

**FATORES IMPORTANTES PARA A SENSAÇÃO DE FELICIDADE DA
POPULAÇÃO DO ESTADO DE SÃO PAULO**

**FACTORS IMPORTANT FOR THE FEELING OF HAPPINESS OF THE
POPULATION OF THE STATE OF SÃO PAULO**

**FACTORES IMPORTANTES PARA EL SENTIMIENTO DE FELICIDAD DE LA
POBLACIÓN DEL ESTADO DE SÃO PAULO**

SERGIO AUGUSTO RODRIGUES¹

GISLAINE CRISTINA BATISTELA²

Recebido em setembro de 2011. Aprovado em dezembro de 2011.

¹Professor Assistente da Faculdade de Tecnologia de Botucatu. Bacharel em estatística. Mestre em Estatística:
Av.: José Ítalo Bacchi, s/n – Jardim Aeroporto – Botucatu/SP – CEP 18606-855 E-mail: sau02@yahoo.com.br

²Professor Assistente da Faculdade de Tecnologia de Botucatu. Graduado em Matemática. Mestre em Estatística:
Av.: José Ítalo Bacchi, s/n – Jardim Aeroporto – Botucatu/SP – CEP 18606-855 E-mail:
gi_batistela@yahoo.com.br

FATORES IMPORTANTES PARA A SENSAÇÃO DE FELICIDADE DA POPULAÇÃO DO ESTADO DE SÃO PAULO

RESUMO

Este trabalho analisa a sensação de felicidade, não como uma característica complexa e subjetiva, mas a felicidade declarada, relacionando-a com fatores socioeconômicos. Para isto, utilizou-se o modelo de regressão logístico múltiplo a partir de um conjunto de dados obtidos de uma pesquisa realizada na população com mais de 16 anos do Estado de São Paulo. Através de uma amostra aleatória de 300 indivíduos, verificou-se que as variáveis faixa etária e tipo de moradia estão conjuntamente associadas à sensação de felicidade dos paulistas. Outro resultado observado foi que os indivíduos com menos de 40 anos e com moradia própria tem 96,4% de chance de se considerar feliz, enquanto que este percentual chega a aproximadamente 65% para indivíduos com mais de 60 anos e sem moradia própria.

Palavras-chave: Regressão logística. Relação entre variáveis. Felicidade.

FACTORS IMPORTANT FOR THE FEELING OF HAPPINESS OF THE POPULATION OF THE STATE OF SÃO PAULO

ABSTRACT

This study examines the feeling of happiness, not as a complex trait and subjective, but the reported happiness, relating it to socioeconomic factors. For this, we used the multiple logistic regression model from a data set obtained from a survey in the population over 16 years of the State of Sao Paulo. Through a random sample of 300 individuals found that the variables of age and state of residence are jointly associated with the feeling of happiness of Sao Paulo. Another result has been observed that individuals with less than 40 years and home ownership is 96.4% chance of considering happy, while this percentage reaches about 65% for individuals over 60 years and without proper housing.

Keywords: Logistic regression. Relation between variable. Felicity.

FACTORES IMPORTANTES PARA EL SENTIMIENTO DE FELICIDAD DE LA POBLACIÓN DEL ESTADO DE SÃO PAULO

RESUMEN

Este estudio analiza el sentimiento de felicidad, no como un rasgo complejo y subjetivo, pero la felicidad declarada, en relación a los factores socioeconómicos. Para ello, se utilizó el modelo de regresión logística múltiple a partir de un conjunto de datos obtenidos de una encuesta en la población mayor de 16 años del Estado de Sao Paulo. A través de una muestra aleatoria de 300 individuos se verificó que las variables de edad y tipo de hogar, de manera conjunta, asocia con el sentimiento de felicidad de São Paulo. Otro de los resultados se ha observado que los individuos con menos de 40 años y propiedad de la vivienda son del 96,4% de probabilidad de considerarse feliz, mientras que este porcentaje alcanza el 65% de las personas mayores de 60 años y sin una vivienda propia.

