University of North Carolina Highway Safety Research Center

e-archives

access alcohol impairment bicycles child passenger safety crashes crosswalks data driver distraction driver behavior engineering evaluation graduated drivers licensing highways injury prevention medians motor vehicles occupant protection older drivers pedestrians public health research roadway design safety school travel seat belts sidewalks traffic transportation walking

Yosef Hochberg, and Donald W. Reinfurt. (1975). An Investigation of Safety Belt Usage and Effectiveness. Chapel Hill, NC: University of North Carolina Highway Safety Research Center.

UNC
Highway
Safety
Research
Center

AN INVESTIGATION OF SAFETY BELT USAGE AND EFFECTIVENESS

Yosef Hochberg

Donald W. Reinfurt

March 1975

~~December 1974~~

UNC/HSRC-75/3/1

UNC HSRC
University of North Carolina
Highway Safety Research Center
Chapel Hill, N.C. 27514

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT-HS-801-427		

4. Title and Subtitle	5. Report Date
AN INVESTIGATION OF SAFETY BELT USAGE AND EFFECTIVENESS	March 1975
	6. Performing Organization Code
7. Author(s) Yosef Hochberg, Donald W. Reinfurt	8. Performing Organization Report No.
9. Performing Organization Name and Address The Highway Safety Research Center University of North Carolina Chapel Hill, N.C. 27514	10. Work Unit No. (TRAIS)
	11. Contract or Grant No. DOT-HS-4-00897(SY085)
12. Sponsoring Agency Name and Address Department of Transportation National Highway Traffic Safety Administration 400 seventh street S.W. Washington D.C. 20590	13. Type of Report and Period Covered Interim Report July 1, 1974-Dec. 31, 1974
	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

This interim report contains primarily a theoretical discussion of problems of inference in studies on seat belt utilization and effectiveness.

Seat belt effectiveness in accidents is initially discussed from a population paramenter point of view. Three measures are introduced and their features explored. Several inferential problems are discussed and certain simplifications obtained.

This report includes some exploratory studies of the degree of misclassification errors involved in the police reports on belt usage and degree of injury along with some tools to partially resolve such problems.

Methodology is developed for investigating the bivariate injury distribution for belted and unbelted drivers to provide insight into the mechanism of the injury-reducing potential of belts in accidents.

The final report will include the analysis of a stratified random sample of 21,000 passenger car observations (taken in N.C. in October 1974) where belt utilization was recorded along with age(approximate), sex, and race of driver and vehicle license plate number.

Regarding effectiveness, the final report will have the analysis of frequency tables of the form: belt usage x injury, adjusted to match desired margins based on external information that will be collected.

17. Key Words	18. Distribution Statement
Safety belt effectiveness Utilization, Relative risk Ridit, Odds ratio	Document is available to the public through the National Technical Information Service, Springfield, Virginia 22151

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	68	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

★ U. S. GOVERNMENT PRINTING OFFICE : 1973 725-504/328

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (cont.)

LIST OF FIGURES

I. INTRODUCTION

The injury-reducing potential of safety belts has long been an issue for research and controversy. For a recent comprehensive review of the literature, see Griffin (1973).

There are several major problems which make a precise measurement of seat belt effectiveness very difficult. Some of these difficulties are discussed in Mela (1974a) where the current situation is summarized as follows:

> In practice, precise measurement of safety belt effectiveness is very difficult and extensive research and data collection would be needed to obtain any major improvements in this situation. Until definitive research is accomplished and standard methods of measuring effectiveness are adopted, there will continue to be a wide range of published estimates of safety belt effectiveness. For example, in the proceedings of a recent international meeting on automotive safety there appear several estimates of lap/shoulder belt effectiveness in reducing fatalities ranging from 31 percent to 80 percent.

It is the purpose of this study to identify the difficulties associated with a precise study of seat belt effectiveness; to develop techniques for partially or fully resolving these difficulties; and to arrive at a better understanding of the injury-reducing potential of seat belts.

In this report, we first identify the difficulties associated with studies of seat belt effectiveness. These difficulties are then examined individually in subsequent chapters.

Now, effectiveness of seat belts in reducing injuries in automobile accidents is usually inferred from contingency tables displaying belt usage and degree of injury as provided by the police report. A variety of problems regarding the inferential statistical techniques to be used and the proper interpretations are associated with such data. The following is a list of some of the troublesome issues involved in most of the studies devoted to this subject:

(1) Loose definitions (or no definitions) of what is meant by "seat belt effectiveness in injury reduction."

(2) Non-invariance of the uni-dimensional parameter ordinarily used as a measure for "seat belt effectiveness" to different breakdowns of the injury scale (i.e., various injury definitions for "severe," "moderate," etc., are likely to produce different assessments of the effectiveness of belts as measured by the ordinary relative risk parameter).

(3) Bias resulting from using data based on the police report. Misclassification errors in reporting seat belt usage by the policeman might seriously mislead potential inference on effectiveness, especially if such errors are confounded with degree of injury. (A higher "positive" bias in classifying belt usage for low injury than for serious injury will tend to increase most measures of association between belt usage and injury reduction and vice-versa).

Similarly, parallel misclassification errors probably occur on the injury scale. Such errors might be very severe (regarding their effect on seat belt effectiveness measured by any reasonable measure of association) when interacting with the levels of belt usage. At the extreme, one should consider the possibility that in a "large" number of accidents the belted driver suffers an internal injury caused by the restraint which is not revealed to the police investigator.

(4) Interactions and confounding in the data. Ignoring the interactive nature and/or confounding in the data of the effects of factors such as age, type of car, sex, etc., might bring about misleading interpretations of the results concerning the association between seat belt usage and injury reduction.

(5) Discussion of sample structure and its possible effects on inference. In most works, estimates based on samples generally lack estimates of standard errors or corresponding confidence intervals. These issues will be considered in detail in the sections following.

(6) Superficial estimates of belt effectiveness. Questions such as, "What proportion of moderate injuries when using the belt would have resulted in more severe injuries had the belt not been used?" or "What is the proportion of belted drivers injured in accidents which should be attributed to the use of the belt?" cannot be answered properly by the usual approach of comparing marginal frequencies on the injury scale for belted and unbelted drivers.

II. THE CONCEPT OF "SEAT BELT EFFECTIVENESS" AND SOME USEFUL MEASURES

Consider a given population of accidents. Let F(x) denote the cumulative distribution function of injury in an imaginary setup where all occupants are wearing their belts. Similarly, let G(x) denote the corresponding function when all occupants do not wear their belts. Let C(x) and D(x) denote the actual conditional injury distributions for belted and unbelted occupants, respectively.

It is generally accepted that measures for "seat belt effectiveness" are to be based on a comparison of F(x) with G(x).

> To measure safety belt injury-reducing effectiveness involves comparing the estimated numbers of injuries to be expected in a population of drivers and passengers using safety belts with the estimated numbers of injuries in the same or a similar population without safety belts. The overall effect of the belts depends upon the probability of belt use in a crash and upon the injury-reducing capability of the belts if worn in a crash." (Mela, 1974a.)

Since the actual belted population might be quite different from the unbelted one with respect to several important factors, comparisons based on C(x) and D(x) are not meaningful. This point is well emphasized in Griffin (1973) and Mela (1974a).

In this section, we confine the discussion to a well-defined sub-population in which the conditional distribution of injury among belted occupants is the same as the injury distribution obtained if all occupants were wearing their belts; likewise for unbelted occupants. Thus, F(x) = C(x) and G(x) = D(x), for all x. In this simplified setup, we now explore several definitions and corresponding measures (population parameters) for "seat belt effectiveness."

To date, the general inclination has been to use a uni-dimensional parameter as a measure to study the "injury reduction potential" or "effectiveness" of the belts. Mostly this parameter was the so-called relative risk, which for a given level of injury j, is defined as the ratio:

$$R_j = \frac{\Pr(\text{at least a } j \text{ level of injury} \mid \text{no belt})}{\Pr(\text{at least a } j \text{ level of injury} \mid \text{belt})}$$

Note that R_j is a parameter depending on j and specific to some well-defined population of accidents. With this understanding, $100\ (R_j - 1)$ can be thought of as the percentage reduction in injuries of at least level j if all passengers were using their belts compared with the situation where no one uses belts. Alternatively, $100\ (R_j - 1)$ can be interpreted as the reduction of the individual's risk of level at least j obtained by using his belt.

Actually, this is a multi-dimensional measure, R_j, $j = 2, 3, \ldots, J$. Surprisingly enough, most writers have not sufficiently emphasized the dependence of R_j on j. This simple fact is probably one of the main reasons for the heterogeneous set of relative-risk measures that appears in the literature (see Mela, 1974a). In different studies, "same" levels j are not equally defined, resulting in different R's for apparently the same levels of j. Clearly this is so even when $J = 2$; i.e., the injury scale is divided into two groups.

A number of other measures for seat belt effectiveness can be defined. It is the purpose of this section to study carefully the features of three such measures. These measures will first be introduced as population parameters and some natural properties (such as invariance) will be examined.

Let π_j^B denote the conditional probability of a belted driver sustaining an injury of level j given an accident, and similarly define the conditional risks for the unbelted driver, π_j^U, j = 1, ..., J. Let

$$\gamma_i^B = \sum_{j=1}^{i-1} \pi_j^B \quad , \qquad i = 2, 3, \ldots, J \text{ and } \gamma_1^B \equiv 0$$

$$\gamma_i^U = \sum_{j=1}^{i-1} \pi_j^U \quad , \qquad i = 2, 3, \ldots, J \text{ and } \gamma_1^U \equiv 0$$

The measures of belt effectiveness include the following:

(i) The probability that an unbelted occupant gets more severely injured than a belted occupant when undergoing the "same" accident. This measure is the ridit introduced by Bross (1958). We denote it by T, where:

$$T = \sum_{i=1}^{J} [\gamma_i^B + \tfrac{1}{2}\pi_i^B]\, \pi_i^U \; .$$

This measure is independent of j.

