



UNIVERSIDADE ESTADUAL PAULISTA  
"JÚLIO DE MESQUITA FILHO"

**NILO AMARAL MARTIN**

Modelo de Regressão Logística aplicado ao estudo  
epidemiológico do HIV/AIDS na cidade de Campinas-SP.

Botucatu  
Estado de São Paulo – Brasil  
2010

**NILO AMARAL MARTIN**

**Modelo de Regressão Logística aplicado ao estudo epidemiológico do  
HIV/AIDS na cidade de Campinas-SP.**

Trabalho de Conclusão de  
Curso apresentado ao  
Instituto de Biociências da  
Universidade Estadual  
Paulista “Júlio de Mesquita  
Filho” para obtenção do  
Título de Bacharel em Física  
Médica

Orientador: IBB: Prof. Dr. José Raimundo de Souza Passos

Botucatu  
Estado de São Paulo – Brasil  
2010

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.  
DIVISÃO DE BIBLIOTECA E DOCUMENTAÇÃO - CAMPUS DE BOTUCATU - UNESP  
BIBLIOTECÁRIA RESPONSÁVEL: *ROSEMEIRE APARECIDA VICENTE*

Martin, Nilo Amaral.

Modelo de regressão logística aplicado ao estudo epidemiológico do HIV/AIDS na cidade de Campinas-SP / Nilo Amaral Martin. - Botucatu, 2010

Trabalho de conclusão de curso (bacharelado - Física Médica) - Instituto de Biociências de Botucatu, Universidade Estadual Paulista, 2010

Orientador: Prof. Dr. José Raimundo de Souza Passos

Capes: 40601005

1. Medicina - Saúde Coletiva. 2. Epidemiologia – Métodos estatísticos.  
3. Análise de regressão.

Palavras-chave: Modelo de regressão, Razão de chances, Resposta binomial.

*À minha família,  
por terem sempre me mostrado os reais valores da vida e a todos que amo,  
simplesmente por fazerem parte de tudo.*

## ***Agradecimentos***

*Primeiramente agradeço ao meu orientador, o Professor Doutor José Raimundo de Souza Passos e à sua esposa, a Professora Doutora Liciania Vaz de Arruda Silveira por terem me dado a oportunidade e o incentivo para o desenvolvimento deste trabalho, e por terem aberto novas portas para o meu desenvolvimento futuro. Agradeço também à Professora Doutora Maria Rita, por ceder os dados utilizados para este estudo, pois sem eles este trabalho nunca teria existido.*

*Gostaria de agradecer também aos meus pais, Carlos Roberto A. Martin e Ana Silvia M. Amaral pelo imensurável apoio e disposição em possibilitar que eu chegasse até aqui e à minha namorada e companheira de cada momento Juliana A. Fernandes, que durante praticamente todo o processo de desenvolvimento deste trabalho esteve sempre disposta a ajudar e incentivar as minhas decisões.*

*Por fim agradeço a todos que fizeram e fazem parte do meu convívio em Botucatu durante todo o período da graduação, no qual juntos aprendemos e crescemos como pessoas, não só em relação à universidade, mas também como seres humanos. Aos integrantes da banda Grau Etilico, que proporcionaram momentos e oportunidades inesquecíveis e aos membros da república BGD pelo imenso companheirismo. Considero todos como uma família para mim.*

*“Vivemos pelo que acreditamos.  
Nosso limite está nisso.  
Portanto, se cremos no que ilimitado é, sem limites viveremos”.*  
*Allan Kardec*

## **Resumo**

A compreensão da epidemia do HIV e da AIDS em suas diferentes dimensões socioeconômica, cultural, biológica e política entre outras, representa um grande desafio para a Saúde Pública no Brasil e no mundo, desde o seu surgimento a cerca de duas décadas.

Tanto na magnitude da propagação do HIV quanto na mortalidade por AIDS, reflete-se a crescente desigualdade entre países desenvolvidos e em desenvolvimento.

Na medida em que se acentuam as diferenças de acesso ao tratamento, diminui a mortalidade por AIDS nos países mais ricos e aumenta nos países mais pobres, exceção feita ao Brasil, um dos poucos países que adotaram a política de distribuição gratuita de anti-retrovirais.

Segue, no desenvolvimento deste trabalho, a análise do perfil epidemiológico da doença no município de Campinas-SP de acordo com a evolução do tratamento disponibilizado pelo Sistema Único de Saúde (SUS), utilizando o método de Regressão Logística segundo variáveis consideradas relevantes para o estudo e com base em parâmetros estimados a partir dessas variáveis.

## SUMÁRIO

1. Introdução .....	p. 09
2. Metodologia .....	p. 13
2.1 Regressão Logística .....	p. 13
2.2 Teste de Hipóteses .....	p. 20
3. Resultados e Discussão .....	p. 23
4. Referências Bibliográficas .....	p. 33

## 1. Introdução

A AIDS, quando começou, parecia ser uma doença de Primeiro Mundo e de gente rica, e talvez por isso tenha despertado tanto investimento - talvez insuficiente - em pesquisas e tanto interesse na mídia. Com o tempo, verificou-se que a AIDS era uma epidemia mundial, que se deslocava do Primeiro para o Terceiro Mundo, inclusive mostrando-se uma tragédia em muitos países da África. A AIDS tornou-se mundial e ligada principalmente à pobreza.

Entretanto, o tratamento da AIDS em qualquer país exige muita atenção médica, além de ser caro. São caros os remédios para prevenir ou combater infecções, é caro internar um doente com AIDS. Enfim, a AIDS é muito cara, e não prevenir a doença fica mais caro ainda.

No Brasil não é diferente. A maioria das pessoas infectadas com o vírus, ou doentes, é pobre e não consegue recursos públicos ou particulares para ter o atendimento de que necessitam, nem para a prevenção nem para o tratamento. Além disso, o acesso a informações sobre a prevenção também é limitado em virtude da baixa qualidade da educação a qual essa parcela da população tem acesso (CASTILHO E. A. et al, 2000).

Por outro lado, o Brasil apresenta um excelente programa de assistência às pessoas infectadas pelo HIV, disponibilizando sem qualquer custo os medicamentos necessários para o tratamento, possibilitando assim uma melhora na qualidade de vida dos indivíduos infectados, desde que o indivíduo saiba que está infectado e saiba também como recorrer ao estar ciente da infecção.

Percebe-se, além disso, que não é suficiente oferecer conhecimento, informando as pessoas precisamente sobre como se comportar e sobre o que evitar. O desafio real então é produzir meios de intervenção que ultrapassem os aspectos quase que exclusivamente informativos e sejam capazes de mobilizar transformações no modo de ser das pessoas (CHAVES LIMA M. A., 2008).

Ainda assim, em alguns hospitais públicos estão os pobres, onde além do atendimento médico solidário e humano, necessitam de meios para ter acesso ao que a ciência moderna já colocou à disposição de todos, em termos de diagnósticos sofisticados e remédios eficientes.

Quando a AIDS surgiu, nos anos 80, pouco se sabia sobre as diferenças entre contaminado e doente, pois ambos morriam em muito pouco tempo. Era questão de um ano entre a notícia e a morte. Com o surgimento das primeiras drogas que controlavam o desenvolvimento do vírus e o acúmulo do conhecimento clínico que acelerava o diagnóstico e tratamento das infecções oportunistas, o tempo de vida foi aumentando de forma muito significativa (CASTILHO E. A. et al, 2000).

Hoje se sabe que uma pessoa pode ficar contaminada sem manifestar a doença por muito tempo, entre 10 e 15 anos, e que alguns podem viver sem que a doença se manifeste. Estar com o vírus não é mais sinônimo de ficar doente. Sabe-se também que uma pessoa doente que recebe todos os tratamentos e cuidados necessários pode sobreviver com a doença por muitos ou vários anos, três a quatro vezes mais do que se vivia antes. Resumindo, viver ou morrer, em grande medida, depende do tratamento existente.

Durante todo o período de existência e propagação da epidemia de HIV/AIDS, algo foi imprescindível para estabelecer medidas de controle sobre a disseminação do vírus: conhecer suas formas e seus meios de transmissão, bem como as maneiras de prevenir-se. Desde o início, essa epidemia mobilizou esforços da comunidade científica no sentido de descobrir, de maneira rápida, os meios de reagir contra a contaminação pelo HIV (HENN, C. A., 2007).

