مجله علوم آماری، پاییز و زمستان ۱۳۹۴
جلد ۹، شماره ۲، ص ۲۲۷-۲۳۹

یک روش پاسخ تصادفیده جدید و مقایسه آن با روش سیمونس

سید محمدرضا علوی، محبوبه تاج‌الدینی
گروه آمار، دانشگاه شهید چمران اهواز

تاریخ دریافت: ۱۳۹۲/۱۱/۲۹ تاریخ آخرین بازنگری: ۱۳۹۳/۱۲/۲۳

چکیده : معمولاً درنمونه‌گیری پاسخگو به سؤال‌های حساس پاسخ واقعی را نمی‌دهد. روش‌های پاسخ‌های تصادفیده برای حفاظت از محرمانگی پاسخ پاسخگو طراحی شده‌اند. در این مقاله تمرکز بر روش پاسخ تصادفیده برای متغیرهای کیفی بر اساس روش سیمونس است. با استفاده از ایده تکرار، روش جدید پاسخ تصادفیده مکرر معرفی و کارایی آن با روش سیمونس مقایسه شده است. سپس با این روش نسبت تقلب دانشجویان در امتحانات در دانشگاه شهید چمران اهواز برآورد گردیده است.

واژه‌های کلیدی : پاسخ تصادفیده، متغیر حساس، سؤال نامرتبط توزیع پواسون اریب اندازه، نمونه‌گیری طبقه‌بندی.

۱ مقدمه

در بسیاری از طرح‌های نمونه‌گیری از جامعه‌های انسانی دستیابی به پاسخ‌های صادقانه سؤال‌هایی که به‌طور مستقیم مطرح می‌شوند، بسیار مشکل است و در برخی موارد پاسخگویان به دلیل حساس بودن بعضی از سؤال‌ها حاضر به همکاری

آدرس الکترونیک مسئول مقاله: سید محمد رضا علوی، alavi-m@scu.ac.ir
کد موضوع‌بندی ریاضی (۲۰۱۰): ۶۲D۰۵

نمی‌شوند. این قبیل سؤالات ممکن است برای بعضی افراد ترساننده یا عذاب آور باشند و پاسخگو به راحتی پاسخ ندهد یا پاسخ واقعی را بیان نکند. برخی از سؤالات حساس فرار از پرداخت مالیات، رانندگی بی‌دقت، قماربازی، اعتیاد به الکل، تقلب در امتحانات و غیره است. در چنین وضعیت‌هایی، به‌جای تلاش برای نشان دادن یک پاسخ مستقیم[۱]، روش جالب توجه برای حفاظت از حریم شخصی فرد پاسخگو استفاده از روش پاسخ تصادفیده[۲] است.

ایده اصلی پاسخ‌های تصادفیده، به منظور تشویق به همکاری و پاسخ دادن صادقانه، توسط وارنر (۱۹۶۵) مطرح شد. روش پاسخ‌های تصادفیده با ادغام سؤال حساس در یک سؤال دیگر این امکان را فراهم می‌آورد که پاسخگو اطمینان داشته باشد که پاسخ واقعی علیه او به‌کار نمی‌رود.

سؤال دوم ممکن است مرتبط با سؤال حساس یا نامرتبط با آن مطرح شود. وارنر (۱۹۶۵) سؤال دوم را نقیض سؤال حساس مطرح کرد و پاسخ دهنده پس از انتخاب سؤال با استفاده از یک ابزار تصادفی کردن مانند پرتاب تاس، پرتاب سکه، دسته کارت یا غیره به آن پاسخ «بلی» یا «خیر» می‌دهد.

