ESTIMATION OF ODDS RATIO AND ATTRIBUTABLE RISK USING RANDOMIZED RESPONSE TECHNIQUES

Cheon-sig Lee[1], Stephen A. Sedory[2], Sarjinder Singh[2]
[1] Department of Mathematics, Coastal Bend College, Beeville, TX 78102
[2] Department of Mathematics, Texas A&M University-Kingsville, Kingsville, TX 78363

ABSTRACT

In this paper, we first define odds ratio and attributable risk while considering investigating two sensitive attributes in real practice. Then we define two estimators of odds ratio and two estimators of attributable risk based on data collected either using the simple model or crossed model proposed by Lee, Sedory and Singh (2013). We define expressions for biases and variances of the resultant estimators. We investigate the performance of crossed model over the simple model under the same choice of parameters as discussed in Lee et al (2013). Also the values of odds ratio and attributable risk are reported based on a real data set.

Key words: Sensitive characteristics, estimation of proportion, crossed model, simple model.

1. INTRODUCTION

In 1965, S. L. Warner proposed the first research method in structured survey interview. Lee, Sedory and Singh (2013) introduced a new methodology for estimating the proportions of persons in a population possessing each of two sensitive characteristics, say *A* and *B*, along with the proportion possessing both, $A \cap B$, by using two different randomized response models: Simple model and Crossed model. There are many situations where their proposed models could be implemented in real practice. For example, (1) assume *A* is a group of smokers, *B* is a group of drinkers, then $A \cap B$ will be a group of both smokers and drinkers; (2) assume *A* is a group of smack users, *B* is a group of people involved criminally, then $A \cap B$ will be a group of both smack users and criminally active people; and (3) assume *A* represents hidden membership in a terrorist group-I, *B* represents a hidden membership in a terrorist group-II, then $A \cap B$ will be a hidden membership in both terrorist groups. Their models also allows one to estimate several other parameters, such as correlation coefficient, conditional proportions, and relative risk, etc. A pictorial representation of such a population is shown in Figure 1.1. Let π_A, π_B and $\pi_{A\cap B}$ be the true proportions of respondents possessing the sensitive characteristics *A*, *B*, and both $A \cap B$. Also note that $\pi_{A\cap B} \le Min(\pi_A, \pi_B)$.

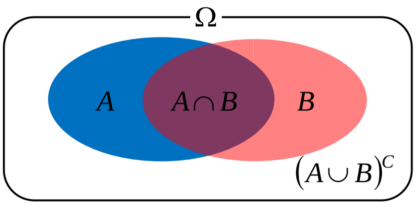

Fig.1.1. Populaton under study

Assume a simple random and with replacement sample (SRSWR) of n respondents is selected from the population Ω. The authors suggest above two different randomized response models, which are described in brief in the following sub-sections named as Simple Model, and as Crossed Model.

1.1 Simple Model

In the simple model proposed by Lee, Sedory and Singh (2013), they suggest to using a pair of decks of cards in order: say Deck-I and Deck-II. Each respondent, selected in a simple random with replacement sample of size n, is requested to draw two cards, one card from each deck of cards and keep the response in order from Deck-I and Deck-II respectively. Deck-I consists of cards, each bearing one of two mutually exclusive statements: "I belong to the sensitive group A", with probability P, and "I belong to the non-sensitive group A^c", with probability $(1-P)$. Deck-II also consists of cards, each bearing one of two mutually exclusive statements: " I belong to the sensitive group B", with probability T, and "I belong to the non-sensitive group B^c", with probability $(1-T)$. By following the notation of Lee, Sedory and Singh (2013) for the simple model, the probabilities of obtaining, from a given respondent, each of the following responses (*Yes*, *Yes*), (*Yes*, *No*), (*No*, *Yes*) and (*No*, *No*) are, respectively, given by:

$$\theta_{11} = (2P-1)(2T-1)\pi_{AB} + (2P-1)(1-T)\pi_A + (1-P)(2T-1)\pi_B + (1-P)(1-T)\,, \quad (1.1)$$

$$\theta_{10} = -(2P-1)(2T-1)\pi_{AB} + (2P-1)T\pi_A - (1-P)(2T-1)\pi_B + (1-P)T\,, \quad (1.2)$$

$$\theta_{01} = -(2P-1)(2T-1)\pi_{AB} - (2P-1)(1-T)\pi_A + P(2T-1)\pi_B + P(1-T)\,, \quad (1.3)$$

and

$$\theta_{00} = (2P-1)(2T-1)\pi_{AB} - T(2P-1)\pi_A - P(2T-1)\pi_B + PT\,. \quad (1.4)$$

Let $\hat{\theta}_{11} = n_{11}/n$, $\hat{\theta}_{10} = n_{10}/n$, $\hat{\theta}_{01} = n_{01}/n$ and $\hat{\theta}_{00} = n_{00}/n$, be the observed proportions of (*Yes*, *Yes*), (*Yes*, *No*), (*No*, *Yes*) and (*No*, *No*) responses, so that $n_{11} + n_{10} + n_{01} + n_{00} = n$. Then Lee, Sedory and Singh (2013) obtained unbiased estimators for the simple model as following:

$$\hat{\pi}_A = \frac{\hat{\theta}_{11} + \hat{\theta}_{10} - \hat{\theta}_{01} - \hat{\theta}_{00} + (2P-1)}{2(2P-1)}\,, \quad (1.5)$$

$$\hat{\pi}_B = \frac{\hat{\theta}_{11} - \hat{\theta}_{10} + \hat{\theta}_{01} - \hat{\theta}_{00} + (2T-1)}{2(2T-1)}\,, \quad (1.6)$$

and

$$\hat{\pi}_{AB} = \frac{(P+T)\hat{\theta}_{11} + (T-P)\hat{\theta}_{10} + (P-T)\hat{\theta}_{01} + (2-P-T)\hat{\theta}_{00} - T(1-P) - P(1-T)}{2(2P-1)(2T-1)}\,, \quad (1.7)$$

for $P \neq 0.5$ and $T \neq 0.5$.

Defining

$$\varepsilon_{AB} = \frac{\hat{\pi}_{AB}}{\pi_{AB}} - 1, \quad \varepsilon_A = \frac{\hat{\pi}_A}{\pi_A} - 1, \quad \varepsilon_{AB} = \frac{\hat{\pi}_B}{\pi_B} - 1$$

such that

$$E(\varepsilon_{AB}) = E(\varepsilon_A) = E(\varepsilon_B) = 0$$

$$E(\varepsilon_{AB}^2) = \frac{V(\hat{\pi}_{AB})}{\pi_{AB}^2}, \quad E(\varepsilon_A^2) = \frac{V(\hat{\pi}_A)}{\pi_A^2}, \quad E(\varepsilon_B^2) = \frac{V(\hat{\pi}_B)}{\pi_B^2}$$

$$E(\varepsilon_{AB}\varepsilon_A) = \frac{Cov(\hat{\pi}_{AB}, \hat{\pi}_A)}{\pi_{AB}\pi_A}, \quad E(\varepsilon_{AB}\varepsilon_B) = \frac{Cov(\hat{\pi}_{AB}, \hat{\pi}_B)}{\pi_{AB}\pi_B}, \quad E(\varepsilon_A\varepsilon_B) = \frac{Cov(\hat{\pi}_A, \hat{\pi}_B)}{\pi_A\pi_B}$$

where

$$V(\hat{\pi}_A) = \frac{\pi_A(1-\pi_A)}{n} + \frac{P(1-P)}{n(2P-1)^2}, \quad (1.8)$$

$$V(\hat{\pi}_B) = \frac{\pi_B(1-\pi_B)}{n} + \frac{T(1-T)}{n(2T-1)^2}, \quad (1.9)$$

$$V(\hat{\pi}_{AB}) = \frac{\pi_{AB}(1-\pi_{AB})}{n} + \frac{(2P-1)^2 T(1-T)\pi_A + P(1-P)(2T-1)^2\pi_B + PT(1-P)(1-T)}{n(2P-1)^2(2T-1)^2}, \quad (1.10)$$

$$Cov(\hat{\pi}_A, \hat{\pi}_B) = \frac{\pi_{AB} - \pi_A\pi_B}{n}, \quad (1.11)$$

$$Cov(\hat{\pi}_{AB}, \hat{\pi}_A) = \frac{\pi_{AB}(1-\pi_A)}{n} + \frac{P(1-P)\pi_B}{n(2P-1)^2}, \quad (1.12)$$

and

$$Cov(\hat{\pi}_{AB}, \hat{\pi}_B) = \frac{\pi_{AB}(1-\pi_B)}{n} + \frac{T(1-T)\pi_A}{n(2T-1)^2}. \quad (1.13)$$

