

## برآورد بازه‌ای در توزیع‌های پواسن و آبل (پواسن تعمیم یافته)

محسن عارفی<sup>۱</sup>، غلامرضا محتشمی‌برزادران<sup>۲\*</sup> و یدالله واقعی<sup>۳</sup>

\*پست الکترونیکی: [gmohtashami@birjand.ac.ir](mailto:gmohtashami@birjand.ac.ir)

تاریخ دریافت: ۸۶/۳/۱ تاریخ پذیرش: ۸۷/۱۱/۶

### چکیده

در این مقاله، برآورد بازه‌ای برای میانگین جامعه در توزیع‌های پواسن و آبل (پواسن تعمیم یافته) مورد بررسی قرار گرفته است. توزیع پواسن جزء خانواده توزیع‌های نمایی طبیعی با تابع واریانس درجه دوم بر حسب میانگین است که این خانواده شامل شش توزیع «دوجمله‌ای، دوجمله‌ای منفی، پواسن، نرمال، گاما و سکانت هندلولی تعمیم یافته» می‌باشد. توزیع آبل نیز جزء خانواده توزیع‌های نمایی طبیعی با تابع واریانس درجه سوم بر حسب میانگین است که این خانواده شامل نیز شش توزیع «تاکاس، آبل، استریک آرکسینوس، لارج آرکسینوس، رسل و گوسین معکوس» می‌باشد. برای این دو توزیع، بازه‌های ممکن «والد، اسکور، اگرتی-کول، نسبت درست‌نمایی و جفریز» محاسبه و از نظر طول مورد انتظار و احتمال پوشش مورد مقایسه قرار گرفته‌اند. از مقایسه این بازه‌های اطمینان با یکدیگر معلوم می‌گردد که دو بازه جفریز و نسبت درست‌نمایی دارای طول کم‌تری هستند و احتمال پوشش آن‌ها نیز به سطح اطمینان نزدیک‌تر است.

واژه‌های کلیدی: احتمال پوشش، تابع واریانس درجه دوم، تابع واریانس درجه سوم، توزیع آبل، توزیع پواسن، خانواده نمایی طبیعی، بازه اسکور، بازه اگرتی-کویل، بازه جفریز، بازه نسبت درست‌نمایی، بازه والد

### مقدمه

اهمیت است. برآورد بازه‌ای برای خانواده توزیع‌های نمایی طبیعی در مقالات متعددی از جمله سنتنر [۱]، اگرتی-کول [۲]، کای [۳]، لویی [۴] و سون و همکاران [۵] مورد مطالعه و بررسی قرار گرفته است. انواع بازه‌های اطمینان والد، اسکور، اگرتی-کول، نسبت درست‌نمایی

از آن جایی که برآورد نقطه‌ای مقدار مناسبی از پارامتر را گزارش نمی‌کند، محاسبه برآورد بازه‌ای بسیار حائز اهمیت است. برآورد بازه‌ای برای خانواده توزیع‌های

۱- دانشجوی دکتری دانشکده علوم ریاضی - دانشگاه صنعتی اصفهان

۲- دانشیار گروه آمار - دانشگاه فردوسی مشهد

۳- استادیار گروه آمار - دانشگاه بیرجند

با یکدیگر مورد مقایسه قرار می‌گیرد. نکته قابل توجه در محاسبه برخی از بازه‌های اطمینان برای توزیع آبل، حل معادلات درجه سوم بر حسب میانگین توزیع می‌باشد.

### خانواده توزیع‌های نمایی طبیعی

خانواده توزیع‌های نمایی طبیعی اولین بار توسط فیشر [۲۲] بررسی گردید و در این زمینه مقالات و کتب بسیاری به چاپ رسیده است. خانواده توزیع‌های نمایی طبیعی با تابع چگالی (یا تابع احتمال)

$$f(x) = \exp\{\xi x - \psi(\xi)\} h(x),$$

نشان داده می‌شوند، که در آن دامنه تغییرات  $x \in A \subset \mathbb{R}$  به فضای پارامتر  $\Theta$  بستگی ندارد. همچنین  $\xi$  و  $\psi(\xi)$  توابعی بر حسب پارامتر  $\theta$  می‌باشند. (برای اطلاع بیشتر به لهن [۲۳] و بارندروف-نیلسون [۲۴] مراجعه نمایید).  
برای خانواده توزیع‌های نمایی طبیعی فوق

$$\begin{cases} \mu = E(X) = \psi'(\xi), \\ \sigma^2 = \text{var}(X) = \psi''(\xi), \end{cases}$$

به ترتیب میانگین و تابع واریانس نامیده می‌شوند. خانواده توزیع‌های نمایی طبیعی با تابع واریانس درجه دوم بر حسب میانگین و برخی از خواص آن توسط موریس [۲۵] و [۲۶] و لتاک [۲۷] و با تابع واریانس درجه سوم بر حسب میانگین توسط لتاک و مورا [۲۱] و کونسول [۲۸] مورد بررسی قرار گرفته است، که تابع واریانس آن‌ها بر حسب میانگین به صورت

$$\sigma^3 = \text{var}(X) = a_0 + a_1 \mu + a_2 \mu^2 + a_3 \mu^3$$

است، که در آن ضرایب  $a_i$  مقادیر ثابتی هستند. خانواده توزیع‌های نمایی طبیعی با تابع واریانس درجه دوم شامل شش توزیع «دوجمله‌ای، دوجمله‌ای منفی، پواسن، نرمال، گاما و سکانت هذلولی تعمیم یافته» و با تابع واریانس درجه سوم نیز شامل شش توزیع «آبل،

جفریز در این منابع بررسی شده است. یک منبع مناسب برای مقایسه این بازه‌های اطمینان برای خانواده توزیع‌های نمایی طبیعی با تابع واریانس درجه دوم بر حسب میانگین برون و همکاران [۶] است. احتمال‌های پوشش و طول مورد انتظار، از بازه‌های اطمینان مختلف برای توزیع‌های دوجمله‌ای، دوجمله‌ای منفی، پواسن، نرمال، گاما و سکانت هذلولوی تعمیم یافته نیز توسط برون و همکاران [۶] مطالعه گردیده است. این توزیع‌ها، اعضای خانواده نمایی طبیعی با تابع واریانس درجه دوم بر حسب میانگین هستند. برآورد بازه‌ای برای توزیع دوجمله‌ای با یک یا دو نمونه تصادفی نیز در سال‌های اخیر مورد توجه قرار گرفته است. برای بررسی و مطالعه در این زمینه به برون و همکاران [۷-۹]، ریزگل [۱۰]، زو و همکاران [۱۱]، برون و لی [۱۲]، وانگ [۱۳] و روسس و تبی [۱۴] مراجعه نمایید. بازه‌های اطمینان برای توزیع گوسین معکوس و شبیه‌سازی احتمال‌های پوشش برای این بازه‌ها، توسط عارفی و همکاران [۱۵] مطالعه شده است. همچنین، توزیع آبل (پواسن تعمیم یافته) و برخی خصوصیات آن به وسیله نویسندگان بسیاری مورد ارزیابی و بررسی قرار گرفته است. برای بررسی برخی از این منابع می‌توان به کونسول [۱۶]، کونسول و جاین [۱۷]، کونسول و شوکری [۱۸]، ورنیک [۱۹] و امبگاسفیا و بلکریشنان [۲۰] مراجعه کرد. باید توجه داشت که توزیع گوسین معکوس و آبل دو عضو از خانواده نمایی طبیعی با تابع واریانس درجه سوم بر حسب میانگین هستند (برای بررسی بیشتر به لتاک و مورا [۲۱] مراجعه نمایید).

