

یک مدل پاسخ تصادفیده مرکب جدید

سید محمدرضا علوی، محمد جوهرزاده و رحیم چینی پرداز

گروه آمار، دانشگاه شهید چمران اهواز

چکیده: معمولاً در بررسی‌های نمونه‌گیری هنگامی که سؤال حساس مستقیمی پرسیده شود، پاسخگو پاسخ واقعی را ارائه نمی‌کند. روش‌های پاسخ تصادفیده برای حفاظت از محرمانگی پاسخ‌ها مطرح شده‌اند. تمرکز این مقاله بر روش پاسخ تصادفیده در متغیرهای کیفی بر پایه روش سیمونس است. برای حفظ محرمانگی بیشتر با ترکیب دو روش جداگانه سیمونس، یک روش پاسخ تصادفیده مرکب جدید معرفی می‌شود و با استفاده از شبیه‌سازی در نرم افزار R کارایی روش تصادفیده پیشنهادی با روش سیمونس و روش تصادفیده علوی و تاج‌الدینی (۱۳۹۴) مقایسه می‌شود. با استفاده از روش تصادفیده پیشنهادی، نسبت تقلب دانشجویان در امتحانات دانشگاه شهید چمران اهواز برآورد شده‌است.

واژه‌های کلیدی: پاسخ تصادفیده، روش سیمونس، سؤال نامرتبط، پاسخ تصادفیده مرکب، نسبت حساس.

۱ مقدمه

در بسیاری از بررسی‌های نمونه‌گیری، آماردانان سعی دارند تا اطلاعاتی درباره ویژگی‌هایی از جوامع انسانی به‌دست آورند. اگر ویژگی‌های حساسی مستقیماً از افراد پرسیده شوند ممکن است با عدم پاس یا پاسخ‌های غیر واقعی مواجه شوند. زیرا این‌گونه سؤالات امکان دارد برای برخی افراد خجالت‌آور یا عذاب‌آور باشند.

آدرس الکترونیک مسئول مقاله: سید محمدرضا علوی، alavi_m@scu.ac.ir
کد موضوع‌بندی ریاضی (۲۰۱۰): 62D05

بعضی از موضوعات حساس مورد پرسش عبور از چراغ قرمز، قماربازی، فر مالیاتی، مصرف مشروبات الکلی، استعمال مواد مخدر، تقلب در امتحانات و غیره است. به منظور حفظ محرمانگی و کاهش آریبی در پاسخ‌ها، روش‌های پاسخ تصادفیده^۱ به‌عنوان جایگزین روش سؤالات مستقیم^۲ مطرح شده‌اند. ایده اصلی این روش‌ها نخستین بار توسط وارنر (۱۹۶۵) به‌منظور ترغیب به‌همکاری و بیان پاسخ‌های صادقانه مطرح شد. در این روش پاسخ تصادفیده، هر فرد با انجام یک آزمایش تصادفی مانند پرتاب سکه یا ریختن تاس ضمن محرمانه ماندن پاسخ، یکی از دو سؤال حساس یا مکمل حساس را برای پاسخگویی انتخاب می‌کند. در این پژوهش‌ها معمولاً هدف، برآورد نسبت حساس با استفاده از پاسخ‌های تصادفیده است. گرینبرگ و همکاران (۱۹۶۹) برای افزایش حفظ محرمانگی استفاده از سؤال نامرتبط غیرحساس را در پرسش دوم مطرح کردند که روش سیمونس نامیده می‌شود. مدل‌های تصادفیده زیادی توسط افرادی نظیر مورس (۱۹۷۱)، راجاوارائو (۱۹۷۸)، منگات و سینگ (۱۹۹۰)، کاک (۱۹۹۰)، منگات (۱۹۹۴)، سینگ (۲۰۰۲)، هانگ (۲۰۰۴)، چانگ و همکاران (۲۰۰۴)، کیم و وارده (۲۰۰۵)، جست وانگ و سینگ (۲۰۰۶)، حسین و شبیر (۲۰۰۷)، مهتا و همکاران (۲۰۱۲)، یزاری و علوی (۱۳۹۳) و علوی و تاج‌الدینی (۱۳۹۴ و ۲۰۱۶) معرفی شده‌اند. در این مقاله براساس روش سیمونس یک روش جدید پاسخ تصادفیده مرکب برای افزایش محرمانگی پاسخ‌ها، معرفی شده است. مقاله از پنج بخش تشکیل شده است. در بخش ۲ روش پاسخ تصادفیده سیمونس و گونه مکرر آن یعنی روش علوی و تاج‌الدینی (۱۳۹۴) بیان شده است. در بخش ۳ با ترکیب دو روش جداگانه سیمونس، یک روش پاسخ تصادفیده مرکب جدید معرفی و برآورد نسبت حساس توسط این روش پیشنهادی ارائه و در بخش ۴ کارایی^۳ این روش از طریق شبیه‌سازی با استفاده از بسته نرم‌افزاری R با روش سیمونس و روش علوی و تاج‌الدینی (۱۳۹۴) مقایسه شده است. سرانجام در بخش ۵ با استفاده از روش پیشنهادی نسبت تقلب دانشجویان در امتحانات دانشگاه شهید چمران اهواز آورده شده است.