1.2 Crossed Model

In the crossed model, while the rest of the procedure remains the same as for the simple model but the composition of the decks is different. Deck-I consists of cards, each bearing one of two mutually exclusive statements: "I belong to the sensitive group A", with probability P and "I belong to the non-sensitive group B^c", with probability $(1-P)$ respectively. Deck-II also consists of cards, each bearing one of two mutually exclusive statements: "I belong to the sensitive group B" with probability T and "I belong to the non-sensitive group A^c"with probability $(1-T)$ respectively. By following the notation of Lee, Sedory and Singh (2013) for the crossed model, the probabilities of obtaining, from a given respondent, each of the following responses, (*Yes*, *Yes*), (*Yes*, *No*), (*No*, *Yes*) and (*No*, *No*) are, respectively, given by:

$$\theta_{11}^* = \pi_{AB}\{PT + (1-P)(1-T)\} - \pi_A(1-P)(1-T) - \pi_B(1-P)(1-T) + (1-P)(1-T), \quad (1.14)$$

$$\theta_{10}^* = -\pi_{AB}\{PT + (1-P)(1-T)\} - \pi_A\{(1-P)T - 1\} - \pi_B(1-P)T + (1-P)T, \quad (1.15)$$

$$\theta_{01}^* = -\pi_{AB}\{PT + (1-P)(1-T)\} - \pi_A P(1-T) - \pi_B\{P(1-T) - 1\} + P(1-T), \quad (1.16)$$

and

$$\theta_{00}^* = \pi_{AB}\{PT + (1-P)(1-T)\} - \pi_A PT - \pi_B PT + PT. \quad (1.17)$$

Let $\hat{\theta}_{11}^{*}=n_{11}^{*}/n$, $\hat{\theta}_{10}^{*}=n_{10}^{*}/n$, $\hat{\theta}_{01}^{*}=n_{01}^{*}/n$ and $\hat{\theta}_{00}^{*}=n_{00}^{*}/n$, be the observed proportions of (*Yes*, *Yes*), (*Yes*, *No*), (*No*, *Yes*) and (*No*, *No*) responses so that $n_{11}^{*}+n_{10}^{*}+n_{01}^{*}+n_{00}^{*}=n$. Lee, Sedory and Singh (2013) obtained unbiased estimators for the crossed model as following:

$$\hat{\pi}_{A}^{*}=\frac{1}{2}+\frac{(T-P+1)(\hat{\theta}_{11}^{*}-\hat{\theta}_{00}^{*})+(P+T-1)(\hat{\theta}_{10}^{*}-\hat{\theta}_{01}^{*})}{2(P+T-1)}, \tag{1.18}$$

$$\hat{\pi}_{B}^{*}=\frac{1}{2}+\frac{(P-T+1)(\hat{\theta}_{11}^{*}-\hat{\theta}_{00}^{*})+(P+T-1)(\hat{\theta}_{01}^{*}-\hat{\theta}_{10}^{*})}{2(P+T-1)}, \tag{1.19}$$

and

$$\hat{\pi}_{AB}^{*}=\frac{PT\hat{\theta}_{11}^{*}-(1-P)(1-T)\hat{\theta}_{00}^{*}}{\{PT+(1-P)(1-T)\}(P+T-1)}, \tag{1.20}$$

with $P+T\neq 1$.

Defining

$$\varepsilon_{AB}^{*}=\frac{\hat{\pi}_{AB}^{*}}{\pi_{AB}}-1,\quad \varepsilon_{A}^{*}=\frac{\hat{\pi}_{A}^{*}}{\pi_{A}}-1,\quad \varepsilon_{B}^{*}=\frac{\hat{\pi}_{B}^{*}}{\pi_{B}}-1$$

such that

$$E\left(\varepsilon_{AB}^{*}\right)=E\left(\varepsilon_{A}^{*}\right)=E\left(\varepsilon_{B}^{*}\right)=0$$
$$E\left(\varepsilon_{AB}^{*2}\right)=\frac{V\left(\hat{\pi}_{AB}^{*}\right)}{\pi_{AB}^{2}},\quad E\left(\varepsilon_{A}^{*2}\right)=\frac{V\left(\hat{\pi}_{A}^{*}\right)}{\pi_{A}^{2}},\quad E\left(\varepsilon_{B}^{*2}\right)=\frac{V\left(\hat{\pi}_{B}^{*}\right)}{\pi_{B}^{2}}$$
$$E\left(\varepsilon_{AB}^{*}\varepsilon_{A}^{*}\right)=\frac{Cov\left(\hat{\pi}_{AB}^{*},\hat{\pi}_{A}^{*}\right)}{\pi_{AB}\pi_{A}},\quad E\left(\varepsilon_{AB}^{*}\varepsilon_{B}^{*}\right)=\frac{Cov\left(\hat{\pi}_{AB}^{*},\hat{\pi}_{B}^{*}\right)}{\pi_{AB}\pi_{B}},\quad E\left(\varepsilon_{A}^{*}\varepsilon_{B}^{*}\right)=\frac{Cov\left(\hat{\pi}_{A}^{*},\hat{\pi}_{B}^{*}\right)}{\pi_{A}\pi_{B}}$$

where

$$V(\hat{\pi}_{A}^{*})=\frac{\pi_{A}(1-\pi_{A})}{n}+\frac{(1-P)T\{PT+(1-P)(1-T)\}(1-\pi_{A}-\pi_{B}+2\pi_{AB})}{n(P+T-1)^{2}}, \tag{1.21}$$

$$V(\hat{\pi}_{B}^{*})=\frac{\pi_{B}(1-\pi_{B})}{n}+\frac{(1-T)P\{PT+(1-P)(1-T)\}(1-\pi_{A}-\pi_{B}+2\pi_{AB})}{n(P+T-1)^{2}}, \tag{1.22}$$

$$V(\hat{\pi}_{AB}^{*})=\frac{\pi_{AB}(1-\pi_{AB})}{n}+\frac{\pi_{AB}\left[P^{2}T^{2}+(1-P)^{2}(1-T)^{2}-\{PT+(1-P)(1-T)\}(P+T-1)^{2}\right]}{n\{PT+(1-P)(1-T)\}(P+T-1)^{2}}$$
$$+\frac{PT(1-P)(1-T)(1-\pi_{A}-\pi_{B})}{n\{PT+(1-P)(1-T)\}(P+T-1)^{2}}. \tag{1.23}$$

$$Cov\left(\hat{\pi}_{AB}^{*},\hat{\pi}_{A}^{*}\right)=\frac{\pi_{AB}\left(1-\pi_{A}\right)}{n}+\frac{\pi_{AB}\left\{PT+\left(1-P\right)\left(1-T\right)\right\}T\left(1-P\right)\left(P-T+1\right)}{n\left\{PT+\left(1-P\right)\left(1-T\right)\right\}\left(P+T-1\right)^{2}}$$
$$+\frac{PT\left(1-P\right)\left(1-T\right)\left(T-P+1\right)\left(1-\pi_{A}-\pi_{B}\right)}{n\left\{PT+(1-P)(1-T)\right\}(P+T-1)^{2}} \tag{1.24}$$

$$Cov\left(\hat{\pi}_{AB}^{*},\hat{\pi}_{B}^{*}\right)=\frac{\pi_{AB}\left(1-\pi_{B}\right)}{n}+\frac{\pi_{AB}P(1-T)(T-P+1)}{n(P+T-1)^{2}}+\frac{PT(1-P)(1-T)(P-T+1)(1-\pi_{A}-\pi_{B})}{n\{PT+(1-P)(1-T)\}(P+T-1)^{2}} \tag{1.25}$$

and

$$Cov\left(\hat{\pi}_{A}^{*},\hat{\pi}_{B}^{*}\right)=\frac{\pi_{A}(1-\pi_{A})-\pi_{AB}\{PT+(1-P)(1-T)\}}{n}-\frac{\{PT+(1-P)(1-T)+(P-T)(1-2\pi_{A})\}(1-\pi_{A}-\pi_{B})}{2n}+\frac{\{PT+(1-P)(1-T)\}^{2}\{1-\pi_{A}-\pi_{B}+2\pi_{AB}\}}{2n(P+T-1)^{2}} \tag{1.26}$$

In the next section, we consider two estimators of odds ratio (OR); one based on the simple model and the other based on crossed model.