در این مقاله، علاوه بر بررسی بازه‌های اطمینان و خواص آن برای توزیع پواسن، این بازه‌ها را برای توزیع آبل نیز که جزء خانواده توزیع‌های نمایی طبیعی با تابع واریانس درجه سوم است، تعمیم می‌دهیم. از طرف دیگر، احتمال پوشش این بازه‌های اطمینان به وسیله شبیه‌سازی برآورد و

بازه اگوستی- کول  $CI_{AC}$ : این بازه برای خانواده توزیع‌های نمایی طبیعی با تابع واریانس درجه دوم، بر

اساس تصحیح  $\hat{\mu} = \frac{n\hat{\mu} + k^2 a_0}{n - k^2 a_0}$  و  $\hat{\sigma}^2 = \frac{n\hat{\sigma}^2 + k^2 a_1}{n - k^2 a_1}$  در بازه والد به جای  $\hat{\mu}$  و  $n$  به دست می‌آید:

$$CI_{AC} = \hat{\mu} \pm k (\text{var}(\hat{\mu}))^{1/2} \hat{\sigma}^{-1/2} = \hat{\mu} \pm k (a_0 + a_1 \hat{\mu} + a_2 \hat{\mu}^2)^{1/2} \hat{\sigma}^{-1/2}. \quad (3)$$

بازه نسبت درست‌نمایی  $CI_{LR}$ : این بازه بر مبنای وارونی از آزمون نسبت درست‌نمایی، با پذیرش فرض صفر  $H_0: \mu = \mu_0$  تحت رابطه  $-2 \log \Lambda_n \leq \chi_{\alpha, 1}^2$  محاسبه می‌شود.  $\Lambda_n$  برابر است با:

$$\Lambda_n = \frac{L(\mu_0)}{\sup_{\mu} L(\mu)},$$

برای بررسی بیشتر به راثو [۲۹] و سرفلینگ [۳۰] مراجعه نمایید.

بازه جفریز  $CI_J$ : اطلاع فیشر برای خانواده توزیع نمایی طبیعی به صورت زیر محاسبه می‌شود:

$$I(\mu) = -E \left( \frac{\partial^2 \ln f(\underline{X}, \mu)}{\partial \mu^2} \right) = n(\psi''(\xi))^{-1}.$$

جفریز [۳۱] بر اساس اطلاع فیشر، تابع پیشینی را متناسب با  $P(\mu) \propto \sqrt{I(\mu)}$  معرفی نمود. بر این اساس، تابع پسین متناسب است با:

$$f(\mu | \underline{x}) \propto \exp\{n \bar{x} \xi - n \psi(\xi) - \frac{1}{\gamma} \ln(a_0 + a_1 \mu + a_2 \mu^2 + a_3 \mu^3)\}.$$

بنابراین، بازه جفریز چندک‌های  $\frac{\alpha}{\gamma}$  و  $1 - \frac{\alpha}{\gamma}$  از تابع پسین فوق است:

$$CI_J = [J_{\alpha/\gamma}, J_{1-\alpha/\gamma}].$$

تاکاس، استریک آرکسینوس، لارج آرکسینوس، رسل و گوسین معکوس» علاوه بر خانواده نمایی طبیعی با تابع واریانس درجه دوم می‌باشد. همچنین برآورد درست‌نمایی ماکزیمم برای  $\mu$  برابر  $\hat{\mu} = \bar{x}$  است. هدف از این مقاله، محاسبه بازه‌های اطمینان برای دو توزیع پواسن و آبل (پواسن تعمیم یافته) می‌باشد.

### انواع بازه‌های اطمینان

محاسبه برآورد بازه‌ای برای خانواده توزیع‌های نمایی طبیعی مقوله بسیار مهمی است که در مقالات بی‌شماری برای توزیع‌های گسسته و پیوسته به آن پرداخته شده است، که از آن جمله می‌توان به سنتنر [۱]، اگوستی- کول [۲]، کای [۳] و برون و همکاران [۶] اشاره کرد. اکنون انواع مختلف بازه‌های اطمینان را می‌توان به صورت زیر محاسبه نمود:

بازه والد  $CI_W$ : این بازه بر اساس کمیت اسلوتسکی

$$W_n = \frac{\sqrt{n}(\hat{\mu} - \mu)}{\sqrt{a_0 + a_1 \hat{\mu} + a_2 \hat{\mu}^2 + a_3 \hat{\mu}^3}},$$

به صورت زیر محاسبه می‌شود:

$$CI_W = \hat{\mu} \pm k \hat{\sigma} n^{-1/2} = \hat{\mu} \pm k (a_0 + a_1 \hat{\mu} + a_2 \hat{\mu}^2 + a_3 \hat{\mu}^3)^{1/2} n^{-1/2}. \quad (1)$$

بازه اسکور  $CI_S$ : این بازه بر اساس قضیه حد مرکزی تحت کمیت زیر و با حل معادلات درجه سوم  $-k \leq Z \leq k$  بر حسب  $\mu$  به دست می‌آید:

$$Z = \frac{\sqrt{n}(\hat{\mu} - \mu)}{\sqrt{a_0 + a_1 \hat{\mu} + a_2 \hat{\mu}^2 + a_3 \hat{\mu}^3}}, \quad (2)$$

که در حالت با تابع واریانس درجه دوم به حل معادلات درجه دوم تبدیل می‌شود.

## ۳-۱ بازه آگرستی-کول

این بازه بر اساس رابطه (۳) به صورت زیر محاسبه می‌شود:

$$CI_{AC}: \hat{\lambda} + \frac{k^2}{2n} \pm k n^{-1/2} \left( \hat{\lambda} + \frac{k^2}{2n} \right)^{1/2}.$$

## ۴-۱ بازه نسبت درست‌نمایی

نسبت درست‌نمایی برای توزیع پواسن برابر است با:

$$\Lambda_n = \frac{\lambda^{n\bar{x}} e^{-n\lambda}}{\bar{x}^{n\bar{x}} e^{-n\bar{x}}}.$$

معادله  $-\ln \Lambda_n \leq k^2/2$  دارای دو ریشه بر حسب  $\lambda$

است. با تغییر متغیر  $t = \frac{\lambda}{\bar{x}} - 1$  معادله فوق به صورت

$$t - \ln(1+t) = \frac{k^2}{2n\bar{x}},$$

تبدیل می‌شود. بر اساس بسط مکلورن بر حسب  $t$  معادله بالا به صورت زیر تبدیل می‌شود:

$$\frac{t^2}{2} - \frac{t^3}{3} + \frac{t^4}{4} - \frac{k^2}{2n\bar{x}} = O(n^{-2}). \quad (\varepsilon)$$

اکنون فرض کنید  $t = b_0 + b_1 n^{-1/2} \pm k b_2 n^{-1} + b_3 n^{-3/2}$

با  $b_0 = \pm \frac{k b_1 b_2}{12} - b_3$  باشد. با قرار دادن  $t$  در معادله (۴)

و حل نسبت به  $b_0, b_1, b_2, b_3$  داریم:

$$b_0 = 0, \quad b_1 = \pm \frac{k}{\bar{x}^{1/2}},$$

$$b_2 = \pm \frac{k}{3\bar{x}}, \quad b_3 = \pm \frac{k^2}{36\bar{x}^{3/2}}.$$

در نتیجه حدود بازه اطمینان بر حسب  $t$  برابر است با:

$$\begin{cases} \underline{t} = b_0 + b_1 n^{-1/2} - k b_2 n^{-1} + b_3 n^{-3/2} + R_n, \\ \bar{t} = b_0 + b_1 n^{-1/2} + k b_2 n^{-1} + b_3 n^{-3/2} + R_n. \end{cases}$$

بنابراین بازه نسبت درست‌نمایی برای  $\lambda$  به صورت زیر محاسبه می‌شود:

تذکر ۱: بازه آگرستی-کول فقط در خانواده توزیع‌های نمایی طبیعی با تابع واریانس درجه دوم قابل محاسبه می‌باشد.