(ii) The relative risk. This measure is usually used when J = 2, and is defined as a vector of J - 1 values.

$$R_j = \frac{1 - \gamma_j^U}{1 - \gamma_j^B} \; ; \; j = 2, \ldots, J \; .$$

(iii) The odds ratio. As with the second measure, it depends on j. It is given by

$$D_j = R_j \frac{\gamma_j^B}{\gamma_j^U}, \quad j = 2, \ldots, J$$

and is interpreted, for a given j, as the ratio (belted/unbelted) of the odds of a less severe injury than j versus a more severe one.

These three measures are mathematically independent, i.e., no one is an explicit function of any other. This is so even in the simplest case when J = 2, i.e., the injury scale has just two levels. Note that the ridit measure, being an overall measure across all levels j of injury, is a "final product" as far as risk is concerned. The other measures are not "final" or "complete" in the sense that they do not summarize the overall risk comparison between belted and unbelted occupants but rather are confined to the upper levels of risk. This is so even when J = 2. Since in most cases π_1^B, and π_1^U will be "large," this means that T will usually not be far from 0.5. (We will see that in most cases T will lie within the range of 0.5 - 0.6.) The other two measures will have a much wider range.

In some studies the measure T is replaced by

$$T^* = \frac{T}{1-T}$$

where T* is referred to as "the overall odds ratio risk."

Various studies (see Griffin, 1973, and Campbell et al., 1974) indicate that one usually finds π_j^U / π_j^B increasing with j. This seems to be basic to the nature of the belt effect. One might then wish to have a measure that reflects this feature. It is shown in Appendix A that the measure R_j increases with j under such conditions. The odds ratio, D_j, is not necessarily increasing with j as shown by the hypothetical example in Table 1.

Table 1. Comparison of R and D for hypothetical example

j	π_j^B	π_j^U	R_j	D_j
1	.500	.300	---	---
2	.300	.400	1.400	2.333
3	.100	.140	1.500	1.714
4	.060	.090	1.600	1.714
5	.040	.070	1.750	1.806

where, e.g.,

$$R_3 = \frac{1-(\pi_1^U + \pi_2^U)}{1-(\pi_1^B + \pi_2^B)} = \frac{1-(.30+.40)}{1-(.50+.30)} = 1.50$$

$$D_3 = (1.50)\left[\frac{.50+.30}{.30+.40}\right] = 1.714$$

It is worthwhile to comment that, for various situations where studies of association in 2 x 2 tables are concerned, the measure D has an invariance property which is not shared by any other measure. We are referring to Plackett (1965) and Mosteller (1968) where the use of this measure is advocated. This is now discussed in the following example.

Consider three hypothetical populations of accidents A, B, C, with the following 2 x 2 tables:

Table 2. Belt usage by injury distributions for three populations of accidents

	A		B		C	
	Belt	No Belt	Belt	No Belt	Belt	No Belt
No injury	2	3	4	30	12	90
Injury	1	4	2	40	1	20

Note that Table B was obtained from Table A by multiplying the first column by 2 and the second one by 10. Table C was obtained from Table B by multiplying the first row by 3 and the second row by .5. We now give four measures of association between belt usage and reduction of injury for each of these three tables (see Table 3).

Table 3. Four measures of association for accident populations A, B, C

Table	A	B	C
ρ	.350	.290	.250
T	.619	.619	.552
R	1.555	1.555	1.128
D	2.667	2.667	2.667

$$-1 \leq \rho \leq 1$$

$$0 \leq T \leq 1$$

$$0 \leq R \leq \infty$$

$$0 \leq D \leq \infty$$

The measure ρ is the correlation coefficient obtained when fitting a bivariate normal distribution to the tables (see Mosteller, 1968). The only measure which is invariant under both multiplication of columns (i.e., change in the marginal distribution of belt usage) and/or rows (i.e., a change in the marginal distribution of injury) is the odds ratio.

However, invariance of the measure under changes in the injury distribution is not necessarily a desired feature in our case. Multiplication of rows as above may change non-proportionally the conditional distributions of injury for belted and unbelted accidents. In the present application, the measures should only be restrained to have invariance under column multiplications (i.e., with respect to belt status). The measure ρ is not invariant under either column or row multiplication and, therefore, omitted from further consideration. All the other three measures will be examined further.

Clearly, if one introduces a more specific functional relation between the conditional distribution of injury for belted drivers and that for unbelted drivers (e.g., some location translation model), usually much more can be said about the performance of the three measures given above as summary statistics for belt effectiveness. However, any overall specification of such a model in accident data would be rather arbitrary. Such models probably do exist but are highly dependent in structure on the specific population. As an example, we consider in Appendix B the case when the conditional distributions of injury (for belted and unbelted occupants) are taken to be continuous and members of the exponential family. The general shape of this distribution fits automotive accident data. The effect of the belt is assumed to be such as to produce an increase in the mode of the probability density function (see Figure 1).

In this study, the following situations are considered:

1. In two different populations, $\theta_1^U = \theta_2^U$ but $d_1 \neq d_2$.

2. In two different populations, $\theta_1^U \neq \theta_2^U$ but $d_1 = d_2$.

Under the exponential model, one would require such a measure that in Case 1 the difference between the two populations is only a function of $d_1 - d_2$, and in Case 2 the measures for both populations should be the same. Among the three measures, only the relative risk parameter satisfies these conditions. This is shown in Appendix B.

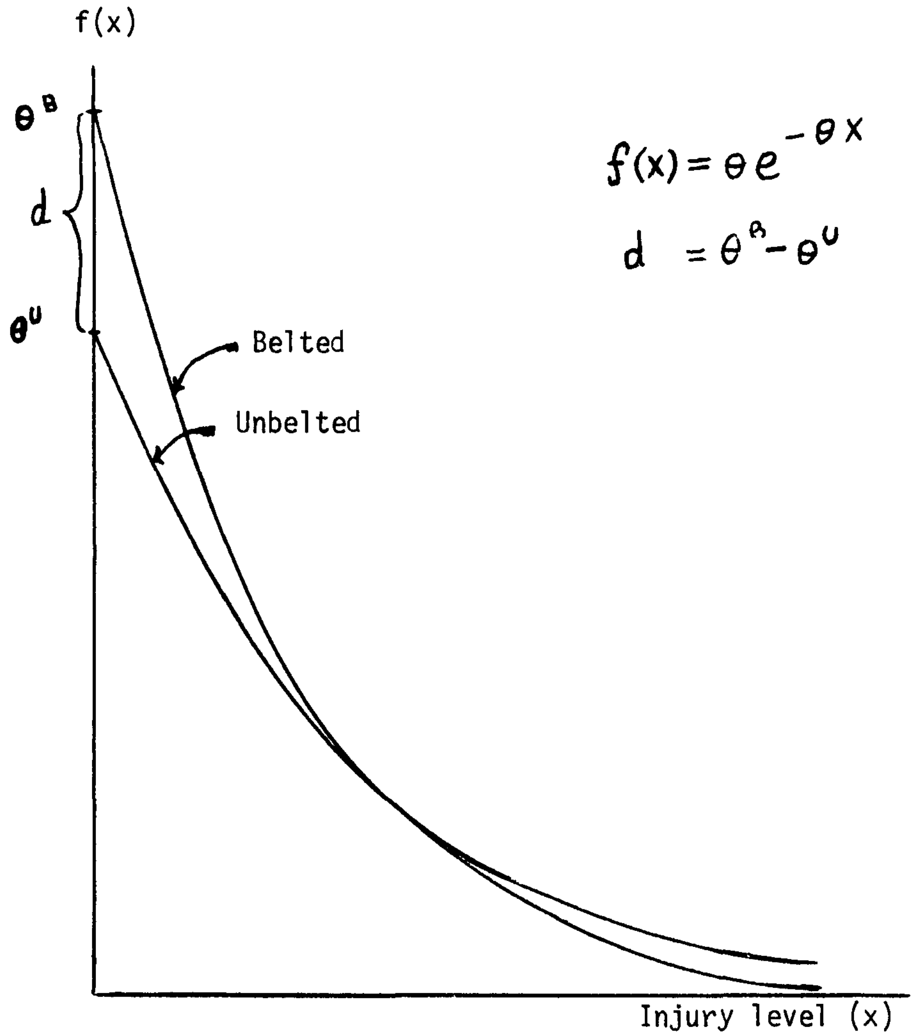

Figure 1. Probability density functions for belted and unbelted occupants.

III. INFERENCE PROBLEMS

The Dependence of Seat Belt Effectiveness on Other Factors

Effectiveness of seat belts has been studied using different sources of data which have produced a wide range of published estimates of this parameter. (See, for example, Griffin, 1973, and the *Proceedings of the Third International Congress on Automotive Safety*, 1974.)

Some of the reasons for this heterogeneity have to do with the fact that usually the effect of belt wearing on injury reduction is probably both interacting with and confounded with other factors. Some of these factors are: vehicle size and model year, age and sex of occupants, and accident configuration, including speed of vehicle.

Mela (1974a) addresses this point in the following:

> 1. Belt users are not involved in the same kinds and severities of accidents as non-users.
>
> In general, belt usage is greater in high-speed rural driving than under urban driving conditions; accident statistics show much greater usage on the interstate system than on local roads. Belt usage is much lower at night, particularly during the midnight to 6:00 a.m. "drinking hours" than during the noon to 6:00 p.m. period. In general, the more severe the crash, the less the likelihood of belt use.
>
> Because of these differences, comparisons of injury rates for belt-users and non-users may be quite misleading unless some type of statistical standardization or normalization is carried out. A few of the published studies have included this, but most have not.
>
> 2. Estimates of safety belt effectiveness are affected by the manner of selecting the population of accident victims. The differences in belt effectiveness estimates in various studies are to some extent attributable to the population differences. For example, some studies are done with sets of accidents in which

> the criterion for inclusion is that at least one injury shall have occurred in the accident. In other cases, other studies include property damage accidents for which the reporting criterion is an estimated damage level.