2. ESTIMATION OF ODDS RATIO

The use of estimation of odds ratio and its problem are well known to statisticians who dealing with the problem of estimation of proportion of characteristics. In case of two sensitive characteristics A and B, the four cells of the 2×2 contingency table can be labeled as:

Attributes	B	B^c	Total
A	π_{AB}	$(\pi_A-\pi_{AB})$	π_A
A^c	$(\pi_B-\pi_{AB})$	$(1-\pi_A-\pi_B+\pi_{AB})$	$(1-\pi_A)$
Total	π_B	$(1-\pi_B)$	1

Thus, we consider a measure of odds ratio (OR) in case of two sensitive variables A and B as:

$$\text{OR}=\frac{\pi_{AB}\left(1-\pi_{A}-\pi_{B}+\pi_{AB}\right)}{\left(\pi_{A}-\pi_{AB}\right)\left(\pi_{B}-\pi_{AB}\right)} \tag{2.1}$$

In the following sub-sections, we consider estimators of the odds ratio (OR) defined in (2.1) by using the simple model and the crossed model.

2.1 ESTIMATION OF ODDS RATIO USING SIMPLE MODEL

By using the same notations for the simple model from Lee et al. (2013), we consider first estimator of the odds ratio (OR) as:

$$\hat{\text{OR}}_{1}=\frac{\hat{\pi}_{AB}\left(1-\hat{\pi}_{A}-\hat{\pi}_{B}+\hat{\pi}_{AB}\right)}{\left(\hat{\pi}_{A}-\hat{\pi}_{AB}\right)\left(\hat{\pi}_{B}-\hat{\pi}_{AB}\right)} \tag{2.2}$$

Now, we have the following theorems:

Theorem 2.1. The bias in the estimator $\hat{OR}_1$ of the odds ratio (OR) is given by:

$$B(\hat{OR}_1) = OR\left[\left\{\frac{(1-\pi_B)(\pi_A\pi_B-\pi_{AB}^2)}{\pi_{AB}(\pi_A-\pi_{AB})^2(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}+\frac{\pi_B}{\pi_{AB}(\pi_B-\pi_{AB})^2}\right\}V(\hat{\pi}_{AB})\right.$$
$$+\frac{(1-\pi_B)V(\hat{\pi}_A)}{(\pi_A-\pi_{AB})^2(1-\pi_A-\pi_B+\pi_{AB})}+\frac{(1-\pi_A)V(\hat{\pi}_B)}{(\pi_B-\pi_{AB})^2(1-\pi_A-\pi_B+\pi_{AB})}-\left\{\frac{1}{(\pi_A-\pi_{AB})(\pi_B-\pi_{AB})}\right.$$
$$\left.+\frac{\pi_B}{\pi_{AB}(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}+\frac{\pi_A+\pi_{AB}-2\pi_{AB}\pi_B}{\pi_{AB}(\pi_A-\pi_{AB})^2(1-\pi_A-\pi_B+\pi_{AB})}\right\}Cov(\hat{\pi}_{AB},\hat{\pi}_A)$$
$$-\frac{1-\pi_A}{\pi_{AB}(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}\left(\frac{\pi_A}{\pi_A-\pi_{AB}}+\frac{2\pi_{AB}}{\pi_B-\pi_{AB}}\right)Cov(\hat{\pi}_{AB},\hat{\pi}_B)$$
$$\left.+\frac{1-\pi_{AB}}{(\pi_A-\pi_{AB})(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}Cov(\hat{\pi}_A,\hat{\pi}_B)\right] \tag{2.3}$$

Proof. The estimator $\hat{OR}_1$ of the odds ratio can be approximated as:

$$\hat{OR}_1 = \frac{\hat{\pi}_{AB}(1-\hat{\pi}_A-\hat{\pi}_B+\hat{\pi}_{AB})}{(\hat{\pi}_A-\hat{\pi}_{AB})(\hat{\pi}_B-\hat{\pi}_{AB})}$$
$$=\frac{\pi_{AB}(1+\varepsilon_{AB})\{1-\pi_A(1+\varepsilon_A)-\pi_B(1+\varepsilon_B)+\pi_{AB}(1+\varepsilon_{AB})\}}{\{\pi_A(1+\varepsilon_A)-\pi_{AB}(1+\varepsilon_{AB})\}\{\pi_B(1+\varepsilon_B)-\pi_{AB}(1+\varepsilon_{AB})\}}$$
$$=OR\left[1+\left\{\frac{\pi_{AB}(1-\pi_B)}{(\pi_A-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}+\frac{\pi_B}{\pi_B-\pi_{AB}}\right\}\varepsilon_{AB}\right.$$
$$-\frac{\pi_A(1-\pi_B)\varepsilon_A}{(\pi_A-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}-\frac{\pi_B(1-\pi_A)\varepsilon_B}{(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}$$
$$+\left\{\frac{\pi_{AB}(1-\pi_B)(\pi_A\pi_B-\pi_{AB}^2)}{(\pi_A-\pi_{AB})^2(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}+\frac{\pi_{AB}\pi_B}{(\pi_B-\pi_{AB})^2}\right\}\varepsilon_{AB}^2$$
$$+\frac{\pi_A^2(1-\pi_B)}{(\pi_A-\pi_{AB})^2(1-\pi_A-\pi_B+\pi_{AB})}\varepsilon_A^2+\frac{\pi_B^2(1-\pi_A)}{(\pi_B-\pi_{AB})^2(1-\pi_A-\pi_B+\pi_{AB})}\varepsilon_B^2$$
$$-\left\{\frac{\pi_A\pi_B}{(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}+\frac{\pi_{AB}\pi_A}{(\pi_A-\pi_{AB})(\pi_B-\pi_{AB})}+\frac{\pi_A(\pi_A+\pi_{AB}-2\pi_{AB}\pi_B)}{(\pi_A-\pi_{AB})^2(1-\pi_A-\pi_B+\pi_{AB})}\right\}\varepsilon_{AB}\varepsilon_A$$
$$-\frac{\pi_B(1-\pi_A)}{(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}\left(\frac{\pi_A}{\pi_A-\pi_{AB}}+\frac{2\pi_{AB}}{\pi_B-\pi_{AB}}\right)\varepsilon_{AB}\varepsilon_B$$
$$\left.+\frac{\pi_A\pi_B(1-\pi_{AB})}{(\pi_A-\pi_{AB})(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}\varepsilon_A\varepsilon_B+O(\varepsilon^2)\right]$$

By the definition of bias that is

$$B\left(\hat{OR}_1\right) = E(\hat{OR}_1) - OR$$

we have the theorem.