۲ روش پاسخ تصادفیده سیمونس و گونه مکرر آن

در این بخش ابتدا روش پاسخ تصادفیده سیمونس شرح داده می‌شود و سپس گونه مکرر آن یعنی روش علوی و تاج‌الدینی (۱۳۹۴) بیان می‌شود.

¹Randomized response techniques

²Direct questions

³Efficiency

۱.۲ روش پاسخ تصادفیده سیمونس

در روش سیمونس از فرد i ام نمونه تقاضا می‌شود یک آزمایش برنولی با احتمال موفقیت p انجام داده در صورت پیروزی به سؤال حساس و در صورت شکست به یک سؤال نامرتبب غیر حساس پاسخ دهد. اگر متغیرهای مستقل برنولی Y_i ، X_i و T_i به‌ازای $i = 1, \dots, n$ به‌ترتیب معرف پاسخ سؤال حساس، پاسخ سؤال نامرتبب و نتیجه آزمایش برنولی با احتمال‌های به‌ترتیب θ ، π_B و p باشند، پاسخ تصادفیده به‌صورت

$$Z_i = Y_i T_i + X_i (1 - T_i), \quad i = 1, \dots, n \quad (1)$$

است. واضح است که Z_i یک متغیر تصادفی برنولی با احتمال پیروزی به‌صورت

$$\phi = E_P E_R(Z_i) = E_P E_R[Y_i T_i + X_i (1 - T_i)] = p\theta + (1 - p)\pi_B, \quad i = 1, \dots, n$$

است، که در آن E_P بیانگر امید ریاضی روی تمام نمونه‌های ممکن و E_R امید ریاضی تحت عمل تصادفی کردن است. بر پایه یک نمونه تصادفی از پاسخ‌های تصادفیده، برآورد ناریب θ و واریانس آن به‌ترتیب به‌صورت

$$\hat{\theta} = \frac{\bar{Z} - (1 - p)\pi_B}{p} \quad (2)$$

$$Var(\hat{\theta}) = \frac{\theta(1 - \theta)}{n} + \frac{\theta(1 - p)(1 - \pi_B)}{np} + \frac{\pi_B(1 - p)[1 - \pi_B(1 - p)(1 - p)]}{np^2}$$

هستند که در آن $\bar{Z} = \frac{\sum_{i=1}^n Z_i}{n}$ میانگین پاسخ بله در نمونه تصادفی به حجم n یک برآورد ناریب برای ϕ است. در ادامه برآوردی برای $Var(\hat{\theta})$ به‌صورت

$$\hat{V}ar(\hat{\theta}) = \frac{Var(\bar{Z})}{p^2} = \frac{\bar{Z}(1 - \bar{Z})}{np^2}$$

معرفی می‌شود.

