MONITORAMENTO DE ELEIÇÕES MUNICIPAIS NO ESTADO DO ESPÍRITO SANTO - AMOSTRAGEM CLÁSSICA E BAYESIANA

Gutemberg Hespanha Brasil e Antonio Fernando Pêgo e Silva
DEST
UFES - Vitória, ES
Jose Weber Freire Macedo
DPSI
UFES - Vitória, ES
Erly Euzébio dos Anjos
DCSO
UFES - Vitória, ES

RESUMO

Neste trabalho apresentamos a utilização de dois procedimentos para monitorização de uma pesquisa eleitoral: a abordagem clássica de amostragem comumente adotada e o modelo bayesiano para previsão de resultados eleitorais introduzido por Bernardo (1984), juntamente com a metodologia para seleção do número de locais descrita em Brasil & Pego e Silva (1994); nestes uma amostra de mesmo tamanho é selecionada

nos locais de comportamento mais similar a toda a região sendo pesquisada, no caso, o município. No período de um ano antecedendo as eleições municipais de 03/10/92, realizamos cinco pesquisas no município de Guarapari-ES, incluindo a pesquisa de "boca-de-urna", das quais, quatro via metodologia clássica e três, via metodologia bayesiana. Em duas ocasiões, agosto e setembro de 92, as duas metodologias foram aplicadas simultaneamente sendo que algumas conclusões extraidas dessa comparação são aqui discutidas.

PALAVRAS CHAVE: Previsão Eleitoral. Pesquisa de Opinião. Metodologias Clássica e Bayesiana. Campanhas Políticas.

1. INTRODUÇÃO

Apresentamos neste trabalho a utilização de dois procedimentos para monitoração de uma pesquisa eleitoral: a abordagem clássica de amostragem comumente adotada e o modelo Bayesiano para previsão de resultados eleitorais introduzido por Bernardo (1984), juntamente com a metodologia para escolha do número de locais descrita em Brasil & Pêgo e Silva (1994).

No período de um ano antecedendo as eleições municipais de 03/10/92, realizamos cinco pesquisas no município de Guarapari-ES, incluindo a pesquisa de "boca-de-urna". Em quatro dessas pesquisas utilizamos a metodologia clássica e em três, a metodologia Bayesiana. Em duas ocasiões, agosto e setembro de 92, as duas metodologias foram aplicadas simultaneamente sendo extraídas algumas conclusões dessa comparação. A organização do trabalho é a seguinte: na seção 2 descrevemos as duas metodologias utilizadas; na seção 3 apresentamos alguns resultados e, finalmente na seção 4 é realizada uma discussão dos resultados.

2. MONITORAMENTO DE ELEIÇÕES MUNICIPAIS: METODOLOGIAS ADOTADAS

2.1. Geral

Usualmente em uma pesquisa eleitoral, pesquisa de intenção de voto com o intuito de monitorar a campanha de algum candidato, são avaliados os seguintes grupos de questões: (1) perfil sócio-econômico do entrevistado (sexo, faixa etária, classe social, renda, etc); (2) perfil político (simpatia por partidos, voto na ultima eleição, avaliação das administrações, etc); (3) atitudes e preferências políticas (voto espontâneo, voto estimulado, avaliação dos índices de rejeição dos candidatos, avaliação das preferências por chapas simuladas, principais lideranças políticas, etc) e, (4) avaliação dos temas julgados mais importantes pelos eleitores (medição do alcance da publicidade, penetração da mensagem eleitoral, etc).

Neste monitoramento eleitoral, realizado através das cinco pesquisas de opinião, todas estas questões e outras mais, foram avaliadas, o que propiciou um bom direcionamento da campanha por parte dos estrategistas de marketing. Do ponto de vista da amostragem estatística utilizamos dois desenhos amostrais de natureza bastante distinta, mas cujos resultados são similares. Estes dois desenhos amostrais são sucintamente descritos a seguir.

