



Available online at <http://scik.org>

J. Math. Comput. Sci. 11 (2021), No. 6, 7857-7871

<https://doi.org/10.28919/jmcs/5769>

ISSN: 1927-5307

A RANDOMIZED RESPONSE MODEL BASED ON BLANK CARD AND GEOMETRIC DISTRIBUTION TO ESTIMATE SENSITIVE PROPORTION

MUSTAFA KAMAL*

Department of Basic Sciences, College of Science and Theoretical Studies, Saudi Electronic University, Dammam,
Kingdom of Saudi Arabia

Copyright © 2021 the author(s). This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract: Using geometric probability (GP) distribution as a randomization device, a new randomized response (RR) model has been proposed which has an immense potential to estimate the human population proportion that possess a stigmatized character. Privacy protection measure of proposed model and some of its properties have been investigated. In addition, empirical experiments are conducted to validate the theoretical results, which demonstrate the better performance of the suggested estimators over their direct competitors. Finally, results are analyzed and appropriate suggestions are made available to survey practitioners when dealing with sensitive aspects.

Keywords: randomized response model; geometric distribution; sensitive attribute; privacy protection.

2010 AMS Subject Classification: 62P25.

1. INTRODUCTION

GP distribution has key importance in several scientific disciplines, including reliability theory, quality control, psychiatry, biology, and ecology. Many researchers, including Balakrishnan et al.

*Corresponding author

E-mail addresses: m.kamal@seu.edu.sa; kamal19252003@gmail.com

Received March 25, 2021

[2], Koutras et al. [12], Shannon [17], and others, have employed GP distribution in a variety of applications, such as statistical process control, optimum control, signal transmission, and modulation codes. When dealing with delicate or personal matters in socioeconomic surveys, it is extremely difficult to get honest responses from respondents. The basic topics of socioeconomic surveys include demographic characteristics of a group of individuals, their social environment as it relates to social and economic variables, their opinions and attitudes, and so on. If the subject matter under research does not involve any personal, sensitive, or contentious concerns, the data acquired from such surveys is more consistent and less prone to inaccuracy. However, if the subject matter of these surveys includes sensitive issues such as questions about their erroneous income tax payment, illegal use of drugs, alcohol, or smoking habits, suffering from some kind of mental disorder, marginalized disease such as Aids, and so on, the actual answers to these questions are masked or misinformed by the individuals who participate in the surveys.

To prevent such instances in practice, Warner [23] presented an intriguing solution known as the RR technique to collect information on stigmatized topics. Greenberg et al. [6], Moors [15], Cochran [4], Fox and Tracy [5], Chaudhuri and Mukherjee [3], Tracy and Osahan [22], Singh et al. [18], Kim and Elam [9], Kim and Chae [10], Singh and Tarray [19] and Singh et al. [21] further modified the RR mechanism for diverse practical scenarios.

Kuk [11] suggested a RR model that included two randomization devices. Singh and Grewal [20] enhanced the Kuk [11] approach by employing GP distribution as a random device to acquire sensitive data from respondents. Hussain et al. [8] and Pervez et al. [16] developed a randomization mechanism by utilizing negative binomial distribution. A k order GP distribution was employed by Hussain et al. [7] to calculate the population proportion of a sensitive attribute. Yennum et al. [24] developed a two-stage randomization mechanism to improve Singh and Grewal [20] and Hussain et al. [8] models.

It is to be mentioned that the RR models which are based on the GP distribution have not considered the cases of non-sensitive statements and blank cards. Motivated by these arguments and the previously cited works in this article, we suggested a three-stage randomization device to

produce more precise estimate of sensitive proportion using the blank card approach and GP distribution. Measure of privacy protection of proposed RR model and its properties have been investigated and quantified. Empirical studies have been performed and it has been found that the proposed RR model dominates Singh and Grewal [20] and Yennum et al. [24] models. The findings have been analyzed, and appropriate recommendations for future research and practical implications have been made.

2. PROPOSED MODEL

In this paper, we propose Z_i , a new alternative RR model. Assume that a sample of n individuals is collected from a population using simple random sampling without replacement. Now, selected respondents are provided with three randomization devices. A deck of red and green cards is used in the first randomization device with an assumption that the probability of selecting a red card being R and the probability of getting a green card being $(1 - R)$.

There are two decks of cards in the second randomization device. Cards in the first deck of the second randomization device are with three statements: (i) Are you an individual of a sensitive category A with probability P_1 (ii) Are you an individual of unrelated category B with probability P_2 (iii) Blank card with probability P_3 . The second deck of second randomization device contains three statements: (i) Are you an individual of a sensitive category A with probability T_1 (ii) Are you an individual of unrelated category B with probability T_2 (iii) Blank card with probability T_3 .

Similarly, the third randomization device also uses two decks of cards. The first deck of third randomization device contains cards with the statement (i) Are you an individual of a sensitive category A with probability Q_1 and (ii) Are you an individual of a non-sensitive category A^c with probability $(1 - Q_1)$. Cards in the second deck of third device contains the statements (i) Are you an individual of a non-sensitive category A^c with probability δ_1 and (ii) Are you an individual of a sensitive category A with probability $(1 - \delta_1)$.

Now at the first stage of response, the respondents requested to select a card from the first randomization device. If the selected card is red, then the respondent at the second stage is invited

to select cards one by one from the first deck of the second randomization device. If the chosen card is blank, the responder is requested to draw cards one at a time with replacement from the first deck of the third randomization device until the first card that represents the responder status statement. Likewise, if the participant acquires a green card, then the respondent at the second stage is invited to select cards from the second deck of the second randomization device. If the chosen card is blank, the responder is requested to draw cards one at a time with replacement from the second deck of the third randomization device until the first card that represents the responder's status statement.

