AN ABSTRACT OF THE THESIS OF

Md. Abusaleh Abdullah for the degree of Doctor of Philosophy

in Statistics presented on April 27, 1979

Title: Classification Errors in the Analysis of Fourfold Tables.

Redacted for privacy

Abstract approved: G. David Faulkenberry

Misclassification introduced by fallible measurements affects the estimate of the proportion in a class as well as the comparison of proportions in different classes. In this thesis the magnitude of effects of misclassification and the importance of misclassification error rates on estimates of proportions and variances of these estimates were examined. Studies on the values of the false negative and false positive rates associated with medical screening were reviewed to determine typical levels of error rates occurring in practice. A lack of a consistent relationship between these rates was found and the common assumption of a small constant error rate in all groups being compared was violated in almost all studies.

An attempt was made to determine how robust the usual statistical procedure for analyzing a given set of data is against these classification errors. The study was carried out for the case of two independent binomial samples (very common in epidemiologic research) with the conditional model (Fisher's exact test) considered in detail under various error rates.

Substantial effects of misclassification on the estimation of parameters as well as on hypothesis testing showed the importance of estimating the values of the misclassification rates in a particular study. The randomized response technique was used to estimate error rates and the prevalence rate, π, for a situation where the true classification can only be obtained directly from the respondent, but the response has a stigmatizing nature. An unbiased estimate of π was obtained along with an expression for the variance of $\hat{\pi}$. Formulas for sample size determination for fixed cost and fixed variance problems are given.

Classification Errors in the Analysis of Fourfold Tables

by

Md. Abusaleh Abdullah

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Completed (April 27, 1979)

Commencement June 1979

APPROVED:

Redacted for privacy

Associate Professor of Statistics
in charge of major

Redacted for privacy

Chairman of Department of Statistics

Redacted for privacy

Dean of Graduate School

Date thesis is presented April 27, 1979

Typed by Jean L. Harvey for Md. Abusaleh Abdullah

ACKNOWLEDGMENT

The author expresses his indebtedness to Dr. G. David Faulkenberry who initially suggested the problem and whose guidance has been very precious all throughout this research.

To all members of the Department of Statistics the author is grateful, especially to Dr. David Birkes for his help and his kindness and to Dr. Justus Seely for his guidance in the graduate program.

The author is also grateful to Dr. Lyle D. Calvin, a friend and a teacher, for his constant help and to the computer center. Without their support this thesis could not have been completed.

Finally, I would like to thank my wife, Rashida, and my parents for their moral support.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONT'D.)

LIST OF TABLES

LIST OF TABLES (CONT'D.)

CLASSIFICATION ERRORS IN THE ANALYSIS OF FOURFOLD TABLES

CHAPTER I

INTRODUCTION, LITERATURE REVIEW, AND OBJECTIVE OF THESIS

1.1 Introduction

The effects of misclassification of categorical responses have been considered by investigators in many different fields. The effects of classification errors in sample surveys was discussed by Hansen, Hurwitz and Bershad (1961) and Madow (1965). Dunn and Buell (1950) considered classification errors in medical screening tests. While the methods used to study classification errors range from analyses in sociology (Sutcliffe [1965]), to cost utility methods (Berkson [1947]), to operations research techniques (Blumberg [1957]), the problem reduces to that of somehow evaluating the suspected extent of the errors in the data and then determining necessary adjustments to make in the analyses. The diverse models available for the study of errors of measurement have been discussed in a review paper by Cochran (1968). Attention was given to the type of mathematical model used to represent errors of

measurements, the extent to which standard techniques of analysis become erroneous and misleading if certain types of errors are present, and the techniques that are available for the numerical study of errors of measurements.

1.2 Misclassification in the Fourfold Table-- Errors in One Direction

(a) Bross Model

Bross (1954) examined misclassification for the case of the 2 x 2 table with one axis subject to errors of measurement. The model assumed that a fallible classifier is used to separate each of two independent samples into two groups. Two types of error, false negative and false positive, are ideally obtained from an evaluation study in which the fallible measurement is compared to a standard of "truth" as defined by true measurement. In this situation, we observe the following:

In sample i, we have

		Fallible measure +	Fallible measure -	
True measure	+	a_i	c_i	$a_i + c_i$
	-	b_i	d_i	$b_i + d_i$
		$a_i + b_i$	$c_i + d_i$	n_i

Then, the false negative rate, $\theta_i = E\left(\frac{c_i}{a_i + c_i}\right)$, is the probability of a true positive being incorrectly classified

negative in the ith sample and the false positive rate, $\phi_i = E(\frac{b_i}{b_i + d_i})$, is the probability of a true negative being incorrectly classified positive in the ith sample.

Let π_i = Expected proportion, in sample i, of positive results on the factor if no misclassification were present (may be considered as the prevalence rate in the epidemiologic situation)

P_i = Expected proportion, in sample i, of positive results on the factor under fallible classification.

If in sample i of size n_i, x_i individuals screen positive, the ratio $\frac{x_i}{n_i}$ is used to estimate the true proportion of positive results in the ith sample. Then under the assumption that $\theta_1 = \theta_2$, and $\phi_1 = \phi_2$,

$$E\left(\frac{x_1}{n_1} - \frac{x_2}{n_2}\right) = P_1 - P_2 = (\pi_1 - \pi_2)(1 - \theta - \phi)$$

where $P_i = \pi_i (1 - \theta_i) + (1 - \pi_i) \phi_i$ and

$x_i = a_i + b_i$, $i = 1,2$.

The significance level of the Chi-square test that $\pi_1 = \pi_2$ remains unchanged, but the power is reduced (Cochran [14]).

(b) Limitations

The limitations of this approach have been pointed out by Rubin, Rosenbaum and Cobb (1956) who studied this

scheme in order to tabulate the ratio of sample sizes required to regain the power of the no-misclassification case using cost considerations versus screening interviews. They pointed out that:

1. θ and ϕ should both be small.

2. ϕ is more important in the determination of sample size in the ranges of prevalences found in disease studies.

3. θ and ϕ are usually unknown and must be estimated.

4. Assuming the equality of error rates in the two samples is unrealistic. This is also noted by Cochran (14), Diamond and Lilienfeld (17), Newel (41), Feldman (21) and Keys and Kihlberg (29).

Walsh (1963) exploits this model further in his comparison of the null hypothesis $\pi_1 = \pi_2$ with the alternative hypothesis $\pi_1 \neq \pi_2$ rather than Bross' alternative of $P_1 \neq P_2$. The conclusion reached is that by using two samples of sizes kn_1 and kn_2 when misclassification is present, instead of samples of sizes n_1 and n_2 without misclassification, the power of the one-sided X^2 test remains the same and the efficiency is

$$\frac{1}{k} \simeq \frac{(1-\theta-\phi)^2\pi(1-\pi)}{[(1-\theta-\phi)\pi+\phi][(1-\theta-\phi)(1-\pi)+\theta]} \text{ for } \pi_1 \simeq \pi_2 \simeq \pi,$$

independent of n_1 and n_2 with θ and ϕ small.

The general result for this model that misclassification damps the observed difference between the two proportions has been contradicted by Diamond and Lilienfeld (17) who claim that an observed difference between proportions and a relative risk of significant magnitude can result in whole or in part from misclassification errors. In essence they analyze their data under two alternative assumptions:

1. Prob. [true positive/stated positive] and Prob. [true negative/stated negative] are equal in the two groups being compared. This leads to differences in proportions and relative risks that are greater than or equal to the true differences.

2. Prob.[stated positive/true positive] = $1 - \theta$ and Prob.[stated negative/true negative] = $1 - \phi$ are equal for cases and controls. Buell and Dunn (1964) point out that only this later assumption is reasonable and is indeed the Bross assumption. The inconsistent finding of Diamond and Lilienfeld is the result of confusion of the definitions of false positive and false negative (Newell [41]). That this contradiction should occur in this model is a consequence of the lack of independence between Prob.[true positive/stated positive] and the disease prevalence which may differ from one study to another. Diamond and Lilienfeld reply to this criticism

by noting that unless the relationship between true and observed status is determined separately for cases and controls, there is no way to ascertain which assumption is appropriate.

Further clarification of the issues raised has been attempted by Keys and Kihlberg (1963) who point out that the problems in logic are caused by estimating the relative prevalence of an attribute in one population group as compared with another population group when the error rates are not necessarily equal in the two groups (the same criticism noted above for Bross's model). In particular these authors suggest that instead of examining the relative risk as Diamond and Lilienfeld do, one should compare π_i with P_i by considering the rate $\frac{P_i}{\pi_i}$ for each of two groups. Or they advise the comparison of the relative risk with misclassification to that without misclassification by means of the ratio of the two risks.

1.3 One Direction Misclassification in the r x c Table

The model considered so far for the 2 x 2 table with errors of measurement along one axis has been generalized to the 2 x c table by Mote and Anderson (1965).

If there is no misclassification, the usual test criterion is the X^2_{c-1}. The rejection of the null

hypothesis on the P's is equivalent to rejection of the null hypothesis on the π's without misclassification.

Mote and Anderson extend their results to two-way contingency tables with one-way errors (known) only in the jth categories for the following situations:

i) stratified sampling

Let, r = number of independent samples

n_i = number of observations in the ith sample

c = number of categories j = 1,2,...c

$\theta_{ijj'}$ = probability of the observation in cell (i,j) being misclassified into cell (i,j'); $\sum_{j'=1}^{c} \theta_{ijj'} = 1$.

The $\theta_{ijj'}$ are independent of i and j' and therefore equal for all i samples. π_{ij} and p_{ij} are the expected probabilities of being observed in cell (ij) without and with misclassification.

For the null hypothesis that $\pi_{ij} = \pi_{.j}$, the same X^2 criterion is used as if there were no errors of measurement. The power is reduced.

ii) random sampling

If N individuals are taken at random from one population and classified with respect to two variables with errors in only one of the variables, the test procedure is the same as that given above to test independence when errors are assumed known.

1.4 Misclassification in Two Directions in the 2 x 2 Table Under Simplifying Assumptions

We again consider the fourfold table with errors of measurement now permitted in both variables. This situation has been examined by Rogot (1961) who compared morbidity rates in two samples with varying degrees of misclassification. His results support those of Bross in that the power of the test to detect differences in the two proportions is reduced. That the difference between the two proportions or prevalences may be either under or overestimated when errors are present has been shown by Keys and Kihlberg (1963), Gullen et al. (1968) and others. These latter authors provide a two-directional analogue to the Bross result. Under the assumption of independence between factor and disease, $P_1 - P_2 = (\pi_1 - \pi_2)\, k$, where k is a function of the four misclassification rates and the probability of being in the disease or disease-free groups. Under the assumption of equality of error rates for each of the categories in the two directions, they show that the difference between the two proportions is always damped, the significance level of the test remains unchanged, and power is lost.

1.5 Two-Directional Misclassification in the r x c Table

The general theoretical framework for errors of misclassification in the r x c contingency has been given by Assakul and Proctor (1967). When the joint density of the cells is multinomial, $X^2_{(r-1)(c-1)}$ is the traditional significance test used. If we let $\Theta = (\theta_{sikj})$ be the error matrix, nonsingular and known, then two situations may be considered, where θ_{sikj} is the probability of an individual truly in cell (s,k) being classified in (i,j).

(i) independent errors

The first case is that of independent errors and the hypothesis being tested is $H_0 : \pi_{ij} = \pi_i \cdot \pi_j$

As before, the appropriate test of H_0 is equivalent to a test of $H_0 : P_{ij} = P_i \cdot P_j$, with the result that the test statistic and significance level remain unchanged. The ratio of the noncentrality parameter with misclassification to that without misclassification is used to provide a measure of the effects of errors on the power of the X^2 test.

(ii) non-independent errors

For the case of non-independent errors, the significance level of the X^2 increases and may be found as the tail of a non-central X^2. If Θ is known, the π_{ij} may be

expressed as functions of r x c - 2 parameters which are estimated by minimum X^2 method. If some of the θ's are unknown, they may then be estimated.

Assakul and Proctor point out that the effects of misclassification become more serious as the significance level (α) decreases.

1.6 Alternative Model for Errors of Measurement in 2 x 2 Tables

Sample Survey Approach

Using a regression framework, the Census Bureau developed a model for incorporating sampling variability from repeated measurements into the study of misclassification of qualitative data. The basic model developed by Hansen, Hurwitz and Bershad (1961) considers the case of continuous data with the mean square error taken as the appropriate measure of variability rather than the standard error. When this model is reinterpreted for binomial data, the results are those of Bross (Cochran [14]). The error variance and the index of inconsistency were given.

The advantages of this model are its generality, since no assumptions are made about the distributions of the parameters, and the use of MSE (Koch [1969]). This model has been further considered by Fellegi (1964) who

uses this approach to include interviewer effects combining re-enumeration and interpenetrating sample methods and is therefore able to estimate more parameters than either method alone.

For the effects of errors of measurement in qualitative data a model that involves no assumptions about the distribution of measure of association has been postulated by Koch (1969). The methods involve decomposing the mean square errors of measures of association into bias, response variance, and sampling variance with the index of

$$\text{inconsistency} = \frac{\text{response variance}}{\text{response variance + sampling variance}}$$

1.7 Objectives of the Thesis

As we can see from the literature review given above, many aspects of the problem of misclassification are still open. We have in this thesis

(a) examined the effects of misclassification on estimates of different levels of prevalence rates and considered the relative importance of false positive and false negative rates;

(b) studied the values of error rates actually occurring in research and investigated the validity of the assumptions made about these error parameters in the analysis of data with misclassification;

(c) determined how robust the usual statistical procedure for analyzing a given set of data is against

these classification errors. (This study was carried out for the case of two independent binomial samples (very common in epidemiologic research) with the conditional model (Fisher's exact test) considered in detail under various error assumptions. The values of the error parameters studied were chosen from levels found in Chapter 2 of this thesis; levels of prevalence were selected from values for diseases in which screening is of concern.); and

(d) estimated binomial proportions from data subject to misclassification using the randomized response technique. This problem was considered as a special type of misclassification where the true classification can only be obtained directly from the respondent, but the response has a stigmatizing nature. The maximum likelihood estimation procedure was used and an expression for the variance of the estimator is given. Sample size determination was also considered for fixed cost and for fixed variance.

CHAPTER II

FALSE NEGATIVE AND FALSE POSITIVE RATES IN MEDICAL RESEARCH

2.1 Introduction

Whenever some fallible measurement or screening test is used in place of a true measurement in epidemiologic research, the misclassification introduced by the fallible measurement affects the estimates of disease prevalence as well as comparisons of prevalence rates in different groups. The extent to which these measures are affected is a function of the true prevalence rate, π, the false negative rate, θ, and the false positive rate, ϕ. A starting point in the study of misclassification and its effects would be (a) to examine the effects of different levels of error rates on the estimate of π; (b) to study the values of error rates actually occurring in research and to investigate the validity of assumptions made about the error parameters in the analysis of data subject to misclassification.

We first present some straightforward calculations to show the effect of θ and ϕ on the observed prevalence rate, x/n, and on the variance of x/n. Then we have a report on a few representative illustrations of screening situations (diabetes, heart disease, etc.) found in the literature. For each of the illustrated examples we have demonstrated

actual values of θ and ϕ occurring and the validity of assumptions that may be made about them.

2.2 Magnitude of Effects of Misclassification on Estimation of π

The effects of misclassification on estimates of π were examined for different combinations of π, θ, and ϕ with the hope of clarifying the extent to which misclassification affects prevalence estimates. This is done for the values of π ranging from .05 to .95 and values of θ and ϕ from 0.00 to 1.00, i.e. from perfect classification to complete misclassification.

For π= .05, .25 .50 .95, the effects of misclassification on estimation of π for different values of θ and ϕ are given below. (For other values of π see Appendix B, Tables 1- 9 .)

The relative importance of θ and ϕ depends very much on the magnitude of π, the true prevalence of the disease. When π is small (i.e. a rare disease), the false positive rate, ϕ, has a very marked effect on P, which is the expected prevalence rate based on the results of fallible measurements. Even a small false positive rate may lead to a severe over-estimate of the true prevalence. For example when π = .05 (Table 2-a) P is usually greater than π for all false positive rates greater than .05. There is a linear increase in the amount of

Table 2-a

Expected Value of Observed Proportion for Different Error Rates, When $\pi = .05$

$$E\left(\frac{x}{n}\right) = \pi(1 - \theta) + (1 - \pi)\phi$$

$$\pi = .05$$

ϕ	θ								
	.000	.025	.05	.10	.25	.45	.50	.75	1.00
.00	.0500	.0480	.0475	.0450	.0375	.0275	.0250	.0125	.0000
.01	.0595	.0583	.0570	.0545	.0470	.0370	.0345	.0220	.0095
.02	.0690	.0678	.0665	.0640	.0565	.0465	.0440	.0315	.0190
.05	.0975	.0963	.0950	.0925	.0850	.0750	.0725	.0600	.0475
.10	.1450	.1438	.1425	.1400	.1325	.1225	.1200	.1075	.0950
.15	.1925	.1913	.1900	.1875	.1800	.1700	.1675	.1550	.1425
.20	.2400	.2388	.2375	.2350	.2275	.2175	.2150	.2025	.1900
.30	.3350	.3338	.3325	.3300	.3225	.3125	.3100	.2975	.2850
.40	.4300	.4288	.4275	.4250	.4175	.4075	.4050	.3925	.3800
.50	.5250	.5238	.5200	.5200	.5125	.5025	.5000	.4875	.4750
.75	.7625	.7613	.7575	.7575	.7500	.7400	.7375	.7250	.7125
.90	.9050	.9038	.9025	.9000	.8925	.8825	.8800	.8675	.8550
1.00	1.0000	.9988	.9975	.9750	.9825	.9775	.9750	.9625	.9500

Table 2-b

Variance of Observed Proportion for Different Error Rates When $\pi = .05$

$$\mathrm{Var}(x/n) = \pi \frac{(1 - \pi)}{n} \left[(1 - \theta - \phi)^2 + \frac{\theta(1 - \theta)}{1 - \pi} + \frac{\phi(1 - \phi)}{\pi}\right]$$

$n = 10 \quad \pi = .05$

ϕ \ θ	.00	.025	.05	.10	.25	.45	.50	.75	1.00
.00	.004750	.004642	.004524	.004297	.003609	.002674	.002437	.001234	.000000
.01	.005596	.005490	.005375	.005153	.004479	.003569	.003331	.002152	.000941
.02	.006426	.006320	.006208	.005990	.005331	.004434	.004206	.003051	.001864
.05	.008799	.008703	.008597	.008394	.007777	.006938	.006724	.005640	.004524
.10	.012397	.012312	.012219	.012040	.011494	.010749	.010560	.009594	.008597
.15	.018240	.015624	.015390	.015234	.014760	.014110	.013944	.013098	.012219
.20	.020484	.018177	.018109	.017977	.017574	.017019	.016877	.016149	.015390
.30	.022277	.022238	.022194	.022110	.021849	.021484	.021390	.020899	.020377
.40	.024510	.024493	.024474	.024437	.024319	.024144	.024097	.023844	.023560
.50	.024937	.024943	.024949	.024960	.024984	.024999	.025000	.024984	.024930
.75	.018109	.018172	.018240	.018369	.018750	.019240	.019359	.019938	.020484
.90	.008598	.008695	.008799	.009000	.009594	.010369	.010560	.011494	.012397
1.00	.000000	.000120	.000249	.000497	.001234	.002199	.002437	.003609	.004750

Table 2-c

Expected Value of Observed Proportion for Different Error Rates When $\pi = .25$

$$E\left(\frac{x}{n}\right) = \pi(1 - \theta) + (1 - \pi)\phi$$

$\pi = .25$

ϕ \ θ	.00	.025	.05	.10	.25	.45	.50	.75	1.00
.00	.2500	.2438	.2375	.2250	.1875	.1375	.1250	.0625	.0000
.01	.2575	.2513	.2450	.2325	.1950	.1450	.1325	.0700	.0075
.02	.2650	.2588	.2525	.2400	.2025	.1525	.1400	.0775	.0150
.05	.2875	.2813	.2750	.2625	.2250	.1750	.1625	.1600	.0375
.10	.3250	.3188	.3125	.3000	.2625	.2125	.2000	.1375	.0750
.15	.3625	.3563	.3500	.3375	.3000	.2500	.2375	.1750	.1125
.20	.4000	.3938	.3875	.3750	.3375	.2875	.2750	.2125	.1500
.30	.4750	.4688	.4625	.4500	.4125	.3625	.3500	.2875	.2250
.40	.5500	.5438	.5375	.5250	.4875	.4375	.4250	.3625	.3000
.50	.6250	.6188	.6125	.6000	.5625	.5125	.5000	.4375	.3750
.75	.8125	.8063	.8000	.7875	.7500	.7000	.5000	.6250	.5625
.90	.9250	.9188	.9125	.9000	.8625	.8125	.8000	.7375	.6750
1.00	1.0000	.9938	.9875	.9750	.9375	.8875	.8750	.8125	.7500

Table 2-d

Variance of Observed Proportion for Different Error Rates When $\pi = .25$

$$\mathrm{Var}(x/n) = \frac{\pi(1-\pi)}{n}\left[(1-\theta-\phi)^2 + \frac{\theta(1-\theta)}{1-\pi} + \frac{\phi(1-\phi)}{\pi}\right]$$