۲.۲ روش پاسخ تصادفیده مکرر سیمونس

در صورتی‌که سؤال غیرمرتبب در روش سیمونس یک آزمایش تصادفی باشد که با تکرار آن نتایج بتوانند تغییر کنند، علوی و تاج‌الدینی (۱۳۹۴) با تکرار روش سیمونس یک روش جدید برای حفظ محرمانگی بیشتر معرفی کردند. لازم به ذکر است اگر سؤال غیرمرتبب غیرتصادفی باشد ممکن است پاسخگو احساس کند تکرار پاسخ

تصادفیده سیمونس محرمانگی را حفظ نمی‌کند. در این روش از فرد i ام نمونه تقاضا می‌شود روش پاسخ تصادفیده سیمونس را f_i بار تکرار کند که f_i دارای توزیع اریب اندازه پواسن با میانگین دلخواه مثلاً ۳ است. اگر m_i نسبت تعداد بله پاسخ‌های تصادفیده فرد i ام نمونه باشد، در آن صورت یک برآورد نااریب برای نسبت حساس θ با استفاده از پاسخ‌های تصادفیده فرد i ام به صورت

$$\hat{\theta}_i = \frac{m_i - (1-p)\pi_B}{p}, \quad i = 1, \dots, n$$

است. پس یک برآورد نااریب براساس تمام پاسخ‌های تصادفیده نمونه از رابطه

$$\hat{\theta} = \frac{\sum_{i=1}^n \hat{\theta}_i}{n} \quad (3)$$

به دست می‌آید (علوی و تاج‌الدینی، ۱۳۹۴).

۳ روش پاسخ تصادفیده مرکب جدید

در روش پاسخ تصادفیده از هر فرد در نمونه تصادفی تقاضا می‌شود اگر دارای صفت غیر حساس B است، روش پاسخ تصادفیده سیمونس اول و در غیر این صورت روش پاسخ تصادفیده سیمونس دوم را انجام دهد. روش پاسخ تصادفیده سیمونس اول: اگر فرد دارای صفت غیرحساس B است یک آزمایش برنولی با احتمال موفقیت p_1 انجام داده در صورت پیروزی فقط به سؤال "آیا شما دارای صفت حساس A هستید؟" و در صورت شکست فقط به سؤال "آیا شما دارای صفت غیرحساس B هستید؟" پاسخ داده و نتیجه را در پاسخ‌نامه ثبت کند.

روش پاسخ تصادفیده سیمونس دوم: اگر فرد دارای صفت غیرحساس B^c (مکمل B) است یک آزمایش برنولی با احتمال موفقیت p_2 انجام داده در صورت پیروزی فقط به سؤال "آیا شما دارای صفت حساس A هستید؟" و در صورت شکست فقط به سؤال "آیا شما دارای صفت غیرحساس B^c هستید؟" پاسخ داده و نتیجه را در پاسخ‌نامه ثبت کند.

اگر متغیرهای مستقل برنولی X, Y, T_1 و T_2 به ترتیب پاسخ سؤال حساس A ، پاسخ سؤال غیر حساس B ، نتیجه آزمایش تصادفی روش اول و روش دوم با احتمال‌های به ترتیب θ, π_B, p_1 و p_2 باشند، آنگاه پاسخ تصادفیده Z به صورت

$$Z = X[T_1 Y + (1 - T_1) X] + (1 - X)[T_2 Y + (1 - T_2)(1 - X)]$$

است. با توجه به اینکه توان دوم یک متغیر برنولی معادل خودش است، Z به صورت

$$Z = [T_1 X + T_2(1 - X)]Y + (1 - T_1)X + (1 - T_2)(1 - X) \quad (۴)$$

ساده می‌شود. امید ریاضی پاسخ تصادفیده برابر

$$\begin{aligned} \phi &= E_P E_R [T_1 X + T_2(1 - X)]Y + (1 - T_1)X + (1 - T_2)(1 - X) \\ &= [p_1 \pi_B + p_2(1 - \pi_B)\theta] + \pi_B(1 - p_1) + (1 - \pi_B)(1 - p_2) \end{aligned}$$