Palabras-clave: Regresión Logística. Relación entre las Variables. Felicidad.

1 INTRODUÇÃO

Dentre os modelos de regressão observados na literatura estatística, um modelo bastante utilizado é o modelo de regressão logística. Neste modelo, diferente do modelo de regressão linear (modelo mais popular), em que a variável resposta é um valor contínuo, sua variável resposta é uma chance de ocorrer um determinado fato.

A regressão logística é caracterizada por suas variáveis dependentes serem qualitativas ou categorizadas. Atualmente, a regressão logística está sendo aplicada em diversas áreas, entre elas pode-se destacar: a determinação de risco de crédito (BRITO e ASSAF NETO, 2007), da chance de um aluno desistir de seu curso (CARRER et al., 2007), a chance de ocorrer uma determinada doença (GIMENO e SOUZA, 1995), entre outras.

Segundo Rodrigues et al. (2006) relacionar a felicidade e bem estar de uma população com dados econômicos de uma região ou país vem ganhando espaço na literatura econômica de alguns países, porém é um assunto pouco explorado no Brasil.

Hirata (2005) *apud* Rodrigues et al. (2006) estuda a felicidade como um indicador para desenvolvimento de políticas públicas. Este trabalho sugere a utilização do FIB (Felicidade Interna Bruta) como um indicador de desenvolvimento de uma nação.

Segundo Corbi (2006), uma pesquisa sobre os determinantes da felicidade pode também contribuir para a resolução de paradoxos empíricos que a teoria econômica convencional tem dificuldades para explicar. Um destes paradoxos mostra, por exemplo, que em muitos países ocidentais a renda real se elevou drasticamente a partir da Segunda Guerra mundial, ao mesmo tempo em que medidas indicam que o bem-estar subjetivo manteve-se constante ou sofreu até pequenas quedas durante o período.

2 MATERIAIS E MÉTODOS

Um levantamento bibliográfico foi realizado a fim de descrever a metodologia do modelo de regressão logístico simples e múltiplo. Além disso, realizou-se um estudo de caso, aplicando esta metodologia em um conjunto de dados reais, utilizando o modelo de regressão logístico múltiplo para determinar alguns fatores sócio-econômicos que podem influenciar na felicidade momentânea de um indivíduo.

2.1 Regressão logística

No modelo de regressão logística a variável resposta ou dependente assume um valor categorizado.

Quando a variável resposta está disposta em categorias, existem três

procedimentos de regressão logística distintos para relacionar variáveis: regressão logística binária, ordinária e nominal. A escolha depende do número de categorias e característica da variável resposta.

A regressão logística binária é utilizada quando a variável apresenta duas categorias e dois níveis de resposta.

Já a regressão logística ordinária é utilizada quando apresenta três ou mais categorias e existe ordenação natural dos níveis de resposta.

Por último, a regressão logística nominal é utilizada quando apresenta três ou mais categorias e não existe a ordenação natural nos níveis de respostas.

Para o caso da regressão logística binária, as variáveis explicativas assumem qualquer valor, porém a variável resposta será do tipo de Bernoulli, ou seja, assumem valores 0 ou 1 (Toledo, Ovalle: 1995).

2.2 Regressão Logística Simples

Segundo Hair et al (2005), a regressão logística é um método estatístico que descreve as relações entre uma variável resposta e uma ou mais variáveis independentes. A variável resposta (y) é dicotômica, ou seja, é uma variável que apresenta duas possibilidades de resposta (sucesso e fracasso), atribuindo o valor 1 para o acontecimento de interesse (sucesso)

e o valor 0 para o acontecimento complementar (fracasso).

Desta forma, Hair et. al. (2005) descreve a probabilidade de sucesso do modelo logístico simples por:

$\pi_i = \pi(x_i) = P(y = 1 x = x_i) = \frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)}$	(1)
---	-----

e a probabilidade de fracasso:

$1 - \pi_i = 1 - \pi(x_i) = P(y = 0 x = x_i) = \frac{1}{1 + \exp(\beta_0 + \beta_1 x_i)}$	(2)
---	-----

onde β_0 e β_1 são parâmetros desconhecidos.

