



## As crises financeiras mundiais e o PIB Brasileiro: uma aplicação da Regressão Logística

### Worldwide financial crisis and Brazilian GDP: an application of Logistic Regression

Bruno Magalhães Bourgard<sup>1</sup>

Carlos Francisco Simões Gomes<sup>2</sup>

**Resumo:** Após a década de 80, as crises financeiras mundiais causaram impactos nas variáveis econômicas e consequentemente no crescimento econômico brasileiro. Este trabalho tem como objetivo geral desenvolver uma metodologia que seja capaz de avaliar de maneira quantitativa o impacto das crises econômicas mundiais no crescimento econômico brasileiro através da análise de variáveis macroeconômicas. O trabalho utilizou 14 variáveis determinantes para o crescimento econômico: "exportações", "inflação IPCA", "crises", "M1", "M2", "M3", "M4", "taxa de desemprego", "produção industrial", "balança comercial", "taxa de juros", "importações", "consumo de energia elétrica nas indústrias" e a "taxa de câmbio". As estatísticas de ajuste do modelo, o pseudo R2 de Nagelkerke apresentou um poder de explicação 25,9% e a medida de Hosmer e Lemeshow de ajuste geral através de um teste estatístico indica que não houve diferença estatisticamente significativa entre as classificações observadas e previstas para o modelo final. A tabela de classificação do modelo mostrou uma taxa de acerto de 72,5%. Além da validação dessas estatísticas, com a finalidade de verificar a precisão das estimativas, foi realizada a avaliação cruzada (*cross validation*). Utilizando o método *k folders* e *leave one out* as taxas de acerto da validação cruzada ficaram em 71,6%, valores bem próximos aos obtidos na etapa de modelagem. Esse resultado demonstra a boa capacidade de generalização do modelo e a inexistência da necessidade de reajuste.

**Palavras-chave:** Crises Financeiras; Econometria; PIB.

<sup>1</sup> UFF – Universidade Federal Fluminense

<sup>2</sup> UFF – Universidade Federal Fluminense

**Abstract:** After the 1980s, global financial crises caused impacts on economic variables and consequently on Brazilian economic growth. The objective of this work is to develop a methodology that can quantitatively assess the impact of global economic crises on Brazilian economic growth through the analysis of macroeconomic variables. The study used 14 determinants of economic growth: "exports", "IPCA inflation", "crises", "M1", "M2", "M3", "M4", "unemployment rate", "industrial production", "Trade balance", "interest rate", "imports", "electricity consumption in industries" and "exchange rate". The adjustment statistics of the model, the pseudo R2 of Nagelkerke presented an explanatory power of 25.9% and the Hosmer and Lemeshow measure of general adjustment through a statistical test indicates that there was no statistically significant difference between the observed and predicted classifications for the final model. The model rating table showed a hit rate of 72.5%. In addition to the validation of these statistics, in order to verify the accuracy of the estimates, a cross-validation was performed. Using the k folders and leave one out method, the cross validation accuracy rates were 71.6%, values close to those obtained in the modeling step. This result demonstrates the good capacity of generalization of the model and the absence of the necessity of readjustment.

**Keywords:** Financial Crises; Econometrics; GDP.

---

## **1 Introdução**

### **1.1 Contextualização**

Uma crise econômica é uma perturbação, atribuída pela economia clássica a um desequilíbrio entre produção e consumo localizado em setores isolados da produção (Gomes *et al.*, 2016). Após a década de 80, as crises externas causaram impactos nas variáveis econômicas e como consequência no crescimento econômico dos países.

De acordo com Vidal (2011), uma crise financeira pode ser caracterizada como os desequilíbrios e recessões oriundos da busca pela riqueza, seja em termos de crescimento econômico, seja em patrimônio do investidor. Qualquer revés que ocorra em uma economia pode ser transmitido a outras. As crises podem ser: de inflação, choques cambiais, estouro de bolha de ativos, crises bancárias, crises de dívidas internas e externas (Reinhart e Rogoff, 2010).

Desequilíbrios iniciados em um país são rapidamente transmitidos para a economia dos outros, muitas vezes de maneira intensa e imprevisível. Esta situação pode ser caracterizada como crise financeira.

Fatores como o desenvolvimento da Tecnologia da Informação (TI) e a padronização dos mercados financeiros contribuem para que os mercados acionários internacionais estejam cada vez mais integrados.

Assim, as crises financeiras têm grande importância tanto na esfera econômica quanto na comercial e política, pois estas alteram a situação existente e forçam os governantes e investidores a tomarem decisões cruciais que terão novamente impacto na estrutura de outros países.

Acredita-se que a quantificação do impacto das crises financeiras mundiais sobre as variáveis macroeconômicas mais relevantes e consequentemente sobre o crescimento econômico brasileiro (mensurado pela variação do PIB) possa contribuir para o melhor entendimento sobre o tema e a formulação de políticas econômicas que minimizem os impactos das crises na economia nacional.

