شاپا چاپی: ۲۵۸۸-۲۵٤٦
شاپا الکترونیکی: ۲۵۸۸-۲۵٥٤
پژوهش‌های ریاضی
Homepage: https://mmr.khu.ac.ir/

Efficiency of shrinkage pretest estimator for the intercept parameter in simple linear regression model

Z. Mahdizadeh[1], M. Naghizadeh Qomi[2] *

1. Department of Statistics, University of Mazandaran, Babolsar, Iran.

2. Corresponding Author, Department of Statistics, University of Mazandaran, Babolsar, Iran.

* E-mail: m.naghizadeh@umz.ac.ir

Article Info

Article type:
Research Article

Article history:
Received:
29 November 2020
Received in revised form:
22 November 2021
Accepted:
10 December 2021
Published online:
3 December 2023

Keywords:
Non-sample Prior information,
Shrinkage pretest estimator,
Reflected normal loss function.

ABSTRACT

Introduction

Traditionally the classical estimators of unknown parameters of the linear regression model are based exclusively on the sample information, like the maximum likelihood method. Sometimes, in practice the researcher has some prior information about the unknown intercept parameter as a guess that is called non-sample prior information. In this case, the shrinkage and shrinkage preliminary test estimators are introduced by a linear combination of sample and non-sample prior information. Now, if the non-sample prior information is available according to the experimenter's experiences or the results of previous experiments, it can be expressed in the form of a preliminary hypothesis test and the desired parameter can be estimated using the shrinkage pretest estimator. The results show that the closer the guessed value to the real parameter, the better shrinkage pretest estimator performs compared to the maximum likelihood estimator.

In the past years, in the regression estimation problem, the behavior of the shrinkage and shrinkage pretest estimators under the squared error loss (SEL) and LINEX loss functions was investigated. The squared error loss (SEL) function is popularly used for estimating the unknown parameter in decision theory because of its mathematical and interpretational convenience. Due to the symmetric nature, the use of SEL function may not be appropriate, when positive and negative errors have different consequences. The SEL and LINEX loss functions are symmetric and asymmetric loss functions, respectively, but both are unbounded. Also, The SEL and LINEX loss functions have an infinite maximum value which isn't always appropriate. Sometimes in practice it is necessary to use bounded loss functions to estimate parameters. So we used the reflected normal loss function to estimate parameters, which is the bounded loss function.

Material and methods

In this article, the shrinkage and shrinkage pretest estimators are introduced for the intercept parameter of the simple linear regression according to the non-sample prior information and their risk functions and relative efficiencies are derived under the reflected normal loss function. The behavior of shrinkage pretest estimator is compared with respect to the shrinkage and

maximum likelihood estimator using graphical method. The intervals where the shrinkage pretest estimator has less risk compared to the maximum likelihood estimator are presented. The optimum value of the significant level of test is determined using the max-min method. Besides, several methods of finding shrinkage coefficient of the shrinkage pretest estimators are proposed. Then, the application of the proposed estimators is shown using a real data set.

Results and discussion

In comparison between the desired estimators, our findings show that the shrinkage and shrinkage pretest estimator are better than the maximum likelihood estimator when non-sample prior information is close to the real value. Also, the shrinkage pretest estimator with smaller level of significance has higher efficiency. An important issue for the shrinkage pretest estimator is the proper selection of the shrinkage coefficient. So, we applied several methods for finding the shrinkage coefficient to obtain shrinkage pretest estimator.

Conclusion

According to the reported results, the shrinkage pretest estimator has smaller risk than the maximum likelihood estimator in neighborhood of null hypothesis. Therefore, in practice, if the researcher can have prior information of the unknown intercept parameter from previous knowledge or experience that is not far from the real value, the shrinkage pretest estimator would be the best choice as intercept estimator.

The effective intervals where the shrinkage pretest estimator has less risk compared to the maximum likelihood estimator was presented. By increasing the significant level of the test and the sample size, the length of the effective interval decreases. The optimal value of the significant level of the test was obtained using the max-min method. Then, the application of the proposed estimators is shown using a real data set.

How to cite: Z. Mahdizadeh,, M. Naghizadeh Qomi, (2023). Efficiency of shrinkage pretest estimator for the intercept parameter in simple linear regression model. Mathematical Researches, 9 (2), 125 – 145.

Publisher: Kharazmi University

پژوهش‌های ریاضی
شاپا چاپی: ۲۵۸۸-۲۵٤٦
شاپا الکترونیکی: ۲۵۸۸-۲۵٥٤
Homepage: https://mmr.khu.ac.ir/

کارایی براوردگر پیش‌آزمون انقباضی پارامتر عرض از مبدأ در مدل رگرسیونی خطی ساده

زهره مهدی‌زاده [1]، **مهران نقی‌زاده قمی** [2] *

1. گروه آمار دانشگاه مازندران، ایران.
2. گروه آمار دانشگاه مازندران، ایران. رایانامه: m.naghizadeh@umz.ac.ir

اطلاعات مقاله

نوع مقاله: مقاله پژوهشی

تاریخ دریافت: 1399/9/9
تاریخ بازنگری: 1400/9/1
تاریخ پذیرش: 1400/9/10
تاریخ انتشار: 1402/9/12

واژه‌های کلیدی:
اطلاع پیشین،
براوردگر پیش‌آزمون انقباضی،
تابع زیان نرمال بازتابیده

چکیده

معمولا برای براورد پارامترهای یک مدل رگرسیونی از روش های مرسوم براوردیابی مانند روش ماکسیمم درستنمایی استفاده می‌شود. گاهی اوقات محقق دارای اطلاع پیشین درمورد پارامتر عرض از مبدأ به صورت یک حدس است که به آن اطلاع پیشین غیرنمونه‌ای می‌گویند. در این مقاله براوردگرهای انقباضی و پیش آزمون انقباضی برای پارامتر عرض از مبدأ با توجه به اطلاع پیشین غیرنمونه ای معرفی و کارایی نسبی آن ها تحت تابع زیان نرمال بازتابیده مورد بررسی قرار می‌گیرد. رفتار براوردگر پیش‌آزمون انقباضی در مقایسه با براوردگرهای انقباضی و ماکسیمم درستنمایی به صورت گرافیکی ارزیابی می‌شود. بازه‌هایی که براوردگر پیش آزمون انقباضی دارای مخاطره کمتری نسبت به براوردگر ماکسیمم درستنمایی است ارائه می‌شود. نتایج نشان می‌دهد هرچه مقدار حدس زده شده به پارامتر واقعی نزدیکتر باشد براوردگر پیش‌آزمون انقباضی عملکرد بهتری نسبت به برآوردگر ماکسیمم درستنمایی دارد. همچنین با استفاده از روش ماکس‌مین مقدار بهینه سطح معنی‌داری آزمون تعیین می‌شود. سپس با استفاده از یک مجموعه داده واقعی، کاربرد براوردگرهای پیشنهادی نشان داده می‌شود.

استناد: زهره مهدی‌زاد، مهران نقی‌زاده قمی (1402). کارایی براوردگر پیش‌آزمون انقباضی پارامترعرض از مبدأ درمدل رگرسیونی خطی ساده. *پژوهش‌های ریاضی*، 9 (2)، 125 – 145.

ناشر: دانشگاه خوارزمی

مقدمه

در روش‌های کلاسیک آمار، براساس اطلاعات موجود در نمونه و با روش‌های معمول مانند روش ماکسیمم درستنمایی به براوردیابی پارامتر نامعلوم می‌پردازند. گاهی در عمل، محقق دارای اطلاعاتی درباره پارامتر نامعلوم به صورت یک حدس یا گمان است که اطلاعات غیرنمونه‌ای نامیده می‌شوند. در این حالت، براوردگر پیش‌آزمون انقباضی[1] با ترکیب خطی اطلاعات غیرنمونه‌ای و اطلاعات موجود در نمونه معرفی می‌شود که دارای کارایی نسبی بیشتری نسبت به براوردگرهای معمول می باشد [1].