است. در نتیجه اگر Z_1, \dots, Z_n یک نمونه تصادفی از پاسخ‌های تصادفیده به روش پیشنهادی باشند، یک برآورد ناریب برای نسبت حساس θ به صورت

$$\hat{\theta} = \frac{\bar{Z} - [\pi_B(1 - p_1) + (1 - \pi_B)(1 - p_2)]}{\pi_B p_1 + (1 - \pi_B)p_2} \quad (۵)$$

به دست می‌آید، که در آن $\bar{Z} = \frac{\sum_{i=1}^n Z_i}{n}$ برآورد ناریب ϕ است. واریانس این برآوردگر برابر

$$Var(\hat{\theta}) = \frac{\phi(1 - \phi)}{nk^2} = \frac{(k\theta + d)k(1 - \theta)}{nk^2} = \frac{\theta(1 - \theta)}{n} + \frac{d(1 - \theta)}{nk}$$

است که در آن $k = \pi_B p_1 + (1 - \pi_B)p_2$ و $d = \pi_B(1 - p_1) + (1 - \pi_B)(1 - p_2)$ برآورد

این واریانس به صورت $\hat{V}ar(\hat{\theta}) = \frac{\bar{Z}(1 - \bar{Z})}{nk^2}$ است.

حالت‌های خاص مدل پیشنهادی:

(الف) عبارت (۵) به ازای $p_1 = 1 - p_2$ ، به صورت $\hat{\theta} = \frac{\bar{Z} - [\frac{\pi_B(1 - 2p_1) + p_1}{\pi_B(2p_1 - 1) + (1 - p_1)}]}$ است.

(ب) اگر $\pi_B = \frac{1}{2}$ ، آنگاه $\hat{\theta} = \frac{\bar{Z} - [\frac{1 - p_1 + p_2}{p_1 + p_2}]}$ ، که به ازای $p = \frac{p_1 + p_2}{2}$ داریم: $\hat{\theta} = \frac{\bar{Z} - [1 - p]}{p}$

(پ) اگر $\pi_B = \frac{1}{2}$ و $p_1 = 1 - p_2$ ، آنگاه

$$\hat{\theta} = 2\bar{Z} - 1, \quad Var(\hat{\theta}) = \frac{4\phi(1 - \phi)}{n} \quad (۶)$$

و یک برآورد برای این واریانس عبارتست از

$$\hat{V}ar(\hat{\theta}) = \frac{4\bar{Z}(1 - \bar{Z})}{n} \quad (۷)$$

۴ مطالعه شبیه‌سازی ب

با توجه به نااریب بودن برآورد نسبت حساس در هر سه روش، برای ارزیابی روش پیشنهادی و مقایسه کارایی آن با دو روش سیمونس و روش علوی و تاج الدینی (۱۳۹۴) با استفاده از بسته نرم افزاری R سه شبیه‌سازی در نظر گرفته شده‌است.

برای شبیه‌سازی روش سیمونس مراحل زیر انجام شده است:

مرحله ۱: سه نمونه تصادفی به حجم n از متغیرهای تصادفی برنولی X, Y و T به ترتیب با احتمال‌های پیروزی معلوم θ, π_B و p تولید و سپس با استفاده از رابطه (۱) داده‌های تصادفیده Z به دست آورده و براساس رابطه (۲) برآوردهای $\hat{\theta}$ برای نمونه به حجم n محاسبه شده‌اند.

مرحله ۲: مرحله ۱، k بار تکرار، میانگین و واریانس این k بار به ترتیب به عنوان امید ریاضی و واریانس شبیه‌سازی شده $\hat{\theta}$ در نظر گرفته شدند.

برای شبیه‌سازی روش پیشنهادی مراحل زیر انجام شده است:

مرحله ۱: چهار نمونه تصادفی به حجم n از متغیرهای برنولی X, Y, T_1 و T_2 به ترتیب با احتمال‌های پیروزی معلوم θ, π_B, p_1 و p_2 تولید شدند. لازم به ذکر است که T_1 و T_2 بیانگر نتیجه آزمایش‌های برنولی مراحل اول و دوم روش پیشنهادی هستند.

حال با استفاده از رابطه (۴) داده‌های تصادفیده Z را به دست آورده و بر اساس رابطه (۵)، برآوردهای θ در نمونه به حجم n محاسبه شدند.