Theorem 2.2. The mean squared error of the estimator $\hat{OR}_1$ of the odds ratio (OR) is given by:

$$MSE(\hat{OR}_1) = OR^2\left[\left\{\frac{1-\pi_B}{(\pi_A-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}+\frac{\pi_B}{\pi_{AB}(\pi_B-\pi_{AB})}\right\}^2 V(\hat{\pi}_{AB})\right.$$
$$+\frac{(1-\pi_B)^2V(\hat{\pi}_A)}{(\pi_A-\pi_{AB})^2(1-\pi_A-\pi_B+\pi_{AB})^2}+\frac{(1-\pi_A)^2V(\hat{\pi}_B)}{(\pi_B-\pi_{AB})^2(1-\pi_A-\pi_B+\pi_{AB})^2}$$
$$-\frac{2(1-\pi_B)}{(\pi_A-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}\left\{\frac{1-\pi_B}{(\pi_A-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}+\frac{\pi_B}{\pi_{AB}(\pi_B-\pi_{AB})}\right\}Cov(\hat{\pi}_{AB},\hat{\pi}_A)$$
$$-\frac{2(1-\pi_A)}{(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}\left\{\frac{1-\pi_B}{(\pi_A-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}+\frac{\pi_B}{\pi_{AB}(\pi_B-\pi_{AB})}\right\}Cov(\hat{\pi}_{AB},\hat{\pi}_B)$$
$$\left.+\frac{2(1-\pi_A)(1-\pi_B)Cov(\hat{\pi}_A,\hat{\pi}_B)}{(\pi_A-\pi_{AB})(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})^2}\right] \quad (2.4)$$

Proof. By the definition of mean squared error, we have

$$MSE\left(\hat{OR}_1\right)=E\left(\hat{OR}_1-OR\right)^2$$
$$\cong OR^2E\left[\left\{\frac{\pi_{AB}(1-\pi_B)}{(\pi_A-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}+\frac{\pi_B}{\pi_B-\pi_{AB}}\right\}\varepsilon_{AB}\right.$$
$$\left.-\frac{\pi_A(1-\pi_B)\varepsilon_A}{(\pi_A-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}-\frac{\pi_B(1-\pi_A)\varepsilon_B}{(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}\right]^2$$

Expanding and taking the expected value, we have the theorem.

2.2 ESTIMATION OF ODDS RATIO USING CROSSED MODEL

By using the same notations for the crossed model from Lee et al. (2013), we consider second estimator of the odds ratio (OR) as:

$$\hat{OR}_2=\frac{\hat{\pi}^*_{AB}\left(1-\hat{\pi}^*_A-\hat{\pi}^*_B+\hat{\pi}^*_{AB}\right)}{\left(\hat{\pi}^*_A-\hat{\pi}^*_{AB}\right)\left(\hat{\pi}^*_B-\hat{\pi}^*_{AB}\right)} \quad (2.5)$$

Now, we have the following theorems:

Theorem 2.3. The bias in the estimator $\hat{OR}_2$ of the odds ratio (OR) is given by:

$$B(\hat{OR}_2)=OR\left[\left\{\frac{(1-\pi_B)(\pi_A\pi_B-\pi_{AB}^2)}{\pi_{AB}(\pi_A-\pi_{AB})^2(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}+\frac{\pi_B}{\pi_{AB}(\pi_B-\pi_{AB})^2}\right\}V(\hat{\pi}^*_{AB})\right.$$

$$
+\frac{(1-\pi_B)V\left(\hat{\pi}_A^*\right)}{(\pi_A-\pi_{AB})^2(1-\pi_A-\pi_B+\pi_{AB})}+\frac{(1-\pi_A)V\left(\hat{\pi}_B^*\right)}{(\pi_B-\pi_{AB})^2(1-\pi_A-\pi_B+\pi_{AB})}-\left\{\frac{1}{(\pi_A-\pi_{AB})(\pi_B-\pi_{AB})}\right.
$$

$$
\left.+\frac{\pi_B}{\pi_{AB}(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}+\frac{\pi_A+\pi_{AB}-2\pi_{AB}\pi_B}{\pi_{AB}(\pi_A-\pi_{AB})^2(1-\pi_A-\pi_B+\pi_{AB})}\right\}Cov\left(\hat{\pi}_{AB}^*,\hat{\pi}_A^*\right)
$$

$$
-\frac{1-\pi_A}{\pi_{AB}(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}\left(\frac{\pi_A}{\pi_A-\pi_{AB}}+\frac{2\pi_{AB}}{\pi_B-\pi_{AB}}\right)Cov\left(\hat{\pi}_{AB}^*,\hat{\pi}_B^*\right)
$$

$$
\left.+\frac{1-\pi_{AB}}{(\pi_A-\pi_{AB})(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}Cov\left(\hat{\pi}_A^*,\hat{\pi}_B^*\right)\right] \tag{2.6}
$$

Proof. The estimator $\hat{OR}_2$ of the odds ratio can be approximated as:

$$
\hat{OR}_2=\frac{\hat{\pi}_{AB}^*\left(1-\hat{\pi}_A^*-\hat{\pi}_B^*+\hat{\pi}_{AB}^*\right)}{\left(\hat{\pi}_A^*-\hat{\pi}_{AB}^*\right)\left(\hat{\pi}_B^*-\hat{\pi}_{AB}^*\right)}
$$

$$
=\frac{\pi_{AB}\left(1+\varepsilon_{AB}^*\right)\left\{1-\pi_A\left(1+\varepsilon_A^*\right)-\pi_B\left(1+\varepsilon_B^*\right)+\pi_{AB}\left(1+\varepsilon_{AB}^*\right)\right\}}{\left\{\pi_A\left(1+\varepsilon_A^*\right)-\pi_{AB}\left(1+\varepsilon_{AB}^*\right)\right\}\left\{\pi_B\left(1+\varepsilon_B^*\right)-\pi_{AB}\left(1+\varepsilon_{AB}^*\right)\right\}}
$$

$$
=OR\left[1+\left\{\frac{\pi_{AB}(1-\pi_B)}{(\pi_A-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}+\frac{\pi_B}{\pi_B-\pi_{AB}}\right\}\varepsilon_{AB}^*-\frac{\pi_A(1-\pi_B)\varepsilon_A^*}{(\pi_A-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}\right.
$$

$$
-\frac{\pi_B(1-\pi_A)\varepsilon_B^*}{(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}+\left\{\frac{\pi_{AB}(1-\pi_B)\left(\pi_A\pi_B-\pi_{AB}^2\right)}{(\pi_A-\pi_{AB})^2(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}+\frac{\pi_{AB}\pi_B}{(\pi_B-\pi_{AB})^2}\right\}\varepsilon_{AB}^{*2}
$$

$$
+\frac{\pi_A^2(1-\pi_B)}{(\pi_A-\pi_{AB})^2(1-\pi_A-\pi_B+\pi_{AB})}\varepsilon_A^{*2}+\frac{\pi_B^2(1-\pi_A)}{(\pi_B-\pi_{AB})^2(1-\pi_A-\pi_B+\pi_{AB})}\varepsilon_B^{*2}
$$

$$
-\left\{\frac{\pi_A\pi_B}{(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}+\frac{\pi_{AB}\pi_A}{(\pi_A-\pi_{AB})(\pi_B-\pi_{AB})}+\frac{\pi_A(\pi_A+\pi_{AB}-2\pi_{AB}\pi_B)}{(\pi_A-\pi_{AB})^2(1-\pi_A-\pi_B+\pi_{AB})}\right\}\varepsilon_{AB}^*\varepsilon_A^*
$$

$$
-\frac{\pi_B(1-\pi_A)}{(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}\left(\frac{\pi_A}{\pi_A-\pi_{AB}}+\frac{2\pi_{AB}}{\pi_B-\pi_{AB}}\right)\varepsilon_{AB}^*\varepsilon_B^*
$$

$$
\left.+\frac{\pi_A\pi_B(1-\pi_{AB})}{(\pi_A-\pi_{AB})(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}\varepsilon_A^*\varepsilon_B^*+O\left(\varepsilon^{*2}\right)\right]
$$

By the definition of bias that is

$$
B\left(\hat{OR}_2\right)=E(\hat{OR}_2)-OR
$$

we have the theorem.