## **2 Definição das variáveis e Modelagem Estatística**

Foram observadas as variáveis mais comumente utilizadas pela literatura de indicadores relacionados às crises financeiras. Os dados referentes à economia brasileira foram coletados mensalmente e compreendem o período entre janeiro de 1990 a fevereiro de 2016 (devido a disponibilidade dos dados) contemplando assim 313 observações. A Tabela 1 apresenta essas variáveis e suas respectivas fontes de dados.

**Tabela 1** – Resumo das variáveis utilizadas no estudo

Variável	Legenda	Utilização	Tipo	Fonte	Unidade de Medida
PIB	-	Dependente	Binária	Banco Central	R\$ milhões
Crises	CRISES	Independente	Binária	Gerada pelo autor	-
Taxa de Desemprego	TD	Independente	Contínua	IBGE	Percentual
Consumo de energia elétrica das Indústrias	CEEI	Independente	Contínua	IPEA	GWh
Importações	IMP	Independente	Contínua	Banco Central	US\$
Exportações	EXP	Independente	Contínua	Banco Central	US\$
Produção Industrial	PI	Independente	Contínua		
Inflação	INF	Independente	Contínua	IPEA	Percentual
Balança Comercial	BC	Independente	Contínua	Banco Central	US\$ (mil)
Taxa de Juros	TJ	Independente	Contínua	Banco Central	Percentual
Taxa de Câmbio	TC	Independente	Contínua	Banco Central	u.m.c./US\$
M1	M1	Independente	Contínua	Banco Central	Real (mil)
M2	M2	Independente	Contínua	Banco Central	Real (mil)
M3	M3	Independente	Contínua	Banco Central	Real (mil)
M4	M4	Independente	Contínua	Banco Central	Real (mil)

## 2.1 Modelagem Estatística

Foram observadas mensalmente as variáveis anteriormente descritas no período de janeiro de 1990 a fevereiro de 2016. Foram utilizadas para a modelagem estatística as variáveis selecionadas entre diversos estudos sobre o tema, (12 variáveis), o agregado monetário M4 e a variável binária crises (gerada pelo autor, 1 para o caso do período analisado estar ocorrendo alguma crise, ou 0 para o caso do período analisado não estar ocorrendo alguma crise) totalizando 14 variáveis explicativas. A inclusão da variável M4 foi realizada para que pudesse ser feita uma análise de qual o agregado monetário possui maior influência sobre o crescimento econômico brasileiro.

### 2.1.1 Regressão Logística Binária

Uma vantagem da regressão logística é a sua utilização, mesmo em casos onde a resposta de interesse não é originalmente do tipo binária. Dessa maneira, muitos pesquisadores têm dicotomizado a resposta de modo a usar a regressão logística.

Neste estudo a primeira etapa da modelagem logística é dicotomizar a variável resposta PIB da seguinte maneira:

$$\text{PIB} = \begin{cases} 1 : \text{Caso o PIB seja maior que o mês anterior - Crescimento econômico} \\ 0 : \text{Caso contrário - Sem Crescimento econômico} \end{cases}$$

Após a dicotomização da variável resposta, é preciso definir qual o método a ser utilizado para adição ou remoção de variáveis. O critério para adição ou remoção de variáveis em regressão linear é geralmente baseado na estatística F, comparando os modelos com e sem as variáveis em análise. Em regressão logística, os erros seguem uma distribuição binomial sendo baseados no teste de razão de verossimilhança (Silva 2011). Para a execução dessa etapa foi utilizado o software SPSS 21. Optou-se pelo método *Enter* de seleção das variáveis por considerar que todas as variáveis são necessárias para se estimar os parâmetros do modelo.

## 2.2 Avaliação da Multicolineariedade

Segundo Corrar *et al.* (2007), a técnica de regressão logística se destaca pela possibilidade de contornar certas restrições encontradas em outros modelos multivariados. Entretanto, o modelo de regressão logístico é sensível a colinearidade entre as variáveis (Hair *et al.* 1998). Segundo Neter *et al.* (2005), as seguintes técnicas de diagnóstico podem ser utilizadas para avaliar indícios da presença de multicolineariedade;

- Duas ou mais variáveis independentes são correlacionadas;
- O sinal ou ordem de grandeza das estimativas dos coeficientes são diferentes do esperado;
- Ocorrem grandes mudanças nas estimativas dos coeficientes quando variáveis são adicionadas ou excluídas do modelo;

Pela Tabela 2 que apresenta as correlações das variáveis, pode-se verificar valores altos dos coeficientes de correlação tais como, 0,952 entre a taxa de juros e a inflação e os coeficientes entre o consumo de energia elétrica das indústrias e as exportações e importações (0,94 e 0,90 respectivamente). Os coeficientes altos entre os agregados monetários já era esperado por conta da própria definição dos mesmos (os agregados são combinações entre si), porém deseja-se obter a informação de qual desses agregados servirá como melhor preditor para o crescimento econômico.

