INSTITUTO SUPERIOR DE ECONOMIA E GESTÃO

Estatística II - Licenciatura em Gestão - 31 Maio 2011

Parte teórica

Nome: ____ Nº ________

1. **Perguntas de Verdadeiro/Falso** (2 valores) - Para cada afirmação, assinale se esta é Verdadeira (V) ou Falsa (F). Uma resposta certa vale 0.25 e uma resposta errada penaliza em idêntico valor.

	V	F
Num teste de dimensão $\alpha = 0.1$ em que o valor-*p*=0.07, rejeita-se H_0.	X	
$\hat{\theta}$ é estimador não enviesado de θ se e só se $\mathrm{var}(\hat{\theta}) = \mathrm{var}(\theta)$.		X
Se num processo repetitivo se forem construindo intervalos de confiança para θ com base em amostras casuais simples independentes umas das outras e todas de mesma dimensão, é legítimo concluir que a proporção de intervalos que vão incluir θ se vai aproximando do nível de confiança escolhido.	X	
O modelo $z_t = \exp\left(\alpha_1 + \alpha_2 x_{t2} + \alpha_3 x_{t3} + \alpha_4 x_{t3}^2\right) \times v_t$, $t = 1,2,\cdots,n$, é linearizável.	X	
Num teste do χ^2 à bondade do ajustamento a região de rejeição pode ser bilateral		X
Num teste de hipótese simples contra hipótese simples pode acontecer que $P(\text{Rej}\, H_0 \mid H_0) + P(\text{não Rej}\, H_0 \mid H_0) < 1$		X
A eficiência de um estimador é uma propriedade que apenas interessa em grandes amostras.		X
A dimensão de um teste é dada pela probabilidade de rejeitar a hipótese nula quando esta é falsa.		X

2. **Perguntas de resposta múltipla** (2 valores) - Para cada pergunta escolha **a** alternativa correcta. Uma resposta certa vale 0.5 valores e uma resposta errada penaliza em 0.17 valores.

1. A amplitude de um intervalo de confiança para a média de um universo normal de variância conhecida, construído com base no processo habitual, pode ser reduzida por
   - ☐ Um aumento da dimensão da amostra
   - ☐ Uma diminuição do nível de confiança
   - ✓ Qualquer das 2 primeiras alternativas
   - ☐ Nenhuma das 2 primeiras alternativas

2. Considere o MRL estimado pelos mínimos quadrados $\hat{Y}_t = b_1 + b_2 x_{t2} + b_3 x_{t3}$, em que Y representa o preço de uma propriedade agrícola em determinada região, x_2 a sua área e x_3 a distância à cidade mais próxima. Qual a estimativa da elasticidade preço-área?
   - ☐ b_2
   - ✓ $b_2 x_t / \hat{Y}_t$
   - ☐ $(b_2 + b_3) x_t / \hat{Y}_t$
   - ☐ b_3

3. A Desigualdade de Fréchet-Cramer-Rao é um resultado particularmente importante
   - ☐ Na estimação por intervalo
   - ☐ Nos testes de hipóteses
   - ✓ Na estimação por pontos
   - ☐ Para qualquer das alternativas

4. Estimado um MRL pelos mínimos quadrados obteve-se $\widehat{\ln y_t} = 5.2 - 0.32 \ln x_{t2} + 3.2 x_{t3}$. Neste quadro quando x_2 aumenta 0.1%, pode-se afirmar que, tudo o resto constante,
   - ☐ $E(y \mid X)$ diminui de 0.032 unidades aproximadamente
   - ☐ $E(y \mid X)$ decresce de 0.32% aproximadamente
   - ☐ $E(y \mid X)$ decresce de 3.2% aproximadamente
   - ✓ $E(y \mid X)$ decresce de 0.032% aproximadamente

3. Perguntas de desenvolvimento (2 valores) **–** Cada resposta certa vale 1 valor.

a. Defina o conceito de variável fulcral para o parâmetro θ e diga qual a sua utilidade.