تذکر ۲: در این مقاله، فقط میانگین خانواده نمایی  $\mu$  و هر تابع از این میانگین به عنوان پارامتر مجهول در نظر گرفته می‌شود. بنابراین اگر پارامتری غیر از میانگین توزیع وجود داشته باشد، آن را با یک برآورد مناسب جایگزین می‌کنیم.

## توزیع پواسن

توزیع  $Poisson(\lambda)$  عضوی از خانواده توزیع‌های نمایی طبیعی با تابع واریانس درجه دوم بر حسب میانگین است. تابع احتمال آن به صورت زیر است:

$$f(x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, \dots, \lambda > 0.$$

واریانس در این توزیع برابر است با  $Var(X) = \lambda = \mu$  بنابراین  $a_0 = a_1 = 0$  و  $a_2 = 1$ . همچنین برآورد درست‌نمایی ماکزیمم برای  $\lambda$  برابر  $\hat{\lambda} = \bar{x}$  است.

## ۱- انواع بازه‌های اطمینان برای میانگین توزیع پواسن

## ۱-۱ بازه والد

این بازه بر اساس رابطه (۱) به صورت زیر محاسبه می‌شود:

$$CI_W: \hat{\lambda} \pm k \sqrt{\hat{\lambda}/n}.$$

## ۲-۱ بازه اسکور

این بازه بر اساس رابطه (۲) و حل یک معادله درجه دوم بر حسب  $\mu$  به صورت زیر به دست می‌آید:

$$CI_S: \hat{\lambda} + \frac{k^2}{2n} \pm k n^{-1/2} \left( \hat{\lambda} + \frac{k^2}{2n} \right)^{1/2}.$$

## توزیع آبل (پواسن تعمیم یافته)

توزیع آبل  $Abel(\beta, \theta)$  عضوی از خانواده توزیع‌های نمایی طبیعی با تابع واریانس درجه سوم بر حسب میانگین است. تابع احتمال آن به صورت زیر است:

$$f(x) = \frac{(1 + \beta x)^{x-1}}{x!} \theta^x e^{-\theta(1+\beta x)}, \quad x = 0, 1, \dots, \\ \theta > 0, \quad 0 \leq \beta < \theta^{-1}.$$

میانگین و تابع واریانس برای این توزیع برابر است با:

$$\begin{cases} \mu = \frac{\theta}{1 - \beta\theta}, \\ \text{var}(X) = \frac{\theta}{(1 - \beta\theta)^2} = \mu(1 + \beta\mu), \end{cases}$$

بنابراین ضرایب تابع واریانس برابر  $a_0 = 0$ ،  $a_1 = 1$ ،  $a_2 = 2\beta$  و  $a_3 = \beta^2$  است. همچنین برآورد درست‌نمایی ماکزیمم برای پارامتر  $\mu$  برابر  $\hat{\mu} = \bar{x}$  است. برای  $\beta = 0$ ، توزیع پواسن با میانگین  $\theta$  نتیجه می‌شود.

## ۱- انواع بازه‌های اطمینان برای میانگین در توزیع آبل

## ۱-۱ بازه والد

این بازه بر اساس رابطه (۱) به صورت زیر محاسبه می‌شود:

$$CI_W : \hat{\mu} \pm k n^{-1/2} \hat{\mu}^{1/2} (1 + \beta \hat{\mu}).$$

## ۱-۲ بازه اسکور

بر اساس کمیت  $-k \leq \frac{\sqrt{n}(\hat{\mu} - \mu)}{\sqrt{\mu(1 + \beta\mu)}} \leq k$  و با تغییر متغیر

$y = \mu$ ، دو معادله درجه سوم زیر به دست می‌آید:

$$\begin{cases} y^3 + \frac{\sqrt{n}}{\beta k} y^2 + \frac{1}{\beta} y - \frac{\sqrt{n}}{\beta k} \bar{x} \geq 0, \\ y^3 - \frac{\sqrt{n}}{\beta k} y^2 + \frac{1}{\beta} y + \frac{\sqrt{n}}{\beta k} \bar{x} \geq 0. \end{cases}$$

با توجه به حل معادلات درجه سوم، مشخصه‌های زیر را داریم:

$$CI_{LR} : \pm k \bar{x}^{-1/2} n^{-1/2} + \frac{1}{3} k^2 n^{-1} \\ \pm \frac{1}{36} k^2 \bar{x}^{-1/2} n^{-1/2} + \bar{x} + R_n.$$

## ۱-۵ بازه جفریز

تابع پیشین برای توزیع پواسن متناسب با  $\lambda^{-1/2}$  و تابع احتمال پسین به صورت  $Gamma(n\bar{x} + 1/2, 1/n)$  محاسبه می‌شود. بنابراین بازه جفریز در سطح اطمینان  $1 - \alpha$  برای  $\lambda$  برابر است با:

$$CI_J : [Gamma_{\{\alpha/2, n\bar{x} + 1/2, 1/n\}}, \\ Gamma_{\{1 - \alpha/2, n\bar{x} + 1/2, 1/n\}}].$$

## ۲- مقایسه انواع بازه‌های اطمینان از نظر طول مورد انتظار

## در توزیع پواسن

طول مورد انتظار برای توزیع پواسن به صورت زیر محاسبه می‌شود:

$$E(L_W) = 2k\lambda^{1/2} n^{-1/2} \left(1 - \frac{9}{\sqrt{2n\lambda}}\right) + O(n^{-1}), \\ E(L_{LR}) = 2k\lambda^{1/2} n^{-1/2} \left(1 + \frac{9(k^2 - 1) - \sqrt{k^2}}{\sqrt{2n\lambda}}\right) + O(n^{-1}), \\ E(L_J) = 2k\lambda^{1/2} n^{-1/2} \left(1 + \frac{9(k^2 - 1) + 4 - \sqrt{k^2}}{\sqrt{2n\lambda}}\right) + O(n^{-1}), \\ E(L_R) = 2k\lambda^{1/2} n^{-1/2} \left(1 + \frac{9(k^2 - 1)}{\sqrt{2n\lambda}}\right) + O(n^{-1}), \\ E(L_{AC}) = 2k\lambda^{1/2} n^{-1/2} \left(1 + \frac{9(2k^2 - 1)}{\sqrt{2n\lambda}}\right) + O(n^{-1}).$$