گرین برگ و همکاران (۱۹۶۹) استفاده از سؤال نامرتبط غیر حساس را برای پرسش دوم معرفی کردند که روش سیمونس نامیده می‌شود. گرین برگ و همکاران (۱۹۷۱) کاربرد روش پاسخ تصادفیده را در به‌دست آوردن داده‌های کمی مورد بررسی قرار دادند. آرناب (۱۹۹۹) در جریان نمونه تصادفی ساده با جایگذاری براساس تکرار افراد و با فرض مستقل بودن این تکرارها، برآوردکننده‌هایی برای نسبت حساس معرفی کرد. کریستوفیدز (۲۰۰۵) روش پاسخ تصادفیده را در نمونه‌گیری طبقه‌بندی[۳] به‌کار برد. چادری و پال (۲۰۰۸) پاسخ تصادفیده وارنر (۱۹۶۵) را بر اساس افراد متمایز نمونه به‌کار بردند. چادری و همکاران (۲۰۱۱) چند روش تصادفیده را با هم مقایسه کردند. علوی (۱۳۸۶) روش پاسخ‌های تصادفیده چند گزینه‌ای و علوی و چینی پرداز (۱۳۸۴) روش پاسخ‌های تصادفیده را

۱ Direct response
۲ Randomized response technique
۳ Stratification sampling

برای برآورد نسبت تقلب در دانشگاه استفاده کردند. در بخش ۲ روش پاسخ تصادفیده سیمونس معرفی می‌شود. در بخش ۳ با استفاده از ایده تکرار پاسخ تصادفیده توسط فرد پاسخگو، روشی جدید به نام تصادفیده مکرر معرفی و برآورد نسبت حساس توسط این روش پیشنهادی ارائه می‌شود. در بخش۴ کارایی[۴] گونه‌ای خاص از آن با نام روش تصادفیده مکرر پواسون اریب اندازه[۵] با روش سیمونس مقایسه می‌شود و در بخش ۵ با روش پیشنهادی نسبت تقلب دانشجویان دانشگاه شهید چمران برآورد می‌شود.

۲ روش پاسخ تصادفیده سیمونس

در روش سیمونس از فرد iام نمونه تقاضا می‌شود یک آزمایش برنولی با احتمال پیروزی p انجام دهد. اگر پیروزی رخ دهد، پاسخ سؤال حساس و اگر شکست رخ دهد، پاسخ سؤال نامرتبط را ارائه دهد. اگر متغیرهای مستقل برنولی Y_i ، X_i و T_i برای $i = 1, \ldots, n$ به ترتیب بیانگر پاسخ سؤال حساس، پاسخ سؤال نامرتبط و نتیجه آزمایش برنولی با احتمال‌های به ترتیب θ، μ_X و p باشند، واضح است که پاسخ تصادفیده I_i به‌صورت

$$I_i = Y_i(T_i) + X_i(1 - T_i) \quad i = 1, \ldots, n$$

است، که یک متغیر برنولی با احتمال پیروزی γ به‌صورت

$$\gamma = E(I_i) = E_1 E_2(I_i) = E_1(pY_i + (1 - p)X_i) = p\theta + (1 - p)\mu_X$$

می‌باشد، که در آن E_1 بیانگر امید ریاضی روی تمام نمونه‌های ممکن و E_2 امیدگیری تحت ابزار تصادفی کردن است. بر اساس یک نمونه تصادفی از پاسخ‌های تصادفیده، برآورد نااریب θ و واریانس آن به‌ترتیب به‌صورت

$$\hat{\theta} = \frac{\bar{I} - (1 - p)\mu_X}{p}$$

۴ Efficiency

۵ Size biased Poisson-randomized response technique

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

معرفی شد، که در آن $\bar{I}$ میانگین پاسخ بله در نمونه و برآوردی نااریب برای γ است، یعنی

$$\hat{\gamma} = \frac{n'}{n} = \bar{I}$$

که در آن n حجم نمونه و n' تعداد پاسخ بله در نمونه است. برآوردی برای $V(\hat{\theta})$ به‌صورت

$$\hat{V}(\hat{\theta}) = \frac{V(\bar{I})}{p^2} = \frac{\hat{\gamma}(1-\hat{\gamma})}{np^2}$$

قابل حصول است.