مرحله ۲: مرحله ۱، k بار تکرار، سپس میانگین و واریانس این k بار به ترتیب به عنوان امید ریاضی و واریانس شبیه‌سازی شده $\hat{\theta}$ در نظر گرفته شدند.

برای شبیه‌سازی روش علوی و تاج الدینی (۱۳۹۴) مراحل زیر انجام شده است:

مرحله ۱: دو نمونه تصادفی به حجم n از متغیرهای تصادفی برنولی X و Y به ترتیب با احتمال‌های پیروزی معلوم θ و π_B تولید و سپس برای واحد i ام، روش پاسخ تصادفیده سیمونس f_i بار تکرار شد که f_i دارای توزیع پواسن اریب اندازه با میانگین ۳ است. اگر m_i نسبت تعداد بله پاسخ‌های تصادفیده واحد i ام باشد، در آن صورت یک برآورد نااریب برای نسبت حساس θ با استفاده از پاسخ‌های تصادفیده واحد i ام به صورت

$$\hat{\theta}_i = \frac{m_i - (1-p)\pi_B}{p}, \quad i = 1, \dots, n$$

است، سپس یک برآورد ناریب براساس تمام پاسخ‌های تصادفیده نمونه از رابطه (۳) محاسبه شده است (علوی و تاج‌الدینی، ۱۳۹۴).

مرحله ۲: مرحله ۱، k بار تکرار و سپس میانگین و واریانس این k بار به ترتیب به عنوان امید ریاضی و واریانس شبیه‌سازی شده θ در نظر گرفته شدند.

جدول ۱: واریانس روش‌های سیمونس، پیشنهادی و علوی و تاج‌الدینی به ازای $\pi_B = 0/3$

$Var(AT)$	$Var(P)$	$Var(S)$	p_2	p_1	p	θ	n
0/029	0/083	0/096	0/25	0/5	0/33	0/3	20
0/033	0/064	0/107	0/25	0/5	0/33	0/5	
0/035	0/041	0/112	0/25	0/5	0/33	0/7	
0/014	0/032	0/038	0/25	0/5	0/33	0/3	50
0/016	0/025	0/042	0/25	0/5	0/33	0/5	
0/017	0/016	0/045	0/25	0/5	0/33	0/7	
0/007	0/016	0/019	0/25	0/5	0/33	0/3	100
0/5	0/33	0/5	0/25	0/021	0/013	0/008	
0/008	0/008	0/022	0/25	0/5	0/33	0/7	
0/003	0/008	0/010	0/25	0/5	0/33	0/3	200
0/004	0/006	0/011	0/25	0/5	0/33	0/5	
0/004	0/004	0/011	0/25	0/5	0/33	0/7	

جداول ۱ تا ۳ واریانس روش سیمونس (S)، پیشنهادی (P) و علوی و تاج‌الدینی (AT) را تا سه رقم اعشار برای اندازه‌های نمونه ۲۰، ۵۰، ۱۰۰ و ۲۰۰ و مقادیر مختلف θ و π_B به ازای $k = 10000$ بار شبیه‌سازی، نشان می‌دهد. قابل ذکر است که θ احتمال نسبت حساس، π_B احتمال نسبت غیر حساس، p احتمال پیروزی در آزمایش برنولی به روش سیمونس و p_1 و p_2 احتمالات پیروزی در آزمایش‌های برنولی مرحله اول و دوم روش پیشنهادی هستند و مقادیر p ، p_1 و p_2 جهت حفظ محرمانگی نباید خیلی کم و یا زیاد باشند،

جدول ۲: واریانس روش‌های سیمونس، پیشنهادی و علوی و تاج الدینی به ازای $\pi_B = 0/5$

