

APLICAÇÃO DO MODELO PROBIT PARA ANÁLISE DA RELAÇÃO ENTRE VARIÁVEIS AMBIENTAIS E O GRAU DE SATISFAÇÃO DE USUÁRIOS DE EDIFICAÇÕES

Eduardo Henrique Silveira de Araújo, Especialista em Estatística e Engenharia de Produção
Departamento de Estatística / CONSULEST - CCET - UFRN
Campus Universitário, BR 101 - Lagoa Nova. CEP 59.072-970 Natal/RN
Tel. (084)215.3789, (084)215.3787 FAX (084)215.3781 E-mail: ehisa@digi.com.br

RESUMO

O presente trabalho trata da aplicação de um método estatístico de análise de dados de variáveis que estão relacionadas entre si. O método é baseado em um modelo probabilístico que tem a característica de transformar a distribuição de probabilidade da variável. Esta transformação possibilita a interpretação da relação entre variáveis.

ABSTRACT

This paper treat the application of a estatistic method to analyse the variables that are relations between them. The method is based in a probabilistic model that has the characteristic to transform the probability distribution of the variable. This transformation make possible the interpretation of variables relations.

INTRODUÇÃO

Este trabalho apresenta a aplicação de um modelo estatístico que foi usado para representar a relação existente entre o estímulo biológico de indivíduos e as variações ocorridas em variáveis que representam condições climáticas. O modelo foi aplicado aos dados oriundos de uma pesquisa de campo aplicada a indivíduos que estavam realizando uma atividade sedentária e que estavam submetidos as variáveis climáticas. Estas variáveis foram medidas internamente nas edificações (ARAÚJO,1996).

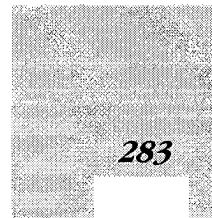
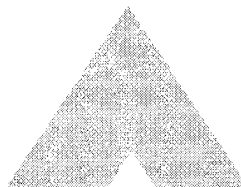
As condições ambientais internas das edificações devem sempre atender as exigências higrotérmicas dos usuários. Em regiões onde o clima sofre várias mudanças durante um período diário, mensal ou anual, o interior das edificações sofre com essas mudanças. Portanto, as condições ambientais no interior das edificações variam de acordo com as condições climáticas externas. Como consequência, os usuários também sentirão os efeitos dessas variações. Por exemplo, quando um estímulo térmico é aplicado a vários seres humanos, com o mesmo nível de intensidade, a resposta pode variar de um indivíduo para outro. Portanto, estímulos diferentes podem ser comparados em termos da magnitude das respostas que produzem em vários indivíduos.

Os estímulos térmicos são medidos através de equipamentos que indicam um resultado numérico e a resposta observada nos indivíduos são identificadas por categorias. Portanto, ao se realizar um estudo que leve em conta o comportamento da relação entre o estímulo e a resposta, torna-se imprescindível o uso de algumas técnicas estatísticas específicas de modelagem de variáveis.

O modelo retrata um princípio físico de causa e efeito. Este princípio na prática pode ser observado, por exemplo, quando se verifica que há um aumento ou diminuição nos valores medidos da temperatura do ar de uma edificação, observa-se mudanças nas sensações térmicas dos usuários. Estas sensações são representadas por categorias que retratam o que os indivíduos sentem sobre o ambiente construído. Quando na referida pesquisa foram medidas a temperatura, as categorias definida "apriori" classificaram o mesmo em: muito quente, quente, confortável, frio, muito frio.

CARACTERIZAÇÃO DAS VARIÁVEIS

As variáveis objetos de estudo na pesquisa, foram classificadas em duas categoria: a) Variável Contínua e b) Variável Categórica. As variáveis contínuas estudadas foram: Temperatura do ar, Temperatura radiante média, Umidade relativa do ar e Velocidade do ar no ambiente. A variável categórica estudada foi representada pela sensibilidade percebida pelos indivíduos sobre as condições do ambiente. As unidades de medida das variáveis contínuas foram *graus celsius*, *percentual* e *metros por segundo*. A variável categórica sensibilidade foi identificada nas seguintes categorias: muito quente, quente, confortável, frio, muito frio.