۳ روش پاسخ تصادفیده مکرر

در این بخش بر مبنای روش سیمونس روش پیشنهادی معرفی می‌شود، که در آن از هر فرد در نمونه تصادفی خواسته می‌شود که چند بار روش تصادفیده سیمونس را تکرار کند. فرض کنید f_i تعداد تکرار پاسخ تصادفیده فرد iام، از متغیر تصادفی گسسته N با تابع چگالی احتمال $f_N(f_i)$ پیروی کند. اگر X_{ij}، Y_{ij} و T_{ij} به ترتیب بیانگر پاسخ سؤال حساس، پاسخ سؤال نامرتبط و نتیجه آزمایش برنولی فرد iام در تکرار jام با احتمال‌های به‌ترتیب θ، μ_X و p باشند، در آن صورت پاسخ تصادفیده فرد iام در تکرار jام (I_{ij}) دارای توزیع برنولی با احتمال پیروزی

$$\gamma = p\theta + (1-p)\mu_X$$

است. واضح است که توزیع شرطی تعداد پاسخ بله فرد iام، دوجمله‌ای است یعنی

$$\sum_{j=1}^{f_i} I_{ij} \mid_N = f_i \sim B(f_i, \gamma)$$

در نتیجه $m_i = \frac{\sum_{j=1}^{f_i} I_{ij}}{f_i}$ یک برآورد نااریب برای γ است. بنابراین یک برآورد نااریب برای نسبت حساس بر اساس پاسخ‌های تصادفیده فرد iاُم نمونه به‌صورت

$$\hat{\theta}_i = \frac{m_i - (1-p)\mu_X}{p} \quad i = 1, \ldots, n \tag{۲}$$

به‌دست خواهد آمد. واریانس این برآوردکننده برابر است با:

$$V(\hat{\theta}_i) = \frac{1}{p^2}[p(1-p)(\theta + \mu_X + 2\theta\mu_X) + (1-p)^2\mu_X(1-\mu_X)]E(\frac{1}{N}) + \theta(1-\theta)$$

و برآورد این واریانس به‌صورت

$$\hat{V}(\hat{\theta}_i) = \frac{\hat{V}(m_i)}{p^2} = \frac{m_i(1-m_i)}{f_i p^2} \tag{۳}$$

است. با استفاده از تمام مشاهدات نمونه برآوردی دیگر برای θ به‌صورت

$$\hat{\theta_P} = \frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_i \tag{۴}$$

حاصل می‌شود، که واریانس آن برابر است با:

$$V(\hat{\theta_P}) = \frac{1}{np^2}[p(1-p)(\theta + \mu_X + 2\theta\mu_X) + (1-p)^2\mu_X(1-\mu_X)]E(\frac{1}{N}) + \frac{\theta(1-\theta)}{n} \tag{۵}$$

با فرض مستقل بودن X، Y و N برآورد واریانس برآوردکننده θ به‌صورت

$$\hat{V}(\hat{\theta_P}) = \frac{1}{n^2p^2}\sum_{i=1}^{n}\hat{V}(\hat{\theta}_i) = \frac{1}{n^2p^2}\sum_{i=1}^{n}\frac{m_i(1-m_i)}{f_i} \tag{۶}$$

حاصل می‌شود. چون هر فرد در نمونه حداقل یک بار لازم است که روش تصادفیده سیمونس را انجام دهد، یکی از توزیع‌های مناسب برای N توزیع پواسون اریب اندازه (علوی و چینی پرداز، ۲۰۰۹) با تابع چگالی احتمال

$$f_N(f_i) = \frac{\mu^{f_i - 1}e^{-\mu}}{(f_i - 1)!} \qquad f_i = 1, \ldots$$

است و داریم

$$E(\frac{1}{N}) = \frac{(1-e^{-\mu})}{\mu}$$

که با جایگذاری آن در (۵) داریم:

$$V(\hat{\theta_P}) = \frac{1}{np^2}[p(1-p)(\theta+\mu_X+2\theta\mu_X) \qquad (7)$$
$$+(1-p)^2\mu_X(1-\mu_X)]\frac{(1-e^{-\mu})}{\mu} + \frac{\theta(1-\theta)}{n}$$