$Var(AT)$	$Var(P)$	$Var(S)$	p_2	p_1	p	θ	n
0/039	0/069	0/114	0/25	0/5	0/33	0/3	20
0/041	0/055	0/117	0/25	0/5	0/33	0/5	
0/040	0/036	0/114	0/25	0/5	0/33	0/7	
0/019	0/027	0/045	0/25	0/5	0/33	0/3	50
0/020	0/022	0/046	0/25	0/5	0/33	0/5	
0/020	0/014	0/045	0/25	0/5	0/33	0/7	
0/009	0/014	0/022	0/25	0/5	0/33	0/3	100
0/010	0/011	0/023	0/25	0/5	0/33	0/5	
0/010	0/007	0/022	0/25	0/5	0/33	0/7	
0/005	0/007	0/011	0/25	0/5	0/33	0/3	200
0/005	0/005	0/012	0/25	0/5	0/33	0/5	
0/005	0/004	0/011	0/25	0/5	0/33	0/7	

بنابراین در این شبیه‌سازی برای اختصار مقادیر $33/0$ ، $5/0$ و $25/0$ به ترتیب برای p ، p_1 و p_2 در نظر گرفته شد. همان‌طور که از جداول ۱ تا ۳ ملاحظه می‌شود کارایی روش پیشنهادی بیشتر از سیمونس است. اما کارایی آن از روش علوی و تاج الدینی (۱۳۹۴) کمتر است. با افزایش حجم نمونه کارایی هر سه روش افزایش می‌یابد. به نظر می‌رسد با افزایش حجم نمونه کارایی روش پیشنهادی و روش علوی و تاج الدینی (۱۳۹۴) یکسان شود. با افزایش θ کارایی روش پیشنهادی افزایش می‌یابد.

۵ برآورد نسبت تقلب با روش پیشنهادی

تقلب از جمله روش‌های نامطلوبی است که باعث ایجاد رقابت ناسالم و گاه برتری کاذب متقلب بر دیگران می‌شود. از مصادیق تقلب در امتحانات می‌توان به استفاده از هر گونه یادداشت، نوشته، کتاب و جزوه غیر مجاز،

جدول ۳: واریانس روش‌های سیمونس، پیشنهادی و علوی و تاج‌الدینی به ازای $\pi_B = 0/7$

$Var(AT)$	$Var(P)$	$Var(S)$	p_2	p_1	p	θ	n
0/031	0/057	0/110	0/25	0/5	0/33	0/3	20
0/033	0/046	0/109	0/25	0/5	0/33	0/5	
0/034	0/031	0/098	0/25	0/5	0/33	0/7	
0/015	0/023	0/046	0/25	0/5	0/33	0/3	50
0/016	0/019	0/042	0/25	0/5	0/33	0/5	
0/016	0/012	0/038	0/25	0/5	0/33	0/7	
0/008	0/012	0/022	0/25	0/5	0/33	0/3	100
0/008	0/009	0/021	0/25	0/5	0/33	0/5	
0/003	0/006	0/011	0/25	0/5	0/33	0/7	
0/003	0/006	0/011	0/25	0/5	0/33	0/3	200
0/004	0/005	0/011	0/25	0/5	0/33	0/5	
0/004	0/003	0/010	0/25	0/5	0/33	0/7	

رد و بدل کردن هر گونه اطلاعات کتبی یا شفاهی با سایر دانشجویان بدون هماهنگی با مراقبین و استفاده از تلفن همراه به هر دلیل اشاره کرد. یکی از آسیب‌های هر نظام آموزشی تقلب است که از جمله عوامل پدیدار شدن افت تحصیلی در محیط‌های آموزشی است. نسبت تقلب دانشجویان در امتحانات یکی از معیارهای محاسبه میزان سلامت اهداف آموزشی هر دانشگاه محسوب می‌شود. با توجه به این‌که اگر مستقیماً از دانشجویان درباره ارتکاب تقلب آن‌ها در امتحانات سؤال شود ممکن است پاسخ واقعی را ارائه نکنند، بنابراین در این بخش با استفاده از پاسخ‌های تصادفیه به روش پیشنهادی برای به‌دست آوردن نسبت تقلب دانشجویان در دانشگاه شهید چمران اهواز استفاده شده است.