2.2. Amostragem Clássica

A amostragem "clássica" adotada é suficientemente bem conhecida; ver por exemplo, Raj (1972) e Cochran (1977). Apresentamos aqui apenas uma breve descrição do procedimento adotado.

De acordo com os dados censitários disponíveis (Censo-1991), a população residente no município de Guarapari era de 61.597 habitantes, sendo 39.794 eleitores, segundo o TRE (Tribunal Regional Eleitoral). Para a

amostragem clássica, o município foi dividido em 11 áreas, sendo 10 urbanas perfazendo 86,33% do eleitorado e, uma interiorana, com 13,67% do eleitorado, predominantemente rural.

Adotou-se uma amostragem aleatória estratificada por microregiões (áreas) com controle de quotas, abrangendo sexo, faixa etária e classe sócio-econômica (critério ABA-ABIPEME). Observe-se que essas microrregiões de Guarapari são compostas por bairros geograficamente próximos, e possuem, aparentemente, uma razoável homogeneidade quanto a população residente.

A amostra foi dimensionada utilizando-se as relações usuais para levantamentos dessa natureza adotando-se um intervalo de confiança de 95%. Os pesos de cada microrregião foram estabelecidos segundo o peso eleitoral correspondente, utilizando-se para isso os dados do cadastro eleitoral do TRE. Todas as Microrregiões (áreas) foram consideradas como auto-representativas, i.e. com probabilidade igual a 1 de pertencerem a amostra (na microrregião Interior, de dificil acesso, as entrevistas foram concentradas em cinco locais). Desse modo a amostra foi dimensionada para um total de 590 entrevistas conduzindo a um erro amostral de 4,0%.

2.3. Amostragem Bayesiana

Diferentemente dos desenhos tradicionais de planejamento amostral, que supoem que o voto esteja relacionado fortemente a características da população tais como distribuições etárias, de renda e sexo, a metodologia proposta por Bernardo (1984), procura selecionar aqueles locais que sejam mais representativos do comportamento político da população. Uma vez identificados esses locais, uma amostra aleatória de mesma dimensão de eleitores é selecionada em cada um deles.

Na presente aplicação identificou-se os locais (bairros e/ou agregações de bairros) mais representativos da população votante, no sentido de que o comportamento do voto assumido pelos eleitores destes locais, fosse o mais similar possivel ao comportamento da população de todo o município. Para isso foram utilizados os resultados eleição imediatamente anterior (1988).

Adotando-se como critério a medida de divergência direta de Kullback-Leibler, Kullback (1968) e, através da minimização das "distâncias" entre esses locais e o universo eleitoral (o município), pode-se construir uma "hierarquia de similitudes", para selecionar os locais "mais representativos" do município.

A medida de Divergência de Kullback-Leibler é dada por:

$$D_i = \sum_{j=1}^{m} \theta_j . Ln\left(\frac{\theta_j}{\hat{\theta}_{ij}}\right) \qquad i = 1,2,\ldots,k. \qquad (1)$$

onde: $\hat{\theta}_{ij}$ é o estimador Bayesiano da probabilidade de um eleitor do i-ésimo local (i.e. com características similares às daqueles que vivem na região abrangida pelo local i) votar no j-ésimo candidato; m - número de candidatos existentes e θ_j - é a proporção de votos obtidos pelo candidato j (na área total) na eleição imediatamente anterior; j = 1, 2,...m.

Quanto menor for a medida D_i , mais representativo será o local i ou seja, é mais similar ao município como um todo, no sentido descrito anteriormente. Lembramos que foi adotada a suposição de que esses "locais representativos" de toda a região permanecem os mesmos de uma eleição para a seguinte.