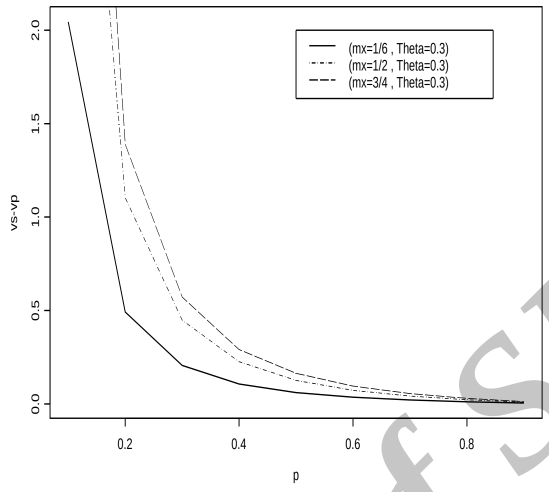
۴ مقایسه کارایی دو روش پیشنهادی و سیمونس

با توجه به نااریب بودن برآورد نسبت حساس در هر دو روش، برای ارزیابی و مقایسه کارایی دو روش پیشنهادی و سیمونس از تفاضل عبارات (۷) و (۱) داریم:

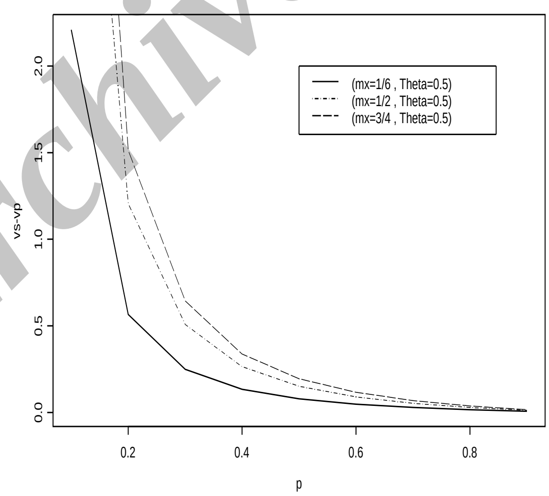
$$V_S - V_P = (1 - \frac{1-e^{-\mu}}{\mu})[\frac{\theta(1-p)}{np} + \frac{(1-p)\mu_X}{np^2}(1-(1-p)\mu_X)]$$
$$-[\frac{2\theta(1-p)\mu_X}{np}(1+\frac{1-e^{-\mu}}{\mu}) + \frac{(1-p)^2\mu_X}{np^2}\frac{(1-e^{-\mu})}{\mu}]$$

که در آن V_S و V_P به‌ترتیب واریانس برآوردکننده‌ها با روش‌های سیمونس و پیشنهادی هستند. شکل‌های ۱ تا ۳ تفاوت واریانس دو روش $(V_S - V_P)$ را برای اندازه نمونه $n = 100$ و $\mu = 3$ به ازای مقادیر مختلف p، μ_X و θ نشان می‌دهند. همان‌طور که ملاحظه می‌شود اختلاف واریانس روش سیمونس از روش پیشنهادی، همیشه مثبت است و با افزایش p این اختلاف کاهش می‌یابد. با افزایش اندازه نمونه و مقدار μ همین روند نیز مشاهده می‌شود که از رسم نمودارهای آن‌ها برای اختصار صرف نظر شده است. بنابراین کارایی روش پیشنهادی از روش سیمونس بیشتر است.

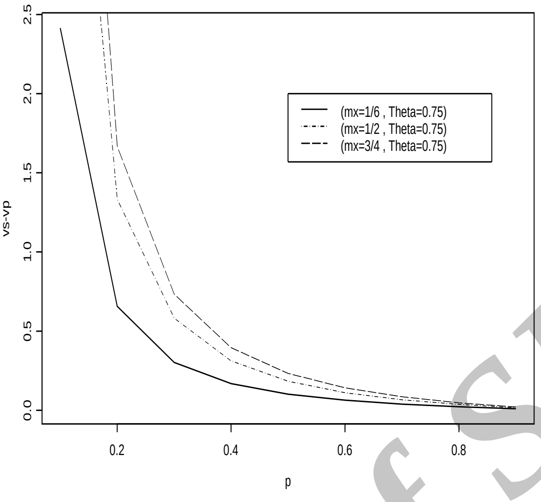
از طرفی بیشترین احساس حفاظت از محرمانگی برای پاسخگو وقتی حاصل می‌شود که احتمال انتخاب بین سؤال حساس و نامرتبط یکسان، یعنی p برابر ۰/۵ باشد. به همین جهت در بخش کاربرد برای اجرای روش پیشنهادی از انداختن سکه برای انتخاب سؤال حساس و نامرتبط استفاده شده است.