Definição: ver defifinição 7.9 do livro.
Dois pontos importantes:

1. A variável fulcral só depende da amostra e do parâmetro para o qual se quer fazer inferência
2. A distribuição da variável fulcral não depende do parâmetro para o qual se quer fazer inferência

A utilidade da variável fulcral reside na construção de intervalos de confiança para o parâmetro.

b. Mostre que, quando se comparam dois modelos com a mesma variável dependente e estimados com base no mesmo conjunto de observações desta, é equivalente escolher aquele que tem maior coeficiente de determinação ajustado ou aquele que tem menor erro padrão da regressão (s^2). Recorda-se que para o modelo $y_t = \beta_1 + \beta_2 x_{t2} + \cdots + \beta_k x_{tk} + u_t$, $t = 1,2,\cdots,n$, $\overline{R}^2 = 1 - \frac{VR/(n-k)}{VT/(n-1)}$, sendo VR e VT as variações residuais e totais respectivamente, e $s^2 = VR/(n-k) = \sum_{t=1}^{n} \hat{u}_t^2 /(n-k)$

Tem-se assim $\overline{R}^2 = 1 - \frac{s^2}{VT/(n-1)}$.

Compare-se então o modelo A com o modelo B sabendo que $n_A = n_B = n$ e $VT_A = VT_B = VT$ (mesmo conjunto de observações do regressando).

$$\overline{R}_A^2 > \overline{R}_B^2 \Leftrightarrow 1 - \frac{s_A^2}{VT/(n-1)} > 1 - \frac{s_B^2}{VT/(n-1)} \Leftrightarrow \frac{s_B^2}{VT/(n-1)} > \frac{s_A^2}{VT/(n-1)} \Leftrightarrow s_B^2 > s_A^2$$

O que prova a a afirmação feita.

INSTITUTO SUPERIOR DE ECONOMIA E GESTÃO
Estatística II - Licenciatura em Gestão - 31 de Maio 2011

Nome:____ Nº ________

Alínea	**1a**	**1b**	**2**	**3**	**4a**	**4b**	**4c**	**4d**	**4e**	**4f**
Cotação	**15**	**15**	**15**	**15**	**10**	**10**	**15**	**15**	**15**	**15**
Classificação										

T:
P:

TOTAL:

Em todos os testes de hipóteses que fizer formule as hipóteses, indique a estatística de teste e sua distribuição. Assuma por defeito uma dimensão $\alpha = 0.05$.

1. A distância percorrida por um veículo pesado de mercadorias num dia de trabalho, em centenas de quilómetros, é uma variável aleatória, X, com média $\mu = \theta + 2$, variância $\sigma^2 = \theta^2$ e distribuição dada por

$f(x \mid \theta) = \frac{1}{\theta} e^{-\frac{x-2}{\theta}}$ para $x > 2$, onde θ é um parâmetro desconhecido ($\theta > 0$).

a) Mostre que o estimador de máxima verosimilhança para θ é $\hat{\theta} = \overline{X} - 2$ e analise a sua consistência.

$$L(\theta) = \prod_{i=1}^{n} f(x_i \mid \theta) = \prod_{i=1}^{n} \frac{1}{\theta} e^{-\frac{x_i - 2}{\theta}} = \frac{1}{\theta^n} e^{-\frac{\sum x_i - 2n}{\theta}} \quad \text{com } (x_i > 2) \text{ para i=1, 2, ..., n}$$

$$l(\theta) = \ln\left(\frac{1}{\theta^n} e^{-\frac{\sum x_i - 2n}{\theta}}\right) = -n \ln(\theta) - \frac{\sum x_i - 2n}{\theta}$$

$$\frac{d\, l(\theta)}{d\theta} = 0 \Leftrightarrow -\frac{n}{\theta} + \frac{\sum x_i - 2n}{\theta^2} = 0 \Leftrightarrow -n\theta + \sum x_i - 2n = 0 \Leftrightarrow \hat{\theta} = \frac{\sum x_i - 2n}{n} = \bar{x} - 2$$

com $\left[\frac{d^2\, l(\theta)}{d\theta^2}\right]_{\theta = \bar{x} - 2} < 0$. Assim o estimador de máxima verosimilhança será $\hat{\theta} = \overline{X} - 2$.