در یک مدل رگرسیونی خطی ساده برای براورد ضرایب رگرسیونی معمولا از روش‌های کلاسیک با استفاده از اطلاعات نمونه ای مانند ماکسیمم درستنمایی استفاده می‌شود. حال اگر اطلاعات پیشین غیرنمونه‌ای با توجه به تجربیات آزمایشگر یا نتایج آزمایش‌های قبلی در اختیار باشد می‌توان این اطلاعات غیرنمونه‌ای را در قالب یک آزمون فرض مقدماتی بیان کرد و به کمک براوردگر پیش‌آزمون انقباضی به براورد پارامتر مورد نظر پرداخت. در مساله براورد رگرسیونی، [5] و [13] رفتار تعدادی از براوردگرها از جمله براوردگرهای انقباضی و پیش‌آزمون را برای پارامتر عرض از مبدأ مدل رگرسیونی خطی ساده تحت تابع زیان توان دوم خطا بررسی کردند. [2] کارایی نسبی براوردگر پیش‌آزمون را تحت تابع زیان لاینکس[2] مورد مطالعه قرار دادند. تابع زیان توان دوم خطا به علت راحتی در تحلیل و محاسبات ریاضی برای براورد پارامتر نامعلوم در نظریه تصمیم استفاده می‌شود. از طرفی به دلیل متقارن بودن تابع زیان توان دوم خطا ممکن است، تابع زیان مناسبی نباشد زیرا به بیش براوردی و کم براوردی جریمه‌های یکسانی نسبت می‌دهد. تابع زیان‌های توان دوم خطا و لاینکس غیرکراندار هستند. گاهی در عمل زیان ناشی از براورد پارامترها باید کراندار باشد لذا لازم است که از توابع زیان کراندار استفاده کرد. در این مقاله، به بررسی کارایی نسبی براوردگر پیش‌آزمون انقباضی پارامتر عرض از مبدأ در مدل رگرسیونی خطی ساده تحت تابع زیان نرمال بازتابیده می‌پردازیم. تابع زیان نرمال بازتابیده که توسط [9] معرفی شد، دارای فرمی به صورت

$$L(D) = k(1- e^{-\frac{D^2}{2l^2}}) \qquad (1)$$

است که در آن $D=d-q$ ، d براوردگر پارامتر q، $k>0$ ماکسیمم تابع زیان و $l>0$ پارامتر شکل است. تابع زیان (1) کراندار و اساسا برگردانی از تابع چگالی احتمال نرمال از بالا می‌باشد و بر همین اساس تابع زیان نرمال بازتابیده نامیده می‌شود. شکل 1 تابع زیان نرمال بازتابیده را به ازای $k=1$ و مقادیر مختلف $l=1,5,10$ نشان می‌دهد. ویژگی‌های تابع زیان (1) عبارتند از : $L(D)$

۱) تابعی کراندار است.

2) $L(D)$ به‌صورت نامتناهی مشتق‌پذیر و در نتیجه پیوسته است.

[1] Shrinkage pretest estimator

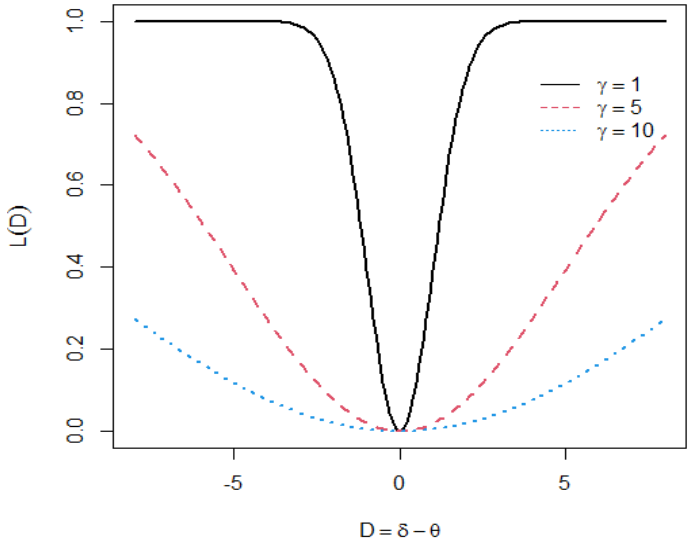
[2] LINEX loss function

3) $L(D)$ دارای مینیمم یکتای صفر به ازای $D=0$ است و در بازه $(-\infty,0)$ یک تابع نزولی و در بازه $(0,\infty)$ یک تابع صعودی است. همچنین $\lim_{D\to\pm\infty} L(D)=k$.

4) تابع زیان (1) حالت خاصی از فاصله هلینگر[1] به ازای $k=1$ و $g=2s$ به فرم

$$H^2(f(x|\theta),f(x|\delta))=\frac{1}{2}E_\theta\left\{\left(\frac{\sqrt{f(X|\theta)}-\sqrt{f(X|\delta)}}{\sqrt{f(X|\theta)}}\right)^2\right\},$$

است که در آن $f(x|\theta)$ و $f(x|\delta)$ به ترتیب تابع‌های چگالی احتمال توزیع نرمال $N(\theta,s^2)$ و $N(\delta,s^2)$ است. تابع زیان (1) توسط [11]،[8] و [7] در مسائل مختلف براوردیابی مورد استفاده قرار گرفت. از آنجایی که مقدار $k>0$ هیچ تأثیری در نتایج ندارد بنابراین در این مقاله بدون از دست دادن کلیت مسئله $k=1$ در نظر گرفته می‌شود.



شکل 1: تابع زیان نرمال بازتابیده برای $k=1$.

در این مقاله، در بخش 2، مدل رگرسیونی خطی ساده، براوردگرهای انقباضی و پیش‌آزمون انقباضی پارامتر عرض از مبدا معرفی می‌شوند. تابع مخاطره براوردگرهای ماکسیمم درستنمایی، انقباضی و پیش‌آزمون انقباضی پارامتر عرض از مبدا مدل رگرسیونی در بخش 3 محاسبه می‌شوند. در بخش 4 کارایی نسبی این سه براوردگر به صورت گرافیکی مورد بررسی قرار می

[1] Hellinger distance

گیرند. بازه‌های موثر که به ازای آن براوردگر پیش‌آزمون انقباضی نسبت به براوردگر ماکسیمم درستنمایی دارای مخاطره کمتری است ارائه می‌شود. همچنین مقدار بهینه سطح آزمون معنی‌داری با استفاده از روش ماکس‌مین به دست می‌آید. در بخش 5 یک مجموعه داده واقعی برای کاربرد براوردگرهای انقباضی و پیش‌آزمون انقباضی ارائه خواهد شد. در بخش پایانی به بحث و نتیجه گیری پرداخته می‌شود.

2 معرفی مدل و براوردگر پیش‌آزمون انقباضی

فرض کنید یک نمونه n تایی از مدل رگرسیونی خطی ساده به صورت

$$Y_i = b_0 + b_1 x_i + e_i \qquad i = 1,2,\ldots,n$$

در اختیار داریم که در آن x_i متغیر پیشگو، Y_i متغیر تصادفی پاسخ، b_0 پارامتر عرض از مبدأ، b_1 پارامتر شیب و $e_1,\ldots,e_n$ یک نمونه تصادفی n تایی از $N(0,s^2)$ می‌باشد.

مطابق با [6] براوردهای ماکسیمم درستنمایی پارامتر عرض از مبدأ b_0 و شیب b_1 در مدل رگرسیون خطی ساده به صورت $\hat{b}_0 = \bar{y} - \hat{b}_1\bar{x}$ و $\hat{b}_1 = S_{xx}^{-1}S_{xy}$ هستند که در آن $S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2$، $S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$، $\bar{x} = n^{-1}\sum_{i=1}^{n} x_i$ و $\bar{y} = n^{-1}\sum_{i=1}^{n} y_i$ است. $\hat{b}_0$ به عنوان براوردگر دارای توزیع نرمال با میانگین b_0 و واریانس s^2H است که در آن $H = \frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}$ می‌باشد، به عبارت دیگر

$$Z = \frac{\hat{b}_0 - b_0}{s\sqrt{H}} : N(0,1). \qquad (2)$$

فرض می‌کنیم محقق دارای اطلاع پیشین (باور یا عقیده) غیرنمونه‌ای در خصوص پارامتر b_0 است که با b_{00} نشان می‌دهیم که از طریق تجربه یا مطالعات گذشته قابل دسترس باشد. بر اساس [10] براوردگر انقباضی با ترکیب خطی از مقدار حدس زده شده و براوردگر نمونه‌ای به صورت

$$\hat{b}_0^{SE} = d\hat{b}_0 + (1-d)b_{00}$$

تعریف می‌شود، که در آن $0 \le d \le 1$ ضریب انقباضی است و توسط محقق با توجه به عقیده او نسبت به مقدار حدس زده شده تعیین می‌شود. اگر $d = 0$ تنها از مقدار حدسی و اگر $d = 1$ از براوردگر نمونه‌ای استفاده می‌شود. اطلاع پیشین غیرنمونه‌ای می‌تواند به کمک فرض صفر $H_0 : b_0 = b_{00}$ در مقابل $b_0 \neq b_{00}$ آزمون شود. براوردگر پیش‌آزمون انقباضی پارامتر b_0 از ترکیب اطلاع پیشین غیرنمونه‌ای (b_{00}) و اطلاعات نمونه‌ای ($\hat{b}_0$) به صورت

$$\hat{\beta}_0^{SP} = \begin{cases} \delta\hat{\beta}_0 + (1-\delta)\beta_{00} & F_{1,n} < F_{1-\alpha} \\ \hat{\beta}_0 & F_{1,n} > F_{1-\alpha} \end{cases}$$
$$= \hat{\beta}_0 - (1-\delta)(\hat{\beta}_0 - \beta_{00})I(A)$$

تعریف می‌شود، که در آن $I(A)$ تابع نشانگر مجموعه $A = \{F_{1,n} < F_{1-\alpha}\}$ ، $F_{1,n}$ آماره آزمون به صورت $F_{1,n} = \frac{(\hat{\beta}_0 - \beta_{00})^2}{S_n^2 H}$ با $S_n^2 = \frac{1}{(n-2)}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$ که براوردگری ناریب برای σ^2 است و در آن $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ می‌باشد و $F_{1-\alpha} = F_{1,n}(1-\alpha)$ مقدار چندک مرتبه $100(1-\alpha)$ درصد توزیع فیشر با درجات آزادی به ترتیب 1 و $\nu = n-2$ است.