شکل ۱: تفاوت واریانس‌های دو روش به‌ازای مقادیر مختلف p، μ_X و $\theta = ۰/۳$



شکل ۲: تفاوت واریانس‌های دو روش به‌ازای مقادیر مختلف p، μ_X و $\theta = ۰/۵$



شکل ۳: تفاوت واریانس‌های دو روش به‌ازای مقادیر مختلف p، μ_X و $\theta = ۰/۷۵$

۵ برآورد نسبت تقلب با روش پیشنهادی

تقلب در امتحانات از جمله آفت‌های آموزشی در دانشگاه‌ها و مراکز آموزشی است که باعث افت تحصیلی نیز می‌شود. درصد تقلب دانشجویان در دانشگاه می‌تواند معیاری برای سلامت برنامه‌های آموزشی هر دانشگاهی محسوب شود. پرسش درباره تقلب در امتحانات برای اکثر دانشجویان، موضوعی بسیار حساس است. در این بخش از روش تصادفیده مکرر پواسون اریب اندازه برای برآورد نسبت تقلب در دانشگاه استفاده شده است. در این پژوهش دانشجویان دانشگاه شهید چمران اهواز به عنوان جامعه آماری در نظر گرفته شده و از نمونه‌گیری تصادفی طبقه‌بندی استفاده شده است. دانشکده‌ها به عنوان طبقه‌ها منظور شده و از هر طبقه نمونه‌ای تصادفی انتخاب شده است.

گردآوری داده‌ها در ماه اردیبهشت سال ۱۳۹۲ با مراجعه به دانشکده‌های مختلف صورت گرفته است. از هر دانشکده چندین کلاس به‌طور تصادفی انتخاب و پرسش نامه‌ها میان دانشجویان توزیع شده‌اند. سؤال نامرتبط در پرسش‌نامه می‌بایست از دو

ویژگی برخوردار می‌بود. اول معلوم بودن احتمال پاسخ بله آن سؤال، برای مجریان طرح نمونه‌گیری و دوم مخفی ماندن پاسخ آن سؤال از مصاحبه کننده. به همین جهت آمدن عدد شش در انداختن تاس به عنوان سؤال نامرتبط در پرسش‌نامه در نظر گرفته شد. به همراه هر پرسش‌نامه یک عدد سکه و یک تاس به دانشجو تحویل داده شده و از او درخواست شده بود به تعداد f_i بار روش تصادفیده مطرح شده در پرسش‌نامه را که شامل دو مرحله بود انجام دهد و پاسخ‌های تصادفیده را در محل مورد نظر درج نماید. اعداد f_i درج شده در پرسش‌نامه‌ها قبلاً توسط محقق از توزیع پواسون اریب اندازه با پارامتر $\mu = ۳$ با استفاده از نرم افزار R تولید شده بودند. دو مرحله روش تصادفیده به شرح زیر در پرسش‌نامه مطرح شده بودند:

مرحله ۱- تاس را به دور از چشم دیگران پرتاب کرده و شماره ظاهر شده را به ذهن خود بسپارید.

مرحله ۲- سکه را به تصادف بیندازید، اگر شیر آمد فقط به سؤال الف و اگر خط آمد فقط به سؤال ب پاسخ دهید و نتیجه را در پاسخ‌نامه درج شده در انتهای پرسش‌نامه وارد نمایید.

الف- آیا شما جزء دانشجویانی هستید که تا کنون در امتحانات رسمی تقلب کرده‌اند؟

ب-آیا نتیجه انداختن تاس شما عدد ۶ است؟

مراحل ۱ و ۲ را به تعداد f_i بار تکرار کنید و هر بار نتیجه را در پاسخ‌نامه علامت گذاری کنید.