PROCEDIMENTO DE MEDIÇÃO DAS VARIÁVEIS

As variáveis contínuas foram medidas no ambiente através de equipamentos próprios. A variável categórica foi identificada por cada um dos indivíduos que se encontravam no ambiente através de sua resposta. A *proporção* de indivíduos que responderam se identificando em cada uma das categorias, caracterizou a situação observada do ambiente.

METODOLOGIA DE ANÁLISE DOS DADOS

Observando o banco de dados das variáveis estudadas, notou-se visualmente uma grande variação nos valores das variáveis medidas nos ambientes e os identificados nas respostas dos indivíduos. Os valores das *proporções* mostraram que realmente existe um efeito, por exemplo, da temperatura na sensibilidade de cada indivíduo. Isto reproduz a situação em que para uma intensidade qualquer da temperatura em graus no ambiente, uma resposta do indivíduo é obtida, isto é, um estímulo é aplicado e uma resposta é obtida. Nesse caso a resposta (do indivíduo) depende da intensidade do estímulo (temperatura). Assim pode-se dizer que as variáveis contínuas são independentes e aleatórias e a variável categórica é dependente. Por outro lado, para todo indivíduo, haverá sempre um certo nível de intensidade do estímulo (um limite) onde ele muda a sua categoria de resposta. Este limite é chamado de tolerância para a resposta do indivíduo. Essa tolerância pode variar de um indivíduo para outro e ela tem influência nos valores das *proporções* observadas em cada categoria da sensibilidade. Como existem várias tolerâncias, pois limite máximo de tolerância para uma categoria da sensibilidade pode passar a ser o limite mínimo de tolerância de outra categoria, os seus valores formam uma certa distribuição de probabilidade.

O método de análise dos dados diante da descrição acima, baseou-se num estudo de um modelo que representasse bem a relação entre as variações dos valores das *proporções* e os valores medidos da temperatura (tolerâncias) e também das outras variáveis. Este modelo escolhido chama-se PROBIT (FINNEY, 1964).

Para melhor interpretação da relação acima citada, as *proporções* observadas foram transformadas em probabilidade e representadas por P_i . Os valores medidos da temperatura foram representados por T_i . Verificou-se graficamente (Figura 1), que a relação entre P_i e T_i tem a forma de uma curva do tipo sigmoidal.

No referido gráfico, por exemplo, a curva pode representar a relação entre os valores das probabilidades na categoria muito quente da sensibilidade com os valores medidos da temperatura. Verifica-se que probabilidades quase nulas, estão relacionadas com os menores valores da temperatura e que as probabilidades quase iguais a um, estão relacionadas com os maiores valores da temperatura (tolerância).