3 محاسبه تابع مخاطره $\hat{\beta}_0$، $\hat{\beta}_0^{SE}$ و $\hat{\beta}_0^{SP}$

در این بخش، تابع مخاطره $\hat{\beta}_0$، $\hat{\beta}_0^{SE}$ و $\hat{\beta}_0^{SP}$ تحت تابع زیان نرمال بازتابیده محاسبه می‌شوند. برای سهولت در محاسبه تابع مخاطره از لم‌های زیر استفاده می‌کنیم.

لم 1. اگر $Z : N(0,1)$ و $S : \chi_k^2$ مستقل از هم باشند، آنگاه در صورت وجود امید ریاضی $f : \mathbb{R} \times (0,\infty) \to \mathbb{R}$ و برای هر $c < 0/5$ رابطه زیر را داریم

$$E[e^{cZ^2} f(Z,S)] = \frac{1}{\sqrt{1-2c}} E\left[f\left(\frac{Z}{\sqrt{1-2c}}, S\right)\right].$$

برهان: با استفاده از خاصیت امید ریاضی داریم

$$E[e^{cZ^2} f(Z,S)] = E[E[e^{cZ^2} f(Z,S) \mid S]]$$
$$= E[\int_{-\infty}^{+\infty} e^{cz^2} f(z,S) \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz]$$
$$= E[\int_{-\infty}^{+\infty} f(z,S) \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}(1-2c)} dz]$$
$$= E[\int_{-\infty}^{+\infty} f(z,S) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z\sqrt{1-2c})^2} dz].$$

داریم $J = \frac{1}{\sqrt{1-2c}}$ و ژاکوبی تبدیل $u = z\sqrt{1-2c}$ با تغییر متغیر

$$E[e^{cZ^2} f(Z,S)] = \frac{1}{\sqrt{1-2c}} E[\int_{-\infty}^{+\infty} f\left(\frac{u}{\sqrt{1-2c}}, S\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du]$$
$$= \frac{1}{\sqrt{1-2c}} E\left[f\left(\frac{Z}{\sqrt{1-2c}}, S\right)\right]. \quad \square$$

لم 2. [3]. اگر $Z : N(0,1)$ و $S : c_k^2$ مستقل از هم باشند، آنگاه در صورت وجود امید ریاضی $f : Â´ (0,¥) ® Â$ و برای هر $cÎ Â$ رابطه زیر را داریم

$$E[e^{cZ} f(Z,S)] = e^{\frac{c^2}{2}} E[f(Z+c,S)].$$

برای محاسبه تابع مخاطره $\hat{b}_0$ تحت تابع زیان (1) با توجه به هم توزیع بودن $(\hat{b}_0 - b_0)^2$ و $Z^2 s^{-2} H$ و با استفاده از لم 1 داریم

$$\begin{aligned} R(\hat{b}_0, b_0) &= E[L(\hat{b}_0, b_0)] = 1 - E[e^{-\frac{1}{2g^2}(\hat{b}_0 - b_0)^2}] \\ &= 1 - E[e^{-\frac{1}{2g^2} Z^2 s^{-2} H}] \\ &= 1 - E[e^{-aZ^2}] = 1 - \frac{1}{\sqrt{1+2a}} = 1 - (1+2a)^{-\frac{1}{2}}, \end{aligned}$$

که در آن $a = \frac{s^{-2} H}{2g^2}$. تابع مخاطره $\hat{b}_0^{SP}$ تحت تابع زیان نرمال بازتابیده برابر است با

$$\begin{aligned} R(\hat{b}_0^{SP}, b_0) &= E[1 - e^{-\frac{1}{2g^2}(\hat{b}_0^{SP} - b_0)^2}] \\ &= 1 - E[e^{-\frac{1}{2g^2}(d\hat{b}_0 + (1-d) b_{00} - b_0)^2} I(F_{1,n} < F_{1-a})] \\ &\quad - E[e^{-\frac{1}{2g^2}(\hat{b}_0 - b_0)^2} I(F_{1,n} > F_{1-a})]. \end{aligned} \qquad (3)$$

آماره آزمون $F_{1,n}$ با توجه به رابطه (2) به صورت

$$\begin{aligned} F_{1,n} &= \frac{[H^{-\frac{1}{2}}(\hat{b}_0 - b_{00}) s^{-1}]^2}{S/n} \\ &= \frac{[H^{-\frac{1}{2}}(\hat{b}_0 - b_0) s^{-1} + H^{-\frac{1}{2}}(b_0 - b_{00}) s^{-1}]^2}{S/n} \\ &= \frac{(Z+D)^2}{S/n} \end{aligned}$$

است که در آن $D = H^{-\frac{1}{2}}(b_0 - b_{00}) s^{-1}$ و $S = \frac{n S_n^2}{s^2}$ است. بنابراین اولین امید ریاضی در رابطه (3) به صورت

$$\begin{aligned} E[e^{-\frac{1}{2g^2}(d\hat{b}_0 + (1-d) b_{00} - b_0)^2} I(F_{1,n} < F_{1-a})] &= E[e^{-\frac{1}{2g^2}(ds\sqrt{H} Z - (1-d) Ds\sqrt{H})^2} I(F_{1,n} < F_{1-a})] \\ &= e^{-(1-d)^2 a D^2} E[e^{c_1 Z^2} f(Z,S)] \end{aligned}$$

محاسبه می‌شود، که در آن $c_1 = -ad^2$ و $f(Z,S) = e^{2ad(1-d)DZ} I(\frac{(Z+D)^2}{S/n} < F_{1-a})$. با استفاده از لم 1 داریم

$$e^{-(1-d)^2aD^2}E[e^{c_1Z^2}f(Z,S)]=\frac{e^{-(1-d)^2aD^2}}{\sqrt{1+2ad^2}}E[e^{\frac{2ad(1-d)D}{\sqrt{1+2ad^2}}Z}I(\frac{(\frac{Z}{\sqrt{1+2ad^2}}+D)^2}{S/n}<F_{1-a})]$$

$$=\frac{e^{-(1-d)^2aD^2}}{\sqrt{1+2ad^2}}E[f^{*}(Z,S)e^{c_2Z}],$$

که در آن $c_2=\frac{2ad(1-d)D}{\sqrt{1+2ad^2}}$ و $f^{*}(Z,S)=I(\frac{(Z+\sqrt{1+2ad^2}D)^2}{S/n}<(1+2ad^2)F_{1-a})$. در نتیجه با استفاده از لم 2 داریم

$$E[e^{-\frac{1}{2g^2}(\hat{b}_0^{SE}-b_0)^2}I(F_{1,n}<F_{1-a})]=\frac{e^{-(1-d)^2aD^2}}{\sqrt{1+2ad^2}}e^{\frac{c_2^2}{2}}E[I(\frac{(Z+c_2+\sqrt{1+2ad^2}D)^2}{S/n}<(1+2ad^2)F_{1-a})]$$

$$=\frac{e^{-(1-d)^2aD^2+\frac{c_2^2}{2}}}{\sqrt{1+2ad^2}}P(\frac{(Z+c_2+\sqrt{1+2ad^2}D)^2}{S/n}<(1+2ad^2)F_{1-a})$$

$$=\frac{e^{-(1-d)^2aD^2+\frac{c_2^2}{2}}}{\sqrt{1+2ad^2}}G_{1,n}[(1+2ad^2)F_{1-a};(c_2+\sqrt{1+2ad^2}D)^2], \qquad (4)$$

که در آن $G_{i,j}(q;q)$ تابع توزیع تجمعی توزیع F نامرکزی با (i,j) درجه آزادی و پارامتر غیرمرکزی q است. دومین امید ریاضی در رابطه (3) برابر است با

$$E[e^{-\frac{1}{2g^2}(\hat{b}_0^{PT}-b_0)^2}I(F_{1,n}>F_{1-a})]=E[e^{-\frac{s^2H}{2g^2}Z^2}I(F_{1,n}>F_{1-a})]$$

$$=E[e^{-aZ^2}I(\frac{(Z+D)^2}{S/n}>F_{1-a})]$$

$$=\frac{1}{\sqrt{1+2a}}E[I(\frac{(\frac{Z}{\sqrt{1+2a}}+D)^2}{S/n}>F_{1-a})]$$

$$=\frac{1}{\sqrt{1+2a}}P(I(\frac{(Z+\sqrt{1+2a}D)^2}{S/n}>(1+2a)F_{1-a}))$$

$$=\frac{1}{\sqrt{1+2a}}(1-G_{1,n}((1+2a)F_{1-a};(1+2a)D^2)). \qquad (5)$$

با جای‌گذاری روابط (4) و (5) در (3) نتیجه می‌شود.