O procedimento para a seleção do número de locais foi aquele exposto em Brasil & Pego e Silva (1993). Através do uso de (1) ordenou-se os locais em ordem de similitude, i.e. do mais similar ao menos similar para todo o município. A seguir, através das relações (2) e (3) abaixo, construiu-se os gráficos do Erro Quadrático Médio versus número de locais (Figuras 1a e

1b) para selecionar finalmente, os locais a serem pesquisados (dentre os k disponíveis). Temos, para i = 1, 2,...,k e j = 1, 2, ..., m:

$$\bar{\theta}_{ij} = \frac{1}{i} \cdot \sum_{l=1}^{i} \hat{\theta}_{lj} \qquad (2)$$

$$EQM_i = \frac{1}{m} \cdot \sum_{j=1}^{m} \left(\theta_j - \bar{\theta}_{ij}\right)^2 \quad (3).$$

Analisando-se os gráficos obtidos, pode-se realizar uma escolha do número de locais a serem pesquisados. Uma completa descrição do modelo Bayesiano para previsão das eleições pode ser encontrada em Bernardo (1984) e Souza & Brasil (1989).

A Tabela 1 apresenta a ordenação dos locais com os dados de 1988 e os resultados finais da eleição de 1992. Pode-se constatar que sete locais permanecem os mesmos entre os dez mais similares. Na Figura 1a, vemos que o mínimo ocorreu para 4-6 locais no ano eleitoral de 1988. Utilizando-se o resultado final de 92 o mínimo é obtido entre 6 a 10 locais.

Desse modo, para o município de Guarapari optou-se por realizar, por razões de custo, aproximadamente 40 entrevistas nos seis locais mais similares, totalizando uma amostra de 240 individuos. Para a pesquisa de "boca-de-urna" a amostra foi de aproximadamente 50 entrevistas em cada um dos seis locais. Observe-se que foram adotados os locais de acordo com o resultado de 1988 (Figura 1a).

3. RESULTADOS DAS PESQUISAS

A Tabela 2 sintetiza as pesquisas realizadas com as duas metodologias expostas. A Figura 2 apresenta a evolução do voto (menção estimulada por cedula eleitoral) ao longo do periodo monitorado; contém tambem o percentual de indecisos.

