

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE COMPUTAÇÃO

**Distribuição Urnas-por-Votos
nas Eleições Presidenciais de 2010**

Jorge Stolfi

Technical Report - IC-12-?? - Relatório Técnico
October - 2012 - Outubro

The contents of this report are the sole responsibility of the authors.
O conteúdo do presente relatório é de única responsabilidade dos autores.

Distribuição Urnas-por-Votos nas Eleições Presidenciais de 2010

Jorge Stolfi

Instituto de Computação, Universidade Estadual de Campinas
Caixa Postal 6176, 13083-971 Campinas, SP.

stolfi@ic.unicamp.br

Última edição 2012-10-03 13:46:57 by stolfilocal

Resumo

Apresentamos estatísticas de número de urnas em função do número de votos obtidos por determinado candidato nas eleições presidenciais de 2010, e comparamos com a distribuição esperada na hipótese de preferência uniforme por estado.

1 Definições

Suponhamos que existam K urnas numa determinada região (por exemplo, um estado), numeradas de 1 a K ; e que na urna número k votaram p_k eleitores, dos quais v_k votaram em determinado candidato. Definimos a *distribuição de urnas por voto*, para esse candidato e região, como sendo a função y que, a cada inteiro x , associa o número $y(x)$ de urnas onde o candidato obteve x votos. isto é.

$$y(x) = \sum_{k=1}^K (v_k = x)$$

onde a fórmula ' $v_k = x$ ' vale 1 se os dois números são iguais, e 0 se são diferentes.

Sejam agora N o número de eleitores que votaram nessa região, e n o número de votos obtidos nessa região por esse candidato. Na *hipótese de homogeneidade*, esses n eleitores são alocados aleatoriamente dentre todas as urnas do estado. Sob esta hipótese, a probabilidade de que uma urna que recebeu um total de p votos tenha exatamente x votos para esse candidato é

$$P(p, x, \mu) = \binom{p}{x} \mu^x (1 - \mu)^{p-x}$$

onde μ é a fração n/N de votos do candidato. Portanto, nessa hipótese, o valor esperado de $y(x)$ é

$$\tilde{y}(x) = \sum_{k=1}^K P(p_k, x, \mu) = \sum_{k=1}^K \binom{p_k}{x} \mu^x (1 - \mu)^{p_k-x}$$

2 Dados

Os dados abaixo se referem às eleições de primeiro turno de 2010 para Presidente, nos estados de SP ($K = 80.220$, $N = 25.310.267$), MG ($K = 43.851$, $N = 11.838.442$), PR ($K = 23.785$, $N = 6.347.623$) e RJ ($K = 30.911$, $N = 9.572.486$).

A tabelas 1 e 2 mostram, para cada candidato, o número de votos n , a preferência $\mu = n/N$, e o número máximo v_{\max} de votos obtidos em uma urna nos estados considerados.

As figuras 1 e 2 mostram a distribuição real de votos $y(x)$ e a distribuição $\tilde{y}(x)$ esperada pela hipótese de homogeneidade, para cada candidato e em cada estado.

3 Discussão

No caso das eleições de 2010, a maioria das urnas tem p_k entre 300 e 400, mas há um numero significativo de urnas com p_k bem menor. A distribuição $\tilde{y}(x)$ é portanto parecida com $K \cdot P(350, x, \mu)$ mas um pouco mais larga e com uma “cauda” na região de $0 < x < 300\mu$.

Em geral, a distribuição real $y(x)$ é bem diferente da distribuição esperada $\tilde{y}(x)$. Isso ocorre porque a hipótese de homogeneidade está longe de ser válida. Em geral, os eleitores que votam num determinado candidato estão distribuídos de maneira muito desigual dentro do estado, e portanto o parâmetro μ varia de área para área. Portanto a distribuição real y é em geral bem mais espalhada que \tilde{y} .

A heterogeneidade geográfica e o conseqüente espalhamento da distribuição y é evidente na maioria dos casos para os candidatos com votação significativa (13-Dilma, 43-Marina, 45-Serra). Ela é menor, mas ainda perceptível, no caso dos votos em branco e nulos, e do candidato 50-Plínio, especialmente em SP e RJ.

A heterogeneidade é evidente também no caso do candidato 16-ZeMaria, que, apesar de ter apenas 0.069% dos votos em SP, chegou a obter 75 votos em uma mesma urna desse estado, e mais de 15 votos em outras 15 urnas (quando, se valesse a hipótese de homogeneidade, o número máximo de votos que ele poderia ter na mesma urna seria por volta de 7). Já nos outros três estados, onde sua preferência geral era igual ou maior (MG $\mu = 0.078\%$, PR $\mu = 0.062\%$ e RJ $\mu = 0.153\%$), a distribuição y ficou bastante próxima à homogênea \tilde{y} , e ele não obteve mais que 7 votos na mesma urna.