Sabe-se que o HIV, embora possa ser encontrado em todos os fluidos corpóreos, só tem poder de contaminação quando no sangue, nas secreções sexuais (esperma e secreção vaginal) e no leite materno, que são suas vias de contaminação. Quando o indivíduo se expõe a alguma situação de risco e contrai o HIV, o sistema imunológico reconhece sua presença e dá início à produção de anticorpos anti-HIV, na tentativa de neutralizar seus efeitos. O HIV, porém, vai se alojar dentro de algumas células do organismo, inclusive dentro das células coordenadoras do sistema imunológico (os linfócitos). Em geral, o processo de produção de anticorpos pode ser iniciado com até duas semanas de retardo, contadas a partir do momento da contaminação, estendendo-se até três meses após o momento do contágio. Essa faixa de tempo ocorre em 98% dos casos, sendo denominada soroconversão ou "janela imunológica". Em casos mais raros esse período pode se estender até seis meses (Disponível em: <http://www.mundodoquimico.hpg.ig.com.br/aids2.htm>, acesso em 03/09/2010).

Desde o momento de entrada do vírus no organismo humano ocorre uma reação das células de defesa tão logo elas tenham identificado e localizado a presença do vírus. Inicia-se um processo de embate entre a carga viral e as células imunológicas na tentativa de manter o funcionamento da imunidade corporal em equilíbrio. Esse ciclo se repete até que o sistema imunológico, desorganizado, não consiga mais desempenhar seu papel de defesa do organismo contra o ataque dos agentes causadores de doenças, o que se traduz na condição de imunodeficiência. O organismo, incapaz de se proteger contra o ataque dos agentes causadores de doenças, fica vulnerável a uma série de infecções oportunistas, assim chamadas porque se aproveitam da oportunidade do sistema imunológico estar desorganizado para se

manifestarem, caracterizando o desenvolvimento da AIDS (Disponível em: <http://www.mundodoquimico.hpg.ig.com.br/aids2.htm>, acesso em 03/09/2010).

A Organização Mundial da Saúde (OMS) estimou até 2001 a existência de 40 milhões de pessoas infectadas pelo HIV no mundo e um total de 25 milhões de óbitos em consequência da epidemia do vírus. Segundo estimativas, também da OMS, três milhões de pessoas morreram em consequência da AIDS em 2001 (HENN, C. A., 2007).

Segundo dados obtidos no site do Governo Federal, de 1980 a junho de 2007 foram notificados 474.273 casos de AIDS no Brasil – 289.074 no Sudeste, 89.250 no Sul, 53.089 no Nordeste, 26.757 no Centro Oeste e 16.103 no Norte. No Brasil e nas regiões Sul, Sudeste e Centro Oeste, a incidência de AIDS tende à estabilização, porém no Norte e Nordeste a tendência é de crescimento. Segundo critérios da Organização Mundial de Saúde (OMS), o Brasil tem uma epidemia concentrada, com taxa de prevalência da infecção pelo HIV de 0,6% na população de 15 a 49 anos.

Ainda segundo o site do Governo, em 2006 foram registrados 32.628 casos da doença, e em 2005 foram identificados 35.965 casos, o que representa uma taxa de incidência de 19,5 casos de AIDS a cada 100 mil habitantes, ainda segundo a OMS.

Considerando a sobrevivência de indivíduos após a notificação da doença, os dados apontam que, cinco anos após serem diagnosticadas, 90% das pessoas com AIDS no Sudeste (região na qual está contida o estado de São Paulo) estavam vivas em 2000. Nas outras regiões, os percentuais foram de 78%, no Norte; 80%, no Centro Oeste; 81%, no Nordeste; e 82%, no Sul, o que representa uma gradual melhora no tratamento disponibilizado pelo governo nacional.

Porém, no mesmo ano, 20,5% dos indivíduos diagnosticados com AIDS no Norte haviam morrido em até um ano após a descoberta da doença, no Centro Oeste o percentual foi de 19,2% e no Nordeste, de 18,3%. Na região Sudeste, o indicador cai para 16,8% e, no Sul, para 13,5%. A média do Brasil foi de 16,1%. Em números absolutos, o Brasil registrou 192.709 óbitos por AIDS de 1980 a 2006 (Disponível em <http://www2.aids.gov.br/data/Pages/LUMIS13F4BF21PTBRIE.htm>, acesso em 18/10/2010).

No total, até hoje, foram identificados 314.294 casos de AIDS em homens e 159.793 em mulheres. Ao longo do tempo, a razão entre os sexos vem diminuindo de forma progressiva. Em 1985, havia 15 casos da doença em homens para 1 em mulher, já atualmente a relação é de 1,5 para 1. Na faixa etária de 13 a 19 anos, houve inversão na razão de sexo, a partir de 1998

(Disponível em <http://www2.aids.gov.br/data/Pages/LUMIS13F4BF21PTBRIE.htm>, acesso em 18/10/2010).

Em ambos os sexos, a maior parte dos casos se concentra na faixa etária de 25 a 49 anos. Porém, nos últimos anos, tem-se verificado aumento percentual de casos na população acima de 50 anos, em ambos os sexos.

Do total de 192.709 óbitos por AIDS identificados no Brasil, a maioria foi no Sudeste, região na qual se encontra o estado de São Paulo, com 131.840 mortes em decorrência da doença, pelo fato de a epidemia na região ser a mais antiga do país e conseqüentemente apresentar o maior número de notificações. Em seguida, vêm Sul (28.784), Nordeste (18.379), Centro Oeste (8.738) e Norte (4.968) (Disponível em <http://www2.aids.gov.br/data/Pages/LUMIS13F4BF21PTBRIE.htm>, acesso em 18/10/2010).

No ano de 2004 foi feita uma pesquisa de abrangência nacional, que estimou que de 593 mil pessoas, entre 15 a 49 anos de idade viviam com HIV e AIDS (0,61%) no Brasil. Deste número, cerca de 208 mil são mulheres (0,42%) e 385 mil são homens (0,80%).

A pesquisa mostra ainda que quase 91% da população brasileira de 15 a 54 anos citaram a relação sexual como forma de transmissão do HIV e 94% citaram o uso de preservativo como forma de prevenção da infecção. O conhecimento é maior entre as pessoas de 25 a 39 anos, entre os mais escolarizados e entre as pessoas residentes nas regiões Sul e Sudeste. Hoje, estima-se que aproximadamente 600 mil pessoas vivem com HIV/AIDS no Brasil (Disponível em <http://www2.aids.gov.br/data/Pages/LUMIS13F4BF21PTBRIE.htm>, acesso em 18/10/2010).

Os primeiros casos de AIDS no Brasil foram detectados nas cidades de São Paulo e Rio de Janeiro, no início da década de 1980 (CASTILHO et al., 1994). Porém, acredita-se que o vírus HIV foi introduzido no país durante a década de 1970, sendo disseminado de maneira progressiva por todo o território nacional nas décadas seguintes (DIAS E NOBRE, 2001).

Já em 1985, o município de Campinas estava entre os seis municípios com cinco ou mais casos de AIDS no Estado de São Paulo.

A assistência à AIDS no Brasil e conseqüentemente no município de Campinas oferece acesso universal à terapia anti-retroviral (TARV). Baseando-se nos princípios do Sistema Único de Saúde, o Ministério da Saúde iniciou a distribuição de drogas profiláticas e terapêuticas para o tratamento das doenças oportunistas em meados dos anos 80.

Já em 1991 a primeira droga anti-retroviral (ARV), o AZT (zidovudina), foi disponibilizada para um número significativo de pacientes vivendo com AIDS, mas não para a

totalidade da população específica. No ano de 1996, foi promulgada a lei que garante a todos os indivíduos vivendo com HIV/AIDS o direito de receber, sem custo, todas as drogas necessárias ao seu tratamento pelo SUS. De 1996 e até o presente momento, pacientes HIV e AIDS recebem gratuitamente a TARV resultando na melhora da qualidade de vida, no aumento da sobrevida e em grande redução das taxas de hospitalização no país (Brasil, MS; CHEQUER et al, 1992).