Figura 1 - Erro Quadrático Médio (EQM) x Número de Locais

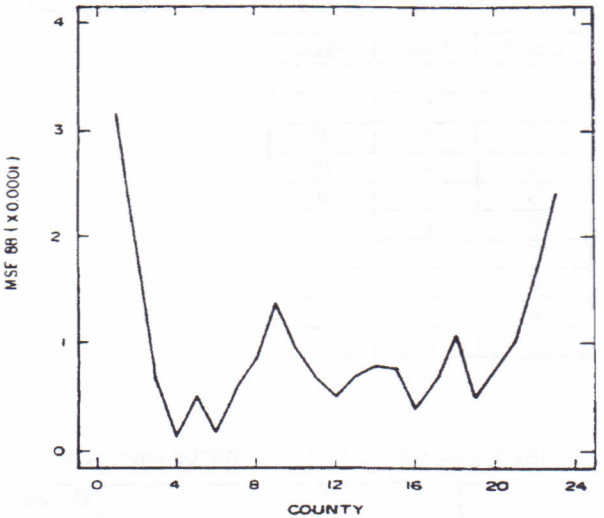


Figura 1a - Ano eleitoral de 1988

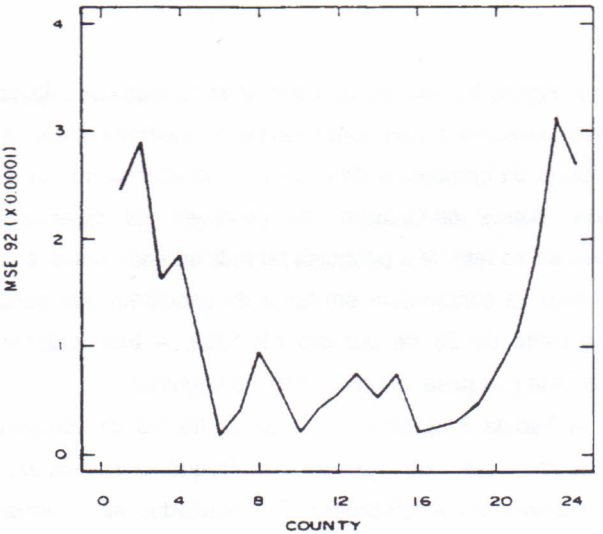


Figura 1b - Ano eleitoral de 1992

Tabela 1: Ordenação dos locais segundo a medida de divergência

Local	1988	1992
1	1	4
2	2	X
3	3	X
4	4	10
5	5	8
6	6	2
7	7	1
8	9	7
9	9	X
10	10	3

Tabela 2: Pesquisas realizadas - Datas e tamanhos amostrais

Metodologia	Set./1991	Mar./1992	Ag./1992	Set./1992	Out./1992 Boca-Urna
Clássica	520	592	593	593	-------
Bayesiana	-----	------	266	236	311

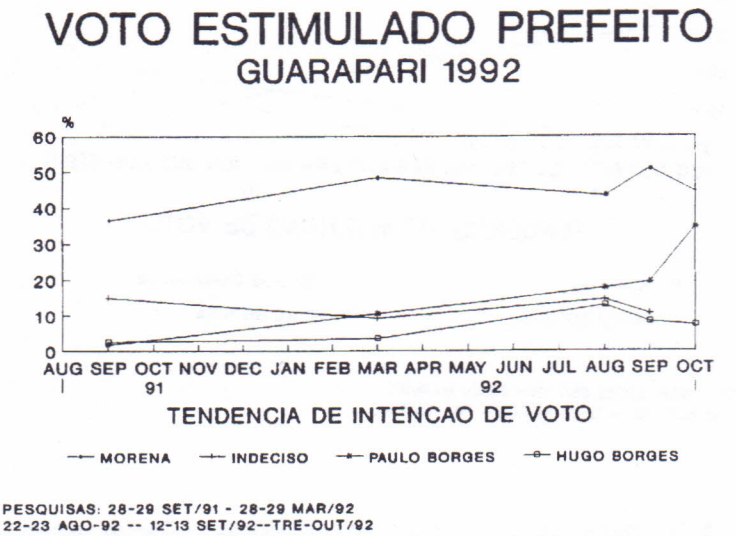
A Figura 3 contém o percentual de "outros candidatos". Ambos s gráficos são ilustrativos porquanto revelam acontecimentos e atitudes tomadas ao longo da campanha. Por exemplo, o acentuado decréscimo do percentual de "outros candidatos" da pesquisa de setembro/91 para março/92, deve-se ao fato de a pesquisa ter indicado que um dos candidatos deveria abrir mão da candidatura em favor da candidata que consagrou-se vencedora no pleito de 03 de outubro de 1992. A transferência do voto detetetada em uma pesquisa foi confirmada na seguinte.

A Tabela 3 apresenta resultados típicos da comparação das duas metodologias, caso de Tabelas de Frequências globais, para as pesquisas realizadas em setembro/92. Os resultados são aproximados, a

despeito dos tamanhos das amostas utilizados (bem menor no caso bayesiano).

As Tabelas 4 e 5 apresentam alguns resultados de cruzamentos entre variáveis. Outras tabelas cruzadas de características globais de toda a região sendo pesquisada, não apresentadas nesta versão resumida do artigo, também são similares.

Figura 2: Evolução da intenção de voto (incluindo indecisos)



Um resultado bastante interessante foi com relação a questão: "Qual o seu candidato a vereador ?" Parece anti-intuitivo que as preferências dos eleitores fiquem sequer parecidas relativamente aos dois procedimentos amostrais. No entanto, não foi esse o caso. Na pesquisa de agosto de 1992, por exemplo, dentre os 12 candidatos a vereador mais mencionados em ambas as pesquisas, nove são coincidentes.