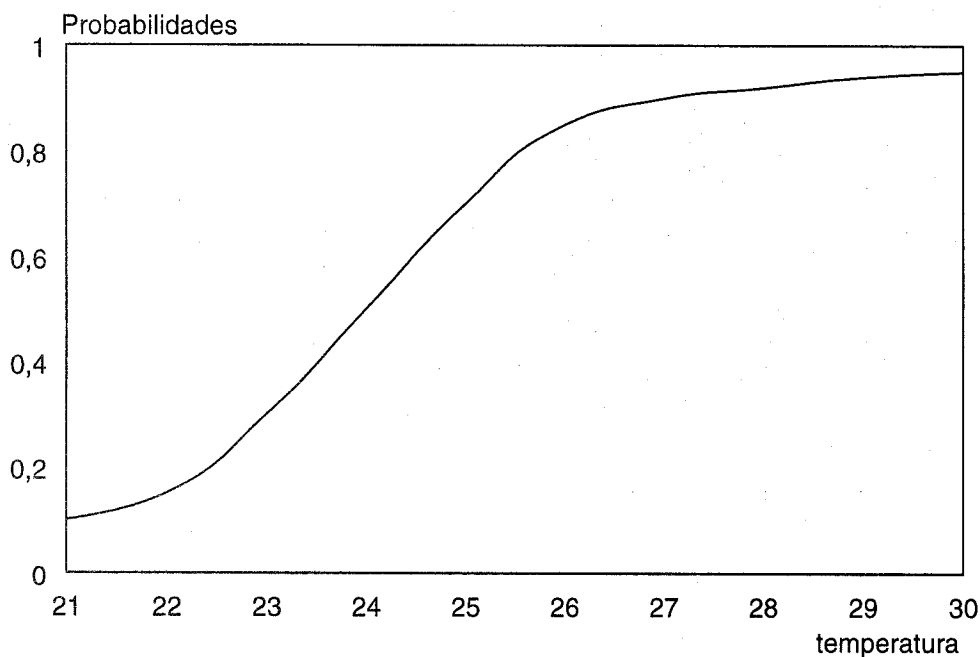


Figura 01 Curva sigmoidal entre temperatura e probabilidade

Visualizando o gráfico da Figura 1 nota-se que os valores maiores de P_i são encontrados por uma função estritamente crescente de T_i . Esta é uma função contínua em T e que tem as propriedades matemáticas de ser uma *função de distribuição contínua acumulada* com uma forma sigmoidal. Os valores das probabilidades na abscissa são encontradas através da relação abaixo, onde T_i são os valores medidos da temperatura:

$$P_i = P(Z < T_i) = \int_{-\infty}^{t_i} f(z) dz, \text{ onde } P_i \text{ é a probabilidade desejada} \quad \text{Eq. 1}$$

$f(z)$ função que representa a curva senoidal.

Existem várias funções $f(z)$ que podem representar o comportamento da trajetória da curva. Neste trabalho optou-se por trabalhar com a função que geralmente é conhecida como Probit.

O MODELO "PROBIT" (PROBABILITY UNIT)

A função $f(z)$ da equação 1 é contínua e não linear. Supondo que os valores encontrados da temperatura (tolerância) pela equação 2 tenha distribuição de probabilidade normal, com uma média m e variância s^2 , a função densidade de probabilidade que representa a distribuição dos dados é dada por:

$$f_p(t; m; s^2) = (1 / \sqrt{2\pi s}) * \exp[-(t - m) / s]^2 \quad \text{com } m \in \mathbb{R} \text{ e } s^2 > 0 \quad \text{Eq. 2}$$

Se a probabilidade (m_1) da temperatura assumir um valor menor do que um certo t_1 , ou seja, $m_1 = P(T \leq t_1)$ pode ser calculada através da transformação da variável T para uma variável Z , que possui uma distribuição normal padronizada. Dessa forma, m_1 pode ser calculada pela equação:

$$m_1 = P\left\{ \left[\frac{(T - m)}{s} \right] \leq \left[\frac{(t_1 - m)}{s} \right] \right\} = P\left[Z \leq \left(\frac{m}{s} \right) + \left(\frac{1}{s} \right) * t_1 \right] \quad \text{Eq.3}$$

$$\text{ou, } m_1 = F(b_1 + b_2 * t_1), \text{ pois } Z \sim N(0, 1) \text{ onde } b_1 = -m/s \text{ e } b_2 = 1/s \quad \text{Eq.4}$$

Portanto, uma transformação inversa em F produz os valores probites que possibilitam encontrar probabilidades da variável sensibilidade, ou seja, o probito (m_1) = $F^{-1}(m_1) = b_1 + b_2 * t_1$. Os valores dos probites são encontrados em FISHER & YATES (1971).

RESULTADOS

Os valores dos probites estimados, formam com os valores da temperatura uma reta linear. Através do método de regressão linear simples é possível se estimar para qualquer valor da temperatura o valor corresponde do probito. Em consequência com esse valor, se estima o valor da probabilidade (proporção) de um indivíduo pertencer a uma categoria da variável sensibilidade.

Tomando-se, como exemplo $m = 5$ e $s = 2,17$, os gráficos das figuras 2, 3 e 4 mostram as curvas da $f(z)$, a curva da probabilidade m_1 e a reta dos valores dos probitos (m_1).