Consistência

$$E(\hat{\theta}) = E(\overline{X} - 2) = E(\overline{X}) - 2 = \mu - 2 = \theta + 2 - 2 = \theta$$

$$Var(\hat{\theta}) = Var(\overline{X} - 2) = Var(\overline{X}) = \frac{\sigma^2}{n} = \frac{\theta^2}{n}$$

Como $\lim_{n \to \infty} E(\hat{\theta}) = \theta$ e $\lim_{n \to \infty} Var(\hat{\theta}) = \lim_{n \to \infty} \frac{\theta^2}{n} = 0$ verificam-se as condições suficientes para a consistência

Então $\hat{\theta} = \overline{X} - 2$ é estimador consistente para θ.

b) Observada uma amostra casual de dimensão 50 obteve-se $\sum_{i=1}^{50} x_i = 350$. Apresente, justificando, a estimativa da máxima verosimilhança para a percentagem de dias de trabalho em que um veículo desse tipo percorre uma distância superior a 500 quilómetros.

Como $\sum_{i=1}^{50} x_i = 350$ então $\bar{x} = 7$

$$P(X > 5) = \int_5^{+\infty} \frac{1}{\theta} e^{-\frac{x-2}{\theta}} dx = \left[-e^{-\frac{x-2}{\theta}} \right]_5^{+\infty} = e^{-\frac{3}{\theta}}$$

Logo $\widehat{P(X>5)} = e^{-\frac{3}{\hat{\theta}}} = e^{-\frac{3}{\bar{x}-2}} = e^{-\frac{3}{5}} = 0.5488$

pela propriedade da invariância dos EMV

2. Numa empresa de aluguer de veículos o parâmetro fundamental para estabelecer a tarifa diária de aluguer dos veículos ligeiros de passageiros, na categoria de quilometragem ilimitada, é o número médio de quilómetros percorridos diariamente que, na tarifa actualmente praticada, se pressupõe não ultrapassar 275. Para avaliar se é necessário rever essa tarifa recolheu-se uma amostra casual de 500 alugueres nesta categoria, tendo-se observado uma média de 278.5 e uma variância corrigida de 6430.5. Recorra a um teste de hipóteses de dimensão $\alpha = 0.05$ para avaliar se é de proceder à revisão da tarifa diária deste tipo de aluguer.

$X -$ número médio de quilómetros percorridos diariamente, por aluguer

$H_0 : \mu \leq 275$ (não é necessário rever a tarifa) vs $H_1 : \mu > 275$ (necessário rever a tarifa).

Amostra casual : n = 500 (Grandes amostras) $\bar{x} = 278.5$ e $s'^2 = 6430.5$

Teste unilateral na aba direita $\Rightarrow W_{5\%} = \{z : z > 1{,}645\}$ onde $Z = \frac{\bar{X} - 275}{S'/\sqrt{n}} \overset{a}{\sim} N(0,1)$.

Como $z_{obs} = \frac{278.5 - 275}{\sqrt{\frac{6430.5}{500}}} = 0.975968 \notin W_{5\%}$ (ou pelo valor-p = $P(Z \geq 0.97596) = 0.16454 > 0.05$)

Não se rejeita Ho, logo não é necessário proceder à revisão da tarifa.

3. O técnico que está a realizar o estudo defende que existe associação entre a quilometragem média diária percorrida e o número de dias de aluguer. Para testar esta afirmação recorreu a um programa de estatística tendo obtido os seguintes resultados

Dias * Kms Crosstabulation

			Kms					Total
			até 220	220,0 – 250	250,0 – 300	300,0 - 375	mais de 375	
Dias	até 7	Count	18	14	55	86	36	209
		Expected Count	35.5	62.7	50.2	**? = 45.144**	**? = 15.466**	209.0
	8 ou +	Count	67	136	65	22	1	291
		Expected Count	49.5	87.3	69.8	**? = 62.856**	**? = 21.534**	291.0
Total		Count	85	150	120	108	37	500
		Expected Count	85,0	150.0	120.0	108.0	37.0	500.0

Chi-Square Tests

	Value	Df	Asymp. Sig.
Pearson Chi-Square	191.031[a]	4	.000
N of Valid Cases	500		

a. 0 cells (0%) have expected count less than 5. The minimum expected count is 15,47.

Complete, no próprio quadro, os valores em falta na tabela de contingência.