Now suppose one considers d factors with levels $I_1, I_2, \ldots, I_d$. Let $\underset{\sim}{i} = (i_1, \ldots, i_d)$ index a specific combination of levels and denote by $w_{\underset{\sim}{i}}$ the proportion of this cell among all accidents. Viewing the $I_1 \times I_2 \ldots \times I_d$ sub-populations as post-stratified by the sample (with corresponding weights $w_{\underset{\sim}{i}}$), one proceeds to obtain

$$p(B) = \sum_{\underset{\sim}{i}} w_{\underset{\sim}{i}} \, p_{\underset{\sim}{i}}(B) \qquad p(1,B) = \sum_{\underset{\sim}{i}} w_{\underset{\sim}{i}} \, p_{\underset{\sim}{i}}(1,B)$$

$$p(U) = \sum_{\underset{\sim}{i}} w_{\underset{\sim}{i}} \, p_{\underset{\sim}{i}}(U) \qquad p(1,U) = \sum_{\underset{\sim}{i}} w_{\underset{\sim}{i}} \, p_{\underset{\sim}{i}}(1,U)$$

where $p_{\underset{\sim}{i}}(B)$, $p_{\underset{\sim}{i}}(1,B)$ are the sample proportions of belted occupants and of uninjured belted occupants, respectively, in the sample of accidents under study. Similarly, $p_{\underset{\sim}{i}}(U)$, $p_{\underset{\sim}{i}}(1,U)$ refer to uninjured occupants.

Then the standardized rates based on the assumed post-stratification are given by

$$p(1|B) = \frac{p(1,B)}{p(B)} \qquad p(1|U) = \frac{p(1,U)}{p(U)}$$

from which standardized measures of belt effectiveness can be obtained.

A major risk in such a standardization involves sampling procedures where sample post-stratification weights $w_{\underset{\sim}{i}}$ are biased estimates of the true population parameters. In such cases, one should work at getting better estimates of the true sub-population weights. Occasionally, it is difficult to find such a set of estimates. Rather, one usually finds it feasible to provide educated guesses of the marginal distributions of the factors considered. In such a case, the multi-dimensional table of $I_1 \times I_2 \times \ldots I_d \times 2 \times 2$ cells (i.e., factors x usage x injury) must be adjusted to pertain to the desired d margins. This is accomplished by using the Iterative Proportional Fitting Procedure (IPFP) (see Fienberg, 1970).

In this study we will attempt various standardized comparisons between injury rates for belted and unbelted occupants based on the post-stratification approach of Cochran (1954) and on the Mantel-Haenszel method (see Mantel, 1963).

There is another major risk involved when using such standardized comparisons which might be designated as "over-standardization" and is illustrated by the following example. In Campbell et al. (1974), injuries and seat belt effectiveness specific to five market class groups were studied. Since the factor, "pattern of accidents" (type speed, etc.), is confounded with market class, direct comparisons among the five groups might have been misleading. The authors undertook a "standardized comparison" in which weighted injury distributions for each of the five groups were obtained by adjusting the crash pattern distribution in each group to match that of the reference population of crashes. The problem is then that of interpreting effects of factors that are confounded with other factors. We cite from Campbell et al. (1974) to clarify the difficulty of "over-standardization" in this case.

> Each group is compared to the reference group as if it had the reference group's mix of crashes when in fact it has its "own." Normally the differences are small, but a few cars have a rather distinctive pattern of crashes.

> Then the question is whether to compare the car on the basis of a "standard" array of crashes, or on the basis of the array it actually manifested.

This difficulty is almost unavoidable when dealing with such highly confounded effects as those involved in studying belt-effectiveness.

In addition to a variety of standardized measures, we will also use statistical regression models for categorical data (Grizzle et al., 1969 Goodman, 1971). By using such models, we can summarize the variability of the various $I_1 \times I_2 \times \ldots \times I_d$ measures in a compact form, get estimates of effects and their standard errors, and have, to some degree, interpretable simultaneous assessments of significance or non-significance of sets of effects.

Another problem is that of inference from a given data set (N.C., 1974, for example) to a larger reference population (U.S.A., 1974). As mentioned above, accident data might vary from one state to another with respect to distributions of factors confounded with belt usage. This seems to limit greatly the extent to which projections from one source of data to a larger population of a possibly "different" distribution can be made. Quoting from Griffin (1973):

> Any attempt to extrapolate from any one of these studies, or from any combination of these studies, to the nation's savings due to lap belts is really a guess. There is no mathematical formula which can determine how effective lap belts are nationwide. Until such time as the nation is willing to invest in a random sampling of accidents across the country, educated guessing will be the domain of investigators in the field of highway safety.

Part of the current study will involve an investigation of a statistically-sound method for making such desired inferences.

Sample Invariance of Measures of Association; Effect of Sample Structure on Inference; Standard Errors

Next consider the problem of the "sample-invariance" of the measures. Even though a distinction has been made between the concept of a "population-invariance" and "sample invariance" of the measures, an additional emphasis is in order. If two different studies of belt effectiveness in a given population are undertaken with different breakdowns of the injury scale, then the resulting parametric measures for effectiveness of belts may well indicate different numerical values. There is no conceptual difficulty associated with a heterogeneous set of measures so long as they are accompanied by interpretations according to the corresponding breakdown of the injury scale. However, if the same conceptual breakdown of the injury scale is used and then the three measures produce the same parametric quantities in both studies, there is still the problem of "sample invariance."

It is reasonable to look for a statistic, a function of the sample frequencies, which is invariant under some transformations of the sample data. From previously, it is clear that the only measure which is invariant under both multiplications of rows and columns of our two-way table is D, the odds ratio. A related type of invariance also shared only by the odds ratio is illustrated by an example. Suppose that a sample of 50 occupants involved in accidents is classified in the following table.

Table 4A. Hypothetical example

	Belt	No Belt	Total
No Injury	30	10	40
Injury	5	5	10
Total	35	15	50

We are now concerned with the problem that the very same sample could have been differently classified by someone who had different ideas about what injured and uninjured meant. Supose, thus, that a more careful observer would have classified eight more cases as injured while preserving the association between belt usage and injury. This means that to obtain Table 4B we must use an iterative proportional fitting procedure (see Mosteller, 1968) to adjust the original table to fit the new margins (see Table 4B) while retaining the interaction structure of the original table.

Table 4B. Required margins

	Belt	No Belt	Total
No Injury			32
Injury			18
Total	35	15	50

After four steps we obtain Table 4C,

Table 4C. Adjusted table

	Belt	No Belt	Total
No Injury	25.14	6.86	32
Injury	9.86	8.14	18
Total	35	15	50

or, when approximating by integers, we get Table 4D.

Table 4D. Approximated integer table

	Belt	No Belt	Total
No Injury	25	7	32
Injury	10	8	18
Total	35	15	50

The three measures for belt effectiveness for these three tables are given in Table 4E.

Table 4E. Measures of association

Table No.	4A	4C	4D
T	.595	.631	.624
R	1.286	1.563	1.531
D	3.000	≈3.000	2.857

The direction of change in the three measures is the same but the odds ratio (D) is best at preserving the association.

Next we consider the problem of the effect the sampling scheme has on inference. Since belt effectiveness is often studied from two-way tables of (Belt Usage) x (Injury), a key issue concerns the way in which the inference depends on the sampling scheme. The underlying structure of a sample producing such two-way tables might be any one of an infinity of possible structures! The usual, well-known structures discussed in text books (i.e., only total sampling fixed, one margin fixed, both margins fixed) are only a few examples. Other examples include a sample structure where sampling is continued until a pre-specified number N°_{11}, for example, is reached in the upper left entry, etc. Most suitable for accident data seem to be the so-called "conditional Poisson" models where events (generated by a Poisson process [in time] of a given intensity) are classified according to the structure of a two-way table. It is well known that the conditional distributions of the frequencies N_{ij} in the table, when the overall sample, one margin, or both margins are fixed are the multinomial and hypergeometric distributions. However, if another stopping rule is used, the resulting distributions are different.

Must one really consider each sampling structure by itself and study the appropriate distribution? Samples are usually quite large in accident research. In most or all sampling schemes, one may base inference on the large sample distribution of the N_{ij}'s which is multivariate normal by the central limit theorem for any of the multinomials, hypergeometric, and conditional Poisson distributions from which we sample (see Haberman, 1970).

However, this does not completely specify our table's distribution even for large samples and one has to select carefully a proper variance-covariance matrix for any given table. Two basic structures will be used in this study corresponding to: (1) an overall multinomial, i.e., only the total sampling is fixed, (2) independent multinomials for the respective levels of belt usage, i.e., the marginal totals for levels of belt usage are fixed. Then, each of the three measures can be expressed as a compound exponential-logarithmic function of the N_{ij}'s (see Forthofer and Koch, 1973). Expressions for the standard errors of the measures along with their estimates will be obtained as in Forthofer and Koch (1973).

IV. ERRORS OF MISCLASSIFICATION IN 2 x 2 CONTINGENCY TABLES

We are concerned with errors of systematic bias in 2 x 2 tables classifying occupants involved in accidents by belt usage and injury sustained. The effect of misclassification errors of a random nature (response errors) with no systematic bias on measures of association has been studied elsewhere but is of minor interest in our data.

Consider a 2 x 2 table of proportions of belt usage by injury level. Let $\pi(1,B)$ denote the population proportion of uninjured belted occupants in a given population of accidents. Similarly, we use $\pi(2,B)$ for "belted-injured", $\pi(1,U)$ for "unbelted-uninjured" and $\pi(2,U)$ for "unbelted-injured". The use of $\pi(A|B)$ indicates the conditional probability of A given B, e.g., $\pi(1|B)$ is the population proportion of non-injury among belted occupants.

Theoretically there is a total of 12 independent misclassification errors that might take place when classifying individuals into a 2 x 2 table. For example, for any occupant there is the error of classifying "belted-injured" as "unbelted-injured" or as "belted-uninjured", etc. Suppose we denote these error probabilities by $\alpha_{ij,mn}$, $(i,j) \neq (m,n)$ $i,j,m,n = 1, 2$, where, for example, $\alpha_{21,12}$ refers to the error of classifying a "belted-injured" case as "uninjured-unbelted", etc. If $\gamma(1,B)$, $\gamma(2,B)$, $\gamma(1,U)$, $\gamma(2,U)$ represent the proportions in the 2 x 2 tables of (belt usage) x (injury) with the errors incorporated, one may relate the γ's to the α's and π's by easily derivable but complicated expressions.