$$R(\hat{b}_0^{SP},b_0)=1-\frac{e^{-(1-d)^2aD^2+\frac{c_2^2}{2}}}{\sqrt{1+2ad^2}}G_{1,n}((1+2ad^2)F_{1-a};(c_2+\sqrt{1+2ad^2}D)^2)$$

$$-\frac{1}{\sqrt{1+2a}}(1-G_{1,n}((1+2a)F_{1-a};(1+2a)D^2)). \qquad (6)$$

تابع مخاطره $\hat{b}_0^{SE}$ با استفاده از لم 1 برابر است با

$$
\begin{aligned}
R(\hat{b}_0^{SE}, b_0) &= E[1 - e^{-\frac{1}{2g^2}(\hat{b}_0^{SE} - b_0)^2}] \\
&= 1 - E[e^{-\frac{1}{2g^2}(d\hat{b}_0^{SP} + (1-d)b_{00} - b_0)^2}] \\
&= 1 - E[e^{-\frac{1}{2g^2}(d(\hat{b}_0^{SP} - b_0) - (1-d)(b_0 - b_{00}))^2}] \\
&= 1 - E[e^{-\frac{1}{2g^2}(ds\sqrt{H}Z - (1-d)Ds\sqrt{H})^2}] \\
&= 1 - e^{-(1-d)^2aD^2} E[e^{-ad^2Z^2} e^{(2ad(1-d)D)Z}] \\
&= 1 - \frac{1}{\sqrt{1+2ad^2}} e^{-(1-d)^2aD^2} E[e^{\frac{2ad(1-d)D}{\sqrt{1+2ad^2}}Z}] \\
&= 1 - \frac{1}{\sqrt{1+2ad^2}} e^{-(1-d)^2aD^2} M_Z\left(\frac{2ad(1-d)D}{\sqrt{1+2ad^2}}\right) \\
&= 1 - \frac{1}{\sqrt{1+2ad^2}} e^{-(1-d)^2aD^2\left(1-\frac{2ad^2}{1+2ad^2}\right)} \qquad (7)
\end{aligned}
$$

که در آن $M_Z(\lambda)$ تابع مولد گشتاور متغیر تصادفی Z است.

4 مقایسه $\hat{b}_0^{SP}$ نسبت به $\hat{b}_0^{SE}$ و $\hat{b}_0^{SP}$

در این بخش ابتدا کارایی نسبی $\hat{b}_0^{SP}$ نسبت به $\hat{b}_0^{SE}$ و $\hat{b}_0^{SE}$ محاسبه شده و سپس به صورت گرافیکی مورد بررسی بیشتر قرار می‌گیرند.

1.4 کارایی نسبی $\hat{b}_0^{SP}$ نسبت به $\hat{b}_0^{SE}$

برای هر مقدار D مخالف صفر از رابطه (6)، تابع مخاطره براوردگر $\hat{b}_0^{SP}$ را می‌توان به صورت

$$R(\hat{b}_0^{SP}, b_0) = R(\hat{b}_0^{SE}, b_0) + g(D),$$

نوشت که در آن $g(D)$ عبارت است از

$$
\begin{aligned}
g(D) &= \frac{1}{\sqrt{1+2a}} G_{1,n}((1+2a)F_{1-a}; (1+2a)D^2) \\
&\quad - \frac{e^{-(1-d)^2aD^2 + \frac{c_2^2}{2}}}{\sqrt{1+2ad^2}} G_{1,n}((1+2ad^2)F_{1-a}; (c_2 + \sqrt{1+2ad^2}D)^2).
\end{aligned}
$$

بنابراین کارایی نسبی براوردگر $\hat{b}_0^{SP}$ نسبت به براوردگر $\hat{b}_0^{SE}$ به صورت

$$Eff[\hat{b}_0^{SP}, \hat{b}_0^{SE}] = \frac{R(\hat{b}_0^{SE}, b_0)}{R(\hat{b}_0^{SP}, b_0)} = \frac{1 - (1+2a)^{-\frac{1}{2}}}{1 - (1+2a)^{-\frac{1}{2}} + g(D)} \qquad (8)$$

است. تحت فرض $H_0 : b_0 = b_{00}$ داریم $D = 0$ و در نتیجه

$$g(0) = \frac{1}{\sqrt{1+2a}} G_{1,n}((1+2a)F_{1-a}; 0) - \frac{1}{\sqrt{1+2ad^2}} G_{1,n}((1+2ad^2)F_{1-a}; 0) \leq 0.$$

بنابراین به ازای $D=0$ و $0 \le d \le 1$ براوردگر $\hat{b}_0^{SP}$ نسبت به براوردگر $\tilde{b}_0$ کارایی نسبی معادل یا بیشتری دارد. اگر $|D| \to \infty$ آنگاه $g(D)$ همگرا به صفر و در نتیجه کارایی نسبی همگرا به یک می‌باشد ($Eff[\hat{b}_0^{SP}, \tilde{b}_0] \to 1$). بنابراین وقتی $|D|$ به اندازه کافی بزرگ باشد کارایی نسبی تقریبا یک است.

2.4 مقایسه $\hat{b}_0^{SP}$ نسبت به $\hat{b}_0^{SE}$

کارایی نسبی $\hat{b}_0^{SP}$ نسبت به $\hat{b}_0^{SE}$ با استفاده از روابط (7) و (6) به صورت

$$Eff[\hat{b}_0^{SP}, \hat{b}_0^{SE}] = \frac{R(\hat{b}_0^{SE}, b_0)}{R(\hat{b}_0^{SP}, b_0)} = \frac{1-(1+2ad^2)^{-\frac{1}{2}} e^{-(1-d)^2 aD^2 (1-\frac{2ad^2}{1+2ad^2})}}{1-(1+2a)^{-\frac{1}{2}} + g(D)} \qquad (9)$$

است. در همسایگی فرض صفر ($|D| \to 0$)، با توجه به اینکه $1+2ad^2 \le 1+2a$ و بنابر غیرنزولی بودن تابع توزیع $G_{1,n}((1+2ad^2)F_{1-a};0) \le G_{1,n}((1+2a)F_{1-a};0)$ داریم

$$Eff[\hat{b}_0^{SP}, \hat{b}_0^{SE}; D=0] = \frac{1-(1+2ad^2)^{-\frac{1}{2}}}{1-(1+2a)^{-\frac{1}{2}} + g(0)} \le 1,$$

اگر $|D|$ از یک مقداری بزرگتر شود ($|D| \to \infty$)، فرض صفر به ازای هر $0 < a < 1$ رد می‌شود و در نتیجه $\hat{b}_0^{SP} = \tilde{b}_0$ و همینطور $R(\hat{b}_0^{SE}, b_0) \to 1$ آنگاه داریم

$$\lim_{|D| \to \infty} Eff[\hat{b}_0^{SP}, \hat{b}_0^{SE}] = \frac{1}{1-(1+2a)^{-\frac{1}{2}}} > 1.$$

بنابراین در این حالت کارایی نسبی براوردگر $\tilde{b}_0$ ($\hat{b}_0^{SP} = \tilde{b}_0$) نسبت به $\hat{b}_0^{SE}$ بیشتر است.

3.4 مقایسه براوردگرها به صورت گرافیکی

در این بخش کارایی نسبی براوردگر $\hat{b}_0^{SP}$ نسبت به براوردگر $\tilde{b}_0$ و $\hat{b}_0^{SE}$ به صورت گرافیکی در نرم افزار R ارزیابی می‌شود.