Os dados obtidos para a realização do presente estudo são provenientes da cidade de Campinas-SP, num período que vai de 1980 até 2005, apresentando informações como sexo, idade, grau de escolaridade, ano de notificação da infecção e condição de óbito dos indivíduos notificados como portadores do HIV. Essas informações, com exceção da condição de óbito do indivíduo, serão utilizadas como covariáveis no estudo.

Os anos de estudo (1980 a 2005) foram agrupados em dois períodos. O primeiro vai de 1980 até 1995, e o segundo vai de 1996 até 2005. Os períodos foram definidos dessa forma para sumarizar a distribuição temporal dos casos de AIDS em maiores de 13 anos no município e evidenciar tendências da infecção pelo HIV, além de relacionar o numero de óbitos com o processo de distribuição e disponibilização dos medicamentos aos indivíduos infectados pelo vírus.

A partir disso, foi então aplicado o modelo de regressão logística com o objetivo de estudar o efeito das variáveis (gênero, idade, grau de escolaridade e período) com relação à situação do individuo portador do vírus.

## 2. Metodologia

### 2.1 Regressão Logística

O modelo de regressão logística neste trabalho foi utilizado como ferramenta de pesquisa na área epidemiológica. A abordagem foi no sentido de qual a relação de determinadas variáveis estudadas ( $X_1, X_2, \dots, X_k$ ) com uma determinada doença ou estado (D), onde a variável resposta necessariamente deve ser binária.



Figura 1: Representação da relação de determinadas variáveis com a resposta binária.

Fonte: KLEINBAUM, D. G.; KLEIN, M. **Logistic Regression, a Self-Learning Text**, 3ª edição, Londres-ENG, Springer, 2002, p. 5.

No caso, a infecção pelo vírus HIV (com desenvolvimento ou não da AIDS), resultando no óbito ou não do indivíduo infectado pelo vírus foi escolhida como variável resposta, sendo que 1 representa a morte do indivíduo infectado e 0 representa que o indivíduo permaneceu vivo.

A partir daí foram enumeradas as variáveis independentes que estão relacionadas com a variável resposta (dependente) para que em seguida fosse possível estimar os parâmetros relacionados com cada uma das variáveis independentes escolhidas, possibilitando uma análise da relação de cada uma dessas variáveis com o *estado* do indivíduo (morto ou vivo). As variáveis enumeradas foram *idade*, *gênero*, *período* (relacionada com o ano de notificação da infecção) e *grau de escolaridade*.

A variável “*período*” foi denotada da seguinte maneira: o primeiro período representa que o indivíduo obteve a notificação da infecção entre os anos de 1980 e 1995, e o segundo período representa que o indivíduo recebeu a notificação da infecção entre os anos de 1996 e 2005. Essa variável foi incluída no estudo pois existem diferenças significativas no tratamento disponibilizado pelo SUS exatamente nestes períodos, com uma melhora gradual crescente iniciada por volta de 1996, o que acaba por influenciar significativamente na variável resposta.

Já a variável “*escolaridade*” foi categorizada de acordo com os anos de estudo concluídos pelo indivíduo. A classificação é obtida em função da série e do grau que a pessoa estava frequentando ou frequentou, considerando a última série concluída com aprovação. A correspondência é feita de forma que cada série concluída com aprovação corresponde a um ano de estudo, sendo que 1 equivale a nenhum ano de estudo, 2 equivale a um a três anos de estudo, 3 equivale a quatro a sete anos de estudo, 4 equivale a oito a onze anos de estudo, 5 equivale a doze e mais anos de estudo, 6 não se aplica e 9 significa que a informação foi ignorada.

As variáveis *idade* e *gênero* representam a idade do indivíduo no momento da notificação da infecção e se o indivíduo é do sexo masculino ou feminino, respectivamente.

Essas variáveis foram escolhidas com o objetivo de traçar um perfil socioeconômico dos indivíduos que apresentaram a infecção pelo vírus HIV, correlacionando essas informações com a gradual melhora da qualidade do tratamento desde 1980 até 2005 e conseqüentemente com a diminuição do número de óbitos dos indivíduos infectados pelo vírus.

A variável resposta será chamada de  $Y$ , que é função das variáveis *idade*, *gênero*, *período* e *grau de escolaridade* que são as variáveis *independentes*. As variáveis independentes serão denotadas por  $X_1$ ,  $X_2$ ,  $X_3$  e  $X_4$  respectivamente.

Para que as relações entre as variáveis independentes e a variável resposta  $Y$  possam ser descritas com uma precisão tal qual seja possível tirar conclusões confiáveis, será utilizado então o modelo da Regressão Logística.

Outros tipos de modelagens matemáticas poderiam também ser utilizados, mas a regressão logística é de longe o método mais popular e mais utilizado em casos de análise quando a variável resposta é uma resposta binária (KLEINBAUM, D. G et al 2002).

Primeiramente, é interessante se atentar à *função logística*, que descreve a forma matemática na qual o modelo logístico é baseado. Tal função, chamada  $f(z)$ , é denotada por  $f(z) = \frac{1}{1+e^{-z}}$ . Na figura a seguir, foram plotados os valores dessa função com  $z$  variando de  $-\infty$  a  $+\infty$ .

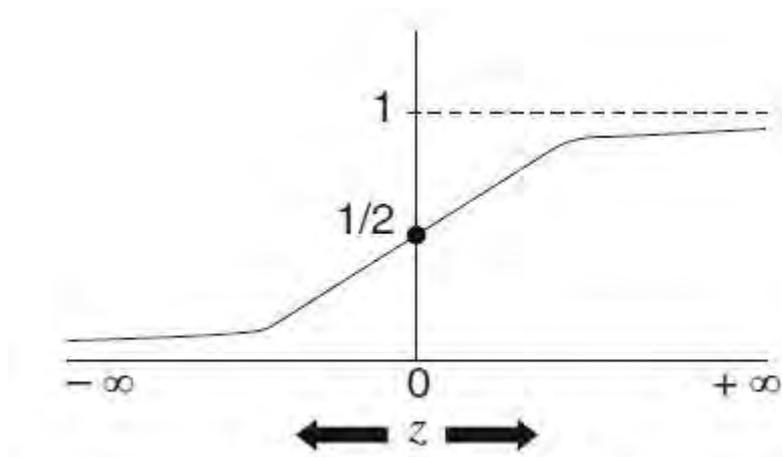


Figura 2: gráfico da função logística com  $z$  variando de  $-\infty$  a  $+\infty$ . Fonte: KLEINBAUM, D. G.; KLEIN, M. **Logistic Regression, a Self-Learning Text**, 3ª edição, Londres-ENG, Springer, 2002, p. 5.

É interessante notar que, do lado esquerdo do gráfico (quando  $z$  é  $-\infty$ ), a função logística  $f(z)$  é igual a zero e, do lado direito do gráfico (quando  $z$  é  $+\infty$ ),  $f(z)$  é igual a um.

Então, como mostra o gráfico, o valor de  $f(z)$  estará sempre entre 0 e 1, independentemente do valor de  $z$ . Isto é, qualquer estimativa que seja obtida através do modelo logístico nunca será menor que 0 e nem maior que 1.

Agora, a atenção deve ser voltada para a maneira de se obter o *modelo logístico* a partir da *função logística*  $f(z) = \frac{1}{1+e^{-z}}$ .

Inicialmente, considera-se a variável  $z$  como sendo a soma linear  $\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$ , onde os termos  $X_s$  são as variáveis independentes, já citadas anteriormente e os termos  $\beta$  são constantes que representam os parâmetros ainda não conhecidos, que serão

estimados baseando-se nos dados obtidos através das variáveis  $X$ s relacionadas com o estado  $Y$ , para o grupo de indivíduos, e o termo  $\beta_0$  representa o erro relacionado ao modelo. Substituindo agora a soma que representa  $z$  na função logística, temos  $f(z) = \frac{1}{1+e^{-(\beta_0+\sum\beta_i X_i)}}$ , aonde  $i$  vai de 1 até  $n$ .

Assim, voltando ao início, o objetivo é utilizar as variáveis independentes observadas ( $X_1, X_2, X_3$  e  $X_4$ ) para determinar sua relação com a ocorrência de óbito nos indivíduos analisados.