Segundo Toledo e Ovalle (1995), em qualquer problema de regressão, a quantidade a ser modelada é o valor médio da variável resposta, dado os valores das variáveis independentes. Esta quantidade é chamada de média condicional, denotada por $E(y/x = x_i)$, sendo que “ y ” é a variável resposta e x_i , os valores das variáveis independentes. Na regressão logística, devido à natureza da variável resposta $0 \leq E(y/x = x_i) \leq 1$, ou seja, a esperança condicional está variando entre 0 e 1.

Na regressão logística, o valor da variável resposta dado x_i , é expresso por $y_i = \pi_i + e_i$, como a quantidade e_i , que pode assumir somente um de dois possíveis valores, isto é, $e_i = 1 - \pi_i$ para $y_i = 1$ ou $e_i = -\pi_i$ para $y_i = 0$, segue que e_i tem

distribuição com média zero e variância dada por $\pi_i(1 - \pi_i)$, de acordo com Hosmer et al. (1989).

Segundo Moore (2006), a regressão logística modela a média “p” em termos de uma variável explicativa “x”. Tentar relacionar “p” e “x” como numa regressão linear simples: $p = \beta_0 + \beta_1 x$. Infelizmente, esse não é um bom modelo. Sempre que $\beta_1 \neq 0$, valores extremos de “x” fornecerão valores para $\beta_0 + \beta_1 x$ que ficam fora do conjunto de possíveis valores de “p”, $0 \leq p \leq 1$. O modelo de regressão logística remove essa dificuldade, operando com o logaritmo natural da chance, $p/(1 - p)$.

Utiliza-se o termo logaritmo (ou log) da chance para essa transformação. À medida que “p” se afasta de 0 em direção a 1, o logaritmo da chance assume todos os valores numéricos possíveis negativos e positivos. Modela-se o logaritmo da chance como função linear da variável explicativa, isto é, o modelo para a regressão logística simples:

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x \quad (3)$$

onde “p” é uma proporção binomial e “x” é a variável explicativa. Os parâmetros do modelo logístico são β_0 e β_1 , os quais devem ser estimados pelo método de máxima verossimilhança.

2.3 Regressão Logística Múltipla

Segundo Hair et al. (2005), a regressão logística múltipla é uma generalização do modelo de regressão logística simples, isto é, na regressão logística simples, trabalha-se com uma variável explicativa e na regressão logística múltipla com mais de uma variável explicativa.

Para modelar o logaritmo da chance como função de duas ou mais variáveis explicativas, ajusta-se um modelo da forma:

$$\ln\left[\frac{\hat{p}}{1-\hat{p}}\right] = \hat{\alpha} + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_q x_q \quad (4)$$

onde x_1, x_2, \dots, x_q são as variáveis explicativas e $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_q$ os parâmetros do modelo logístico.

3 RESULTADOS

Os dados utilizados para este estudo de caso foram obtidos junto à pesquisa intitulada “Mapa da Felicidade” (SAMPLING Consultoria Estatística e Pesquisa de Mercado; LIMITE Pesquisa de Mercado, 2004). Nesta pesquisa, foram abordadas questões como: sexo, idade, renda, escolaridade, estado civil, sensação de felicidade, como considera o povo brasileiro, forma que prefere pagar suas contas, plano de saúde, internet na residência, ocupação, tipo de moradia. A partir da amostra de mais

de cinco mil indivíduos entrevistados, obtida através de uma amostragem estratificada de toda população com mais de 16 anos do Estado de São Paulo, 300 indivíduos foram sorteados aleatoriamente desta base de dados

para comporem o banco de dados deste trabalho.

Na Tabela 1, são apresentadas todas as variáveis consideradas e suas respectivas categorias de respostas.