Uma certa heterogeneidade também pode ser vista nas preferências pelos candidatos 21-Ivan em RJ ($\mu = 0.054\%$, $v_{\max} = 10$), 27-Eymael em PR ($\mu = 0.073\%$, $v_{\max} = 16$), e 28-Levy em PR ($\mu = 0.062\%$, $v_{\max} = 9$). Por outro lado, os candidatos pouco votados tiveram distribuição surpreendentemente homogênea em vários estados:

- 16-ZeMaria: em MG, PR, RJ.
- 21-Ivan: em SP, MG, PR.
- 27-Eymael: em SP, MG, PR.
- 28-Levy: em SP, MG, e RJ.
- 29-Pimenta: em SP, MG, PR e RJ.

Tabela 1: Votação de cada candidato, nos estados de SP e MG.

Candidato		SP			MG		
N.	Nome	n	$\mu(\%)$	v_{\max}	n	$\mu(\%)$	v_{\max}
13	Dilma	8.740.949	34.535	290	5.067.399	42.805	292
16	ZeMaria	17.584	0.069	75	9.196	0.078	7
21	Ivan	7.324	0.029	4	3.921	0.033	4
27	Eymael	14.684	0.058	7	9.706	0.082	7
28	Levy	13.573	0.054	5	6.342	0.054	5
29	Pimenta	3.134	0.012	3	1.868	0.016	3
43	Marina	4.865.828	19.225	177	2.291.502	19.356	219
45	Serra	9.524.050	37.629	323	3.317.872	28.026	294
50	Plinio	239.173	0.945	27	77.782	0.657	18
95	BRANCO	856.433	3.384	49	418.468	3.535	66
96	NULO	1.027.535	4.060	43	634.386	5.359	169

Tabela 2: Votação de cada candidato, nos estados de PR e RJ.

Candidato		PR			RJ		
N.	Nome	n	$\mu(\%)$	v_{\max}	n	$\mu(\%)$	v_{\max}
13	Dilma	2.311.239	36.411	293	3.739.632	39.066	281
16	ZeMaria	3.929	0.062	6	14.654	0.153	6
21	Ivan	2.302	0.036	4	5.158	0.054	10
27	Eymael	4.638	0.073	6	18.058	0.189	16
28	Levy	3.914	0.062	9	7.053	0.074	5
29	Pimenta	461	0.007	2	1.442	0.015	3
43	Marina	944.402	14.878	147	2.693.130	28.134	193
45	Serra	2.607.664	41.081	242	1.925.166	20.111	237
50	Plinio	56.452	0.889	20	140.782	1.471	34
95	BRANCO	163.486	2.576	43	368.193	3.846	43
96	NULO	249.136	3.925	66	659.218	6.887	76