$\pi = .25 \quad n = 10$

ϕ \ θ	.00	.025	.05	.10	.25	.45	.50	.75	1.00
.00	.018750	.018436	.018109	.017437	.015234	.011859	.010937	.005859	.000000
.01	.019119	.018815	.018497	.017844	.015697	.012398	.011494	.006510	.000744
.02	.019477	.019182	.018874	.018240	.016149	.012924	.012040	.007149	.001478
.05	.020484	.020217	.019937	.019359	.017437	.014438	.013609	.009000	.003609
.10	.021937	.021717	.021484	.021000	.019359	.016734	.016000	.011859	.006937
.15	.023109	.022935	.022750	.022359	.021000	.018750	.018109	.014438	.009980
.20	.024000	.023872	.023734	.023437	.022359	.020484	.019937	.016734	.012750
.30	.024937	.024903	.024859	.024750	.024234	.023109	.022750	.020484	.017430
.40	.024750	.024808	.024859	.024937	.024984	.024609	.024437	.023109	.021000
.50	.023438	.023589	.023734	.024000	.024609	.024984	.025000	.024609	.023438
.75	.015234	.015618	.016000	.016734	.018750	.021000	.021084	.023438	.024609
.90	.006938	.007461	.007984	.009000	.011859	.011859	.016000	.019359	.021937
1.00	.000000	.000616	.001234	.002438	.005859	.007984	.010937	.015234	.018750

Table 2-e

Expected Value of Observed Proportion for Different Error Rates When $\pi = .50$

$$E\left(\frac{x}{n}\right) = \pi(1 - \theta) + (1 - \pi)\phi$$

$\pi = .50$

θ

ϕ	.00	.025	.05	.10	.25	.45	.50	.75	1.00
.00	.5000	.4875	.4750	.4500	.3750	.2750	.2500	.1250	.0000
.01	.5050	.4925	.4800	.4550	.3800	.2800	.2550	.1300	.0050
.02	.5100	.4975	.4850	.4600	.3850	.2850	.2600	.1350	.0100
.05	.5250	.5125	.5000	.4750	.4000	.3000	.2750	.1500	.0250
.10	.5500	.5375	.5250	.5000	.4250	.3250	.3000	.1750	.0500
.15	.5750	.5625	.5500	.5250	.4500	.3500	.3250	.2000	.0750
.20	.6000	.5875	.5750	.5500	.4750	.3750	.3500	.2250	.1000
.30	.6500	.6375	.6250	.6000	.5250	.4250	.4000	.2750	.1500
.40	.7000	.6865	.6750	.6500	.5750	.4750	.4500	.3250	.2000
.50	.7500	.7375	.7250	.7000	.6250	.5250	.5000	.3750	.2500
.75	.8750	.8625	.8500	.8250	.7500	.6500	.6250	.5000	.3750
.90	.9500	.9375	.9250	.9000	.8250	.7250	.7000	.5750	.4500
1.00	1.0000	.9875	.9750	.9500	.8750	.7750	.7500	.6250	.5000

Table 2-f

Variance of Observed Proportion for Different Error Rates When $\pi = .50$

$$\text{Var}(x/n) = \frac{\pi(1-\pi)}{n}\left[(1 - \theta - \phi)^2 + \frac{\theta(1-\theta)}{1-\pi} + \frac{\phi(1-\phi)}{\pi}\right]$$

$\pi = .50 \quad n = 10$

ϕ	.00	.025	.05	.10	.25	.45	.50	.75	1.00
.00	.025000	.024984	.024937	.024750	.023438	.019938	.018750	.010938	.000000
.01	.024997	.024994	.024960	.024797	.023560	.020160	.018997	.011310	.000497
.02	.024990	.024999	.024977	.024840	.023677	.020378	.019240	.011678	.000990
.05	.024937	.024986	.025000	.024937	.024000	.021000	.019937	.012750	.002437
.10	.024700	.024859	.024937	.025000	.024437	.021938	.021000	.014438	.006750
.15	.026437	.024609	.024750	.024937	.024750	.022750	.021937	.016000	.006937
.20	.024000	.024234	.020437	.020750	.024937	.023438	.022750	.017438	.009000
.30	.022750	.023109	.023438	.024000	.024937	.024438	.024000	.019938	.012750
.40	.021000	.021484	.021937	.022750	.024437	.024938	.024750	.021938	.016000
.50	.018750	.019359	.019937	.021000	.023638	.024938	.025000	.023438	.018750
.75	.010937	.011859	.012750	.014437	.018750	.022750	.023438	.025000	.023438
.90	.004750	.005859	.006937	.009000	.014437	.019938	.021000	.024438	.024750
1.00	.000000	.001234	.002437	.004750	.010937	.017438	.018750	.023438	.025000

Table 2-g

Expected Value of Observed Proportion for Different Error Rates When $\pi = .95$

$$E\left(\frac{x}{n}\right) = \pi(1 - \theta) + (1 - \pi)\phi$$

$\pi = .95$

ϕ \ θ	.00	.025	.05	.10	.25	.45	.50	.75	1.00
.00	.9500	.9263	.9025	.8550	.7125	.5225	.4750	.2375	.0000
.01	.9505	.9268	.9030	.8555	.7130	.5230	.4755	.2380	.0005
.02	.9510	.9273	.9035	.8560	.7135	.5235	.4760	.2385	.0010
.05	.9525	.9288	.9050	.8575	.7150	.5250	.4775	.2400	.0025
.o0	.9550	.9313	.9075	.8600	.7175	.5275	.4800	.2425	.0050
.15	.9575	.9338	.9100	.8625	.7200	.5300	.4825	.2450	.0075
.20	.9600	.9363	.9125	.8650	.7225	.5325	.4850	.2475	.0100
.30	.9650	.9413	.9175	.8700	.7275	.5375	.4900	.2525	.0150
.40	.9700	.9463	.9225	.8750	.7325	.5425	.4950	.2575	.0200
.50	.9750	.9513	.9275	.8800	.7375	.5475	.5000	.2625	.0250
.75	.9875	.9638	.9400	.8925	.7500	.5600	.5125	.2750	.0375
.90	.9950	.9713	.9475	.9000	.7575	.5675	.5200	.2825	.0450
1.00	1.0000	.9763	.9525	.9050	.7625	.5725	.5250	.2875	.0500

Table 2-h

Variance of Observed Proportion for Different Proportion, When $\pi = .95$

$$\text{Var}(x/n) = \frac{\pi(1-\pi)}{n}\left[(1-\theta-\phi)^2 + \frac{\theta(1-\theta)}{1-\pi} + \frac{\phi(1-\phi)}{\pi}\right]$$

$\pi = .95 \quad n = 10$

ϕ \ θ	.00	.025	.05	.10	.25	.45	.50	.75	1.00
.00	.004750	.006827	.008799	.012397	.020484	.024949	.024937	.018109	.00000
.01	.004705	.006784	.008759	.012362	.020463	.024947	.024940	.018136	.00005
.02	.004660	.006741	.008719	.012326	.020442	.024945	.024942	.018162	.00010
.05	.004524	.006613	.008598	.012219	.020377	.024938	.024949	.018240	.00024
.10	.004298	.006398	.008394	.012040	.020269	.024924	.024960	.018369	.00049
.15	.004069	.006182	.008190	.011859	.020160	.024910	.025969	.018499	.00074
.20	.003840	.005964	.007984	.011678	.020049	.024894	.024977	.018624	.00099
.30	.003378	.005525	.007569	.011310	.019824	.024859	.024990	.018874	.00147
.40	.002910	.005082	.007149	.010938	.019594	.024819	.024977	.019119	.00196
.50	.002438	.004633	.006724	.010560	.019359	.024774	.025000	.019359	.00243
.75	.001234	.003489	.005640	.009594	.018750	.024640	.024984	.019938	.00360
.90	.000498	.002788	.004974	.009000	.018369	.024544	.024960	.020269	.00429
1.00	.000000	.002314	.004524	.008598	.018109	.024474	.024937	.020484	.00475

bias as ϕ increases. We note from the model given above that P can be considered as a linear function of ϕ with slope $(1 - \pi)$ and intercept $\pi(1 - \theta)$. On the other hand, when π is small, the false negative rate, θ, has little effect on the estimate of π. As the prevalence of the disease (π) increases, the false positive rate has a decreasing effect on P and θ becomes more important--a trend that may be clearly seen from Appendix B, Tables 1- 9. When π is large (e.g., .95) (Table 2-g) the estimate of P based on the screening test (fallible) is generally an underestimate of the true prevalence for all false negative rates greater than .05. The estimate is now relatively unaffected by the false positive rate. The same conclusion can be drawn on the effects of error rates on the variance of $\hat{\pi}$ (see tables 2-b, 2-d, 2-f and 2-1).

In practice, one should expect small prevalence to be overestimated with the degree of bias being determined mainly by the false positive rate. Large prevalence rates tend to be underestimated with the false negative rate influencing the extent. For prevalence in the neighborhood of .50, both of the error rates are important and the true prevalence may be underestimated, overestimated or unbiased (Table 2-e). It is apparent that the estimates of π whenever even small amounts of misclassification are present can be strikingly different from the true π. Thus, there is a need for knowledge about the values of θ and ϕ which

actually occur in order to ascertain reasonable ways to adjust for the presence of errors in a particular study. The statistical models given by Bross (6), Mote and Anderson (38), Assakul (1965), and Assakul and Proctor (1967) for handling misclassification generally assume that the parameters are known. But this is generally not the case. Frequently, the investigator makes no effort to evaluate the procedures used to classify subjects in a study. He may have no "truth" against which to compare his procedure. In other instances one error rate may be known or found, but the other error rate remains unknown. We can see this in the following examples from medical literature.

Cooly et al. (1960) studied the barium enema as a diagnostic technique for cancer of the colon and found a false negative rate of .10 but no false positive rate. (This is an instance where surgery provides confirmation of the test, but is obviously not performed unless there is other clinical evidence of disease.) Mateer et al. (1943) evaluate liver function tests with respect to false negative rates but not false positive rates because of the difficulties in dealing with young, healthy subjects. Sosman (1950) reports false positive rates but no false negative rates for gastrointestinal radiologic techniques because no negatives on x-ray are subjected to surgery. Thus it is clear that it is not always easy to find both the false positive and false negative rates associated with a

screening procedure. There is need to make some effort to have the knowledge of both error rates in a particular study. Often this is done by repeated readings of x-rays or repeated use of the screening tests, although at greatly increased costs.

2.3 False Negative and False Positive Rates As Found in Literature for Various Diseases

The literature was reviewed in detail in order to determine the levels of the error parameters and any possible general trends in these levels, and check the validity of the assumptions usually made about the misclassification parameters. These assumptions include:

(i) θ and ϕ are constant and small in all groups being compared (6, 26, 43).

(ii) θ and ϕ are inversely related and this relationship is the same in all subgroups under study (1, 2, 38).

(iii) Equality of error rates in two groups (6, 14, 17, 21, 29, 41).

The data found in the medical literature reflect a variety of procedures used as screening tests. These methods include the simplest tests based on a single dichotomous measure, those tests based on combinations of dichotomous measures and those based on one or more continuous measurements. Continuous data in a screening curve situation in which the cutoff point for dividing the sample into

the well and the ill can be chosen subjectively by the investigator so that the efficiency of the test is optimized. There are situations in which several dichotomous tests are used singly or in combinations in order to vary the stringency of the criteria for calling an individual positive on screening. Repeated independent observations of the same test may be similar to the situation of varying the criterion in a screening test since each reader is applying a slightly different definition for calling an individual positive. Also, the same test may be used at different levels in different studies and if the data from different sources are comparable, a screening curve might be used to determine the best level of the test in general.

2.4 Screening Curves

(a) Diabetes

We shall first consider an example of a screening curve for diabetes. The presence of diabetes may be detected using blood-sugar level (somogyi-Nelson method) compared to the definitive diagnostic procedure which is a complete clinical examination and glucose tolerance test (32). Since this blood sugar test is available at several levels and evaluation is possible at each, a screening curve can be obtained. The "best" cutoff level chosen by these

investigators (for the purposes of dividing the population into two groups) is 130 mg/100 ml, resulting in a false negative rate, θ, of 0.35 and a false positive rate of .02.

Table 2-i below gives the false negative and false positive rates due to different blood sugar levels considered as positive.

Diabetes is one of the diseases for which alternative methods of screening are available. Fasting blood tests (Wilkerson-Heftman with larger θ and ϕ than Somogyi-Nelson), Fasting urine tests (copper reduction, bismuth reduction both of which have larger error rates than the blood tests $\theta > .6$ and $\phi > .07$) and combinations of these tests (Kurlander [32]). If all four tests are combined, θ reduces to .286 and ϕ increases to .190 when the definition of positive on screening is positive on any one of the four tests.

The intervaneous tolbutamide test has also been evaluated against the Somogyi-Nelson method blood sugar test rather than the standard glucose tolerance test. θ was found $\leq .15$ and $\phi \leq .06$. Other tests have been dichotomized for the purpose of distinguishing diabetics and non-diabetics with false negative rates smaller than the false positive rates ($\theta \leq .15$, $\phi \leq .33$)(39). This

Table 2-i

Screening for Diabetes

Blood Sugar Level Somogyi-Nelson Method mg/100 ml	θ	ϕ
80	.00	.59
90	.04	.33
100	.13	.18
110	.26	.09
120	.30	.03
130	.35	.02
140	.44	.01
150	.48	.01
160	.61	.00
170	.70	.00
180	.78	.00
190	.78	.00

situation is the reverse of that found for the diabetes screening tests currently in use. It is a reflection of the general trend toward refinement of techniques as the result of increased knowledge about the nature of the disease and technological advancements which serve to reduce the error rates. Table 2-j gives the false negative and false positive rates due to different screening procedures.

(b) Heart Disease and Hypertension

A similar situation is observed for the case of screening for hypertension and hypertensive heart disease using blood pressure levels (31) compared to clinical examination. Table 2-k gives θ and ϕ values for various

Table 2-j

Diabetes Screening

Test	θ	ϕ	Literature
Urine			
Copper reduction (Clinitest)	.667	.150	Kurlander
Bismuth reduction (Gulatest)	.619	.071	et al. (31)
Blood			
Wilkerson-Heftman	.429	.031	
Somogyi-Nelson >130 mg/100 ml	.333	.022	
Positive on 1 test or negative on all 4 tests	.286	.190	
Intravenous Tolbutamide Response Test evaluated against the Somogyi-Nelson blood sugar			Unger (50)
20-minute test			
Cutoff at 80% fasting blood sugar	.05	.06	
Cutoff at 85% fasting blood sugar	.04	.04	
30-minute test			
Cutoff at 70% fasting blood sugar	.10	.05	
Cutoff at 75% fasting blood sugar	.15	.03	
1 dose oral standard (cutoff at 120 mg/cc)	.346	.010	Moyer and Womack (39)
2 dose one-hour test			
a) Exton-Rose	.00	.363	
b) Gould-Altschuler modification	.03	.00	
Moyer-Womack use of Exton-Rose	.08	.330	
Moyer-Womack use of modification	.15	.35	
1-hr. oral glucose tolerance test	.00	.15	National
100 gm. urine glucose vs. 50 gm.	.375	.350	Health Survey (40b)

Table 2-k

Screening for Heart Disease

Blood Pressure Level	θ	ϕ
180/100	.39	.09
160/100	.27	.10
160/96	.22	.14
150/100	.22	.14
150/96	.19	.18
160/90	.15	.24
150/90	.14	.26
140/90	.09	.29

cutoff points. Both θ and ϕ cover only narrow ranges of the values with $\theta \leq .39$ and $\phi \leq .29$. The best choice for defining the well and the ill was a reading of 150 mm Hg or over systolic or 90 mm Hg or over diastolic to call an individual positive. It can be seen that the θ and ϕ levels are not small even for the best choice. The relationship between the parameters is reciprocal.

Various techniques have been employed to screen for hypertension and hypertensive heart disease. These methods range from medical histories to chest x-rays, blood pressure levels (discussed above) and electrocardiograms used individually and in combinations. The standard use for comparison is the clinical examination. The National Health Survey (40a) used medical histories as a screen and

found the false negative rate $\theta = .620$, the false positive rate $\phi = .086$ which implies that few people would say they do have heart disease when in fact they do not, but that many do not reveal that they are ill or do not know that they are ill. Kurlander et al. (31) found $\theta = .284$, $\phi = .528$ for physician histories and $\theta = .580$ $\phi = .324$ for clerk histories. There is little consistency between physician histories and clerk histories in the same study and between different studies. The National Health Survey (40a) and Wylie (1962) data reveal the same trend of θ and ϕ unlike the Kurander et al. (31) data.

Chest x-rays have also been used by several investigators (12, 31, 53) as a dichotomous screening test for heart disease (see Table 2-1). The results vary with the reader, with the definition of positive results on the test, and with the type of x-ray. In general, θ is fairly large, ϕ is small. The relationship between θ and ϕ is inverse. The variability from study to study may reflect differences in the criteria used to define positives on screening.

Electrocardiograms also have fairly large θ values associated with their use (Table 2-1) and the two studies using ECG show inverse relationship between θ and ϕ. When blood pressure levels are used to screen for heart disease, Kurlander (31) chose the best cutoff for the positive and negative groups at 150 mm Hg systotic or 90 mm Hg diastolic pressure with the result that $\theta = .137$ and $\phi = .260$

Table 2-1

Hypertension and Hypertensive Heart Disease

Medical Histories	θ	ϕ
National Health Survey (40a)	.620	.086
Kurlander et al. (31)		
Physicians	.284	.528
Clerks	.580	.324
Wylie (53) death rates	.753	.060
X-rays		
Chapman et al. (12) - 70 mm photofluorograms		
Reader A	.29	.29
Reader B	.51	.14
Heart enlarged	.70	.05
All CV abnormal	.52	.10
Kurlander et al. (31)		
35 mm	.679	.042
70 mm	.445	.100
All x-ray	.524	.079
Wylie (53) death rates 70 mm	.443	.177
ECG		
Kurlander et al. (31)	.494	.244
Wylie (53) death rates	.516	.121
Blood Pressure		
Kurlander et al. (31)	.137	.260
Wylie (53) death rates	.762	.065

(see Table 2-k). Wylie (53), using death rates, found $\theta = .762, \phi = .065$. There is considerable variability between the different studies with regard to these screening procedures.

(c) Cervical Cancer

Screening for cervical cancer is ordinarily carried out using the Papanicolan smear. In early studies using poorly defined criteria, Meigs et al. (37) evaluated vaginal smears against the definitive diagnosis of cancer provided by biopsy of the cervix, endometrium or uterus. They found false negative rate $\theta = .104$ and false positive rate $\phi = .30$, and for the Papanicolan smear compared with complete clinical diagnosis, $\theta = .14$, and $\phi = .003$. More recent work has focused on refinements in technique, in particular the use of instruments such as the cytoanalyzer to read smears and upon criteria for calling a smear positive.

(d) Rheumatism Symptoms

History of rheumatism or arthritis may be used to screen for the presence of the disease in several different settings such as household survey questions, individual interviews and physician interviews (Cobb et al. [1955]). The differences in the false positive rates and false negative rates obtained with those methods are slight. The false positive rates are smaller than false negative rates with a reciprocal relationship between the two rates.

(e) Syphilis Screening

Several alternative tests are available to screen for syphilis. These include the TPCF which is a complement fixation test, the fluorescent treponemal antibody test (FTA), the VDRL--a slide used for serologic testing with the cardiolipin antigen, the rapid plasma Reagin (RPR) which uses potassium oxalate and the Kolmer-Reiter Protein (KRP) which uses treponemal antigen (Jolly et al. [1960]). Most of the available data for the evaluation of these dichotomous tests involves comparisons of one test with another instead of with a "true standard." Buck and Mayer (1964) studied different groups of diseased subjects in Ethiopia comparing the VDRL with the RPR cord test, considering the VDRL the procedure of choice. They found a wide range of θ (0 - .42) and ϕ(.03 - .50) values for the RPR test (Table 2-m) in different groups. Buck and Spruyt (1964) examined VDRL against the FTA as a standard in different age groups as well as different disease groups also in Ethiopia. The different age groups have varying false negative and false positive rates, having the largest θ = .67 and the smallest ϕ = .04 (Table 2-m).

The National Health Survey evaluated the relative sensitivity and specificity of the VDRL and the KRP tests for males and females as well as for blacks and for whites (Table 2-m). There is greater variability in θ in the KRP test when the VDRL is taken as a standard with blacks having

Table 2-m

Syphilis Screening

A. RPR card test (compared to VDRL as standard) - Buck and Mayer (8)

Subgroup	θ	ϕ
Patients with early syphilis	.021	.50
Healthy controls	.059	.10
Newly registered VD clinic outpatients	.116	.123
Patients with tuberculoid leprosy	.000	.025
General population in area of hyper-endemic leprosy	.226	.368

B. VDRL (compared to FTA as standard) - Buck and Spruyt (10)

Controls	.25	.08
Lepromatous leprosy	.14	.32
Tuberculoid leprosy	.50	.00
General population in area of hyper-endemic malaria	.00	.49
Age groups		
0-4	.67	.04
5-14	.20	.14
15-24	.27	.14
25-34	.30	.34
35-44	.20	.47
45+	.22	.35

C. KRP (VDRL as standard) - National Health Survey (40c)

Males	.54	.023
Females	.78	.025
Negro males	.29	.118
Negro females	.32	.103
White males	.69	.012
White females	.87	.018

VDRL (KRP as standard) National Health Survey (40c)

Males	.49	.028
Females	.64	.036
Negro males	.45	.070
Negro females	.55	.043
White males	.52	.025
White females	.77	.035

(Cont'd. on next page.)

Table 2-m (Cont'd.)

D. RPR, VDRL, TPCF (compared to clinical diagnosis) -

Jolly (28)

Test	θ	ϕ
RPR	.04	.02
VDRL	.16	.02
TPCF	.08	.00
RPR rel. to VDRL	.02	.01
VDRL rel. to RPR	.30	.01

much lower θ values than the other groups. In general, θ is quite high and has a larger range for the KRP than for the VDRL. The false positive rates are lower for the VDRL than for the KRP. The major point to note here is that θ is very high and ϕ is small.