Let X_1 and X_3 , X_4 and X_2 represent the total number of cards pulled whereas a representative of A , A^c and Y respectively. Assume that π , $(1 - \pi)$, and π_y indicate the true proportion of responders who belong to the sensitive, non-sensitive, and unrelated groups, respectively. Thus, the distribution of observed response Z_i from the n respondents is given as

$$Z_i = \begin{cases} X_1 & \text{with prob. } \pi W \\ X_2 & \text{with prob. } \pi_y W \\ X_3 & \text{with prob. } \pi W \\ X_4 & \text{with prob. } (1 - \pi)W \\ Y_1 & \text{with prob. } \pi(1 - W) \\ Y_2 & \text{with prob. } \pi_y(1 - W) \\ Y_3 & \text{with prob. } (1 - \pi)(1 - W) \\ Y_4 & \text{with prob. } \pi(1 - W) \end{cases} \quad (1)$$

Given that

$$\begin{aligned} X_1 &\sim G(P_1); & X_2 &\sim G(P_2); & X_3 &\sim G(Q_1); & X_4 &\sim G(1 - Q_1) \\ Y_1 &\sim G(T_1); & Y_2 &\sim G(T_1); & Y_3 &\sim G(\delta_1); & Y_4 &\sim G(1 - \delta_1) \end{aligned}$$

So, we have

$$\begin{aligned} E(X_1) &= \frac{1}{P_1}; & E(X_2) &= \frac{1}{P_2}; & E(X_3) &= \frac{1}{Q_1}; & E(X_4) &= \frac{1}{(1 - Q_1)} \\ E(Y_1) &= \frac{1}{T_1}; & E(Y_2) &= \frac{1}{T_2}; & E(Y_3) &= \frac{1}{\delta_1}; & E(Y_4) &= \frac{1}{(1 - \delta_1)} \\ E(X_1^2) &= \frac{2 - P_1}{P_1^2}; & E(X_2^2) &= \frac{2 - P_2}{P_2^2}; & E(X_3^2) &= \frac{2 - Q_1}{Q_1^2}; & E(X_4^2) &= \frac{1 + Q_1}{(1 - Q_1)^2} \\ E(Y_1^2) &= \frac{2 - T_1}{T_1^2}; & E(Y_2^2) &= \frac{2 - T_2}{T_2^2}; & E(Y_3^2) &= \frac{2 - \delta_1}{\delta_1^2}; & E(Y_4^2) &= \frac{1 + \delta_1}{(1 - \delta_1)^2} \end{aligned} \quad (2)$$

Now, the expected value of the observed response, Z_i , is obtained as

$$\begin{aligned}
E(Z_i) &= \pi WE(X_1) + \pi_Y WE(X_2) + P_3[\pi WE(X_3) + (1 - \pi)WE(X_4)] + \pi(1 - W)E(Y_1) \\
&\quad + \pi_Y(1 - W)E(Y_2) + T_3[(1 - \pi)(1 - W)E(Y_3) + \pi(1 - W)E(Y_4)] \\
&= \pi \left[\frac{W}{P_1} + \frac{P_3 W}{Q_1} + \frac{(1 - W)}{T_1} + \frac{T_3(1 - W)}{(1 - \delta_1)} \right] + \left[\frac{P_3(1 - \pi)W}{(1 - Q_1)} + \frac{T_3(1 - \pi)(1 - W)}{\delta_1} \right] \\
&\quad + \left[\frac{\pi_Y W}{P_2} + \frac{\pi_Y(1 - W)}{T_2} \right] \tag{3}
\end{aligned}$$

After simplification of equation (3), we get

$$\hat{\pi}_{DD} = \frac{\frac{\{P_1 Q_1 T_1 (1 - \delta_1)(1 - Q_1) \delta_1 P_2 T_2\}}{n} \sum_{i=1}^n Z_i - \pi_Y (\beta_1 - \beta_2)}{(\beta_3 + \beta_4 - \beta_5)} \tag{4}$$

where

$$\begin{aligned}
\beta_1 &= P_1 Q_1 T_1 (1 - \delta_1)(1 - Q_1) \delta_1 T_2 W + P_1 Q_1 T_1 (1 - \delta_1)(1 - Q_1) \delta_1 P_2 (1 - W) \\
\beta_2 &= P_1 Q_1 T_1 (1 - \delta_1) P_2 T_2 \delta_1 P_3 W + P_1 Q_1 T_1 (1 - \delta_1) P_2 T_2 (1 - Q_1) T_3 (1 - W) \\
\beta_3 &= (1 - Q_1) \delta_1 P_2 T_2 W Q_1 T_1 (1 - \delta_1) + (1 - Q_1) \delta_1 P_2 T_2 W P_1 T_1 (1 - \delta_1) \\
\beta_4 &= (1 - Q_1) \delta_1 P_2 T_2 (1 - W) P_1 Q_1 (1 - \delta_1) + (1 - Q_1) \delta_1 P_2 T_2 P_1 Q_1 T_1 T_3 (1 - W) \\
\beta_5 &= P_1 Q_1 T_1 (1 - \delta_1) P_2 T_2 \delta_1 P_3 W + P_1 Q_1 T_1 (1 - \delta_1) P_2 T_2 (1 - Q_1) T_3 (1 - W)
\end{aligned}$$

As a result, we could derive and verify the following theorems.

Theorem 1: The suggested population proportion estimator $\hat{\pi}_{DD}$ is an unbiased estimator, i.e.

$$E(\hat{\pi}_{DD}) = \pi \tag{5}$$

Proof: By applying expectation on both sides of the equation (4), we obtained

$$\begin{aligned}
E(\hat{\pi}_{DD}) &= E \left[\frac{\frac{\{P_1 Q_1 T_1 (1 - \delta_1)(1 - Q_1) \delta_1 P_2 T_2\}}{n} \sum_{i=1}^n Z_i - \pi_Y (\beta_1 - \beta_2)}{(\beta_3 + \beta_4 - \beta_5)} \right] \\
&= \frac{\frac{\{P_1 Q_1 T_1 (1 - \delta_1)(1 - Q_1) \delta_1 P_2 T_2\}}{n} \sum_{i=1}^n E(Z_i) - \pi_Y (\beta_1 - \beta_2)}{(\beta_3 + \beta_4 - \beta_5)}
\end{aligned}$$

Substitution of the value of $E(Z_i)$ will prove the theorem.