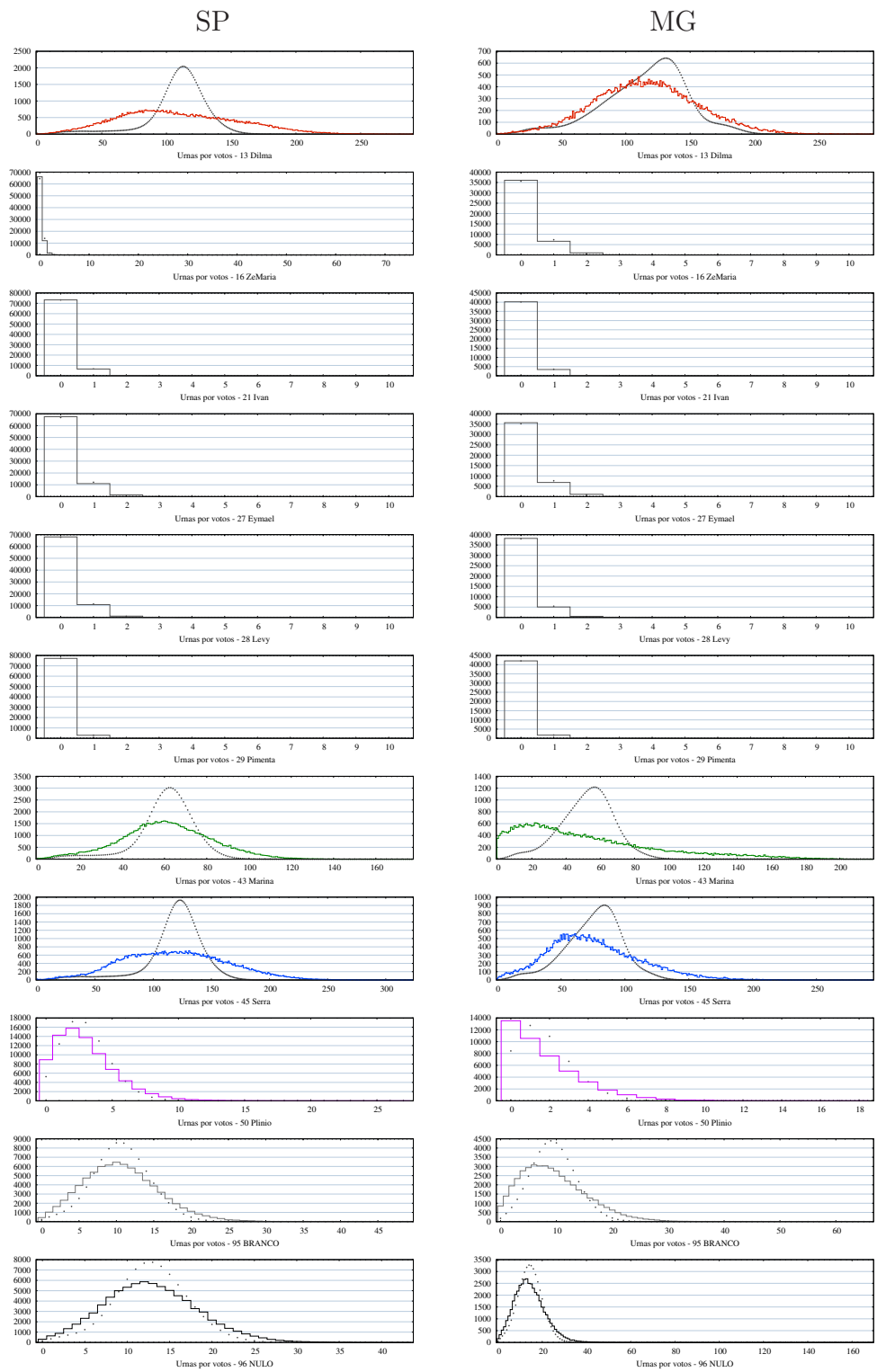


Figura 1: A distribuição real $y(x)$ de urnas por número de votos no candidato (barras), nos estados de SP e MG, e a distribuição $\tilde{y}(x)$ esperada pela hipótese de homogeneidade em cada estado (pontos cinza).