در نتیجه برای هر  $k > 2/\sqrt{7} = 0.76$ ، طول مورد انتظار به صورت زیر مرتب می‌شود:

$$E(L_W) < E(L_{LR}) < E(L_J) < E(L_S) < E(L_{AC}).$$

اگر از بازه والد که بازه ضعیفی است صرف‌نظر کنیم، آن‌گاه بازه نسبت درست‌نمایی و جفریز دارای طول مورد انتظار کم‌تری هستند. برای بررسی بیش‌تر در این زمینه به برون و همکاران [۶] مراجعه شود.

$$\Lambda_n = \frac{\left(\frac{\mu}{1+\beta\mu}\right)^{n\bar{x}} \exp\left\{-\frac{n\mu}{1+\beta\mu}(1+\beta\bar{x})\right\}}{\left(\frac{\bar{x}}{1+\beta\bar{x}}\right)^{n\bar{x}} \exp\left\{-\frac{n\bar{x}}{1+\beta\bar{x}}(1+\beta\bar{x})\right\}},$$

با توجه به رابطه  $-\ln(\Lambda_n) \leq k^\tau$  داریم:

$$\frac{\mu - \bar{x}}{1 + \beta\mu} - \bar{x} \ln\left(\frac{\mu(1 + \beta\bar{x})}{\bar{x}(1 + \beta\mu)}\right) \leq \frac{k^\tau}{2n}.$$

با تغییر متغیر  $t = \frac{\mu(1 + \beta\bar{x})}{\bar{x}(1 + \beta\mu)} - 1$ ، رابطه بالا به صورت

زیر تبدیل می‌شود:

$$t - \ln(1+t) \leq \frac{k^\tau}{2n\bar{x}},$$

که شبیه معادله بازه نسبت درست‌نمایی در توزیع پواسن است. ریشه‌های این معادله بر حسب  $t$  به صورت زیر است:

$$\begin{cases} \underline{t} = -k(n\bar{x})^{-1/\tau} + \frac{1}{3}k^\tau(n\bar{x})^{-1} - \frac{1}{36}k^\tau(n\bar{x})^{-2/\tau} + R_n, \\ \bar{t} = +k(n\bar{x})^{-1/\tau} + \frac{1}{3}k^\tau(n\bar{x})^{-1} + \frac{1}{36}k^\tau(n\bar{x})^{-2/\tau} + R_n. \end{cases}$$

با توجه به تغییر متغیر  $t = \frac{\mu(1 + \beta\bar{x})}{\bar{x}(1 + \beta\mu)} - 1$  داریم:

$$\frac{\bar{x}}{1 + \beta\bar{x}}(\underline{t} + 1) \leq \frac{\mu}{1 + \beta\mu} \leq \frac{\bar{x}}{1 + \beta\bar{x}}(\bar{t} + 1),$$

بنابراین بازه مورد نظر برای  $\mu$  برابر است با:

$$\frac{\bar{x}(\underline{t} + 1)}{1 - \beta\bar{x}\underline{t}} \leq \mu \leq \frac{\bar{x}(\bar{t} + 1)}{1 - \beta\bar{x}\bar{t}}.$$

#### ۱-۴ بازه جفریز

تابع پیشین جفریز متناسب با  $\theta^{-1/\tau}(1 - \beta\theta)^{-1/\tau}$  و تابع احتمال پسین برابر است با:

$$f(\theta | \underline{x}) = \frac{1}{K} \theta^{n\bar{x}-1/\tau} (1 - \beta\theta)^{-1/\tau} \exp\{-n\theta(1 + \beta\bar{x})\}.$$

$$\begin{cases} Q = \frac{n - \beta k^\tau}{\sqrt{n} \sqrt{2n - 9\beta k^\tau - 27\beta^2 k^\tau}}, \\ R = \frac{54\beta^2 k^\tau}{\sqrt{n} \sqrt{2n - 9\beta k^\tau - 27\beta^2 k^\tau}}. \end{cases}$$

در نتیجه داریم:

$$Q^\tau - R^\tau = \frac{-\varepsilon\beta k^\varepsilon + nk^\tau(1 - 18\beta - 27\beta^2) + \varepsilon n^\tau}{108\beta^2 k^\varepsilon}.$$

فرض کنید مشخصه‌های زیر را داشته باشیم:

$$0 \leq \beta \leq \frac{\sqrt{432} - 18}{54} = 0.1052,$$

$$k^\tau \geq \frac{n(1 - 18\beta - 27\beta^2) + \sqrt{n^2(1 - 18\beta - 27\beta^2)^2 + 72\varepsilon n^\tau \beta}}{18\beta}.$$

با توجه به این مشخصه‌ها و به ازای  $n > 20$ ، معادلات بالا هر کدام یک ریشه به صورت زیر دارند:

$$\begin{cases} x = -\left[ (|R| + \sqrt{R^\tau - Q^\tau})^{1/\tau} + (|R| - \sqrt{R^\tau - Q^\tau})^{1/\tau} \right] - \frac{\sqrt{n}}{3\beta k}, \\ x' = +\left[ (|R| + \sqrt{R^\tau - Q^\tau})^{1/\tau} + (|R| - \sqrt{R^\tau - Q^\tau})^{1/\tau} \right] + \frac{\sqrt{n}}{3\beta k}. \end{cases}$$

در نتیجه می‌توان نوشت:

$$\begin{cases} x \leq y \\ x' \leq y \end{cases} \Rightarrow x^\tau \leq y^\tau,$$

با تغییر متغیر  $y^\tau = \mu$  به صورت  $x^\tau \leq \mu$  تبدیل می‌شود که یک بازه یک‌طرفه است.

تذکر ۳: اگر در بازه اسکور، مشخصات فوق تغییر کند، آن‌گاه ممکن است یک بازه اطمینان دوطرفه به دست آید. در مثال‌های عددی این بازه را به طور مستقیم از کمیت

$$\text{محوری } -k \leq \frac{\sqrt{n}(\hat{\mu} - \mu)}{\sqrt{\mu(1 + \beta\mu)}} \leq k \text{ به دست می‌آوریم.}$$

#### ۱-۳ بازه نسبت درست‌نمایی

نسبت درست‌نمایی برای توزیع آبل برابر است با:

$$E(W_n) = -\frac{1}{2}(a_1 + 2a_1\mu + 3a_1\mu^2)\sigma^{-1}n^{-1/2} \\ - \frac{15}{16}(a_1 + 2a_1\mu + 3a_1\mu^2)^2\sigma^{-2}n^{-3/2} \\ + \frac{9}{4}(a_1 + 2a_1\mu + 3a_1\mu^2)(a_1 + 3a_1\mu)\sigma^{-1}n^{-3/2} \\ - \frac{3}{2}a_1n^{-3/2}\sigma + O(n^{-2}).$$

میزان اریبی برای توزیع پواسن برابر است با:

$$E(W_n) = -\frac{1}{2\sqrt{n}\lambda}(1 + \frac{15}{8n\lambda}) + O(n^{-2}),$$

بنابراین میزان اریبی برای توزیع پواسن، به ازای تمام مقادیر  $n$  و  $\lambda$  منفی است. همچنین میزان اریبی برای توزیع آبل برابر است با:

$$E(W_n) = -\frac{1}{2}(1 + \epsilon\beta\mu + 3\beta^2\mu^2)\mu^{-1/2}(1 + \beta\mu)^{-1} \\ n^{-1/2}\left\{1 + \frac{15 + \epsilon 8\beta\mu + 78\beta^2\mu^2 + 72\beta^2\mu^2 + 27\beta^4\mu^4}{8n\mu(1 + \beta\mu)^2} \right. \\ \left. + \frac{3\beta^2\mu(1 + \beta\mu)^2}{n(1 + \epsilon\beta\mu + 3\beta^2\mu^2)}\right\} + O(n^{-2});$$

بنابراین میزان اریبی برای توزیع آبل همواره منفی است.

