	199	107	484
	Std Res	.1	-.8	1.0	39.8%
Column Total		444	527	244	1215
Std Res		36.5%	43.4%	20.1%	100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F. < 5
2.60115	2	.2724	97.198	None

Crosstabulation: ESCIV
By NETRAB

Count					Row
NETRAB->	Std Res	0	1	2	Total
ESCIV					
1	299	255	122	676	
	3.4	-2.3	-1.1	56.1%	
2	139	270	119	528	
	-3.8	2.6	1.3	43.9%	
Column Total	438	525	241	1204	
	36.4%	43.6%	20.0%	100.0%	

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
41.34545	2	.0000	105.688	None

Crosstabulation: IDADE
By NETRAB

Count					Row
NETRAB->	Std Res	0	1	2	Total
IDADE					
0	263	235	115	613	
	2.5	-1.6	-1.0	53.9%	
1	133	202	105	440	
	-2.3	1.1	1.5	38.7%	
2	23	48	14	85	
	-1.5	2.0	-.8	7.5%	
Column Total	419	485	234	1138	
	36.8%	42.6%	20.6%	100.0%	

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
25.13093	4	.0000	17.478	None

Crosstabulation: INGR
By NETRAB

NETRAB->	Count				Row Total	
		Std Res	0	1		2
INGR	1		352	379	183	914
			1.0	-.9	-.1	75.4%
	2		46	80	29	155
			-1.4	1.5	-.4	12.8%
	3		44	68	32	144
			-1.2	.7	.6	11.9%
	Column Total		442	527	244	1213
			36.4%	43.4%	20.1%	100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
8.54152	4	.0736	28.966	None

Crosstabulation: SERAC
By NETRAB

NETRAB->	Count				Row Total	
		Std Res	0	1		2
SERAC	1		236	246	131	613
			.7	-1.0	.6	51.3%
	2		145	199	86	430
			-1.0	1.1	-.1	36.0%
	3		59	68	26	153
			.4	.3	-.9	12.8%
	Column Total		440	513	243	1196
			36.8%	42.9%	20.3%	100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
5.23198	4	.2643	31.086	None

Crosstabulation: SEXO
By ESOUCUR

ESOUCUR->	Count				Row Total
		0	1	2	
SEXO	Std Res				
1		576	117	15	708
		.4	-.1	-1.8	60.2%
2		365	79	25	469
		-.5	.1	2.3	39.8%
Column Total		941	196	40	1177
		79.9%	16.7%	3.4%	100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
9.02072	2	.0110	15.939	None

Crosstabulation: ESCIV
By ESOUCUR

ESOUCUR->	Count				Row Total
		0	1	2	
ESCIV	Std Res				
1		473	152	29	654
		-2.2	4.1	1.4	56.0%
2		459	43	11	513
		2.4	-4.6	-1.6	44.0%
Column Total		932	195	40	1167
		79.9%	16.7%	3.4%	100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
52.97587	2	.0000	17.584	None

Crosstabulation: IDADE
By ESOUCUR

ESOUCUR->	Count	Count			Row Total
		0	1	2	
IDADE	0	424	145	29	598
	Std Res	-2.4	4.5	1.7	54.3%
	1	373	37	9	419
	Std Res	2.1	-3.9	-1.5	38.1%
	2	81	2	1	84
	Std Res	1.7	-3.2	-1.1	7.6%
	Column Total	878	184	39	1101
	Total	79.7%	16.7%	3.5%	100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
66.02743	4	.0000	2.975	1 OF 9 (11.1%)

Crosstabulation: INGR
By ESOUCUR

ESOUCUR->	Count	Count			Row Total
		0	1	2	
INGR	1	676	173	37	886
	Std Res	-1.2	2.1	1.2	75.4%
	2	145	5	1	151
	Std Res	2.2	-4.0	-1.8	12.9%
	3	119	17	2	138
	Std Res	.8	-1.2	-1.2	11.7%
	Column Total	940	195	40	1175
	Total	80.0%	16.6%	3.4%	100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
35.63259	4	.0000	4.698	1 OF 9 (11.1%)

Crosstabulation: SERAC
By ESOUCUR

ESOUCUR->	SERAC	Count				Row
		Std Res	0	1	2	Total
1		438	126	23	587	
		-1.4	2.7	.9	50.7%	
2		355	56	13	424	
		.9	-1.9	-.2	36.6%	
3		131	14	2	147	
		1.3	-2.2	-1.3	12.7%	
Column Total		924	196	38	1158	
		79.8%	16.9%	3.3%	100.0%	

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
22.20046	4	.0002	4.824	1 OF 9 (11.1%)

Crosstabulation: SEXO
By PRESFAM

PRESFAM->	SEXO	Count				Row
		Std Res	0	1	2	Total
1		636	82	11	729	
		1.0	-1.7	-1.9	60.1%	
2		380	82	21	483	
		-1.2	2.1	2.3	39.9%	
Column Total		1016	164	32	1212	
		83.8%	13.5%	2.6%	100.0%	

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
18.45869	2	.0001	12.752	None

Crosstabulation: ESCIV
By PRESFAM

PRESFAM->	Count				Row
ESCIV	Std Res	0	1	2	Total
1	619	48	5	672	56.0%
	2.4	-4.5	-3.0		
2	387	115	27	529	44.0%
	-2.7	5.1	3.4		
Column Total	1006	163	32	1201	100.0%
	83.8%	13.6%	2.7%		

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
80.27933	2	.0000	14.095	None

Crosstabulation: IDADE
By PRESFAM

PRESFAM->	Count				Row
IDADE	Std Res	0	1	2	Total
0	538	66	9	613	54.0%
	.9	-1.6	-1.7		
1	352	68	17	437	38.5%
	-.9	1.4	1.7		
2	67	15	3	85	7.5%
	-.6	1.1	.6		
Column Total	957	149	29	1135	100.0%
	84.3%	13.1%	2.6%		

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
13.99867	4	.0073	2.172	1 OF 9 (11.1%)

Crosstabulation: INGR
By PRESFAM

PRESFAM->	Count	INGR			Row Total	
		Std Res	0	1		2
INGR	1		774	116	22	912
			.4	-.7	-.4	75.4%
	2		130	21	4	155
			.0	-.0	-.0	12.8%
	3		110	27	6	143
			-.9	1.7	1.1	11.8%
Column Total			1014	164	32	1210
			83.8%	13.6%	2.6%	100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
5.88383	4	.2080	3.782	2 OF 9 (22.2%)

Crosstabulation: SERAC
By PRESFAM

PRESFAM->	Count	SERAC			Row Total	
		Std Res	0	1		2
SERAC	1		514	78	16	608
			.2	-.6	.1	51.0%
	2		360	61	11	432
			-.1	.3	-.1	36.2%
	3		125	24	4	153
			-.3	.7	.0	12.8%
Column Total			999	163	31	1193
			83.7%	13.7%	2.6%	100.0%

Chi-Square	D.F.	Significance	Min E.F.	Cells with E.F.< 5
.97238	4	.9140	3.976	1 OF 9 (11.1%)