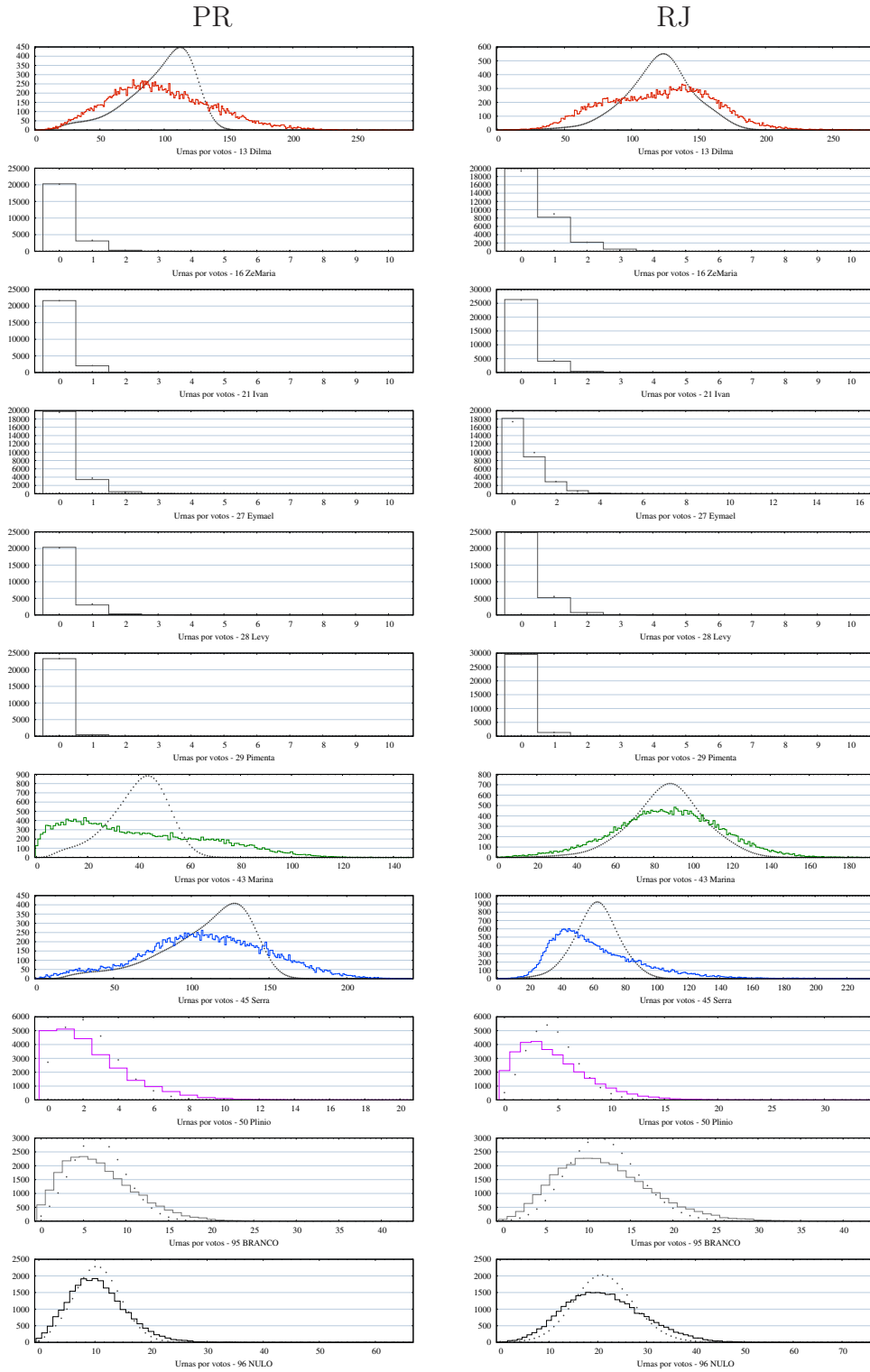


Figura 2: A distribuição real $y(x)$ de urnas por número de votos no candidato (barras), nos estados de PR e RJ, e a distribuição $\tilde{y}(x)$ esperada pela hipótese de homogeneidade em cada estado (pontos cinza).

Tabela 3: Distribuição real $y(x)$ de urnas por número de votos no candidato 16-ZeMaria, e distribuição teórica $\tilde{y}(x)$, nos estados de SP, MG, PR e RJ.

x	SP		MG		PR		RJ	
	$y(x)$	$\tilde{y}(x)$	$y(x)$	$\tilde{y}(x)$	$y(x)$	$\tilde{y}(x)$	$y(x)$	$\tilde{y}(x)$
0	66070	64488.	36059	35612.	20292	20180.	19879	19324.
1	12056	14018.	6610	7354.7	3110	3300.0	8196	8993.6
2	1661	1585.1	1010	817.3	338	286.9	2188	2178.8
3	269	121.9	135	63.5	40	17.2	528	363.0
4	81	7.1	30	3.8	3	0.8	104	46.6
5	17	0.3	4	0.2	1		14	4.9
6	20				1		2	0.4
7	2		3					
8	10							
9	6							
10	3							
11	4							
12	4							
13	1							
14								
15								
16	2							
17								
18								
19								
20	3							
21								
22								
23	1							
24	1							
25								
26	1							
27								
28								
29								
30								
31								
32	1							
33								
34								
35								
36								
37								
38								
39	1							
40								
41								
42								
43								
44								
45								
46	1							
47	1							

Tabela 4: Distribuição real $y(x)$ de urnas por número de votos no candidato 21-Ivan e a distribuição teórica $\tilde{y}(x)$, nos estados de SP, MG, PR e RJ.

x	SP		MG		PR		RJ	
	$y(x)$	$\tilde{y}(x)$	$y(x)$	$\tilde{y}(x)$	$y(x)$	$\tilde{y}(x)$	$y(x)$	$\tilde{y}(x)$
0	73302	73232.	40196	40112.	21630	21597.	26341	26174.
1	6538	6663.4	3407	3563.7	2018	2078.1	4084	4340.0
2	356	314.7	231	169.9	128	106.2	425	373.7
3	22	10.1	16	5.7	8	3.7	46	22.1
4	2	0.2	1	0.1	1	0.1	6	1.0
5							3	
6								
7							3	
8							2	
9								
10							1	