Theorem 2: The variance $V(\hat{\pi}_{DD})$ of the proposed estimator $\hat{\pi}_{DD}$ is given as

$$V(\hat{\pi}_{DD}) = \frac{1}{[\beta_3 + \beta_4 - \beta_5]^2} \left[\frac{\{P_1^2 Q_1^2 T_1^2 (1 - \delta_1)^2 (1 - Q_1)^2 \delta_1^2 P_2^2 T_2^2\}}{n^2} (\sigma_Z^2) \right] \tag{6}$$

Proof: Since all the observed response Z_i are independent, therefore, we have

$$\begin{aligned}
 V(\hat{\pi}_{DD}) &= V \left[\frac{\left\{ \frac{P_1 Q_1 T_1 (1 - \delta_1) (1 - Q_1) \delta_1 P_2 T_2}{n} \sum_{i=1}^n Z_i - \pi_Y (\beta_1 - \beta_2) \right\}}{(\beta_3 + \beta_4 - \beta_5)} \right] \\
 &= \frac{\left\{ \frac{P_1^2 Q_1^2 T_1^2 (1 - \delta_1)^2 (1 - Q_1)^2 \delta_1^2 P_2^2 T_2^2}{n^2} \sum_{i=1}^n V(Z_i) \right\}}{[\beta_3 + \beta_4 - \beta_5]^2} \\
 &= \left[\frac{\left\{ \frac{P_1^2 Q_1^2 T_1^2 (1 - \delta_1)^2 (1 - Q_1)^2 \delta_1^2 P_2^2 T_2^2}{n^2} (\sigma_Z^2) \right\}}{[\beta_3 + \beta_4 - \beta_5]^2} \right] \quad (7)
 \end{aligned}$$

where $\sigma_Z^2 = E(Z_i^2) - [E(Z_i)]^2$ and

$$\begin{aligned}
 E(Z_i^2) &= \pi \left[\left(\frac{2 - P_1}{P_1^2} \right) W + \left(\frac{2 - Q_1}{Q_1^2} \right) P_3 W + \left(\frac{2 - T_1}{T_1^2} \right) (1 - W) + \left(\frac{1 + \delta_1}{(1 - \delta_1)^2} \right) T_3 (1 - W) \right] \\
 &\quad + (1 - \pi) \left[\left(\frac{1 + Q_1}{(1 - Q_1)^2} \right) P_3 W + \left(\frac{2 - \delta_1}{\delta_1^2} \right) T_3 (1 - W) \right] \\
 &\quad + \pi_Y \left[\left(\frac{2 - P_2}{P_2^2} \right) W + \left(\frac{2 - T_2}{T_2^2} \right) (1 - W) \right] \quad (8)
 \end{aligned}$$

Substituting the expressions from equations (3) and (8) in equation (7), the proof of the theorem is straightforward.

3. MEASURE OF PRIVACY PROTECTION

The most essential aspect of RR technique is degree of privacy protection. Therefore, in this section, the respondents' degree of privacy protection is given significant consideration. For the suggested RR model on rare sensitive attribute, an attempt has been made in this regard. Numerous researchers, such as Leysieffer and Warner [14], Anderson [1], Lanke [13], Zhimin and Zaizai [25], have suggested measures to secure respondents' privacy when the study possess the sensitive characteristics. According to Zhimin and Zaizai [25], efficiency and respondents' privacy protection do not have to go in opposite directions. The following is a quick overview of Zhimin and Zaizai [25]:

3.1. Zhimin and Zaizai [25] Privacy measure

Let π_1 and $(1 - \pi_1)$ be the unknown population proportions of the sensitive group A and non-sensitive group A^c respectively. In the dichotomous response model, response is denoted as R which may be “yes” (say y) or “no” (say n). So, the conditional probabilities may be written as $P(R|A)$ and $P(R|A^c)$ for the response R from an individual group A and A^c . By using the Bays’ rule and based on the conditional probabilities, Zhimin and Zaizai [25] described posterior probabilities as

$$P(A|y) = \frac{\pi_1}{\pi_1 + (1-\pi_1)[P(y|A^c)/P(y|A)]} \quad (9)$$

$$P(A|n) = \frac{\pi_1}{\pi_1 + (1-\pi_1)[P(n|A^c)/P(n|A)]} \quad (10)$$

Considering $\tau(y) = P(y|A^c)/P(y|A)$ and $\tau(n) = P(n|A^c)/P(n|A)$

Thus, the measure of degree of privacy is

$$M(R) = \left| 1 - \frac{1}{2} [\tau(y) + \tau(n)] \right| \quad (11)$$

As $M(R)$ approaches zero, more anonymity is preserved. The aforementioned privacy measure is used to compare the performance of the proposed model.

3.2. Proposed model privacy measure

Following the Zhimin and Zaizai [25] privacy measure, the designed probabilities of the proposed RR model are calculated as

$$P(y|A) = WP_1 + WP_2\pi_Y + P_3WQ_1 + (1 - W)T_1 + (1 - W)T_2\pi_Y + (1 - W)(1 - \delta_1)$$

$$P(y|A^c) = WP_2\pi_Y + (1 - W)T_2\pi_Y + W(1 - Q_1) + (1 - W)(1 - \delta_1)$$

$$P(n|A) = 1 - [WP_1 + WP_2\pi_Y + P_3WQ_1 + (1 - W)T_1 + (1 - W)T_2\pi_Y + (1 - W)(1 - \delta_1)]$$

$$P(n|A^c) = 1 - [WP_2\pi_Y + (1 - W)T_2\pi_Y + W(1 - Q_1) + (1 - W)(1 - \delta_1)]$$

By using the Bays’ rule, the posterior probabilities can be described as

$$P(A|y) = \frac{\pi}{\pi + (1-\pi) \left[\frac{\{WP_2\pi_Y + (1 - W)T_2\pi_Y + W(1 - Q_1)\} + (1 - W)(1 - \delta_1)}{\{WP_1 + WP_2\pi_Y + P_3WQ_1 + (1 - W)T_1\} + (1 - W)T_2\pi_Y + (1 - W)(1 - \delta_1)} \right]} \quad (12)$$

$$P(A|n) = \frac{\pi}{\pi + (1-\pi) \left[\frac{1 - \{WP_2\pi_Y + (1-W)T_2\pi_Y + W(1-Q_1)\}}{+(1-W)(1-\delta_1)} \right]} \quad (13)$$

$$\left[\frac{1 - \{WP_1 + WP_2\pi_Y + P_3WQ_1 + (1-W)T_1\}}{+(1-W)T_2\pi_Y + (1-W)(1-\delta_1)} \right]$$