**Tabela 2 – Matriz de correlação das variáveis**

	CRISES	TD	PI	BC	TJ	INF	TC	M1	M2	M3	M4	CEEI	EXP	IMP	
CRISES	1	,166**	,407**	-	-	-	-,098	-,103	-,095	-,095	-,096	-,067	-	-	
TD			-,099	,538**	-	-	-	-	-,117*	-,117*	-,118*	,261**	,051	-,086	
PI				1	-,060	-,072	-,044	-,050	-,041	-,042	-,042	-	-	-	
BC					1	-,024	-,014	,001	,001	,001	,001	,280**	,254**	,004	
TJ						1	,952**	,367**	,381**	,360**	,360**	,360**	-	-	
INF							1	,341**	,353**	,336**	,335**	,336**	-	-	
TC								1	,998**	1,000**	1,000**	,999**	-	-	
M1									1	,996**	,996**	,996**	-	-	
M2										1	1,000**	1,000**	-	-	
M3											1	1,000**	-	-	
M4												1	-	-	
CEEI													1	-	
EXP														1	
IMP															1

CEEI										1	,940**	,901**
EXP											1	,968**
IMP												1

Como essas correlações são possíveis indicativos de multicolineariedade, o estudo optou por realizar uma análise da contribuição de cada variável para a multicolineariedade através do *varuance inflation factor (VIF)*.

A tabela 3 calcula o VIF para cada variável. Valores maiores que 5 significam que o preditor infla a variância dos coeficientes de regressão estimados em comparação à quando as variáveis preditoras não são relacionadas linearmente. Dessa maneira deve ser excluída do modelo a variável em que  $VIF > 5$ .

**Tabela 3** – Análise do *VIF* para todas as variáveis.

VARIÁVEIS	VIF
CRISES	1,37
DESEMPREGO	2,93
PRODUÇÃO_INDUSTRIAL	1,79
BALANÇA_COMERCIAL	98,63
TAXA_DE_JUROS	11,61
INFLAÇÃO	10,38
TAXA_DE_CÂMBIO	3.437,13
M1	389,86
M2	24.664,95
M3	1.976,99
M4	27.082,94
CEEI	19,51
EXPORTAÇÕES	1.524,40
IMPORTAÇÕES	1.481,78

Pela tabela 3 percebe-se que diversas variáveis apresentam alto valor de *VIF*, devendo ser excluídas do modelo. Esse processo de exclusão deve ser realizado de uma em uma variável, priorizando as variáveis que apresentam maior *VIF*. Dessa maneira, deve-se excluir primeiramente a variável M4. Esse processo é realizado até que todas as variáveis tem seu *VIF* menor que 5. O resultado final desse processo é apresentado na tabela 4. Um ponto importante é a definição do ponto de corte. Nesse estudo optou-se por um ponto de corte de 0,5, que é o utilizado quando não se conhece a proporção populacional. (Karam, 2006)

**Tabela 4** – Variáveis selecionadas através da Análise do *VIF*

VARIÁVEIS	VIF
CRISES	1,31
DESEMPREGO	1,96
PRODUÇÃO_INDUSTRIAL	1,54
BALANÇA_COMERCIAL	1,72
INFLAÇÃO	1,80
M3	1,14
IMPORTAÇÕES	1,56

### 2.3 Modelo Proposto

As variáveis selecionadas no item anterior serão utilizadas para a obtenção do modelo que melhor se ajuste aos dados. Na tabela 5 pode-se visualizar o resumo das estimativas do modelo. O método de seleção da variáveis utilizado foi o *enter*, em virtude disso no modelo inicial proposto todas as variáveis são consideradas. Na tabela 5 pode-se verificar que algumas variáveis não são estatisticamente significativas ao nível de confiança de 90%. É necessária a exclusão dessas variáveis uma de cada vez, priorizando a exclusão da variável que possuir o maior Sig. Portanto a variável taxa de desemprego será excluída e o modelo gerado novamente.

**Tabela 5** – Resumo do Modelo Inicial

Variável	B	S.E.	Wald	Df	Sig.	Exp(B)
CRISES	-0,598	0,332	3,246	1	0,072	0,55
DESEMPREGO	-0,019	0,092	0,043	1	0,835	0,981
PRODUÇÃO INDUSTRIAL	0,049	0,013	14,644	1	0	1,05
BALANÇACOMERCIAL	0	0	11,689	1	0,001	1
INFLAÇÃOIGPM	0,098	0,064	2,348	1	0,125	1,103
M3	0	0	1,16	1	0,281	1
IMPORTAÇÕES	0	0	1,61	1	0,204	1
Constante	-5,513	1,797	9,408	1	0,002	0,004

**Tabela 6** – Resumo final do Modelo Inicial

Variável	B	S.E.	Wald	Df	Sig.	Exp(B)
CRISES	-0,641	0,322	3,959	1	0,047	0,527
PRODUÇÃO INDUSTRIAL	0,047	0,012	14,827	1	0	1,048
BALANÇACOMERCIAL	0	0	15,733	1	0	1,037
INFLAÇÃOIGPM	0,18	0,047	14,599	1	0	1,197
IMPORTAÇÕES	0	0	5,131	1	0,024	1,059
Constante	-5,534	1,449	14,589	1	0	0,004

A tabela 6 apresenta o modelo onde todas as variáveis são estatisticamente significativas ao nível de confiança de 95%. Entre todas as variáveis selecionadas para o estudo as estatisticamente significativas ao nível de confiança de 95% são: crises, produção industrial, balança comercial, inflação e importações. Antes da análise dos coeficientes dessa variáveis e suas respectivas estatísticas apresentadas o modelo terá o seu ajuste avaliado.