Tabela 1 – Descrição das variáveis de estudo

Variável	Categorias das variáveis		
Sexo	1-Masculino	2-Feminino	
Faixa etária (Idade)	1-até 40 Anos	2-40 a 60 Anos	3-Mais de 60 Anos
Estado civil	1-casado	2-Solteiro	3-outros (viúvo, amasiado, divorciado)
Tipo de moradia	1-própria	2-não própria	
Nível escolaridade	1- 1º grau (completo e incompleto)		2- 2º grau (compl.e incompl.)
	3-superior		
Profissão exercida	1-indústria/func. público/comércio.	2-serviço/autônomo/ agropecuário	
		3-desempregado/aposentado	4-estudante/do lar
Renda mensal	1-até 400,00	2- 400,00 a 1.000,00	3-1.000,00 a 2.000,00
	4-mais de 2.000,00		
Quantas pessoas dependem da renda	1-até 2 pessoas	2- 2 a 5 pessoas	3-mais de 5 pessoas
Prefere pagar à vista com cartão de débito	0-não	1-sim	
Prefere pagar à vista com cheque	0-não	1-sim	
Prefere pagar à vista com dinheiro	0-não	1-sim	
Prefere pagar cartão de crédito	0-não	1-sim	
Prefere pagar cheque pré-datado	0-não	1-sim	
Prefere pagar crediário	0-não	1-sim	
Pegou dinheiro emprestado	1-sim	2-não	
Tem plano de saúde	1-sim	2-não	
Teve plano de saúde	1-sim	2-não	
Acesso à internet em casa	1-sim	2-não	
Opinião de acha o povo brasileiro feliz	0-não	1-sim	
Felicidade	0-não feliz	1-feliz	

Fonte: Pesquisa “Mapa da Felicidade”

Inicialmente, uma análise bivariada da relação de cada variável com a variável binária “Felicidade” foi realizada através do teste usando a estatística *qui-quadrado* (χ^2).

Este teste pode ser encontrado em Bussab e Morettin (1987). A medida utilizada como conclusão dos testes realizados foi o *p-valor*. Esta medida indica a probabilidade de se

cometer um erro ao rejeitar a hipótese de que as variáveis são independentes. Desta forma, se o valor desta medida é muito pequeno, conclui-se que há evidências estatísticas de que a hipótese de independência das variáveis pode ser rejeitada.

Na Tabela 2, pode-se observar um resumo dos *p-valores* obtidos através do

software SPSS versão 13.0. Considerando um nível de significância de 5%, verifica-se que somente as variáveis: faixa etária (idade), estado civil, tipo de moradia, número de pessoas que dependem da renda, acesso à internet em casa e opinião sobre o povo brasileiro, estão estatisticamente associadas com a variável Felicidade (feliz ou não feliz).

Tabela 2: Resultado do teste *qui-quadrado* para associação entre a variável binária felicidade com as demais variáveis do estudo

Variáveis	P-valor	Resultado com nível de significância de 5%
Sexo	0,49	Não existe relação
Faixa etária (Idade)	<0,001	Existe relação
Estado civil	<0,001	Existe relação
Tipo de moradia	0,009	Existe relação
Nível escolaridade	0,475	Não existe relação
Profissão exercida	0,996	Não existe relação
Renda mensal	0,339	Não existe relação
Quantas pessoas dependem da renda	<0,001	Existe relação
Prefere pagar à vista com cartão de débito	0,339	Não existe relação
Prefere pagar à vista com cheque	0,963	Não existe relação
Prefere pagar à vista com dinheiro	0,100	Não existe relação
Prefere pagar cartão de crédito	0,604	Não existe relação
Prefere pagar cheque pré-datado	0,686	Não existe relação
Prefere pagar crediário	0,128	Não existe relação
Pegou dinheiro emprestado	0,877	Não existe relação
Tem plano de saúde	0,135	Não existe relação
Teve plano de saúde	0,741	Não existe relação
Acesso à internet em casa	0,017	Existe relação
Opinião se acha o povo brasileiro feliz	0,027	Existe relação

Após a análise bivariada, o modelo de regressão logístico múltiplo foi aplicado utilizando as variáveis que apresentaram alguma relação com a variável felicidade, ou

seja, aquelas variáveis que apresentaram um *p-valor* com nível de significância de 5% no teste *qui-quadrado*.

