RESEARCH REPORT

Serial No. 14

June 1992

THE DIFFERENCE IN VOTE SHARES FOR MULTIPLE VOTE SYSTEMS

by

John Bacon-Shone

THE UNIVERSITY OF HONG KONG

DEPARTMENT OF STATISTICS

Pokfulam Road, Hong Kong. Tel: (852) 859 2466 Fax: (852) 858 9041
E-Mail: STATIST@HKUCC.HKU.HK or STATIST@HKUCC.BITNET

THE DIFFERENCE IN VOTE SHARES FOR MULTIPLE VOTE SYSTEMS

John Bacon-Shone
Department of Statistics and
Social Sciences Research Centre
The University of Hong Kong

INTRODUCTION

When individuals only have single votes in a non-transferable vote system, it is quite simple to estimate vote shares and difference in vote shares, together with standard errors. Hence we can assess how certain we are, that given a representative sample, there is a difference between the populations voting intentions or behaviour for two candidates. However, when voters are allowed to select two or more candidates as in the 1991 Legislative Council elections in Hong Kong, then even without transferable votes, the picture is more complex. For example, if we compare two candidates who are a popular combination choice, then many of the votes may be irrelevant in looking at the difference in vote shares, as these voters have voted for both the candidates.

THEORY

We will develop the theory for the situation with voters allowed to select up to two candidates, although the extension to a larger number is straightforward.

Let n_{ij} = the number of voters in the sample who select candidates i and j ($i, j = 1, \ldots, k$, $i \neq j$)
and n_{ii} = the number of voters who select candidate i only.

Then if n is the total sample size of those who vote for at least one candidate and m is the total number of votes cast in the sample:

$$n = \sum_{i=1}^{k} \sum_{j=1}^{i} n_{ij}$$

$$m = \sum_{i=1}^{k} \sum_{j=1}^{k} n_{ij}$$

The vote share for candidate i is then

$\frac{n_{i+}}{m}$, *where* $n_{i+} = \sum_{j=1}^{k} n_{ij}$ and the voter share for candidate i is $\frac{n_{i+}}{n}$ and the difference in vote shares between candidates i and j is

$$\frac{n_{i+} - n_{j+}}{m} = \sum_{l=1}^{k} \frac{n_{il} - n_{jl}}{m}$$

$$= \frac{\sum_{l=1,\, l\neq j}^{k} n_{il} - \sum_{l=1,\, l\neq i}^{k} n_{jl}}{m}$$

Note that the numerator is the number of voters for i and not j less the number of voters for j and not i. Thus if we condition on the total number of voters who use at least one vote, we have a trinomial distribution, where the three disjoint choices are

a) vote for i and not j
b) vote for j and not i
c) vote for both i and j <u>or</u> vote for neither

Thus if $p_{i.}$ is the unknown proportion of voters in the (assumed infinite) population who would make the choice a), and $n_{i.}$ is the number of voters in the sample who make choice a), then Pr ($n_{i.}$ voters for choice a and $n_{j.}$ voters for choice b)

$$= p_{i.}^{n_{i.}}\, p_{j.}^{n_{j.}}\, (1 - p_{i.} - p_{j.})^{n - n_{i.} - n_{j.}}\, \frac{n!}{n_{i.}!\; nj.!\; (n - n_{i.} - n_{j.})!} ,$$

It is then easy to show using standard combinatorics that

$$E\left(\frac{n_{i+} - n_{j+}}{m}\right) = E\left(\frac{n_{i.} - n_{j.}}{m}\right) = \frac{n\,(p_{i.} - p_{j.})}{m} ,$$

$$V\left(\frac{n_{i+} - n_{j+}}{m}\right) = V\left(\frac{n_{i.} - n_{j.}}{m}\right) = \frac{n\,(p_{i.} + p_{j.} - (p_{i.} - p_{j.})^2\,)}{m^2}$$

$$= E\left[\frac{(\, n_i + n_j\,)\;(\, n - n_i - n_j\,) + 4n_i n_j}{m^2\;(\, n - 1\,)}\right]$$

Thus $\frac{n_{i.} - n_{j.}}{m}$ is an unbiased estimator of the mean difference in vote shares, and

$\frac{(n_{i.} + n_{j.})(n - n_{i.} - n_{j.}) + 4n_{i.}n_{j.}}{nm^2}$ is an unbiased estimator of the variance of the difference in vote shares. Note that these formulae can be adjusted for a finite population if necessary.