### 2.4 Avaliação do Ajuste

O resultado do modelo do bloco inicial (tabela 7) apresenta a tabela de classificação considerando o modelo com apenas uma constante (etapa 0), ou seja, se arbitrariamente todos os meses fossem classificados como crescimento econômico, a taxa de acerto seria de 64,2%.

**Tabela 7** – Tabela de classificação do bloco inicial

Observado			Previsto		
			BINARIO		Porcentagem correta
			0	1	
Etapa 0	BINARIO	0	0	112	0,0
		1	0	201	100,0
Porcentagem global					64,2

O modelo de regressão logística que irá estimar a probabilidade de crescimento econômico precisa ser mais assertivo do que o modelo inicial. O resultado do modelo final (tabela 8) apresenta a tabela de classificação do modelo gerado (etapa 1), a taxa de acerto seria de 72,5% o que significa que o modelo final melhorou a classificação aleatória que era de 64%.

**Tabela 8** – Tabela de classificação do bloco do modelo

Observado			Previsto		
			BINARIO		Porcentagem correta
			0	1	
Etapa 1	BINARIO	0	47	65	42,0
		1	21	180	89,6
Porcentagem global					72,5

Outra medida de avaliação do ajuste do modelo são os pseudo  $R^2$  de Cox & Snell e Nagelkerk, conforme apresentado na tabela 9 os indicadores dessas estatísticas são satisfatórios.

**Tabela 9** – Tabela de resumo dos pseudos  $R^2$

Resumo do modelo			
Etapa	Verossimilhança de log -2	R quadrado Cox & Snell	R quadrado Nagelkerke
1	342,850 <sup>3</sup>	0,189	0,259

A medida Hosmer e Lemeshow de ajuste geral tem um teste que indica que não houve diferença estatisticamente significativa entre as classificações observadas e previstas para todos os modelos com duas ou mais variáveis. Para tal é gerada uma tabela de contingência (tabela 10) que compara os valores observados com os valores esperados para cada estado da variável resposta.

**Tabela 10** – Tabela de contingência para o teste de Hosmer e Lemeshow

Tabela de contingência para teste de Hosmer e Lemeshow

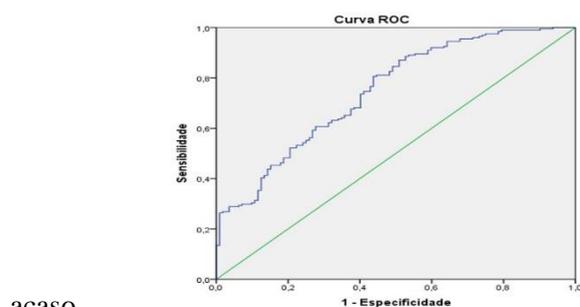
	BINARIO = 0		BINARIO = 1		Total	
	Observado	Esperado	Observado	Esperado		
Etapa 1	1	26	22,108	5	8,892	31
	2	19	17,482	12	13,518	31
	3	13	14,947	18	16,053	31
	4	9	13,721	22	17,279	31
	5	10	12,362	21	18,638	31
	6	12	10,938	19	20,062	31
	7	8	9,203	23	21,797	31
	8	8	7,116	23	23,884	31
	9	6	3,752	25	27,248	31
	10	1	0,37	33	33,63	34

O valor de Hosmer e Lemeshow mede a correspondência dos valores efetivos e previstos da variável dependente. Neste caso, o melhor ajuste do modelo é indicado por uma diferença menor na classificação observada e prevista. Um bom ajuste de modelo é indicado por um valor chi-quadrado não significativo (Hair, 1998). A medida Hosmer e Lemeshow (tabela 11) indica ausência de diferença significativa na distribuição de valores dependentes efetivos e previstos.

Tabela 11 – Teste de Hosmer e Lemeshow

Etapa	Qui-quadrado	df	Sig.
1	9,988	8	0,266

Outra análise do ajuste é realizada através da curva ROC (*Receiver operating characteristic*) (gráfico 1). Quanto mais afastada a curva em relação a uma curva de referência de 45° melhor é o ajuste do modelo. Uma curva muito próxima a curva de referência demonstra que a capacidade do modelo para discriminar entre a ocorrência e a não ocorrência é devido ao



acaso.

Gráfico 1 - Curva da *Receiver Operating Characteristic*

Fonte: Elaborado pelo autor, 2016.

Pela análise gráfica pode-se concluir pela rejeição da hipótese de ajuste do modelo não ser adequado aos dados. Observando a tabela 12, formalmente a hipótese de ajuste do modelo não ser adequado aos dados é rejeitada.