Thus, the measure of protection can be defined as

$$M(R) = \left| 1 - \frac{1}{2} \{ \tau(y) + \tau(n) \} \right| \quad (14)$$

where

$$\tau(y) = \left[\frac{\{WP_1 + WP_2\pi_Y + P_3WQ_1 + (1-W)T_1 + (1-W)T_2\pi_Y + (1-W)(1-\delta_1)\}}{\{WP_2\pi_Y + (1-W)T_2\pi_Y + W(1-Q_1) + (1-W)(1-\delta_1)\}} \right] \quad (15)$$

$$\tau(n) = \left[\frac{1 - \{WP_1 + WP_2\pi_Y + P_3WQ_1 + (1-W)T_1 + (1-W)T_2\pi_Y + (1-W)(1-\delta_1)\}}{1 - \{WP_2\pi_Y + (1-W)T_2\pi_Y + W(1-Q_1) + (1-W)(1-\delta_1)\}} \right] \quad (16)$$

The performance of the proposed model has been evaluated and the values of $M(R)$ are calculated for different parametric choices as given in table 1-4.

4. EFFICIENCY COMPARISONS

To demonstrate the efficacy of the suggested estimators over the Singh and Grewal [20] and Yennum et al. [24] models, empirical investigations have been carried out. The percent relative efficiencies (PRE) of the proposed model are calculated for various parameter values and probabilities in contrast to the models of Singh and Grewal [20] and Yennum et al. [24] using the following formulae:

$$\text{PRE}^1 = \frac{V(\text{Yennum et al. [24] estimator})}{V(\text{Proposed estimator})} \times 100$$

$$\text{PRE}^2 = \frac{V(\text{Singh and Grewal [20] estimator})}{V(\text{Proposed estimator})} \times 100$$

5. INTERPRETATION OF RESULTS

The following interpretations may be read out from Tables 1-4.

- i. From Tables 1-4, it is apparent to observe that the percent relative efficiencies of the proposed RR model with respect to Yennum et al. [24] and Singh & Grewal [20] models are always greater than 100, indicating the proposed model's dominating behaviors over Yennum et al. [24] and Singh & Grewal [20] model.

RANDOMIZED RESPONSE MODEL

Table 1: PRE of the proposed model in comparison to the Yennum et al. [24] estimator along with privacy protection for $\pi_y = 0.10$.

P_1	P_2	T_1	T_2	Q_1	δ_1	W							
						0.2		0.3		0.4		0.5	
						π							
						0.15		0.25		0.35		0.45	
						PRE	PP	PRE	PP	PRE	PP	PRE	PP
0.25	0.65	0.3	0.3	0.2	0.25	828.3	0.28	1074.8	0.10	1345.4	0.06	1639.4	0.22
0.35	0.55	0.3	0.3	0.2	0.25	715.2	0.33	905.0	0.17	1136.8	0.02	1423.0	0.13
0.45	0.45	0.3	0.3	0.2	0.25	647.0	0.37	789.5	0.23	965.4	0.10	1187.5	0.03
0.55	0.35	0.3	0.3	0.2	0.25	597.4	0.42	701.5	0.30	826.1	0.18	978.1	0.07
0.65	0.25	0.3	0.3	0.2	0.25	545.4	0.47	608.9	0.37	677.9	0.27	753.3	0.18
0.75	0.15	0.3	0.3	0.2	0.25	440.0	0.52	438.9	0.44	434.3	0.37	426.3	0.29
0.25	0.65	0.4	0.4	0.2	0.25	1013.7	0.43	1266.7	0.24	1497.5	0.05	1718.3	0.12
0.35	0.55	0.4	0.4	0.2	0.25	881.2	0.48	1107.7	0.30	1358.2	0.13	1638.9	0.03
0.45	0.45	0.4	0.4	0.2	0.25	791.1	0.52	969.9	0.37	1178.7	0.22	1426.3	0.07
0.55	0.35	0.4	0.4	0.2	0.25	716.9	0.57	844.9	0.43	992.2	0.30	1163.8	0.17
0.65	0.25	0.4	0.4	0.2	0.25	628.4	0.62	694.1	0.50	763.2	0.39	836.0	0.28
0.75	0.15	0.4	0.4	0.2	0.25	456.0	0.66	438.0	0.57	422.4	0.48	408.2	0.39
0.25	0.65	0.5	0.5	0.2	0.25	2439.0	0.57	2126.5	0.36	1971.7	0.17	1934.4	0.03
0.35	0.55	0.5	0.5	0.2	0.25	2357.2	0.62	2159.0	0.43	2077.8	0.24	2091.6	0.07
0.45	0.45	0.5	0.5	0.2	0.25	2160.3	0.66	1993.4	0.49	1937.9	0.33	1966.1	0.16
0.55	0.35	0.5	0.5	0.2	0.25	1873.8	0.71	1691.7	0.56	1617.8	0.41	1614.1	0.27
0.65	0.25	0.5	0.5	0.2	0.25	1435.0	0.75	1229.8	0.63	1125.1	0.50	1074.1	0.37
0.75	0.15	0.5	0.5	0.2	0.25	768.3	0.80	601.0	0.70	510.2	0.59	453.8	0.49
0.25	0.65	0.3	0.3	0.3	0.35	2922.9	0.44	3222.4	0.23	3520.3	0.03	3847.0	0.15
0.35	0.55	0.3	0.3	0.3	0.35	2628.9	0.50	2886.2	0.31	3215.4	0.14	3651.5	0.03
0.45	0.45	0.3	0.3	0.3	0.35	2421.4	0.57	2588.1	0.41	2830.0	0.25	3180.7	0.11
0.55	0.35	0.3	0.3	0.3	0.35	2244.9	0.64	2316.3	0.50	2443.2	0.37	2642.2	0.25
0.65	0.25	0.3	0.3	0.3	0.35	2018.2	0.71	1972.8	0.60	1959.7	0.50	1975.4	0.40
0.75	0.15	0.3	0.3	0.3	0.35	1517.2	0.78	1313.7	0.71	1154.9	0.63	1022.6	0.56
0.25	0.65	0.4	0.4	0.3	0.35	3062.1	0.63	3289.6	0.40	3464.3	0.18	3655.1	0.02
0.35	0.55	0.4	0.4	0.3	0.35	2769.4	0.69	3062.0	0.48	3396.1	0.29	3810.6	0.10
0.45	0.45	0.4	0.4	0.3	0.35	2520.6	0.76	2747.6	0.57	3052.0	0.40	3473.4	0.23
0.55	0.35	0.4	0.4	0.3	0.35	2275.6	0.82	2389.4	0.67	2568.6	0.52	2834.4	0.37
0.65	0.25	0.4	0.4	0.3	0.35	1936.7	0.89	1895.1	0.77	1895.6	0.64	1932.4	0.52
0.75	0.15	0.4	0.4	0.3	0.35	1282.9	0.96	1086.7	0.87	950.6	0.78	848.2	0.69
0.25	0.65	0.5	0.5	0.3	0.35	4598.5	0.80	3969.3	0.55	3643.0	0.32	3539.6	0.09
0.35	0.55	0.5	0.5	0.3	0.35	4456.9	0.86	4082.6	0.64	3946.9	0.43	4007.1	0.22
0.45	0.45	0.5	0.5	0.3	0.35	4081.3	0.92	3780.8	0.73	3725.4	0.54	3869.2	0.35
0.55	0.35	0.5	0.5	0.3	0.35	3525.9	0.99	3185.7	0.82	3081.6	0.65	3145.7	0.49
0.65	0.25	0.5	0.5	0.3	0.35	2677.7	0.99	2272.0	0.92	2072.8	0.78	1986.9	0.64
0.75	0.15	0.5	0.5	0.3	0.35	1413.0	0.99	1076.3	0.98	891.9	0.91	775.8	0.80