When the RPR, VDRL and TPCF tests were evaluated against a complete clinical diagnosis, the false negative and false positive rates were quite reasonable with θ less than .16 and ϕ very small ($\leq$.02). The best procedure was the RPR test, although TPCF appeared also to be good. Table 2-m shows that there appears to be an inverse relationship between θ and ϕ within studies, with the exception of Jolly data.

(f) Circumcision Screening

A series of studies has been done to determine the usefulness of yes-no questions to determine the circumcision status of males. A physician examination is considered the criterion for evaluation of the questioning. When wives were questioned with regard to the status of their husbands, Aitken-Swan and Baird (3) found the false negative rate θ = .32 and the false positive rate ϕ = .00; in a similar situation Stern and Lachenbruch (45) found θ = .07, ϕ = .06.

When men were asked about their own status, the results reflect considerable variation from study to study and for Jews versus non-Jews. The Aitken-Swan and Baird data (Table 2-n) showed that false negative rates, based

Table 2-n

Husbands' Statements about Circumcision Status

Study	θ	ϕ
Aitken-Swan and Baird (3) (excluding uncertain)	.17	.01
Partial or complete	.28	.07
Lilienfeld and Graham (34)	.56	.18
Stern and Lachenbruch (46)	.05	.09
Dunn and Buell (19)		
Complete circumcision		
Non-Jews		
Under 40	.57	.08
40-59	.74	.06
60 or over	.73	.10
Jews		
L.A.C. General Hospital	.00	1.00
Cedars of Lebanon	.00	.00
Partial or complete		
Non-Jews		
Under 40	.52	.07
40-59	.48	.06
60 or over	.67	.07

on males' statements are lower than the rates based on wives' statements. In contrast, the Stern and Lachenbruch findings are similar for both husbands and wives when Jews and non-Jews are considered separately (Dunn & Buell [19]), we notice that the rates are different. This is because Jews are generally circumcised and are aware of it. Among the non-Jews, the false negative rate is rather high, and the

false positive rate is fairly low in all age groups.

2.5 Discussion

This study of the values of the false negative and false positive rates associated with medical screening tests was done in order to determine the levels of these parameters in practice and the interrelationships, if any, actually observed between the two error rates. We have reported here a few representative illustrations of screening situations found in the literature. Table 2-o lists the diseases which were investigated and the types of screening measurements available for them.

There are a fair number of situations in disease screening with an inverse relationship between θ and ϕ, although both rates range from 0 to 1. There are also some situations in which there is an inverse relationship with both θ and ϕ falling into narrower ranges. These include blood pressure levels to screen for heart disease, repeated readings of x-rays to screen for tuberculosis, combinations of dichotomous tests to screen for heart disease, combinations of dichotomous questions to screen for rheumatism.

There is a lack of any consistent relationship between θ and ϕ as we move from disease to disease. Also many of the assumptions in the literature are contradicted. The assumption of equal error rates in the two groups for

Table 2-o

False Positive Rates and False Negative Rates for Different Diseases

Disease	Measurement	Method of Screening	Ranges	Relationship
Diabetes	Continuous	Screening curve	.00-.67	Inverse
Hypertension	Continuous	Screening curve	Not small .05-.75	Inverse $\phi > \theta$
Hypertension-Heart Disease	Continuous	Combining Dichot. Contin. tests	Not small,wide	$\theta > \phi$
Cervical Cancer	Continuous	Screening curve	ϕ and θ varies widely on sub-groups	No trend $\theta > \phi$
Rheumatism	Dichotomous	Combo dichot. questions	Moderate	No trend $\phi > \theta$
Syphillis	Dichotomous	Single dichot. tests	Vary both widely for different subgroups	Inverse $\phi > \theta$
Circumcision	Dichotomous	Single dichot. questions	ϕ stable, small vary for differ-ent groups	Inverse

the two sample problems frequently does not hold. For example, the rates with age and religion for circumcision. Syphilis screening tests differ with sex and race; rheumatism screening by symptoms differs with sex, tine tests for tuberculin sensitivity reflect sex differences, etc.

The assumption that both θ and ϕ tend to be small is violated almost everywhere. In fact, the rates cover the entire range possible. This is a reflection of the lack of precision in the screening tests used. The levels of θ and ϕ found in a given study are a reflection of the method used for screening, how good the procedure is, what the goals of the study are, the nature of the population under study and a reflection of the current state of knowledge about the disease. This last point is particularly important in that as knowledge increases one learns which factors are best included in a screening test, and often chooses those factors that may have been unknown at an earlier point of time.

We have also discussed the problems involved in obtaining the values of both false negative and false positive rates in the same study using the same criteria on the same group of subjects. This is one of the reasons for the study of robustness of the usual procedure for the 2 x 2 table when misclassification is present. The precise effects of θ and ϕ on the significance level and power of statistical procedures employed will be studied in

Chapter 3 in order to quantify the effects of using the screening tests in epidemiologic research.

CHAPTER III

THE EFFECT OF MISCLASSIFICATION ON THE EXACT SIGNIFICANCE LEVEL AND POWER IN THE 2 x 2 TABLE

3.1 Introduction

In medical and epidemiologic research, it is often required to compare two groups on the presence or absence of a disease or factor which is subject to misclassification. In these cases, if the presence of errors is suspected, but the levels of the error rates are unknown, one might wish to determine the effect of these errors on the properties of the exact test. Investigators (Cochran [14]) have examined this problem under restrictive assumptions for large samples; the problem has not been examined for the exact tests in the 2 x 2 table. In the fourfold table, a significance test is generally performed on the difference between the two proportions or on the relative odds. Even slight misclassification can cause severe bias in these measures (16), where false positive rates play a more important role than false negative rates for $\pi < .5$ and the situation is reversed for $\pi > .5$.

The problem has been examined for the fourfold table by several investigators. Bross (6) considers the case of the fourfold table with misclassification in each of two

independent binomial samples with underlying prevalences π_1 and π_2. If θ_i and ϕ_i are false negative and false positive rates in ith sample, $E\left(\frac{x_i}{n_i}\right) = \pi_i (1 - \theta_i) + (1 - \pi_i)\phi_i$. Under the assumption that $\theta_1 = \theta_2 = \theta$ and $\phi_1 = \phi_2 = \phi$, the difference between the two proportions is always damped by the presence of error measurements, i.e., $E\left(\frac{x_1}{n_1} - \frac{x_2}{n_2}\right) =$ $(\pi_1 - \pi_2)(1 - \theta - \phi)$ and the difference is unbiased under $\pi_1 = \pi_2$ or $\theta + \phi = 1$. The one-degree of freedom chi-square is then used as a significance test to evaluate this difference.

3.2 Asymptotic Results

Some consideration has been given to an examination of the effects of errors of measurement on the level and power of the X^2 with one degree of freedom. The level of the test is unchanged by the presence of equal misclassification in the two samples, because the difference between the two proportions and the estimated variance of this difference are unbiased under $H_0 : \pi_1 = \pi_2$. However the power is generally reduced. Power curves for this model were presented by Rubin et al. (44) for equal misclassification in two subgroups when interviews were used instead of clinical examinations to classify arthritics. Using calculations similar to Bross, the authors found the ratio

of sample sizes required to regain the power of the no-misclassification cases. Walsh (51) examines the same situation, but considers the alternative hypothesis of interest to be $\pi_1 \neq \pi_2$ rather than Bross' alternative of $P_1 \neq P_2$ where $P_i = \pi_i (1 - \theta) + (1 - \pi_i)\phi$.

This Bross model has also been extended by Gullen et al. (26) and Rogot (43) to permit errors of measurements along both axes of the fourfold table under similar simplifying assumptions with similar results (i.e. equal error rates leave the level of X^2 unchanged, but the power is reduced).

Giesbrecht (1967) provides a general model for measurement errors in both directions of the 2 x 2 table without any restrictive assumptions about the nature of the errors. By using the test statistic

$$\frac{[x_1 (n_2 - x_2) - x_2 (x_1 - n_1)]^2}{n_1 n_2 t (n - t)}$$

where $n = n_1 + n_2$, $x_1 + x_2 = t$,

he evaluates the effects of the error parameters which are varied one at a time. Bross' conclusions are supported for the one direction equal error rates situation. When error rates are present in both directions, there are situations in which the significance level of the X^2 is affected as the result of the errors. Power may increase in some instances. These results are asymptotic and the

distribution of the X^2 is not examined for finite samples.

3.3 Larger Contingency Tables

The properties of the chi-square test for the 2 x c table with misclassification are evaluated by Mote and Anderson (38) who show that the level of $X^2_{(c-1)}$ remains unchanged if false negative and false positive rates are equal in each of the c classes (misclassification only in one direction). On the other hand power is reduced. The change in power occurs whether or not the error rates are known and also for the special case of misclassification only in neighboring classes.

The authors also consider the case of stratified sampling from r independent groups with errors in c categories; in this case the power is sometimes reduced.

Assakul and Proctor (2) extend these results to the case of the r x c contingency table with errors in both directions. For the r x c table with a nonsingular error matrix and independent errors in the two directions, acceptance or rejection of $H_0 : \pi_{ij} = \pi_{i.} \cdot \pi_{.j}$ is equivalent to tests of $H_0 : P_{ij} = P_{i.} P_{.j}$ with the result that the test statistic, $X^2_{(r-1)(c-1)}$, and its significance level remains unchanged. Asymptotic power is reduced. When the errors are not independent, the significance level as well as the power is affected if the presence of error is

ignored. The effects of misclassification become more serious as the significance level (α) decreases.

3.4 Extension of Bross Model for 2 x 2 Table

Once more the question of the validity of assuming a "nice" relationship between false negative and false positive rates in the two samples of the fourfold table arises. Many investigators (14, 17, 18, 21, 41) have cast doubt upon the simplifying assumption that $\theta_1 = \theta_2$, $\phi_1 = \phi_2$. This assumption is unreasonable and invalid in many situations that we have discussed earlier. This consideration leads us to an examination of the two sample problems without this restrictive assumption. Eliminating the assumptions that $\theta_1 = \theta_2$ and $\phi_1 = \phi_2$ in the Bross model, yields:

$$E\left(\frac{x_1}{n_1} - \frac{x_2}{n_2}\right) = \pi_1 (1 - \theta_1 - \phi_1) - \pi_2(1 - \theta_2 - \phi_2) + (\phi_1 - \phi_2). \tag{3.1}$$

Under $H_0 : \pi_1 = \pi_2 = \pi$, (3.1) reduces to:

$$E\left(\frac{x_1}{n_1} - \frac{x_2}{n_2}\right) = \pi [(\theta_2 - \theta_1) + (\phi_2 - \phi_1)] + (\phi_1 - \phi_2) \tag{3.2}$$

Therefore, the expected bias under the null hypothesis is independent of π if

$$[(\theta_2 - \theta_1) + (\phi_2 - \phi_1)] = 0.$$

Under the assumption of two independent binomial samples, the variance of the difference is now:

$$\text{Var}\left(\frac{x_i}{n_1} - \frac{x_2}{n_2}\right) = \frac{\pi_1(1 - \pi_1)}{n_1} k_1' + \frac{\pi_2(1 - \pi_2)}{n_2} k_2'$$

where $k_i' = (1 - \theta_i - \phi_i)^2 + \frac{\theta_i (1 - \theta_i)}{1 - \pi_i} + \frac{\phi_i(1 - \phi_i)}{\pi_i}$

and the MSE $\left(\frac{x_1}{n_1} - \frac{x_2}{n_2}\right)$ becomes

$$\text{Var}\left(\frac{x_1}{n_1} - \frac{x_2}{n_2}\right) + [(1 - \pi_1)\phi_1 - \pi_1\theta_1 + (1 - \pi_2)\phi_2 - \pi_2\theta_2]^2 .$$

Thus the estimated difference between the two proportions is no longer unbiased under $H_0 : \pi_1 = \pi_2$, as it is for the case of equal misclassification.

As a result the test statistic is affected. Therefore, the significance level of $X^2_{(1)}$ will change by the presence of unequal error rates.

3.5 Exact Model for the 2 x 2 Table with Misclassification

All of the results given earlier in the chapter are for misclassification effects on the asymptotic X^2. With the use of computers it is feasible to consider the appropriate exact tests and their robustness against misclassification in the fourfold table. This is particularly relevant in epidemiologic research in which small pilot studies may be performed to determine the feasibility of continuing research on a larger scale.

The fourfold table can arise from different situations. If we take a random sample of size n from a population and classify the elements with respect to a factor present (+) or absent (-) and belonging to group I or group II, then the distribution of the observed table is given by the multinomial distribution.

On the other hand if we observe the same table by taking a random sample of size n_1 from binomial population I and another random sample of size n_2 from a different binomial population II and then classify the elements in each sample with respect to presence or absence of the factor of interest, the distribution of the observed table is given by the product of two independent binomial distributions.

In our study we consider the distribution of 2 x 2 tables obtained by conditioning on the sum $x_1 + x_2 = t$ in the product binomial model. From the following observed table (with underlying fallible proportions P_1 and P_2 with respect to the attribute under study in binomial samples I and II), we have

		Attribute Status		
		+	-	
Sample	I	x_1	n_1-x_1	n_1
	II	x_2	n_2-x_2	n_2
		t	n-t	n

$$\text{Then } P\ (x_1/x_1 + x_2 = t) = \frac{P(x_1, x_1 + x_2 = t)}{P(x_1 + x_2 = t)}$$

$$= \frac{\binom{n_1}{x_1}\binom{n_2}{t-x_1} P_1^{x_1} Q_1^{n_1-x_1} P_2^{t-x_1} Q_2^{n_2-(t-x_1)}}{\sum_{k=0}^{t} \binom{n_1}{k}\binom{n_2}{t-k} P_1^{k} Q_1^{n_1} Q_2^{n_1-k} P_2^{t-k} Q_2^{n_2-(t-k)}}$$

$$= \frac{\binom{n_1}{x_1}\binom{n_2}{t-x_1} Q_1^{n_1} Q_2^{n_2} \left(\frac{P_1Q_2}{P_2Q_1}\right)^{x_1} \left(\frac{P_2}{Q_2}\right)^{t}}{Q_1^{n_1} Q_2^{n_2} \left(\frac{P_2}{Q_2}\right)^{t} \sum_{K=0}^{t} \binom{n_1}{k}\binom{n_2}{t-k}\left(\frac{P_1Q_2}{P_2Q_1}\right)^{k}}$$

$$= \frac{\binom{n_1}{x_1}\binom{n_2}{t-x_1}\left(\frac{P_1Q_2}{P_2Q_1}\right)^{x_1}}{\sum_{k=0}^{t}\binom{n_1}{k}\binom{n_2}{t-k}\left(\frac{P_1Q_2}{P_2Q_1}\right)^{k}} \quad ; \text{ for } x_1 = 0,1,2\ldots t \tag{4.1}$$

where $P_i = \pi_i (1 - \theta_i) + (1 - \pi_i)\ \phi_i$ and $Q_i = 1 - P_i$.

This model is analogous, in epidemiology, to a retrospective study subject to misclassification with respect to the attribute under study, or to a prospective study subject to misclassification with respect to disease. The advantages of considering the conditional model (4.1) are several:

(a) that the test based on this distribution is the uniformly most powerful unbiased level α test for the

exponential family of distributions with the odds ratio as the parameter about which the hypotheses are formulated (Lehmann[33]),

(b) that the probability of observing a zero sum in one of the margins is non-existent unless specified, and

(c) that the probability distribution under the null hypothesis is independent of values of π_1 and π_2 (relative odds is 1) unlike the joint binomial model.

Therefore the test based on this model is performed on the observed table without any assumptions about the underlying probabilities and error rates. By varying prevalence rates and error parameters, we may examine the distortion in the results of the analysis introduced by misclassifications. On the other hand, even under the null hypothesis, the binomial model depends on the value of the probability specified.

We are interested here in both the change in significance level and the change in power of the exact test resulting from the presence of misclassification in the fourfold table. The change in significance level may be found by comparing the level under the null hypothesis $H_0 : \pi_1 = \pi_2$ without errors with the level for this hypothesis when the samples are subject to misclassification. The change in power under the alternative hypothesis may be found by comparing the power without

misclassification for the significance level chosen to the power in the presence of misclassification against the same $H_0 : \pi_1 = \pi_2$.

The levels of misclassification chosen for study here range from small to moderate values of false negative and false positive rates (θ_i, ϕ_i = 0.0, .05, .10, .15, .20, .30, .40 in all combinations). These levels of the error parameters were chosen so as to be reasonable to our previous discussions and to be small enough so that there is some justification for using screening tests with associated errors of these magnitudes. No further assumptions are made about the values of these parameters or the relationship between them. Samples of sizes n_i = 10, 15 and 25 were considered with the latter providing some measure of comparison to asymptotic results. The values of $x_1 + x_2 = t$ were chosen such that approximately $t = n\pi$ under $H_0 : \pi_1 = \pi_2$.

3.6 Effects of Errors on Level of Fisher's Exact Test

We will first consider the effects of misclassification on the nominal significance level of Fisher's exact test in small samples.

(a) One-Sided Tests

What we are concerned with here is the change in significance level for testing $H_0: \pi_1 = \pi_2$ vs. $H_1: \pi_1 > \pi_2$ when

errors of measurement are present. Table 3-a shows the effects of misclassification on level for selected error combinations under $H_0 : \pi_1 = \pi_2 = 0.10$, $n_1 = n_2 = 10$, $x_1 + x_2 = 3$. Without misclassification, the distribution of x_1 is symmetric, and therefore, the upper and lower tails contain the same proportion of distribution. This symmetry is a property of Fisher's exact test under the null hypothesis with $n_1 = n_2$.

The significance level is independent of π when there is no misclassification. In addition, if $\theta_1 = \theta_2$ and $\phi_1 = \phi_2$, the symmetry remains and the nominal level is unchanged.

For a particular critical value, notice what happens to the nominal level that we think we are operating at when misclassification is present. For example in Table 3-a, if $\alpha = .105263$, the critical value of x_1 is 3. If ϕ_2, the false positive rate in the second sample, increases, the true level for this critical value decreases when compared to the fixed nominal level of the test. That is, the null hypothesis is not rejected as often as it should be. If the error set is (0, 0, 0, 0.30), the significance level for a critical value of 3 is .003659 which means H_0 is rejected only .37% of the time when in fact it should be rejected 10.5% of the time.

On the other hand, as ϕ_1 increases, all other parameters constant, the actual level increases when

Table 3-a

Effects of Misclassification on Significance Level of Fisher's Exact Test-Selected Error Rates

$n_1 = n_2 = 10 \quad x_1 + x_2 = 3 \quad H_0: \pi_1 = \pi_2 = .10$

Exact probabilities of observing x_1 or greater values of x_1

				X^2=3.529	X^2=.392	X^2=.392	X^2=3.529
θ_1	ϕ_1	θ_2	ϕ_2	x_1=0	x_1=1	x_1=2	x_1=3
00	00	00	00	1.000000	.894737	.500000	.105263
00	00	00	.10	1.000000	.731472	.254483	.028540
00	00	00	.30	1.000000	.459592	.076194	.003659
00	00	.05	.10	1.000000	.740352	.263746	.030452
00	00	.05	.30	1.000000	.465878	.078700	.003858
00	00	.05	00	1.000000	.903443	.520181	.114506
00	00	.30	00	1.000000	.944210	.635371	.179467
00	00	.10	.05	1.000000	.831402	.381289	.060837
00	00	.30	.10	1.000000	.785501	.316286	.042577
00	00	.40	.05	1.000000	.885907	.480696	.096924
00	00	.05	.40	1.000000	.362190	.043869	.001509
00	00	.40	.40	1.000000	.400227	.055120	.002167
00	.05	00	.05	1.000000	.894737	.500000	.105263
00	.05	00	00	1.000000	.947412	.646412	.186992
00	.05	.30	00	1.000000	.974953	.763324	.286495
00	.10	00	.30	1.000000	.683510	.209488	.020164
00	.10	.05	.40	1.000000	.578385	.134163	.009351
00	.10	.05	.30	1.000000	.689708	.214864	.021087
00	.40	00	.40	1.000000	.894737	.500000	.105263
.05	.05	.05	.05	1.000000	.894737	.500000	.105263
.05	.30	.05	.30	1.000000	.894737	.500000	.105263
.05	00	.30	00	1.000000	.983409	.616297	.167069
.10	.05	.10	.05	1.000000	.894737	.500000	.105263
.10	.05	.10	.40	1.000000	.461991	.077144	.003734
.10	.40	.30	.05	1.000000	.997635	.941803	.590260
.30	00	.30	00	1.000000	.894737	.500000	.105263
.30	00	.30	.05	1.000000	.783771	.314085	.042024
.30	.05	.30	.30	1.000000	.545571	.115720	.007293
.40	.10	.40	.05	1.000000	.946031	.641603	.183692
.40	.05	.40	.10	1.000000	.816318	.358397	.043969
.40	.20	.40	.05	1.000000	.982689	.807953	.337610
.40	.30	.40	.30	1.000000	.894737	.500000	.105263

compared with the nominal chosen level. If we consider the error set (0, .05, 0 0), the significance level for critical value 3 is .186992 which is an increase over the nominal .105263. In this case the test is actually rejecting H_0 18.6% of the time when it should be rejecting only 10.5% of the time.

The false negative rates θ_1 and θ_2 have less effect than the false positive rates with θ_2 causing increases in the nominal level (e.g. for error set 0, 0, .05, 0 the significance level for $x_1 = 3$ is .114506) and increases in θ_1 causes decrease in significance level when compared to the nominal level (e.g. for the error set .05, 0, 0, 0 significance level for $x_1 = 3$ is .096557).