Tabela 12 – Teste de Significância da Curva ROC

Variável(eis) de resultado de teste: Probabilidade prevista

Área	Modelo padrão <sup>a</sup>	Sig. assintótico <sup>b</sup>	Intervalo de confiança assintótico 95%	
			Limite inferior	Limite superior
0,746	0,029	0	0,69	0,802

## 2.5 Precisão da Estimativa

Após o modelo de regressão logístico ser validado, se faz necessário a avaliação da precisão da estimativa, ou seja, o quão seria preciso o modelo gerado se aplicado em um novo conjunto de dados. Para atingir esse objetivo foi utilizado a técnica de *cross validation* com os métodos *k-folders* e *leave one out*.

- *CROSS VALIDATION*: De acordo com o estudo de Silva et.al (2016), o método consiste em dividir em grupos aleatórios e uniformes a base de dados (em k grupos), sendo todos de mesmo tamanho aproximadamente. Neste método o primeiro grupo fica como grupo de teste ou validação e os demais para treinamento. Em um segundo momento o segundo grupo é requerido para validação enquanto os demais são utilizados para treinamento e assim sucessivamente de forma que se obtém uma estimativa do erro para o conjunto de amostras. Quando **igual**-se o valor de k (subconjuntos) a um **tem**-se o método *leave one out*

**Tabela 13** – Taxa de acerto *Cross Validation k folders = 10*

Folder	Percentual de Acerto
1	81,3%
2	71,0%
3	74,2%
4	74,2%
5	78,1%
6	64,5%
7	64,5%
8	67,7%
9	77,4%
10	71,9%
Total	72,5%

A tabela 13 apresenta o percentual de acerto global utilizando o *k folders* de tamanho igual a 10, por essa tabela **pode**-se observar que a taxa de acerto global não se altera quando comparados ao modelo final proposto.

**Tabela 14** – Alocação das observações no *Cross Validation k folders = 10*

Folder	Observações Presentes
1	2, 12, 22, 35, 64, 76, 83, 88, 95, 111, 123, 130, 138, 142, 143, 157, 163,

	176, 183, 190, 193, 207, 220, 233, 234, 242, 254, 264, 285, 288, 305, 312
2	6, 11, 24, 31, 44, 46, 55, 60, 61, 71, 72, 75, 85, 97, 99, 100, 104,
	105, 110, 129, 140, 144, 188, 192, 219, 239, 244, 248, 266, 278, 310,
3	1, 3, 4, 14, 23, 39, 62, 66, 80, 91, 108, 116, 133, 158, 159, 161, 199,
	200, 205, 226, 228, 246, 263, 271, 275, 277, 281, 286, 289, 298, 308,
4	7, 8, 19, 28, 45, 63, 82, 86, 107, 114, 121, 141, 149, 173, 196, 203, 204,
	209, 214, 223, 224, 249, 252, 253, 255, 256, 262, 287, 294, 299, 313,
5	5, 15, 16, 20, 25, 33, 43, 52, 53, 67, 69, 73, 77, 79, 84, 106, 115,
	125, 134, 164, 171, 172, 185, 197, 208, 222, 225, 267, 284, 291, 302, 309,
6	48, 57, 59, 96, 119, 122, 124, 135, 145, 147, 150, 154, 162, 166, 179, 181, 198,
	211, 217, 229, 251, 257, 260, 261, 274, 276, 280, 293, 304, 306, 206,
7	9, 40, 42, 58, 65, 70, 74, 89, 90, 94, 101, 103, 117, 152, 155, 156, 178,
	184, 202, 227, 236, 237, 245, 250, 259, 269, 282, 283, 295, 300, 311,
8	30, 36, 37, 47, 50, 54, 56, 92, 113, 118, 127, 131, 148, 167, 168, 169, 170,
	186, 187, 191, 194, 195, 212, 221, 230, 235, 240, 243, 279, 292, 303,
9	13, 21, 26, 29, 32, 34, 38, 51, 68, 78, 81, 102, 120, 126, 136, 137, 139,
	151, 175, 180, 182, 210, 213, 215, 241, 247, 258, 272, 273, 290, 307,
10	10, 17, 18, 27, 41, 49, 87, 93, 98, 109, 112, 128, 132, 146, 153, 160, 165,
	174, 177, 189, 201, 216, 218, 231, 232, 238, 265, 268, 270, 296, 297, 301,

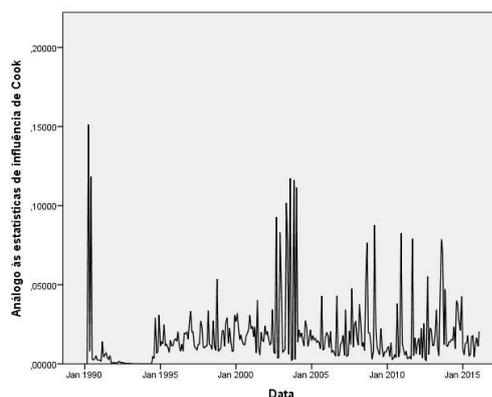
**Tabela 15** – Taxa de acerto Global do *Cross Validation - leave one out*