Table 2: PRE of the proposed model in comparison to the Yennum et al. [24] estimator along with privacy protection for $\pi_y = 0.20$.

P_1	P_2	T_1	T_2	Q_1	δ_1	W							
						0.2		0.3		0.4		0.5	
						π							
						0.15		0.25		0.35		0.45	
						PRE	PP	PRE	PP	PRE	PP	PRE	PP
0.25	0.65	0.3	0.3	0.2	0.25	670.0	0.21	886.3	0.08	1134.6	0.05	1421.4	0.17
0.35	0.55	0.3	0.3	0.2	0.25	578.7	0.25	746.4	0.13	959.5	0.01	1237.4	0.10
0.45	0.45	0.3	0.3	0.2	0.25	523.6	0.28	650.6	0.18	812.8	0.08	1026.9	0.02
0.55	0.35	0.3	0.3	0.2	0.25	483.6	0.32	578.0	0.23	693.8	0.14	840.2	0.06
0.65	0.25	0.3	0.3	0.2	0.25	441.7	0.35	501.2	0.28	567.1	0.21	641.0	0.14
0.75	0.15	0.3	0.3	0.2	0.25	354.9	0.39	358.1	0.33	357.6	0.28	353.5	0.22
0.25	0.65	0.4	0.4	0.2	0.25	785.9	0.33	1000.0	0.19	1207.5	0.04	1421.4	0.10
0.35	0.55	0.4	0.4	0.2	0.25	683.0	0.37	874.5	0.23	1096.5	0.10	1362.1	0.02
0.45	0.45	0.4	0.4	0.2	0.25	613.2	0.40	765.2	0.28	950.2	0.17	1182.4	0.05
0.55	0.35	0.4	0.4	0.2	0.25	555.9	0.44	666.5	0.33	798.4	0.23	959.8	0.13
0.65	0.25	0.4	0.4	0.2	0.25	487.6	0.47	547.3	0.38	612.2	0.30	683.4	0.21
0.75	0.15	0.4	0.4	0.2	0.25	353.5	0.51	343.9	0.44	335.4	0.37	327.4	0.30
0.25	0.65	0.5	0.5	0.2	0.25	1824.3	0.45	1624.9	0.29	1538.9	0.13	1546.3	0.02
0.35	0.55	0.5	0.5	0.2	0.25	1757.6	0.48	1644.3	0.33	1619.3	0.19	1675.0	0.05
0.45	0.45	0.5	0.5	0.2	0.25	1609.9	0.51	1515.9	0.38	1507.3	0.25	1571.3	0.13
0.55	0.35	0.5	0.5	0.2	0.25	1398.6	0.55	1287.1	0.43	1256.8	0.32	1284.8	0.21
0.65	0.25	0.5	0.5	0.2	0.25	1075.5	0.58	938.0	0.48	873.5	0.39	849.9	0.29
0.75	0.15	0.5	0.5	0.2	0.25	579.9	0.62	460.0	0.54	395.2	0.45	355.6	0.37
0.25	0.65	0.3	0.3	0.3	0.35	2346.8	0.30	2641.7	0.16	2957.7	0.02	3332.0	0.11
0.35	0.55	0.3	0.3	0.3	0.35	2112.6	0.34	2371.3	0.22	2715.4	0.10	3197.9	0.02
0.45	0.45	0.3	0.3	0.3	0.35	1946.4	0.38	2125.8	0.28	2386.3	0.17	2776.2	0.07
0.55	0.35	0.3	0.3	0.3	0.35	1805.1	0.43	1901.2	0.34	2053.6	0.25	2287.1	0.17
0.65	0.25	0.3	0.3	0.3	0.35	1621.9	0.47	1615.0	0.40	1635.7	0.34	1684.1	0.27
0.75	0.15	0.3	0.3	0.3	0.35	1210.7	0.52	1061.0	0.47	942.4	0.42	841.5	0.38
0.25	0.65	0.4	0.4	0.3	0.35	2356.7	0.44	2581.3	0.28	2780.1	0.13	3014.6	0.02
0.35	0.55	0.4	0.4	0.3	0.35	2131.3	0.48	2404.9	0.34	2737.0	0.20	3178.6	0.07
0.45	0.45	0.4	0.4	0.3	0.35	1939.9	0.52	2157.3	0.40	2458.5	0.28	2898.0	0.16
0.55	0.35	0.4	0.4	0.3	0.35	1752.1	0.57	1875.2	0.46	2063.9	0.36	2349.2	0.26
0.65	0.25	0.4	0.4	0.3	0.35	1491.9	0.61	1485.3	0.53	1514.6	0.44	1579.0	0.36
0.75	0.15	0.4	0.4	0.3	0.35	986.9	0.65	846.4	0.59	749.0	0.53	675.5	0.47
0.25	0.65	0.5	0.5	0.3	0.35	3423.9	0.57	3015.2	0.40	2826.3	0.23	2814.6	0.07
0.35	0.55	0.5	0.5	0.3	0.35	3307.6	0.61	3090.7	0.46	3059.2	0.31	3200.3	0.16
0.45	0.45	0.5	0.5	0.3	0.35	3027.0	0.65	2857.6	0.52	2882.2	0.38	3087.6	0.25
0.55	0.35	0.5	0.5	0.3	0.35	2619.3	0.70	2409.1	0.58	2380.8	0.46	2498.7	0.35
0.65	0.25	0.5	0.5	0.3	0.35	1997.7	0.74	1722.6	0.64	1599.6	0.55	1565.3	0.45
0.75	0.15	0.5	0.5	0.3	0.35	1061.9	0.78	818.8	0.71	686.4	0.63	603.8	0.55