That the major effect on the nominal significance level for the test of $H_0 : \pi_1 = \pi_2$ vs $H_1 : \pi_1 > \pi_2$ is caused by the difference between two false positive rates with $\phi_1 > \phi_2$ causing increases in level and $\phi_1 < \phi_2$ causing decreases may be verified from able 3-b. Table 3-b shows the one-sided significance levels for $n_i = 10$, $t = 3$ $\pi = .10$ for Fisher's exact test averaged over 16 combinations of four false positive rates for each set of false negative rates and averaged over four false negative rates for each set of false positive rates. From this it is clear that the level of the test is quite stable against false negative rates, but quite unstable with respect to false positive rates.

Table 3-b

Effects of Misclassification on One Sided Significance Level of Fisher's Exact Test

$n_1 = n_2 = 10$ $H_0: \pi_1 = \pi_2 = .10$ $x_1 + x_2 = 3$

A: Averaged Over 16 Combinations of False Positive Rates

θ_1 \ θ_2	.00	.05	.10	.30
.00	.1556	.1626	.1700	.2074
.05	.1503	.1572	.1645	.1985
.10	.1453	.1519	.1574	.1923
.30	.1255	.1314	.1322	.1679

B: Averaged Over 16 Combinations of False Negative Rates

ϕ_1 \ ϕ_2	.00	.05	.10	.30
.00	.1499	.0506	.0263	.0031
.05	.200	.1051	.0637	.0092
.10	.2875	.1711	.1104	.0933
.30	.5658	.4281	.3581	.1057

Tables 3-c and 3-d give selected error combinations for other values of n_i, t, π. As before, as ϕ_2 increases, the true level for a particular critical value decreases when compared to the fixed nominal level of the test. As ϕ_1 increases (other parameters constant) the actual level increases when compared with the nominal chosen level. The false negative rates θ_1 and θ_2 are less important. An increase in θ_2 causes an increase over the nominal level and an increase in θ_1 causes a slight decrease in the significance level when compared to the nominal level.

We now consider the one tailed test for $H_0 : \pi_1 = \pi_2$ versus the alternative $H_1 : \pi_1 < \pi_2$ using Fisher's exact test. For this test, the change in nominal significance level may be found from the lower tail of the distributions for x_1 in Tables 3-a, 3-c and 3-d. In this case the critical value of x_1 such that the probability of observing x_1 or smaller value of x_1 is α. The effects in significance level with respect to the error parameters are the reverse of those discussed earlier for the alternative hypothesis $H_1 : \pi_1 > \pi_2$. That is, ϕ_2 increases, cause increases over the nominal level and ϕ_1 increases cause decreases over nominal level. Increases in θ_2 causing decreases over the nominal level and increases in θ_1 causes increases in significance level compared to nominal level.

Table 3-c

Effects of Misclassification on Significance Level of Fisher's Exact Test as $x_1 + x_2 = t$ Changes

$n_1 = n_2 = 10$ $\pi = .25$ (Probability of observing the value of x_1 or a greater value)

						$x_1+x_2=5$				$x_1+x_2=7$						
				$X^2=6.67$	2.40	2.67	2.40	6.67	16.364	9.90	5.050	1.818	.202	.202	1.818	5.050
θ_1	ϕ_1	θ_2	ϕ_2	$x_1=0$	$x_1=1$	$x_1=2$	$x_1=4$	$x_1=5$	$x_1=0$	$x_1=1$	$x_1=2$	$x_1=3$	$x_1=4$	$x_1=5$	$x_1=6$	$x_1=7$
0	0	0	0	1.00000	.983746	.84829	.15170	.01625	1.00000	.99964	.99011	.91509	.67504	.32496	.08490	.00988
0	0	0	.05	1.00000	.97426	.79708	.10978	.00990	1.00000	.99925	.98273	.87531	.59133	.24923	.05543	.00541
0	0	0	.15	1.00000	.94629	.68391	.05657	.00372	1.00000	.997472	.95795	.77552	.43332	.14077	.02318	.00164
0	0	0	.40	1.00000	.813045	.38815	.00890	.00028	1.00000	.97729	.81219	.45760	.14813	.02505	.00203	.00007
0	0	.05	0	1.00000	.98631	.86424	.168828	.01921	1.00000	.99973	.99197	.92666	.70300	.35393	.09777	.01211
0	0	.15	0	1.00000	.99062	.89412	.208863	.02696	1.00000	.99985	.99494	.94714	.75821	.41786	.12955	.01826
0	0	.05	.30	1.00000	.88746	.52418	.02210	.00098	1.00000	.99095	.89754	.612578	.25932	.05975	.00673	.00032
0	.10	.05	.05	1.00000	.99129	.89920	.21705	.02869	1.00000	.99986	.99537	.95045	.76801	.43027	.13630	.01969
0	.10	0	.10	1.00000	.98374	.84829	.151170	.01625	1.00000	.99964	.99011	.91509	.67504	.32496	.08990	.00988
.05	0	.30	0	1.00000	.99421	.92342	.26362	.03943	1.00000	.99993	.99718	.96548	.81653	.49755	.17637	.02599
0	.05	.10	.30	1.00000	.92545	.61908	.03892	.00217	1.00000	.99558	.93755	.71218	.35650	.10069	.01418	.00085
.05	.30	.05	.30	1.00000	.983746	.84829	.15170	.01625	1.00000	.99964	.99011	.91509	.67504	.32496	.08490	.00988
.10	.30	.05	.30	1.00000	.981587	.835687	.13974	.01431	1.00000	.99956	.98850	.905662	.65360	.30408	.07618	.00847
.10	.40	.10	.40	1.00000	.983746	.84829	.15170	.0162	1.00000	.9984	.99011	.91509	.67504	.32496	.08490	.00988
.30	.05	.30	.05	1.00000	.98374	.84829	.15170	.01625	1.00000	.99964	.99011	.91509	.67504	.32496	.08490	.00988
.30	.05	.40	.05	1.00000	.98913	.88330	.19295	.0237	1.00000	.99981	.99394	.93392	.73776	.39305	.11665	.01565
.40	.40	.40	.40	1.00000	.98374	.84829	.15170	.0162	1.00000	.99964	.99011	.91509	.67504	.32496	.08090	.00988

Table 3-d

Effects of Misclassification on Significance Level of Fisher's Exact Test as ample Size Changes

				$n_1=n_2=15$		$x_1+x_2=3$		$\pi_1=\pi_2=.10$		$n_1=n_2=25$		$x_1+x_2=5$	
θ_1	ϕ_1	θ_2	ϕ_2	$x_1=0$	$x_1=1$	$x_1=2$	$x_1=3$	$x_1=0$	$x_1=1$	$x_1=2$	$x_1=3$	$x_1=4$	$x_1=5$
0	0	0	0	1.00000	.88793	.50000	.11207	1.00000	.97492	.82566	.5000	.17438	.02507
0	0	0	.05	1.00000	.80139	.35096	.05585	1.00000	.93483	.68067	.31737	.07934	.00786
0	0	0	.15	1.00000	.63666	.18174	.01719	1.00000	.82062	.43441	.12962	.01938	.00111
0	0	0	.40	1.00000	.33889	.04040	.00146	1.00000	.50688	.12403	.01508	.00087	.00001
0	0	.05	0	1.00000	.89720	.52056	.12190	1.00000	.97828	.84147	.52552	.19109	.02887
0	0	.15	0	1.00000	.91528	.56451	.14497	1.00000	.98426	.87223	.57988	.23021	.03851
0	0	.05	.30	1.00000	.44619	.07563	.00400	1.00000	.63480	.21460	.03739	.00313	.00009
0	.10	.05	.05	1.00000	.93814	.62962	.18505	1.00000	.99067	.91072	.65934	.29712	.05791
0	.10	0	.10	1.00000	.88793	.50000	.11207	1.00000	.97492	.82566	.50000	.17434	.02507
0	0	.30	.10	1.00000	.77250	.31312	.04516	1.00000	.91820	.63420	.27274	.06189	.00552
0	.05	.10	.30	1.00000	.58083	.14338	.01148	1.00000	.77183	.36235	.09299	.01179	.00056
.05	.30	.05	.30	1.00000	.88793	.50000	.11207	1.00000	.97492	.82566	.50000	.17434	.02507
.10	.30	.05	.30	1.00000	.88425	.49216	.10847	1.00000	.97353	.81939	.49027	.16821	.02375
.10	.30	.10	.30	1.00000	.88793	.50000	.11207	1.00000	.97492	.82566	.50000	.17434	.02507
.30	.05	.30	.05	1.00000	.88793	.50000	.11207	1.00000	.97492	.82566	.50000	.07934	.02507
.30	.05	.40	.05	1.00000	.90421	.53695	.13016	1.00000	.98069	.85341	.54582	.20513	.03218
.40	.40	.40	.40	1.00000	.88793	.50000	.11207	1.00000	.97492	.82566	.50000	.17434	.02507

In all cases discussed above it is clear that the null hypothesis is being rejected the proper number of times only when the false negative rates (θ_1, θ_2) are equal and the false positive rates (ϕ_1, ϕ_2) are equal (Bross result).

When the error rates are not equal we observe in Table 3-c that regardless of what the fixed margin, t, is the effects on the level of the test are similar. Also, in Table 3-d we can see that the effects of misclassification on the distribution of x_1 are similar as n_i increases. The distribution of $n_i = 15$, $\pi = .10$ (Table 3-c) may be compared with that for $n_i = 10$, $\pi = .10$ (Table 3-a) for $t = 3$. If we take critical value $x_1 = 3$ in Table 3-a, $P = .10526$ compared to .1120 in Table 3-d without misclassification in both cases. The differences due to sample size are very slight when misclassification is present.

If we have a desired significance level, say $\alpha = \alpha_o$, in order to obtain this exact level α_0, it is necessary to randomize at the critical value, c. That is $\alpha_0 =$ P (reject $H_0/x_1 > c$) $\cdot$ P $(x_1 > c)$ + P (reject $H_0/x_1 = c$) P $(x_1 = c)$ or $\alpha_0 = P$ $(x_1 > c) + \gamma P(x_1 = c)$, where γ = prob (randomize).

From Table 3-a (page 54), for $x_1 = 3$, if we perform the randomized test, we reject the null hypothesis 95% of the time at this value to achieve a nominal level $\alpha = .10$.

For $n_i = 15$ and $t = 3$, $x_1 = 3$, for the randomization test we reject the null hypothesis 89.22% of the time to achieve a nominal level $\alpha = .10$ (for $n_i = 25$, $t = 5$, $x_1 = 4$ reject 50.19% of the time, for $n_i = 10$, $t = 5$, $x_i = 4$ reject 62% of the time) (see Table 3-e).

(b) Two-Sided Tests

We now consider the two-sided Fisher's exact test for $H_0 : \pi_1 = \pi_2$ vs. $H_1 : \pi_1 \neq \pi_2$. The exact level may be found by combining the upper and lower tails of the distribution of x_1 to obtain a fixed level α. Without misclassification, because of the symmetry of the distribution of x_1, this is equivalent to doubling the single tail probabilities. However, if the error rates in two samples are unequal, the distribution of x_1 is no longer symmetric.

If we turn to Table 3-f which shows the level of the two-sided tests for selected error combinations, we can see that the actual level generally exceeds the nominal level with the greatest increase occurring when the difference between the false positive rates is large as well as when the difference between the false negative rates is large. One generally rejects the null hypothesis too often when this happens.

Table 3-e

Effects of Large Misclassification Rates on One Sided Significance Level of Fisher's Exact Test

$H_0: \pi_1 = \pi_2$ vs $H_1: \pi_1 > \pi_2$

				$\pi_1=\pi_2=.10$			$\pi_1=\pi_2=.25$
				$n_i=10$	$n_i=15$	$n_i=25$	$n_i=10$
				Critical $t=3$	$t=3$	$t=5$	$t=5$
θ_1	ϕ_1	θ_2	ϕ_2	value x_1 3	3	4	4
0	0	0	0	.1052	.1120	.1743	.1517
0	0	0	.05	.0526	.0559	.0793	.1098
0	0	0	.10	.0285	.0301	.0383	.0790
0	0	0	.30	.0037	.0038	.0029	.0195
0	0	0	.40	.0014	.0014	.0008	.0089
0	0	.10	0	.1247	.1383	.2096	.1878
0	0	.30	0	.1794	.1907	.3063	.2868
0	0	.40	0	.2182	.2314	.3720	.3538
0	.30	.40	.30	.1356	.1443	.2291	.2726
.10	.30	.40	.30	.1277	.1359	.2149	.2393
.40	.40	.40	.40	.1052	.1120	.1743	.1517
Randomized test $\alpha = .10$				Reject 95% of time	Reject 89% of time	Reject 51% of time	Reject 62% of time
0	0	0	0	.1	.1	.1	.1
0	0	0	.05	.05	.0498	.0437	.0717
0	0	0	.10	.0271	.0269	.0206	.0420
0	0	0	.30	.0035	.0034	.0015	.0121

Table 3-f

Level and Power of Two Sided Fisher's Exact Test with Misclassification

				$H_0: \pi_1=\pi_2=.25$ $n_1=n_2=10$ $x_1+x_2=5$		$\pi_1=\pi_2=.10$ $n_1=n_2=15$ $x_1+x_2=3$		$H_0: \pi_1=\pi_2=.10$ $n_1=n_2=25$ $x_1+x_2=5$	
				$H_1: \pi_1=.25$	$H_1: \pi_1=.50$	$H_1: \pi_1=.10$	$H_1: \pi_1=.10$	$H_1: \pi_1=.10$	$H_1: \pi_1=.10$
θ_1	ϕ_1	θ_2	ϕ_2	$\pi_2=.50$	$\pi_2=.25$	$\pi_{2=}.25$	$\pi_2=.50$	$\pi_2=.25$	$\pi_2=.50$
.00	.00	.00	.00	.5503	.550	.4027	.7009	.2014	.5434
.00	.00	.00	.30	.7546	.306	.6773	.8201	.5129	.7115
.00	.00	.05	.30	.7227	.309	.6650	.8025	.4971	.6866
.00	.00	.10	.05	.5152	.627	.4208	.6773	.2190	.5122
.00	.00	.30	.10	.4156	.550	.4027	.5989	.2014	.4136
.00	.05	.30	.00	.3244	.271	.2328	.4184	.0565	.2167
.00	.10	.05	.30	.6007	.335	.4593	.6434	.2578	.4687
.05	.30	.10	.05	.3041	.723	.2628	.2587	.0795	.0765
.10	.30	.30	.00	.3468	.817	.3695	.2245	.1703	.0504
.30	.05	.30	.30	.6264	.306	.5555	.6635	.3626	.4943
.30	.30	.30	.30	.5403	.550	.4027	.7009	.2014	.5434

Significance level: $x_1 \leq 1$ or $x_1 \geq 4$

3.7 Power of Fisher's Exact Test with Misclassification

(a) One Sided Power

We now examine the change in power resulting from the presence of errors of measurement. Table 3-f and Table 3-g show the cumulative distribution of x_1 for selected error rates under various alternative hypotheses. For the one-sided test of $H_0 : \pi_1 = \pi_2$ vs. $H_1 : \pi_1 < \pi_2$ the power may be obtained for a given level α from the upper tail of these cumulative distributions. Similarly we may find the power of the one-sided alternative $H_1 : \pi_1 > \pi_2$ from the lower tails of these distribution for the reverse error sets. The effects of misclassification on these one-sided powers for Fisher's exact test are similar to the effects on the corresponding one-sided significance levels discussed earlier.

(b) Two-Sided Power

In order to find the two-sided power for fixed level α, we must consider the alternatives $\pi_1 < \pi_2$ and $\pi_1 > \pi_2$ separately and then combine the appropriate tails of the separate distributions. Table 3-h gives the two-sided power for $H_0 : \pi_1 = \pi_2 = .25$ vs. $H_1 : \pi_1 - \pi_2 = .25$ ($n_i = 10$, $t = 5$) found from the lower tails of $\pi_1 = .25$, $\pi_2 = .50$ and the upper tails of $\pi_1 = .50$ and $\pi_2 = .25$. The power of the two-sided test may also be found by

Table 3-g

Exact Probabilities of Observing the Value of the Test Statistic or More Extreme Value - Fisher's Exact Test for Selected Error Rates under Alternative Hypotheses: $\pi_1 < \pi_2$

A. $H_0: \pi_1 = \pi_2 = .10$ $H_1: \pi_1 = .10, \pi_2 = .25$

$n_1 = n_2 = 10$ $t = 3$

θ_1	ϕ_1	θ_2	ϕ_2	$X^2 = 3.529$ $x_1 = 0$	.392 1	.392 2	3.529 3
.00	.00	.00	.00	1.000	.6301	.1678	.0137
.00	.00	.00	.30	1.000	.3416	.0384	.0012
.00	.00	.05	.30	1.000	.3544	.0418	.0014
.00	.00	.10	.05	1.000	.6104	.1542	.0118
.00	.00	.30	.10	1.000	.6301	.1678	.0137
.00	.05	.30	.00	1.000	.8552	.4211	.0740
.00	.10	.05	.30	1.000	.5692	.1288	.0087
.05	.30	.10	.05	1.000	.9525	.6648	.2001
.10	.30	.30	.00	1.000	.9821	.8044	.3332
.30	.05	.30	.30	1.000	.4683	.0797	.0039

(Cont'd. on next page.)

Table 3-g (Cont'd.)

B. $H_0: \pi_1=\pi_2=.10$ $H_1: \pi_1=.10,$ $\pi_2=.25$

$n_1=n_2=15$ $t=3$

θ_1	ϕ_1	θ_2	ϕ_2	$X^2=3.333$ $x_1=0$	.370 1	.370 2	3.333 3
.00	.00	.00	.00	1.000	.6117	.1637	.0144
.00	.00	.00	.30	1.000	.3239	.0365	.0012
.00	.00	.05	.30	1.000	.3364	.0397	.0014
.00	.00	.10	.05	1.000	.5916	.1502	.0124
.00	.00	.30	.10	1.000	.6117	.1637	.0144
.00	.05	.30	.00	1.000	.8459	.4196	.0787
.00	.10	.05	.30	1.000	.5498	.1249	.0091
.05	.30	.10	.05	1.000	.9496	.6677	.2124
.10	.30	.30	.00	1.000	.9812	.8085	.3507
.30	.05	.30	.30	1.000	.4486	.0766	.0041

(Cont'd. on next page.)

Table 3-g (Cont'd.)

C. $H_0: \pi_1 = \pi_2 = .10$ $H_1: \pi_1 = 10,$ $\pi_2 = .50$

$n_1 = n_2 = 15$ $t = 3$

θ_1	ϕ_1	θ_2	ϕ_2	$X^2 = 3.333$ $x_1 = 0$	.370 1	.370 2	3.333 3
.00	.00	.00	.00	1.000	.3001	.0309	.0010
.00	.00	.00	.30	1.000	.1801	.0103	.0002
.00	.00	.05	.30	1.000	.1977	.0126	.0002
.00	.00	.10	.05	1.000	.3239	.0365	.0012
.00	.00	.30	.10	1.000	.4039	.0601	.0028
.00	.05	.30	.00	1.000	.5943	.1520	.0127
.00	.10	.05	.30	1.000	.3584	.0458	.0018
.05	.30	.10	.05	1.000	.7939	.3407	.0528
.10	.30	.30	.00	1.000	.8951	.5158	.1196
.30	.05	.30	.30	1.000	.3379	.0401	.0014

combining the lower tail of the distribution for $\pi_1 < \pi_2$ for a given set of $(\theta_1, \phi_1, \theta_2, \phi_2)$ with the lower tail for $\pi_1 > \pi_2$ with reverse error set $(\theta_2, \phi_2, \theta_1, \phi_1)$. For example, if $\theta_1 = .05$, $\phi_1 = .30$, $\theta_2 = .10$ and $\phi_2 = .05$, for the alternative $\pi_1 < \pi_2$, the lower tail probability is .1643 from Table 3-g. For the alternative $\pi_1 > \pi_2$, we take the lower tail probability for the error set $\theta_1 = .10$, $\phi_1 = .05$, $\theta_2 = .05$, $\phi_2 = .30$, which is .6978. Adding the two probabilities, we obtain .8621, which is then randomized to be compatible with $\alpha = .10$ in the (0, 0, 0, 0) case: that is, power = .287, which is the same power as given in Table 3-i. Notice that the power of this without misclassification is 35%.