جمع آوری داده‌ها در فروردین و اردیبهشت ماه سال ۱۳۹۵ با مراجعه تصادفی به دانشجویان مختلف صورت گرفته و مراحل انجام آزمایش برای آن‌ها شرح داده شده است. سؤال نامرتبط در پرسشنامه می‌بایست

از دو ویژگی برخوردار باشد: یکی معلوم بودن احتمال پاسخ بله آن سؤال برای طراح نمونه‌گیری و دیگری مخفی ماندن پاسخ آن از مصاحبه‌کننده. از این رو سؤال نامرتب، زوج یا فرد بودن یکان شماره شناسنامه دانشجوی در نظر گرفته شد. به همراه پرسشنامه یک عدد سکه به دانشجوی تحویل و از او درخواست شد روش تصادفیه مطرح شده در پرسشنامه را طبق دستورالعمل زیر در محل مورد نظر درج نماید:

در صورتی که یکان شماره شناسنامه شما زوج است یک سکه انداخته اگر شیرآمد، فقط به سؤال الف و اگر خط آمد فقط به سؤال ب زیر پاسخ داده و نتیجه را در پاسخنامه گزارش نمایید.

الف) آیا شما در امتحانات رسمی دانشگاه تقلب کرده‌اید؟

ب) آیا رقم یکان شماره شناسنامه شما زوج است؟

اما در صورتی که یکان شماره شناسنامه شما فرد است یک سکه انداخته اگر شیرآمد فقط به سؤال الف و اگر خط آمد فقط به سؤال ب زیر پاسخ داده و نتیجه را در پاسخنامه گزارش نمایید.

الف) آیا شما در امتحانات رسمی دانشگاه تقلب کرده‌اید؟

ب) آیا رقم یکان شماره شناسنامه شما فرد است؟

پاسخنامه: بله خیر

در این پرسشنامه چون از انداختن سکه برای انتخاب سؤالها استفاده شده احتمال موفقیت آزمایش‌های برنولی در مرحله اول و دوم یعنی p_1 و p_2 برابر $\frac{1}{2}$ و چون از زوج یا فرد بودن یکان شناسنامه برای سؤال غیرحساس استفاده شده، π_B برابر با $\frac{1}{2}$ در نظر گرفته شده است. بنابراین از روابط (۶) و (۷) برای برآورد نسبت تقلب و برآورد واریانس آن استفاده شده است. در جدول ۴ خلاصه‌ای از نتایج نمونه‌گیری آمده است.

جدول ۴: نتایج نمونه‌گیری برای برآورد نسبت تقلب دانشجویی در دانشگاه

اندازه نمونه	تعداد بله	نسبت بله	برآورد نسبت تقلب	خطای معیار	بازه اطمینان ۹۵%
۱۷۰	۱۲۰	۰/۷۰۶	۰/۴۱۲	۰/۰۷۵	(۰/۲۶۵ و ۰/۵۵۹)

۶ بحث و نتیجه‌گیری

در این مقاله براساس روش پاسخ تصادفیه سیمونس، برای حفظ محرمانگی بیشتر یک روش پاسخ تصادفیه مرکب جدید برای برآورد نسبت حساس معرفی گردید و با شبیه سازی با استفاده از نرم‌افزار R کارایی آن با

روش سیمونس و روش پاسخ تصادفیده علوی و تاج‌الدینی (۱۳۹۴) مقایسه شد. نشان داده شد که کارایی آن از روش سیمونس بیشتر اما از روش علوی و تاج‌الدینی کمتر است. با استفاده از داده‌های تصادفیده با این روش، نسبت تقلب دانشجویان در امتحانات دانشگاه شهید چمران اهواز در نیمسال دوم تحصیلی ۹۵-۱۳۹۴ بر اساس یک نمونه تصادفی به حجم ۱۷۰ دانشجو، ۴۱۲/۰ با خطای معیار ۷۵/۰ برآورد و فاصله اطمینان ۹۵٪ برای نسبت تقلب (۵۵۹/۰، ۲۶۵/۰) محاسبه شد.

تقدیر و تشکر

نویسندگان از پیشنهادات داوران و ویراستار محترم مجله که موجب ارائه بهتر مقاله شدند تشکر می‌نمایند.