Observado		Previsto			
		BINARIO		Porcentagem correta	
		0	1		
Etapa 1	BINARIO	0	46	66	41,1%
		1	23	178	88,6%
Porcentagem global					71,6%

Na tabela 14 é apresentada a alocação de cada observação da base de dados nos 10 *folders*. A tabela 15 apresenta o percentual de acerto global utilizando o *leave one out*. Por essa tabela pode-se observar que a taxa de acerto global e individuais de acerto não se alteraram muito quando comparados ao modelo final proposto. A taxa de acerto muda de 72,5% (modelo final) para 71,6% (*leave one out*), as taxas de acerto quando a variável binária assume o valor 0 se altera de 42% (modelo final) para 41,1% (*leave one out*) e as taxas de acerto quando a variável binária assume o valor 1 se altera de 89,6% ( modelo final) para 88,6% (*leave one out*). Pelos valores apresentados, pode-se concluir que o modelo possui uma boa capacidade de generalização podendo ser utilizado para produzir boas estimativas.

## 2.6 Avaliação dos *Outliers*

Conforme foi mencionado na análise descritiva das variáveis, alguns preditores possuem indícios de *outliers*. Essa etapa da modelagem consistirá na avaliação do impacto desses valores atípicos para a geração do modelo. Para verificar se esses valores são influentes para a regressão esse estudo utilizou a Distância de Cook.



**Gráfico 2** – Distância de Cook.

Fonte: Elaborado pelo autor, 2016.

De acordo com o gráfico 2, o modelo final proposto não apresentou nenhum valor da Distância de Cook maior que 1, por esse fato **pode-se** concluir que os possíveis valores discrepantes observados não possuem grande influência sobre a elaboração do modelo.

## 2.7 Interpretação dos Resultados

O modelo de regressão logístico selecionou as 5 variáveis que melhor explicam o crescimento econômico. Gerado o modelo, atribui-se uma probabilidade de crescimento econômico para cada período da base de dados. Através do modelo de regressão logística pode-se observar que em períodos de crise a probabilidade de crescimento econômico se reduz em 47,3%. As variáveis mais impactantes no crescimento econômico brasileiro são: crises, inflação, produção industrial, balança comercial e importações. Outros resultados interessantes podem ser obtidos analisando os coeficientes estimados para cada variável do modelo apresentado na tabela 16.

**Tabela 16** – Resumo das variáveis do modelo final

	B	S.E.	Wald	df	Sig.	Exp(B)	Intervalo de Confiança	
							Inferior	Superior
CRISES	-,641	,322	3,959	1	,047	,527	,280	,990
PRODUÇÃO INDUSTRIAL	,047	,012	14,827	1	,000	1,048	1,023	1,073
BALANÇA COMERCIAL	,000	,000	15,733	1	,000	1,037	1,000	1,000
INFLAÇÃO IGPM	,180	,047	14,599	1	,000	1,197	1,092	1,313
IMPORTAÇÕES	,000	,000	5,131	1	,024	1,059	1,000	1,000

Constante	-5,534	1,449	14,589	1	,000	,004	
-----------	--------	-------	--------	---	------	------	--

A primeira coluna da tabela 16 fornece a estimativa dos coeficientes das variáveis explicativas obtidas. Através do sinal de cada coeficiente pode-se avaliar o sentido do impacto na variável resposta. O sinal negativo no coeficiente estimado da variável crises significa que as crises financeiras contribuem para a redução do crescimento econômico. Esse foi o único coeficiente estimado com sinal negativo.

A segunda coluna da tabela 16 fornece o desvio padrão das estimativas dos coeficientes, quanto menor for o desvio padrão melhor é a estimativa. A estatística de Wald apresentada na tabela 16 fornece a significância estatística para cada coeficiente estimado, de modo que o teste de hipóteses pode ocorrer como acontece na regressão múltipla (Hair *et al.*2009)

Os graus de liberdade que também são apresentados na tabela 16 fornecem uma medida de quão restritos estão os dados para alcançar um certo nível de previsão. Se o número de graus de liberdade for pequeno a previsão resultante pode ser menos generalizável (Hair *et al.*2009)

Em seguida a tabela 16 fornece a Razão de Chances que é a forma de deduzir a interpretação dos coeficientes estimados no modelo de regressão logística. Ela indica, para o caso de variáveis dicotômicas, qual a chance do evento de interesse ocorrer quando a variável preditora tem a característica testada em relação à chance do evento ocorrer caso a variável preditora não tenha tal característica (Fávero *et al.*, 2009). Para variáveis nominal e numérica, a Razão de Chances indica qual o aumento (ou a diminuição) da chance do evento ocorrer, após o aumento de uma unidade (ou nível) na variável explicativa

A análise da razão de chances se aplica a variável dicotômica crises. Com razão de chances de 0,527 pode-se concluir que em períodos de crises a probabilidade de crescimento econômico brasileiro se reduz em 47,3 %. Em seguida, pode-se analisar a razão de chances da variável produção industrial, um aumento em 1pp na produção industrial ocasiona um aumento de 0,5% na probabilidade de crescimento econômico brasileiro. A balança comercial também se mostrou significativa para o crescimento econômico com sua razão de chances apontando um aumento de 3,7% na probabilidade de crescimento econômico para cada variação de US\$ 1 no saldo da balança comercial. Outro resultado obtido mediante a razão de chances é o da variável inflação, o aumento em 1pp na taxa de inflação acarreta um aumento de 19,7% no crescimento econômico brasileiro.