A probabilidade que está sendo modelada neste caso pode ser denotada então da seguinte maneira:  $P(Y=1|X_1, X_2, X_3, X_4)$ , ou  $P(\mathbf{X})$ , onde  $\mathbf{X} = X_1, X_2, X_3, X_4$  (KLEINBAUM, D. G et al, 2002). Daí, o modelo logístico passa a ser escrito como  $P(\mathbf{X}) = \frac{1}{1+e^{-(\beta_0+\sum\beta_i X_i)}}$ .

O próximo passo a partir desta expressão é linearizar a função  $P(\mathbf{X})$ , de modo que o eixo no qual está contida a variável  $z$  seja parte do conjunto dos *números reais*, ou seja, tendo seu valor contido necessariamente no intervalo  $[0,1]$ . A linearização é obtida encontrando-se o valor da variável  $z$ . Este processo é denominado *transformação logit*, a partir do qual irá ser descrito o formato *logit* da função logística.

Para início do processo, sabe-se que  $P(\mathbf{X}) = \frac{1}{1+e^{-z}}$ , onde  $z = (\beta_0 + \sum\beta_i X_i)$ . Agora, o objetivo é encontrar o valor de  $z$ :

$$P(\mathbf{X}) = \frac{1}{1 + e^{-z}} \Rightarrow P(\mathbf{X})(1 + e^{-z}) = 1$$

$$1 + e^{-z} = \frac{1}{P(\mathbf{X})} \Rightarrow e^{-z} = \frac{1}{P(\mathbf{X})} - 1 \Rightarrow e^{-z} = \frac{1 - P(\mathbf{X})}{P(\mathbf{X})}$$

$$e^z = \frac{P(\mathbf{X})}{1 - P(\mathbf{X})} \Rightarrow z = \ln \frac{P(\mathbf{X})}{1 - P(\mathbf{X})}$$

Daí tem-se:

$$\beta_0 + \sum\beta_i X_i = \ln \frac{P(\mathbf{X})}{1 - P(\mathbf{X})}, \text{ ou } Y = \beta_0 + \sum\beta_i X_i,$$

que consiste no *formato logit* da função logística.

A partir disso, este modelo será aplicado às variáveis independentes de interesse neste estudo  $X_1, X_2, X_3$  e  $X_4$ , onde:

$Y$  = morte ou não do indivíduo infectado  $(0,1)$  (variável resposta)

$X_1$  = idade

$X_2$  = gênero

$X_3$  = período

$X_4$  = grau de escolaridade

Então, a função logística para este caso será escrita da seguinte maneira:

$$P(\mathbf{X}) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4)}}.$$

A partir daí, o objetivo é utilizar os dados para estimar os parâmetros desconhecidos  $\beta_0, \beta_1, \beta_2, \beta_3$  e  $\beta_4$  que, utilizando a notação estatística, são denotados como  $\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_2, \widehat{\beta}_3$ , e  $\widehat{\beta}_4$  que são os estimadores dos valores reais de  $\beta_0, \beta_1, \beta_2, \beta_3$  e  $\beta_4$ .

Para encontrar os valores desses estimadores, utiliza-se o *método dos mínimos quadrados*. Segue uma explicação deste método de uma maneira generalizada, para que em seguida as deduções possam ser aplicadas de acordo com os objetivos deste trabalho.

Porém, antes de começar, é interessante levar em consideração algumas pressuposições:

- I) A variável dependente  $Y$  é função linear das variáveis independentes  $X_i, i = 1, \dots, k$ ;
  - II) Os valores das variáveis independentes são fixos;
  - III)  $E(e_i) = 0$ , ou seja,  $E(\mathbf{e}) = \mathbf{0}$ , onde  $\mathbf{0}$  representa um vetor de zeros;
  - IV) Os erros são homocedásticos, isto é,  $E(e_j^2) = \sigma^2$ ;
  - V) Os erros são não-correlacionados entre si, isto é,  $E(e_j e_h) = 0$  para  $j \neq h$ ;
  - VI) Os erros têm distribuição normal;
- (HOFMANN, R. et al 1987).

Feito isso, vamos supor o modelo da função logística de acordo com os parâmetros determinados:  $Y_i = \beta_0 + \beta_1 X_1 + \dots + \beta_n X_n + e_i$ , onde  $e_i$  representa o valor do erro total da amostra.

Essa expressão caracteriza uma regressão linear múltipla, já que o valor da variável dependente é função linear de duas ou mais variáveis independentes, pois neste caso (HOFMANN, R. et al 1987). Então, pode-se escrever o modelo da seguinte maneira:

$$Y_j = \beta_0 + \sum \beta_{ij} X_{ij} + e_j$$

Utilizando notação matricial, pois esta é uma regressão de múltiplas variáveis, temos

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$$

onde

$$\mathbf{y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{k1} \\ 1 & X_{12} & X_{22} & \cdots & X_{k2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{kn} \end{bmatrix} \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} \quad \mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

Agora, o próximo passo é determinar os valores de  $\beta_0, \beta_1, \dots, \beta_n$ , ou seja, o valor de  $\boldsymbol{\beta}$  a partir do *método dos mínimos quadrados*. Este método tem como objetivo encontrar os valores  $\boldsymbol{\beta}$  que tornem os valores dos erros  $\mathbf{e}$  os menores possíveis, pois quanto menor for o erro total melhor será a estimativa (HOFMANN, R. et al 1987). Seguindo este raciocínio, temos

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \Rightarrow \mathbf{e} = \mathbf{y} - \mathbf{X}\boldsymbol{\beta}$$

A soma dos quadrados dos erros é dada por

$$Z = \mathbf{e}'\mathbf{e} = (\mathbf{y}' - \mathbf{X}'\boldsymbol{\beta}')(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) = \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}\boldsymbol{\beta} - \mathbf{X}'\boldsymbol{\beta}'\mathbf{y} + \mathbf{X}'\boldsymbol{\beta}'\mathbf{X}\boldsymbol{\beta}$$

Nota-se que as matrizes  $\mathbf{y}'\mathbf{X}\boldsymbol{\beta}$  e  $\mathbf{X}'\boldsymbol{\beta}'\mathbf{y}$  são iguais, pois uma é a transposta da outra e cada uma tem apenas um elemento. Então:

$$Z = \mathbf{y}'\mathbf{y} - 2\mathbf{X}'\boldsymbol{\beta}'\mathbf{y} + \mathbf{X}'\boldsymbol{\beta}'\mathbf{X}\boldsymbol{\beta}$$

A função  $Z$  apresenta ponto de mínimo para os valores de  $\beta$  que a tornem uma diferencial identicamente nula, isto é:

$$dZ = -2(d\beta')X'y + (d\beta')X'X\beta + \beta'X'X(d\beta) \equiv 0$$

Como  $(d\beta')X'X\beta = \beta'X'X(d\beta)$ , por serem matrizes com apenas um elemento e uma ser a transposta da outra, segue-se que

$$-2(d\beta')X'y + 2(d\beta')X'\beta \equiv 0$$

ou

$$(d\beta')(X'X\beta - X'y) \equiv 0$$

Portanto, a diferencial de  $Z$  será identicamente nula para

$$X'X\beta = X'y$$

que é o sistema de equações normais (HOFMANN, R. et al 1987).

Se  $X'X$  é não-singular, existe a matriz inversa  $(X'X)^{-1}$ . Pré-multiplicando os dois membros da equação  $X'X\beta = X'y$  por  $(X'X)^{-1}$ , obtém-se:

$$\beta = (X'X)^{-1}X'y$$

Esta expressão irá fornecer os valores estimados dos parâmetros  $\beta$ . É interessante se atentar também ao formato das matrizes  $X'X$  e  $X'y$ :

$$\mathbf{X}'\mathbf{X} = \begin{bmatrix} n & \sum X_{1j} & \sum X_{2j} & \dots & \sum X_{kj} \\ \sum X_{1j} & \sum X_{1j}^2 & \sum X_{1j}X_{2j} & \dots & \sum X_{1j}X_{kj} \\ \sum X_{2j} & \sum X_{1j}X_{2j} & \sum X_{2j}^2 & \dots & \sum X_{2j}X_{kj} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum X_{kj} & \sum X_{1j}X_{kj} & \sum X_{2j}X_{kj} & \dots & \sum X_{kj}^2 \end{bmatrix} \quad \text{e} \quad \mathbf{X}'\mathbf{y} = \begin{bmatrix} \sum Y_j \\ \sum X_{1j}Y_j \\ \sum X_{2j}Y_j \\ \vdots \\ \sum X_{kj}Y_j \end{bmatrix}$$

Porém, no caso deste trabalho, os valores da matriz  $\beta$  foram obtidos através do programa SAS, no qual os dados foram inseridos e todo este processo descrito acima foi realizado pelo próprio programa, que por sua vez forneceu as saídas com os valores de interesse.