Com isso, para o ajuste do modelo de regressão logístico múltiplo, foi considerada a variável binária Felicidade como variável dependente e as variáveis: faixa etária, estado civil, tipo de moradia, número de pessoas na família, número de pessoas que dependem da renda, acesso da internet em casa, opinião sobre o povo brasileiro como variáveis explicativas.

Para a estimação deste modelo, utilizou-se o software SPSS versão 13, com o método *Forward Stepwise: Wald*. Neste método, cada variável é incluída separadamente no modelo e verificado seu desempenho em estimar a variável binária “felicidade”. O processo termina até as variáveis que, conjuntamente, explicam melhor à variável dependente forem selecionadas.

Para verificar a qualidade do modelo ajustado, utilizou-se primeiramente o teste de *Omnibus*. Verificou-se que o modelo obtido apresenta adequação estatisticamente significativa, considerando um nível de significância de 2%, ou seja, rejeitou-se a hipótese inicial de que as variáveis explicativas não explicam adequadamente a variável dependente “Felicidade”.

Outro teste importante é o teste de adequação do modelo logístico dado pela estatística “ $-2 \log Likelihood$ ”. De acordo com Hair et al. (2005), um modelo é bem ajustado quanto menor for o valor desta

estatística. Verificou-se que a cada passo (inclusão de variáveis no modelo) esta estatística foi reduzindo indicando que conforme as variáveis iam sendo introduzidas no modelo, sua adequação foi melhorando.

Outras medidas também foram utilizadas (medidas de adequação de *Cox & Snell e Nagelkerke*) a fim de avaliar se o modelo foi bem ajustado. Segundo Hair et al. (2005), estas medidas de adequação de ajuste comparam as probabilidades estimadas com as probabilidades observadas, sendo que os valores mais altos significam um ajuste melhor.

Para Carpenter (2006), o teste de “*Hosmer and Lemeshow*” é outro teste que pode ser utilizado para verificar a qualidade do modelo ajustado. Esse teste mede a correspondência entre os valores reais e os previstos da variável dependente, através da estatística *qui-quadrado*. Para o modelo ajustado, observou-se que não há evidências estatísticas para rejeitar a hipótese de que os valores previstos pelo modelo estão próximos dos valores observados, indicando que o modelo foi bem ajustado.

Para avaliar o modelo logístico ajustado em relação a sua capacidade de prever as chances de um indivíduo se

considerar feliz ou não feliz, utilizou-se o método de matrizes de Classificação.

A Tabela 3 mostra a classificação dos casos previstos pelo modelo comparando

com os dados observados da variável “Felicidade”.

Tabela 3 – Matriz de classificação

Situação Observada na amostra		Valores Previstos		% Correto
		0 (não feliz)	1 (feliz)	
Passo 2	0 (não-feliz)	4	37	9,8
	1 (feliz)	8	245	96,8
Percentual de acerto				84,7

Observa-se que 84,7% dos indivíduos são classificados de uma forma correta em um dos dois grupos (feliz ou não feliz). O percentual de classificação correto de um indivíduo não feliz foi de 9,8% e o percentual correto para um indivíduo feliz é de 96,8%. Portanto, pode-se dizer que o modelo ajustado possui uma capacidade

melhor para classificar corretamente indivíduos felizes.

A Tabela 4 mostra as variáveis que entraram no modelo, gerada pela ferramenta de regressão logística múltipla bem como os seus coeficientes e os resultados do teste de significância das variáveis independentes.

Tabela 4 – Modelo de regressão logístico múltiplo ajustado

		β (Coeficientes)	Erro Padrão	Est. Wald	G.l.	P-valor	Exp(β)
Passo 2	Faixa etária			28,958	2	0,000	
	Faixa etária (1)	-1,808	0,534	11,458	1	0,001	0,164
	Faixa etária (2)	-3,507	0,653	28,827	1	0,000	0,030
	Moradia (1)	0,816	0,491	2,764	1	0,096	2,261
	Constante	3,296	0,455	52,373	1	0,000	27,000

O teste de Wald é usado em regressão logística para obter a significância de cada coeficiente logístico β (colunas *Est. Wald*, *gl* e *p-valor*). De acordo com Hair et al. (2005), esse teste irá identificar o quanto à variável

explicativa participa individualmente da explicação da variável dependente.