مراجع

[۱] علوی، س. م. ر. و تاج‌الدینی، م. (۱۳۹۴)، یک روش پاسخ تصادفیده جدید و مقایسه آن با روش سیمونس، مجله علوم آماری ایران، ۹، ۲۲۷-۲۳۹.

[۲] یزاری، ز و علوی، س. م. ر. (۱۳۹۳)، مدل پاسخ تصادفیده ی کمی اختیاری سه مرحله‌ای، مجله علوم آماری ایران، ۸، ۲۴۵-۲۶۰.

- [3] Alavi, S. M. R. and Tajodini, M. (2016), Maximum Likelihood Estimation of Sensitive Proportion Using Repeated Randomized Response Techniques, *Journal of Applied Statistics*, **43**, 563-571.
- [4] Chang, H., Wang, C. and Haung, K. (2005), On Estimating the Proportion of a Qualitative Sensitive Character Using Randomized Response Sampling, *Quality and Quantity*, **38**, 675-680.
- [5] Greenberg, B. G., Kuebler, R. R., Abernathy, J. R. and Horvitz, D. G. (1969), The Unrelated Question Randomized Response Model: Theoretical Framework, *Journal of the American Statistical Association*, **64**, 520-539.

- [6] Gjestvang, C. R. and Singh, S. (2006), A New Randomized Response Model, *Journal of the Royal Statistical Society, B*, **68**, 523-530.
- [7] Haug, K. (2004), A Survey Technique for Estimating the Proportion and Sensitivity in a Dichotomous Finite Population, *Statistica Neerlandica*, **58**, 75-82.
- [8] Hussain, Z. and Shabbir, J. (2007), Randomized Use of Warner's Randomized Response Model, *InterStat*: April 7, <http://interstat.statjournals.net/INDEX/Apr07.html>.
- [9] Kim, J. M., Warde, D. W. (2005), A Mixed Randomized Response Model, *Journal of Statistical Planning and Inference*, **133**, 211-221.
- [10] Kuk, A. Y. C. (1990), Asking Sensitive Questions Directly, *Biometrika*, **77**, 436-438.
- [11] Mangat, N. S. and Singh, R. (1990), An Alternative Randomized Response Procedure, *Biometrika*, **77**, 439-772.
- [12] Mangat, N. S. (1994), An Improved Randomized Response Strategy, *Journal of the Royal Statistical Society, B*, **56**, 93-95.
- [13] Moors, J. J. A. (1971), Optimization of the Unrelated Question Randomized Response Model, *Journal of the American Statistical Association*, **66**, 627-629.
- [14] Raghavarao, D. (1978), On an Estimation Problem in Warner's Randomized Response Technique, *Biometrics*, **34**, 87-90.
- [15] Mehta, S., Dass, B. K., Shabbir, J. and Gupta, G. (2012), A Three Stage Optional Randomized Response Model, *Journal of Statistical Theory and Practice*, **6**, 417-427.

- [16] Singh, S. (2002), Randomized Response Model, *Metrika*, **56**, 131-142.
- [17] Warner, S. L. (1965), Randomized Response: a Survey Technique for Eliminating Evasive Answer Bias, *Journal of the American Statistical Association*, **60**, 63-69.

A New Combined Randomized Response Model

Alavi, S. M. R., Joharzadeh, M., Chinipardaz, R.

Department of Statistics, Shahid Chamran University of Ahvaz, Ahvaz , Iran.

Abstract: Usually in survey sampling when the sensitive questions are asked directly, the respondents do not provide true answers. The randomized response techniques have been introduced to protect the privacy responses. In this article we focus on Simons randomized response technique for qualitative variables. Using the combination of the two different Simons' models, a new combined randomized response technique is introduced to increase protection of privacy. Using simulation in R package, efficiency of the proposed model is compared to the Simmons' and Alavi and Tajodini's (1394) models. Finally, the proposed model has been employed for estimating the proportion of student cheating in Shahid Chamran University.

Keywords: Randomized Rresponse, Simmons' method, Unrelated Questions, Combined Randomized Response, Proportion of Sensitive Attribute.

Mathematics Subject Classification (2010): 62D05.