To study the effect of misclassification errors on the effectiveness measures, one could write down expressions for T,R,D in terms of the α's and π's through the γ's. However, even though the algebra is simple, the expressions become quite involved and do not simplify reasonably. There seems to be no feasible way of studying the effects of these errors algebraically unless one makes some extremely simplifying assumptions. For example, one could fairly easily obtain simple expressions relating the percentages of errors in T,R, and D to the α's and π's under each of the following two assumptions:

(i) No error in classifying belt usage (both when injured and when uninjured);

(ii) No error in classifying injury <u>and</u> no error in classifying belt usage when injured.

The first case is the simplest case. This is the one usually considered (see Fleiss (1973)). The second case is somewhat more complicated and is the case considered by means of a numerical example by Mela (1974a).

These simplifying cases may, of course, be used to demonstrate (as in Mela (1974a)) just how severe misclassification errors might be on a given measure of "belt effectiveness". They cannot, however, be accepted as pertaining to the situation prevailing with accident data. On the other hand, a "saturated model" including all 12 independent parameters is too cumbersome. The dimensionality of errors of misclassification in our 2 x 2 tables can reasonably be reduced to 6 parameters (instead of 12) based on the following assumptions:

(i) In no case will an uninjured person be classified as injured. (Note that this does not preclude errors of classifying injured people as uninjured both when belted or when unbelted).

(ii) Probabilities of errors in the two characteristics (e.g., the probability of classifying "belted-injured" as "unbelted-uninjured") are given by multiplying the corresponding conditional error probabilities (e.g., the probability of classifying "belted-injured" as "unbelted-uninjured" is given by multiplying the probability of classifying as "unbelted-injured" by the probability of classifying as "belted-uninjured" when actually "belted-injured").

The six error parameters obtained under these assumptions are:

$$\alpha(1|2;B),\ \alpha(1|2;U),\ \alpha(B|U;1),\ \alpha(U|B;1),\ \alpha(B|U;2),\ \alpha(U|B;2)$$

Where, for example, $\alpha(1|2;B)$ represents the probability of classifying as "uninjured-belted" an "injured-belted" occupant, and $\alpha(U|B;2)$ refers to the error of classifying an "injured-belted" occupant as "injured-unbelted", etc. With this notation and assumptions (i), (ii), we now have the following expressions for the γ's:

$$\gamma(1,B) = \pi(1,B)\left[1-\alpha(U|B;1)\right] + \pi(1,U)\,\alpha(B|U;1) + \pi(2,B)\,\alpha(1|2;B) + \pi(2,U)\,\alpha(1|2;U)\,\alpha(B|U;2)$$

$$\gamma(2,B) = \pi(2,U)\,\alpha(B|U;2) + \pi(2,B)\left[1-\alpha(1|2;B)-\alpha(U|B;2)-\alpha(1|2;B)\,\alpha(U|B;2)\right]$$

$$\gamma(1,U) = \pi(1,U)\left[1-\alpha(B|U;1)\right] + \pi(1,B)\,\alpha(U|B;1) + \pi(2,U)\,\alpha(1|2;U) + \pi(2,B)\,\alpha(1|2;B)\,\alpha(U|B;2)$$

$$\gamma(2,U) = 1-\gamma(1,B)-\gamma(2,B)-\gamma(1,U)$$

We now define the conditional probabilities

$$\gamma(1|B) = \frac{\gamma(1,B)}{\gamma(1,B) + \gamma(2,B)} \qquad \gamma(1|U) = \frac{\gamma(1,U)}{\gamma(1,U) + \gamma(2,U)}$$

It is possible to evaluate, for any initial table of π's, the true measures of effectiveness T( ), R( ), D( ) where

$$T(\pi) = 1/2 \left[1 + \pi(1|B) + \pi(1|U)\right]$$

$$R(\pi) = \frac{1 - \pi(1|U)}{1 - \pi(1|B)} \qquad D(\pi) = R(\pi) \frac{\pi(1|B)}{\pi(1|U)}$$

That is, for any given 6 error probabilities, calculate the resulting table (γ), and the corresponding T(γ), R(γ), D(γ), and then compare them to the "true" measures.

For the following simulation study, injury is dichotomized as follows:

A. Killed vs. not killed

B. Serious plus fatal vs. other

C. Any injury vs. property damage only.

One must, however, realize that the measure T is not a fair competitor to the other two when studying robustness to errors of misclassification. This is because of the basic distinction between T and either R or D discussed earlier. The measure T gives a 'final' comparison between the belted and unbelted driver, i.e., T automatically gives lower weights to some differences $\pi(1|B) - \pi(1|U)$ for larger values of $\pi(1|B)$ and $\pi(1|U)$. For this reason the only meaningful comparison is between R and D.

In all three cases, the range of values for $\alpha(B|U;1)$, $\alpha(U|B;1)$, and $\pi(B)$ chosen are as follows:

$$\alpha(B|U;1) = .01,\ .15,\ .30$$

$$\alpha(U|B;1) = .01,\ .15,\ .30$$

$$\pi(B) = .01,\ .15,\ .30$$

Values were allocated to the other parameters as follows:

Case A. $[\pi(1|B), \pi(1|U)] = [.9975, .9950]$

$\alpha(1|2;B) = .01,\ .10$ $\qquad$ $\alpha(B|U;2) = .00,\ .05$

$\alpha(1|2;U) = .01,\ .10$ $\qquad$ $\alpha(U|B;2) = .00,\ .05$

Case B. $[\pi(1|B), \pi(1|U)] = [.97, .95]$

$\alpha(1|2;B) = .01,\ .20$ $\qquad$ $\alpha(B|U;2) = .01,\ .10$

$\alpha(1|2,U) = .01,\ .20$ $\qquad$ $\alpha(U|B;2) = .01,\ .10$

Case C. $[\pi(1|B), \pi(1|U)] = [.84, .78]$

$\alpha(1|2;B) = .01,\ .30$ $\qquad$ $\alpha(B|U;2) = .01,\ .15$

$\alpha(1|2;U) = .01,\ .30$ $\qquad$ $\alpha(U|B;2) = .01,\ .15$.

Thus, for each of the three cases there were 432 simulated cases. The reported quantities for each case are the percentage changes

$$\mathring{T} = 100\left[\frac{T(\gamma)}{T(\pi)} - 1\right] \qquad \mathring{R} = 100\left[\frac{R(\gamma)}{R(\pi)} - 1\right] \qquad \mathring{D} = 100\left[\frac{D(\gamma)}{D(\pi)} - 1\right]$$

Due to space limitations, only about a sixth of the combinations for Case A are reported here. Specifically for $\pi(B)=0.15$, $\pi(1|B)=0.9975$, $\pi(1|U)=0.995$, we have the following results (see Table 5).

Table 5. Errors in measures due to misclassification errors.

$\alpha(1\|2;A)$	$\alpha(1\|2;U)$	$\alpha(B\|U;1)$	$\alpha(U\|B;1)$	$\alpha(B\|U;2)$	$\alpha(U\|B;2)$	$\mathring{T}$	$\mathring{R}$	$\mathring{D}$
.01	.01	.01	.01	.00	.00	.01	5.50	5.52
					.05	.03	11.66	11.69
				.05	.00	-.15	-36.22	-36.32
					.05	-.13	-33.78	-33.87
			.15	.00	.00	-.04	-10.74	-10.82
					.05	-.02	-5.59	-5.61
				.05	.00	-.21	-46.06	-46.18
					.05	-.20	-44.00	-44.11
		.15	.01	.00	.00	.20	115.34	115.76
					.05	.20	127.90	128.37
				.05	.00	.09	30.10	30.21
					.05	.10	35.08	35.21
			.15	.00	.00	.17	93.37	93.69
					.05	.18	104.65	105.01
				.05	.00	.06	16.84	16.90
					.05	.07	21.30	21.39
	.10	.01	.01	.00	.00	-.03	-4.09	-4.12
					.05	-.02	-1.55	-1.53
				.05	.00	-.19	-42.63	-42.74
					.05	-.18	-40.40	-40.51
			.15	.00	.00	-.08	-18.90	-18.97
					.05	-.06	-14.14	-14.19
				.05	.00	-.26	-51.48	-51.60
					.05	-.25	-49.60	-49.72
		.15	.01	.00	.00	.14	95.78	96.04
					.05	.15	107.28	107.60
				.05	.00	.03	17.03	17.07
					.05	.04	21.58	21.62
			.15	.00	.00	.12	75.79	75.99
					.05	.13	86.13	86.36
				.05	.00	.00	5.09	5.10
					.05	.01	9.17	9.19
.10	.01	.01	.01	.00	.00	.03	16.05	16.10
					.05	.05	24.14	24.20
				.05	.00	-.13	-32.31	-32.39
					.05	-.11	-29.35	-29.43
			.15	.00	.00	-.01	-1.87	-1.88
					.05	.01	4.96	4.97
				.05	.00	-.19	-42.75	-42.86
					.05	-.17	-40.25	-40.36

Table 5. Errors in measures due to misclassification errors. (Con't)

$\alpha(1\|2;B)$	$\alpha(1\|2;U)$	$\alpha(B\|U;1)$	$\alpha(U\|B;1)$	$\alpha(B\|U;2)$	$\alpha(U\|B;2)$	$\check{T}$	$\check{R}$	$\check{D}$
		.15	.01	.00	.00	.21	136.87	137.37
					.05	.22	153.39	153.94
				.05	.00	.10	38.09	38.23
					.05	.11	44.13	44.29
			.15	.00	.00	.18	112.70	113.09
					.05	.19	127.53	127.97
				.05	.00	.07	24.00	24.09
					.05	.08	29.43	29.53
	.10	.01	.01	.00	.00	.01	5.50	5.50
					.05	.00	12.91	12.91
				.05	.00	-.17	-39.11	-39.21
					.05	-.16	-36.41	-36.52
			.15	.00	.00	-.06	-10.80	-10.84
					.05	-.04	-4.54	-4.57
				.05	.00	-.24	-48.50	-48.62
					.05	-.22	-46.22	-46.34
		.15	.01	.00	.00	.15	115.34	115.67
					.05	.16	130.45	130.84
				.05	.00	.05	24.21	24.27
					.05	.06	29.71	29.78
			.15	.00	.00	.13	93.37	93.62
					.05	.14	106.94	107.23
				.05	.00	.02	11.54	11.56
					.05	.03	16.48	16.51

For Case A, the "true" measures are

T=.5012

R=2.000

D=2.005

From the definitions of R and D, it is clear that these are very close whenever $\pi(1|B)$ and $\pi(1|U)$ are close to unity.

