USING STRENGTH OF OPINION TO TEST FOR NONRESPONSE BIAS IN SAMPLE SURVEYS

Dennis K. Pearl and David Fairley, Ohio State University

1. INTRODUCTION

This paper discusses the use of a statistical hypothesis test to investigate nonresponse bias. The test should be most useful for mail surveys conducted under a rigid time constraint. Mail surveys to estimate a population preference are often criticized because the typically low response rates have great potential for nonresponse bias. Estimates and confidence statements can be misleading when based only on information from sampled individuals who give a response. Nonresponse bias will occur if their preferences differ from those of individuals who fail to respond. This makes the relation between preference and response crucial to examine. To investigate this relation, the researcher may utilize two kinds of measured variables:
1) variables **D** for which the population distribution is known such as demographic information from a census or the results of a previous election, and 2) variables **S** for which the population distribution is unknown. Some of the potential for nonresponse bias may be alleviated by using ratio estimation to align the sample with the known distribution of **D** in the population. This method will be most effective when **D** is highly correlated with both preference and response, since in such cases **D** will explain much of the preference/response association.

Procedures to weight the sample because of an assumed correlation between preference and **S** have less validity since the relationship between **S** and response is unknown. This is especially true when a rigid deadline makes multiple mailings to resample the nonrespondents impossible. However, we can envision an ordinal function of **S** known to be correlated with response - even if the level of that correlation is not estimatable. A clear example of such a function is one which measures the strength of opinion about the issue of interest. People who feel strongly about the issues in a survey are more likely to register their preferences by responding. When strength of opinion is also related to reply on issue preference then some nonresponse bias can be expected (see for example, Armstrong & Overton, 1977; Baur, 1947; Benson, 1946; Donald, 1960; and Scott, 1961). Thus we have suggested (Pearl & Fairley, 1985) using a statistical hypothesis test to find correlation between the respondents' strength of opinion and preference as a check for this source of nonresponse bias. This paper discusses some of the properties of using one such test based on Kendall's tau (Kendall, 1938).

For simplicity of exposition, we begin in sections 2 and 3 by neglecting the variables **D**. Section 2 gives assumptions for the test to have power in detecting nonresponse bias. In section 3 we show that Kendall's tau, when appropriately scaled, gives an estimate of the size of the nonresponse bias for a particular circumstance regarding the strength of opinion of the nonrespondents. Section 4 discusses the generalizations necessary to deal with adjustments for knowledge about **D** and the final section provides some general discussion.

2. ASSUMPTIONS AND NOTATION

We assume we have survey results for a question with a binary response that we will call 'preference' (yes/no, for/against, etc.), along with an ordinal variable we call 'strength of opinion' and label S. S may be a function of several measured quantities whose population distribution is unknown but which is known to have a negative association with response. Here a value of S = 1 is given to those with the strongest opinions, presumed to be most likely to respond, while S = J corresponds to those with the least resolute opinions who are least likely to respond to a survey. Thus, the results of the poll fall in a 2 x J table:

Preference	Strength of Opinion 1	2	3	. . .	J	Total
Yes	n_1	n_2	n_3	. . .	n_J	n
No	m_1	m_2	m_3	. . .	m_J	m
	t_1	t_2	t_3	. . .	t_J	t

where n_s = number of respondents in strength category s with preference yes
m_s = number of respondents in strength category s with preference no
and $t_s = n_s + m_s$, s = 1, 2, . . . , J.

We will make three probabilistic assumptions underlying the properties of the hypothesis test to be suggested.

Assumption I. We assume that the original sample was a simple random sample.

This is merely a convenience. The results presented below can be adapted to other probability sampling methods.

Assumption II. The probability of response within a strength of opinion category is constant. In particular, the probability does not depend on people's preference within the category.

Assumption II implies $P(R|Yes,S=s) = P(R|No, S=s)$ or equivalently, $P(Yes \mid R,S=s)= P(Yes \mid NR,S=s)$, where R and NR denote Respond and does Not Respond, respectively.

Assumption II essentially says that strength of opinion explains all the bias there may be. This assumption will be relaxed in section 4 when we consider procedures which adjust for the known distribution of some variables **D**. Assumptions I and II together allow us to assume that, conditional on $t_1, t_2, \ldots, t_J$, we have a random sample within each strength of opinion category. This makes finding the expected value and distributution of the test statistic much easier.