### مثال‌های عددی

مثال ۱: تعداد تصادفات در یک جاده بین ساعت ۱۰ تا ۱۲ برای ۵۰ روز گذشته به صورت زیر است:

داده	داده	داده	داده	داده	داده	داده	داده	داده	داده
۸	۷	۱۵	۶	۱۲	۱۲	۷	۹	۸	۶
۹	۱۰	۷	۵	۱۵	۱۳	۶	۶	۹	۱۲
۱۵	۶	۱۰	۱۴	۹	۱۱	۱۲	۱۱	۵	۱۰
۱۵	۱۱	۱۲	۷	۱۴	۸	۷	۹	۱۶	۱۲
۱۲	۱۱	۸	۱۱	۷	۴	۱۳	۹	۱۱	۶

فرض کنید، این نمونه تصادفی دارای توزیع پواسن  $Poisson(\lambda)$  باشند. برآورد ماکزیمم درست‌نمایی به

در نتیجه چندک‌های  $\frac{\alpha}{2}$  و  $1 - \frac{\alpha}{2}$  از تابع پسین فوق، بازه جفریز برای  $\theta$  است.

تذکر ۴: محاسبه طول مورد انتظار برای بازه‌های جفریز، نسبت درست‌نمایی و اسکور بر اساس بسط مکولرن کار پیچیده‌ای است و نیاز به محاسبات عددی دارد.

تذکر ۵: در بازه‌های اطمینان محاسبه شده برای میانگین توزیع آبل، فرض بر این است که پارامتر  $\beta$  معلوم باشد. در غیر این صورت، ابتدا مقدار برآوردی را به جای پارامتر  $\beta$  جایگزین می‌کنیم و آن‌گاه بازه‌های اطمینان را محاسبه می‌نماییم. برای مثال، برآورد گشتاوری برای پارامتر  $\beta$  برابر است با  $\hat{\beta} = \max[0, \frac{1}{\bar{x}}(\frac{s}{\sqrt{\bar{x}}} - 1)]$  که

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

تذکر ۶: برای به دست آوردن برخی از محاسبات عددی و انتگرال‌ها، می‌توان از نرم‌افزار Maple استفاده کرد.

### میزان اریبی در توزیع پواسن و آبل

بسط کمیت اسلوتسکی  $W_n$  بر اساس  $Z$ ، برای خانواده توزیع‌های نمایی طبیعی با تابع واریانس درجه سوم (یا با تابع واریانس درجه دوم با  $a_0 = 0$ ) به صورت زیر است:

$$W_n = \frac{\sqrt{n}(\hat{\mu} - \mu)}{\sqrt{a_0 + a_1\hat{\mu} + a_2\hat{\mu}^2 + a_3\hat{\mu}^3}} \\ = Z[1 + (a_1 + 2a_1\mu + 3a_1\mu^2)\sigma^{-1}n^{-1/2}Z \\ + (a_2 + 3a_2\mu)n^{-1}Z^2 + a_3n^{-3/2}\sigma Z^3]^{-1/2}.$$

بر اساس بسط مکولرن بر حسب  $Z$  و با استفاده از گشتاورهای توزیع نرمال، میزان اریبی برابر است با:

ابتدا بر اساس تذکر ۴، برآورد پارامتر  $\beta$  به صورت  $\hat{\beta} = 0.002$  به دست می‌آید. اکنون بازه‌های مختلف برای پارامتر  $\theta$  در سطح اطمینان  $1 - \alpha = 0.95$  به صورت زیر به دست می‌آیند (باید توجه داشت که، ابتدا بازه‌ها برای میانگین توزیع به دست آمده است و آن‌گاه از روی آن بازه‌های مختلف برای پارامتر  $\theta$  به دست می‌آید):

$$CI_W : [7/9522, 9/5796],$$

$$CI_S : [7/9910, 9/6202],$$

$$CI_{LR} : [7/9784, 9/6063],$$

$$CI_J : [7/8683, 9/4424].$$

طول این بازه‌ها برابر است با:

$$L_W = 1/6274, \quad L_S = 1/6292,$$

$$L_{LR} = 1/6279, \quad L_J = 1/5741.$$

بنابراین، طول‌ها برای بازه‌های مختلف به صورت  $L_J < L_W < L_{LR} < L_S$  مرتب می‌شوند.

### بررسی احتمال پوشش برای برآورد بازه اطمینان به کمک شبیه‌سازی

مسئله شبیه‌سازی احتمال پوشش در برآورد بازه‌های مختلف و مقایسه بازه‌های اطمینان از نظر نزدیکی احتمال پوشش به سطح اطمینان، بسیار حائز اهمیت می‌باشد. فرض کنید یک پارامتر ثابت داشته باشیم. تحت شرایط حدی یک بازه اطمینان در سطح اطمینان  $1 - \alpha$  به دست می‌آوریم. اگر بازه، پارامتر مورد نظر را در برداشته باشد، بازه مذکور پارامتر را پوشش می‌دهد و گرنه آن را پوشش نمی‌دهد. بنابراین پوشش یک متغیر برنولی با احتمال پوشش واقعی  $P$  می‌باشد. محاسبه احتمال پوشش واقعی به طور نظری کار پیچیده‌ای است، ولی می‌توان آنرا به کمک شبیه‌سازی برآورد کرد. بدین منظور با یک اندازه نمونه تصادفی و یک پارامتر معین، یک بازه را در شرایط

صورت  $\hat{\lambda} = \bar{x} = 9/76$  است. بنابراین، برآوردهای بازه‌ای مختلف برای پارامتر  $\lambda$  در سطح اطمینان  $1 - \alpha = 0.95$  به صورت زیر به دست می‌آیند:

$$CI_W : [8/8941, 10/6259],$$

$$CI_S : [8/9316, 10/6652],$$

$$CI_{AC} : [8/9199, 10/6542],$$

$$CI_{LR} : [8/9195, 10/6517],$$

$$CI_J : [8/9227, 10/6552].$$

طول این بازه‌ها برابر است با:

$$L_W = 1/7318, \quad L_S = 1/7336, \quad L_{AC} = 1/7343,$$

$$L_{LR} = 1/7322, \quad L_J = 1/7325.$$

بنابراین، طول‌ها برای بازه‌های مختلف به صورت  $L_W < L_{LR} < L_J < L_S < L_{AC}$  مرتب می‌شوند، که با نتیجه به دست آمده برای طول مورد انتظار در زیربخش ۴-۲ برابر است.

تذکر ۷: اگر فرض کنیم که داده‌های مثال ۱ دارای توزیع آبل  $Abel(\beta, \theta)$  هستند، آن‌گاه با توجه به تذکر ۴، برآورد پارامتر  $\beta$  برابر است با

$$\hat{\beta} = \max(0, -0.014) = 0.$$

در این صورت توزیع آبل با توزیع پواسن برابر خواهد بود.