From Table 5, the erroneous measures due to misclassification errors are easily obtained. For example, consider the case when there is no error in classifying belt usage when "killed", $\alpha(B|U;2) = \alpha(U|B;2) = 0$. Suppose further that $\alpha(B|U;1) = .15$, $\alpha(U|B;1) = .01$

and there are only minor errors in classifying "injury". From the table we get a percentage error of 115.34% for the relative risk measure. Thus, the apparent measure in this case is R=(1.1534)x2 + 2=4.3068.

It is clear from the table which combinations of errors produce large biases. Clearly, low errors produce a lesser bias, but it depends which components of errors are considered. Thus, note that positive biases are larger than negative biases! Also, various error components work in opposite directions and tend to cancel their individual effects. For example, the effect of an error $\alpha(B|U;2)$ is reduced when accompanied by a large error of $\alpha(U|B;2)$. Similarly for the injury scale. Note, however, that in this case one may roughly state that the classification errors of belt usage are the more troublesome, with the extremist cases corresponding to high values of $\alpha(B|U;1)$, $\alpha(U|B;2)$ and low values for $\alpha(B|U;2)$, $\alpha(U|B;1)$.

V. AN INVESTIGATION OF THE JOINT INJURY DISTRIBUTION OF BELTED AND UNBELTED OCCUPANTS

It is presumed that a better insight into the mechanism underlying the injury-reducing potential of seat belts might be obtained by constructing a bivariate injury distribution in a hypothetical population of accidents where each occupant "undergoes" the same accident twice, belted and unbelted, in an independent manner. To be meaningful, the constructed bivariate distribution is restrained to reflect the actual accident population.

Let i be a generic index for a combination of levels of the totality of factors that might affect the (belt usage) x (injury) distribution, and let w_i be the weight for that i-th sub-population.

To achieve such a bivariate distribution, we confined attention to drivers only and assumed that within a given sub-population the hypothetical match for a given accident with a belted driver is given by any of the unbelted driver accidents in that sub-population. To this aim, i ranged over 396 configurations corresponding to type of crash (11 levels), severity of crash (3 levels), car size (3 levels), car age (2 levels), and driver age (2 levels).

Let $N_i(j,B)$ denote the number of belted drivers in sub-population i involved in accidents that suffered a j-th degree of injury. Similarly, let $N_i(j,U)$ correspond to unbelted drivers with

$$N_i(B)=\sum_j N_i(j,B) \qquad N_i(U)=\sum_j N_i(j,U)$$

We are looking for an estimate of a hypothetical distribution

$\pi^{(i)}_{B,U}(j,k)$ classifying imaginary belted-unbelted pairs of accidents, where the j-th degree corresponds to a belted driver and k-th to an unbelted one. Within a given level i , we make the following assumption of independence:

$$\pi^{(i)}_{B,U}(j,k) = \pi^{(i)}_{B}(j)\ \pi^{(i)}_{U}(k),\ \text{for all } j,k\ .$$

The margins $\pi^{(i)}_{B}(j)$, $\pi^{(i)}_{U}(k)$ are assumed to be equal to the actual conditional distributions of injury for belted and unbelted drivers in sub-population i. Under this assumption, we have the estimates

$$\pi^{(i)}_{B}(j) \triangleq \frac{N_i(j,B)}{N_i(B)} \qquad \pi^{(i)}_{U}(j) \triangleq \frac{N_i(j,U)}{N_i(U)}$$

Note that on forming, for a given i, all the actual pairs of "belted-unbelted" accidents ($N_i(B) \times N_i(U)$ in number), the hypothetical bivariate injury distribution is estimated by

$$\pi^{(i)}_{B,U}(j,k) \triangleq \frac{N_i(j,B)\ N_i(k,U)}{N_i(B)\ N_i(U)}$$

However, we are interested in the overall bivariate injury distribution

$$\pi_{B,U}(j,k) = \sum_i w_i\ \pi^{(i)}_{B,U}(j,k)$$

of which we have the estimate

$$\sum_i w_i \frac{N_i(j,B)\ N_i(k,U)}{N_i(B)\ N_i(U)}$$

The resulting distributions corresponding to the 1973 and 1974 North Carolina accident files are now being processed on the computer.

VI. SHOULDER BELT UTILIZATION STUDY

The introduction of the interlock (1974) in addition to the buzzer system (1972) has prompted the need for current information along the lines of Anderson (1971) and Hunter and Lacey (in press) on seat belt usage rates in North Carolina. Anderson's data was collected in June, 1970, while Hunter and Lacey made their observations during the summer months of 1971.

Initially, plans included obtaining both lap and shoulder belt wearing rates. Trial observations indicated, however, that the lap belt information was quite unreliable. Difficulties in seeing lap belts when positioned on high ground or in an elevated vehicle included a lack of color contrast between the belt and the driver's clothing and also vision obstructions due to women's purses, large packages on the seat, sun glare on the window, etc. As an illustration of this, Table 6 shows the discrepancies that were found between pairs of observers collecting lap belt data from an elevated van moving in the same direction as the vehicle being observed. All observations were made on an interstate-type highway.

In view of these problems, it was decided that the study should focus on shoulder belt usage. The five stratifying factors considered in the sampling procedure were as follows:

1. Geographic area
    a. Coast
    b. Piedmont (Central)
    c. Mountains

Table 6. Results of pilot shoulder belt study

Date	Observing Group	No. of Observations	Agree (N,N)*	Agree (L,L)	Agree (B,B)	Disagree (N,L)	Disagree (N,B)	Disagree (L,N)	Disagree (L,B)	Disagree (B,N)	Disagree (B,L)
10/11/74	(I,II)	126	106	4	9	6	0	0	0	1	0
10/22/74	ALL	110	79	4	9			18			
	(IV,II)		87	4	9	6	0	1	1	1	1
	(III,II)		81	6	10	4	0	7	0	1	1
	(IV,III)		83	4	10	9	1	1	1	1	0

*N = no belt
L = lap belt only
B = lap and shoulder belts

where, for example,
(N,N) = both observers agree--no belts
(L,N) = first observer records lap belt--second records no belt

2. Road type
   a. Interstate
   b. Primary
   c. Secondary

3. Location
   a. Urban
   b. Rural

4. Time of day
   a. Commuting hours
   b. Other

5. Day of week
   a. Weekday
   b. Weekend

The vicinities of Rocky Mount, Durham, and Asheville represented the three geographic areas, respectively. Observations during the commuting hours were taken from 7:00 am to 9:30 am and from 4:00 pm to 7:00 pm. A total of three hours of observations were made for both the commuting and the other time periods. Weekends, however, did not include the commuting vs. other distinction. Weekday observations ended at 3:00 pm on Friday as late Friday traffic was thought to resemble weekend travel.

Within each sampling period, information on driver age (approximate), race and sex, and vehicle license plate number (O.S. for out-of-state) was recorded along with whether or not the driver was wearing his shoulder belt (see Figure 2).

The data thus collected (approximately 21,000 cases) have been processed and put on tape. The current effort involves obtaining make and model year information from the license plate number through the registration file. See Chapter 7 for a description of the future plans for analyzing this data.

Figure 2. Data form for shoulder belt utilization study

(1) Geographical area
1. Coast
2. Piedmont
3. Mountains

(2) Road type
1. Interstate
2. Primary
3. Secondary

(3) Location
1. Urban
2. Rural

(4) Time of day
1. Commuting home
2. Other

(5) Day of week
1. Weekday
2. Weekend

Recorder's name:______________________Date:_______________

RACE 1. White 2. Non-white	SEX 1. Male 2. Female	AGE 1. 16-35 2. 36-55 3. 56+	LICENSE PLATE NO.	SHOULDER BELT UTILIZATION 0. No 1. Yes

Some cross-tabulations describing the nature of this data include the following:

Table 7A. Frequency distribution by geographical region

	Frequency	%
Coast	5893	27.7
Piedmont	7591	35.7
Mountains	7772	36.6
Total	21256	100.0

Table 7B. Frequency distribution by part of week

	Frequency	%
Weekday	13835	65.1
Weekend	7422	34.9
Total	21257	100.0

Table 7C. Frequency distribution by part of day (weekdays only)

	Frequency	%
Commuting	7527	54.5
Other	6288	45.5
Total	13815	100.0

Table 7D. Frequency distribution for road type by location

Frequency Row % Col. % Total %	Primary	Secondary	Interstate	Total
Urban	7452 53.3 68.7 35.0	6521 46.7 71.8 30.6	0 0.0 0.0 0.0	13973 65.6
Rural	3403 46.5 31.3 16.0	2564 35.0 28.2 12.0	1358 18.5 100.0 6.4	7325 34.4
Total	10855 51.0	9085 42.7	1358 6.4	21298 100.0

VII. FUTURE RESEARCH

Effectiveness of Seat Belts.

Being aware of the sensitivity of our measures to errors of misclassification, we have conducted a small scale survey using telephone interviews of people involved in accidents.

This survey aimed at verifying that sometimes the police report on belt usage doesn't agree with the driver's response. There were a total of 48 drivers contacted and a total of 67 occupants. The results are indicated in the following tables:

Table 8.A. Police report vs. driver response-drivers only

Drivers' Response

	Yes	No	Total
Yes	8 (16.7%)	0	8 (16.7%)
No	14 (29.2%)	26 (54.2%)	40 (83.3%)
Total	22 (45.8%)	26 (54.2%)	48 (100%)

Table 8.B. Police report vs. driver response-all occupants

Drivers' Response

	Yes	No	Total
Yes	12 (17.9%)	0	12 (17.9%)
No	17 (25.4%)	38 (56.7%)	55 (82.1%)
Total	29 (43.3%)	38 (56.7)	67 (100%)

Other results indicate that the officer did not ask or investigate seat belt usage 53.3% of the time, did investigate 29.2% of the time, with 12.5% no response from the driver.

