



Universidade de Brasília

IE – Departamento de Estatística

**IDOSOS RESPONSÁVEIS PELOS DOMICÍLIOS NA  
ÁREA METROPOLITANA DE BRASÍLIA (AMB):  
UMA ANÁLISE UTILIZANDO DADOS DO CENSO  
2010**

**Bruna Rejane Freitas Alves**

Brasília

Dezembro de 2013

**Bruna Rejane Freitas Alves**

**IDOSOS RESPONSÁVEIS PELOS DOMICÍLIOS NA  
ÁREA METROPOLITANA DE BRASÍLIA (AMB):  
UMA ANÁLISE UTILIZANDO DADOS DO CENSO  
2010**

Monografia apresentada junto ao Departamento de Estatística da Universidade de Brasília, como requisito parcial à obtenção do título de Bacharel em Estatística.

Orientadora: Prof<sup>ª</sup> Dra. Marília Miranda Forte Gomes

Co-orientadora: Prof<sup>ª</sup> Dra. Ana Maria Nogales Vasconcelos

Brasília

Dezembro de 2013

**Bruna Rejane Freitas Alves**

**IDOSOS RESPONSÁVEIS PELOS DOMICÍLIOS NA  
ÁREA METROPOLITANA DE BRASÍLIA (AMB):  
UMA ANÁLISE UTILIZANDO DADOS DO CENSO  
2010**

Monografia apresentada junto ao Departamento de Estatística da Universidade de Brasília, como requisito parcial à obtenção do título de Bacharel em Estatística.

Banca Examinadora

---

Professora Dra. Marília Miranda Forte Gomes

---

Professora Dra. Ana Maria Nogales Vasconcelos

---

Professor PhD Lúcio José Vivaldi

Brasília

Dezembro de 2013

# DEDICATÓRIA

Aos meus familiares e amigos.

# AGRADECIMENTOS

Agradeço primeiramente à Deus, o detentor de toda a sabedoria, pela dádiva da vida e por sempre me fortalecer nos momentos difíceis.

Agradeço aos meus avós, pelo exemplo de vida e por fazerem dos meus sonhos o sonho deles e dos meus objetivos o objetivo deles.

Agradeço à minha mãe, que sempre me incentivou na continuação do curso, e que hoje sorri orgulhosa e comemora junto a mim essa conquista que é nossa. Muito obrigada por abdicar dos seus desejos e metas para se doar completamente aos meus cuidados. Você sempre cumpriu e cumpre com excelência o dom divino de ser mãe.

Agradeço aos meus tios pelo apoio, preocupação, cuidado e amor. Agradeço em especial à minha tia Josiane, exemplo de professora em minha vida, que sempre me mostrou que o estudo é a maior riqueza que um ser humano pode ter.

Agradeço aos meus “amigos irmãos” Géssica Carla, Marcos Vinnicius, Aline Cristina, Cristina Dias, Eliane Soares, Lucas Mendonça, Jéssica Lillian e Thaís Bianca por estarem ao meu lado nessa longa caminhada desde o primeiro passo: o resultado do vestibular. Obrigada pelas inúmeras alegrias e pelo respeito nos momentos tristes. Vocês foram essenciais nessa jornada.

Agradeço aos meus colegas de curso Amanda Ferraz, Caio Felipe, Camyla Serpa, Fátima Lira, Jaqueline Lopes e todos aqueles com quem passei horas estudando para as provas, conversando e me divertindo. Obrigada pelas caronas, lanches e reuniões de comemoração de fim de semestre, todos esses momentos foram únicos e estão guardados pra sempre em minha memória e coração.

Agradeço às minhas amigas Lívia Sant’Anna e Tatiana Santos pelo companheirismo e amizade ao longo desses intermináveis semestres. Obrigada por serem minhas professoras particulares e por não me deixarem desanimar ou desistir do curso.

Agradeço à minha orientadora Professora Marília Miranda, por aceitar fazer parte desse momento tão importante em minha vida. Obrigada por dedicar seu tempo e compartilhar seu conhecimento e experiência para a conclusão dessa monografia. Serei eternamente grata.

Agradeço a todos que de forma direta ou indiretamente fizeram parte da realização desse sonho.

“Nascer é uma possibilidade.  
Viver é um risco.  
Envelhecer será sempre um privilégio.”

Autor desconhecido

# RESUMO

O envelhecimento populacional é um fenômeno mundial e tem ocorrido de forma mais rápida entre os países em desenvolvimento. O intenso processo de redução dos níveis de fecundidade, combinado com a queda da mortalidade, vêm determinando mudanças significativas na estrutura etária dos países, caracterizada por um aumento progressivo e acentuado da população adulta e principalmente idosa (Carvalho & Garcia, 2003; Camarano et al, 2004).

Com o tema envelhecimento cada vez mais presente nas notícias e estudos pelo mundo, o presente trabalho tem o objetivo de estudar fatores relacionados aos idosos responsáveis pelo domicílio na Área Metropolitana de Brasília (AMB).

Para alcançar o objetivo proposto foram utilizados os dados do Censo Demográfico de 2010 e estimado modelos de regressão logística.

Os resultados evidenciam que idosos do sexo masculino, com ensino superior completo, viúvos, sem deficiência mental/intelectual e com idade entre 60 e 69 anos, têm maiores chances de serem os responsáveis pelo domicílio na Área Metropolitana de Brasília. Com esse trabalho o governo poderá ter um maior conhecimento do perfil dos seus idosos e assim planejar melhor políticas públicas a fim de proporcionar melhorias nas áreas de saúde, lazer e segurança a todos de forma igual e digna. Afinal, todos envelhecem um pouco a cada dia que passa.

Palavras-chave: Envelhecimento, AMB, idoso responsável.

# ABSTRACT

Population aging is a global phenomenon and has occurred more quickly among developing countries. The intense process of reduction of fertility levels, combined with the decrease in mortality, have caused significant changes in the age structure of countries, characterized by an increase in the adult population and elderly (Carvalho & Garcia, 2003; Camarano et al., 2004). This paper aims to study factors related to elderly household heads in Brasília Metropolitan Area (AMB). We used data from the 2010 Census and estimated logistic regression models. The results show that elderly males, with higher education, widowers, without mental / intellectual disabilities and aged between 60 and 69 years are more likely to be the head of household in the AMB. With this work the government may have a greater knowledge of the profile of their elderly population and thus plan better public policies to provide improvements in health, leisure and security to all equally and dignified. After all, all ages a bit with each passing day.

Keywords: population aging, Brasília Metropolitan Area (AMB), household heads.



# LISTA DE TABELAS E ILUSTRAÇÕES

Figura 1 Representação do processo de regressão.....	9
Figura 2 Pirâmide etária da AMB de 1970.....	21
Figura 3 Pirâmide etária da AMB de 1980.....	21
Figura 4 Pirâmide etária da AMB de 1991.....	21
Figura 5 Pirâmide etária da AMB de 2000.....	22
Figura 6 Pirâmide etária da AMB de 2010.....	22
Tabela 1 Distribuição dos idosos responsáveis pelo domicílio por faixa etária, sexo, cor ou raça, nível de instrução, estado civil e região da AMB.....	23
Tabela 2 Distribuição dos idosos responsáveis pelo domicílio por variáveis de condições de saúde.....	25
Tabela 3 Categorias e codificação das variáveis que entraram no modelo .....	26
Tabela 4 Estimativas e inferências das variáveis que entraram no modelo.....	27
Tabela 5 Relação de acertos entre valores estimados e preditos nos dados da modelagem	29
Tabela 6 Composição da AMB durante os anos de 1970 a 2010 .....	34

# **LISTA DE SIGLAS E ABREVIATURAS**

AMB - Área Metropolitana de Brasília

IBGE - Instituto Brasileiro de Geografia e Estatística

LISA - Análise de Interdependência Espacial

OMS - Organização Mundial de Saúde

ONU - Organização das Nações Unidas

PNI - Política Nacional do Idoso

TAB – Tabela

# SUMÁRIO

1.	INTRODUÇÃO E JUSTIFICATIVA.....	1
2.	OBJETIVOS .....	4
2.1.	GERAL:.....	4
2.2.	ESPECÍFICOS: .....	4
3.	MATERIAL E MÉTODOS .....	5
3.1.	FONTE DE DADOS: CENSO DEMOGRÁFICO 2010 .....	5
3.1.1.	O QUE É O CENSO .....	5
3.1.2.	METODOLOGIA DA COLETA DE DADOS DO CENSO.....	7
3.2.	VARIÁVEIS SELECIONADAS .....	8
3.3.	ANÁLISE DE REGRESSÃO .....	9
3.3.1.	REGRESSÃO LOGÍSTICA .....	10
4.	RESULTADOS.....	20
4.1.	CARACTERÍSTICAS GERAIS DOS IDOSOS RESIDENTES NA ÁREA METROPOLITANA DE BRASÍLIA .....	20
4.2.	ESTIMAÇÃO DO MODELO.....	25
4.3.	AVALIAÇÃO DO MODELO .....	28
5.	CONCLUSÕES .....	30
6.	REFERÊNCIAS BIBLIOGRÁFICAS.....	31
7.	ANEXOS .....	34

# 1. INTRODUÇÃO E JUSTIFICATIVA

O envelhecimento populacional é um fenômeno mundial e tem ocorrido de forma mais rápida entre os países em desenvolvimento. O intenso processo de redução dos níveis de fecundidade, combinado com a queda da mortalidade, vêm determinando mudanças significativas na estrutura etária dos países, caracterizada por um aumento progressivo e acentuado da população adulta e principalmente idosa<sup>1</sup> (Carvalho & Garcia, 2003; Camarano et al, 2004).