Theorem 2.4. The mean squared error of the estimator $\hat{\mathrm{OR}}_2$ of the odds ratio (OR) is given by:

$$MSE(\hat{OR}_2) = OR^2\left[\left\{\frac{1-\pi_B}{(\pi_A-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}+\frac{\pi_B}{\pi_{AB}(\pi_B-\pi_{AB})}\right\}^2 V\left(\hat{\pi}^*_{AB}\right)\right.$$
$$+\frac{(1-\pi_B)^2 V\left(\hat{\pi}^*_A\right)}{(\pi_A-\pi_{AB})^2(1-\pi_A-\pi_B+\pi_{AB})^2}+\frac{(1-\pi_A)^2 V\left(\hat{\pi}^*_B\right)}{(\pi_B-\pi_{AB})^2(1-\pi_A-\pi_B+\pi_{AB})^2}$$
$$-\frac{2(1-\pi_B)}{(\pi_A-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}\left\{\frac{1-\pi_B}{(\pi_A-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}+\frac{\pi_B}{\pi_{AB}(\pi_B-\pi_{AB})}\right\}Cov\left(\hat{\pi}^*_{AB},\hat{\pi}^*_A\right)$$
$$-\frac{2(1-\pi_A)}{(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}\left\{\frac{1-\pi_B}{(\pi_A-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}+\frac{\pi_B}{\pi_{AB}(\pi_B-\pi_{AB})}\right\}Cov\left(\hat{\pi}^*_{AB},\hat{\pi}^*_B\right)$$
$$\left.+\frac{2(1-\pi_A)(1-\pi_B)Cov\left(\hat{\pi}^*_A,\hat{\pi}^*_B\right)}{(\pi_A-\pi_{AB})(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})^2}\right] \tag{2.7}$$

Proof. By the definition of mean squared error, we have

$$MSE\left(\hat{OR}_2\right) = E\left(\hat{OR}_2 - OR\right)^2$$
$$\cong OR^2 E\left[\left\{\frac{\pi_{AB}(1-\pi_B)}{(\pi_A-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}+\frac{\pi_B}{\pi_B-\pi_{AB}}\right\}\varepsilon^*_{AB}\right.$$
$$\left.-\frac{\pi_A(1-\pi_B)\varepsilon^*_A}{(\pi_A-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}-\frac{\pi_B(1-\pi_A)\varepsilon^*_B}{(\pi_B-\pi_{AB})(1-\pi_A-\pi_B+\pi_{AB})}\right]^2$$

Expanding and taking the expected value, we have the theorem.

In the next section, we consider the problem of estimation of attributable risk.

3. ESTIMATION OF ATTRIBUTABLE RISK

In order to define an attributable risk, we have the following theorem.

Theorem 3.1. The attributable risk $AR(A\,|\,B)$ is given by:

$$AR(A\,|\,B) = \frac{\pi_{AB}-\pi_A\pi_B}{\pi_A(1-\pi_B)} \tag{3.1}$$

Proof. We know that the relative risk (RR) of being in group A given a respondent belongs to group B is defined as:

$$RR(A\,|\,B) = \frac{P(A\,|\,B)}{P(A\,|\,B^C)} = \frac{\dfrac{P(A\cap B)}{P(B)}}{\dfrac{P(A\cap B^C)}{P(B^C)}} = \frac{P(A\cap B)[1-P(B)]}{P(B)[P(A)-P(A\cap B)]} = \frac{\pi_{AB}(1-\pi_B)}{\pi_B(\pi_A-\pi_{AB})}$$

Following Rosner (2016), by the definition of attributable risk, we have

$$AR(A|B) = \frac{[RR(A|B)-1]\pi_B}{[RR(A|B)-1]\pi_B + 1} = \frac{\left[\frac{\pi_{AB}(1-\pi_B)}{\pi_B(\pi_A - \pi_{AB})} - 1\right]\pi_B}{\left[\frac{\pi_{AB}(1-\pi_B)}{\pi_B(\pi_A - \pi_{AB})} - 1\right]\pi_B + 1}$$

$$= \frac{[\pi_{AB}(1-\pi_B) - \pi_B(\pi_A - \pi_{AB})]\pi_B}{[\pi_{AB}(1-\pi_B) - \pi_B(\pi_A - \pi_{AB})]\pi_B + \pi_B(\pi_A - \pi_{AB})}$$

$$= \frac{[\pi_{AB} - \pi_{AB}\pi_B - \pi_A\pi_B + \pi_{AB}\pi_B]\pi_B}{[\pi_{AB} - \pi_{AB}\pi_B - \pi_A\pi_B + \pi_{AB}\pi_B]\pi_B + \pi_B(\pi_A - \pi_{AB})}$$

$$= \frac{[\pi_{AB} - \pi_A\pi_B]\pi_B}{[\pi_{AB} - \pi_A\pi_B]\pi_B + \pi_B(\pi_A - \pi_{AB})}$$

$$= \frac{[\pi_{AB} - \pi_A\pi_B]\pi_B}{[\pi_{AB} - \pi_A\pi_B + \pi_A - \pi_{AB}]\pi_B} = \frac{\pi_{AB} - \pi_A\pi_B}{-\pi_A\pi_B + \pi_A} = \frac{\pi_{AB} - \pi_A\pi_B}{\pi_A(1-\pi_B)}$$

which proves the theorem.

3.1 ESTIMATION OF ATTRIBUTABLE RISK USING SIMPLE MODEL

By using the same notations for the simple model from Lee et al. (2013), we consider first estimator of the attributable risk as:

$$\hat{AR}_1(A|B) = \frac{\hat{\pi}_{AB} - \hat{\pi}_A\hat{\pi}_B}{\hat{\pi}_A(1-\hat{\pi}_B)} \tag{3.2}$$

Now, we have the following theorems:

Theorem 3.2. The bias in the estimator $\hat{AR}_1(A|B)$ of the attributable risk $AR(A|B)$ is given by:

$$B(\hat{AR}_1(A|B)) = AR(A|B)\left[\frac{\pi_{AB}V(\hat{\pi}_A)}{\pi_A^2(\pi_{AB} - \pi_A\pi_B)} - \frac{(\pi_A - \pi_{AB})V(\hat{\pi}_B)}{(1-\pi_B)^2(\pi_{AB} - \pi_A\pi_B)} - \frac{Cov(\hat{\pi}_{AB}, \hat{\pi}_A)}{\pi_A(\pi_{AB} - \pi_A\pi_B)} + \frac{Cov(\hat{\pi}_{AB}, \hat{\pi}_B)}{(1-\pi_B)(\pi_{AB} - \pi_A\pi_B)} - \frac{\pi_{AB}Cov(\hat{\pi}_A, \hat{\pi}_B)}{\pi_A(1-\pi_B)(\pi_{AB} - \pi_A\pi_B)}\right] \tag{3.3}$$

Proof. The estimator $\hat{AR}_1(A|B)$ can be approximated as.

$$\hat{AR}_1(A|B) = \frac{\hat{\pi}_{AB} - \hat{\pi}_A\hat{\pi}_B}{\hat{\pi}_A(1-\hat{\pi}_B)} = \frac{\pi_{AB}(1+\varepsilon_{AB}) - \pi_A(1+\varepsilon_A)\pi_B(1+\varepsilon_B)}{\pi_A(1+\varepsilon_A)[1-\pi_B(1+\varepsilon_B)]}$$

$$= AR(A|B)\left[1 + \frac{\pi_{AB}\varepsilon_{AB}}{\pi_{AB} - \pi_A\pi_B} - \frac{\pi_{AB}\varepsilon_A}{\pi_{AB} - \pi_A\pi_B} - \frac{\pi_B(\pi_A - \pi_{AB})\varepsilon_B}{(1-\pi_B)(\pi_{AB} - \pi_A\pi_B)}\right.$$

$$+\frac{\pi_{AB}\varepsilon_A^2}{\pi_{AB}-\pi_A\pi_B}-\frac{\pi_B^2(\pi_A-\pi_{AB})\varepsilon_B^2}{(1-\pi_B)^2(\pi_{AB}-\pi_A\pi_B)}-\frac{\pi_{AB}\varepsilon_{AB}\varepsilon_A}{\pi_{AB}-\pi_A\pi_B}$$
$$+\frac{\pi_{AB}\pi_B\varepsilon_{AB}\varepsilon_B}{(1-\pi_B)(\pi_{AB}-\pi_A\pi_B)}-\frac{\pi_{AB}\pi_B\varepsilon_A\varepsilon_B}{(1-\pi_B)(\pi_{AB}-\pi_A\pi_B)}+O(\varepsilon^2)\Bigg]$$