It is clear from that survey that there is far from complete agreement between these two sources of information. Misclassification errors on the injury scale by the police report is likewise well anticipated. (See McLean (1973), for example.)

In order to resolve these problems, the misclassification errors discussed above must be estimated and the (belt usage) x (injury) tables should be correspondingly adjusted before attempting a study of seat belt effectiveness. (See Appendix C for samples of unadjusted tables using recent North Carolina accident data.)

Efforts are currently being made to come up with estimates of misclassification errors of belt usage for given degrees of injury. To this aim the randomized response technique (see Appendix D) will be used.

In a later stage, the injury misclassification errors will be estimated using information from a more precise source than the police report (such as hospitals).(see McLean (1973).

Based on these sources of information, the two-way tables (belt usage) x (injury) will be adjusted and then, the issue of belt effectiveness will be studied by appropriate statistical methodologies applied to these adjusted tables.

Usage of Shoulder Belts in the Population at Risk.

We have approximately 21,000 observations of shoulder belt utilization in North Carolina during October, 1974, obtained using a stratified sampling scheme. Future plans here include modelling the variation in the proportion of usage as being dependent on factors like: Urban vs. rural driving, type of road (Interstate, primary, secondary), time (part of day x part of week), car characteristics (e.g., model year) and driver characteristics (age, sex, race). The statistical modelling will be along the lines of Grizzle et al (1969) and that of Goodman (1971).

Estimates of the conditional usage probabilities for different levels of model year, type of road, etc., including the corresponding standard errors and/or confidence intervals will be obtained.

APPENDIX A

Monotonicity of the Relative Risk with Severity of Injury

Viewing injury on a continuous scale, let F(x) and G(x) denote the cumulative injury distribution functions for belted and unbelted occupants, respectively. Assume that these distribution functions have corresponding densities f(x) and g(x).

If $h(x) = g(x)/f(x)$ is increasing with x, it follows that

$R(x) = \dfrac{1 - G(x)}{1 - F(x)}$ is also increasing with x. To show this, we must verify that

$$R'(x) = \frac{dR(x)}{d(x)} = \frac{f(x)\left[1 - G(x)\right] - g(x)\left[1 - F(x)\right]}{\left[1 - F(x)\right]^2}$$

is non-negative for all x; i.e., that

$$\frac{g(x)}{f(x)} \leq \frac{1 - G(x)}{1 - F(x)} \qquad \text{for all } x.$$

We have that

$$\frac{1 - G(x)}{1 - F(x)} = \frac{\int_{t=x}^{\infty} g(t)\,dt}{\int_{t=x}^{\infty} f(t)\,dt} = \frac{\int_{t=x}^{\infty} h(t) f(t)\,dt}{\int_{t=x}^{\infty} f(t)\,dt} \geq \frac{h(x)\int_{t=x}^{\infty} f(t)\,dt}{\int_{t=x}^{\infty} f(t)\,dt} = h(x)$$

since $h(t) = g(t)/f(t)$ increases with t.

The monotone increase in R(x) holds also when G and/or F are not continuous under the condition that

$$h(x) = \frac{dG(x)}{dF(x)}$$

is an increasing function of x. The proof in that setup is similar.

APPENDIX B

Measures for Seat Belt Effectiveness Under

An Exponential Distribution of Injury

Suppose that the injury distributions are exponential f(x) and g(x) with modes θ_B, θ_U for belted and unbelted occupants, respectively. Let $\theta_B^{(i)}, \theta_U^{(i)}$ denote the i-th population mode, i=1,2, for belted and unbelted occupants. Thus,

$$f_i(x) = \theta_B^{(i)} e^{-\theta_B^{(i)} x}, \qquad x \geq 0$$

$$g_i(x) = \theta_U^{(i)} e^{-\theta_U^{(i)} x}, \qquad x \geq 0$$

If $d_i = \theta_B^{(i)} - \theta_U^{(i)}$, i=1,2, then the following expressions provide the three measures of belt effectiveness under this model:

$$T_i = \int_{t=0}^{\infty} (1 - G_i(t))\, dF_i(t) = \frac{\theta_U^{(i)} + d_i}{2\theta_U^{(i)} + d_i}, \qquad i = 1, 2.$$

$$R_i(x) = \frac{1 - G_i(x)}{1 - F_i(x)} = e^{d_i x}, \quad x \geq 0, \qquad i = 1, 2.$$

$$D_i(x) = e^{d_i x} \left[\frac{1 - e^{-(\theta_U^{(i)} + d_i)\, x}}{1 - e^{-\theta_U^{(i)} x}} \right], \quad x \geq 0, \qquad i = 1, 2.$$

Note: When $\theta_U^{(1)} \neq \theta_U^{(2)}$ and $d_1 = d_2$ (i.e., same belt effect--the difference is only in the initial conditions), only $R_i(x)$, the relative risk measure, is independent of $\theta_U^{(i)}$ and depends for any given x only on $d_1 = d_2$. The ridit measure, T_i , gives a higher value for populations where accidents are more severe. The odds ratio measure also depends on $\theta_U^{(i)}$ in a way which is dependent on the value of x. Similar considerations point out the advantages of R(x) when $d_1 \neq d_2$ and $\theta_U^{(1)} = \theta_U^{(2)}$.

APPENDIX C

INITIAL CROSS-TABULATIONS

A large number of cross-tabulations of (usage of belt) x (injury) x (factor of interest) regarding the driver only, were obtained from the 1973, and first half of 1974 accident data. Only a small sample will appear in this appendix, the complete tables have been supplied to the contractor, November, 1974. The factors and their levels which were cross-tabulated with (usage of belt) x (injury) are:

Factor	Levels
Vehicle model year	1969-71, 72, 73, 74
Driver's age	16-35, 36-55, 56+
Vehicle maneuver just before accident	The 16 levels as in N.C. accident file
Crash configuration	12 levels formed by the variables point of impact and speed
Sobriety	4 levels
Violations	18 levels
Accident type	23 levels
Severity of accident as measured by TAD	Two TAD ratings each with seven severity levels
position in car	6 levels

Besides these tables, a number of two way tables of the form: (belt usage) x (factor) were obtained. These factors were:

Factor	Levels
Race	White, non-white
Highway class	Interstate, U.S., N.C., rural road, city street
Speed limit	5-20, 21-35, 36-50, 51+
Time	20 levels. (5 parts of the day) x (2 parts of week) x (2 parts of year)

The tables were presented in two formats--first, with three levels for both belt usage and injury, and second, after collapsing into 2 x 2 tables.

The following tables are but a sample of the above listed tabulations.

Table C.1 Vehicle model years 1969-71

	Belt	No belt		
No injury	7408	41500	48908	T= .51
Injury	1202	8638	9840	R=1.23
	8610	50138	58748	D=1.28

Table C.2 Vehicle model year 1974

	Belt	No belt		
No injury	1016	1074	2090	T= .52
Injury	164	232	396	R=1.28
	1180	1306	2486	D=1.34

Table C.3 TAD severity 1 for frontal collisions

	Belt	No belt		
No injury	1662	10875	12537	T= .51
Injury	54	613	667	R=1.70
	1716	11488	13204	D=1.73

Table C.4 TAD severity 7 for frontal collisions

	Belt	No belt		
No injury	72	349	421	T= .58
Injury	108	1084	1192	R=1.26
	180	1433	1613	D=2.07

APPENDIX D

TWO-STAGE RANDOMIZED RESPONSE SCHEMES FOR ESTIMATING A MULTINOMIAL DISTRIBUTION

Introduction

Warner's (1965) original randomized response technique for dichotomous data was extended by Abul-Ela et al. (1967) to the case of a multinomial with $t \geq 3$ groups, r of which are stigmatizing. A competing approach for estimating a multinomial via the randomized response technique is discussed by Warner (1971). Yet a third approach for the multinomial is considered by Greenberg et al. (1969) and is based on using an alternate independent question.

Here we outline some alternative schemes for estimating the t group proportions when $r < t-1$, using only one sample. Their realizations for any sampled individual constitute two-stage schemes. The second stage is conditional on the individual's response in the first stage.

In the next section we briefly summarize the three existing methods. The following section describes the new procedures and the resulting estimators. Next we identify the new schemes as special cases of Warner's (1974) general linear randomized response model thereby obtaining alternative estimators based on a modified generalized least squares method due to Zellner (1962).

Existing Methods

(i) The approach in Abul-Ela et al. (1967). Let π_i be the population proportion of the i^{th} group, i = 1, ..., t, $\Sigma \pi_i = 1$. Put $\underset{\sim}{\pi} : (t-1)\times 1 = (\pi_1, \ldots, \pi_{t-1})'$. This approach uses t-1 samples as follows. A matrix $(p_{ij}) : (t-1)\times t = (\underset{\sim}{p}_1, \ldots, \underset{\sim}{p}_t)$ must be determined such that $\sum_{j=1}^{t} p_{ij}$ for $i = 1, \ldots, t-1$ and $\underset{\sim}{P} : (t-1)\times(t-1) = (\underset{\sim}{p}_1 - \underset{\sim}{p}_t, \ldots, p_{t-1} - p_t)$ is non-singular. The p_{ij}'s determine the randomized scheme, where p_{ij} is the chance for an individual in the i^{th} sample to randomly select the question: "Are you in group j?" i = 1, ..., t-1; j = 1, . . . , t. Let n_i denote the i^{th} sample size, n_{i1} denote the number of "yes" responses in the i^{th} sample and let $\underset{\sim}{d} = \left(\frac{n_{i1}}{n_1}, \ldots, \frac{n_{t-1,1}}{n_{t-1}}\right)$. It is easily verified that the MLE's of the π_i's are given by

$$\hat{\underset{\sim}{\pi}} = \underset{\sim}{P}^{-1} \underset{\sim}{c} \qquad \hat{\pi}_t = 1 - \sum_{i=1}^{t-1} \hat{\pi}_i$$

where $\underset{\sim}{c} = \underset{\sim}{d} - \underset{\sim}{p}_t$.