Estimativas da ONU mostram que a população idosa no mundo irá triplicar nos próximos 40 anos. A Europa é o continente com maior número de pessoas com 60 anos ou mais no mundo. Para o Japão, as mesmas estimativas indicam que em 2050 haverá tantos trabalhadores quanto idosos já aposentados (United Nations, 2011). Particularmente, o Brasil está entre as sociedades que envelhece mais rapidamente no mundo. Segundo Gomes (2010), nos últimos 70 anos, o número absoluto de pessoas com 60 anos e mais aumentou doze vezes. Dados do IBGE mostram que, em 1940 foram recenseados 1,7 milhões de idosos no Brasil e, em 2010, 20,5 milhões. A participação relativa dessa população que era de 5,1%, em 1970, passou para 8,6%, em 2000 e, segundo o último Censo, já representa 10,8% da população total. Para 2050, estima-se que essa proporção seja de aproximadamente 30% (United Nations, 2011).

Conforme Anselmi (2010), esse envelhecimento acelerado força os governos de todos os países a repensarem num modelo de sociedade ideal para essa nova população. Não só em termos sociais, culturais, mas também econômicos, já que os idosos estão se tornando maioria em muitos países e já deixaram de ser, há muito tempo, apenas os “avós” para se tornarem também provedores de muitos lares. A autora destaca ainda que os idosos têm se tornado cada vez mais economicamente ativos e socialmente atuantes.

Conhecer as características relacionadas a esses idosos, especialmente daqueles chefes dos domicílios, tem sido o foco de muitos estudos demográficos e pesquisas acadêmicas.

---

<sup>1</sup> A Política Nacional do Idoso (PNI), Lei nº 8.842, de 4 de janeiro de 1994, e o estatuto do Idoso, Lei nº 10.741, de 1º de outubro de 2003, define Idoso pessoas com 60 anos ou mais. Já a Organização Mundial da Saúde (OMS) (2002) define o idoso a partir da idade cronológica, portanto, idosa é aquela pessoa com 60 anos ou mais, em países em desenvolvimento e com 65 anos ou mais em países desenvolvidos.

Camargos *et al* (2007) fizeram um estudo sobre a relação entre renda e morar sozinho para idosos paulistanos no ano de 2000 utilizando a base de dados do Projeto Sabe (Saúde, Bem-estar e Envelhecimento na América Latina e Caribe), e concluíram que as probabilidades de o idoso morar sozinho crescem à medida que aumenta a renda.

França *et al* (2008), com o objetivo de identificar os determinantes da condição do idoso morar sozinho no estado do Rio Grande do Norte, utilizando os dados do Censo Demográfico 2000, mostraram que idosos homens, sem nenhuma das paralisias permanentes consideradas, que moram em municípios com menor população, com idades mais avançadas, branco, com maior escolaridade, que moram em domicílios com menos cômodos e que são solteiros têm maior chance de morar sozinho.

O trabalho de Longo *et al* (2012) teve como objetivo identificar uma situação de vulnerabilidade espacial considerando a concentração de idosos e a oferta dos serviços de saúde, principalmente serviços de saúde preventiva, utilizando a análise de interdependência espacial (LISA). A hipótese por trás desse estudo é que as regiões carentes de uma oferta básica de serviços de saúde podem representar um risco maior para a população idosa, principalmente para aqueles que moram sozinhos. Os dados do Censo Demográfico 2010 e de outras fontes de informação para saúde para o estado de Minas Gerais foram utilizados nesse estudo e confirmaram a hipótese de que as regiões carentes de uma oferta básica de serviços de saúde podem representar um risco maior para a população idosa, principalmente para aqueles que moram sozinhos. Nesse caso, a “independência” de morar sozinho pode ser, na verdade, um indicativo de vulnerabilidade.

Paulo *et al* (2008) investigaram a relação entre renda e arranjo domiciliar do idoso. Nesse estudo, realizou-se uma análise descritiva dos arranjos familiares dos idosos utilizando os dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) em 1985, 1995 e 2005 e investigou-se o impacto do Benefício da Prestação Continuada (BPC) sobre os arranjos domiciliares dos idosos utilizando os dados da PNAD em 2002 e 2004. A proposta do trabalho foi verificar se o recebimento do BPC impulsionaria os idosos a morarem sozinhos ou se, devido à situação de pobreza na qual se inserem esses indivíduos, a nova renda do idoso provocaria a permanência (ou mesmo maior atração) de familiares em torno desse idoso. O estudo concluiu que, para o caso brasileiro, o recebimento do BPC aumentou a probabilidade de o idoso morar sozinho.

Apesar dos estudos mencionados acima terem como foco os domicílios unipessoais, se faz muito importante estudar os novos arranjos familiares dos idosos brasileiros, sejam eles unipessoais ou não, uma vez que projeções populacionais apontam que em alguns anos a proporção de idosos irá ultrapassar a de jovens, causando assim mudanças significativas nas estruturas familiares.

Tendo em vista o processo de envelhecimento da população residente na Capital Federal, acompanhado de uma série de mudanças socioeconômicas, comportamentais e culturais (França *et al*, 2008), esse trabalho contribuirá não só para um maior conhecimento do perfil dos seus idosos, como também para que o governo possa planejar melhor políticas públicas a fim de proporcionar melhorias nas áreas de saúde, lazer e segurança a todos de forma igual e digna. Afinal, todos envelhecem um pouco a cada dia que passa.

## **2. OBJETIVOS**

### **2.1. GERAL:**

- Analisar os fatores relacionados com os idosos responsáveis pelos domicílios na Área Metropolitana de Brasília (AMB).

### **2.2. ESPECÍFICOS:**

- Descrever a população idosa da AMB, segundo variáveis demográficas, socioeconômicas e de condições de saúde;
- Estudar e aplicar o modelo de regressão logística;
- Identificar variáveis que mais contribuem para o idoso ser o responsável pelo domicílio.

### **3. MATERIAL E MÉTODOS**

Para atender os objetivos propostos foram utilizados os dados do Censo 2010. Primeiramente foi feita uma seleção de variáveis do Censo, posteriormente uma análise descritiva dos dados. Após essa etapa, foi aplicada a técnica de regressão logística binária múltipla utilizando o pacote estatístico SPSS para estudar a relação das variáveis em estudo com o fato do idoso ser o responsável pelo domicílio.

Todas as análises aqui propostas tiveram como foco a Área Metropolitana de Brasília (AMB). Segundo Vasconcelos (2010), devido à relação cotidiana estreita e intensa entre a população residente nos municípios adjacentes ao Distrito Federal com a área central da Capital Federal, torna-se necessário considerar no planejamento da cidade e nos estudos demográficos e socioeconômicos toda a área denominada Área Metropolitana de Brasília (AMB), constituída pelo DF, com todas as suas Regiões Administrativas, e dez municípios goianos limítrofes: Águas Lindas de Goiás, Cidade Ocidental, Cristalina, Formosa, Luziânia, Novo Gama, Padre Bernardo, Planaltina, Santo Antônio do Descoberto e Valparaíso de Goiás.

#### **3.1. FONTE DE DADOS: CENSO DEMOGRÁFICO 2010**

##### **3.1.1. O QUE É O CENSO**

O Censo Demográfico é uma pesquisa realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE) a cada dez anos. O primeiro Censo brasileiro em 1872 e recebeu o nome de Recenseamento da População do Império do Brasil. O mais recente foi o XII Censo Demográfico, realizado em 2010.

O Censo Demográfico representa um grande retrato em extensão e profundidade da população brasileira e das suas características socioeconômicas e, ao mesmo tempo, é a base sobre a qual deverá se assentar todo o planejamento público e privado da próxima década. A partir das informações censitárias pode-se traçar um perfil completo da população residente no Brasil e as características de seus domicílios (Hakkert, 1996).



Segundo o IBGE, entre as principais utilizações dos resultados censitários estão as de:

1. Acompanhar o crescimento, a distribuição geográfica e a evolução de outras características da população ao longo do tempo;

2. Identificar áreas de investimentos prioritários em saúde, educação, habitação, saneamento básico, transporte, energia, programas de assistência à infância e à velhice, possibilitando a avaliação e revisão da alocação de recursos público e privado;

3. Selecionar locais que necessitam de programas de estímulo ao crescimento econômico e desenvolvimento social;

4. Fornecer as referências para as projeções populacionais com base nas quais o Tribunal de Contas da União define as cotas do Fundo de Participação dos Estados e do Fundo de Participação dos Municípios;

5. Fornecer as referências para as projeções populacionais com base nas quais é definida a representação política do País: o número de deputados federal, estadual e vereadores de cada estado e município;

6. Fornecer parâmetros para conhecer e analisar o perfil da mão de obra em nível municipal, informação esta de grande importância para organizações sindical, profissional e de classe, assim como para decisões de investimentos do setor privado;

7. Fornecer parâmetros para selecionar locais para a instalação de fábricas, *shopping centers*, escolas, creches, cinemas, restaurantes, etc.;

8. Fundamentar diagnósticos e reivindicações pelos cidadãos, de maior atenção dos governos estadual ou municipal para problemas local e específico, como de insuficiência das redes de água e esgoto, de atendimento médico ou escolar, etc.; e

9. Subsidiar as comunidades acadêmicas e técnico-científicas em seus estudos e projetos.

### **3.1.2. METODOLOGIA DA COLETA DE DADOS DO CENSO**

De acordo com o IBGE, a coleta de dados para o Censo inicia-se em uma data estipulada, e essa data será a mesma em todos os censos. Serão visitados todos os domicílios do país e qualquer morador capaz de responder às perguntas do questionário responderá ao recenseador por todos os demais moradores daquele domicílio.