Table 3-h

Exact Probabilities of Observing That Value or More Extreme Value of Test Statistic for Selected Error Rates Under Alternative Hypotheses

A. $H_0: \pi_1=\pi_2=.25 \quad n_1=n_2=10 \quad x_1+x_2=5$

					$H_1:\pi_1=.25$ $\pi_2=.50$						$H_1:\pi_1=.50$ $\pi_2=.25$					
				X^2	6.667	2.400	.267	.267	2.400	6.667	6.667	2.400	.267	.267	2.400	6.667
θ_1	ϕ_1	θ_2	ϕ_2	x_1	0	1	2	3	4	5	0	1	2	3	4	5
0	0	0	0	1.000	.8583	.4649	.1276	.0152	.0006		1.000	.999	.985	.872	.535	.1 2
0	0	0	.30	1.000	.6987	.2480	.0400	.0027	.0001		1.000	.987	.871	.542	.177	.021
0	0	.05	.30	1.000	.7304	.2810	.0499	.0037	.0001		1.000	.989	.882	.563	.191	.023
0	0	.10	.05	1.000	.8782	.5043	.1500	.0195	.0008		1.000	.999	.982	.858	.509	.128
0	0	.30	.10	1.000	.7600	.3155	.0614	.0050	.0001		1.000	.999	.985	.872	.535	.142
0	.05	.30	0	1.000	.9679	.7675	.3813	.0919	.0076		1.000	1.000	.997	.954	.732	.282
0	.10	.05	.30	1.000	.8268	.4097	.0999	.0104	.0003		1.000	.994	.918	.643	.253	.037
.05	.30	.10	.05	1.000	.9816	.8357	.4790	.1398	.0143		1.000	1.000	.995	.943	.698	.251
.10	.30	.30	0	1.000	.9948	.9284	.6691	.2752	.0424		1.000	1.000	.999	.976	.816	.374
.30	.05	.30	.30	1.000	.8104	.3842	.0884	.0086	.0003		1.000	.979	.821	.456	.127	.012

B. $H_0:\pi_1=\pi_2=.10 \quad n_1=n_2=25 \quad x_1+x_2=5$

					$H_1:\pi_1=.10$ $\pi_2=.25$						$H_0:\pi_1=.10$ $\pi_2=.50$					
0	0	0	0	1.000	.7994	.4015	.1120	.0156	.0008		1.000	.4566	.0973	.0102	.0005	.0000
0	0	0	.30	1.000	.4878	.1133	.0130	.0007	.0000		1.000	.2885	.0351	.0020	.0000	.0000
0	0	.05	.30	1.000	.5073	.1222	.0147	.0008	.0000		1.000	.3144	.0423	.0027	.0001	.0000
0	0	.10	.05	1.000	.7816	.3757	.0993	.0130	.0006		1.000	.4878	.1133	.0130	.0007	.0000
0	0	.30	.10	1.000	.7994	.4015	.1120	.0156	.0008		1.000	.5864	.1760	.0268	.0019	.0000
0	.05	.30	0	1.000	.9574	.7545	.4006	.1176	.0139		1.000	.7840	.3792	.1009	.0133	.0007
0	.10	.05	.30	1.000	.7426	.3250	.0766	.0088	.0004		1.000	.5313	.1387	.0181	.0011	.0000
.05	.30	.10	.05	1.000	.9934	.9295	.7047	.3416	.0729		1.000	.9307	.6686	.3052	.0744	.0072
.10	.30	.30	0	1.000	.9987	.9784	.8606	.5485	.1690		1.000	.9775	.8379	.5196	.1871	.0279
.30	.05	.30	.30	1.000	.6375	.2169	.0381	.0032	.0001		1.000	.5057	.1233	.0149	.0009	.0000

Table 3-i

Two-Sided Power of Fisher's Exact Test with Misclassification Randomized to $H_0: \alpha = .10$

				$H_0: \pi_1 = \pi_2 = .25$			$H_1: \pi_1 - \pi_2 = .25$
	$n_1 = n_2 = 10$ t=5			Level $x_1 \leq 1$ or $x_1 \geq 4$	Lower[1] tail	Upper[2] tail	Power two-sided for $\alpha = .10$
θ_1	ϕ_1	θ_2	ϕ_2				
.00	.00	.00	.00	.3034	.535	.535	.353
.00	.00	.05	.30	.4979	.719	.191	.300
.00	.00	.10	.05	.1821	.496	.509	.330
.00	.00	.30	.10	.3034	.376	.535	.301
.00	.05	.30	.00	.4013	.233	.732	.318
.00	.10	.05	.30	.3845	.500	.253	.248
.05	.30	.10	.05	.4767	.164	.698	.287
.10	.30	.30	.00	.6377	.072	.816	.293
.30	.05	.30	.30	.4830	.616	.127	.278

[1] Lower tail from $\pi_1 = .25, \pi_2 = .50$

[2] Upper tail from $\pi_1 = .50, \pi_2 = .25$

CHAPTER IV

ESTIMATION OF BINOMIAL PROPORTION, FALSE NEGATIVE AND FALSE POSITIVE RATES USING THE RANDOMIZED RESPONSE TECHNIQUE

4.1 Introduction

We have shown that the presence of misclassification can have substantial effects on the estimation of parameters as well as on hypothesis testing. This effect on the sample estimate of the binomial proportion was first studied by Bross (6) under the assumption of a probabilistic misclassification model. The statistical model discussed by Bross (6), Mote and Anderson (38) and Assakul and Proctor (2) for handling misclassification assumes that the error rates are known. But this is not generally the case as we have seen earlier. The previous discussions on the effects of misclassification lead us to the awareness of the importance of estimating the values of θ and ϕ in a particular study. Since the error rates vary in nature from one study to the other, and from one group of individuals to another group with respect to age, sex, religion, etc., it is necessary to estimate these error rates in each separate independent study before making valid conclusions on the study. In order to adjust the bias due to misclassification,

the amount of misclassification that is present in the data must be available. One method of obtaining information on the extent of misclassification is to compare the results obtained by two or more measuring devices on the same group of sampling units. Suppose an investigator has a true and fallible measuring device to classify sampling units into one of two categories, denoted as "0" and "1" respectively. The fallible device is a relatively inexpensive procedure which tends to misclassify units, whereas the true device is a more expensive device which is subject to no misclassification. Using only the fallible classifier on all N units in the sample results in a biased estimate of π. A better estimate of π could be obtained if the true classifier were used; however, the expense of using the true classifier on all N units may be too high. As a compromise between these two extremes, Tenenbein (1970) introduced a double sampling scheme for estimating from binomial data with misclassification.

The method introduced by Tenenbein requires a true device subject to no errors of measurement. But experimental situations arise where there is no exact device for measuring a true response. In some practical situations only the respondent knows the true response. If the response has a stigmatizing nature, estimates based on a direct questionnaire may be biased. In this special

situation we propose the use of a randomized response technique to reduce the bias in estimating the population proportion.

The randomized response technique was originally proposed by S. L. Warner (1965). He developed a design for estimating the proportion, π_A, of individuals with a sensitive attribute, A, without requiring the individual respondent to report his actual classification, whether it be A or not A, to the interviewer. A simple random sample of n individuals is drawn with replacement from the population and provisions made for each person to be interviewed. Before the interviews, each interviewer is furnished with an identical spinner with a face marked so that the spinner points to the letter "A" with probability δ and to the letter "B" with probability $1-\delta$. Then in each interview the interviewee is asked to spin the spinner unobserved by the interviewer and report only whether or not the spinner points to the letter representing the group to which the interviewee belongs. That is, the interviewee is required only to say "yes" or "no" according to whether or not the spinner points to the correct group, without reporting to the interviewer the group to which the spinner points. Under the assumption that these "yes" or "no" reports were made truthfully, we could estimate the misclassification rates by applying the technique to a subsample of n units from a main sample of N units drawn from

the population.

4.2 Model

Our proposed sampling technique is as follows: (a) Take a sample of N individuals from the population and obtain their fallible classification; (b) Take a subsample of n units from the main sample of N units and obtain the responses by the randomized response technique; (c) Combine (a) and (b) to obtain estimates of the true proportion, π, false negative rate, and false positive rate. Under the above scheme, the data on n observations can be summarized as

	Fallible Response "NO" ("0")	Fallible Response "YES" ("1")	
Randomized "NO" ("0")	n_{00}	n_{01}	$n_{0\cdot}$
Response "YES" ("1")	n_{10}	n_{11}	$n_{1\cdot}$
	$n_{\cdot 0}$	$n_{\cdot 1}$	n

where the probabilities for each classification are

	Fallible Response "NO" ("0")	Fallible Response "YES" ("1")	
Randomized "NO" ("0")	P_{00}	P_{01}	$1 - \lambda$
Response "YES" ("1")	P_{10}	P_{11}	λ
	$1-\alpha$	α	1

For the N - n remaining individuals information is

only available from the fallible measurement.

	Fallible Response		
	"NO"	"YES"	Total
Frequency	Y	X	N-n
Probabilities	$1-\alpha$	α	1

A fallible "yes" or "no" response means belonging or not belonging to the group under study. A randomized response "yes" or "no" means whether or not the spinner points to the correct group to which the respondent belongs. Thus, for each sampling unit we define,

$P(S)$ = Probability that the spinner points to $A=\delta$

$P(\bar{S})$ = Probability that the spinner does not point to $A = 1 - \delta$

$P(A)$ = Probability that the individual belongs to Group $A = \pi$

$P(\bar{A})$ = Probability that the individual not belonging to $A = 1 - \pi$.

Let $R = \begin{cases} 1 & \text{if the sample element says "yes" for the randomized response} \\ 0 & \text{if the sample element says "no" for the randomized response} \end{cases}$

and $F = \begin{cases} 1 & \text{if the sample element gives "yes" for the fallible measurement} \\ 0 & \text{if the sample element gives "no" for the fallible measurement.} \end{cases}$

Also let $\theta_{01} = P[F = 0/A]$ and $\theta_{10} = P[F = 1/\bar{A}]$

In our previous terminology, θ_{01} and θ_{10} can be described as false negative and false positive rates respectively. Then,

$$\lambda = P(R=1) = P(S \cap A \text{ or } \bar{S} \cap \bar{A})$$

$$= P[A \cap S] + P[\bar{A} \cap \bar{S}]$$

$$= P(A) \cdot P(S) + P(\bar{A}) \cdot P(\bar{S})$$

$$\lambda = \pi\delta + (1 - \pi)(1 - \delta)$$

and $P[R = 0] = 1 - P[R = 1] = 1 - \lambda$

$$= \pi(1 - \delta) + (1 - \pi)\delta.$$

The cell probabilities can be expressed in terms of the parameters above in the following manner:

$$P_{00} = P[R = 0, F = 0]$$

$$= P(R = 0, F = 0, S, A) + P(R = 0, F = 0, S, \bar{A})$$

$$+ P(R = 0, F = 0, \bar{S}, A) + P(R = 0, F = 0, \bar{S}, \bar{A})$$

$$= P(A_1) + P(B_1) + P(C_1) + P(D_1)$$

Now,

$$P(A_1) = P(R = 0, F = 0, S, A)$$

$$= P(R = 0 \mid S,A)P(F = 0 \mid S,A)P(S)P(A)$$

$$= (0)(\theta_{01})\delta\pi$$

$$= 0$$

$$P(B_1) = P(R = 0, F = 0, S, \bar{A})$$

$$= P(R = 0 \mid S,\bar{A})P(F = 0 \mid S,\bar{A})P(S)P(\bar{A})$$

$$= (1)(1 - \theta_{10})\delta(1 - \pi)$$

$$= \delta(1 - \pi)(1 - \theta_{10})$$

$$P(C_1) = P(R = 0, F = 0, \bar{S}, A)$$

$$= P(R = 0 \mid \bar{S},A)P(F = 0 \mid \bar{S},A)P(\bar{S})P(A)$$

$$= (1)\theta_{01}\ (1 - \delta)\pi$$

$$= \pi(1 - \delta)\theta_{01}$$

$$P(D_1) = P(R = 0,\ F = 0, \bar{S}, \bar{A})$$

$$= P(R = 0/\bar{S}, \bar{A})P(F = 0/\bar{S}, \bar{A})P(\bar{S})P(\bar{A})$$

$$= (0)(1 - \theta_{10})(1 - \delta)(1 - \pi)$$

$$= 0$$

Using above four probabilities, we obtain

$$P_{00} = \delta(1 - \pi)(1 - \theta_{10}) + \pi(1 - \delta)\theta_{01}$$

Similarly,

$$P_{01} = \delta(1 - \pi)\theta_{10} + \pi(1 - \delta)(1 - \theta_{01})$$

$$P_{10} = \theta_{01}\delta\pi + (1 - \theta_{10})(1 - \delta)(1 - \pi)$$

$$P_{11} = (1 - \theta_{01})\delta\pi + (1 - \delta)(1 - \pi)\theta_{10} \qquad (2.1)$$

so n_{00}, n_{01}, n_{10}, n_{11} are quadrinomial with cell probabilities P_{00}, P_{01}, P_{10} and P_{11} and sample size, n, where the P_{ij}'s are given by (2·1) i,j = 0,1. Also, the distribution of X is binomial with probability α and sample size (N - n).

Since the measurements in the second sample of (N - n) units are independent of the measurements in the first sample of n units, we can write the joint distribution of n_{00}, n_{01}, n_{10}, n_{11}, X, and Y as

$$L = P(n_{00}, n_{01}, n_{10}, n_{11}, X, Y/N, n)$$

$$= C\ P_{00}^{n_{00}}\ P_{01}^{n_{01}}\ P_{10}^{n_{10}}\ P_{11}^{n_{11}}\alpha^{X}\ (1 - \alpha)^{Y} \qquad (2.2)$$

4.3 M.L.E. of π, θ_{01}, θ_{10}

Let us define,

γ_1 = Conditional probability that randomized response gives "yes" when fallible response gives "yes"

$= P[R = 1|F = 1] = \frac{P_{11}}{\alpha}$

γ_2 = Conditional probability that randomized response gives "yes" when fallible response gives "no"

$= P[R = 1|F = 0] = \frac{P_{10}}{1-\alpha}$

so $(1 - \gamma_1) = P[R = 0|F = 1] = \frac{P_{01}}{\alpha}$

and $(1 - \gamma_2) = P[R = 0|F = 0] = \frac{P_{00}}{1-\alpha}$

Using the definition (2.1) becomes

$$P_{00} = (1 - \alpha)(1 - \gamma_2) \qquad (3.1)$$

$$P_{01} = (1 - \gamma_1)\alpha$$

$$P_{10} = (1 - \alpha)\gamma_2$$

$$P_{11} = \gamma_1\alpha$$

so the likelihood function in (2.2) can be written in terms of γ_1, γ_2, α as

$$L = C \; \gamma_1^{n_{11}}(1-\gamma_1)^{n_{01}} \gamma_2^{n_{10}} (1-\gamma_2)^{n_{00}} \alpha^{n_{01}+n_{11}+X} (1-\alpha)^{n_{00}+n_{10}+Y}$$

The maximum likelihood estimators (MLE) of α, γ_1, γ_2 are given by $\hat{\alpha}$, $\hat{\gamma}_1$, $\hat{\gamma}_2$ respectively, where

$$\hat{\gamma}_1 = \frac{n_{11}}{n_{\cdot 1}}, \quad \hat{\alpha} = \frac{X + n_{\cdot 1}}{N} \tag{3.2}$$

$$\hat{\gamma}_2 = \frac{n_{10}}{n_{\cdot 0}}$$

Note that $\lambda = \gamma_2 (1 - \alpha) + \gamma_1 \alpha$; so we can obtain $\hat{\lambda}$ as

$$\hat{\lambda} = \hat{\gamma}_2(1 - \hat{\alpha}) + \hat{\gamma}_1 \hat{\alpha}$$

$$= \frac{n_{10}}{n_{\cdot 0}} \left(1 - \frac{X + n_{\cdot 1}}{N}\right) + \frac{n_{11}}{n_{\cdot 1}} \left(\frac{X + n_{\cdot 1}}{N}\right)$$

$$= \frac{n_{10}}{n_{\cdot 0}} \left(\frac{Y + n_{\cdot 0}}{N}\right) + \frac{n_{11}}{n_{\cdot 1}} \left(\frac{X + n_{\cdot 1}}{N}\right) \tag{3.3}$$

In the estimate of λ given by (3.3), the proportion of the N units which has been classified in category "1" and "0" by the fallible measuring device, namely $\frac{X + n_{\cdot 1}}{N}$ and $\frac{Y + n_{\cdot 0}}{N}$, are corrected by multiplication by the ratios $\frac{n_{11}}{n_{\cdot 1}}$ and $\frac{n_{10}}{n_{\cdot 0}}$, respectively. The former ratio is an estimate of the proportion of randomized category "1" measurements which are also in fallible category "1"; the latter ratio is an estimate of the proportion of randomized category "1" measurements which are fallible category "0." Summing these two products of ratios yields an estimate of λ.

From $\lambda = \delta\pi + (1 - \delta)(1 - \pi)$

we get $\pi = \frac{\lambda - (1 - \delta)}{2\delta - 1}$ which in turn gives MLE of π

as $\hat{\pi} = \frac{\hat{\lambda} - (1 - \delta)}{2\delta - 1}$.

To obtain the MLE's of θ_{01}, θ_{10} we write

$$\gamma_2 = \frac{1}{1 - \alpha} [\theta_{01} \delta\pi + (1 - \theta_{10})(1 - \delta)(1 - \pi)]$$

and $1 - \gamma_1 = \frac{1}{\alpha} [\delta (1 - \pi)\theta_{10} + \pi(1 - \delta)(1 - \theta_{01})]$

or $(1 - \gamma_1) - \frac{(1 - \delta)\pi}{\alpha}$

$$= - \frac{\theta_{01}(1 - \delta)\pi}{\alpha} - \frac{\delta(1 - \pi)}{\alpha}\theta_{10}$$

and $\gamma_2 - \frac{(1 - \pi)(1 - \delta)}{1 - \alpha} = \frac{\delta\pi}{1 - \alpha}\theta_{01}$

$$+ \frac{(1 - \delta)(1 - \pi)}{1 - \alpha} \theta_{10}$$

or

$$\begin{pmatrix} (1 - \gamma_1) - \frac{(1 - \delta)\pi}{\alpha} \\ \gamma_2 - \frac{(1 - \pi)(1 - \delta)}{1 - \alpha} \end{pmatrix} =$$

$$\begin{pmatrix} -\frac{(1 - \delta)\pi}{\alpha} & \frac{\delta(1 - \pi)}{\alpha} \\ \frac{\delta\pi}{1 - \alpha} - \frac{(1 - \delta)(1 - \pi)}{1 - \alpha} \end{pmatrix} \begin{pmatrix} \theta_{01} \\ \theta_{01} \end{pmatrix}$$

so,

$$\begin{bmatrix} \theta_{01} \\ \theta_{10} \end{bmatrix} = \begin{bmatrix} -\frac{\pi(1 - \delta)}{\alpha} & \frac{\delta(1 - \pi)}{\alpha} \\ \frac{\delta\pi}{1 - \alpha} - \frac{(1 - \delta)(1 - \pi)}{1 - \alpha} \end{bmatrix}^{-1} \begin{bmatrix} (1 - \gamma_1) - \frac{(1 - \delta)\pi}{\alpha} \\ \gamma_2 - \frac{(1 - \pi)(1 - \delta)}{1 - \alpha} \end{bmatrix}$$

It is easily verified that the above matrix is non-singular provided $\delta \neq 1/2$.

With the randomized response technique, $\delta \neq 1/2$. Therefore, the MLE of π, θ_{01}, θ_{10} are given by

$$\hat{\pi} = \frac{\hat{\lambda} - (1 - \delta)}{2\delta - 1}$$

$$\begin{pmatrix} \hat{\theta}_{01} \\ \hat{\theta}_{10} \end{pmatrix} = \begin{pmatrix} -\frac{\hat{\pi}(1-\delta)}{\hat{\alpha}} & \frac{\delta(1-\hat{\pi})}{\hat{\alpha}} \\ \frac{\delta\hat{\pi}}{1-\hat{\alpha}} & -\frac{(1-\delta)(1-\hat{\pi})}{1-\hat{\alpha}} \end{pmatrix}^{-1} \begin{pmatrix} (1-\hat{\gamma}_1) - \frac{(1-\delta)\hat{\pi}}{\hat{\alpha}} \\ \hat{\gamma}_2 - \frac{(1-\hat{\pi})(1-\delta)}{1-\hat{\alpha}} \end{pmatrix} \tag{3.4}$$

4.4 $E(\hat{\pi}) = \pi$ and Asymptotic var $(\hat{\pi})$

Lemma 1: Let $\hat{\gamma}_1$, $\hat{\gamma}_2$, $\hat{\alpha}$ be as (2.2). Then

$$\text{var}(\hat{\gamma}_1) \doteq \frac{\gamma_1(1-\gamma_1)}{n\alpha}$$

$$\text{var}(\hat{\gamma}_2) \doteq \frac{\gamma_2(1-\gamma_2)}{n(1-\alpha)}$$

$$\text{var}(\hat{\alpha}) \doteq \frac{\alpha(1-\alpha)}{N}$$

$$\text{cov}(\hat{\gamma}_1\ \hat{\gamma}_2) = \text{cov}(\hat{\gamma}_1, \hat{\alpha}) = \text{cov}(\hat{\gamma}_2, \hat{\alpha}) = 0$$

Lemma 2:

(a) $E(\hat{\lambda}) = \lambda$ where $\hat{\lambda}$ is given in (3.3).

(b) $$\text{var}(\hat{\lambda}) \doteq \frac{\alpha\gamma_1(1-\gamma_1)}{n} + (1-\alpha)\frac{\gamma_2(1-\gamma_2)}{n} + (\gamma_1-\gamma_2)^2\frac{\alpha(1-\alpha)}{N}$$

Lemma 3: $E(\hat{\pi}) = \pi$.

$$\text{var}(\hat{\pi}) \doteq \frac{1}{(2\delta-1)^2}\left[\frac{\gamma_1(1-\gamma_1)\alpha + \gamma_2(1-\gamma_2)(1-\alpha)}{n} + \frac{(\gamma_1-\gamma_2)^2\alpha(1-\alpha)}{N}\right] \tag{4.1}$$

Lemma 1 can easily be established by noting that

$$\frac{\partial^2 \text{Log } L}{\partial \gamma_1^2} = \frac{-n_{11}}{\gamma_1^2} - \frac{n_{01}}{(1-\gamma_1)^2}$$

$$\frac{\partial^2 \text{Log } L}{\partial \gamma_2^2} = \frac{-n_{10}}{\gamma_2^2} - \frac{n_{00}}{(1-\gamma_2)^2}$$

$$\frac{\partial^2 \text{Log } L}{\partial \alpha^2} = -\frac{X + n_{\cdot 1}}{\alpha^2} - \frac{Y + n_{\cdot 0}}{(1-\alpha)^2}$$

and $$\frac{\partial^2 \text{Log } L}{\partial \gamma_1 \partial \gamma_2} = \frac{\partial^2 \text{Log} L}{\partial \gamma_1 \partial \alpha} = \frac{\partial^2 \text{Log} L}{\partial \gamma_1 \partial \alpha} = 0$$

and getting the inverse of the information matrix obtained by taking expectations of the above derivatives.