جدول ۱: نمونه پاسخ‌نامه درج شده در انتهای پرسش‌نامه

	شماره پاسخ نامه							
پاسخ	۱	۲	۳	.	.	.	۱۲	۱۳
بله								
خیر								

در این پرسش‌نامه چون از ابزار انداختن سکه برای انتخاب سؤال حساس استفاده شده است، لذا $p = \frac{1}{2}$ در نظر گرفته شد. با توجه به اینکه بیشترین عدد

تولید شده توسط نرم‌افزار از توزیع پواسون اریب اندازه، عدد ۱۳ بوده، لذا در پاسخ‌نامه ۱۳ ستون در نظر گرفته شده است. خلاصه‌ای از پاسخ‌های تصادفیده دانشکده‌ها در جدول ۲ آمده است. با جایگذاری $p = \frac{۱}{۲}$ و $\mu_X = \frac{۱}{۶}$ در رابطه‌های (۲) و (۳) برآورد نسبت تقلب و برآورد واریانس آن براساس پاسخ‌های تصادفیده فرد iام به‌صورت

$$\hat{V}(\hat{\theta}_i) = \frac{۴m_i(۱-m_i)}{f_i}, \quad \hat{\theta}_i = ۲m_i - \frac{۱}{۶}$$

حاصل شده‌اند. براساس یک نمونه تصادفی برآورد نااریب کاراتر و برآورد واریانس آن بنابر روابط (۴) و (۶) به‌ترتیب به‌صورت

$$\begin{aligned} \hat{\theta}_P &= \frac{۱}{n}\sum_{i=۱}^{n}\hat{\theta}_i \\ \hat{V}(\hat{\theta}_P) &= \frac{۴}{n^۲}\sum_{i=۱}^{n}\frac{m_i(۱-m_i)}{f_i} \end{aligned}$$

به‌دست می‌آیند. با به‌کار بردن این روابط برای طبقه (دانشکده) hام داریم:

$$\begin{aligned} \hat{\theta}_{hi} &= ۲m_{hi} - \frac{۱}{۶} \\ \hat{\theta}_{Ph} &= \frac{۱}{n_h}\sum_{i=۱}^{n_h}\hat{\theta}_{hi} \\ \hat{V}(\hat{\theta}_{Ph}) &= \frac{۴}{n_h^۲}\sum_{i=۱}^{n_h}\frac{m_{hi}(۱-m_{hi})}{f_{hi}} \end{aligned}$$

که در آن n_h حجم نمونه انتخابی از طبقه h و m_{hi} نسبت پاسخ‌های بله فرد iام از طبقه hام می‌باشد. اگر تعداد طبقات برابر L در نظر گرفته شود، برآورد نسبت تقلب برای کل جامعه به‌صورت

$$\hat{\theta}_P = \sum_{h=۱}^{L}\frac{N_h}{N}\hat{\theta}_{Ph} = \sum_{h=۱}^{L}W_h\hat{\theta}_{Ph}$$

به‌دست خواهد آمد، که در آن N_h حجم طبقه h، N حجم کل جامعه ($\sum_{h=۱}^{L}N_h$) و $W_h = \frac{N_h}{N}$ وزن طبقه h هستند. واریانس $\hat{\theta}_P$ نیز به‌صورت

$$V(\hat{\theta}_P) = \sum_{h=۱}^{L}W_h^۲V(\hat{\theta}_{Ph})$$

است و برآورد آن عبارت است از:

$$\hat{V}(\hat{\theta}_P) = \sum_{h=۱}^{L} W_h^۲ \hat{V}(\hat{\theta}_{Ph})$$

برآورد نسبت تقلب دانشکده‌های مختلف در جدول ۲ ارائه شده است. نسبت تقلب دانشجویان دانشگاه، براساس روش پیشنهادی به کار رفته در این پژوهش، ۰/۶۴ برآورد گردید. خطای معیار این برآورد ۰/۰۶ برآورد شد.

بحث و نتیجه‌گیری

در این مقاله بر اساس ایده روش پاسخ تصادفیده سیمونس، روش پاسخ تصادفیده مکرر پواسون اریب اندازه برای برآورد نسبت حساس معرفی گردید و نشان داده شد که کارایی آن از روش سیمونس بیشتر است. نسبت تقلب دانشجویان با استفاده از این روش در دانشکده‌های مختلف دانشگاه شهید چمران اهواز محاسبه و در نهایت نسبت تقلب در این دانشگاه ۰/۶۴ با خطای معیار ۰/۰۶ برآورد گردید.