Devido à extensão continental do Brasil, não é possível aplicar um questionário extenso em todos os domicílios do país, pois isso envolveria custos e prazos grandes, sendo assim, a opção viável é utilizar dois tipos de questionários:

- um questionário grande, para ser aplicado em uma amostra de domicílios (e cada um de seus moradores), chamado de questionário da amostra; e
- um questionário pequeno, para ser aplicado aos domicílios (e cada um de seus moradores) não selecionados para a amostra, chamado de questionário básico.

O conjunto Amostra provém dos dados dos questionários da amostra. Esse conjunto contém características mais detalhadas da população, essas informações serão devidamente ponderadas para fornecer estimativas para a população como um todo. O plano amostral do Censo baseia-se na técnica de amostragem complexa.

Já o conjunto Universo provém dos dados dos dois questionários, uma vez que todas as perguntas do questionário básico estão contidas no questionário da amostra, de forma que essas variáveis comuns são investigadas censitariamente, ou seja, para todos os domicílios e pessoas. Isto permite a junção dos registros dos dois tipos de questionários formando o conjunto Universo, ou seja, informações básicas para o conjunto da população recenseada.

O banco de dados utilizado neste trabalho tem como base os dados oriundos do conjunto Amostra para os domicílios particulares permanentes do Distrito Federal e dos dez municípios que formam a AMB.

### 3.2. VARIÁVEIS SELECIONADAS

A principal variável deste estudo é “responsável pelo domicílio”. Ela é uma variável binária, pois pode assumir os valores “sim” (sucesso) ou “não” (fracasso), atribuindo-se os valores 1 e 0, respectivamente.

Outras variáveis deste estudo são: nível de instrução, sexo, idade, cor ou raça, estado civil, condições de saúde e regiões da AMB. Estas variáveis estão assim categorizadas:

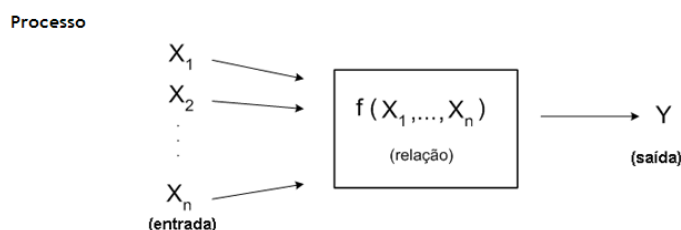
- Nível de instrução: Sem instrução e fundamental incompleto; Fundamental completo e médio incompleto; Médio completo e superior incompleto; Superior completo; e Não determinado.
- Sexo: Homem e Mulher.
- Idade: de 60 a 69 anos, de 70 a 79 anos e 80 ou mais anos.
- Cor ou raça: Branca, Preta e Parda.
- Estado civil: Casado(a); Desquitado(a), separado(a) judicialmente ou divorciado(a); Viúvo(a); e Solteiro(a).
- Variáveis de condições de saúde: (1) Dificuldades em enxergar, (2) Dificuldades em ouvir, (3) Dificuldades em caminhar e (4) Deficiência mental. As variáveis (1), (2) e (3) estão assim categorizadas: Sim (alguma dificuldade, grande dificuldade ou não consegue de modo algum), e Não (nenhuma dificuldade); e a variável (4) está categorizada em Sim e Não.
- Região da AMB: Foram considerados quatro grupos de localidades de residência segundo estratos de renda. O Distrito Federal foi dividido em três regiões: Região 1 – localidades com renda domiciliar média elevada (Plano Piloto, Sudoeste/Octogonal, Lago Norte e Lago Sul); Região 2 – localidades com renda domiciliar média intermediária (Cruzeiro, Candangolândia, Núcleo Bandeirante, Guará, Gama, Taguatinga, Águas Claras, Vicente Pires, Riacho Fundo I e São Sebastião); Região 3 – localidades com renda domiciliar média baixa (Brazlândia, Ceilândia, Itapoã, Planaltina, Santa Maria, Recanto das Emas, Riacho Fundo II,

Samambaia e áreas rurais). O quarto grupo de localidades é formado pelos dez municípios goianos, cuja renda domiciliar média é ainda inferior àquela da Região 3 do DF, segundo o Censo de 2010.

### 3.3. ANÁLISE DE REGRESSÃO

A análise de regressão é uma técnica estatística utilizada quando o interesse é verificar se duas ou mais variáveis estão relacionadas de alguma forma. Esta técnica é muito utilizada em diversos problemas nas áreas médica, biológica, industrial, química entre outras.

Esta relação pode ser analisada como um processo. Nesse processo, os valores de  $X_1, X_2, \dots, X_n$  são chamados de variáveis de entrada ou regressoras (inputs) e  $Y$  de Variável de saída ou resposta (output). (Neter et al, 1985; Myers, 1990; Montgomery & Peck, 1992).



**Figura 1 Representação do processo de regressão**

Modelos de regressão são modelos matemáticos que relacionam o comportamento de uma variável  $Y$  com outra variável  $X$ , esse tipo de modelagem ajuda a entender como determinadas variáveis influenciam outra variável.

Se o interesse está na relação de apenas uma variável de entrada com a variável resposta, o caso pode ser representado por uma regressão linear simples. Mas se o objetivo é relacionar a variável resposta com mais de uma variável regressora, a regressão linear múltipla é utilizada. (Neter et al, 1985; Myers, 1990; Montgomery & Peck, 1992).

A função de regressão linear é dada por:  $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, i = 1, 2, \dots, n$  onde  $Y_i$  é a variável dependente;  $x_i$ , a variável independente;  $\beta_0$  e  $\beta_1$  são os parâmetros

desconhecidos, chamados de coeficientes de regressão; e  $\varepsilon_i$ , o erro (resíduo) inerente à análise estatística.

Caso a variável resposta seja uma variável categórica, ou seja, a variável apresenta como possíveis realizações uma qualidade (ou atributo) e não mais uma mensuração, o modelo de regressão logística passa a ser mais apropriado. (Neter et al, 1985; Myers, 1990; Montgomery & Peck, 1992).

### 3.3.1. REGRESSÃO LOGÍSTICA

O modelo de regressão logístico é utilizado quando a variável resposta é qualitativa com dois possíveis resultados, também chamada de variável *dummy* (indicadora, dicotômica ou binária). Dessa forma, a variável será assim representada:

$Y_i = 1$ , a variável possui o atributo a ser estudado e

$Y_i = 0$ , a variável não possui o atributo a ser estudado.

Será considerado “sucesso” quando  $Y_i = 1$  e considerado “fracasso” quando  $Y_i = 0$ .

Considerando o modelo de regressão linear simples:

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \text{ onde } Y_i \in \{0,1\}.$$

Assumindo que  $E(\varepsilon_i) = 0$ , tem-se que

$$E(Y_i) = \beta_0 + \beta_1 x_i. \quad (1)$$

A variável resposta  $Y$  tem distribuição *Bernoulli*  $(1, \pi)$ , com probabilidade de sucesso  $P(Y_i = 1) = \pi_i$  e de fracasso  $P(Y_i = 0) = 1 - \pi_i$ . Desta forma

$$E(Y_i) = \pi_i \quad (2)$$

E a variância de  $Y_i$  é dada por

$$Var(Y_i) = \pi_i(1 - \pi_i)$$

Igualando (1) e (2), tem-se

$$E(Y_i) = \pi_i = \beta_0 + \beta_1 x_i.$$

Essa igualdade viola as suposições do modelo linear. De fato,

i) Os erros não são normais, pois:

$$\varepsilon_i = Y_i - \pi_i,$$

ou seja

$$Y_i = 1 \Rightarrow \varepsilon_i = 1 - \beta_0 - \beta_1 x_i$$

$$Y_i = 0 \Rightarrow \varepsilon_i = 0 - \beta_0 - \beta_1 x_i$$

Assim não faz sentido assumir a normalidade dos erros.

ii) Não homogeneidade da variância.

A variância de  $Y_i$  é dada por

$$\text{Var}(Y_i) = E[(Y_i - E(Y_i))^2] = (1 - \pi_i)^2 \pi_i + (0 - \pi_i)^2 (1 - \pi_i)$$

$$\text{Var}(Y_i) = \pi_i(1 - \pi_i) = (\beta_0 + \beta_1 x_i)(1 - \beta_0 - \beta_1 x_i),$$

então a variância de  $Y_i$  depende de  $x_i$ , e conseqüentemente, não é constante.

E a variância do erro é:

$$\text{Var}(\varepsilon_i) = \text{Var}(Y_i - \pi_i) = \text{Var}(Y_i)$$

Portanto, os resíduos são heterocedásticos.

iii) Restrição para a resposta média  $E(Y_i)$ . Como a resposta média é obtida em probabilidades temos que  $0 \leq E(Y_i) = \pi_i = \beta_0 + \beta_1 x_i \leq 1$ . Entretanto, esta restrição é inapropriada para resposta em um modelo linear, que assume valores no intervalo  $(-\infty, \infty)$ . (Neter et al, 1985; Myers, 1990; Montgomery & Peck, 1992).