(ii) The approach in Warner (1971). This is a special case of Warner's (1971) general linear randomized response model. In the general setup a simple random sample of size m is drawn and for the i^{th} individual in the sample the realization $\underset{\sim}{x}_i : t\times 1$ of a random vector $\underset{\sim}{X}$ cannot be observed. Rather, one observes

$$\underset{\sim}{y}_i : q\times 1 = \underset{\sim}{T}_i \underset{\sim}{x}_i , \qquad i = 1, \ldots, m$$

where the $\underset{\sim}{T}_i : q\times p$'s are random independent matrices and independent of the $\underset{\sim}{X}_i$'s. Assuming that the expectations of the $\underset{\sim}{T}_i$'s are known, the problem is that of estimating the expectation vector of $\underset{\sim}{X}$. Warner (1971) discusses generalized least squares estimates of the mean vector of $\underset{\sim}{X}$. For the problem of estimating a multinomial, Warner proposes the following scheme. Let $\underset{\sim}{x}_i = (x_{i1}, \ldots, x_{it})'$, where $x_{ij} = 1$ or 0 according as the i^{th} individual belongs to the j^{th} group or not. Here q = t, where $\underset{\sim}{T}_i$ for all $i = 1, \ldots, m$ is the random $t \times t$ matrix whose possible configurations are formed by permuting the columns of the t x t identity matrix. Then the random tranform amounts to directing the individual's response according to which group he is in depending on the random (unknown to the interviewer) realizations of the $\underset{\sim}{T}_i$'s.

(iii) The approach in Greenberg et al. (1969). This approach is very similar to the first. It is obtained from the first by giving a zero chance for an individual in any of the samples to be faced with the question: "Are you in group t?" and instead replaces this question by an alternate question "Do you possess characteristic Y?" where Y is unrelated to the characteristic by which $\pi_1, \ldots, \pi_t$ were formed. This is a modification of the dichotomous unrelated question randomized response model. If π_Y (proportion in population of individuals with characteristic Y) is known, only t - 1 samples are necessary to estimate the π_i's; otherwise, t samples are required.

Several observations are in order. All three procedures are quite involved. First, in all of them inversion of matrices must take place. Also, the task of choosing the p_{ij}'s and n_i's in the first and third methods and the chance probabilities attached to the various realizations of the $\underset{\sim}{T}_i$'s in the second method is difficult since no guiding theory exists. Also, in all three schemes there is a loss in efficiency resulting from their low sensitivity to the relation between r and t. Clearly, good randomized response schemes give protection to individuals in stigmatizing groups while minimizing protection (i.e., remove uncertainty) for individuals in non-stigmatizing groups. The ability to design such efficient schemes depends very much on the relation between r and t. This is clarified in the next section where we outline two new schemes with a double stage interview of an individual where the second scheme depends on the individual's response in the first scheme. Thus, these procedures will be referred to as Two Stage Randomized Response Schemes (TSRRS).

Two Stage Randomized Response Schemes (TSRRS)

We first treat the case $r = 1$, and, without loss in generality, assume that the first group is the stigmatizing one.

Scheme 1. Take a simple random sample of n individuals. Use a randomization device which gives a chance p for an individual to be faced with the question, "Are you in group 1?" and chance $1 - p$ to be requested to say the word "yes." This is the first stage. All individuals who answered "no" in this stage, say, n_0 of them, are directly asked in the second stage: "In what group are you?" since, clearly, they do not belong to the stigmatizing group. All other $n - n_0$ individuals who answered "yes" are protected. For these individuals there is no second stage. Let n_{oi} denote the number of individuals among the n_0 who responded "no" in the first stage, who identified themselves as belonging to group $i = 2, 3, \ldots, t$. Let λ denote the probability of a "no" response in the first stage.

$$\lambda = p(1-\pi_1)$$

The MLE estimate of λ is n_0/n and thus we get the MLE of π_1, as

$$\hat{\pi}_1 = 1 - \frac{n_0}{np}$$

Now, conditional on n_0 we have that $(n_{02}, n_{03}, \ldots, n_{0t})$, $\sum_{i=2}^{t} n_{0i} = n_0$, is a multinomial with cell probabilities.

$$\frac{\pi_2}{1-\pi_1}, \frac{\pi_3}{1-\pi_1}, \ldots, \frac{\pi_t}{1-\pi_1}$$

Thus our estimates of the population proportions of the non-stigmatizing groups are given by

$$\hat{\pi}_i = \frac{n_{0i}}{n_0}(1-\hat{\pi}_1) = \frac{n_i}{pn}, \quad i = 2, 3, \ldots, t.$$

These estimates are unbiased

$$E(\hat{\pi}_1) = 1 - \frac{n(1-\pi_1)p}{np} = \pi_1$$

$$E\left[E(\hat{\pi}_i|n_0)\right] = E\left[\frac{n_0\pi_i}{pn(1-\pi_1)}\right] = \pi_i, \quad i = 2, 3, \ldots, t.$$

The variances are given by

$$VAR(\hat{\pi}_1) = \frac{(1-\pi_1)(1-p+p\pi_1)}{np}$$

$$VAR(\hat{\pi}_i) = E\left[VAR(\hat{\pi}_i | n_0)\right] + VAR\left[E(\hat{\pi}_i | n_0)\right]$$

$$= \frac{\pi_i(1-\pi_1-\pi_i)}{np(1-\pi_1)} + \frac{\pi_i^2(1-p+p\pi_1)}{np(1-\pi_1)}$$

$$= \frac{\pi_i(1-\pi_i p)}{np} \quad , \quad i = 2, 3, \ldots, t.$$

Scheme 2. In the first stage we present a direct question to all individuals in the sample, "Are you in group 3 or 4 or, ..., or t or are you in either the first or the second?" Let n_i be those who fell into group i, i = 3, 4, ..., 5 and $n_{1,2}$ be the number of those who are in either the first or second group. All $n_{1,2}$ individuals then undergo a second stage in which a randomized scheme is used in order to estimate $\pi_j/(\pi_1+\pi_2)$, j = 1,2. The second stage is then a conditional randomized response scheme for dichotomous data and one may use any of the available techniques for this stage. Here we will use the scheme where one chooses p_1, p_2, p_3 for the chances that an individual undertaking the second stage will have to answer honestly the question "Are you in group 1?", will have to say "yes," or will have to say "no," respectively. Clearly $p_1 + p_2 + p_3 = 1$. Here

$$\hat{\pi}_i = \frac{n_i}{n} \quad , \quad i = 3, 4, \ldots, t$$

which are the best possible estimates of these unknown quantities. Let m_1 denote the number of individuals who say "yes" in the second stage.

We estimate $\pi_1/(\pi_1+\pi_2)$ from the relation

$$p_1 \frac{\pi_1}{\pi_1+\pi_2} + p_2 \hat{=} \frac{m_1}{n_{1,2}}$$

Thus $$\hat{\pi}_1 = \frac{n_{1,2}}{n}\left(\frac{m_1}{n_{1,2}} - p_2\right)\Big/ p_1 = \frac{m_1 - n_{1,2}p_2}{n p_1}$$

$$\hat{\pi}_2 = \frac{n_{1,2}}{n} - \hat{\pi}_1 .$$

The $\hat{\pi}_i$'s for i = 3, 4, ..., t are minimum variance unbiased estimates. The $\hat{\pi}_i$, i = 1, 2, are unbiased.

$$E\left[E(\hat{\pi}_1|n_{1,2})\right] = E\left[\frac{\left(\frac{p\pi_1}{\pi_1+\pi_2} + p_2\right)n_{1,2} - n_{1,2}p_2}{n p_1}\right] = \pi_1$$

hence $E(\hat{\pi}_2) = \pi_2$.

$$VAR(\hat{\pi}_1) = E\left[VAR(\hat{\pi}_1|n_{1,2})\right] + VAR\left[E(\hat{\pi}_1|n_{1,2})\right]$$

$$= \frac{\pi_1+\pi_2}{n p_1^2}\left[p_1\frac{\pi_1}{\pi_1+\pi_2} + p_2\right]\left[1 - p_1\frac{\pi_1}{\pi_1+\pi_2} - p_2\right]$$

$$+ \frac{\pi_1^2}{n(\pi_1+\pi_2)}(1-\pi_1-\pi_2)$$

We now turn to generalize the TSRRS's to $1 < r < t-1$.

<u>Scheme 1</u>. Without loss of generality, we assume that the r stigmatizing groups are the first r. Let p_i be the probability provided by some randomizing mechanism that an individual will be asked, "Are you in group i?" and 1 - p_i is the chance that he will be asked to say: "yes," i = 1, ..., r. Let n_0 denote the number of individuals in the

sample who responded with a "no" to all r questions. These n_0 individuals are clearly not in any stigmatizing group and thus a second stage in which they are asked to reveal their group is appropriate. All other $n - n_0$ individuals are protected.

Let λ_i denote the probability of a "yes" response on the i^{th} question, i = 1, ..., r, and let n_{yi} be the number of "yes" responses for the i^{th} question

$$\lambda_i = \pi_i \rho_i + 1 - \rho_i \; , \quad i = 1, \ldots, r$$

A possible estimate of λ_i is $\frac{n_{yi}}{n}$, i = 1, ..., r, we then get

$$\hat{\pi}_i = \frac{n_{yi} + n(\rho_i - 1)}{n \rho_i} \; , \quad i = 1, \ldots, r$$

(Note, $\frac{n_{yi}}{n}$ as an estimate of λ_i is unbiased and asymptotically best, based on the asymptotic normality of the n_{yi}'s. For finite samples the common distribution of the n_{yi}'s is not simple.)

We let n_{0i} denote the number of individuals among the n_0 who reported that they belong to the i^{th} group in the second stage, i = r+1, ..., t. Then, since, conditional on n_0, $(n_{0\,r+1}, \ldots, n_{0t})$ is a multinomial with cell proportions $\pi^*_{r+1}, \ldots, \pi^*_t$,

where

$$\pi_\ell^* = \frac{\pi_\ell}{\sum_{v=r+1}^{t} \pi_v} \; , \quad \ell = r+1, \ldots, t$$

it follows that

$$\hat{\pi}_i = \frac{n_{0i}}{n_0}\left(1 - \sum_{j=1}^{r} \hat{\pi}_j\right), \quad i = r+1, \ldots, t$$

These estimates are unbiased. The individual variances can be computed in a manner similar to the above. However, care must be taken for the covariances among the $\hat{\pi}_j$'s.