Table 3: PRE of the proposed model in comparison to the Singh & Grewal [20] estimator along with privacy protection for $\pi_y = 0.10$.

P_1	P_2	T_1	T_2	Q_1	δ_1	W							
						0.2		0.3		0.4		0.5	
						π							
						0.15		0.25		0.35		0.45	
						PRE	PP	PRE	PP	PRE	PP	PRE	PP
0.85	0.15	0.20	0.20	0.65	0.25	154.85	0.26	174.77	0.22	222.86	0.17	371.2	0.12
0.75	0.25	0.20	0.20	0.65	0.25	174.32	0.25	209.11	0.20	285.18	0.16	513.5	0.12
0.65	0.35	0.20	0.20	0.65	0.25	182.85	0.23	225.66	0.19	318.19	0.15	596.5	0.11
0.55	0.45	0.20	0.20	0.65	0.25	190.67	0.22	241.27	0.17	350.08	0.13	678.3	0.10
0.45	0.55	0.20	0.20	0.65	0.25	200.82	0.20	261.61	0.15	391.66	0.11	784.6	0.08
0.35	0.65	0.20	0.20	0.65	0.25	216.35	0.18	292.34	0.13	453.52	0.09	939.2	0.06
0.85	0.15	0.30	0.30	0.65	0.25	110.10	0.38	131.18	0.30	175.77	0.23	306.6	0.16
0.75	0.25	0.30	0.30	0.65	0.25	134.66	0.37	178.34	0.29	267.45	0.22	527.2	0.15
0.65	0.35	0.30	0.30	0.65	0.25	146.09	0.35	202.90	0.27	321.26	0.20	674.8	0.15
0.55	0.45	0.30	0.30	0.65	0.25	155.86	0.34	224.18	0.26	368.45	0.19	805.7	0.13
0.45	0.55	0.30	0.30	0.65	0.25	167.86	0.32	249.90	0.24	423.97	0.17	953.4	0.12
0.35	0.65	0.30	0.30	0.65	0.25	185.62	0.31	286.53	0.22	498.26	0.15	1132.6	0.10
0.85	0.15	0.40	0.40	0.65	0.25	121.71	0.52	139.30	0.40	181.30	0.29	309.8	0.20
0.75	0.25	0.40	0.40	0.65	0.25	166.82	0.51	219.81	0.38	328.17	0.28	643.3	0.19
0.65	0.35	0.40	0.40	0.65	0.25	188.90	0.50	266.02	0.37	426.92	0.27	907.0	0.19
0.55	0.45	0.40	0.40	0.65	0.25	205.65	0.49	301.87	0.35	505.03	0.25	1119.4	0.17
0.45	0.55	0.40	0.40	0.65	0.25	223.92	0.47	339.37	0.34	581.53	0.24	1308.3	0.16
0.35	0.65	0.40	0.40	0.65	0.25	248.67	0.46	385.46	0.32	662.05	0.21	1464.6	0.14
0.85	0.15	0.20	0.20	0.75	0.45	127.27	0.03	182.59	0.00	329.08	0.05	1242.2	0.11
0.75	0.25	0.20	0.20	0.75	0.45	149.36	0.03	227.11	0.00	434.92	0.03	1760.3	0.08
0.65	0.35	0.20	0.20	0.75	0.45	157.67	0.03	245.85	0.01	484.66	0.02	2033.0	0.05
0.55	0.45	0.20	0.20	0.75	0.45	163.57	0.04	260.18	0.01	525.04	0.01	2265.4	0.04
0.45	0.55	0.20	0.20	0.75	0.45	170.11	0.04	276.45	0.02	571.65	0.00	2536.0	0.02
0.35	0.65	0.20	0.20	0.75	0.45	179.15	0.03	298.56	0.02	633.70	0.00	2887.1	0.01
0.85	0.15	0.30	0.30	0.75	0.45	98.68	0.06	142.33	0.02	259.15	0.04	993.3	0.10
0.75	0.25	0.30	0.30	0.75	0.45	133.61	0.07	213.60	0.02	429.47	0.02	1826.3	0.07
0.65	0.35	0.30	0.30	0.75	0.45	148.70	0.07	249.10	0.03	526.75	0.01	2374.1	0.05
0.55	0.45	0.30	0.30	0.75	0.45	158.51	0.07	273.50	0.03	596.96	0.00	2787.6	0.03
0.45	0.55	0.30	0.30	0.75	0.45	167.95	0.07	296.84	0.04	662.93	0.01	3163.7	0.02
0.35	0.65	0.30	0.30	0.75	0.45	179.54	0.07	323.22	0.04	730.08	0.02	3497.5	0.00
0.85	0.15	0.40	0.40	0.75	0.45	79.79	0.11	112.84	0.05	204.90	0.02	792.8	0.09
0.75	0.25	0.40	0.40	0.75	0.45	125.23	0.11	200.89	0.05	408.50	0.00	1766.2	0.06
0.65	0.35	0.40	0.40	0.75	0.45	148.79	0.11	254.95	0.06	554.05	0.01	2574.7	0.04
0.55	0.45	0.40	0.40	0.75	0.45	163.83	0.11	291.35	0.06	656.49	0.02	3167.0	0.02
0.45	0.55	0.40	0.40	0.75	0.45	176.67	0.11	321.02	0.07	734.36	0.03	3576.0	0.01
0.35	0.65	0.40	0.40	0.75	0.45	190.15	0.11	346.67	0.07	785.35	0.03	3749.3	0.01

Table 4: PRE of the proposed model in comparison to the Singh & Grewal [20] estimator along with privacy protection for $\pi_y = 0.20$.