Na coluna “ β (coeficientes)”, são apresentados, no passo 2, em que todas as variáveis já foram selecionadas, os coeficientes do modelo para cada variável

explicativa. Na coluna Erro Padrão, apresenta-se uma medida de variabilidade destes coeficientes.

A coluna *Exp. (β)* é o exponencial dos coeficientes de cada categoria das variáveis do modelo e indica chance do indivíduo ser feliz se comparado com a chance da linha da categoria anterior.

Analisando a Tabela 4, verifica-se que a chance de um indivíduo ser feliz pode ser explicada pelas variáveis: “Faixa Etária” e “Tipo de moradia”. Utilizando somente essas variáveis como variáveis explicativas do modelo de regressão logístico, pode-se prever a chance de um indivíduo ser considerado uma pessoa feliz.

Logo, o modelo de regressão logístico múltiplo é dado por:

$$\log(\text{chance}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n \quad (5)$$

Considerando os coeficientes estimados do modelo apresentados na Tabela 4, com poder explicativo de 96,8%, o modelo de regressão logístico que melhor determina as chances de felicidade de um indivíduo pode ser estimado por:

$$\begin{aligned} \log(\text{chance}) = & 3,296 + 0(\text{faixa etária}) - 1,808(\text{faixa etária1}) \\ & - 3,507(\text{faixa etária2}) + 0(\text{tipo de moradia}) \\ & + 0,816(\text{tipo de moradia1}) \end{aligned} \quad (6)$$

sendo:

- (faixa etária) = a variável faixa etária na categoria 1, ou seja, menos de 40 anos;
- (faixa etária1) = a variável faixa etária na categoria 2, ou seja, de 40 a 60 anos;
- (faixa etária2) = a variável faixa etária na categoria 3, ou seja, mais de 60 anos;
- (tipo de moradia) = a variável tipo de moradia na categoria 1, ou seja, casa própria;
- (tipo de moradia1) = a variável tipo de moradia na categoria 2, ou seja, não própria.

Nas categorias de resposta da idade, os coeficientes são negativos, indicando que quanto maior a idade dos paulistas, menor a chance de um indivíduo ser feliz para um dado tipo de moradia.

Assim, se existe um indivíduo de 35 anos e com moradia não própria, sua chance de ser feliz é de 98,39%.

Com isto, observando a Tabela 4, na linha da faixa etária 1 (indivíduos com idade entre 40 e 60 anos), a chance de um indivíduo desta faixa etária ser feliz é 16,4% da chance de um indivíduo da faixa etária anterior (idade inferior a 40 anos), ou melhor, pode-se dizer que para um indivíduo com menos de 40 anos, sua chance de ser feliz é 6,1 (1/0,164) vezes a chance de um indivíduo com idade entre 40 e 60 anos. Já a

chance do grupo dos mais velhos, faixa etária 2 (mais de 60 anos), a chance de ser feliz é 3% da chance do grupo dos mais jovens, ou melhor, indivíduos com menos de 40 anos, sua chance de ser feliz é aproximadamente 33 vezes (1/0,03) a chance de um indivíduo com mais de 60 anos.

A Tabela 5 apresenta a chance de um indivíduo ser feliz de acordo com o perfil de

idade e tipo de moradia. Verifica-se que 39,33% dos indivíduos da amostra têm menos de 40 anos, com moradia própria, sendo que sua chance de ser feliz é de 96,43%. Já os indivíduos com mais de 60 anos sem moradia própria, os quais representam 5% da amostra, sua chance de ser feliz é de 64,66%.