Quando a variável resposta é binária, a forma da função resposta será curvilínea com o formato sigmoidal, ou seja, a função resposta tem a forma de S e tem assíntotas em 0 e 1. Uma forma de resolver esse problema é utilizar a distribuição logística, dando origem ao modelo de regressão logística.

As qualidades do modelo de regressão logística estão no fato de seus resultados poderem ser explicados em termos de probabilidades e a possível interpretação para os parâmetros estimados, além do que, sob o ponto de vista matemático, uma função de distribuição logística é flexível e fácil de ser usada (Machado, 2010).

A expressão do modelo logístico simples é dada por

$$E(Y_i) = \pi_i = \frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)} \quad (3)$$

Ou de forma equivalente

$$E(Y_i) = [1 + \exp(-\beta_0 + \beta_1 x_i)]^{-1}$$

Uma propriedade importante é que a função logística pode ser linearizada. Denotando  $E(Y_i)$  por  $\pi_i$ , pois, quando a variável resposta é binária, a resposta média é a probabilidade. Fazendo a transformação:

$$\pi_i' = \ln\left(\frac{\pi_i}{1 - \pi_i}\right) \quad (4)$$

$$\pi_i' = \beta_0 + \beta_1 x_i \quad (5)$$

Essa transformação é denominada transformação logit da probabilidade  $\pi_i$ .

A razão  $\frac{\pi_i}{1 - \pi_i}$  na transformação logit é chamada de *odds* (chance). A função resposta transformada é chamada de função resposta logit e,  $\pi_i'$  é chamada de resposta média logit. Na função resposta logit, tanto  $\pi_i'$  quanto  $x_i$  estão definidos no intervalo de  $-\infty$  a  $\infty$ .

Os parâmetros do modelo de regressão logística simples são estimados pelo método da máxima verossimilhança.

Considerando que  $P(Y_i = 1) = \pi_i$  e  $P(Y_i = 0) = 1 - \pi_i$ , e que a distribuição de probabilidade de Bernoulli é dada por:

$$f_i(Y_i) = \pi_i^{Y_i} (1 - \pi_i)^{1 - Y_i}, \quad Y_i = 0, 1; i = 1, 2, \dots, n$$

Como as observações  $Y_i$  são independentes, a densidade conjunta é dada por:

$$g(Y_1, \dots, Y_n) = \prod_{i=1}^n f_i(Y_i) = \prod_{i=1}^n \pi_i^{Y_i} (1 - \pi_i)^{1 - Y_i}$$

Aplicando o logaritmo, tem-se:

$$\ln g(Y_1, \dots, Y_n) = \sum_{i=1}^n \left[ Y_i \ln \left( \frac{\pi_i}{1 - \pi_i} \right) \right] + \sum_{i=1}^n \ln(1 - \pi_i)$$

Sabendo-se que  $E(Y_i) = \pi_i$  para uma variável binária e da expressão (3), tem-se que:

$$1 - \pi_i = [(1 + \exp(\beta_0 + \beta_1 x_i))]^{-1}$$

Além disso, considerando (4) e (5), a função de verossimilhança é dada por:

$$\ln L(\beta_0, \beta_1) = \sum_{i=1}^n Y_i (\beta_0 + \beta_1 x_i) - \sum_{i=1}^n \ln[1 + \exp(\beta_0 + \beta_1 x_i)] \quad (6)$$

No entanto, não existe uma solução analítica para os valores  $\beta_0$  e  $\beta_1$  que maximizam a função de verossimilhança (6). É preciso recorrer a métodos numéricos para encontrar as estimativas de máxima verossimilhança de  $\hat{\beta}_0$  e  $\hat{\beta}_1$ .

Encontradas as estimativas de  $\hat{\beta}_0$  e  $\hat{\beta}_1$ , devem-se substituir os valores na equação (3) para encontrar os valores ajustados. Dessa forma, o valor ajustado para o  $i$ -ésimo valor é dado por:

$$\hat{\pi}_i = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 x_i)}{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 x_i)}$$

E a função resposta ajustada é dada por:

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

Usando a transformação logit (4), a função resposta ajustada é dada por:

$$\hat{\pi}' = \hat{\beta}_0 + \hat{\beta}_1 x \quad (7)$$



onde,

$$\hat{\pi}' = \ln\left(\frac{\hat{\pi}}{1 - \hat{\pi}}\right) \quad (8)$$

Considerando o valor da função resposta ajustada em (7) em  $x = x_j$ , tem-se que

$$\hat{\pi}'(x_j) = \hat{\beta}_0 + \hat{\beta}_1 x_j \quad (9)$$

E em  $x = x_j + 1$ , tem-se que

$$\hat{\pi}'(x_j + 1) = \hat{\beta}_0 + \hat{\beta}_1(x_j + 1) \quad (10)$$

A diferença entre os dois valores é dada por:

$$\hat{\pi}'(x_j + 1) - \hat{\pi}'(x_j) = \hat{\beta}_1$$

De acordo com (8), a expressão (9) é o logaritmo da chance (*odds*) estimada quando  $x = x_j$  e é denominado por  $\ln(chance_1)$ . Da mesma forma, (10) é o logaritmo da chance (*odds*) estimada quando  $x = x_j + 1$ , e é denominada por  $\ln(chance_2)$ . Assim, a diferença é dada por:

$$\ln(chance_2) - \ln(chance_1) = \ln\left(\frac{chance_2}{chance_1}\right) = \hat{\beta}_1$$

Aplicando o anti-logaritmo em cada lado da equação, vê-se que a razão das chances estimadas, denominada de razão das chances (*odds ratio*), é dada por:

$$\widehat{OR} = \frac{chance_2}{chance_1} = \exp(\hat{\beta}_1)$$

Em geral, a razão das chances estimadas quando existe uma diferença de  $c$  unidades em  $x$ , é igual a  $\exp(c\hat{\beta}_1)$ .

## Regressão Logística Múltipla

O modelo de regressão logística múltipla, é obtido substituindo-se, na equação (3),  $\beta_0 + \beta_1 x_i$  por  $\beta_0 + \beta_1 x_1 + \dots + \beta_{p-1} x_{p-1}$ , onde,  $p-1$  = número de variáveis explicativas.

Em termos matriciais tem-se:

$$\boldsymbol{\beta}_{(px1)} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \cdot \\ \cdot \\ \beta_{p-1} \end{bmatrix}; \mathbf{X}_{(px1)} = \begin{bmatrix} 1 \\ X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_{p-1} \end{bmatrix}; \mathbf{X}_{i(p \times 1)} = \begin{bmatrix} 1 \\ X_{i1} \\ X_{i2} \\ \cdot \\ \cdot \\ X_{i,(p-1)} \end{bmatrix}$$

em que

$$\boldsymbol{\beta}'\mathbf{X} = \beta_0 + \beta_1 X_1 + \dots + \beta_{p-1} X_{p-1}$$

$$\boldsymbol{\beta}'\mathbf{X}_i = \beta_0 + \beta_1 X_{i1} + \dots + \beta_{p-1} X_{i,(p-1)}$$

A função definida pela equação (3), pode ser generalizada como:

$$E(Y) = \frac{\exp(\boldsymbol{\beta}'\mathbf{X})}{1 + \exp(\boldsymbol{\beta}'\mathbf{X})}$$

Ou de forma equivalente:

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

A transformação logit resulta em:

$$\pi' = \boldsymbol{\beta}'\mathbf{X}$$

Sejam  $Y_i$  variáveis aleatórias independentemente distribuídas segundo uma *Bernoulli* com valores esperados  $E(Y_i) = \pi_i$ , a formulação do modelo é dada por:

$$E(Y_i) = \pi_i = \frac{\exp(\boldsymbol{\beta}'\mathbf{X}_i)}{1 + \exp(\boldsymbol{\beta}'\mathbf{X}_i)}$$

As variáveis X (preditoras) podem ser quantitativas ou qualitativas representadas por variáveis indicadoras. Essa flexibilidade torna o modelo de regressão logística múltiplo bastante utilizado em análises estatísticas.

A função log-verossimilhança expressa em (6), pode ser estendida para o modelo múltiplo, logo, tem-se que

$$\ln L(\boldsymbol{\beta}) = \sum_{i=1}^n Y_i (\boldsymbol{\beta}' \mathbf{X}_i) - \sum_{i=1}^n \ln[1 + \exp(\boldsymbol{\beta}' \mathbf{X}_i)] \quad (11)$$

Novamente, métodos numéricos devem ser utilizados para encontrar os valores de  $\beta_0 + \beta_1 X_1 + \dots + \beta_{p-1} X_{p-1}$  que maximizam (11). As estimativas de máxima verossimilhança serão denotadas por  $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_{p-1}$ .