By the definition of bias, we have

$$B\left(\hat{AR}_1(A\,|\,B)\right)=E\left(\hat{AR}_1(A\,|\,B)\right)-AR(A\,|\,B)$$
$$=AR(A\,|\,B)\Bigg[\frac{\pi_{AB}E(\varepsilon_A^2)}{\pi_{AB}-\pi_A\pi_B}-\frac{\pi_B^2(\pi_A-\pi_{AB})E(\varepsilon_B^2)}{(1-\pi_B)^2(\pi_{AB}-\pi_A\pi_B)}-\frac{\pi_{AB}E(\varepsilon_{AB}\varepsilon_A)}{\pi_{AB}-\pi_A\pi_B}$$
$$+\frac{\pi_{AB}\pi_B E(\varepsilon_{AB}\varepsilon_B)}{(\pi_{AB}-\pi_A\pi_B)(1-\pi_B)}-\frac{\pi_{AB}\pi_B E(\varepsilon_A\varepsilon_B)}{(1-\pi_B)(\pi_{AB}-\pi_A\pi_B)}\Bigg]$$
$$=AR(A\,|\,B)\Bigg[\frac{\pi_{AB}V(\hat{\pi}_A)}{\pi_A^2(\pi_{AB}-\pi_A\pi_B)}-\frac{(\pi_A-\pi_{AB})V(\hat{\pi}_B)}{(1-\pi_B)^2(\pi_{AB}-\pi_A\pi_B)}-\frac{Cov(\hat{\pi}_{AB},\hat{\pi}_A)}{\pi_A(\pi_{AB}-\pi_A\pi_B)}$$
$$+\frac{Cov(\hat{\pi}_{AB},\hat{\pi}_B)}{(1-\pi_B)(\pi_{AB}-\pi_A\pi_B)}-\frac{\pi_{AB}Cov(\hat{\pi}_A,\hat{\pi}_B)}{\pi_A(1-\pi_B)(\pi_{AB}-\pi_A\pi_B)}\Bigg]$$

which proves the theorem.

Theorem 3.2. The mean square of the estimator $\hat{AR}_1(A\,|\,B)$ of the attributable risk $AR(A\,|\,B)$ is given by:

$$MSE\left(\hat{AR}_1(A\,|\,B)\right)=\{AR(A\,|\,B)\}^2\Bigg[\frac{V(\hat{\pi}_{AB})}{(\pi_{AB}-\pi_A\pi_B)^2}+\frac{\pi_{AB}^2V(\hat{\pi}_A)}{\pi_A^2(\pi_{AB}-\pi_A\pi_B)^2}$$
$$+\frac{(\pi_A-\pi_{AB})^2V(\hat{\pi}_B)}{(1-\pi_B)^2(\pi_{AB}-\pi_A\pi_B)^2}-\frac{2\pi_{AB}Cov(\hat{\pi}_{AB},\hat{\pi}_A)}{\pi_A(\pi_{AB}-\pi_A\pi_B)^2}$$
$$-\frac{2(\pi_A-\pi_{AB})Cov(\hat{\pi}_{AB},\hat{\pi}_B)}{(1-\pi_B)(\pi_{AB}-\pi_A\pi_B)^2}+\frac{\pi_{AB}(\pi_A-\pi_{AB})Cov(\hat{\pi}_A,\hat{\pi}_B)}{\pi_A(1-\pi_B)(\pi_{AB}-\pi_A\pi_B)^2}\Bigg] \quad (3.4)$$

Proof. By the definition of mean square error, we have

$$MSE\left(\hat{AR}_1(A\,|\,B)\right)=E\left(\hat{AR}_1(A\,|\,B)-AR(A\,|\,B)\right)^2$$
$$\cong\{AR(A\,|\,B)\}^2E\left(\frac{\pi_{AB}\varepsilon_{AB}}{\pi_{AB}-\pi_A\pi_B}-\frac{\pi_{AB}\varepsilon_A}{\pi_{AB}-\pi_A\pi_B}-\frac{\pi_B(\pi_A-\pi_{AB})\varepsilon_B}{(1-\pi_B)(\pi_{AB}-\pi_A\pi_B)}\right)^2$$

Expanding and taking expected value, we have the theorem.

3.2 ESTIMATION OF ATTRIBUTABLE RISK USING CROSSED MODEL

By using the same notations for the crossed model from Lee et al. (2013), we consider first estimator of the attributable risk as:

$$\hat{AR}_2(A \mid B) = \frac{\hat{\pi}^*_{AB} - \hat{\pi}^*_A \hat{\pi}^*_B}{\hat{\pi}^*_A\left(1 - \hat{\pi}^*_B\right)} \tag{3.5}$$

Now, we have the following theorems:

Theorem 3.3. The bias in the estimator $\hat{AR}_2(A \mid B)$ of the attributable risk $AR(A \mid B)$ is given by:

$$B(\hat{AR}_2(A \mid B)) = AR(A \mid B)\left[\frac{\pi_{AB} V\left(\hat{\pi}^*_A\right)}{\pi^2_A\left(\pi_{AB} - \pi_A\pi_B\right)} - \frac{\left(\pi_A - \pi_{AB}\right)V\left(\hat{\pi}^*_B\right)}{\left(1-\pi_B\right)^2\left(\pi_{AB} - \pi_A\pi_B\right)} - \frac{Cov\left(\hat{\pi}^*_{AB}, \hat{\pi}^*_A\right)}{\pi_A\left(\pi_{AB} - \pi_A\pi_B\right)} + \frac{Cov\left(\hat{\pi}^*_{AB}, \hat{\pi}^*_B\right)}{\left(1-\pi_B\right)\left(\pi_{AB} - \pi_A\pi_B\right)} - \frac{\pi_{AB} Cov\left(\hat{\pi}^*_A, \hat{\pi}^*_B\right)}{\pi_A\left(1-\pi_B\right)\left(\pi_{AB} - \pi_A\pi_B\right)}\right] \tag{3.6}$$

Proof. The estimator $\hat{AR}_2(A \mid B)$ can be approximated as.

$$\hat{AR}_2(A \mid B) = \frac{\hat{\pi}^*_{AB} - \hat{\pi}^*_A \hat{\pi}^*_B}{\hat{\pi}^*_A\left(1 - \hat{\pi}^*_B\right)} = \frac{\pi_{AB}\left(1+\varepsilon^*_{AB}\right) - \pi_A\left(1+\varepsilon^*_A\right)\pi_B\left(1+\varepsilon^*_B\right)}{\pi_A\left(1+\varepsilon^*_A\right)\left[1 - \pi_B\left(1+\varepsilon^*_B\right)\right]}$$

$$= AR(A \mid B)\left[1 + \frac{\pi_{AB}\varepsilon^*_{AB}}{\pi_{AB} - \pi_A\pi_B} - \frac{\pi_{AB}\varepsilon^*_A}{\pi_{AB} - \pi_A\pi_B} - \frac{\pi_B\left(\pi_A - \pi_{AB}\right)\varepsilon^*_B}{\left(1-\pi_B\right)\left(\pi_{AB} - \pi_A\pi_B\right)} + \frac{\pi_{AB}\varepsilon^{*2}_A}{\pi_{AB} - \pi_A\pi_B} - \frac{\pi^2_B\left(\pi_A - \pi_{AB}\right)\varepsilon^{*2}_B}{\left(1-\pi_B\right)^2\left(\pi_{AB} - \pi_A\pi_B\right)} - \frac{\pi_{AB}\varepsilon^*_{AB}\varepsilon^*_A}{\pi_{AB} - \pi_A\pi_B} + \frac{\pi_{AB}\pi_B\varepsilon^*_{AB}\varepsilon^*_B}{\left(1-\pi_B\right)\left(\pi_{AB} - \pi_A\pi_B\right)} - \frac{\pi_{AB}\pi_B\varepsilon^*_A\varepsilon^*_B}{\left(1-\pi_B\right)\left(\pi_{AB} - \pi_A\pi_B\right)} + O\left(\varepsilon^{*2}\right)\right]$$