As saídas obtidas foram as porcentagens de óbitos segundo cada uma das variáveis já citadas anteriormente (*idade, gênero, período e nível de escolaridade*), ilustrando a relação de cada uma delas com a condição de óbito ou não do indivíduo, o que representa a variável resposta que está sendo estudada.

Além disso, o programa também executa um teste de hipóteses para verificar a significância de cada um dos parâmetros avaliados, fornecendo também o resultado do teste que irá mostrar se determinado parâmetro é significativo ou não. O procedimento será explicado na sessão a seguir.

Foi estudado um total de 2432 indivíduos que tiveram seus casos notificados na cidade de Campinas-SP, entre os anos de 1980 e 2005, e para cada individuo foram inseridos os dados que foram considerados como variáveis independentes para este estudo.

## 2.2 Teste de Hipóteses

Existem diferentes tipos de testes de hipóteses baseados em diferentes tipos de distribuição dos dados obtidos, sendo que o mais indicado para esta situação é o tipo de teste que se baseia na distribuição *Qui-Quadrado*. Este tipo de teste, relacionado à inferência estatística, é indicado em situações nas quais se deseja testar se uma variável se aplica a um determinado modelo. Por isso a distribuição *Qui-Quadrado* mostrou-se a mais indicada para o presente estudo, já que o intuito do mesmo é de propor um modelo estatístico de acordo com parâmetros estimados através de uma amostra de dados.

Falando de uma maneira mais sucinta, o objetivo do teste *Qui-Quadrado* é estabelecer um procedimento para determinar se um parâmetro se aplica ou não ao modelo proposto.

No presente estudo foram consideradas quatro variáveis (*idade, gênero, período e nível de escolaridade*), mas a teoria a seguir será desenvolvida de forma que se deve aplicá-la a uma variável por vez. Então, para esta explicação, será utilizada uma variável genérica X.

Considerando a variável X, da qual se tem uma amostra de valores, o intuito é verificar sua adequação ao modelo proposto. Os valores observados devem ser divididos em  $k$  categorias (neste caso as categorias são o óbito e não óbito do indivíduo, com  $k=2$ ), sendo que as mesmas e a frequência observada para cada uma das categorias devem ser apresentadas em tabelas, como nas tabelas mais adiante na sessão de dados e resultados. A frequência de cada variável deve estar relacionada às categorias. Depois, devem-se calcular as frequências esperadas de cada variável para cada categoria e então é feita uma comparação entre as frequências esperadas e as frequências observadas.

A partir daí é então aplicado o chamado *Teste de Aderência*, que vai criar o critério para decidir se o modelo em questão será aceito ou não. Em outras palavras, se os dados amostrais *aderem* ou não ao modelo.

As hipóteses do teste são:

$H_0$ : a variável X segue o modelo proposto;

$H_a$ : a variável X não segue o modelo.

A quantidade que vai dizer se o modelo é adequado ou não é baseada na diferença entre os valores esperados sob  $H_0$  e aqueles observados na amostra. Pode-se dizer que a diferença  $o_i - e_i$  dá uma idéia da compatibilidade entre os valores observados e o modelo proposto. Assim, se as diferenças forem muito grandes, é razoável admitir que o modelo não seja adequado. Por outro lado pequenas diferenças podem ser aceitas, pois sempre se está sujeito a flutuações ao trabalhar com variáveis aleatórias (MAGALHÃES, M. N. et al, 2004).

Baseando-se nessa idéia, a quantidade utilizada no teste será:

$$Q^2 = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i}$$

Sendo que  $k$  representa o número de categorias,  $o_i$  a frequência observada e  $e_i$  a frequência esperada para a categoria  $i$ . O termo  $o_i - e_i$  indica a diferença, para a categoria  $i$ , entre a frequência observada e a esperada, ou seja, o desvio em relação ao modelo proposto. É interessante notar que é tomado o quadrado dos desvios para que a soma total não se anule, visto que o total de dados é o mesmo. Para evitar isso, é tomado o quadrado dos desvios.

Feito isso para cada uma das variáveis, a decisão sobre a *aderência* ou não dos ao modelo é tomada baseando-se no valor obtido de  $Q^2$ . Considerando o tamanho da amostra como sendo grande, a distribuição de  $Q^2$  pode ser aproximada pela *Qui-Quadrado* (Figura 2), com número de graus de liberdade ( $k-1$ ) que depende de quantas categorias serão estabelecidas.

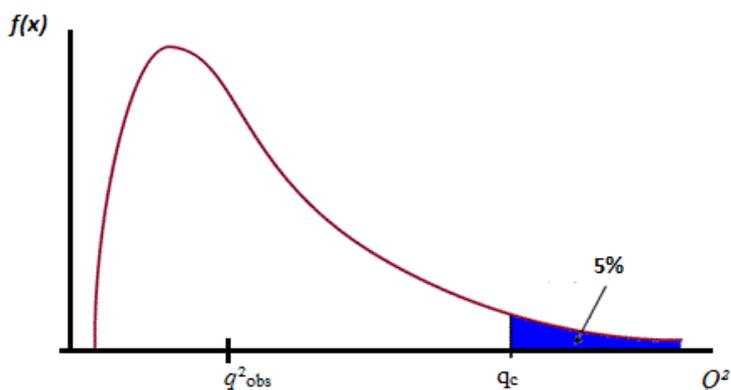


Figura 3: Modelo de distribuição *Qui-Quadrado*, mostrando o valor crítico  $q_c$  e o nível de significância de 5%.

A região crítica é constituída de valores grandes de  $Q^2$  com  $q_c$  (valor crítico) sendo determinado pelo nível de significância do teste (definido como sendo 5% no presente caso), ou seja:

$$\alpha = P(Q^2 \geq q_c | H_0 \text{ verdadeira}).$$

Escolhendo  $\alpha = 5\%$  é então possível determinar o valor de  $q_c$  utilizando a distribuição *Qui-Quadrado* com  $k$  graus de liberdade:

$$\alpha = P(Q^2 \geq q_c | H_0) = \alpha \Rightarrow P(Q^2 \geq q_c | H_0) = 0,05.$$

Feito isso, consulta-se a tabela de distribuição *Qui-Quadrado* na linha correspondente a  $k$  graus de liberdade e na coluna referente a 5% para se obter o valor de  $q_c$ .

Então, depois de obtido o valor crítico  $q_c$  é possível definir se determinada variável adere ou não a um dado modelo. Se o valor  $q_{obs}^2$ , que corresponde ao valor de  $Q^2$  observado, for menor que o valor  $q_c$ , então a hipótese  $H_0$  é considerada verdadeira, caso contrário não.

### 3. Resultados e Discussão

Numerosos estudos acerca da dinâmica da epidemia da AIDS levantam a hipótese de a expansão da epidemia no Brasil vir se fazendo acompanhar de mudanças nas condições sociais dos indivíduos com HIV/AIDS, com disseminação que afeta progressiva e mais profundamente as classes menos favorecidas (FONSECA et al., 2002). A seguir, serão apresentados os resultados obtidos para que esse processo seja ilustrado de uma maneira prática na cidade de Campinas, o que pode ser considerado como um reflexo do que acontece em todo o país.

No estudo da variável *gênero* relacionada com o óbito do indivíduo, foram obtidos os seguintes dados, que podem ser observados na Tabela 1: 715 indivíduos (29,4%) dos casos registrados foram do sexo feminino, sendo que destes, 549 indivíduos (22,49%) permaneceram vivos e 168 indivíduos (6,91%) morreram. Do sexo masculino foram registrados 1717 indivíduos (70,6%) dos casos, dos quais 1078 (44,33%) não vieram a óbito e 639 indivíduos (26,27%) faleceram. Portanto, 1625 indivíduos (66,82%) não vieram a óbito e 807 indivíduos (33,18%) faleceram.