در شکل 2 کارایی نسبی براوردگر $\hat{b}_0^{SP}$ نسبت به براوردگر $\tilde{b}_0$ با استفاده از رابطه (8) به ازای مقادیر $n=25$، $d=0/2,0/4,0/6,0/8$، $g=2$، $a=0/01,0/05,0/1,0/2$ و $s^2H=\frac{1}{n}$ نسبت به D رسم شده است. از شکل 2 ملاحظه می‌شود که براوردگر $\hat{b}_0^{SP}$ در همسایگی فرض صفر ($|D| \to 0$) به دلیل نزدیک بودن مقدار اولیه b_{00} به مقدار واقعی b_0، مقدار مخاطره براوردگر $\hat{b}_0^{SP}$ نسبت به $\tilde{b}_0$ کمتر است بنابراین در این حالت براوردگر $\hat{b}_0^{SP}$ دارای کارایی نسبی بیشتری

نسبت به $\hat{b}_0^{\%}$ است. علاوه بر این مشاهده می‌شود براوردگرهای $\hat{b}_0^{SP}$ با سطح معنی‌داری کوچکتر $(a=0/01)$ دارای کارایی نسبی بیشتری نسبت به سایر حالات متناظر هستند زیرا در این حالت فرض صفر با اطمینان بیشتری پذیرفته می‌شود که اعتماد بیشتر آزمایشگر نسبت به مقدار اولیه b_{00} را نشان می‌دهد. همچنین با کوچک شدن ضریب انقباضی (d) یا به عبارتی بزرگ شدن ضریب b_{00}، کارایی نسبی براوردگر $\hat{b}_0^{SP}$ افزایش می‌یابد. در شکل 3 کارایی نسبی براوردگر $\hat{b}_0^{SP}$ نسبت به d به ازای $D=0$ برای مقادیر مشخص $n=10,20,30,40$، $a=0/01,0/05,0/1,0/2$ و $s^2H=\frac{1}{n}$ رسم شده است. از شکل 3 ملاحظه می‌شود که کارایی نسبی تحت فرض صفر $(D=0)$ نسبت به d، تابعی نزولی است، زیرا با بزرگ شدن مقدار d ضریب اعتماد براوردگر $\hat{b}_0^{\%}$ افزایش می‌یابد، به طوری که وقتی d به یک میل کند $(d \to 1)$، آنگاه $\hat{b}_0^{SP}$ به $\hat{b}_0^{\%}$ نزدیک می‌شود، در این حالت کارایی نسبی براوردگر $\hat{b}_0^{SP}$ نسبت به $\hat{b}_0^{\%}$ به یک میل می‌کند. با افزایش حجم نمونه (n) کارایی نسبی براوردگر $\hat{b}_0^{SP}$ نسبت به $\hat{b}_0^{\%}$ برای $n=30,40$ به دلیل اعتماد بیشتر به براوردگر نمونه‌ای $\hat{b}_0^{\%}$، در مقایسه با $n=10,20$ کاهش می‌یابد. همچنین برای $a=0/01$ کارایی نسبی $\hat{b}_0^{SP}$ نسبت به $\hat{b}_0^{\%}$، بیشتر از سایر براوردگرها است.

در شکل 4 کارایی نسبی $\hat{b}_0^{SP}$ نسبت به $\hat{b}_0^{SE}$ با استفاده از رابطه (9) به ازای مقادیر $n=20$، $d=0/4$، $g=2$، $a=0,0/01,0/05,0/1,0/2$ و $s^2H=\frac{1}{n}$ نسبت به D رسم شده است. همانطور که در شکل 4 ملاحظه می‌شود در همسایگی فرض صفر کارایی نسبی $\hat{b}_0^{SP}$ کمتر از یک است در نتیجه $\hat{b}_0^{SE}$ کاراتر از $\hat{b}_0^{SP}$ می‌باشد. اما وقتی $|D|$ از یک مقداری بزرگتر شود که نشان دهنده‌ی غلط بودن اطلاع پیشین b_{00} است، فرض صفر به ازای هر $0<a<1$ رد می‌شود بنابراین $\hat{b}_0^{SP}=\hat{b}_0^{\%}$ و کارایی نسبی $\hat{b}_0^{SP}$ نسبت به $\hat{b}_0^{SE}$ طبق شکل 4 بزرگتر از یک و نسبت به $|D|$ صعودی است. با توجه به شکل برای $a=0$ به دلیل برابر بودن براوردگرهای $\hat{b}_0^{SP}$ و $\hat{b}_0^{SE}$، کارایی نسبی $\hat{b}_0^{SP}$ نسبت به $\hat{b}_0^{SE}$ روی خط یک منطبق می‌باشد.

در جدول 1 بازه‌های موثر یعنی دامنه‌ای از مقادیر D در $g=2$ که به ازای آن براوردگر $\hat{b}_0^{SP}$ نسبت به براوردگر $\hat{b}_0^{\%}$ دارای کارایی نسبی بیشتر از یک است، خلاصه شده است. با توجه به نتایج به دست آمده در جدول 1، با افزایش a، طول بازه موثر کاهش می‌یابد. همچنین با افزایش n، دامنه‌ای از مقادیر D که کارایی نسبی براوردگر $\hat{b}_0^{SP}$ نسبت به $\hat{b}_0^{\%}$ بزرگتر از یک است، کاهش می‌یابد. از طرفی با افزایش ضریب انقباضی d، طول بازه موثر افزایش می‌یابد.

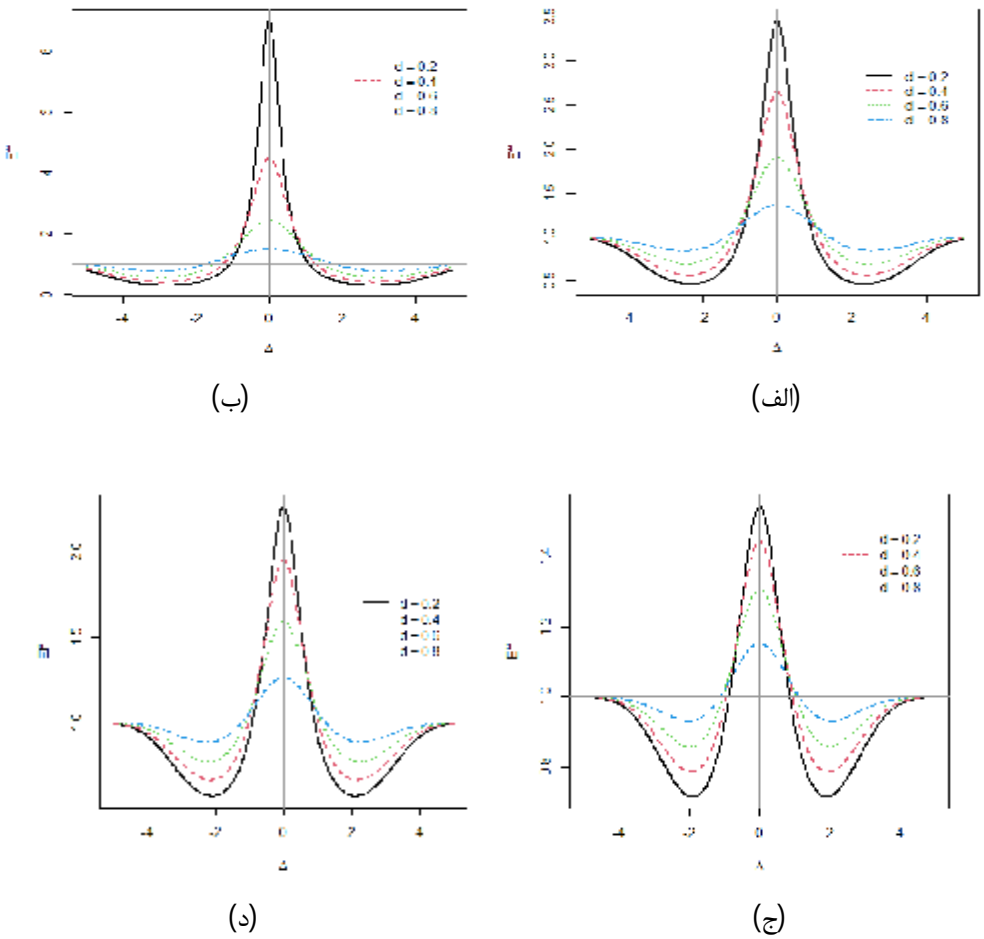
4.4 تعیین سطح معنی‌داری بهینه برای براوردگر $\hat{b}_0^{SP}$

تابع مخاطره براوردگر $\hat{b}_0^{SP}$ به سطح معنی‌داری پیش‌آزمون (a)، ضریب انقباضی (d)، پارامتر D و پارامتر شکل تابع زیان نرمال بازتابیده (g) بستگی دارد. بدین منظور، کارایی نسبی براوردگر $\hat{b}_0^{SP}$ نسبت به براوردگر $\hat{b}_0^{\%}$ را به‌عنوان تابعی از a و D در نظر می‌گیریم. بنابراین داریم

$$Eff\,[\hat{b}_0^{SP}, b_0^{*};a, D] = \frac{1-(1+2a)^{-\frac{1}{2}}}{1-(1+2a)^{-\frac{1}{2}} + g(D)}. \qquad (10)$$