A função resposta logística ajustada e os valores ajustados são dados por:

$$\hat{\pi} = \frac{\exp(\hat{\boldsymbol{\beta}}' \mathbf{X})}{1 + \exp(\hat{\boldsymbol{\beta}}' \mathbf{X})} = [1 + \exp(-\hat{\boldsymbol{\beta}}' \mathbf{X})]^{-1}$$

$$\hat{\pi}_i = \frac{\exp(\hat{\boldsymbol{\beta}}' \mathbf{X}_i)}{1 + \exp(\hat{\boldsymbol{\beta}}' \mathbf{X}_i)} = [1 + \exp(-\hat{\boldsymbol{\beta}}' \mathbf{X}_i)]^{-1}$$

### Inferências para o modelo logístico

É interesse em análise de regressão fazer inferências sobre os parâmetros estimados e suas relações com os valores reais, a fim de se extrair informações a cerca do comportamento da variável de resposta na população e, também, para testar se realmente as variáveis explicativas que estão no modelo são significativas. Além do que, procura-se fazer testes para comprovar a adequabilidade do modelo aos dados. (Pereira, 2006)

Para testar a significância do modelo, podem ser utilizados três testes: Teste de Wald (ou Teste Z), Teste dos Escores e o Teste da Razão de Verossimilhança. Serão comentados os Testes de Wald e da Razão de Verossimilhança.

As hipóteses do Teste de Wald são as seguintes:

$$\begin{cases} H_0: \beta_k = 0 \\ H_1: \beta_k \neq 0 \end{cases}$$

A estatística do Teste de Wald é:

$$Z^2 = \left( \frac{\hat{\beta}_k}{\text{var}(\hat{\beta}_k)} \right)^2 \sim \chi^2 \text{ com 1 grau de liberdade (1 g.l.)}$$

ou equivalentemente, a estatística do Teste Z é:

$$Z = \frac{\hat{\beta}_k}{\text{var}(\hat{\beta}_k)} \sim N(0,1)$$

O cálculo do intervalo de confiança para os parâmetros do modelo é feito através da seguinte fórmula:

$$IC: \left\{ \hat{\beta}_k \pm z_{\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\beta}_k)} \right\}$$

O intervalo encontrado fornece em quanto os valores de cada parâmetro podem variar na população. Quando se calcula o intervalo de confiança utilizando  $e^{\hat{\beta}_k}$  no lugar de  $\hat{\beta}_k$  tem-se como resultado o intervalo em que pode variar a *odds ratio* (razão de chance) de cada parâmetro na população (Pereira,2006).

Quando o Teste de Wald não rejeitar a hipótese nula, recomenda-se utilizar o Teste da Razão de Verossimilhança para testar se realmente o coeficiente não é significativo.

O Teste da Razão de Verossimilhança é utilizado como um teste para se verificar a adequabilidade de ajustamento, nele, as hipóteses são:

$$\begin{cases} H_0: O \text{ modelo se ajusta aos dados} \\ H_1: O \text{ modelo não se ajusta aos dados} \end{cases}$$

A estatística do Teste da Razão de Verossimilhança é:

$$G^2 = -2(L_0 - L_1)$$

onde  $L_0$  é o máximo do log da função de verossimilhança quando a hipótese nula é verdadeira e  $L_1$  é o máximo do log da função de verossimilhança quando a hipótese alternativa é verdadeira. Sendo que  $G^2$  se aproxima da distribuição  $\chi^2$  com c graus de liberdade, onde c é a diferença entre o número de parâmetros definidos para cada hipótese.

A estatística Deviance é uma medida de ajuste do modelo aos dados onde se compara um modelo mais simples (com menos variáveis explicativas) com um modelo mais complexo (com mais variáveis explicativas) onde se busca usar o modelo mais simples (Pereira,2006).

As hipóteses para o teste da Deviance são:

$$\begin{cases} H_0: O \text{ modelo 1 se ajusta aos dados} \\ H_1: O \text{ modelo 2 se ajusta aos dados,} \end{cases} \quad \text{onde } M_1 \subset M_2$$

Para o teste de adequabilidade utiliza-se:

$$G^2 = 2 \sum (\text{observado}) \ln \left( \frac{\text{observado}}{\text{esperado}} \right) \sim \chi_{g-k}^2$$

onde k é o número de parâmetros do modelo e g é o número de conjuntos de valores distintos das variáveis explicativas.

O cálculo da estatística  $G^2$  para os modelos que estão sendo comparados é

$$\text{deviance } (M_1) = -2 (L_{M_1} - L_C)$$

$$\text{deviance } (M_2) = -2 (L_{M_2} - L_C)$$

onde  $M_1$  é o modelo 1, com menos variáveis que  $M_2$  (o modelo 2) e  $L_C$  é o máximo do log da função de verossimilhança para o modelo completo (saturado), ou seja, com todas as variáveis. A estatística do teste é calculada pela diferença entre a *deviance* ( $M_1$ ) e *deviance* ( $M_2$ ) (Pereira,2006).

$$G^2 = \text{deviance } (M_1) - \text{deviance } (M_2)$$

Se  $G^2$  for significativa, utiliza-se o modelo 2, caso contrário, utiliza-se o modelo 1.

O próximo passo na análise de regressão é a seleção de variáveis, que consiste em escolher um conjunto de variáveis que melhor possam descrever a variável em estudo. Um conhecimento prévio das variáveis explicativas contribui para essa seleção, no entanto, nem sempre esse conhecimento é disponível, sendo assim, para se fazer a seleção os procedimentos mais utilizados são o Stepwise Regression, Forward Regression e o Backward Regression. Esses métodos são os mesmos utilizados em regressão linear.

O procedimento Stepward consiste basicamente em adicionar ou remover variáveis explicativas do modelo. A cada adição ou retirada de uma variável cria-se um novo modelo e para cada novo modelo calcula-se a estatística Deviance a fim de verificar se o modelo com a nova variável inserida possui maior poder de predição que o modelo sem esta variável. Então, as hipóteses do teste são:

$$\begin{cases} H_0: \beta_k = 0 \\ H_1: \beta_k \neq 0 \end{cases}$$

E a estatística do teste: 
$$G^2 = \frac{\hat{\beta}_k}{\text{var}(\hat{\beta}_k)}$$

Como regra de decisão, não entrará no modelo as variáveis explicativas para as quais se aceita a hipótese nula do teste, enquanto que, aquelas que tiverem o maior valor de  $G^2$  são fortes candidatas a entrar no modelo (Pereira,2006).

A cada iteração todas as variáveis são testadas novamente e podem tanto continuar quanto sair do modelo. O usual é determinar um valor de corte (teórico ou tabelado) e, quando o valor encontrado de  $G^2$  for maior que o determinado, a variável entrará no modelo, caso contrário, o modelo termina sem nenhuma variável explicativa selecionada. O procedimento acaba quando não for mais possível incluir ou eliminar nenhuma variável (Pereira,2006).

O procedimento Forward é idêntico ao Stepward, porém não há retirada de variáveis. O procedimento inicia-se com um modelo sem nenhuma variável explicativa e a cada novo passo uma variável explicativa é incluída, até que no final todas as variáveis estejam no modelo e não houve retirada de nenhuma durante o processo.

O procedimento Backward é o oposto do Forward, nele o procedimento inicia-se com o modelo completo e a cada novo passo uma variável explicativa é analisada e pode ser retirada.

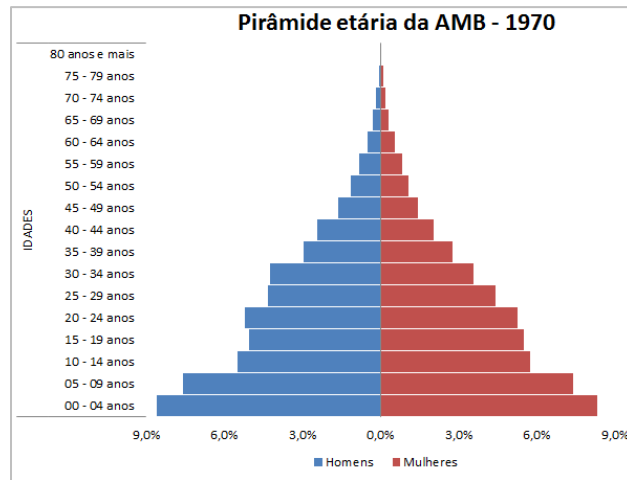
É importante observar que esses procedimentos pressupõem a existência de um conjunto de variáveis explicativas consideradas “ótimas” e é possível que não exista apenas um conjunto dessas variáveis. O valor de corte determinado não possui sentido estatístico, pois este é sempre fixo para as iterações, enquanto que os graus de liberdade variam de acordo como número de variáveis eu estão no modelo.