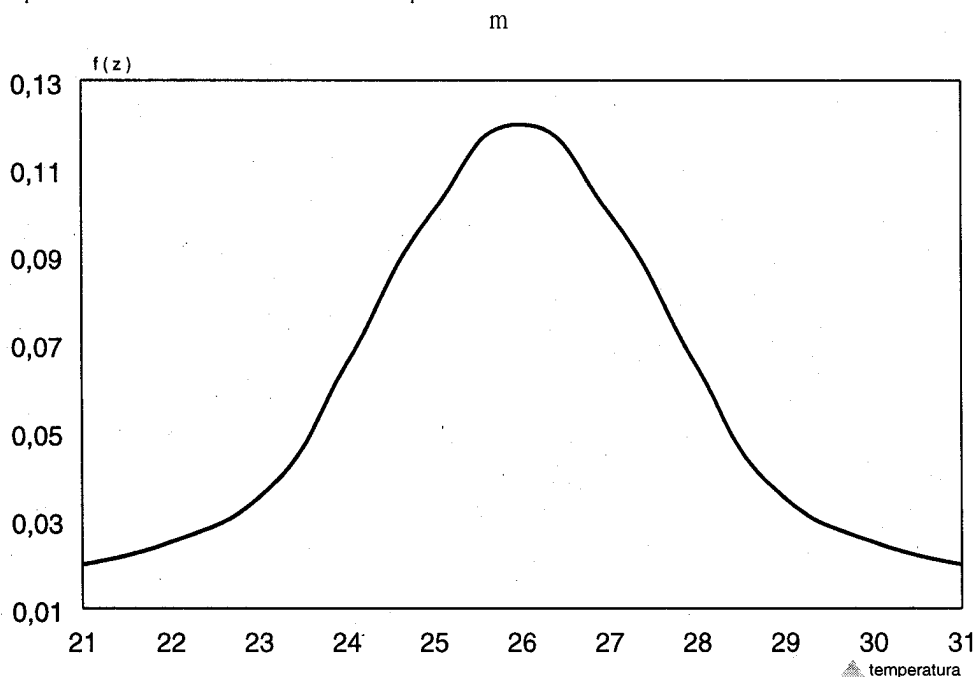


Figura 2. Curva normal

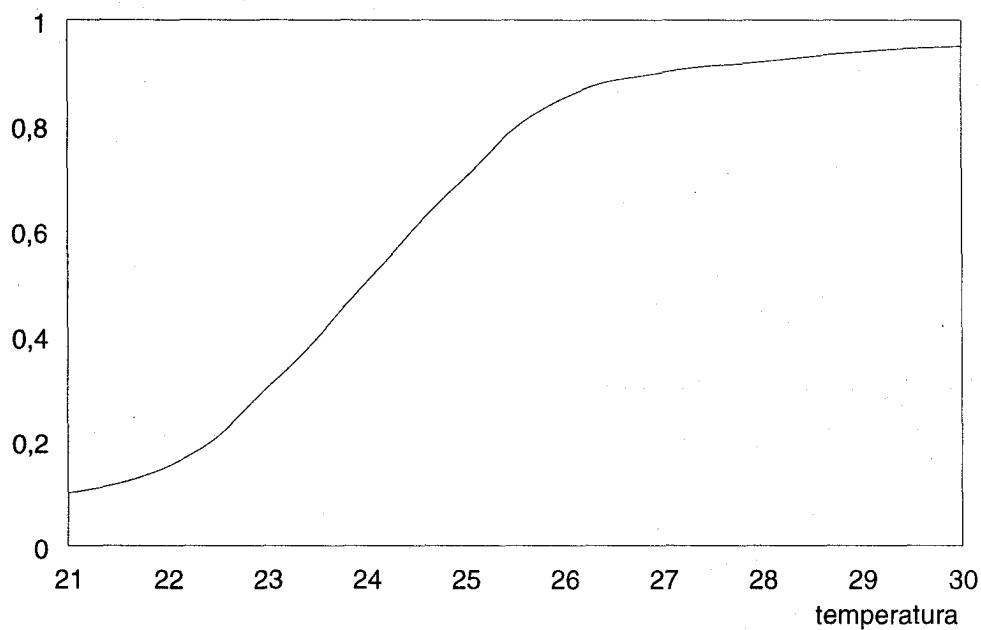


Figura 3 . Probabilidade m

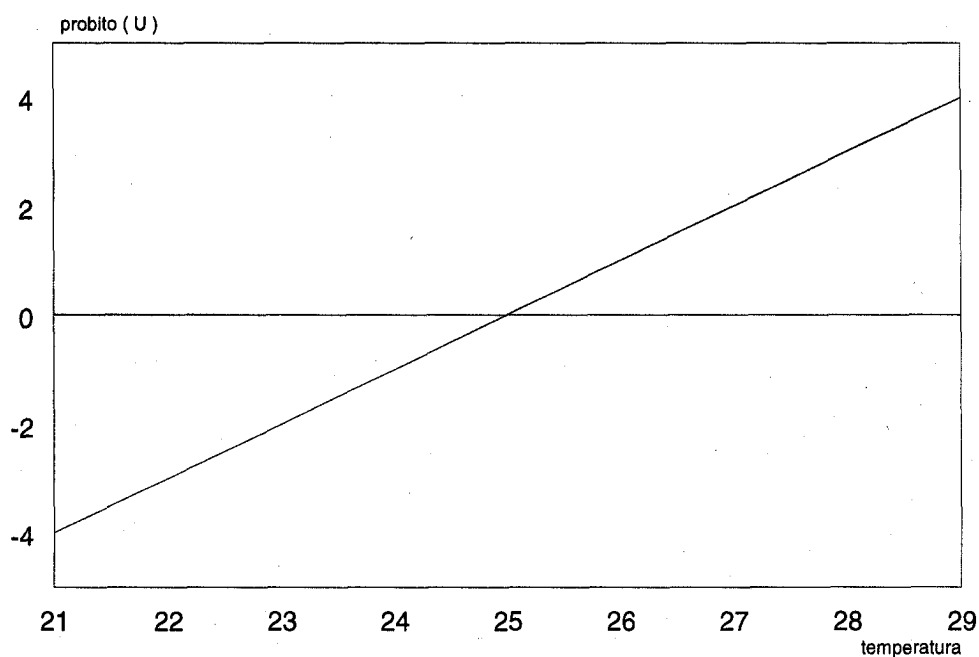


Figura 4. Reta estimada dos probites

Os dados medidos da temperatura nos ambientes durante a referida pesquisa estão na coluna TBS(°C) na tabela 1 do Anexo. Nota-se que os valores das proporções para alguns valores da temperatura não existem, ou seja, não foram identificados pelos indivíduos na pesquisa. Através da transformação das proporções pelo modelo PROBIT foi possível se estimar as proporções para os valores das temperaturas onde não existiam os valores das mesmas.

As figuras 5 e 6 mostram as estimativas dos probites usando-se um modelo de regressão linear simples. Em função disso, foi possível encontrar os valores das tolerâncias para cada uma das categorias da variável sensibilidade. A figura 7 mostra os limites para a categoria "Confortável" da variável sensibilidade.