BAYESIAN APPROACH

If we assume a prior on the probabilities s.t.

$$p\,(P_{i.},\, p_{j.}) \;\alpha\; p_{i.}^{\alpha-1}\, p_{j.}^{\beta-1}(1 - p_{i.} - p_{j.})^{\gamma-1}$$

then it is easy to show that the posterior density for the difference in vote shares (denoted by x) is :

$$p(x\,|\,data) \;\alpha\; \int_0^{(1-\frac{mx}{n})/2} (\frac{mx}{n}+y)^{\alpha^1-1}\, y^{\beta^1-1}(1 - \frac{mx}{n} - 2y)^{\gamma^1-1}\, dy$$

$$\text{where}\quad \begin{aligned} \alpha^1 &= \alpha + n_{i.} \\ \beta^1 &= \beta + n_{j.} \\ \gamma^1 &= \gamma + n - n_{i.} - n_{j.} \end{aligned}$$

Given the conditional independence structure of the Dirichlet distribution, we can easily find

$$P_r\,(\,x > 0\,|\,data\,) = \frac{\int_{\frac{1}{2}}^{1} t^{\alpha^1-1}\,(\,1 - t)^{\beta^1-1}dt}{\int_o^1 t^{\alpha^1-1}\,(\,1 - t\,)^{\beta^1-1}\,dt}$$

which can be found in incomplete Beta tables,

while $E\,(\,x\,|\,data\,) = \frac{n}{m}\,\frac{(\alpha^1 - \beta^1)}{(\alpha^1 + \beta^1 + \gamma^1)}$

and $V\,(\,x\,|\,data\,) = \frac{n^2}{m^2}\,\frac{[4\alpha^1\beta^1 + \gamma^1(\alpha^1 + \beta^1)]}{(\alpha^1 + \beta^1 + \gamma^1)^2\,(\alpha^1 + \beta^1 + \gamma^1 + 1)}$.

EXAMPLE

One of the 1991 Hong Kong Legislative Council election telephone surveys done by the Social Sciences Research Centre at The University of Hong Kong in collaboration with Asia Television Limited gave the number of sample voters in the New Territories South constituency for the combinations as shown in Table 1. Table 2 shows the differences in vote shares together with standard errors and the posterior probabilities that the differences are positive, assuming a uniform prior. The candidates are shown for simplicity in decreasing order of vote share, although the analysis has not been made conditional on the ordering. This means that the standard errors are conservative, which is preferable, given that this table ignores non-sampling errors. (Table 1 and 2 here)

Table 2 Difference in vote shares for New Territories South

	Vote share(%)	Difference(%)	Standard error	Probability
LEE	31.4			
		1.7	5.1	0.629
CHAN	29.7			
		9.4	5.7	0.944
YEUNG	20.3			
		1.7	5.1	0.629
LEUNG	18.6			

Note: 1) Standard error refers to the estimated standard error of the difference.
2) Probability is the posterior probability that the difference is greater than zero, for a uniform prior.

Table 1 Number of sample votes for candidate combinations in New Territories South

	LEUNG	CHAN	LEE	YEUNG	TOTAL VOTES
LEUNG	4	8	5	5	22
CHAN	8	3	18	6	35
LEE	5	18	7	7	37
YEUNG	5	6	7	6	24
TOTAL VOTES	22	35	37	24	118

Total voters = 69