P_1	P_2	T_1	T_2	Q_1	δ_1	W							
						0.2		0.3		0.4		0.5	
						π							
						0.15		0.25		0.35		0.45	
						PRE	PP	PRE	PP	PRE	PP	PRE	PP
0.25	0.65	0.3	0.3	0.2	0.25	312.54	0.01	310.97	0.01	306.23	0.02	299.35	0.03
0.35	0.55	0.3	0.3	0.2	0.25	323.94	0.00	327.15	0.01	327.71	0.02	327.13	0.02
0.45	0.45	0.3	0.3	0.2	0.25	332.75	0.00	340.19	0.01	345.62	0.02	351.01	0.02
0.55	0.35	0.3	0.3	0.2	0.25	340.72	0.00	352.31	0.01	362.63	0.01	373.98	0.01
0.65	0.25	0.3	0.3	0.2	0.25	348.88	0.00	364.82	0.01	380.24	0.01	397.73	0.00
0.75	0.15	0.3	0.3	0.2	0.25	357.95	0.00	378.59	0.01	399.31	0.01	422.90	0.01
0.25	0.65	0.4	0.4	0.2	0.25	294.44	0.01	302.69	0.03	309.41	0.04	315.41	0.05
0.35	0.55	0.4	0.4	0.2	0.25	316.16	0.01	333.51	0.03	350.28	0.04	368.31	0.04
0.45	0.45	0.4	0.4	0.2	0.25	332.38	0.01	357.21	0.03	382.56	0.04	411.18	0.04
0.55	0.35	0.4	0.4	0.2	0.25	345.87	0.01	377.06	0.03	409.57	0.03	446.83	0.03
0.65	0.25	0.4	0.4	0.2	0.25	358.30	0.01	394.89	0.03	433.01	0.03	476.37	0.02
0.75	0.15	0.4	0.4	0.2	0.25	370.62	0.01	411.40	0.03	452.91	0.03	498.72	0.01
0.25	0.65	0.5	0.5	0.2	0.25	254.63	0.03	263.92	0.05	273.87	0.07	284.91	0.07
0.35	0.55	0.5	0.5	0.2	0.25	297.63	0.03	321.52	0.05	346.78	0.07	375.67	0.07
0.45	0.45	0.5	0.5	0.2	0.25	307.96	0.03	337.90	0.05	370.10	0.06	407.43	0.06
0.55	0.35	0.5	0.5	0.2	0.25	326.62	0.04	363.88	0.05	403.59	0.06	449.38	0.06
0.65	0.25	0.5	0.5	0.2	0.25	342.65	0.04	384.84	0.05	428.50	0.06	477.59	0.05
0.75	0.15	0.5	0.5	0.2	0.25	356.98	0.04	400.96	0.05	444.05	0.05	490.43	0.04
0.25	0.65	0.3	0.3	0.3	0.35	435.76	0.08	417.07	0.06	398.95	0.04	415.36	0.02
0.35	0.55	0.3	0.3	0.3	0.35	450.90	0.08	437.50	0.05	425.25	0.03	412.95	0.01
0.45	0.45	0.3	0.3	0.3	0.35	462.41	0.07	453.74	0.05	446.96	0.02	441.35	0.01
0.55	0.35	0.3	0.3	0.3	0.35	472.66	0.07	468.66	0.04	467.41	0.02	468.58	0.01
0.65	0.25	0.3	0.3	0.3	0.35	482.97	0.06	483.91	0.03	488.48	0.01	496.66	0.00
0.75	0.15	0.3	0.3	0.3	0.35	494.30	0.05	500.53	0.03	511.15	0.01	526.36	0.00
0.25	0.65	0.4	0.4	0.3	0.35	405.62	0.10	394.99	0.07	388.50	0.04	383.76	0.02
0.35	0.55	0.4	0.4	0.3	0.35	434.29	0.09	432.83	0.06	436.65	0.03	444.50	0.01
0.45	0.45	0.4	0.4	0.3	0.35	455.38	0.09	461.61	0.05	474.34	0.03	493.45	0.01
0.55	0.35	0.4	0.4	0.3	0.35	472.66	0.08	485.46	0.05	505.78	0.02	534.24	0.00
0.65	0.25	0.4	0.4	0.3	0.35	488.28	0.08	506.69	0.04	533.06	0.02	568.35	0.00
0.75	0.15	0.4	0.4	0.3	0.35	503.48	0.07	526.16	0.04	556.27	0.01	594.66	0.00
0.25	0.65	0.5	0.5	0.3	0.35	331.63	0.11	327.01	0.07	328.16	0.04	332.95	0.01
0.35	0.55	0.5	0.5	0.3	0.35	370.15	0.10	376.65	0.06	390.16	0.03	410.04	0.00
0.45	0.45	0.5	0.5	0.3	0.35	393.07	0.10	408.15	0.06	431.50	0.02	463.77	0.00
0.55	0.35	0.5	0.5	0.3	0.35	422.52	0.09	445.88	0.05	478.00	0.02	520.62	0.00
0.65	0.25	0.5	0.5	0.3	0.35	442.27	0.08	470.49	0.04	506.80	0.01	553.19	0.01
0.75	0.15	0.5	0.5	0.3	0.35	459.61	0.08	489.32	0.04	524.98	0.01	568.56	0.01

- ii. From Tables 1-4, it may also be seen that some values are much closer to zero or equal to zero. Following Zhimin and Zaizai [25], we may interpret that our proposed model protected much privacy of the respondent and there are high chance to get true response from the respondents. Hence, the suggested estimation procedure has been found to be more efficient as well as more protective (high degree of privacy protection) than existing estimators in survey literature.