تابع $Eff[\hat{b}_0^{SP}, b_0^{*};a, D]$ در $D=0$ به ازای تمام $0<a<1$، بزرگتر یا مساوی با یک می‌باشد. همچنین اگر $|D|$ از مقدار صفر دور شود $Eff[\hat{b}_0^{SP}, b_0^{*};a, |D|]$ به طور یکنواخت کاهش می‌یابد و با عبور از خط یک به کمترین مقدار در نقطه $|D|=|D_0|$ می‌رسد. سپس وقتی $|D| \to \infty$، کارایی نسبی به طور یکنواخت از نقطه مینیمم به سمت یک افزایش می‌یابد. علاوه بر این برای $D=0$ و $a \in [0,1]$ داریم

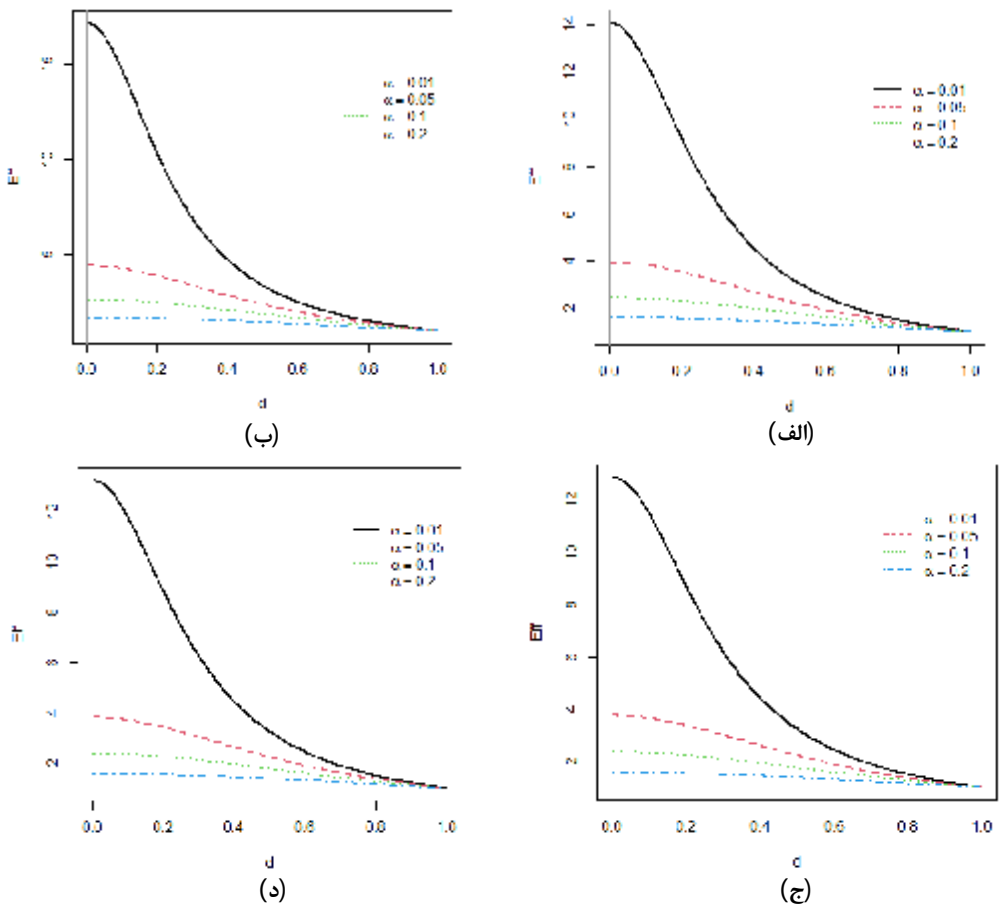
$$\max_a Eff[\hat{b}_0^{SP}, b_0^{*};a, D=0] = Eff[\hat{b}_0^{SP}, b_0^{*};a=0, D=0] = \frac{1-(1+2a)^{-\frac{1}{2}}}{1-(1+2ad^2)^{-\frac{1}{2}}} \geq 1.$$



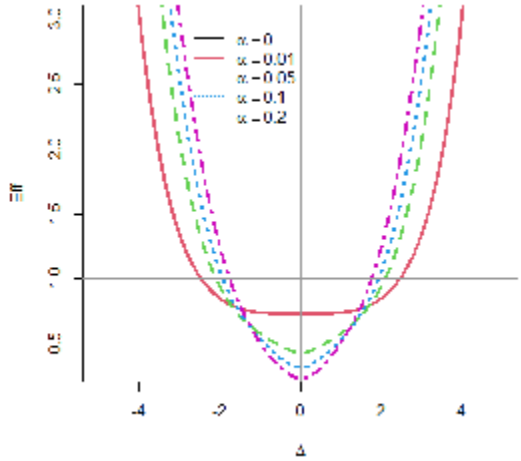
(ب) (الف)

(د) (ج)

شکل 2 : کارایی نسبی براوردگر $\hat{b}_0^{SP}$ نسبت به b_0^{*} : الف- $a=0/01$، ب- $a=0/05$، ج- $a=0/1$، د- $a=0/2$.



شکل 3 : کارایی نسبی براوردگر $\hat{b}_0^{SP}$ نسبت به $\hat{b}_0^{\%}$: الف- $n=10$، ب- $n=20$، ج- $n=30$، د- $n=40$.



شکل 4: کارایی نسبی براوردگر $\hat{b}_0^{SP}$ نسبت به $\hat{b}_0^{SE}$ به ازای مقادیر $n=20$، $d=0/4$، $g=2$ و $a=0,0/01,0/05,0/1,0/2$ نسبت به D.

جدول1: دامنه‌ای از مقادیر γ که به ازای آن براوردگر $\hat{\beta}_0^{SP}$ کاراتر از $\tilde{\beta}_0$ است.

$\alpha=0/2$	$\alpha=0/1$	$\alpha=0/05$	$\alpha=0/01$	γ	n
(-0/8831, 0/8831)	(-0/9502, 0/9502)	(-1/0144, 1/0144)	(-1/1294, 1/1294)	0/2	10
(-0/9782, 0/9782)	(-1/0708, 1/0708)	(-1/1636, 1/1636)	(-1/3453, 1/3453)	0/4	
(-1/0736, 1/0736)	(-1/1978, 1/1978)	(-1/3288, 1/3288)	(-1/6132, 1/6132)	0/6	
(-1/1704, 1/1704)	(-1/3335, 1/3335)	(-1/5152, 1/5152)	(-1/9590, 1/9590)	0/8	
(-0/8644, 0/8644)	(-0/9221, 0/9221)	(-0/9795, 0/9795)	(-1/0927, 1/0927)	0/2	20
(-0/9529, 0/9529)	(-1/0309, 1/0309)	(-1/1112, 1/1112)	(-1/2807, 1/2807)	0/4	
(-1/0404, 1/0404)	(-1/1425, 1/1425)	(-1/2516, 1/2516)	(-1/4991, 1/4991)	0/6	
(-1/1278, 1/1278)	(-1/2586, 1/2586)	(-1/4033, 1/4033)	(-1/7560, 1/7560)	0/8	
(-0/8591, 0/8591)	(-0/9140, 0/9140)	(-0/9691, 0/9691)	(-1/0803, 1/0803)	0/2	30
(-0/9459, 0/9459)	(-1/0195, 1/0195)	(-1/0959, 1/0959)	(-1/2599, 1/2599)	0/4	
(-1/0313, 1/0313)	(-1/1272, 1/1272)	(-1/2299, 1/2299)	(-1/4646, 1/4646)	0/6	
(-1/1162, 1/1162)	(-1/2382, 1/2382)	(-1/3729, 1/3729)	(-1/6998, 1/6998)	0/8	

$\mathrm{Eff}[\hat{\beta}_0^{SP}, \tilde{\beta}_0; \alpha, D=0]$ به عنوان تابعی از α، با افزایش α، کاهش می‌یابد. در حالت کلی برای دو مقدار α_1 و α_2، دو منحنی $\mathrm{Eff}[\hat{\beta}_0^{SP}, \tilde{\beta}_0; \alpha_1, |D|]$ و $\mathrm{Eff}[\hat{\beta}_0^{SP}, \tilde{\beta}_0; \alpha_2, |D|]$ در پایین خط یک، یک دیگر را قطع می‌کنند. با افزایش α_1 و α_2، مقدار $|D|$ در برخورد دو منحنی کاهش می‌یابد. در نتیجه برای انتخاب یک سطح معنی‌داری بهینه با ماکسیمم کارایی نسبی از روش ماکس‌مین استفاده می‌شود. اگر بدانیم $|D|$ طبق جدول 1 متعلق به بازه موثر است، براوردگر $\hat{\beta}_0^{SP}$ انتخاب می‌شود زیرا در این بازه کارایی نسبی بزرگتر از یک می‌باشد. همانطور که در تحلیل کارایی نسبی براوردگر $\hat{\beta}_0^{SP}$ نسبت به $\tilde{\beta}_0$ بررسی شد، براوردگر $\hat{\beta}_0^{SP}$ به طور یکنواخت عملکرد بهتری نسبت به براوردگر $\tilde{\beta}_0$ ندارد. همچنین، مقدار D برای محقق نامشخص است. بنابراین یک مقدار برای کمترین کارایی نسبی مانند Eff_0 درنظر می‌گیریم. مجموعه

$$A_\alpha = \{\alpha \mid \mathrm{Eff}[\hat{\beta}_0^{SP}, \tilde{\beta}_0; \alpha, |D|] \geq \mathrm{Eff}_0\},$$