By the definition of bias, we have

$$B\left(\hat{AR}_2(A \mid B)\right) = E\left(\hat{AR}_2(A \mid B)\right) - AR(A \mid B)$$

$$= AR(A \mid B)\left[\frac{\pi_{AB}E\left(\varepsilon^{*2}_A\right)}{\pi_{AB} - \pi_A\pi_B} - \frac{\pi^2_B\left(\pi_A - \pi_{AB}\right)E\left(\varepsilon^{*2}_B\right)}{\left(1-\pi_B\right)^2\left(\pi_{AB} - \pi_A\pi_B\right)} - \frac{\pi_{AB}E\left(\varepsilon^*_{AB}\varepsilon^*_A\right)}{\pi_{AB} - \pi_A\pi_B} + \frac{\pi_{AB}\pi_B E\left(\varepsilon^*_{AB}\varepsilon^*_B\right)}{\left(\pi_{AB} - \pi_A\pi_B\right)\left(1-\pi_B\right)} - \frac{\pi_{AB}\pi_B E\left(\varepsilon^*_A\varepsilon^*_B\right)}{\left(1-\pi_B\right)\left(\pi_{AB} - \pi_A\pi_B\right)}\right]$$

$$= AR(A \mid B)\left[\frac{\pi_{AB} V\left(\hat{\pi}^*_A\right)}{\pi^2_A\left(\pi_{AB} - \pi_A\pi_B\right)} - \frac{\left(\pi_A - \pi_{AB}\right)V\left(\hat{\pi}^*_B\right)}{\left(1-\pi_B\right)^2\left(\pi_{AB} - \pi_A\pi_B\right)} - \frac{Cov\left(\hat{\pi}^*_{AB}, \hat{\pi}^*_A\right)}{\pi_A\left(\pi_{AB} - \pi_A\pi_B\right)}\right.$$

$$+\frac{Cov\left(\hat{\pi}^*_{AB}, \hat{\pi}^*_B\right)}{(1-\pi_B)(\pi_{AB}-\pi_A\pi_B)} - \frac{\pi_{AB}Cov\left(\hat{\pi}^*_A, \hat{\pi}^*_B\right)}{\pi_A(1-\pi_B)(\pi_{AB}-\pi_A\pi_B)}\Bigg]$$

which proves the theorem.

Theorem 3.2. The mean squared of the estimator $\hat{AR}_2(A|B)$ of the attributable risk $AR(A|B)$ is given by:

$$MSE\left(\hat{AR}_2(A|B)\right) = \{AR(A|B)\}^2\Bigg[\frac{V\left(\hat{\pi}^*_{AB}\right)}{(\pi_{AB}-\pi_A\pi_B)^2} + \frac{\pi^2_{AB}V\left(\hat{\pi}^*_A\right)}{\pi^2_A(\pi_{AB}-\pi_A\pi_B)^2} + \frac{(\pi_A-\pi_{AB})^2V\left(\hat{\pi}^*_B\right)}{(1-\pi_B)^2(\pi_{AB}-\pi_A\pi_B)^2} - \frac{2\pi_{AB}Cov\left(\hat{\pi}^*_{AB}, \hat{\pi}^*_A\right)}{\pi_A(\pi_{AB}-\pi_A\pi_B)^2} - \frac{2(\pi_A-\pi_{AB})Cov\left(\hat{\pi}^*_{AB}, \hat{\pi}^*_B\right)}{(1-\pi_B)(\pi_{AB}-\pi_A\pi_B)^2} + \frac{\pi_{AB}(\pi_A-\pi_{AB})Cov\left(\hat{\pi}^*_A, \hat{\pi}^*_B\right)}{\pi_A(1-\pi_B)(\pi_{AB}-\pi_A\pi_B)^2}\Bigg] \quad (3.7)$$

Proof. By the definition of mean squared error, we have

$$MSE\left(\hat{AR}_2(A|B)\right) = E\left(\hat{AR}_2(A|B) - AR(A|B)\right)^2$$

$$\cong \{AR(A|B)\}^2 E\left(\frac{\pi_{AB}\varepsilon^*_{AB}}{\pi_{AB}-\pi_A\pi_B} - \frac{\pi_{AB}\varepsilon^*_A}{\pi_{AB}-\pi_A\pi_B} - \frac{\pi_B(\pi_A-\pi_{AB})\varepsilon^*_B}{(1-\pi_B)(\pi_{AB}-\pi_A\pi_B)}\right)^2$$

Expanding and taking expected value, we have the theorem.

4. RELATIVE EFFICIENCY

We define the percent relative efficiency of the estimator $\hat{OR}_2$ with respect to the estimator $\hat{OR}_1$ as:

$$\text{RE(OR)} = \frac{MSE(\hat{OR}_1)}{MSE(\hat{OR}_2)} \times 100\% \quad (4.1)$$

We define the percent relative efficiency of the estimator $\hat{AR}_2(A|B)$ with respect to the estimator $\hat{AR}_1(A|B)$ as:

$$\text{RE(AR)} = \frac{MSE(\hat{AR}_1(A|B))}{MSE(\hat{AR}_2(A|B))} \times 100\% \quad (4.2)$$

We wrote FORTRAN codes, given in APPENDIX, to compute the percent relative efficiency values. We used $P = T = 0.7$ which is same choice as in Lee et al. (2013). The percent relative efficiency values so obtained for different choices of π_{AB}, π_A and π_B are presentedin Table 4.1.

Table 4.1. Percent Relative Efficiency values.

π_{AB}	π_A	π_B	RE(OR)	RE(AR)
0.1	0.2	0.2	987.5	1037.4
0.1	0.2	0.3	1065.0	1107.2
0.1	0.2	0.4	1152.7	1178.1
0.1	0.2	0.6	1320.9	1284.7
0.1	0.2	0.7	1345.6	1275.0
0.1	0.3	0.2	1013.1	1000.4
0.1	0.3	0.3	1057.1	1062.2
0.1	0.3	0.4	1126.7	1106.7
0.1	0.3	0.5	1183.5	1119.4
0.1	0.3	0.6	1167.8	1084.5
0.1	0.4	0.2	1067.2	1057.7
0.1	0.4	0.3	1095.1	1082.4
0.1	0.4	0.4	1134.7	1070.6
0.1	0.4	0.5	1098.7	1017.6
0.1	0.5	0.3	1133.2	1091.0
0.1	0.5	0.4	1084.6	1014.0
0.1	0.6	0.2	1198.6	1200.6
0.1	0.6	0.3	1120.2	1078.6
0.1	0.7	0.2	1242.2	1267.7
0.2	0.3	0.3	858.5	966.4
0.2	0.3	0.4	897.8	982.9
0.2	0.3	0.5	932.3	986.4
0.2	0.3	0.6	944.6	965.4
0.2	0.4	0.3	864.0	886.3
0.2	0.4	0.4	851.7	875.0
0.2	0.5	0.3	885.8	872.4
0.2	0.6	0.3	905.7	888.0
0.3	0.4	0.4	796.3	896.4
0.3	0.4	0.5	803.7	883.0
0.3	0.5	0.4	792.4	795.3

From the Table 4.1, one can conclude that the use of crossed model also remains more efficient than the simple model in case of estimating odds raio and attributable risk. The results are consistent with the results obtaind by the use of crossed model while estimating other parameters, such as the relative risk, the correlation coefficient, etc. Thus, we conclude that the crossed model is better than the simple model for all situations we have investigated.