## **4. RESULTADOS**

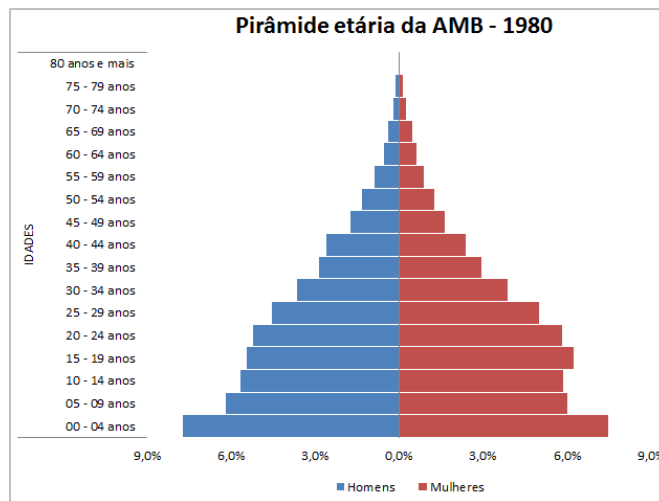
### **4.1. CARACTERÍSTICAS GERAIS DOS IDOSOS RESIDENTES NA ÁREA METROPOLITANA DE BRASÍLIA**

Na amostra do Censo Demográfico de 2010 foram coletadas informações de 206.283 moradores da AMB, desses, 14.314 eram idosos. Após a expansão da amostra, obteve-se uma população de mais de 3,5 milhões de pessoas na AMB, desses, 7,2% são idosos. Para se ter um melhor conhecimento da distribuição dessa população ao longo do tempo, pode-se observar suas pirâmides etárias ilustradas nas FIGURAS 2 à 6. A pirâmide de 1970 mostra uma estrutura típica de população jovem, com elevada natalidade e uma esperança média de vida relativamente baixa, o mesmo pode ser observado na pirâmide de 1980. Já a pirâmide de 1991 mostra uma pequena queda da natalidade (visível pelo estreitamento na base da pirâmide) e um pequeno aumento na esperança de vida. Na pirâmide do ano 2000 pode-se observar um aumento na esperança de vida (alargamento no topo da pirâmide) e um maior estreitamento da base, caracterizando a transição de uma população jovem para uma população predominantemente adulta e idosa. Por fim, em 2010, observa-se que a população da AMB é predominantemente adulta (67,6%), pois sua pirâmide tem o corpo largo, além disso, nota-se que a proporção de jovens está diminuindo, pois o estreitamento na base da pirâmide é cada vez maior. Devido à contínua queda das taxas de mortalidade e de natalidade, e à crescente expectativa de vida, pode-se notar que há uma proporção grande de idosos no topo da pirâmide, o que não ocorre nas pirâmides nos anos anteriores. Pode-se prever que daqui a algumas décadas a pirâmide da AMB se assemelhará às pirâmides dos países desenvolvidos.

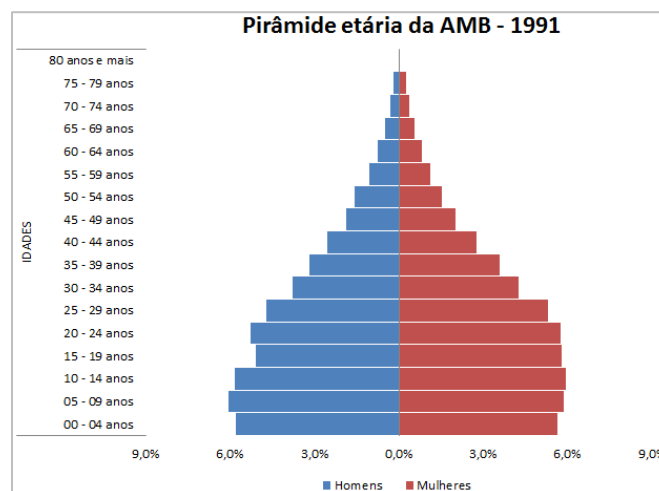
Vale ressaltar que em 1970 a AMB era composta pelo DF e mais cinco municípios goianos (Cristalina, Formosa, Luziânia, Padre Bernardo e Planaltina). Em 1991 foi incluído o município de Santo Antônio do Descoberto. Por fim, em 2000, as cidades de Águas Lindas de Goiás, Cidade Ocidental, Novo Gama e Valparaíso de Goiás também foram incluídas na AMB. Muitos desses novos municípios incluídos ao longo dos anos na AMB são resultados de desmembramentos dos municípios já existentes. Esses desmembramentos tem como uma das causas a forte migração para o DF e entorno.



**Figura 2 Pirâmide etária da AMB de 1970**  
**Fonte: Censo Demográfico 1970 - IBGE**

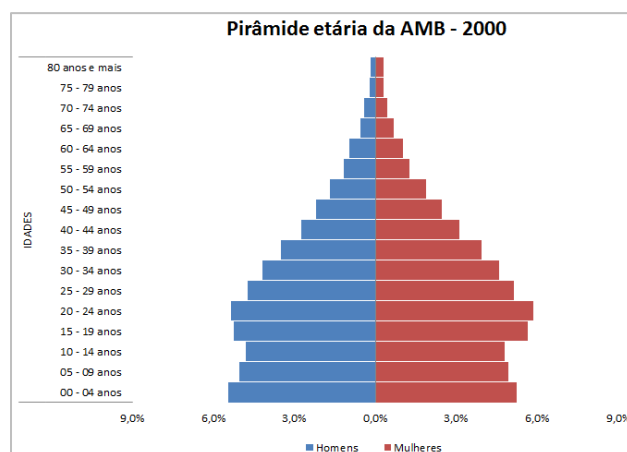


**Figura 3 Pirâmide etária da AMB de 1980**  
**Fonte: Censo Demográfico 1980 - IBGE**



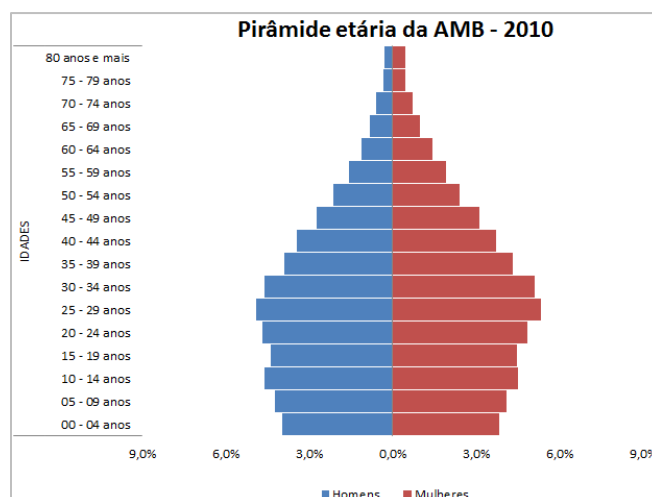
**Figura 4 Pirâmide etária da AMB de 1991**  
**Fonte: Censo Demográfico 1991 - IBGE**





**Figura 5 Pirâmide etária da AMB de 2000**

**Fonte: Censo Demográfico 2000 - IBGE**



**Figura 6 Pirâmide etária da AMB de 2010**

**Fonte: Censo Demográfico 2010 - IBGE**

Segundo dados do último Censo, dentre a população responsável pelo domicílio, 14% são idosos, e dentre os idosos 60% são responsáveis pelos domicílios. Esses números mostram que a população idosa representa uma camada importante da população da AMB, uma vez que a maioria deles são os provedores de muitos lares dessa região.

Como a população de estudo são os idosos, indivíduos com 60 anos ou mais, é de grande interesse conhecer a distribuição desses idosos de acordo com a principal variável desse estudo, responsável ou não pelo domicílio.

Como pode ser visto na TAB. 1, 60% dos idosos da AMB estão entre a faixa etária de 60 a 69 anos, além disso, mais de 10% dos idosos estão acima dos 80 anos de idade, o

que reflete que além do envelhecimento populacional, também está ocorrendo um aumento da longevidade, uma vez que as pessoas estão vivendo mais.

**Tabela 1 Distribuição dos idosos responsáveis pelo domicílio por faixa etária, sexo, cor ou raça, nível de instrução, estado civil e região da AMB.**

FAIXA ETÁRIA	IDOSO RESPONSÁVEL PELO DOMICÍLIO		
	SIM	NÃO	TOTAL
60-69	62,70%	37,30%	100,00%
70-79	62,90%	37,10%	100,00%
80+	49,60%	50,40%	100,00%
SEXO			
Homem	73,30%	26,70%	100,00%
Mulher	52,00%	48,00%	100,00%
COR OU RAÇA			
Branca	60,90%	39,10%	100,00%
Preta	62,70%	37,30%	100,00%
Parda	61,70%	38,30%	100,00%
NÍVEL DE INSTRUÇÃO			
Sem instrução e fundamental incompleto	59,40%	40,60%	100,00%
Fundamental completo e médio incompleto	59,70%	40,30%	100,00%
Médio completo e superior incompleto	62,30%	37,70%	100,00%
Superior completo	69,20%	30,80%	100,00%
Não determinado	59,10%	40,90%	100,00%
ESTADO CIVIL			
Casado(a)	55,20%	44,80%	100,00%
Desquitado(a) ou separado(a) judicialmente ou Divorciado(a)	75,50%	24,50%	100,00%
Viúvo(a)	69,20%	30,80%	100,00%
Solteiro(a)	59,90%	40,10%	100,00%
REGIÃO DA AMB			
Região 1	62,60%	37,40%	100,00%
Região 2	60,90%	39,10%	100,00%
Região 3	60,50%	39,50%	100,00%
Região 4	62,50%	37,50%	100,00%
<b>TOTAL - n (%)</b>	<b>8.791 (61,4%)</b>	<b>5.523 (38,6%)</b>	<b>14.314 (100%)</b>

Fonte: IBGE, Censo Demográfico 2010.

Como já mencionado, aproximadamente 62% dos idosos são responsáveis pelo domicílio, essa proporção se mantém aparentemente constante na faixa etária dos 60 aos 79 anos, já na faixa dos 80 anos ou mais há uma queda dessa proporção e passa a ter um equilíbrio entre as proporções.

A maioria dos idosos é mulher. Dos idosos homens, 73,3% são responsáveis pelo domicílio, e das idosas, 52% são responsáveis pelo domicílio.

Observando a distribuição por cor ou raça, vê-se que a maioria dos idosos é da cor branca. Nas três categorias, mais de 60% dos idosos são responsáveis pelo domicílio.