را درنظر بگیرید. برآوردگر $\hat{\beta}_0^{SP}$ که $\mathrm{Eff}[\hat{\beta}_0^{SP}, \tilde{\beta}_0; \alpha, |D|]$ را به ازای هر $\alpha \in A_\alpha$ و $|D|$ ماکزیمم کند انتخاب می‌شود. بنابراین معادله زیر را حل می‌کنیم

$$\max_{\alpha} \min_{|D|} \mathrm{Eff}[\hat{\beta}_0^{SP}, \tilde{\beta}_0; \alpha, |D|] = \mathrm{Eff}_0. \qquad (10)$$

بیشترین و کمترین مقدار کارایی نسبی براوردگر $\hat{b}_0^{SP}$ نسبت به براوردگر $\tilde{b}_0$ با حل رابطه (10) نسبت به $|D|$ به دست می‌آید. بیشترین میزان کارایی نسبی (Eff^*) و کمترین میزان کارایی نسبی (Eff_0) و مقدار $|D|$ که برای آن کارایی نسبی، مینیمم می‌شود $|D_0|$ به ازای مقادیر انتخابی a و n، $d=0/5$ و $g=2$ در جدول 2 خلاصه شده است. برای مثال اگر $d=0/5$ و $n=20$ باشد آنگاه برای رسیدن به کارایی نسبی به اندازه حداقل 0/8174 مقدار بهینه سطح معنی‌داری برابر $a=0/2$ به ازای $|D_0|=1/9785$ خواهد بود. از جدول 2 ملاحظه می‌شود که Eff^* یک تابع نزولی نسبت به a است در حالی‌که Eff_0 یک تابع صعودی است.

5.4 تعیین ضریب انقباضی مناسب

یک مسئله مهم برای براوردگر $\hat{b}_0^{SP}$ انتخاب مناسب مقدار ضریب انقباضی (d) است. در این بخش دو روش برای انتخاب ضریب انقباضی به صورت زیر ارائه می‌شود.

1. مقدار (d) از مینیمم کردن تابع مخاطره براوردگر $\hat{b}_0^{SP}$ در رابطه (6) نسبت به d از روش‌های عددی حاصل می‌شود. به طور کلی مقدار D مجهول است بنابراین در این روش براورد آن به صورت $\hat{D}=H^{-\frac{1}{2}}(\tilde{b}_0-b_{00})\hat{s}^{-1}$ که در آن $\hat{s}=S_n$، در رابطه (6) جایگزین می‌شود. مقدار ضریب انقباضی حاصل شده از این روش را با d_1 و براوردگر $\hat{b}_0^{SP}$ متناظر با آن را با $\hat{b}_{01}^{SP}$ نشان می‌دهیم.
2. برای آزمون فرض $H_0: b_0=b_{00}$ در مقابل $H_1: b_0 \neq b_{00}$ فرض صفر را رد می‌کنیم اگر و تنها اگر $F_{1,n} > F_{1-a}$. مقدار معنی‌داری آزمون به صورت مقدار $P = P(F_{1,n} > f \mid H_0) = 1 - G_{1,n}(f)$ – به دست می‌آید که f مقدار مشاهده شده آماره آزمون $F_{1,n}$ و $G_{1,n}(f)$ تابع توزیع تجمعی توزیع فیشر با 1 و n درجه آزادی است. P مقدار[1] این آزمون نشان می‌دهد که چقدر فرض صفر توسط داده‌ها پشتیبانی می‌شود. مقادیر بزرگ P -مقدار در تایید نزدیک بودن b_{00} به مقدار واقعی b_0 است. بنابراین در این روش که توسط [12] ارائه شد ضریب انقباضی به صورت

$$d = 1 - (\text{مقدار} - P)$$

درنظر گرفته می‌شود. ضریب انقباضی حاصل شده از این روش با d_2 و براوردگر $\hat{b}_0^{SP}$ متناظر با آن با $\hat{b}_{02}^{SP}$ نشان داده می‌شود. P -مقدار با افزایش حجم نمونه به صفر میل می‌کند، به عبارت دیگر d_2 به 1 نزدیک می‌شود.

همچنین می‌توان ریشه دوم P -مقدار که توسط [4] ارائه شد را برای پشتیبانی قوی‌تر از b_{00} پیشنهاد کرد. بنابراین ضریب انقباضی به صورت $d = 1 - \sqrt{\text{مقدار} - P}$ در نظر گرفته می‌شود. ضریب انقباضی تعیین شده از این روش با d_3 و براوردگر $\hat{b}_0^{SP}$ متناظر با آن با $\hat{b}_{03}^{SP}$ نشان داده می‌شود.

[5] P-value

جدول 2: ماکسیمم و مینیمم کارایی نسبی براوردگر $\hat{b}_0^{SP}$ نسبت به $\hat{b}_0^{\%}$ برای $g=2$ و $d=0/5$

a		n				
		۱۰	۲۰	۳۰	۴۰	۵۰
	Eff^*	2/3693	2/2670	2/2369	2/2226	2/2142
0/05	Eff_0	0/5859	0/62036	0/6298	0/6342	0/6368
	$\lvert D_0 \rvert$	2/6601	2/4422	2/3856	2/3595	2/3445
۰/۱۰	Eff^*	1/8579	1/7867	1/7664	1/756774	1/7511
	Eff_0	0/6853	0/7105	0/7175	0/7208	0/7227
	$\lvert D_0 \rvert$	2/3446	2/2032	2/1657	2/1484	2/1385
0/15	Eff^*	1/5899	1/5394	1/5252	1/5184	1/5145
	Eff_0	0/7513	0/7711	0/7767	0/7793	0/7808
	$\lvert D_0 \rvert$	2/1722	2/0690	2/0415	2/0287	2/0213
0/20	Eff^*	1/4238	1/3871	1/3768	1/3720	1/3691
	Eff_0	0/8014	0/8174	0/8218	0/8239	0/8252
	$\lvert D_0 \rvert$	2/0574	1/9785	1/9573	1/9474	1/9417
۰/۳۰	Eff^*	1/2313	1/2110	1/2054	1/2027	1/2012
	Eff_0	0/8743	0/8847	0/8877	0/8890	0/8898
	$\lvert D_0 \rvert$	1/9098	1/8608	1/8475	1/8413	1/8377

5 مثال عددی

در این بخش یک مثال عددی برای تشریح براوردگرهای پیشنهادی ارائه می‌شود. داده‌ها مربوط به سیستم جلو برنده یک موتور راکت از [6] است که وابستگی مقاومت برش به عنوان متغیر پاسخ به زمان هفتگی کارکرد سیستم به عنوان متغیر

پیشگو بررسی شد. 20 مشاهده روی مقاومت برشی و زمان هفتگی کارکرد سیستم جمع‌آوری شد و در جدول 3 نشان داده شده است. نتایج بررسی شده پیش فرض‌های مناسبت مدل را تایید می‌کند. براوردهای ماکسیمم درستنمایی برای پارامترهای شیب و عرض از مبدا به ترتیب $\hat{b}_1^{\%} = -37/15$ و $\hat{b}_0^{\%} = 2627/82$ به دست آمد.

براوردهای $\hat{b}_{01}^{SP}$، $\hat{b}_{02}^{SP}$ و $\hat{b}_{03}^{SP}$ به ترتیب به ازای d_1، d_2 و d_3 و $\hat{b}_0^{SE}$ به ازای d که طبق روش اول تعیین ضریب انقباضی از مینیمم کردن تابع مخاطره $\hat{b}_0^{SE}$ حاصل می‌شود، برای چهار مقدار اولیه b_{00} در $g = 2$ محاسبه شده‌اند.