By assumption II we can write $\beta_S = P(Yes \mid S = s, R)$, [$=P(Yes \mid S = s, NR)$] for $s = 1, \ldots, J$. Also let $\beta = P(Yes)$, $\beta_R = P(Yes \mid R)$, $\beta_{NR} = P(Yes \mid NR)$, $\alpha_{s|R} = P(S = s \mid R)$ and $\alpha_{s|NR} = P(S = s \mid NR)$ for $s = 1, \ldots, J$.

We are interested in estimating β. Define the naive estimator of β to be:

$$\hat{\beta} = n/t = t^{-1}\sum_s n_s = t^{-1}\sum_s t_s\hat{\beta}_s = \sum_s \hat{\alpha}_{s|R}\hat{\beta}_s,$$

where $\hat{\beta}_s = n_s/t_s$ and $\hat{\alpha}_{s|R} = t_s/t$.

This naive estimator is an unbiased estimator of β_R:

$$E(\hat{\beta} \mid t_1, \ldots, t_J) = t^{-1}\sum_s E(n_s \mid t_1, \ldots, t_J)$$

$$= t^{-1}\sum_s t_s\beta_s$$

since assumptions I and II imply that $n_s|t_s \sim \text{Binomial}(t_s, \beta_s)$. The bias of the naive estimator is

$$E(\hat{\beta}) - \beta = \beta_R - \beta = P(NR)[\beta_R - \beta_{NR}] = P(NR)\sum_s(\alpha_{s|R} - \alpha_{s|NR})\beta_s$$

We wish to know which conditions imply $\beta_R \neq \beta_{NR}$ and hence bias in the naive estimator. In general this is messy, but in the important case where the β_s's are strictly increasing or decreasing, it is not too hard. First we make

Assumption III. $P(NR \mid S=s)/P(R \mid S=s)$ is an increasing function of s.

This is the key assumption in what follows. It says that those who feel more strongly are more likely to respond. Now,

$$\frac{P(NR \mid S=s)}{P(R \mid S=s)} = \frac{P(S=s \mid NR)/P(NR)}{P(S=s \mid R)/P(R)} = \frac{\alpha_{s|NR}/P(NR)}{\alpha_{s|R}/P(R)}.$$

So assumption III is equivalent to $\alpha_{s|NR}/\alpha_{s|R}$ increasing in s.

3. THE HYPOTHESIS TEST

3.1 CONDITION FOR BIAS

Assumption III and decreasing (increasing) β_s's can imply bias ($\beta_R \neq \beta_{NR}$):

Theorem 1. If $\beta_1 \geq \beta_2 \geq \ldots \geq \beta_J$, $\beta_1 > \beta_J$ and $\alpha_{1|NR}/\alpha_{1|R} < \alpha_{J|NR}/\alpha_{J|R}$ then $\beta_R > \beta_{NR}$.

proof: Let the random variables

$X_R = \beta_s$ with probability $\alpha_{s|R}$,

$X_{NR} = \beta_s$ with probability $\alpha_{s|NR}$, $s = 1, \ldots, J$

(with the obvious modification if some of the β_s's are equal).

Since $\beta_s \searrow$ in s, and since Assumption III implies

$$\sum_{s=1}^{j} \alpha_{s|NR} \geq \sum_{s=1}^{j} \alpha_{s|R}$$

for $1 \leq j < J$, X_R is stochastically larger than X_{NR}.

Thus $\beta_R = E(X_R) > E(X_{NR}) = \beta_{NR}$. A similar theorem applies if the β_s 's are increasing.

3.2 HYPOTHESES

In view of the above discussion, and since the alternatives have a potential for large bias, we consider the following hypotheses:

$H_0: \beta_1 = \beta_2 = \ldots = \beta_J$

versus

$H_1: \beta_1 \geq \beta_2 \geq \ldots \geq \beta_J$ or $\beta_1 \leq \beta_2 \leq \ldots \leq \beta_J$

with $\beta_1 \neq \beta_J$. In words, H_0 says preference is constant across strength of opinion while H_1 says preference and strength of opinion are correlated.