6. CONCLUSIONS AND RECOMMENDATIONS

In this paper, a new RR model to produce more precise estimate of sensitive proportion using the blank card approach and GP distribution has been proposed. Measure of privacy protection of proposed RR model and its properties have been investigated and quantified through empirical studies. Based on the results and interpretations, it can be concluded that the suggested RR model is rewarding in terms of percent relative efficiencies and dominates the randomized response models presented by Yennun et al. [24] and Singh & Grewal [20]. Following an analysis of privacy protection measures, it was observed that the suggested model is suitable for dealing with very sensitive concerns while maintaining a high level of privacy protection. As a result, by employing this approach, one may assure maximum privacy protection for respondents whilst still getting more precise estimates. As an outcome of the extremely encouraging results, the suggested model and estimation approach may then be recommended to survey practitioners whenever non-response is unavoidable owing to the stigmatized nature of attributes, notably in socioeconomic and health-related concerns.

FUNDING

“This research received no external funding”.

ACKNOWLEDGMENTS

The author would like to thank Saudi Electronic University, KSA for providing the necessary facilities for the current study.

CONFLICT OF INTERESTS

The author(s) declare that there is no conflict of interests.

REFERENCES

- [1] H. Anderson, Estimation of a proportion through randomized response. *Int. Stat. Rev./Rev. Int. Stat.* 44 (1976), 213-217.
- [2] N. Balakrishnan, S. Bersimis, M.V. Koutras, Run and frequency quota rules in process monitoring and acceptance sampling. *J. Qual. Technol.* 41 (2009), 66-81.
- [3] A. Chaudhuri, R. Mukerjee, *Randomized response: Theory and techniques*. Routledge. (2020).
- [4] W. G. Cochran, *Sampling Techniques: 3d Ed.* New York: Wiley. (1977).
- [5] J. A. Fox, P. E. Tracy, *Randomized response. A method for sensitive surveys*. Beverly Hills: Sage, *Quantitative Applications in the Social Sciences*. (1986).
- [6] B. G. Greenberg, A. L. A. Abul-Ela, W. R. Simmons, D. G. Horvitz, The unrelated question randomized response model: Theoretical framework. *J. Amer. Stat. Assoc.* 64(326) (1969), 520-539.
- [7] Z. Hussain, J. Shabbir, Z. Pervez, S. F. Shah, M. Khan, Generalized geometric distribution of order k: A flexible choice to randomize the response. *Commun. Stat.-Simul. Comput.* 46(6) (2017), 4708-4721.
- [8] Z. Hussain, B. Al-Zahrani, J. Shabbir, M. Khan, On using negative binomial distribution as a randomization device in sensitive surveys. *Commun. Stat.-Simul. Comput.* 45(10) (2016), 3584-3596.
- [9] J. M. Kim, M. E. Elam, A stratified unrelated question randomized response model. *Stat. Papers*, 48(2) (2007), 215-233.
- [10] J. M. Kim, S. S. Chae, Three-Stage Stratified Randomized Response Model. *Korean J. Appl. Stat.* 23(3) (2010), 533-543.
- [11] A. Y. Kuk, Asking sensitive questions indirectly. *Biometrika*, 77(2) (1990), 436-438.
- [12] M. V. Koutras, S. Bersimis, P. E. Maravelakis, Statistical process control using Shewhart control charts with supplementary runs rules. *Methodol. Comput. Appl. Probab.* 9(2) (2007), 207-224.
- [13] J. Lanke, On the degree of protection in randomized interviews. *Int. Stat. Rev./Rev. Int. Stat.* 44 (1976), 197-203.

RANDOMIZED RESPONSE MODEL

- [14] F. W. Leysieffer, S. L. Warner, Respondent jeopardy and optimal designs in randomized response models. *J. Amer. Stat. Assoc.* 71(355) (1976), 649-656.
- [15] J. J. A. Moors, Optimization of the unrelated question randomized response model. *J. Amer. Stat. Assoc.* 66(335) (1971), 627-629.
- [16] Z. Pervez, S. F. Shah, Z. Hussain, A. N. Philippou, An application of negative binomial distribution of order k in sensitive surveys. *Commun. Stat.-Simul. Comput.* 46(8) (2017), 6654-6660.
- [17] C. E. Shannon, Communication theory of secrecy systems. *Bell Syst. Techn. J.* 28(4) (1949), 656-715.
- [18] S. Singh, S. Horn, R. Singh, N. S. Mangat, On the use of modified randomization device for estimating the prevalence of a sensitive attribute. *Stat. Transition*, 6(4) (2003), 515-522.
- [19] H. P. Singh, T. A. Tarray, A stratified unknown repeated trial in randomized response sampling. *Commun. Stat. Appl. Meth.* 19(6) (2012), 751-759.
- [20] S. Singh, I. S. Grewal, Geometric distribution as a randomization device: implemented to the Kuk's model. *Int. J. Contemp. Math. Sci.* 8(5) (2013), 243
- [21] C. Singh, M. Kamal, G. N. Singh, J. M. Kim, Study to Alter the Nuisance Effect of Non-Response Using Scrambled Mechanism. *Risk Manage. Healthcare Policy*, 14 (2021), 1595-1613.
- [22] D. S. Tracy, S. S. Osahan, An improved randomized response technique. *Pakistan J. Stat.-All Ser.* 15(1) (1999), 1-6.
- [23] S. L. Warner, Randomized response: A survey technique for eliminating evasive answer bias. *J. Amer. Stat. Assoc.* 60(309) (1965), 63-69.
- [24] N. Yennun, S. A. Sedory, S. Singh, Improved strategy to collect sensitive data by using geometric distribution as a randomization device. *Commun. Stat.-Theory Meth.* 48(23) (2018), 5777-5795.
- [25] H. Zhimin, Y. Zaizai, Measure of privacy in randomized response model. *Qual. Quant.* 46(4) (2012), 1167-1180.