Tabela 5 – Chances de felicidade dos indivíduos da amostra

Faixa etária	Tipo de moradia	Chances (%)	% da amostra
menos 40 anos	própria	96,43%	39,33%
menos 40 anos	não-própria	98,39%	8,33%
de 40 a 60 anos	própria	81,58%	28%
de 40 a 60 anos	não-própria	90,91%	8%
mais 60 anos	própria	44,44%	11,33%
mais 60 anos	não-própria	64,66%	5%
Total			99,99%

4. CONCLUSÃO

Analisando individualmente a associação de cada variável com a variável binária Felicidade através do teste *qui-quadrado*, verificou-se que as variáveis: faixa etária, estado civil, tipo de moradia, acesso à internet em casa, número de pessoas que dependem da renda e opinião se o povo brasileiro é feliz, foram às únicas variáveis com associação significativa com a variável Felicidade.

No entanto, uma análise conjunta das variáveis anteriores, utilizando o modelo de regressão logístico múltiplo, verificou-se que o melhor modelo para representar a felicidade, é o modelo com as seguintes variáveis explicativas: faixa etária e tipo de moradia.

Desta forma, basta saber em qual das categorias de faixa etária e de tipo de moradia o indivíduo se encaixa para determinar suas chances de serem felizes.

REFERÊNCIAS

- BRITO, G. A. S; ASSAF NETO, A. **Modelo de classificação de risco de crédito de grandes empresas.** São Paulo, 2006. Disponível em <www.congressoeac.locaweb.com.br/artigos52005/383>. Acesso em: 03 março 2007.
- BUSSAB, O; MORETTIN, P.A **Métodos quantitativos: estatística básica** São Paulo: Atual, 1987.
- CARPENTER, E. M. L. **Um modelo de análise de risco de crédito de clientes em relações B2B.** Tese (Doutorado em Administração de Empresas) da PUC Rio, Rio de Janeiro, 2006.
- CARRER, J. T; SILVA, M. C; BELMAR, A. S. **Regressão logística: análise da desistência dos alunos de Gestão da Fundação Educacional de Ituverava.** Trabalho de Conclusão de Curso (Graduação em Licenciatura em Matemática) – Fundação Educacional de Ituverava Faculdade de Filosofia Ciências e Letras, Ituverava, 2007.
- CORBI, R.B; MENEZES, N. A. Os determinantes empíricos da Felicidade no Brasil. São Paulo: **Revista Economia Política**, 2006.
- GIMENO, S. G. A; SOUZA, J. M. P. Utilização de estratificação e modelo de regressão logística na análise de dados de estudo caso-controle. São Paulo: **Revista Saúde Pública**, nº 29, 1995.
- HAIR, J. et al. **Análise multivariada de dados**, 5º ed. Porto Alegre: Bookman, 2005.
- HOSMER, D. W.; LEMESHOW, S. **Applied Logistic Regression.** New York: John Wiley, 1989.
- MAI, A. F. **O perfil do empreendedor versus a mortalidade das micro e pequenas empresas comerciais do município de Aracruz/ES.** Dissertação (Mestrado em Ciências Contábeis) – Fundação Instituto Capixaba de Pesquisas em Contabilidade, Economia e Finanças, Vitória, 2006.
- MOORE, D. et al. **A prática da Estatística Empresarial: como usar dados para tomar decisões**, Rio de Janeiro, LTC, 2006.
- RODRIGUES, S. A; BATISTELA, G. C; BARRETO, M. C. Percepção de felicidade da população do Estado de São Paulo: fatores determinantes. Valinhos: **Revista Ciências Gerenciais**, nº 12, vol. X, ano 2006.
- SAMPLING CONSULTORIA ESTATÍSTICA E PESQUISA DE MERCADO; LIMITE PESQUISA DE MERCADO (Brasil). **Mapa da Felicidade.** Disponível em: <<http://consumoemarcas.com.br/mapadafelicidade.html>>. Acesso em: 01 jan. 2007.
- TOLEDO, L. G; OVALLE, I.I **Estatística básica.** São Paulo: Atlas, 1995.