5. APPLICATION BASED ON REAL DATASET

Lee et al. (2013) collected real data from 75 respondents at the Joint Statistical Meeting (2011), Miami, FL by using crossed model with $P = T = 0.7$ on smoking and drinking. Let π_{AB} , π_A and π_B be the true proportions of smokers, drinkers, and smokers and drinkers, respectively. Lee et al. (2013) reported respective estimates as $\hat{\pi}^*_{AB} = 0.2367816$, $\hat{\pi}^*_A = 0.24$, and $\hat{\pi}^*_B = 0.36$. These estimates are used for estimating estimators of odds ratio and attributable risk. With the crossed model, the estimator of odds is obtained as $\hat{OR}_2 = 380.21$ and the attributable risk is found as 0.9790 for $\hat{AR}_2(A \mid B)$ and 0.5496 for $\hat{AR}_2(B \mid A)$. The high value of $\hat{OR}_2 = 380.21$ indicates that smoking and drinking are highly associated to each other. Lee et al. (2013) have shown that there is high correlation between smoking and drinking. The estimate of the attributable risk of a drinker to be a smoker is 0.9790, which mean a smoker has 97.70% chance to be a drinker than non-user of both; whereas the estimate of the attributable risk of a smoker to be a drinker is 0.5496, which implies a drinker has 54.96% chance to be a smoker than non-user of both.

REFERENCES

Lee, C.S., Sedory, S.A. and Singh, S. (2013). Estimating at least seven measures of qualitative variables from a single sample using randomized response technique. *Statistics and Probbaility Letters*, 83(1), 399-409.

Rosner, B. (2015). Fundamentals of Biostatistics. 8th Edition, Thompson, Brooks/Cole.

Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association* 60:63–69.

APPENDIX

```
! FILE NAME LEEARR.F95
       IMPLICIT NONE
       REAL P,T,PIA,PIB,PIAB,SUM
       DOUBLE PRECISION VARPIA,VARPIB,VARPIAB,CPIABPIA,
      1CPIABPIB,CPIAPIB,VOR1,VOR2,VARPIAS,VARPIBS,VARPIABS,
      1CPIABPAS,CPIABPBS,CPIAPIBS,G1,G2,G3,G4,G5,RE_OR,
      1F1,F2,F3, T1, T2, T3, V_ARR1, V_ARR2, RE_ARR
       CHARACTER*20 OUT_FILE
      WRITE(*,'(A)') 'NAME OF THE OUTPUT FILE'
       READ(*,'(A20)') OUT_FILE
      OPEN(42, FILE=OUT_FILE, STATUS='UNKNOWN')
       P = 0.70
       T = 0.70
      WRITE(42,107)P,T
107    FORMAT(2X,'P=',F6.3,2X,'T=',F6.3)
      WRITE(42,108)
108    FORMAT( 4X,'PIAB',6X,'PIA',6X,'PIB',6X,'RE_OR',6X,'RE_ARR')
       DO 10 PIAB = 0.10, 0.99, 0.10
       DO 10 PIA = 0.10, 0.991, 0.10
```

```
         DO 10 PIB = 0.10, 0.991, 0.10
         SUM = PIA+PIB
         IF ( (PIA*PIB).NE.(PIAB) ) THEN
         IF((PIAB.LT.PIA).AND.(PIAB.LT.PIB).AND.(SUM.LT.0.999)) THEN
         VARPIA = PIA*(1-PIA)+P*(1-P)/(2*P-1)**2
         VARPIB = PIB*(1-PIB)+T*(1-T)/(2*T-1)**2
         VARPIAB = PIAB*(1-PIAB)
        1+( (2*P-1)**2*T*(1-T)*PIA+P*(1-P)*(2*T-1)**2*PIB+P*T*(1-P)*(1-T))
        1/((2*P-1)**2*(2*T-1)**2)
         CPIABPIA = PIAB*(1-PIA)+P*(1-P)*PIB/(2*P-1)**2
         CPIABPIB = PIAB*(1-PIB)+T*(1-T)*PIA/(2*T-1)**2
         CPIAPIB = PIAB-PIA*PIB
         F1=(1-PIB)/((PIA-PIAB)*(1-PIA-PIB+PIAB))+PIB/(PIAB*(PIB-PIAB))
         F2 = (1-PIB)/((PIA-PIAB)*(1-PIA-PIB+PIAB))
         F3 = (1-PIA)/((PIB-PIAB)*(1-PIA-PIB+PIAB))
        VOR1 = F1**2*VARPIAB + F2**2*VARPIA + F3**2*VARPIB
      1 - 2*F1*F2*CPIABPIA - 2 *F1*F3*CPIABPIB + 2*F2*F3*CPIAPIB
        T1 = 1.0/(PIAB-PIA*PIB)
        T2 = PIAB/(PIA*(PIAB-PIA*PIB))
        T3 = (PIA-PIAB)/((1-PIB)*(PIAB-PIA*PIB))
        V_ARR1 = T1**2*VARPIAB + T2**2*VARPIA + T3**2*VARPIB
      1 - 2*T1*T2*CPIABPIA - 2 *T1*T3*CPIABPIB + 2*T2*T3*CPIAPIB
        G1 = P*T+(1-P)*(1-T)
        G2 = 1.0-PIA-PIB+2*PIAB
        VARPIAS = PIA*(1-PIA) + (1-P)*T*G1*G2/(P+T-1)**2
        VARPIBS = PIB*(1-PIB) + (1-T)*P*G1*G2/(P+T-1)**2
        G3 = PIAB*(P**2*T**2+(1-P)**2*(1-T)**2)
        G4 = P*T*(1-P)*(1-T)*(1-PIA-PIB)
       G5 = (P*T+(1-P)*(1-T))*(P+T-1)**2
       VARPIABS = (G3+G4)/G5-PIAB**2
       CPIABPAS=PIAB*(1-PIA) + PIAB*T*(1-P)*(P-T+1)/(P+T-1)**2
      1+P*T*(1-P)*(1-T)*(T-P+1)*(1-PIA-PIB)/(G1*(P+T-1)**2)
       CPIABPBS = PIAB*(1-PIB)+PIAB*P*(1-T)*(T-P+1)/(P+T-1)**2
      1+P*T*(1-P)*(1-T)*(P-T+1)*(1-PIA-PIB)/(G1*(P+T-1)**2)
       CPIAPIBS = PIA*(1-PIA)-PIAB*G1
      1 - ( G1 +(P-T)*(1-2*PIA) )*(1-PIA-PIB)/2
      1 +G1**2*(1-PIA-PIB+2*PIAB)/(2*(P+T-1)**2)
       VOR2 = F1**2*VARPIABS + F2**2*VARPIAS + F3**2*VARPIBS
     1 - 2*F1*F2*CPIABPAS - 2 *F1*F3*CPIABPBS + 2*F2*F3*CPIAPIBS
       V_ARR2 = T1**2*VARPIABS + T2**2*VARPIAS + T3**2*VARPIBS
     1 - 2*T1*T2*CPIABPAS - 2 *T1*T3*CPIABPBS + 2*T2*T3*CPIAPIBS
        RE_OR = VOR1*100/VOR2
        RE_ARR = V_ARR1*100/V_ARR2
        WRITE(42, 101)PIAB, PIA, PIB, RE_OR, RE_ARR
101    FORMAT(2X,F8.4,2X,F8.4,2X,F8.4,2X,F9.2,2X,F9.2)
       ENDIF
       ENDIF
10     CONTINUE
       STOP
       END
```