Além disso, o valor obtido de  $q_c$  para a variável *gênero* foi de 0.7289 e o valor de  $q_{obs}^2$  foi de 0.3933, mostrando que a variável tem significância no óbito do indivíduo.

Nota-se também que, dos indivíduos que morreram 20,82% foram do sexo feminino e 79,18% foram do sexo masculino, e dos que não morreram 33,66% eram mulheres e 66,34% eram do sexo masculino. Além disso, observa-se que, considerando apenas os 715 indivíduos do sexo feminino, 76,5% permaneceram vivos e 23,5% faleceram. Já com relação aos 1717 indivíduos do sexo masculino 62,78% dos indivíduos permaneceram vivos e 37,22% dos casos registrados foram de óbito.

Portanto, percebe-se que a maioria dos casos registrados de HIV na cidade de Campinas entre 1980 e 2005 é creditada aos indivíduos do sexo masculino (que representam 70,6% do total). Porém, nota-se um processo de feminização da epidemia de HIV/AIDS em Campinas, assim como em todo o Brasil, como poderá ser observado nos dados da Tabela 2, a seguir. As possíveis causas deste processo serão também discutidas mais a frente.

Tabela 1: Saídas obtidas relacionando-se a variável *gênero* com a variável *óbito*.

Frequency			
Percent			
Row Pct			
Col Pct	F	M	Total
0	547	1078	1625
	22.49	44.33	66.82
	33.66	66.34	
	76.50	62.78	
1	168	639	807
	6.91	26.27	33.18
	20.82	79.18	
	23.50	37.22	
Total	715	1717	2432
	29.40	70.60	100.00

Entre os usuários de drogas e homossexuais masculinos, para os quais se dirigiram o grosso das campanhas de prevenção inicialmente, a epidemia está estabilizada. Nesses grupos, a taxa de transmissão do HIV diminui a cada ano. Porém, no universo heterossexual ela vem crescendo nos últimos anos. Para se ter uma idéia, as últimas estatísticas do ano de 1999 mostraram um aumento da ordem de 34% em relação ao ano de 1998. E a contaminação feminina contribuiu em muito para fermentar esse número (POLES, C., 2000).

Um dos dados mais estarrecedores divulgados pelo Ministério da Saúde é que as brasileiras contaminadas estão distribuídas em proporções equivalentes por todas as classes sociais e graus de instrução. A quantidade de novas pacientes com segundo e terceiro graus completos cresce, em média, 25% ao ano. Já entre os homens com o mesmo nível de escolaridade, apenas 5% (POLES, C., 2000).

É interessante ressaltar que ocorre um aumento importante do número de casos de AIDS em mulheres a partir de 1996, como pode ser observado na Tabela 2, que mostra a variável *período* relacionada com o *gênero* do indivíduo. A tabela mostra que realmente está havendo um processo de feminização da epidemia de HIV/AIDS, visto que no primeiro período definido,

que vai de 1980 até 1995 foram 100 indivíduos (4,11%) do sexo feminino de um total de 559 indivíduos infectados, e no segundo período, que vai de 1996 até 2005, foram 615 indivíduos (25,29%) do sexo feminino de um total de 1873 indivíduos infectados.

Além disso, observa-se também que apenas para os indivíduos do sexo feminino, no período de 1980 até 1995 foram 13,99% infectados, enquanto que os 86,01% restantes foram infectados no período que vai de 1996 até 2005, para um total de 715 indivíduos.

Pode-se notar também que no primeiro período a razão homem/mulher foi de 4,59 homens para uma mulher, enquanto que no segundo período essa razão foi de 2,06 homens para cada mulher, ou seja, aproximadamente 2,3 vezes mais.

Tabela 2: Saídas obtidas relacionando-se as variáveis *gênero* e *período*.

Frequency			
Percent			
Row Pct			
Col Pct	95	2005	Total
F	100	615	715
	4.11	25.29	29.40
	13.99	86.01	
	17.89	32.84	
M	459	1258	1717
	18.87	51.73	70.60
	26.73	73.27	
	82.11	67.16	
Total	559	1873	2432
	22.99	77.01	100.00

A tendência desse aumento no número de mulheres infectadas se deve provavelmente à maior vulnerabilidade feminina em relação à menor capacidade de negociar *sexo seguro* e, em determinados níveis sociais, à menor possibilidade de acesso aos serviços de saúde reprodutiva (RODRIGUES-JUNIOR, A. L. et al, 2004).

“A verdade é que, apesar de todas as conquistas do sexo feminino, até hoje muitas mulheres não se sentem seguras o bastante para pedir que o parceiro use preservativo. ‘A submissão feminina nesse ponto ainda está presente nas diversas classes sociais’, afirma a infectologista Rosana Del Bianco, médica do Hospital Emílio Ribas, em São Paulo, centro de referência para tratamento da doença” (POLES, C., 2000).

Na atualidade, a crescente feminização da AIDS é um aspecto grave, que precisa ser tratado com mais atenção, uma vez que, se em 1985 a proporção era de 26,5 homens para cada mulher com AIDS, no ano de 2005, a razão era de 1,5 homens para cada mulher (Brasil, 2006).

Segundo o filósofo Sócrates Nolasco, algumas mulheres de grandes centros urbanos associam sua ascensão social ou símbolos de poder, como carros ou cargos importantes, ao seu “lado masculino”. Por outro lado, os homens identificam suas necessidades afetivas como uma característica feminina. Percebe-se que esses discursos mantêm a concepção de um padrão estabelecido do que é ser homem e mulher na atualidade. Ou seja, apesar das várias conquistas femininas ocorridas, ainda se estabelecem fortes relações desiguais de gênero, sendo a falta de autonomia da mulher para negociar o uso do preservativo com o parceiro um exemplo disso (CHAVES LIMA M. A et al. 2008).

Por r outro lado, olhando de uma maneira global, nota-se que houve um aumento no número de casos a partir de 1996, porem houve também uma diminuição no óbito dos indivíduos infectados, como pode ser observado na Tabela 3 a seguir, que relaciona a variável *período* com o *óbito*. Percebe-se que no primeiro período que vai de 1980 até 1995 foram notificados 559 casos, e destes 458 indivíduos (81,93%) vieram a óbito e 101 indivíduos (18,07%) não morreram. Ou seja, as chances de morte eram aproximadamente 4,5 maiores do que de sobrevivência.

Porém, olhando para o segundo período de estudo, que vai de 1996 até 2005 nota-se uma inversão neste cenário. Dos 1873 indivíduos infectados, apenas 349 (18,63%) vieram a óbito, enquanto que 1524 indivíduos (81,37%) não morreram, mostrando que as chances de sobrevivência em relação à morte passaram a ser de aproximadamente 4,4 (isto é, o indivíduo tem 4,4 vezes mais chances de sobreviver).

Esses dados podem ser considerados como uma evidência da significativa melhora do tratamento disponibilizado pelo SUS no Brasil, levando em conta o fato de que no ano de 1996 foi promulgada a lei que garante a todos os indivíduos notificados com HIV o direito de receber sem qualquer custo todos os medicamentos necessários ao seu tratamento.

A variável *período* foi considerada significativa ao modelo, pois o valor obtido para  $q_c$  foi de 2.5584 e o valor de  $q_{obs}^2$  foi de 0.1097.

Tabela 3: Saídas obtidas relacionando-se a variável *período* com o *óbito*.

	95	2005	Total
0	101 4.15 6.22 18.07	1524 62.66 93.78 81.37	1625 66.82
1	458 18.83 56.75 81.93	349 14.35 43.25 18.63	807 33.18
Total	559 22.99	1873 77.01	2432 100.00

Segundo os dados obtidos para a variável *escolaridade* relacionada ao *óbito*, apresentados na Tabela 4 a seguir, 79 indivíduos (3,29%) apresentaram nenhum grau de escolaridade (Grupo 1). Destes, 1,87% não foram a óbito e, portanto, 1,42% apresentaram a condição de óbito.

Com relação aos casos de escolaridade entre um a três anos completos, correspondentes ao Grupo 2, 772 indivíduos (32,13%) apresentaram esta característica, sendo que 14,48% permaneceram vivos e 17,65% vieram a óbito.