Com base no teste de hipóteses adequado diga, apresentando os procedimentos seguidos, se concorda ou não com a afirmação do técnico.

freq esp classe 14 – "dias até 7" e "300,0 a 375"; ...

$$fe_{14} = \frac{n_{1\circ}n_{\circ 4}}{n} = \frac{209 \times 108}{500} = 45.144 \qquad ; \; fe_{15} = 15.466 \, ; \; fe_{24} = 62.856 \, ; \; fe_{25} = 21.534$$

Teste de independência

$H_o : p_{ij} = p_{io} \times p_{oj}$ $(i = 1,2; j = 1,2,3,4,5)$ (independência) $H_1 : H_o$ *falsa* (existe associação).

$$Q = \sum_{i=1}^{2} \sum_{j=1}^{5} \frac{(N_{ij} - fe_{ij})^2}{fe_{ij}} \overset{a}{\sim} \chi^2_{(4)} \; (N_{ij} \text{ - freq da classe } ij; \; fe_{ij} = \frac{N_{i\circ}N_{\circ j}}{n} \text{ freq esperada da classe } ij)$$

$W_{5\%} = \{q : q > 9.488\}$ tabela $\chi^2_{(4)}$ $Q_{obs} = 191.031$ (output do SPSS – Pearson Chi-Square) $\in W_{5\%}$

Ou valor-p=0.000 < 0.05

Rejeita-se Ho logo,confirma-se a opinião do técnico da existência de associação entre a quilometragem média diária percorrida e o número de dias de aluguer.

4. A empresa "RentMais" com o propósito de explicar a quilometragem média diária percorrida pelos seus clientes de aluguer de veículos ligeiros de passageiros, construiu o seguinte modelo:

$$KMD_t = \beta_1 + \beta_2\, ID_t + \beta_3\, ND_t + \beta_4\, NAC_t \times ND_t + \beta_5\, NAC_t + U_t$$

em que, para o $t-ésimo$ aluguer da amostra, KMD representa a quilometragem média diária, ID a idade do condutor, ND o número de dias contratados e NAC uma variável artificial que assume o valor 1 se o cliente reside em Portugal e 0 caso contrário. Todas as alíneas se referem a este modelo embora possa utilizar os resultados em anexo sempre que necessário.

Em todas as regressões apresentadas o regressando é a quilometragem média diária (KMD).

Supondo satisfeitas as hipóteses do modelo de regressão linear obtiveram-se, utilizando o Excel, os seguintes resultados,

Regression Statistics	
Multiple R	0.7505766
R Square	0.5633653
Adjusted R Square	0.5598369
Standard Error	45.6817893
Observations	500

ANOVA

	Df	*SS*	*MS*	*F*	*Significance F*
Regression	4	1332794.623	333198.6559	159.6677	1.19327E-87
Residual	495	1032978.809	2086.8259		
Total	499	2365773.433			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	433.1389840	8.5095656	50.9002462	7.7136E-199
ID	-3.4989469	0.2009155	-17.4150212	2.47811E-53
ND	-2.4044767	0.6033652	-3.9851097	7.75984E-05
NDxNAC	-2.9164748	0.7348825	-3.9686275	8.29862E-05
NAC	49.6963674	8.4148690	5.9057803	6.52846E-09

a) Analise, com base no valor-p, a relevância estatística individual de cada um dos regressores do modelo. Construa um intervalo de confiança a 95% para β_5.

$H_o : \beta_j = 0$ contra $H_1 : \beta_j \neq 0$ Estatística de teste $T = \dfrac{b_j}{s_{b_j}} \sim t(495)$

Os valores apresentados na coluna do *P-value* são todos <0.05 (o mais elevado será 0.000082986), logo pode afirmar-se que todos os regressores são, individualmente, estatisticamente relevantes.

$$T = \frac{b_5 - \beta_5}{s_{b_5}} \sim t(495)$$

IC a 95% para $\beta_5 = (b_5 \pm t_{0.025} \times s_{b_5}) = (49.6963674 \pm 1.9648 \times 8.414869) = (33.1631,\ 66.2296)$

b) Interprete o valor das estimativas obtidas para β_2 e β_3.

$b_2 = -3.4989$. Um acréscimo de 1 ano na idade do condutor implica em média um decréscimo estimado de 3.4989, na quilometragem média percorrida diariamente, tudo o resto constante.

$b_3 = -2.4045$. **Para um condutor não residente em Portugal**, por cada dia a mais nos dias contratados no aluguer resulta em média um decréscimo estimado de 2.4045, na quilometragem média percorrida diariamente, tudo o resto constante.

c) Efectue o teste estatístico $H_o : \beta_4 = \beta_3$ contra $H_1 : \beta_4 \neq \beta_3$. Interprete o seu significado.