3.3 KENDALL'S TAU

A natural statistic to use in testing H_0 versus H_1 would be of a form which provides an estimate of bias:

$$\widehat{\text{bias}} = \frac{T-t}{T} \sum_s (\hat{\alpha}_{s|R} - \alpha^*_{s|NR})\hat{\beta}_s \qquad [1]$$

where T is the original random sample size so that (T-t)/T estimates P(NR) and $\alpha^*_{s|NR}$ is some particular choice of $\alpha_{s|NR}$ satisfying $\sum \alpha^*_{s|NR} = 1$ and assumption III. One such choice is

$$\alpha^*_{s|NR} = \hat{\alpha}_{s|R} \left(\sum_{s' \leq s} \hat{\alpha}_{s'|R} + \sum_{s'<s} \hat{\alpha}_{s'|R} \right)$$

$$= \frac{t_s}{t} \left(1 - \sum_{s'>s} t_{s'}/t + \sum_{s'<s} t_{s'}/t \right).$$

It is easy to verify that these do sum to unity and that $\alpha^*_{s|NR}/\hat{\alpha}_{s|R}$ is increasing in s. For these values of $\alpha^*_{s|NR}$, [1] becomes:

$$\widehat{\text{bias}} = \frac{T-t}{T} \sum_s \left[\frac{t_s}{t} - \frac{t_s}{t}\left(1 - \sum_{s'>s} \frac{t_{s'}}{t} + \sum_{s'<s} \frac{t_{s'}}{t}\right)\right]\frac{n_s}{t}$$

$$= \frac{T-t}{Tt^2} \left\{ \sum_s \left(\sum_{s'>s} n_s m_{s'} - \sum_{s'<s} n_s m_{s'} \right) \right\} = \frac{T-t}{Tt^2} \{ \tau \}.$$

The τ in this last expression is just Kendall's tau statistic measuring the correlation between repondents' preference and strength of opinion.

Thus Kendall's tau provides an estimate of the size of the bias under certain assumptions about the nonrespondents. Of course, the quality of this estimate is unknown and we are suggesting its use only for the hypothesis testing purpose.

In using Kendall's tau to test H_0 versus H_1 , we assume conditioning on both sets of marginals. With this conditioning, Burr (1960) has found τ to be approximately Normal under H_0 with mean 0 and variance

$$\sigma_\tau^2 = \frac{mn(t^3 - \sum_s t_s^3)}{3t(t-1)} .$$

Under H_1, $\beta_s \searrow s$ implies $E(\beta_s | t_1,\ldots,t_J,n) \searrow s$, so that Theorem 1 implies that the conditional expectation of τ will be positive with decreasing β_s (Similarly, it will be negative for increasing β_s). This is true since $\alpha^*_{1|NR}/\hat{\alpha}_{1|R} > \alpha^*_{J|NR}/\hat{\alpha}_{J|R}$ except in the degenerate case where everyone is in one strength of opinion category. Therefore τ should have power against the alternatives H_1 . Also, if the test concludes that nonresponse bias is likely, then the sign of Kendall's tau indicates the likely direction of the bias, a quantity of interest in itself (Armstrong and Overton, 1977).

3.4 APPROXIMATE POWER OF τ

The exact power of τ is difficult to compute, but we can get good approximations in our case where we have large samples, at least for alternatives near H_0 .

Burr (1960), states that, in the 2 x J case, Kendall's tau is equivalent to the Mann-Whitney statistic with ties. Thus τ is asymptotically Normal under a sequence of Pitman alternatives. Again taking the values $t_1,\ldots,t_J$, n as fixed (which also fixes $\hat{\alpha}_{s|R}$ and $\alpha^*_{s|NR}$), the expected value of τ when H_0 is not true is:

$$E_{alt}(\tau) = t^2 \sum_s (\hat{\alpha}_{s|R} - \alpha^*_{s|NR}) E_{alt}(\hat{\beta}_s)$$

$$\approx t^2 \sum_s (\hat{\alpha}_{s|R} - \alpha^*_{s|NR}) \beta_s$$

Hence a large sample estimate of the power of Kendall's tau under alternative hypotheses close enough to H_0 for the Normal approximation (mean: E_{alt} , variance: σ_τ^2) to be applicable is given by

$$P(\text{reject} \mid t_s, \beta_s \text{ for } s = 1,\ldots,J)$$
$$\approx 1 - \Phi(z_{\gamma/2} - E_{alt}/\sigma_\tau) + \Phi(-z_{\gamma/2} - E_{alt}/\sigma_\tau) \quad [2]$$

where Φ is the distribution function of the standard Normal, $z_{\gamma/2}$ is given by $\Phi(z_{\gamma/2}) = 1-\gamma/2$ and γ is the chosen significance level. The accuracy of this approximation may be checked through computer simulation.