A última variável significativa para o modelo é importações, sua razão de chances permite concluir que o aumento em 0,6% na probabilidade de crescimento econômico para o aumento de US\$ (mil) em importações. A última coluna da tabela 16 fornece o intervalo de

confiança para razão de chances calculadas. Uma menor amplitude no intervalo de confiança é desejada para a estimativa.

Utilizando o modelo de regressão logístico nos períodos das crises destacadas para o estudo pode-se obter as probabilidades ou chances do crescimento econômico brasileiro para cada crise. Esse resultado pode ser interpretado como um indicativo do tamanho do impacto ocasionado por cada crise estudada no crescimento do PIB. Quanto menor for a probabilidade ou chance de crescimento econômico no período, mais impactante foi a crise estudada.

Para facilitar o entendimento foi elaborado pelo autor a tabela 17, que apresenta um resumo das crises estudadas e o resultado obtido pelo modelo de regressão logística (probabilidade de crescimento econômico).

**Tabela 17** – Tabela de Probabilidades de crescimento econômico das crises

Crise	País de origem	Marco de início da crise	Período de tranquilidade	Período de crise	Probabilidade de crescimento econômico
Crise da Ásia	Hong Kong	1997	01/06/1996 a 16/10/1997	17/10/1997 a 05/03/1998	37,6%
Crise da Rússia	Rússia	Primeira queda do índice da bolsa no período. Moratória Russa	01/01/1998 a 26/07/1998	27/07/1998 a 16/09/1999	38,1%
Crise Brasileira	Brasil	1999 Colapso Cambial no Brasil	20/09/1998 a 26/07/1998	04/01/1999 a 15/03/1999	43,3%
Crise da bolha da Internet	EUA	Primeira queda do índice NASDAQ no período	04/01/1999 a 10/03/2000	11/03/2000 a 23/04/2001	52,1%
Crise turca	Turquia	Rápida corrosão de sua moeda	Anterior a 1996	2000-2001	56,4%
11 de setembro	EUA	Ataque terrorista ao WTC	25/04/2001 a 10/09/2001	11/09/2001 a 08/04/2003	65,9%
Crise da Argentina	Argentina	Primeira queda do índice da bolsa no período. Quebra do sistema bancário	01/01/2001 a 11/0/2001	12/07/2001 a 24/05/2002	57,4%
Crise Brasileira de 2002	Brasil	Primeira queda do índice da bolsa no período	01/01/2001 a 12/09/2001	13/09/2001 a 02/02/2003	65,5%
Crise do Subprime	EUA	Primeira queda do índice da bolsa no período	09/10/2005 a 25/07/2007	26/07/2007 a 17/03/2009	50,2%

De acordo com a tabela 17 pode-se concluir que dentre as crises observadas neste estudo a crise mais impactante foi a crise da Ásia, esse resultado pode ser comparado ao resultado obtido por Vidal (2011) que ressalta que essa afirmação deve ser ponderada pelo período vivenciado pela economia internacional. Logo em seguida a crise da Rússia se mostrou a segunda crise mais impactante, o seu período de crise é muito próximo ao período de crise asiática, talvez os impactos da crise asiática ainda permanecessem e ajudaram a intensificar a crise

rusa. Em seguida as crises internacionais mais impactantes foram: a Crise do Subprime ocorrida nos EUA, seguida pela crise da bolha da internet.

### 3 Conclusão

Esta pesquisa teve como objetivo geral desenvolver uma metodologia que seja capaz de avaliar de maneira quantitativa o impacto das crises mundiais no crescimento econômico brasileiro, mediante a análise de variáveis macroeconômicas. Foi proposto no presente trabalho o desenvolvimento de um modelo de regressão logística binária

As variáveis econômicas mais relevantes para o estudo: a taxa de desemprego, exportações, inflação (medida pelo IPCA), os agregados monetários M1, M2 e M3, produção industrial, balança comercial, taxa de câmbio, importações, consumo de energia elétrica nas indústrias e a taxa de juros. Foram incluídas mais duas variáveis, a variável crises (binária, 1 em períodos de crises) e o agregado monetário M4.

O modelo estatístico de regressão logística se mostrou capaz de indicar a contribuição de cada uma das variáveis selecionadas no crescimento econômico brasileiro a saber: produção industrial, balança comercial, inflação (medida pelo IPCA) e importações.

Mostrou-se que a taxa de acerto global do modelo de regressão logística é de 72,5%, que é maior do que o modelo com apenas uma constante, ou seja, se arbitrariamente todos os meses fossem classificados como crescimento econômico, a taxa de acerto seria de 64,2%.