Lemma 2a can be shown as follows:

We have

$$\hat{\lambda} = \frac{n_{10}}{n_{\cdot 0}} \left(\frac{Y}{N}\right) + \frac{n_{10}}{N} + \frac{n_{11}}{n_{\cdot 1}} \left(\frac{X}{N}\right) + \frac{n_{11}}{N}$$

$$\text{so } E(\hat{\lambda}) = E\left(\frac{n_{10}}{n_{\cdot 0}}\right) E\left(\frac{Y}{N}\right) + E\left(\frac{n_{10}}{N}\right) + E\left(\frac{n_{11}}{n_{\cdot 1}}\right) E\left(\frac{X}{N}\right) + E\left(\frac{n_{11}}{N}\right)$$

$$= \frac{P_{10}}{1-\alpha} \frac{(N-n)(1-\alpha)}{N} + \frac{nP_{10}}{N} + \frac{P_{11}}{\alpha} \frac{(N-n)\alpha}{N} + \frac{nP_{11}}{N}$$

$$= P_{10} + P_{11} = \lambda$$

To establish lemma 2b, we write $\hat{\lambda} = \hat{\alpha}\hat{\gamma}_1 + (1-\hat{\alpha})\hat{\gamma}_2 = g$. Since the covariances are 0 by lemma 1, then

$$\text{var}(\hat{\lambda}) = \text{var}(\hat{\gamma}_1)\left(\frac{\delta g}{\delta \hat{\gamma}_1}\right)^2 + \text{var}(\hat{\gamma}_2)\left(\frac{\delta g}{\delta \hat{\gamma}_2}\right)^2 + \text{var}(\hat{\alpha})\left(\frac{\delta g}{\delta \hat{\alpha}}\right)^2 \tag{4.2}$$

where the partial derivatives are evaluated at γ_1, γ_2 and α respectively.

Use of 4.2 and lemma 1 establishes lemma 2b.

Since $\hat{\pi} = \frac{\hat{\lambda}-(1-\delta)}{(2\delta-1)}$, lemma 2a,b establish lemma 3.

To express variance of $\hat{\pi}$ in terms of false negative and false positive rate, let us first define:

$$\beta_{01} = P[F = 0 | R = 1]$$

$$\beta_{10} = P[F = 1 | R = 0]$$

The cell probabilities of page 77 can be written as

$$P_{00} = (1-\lambda)(1-\beta_{10})$$

$$P_{01} = \beta_{10}(1-\lambda)$$

$$P_{10} = \lambda\beta_{01}$$

$$P_{11} = \lambda(1-\beta_{01})$$

Lemma 4: Let $\gamma_1, \gamma_2, \alpha, \lambda$ be defined as before. Then the following identities hold: (Proof given in Appendix A.)

$$\text{(a)}\ \gamma_1(1-\gamma_1)\alpha + \gamma_2(1-\gamma_2)(1-\alpha) = \lambda(1-\lambda)\left[1 - \frac{\lambda(1-\lambda)(1-\beta_{01}-\beta_{10})^2}{\alpha(1-\alpha)}\right]$$

$$\text{(b)}\quad (\gamma_1-\gamma_2)^2\alpha(1-\alpha) = \frac{\lambda^2(1-\lambda)^2}{\alpha(1-\alpha)}[1-\beta_{01}-\beta_{10}]^2$$

where β_{01} and β_{10} are given above.

Using Lemma 4 in (4.1),

$$\text{var}(\hat{\pi}) \doteq \frac{\lambda(1-\lambda)}{(2\delta-1)^2 n}\left[1 - \frac{\lambda(1-\lambda)(1-\beta_{01}-\beta_{10})^2}{\alpha(1-\alpha)}\right] + \frac{\lambda^2(1-\lambda)^2(1-\beta_{01}-\beta_{10})^2}{(2\delta-1)^2 N\alpha(1-\alpha)} \tag{4.3}$$

In order to express variance of $\hat{\pi}$ in terms of the false negative rate, θ_{01}, and the false positive rate, θ_{10}, we expressed $(1-\beta_{01}-\beta_{10}) = \frac{1}{\lambda(1-\lambda)}\,\pi(1-\pi)(2\delta-1)(1-\theta_{01}-\theta_{10})$ (see Appendix A), which in turn gives

$$\text{var}(\hat{\pi}) \doteq \frac{\lambda(1-\lambda)}{n}\left\{\frac{1}{(2\delta-1)^2} - \frac{\pi^2(1-\pi)^2}{\lambda(1-\lambda)\alpha(1-\alpha)}\right\}(1-\theta_{01}-\theta_{10})^2$$

$$+ \frac{\pi^2(1-\pi)^2}{N\alpha(1-\alpha)}(1-\theta_{01}-\theta_{10})^2 \tag{4.4}$$

Thus $\hat{\pi}$ is an unbiased estimate of π with asymptotic variance given by the expression (4.4).

4.5 Square of Correlation Coefficient between Randomized and Fallible Measurement

Let ρ^2 be defined as the square of the correlation coefficient between R and F. From the distribution of R and F, we have $E(R) = E(R^2), = \lambda$, $E(F) = E(F^2) = \alpha$, $\text{var}(R) = \lambda(1-\lambda)$ $\text{var}(F) = \alpha(1-\alpha)$ and

$$\begin{aligned}\text{cov}(R,F) &= P_{11} - \lambda\alpha \\ &= \lambda(1-\beta_{01}) - \lambda[\beta_{10}(1-\lambda) + \lambda(1-\beta_{01})]. \\ &= \lambda[(1-\beta_{01})(1-\lambda) - \beta_{10}(1-\lambda)] \\ &= \lambda(1-\lambda)(1 - \beta_{01} - \beta_{10})\end{aligned}$$

$$\rho^2 = \frac{(\text{cov}(R,F))^2}{\text{var}(R)\,\text{var}(F)} = \frac{\lambda^2(1-\lambda)^2(1-\beta_{01}-\beta_{10})^2}{\alpha(1-\alpha)\lambda(1-\lambda)}$$

$$= \frac{\lambda(1-\lambda)(1-\beta_{01}-\beta_{10})^2}{\alpha(1-\alpha)}$$

By definition ρ^2 measures the strength of the relationship between randomized and fallible measurements and $0 \leq \rho^2 \leq 1$.

We can write the $\text{var}(\hat{\pi})$ from equation (4.4) as

$$\text{var}(\hat{\pi}) \doteq \frac{\lambda(1-\lambda)}{(2\delta-1)^2 n}(1-\rho^2) + \frac{\lambda(1-\lambda)}{(2\delta-1)^2 N}\rho^2$$

$$= \frac{\lambda(1-\lambda)}{(2\delta-1)^2}\left[\frac{1}{n}(1-\rho^2) + \frac{\rho^2}{N}\right]$$

Thus, the asymptotic variance of $\hat{\pi}$ is the weighted average of:

a. the variance of a binomial estimate of π based on n randomized measurements and

b. the variance of a binomial estimate of π based on N randomized measurements.

4.6 Special Cases for $\rho^2 = 0,1$

(a) When $\rho^2 = 0$, $\text{var}(\hat{\pi}) \doteq \frac{1}{(2\delta-1)^2}\frac{\lambda(1-\lambda)}{n}$ which can be expressed as the sum of the variance due to sampling and variance due to randomized device as follows:

Note, $\lambda = \pi\delta + (1-\delta)(1-\pi)$

$\Rightarrow$ $(1-\lambda) = \pi(1-\delta) + \delta(1-\pi)$

$$\text{so, } \lambda(1-\lambda) = [\pi\delta + (1-\delta)(1-\pi)][\pi(1-\delta) + \delta(1-\pi)]$$

$$= \pi^2\delta(1-\delta) + \pi(1-\pi)\delta^2 + \pi(1-\pi)(1-\delta)^2 + (1-\pi)^2\delta(1-\delta)$$

$$= 4\pi\delta^2 - 4\delta^2\pi^2 + 4\pi^2\delta + \pi + \delta - 4\delta\pi - \delta^2 - \pi^2$$

$$= 2\delta^2(-2\pi^2 + 2\pi - \tfrac{1}{2}) - 2\delta(-2\pi^2 + 2\pi - \tfrac{1}{2}) + \tfrac{1}{2}(-2\pi^2 + 2\pi - \tfrac{1}{2}) + \tfrac{1}{4}$$

$$= \tfrac{1}{4} + (2\delta^2 - 2\delta + \tfrac{1}{2})(-2\pi^2 + 2\pi - \tfrac{1}{2})$$

$\therefore$ when $\rho^2 = 0$,

$$\text{var}(\hat{\pi}) \doteq \frac{\tfrac{1}{4} + (2\delta^2 - 2\delta + \tfrac{1}{2})(-2\pi^2 + 2\pi - \tfrac{1}{2})}{4(\delta-\tfrac{1}{2})^2 n}$$

$$\doteq \frac{1}{16n}(\delta-\tfrac{1}{2})^{-2} + \frac{2(\delta^2 - \delta + \tfrac{1}{4})2(-\pi^2 + \pi - \tfrac{1}{4})}{4(\delta-\tfrac{1}{2})^2 n}$$

$$\doteq \frac{1}{16n}(\delta-\tfrac{1}{2})^{-2} - \frac{(\pi-\tfrac{1}{2})^2}{n}$$

$$\doteq \frac{1}{n}\left[\frac{1}{16(\delta-\frac{1}{2})^2} - (\pi-\tfrac{1}{2})^2\right]$$

$$\doteq \frac{1}{n}\left[\frac{1}{16(\delta-\tfrac{1}{2})^2} - \tfrac{1}{4}\right] + \frac{\pi(1-\pi)}{n}$$

which equals the variance due to randomized device plus variance due to sampling. $\rho^2 = 0$ implies the precision in the estimate of π, which is attained by using the proposed sampling scheme, is no better than the precision of a

binomial estimate based on n randomized measurements only. N-n additional fallible measurements do not yield any additional information concerning π and false negative and false positive rates. This is to be expected as the fallible and randomized measurements on a given unit are uncorrelated.

(b) When $\rho^2 = 1$, $\text{var}(\hat{\pi}) = \dfrac{\lambda(1-\lambda)}{(2\delta-1)^2 N}$

$$= \frac{\dfrac{1}{16(\delta-\frac{1}{2})^2} - \frac{1}{4}}{N} + \frac{\pi(1-\pi)}{N}$$

That is, in the case $\rho^2 = 1$, we obtain the same precision on N fallible measurements as compared to a binomial estimate of π based on N randomized measurements, since fallible measurements are as good as randomized measurements.

Moreover, it is to be noted that $\text{var}(\hat{\pi})$ is a decreasing function of ρ^2 which can be seen as follows:

$$\text{var}(\pi) \doteq \frac{\lambda(1-\lambda)}{(2\delta-1)^2}\left[\frac{1}{n}(1-\rho^2) + \frac{1}{N}\rho^2\right]$$

$$\doteq \frac{\lambda(1-\lambda)}{(2\delta-1)^2}\left[\frac{1}{n} + \left(\frac{1}{N} - \frac{1}{n}\right)\rho^2\right.$$

$$\doteq \frac{\lambda(1-\lambda)}{n(2\delta-1)^2} + \frac{\lambda(1-\lambda)}{(2\delta-1)^2}\left(\frac{1}{N} - \frac{1}{n}\right)\rho^2$$

For fixed n, N with $n < N$, $\left(\frac{1}{N} - \frac{1}{n}\right)\rho^2$ is negative which establishes the above proposition.

Therefore, when $\rho^2 > 0$ and since $n<N$, the subsampling

scheme (randomized device on subsample) is more efficient than the randomized response on n observations alone when the cost of N fallible measurements is not considered (approximately the case when the total sample of misclassified responses is already available on file).

The efficiency of the randomized response on n observations only with respect to the randomized response on subsampling is given by

$$E = \frac{\text{var}(\hat{\pi}_R)_{n,N}}{\text{var}(\hat{\pi}_R)_n} = 1 - \rho^2(1 - \frac{n}{N})$$

However, ρ^2 is a function of $\pi, \beta_{01}, \beta_{10}$ and δ. In any specific problem the probability of biased estimates from randomized responses should be considered too. The researcher may make a decision based on an appropriate pilot study as to which course of action should be taken: (i) use only fallible responses or (ii) use the subsampling scheme with randomization device on a subsample. Such a decision will be based on best guesses of cost, error rates, π, bias in the randomized response estimator, and δ.

4.7 Sample Size Determination

To develop criteria for selecting n and N, we must have an idea as to the cost of measurement; this will certainly be a consideration for determining the sample sizes. We assume that the total cost of measurements, defined as C,

is a linear function of n and N. That is, $C = C_1 n + C_2 N = C_1 n + C_2(n + m)$, where C_1 = cost of obtaining one randomized measurement, C_2 = cost of obtaining one fallible measurement, and $m = N-n$.

In practice two situations may arise. An investigator might have a certain budget C_o for measurement costs and he might choose n and N to minimize the variance of estimator, or he might want to obtain a given precision of the estimator, v_o, at minimum cost. We can state these two different situations in the following manner.

A. Fixed Cost

We choose n and N to minimize $v(\hat{\pi}) \doteq \frac{\lambda(1-\lambda)}{(2\delta-1)^2}\left[\frac{1}{n}(1-\rho^2) + \frac{1}{N}\rho^2\right]$ subject to $C_1 n + C_2(n + m) = C_o$.

B. Fixed Variance

We choose n and N to minimize $C_1 n + C_2(n + m)$ subject to $v(\hat{\pi}) \doteq \frac{\lambda(1-\lambda)}{(2\delta-1)^2}\left[\frac{1}{n}(1-\rho^2) + \frac{1}{N}\rho^2\right] = v_o$.

Special Cases

(i) When $\rho^2 = 0$, we have mentioned earlier that additional m fallible measurements do not yield any additional information concerning π, β_{01}, and β_{10}, which implies that $C_1 n + C_2 n = C_o$ since $m = 0$, i.e.

$$n = \frac{C_o}{C_1 + C_2}$$

In this situation both fallible and randomized measurements are taken on n units.

(ii) When $\rho^2 = 1$, we obtain the same precision on N fallible measurements as compared to the binomial estimate of π based on N randomized measurements. So randomization on subsampling is not necessary in terms of the precision of the estimate and cost.

(iii) When $0 < \pi^2 < 1$ both of these problems can be solved by the method of undetermined multipliers. The solutions are given in the following tabulations.

Optimum values of n and N

Fixed Cost Problem

$$n = \frac{C_o}{C_1} \frac{R_c f}{1+R_c f}$$

$$N = \frac{C_o - nc_1}{C_2}$$

where $R_c = \frac{C_1}{C_2}$

Fixed Variance Problem

$$n = \frac{A}{v_o} [(1-\rho^2) + \rho^2 f]$$

$$N = \frac{nn_\sigma \rho^2}{n-n_\sigma(1-\rho^2)}$$

$$n_\sigma = \frac{A}{v_o}$$

$$A = \frac{\lambda(1-\lambda)}{(2\delta-1)^2}$$

$$f = \left(\frac{1 - \rho^2}{\rho^2 \frac{C_1}{C_2}} \right)^{1/2}$$

C_1 = Cost for obtaining one randomized measurement

C_2 = Cost for obtaining one fallible measurement

C_o = Total budget for measurement costs

v_o = Given precision of the estimate

To apply the preceding optimum formulae, the square of the correlation coefficient, ρ^2, must be known. But ρ^2 depends on λ, α, and the probabilities of misclassifications. Thus, n and N cannot be determined exactly in advance.

One method of obtaining estimates of n and N is to take a pilot sample of k units and obtain the fallible and randomized responses for each of the k units. Estimates of λ, α, β_{01}, and β_{10} can be obtained and then n and N can be determined using the estimate of the square of the correlation coefficient between randomized and fallible measurements. The problem of these methods is the choice of initial sample size k. On one hand k should be sufficiently large so that reasonable estimates of λ, β_{01}, β_{10} and α can be obtained. On the other hand, if k is chosen too large, we run the risk of too many randomized measurements. This problem is not

studied here and is subject to further research in this area.

CHAPTER V

GENERAL CONCLUSIONS AND COMMENTS

One of the strong motivations for the study reported in this thesis was to determine how serious a problem misclassification presents to the analysis of data in research. The presence of misclassification may be adjusted for in analysis if the error rates are known in the two samples of the fourfold table. But the rates are often unknown and the rates often vary from study to study making it very doubtful that a rate known from one study would be applicable to a different sets of circumstances. The situation in which error rates are unequal and vary from study to study is very realistic. For example, these rates vary on screening tests with age and religion for circumcision status, with sex and race for syphilis, with sex for rheumatism and with sex for tuberculin sensitivity. Thus in practice the niceties of the Bross' model do not apply and the situation is in fact much more complicated.

For the range of π common to epidemiologic research, and for error rates chosen from practice, we have attempted to provide a table which shows the change in the nominal significance level and power of the usual tests when the data in a fourfold table are analyzed under the assumption that no

misclassification errors are present, and under various assumptions about the error structure.

For the exact significance test in the fourfold table, i.e., Fisher's exact test of the null hypothesis $H_0: \pi_1 = \pi_2$ versus the alternative $H_1 : \pi_1 > \pi_2$, the nominal level α which is set for the no-misclassification case is inflated if $\phi_1 > \phi_2$ and deflated if $\phi_1 < \phi_2$. The magnitude of the difference between the two false positive rates determines the extent of the change over the nominal level. The effects of the false positive rates are reversed for $H_1 : \pi_1 < \pi_2$. The trends for the power of this one sided test are similar.

For the two sided test of $H_0 : \pi_1 = \pi_2$ against $H_1 : \pi_1 \neq \pi_2$, the level increases over the nominal level α of the Fisher's exact test. The power associated with this two sided test generally decreases from that associated with the alternative without misclassification for fixed level.

We have shown that the presence of misclassification, even if it is small, can have substantial effects on estimation of parameters as well as on significance level. The discussions in Chapter 2 lead us to the awareness of the importance of estimating the error rates in a particular study. Since those error rates are quite varying in nature, it is necessary to estimate these error rates in each separate independent study wherever its presence is suspected before continuing the research on large scale.

We have proposed the randomized response technique as a useful approach to estimate the error rates and the prevalence rate, π , in certain types of studies. An unbiased estimate of π was given with a method for the estimation of variance of $\hat{\pi}$. Formulae for sample size determinations for fixed cost and fixed variance were given.

We may therefore come to some general conclusions about the effects of misclassification on the analysis of 2 x 2 tables. The false positive rates have more effect on the estimation of parameters as well as on hypothesis testing than false negative rates suggesting that attention is better paid to eliminating false positives than false negatives in a study. The test becomes more stable against errors of measurement as the null hypothesis approaches 0.5 since this value of π is most stable against the bias produced by the errors.

In medical screening, in many situations θ is larger than ϕ; however, often the false positive rates are larger. If the error rates are large enough, it might no longer be suitable to perform analysis on a study before giving a closer look to the misclassification present in the data. It is far more important to spend fixed sums of money on reducing error rates, than on increasing sample sizes. This may be an advantage to the investigator in preventing him from doing research from which no legitimate conclusions may be drawn.

Our proposed randomized response procedure is useful for practical situations where there is no exact device for measuring a true response. Of course, the randomized response technique is not free from bias and we suggest to use the double sampling scheme proposed by Tenenbien if a true device of measurement is possible. For future study, our proposed randomized response technique can be extended to multinomial data with misclassification or one can perform a comparative study of different randomized response designs to estimate error rates and proportions.

BIBLIOGRAPHY

1. Assakul, K.: Testing hypotheses with categorical data subject to misclassification. Mimeo series no. 448, Institute of Statistics, Raleigh, N. C., 1965.

2. Assakul, K. and C. H. Proctor: Testing independence in two way contingency tables with data subject to misclassification. Psychometrika 32: 67-76, 1967.

3. Aitken-Swan, J. and D. Baird: Circumcision and cancer of the cervix. Brit. J. Cancer 19: 217-227, 1965.

4. Berkson, J.: Cost-utility as a measure of the efficiency of a test. J. Amer. Stat. Ass. 42: 246-255, 1947.

5. Blumberg, M. S.: Evaluating health screening procedures. Operations Research 5: 351-360, 1957.

6. Bross, I.: Misclassification in 2 x 2 tables. Biometrics 10: 478-486, 1954.

7. Buck, A. A. and J. J. Gart: Comparison of a screening test and a reference test in epidemiologic studies: I: Indices of agreement and their relation to prevalence. Amer. J. Epid. 83: 586-592, 1966.

8. Buck, A. A. and H. Mayer: Comparative studies of the Rapid Plasma Reagin Card Test for syphillis and the VDRL slide test in Ethiopia. Am. J. Hyg. 80: 85-90, 1964.

9. Buck, A. A., E. H. Sadur, R. I. Anderson and E. Schaffa: Comparative studies of some immunologic screening tests for schistosmiasis in Ethiopia. Am. J. Hyg. 80: 75-84, 1964.

10. Buck, A. A. and D. J. Spruyt: Sereoreactivity in the venereal disease research laboratory slide test and the Fluorescent Treponemal Antibody Test: A study of pallerus in selected disease and control groups in Ethiopia. Am. J. Hyg. 80: 91-102, 1964.

11. Buell, P. and J. E. Dunn: The dilution effect of misclassification. Amer. J. Public Health 54: 598-602, 1964.

12. Chapman, J. M., D. B. Loveland, L. S. Goerka, G. Jacobson and W. J. Rothrock: Validity of 70 mm. photofluorograms of the chest in screening a population for heart disease. J. Chron. Dis. 12: 521-533, 1966.

13. Cobb, S., W. R. Merchant and J. E. Warren: An epidemiologic look at the problem of classification in the field of arthritis. J. Chron. Dis. 2: 50-54, 1955.

14. Cochran, W. G.: Errors of measurement in statistics. Technometrics 10: 637-666, 1968.

15. Cooley, R. N., C. H. Agnew and G. Rios: Diagnostic accuracy of the barium enema study in carcinoma of the colon and rectum. Am. J. Roentgen 84: 316-331, 1960.