It is interesting to examine the term E_{alt}/σ_τ that comes into our approximation to the power of Kendall's tau. After some algebra we have

$$E_{alt}/\sigma_\tau \approx \frac{\text{bias*}}{\sqrt{\hat{\beta}(1-\hat{\beta})/t}} \quad \frac{3}{P(NR)}$$

where bias* is the level of nonresponse bias corresponding to the strength of opinion distributions $\hat{\alpha}_{s|R}$ and $\alpha^*_{s|NR}$ for the respondents and nonrespondents respectively. Now, $\sqrt{\hat{\beta}(1-\hat{\beta})/t}$ is just the standard error that would be reported by a researcher using the naive estimator of population preference. For mail surveys, where P(NR) is commonly in the range .6 to .8, we discover from [2] that Kendall's tau will begin to detect bias about when it is of the same order as the proclaimed sampling error. Thus, since confidence statements about a poll's accuracy are an important part of presenting a poll, the test appears to detect practical as well as statistical significance.

4.DEALING WITH DEMOGRAPHIC ADJUSTMENT

The discussion of sections 2 and 3 can be extended to cases where the researcher is able to lessen the impact of nonresponse bias by aligning the sample with the known population distribution of "demographic" variables **D**. The extension is easy since the bias in a demographically adjusted estimate is just a weighted average of the biases within each demographic group. For concreteness suppose this alignment is done over the values of a categorical quantity we call D. The data then come to us as a three-dimensional contingency table for which we expand the notation in section 2 to let

n_{sd} = number of respondents with preference yes for whom S=s and D=d,

m_{sd} = number of respondents with preference no for whom S=s and D=d

and $t_{sd} = n_{sd} + m_{sd}$, $s = 1,\ldots,J$ $d = 1,\ldots,K$.

The demographically adjusted estimator of β is then defined by

$$\hat{\beta}_D = N^{-1}\sum_d N_d \hat{\beta}_{.d} \quad \text{where}$$

N_d/N = the proportion of individuals in the population for whom D = d and $\hat{\beta}_{.d} = n_{.d}/t_{.d}$ with $n_{.d} = \sum_s n_{sd}$ and $t_{.d} = \sum_s t_{sd}$.

In order to test for nonresponse bias in the adjusted estimator, we change assumptions II and III to:

Assumption II* The probability of response is constant within a strength of opinion/demographic category, (s,d).

Assumption III* $\alpha_{s|dNR}/\alpha_{s|dR}$ increases in s for each d = 1, . . ., K where $\alpha_{s|dNR} = P(S=s \mid D=d,NR)$ and $\alpha_{s|dR} = P(S=s \mid D=d, R)$.

III* gives the reasonable assumption that within any demographic category people with stronger opinions are more likely to respond.

Assumption II* implies that

$P(Yes|D=d,S=s,R) = P(Yes|D=d,S=s,NR)$

and we can denote the common value by β_{sd}. This in turn implies that $\hat{\beta}_D$ is an unbiased estimator of β_R for given values of t_{sd} $s = 1,\ldots,J$ and $d = 1,\ldots,K$. Assumption II* is more realistic than assumtion II. How close it is to true depends on the ability of the researcher to formulate questions which isolate information related to response.

Following section 3.1, we find that if

$$\beta_{1d} \geq \beta_{2d} \geq \ldots \geq \beta_{Jd}, \beta_{1d} > \beta_{Jd}$$

and $\alpha_{1|dNR}/\alpha_{1|dR} < \alpha_{J|dNR}/\alpha_{J|dR}$ for each d, then $\beta_R > \beta_{NR}$ (with a similar result if the β_{sd}'s are increasing in s). A correlation between preference and strength of opinion within demographic groups results in bias. Thus, since

$$\beta_R - \beta_{NR} = \sum_d N_d/N \sum_s (\alpha_{s|dR} - \alpha_{s|dNR})\, \beta_{sd}$$

it is reasonable to consider tests of the hypotheses

H_{0D}: $\beta_{1d} = \beta_{2d} = ... = \beta_{Jd}$ for d=1,...,K

versus

H_{1D}: $\beta_{1d} \geq \beta_{2d} \geq ... \geq \beta_{Jd}$ and $\beta_{1d} > \beta_{Jd}$ for d = 1, ... , K or $\beta_{1d} \leq \beta_{2d} \leq ... \leq \beta_{Jd}$ and $\beta_{1d} < \beta_{Jd}$ for d = 1, ... , K.