Já para a variável nível de instrução, observa-se que mais da metade dos idosos não possui instrução ou possui fundamental incompleto. Das categorias da variável nível de instrução, observa-se que em todas a maioria dos idosos é responsável pelo domicílio. Destaca-se a categoria superior completo, onde quase 70% dos idosos são os responsáveis. Nota-se também que a porcentagem de idosos responsáveis aumenta conforme cresce o grau de escolaridade do idoso.

Observando a distribuição dos idosos por estado civil, vê-se que a maioria dos idosos é casada seguida pelos viúvos. Dos idosos casados, 75,5% são responsáveis pelo domicílio.

Como observado na TAB. 1, a maioria dos idosos da AMB residem na região 3, aquela composta pelas cidades com menor renda média do DF. Em todas as regiões, mais de 60% dos idosos são os responsáveis pelo domicílio.

Com relação às variáveis de condições de saúde, observa-se que em todos os casos, do total de idosos, a maioria não possui dificuldades de enxergar, ouvir, caminhar ou deficiência mental. Em todas as categorias, a maioria dos idosos são responsáveis pelo domicílios, exceto a categoria possuir deficiência mental, onde a maioria dos idosos não são os responsáveis.

**Tabela 2 Distribuição dos idosos responsáveis pelo domicílio por variáveis de condições de saúde**

<b>IDOSO RESPONSÁVEL PELO DOMICÍLIO</b>			
<b>VARIÁVEIS DE CONDIÇÕES DE SAÚDE</b>	<b>SIM</b>	<b>NÃO</b>	<b>TOTAL</b>
<b>DIFICULDADES EM ENXERGAR</b>			
Sim (alguma dificuldade, grande dificuldade ou não consegue de modo algum)	61,6%	38,4%	100,00%
Não (nenhuma dificuldade)	61,2%	38,8%	100,00%
<b>DIFICULDADES EM OUVIR</b>			
Sim (alguma dificuldade, grande dificuldade ou não consegue de modo algum)	63,1%	36,9%	100,00%
Não (nenhuma dificuldade)	61,0%	39,0%	100,00%
<b>DIFICULDADES EM CAMINHAR OU SUBIR DEGRAUS</b>			
Sim (alguma dificuldade, grande dificuldade ou não consegue de modo algum)	60,4%	39,6%	100,00%
Não (nenhuma dificuldade)	61,9%	38,1%	100,00%
<b>DEFICIÊNCIA MENTAL / INTELECTUAL</b>			
Sim	36,1%	63,9%	100,00%
Não	62,2%	37,8%	100,00%
<b>TOTAL - n (%)</b>	<b>8.791 (61,4%)</b>	<b>5.523 (38,6%)</b>	<b>14.314 (100%)</b>

Fonte: IBGE, Censo Demográfico 2010.

## 4.2. ESTIMAÇÃO DO MODELO

Foi utilizado o software estatístico IBM SPSS Statistics 22.0 para executar o modelo para estimação das probabilidades de risco. Todas as variáveis estudadas no trabalho foram incluídas no modelo.

Para a modelagem foi utilizado o método de Forward, onde o modelo inicia-se sem nenhuma variável explicativa e a cada etapa entra uma variável explicativa até que o modelo esteja completo. Na TAB. 3, pode-se observar as variáveis utilizadas na análise e

suas categorias. A primeira categoria de todas as variáveis é utilizada como categoria de referência.

**Tabela 3 Categorias e codificação das variáveis que entraram no modelo**

Variável	Categorias	Códigos dos parâmetros			
		(1)	(2)	(3)	(4)
Nível de instrução	Sem instrução e fundamental incompleto	0	0	0	0
	Fundamental completo e médio incompleto	1	0	0	0
	Médio completo e superior incompleto	0	1	0	0
	Superior completo	0	0	1	0
	Não determinado	0	0	0	1
Região da AMB	Região 1	0	0	0	
	Região 2	1	0	0	
	Região 3	0	1	0	
	Região 4	0	0	1	
Estado Civil	Casado	0	0	0	
	Desquitado ou separado judicialmente ou Divorciado	1	0	0	
	Solteiro	0	1	0	
	Viúvo	0	0	1	
Cor ou raça	Branca	0	0		
	Parda	1	0		
	Preta	0	1		
Idade	60-69	0	0		
	70-79	1	0		
	80+	0	1		
Sexo	Homem	0			
	Mulher	1			
Deficiência em enxergar	Não	0			
	Sim	1			
Deficiência em caminhar ou subir degraus	Não	0			
	Sim	1			
Deficiência em ouvir	Não	0			
	Sim	1			
Deficiência mental / intelectual	Não	0			
	Sim	1			

Na tabela abaixo, TAB. 4, são apresentadas as estimativas dos parâmetros  $\beta$ , seus respectivos desvios padrão, além da Estatística de Wald para testar a significância das variáveis. Têm-se também os valores das *odds ratios* e seus respectivos intervalos de confiança.

**Tabela 4 Estimativas e inferências das variáveis que entraram no modelo**

Variável	Categorias	$\beta$	Desvio padrão	Estatística de Wald	g.l.	p-valor	Exp( $\beta$ )	I.C. para Exp ( $\beta$ ) com 95%	
								Inferior	Superior
Nível de instrução	Sem instrução e fundamental incompleto	-	-	56,47	4	<0,0001	-	-	-
	<b>Fundamental completo e médio incompleto</b>	0,06	0,06	1,03	1	<b>0,31</b>	1,07	0,94	1,21
	Médio completo e superior incompleto	0,21	0,06	12,60	1	<0,0001	1,23	1,10	1,38
	Superior completo	0,53	0,07	53,69	1	<0,0001	1,70	1,47	1,96
	<b>Não determinado</b>	-0,13	0,42	0,10	1	<b>0,76</b>	0,88	0,38	2,00
Região da AMB	<b>Região 1</b>	-	-	0,24	3	<b>0,97</b>	-	-	-
	<b>Região 2</b>	0,01	0,07	0,01	1	<b>0,90</b>	1,01	0,88	1,16
	<b>Região 3</b>	0,01	0,07	0,03	1	<b>0,85</b>	1,01	0,88	1,17
	<b>Região 4</b>	0,03	0,07	0,16	1	<b>0,69</b>	1,03	0,89	1,19
Estado civil	Casado	-	-	736,90	3	<0,0001	-	-	-
	Desquitado ou separado judicialmente ou Divorciado	1,13	0,07	267,93	1	<0,0001	3,08	2,69	3,53
	Solteiro	0,47	0,05	76,10	1	<0,0001	1,60	1,44	1,78
	Viúvo	1,29	0,05	616,57	1	<0,0001	3,62	3,27	4,00
Cor ou raça	<b>Branca</b>	-	-	0,64	2	<b>0,73</b>	-	-	-
	<b>Parda</b>	0,01	0,04	0,12	1	<b>0,73</b>	1,01	0,94	1,10
	<b>Preta</b>	0,06	0,07	0,62	1	<b>0,43</b>	1,06	0,92	1,21
Deficiência mental / intelectual	Não	-	-	-	-	-	-	-	-
	Sim	-1,24	0,11	126,41	1	<0,0001	0,29	0,23	0,36
Idade	60-69	-	-	163,44	2	<0,0001	-	-	-
	70-79	-0,10	0,04	5,46	1	0,02	0,90	0,83	0,98
	80+	-0,82	0,06	161,07	1	<0,0001	0,44	0,39	0,50
Deficiência em enxergar	Não	-	-	-	-	-	-	-	-
	Sim	0,12	0,04	9,72	1	<0,0001	1,13	1,05	1,22
Deficiência em ouvir	Não	-	-	-	-	-	-	-	-
	Sim	0,09	0,05	3,77	1	0,05	1,10	1,00	1,21
Deficiência em caminhar ou subir degraus	Não	-	-	-	-	-	-	-	-
	Sim	0,16	0,04	13,81	1	<0,0001	1,18	1,08	1,28
Sexo	Homem	-	-	-	-	-	-	-	-
	Mulher	-1,33	0,04	1.037,78	1	<0,0001	0,26	0,24	0,29
Constante	-	0,65	0,08	65,36	1	<0,0001	1,92	-	-

Verificando a significância das variáveis através da Estatística de Wald, observa-se que nas categorias em negrito a hipótese de que o parâmetro não possui significância estatística não é rejeitada, portanto, essas variáveis não conseguem explicar a variável dependente, ou seja, não explicam o fato do idoso ser o responsável pelo domicílio.

É interessante observar que na variável nível de instrução, os idosos com maior grau de escolaridade têm maiores chances de serem os responsáveis pelo domicílio. Quando comparado à categoria de referência (sem instrução e fundamental incompleto), a chance do idoso com superior completo ser o responsável pelo domicílio é 1,7 vezes a chance do idoso daquela categoria.

Ao se estudar o estado civil dos idosos, observa-se que os idosos separados ou viúvos têm maior chance de serem os responsáveis pelo domicílio do que os idosos casados, suas *odds ratios* chegam a ser maior que três nesses casos, ou seja, o idoso separado ou viúvo tem 3 vezes a chance do idoso casado de ser o responsável pelo domicílio.

À medida que o idoso envelhece sua chance de ser o responsável pelo domicílio decresce. O idoso com idade entre 60 e 69 anos tem 2,27 vezes a chance do idoso com 80 anos ou mais de ser o responsável pelo domicílio.