برای مثال در سطر اول از جدول 4 براورد b_0 را وقتی که مقدار حدس زده شده $b_{00} = 2200$ است درنظر می‌گیریم، برای براورد $\hat{b}_0^{SE}$ مقدار $d = 0/98$ از مینیمم کردن رابطه (7) در $\hat{D} = 9/68$ به دست آمد. بنابراین براورد انقباضی $\hat{b}_0^{SE}$ به صورت

$$\hat{b}_0^{SE} = 0/98(2627/82) + (1-0/98)(2200) = 2623/351,$$

است. در حالتی که مقدار حدس زده شده $b_{00} = 2600$ است برای براورد $\hat{b}_0^{SE}$ مقدار $d = 0/461$ از مینیمم کردن رابطه (7) در $\hat{D} = 0/62$ به دست آمد. بنابراین براورد انقباضی $\hat{b}_0^{SE}$ در مقدار اولیه $b_{00} = 2600$ به صورت

$$\hat{b}_0^{SE} = 0/461(2627/82) + (1-0/461)(2200) = 2612/826,$$

است. برای آزمون فرض $H_0 : b_0 = 2200$ در مقابل $H_1 : b_0 \neq 2200$ در سطح معنی‌داری $a = 0/05$ مقدار آماره آزمون به صورت

$$f = \frac{(\hat{b}_0^{\%} - b_{00})^2}{S_n^2 H} = 93/75,$$

جدول 3: داده های مثال عددی

مشاهده i	مقاومت برشی y_i	سن عامل (هفته) x_i	مشاهده i	مقاومت برشی y_i	سن عامل x_i
1	2158/70	15/50	11	2156/20	13/00
2	1678/15	23/75	12	2399/55	3/75
3	2316/00	8/00	13	1779/82	25/00
4	2061/30	17/00	14	2336/75	9/75
5	2207/50	5/5	15	1765/30	22/00
6	1708/30	19/00	16	2053/50	18/00
7	1784/70	24/00	17	2414/40	6/00
8	2575/00	2/50	18	2200/50	12/50
9	2357/90	7/50	19	2654/20	2/00
10	2265/70	11/00	20	1753/70	21/50

جدول 4: براورد پارامترهای موردنظر در داده های مثال

$\hat{b}_{03}^{SP}$	$\hat{b}_{02}^{SP}$	$\hat{b}_{01}^{SP}$	$\hat{b}_{0}^{SE}$	$\hat{b}_{0}^{\%}$	b_{00}
2627/82	2627/82	2627/82	2623/351	2627/82	2200
2607/436	2612/885	2613/468	2612/826	2627/82	2600
2652/793	2636/46	2627/824	2644/071	2627/82	2700
2627/82	2627/82	2627/82	2637/875	2627/82	2800

با توجه به این که مقدار آماره آزمون $f = 93/75$ از $F_{0/95} = 4/413$ بیشتر است در نتیجه فرض صفر رد می‌شود، بنابراین براورد $\hat{b}_0^{SP}$ برابر $2627/82$ است. در حالتی که مقدار حدس زده شده $b_{00} = 2600$ است از آنجایی که $f = 0/396$ از $F_{0/95} = 4/413$ کوچکتر است، لذا فرض صفر پذیرفته می‌شود. مقدارهای $d_1 = 0/4841$ از مینیمم کردن تابع مخاطره در $\hat{D} = 0/62$ و مقدار $d_2 = 0/4631$ و $d_3 = 0/2673$ از روش P -مقدار محاسبه می‌شود. بنابراین براوردهای $\hat{b}_0^{SP}$ به صورت

$$\hat{b}_{01}^{SP} = 0/4841(2627/82) + (1- 0/4841)(2600) = 2613/468,$$
$$\hat{b}_{02}^{SP} = 0/4631(2627/82) + (1- 0/4631)(2600) = 2612/885,$$
$$\hat{b}_{03}^{SP} = 0/2673(2627/82) + (1- 0/2673)(2600) = 2607/436,$$

به دست می‌آیند. سایر براوردهای ارائه شده در جدول 4 به طور مشابه حاصل می‌شوند. همچنین در این مثال طبق روش ماکس‌مین اگر $n = 20$ و $d = 0/48$ درنظر گرفته شود آنگاه برای رسیدن به کارایی نسبی $\hat{b}_0^{SP}$ نسبت به $\hat{b}_0^{\%}$ حداقل 0/6998، مقدار بهینه سطح معنی‌داری آزمون $a = 0/1$ می‌باشد. در نتیجه فرض $H_0 : b_0 = 2600$ در مقابل $H_1 : b_0 \neq 2600$ در سطح معنی‌داری $a = 0/1$ آزمون می‌کنیم. از آنجایی که

$$F_{1,n} = 0/396 < F_{0/90} = 3/006,$$

فرض صفر رد نمی‌شود و براورد $\hat{b}_0^{SP}$ برابر است با

$$\hat{b}_0^{SP} = 0/48(2627/82) + (1- 0/48)(2600) = 2613/35.$$

6 بحث و نتیجه‌گیری

در این مقاله کارایی نسبی براوردگرهای انقباضی و پیش‌آزمون انقباضی برای پارامتر عرض از مبدأ مدل رگرسیونی خطی ساده تحت تابع زیان نرمال بازتابیده مورد بررسی قرار گرفت. ابتدا تابع زیان، معرفی و سپس مخاطره براوردگرهای ماکسیمم درستنمایی، پیش‌آزمون انقباضی و انقباضی تحت تابع زیان نرمال بازتابیده به دست آمد. با مقایسه براوردگرهای ماکسیمم

درستنمایی و پیش‌آزمون انقباضی می‌توان نتیجه گرفت که هرچه مقدار اطلاع غیرنمونه‌ای (b_{00}) به پارامتر واقعی نزدیک‌تر باشد، کارایی نسبی براوردگر پیش‌آزمون انقباضی نسبت به براوردگر ماکسیمم درستنمایی بیشتر می‌شود. همچنین براوردگر پیش‌آزمون انقباضی به طور یکنواخت عملکرد بهتری نسبت به براوردگر انقباضی ندارد به طوری که در همسایگی $D=0$ کارایی نسبی براوردگر پیش‌آزمون انقباضی نسبت به براوردگر انقباضی کمتر از یک است. اما وقتی $|D|$ از یک مقداری بزرگتر شود براوردگر پیش‌آزمون انقباضی کارایی بیشتری نسبت به براوردگر انقباضی دارد. بازه‌های موثر یا به‌عبارتی دامنه‌ای از مقادیر D که به ازای آن براوردگر پیش‌آزمون انقباضی نسبت به براوردگر ماکسیمم درستنمایی دارای مخاطره کمتری است ارائه شد. با افزایش سطح معنی‌داری آزمون و حجم نمونه، طول بازه موثر کاهش می‌یابد. همچنین مقدار اندازه آزمون بهینه با استفاده از روش ماکس‌مین به دست آمد. در یک مثال واقعی بر اساس چند مقدار انتخابی برای ضریب انقباضی، براوردهای انقباضی و پیش‌آزمون انقباضی محاسبه شدند.

تقدیر و تشکر

نویسندگان، از سردبیر و داوران محترم مقاله برای ارزیابی و بهبود کیفیت آن کمال تشکر را دارند.

References

1. کیاپور، آ.، برآوردیابی انقباضی کلاسیک و بیزی در توزیع رایلی با استفاده از حدس نقطه ای براساس داده های سانسورشده، مجله پژوهش‌های ریاضی، جلد 4، شماره 1 (1397)، 63-74.

2. Z. Hoque, S. Hossain, Improved Estimation in Regression with Varying Penalty, Journal of Statistical Theory and Practice, **6** (2012) 260-273.

3. Z. Hoque, S. Khan, J. Wesolowski, Performance of Preliminary Test Estimator Under Linex Loss Function, Communications in Statistics-Theory and Methods, **38** (2009) 252-261.

4. M. Jabbari Nooghabi, Shrinkage Estimation of $P(Y< X)$ in the Exponential Distribution Mixing with Exponential Distribution, Communications in Statistics-Theory and Methods, **45** (2016) 1477-1486.

5. S. Khan, Z. Hoque, A.K.M.E. Saleh, Estimation of the Intercept Parameter for Linear Regression Model with Uncertain Prior Information, Statistical Papers, **46** (3) (2005) 379-395.

6. D. C. Montgomery, E.A. Peck, Vining G.G., Introduction to Linear Regression Analysis, 5th Edition, Wiley (2012).

7. M. Naghizadeh Qomi, A. Kiapour, Improving the MLE of the Location Parameter of a Normal Population under Reflected Normal Loss, 47th Annual Iranian Mathematics Conference, Kharazmi University, Karaj, Iran (2016).

8. M. Naghizadeh Qomi, N. Nematollahi, A. Parsian, Estimation After Selection Under Reflected Normal Loss Function, Communications in Statistics Theory and Methods, **41** (2012) 1040-1051.

9. F. Spring, The Reflected Normal Loss Function, Canadian Journal of Statistics, **21** (1) (1993) 321-330.

10. J.R. Thompson, Some Shrunken Techniques for Estimating the Mean, Journal of the American Statistician Association, **63** (1968) 113-122.

11. M. Towhidi , J. Behboodian, Estimation of a Location Parameter with a Reflected Normal Loss function, Iranian Journal of Science and Technology, **25** (2001) 183-190.

12. S.K. Tse, Tso G., Shrinkage Estimation of Reliability Distribution Lifetimes, Communications in Statistics - Simulation and Computation, **25** (2) (1996) 415-430.

13. H. Waldl, Some Comments on Statistical Papers 46, 270-395, Statistical Papers, **51** (2010) 241-246.