Dos 637 indivíduos (26,51%) que representam escolaridade entre quatro a sete anos, do Grupo 3, 22,18% não apresentaram óbito, e 4,33% apresentaram. Do grupo de indivíduos com escolaridade entre oito e onze anos completos, representado pelo Grupo 4, que contém 610 indivíduos (25,38%), 19,02% não apresentaram óbito e, portanto, 6,37% vieram a falecer.

O grupo de indivíduos que apresentou maior escolaridade, correspondente ao Grupo 5, de doze ou mais anos completos de estudo, contém 305 indivíduos (12,69%) , sendo 8,95% destes representando a parcela de não óbito e 3,75% representando a parte dos indivíduos que

vieram a óbito. Foi ignorada a escolaridade de 29 indivíduos por falta de informação sobre os mesmos.

Pode-se concluir, então, que no decorrer de todo o período de 1980 até 2005 o grupo que representa os indivíduos com escolaridade entre um e três anos completos apresentou o maior numero de infecções pelo HIV, com 32,13% do total de casos estudados.

É interessante ressaltar também que a variável *escolaridade* foi considerada significativa para o modelo, pois o valor de  $q_{obs}^2$  foi de 0.3493, enquanto que o valor de  $q_c$  foi de 3.2878.

Tabela 4: Saídas obtidas relacionando-se a variável *escolaridade* com o *óbito*.

Frequency						
Percent						
Row Pct						
Col Pct	1	2	3	4	5	Total
0	45	348	533	457	215	1598
	1.87	14.48	22.18	19.02	8.95	66.50
	2.82	21.78	33.35	28.60	13.45	
	56.96	45.08	83.67	74.92	70.49	
1	34	424	104	153	90	805
	1.41	17.64	4.33	6.37	3.75	33.50
	4.22	52.67	12.92	19.01	11.18	
	43.04	54.92	16.33	25.08	29.51	
Total	79	772	637	610	305	2403
	3.29	32.13	26.51	25.38	12.69	100.00

As propostas de análise da epidemia de AIDS no Brasil que tentam incorporar a escolaridade até hoje desenvolvidas se viram às voltas com as dificuldades decorrentes da grande proporção de casos com escolaridade ignorada.

Pôde-se perceber também que o risco a que estão expostos os indivíduos do sexo masculino de maior escolaridade ainda é superior àquele observado nos indivíduos de menor escolaridade. Já para o sexo feminino, o quadro é claramente distinto, com maior risco entre as mulheres de menor escolaridade (FONSECA, M. G. et al, 2000).

“Ainda que reconhecendo a necessidade de utilização conjunta de indicadores diversos para uma mensuração mais precisa do *status* sócio-econômico dos casos de AIDS, cabe ponderar que a escolaridade

está menos sujeita às influências da evolução da própria infecção sobre o acesso ao emprego ou proporção dos gastos com saúde no conjunto da renda individual e familiar” (KASS N.E. et al., 1994).

Além disso, existe uma ligação cronológica bastante notável entre a epidemia de AIDS com as transformações do capitalismo em geral e do modelo econômico brasileiro em particular. Esta ligação pode ser observada através do processo contínuo de reestruturação no mercado de trabalho nacional e suas repercussões sobre a estrutura ocupacional do país (FONSECA, M. G. et al. 2000).

Vários estudos têm apontado que a posição do indivíduo na estrutura social constitui um importante indicador das suas condições de saúde, sendo que o padrão de risco observado é constantemente desvantajoso para os indivíduos pertencentes aos grupos sociais menos privilegiados.

Dentre os indicadores mais importantes para mensurar o nível sócio-econômico associado à saúde da população, citam-se o nível de instrução, a renda e a ocupação. Essas variáveis têm em comum o fato de evidenciarem a estratificação social: os indivíduos ocupam uma posição na hierarquia social de acordo com a sua ocupação, renda e nível educacional, sendo que cada um desses indicadores cobre um aspecto diferente da estratificação social (FONSECA, M. G. et al 2000).

O nível educacional expressa diferenças entre pessoas em termos de acesso à informação e perspectivas e possibilidades de se beneficiar de novos conhecimentos.

Isso mostra uma importante relação da eventual infecção por HIV com o grau de escolaridade dos indivíduos, já que na maioria dos casos, uma maior escolaridade significa também um grau de certeza de um nível mais alto a ser ocupado pelo indivíduo na sociedade do ponto de vista socioeconômico e conseqüentemente um maior acesso a informação e uma maior instrução sobre prevenção e tratamento com relação ao HIV.

Os autores estabelecem ainda que, embora seja preferível examinar todos os indicadores simultaneamente em vez de apenas um de maneira isolada, o nível educacional é considerado como o melhor, por ser este um indicador fácil de ser obtido e ter uma grande importância como determinante de saúde.

A AIDS aparece em todas as camadas sociais. Ao se observar a Tabela 5, que relaciona a variável *período* com a *escolaridade*, nota-se que houve um aumento em todos os níveis de escolaridade. Porém, nos níveis de maior escolaridade o aumento foi mais considerável, visto que no Grupo 3, considerando primeiro período de estudo, que vai de 1980 até 1995, foram

19,98% e no segundo período, que vai de 1996 até 2005 está contido todo o restante dos casos que representa 80,82%, por volta de quatro vezes mais.

Tabela 5: Saídas obtidas relacionando-se as variáveis *escolaridade* e *período*.

Frequency					
Percent					
Row Pct					
Col Pct	1	2	3	4	Total
95	362	4	117	76	559
	14.88	0.16	4.81	3.13	22.99
	64.76	0.72	20.93	13.60	
	41.14	0.63	19.18	24.92	
2005	518	633	493	229	1873
	21.30	26.03	20.27	9.42	77.01
	27.66	33.80	26.32	12.23	
	58.86	99.37	80.82	75.08	
Total	880	637	610	305	2432
	36.18	26.19	25.08	12.54	100.00

Dos casos do Grupo 4, foram 24,92% no primeiro período de estudo e 75,08% no segundo, ou seja, aproximadamente três vezes mais. Isso pode ser um reflexo, na cidade de Campinas, de uma provável lacuna na educação sexual dos brasileiros em geral, com relação à prevenção e transmissão do HIV.

Além disso, constatação de que o número de casos aumentou nos estratos de menor escolaridade remetem à condição de pior cobertura dos sistemas de vigilância e de assistência médica, entre os menos favorecidos economicamente, sob a hipótese de que a escolaridade é uma variável *proxis* importante de estratificação social (RODRIGUES-JUNIOR, A. L. et al 2004).

Um fato interessante de se considerar é que, embora todas as categorias de transmissão apareçam entre as diversas camadas sociais, o uso de drogas injetáveis é proporcionalmente mais importante entre os pacientes com baixa escolaridade e ocupações menos qualificadas, e a

transmissão sexual é proporcionalmente mais importante nos indivíduos de escolaridade mais alta (SANTOS, N.J.S. et al, 2002).

Entretanto, de uma maneira geral, o número de indivíduos infectados com menor grau de instrução é maior tanto no primeiro quanto no segundo períodos de estudo. Considerando os grupos 1 e 2 como os de menor escolaridade, foram 366 indivíduos (15,04%), para o primeiro período e 1151 indivíduos (47,33%) para o segundo período. Considerando os grupos 3 e 4 como os de maior escolaridade tem-se 193 indivíduos (7,94%) para o primeiro período e 722 indivíduos (29,69%) para o segundo período de estudo.

Portanto, comparando os dois períodos de estudo, pode-se notar que o número de indivíduos infectado aumentou para os grupos que foram considerados de menor instrução.

Esta diminuição do grau de instrução dos pacientes com AIDS tem sido largamente citada como indicador de pauperização da epidemia.

Feita esta análise da epidemia de HIV/AIDS na cidade de Campinas, pôde-se notar que o perfil epidemiológico dos 2432 indivíduos, que foram notificados pelo SUS no período de 1980 até 2005, acompanhou a evolução da epidemia no país. A escolaridade tem sido utilizada como marcador da situação socioeconômica, e o aumento na proporção de casos de AIDS naqueles indivíduos com menor escolaridade tem sido um indício da pauperização da epidemia.