Scheme 2. Form the groups $(1, 2, \ldots, r+1)$, $r+2, \ldots, t$. (Other groupings might be used. The idea is to form a group composed of the r stigmatizing and one of the non-stigmatizing groups.

In the first stage ask the individuals directly to which of these t - r groups they belong. (It is understood that π_{r+1} is not too low to bias an individual's response.) Let $n_{1,r+1}$ be the number of individuals who identified themselves in the combined first group. Denote by n_i the number of those who identified themselves in the i^{th} group, $i = r+2, \ldots, t$. Let p_{i1}, p_{i2}, p_{i3} be the chances that an individual among the $n_{i,r+1}$ will have to answer honestly the subsequent question, "Are you in group i?" i = 1, ..., r, say "yes," say "no," respectively. Let θ_i be the probability of a "yes" response to the i^{th} question in the second stage, and let m_{yi} be the number of such "yes" responses to that question. Possible estimators of the θ_i 's are

$$\hat{\theta}_i = \frac{m_{yi}}{n_{1,r+1}}, \quad i = 1, \ldots, r$$

From

$$\theta_i = \frac{\pi_i}{\sum_{j=1}^{r+1} \pi_j} p_{i1} + p_{i2} \ , \quad i = 1, \ldots, r \ ,$$

we get

$$\hat{\pi}_i = \frac{n_{1,r+1}}{n} \left[\frac{\frac{m_{yi}}{n_{1,r+1}} - p_{i2}}{p_{i1}} \right] = \frac{m_{yi} - n_{1,r+1}\, p_{i2}}{n\, p_{i1}} \ , \quad i = 1, \ldots, r.$$

$$\hat{\pi}_{r+1} = \frac{n_{1,r+1}}{n} - \sum_{i=1}^{r} \hat{\pi}_i \ .$$

Examples

On identifying the TSRRS described above as special cases of Warner's (1971) general linear randomized response model, one may obtain alternative estimators based on Zellner's (1962) modified generalized least squares estimates. Consider the following examples.

Example 1. Suppose t = 3, r = 1 and the first group is the stigmatizing one. Demonstrating the second scheme, we use Warner's (1965) original approach for randomizing in the second stage.

$$\underset{\sim}{x}_i = \begin{cases} (1 \ 0 \ 0)' & \pi_1 \\ (0 \ 1 \ 0)' & \pi_2 \\ (0 \ 0 \ 1)' & \pi_3 \end{cases} \qquad i = 1, \ldots, n$$

Choose $\underset{\sim}{T}_i$ as a 2 x 3 random matrix with distribution

$$\underset{\sim}{T}_i = \begin{cases} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} & p \\ \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & 1-p \end{cases} \qquad i = 1, \ldots, n$$

and

$$\underset{\sim}{y}_i = \begin{cases} (0 \ 0) & \pi_1(1-p) + p\pi_2 \\ (1 \ 0) & \pi_1 p + \pi_2(1-p) \\ (0 \ 1) & \pi_3 \end{cases} \qquad i = 1, \ldots, n$$

p is the randomization chance in the second stage, (0,1) denotes identification of an individual with the third group in the first stage, (1,0) is the response "yes," from an individual in the second stage, while (0,0) is a "no" response from individuals who undertook a second stage.

Example 2. Let us now demonstrate the second scheme when r = 1 t = 3 as above but when an alternate question on a characteristic Y is used in the second stage to randomize with probability p ([1 $-p$]), an answer to the question: "Are you in group 1?" ("do you have characteristic Y?"). Using the same $\underset{\sim}{y}_i$'s as in Example 1, we have

$$\underset{\sim}{x}_i = \begin{cases} (1 \ 0 \ 0 \ 0 \ 0 \ 0)' & \pi_Y \pi_1 \\ (0 \ 1 \ 0 \ 0 \ 0 \ 0)' & \pi_Y \pi_2 \\ (0 \ 0 \ 1 \ 0 \ 0 \ 0)' & \pi_Y \pi_3 \\ (0 \ 0 \ 0 \ 1 \ 0 \ 0)' & \pi_1(1-\pi_Y) \\ (0 \ 0 \ 0 \ 0 \ 1 \ 0)' & \pi_2(1-\pi_Y) \\ (0 \ 0 \ 0 \ 0 \ 0 \ 1)' & \pi_3(1-\pi_Y) \end{cases}, \qquad i = 1, \ldots, n$$

$$T_i = \begin{cases} \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} & p \\ \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} & 1-p \end{cases} \qquad i = 1, \ldots, n$$

Example 3. Taking t = 4, r = 2 (the first two), we here demonstrate the first scheme. By letting "1" refer to a "yes" answer, "0" to a "no" answer and the position in a vector refer to the individual's group, we have

$$x_i = \begin{cases} (1\ 0\ 0\ 0)' \\ (0\ 1\ 0\ 0)' \\ (0\ 0\ 1\ 0)' \\ (0\ 0\ 0\ 1)' \end{cases} \qquad y_i = \begin{cases} (1\ 1\ 0\ 0)' & (1-p_1)(1-p_2)+p_1(1-p_2)\pi_1+(1-p_1)p_2\pi_2 \\ (1\ 0\ 0\ 0)' & (1-p_1)(1-\pi_2)p_2+p_1p_2\pi_1 \\ (0\ 1\ 0\ 0)' & (1-p_2)(1-\pi_1)p_1+p_1p_2\pi_2 \\ (0\ 0\ 1\ 0)' & p_1p_2\pi_3 \\ (0\ 0\ 0\ 1) & p_1p_2\pi_4 \end{cases}$$

$$T_i = \begin{cases} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} & p_1p_2 \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & p_1(1-p_2) \\ \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & p_2(1-p_1) \\ \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} & (1-p_1)(1-p_2) \end{cases} \qquad i = 1, \ldots, n.$$

In all TSRRS's the distributions of the $\underset{\sim}{T}_i$'s are the same for all i = 1, ..., n. Hence, as discussed in Warner (1971, p. 885), Zellner's (1962) method applies directly.

REFERENCES

Abul-Ela, A.A., Greenberg, B.G. and Horvitz, D.G. (1967). A multi-proportions randomized response model. J. Amer. Statist. Assoc., 62, 990-1008.

Anderson, T.E. (1971). Shoulder Belt Utilization. Chapel Hill: University of North Carolina Highway Safety Research Center.

Bross, I.D.J. (1958). How to use ridit analysis. Biometrics, 14, 18-38.

Campbell, B.J., O'Neill, B. and Tingley, B. (1974). Comparative injuries to belted and unbelted drivers of sub-compact, compact, intermediate, and standard cars. Presented at the Third International Congress on Automotive Safety, San Francisco, California.

Cochran, W.G. (1954). Some methods of strengthening the common χ^2 tests. Biometrics, 10, 417-451.

Fienberg, E.S. (1970). An iterative procedure for estimation in contingency tables. Annals Math. Statist., 41, 907-917.

Fleiss, J.L. (1973). Statistical Methods for Rates and Proportions. New York: John Wiley and Sons, Inc.

Folsom, R.E., Greenberg, B.G., Horvitz, D.G. and Abernathy, J.R. (1973). The two alternate questions randomized response model for human surveys. J. Amer. Statist. Assoc., 68, 525-530.

Forthofer, R.N. and Koch, G.G. (1973). An analysis of compounded functions of categorical data. Biometrics, 29, 143-157.

Freeman, J., Koch, G.G., Hunter, W.W. and Lacey, J.H. (1975). An Analysis of Shoulder Belt Use in North Carolina. In press. Chapel Hill: University of North Carolina Highway Safety Research Center.

Goodman, L.A. (1971). The analysis of multidimensional contingency tables: Stepwise procedures and direct estimation methods for building models for multiple classifications. Technometrics, 13, 33-61.

Greenberg, B.G., Abul-Ela, A.A., Simmons, W.R. and Horvitz, D.G. (1969). The unrelated question randomized response model: Theoretical framework. J. Amer. Statist. Assoc., 64, 520-539.

Greenberg, B.G., Horvitz, D.G. and Abernathy, J.R. (1974). A comparison of randomized response designs. Reliability and Biometry. Philadelphia: Siam. 787-815.

Griffin, L.I. (1973). Analysis of Benefits Derived from Certain Presently Existing Motor Vehicle Safety Devices: A Review of the Literature. Chapel Hill: University of North Carolina Highway Safety Research Center.

Grizzle, J.E., Starmer, C.F. and Koch, G.G. (1969). Analysis of categorical data by linear models. Biometrics, 25, 489-504.

Haberman, S.J. (1970). The General Log-linear Model. Unpublished Ph.D dissertation. Chicago: The University of Chicago.

Mantel, N. (1963). Chi-square tests with one degree of freedom: Extensions of the Mantel-Haenszel procedure. J. Amer. Statist. Assoc., 58, 690-700.

McLean, A.J. (1973). Collection and Analysis of Collision Data for Determining the Effectiveness of Some Vehicle Systems. Chapel Hill: University of North Carolina Highway Safety Research Center.

Mela, D.G. (1974a). Review of safety belt usage and effectiveness in accidents. A report of the National Highway Traffic Safety Administration.

Mela, D.G. (1974b). How accurate are seat belt statistics? Highway Safety Highlights, vol. 7, no. 12. Chapel Hill: University of North Carolina Highway Safety Research Center.

Mosteller, F. (1968). Association and estimation in contingency tables. J. Amer. Statist. Assoc., 63, 1-23.

Plackett, R.L. (1960). A class of bivariate distributions. J. Amer. Statist. Assoc., 60, 516-522.

Proceedings of the Third International Congress on Automotive Safety (1974). Government Printing Office, Washington, D.C. (Stock No. 5003-00173).

Warner, S.L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. J. Amer. Statist. Assoc., 60, 66-69.

Warner, S.L. (1971). The linear randomized response model. J. Amer. Statist. Assoc., 66, 884-888.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. J. Amer. Statist. Assoc., 57, 348-368.