Figura 3: Evolução da intenção de voto (incluindo "outros candidatos")

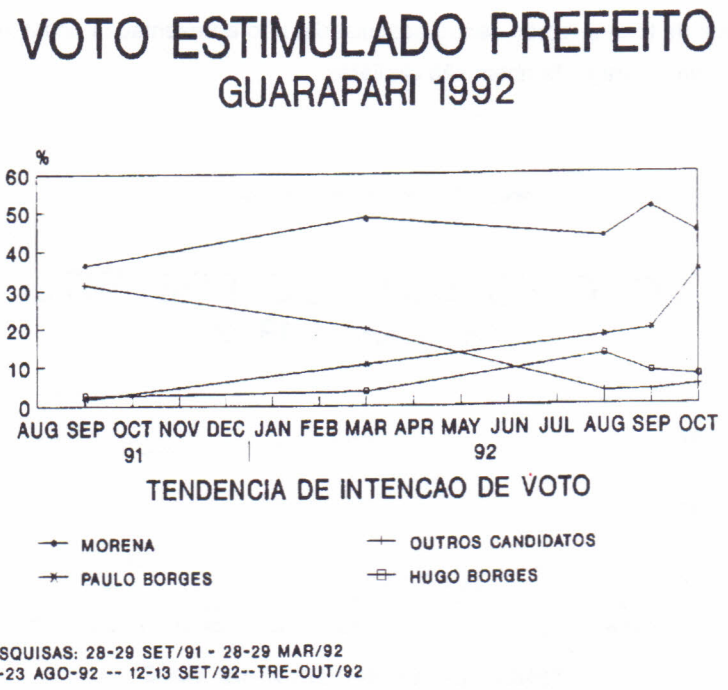


Tabela 3: Comparação de resultados das duas metodologias - Tabelas de freqüência (setembro/92)

Candidato	Candidato de preferência (Resposta estimulada)		Rejeição do voto (Resposta estimulada)	
	Clássica	Bayesiana	Clássica	Bayesiana
Olga V.	0.8	0.8	10.8	16.5
Paulo B.	19.4	23.3	16.4	17.8
Hugo B.	8.3	7.6	14.5	13.1
D. Gotardo	2.7	2.5	7.1	4.7
Morena	50.9	52.1	13.0	16.5
Branco/Nulo	7.5	5.5	5.9	3.8
Não sabe/NR	10.4	8.2	33.2	27.5
Total	100.0	100.0	100.0	100.0

Tabela 4: Candidato de preferência (espontâneo) X classe social
Amostragem Clássica (Setembro/92)

COL PCT Candidato	AB	C	DE	ROW TOTAL
Olga	0.0	0.0	1.1	0.7
Paulo B.	16.1	12.3	17.9	16.2
Hugo B.	3.4	7.1	6.8	6.4
D. Gotardo	2.3	1.3	2.0	1.9
Morena	57.5	56.1	38.7	46.0
Branco-Nulo	9.1	9.1	2.8	5.4
Não sabe/NR	11.5	14.2	30.5	23.4
Column TOTAL	14.7	26.1	59.2	100.0

Tabela 5: Candidato de preferência (Espontâneo) X Classe Social
Amostragem Bayesiana (Setembro/92)

COL PCT Candidato	AB	C	DE	ROW TOTAL
Olga	0.0	0.0	2.1	1.3
Paulo B.	27.8	20.0	15.7	18.6
Hugo B.	5.6	6.7	7.1	6.8
D. Gotardo		1.7	1.4	1.3
Morena	47.2	51.7	42.1	45.3
Branco-Nulo	8.3	6.7	1.4	3.8
Não sabe/NR	11.1	13.3	30.0	22.9
Column TOTAL	15.3	25.4	59.3	100.0

Resta apresentar a pesquisa de "boca-de-urna" realizada via metodologia Bayesiana. Aproximadamente cinquenta entrevistas foram efetivadas em cada um dos seis locais similares. As entrevistas ocorreram no periodo de 08 às 10 horas da manha. Um controle de quotas pouco rigoroso foi realizado (sexo e idade). Observou-se, contudo, um alto indice de não-respostas (16.7%), i.e. de individuos que não declaravam o seu voto

após deposita-lo na urna. A solução adotada foi a substituição do entrevistado. A Tabela 6 mostra os resultados.