$H_0 : \beta_4 = \beta_3$ contra $H_1 : \beta_4 \neq \beta_3$

Estatística de teste $F = \dfrac{(VR_0 - VR_1)/1}{VR_1/495} \sim F_{(1,495)}$

$$F_{obs} = \frac{1033320.869 - 1032978.809}{1032978.809/495} = 0.1639$$

$valor - p = P(F > 0.1639) = 0.68575 > 0.05$, logo não se rejeita H_0 então poder-se-á afirmar que o impacto, na quilometragem média percorrida diariamente, de $+1$ dia contratado para os residentes em Portugal ($\beta_4 + \beta_3 = 2\beta_3$) é o dobro do dos não residentes (β_3).

d) Determine o modelo com a restrição imposta pela seguinte hipótese nula $H_o : \beta_4 = \beta_3 \wedge \beta_5 = 50$. Estimado o modelo reparametrizado obteve-se $R^2 = 0.620425$ e $\sum \hat{u}_t^2 = 1033902.06$. O que conclui sobre a veracidade da hipótese em questão.

Modelo reparametrizado com $H_0 : \begin{cases} \beta_4 = \beta_3 \\ \beta_5 = 50 \end{cases}$

$$KMD_t = \beta_1 + \beta_2\, ID_t + \beta_3\, ND_t + \beta_3\, NAC_t \times ND_t + 50\, NAC_t + U_t$$
$$KMD_t - 50\, NAC_t = \beta_1 + \beta_2\, ID_t + \beta_3\, (ND_t + NAC_t \times ND_t) + U_t$$

Os regressandos do modelo inicial (*KMD*) e do modelo reparametrizado (*KMD-50 NAC*) são diferentes logo para proceder ao teste de Ho tem de utilizar-se a estatística de teste

$$F = \frac{(VR_0 - VR_1)/2}{VR_1/495} \sim F_{(2,495)} \qquad F_{obs} = \frac{(1033902.06 - 1032978.809)/2}{1032978.809/495} = 0.221583$$

$valor - p = P(F > 0.221583) = 0.801329 > 0.05$

Não se rejeita Ho, logo não é de pôr em causa a hipótese em questão.

e) Calcule o intervalo de previsão a 95% para a quilometragem média diária efectuada por um cliente com 44 anos de idade, não residente em Portugal e que contratou um aluguer para 7 dias.

Intervalo de previsão a 95% para KMD_0 (para um caso individual, previsão pontual) de um aluguer em que $ID_0 = 44$, $ND_0 = 7$, $NAC_0 = 0$ e $(ND \times NAC)_0 = 0$.

Que será construído a partir da variável $T = \dfrac{KMD_0 - \hat{KMD}_0}{\sqrt{S_{\hat{\theta}}^2 + S^2}} \sim t_{(495)}$

Utilizando o output apresentado em anexo onde os regressores são $ID - 44$, $ND - 7$, NAC e $(ND \times NAC)$, obtém-se na linha referente a" Intercept": $\hat{KMD}_0 = 262.3539827$ (coluna "Coefficients") e $s_{\hat{\theta}} = 3.1035592$ (Coluna "Standard Error").

$t_{0.025} = 1.9648$ e $\sqrt{s_{\hat{\theta}}^2 + s^2} = \sqrt{3.1035592^2 + 2086.8295} = 45.78709378$

O IP a 95% para KMD_0: $(\hat{KMD}_0 \pm t_{0.025} \times \sqrt{s_{\hat{\theta}}^2 + s^2}) \Rightarrow (262.3539827 \pm 1.9648 \times 45.78709378) \Rightarrow$

$(172.391, 352.317)$

f) Para verificar se a variância da quilometragem média diária é constante estimou-se a regressão auxiliar conducente à realização do teste de White simplificado tendo-se obtido $R^2 = 0.142822$. Explicite qual a regressão referida e diga, formalizando convenientemente, as conclusões a tirar bem como as consequências nas análises feitas anteriormente, se as houver.