Foram analisadas as estatísticas de ajuste do modelo, o pseudo  $R^2$  de Nagelkerke apresentou um poder de explicação 25,9% e a medida de Hosmer e Lemeshow de ajuste geral indicou que não houve diferença estatisticamente significativa entre as classificações observadas e previstas para o modelo final. Com a curva ROC (*receiver operation characteristic*) verificou-se que a hipótese de ajuste do modelo não ser adequado aos dados é rejeitada.

Para verificar a precisão das estimativas foi realizada a avaliação cruzada (*cross validation*), utilizando o método *k folders* e *leave one out*. As taxas de acerto da validação cruzada ficaram em 71,6%, valores bem próximos aos obtidos na etapa de modelagem. Esse resultado demonstram a boa capacidade de generalização do modelo e a inexistência de necessidade de reajustes. Verificou-se que em períodos de crise a probabilidade de crescimento econômico brasileiro se reduz em 47,3%. Devido ao fortalecimento das relações comerciais entre os países, desequilíbrios iniciados em um país são rapidamente transmitidos para a economia dos outros, muitas vezes de maneira intensa.

As variáveis econômicas mais impactantes no crescimento econômico brasileiro são, inflação (medida pelo IPCA), produção industrial, balança comercial, importações e a variável binária crises. Com os coeficientes estimados do modelo pode-se concluir que um aumento 1pp na produção industrial ocasiona um aumento de 0,5% na probabilidade de crescimento econômico. A balança comercial se mostrou significativa para o crescimento econômico com

sua razão de chances apontando um aumento de 3,7% na probabilidade de crescimento econômico para cada variação de US\$ (mil) no saldo da balança comercial. O aumento em 1pp na taxa de inflação acarreta um aumento de 19,7% no crescimento econômico.

Para as importações, permite concluir que aumenta-se em 0,6% a probabilidade de crescimento econômico para o aumento de 1 mil US\$ em importações. Dentre as crises mundiais mais impactantes para o crescimento econômico brasileiro destacaram-se as crises asiática, crise da Rússia seguida do Crise do Subprime ocorrida nos EUA, seguida pela crise do 11 de setembro também nos EUA.

## REFERÊNCIAS

CORRAR, L.J., PAULO, E, DIAS FILHO, J. M. 2007. Análise multivariada para os cursos de administração, ciências contábeis e economia. São Paulo: Atlas, 2007. 542p.

FÁVERO, L.P., BELFIORE, P., SILVA, F.L., CHAN, F.L. 2009. Análise de dados – Modelagem multivariada para tomada de decisões. Rio de Janeiro: Elsevier, 2009.

GOMES, C.F.S., RIBEIRO, P.C.C., DUM, F.A.C., SANSEVERINO, A.M. 2016. As crises econômicas mundiais e as variáveis econômicas no Brasil. Rpep. Relatórios de Pesquisa em Engenharia de Produção. Abr, vol.16, no.1, p.18-36.

HAIR. J.F., BLACK, W.C., BABIN, B.J., ANDERSON, R.E. & TATHAM, R.L. 2009 Análise multivariada de dados. 6. ed. Porto Alegre: Bookman, 2009

HAIR. J. *et al.* 1998 *Multivariate data Analysis with Readings*. 4 ed. New Jersey; Prentice Hall, 1995 .

KARAM, A. K. 2006. Regressão Logística: Um modelo de Risco de Cancelamento de Clientes. (Dissertação Mestrado) Departamento de Administração. PUC-RJ 2006.

Lobão, J.F.M. 2009. Contágio entre Mercados de Acções de Países Desenvolvidos: Um Estudo de Processos de Transmissão de Choques de Rendibilidade US\$ num Contexto de Episódios de Crises Financeiras. Tese (Doutorado em Economia) - Universidade do Minho. Disponível em : <<http://www.eumed.net/tesis/2009/jfssml/index.htm>>. Acesso em 30/05/2016.

NETER, J., KUTNER, M.H., LI, W., NACHTSHEIN, C.J. 2005 *Applied Linear Statistical Models*, McGraw-Hill Higher Education

REINHART, K. S.; ROGOFF, C. M. 2010 *Oito séculos de delírios financeiros: desta vez é diferente*. Rio de Janeiro: Editora Campus

SILVA, A. C. D. 2011. Análise estatística de inquéritos online. Dissertação de mestrado, Universidade do Minho, Portugal.

SILVA, M.F., PEREIRA, E.J.A., FILHO, A.M.S., CASTRO, A.P.N., MIRANDA, J.G.V., ZEBENDE, G.F.2016. Quantifying the contagion effect of the 2008 financial crisis between the G7 countries (by GDP nominal). *Physica A*. Volume 453, 1 July 2016, Pages 1–8

VIDAL, T. L. 2011. Crises financeiras: efeito contágio ou interdependência entre os Países? Evidências utilizando uma abordagem multivariada. 173 p. (Dissertação Mestrado) Faculdade de Economia, Administração e Contabilidade. USP.