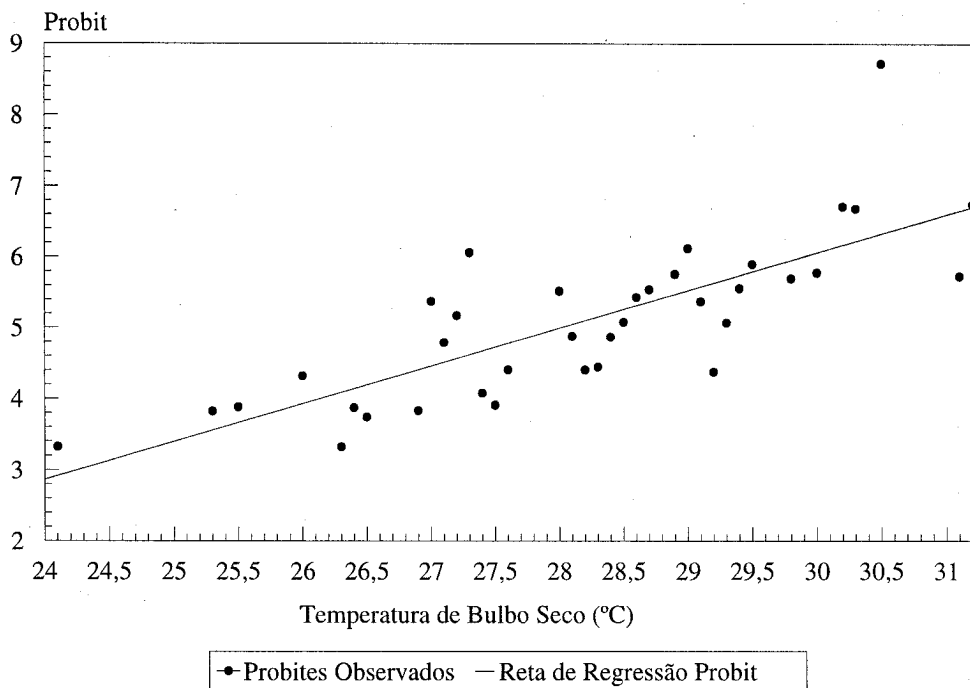


Figura 5. Gráfico do comportamento dos probites observados e estimados relativos à combinação dos graus de satisfação - Muito Quente (1) + Quente (2) - em função da temperatura de bulbo seco (°C).

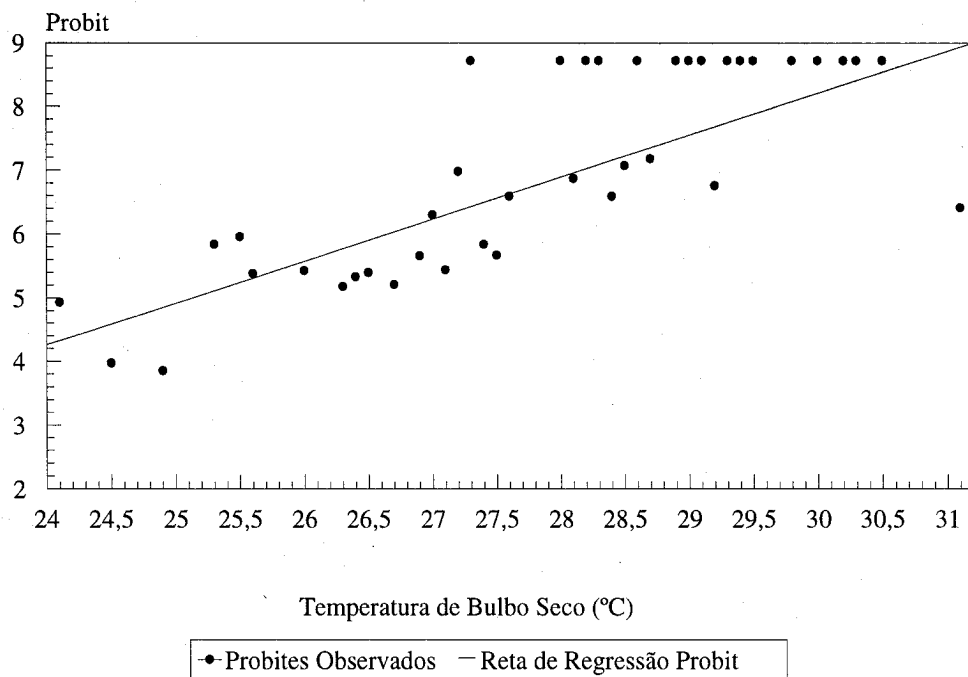


Figura 6. Gráfico do comportamento dos probites observados e estimados relativos à combinação dos graus de satisfação - Muito Quente (1) + Quente (2) + Confortável (3) - em função da temperatura de bulbo seco (°C).