جدول ۲: خلاصه داده‌ها و برآورد نسبت تقلب دانشکده‌های مختلف

دانشکده	تعداد N_h	حجم نمونه n_h	وزن W_h	میانگین بله	برآورد تقلب $\hat{\theta}_h$	برآورد واریانس $\hat{V}(\hat{\theta}_h)$
ادبیات	۷۸۴	۶۶	۰/۰۹۸	۰/۲۸۹	۰/۴۷۱	۰/۰۰۲
اقتصاد	۱۵۰۶	۵۴	۰/۱۸۸	۰/۴۴۴	۰/۶۵۳	۰/۰۰۴
الهیات	۶۰۰	۶۰	۰/۰۷۵	۰/۳۴۹	۰/۵۳۷	۰/۰۰۲
تربیت بدنی	۱۹۰	۴۳	۰/۰۲۴	۰/۴۹۷	۰/۷۷۰	۰/۰۰۲
علوم	۷۶۳	۹۱	۰/۰۹۵	۰/۳۱۰	۰/۴۸۰	۰/۰۰۱
علوم‌آب	۱۸۳	۴۳	۰/۰۲۳	۰/۴۰۰	۰/۶۸۲	۰/۰۰۳
علوم تربیتی	۹۹۰	۷۳	۰/۱۲۳	۰/۴۰۰	۰/۶۶۷	۰/۰۰۲
علوم ریاضی	۸۰۰	۷۷	۰/۰۱۰	۰/۴۲۵	۰/۶۷۴	۰/۰۰۲
کشاورزی	۵۵۵	۵۰	۰/۰۷۰	۰/۳۹۲	۰/۵۹۸	۰/۰۰۲
مهندسی	۱۶۵۲	۸۸	۰/۲۰۶	۰/۴۵۴	۰/۷۶۷	۰/۰۰۱

مراجع

علوی، س. م. ر. (۱۳۸۶)، تصادفی کردن پاسخ سؤالات چند گزینه‌ای و برآورد نسبت گزینه‌های تقلب دانشجویان در دانشگاه شهید چمران اهواز، مجله اندیشه آماری، **۱۲**، ۱۳-۱۹.

علوی، س. م. ر.، چینی پرداز، ر. (۱۳۸۴)، مقایسه تقلب دانشجویان در دانشکده‌ها با استفاده از روش پاسخ تصادفی شده، گزارش نهایی طرح تحقیقاتی شماره ۵۱۲، دانشگاه شهید چمران اهواز.

Alavi, S. M. R. and Chinipardaz, R. (2009), From-Invariance Under Weighted Sampling, *Statistics*, **43**, 81-90.

Arnab, R. (1999), On Use of Distinct Respondents in RR Survey, *Biometrika*, **41**, 507-513.

Chaudhuri, A., Bose, M. and Dihidar, K. (2011), Estimating Sensitive Proportions by Warner's Randomized Response Technique Using Multiple Randomized Response from Distinct Persons Sample, *Statistical Papers*, **52**, 111-124.

Chaudhuri, A. and Pal, S. (2008), Estimating Sensitive Proportions from Warner's Randomized Response in Alternative Ways Restricting to only Distinct Units Sampled, *Metrika*, **68**, 147-156.

Christofides, T. C. (2005), Randomized Response in Stratified Sampling, *Statistical Planning and Inference*, **128**, 303-310.

Greenberg, B. G., Abul-Ela, A. L. A., Simmons, W. R. and Horvitz, D. G. (1969), The Unrelated Question Randomized Responce Model: Theorical Framework, *American Statistical Association*, **64**, 520-539.

Greenberg, B. G., Kuebler Jr, R. R., Abernathy, J. R. and Horvitz, D. G. (1971), Application of the Randomized Response Technique in Obtaining Quantitative Datd, *American Statistical Association*, **66**, 243-250.

Warner, S. L. (1965), Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias, *American Statistical Association*, **60**, 63-69.