16. Diamond, E. L.: The effect of misclassification errors in attribute studies. Mimeo. note, John Hopkins School of Hygiene and Public Health.

17. Diamond, E. L. and A. M. Lilienfeld: Effects of errors in classification and diagnosis in various types of epidemiological studies. Amer. J. Public Health 10: 1137-1144, 1962.

18. Diamond, E. L. and A. M. Lilienfeld: Misclassification errors in 2 x 2 tables with one margin fixed: Some further comments. Amer. J. Public Health 10: 2106-2110, 1962.

19. Dunn, J. E. and P. Buell: Association of cervical cancer with circumcision of sexual partners. J. Nat. Cancer Inst. 22: 749-764, 1959.

20. Dunn, J. E. and S. W. Greenhouse: Cancer diagnostic tests: Principles and criteria for development and evaluation. PHS Publication no. 9, 1950.

21. Feldman, J. J.: The household interview survey as a technique for the collectgio of morbidity data. J. Chron. Dis. 11: 535-557, 1960.

22. Felligi, I. P.: Response variance and its estimation. J. Amer. Stat. Ass. 59: 1016-1041, 1964.

23. Gart, J. J. and A. A. Buck: Comparison of a screening test and a reference test in epidemiologic studies: II: A probabilistic model for the comparison of diagnostic tests. Amer. J. Epid. 83: 593-602, 1966.

24. Giesbrecht, F.: Classification errors and measures of association in contingency tables, Statistics Research Division of Research Triangle Institute, Reprint series no. S-71, reprinted from Social Statistics section Proceedings of the Amer. Stat. Ass., 1967.

25. Greenberg, B. G., D. G. Horvitz and J. R. Abernathy: A comparison of randomized response designs. SIAM Philadelphia: 787-813, 1974.

26. Gullen, W. H., J. E. Bearman and E. A. Johnson: Effects of misclassification in epidemiologic studies. Public Health Report 83: 914-918, 1968.

27. Hansen, M. H., W. N. Hurwitz and M. A. Bershad: Measurement errors in census and surveys. Bull. Int. Stat. lust. 38: 359-374, 1961.

28. Jolly, J. J., W. V. White, J. Portony and J. W. Gutridge: Blood sugar and syphilis serology using a single specimen. PHS Report 75: 115-118, 1960.

29. Keys, A. and J. K. Kohlberg: Effect of misclassification on estimated relative prevalence of a characteristic: Part I. Two populations infallibly distinguished. Part II. Error in two variables. Amer. J. Public Health 53: 1656-1665, 1963.

30. Koch, G. G.: The effect of non-sampling errors on measures of association in 2 x 2 contingency tables. J. Amer. Stat. Ass. 64: 852-863, 1969.

31. Kurlander, A. B., E. H. Hill and P. H. Euterline: An evaluation of some commonly used screening tests for heart disease and hypertension. J. Chron. Dis. 2: 427-439, 1955.

32. Kurlander, A. B., A. P. Iskrant and M. E. Kent: Screening tests for diabetes: A study of specificity and sensitivity. Diabetes 3: 213-219, 1954.

33. Lehmann, E. L.: Testing statistical hypotheses. New York, John Wiley and Sons, 1959.

34. Lilienfeld, A. and S. Graham: Validity of determining circumcision status by questionnaire as related to epidemiological studies of cancer of the cervix. J. Nat. Cancer Iust. 21: 713-720, 1958.

35. Madow, W. G.: On some aspects of response error measurements. ASA Proceedings of the Social Statistics Section 182-192, 1965.

36. Mateer, J. G., J. I. Baltz, D. F. Manon and J. M. MacMillan. Liver function tests. J. A. M. A. 121: 723-728, 1943.

37. Meigs, J. V., R. M. Graham, M. Fremont-Smith, L. T. Jantzen and C. B. Nelson: II. The values of the vaginal smear in the diagnosis of uterine cancer. Surg., Gynec. and Obs. 81: 337-345, 1945.

38. Mote, V. L. and R. N. Anderson: The effect of misclassification on the properties of X^2-tests. Biometrika 52: 95-109, 1965.

39. Moyer, J. H. and S. R. Womack: Glucose tolerance: I. A comparison of four types of diagnostic tests in 103 control subjects and 26 patients with diabetes. Am. J. Med. Sci. 219: 161-173, 1950.

40. National Center for Health Statistics (a) Three views of hypertension and heart disease; Vital and Health Statistics. P.H.S. Publication no. 1000 series 2 no. 22; (b) The one-hour oral glucose tolerance test. P.H.S. Publication no. 1000, series 2, no. 3; (c) Findings on the serologic test for syphilis in adults. P.H.S. Publication no. 1000, series 11, no. 9.

41. Newell, D. J.: Errors in the interpretation of errors in epidemiology. Amer. J. Public Health 52: 1925-1928, 1962.

42. Newell, D. J.: Misclassification in 2 x 2 tables. Biometrics 19: 186-187, 1963.

43. Rogot, E.: A note on measurement errors and detecting real differences. J. Amer. stat. Ass. 56: 314-319, 1961.

44. Rubin, T., J. Rosenbaum and S. Cobb: The use of interview data for the detection of association in field studies. J. chron. Dis. 4: 253-266, 1956.

45. Sosman, M. C.: The specificity and reliability of Roentgenographic diagnosis. New Eng. J. Med. 242 (2): 849-855, 1950.

46. Stern, E. and P. A. Lachenbruch: Circumcision information in a cancer detection center population. J. chron. Dis. 21: 117-124, 1968.

47. Sutcliffe, J. P.: A probability model for errors of classification: II. Particular cases. Psychometrika 30: 129-155, 1965.

48. Tenenbein, A.: A double sampling scheme for estimating from binomial data with misclassification. J. Amer. stat. Ass. 65: 1350-1361, 1970.

49. Tenenbein, A.: A double sampling scheme for estimating from binomial data with misclassifications: sample size determination. Biometrics 27: 935-944, 1971.

50. Unger, R. H. and L. L. Madison: A new diagnostic procedure for mild diabetes mellitus: Evaluation of an intervenous tolbutamide response test. Diabetes 7: 455-461, 1958.

51. Walsh, J. E.: Loss in test efficiency due to misclassification in 2 x 2 tables. Biometrics 19: 158-162, 1963.

52. Warner, S. L.: Randomized response: A survey technique for eliminating evasive answer bias. J. Amer. stat. Ass. 63-69, 1965.

53. Wylie, C. M.: Use of death rates to evaluate cardiovascular screening tests. Am. Heart J. 63: 92-97, 1962.

54. Youden, W. J.: Index for rating diagnostic tests. Cancer 3: 32-35, 1950.

APPENDICES

APPENDIX A

(All notations used here are defined as in the text.)

(i) Show $\gamma_1(1-\gamma_1)\alpha + \gamma_2(1-\alpha_2)(1-\alpha) = \lambda(1-\lambda)\left[1 - \frac{\lambda(1-\lambda)(1-\beta_{01}-\beta_{10})^2}{\alpha(1-\alpha)}\right]$

We have $\gamma_1(1-\gamma_1) + \gamma_2(1-\gamma_2)(1-\alpha)$

$$= \frac{P_{11}P_{01}}{\alpha} + \frac{P_{10}P_{00}}{1-\alpha}$$

since $\gamma_1 = \frac{P_{11}}{\alpha}$

$$\gamma_2 = \frac{P_{10}}{1-\alpha}$$

$$1-\gamma_1 = \frac{P_{01}}{\alpha}$$

$$1-\gamma_2 = \frac{P_{00}}{1-\alpha}$$

$$= \frac{\lambda(1-\beta_{01})\beta_{10}(1-\lambda)}{\alpha} + \frac{\lambda\beta_{01}(1-\lambda)(1-\beta_{10})}{1-\alpha}$$

$$= \frac{\lambda(1-\beta_{01})\beta_{10}(1-\lambda)(1-\alpha) + \lambda\beta_{01}(1-\lambda)(1-\beta_{10})\alpha}{\alpha(1-\alpha)}$$

Numerator $= \lambda(1-\lambda)[(1-\beta_{01})\beta_{10}(1-\alpha) + \beta_{01}(1-\beta_{10})\alpha]$

$$= \lambda(1-\lambda)[(1-\beta_{01})\beta_{10}\{\lambda\beta_{01}+(1-\lambda)(1-\beta_{10})\} +$$

$$\beta_{01}(1-\beta_{10})\{\beta_{10}(1-\lambda) + \lambda(1-\beta_{01})\}]$$

since $\alpha = P_{01} + P_{11} = \beta_{10}(1-\lambda) + \lambda(1-\beta_{01})$

$$1-\alpha = P_{00} + P_{10} = (1-\lambda)(1-\beta_{10}) + \lambda\beta_{01}$$

$$= \lambda(1-\lambda)[\lambda\beta_{01}(1-\beta_{01})\beta_{10} + (1-\lambda)\beta_{10}(1-\beta_{10})(1-\beta_{01}) +$$

$$\lambda(1-\beta_{01})\beta_{01}(1-\beta_{10})$$

$$+ (1-\lambda)\beta_{10}(1-\beta_{10})\beta_{01}]$$

$$=\lambda(1-\lambda)[\lambda\beta_{01}(1-\beta_{01})\{\beta_{10} + 1 - \beta_{10}\}$$

$$+ (1-\lambda)\beta_{10}(1-\beta_{10})\{1 - \beta_{01} + \beta_{01}\}]$$

$$= \lambda(1-\lambda)[\lambda\beta_{01}(1-\beta_{01}) + (1-\lambda)\beta_{10}(1-\beta_{10})]$$

Now $\alpha = \lambda(1-\beta_{01}) + (1-\lambda)\beta_{10}$

or $\alpha\beta_{01} = \lambda\beta_{01}(1-\beta_{01}) + (1-\lambda)\beta_{10}\beta_{01}$

and $(1-\beta_{10})\alpha = \lambda(1-\beta_{01})(1-\beta_{10}) + (1-\lambda)\beta_{10}(1-\beta_{10})$

so $\alpha\beta_{01} + (1-\beta_{10})\alpha = \lambda\beta_{01}(1-\beta_{01}) + (1-\lambda)\beta_{10}\beta_{01} +$

$$\lambda(1-\beta_{01})(1-\beta_{10}) + (1-\lambda)\beta_{10}(1-\beta_{10})$$

or $\lambda\beta_{01}(1-\beta_{01}) + (1-\lambda)\beta_{10}(1-\beta_{10}) = \alpha\beta_{01} + (1-\beta_{10})\alpha$

$$- (1-\lambda)\beta_{10}\beta_{01} - \lambda(1-\beta_{01})(1-\beta_{10})$$

$$= \left\{\lambda(1-\beta_{01}) + (1-\lambda)\beta_{10}\right\}\beta_{01} + \left\{\lambda(1-\beta_{01}) + (1-\lambda)\beta_{10}\right\}(1-\beta_{10})$$

$$- (1-\lambda)\beta_{01}\beta_{10} - \lambda(1 - \beta_{01} - \beta_{10} + \beta_{01}\beta_{10})$$

$$= \lambda\beta_{01} - \lambda\beta_{01}^2 + (1-\lambda)\beta_{10}\beta_{01} + \lambda - \lambda\beta_{01} + (1-\lambda)\beta_{10} -$$

$$\lambda\beta_{10} + \lambda\beta_{01}\beta_{10} - (1-\lambda)\beta_{10}\beta_{10} - (1-\lambda)\beta_{10}\beta_{01} - \lambda +$$

$$\lambda\beta_{01} + \lambda\beta_{10} - \lambda\beta_{01}\beta_{10}$$

$$= -\lambda\beta_{01}^2 + (1-\lambda)\beta_{10} - (1-\lambda)\beta_{10}^2 + \lambda\beta_{01}$$

$$= [-\lambda\beta_{01}^2 + (1-\lambda)\beta_{10} - (1-\lambda)\beta_{10}^2 + \lambda\beta_{01}]\ [\lambda + (1-\lambda)]$$

$$= -\lambda^2\beta_{01}^2 + \lambda(1-\lambda)\beta_{10} - \lambda(1-\lambda)\beta_{10}^2 + \lambda^2\beta_{01} - \lambda(1-\lambda)\beta_{01}^2$$

$$+ (1-\lambda)^2\beta_{10} - (1-\lambda)^2\beta_{10}^2 + (1-\lambda)\lambda\beta_{01}$$

$$= \lambda^2\beta_{01} + \lambda(1-\lambda) - \lambda(1-\lambda)\beta_{10} - \lambda^2\beta_{01}^2 - \lambda(1-\lambda)\beta_{01} +$$

$$\lambda(1-\lambda)\beta_{01}\beta_{10} + \lambda(1-\lambda)\beta_{10}\beta_{01} + (1-\lambda)^2\beta_{10} - (1-\lambda)^2\beta_{10}^2 -$$

$$\lambda(1-\lambda) + \lambda(1-\lambda)\beta_{10} + \lambda(1-\lambda)\beta_{01} - \lambda(1-\lambda)\beta_{10}\beta_{01} - \lambda(1-\lambda)\beta_{10}\beta_{01}$$

$$+ \lambda(1-\lambda)\beta_{10} - \lambda(1-\lambda)\beta_{10}^2 - \lambda(1-\lambda)\beta_{01}^2 + \lambda(1-\lambda)\beta_{01}$$

$$= \lambda[\lambda\beta_{01} + (1-\lambda) - (1-\lambda)\beta_{10}] - \lambda\beta_{01}[\lambda\beta_{01} + (1-\lambda) - (1-\lambda)\beta_{10}]$$

$$+ (1-\lambda)\beta_{10}[\lambda\beta_{01} + (1-\lambda) - (1-\lambda)\beta_{10}] - \lambda(1-\lambda)[1 - \beta_{10} -$$

$$\beta_{01} + \beta_{10}\beta_{01} + \beta_{10}\beta_{01} - \beta_{10} + \beta_{10}^2 + \beta_{01}^2 - \beta_{01}]$$

$$= [\lambda - \lambda\beta_{01} + (1-\lambda)\beta_{10}][\lambda\beta_{01} + (1-\lambda) - (1-\lambda)\beta_{10}]$$

$$- \lambda(1-\lambda)[1 + \beta_{10}^2 + \beta_{01}^2 - 2\beta_{01} - 2\beta_{10} + 2\beta_{10}\beta_{01}]$$

$$= \alpha(1-\alpha) - \lambda(1-\lambda)(1 - \beta_{01} - \beta_{10})^2$$

since $\alpha = P_{01} + P_{11} = \lambda - \lambda\beta_{01} + (1-\lambda)\beta_{10}$

$$1 - \alpha = P_{00} + P_{10} = \lambda\beta_{01} + (1-\lambda) - (1-\lambda)\beta_{10}$$

so we can write

$$\gamma_1(1-\gamma_1)\alpha + \gamma_2(1-\gamma_2)(1-\alpha) =$$

$$\frac{\lambda(1-\lambda)[\alpha(1-\alpha) - \lambda(1-\lambda)(1 - \beta_{01} - \beta_{10})^2]}{\alpha(1-\alpha)}$$

$$= \lambda(1-\lambda)[1 - \frac{\lambda(1-\lambda)(1-\beta_{01} - \beta_{10})^2}{\alpha(1-\alpha)}$$

(ii) Show $(\gamma_1 - \gamma_2)^2 \alpha(1-\alpha) = \frac{\lambda^2(1-\lambda)^2}{\alpha(1-\alpha)} (1-\beta_{01} - \beta_{10})^2$

We have $\alpha(1-\alpha)(\gamma_1 - \gamma_2)^2 = \alpha(1-\alpha)[\frac{P_{11}}{\alpha} - \frac{P_{10}}{1-\alpha}]^2$

$$= \alpha(1-\alpha)[\frac{\lambda(1-\beta_{01})}{\alpha} - \frac{\lambda\beta_{01}}{1-\alpha}]^2$$

$$= \alpha(1-\alpha)[\frac{\lambda - \lambda\alpha - \lambda\beta_{01} + \lambda\alpha\beta_{01} - \lambda\alpha\beta_{01}}{\alpha(1-\alpha)}]^2$$

$$= \frac{(\lambda^2 (1 - \alpha - \beta_{01})^2}{\alpha(1-\alpha)}$$

$$= \frac{\lambda^2}{\alpha(1-\alpha)} \left[\lambda\beta_{01} + (1-\lambda) - (1-\lambda)\beta_{10} - \beta_{01}\right]^2$$

$$= \frac{\lambda^2}{\alpha(1-\alpha)} \left[(1-\lambda) - (1-\lambda)\beta_{10} - \beta_{01}(1-\lambda)\right]^2$$

$$= \frac{\lambda^2(1-\lambda)^2}{\alpha(1-\alpha)} \left[1 - \beta_{10} - \beta_{01}\right]^2$$

(iii) Show $(1-\beta_{01} - \beta_{10}) = \frac{1}{\lambda(1-\lambda)}\ \pi\ (1-\pi)(2\delta - 1)(1 - \theta_{01} - \theta_{10})$

We have $\beta_{01} = \frac{P_{10}}{\lambda}$, $\beta_{10} = \frac{P_{01}}{1-\lambda}$, then

$$\beta_{01} + \beta_{10} = \frac{P_{10}}{\lambda} + \frac{P_{01}}{1-\lambda} = \frac{P_{10}(1-\lambda) + P_{01}\lambda}{\lambda(1-\lambda)}$$

$$= \frac{1}{\lambda(1-\lambda)} \left\{P_{10} - \lambda(P_{10} - P_{01})\right\}$$

Now, $P_{10} - \lambda(P_{10} - P_{01})$

$$= \theta_{01}\delta\pi + (1-\theta_{10})(1-\delta)(1-\pi) - \lambda[\theta_{01}\delta\pi + (1-\theta_{10})(1-\delta)(1-\pi) -$$

$$\delta(1-\pi)\theta_{10} - \pi(1-\delta)(1-\theta_{01})]$$

$$= (1-\delta)(1-\pi) - \lambda(1-\delta)(1-\pi) + \lambda\pi(1-\delta)$$

$$+ \theta_{01}[\delta\pi - \lambda\delta\pi - \lambda\pi(1-\delta)] + \theta_{10}[\lambda(1-\delta)(1-\pi) -$$

$$(1-\delta)(1-\pi) + \lambda\delta(1-\pi)]$$

Now $(1-\delta)(1-\pi) - \lambda(1-\delta)(1-\pi) + \lambda\pi(1-\delta)$ can be expressed as $(\lambda-\delta\pi) - \lambda(\lambda - \delta\pi) + \lambda\pi(1-\delta) = \lambda\pi(1-\delta) + (\lambda - \delta\pi)(1-\lambda)$

$$= \lambda\pi(1-\delta) + \lambda(1-\lambda) - \delta\pi(1-\lambda),$$

and $\delta\pi - \lambda\delta\pi - \lambda\pi(1-\delta)$ can be written as

$$\pi(\delta - \lambda\delta - \lambda + \lambda\delta) = \pi(\delta-\lambda) = \pi[\delta - \delta\pi - (1-\delta)(1-\pi)]$$

$$= \pi[\delta(1-\pi) - (1-\delta)(1-\pi)] = \pi(1-\pi)(2\delta - 1),$$

and finally $\lambda(1-\delta)(1-\pi) - (1-\delta)(1-\pi) + \lambda\delta(1-\pi)$ can be written as

$$(1-\pi)(\lambda - \lambda\delta - 1 + \delta + \lambda\delta) = (1-\pi)[\delta\pi + (1-\delta)(1-\pi) - 1 + \delta]$$

$$= (1-\pi)(\delta\pi - 1 - \delta - \pi + \delta\pi - 1 + \delta)$$

$$= (1-\pi)\,\pi(2\delta - 1)$$

so using the above equation $\beta_{01} + \beta_{10}$ can be written as

$$\beta_{01} + \beta_{10} = \frac{1}{\lambda(1-\lambda)}[\lambda\pi(1-\delta) + (\lambda-\delta\pi)(1-\lambda) + (\theta_{01} + \theta_{10})\pi(1-\pi)(2\delta - 1)]$$

so $$1 - \beta_{01} - \beta_{10} = \frac{1}{\lambda(1-\lambda)}\{\lambda(1-\lambda) - \pi(1-\pi)(2\delta-1)(\theta_{01} + \theta_{10}) - \lambda\pi(1-\delta) - \lambda(1-\lambda) + \delta\pi(1-\lambda)\}$$

$$= \frac{1}{\lambda(1-\lambda)}\{-\pi(1-\pi)(2\delta - 1)(\theta_{01} + \theta_{10}) - \lambda\pi(1-\delta) + \delta\pi(1-\lambda)\}$$

$$= \frac{1}{\lambda(1-\lambda)}\{-\pi(1-\pi)(2\delta-1)(\theta_{01} + \theta_{10}) - \lambda\pi + \delta\pi(\lambda + 1 - \lambda)\}$$

$$= \frac{1}{\lambda(1-\lambda)}\{-\pi(1-\pi)(2\delta-1)(\theta_{01} + \theta_{10}) - [\pi\delta + (1-\delta)(1-\pi)]\pi + \delta\pi\}$$

since $\lambda = \delta\pi + (1 - \delta)(1 - \pi)$

$$= \frac{1}{\lambda(1-\lambda)} \left\{ - \pi(1-\pi)(2\delta - 1)(\theta_{01} + \theta_{10}) - \delta\pi\pi \right.$$

$$\left. - \pi(1-\pi)(1-\delta) + \delta\pi \right\}$$

$$= \frac{1}{\lambda(1-\lambda)} \left\{ -\pi(1-\pi)(2\delta - 1)(\theta_{01} + \theta_{10}) \right.$$

$$\left. + \pi(1-\pi)(2\delta - 1) \right\}$$

$$= \frac{1}{\lambda(1-\lambda)} \pi(1-\pi)(2\delta - 1)(1 - \theta_{01} - \theta_{10})$$