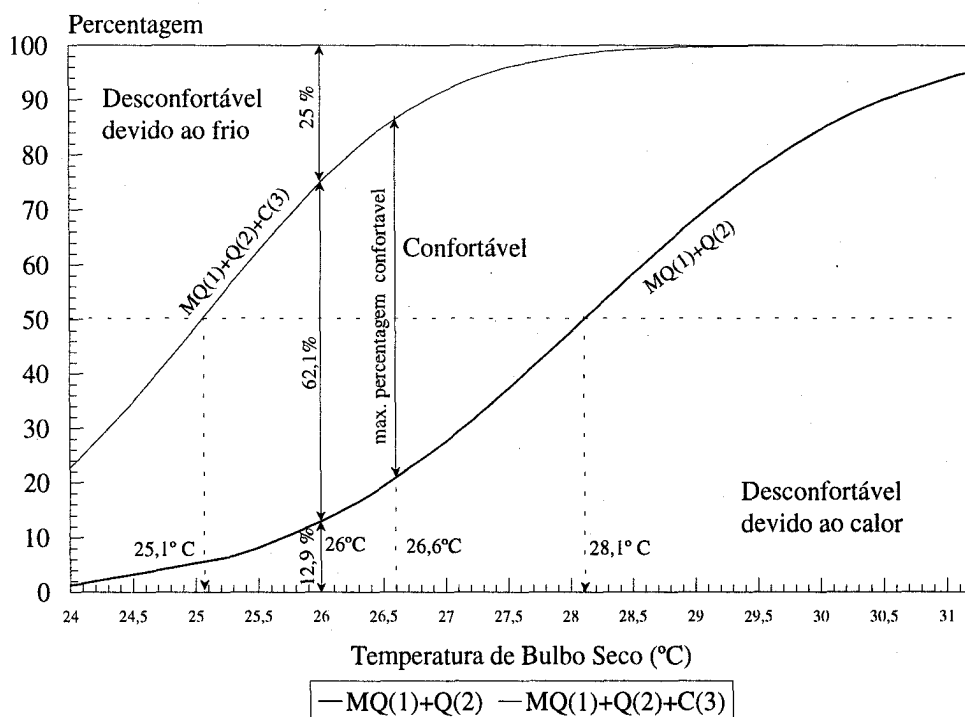


Figura 7. Limites da categoria confortável da temperatura.

As categorias muito frio, frio, confortável e quente também foram analisadas da mesma forma. Verificou-se que o modelo Probit também deveria ser usado para representar a relação. As categorias possuem um interseção em algum valor da temperatura. Os valores da temperatura que produzem valores altos da probabilidade numa determinada categoria, produzem valores baixos na categoria subsequente. Este mesmo resultado observou-se quando migra-se de uma categoria para outra, iniciando-se em muito quente e terminando-se em muito frio.

CONSIDERAÇÕES FINAIS

Observou-se na prática que este método estatístico de análise de dados através do Modelo Probit, demonstrou sua validade científica para medir a relação entre as variáveis ambientais e as sensações térmicas dos usuários das edificações.

REFERÊNCIAS BIBLIOGRÁFICAS

- AITKIN, M. et al. *Statistical Modelling in GLIM*. Oxford: Clarendon, 1989.
- ARANDA-ORDAZ, F.J. *On two families of transformations to additivity for binary response data*. *Biometrika*, 68: 357-363, 1981.
- ARAÚJO, Virgínia M.D. *Parâmetros de conforto térmico para usuários de edificações escolares no litoral nordestino brasileiro*. São Paulo: FAUUSP, 1996, Tese de Doutorado.
- CORDEIRO, G.M. *Modelos Lineares Generalizados*. Campinas: IV SINAPE, 1986.
- D.J. FINNEY, M.A., Sc.D., F.R.S. - *Probit analysis - A statistical treatment of the sigmoid response curve*. 2a. ed.. Cambridge: Cambridge University Press, 1964.
- FANGER, P.O. *Thermal Comfort. Analysis Engineering*. United States: Mc Graw Hill. Book Company, 1972.
- FISHER, R. A., YATES, F. - *Tabelas estatísticas: para pesquisa em biologia, medicina e agricultura*. Tradução de Salvador Licco Haim. São Paulo: Ed. da Universidade de São Paulo e Ed. Polígono, 1971.

ANEXO 1

Tabela 1. Distribuição das frequências absolutas e relativas observadas e estimadas dos probites por categoria do grau de satisfação, segundo a variação de temperatura de bulbo seco (□C).