Tabela 6: Previsão dos resultados finais - Pesquisa de boca-de-urna

Candidato	Olga V.	Paulo B.	Hugo B.	Gotardo	Morena	BR/NU	Total.
Boca-de-Urna	1.0	36.5	6.8	2.9	49.2	3.6	100.0
Resultado	0.9	34.5	7.3	3.7	44.4	9.2	100.0

OBS: Percentuais calculados sobre o total de votantes

O intervalo de probabilidade de 90% para a candidata vencedora ficou em: (38.03 - 60.41), o qual e bastante amplo; mas é devido ao fato do exíguo número de graus de liberdade decorrente do pequeno número de locais selecionados. Ver os comentários apresentados em Brasil & Pego e Silva (1993).

4. DISCUSSÃO

Apresentamos duas metodologias com características distintas e que encerram um alto grau de completitude, pelo menos no sentido de servir aos propósitos descritos na seção 2.1. A metodologia Bayesisana apresentou resultados bastante similares à metodologia clássica, a despeito do tamanho da amostra pesquisada ser bem menor. Isso comprova que esta metodologia é muito boa quanto se trata de captar características globais de toda a região sendo pesquisada, no caso, o município. Quando houver interesse em atuar em áreas específicas, por exemplo, a metodologia clássica parece ser mais adequada.

Analisando-se as Figuras 2 e 3 e, considerando-se que houve um processo de "impeachment" presidencial próximo a data das eleições, observa-se claramente a necessidade de estender-se o período de

acompanhamento das pesquisas até pelo menos uma semana antes do pleito, dada a volatilidade de uma parcela do eleitorado.

A pesquisa de "boca-de-urna" conduzida de acordo com a metodologia Bayesiana apresentou resultados satisfatórios, na medida em que acertou a ordem e a magnitude do percentual de votos de todos os candidatos. Entretanto, acreditamos que os resultados poderiam ter sido melhores caso o período das entrevistas fosse ampliado e se houvesse um tratamento adequado (algo muito difícil) das não-respostas que ocorreram em número excessivo. Apesar do método adotar a hipótese de permutabilidade (exchangeability) entre as diversas seções e entre os indivíduos de cada seção, um controle de quotas contribuiu para a qualidade dos resultados.

Existe um compromisso entre o número de locais selecionados e o número de candidatos disputando o pleito, pois isso influi na determinação dos graus de liberdade e, consequentemente, nos intervalos de probabilidade. Desse modo, o número de locais a serem pesquisados também deve ser controlado pela quantidade de candidatos disputando o pleito, para garantir os graus de liberdade necessários à determinação dos intervalos de probabilidade para as previsões.

5. REFERÊNCIAS

BERNARDO, J. M. (1984), Monitoring the 1982 Spanish Socialist Victory: A Bayesian analysis, J. Amer. Statist. Assoc., 79, 387, 510-515.

BRASIL, Gutemberg H. & PEGO E SILVA, Antonio Fernando (1994), Pesquisas Eleitorais: Algumas Considerações Sobre a Metodologia Bayesiana para Pesquisas Eleitorais com Aplicação as Eleições de 1990 no Estado do Espirito Santo, Revista Brasileira de Estatistica, V 52, N° 197/198, 69-92, 1991.

BOX, G. E. P. & TIAO, G. C. (1973), Bayesian Inference in Statistical Analysis, Reading, Mass., Addison-Wesley.

COCHRAN W. G. (1977), SAMPLING TECHNIQUES, 3 rd ed, Wiley.

KULLBACK, S. (1968), Information Theory and Statistics, New York, Dover Press.

MENDONÇA, Isabel G. S. Furtado de & MIGON, Helio S. (1987), Pesquisa Eleitoral: Uma Analise Bayesiana, R. Bras. Estat., Rio de Janeiro, V. 48, N 189/190, 25-34.

RAJ, D. (1972), THE DESIGN OF SAMPLE SURVEYS, McGraw Hill.

SOUZA, R. C. & BRASIL, G. H. (1989), A Bayesian model to Forecast an Election Outcome: An Application to the Brazilian States Elections of 1986, Estadistica, 41, 136, 13-30.