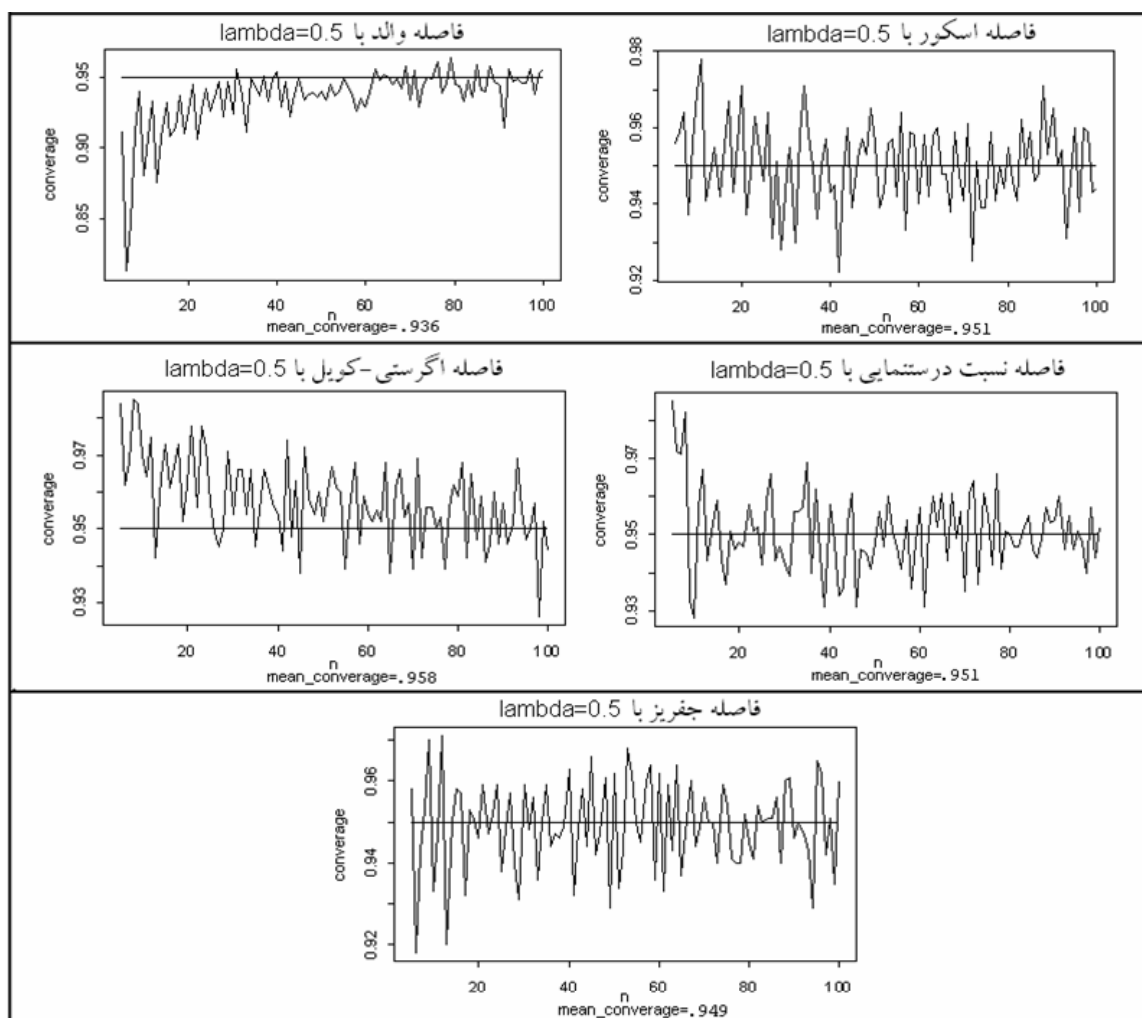
مثال ۲: یک نمونه تصادفی به حجم  $n = 50$  از یک جامعه آبل  $Abel(\beta, \theta)$  گرفته شده است و داده‌های زیر به دست آمده‌اند:

داده	داده	داده	داده	داده	داده	داده	داده	داده	داده
۵	۹	۹	۱۵	۵	۱۱	۶	۱۰	۱۰	۱۱
۱۰	۵	۱۲	۷	۸	۶	۶	۱۲	۱۴	۴
۱۱	۱۱	۱۰	۱۶	۹	۸	۹	۶	۹	۶
۴	۱۳	۹	۶	۶	۱۴	۱۲	۷	۶	۸
۶	۳	۱۴	۸	۱۱	۹	۱۱	۹	۹	۱۱

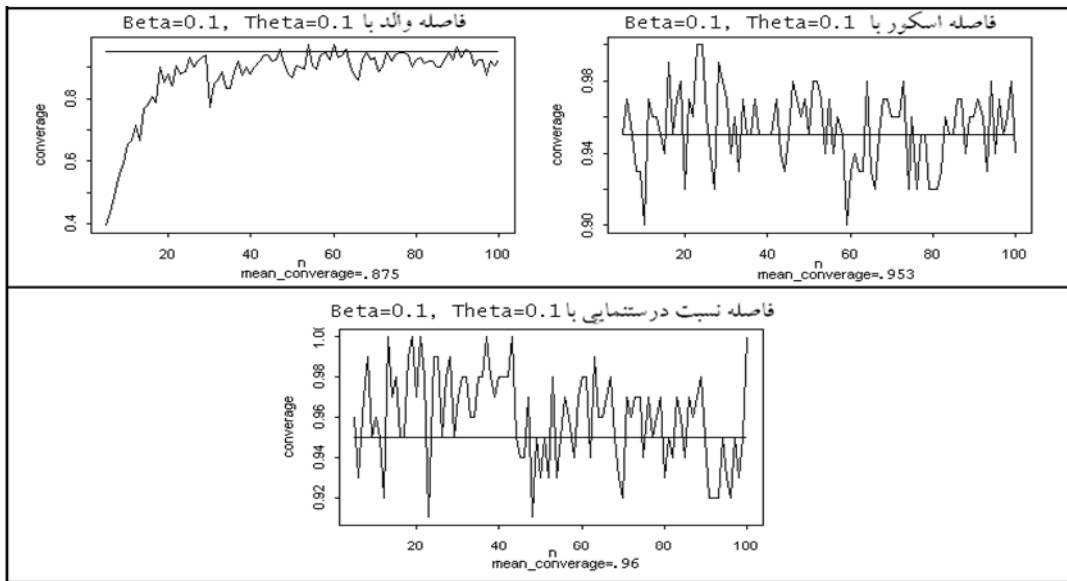


احتمال‌های پوشش در بازه‌های مختلف شبیه‌سازی و با یکدیگر مورد مقایسه قرار می‌گیرند. برآورد احتمال پوشش برای بازه‌های معرفی شده در دو توزیع پواسن و آبل برای حجم نمونه متغیر و پارامتر ثابت در سطح اطمینان اسمی ۹۵ درصد به صورت (نمودارهای ۱ و ۲) شبیه‌سازی شده است.