Em relação às variáveis de condições de saúde, observa-se algo interessante: os idosos com alguma dificuldade em enxergar, ouvir ou caminhar tem maiores chances de serem os responsáveis pelo domicílio do que os idosos sem dificuldades. Porém o mesmo não ocorre com a variável deficiência mental/intelectual: o idoso sem esse tipo de deficiência tem 3,44 vezes a chance do idoso com essa deficiência de ser o responsável pelo domicílio.

Por fim, os idosos homens têm 3,84 vezes a chance das idosas de serem os responsáveis pelos domicílios.

### **4.3. AVALIAÇÃO DO MODELO**

A estatística do teste de Hosmer e Lemeshow testa a hipótese de que os dados observados são significativamente diferentes dos valores previstos pelo modelo. Nesse estudo a estatística encontrada foi de 522,281 com 8 graus de liberdade e com probabilidade aproximadamente zero de se aceitar a hipótese de que existem diferenças significativas, ou seja, o modelo se adequou aos dados, uma vez que os valores preditos não foram significativamente diferentes dos valores observados.

Na TAB. 5 observa-se a relação de acertos entre os valores preditos pelo modelo e os valores observados. O valor de corte para o modelo é de 0,5, sendo assim, o modelo é

melhor para determinar idosos responsáveis do que idosos não responsáveis pelo domicílio.

**Tabela 5 Relação de acertos entre valores estimados e preditos nos dados da modelagem**

Idoso responsável pelo domicílio observado	Idoso responsável pelo domicílio previsto		Porcentagem correta
	Não	Sim	
Não	2784	2725	50,5
Sim	1365	7388	84,4
Porcentagem global			71,3



## 5. CONCLUSÕES

Ao se estudar os idosos responsáveis pelos domicílios e suas características sociodemográficas, estudos já realizados destacam que: com o avançar da idade, diminui a quantidade de idosos responsáveis pelo domicílio; idosos com algum tipo de doença têm menores chances de serem os responsáveis do que aqueles ditos sadios; idosos com maior grau de escolaridade, o que acarreta muitas vezes em uma renda maior, são mais responsáveis pelo domicílio do que aqueles com baixa escolaridade.

Muitos desses resultados foram observados no presente trabalho. Com o uso da técnica estatística de regressão logística e o auxílio do software SPSS, conclui-se que idosos do sexo masculino, com ensino superior completo, viúvos, sem deficiência mental/intelectual e com idade entre 60 e 69 anos, têm maiores chances de serem os responsáveis pelo domicílio na Área Metropolitana de Brasília, com base nos dados do Censo de 2010.

Um resultado observado nesse trabalho e que não era esperado é o fato de que idosos com alguma dificuldade em enxergar, ouvir, caminhar ou subir degraus têm maiores chances de serem os responsáveis pelo domicílio do que idosos sem dificuldades. Essa conclusão pode ter sido afetada pelo tamanho da amostra, seria necessário um estudo aprofundado para obter melhores resultados.

Com esse trabalho o governo poderá ter um melhor conhecimento sobre a distribuição de seus idosos e assim criar políticas públicas para melhor atender essa população que cada vez mais se torna de suma importância na sociedade.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

- ALSEMI, A. **Envelhecimento da população mundial: Desafios e Oportunidades**. Brasil, 2010. Disponível em <<http://portaldoenvelhecimento.org.br/noticias/longevidade/envelhecimento-da-populacao-mundial-desafios-e-oportunidades.html>>. Acesso em 2/5/2013.
- BUARQUE, D. Envelhecimento da população mundial preocupa pesquisadores. São Paulo, 2011. Disponível em <<http://g1.globo.com/mundo/noticia/2011/10/envelhecimento-da-populacao-mundial-preocupa-pesquisadores.html>>. Acesso em 2/5/2013.
- CAMARANO, A.A. Envelhecimento da população brasileira: Uma contribuição demográfica. Rio de Janeiro: IPEA, 2002.
- CAMARANO, A. A.; KANSO, S.; MELLO, J. L. Quão além dos 60 poderão viver os idosos brasileiros? In: CAMARANO, A. M. (Org). Os novos idosos brasileiros: muito além dos 60? Rio de Janeiro: IPEA, 2004. p.411-426.
- CAMARGOS, M.; MACHADO, C.; NASCIMENTO, R.. **A relação entre renda e morar sozinho para idosos paulistanos: 2000**. Revista Brasileira de Estudos Populacionais. São Paulo, v. 24, n. 1, jun. 2007. p.37-51.
- CARVALHO, J. A. M.; GARCIA, R. A. O envelhecimento da população brasileira: um enfoque demográfico. Cadernos de Saúde Pública, Rio de Janeiro, v. 19, n. 3, p. 725-733, maio./jun. 2003.
- FELIX, J. S.. **Economia da Longevidade: uma revisão da bibliografia brasileira sobre o envelhecimento populacional**.
- FRANÇA, M. B. *et al*, **Determinantes das condições dos idosos em domicílios unipessoais: Um estudo para o Rio Grande do Norte – 2000**. (Trabalho apresentado no XVI Encontro Nacional de Estudos Populacionais, realizado em Caxambu- MG – Brasil, de 29 de setembro a 03 de outubro de 2008).

- GOMES, M. M. F. Passado e presente: uma análise dos determinantes da mortalidade entre idosos com base nos dados da SABE 2000-2006. 2011. 189f. Tese (Doutorado em Demografia) – Centro de Desenvolvimento e Planejamento Regional, Universidade Federal de Minas Gerais, Belo Horizonte, 2011.
- HAKKERT, Ralph. **Fontes de Dados Demográficos**. Belo Horizonte, 1996.
- HOSMER, D. W.; LEMESSHOW, S. Applied Logistic Regression, Second Edition, 1985, John Wiley and Sons, Inc., New York.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. Hotsite Censo 2010. Disponível em: <<http://www.censo2010.ibge.gov.br/>>. Acesso em 22/4/2013.
- LONGO, F. A. L. de B. *et al*, **Idosos que moram sozinhos em Minas Gerais: independência ou vulnerabilidade?** (Trabalho apresentado no XVIII Encontro Nacional de Estudos Populacionais, ABEP, realizado em Águas de Lindóia/SP – Brasil, de 20 a 24 de novembro de 2012).
- MACHADO, A. R. **Modelos Estatísticos para Avaliação de Risco em Produtos de Crédito Parcelados**. Brasília, 2010. (Monografia apresentada ao Departamento de Estatística da Universidade de Brasília como parte dos requisitos para a obtenção do grau de Bacharel em Estatística).
- MONTGOMERY, D. C., PECK, E. A. Introduction to Linear Regression Analysis, 2nd ed., 1992, John Wiley and Sons, Inc., New York.
- MYERS, R. Classical and Modern Regression with Applications, 2th Ed., 1990, Duxbury Press, California USA, p.488
- NETER, J.; WASSERMAN, W.; KUTNER, M. Applied linear statistical models: regression, analysis of variance, and experimental designs, 2nd ed., 1985, Homewood, Ill.: R.D. Irwin.
- PAULO, M. A. *et al*, **A relação entre renda e composição domiciliar dos idosos no Brasil: um estudo sobre o impacto do recebimento do Benefício de Prestação**

**Continuada.** (Trabalho apresentado no XVI Encontro Nacional de Estudos Populacionais, ABEP, realizado em Caxambu – MG – Brasil, de 29 de setembro a 03 de outubro de 2008).

- PEREIRA, A. R. De A.. Modelo de *credit scoring* para avaliação de risco na concessão de crédito bancário. Brasília, 2006. (Monografia apresentada ao Departamento de Estatística da Universidade de Brasília como parte dos requisitos para a obtenção do grau de Bacharel em Estatística).

- **PERFIL dos idosos responsáveis pelos domicílios no Brasil 2000.** Rio de Janeiro: IBGE, 2002. (Estudos e pesquisas. Informação demográfica e socioeconômica, n. 9).

- PORTAL Action. São Paulo. Disponível em <<http://www.portalaction.com.br/>>. Acesso em 2/5/2013.

- UNITED NATIONS. Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat. **World Population Prospects: The 2010 Revision.** May, 2011. Disponível em: <<http://esa.un.org/unpd/wpp/index.htm>>. Acesso em: 16 set. 2011.

- VASCONCELOS, A. M. N. Cenários demográficos para a Área Metropolitana de Brasília. In: PAVIANI, A. *et al* (Org.). Brasília 50 anos da capital a metrópole. Brasília: Editora UnB, 2010.

➤ Sites:

- [www.inf.ufsc.br/~ogliari/arquivos/regressao\\_logistica.ppt](http://www.inf.ufsc.br/~ogliari/arquivos/regressao_logistica.ppt). Acesso em 1/7/2013

## 7. ANEXOS

**Tabela 62 Composição da AMB durante os anos de 1970 a 2010**

Municípios	Anos				
	1970	1980	1991	2000	2010
Distrito Federal	x	x	x	x	x
Águas Lindas de Goiás				x	x
Cidade Ocidental				x	x
Cristalina	x	x	x	x	x
Formosa	x	x	x	x	x
Luziânia	x	x	x	x	x
Novo Gama				x	x
Santo Antônio do Descoberto			x	x	x
Padre Bernardo	x	x	x	x	x
Planaltina	x	x	x	x	x
Valparaíso de Goiás				x	x