TBS(°C)	Muito Quente + Quente MQ(1) + Q(2)					Muito Quente+ Quente + Confortável MQ(1) + Q(2) + C(3)				
	A	B	C	D	E	A	B	C	D	E
24,10	1	4,76	3,33	2,84	1,60	10	47,62	4,93	4,32	24,90
24,50				3,04	2,50	4	15,38	3,98	4,60	34,70
24,90				3,25	4,10	1	12,50	3,84	4,89	45,70
25,30	3	12,00	3,82	3,49	6,60	20	80,00	5,84	5,17	57,00
25,50	4	13,33	3,88	3,59	8,10	25	83,33	5,96	5,31	62,50
25,60				3,63	8,60	22	64,71	5,37	5,39	65,20
26,00	3	25,00	4,32	3,86	12,90	8	66,67	5,43	5,67	75,00
26,30	4	4,76	3,33	4,02	16,50	48	57,14	5,17	5,88	81,30
26,40	6	13,04	3,87	4,08	17,90	29	63,04	5,33	5,95	83,20
26,50	8	10,53	3,74	4,13	19,40	50	65,79	5,40	6,03	84,90
26,70				4,22	21,80	7	58,33	5,20	6,17	88,00
26,90	14	12,28	3,83	4,34	25,80	85	74,56	5,66	6,31	90,60
27,00	20	64,52	4,38	4,40	27,50	28	90,32	6,29	6,38	91,70
27,10	28	41,80	4,79	4,45	29,40	45	67,17	5,44	6,45	92,80
27,20	24	57,14	5,17	4,51	31,20	41	97,62	6,97	6,52	93,70
27,30	36	85,71	6,06	4,56	33,20	42	100,00	8,72	6,59	94,50
27,40	10	18,18	4,09	4,61	35,10	44	80,00	5,84	6,67	95,30
27,50	5	13,89	3,91	4,67	37,10	27	75,00	5,67	6,74	96,00
27,60	5	27,78	4,41	4,72	39,20	17	94,45	6,58	6,81	96,50
28,00	7	70,00	5,52	4,94	47,60	10	100,00	8,72	7,09	98,20
28,10	30	45,45	4,42	4,99	49,80	64	96,97	6,88	7,16	98,50
28,20	5	27,78	4,41	5,04	51,90	18	100,00	8,72	7,24	98,74
28,30	7	29,17	4,45	5,10	54,00	24	100,00	8,72	7,31	98,96
28,40	65	45,14	4,87	5,15	56,10	136	94,45	6,58	7,38	99,14
28,50	49	53,26	5,08	5,20	58,30	91	98,91	7,29	7,45	99,29
28,60	26	66,66	5,43	5,26	60,30	39	100,00	8,72	7,52	99,42
28,70	48	70,58	5,54	5,31	62,40	67	58,53	7,17	7,59	99,53
28,90	28	77,78	5,76	5,42	66,40	36	100,00	8,72	7,73	99,69
29,00	20	86,96	6,12	5,47	68,30	23	100,00	8,72	7,80	99,75
29,10	46	64,79	5,37	5,53	70,20	71	100,00	8,72	7,88	99,80
29,20	7	26,92	4,38	5,58	72,00	25	96,15	6,76	7,95	99,84
29,30	52	53,06	4,45	5,63	73,80	98	100,00	8,72	8,02	99,88
29,40	57	71,25	5,55	5,69	75,50	80	100,00	8,72	8,09	99,90
29,50	40	81,63	5,90	5,74	77,20	49	100,00	8,72	8,16	99,92
29,80	19	76,00	5,70	5,90	81,80	25	100,00	8,72	8,37	99,96
30,00	18	78,26	5,78	6,01	84,50	23	100,00	8,72	8,52	99,98
30,20	22	95,65	6,70	6,12	86,90	23	100,00	8,72	8,66	99,99
30,30	42	95,45	6,68	6,17	88,00	44	100,00	8,72	8,73	100,00
30,50	24	100,00	8,72	6,28	90,00	24	100,00	8,72	8,87	100,00
31,10	20	76,93	5,73	6,60	94,60	24	92,31	6,42	9,30	100,00
31,20	71	95,95	6,73	6,65	95,10	74	100,00	8,72	9,37	100,00

Fonte: Dados Coletados.