Pôde-se perceber também que a relação heterossexual é a forma de transmissão que mais tem contribuído para a feminização da epidemia em Campinas, assim como em todo o país. Em todo o Brasil, a incidência de casos entre os heterossexuais foi a que mostrou o maior aumento, influenciando de forma decisiva a expansão da epidemia entre as mulheres, situação que não foi diferente em Campinas.

Normas sociais negam às mulheres conhecimento sobre sexualidade e saúde reprodutiva, o que limita suas chances de terem controle sobre seu corpo e decidirem com quem e como desejam manter relações sexuais. De modo geral, elas não podem exigir o uso de preservativos ou recusar manter relações sexuais, o que pode gerar suspeita de infidelidade (Disponível em <[http://www.scielo.br/scielo.php?pid=S0104-11692005000400008&script=sci\\_arttext](http://www.scielo.br/scielo.php?pid=S0104-11692005000400008&script=sci_arttext)>. Acesso em 21 out. 2010).

Visto isso, pode-se dizer que a atual epidemia com que nos defrontamos é bastante complexa, resultado da existência de subepidemias regionais ou definidas conforme a natureza das diferentes interações sociais. Deve-se, portanto, avaliá-la sob perspectivas diversas e complementares, renovando e refinando com freqüência seus instrumentos de análise e sua estrutura conceitual (HENN, C. A., 2007).

Em virtude das inúmeras e profundas mutações nos estádios evolutivos da infecção pelo HIV, o exame detalhado das tendências da epidemia deve combinar dados provenientes dos casos de AIDS já registrados com aqueles derivados da investigação da ocorrência de novas infecções, aumentando assim a base de referências sociais sobre desenvolvimento da epidemia.

Além disso, pode-se dizer que análise integrada das notificações da doença em diferentes períodos e das condições de sua manifestação serve tanto para as populações sob maior risco de exposição ao HIV quanto à população geral, que hoje se apresenta muito mais vulnerável do que nos primórdios da epidemia. Deve-se também ampliar as tradicionais categorias de análise, incorporando ao atual conhecimento epidemiológico conceitos provenientes da Sociologia e da Antropologia, do desenvolvimento dos métodos quantitativos e, ainda, da correlação entre distintas abordagens sobre o presente tema, tendo em vista a complexidade da situação enfrentada atualmente (CASTILHO E. A. et al, 2000).

Talvez, uma das mais importantes e eficientes contribuições para a prevenção contra a AIDS seria por meio de uma “educação sexual”. Educação não no sentido de normatização, de imposição de uma regra, mas sim no sentido de informação eficaz, que não apenas ensine a usar o preservativo (como freqüentemente se faz demonstrando em uma banana), mas se discuta e dê atenção às manifestações singulares da sexualidade. Esta seria uma maneira de tornar a sexualidade familiar para as pessoas, pois dessa forma elas estariam mais preparadas ao se defrontar com situações de exposição ao risco de contaminação da AIDS (CHAVES LIMA M. A. et al, 2008).

Apesar de todos os avanços conseguidos durante mais de vinte anos de epidemia, em termos de tratamento, melhora da qualidade de vida e prognóstico, não se pode esquecer que a AIDS continua sendo uma doença incurável. A descoberta tardia em relação a ser soropositivo, além de piorar o prognóstico, causou e continua causando danos irreversíveis em termos de não-prevenção, na medida em que o indivíduo infectado permanece longos anos transmitindo o HIV sem estar ciente de sua situação, expondo a risco um número considerável de pessoas. Nesse contexto, a informação e a prevenção da infecção permanecem essenciais. Em relação a interação de profissionais da saúde com indivíduos infectados, o bem de assistir os portadores HIV/AIDS exige enfermeiros e médicos críticos, com competência técnica e conhecimento da política de saúde para lutar por um modelo de sociedade que assegure os direitos dos cidadãos, independente de sexo, cor ou raça.

#### 4. Referências Bibliográficas

Brasil. Ministério da Saúde. (2006). *Boletim epidemiológico: AIDS e DST*, 3 (1), 01ª - 26ª Semanas epidemiológicas. Disponível em < <http://www2.aids.gov.br/data/Pages/LUMIS9A49113DPTBRIE.htm>>. Acesso em 17 set. 2010.

CASTILHO E. A. et al. *A AIDS no Brasil: uma epidemia em mutação*. Cad. Saúde Pública, Rio de Janeiro, 16(Sup. 1):4-5, 2000.

CASTILHO E, CHEQUER P, STRUCHINER C. *A Epidemiologia da AIDS no Brasil*. In: PARKER, R.; BASTOS, C.; GALVÃO, J. e PEDROSA, J. S. (orgs.) - *A AIDS no Brasil (1982 – 1992)*. Rio de Janeiro: Relume Dumará, ABIA: IMS / UERJ, 1994.

CHAVES LIMA M. A.; GUEDES MOREIRA A. C. AIDS e feminização: os contornos da sexualidade. *Revista Mal-estar e Subjetividade*, Fortaleza, Vol. 8, Nº 1, p. 103-118 – mar. 2008. Disponível em < [http://www.unifor.br/joomla/images/pdfs/pdfs\\_notitia/2045.pdf](http://www.unifor.br/joomla/images/pdfs/pdfs_notitia/2045.pdf) >. Acesso em 21 out. 2010.

DIAS P.R.T.P., NOBRE F.F. *Análise dos padrões de difusão espacial dos casos de AIDS por estados brasileiros*. CAD. Saúde Pública; set - out 2001. 17 (5):1173-87.

FONSECA, M. G. et al. *AIDS e grau de escolaridade no Brasil: evolução temporal de 1986 a 1996*. Cad. Saúde Pública, Rio de Janeiro, 16(Sup. 1), p77-87, 2000. Disponível em < <http://www2.aids.gov.br/main.asp?View={E62A8511-7150-4615-9BFA-10FDC4F4E642}&Team=&params=itemID={3F515F13-0892-47A4-ADA9-5F6C27CC7B8E};&UIPartUID={D90F22DB-05D4-4644-A8F2-FAD4803C8898}> >. Acesso em: 12 out. 2010.

HENN, C. A. *Expressão geográfica da Aids no Município de Campinas –SP*. Duas décadas de Epidemia AIDS. Campinas, SP. [s.n.], 2007. Dissertação (Mestrado) Universidade Estadual de Campinas. Faculdade de Ciências Médicas. Campinas, 27-02-2007.

HOFMANN, R.; VIEIRA, S. *Análise de Regressão, uma Introdução à Economia*. 2ª edição, Huitec, 1987, 398p.

KASS, N. E.; MUÑOZ, A. et al. Changes in employment, insurance, and income in relation to HIV status and disease progression. The Multicenter AIDS Cohort Study. *Journal of Acquired Immune Deficiency Syndromes*, v.7, p86-91, 1994. Disponível em <[http://www.scielo.br/scielo.php?pid=S0102-311X2000000700007&script=sci\\_arttext](http://www.scielo.br/scielo.php?pid=S0102-311X2000000700007&script=sci_arttext)>. Acesso em: 20 set. 2010.

KLEINBAUM, D. G.; KLEIN, M. *Logistic Regression, a Self-Learning Text*, 3ª edição, Londres-ENG, Springer, 2002, 701p.

MAGALHÃES, M. N., DE LIMA, A. C. P. *Noções de Probabilidade e Estatística*. 6ª Edição, Editora USP, 2004.

POLES, C. Mulheres em risco. *Revista Veja*, edição 1665, 6 set 2000. Disponível em <[http://veja.abril.com.br/060900/p\\_072.html](http://veja.abril.com.br/060900/p_072.html)>. Acesso em 21 out. 2010.

RACHID M, SCHECHTER M. *Manual de HIV/AIDS*. 7ª edição. Rio de Janeiro: Editora Revinter, 2003.

RODRIGUES-JUNIOR, A. L.; CASTILHO, E. A. *A epidemia de AIDS no Brasil, 1991-2000: descrição espaço-temporal*. Revista da Sociedade Brasileira de Medicina Tropical 37(4), p.312-317, jul-ago, 2004.

SANTOS, N.J.S. et al. *A AIDS no Estado de São Paulo*. As mudanças no perfil da epidemia e perspectivas da vigilância epidemiológica, Revista Brasileira de Epidemiologia. Vol. 5, Nº 3, p. 286-310, 2002.