یکسان  $m = 1000$  بار تولید می‌کنیم، به نسبت بازه‌هایی که پارامتر را شامل است، برآورد احتمال پوشش واقعی  $\hat{P}$  گویند. شبیه‌سازی به دو روش انجام می‌گیرد. در روش اول برای یک پارامتر ثابت از توزیع، با حجم نمونه تصادفی به اندازه  $n = 5, 6, \dots, 100$ ، و در روش دوم با حجم نمونه ثابت و به ازای پارامترهای مختلف،



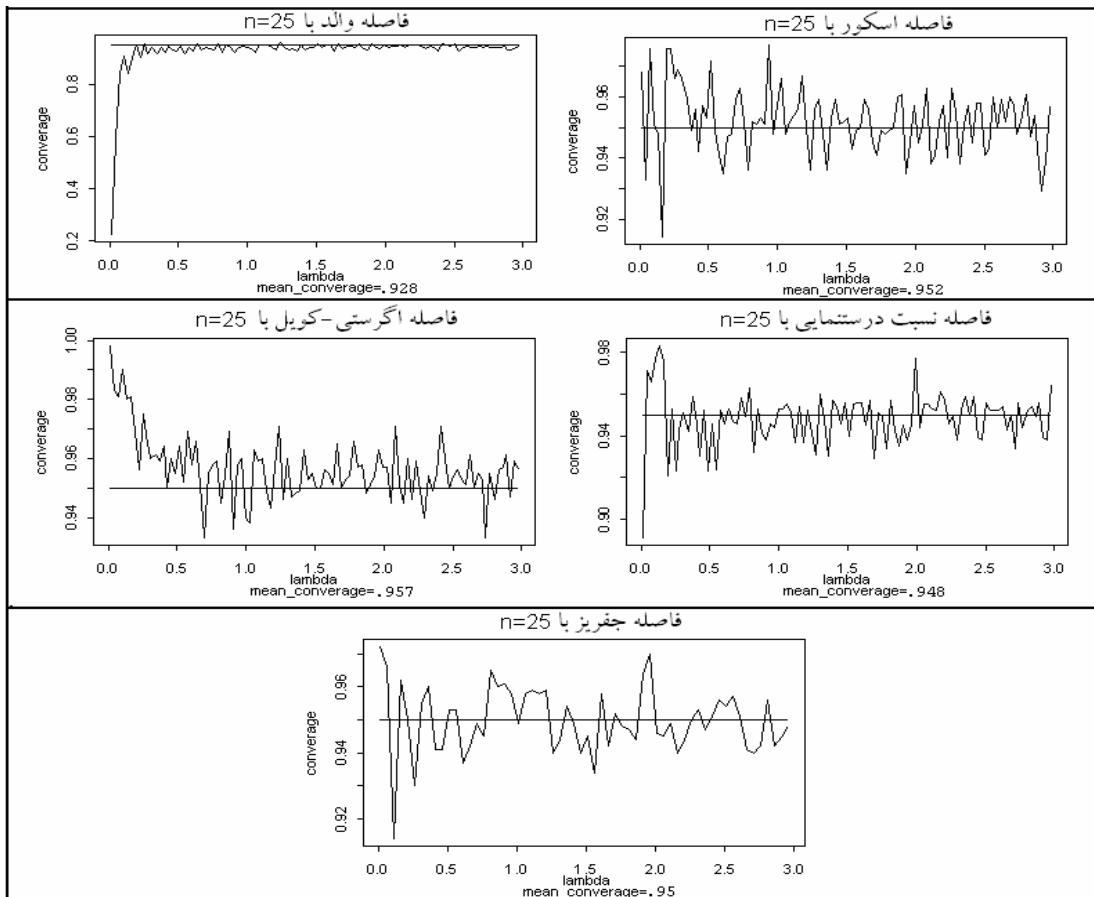
نمودار ۱- احتمال پوشش در توزیع پواسن با نمونه تصادفی متغیر و پارامتر ثابت



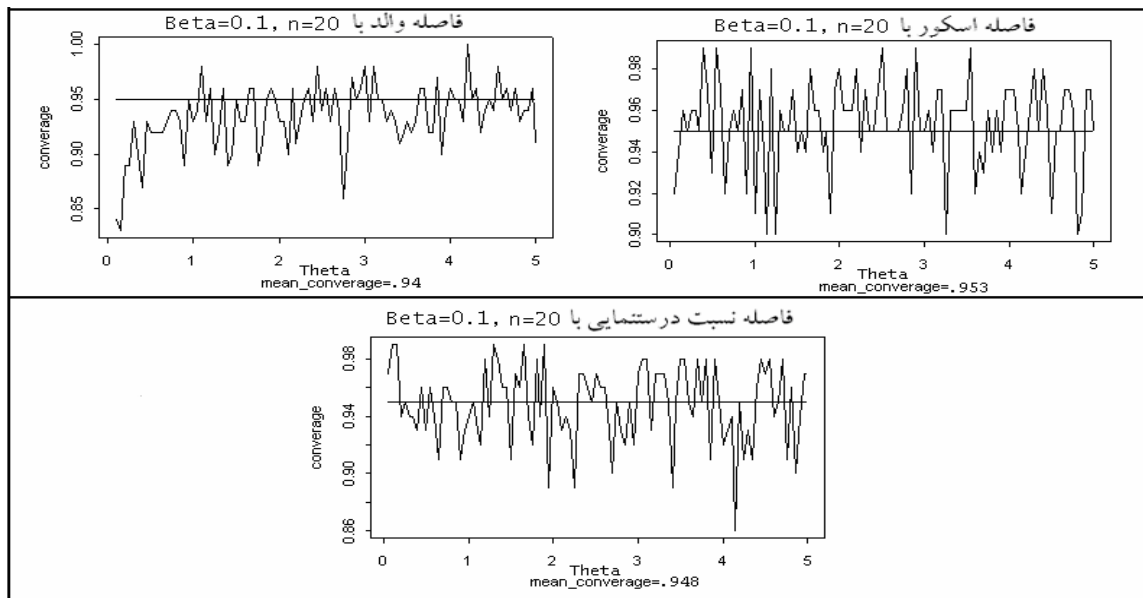
نمودار ۲- احتمال پوشش در توزیع آبل با نمونه تصادفی متغیر و پارامتر ثابت

آبل در سطح اطمینان اسمی ۹۵ درصد به صورت (نمودارهای ۳ و ۴) شبیه‌سازی شده است.

همچنین برآورد احتمال‌های پوشش به ازای یک اندازه نمونه ثابت و پارامترهای مختلف برای دو توزیع پواسن و



نمودار ۳- احتمال پوشش در توزیع پواسن با نمونه تصادفی ثابت و پارامتر متغیر



نمودار ۴- احتمال پوشش در توزیع آبل با نمونه تصادفی ثابت و پارامتر متغیر

توزیع پواسن و آبل مورد بررسی قرار گرفتند. همچنین این بازه‌ها از نظر طول مورد انتظار و برآورد احتمال پوشش، با یکدیگر مقایسه شده‌اند. با توجه به بررسی‌های انجام گرفته، بازه والد در برخی موارد طول کم‌تری دارد ولی برآورد احتمال پوشش آن خیلی کم‌تر از سطح اطمینان اسمی است و به دلیل محاسبه سریع کاربرد زیادی دارد. بازه اسکور دارای طول مناسب و احتمال پوششی در سطح اطمینان اسمی است و سریع نیز محاسبه می‌شود. بازه اگرستی- کول هر چند سریع محاسبه می‌شود، ولی برآورد احتمال پوشش آن بیش از سطح اطمینان اسمی است و فقط برای خانواده توزیع‌های نمایی طبیعی با تابع واریانس درجه دوم قابل محاسبه می‌باشد. بازه‌های نسبت درست‌نمایی و جفریز علاوه بر کم‌ترین طول مورد انتظار، دارای نزدیک‌ترین برآورد احتمال پوشش به سطح اطمینان اسمی است، ولی به علت محاسبه پیچیده آن‌ها کم‌تر مورد استفاده قرار می‌گیرد (در برخی موارد تابع توزیع پسین جفریز به صورت مناسبی به دست نمی‌آید). تقریباً برآورد احتمال پوشش تمام بازه‌ها با افزایش اندازه نمونه به سطح اطمینان اسمی نزدیک‌تر