To decide between these hypotheses, we need a test statistic which examines only that part of the correlation between preference and strength of opinion not explained by demographics. Quade, 1974, gives an excellent discussion of distribution-free measures of partial correlation which can be used to carry out this type of test. We suggest using a statistic of the form:

$$Y = \sum_d w_d \tau_d$$

where τ_d is the Kendall's tau statistic for the individuals in demographic group d and the w_d 's are some weights such that $w_d > 0$, d=1,..., K. Statistics of this form have been suggested by several authors and any such statistic will have power against alternatives H_{1D} . Kendall's tau will have less power for other alternatives when S and preference have positive correlation for some d and negative for others. However, in such cases the bias will also be less. To carry out a test of this form, we would use the fact that when H_{0D} is true, E(Y) = 0 and its variance is given by:

$$\sigma_Y^2 = \sum_d w_d^2 \Big[\frac{m_{.d}\, n_{.d}(t_{.d} - \sum t_{sd})}{3t_{.d}(t_{.d} - 1)} \Big].$$

As before, the statistic $Z_D = Y/\sigma_Y$ would approximately follow a standard normal distribution under H_{0D} when each of the $t_{.d}$'s is large. One logical choice of weights, geared toward this specific application, would take

$$w_d = N_d/N(t_{.d})^2 .$$

This choice makes

$$Y = \sum_d \frac{N_d}{N} \sum_s (\hat{\alpha}_{s|dR} - \alpha^*_{s|dNR})\, \hat{\beta}_{sd}$$

which estimates $\beta_R - \beta_{NR} = \dfrac{\text{bias}}{\text{nonresponse rate}}$ when $\alpha^*_{s|dNR}$ is the correct proportion with S = s in group d among the nonrespondents. By rejecting H_{0D} in favor of H_{1D} one could conclude that the adjusted estimate is biased.

5. DISCUSSION

This paper has shown that Kendall's tau is relevant, natural and of practical importance in detecting one source of nonresponse bias. It is relevant because, under realistic assumptions, when it detects correlation between preference and strength of opinion, it also finds bias. It is natural because tau is an estimate of the bias under certain assumptions about the nonrespondents. Finally, rejecting the null hypothesis is of practical as well as statistical significance because the power of tau becomes high when the bias is on the same order of magnitude as the proclaimed standard error.

Pearl and Fairley (1985) discuss a mail poll taken by *The Columbus Dispatch* , an Ohio newspaper, to ascertain voter preferences a week before the November 1983 election. In the poll, regarding three controversial statewide initiatives, Kendall's tau indicated strength of opinion was significantly correlated with respondent preferences. The *Dispatch* claimed with 95% confidence that the estimates of this poll were accurate to within ±3%. The results of section 3 show how carrying out the hypothesis test provided evidence that this interval was too small. Indeed, the actual errors in predicting the election were in the 10 - 15% range. It is likely that at least some of this error was due to nonresponse bias.

Investigators often assume that nonresponse bias is an ignorable feature of the problem once they have adjusted for demographics. However, surveys with response rates as low as 25% can be greatly biased if this assumption is violated. A purist would then say that such surveys should not be published at all, while a researcher pressed for time might say that a presentation need only make the assumption explicit. We have suggested a compromise. Surveys can be presented by making the less restrictive assumption that nonresponse is ignorable once demographics and strength of opinion have been accounted for and if a test fails to detect bias due to strength of opinion. But beware - many sources of error may still exist and caution should always be used in the interpretation of results.

REFERENCES

Armstrong, J. S. and Overton, T.S. (1977) "Estimating nonresponse bias in mail surveys." *Journal of Marketing Research* **14** 396-402.

Baur, E. J. (1947) "Response bias in a mail survey." *Public Opinion Quarterly* **11** 594-600.

Benson, L. E. (1946) "Mail surveys can be valuable." *Public Opinion Quarterly* **10** 234-241.

Burr, E. J. (1960) "The distribution of Kendall's score S for a pair of tied rankings. "*Biometrika* **47** 151-165.

Donald, M. N. (1960) "Implications of nonresponse for the interpretation of mail questionnaire data." *Public Opinion Quarterly* **24** 99-114.

Kendall, M. G.(1938) "A new measure of rank correlation." *Biometrika* **38** 81-93.

Pearl, D. K. and Fairley, D. (1985) "Testing for the potential for nonresponse bias in sample surveys." *Public Opinion Quarterly*: to appear Winter 1985.

Quade, D. (1974) "Nonparametric partial correlation."In *Measurement in the Social Sciences,* H.M. Blalock (ed.) Chicago: Aldine Publishing.

Scott, C. (1961) "Research on mail surveys." *Journal of the Royal Statistical Society* **124** 143-192.