Verificar se existe homocedasticidade $\Leftrightarrow$ $\text{Var}(u_t) = \sigma^2$ para t = 1,2, ..., 500

Teste de White simplificado:

Regressão auxiliar $\hat{u}_t^2 = \alpha_1 + \alpha_2 K\hat{M}D_t + \alpha_3 K\hat{M}D_t^2 + \varepsilon_t$

$R^2 = 0.00137347$

Testar $H_0 : \alpha_2 = \alpha_3 = 0$ (homocedasticidade) contra $H_1 : \alpha_2 \neq 0 \vee \alpha_3 \neq 0$ (heterocedasticidade)

Estatística de teste: $W = 500R^2 \dot{\sim} \chi^2_{(2)} \rightarrow W_{obs} = 500 \times 0.142822 = 71.411$

$valor - p = P(W \geq 71.411) = 3.1138E - 16 \approx 0.0000$, rejeita-se a nulidade conjunta de α_2 e α_3, logo existe heteroscedasticidade no modelo.

Consequências: Os estimador dos mínimos quadrados para os β_j continuam a ser lineares e não enviesados mas deixam de ter variância mínima. Deixam de ser válidos os erros padrão de b_j bem como toda a inferência feita anteriormente.

ANEXO

Regression Statistics	
Multiple R	0.6589789
R Square	0.4342531
Adjusted R Square	0.4319765
Standard Error	51.8942758
Observations	500

ANOVA

	df	SS	MS	F	Significance F
Regression	2	1027344.548	513672.2738	190.742387	3.36205E-62
Residual	497	1338428.885	2693.0159		
Total	499	2365773.433			

	Coefficients	Standard Error	t Stat	P-value
Intercept	440.4983872	8.9268017	49.3456000	1.118E-193
ID	-4.1913316	0.2147825	-19.5143077	2.2055E-63
NAC	3.4037934	4.6712220	0.7286730	0.46654507

Regression Statistics	
Multiple R	0.7504803
R Square	0.5632207
Adjusted R Square	0.5605789
Standard Error	45.6432712
Observations	500

ANOVA

	df	SS	MS	F	Significance F
Regression	3	1332452.564	444150.8545	213.194982	8.15341E-89
Residual	496	1033320.869	2083.3082		
Total	499	2365773.433			

	Coefficients	Standard Error	t Stat	P-value
Intercept	434.4453685	7.8674256	55.2207785	5.6368E-214
ID	-3.4848862	0.1977242	-17.6249859	2.41362E-54
AUX1	-2.6323025	0.2175133	-12.1018021	1.01588E-29
NAC	47.1091507	5.4701807	8.6119918	9.60024E-17

Em que $AUX1 = ND + ND \times NAC$

V.S.F.F →

Regression Statistics	
Multiple R	0.7505766
R Square	0.5633653
Adjusted R Square	0.5598369
Standard Error	45.6817893
Observations	500

ANOVA

	df	SS	MS	F	Significance F
Regression	4	1332794.623	333198.6559	159.66768	1.19327E-87
Residual	495	1032978.809	2086.8259		
Total	499	2365773.433			

	Coefficients	Standard Error	t Stat	P-value
Intercept	262.3539827	3.1035592	84.5332632	2.5457E-296
ID-44	-3.4989469	0.2009155	-17.4150212	2.47811E-53
ND-7	-2.4044767	0.6033652	-3.9851097	7.75984E-05
NDxNAC	-2.9164748	0.7348825	-3.9686275	8.29862E-05
NAC	49.6963674	8.4148690	5.9057803	6.52846E-09

Nota: A variável $ID-44$ refere-se, naturalmente, a uma variável obtida subtraindo 44 a cada uma das observações de ID. A definição é semelhante para $ND-7$.

Regression Statistics	
Multiple R	0.7505766
R Square	0.5633653
Adjusted R Square	0.5598369
Standard Error	45.6817893
Observations	500

ANOVA

	Df	SS	MS	F	Significance F
Regression	4	1332794.623	333198.6559	159.6676845	1.19327E-87
Residual	495	1032978.809	2086.8259		
Total	499	2365773.433			

	Coefficients	Standard Error	t Stat	P-value
Intercept	291.6350264	4.108122813	70.989851	1.1924E-261
ID-44	-3.498946918	0.200915455	-17.41502123	2.47811E-53
ND-7	-2.404476699	0.603365245	-3.985109714	7.75984E-05
NDxNAC -7	-2.916474823	0.734882489	-3.968627454	8.29862E-05
NAC-1	49.69636743	8.414868961	5.905780312	6.52846E-09

Nota: A variável $ID-44$ refere-se, naturalmente, a uma variável obtida subtraindo 44 a cada uma das observações de ID. A definição é semelhante para $ND-7$, $ND\times NAC-7$ e $NAC-1$.