Tabela 5: Distribuição real $y(x)$ de urnas por número de votos no candidato 27-Eymael e a distribuição teórica $\tilde{y}(x)$, nos estados de SP, MG, PR e RJ.

x	SP		MG		PR		RJ	
	$y(x)$	$\tilde{y}(x)$	$y(x)$	$\tilde{y}(x)$	$y(x)$	$\tilde{y}(x)$	$y(x)$	$\tilde{y}(x)$
0	67527	66843.	35688	35206.	19803	19594.	18140	17348.
1	10973	12151.	6846	7667.5	3416	3775.2	8888	9905.0
2	1499	1148.3	1128	898.7	494	387.1	2887	2946.5
3	182	73.8	163	73.6	59	27.4	733	603.1
4	32	3.6	18	4.7	9	1.5	193	95.1
5	4	0.1	6	0.2	3	0.1	44	12.3
6	2		1		1		11	1.3
7	1		1				7	0.1
8							1	
9							3	
10								
11							1	
12								
13							1	
14								
15							1	
16							1	

Tabela 6: Distribuição real $y(x)$ de urnas por número de votos no candidato 28-Levy e a distribuição teórica $\tilde{y}(x)$, nos estados de SP, MG, PR e RJ.

	SP		MG		PR		RJ	
x	$y(x)$	$\tilde{y}(x)$	$y(x)$	$\tilde{y}(x)$	$y(x)$	$\tilde{y}(x)$	$y(x)$	$\tilde{y}(x)$
0	68075	67769.	38190	37975.	20359	20193.	24841	24629.
1	10864	11393.	5060	5435.0	3016	3289.6	5216	5571.0
2	1150	995.5	535	417.9	357	285.0	736	654.7
3	117	59.1	55	22.4	38	17.1	109	52.9
4	12	2.7	8	0.9	10	0.8	7	3.3
5	2	0.1	3		3		2	0.2
6					1			
7								
8								
9					1			
10								

Tabela 7: Distribuição real $y(x)$ de urnas por número de votos no candidato 29-Pimenta e a distribuição teórica $\tilde{y}(x)$, nos estados de SP, MG, PR e RJ.

	SP		MG		PR		RJ	
x	$y(x)$	$\tilde{y}(x)$	$y(x)$	$\tilde{y}(x)$	$y(x)$	$\tilde{y}(x)$	$y(x)$	$\tilde{y}(x)$
0	77194	77149.	42058	42025.	23334	23329.	29518	29503.
1	2921	3009.7	1719	1784.8	441	451.6	1347	1373.9
2	102	60.9	73	40.6	10	4.6	43	33.2
3	3	0.8	1	0.6			3	0.6

Tabela 8: Distribuição real $y(x)$ de urnas por número de votos no candidato 50-Plínio e a distribuição teórica $\tilde{y}(x)$, nos estados de SP, MG, PR e RJ.

x	SP		MG		PR		RJ	
	$y(x)$	$\tilde{y}(x)$	$y(x)$	$\tilde{y}(x)$	$y(x)$	$\tilde{y}(x)$	$y(x)$	$\tilde{y}(x)$
0	8924	5251.5	13538	8429.1	5000	2721.7	2117	527.5
1	14261	12355.	10551	12737.	5113	5241.1	3472	1826.5
2	15766	17184.	7581	10899.	4424	5789.9	4151	3559.4
3	13743	16992.	5023	6669.3	3264	4611.5	4217	4948.4
4	10253	12986.	3192	3211.2	2298	2889.9	3640	5407.4
5	6836	8082.2	1823	1283.7	1410	1495.5	3249	4901.1
6	4346	4246.9	1021	441.0	961	659.5	2596	3814.1
7	2564	1933.7	553	133.4	604	253.5	2024	2610.6
8	1557	778.0	293	36.2	343	86.3	1564	1599.9
9	864	280.8	120	8.9	167	26.4	1147	890.0
10	463	92.1	88	2.0	103	7.3	859	454.3
11	261	27.7	34	0.4	47	1.9	605	214.6
12	160	7.7	16	0.1	24	0.4	423	94.5
13	90	2.0	9		15	0.1	273	39.0
14	45	0.5	5		4		206	15.2
15	31	0.1	3		3		136	5.6
16	21				3		80	1.9
17	12				1		46	0.6
18	5		1				45	0.2
19	3						19	0.1
20	6				1		13	
21	2						8	
22	2						8	
23	1						5	
24	1						1	
25	2						1	
26								
27	1						2	
28								
29							1	
30							1	
31							1	
32								
33								
34							1	