با توجه به نمودارهای شبیه‌سازی شده نکات زیر استنباط می‌شود:

- برآورد احتمال پوشش برای بازه والد کم‌تر از سطح اطمینان اسمی، برای دو بازه اسکور و نسبت درست‌نمایی در سطح اطمینان اسمی، برای بازه اگرستی- کول بیش از سطح اطمینان اسمی و بازه جفریز نزدیک‌ترین برآورد احتمال پوشش را به سطح اطمینان اسمی دارا می‌باشد.
- برآورد احتمال پوشش تقریباً در هر دو توزیع با افزایش اندازه نمونه به سطح اطمینان اسمی نزدیک‌تر می‌گردد.
- تغییر مقادیر مختلف پارامتر نیز روی برآورد احتمال پوشش تأثیر می‌گذارد. برآورد احتمال با افزایش پارامتر به سطح اطمینان اسمی نزدیک‌تر می‌شود.

### نتیجه‌گیری

در این مقاله، بازه‌های اطمینان والد، اسکور، اگرستی- کول، نسبت درست‌نمایی و جفریز معرفی و برای دو

- [6] Brown, L.D., Cai, T. and DasGupta, A., Interval estimation in exponential families, *Statistics Sinica*, 13 (2003) 19-49.
- [7] Brown, L.D., Cai, T. and DasGupta, A., Confidence interval for a binomial proportion and Edgeworth expansions, Technical Report # 99-18, Department of Statistical, Purdue University, (1999b).
- [8] Brown, L.D., Cai, T. and DasGupta, A., Interval estimation for a binomial proportion (with discussion), *Statistical Science*, 16 (2001) 101-133.
- [9] Brown, L.D., Cai, T. and DasGupta, A., Confidence intervals for a binomial proportion and Edgeworth expansions, *The Annals of Statistics*, 30 (2002) 160-201.
- [10] Reiczigel, J., Confidence intervals for the binomial parameter: some new considerations, *Statistics in Medicine*, 22 (2003) 611-621.
- [11] Zhou, X., Tsao, M. and Qin, Q., New intervals for the difference between two independent binomial proportions, *Journal of Statistical Planning and Inference*, 123 (2004) 97-115.
- [12] Brown, L. and Li, X., Confidence intervals for two sample binomial distribution, *Journal of Statistical Planning and Inference*, 130 (2005) 359-375.
- [13] Wang, H., Exact confidence coefficients of confidence intervals for a binomial proportion, *Statistics Sinica*, 17 (2007) 361-368.
- [14] Roths, S. and Tebbs, J., Revisiting Beal's confidence intervals for the difference of

می‌گردد. بنابراین در حالت کلی، دو بازه اسکور و نسبت درست‌نمایی، بازه‌های مناسبی هستند، هر چند در برخی موارد محاسبه آن‌ها با مشکلاتی همراه است.

### تشکر و قدردانی

نویسندگان مقاله از نظرات و پیشنهادات اساسی داوران محترم که موجب بهبود کیفیت این مقاله گردیده و از حمایت‌های قطب علمی داده‌های ترتیبی و فضایی دانشگاه فردوسی مشهد کمال تشکر و قدردانی را دارند.

### مراجع

- [1] Santner, T.J., A note on teaching binomial confidence intervals, *Teaching Statistics*, 20 (1998) 20-23.
- [2] Agresti, A. and Coull, B.A., Approximate is better than exact for interval estimation of binomial proportions, *The American Statistician*, 52 (1998) 119-126.
- [3] Cai, T. (2003). One-sided confidence intervals in discrete distributions. Technical Report, Department of Statistics, University of Pennsylvania.
- [4] Lui, K.J., Population proportion or prevalence, John Wiley and Sons, Ltd., (2004).
- [5] Sun, X., Zhou, X. and Wang, J., Confidence intervals for the scale parameter of exponential distribution based on Type II doubly censored samples, *Journal of Statistical Planning and Inference*, 138 (2008) 2045-2058.

- in finite samples: one-parameter exponential families, where the MLE is a sufficient statistic, and location scale families, where there are exhaustive ancillary statistics, *The Matter and the Wiley collections*, Section 42 (1934).
- [23] Lehmann, E.L., *Testing statistical hypothesis*, Wiley, New York, (1959).
- [24] Barndorff-Nielsen, O.E., *Information and Exponential Families in Statistical Theory*, Wiley, New York, (1979).
- [25] Morris, C.N., Natural exponential families with quadratic variance functions, *The Annals of Statistics*, 10 (1982) 65-80.
- [26] Morris, C.N., Natural exponential families with quadratic variance functions: statistical theory, *The Annals of Statistics*, 11 (1983) 515-529.
- [27] Letac, G., *Lecture notes on natural exponential families and their variance functions*, Monografia de mathematica, Instituto de mathematica Pura E Aplicada, (1992).
- [28] Consul, P.C., Some characterizations for the exponential class of distributions, *IEEE transactions on reliability*, 44 (1995) 403-407.
- [29] Rao, C.R., *Linear Statistical Inference and its Applications*. Wiley, New York. (1973).
- [30] Serfling, R.J., *Approximation theorems of mathematical statistics*, Wiley, New York, (1980).
- [31] Jeffreys, H.S., *Theory of probability*, 3rd edition, Oxford University Press, (1961).
- two binomial proportions, *Communications in Statistics (Theory and Methods)*, 35 (2006) 1593-1609.
- [15] Arefi, M., Mohtashami Borzadaran, G.R. and Vaghei, Y., A note on interval estimation for the mean of inverse Gaussian distribution, *Statistics and Operations Research Transactions (SORT)*, 32, 1 (2008) 49-56.
- [16] Consul, P.C., *Generalized Poisson Distributions, Properties and Applications*, Marcel Dekker Inc., New York, Basel., (1989).
- [17] Consul, P.C. and Jain, G.C., A generalization of the Poisson distribution, *Technometrics*, 15, 4 (1973) 791-799.
- [18] Consul, P.C. and Shoukri, M.M., The generalized Poisson distribution when the sample mean is larger than the sample variance, *Communications in Statistics-Simulation and Computation*, 14 (1985) 1533-1547.
- [19] Vernic, R., On the bivariate generalized Poisson distribution, *ASTIN Bulletin*, 27, 1 (1997) 23-31.
- [20] Ambagasphiya, R.S. and Balakrishnan, N., On the compound generalized Poisson distributions, *ASTIN Bulletin*, 24 (1994) 255-263.
- [21] Letac, G. and Mora, M., Natural real exponential families with cubic variance functions, *The Annals of Statistics*, 18 (1990) 1-37.
- [22] Fisher, R.A., Concerns two situations when fully efficient estimation is possible