APPENDIX B
Table 1

$E(X/N) = PI(1-THETA) + (1-PI)PHI$

P= .10

THETA

PHI	0	.05	.10	.15	.20	.50	.75	.80	.90	.95	1.00
0	.100000	.095000	.090000	.085000	.080000	.050000	.025000	.020000	.010000	.005000	0
.05	.145000	.140000	.135000	.130000	.125000	.095000	.070000	.065000	.055000	.050000	.045000
.10	.190000	.185000	.180000	.175000	.170000	.140000	.115000	.110000	.100000	.095000	.090000
.15	.235000	.230000	.225000	.220000	.215000	.185000	.160000	.155000	.145000	.140000	.135000
.20	.280000	.275000	.270000	.265000	.260000	.230000	.205000	.200000	.190000	.185000	.180000
.50	.550000	.545000	.540000	.535000	.530000	.500000	.475000	.470000	.460000	.455000	.450000
.75	.775000	.770000	.765000	.760000	.755000	.725000	.700000	.695000	.685000	.680000	.675000
.80	.820000	.815000	.810000	.805000	.800000	.770000	.745000	.740000	.730000	.725000	.720000
.90	.910000	.905000	.900000	.895000	.890000	.860000	.835000	.830000	.820000	.815000	.810000
.95	.955000	.950000	.945000	.940000	.935000	.905000	.880000	.875000	.865000	.860000	.855000
1.00	1.000000	.995000	.990000	.985000	.980000	.950000	.925000	.920000	.910000	.905000	.900000

VARIANCE OF X/N

0	.009000	.008598	.008190	.007778	.007360	.004750	.002438	.001960	.000990	.000497	0
.05	.012398	.012040	.011677	.011310	.010937	.008598	.006510	.006077	.005197	.004750	.004298
.10	.015390	.015078	.014760	.014437	.014110	.012040	.010177	.009790	.009000	.008597	.008190
.15	.017978	.017710	.017438	.017160	.016878	.015078	.013440	.013098	.012398	.012040	.011678
.20	.020160	.019938	.019710	.019477	.019240	.017710	.016298	.016000	.015390	.015078	.014760
.50	.024750	.024798	.024840	.024878	.024910	.025000	.024938	.024910	.024840	.024798	.024750
.75	.017438	.017710	.017978	.018240	.018498	.019938	.021000	.021198	.021578	.021760	.021937
.80	.014760	.015077	.015390	.015697	.016000	.017710	.018997	.019240	.019710	.019937	.020160
.90	.008190	.008597	.009000	.009397	.009790	.012040	.013777	.014110	.014760	.015077	.015390
.95	.004297	.004750	.005197	.005640	.006077	.008597	.010560	.010937	.011677	.012040	.012397
1.00	0	.000497	.000990	.001477	.001960	.004750	.006938	.007360	.008190	.008598	.009000

Table 2

E(X/N)=PI(1-THETA)+(1-PI)PHI

P= .20

THETA

PHI	0	.05	.10	.15	.20	.50	.75	.80	.90	.95	1.00
0	.200000	.190000	.180000	.170000	.160000	.100000	.050000	.040000	.020000	.010000	0
.05	.240000	.230000	.220000	.210000	.200000	.140000	.090000	.080000	.060000	.050000	.040000
.10	.280000	.270000	.260000	.250000	.240000	.180000	.130000	.120000	.100000	.090000	.080000
.15	.320000	.310000	.300000	.290000	.280000	.220000	.170000	.160000	.140000	.130000	.120000
.20	.360000	.350000	.340000	.330000	.320000	.260000	.210000	.200000	.180000	.170000	.160000
.50	.600000	.590000	.580000	.570000	.560000	.500000	.450000	.440000	.420000	.410000	.400000
.75	.800000	.790000	.780000	.770000	.760000	.700000	.650000	.640000	.620000	.610000	.600000
.80	.840000	.830000	.820000	.810000	.800000	.740000	.690000	.680000	.660000	.650000	.640000
.90	.920000	.910000	.900000	.890000	.880000	.820000	.770000	.760000	.740000	.730000	.720000
.95	.960000	.950000	.940000	.930000	.920000	.860000	.810000	.800000	.780000	.770000	.760000
1.00	1.000000	.990000	.980000	.970000	.960000	.900000	.850000	.840000	.820000	.810000	.800000

VARIANCE OF X/N

0	.016000	.015390	.014760	.014110	.013440	.009000	.004750	.003840	.001960	.000990	0
.05	.018240	.017710	.017160	.016590	.016000	.012040	.008190	.007360	.005640	.004750	.003840
.10	.020160	.019710	.019240	.018750	.018240	.014760	.011310	.010560	.009000	.008190	.007360
.15	.021760	.021390	.021000	.020590	.020160	.017160	.014110	.013440	.012040	.011310	.010560
.20	.023040	.022750	.022440	.022110	.021760	.019240	.016590	.016000	.014760	.014110	.013440
.50	.024000	.024190	.024360	.024510	.024640	.025000	.024750	.024640	.024360	.024190	.024000
.75	.016000	.016590	.017160	.017710	.018240	.021000	.022750	.023040	.023560	.023790	.024000
.80	.013440	.014110	.014760	.015390	.016000	.019240	.021390	.021760	.022440	.022750	.023040
.90	.007360	.008190	.009000	.009790	.010560	.014760	.017710	.018240	.019240	.019710	.020160
.95	.003840	.004750	.005640	.006510	.007360	.012040	.015390	.016000	.017160	.017710	.018240
1.00	0	.000990	.001960	.002910	.003840	.009000	.012750	.013440	.014760	.015390	.016000

Table 3

E(X/N)=PI(1-THETA)+(1-PI)PHI

P= .30

THETA

PHI	0	.05	.10	.15	.20	.50	.75	.80	.90	.95	1.00
0	.300000	.285000	.270000	.255000	.240000	.150000	.075000	.060000	.030000	.015000	0
.05	.335000	.320000	.305000	.290000	.275000	.185000	.110000	.095000	.065000	.050000	.035000
.10	.370000	.355000	.340000	.325000	.310000	.220000	.145000	.130000	.100000	.085000	.070000
.15	.405000	.390000	.375000	.360000	.345000	.255000	.180000	.165000	.135000	.120000	.105000
.20	.440000	.425000	.410000	.395000	.380000	.290000	.215000	.200000	.170000	.155000	.140000
.50	.650000	.635000	.620000	.605000	.590000	.500000	.425000	.410000	.380000	.365000	.350000
.75	.825000	.810000	.795000	.780000	.765000	.675000	.600000	.585000	.555000	.540000	.525000
.80	.860000	.845000	.830000	.815000	.800000	.710000	.635000	.620000	.590000	.575000	.560000
.90	.930000	.915000	.900000	.885000	.870000	.780000	.705000	.690000	.660000	.645000	.630000
.95	.965000	.950000	.935000	.920000	.905000	.815000	.740000	.725000	.695000	.680000	.665000
1.00	1.000000	.985000	.970000	.955000	.940000	.850000	.775000	.760000	.730000	.715000	.700000

VARIANCE OF X/N

0	.021000	.020378	.019710	.018998	.018240	.012750	.006938	.005640	.002910	.001477	0
.05	.022278	.021760	.021197	.020590	.019937	.015078	.009790	.008597	.006077	.004750	.003377
.10	.023310	.022897	.022440	.021937	.021390	.017160	.012397	.011310	.009000	.007767	.006510
.15	.024097	.023790	.023437	.023040	.022597	.018997	.014760	.013777	.011677	.010560	.009397
.20	.024640	.024438	.024190	.023898	.023560	.020590	.016877	.016000	.014110	.013097	.012040
.50	.022750	.023177	.023560	.023898	.024190	.025000	.024437	.024190	.023560	.023177	.022750
.75	.014437	.015390	.016297	.017160	.017977	.021937	.024000	.024277	.024697	.024840	.024937
.80	.012040	.013097	.014110	.015077	.016000	.020590	.023177	.023560	.024190	.024437	.024640
.90	.006510	.007777	.009000	.010177	.011310	.017160	.020797	.021390	.022440	.022897	.023310
.95	.003377	.004750	.006077	.007360	.008597	.015077	.019240	.019937	.021197	.021760	.022277
1.00	0	.001477	.002910	.004298	.005640	.012750	.017438	.018240	.019710	.020377	.021000

Table 5

E(X/N)=PI(1-THETA)+(1-PI)PHI

P= .45

THETA

PHI	0	.05	.10	.15	.20	.50	.75	.80	.90	.95	1.00
0	.450000	.427500	.405000	.382500	.360000	.225000	.112500	.090000	.045000	.022500	0
.05	.477500	.455000	.432500	.410000	.387500	.252500	.140000	.117500	.072500	.050000	.027500
.10	.505000	.482500	.460000	.437500	.415000	.280000	.167500	.145000	.100000	.077500	.055000
.15	.532500	.510000	.487500	.465000	.442500	.307500	.195000	.172500	.127500	.105000	.082500
.20	.560000	.537500	.515000	.492500	.470000	.335000	.222500	.200000	.155000	.132500	.110000
.50	.725000	.702500	.680000	.657500	.635000	.500000	.387500	.365000	.320000	.297500	.275000
.75	.862500	.840000	.817500	.795000	.772500	.637500	.525000	.502500	.457500	.435000	.412500
.80	.890000	.867500	.845000	.822500	.800000	.665000	.552500	.530000	.485000	.462500	.440000
.90	.945000	.922500	.900000	.877500	.855000	.720000	.607500	.585000	.540000	.517500	.495000
.95	.972500	.950000	.927500	.905000	.882500	.747500	.635000	.612500	.567500	.545000	.522500
1.00	1.000000	.977500	.955000	.932500	.910000	.775000	.662500	.640000	.595000	.572500	.550000

VARIANCE OF X/N

0	.024750	.024474	.024097	.023619	.023040	.017437	.009984	.008190	.004297	.002199	0
.05	.024949	.024797	.024544	.024190	.023734	.018874	.012040	.010369	.006724	.004750	.002674
.10	.024997	.024969	.024840	.024609	.024277	.020160	.013944	.012397	.009000	.007149	.005197
.15	.024894	.024990	.024984	.024877	.024669	.021294	.015697	.014274	.011124	.009397	.007569
.20	.024640	.024859	.024977	.024994	.024910	.022277	.017299	.016000	.013097	.011494	.009790
.50	.019937	.020899	.021760	.022519	.023177	.025000	.023734	.023177	.021760	.020899	.019937
.75	.011859	.013440	.014919	.016297	.017574	.023109	.024937	.024999	.024819	.024577	.024234
.80	.009790	.011494	.013097	.014599	.016000	.022277	.024724	.024910	.024977	.024859	.024640
.90	.005197	.007149	.009000	.010749	.012397	.020160	.023844	.024277	.024840	.024969	.024997
.95	.002674	.004750	.006724	.008597	.010369	.018874	.023177	.023734	.024544	.024797	.024949
1.00	0	.002199	.004298	.006294	.008190	.017437	.022359	.023040	.024097	.024474	.024750

Table 6

E(X/N)=PI(1-THETA)+(1-PI)PHI

P= .55

THETA

PHI	0	.05	.10	.15	.20	.50	.75	.80	.90	.95	1.00
0	.550000	.522500	.495000	.467500	.440000	.275000	.137500	.110000	.055000	.027500	0
.05	.572500	.545000	.517500	.490000	.462500	.297500	.160000	.132500	.077500	.050000	.022500
.10	.595000	.567500	.540000	.512500	.485000	.320000	.182500	.155000	.100000	.072500	.045000
.15	.617500	.590000	.562500	.535000	.507500	.342500	.205000	.177500	.122500	.095000	.067500
.20	.640000	.612500	.585000	.557500	.530000	.365000	.227500	.200000	.145000	.117500	.090000
.50	.775000	.747500	.720000	.692500	.665000	.500000	.362500	.335000	.280000	.252500	.225000
.75	.887500	.860000	.832500	.805000	.777500	.612500	.475000	.447500	.392500	.365000	.337500
.80	.910000	.882500	.855000	.827500	.800000	.635000	.497500	.470000	.415000	.387500	.360000
.90	.955000	.927500	.900000	.872500	.845000	.680000	.542500	.515000	.460000	.432500	.405000
.95	.977500	.950000	.922500	.895000	.867500	.702500	.565000	.537500	.482500	.455000	.427500
1.00	1.000000	.972500	.945000	.917500	.890000	.725000	.587500	.560000	.505000	.477500	.450000

VARIANCE OF X/N

0	.024750	.024949	.024997	.024894	.024640	.019937	.011859	.009790	.005197	.002674	0
.05	.024474	.024797	.024969	.024998	.024859	.020899	.013440	.011494	.007149	.004750	.002199
.10	.024097	.024544	.024840	.024984	.024977	.021760	.014919	.013097	.009000	.006724	.004297
.15	.023619	.024190	.024609	.024877	.024994	.022519	.016297	.014599	.010749	.008597	.006294
.20	.023040	.023734	.024277	.024669	.024910	.023177	.017574	.016000	.012397	.010369	.008190
.50	.017437	.018874	.020160	.021294	.022277	.025000	.023109	.022277	.020160	.018874	.017437
.75	.009984	.012040	.013944	.015697	.017299	.023734	.024937	.024724	.023844	.023177	.022359
.80	.008190	.010369	.012397	.014274	.016000	.023177	.024999	.024910	.024277	.023734	.023040
.90	.004297	.006724	.009000	.011124	.013097	.021760	.024819	.024977	.024840	.024544	.024097
.95	.002199	.004750	.007149	.009397	.011494	.020899	.024577	.024859	.024969	.024797	.024474
1.00	0	.002674	.005198	.007569	.009790	.019937	.024234	.024640	.024997	.024949	.024750

Table 7

E(X/N)=PI(1-THETA)+(1-PI)PHI

P= .60

THETA

PHI	0	.05	.10	.15	.20	.50	.75	.80	.90	.95	1.00
0	.600000	.570000	.540000	.510000	.480000	.300000	.150000	.120000	.060000	.030000	0
.05	.620000	.590000	.560000	.530000	.500000	.320000	.170000	.140000	.080000	.050000	.020000
.10	.640000	.610000	.580000	.550000	.520000	.340000	.190000	.160000	.100000	.070000	.040000
.15	.660000	.630000	.600000	.570000	.540000	.360000	.210000	.180000	.120000	.090000	.060000
.20	.680000	.650000	.620000	.590000	.560000	.380000	.230000	.200000	.140000	.110000	.080000
.50	.800000	.770000	.740000	.710000	.680000	.500000	.350000	.320000	.260000	.230000	.200000
.75	.900000	.870000	.840000	.810000	.780000	.600000	.450000	.420000	.360000	.330000	.300000
.80	.920000	.890000	.860000	.830000	.800000	.620000	.470000	.440000	.380000	.350000	.320000
.90	.960000	.930000	.900000	.870000	.840000	.660000	.510000	.480000	.420000	.390000	.360000
.95	.980000	.950000	.920000	.890000	.860000	.680000	.530000	.500000	.440000	.410000	.380000
1.00	1.000000	.970000	.940000	.910000	.880000	.700000	.550000	.520000	.460000	.430000	.400000

VARIANCE OF X/N

0	.024000	.024510	.024840	.024990	.024960	.021000	.012750	.010560	.005640	.002910	0
.05	.023560	.024190	.024640	.024910	.025000	.021760	.014110	.012040	.007360	.004750	.001960
.10	.023040	.023790	.024360	.024750	.024960	.022440	.015390	.013440	.009000	.006510	.003840
.15	.022440	.023310	.024000	.024510	.024840	.023040	.016590	.014760	.010560	.008190	.005640
.20	.021760	.022750	.023560	.024190	.024640	.023560	.017710	.016000	.012040	.009790	.007360
.50	.016000	.017710	.019240	.020590	.021760	.025000	.022750	.021760	.019240	.017710	.016000
.75	.009000	.011310	.013440	.015390	.017160	.024000	.024750	.024360	.023040	.022110	.021000
.80	.007360	.009790	.012040	.014110	.016000	.023560	.024910	.024640	.023560	.022750	.021760
.90	.003840	.006510	.009000	.011310	.013440	.022440	.024990	.024960	.024360	.023790	.023040
.95	.001960	.004750	.007360	.009790	.012040	.021760	.024910	.025000	.024640	.024190	.023560
1.00	0	.002910	.005640	.008190	.010560	.021000	.024750	.024960	.024840	.024510	.024000

Table 8

E(X/N)=PI(1-THETA)+(1-PI)PHI

P= .80

THETA

PHI	0	.05	.10	.15	.20	.50	.75	.80	.90	.95	1.00
0	.800000	.760000	.720000	.680000	.640000	.400000	.200000	.160000	.080000	.040000	0
.05	.810000	.770000	.730000	.690000	.650000	.410000	.210000	.170000	.090000	.050000	.010000
.10	.820000	.780000	.740000	.700000	.660000	.420000	.220000	.180000	.100000	.060000	.020000
.15	.830000	.790000	.750000	.710000	.670000	.430000	.230000	.190000	.110000	.070000	.030000
.20	.840000	.800000	.760000	.720000	.680000	.440000	.240000	.200000	.120000	.080000	.040000
.50	.900000	.860000	.820000	.780000	.740000	.500000	.300000	.260000	.180000	.140000	.100000
.75	.950000	.910000	.870000	.830000	.790000	.550000	.350000	.310000	.230000	.190000	.150000
.80	.960000	.920000	.880000	.840000	.800000	.560000	.360000	.320000	.240000	.200000	.160000
.90	.980000	.940000	.900000	.860000	.820000	.580000	.380000	.340000	.260000	.220000	.180000
.95	.990000	.950000	.910000	.870000	.830000	.590000	.390000	.350000	.270000	.230000	.190000
1.00	1.000000	.960000	.920000	.880000	.840000	.600000	.400000	.360000	.280000	.240000	.200000

VARIANCE OF X/N

0	.016000	.018240	.020160	.021760	.023040	.024000	.016000	.013440	.007360	.003840	0
.05	.015390	.017710	.019710	.021390	.022750	.024190	.016590	.014110	.008190	.004750	.000990
.10	.014760	.017160	.019240	.021000	.022440	.024360	.017160	.014760	.009000	.005640	.001960
.15	.014110	.016590	.018750	.020590	.022110	.024510	.017710	.015390	.009790	.006510	.002910
.20	.013440	.016000	.018240	.020160	.021760	.024640	.018240	.016000	.010560	.007360	.003840
.50	.009000	.012040	.014760	.017160	.019240	.025000	.021000	.019240	.014760	.012040	.009000
.75	.004750	.008190	.011310	.014110	.016590	.024750	.022750	.021390	.017710	.015390	.012750
.80	.003840	.007360	.010560	.013440	.016000	.024640	.023040	.021760	.018240	.016000	.013440
.90	.001960	.005640	.009000	.012040	.014760	.024360	.023560	.022440	.019240	.017160	.014760
.95	.000990	.004750	.008190	.011310	.014110	.024190	.023790	.022750	.019710	.017710	.015390
1.00	0	.003840	.007360	.010560	.013440	.024000	.024000	.023040	.020160	.018240	.016000

Table 9

E(X/N)=PI(1-THETA)+(1-PI)PHI

P= .90

THETA

PHI	0	.05	.10	.15	.20	.50	.75	.80	.90	.95	1.00
0	.900000	.855000	.810000	.765000	.720000	.450000	.225000	.180000	.090000	.045000	0
.05	.905000	.860000	.815000	.770000	.725000	.455000	.230000	.185000	.095000	.050000	.005000
.10	.910000	.865000	.820000	.775000	.730000	.460000	.235000	.190000	.100000	.055000	.010000
.15	.915000	.870000	.825000	.780000	.735000	.465000	.240000	.195000	.105000	.060000	.015000
.20	.920000	.875000	.830000	.785000	.740000	.470000	.245000	.200000	.110000	.065000	.020000
.50	.950000	.905000	.860000	.815000	.770000	.500000	.275000	.230000	.140000	.095000	.050000
.75	.975000	.930000	.885000	.840000	.795000	.525000	.300000	.255000	.165000	.120000	.075000
.80	.980000	.935000	.890000	.845000	.800000	.530000	.305000	.260000	.170000	.125000	.080000
.90	.990000	.945000	.900000	.855000	.810000	.540000	.315000	.270000	.180000	.135000	.090000
.95	.995000	.950000	.905000	.860000	.815000	.545000	.320000	.275000	.185000	.140000	.095000
1.00	1.000000	.955000	.910000	.865000	.820000	.550000	.325000	.280000	.190000	.145000	.100000

VARIANCE OF X/N

0	.009000	.012397	.015390	.017977	.020160	.024750	.017438	.014760	.008190	.004297	0
.05	.008597	.012040	.015077	.017710	.019937	.024797	.017710	.015077	.008597	.004750	.000497
.10	.008190	.011677	.014760	.017437	.019710	.024840	.017977	.015390	.009000	.005197	.000990
.15	.007777	.011310	.014437	.017160	.019477	.024878	.018240	.015697	.009397	.005640	.001477
.20	.007360	.010937	.014110	.016877	.019240	.024910	.018497	.016000	.009790	.006077	.001960
.50	.004750	.008597	.012040	.015077	.017710	.025000	.019937	.017710	.012040	.008597	.004750
.75	.002437	.006510	.010177	.013440	.016297	.024937	.021000	.018997	.013777	.010560	.006937
.80	.001960	.006077	.009790	.013097	.016000	.024910	.021197	.019240	.014110	.010937	.007360
.90	.000990	.005197	.009000	.012397	.015390	.024840	.021577	.019710	.014760	.011677	.008190
.95	.000497	.004750	.008597	.012040	.015077	.024797	.021760	.019937	.015077	.012040	.008597
1.00	0	.004298	.008190	.011678	.014760	.024750	.021937	.020160	.015390	.012397	.009000