

مقایسه مونت کارلویی فاصله‌های تحمل تقریبی برای توزیع پواسون

*مهران نقی‌زاده قمی، آریتا نوروزی فیروز؛

دانشگاه مازندران، گروه آمار

پذیرش ۹۵/۵/۲

دریافت ۹۴/۳/۱۹

چکیده

بسیاری از محققان به مسئله یافتن فاصله‌های تحمل توجه دارند و به‌طور گسترده‌ای در بسیاری از زمینه‌های مختلف آماری مانند زیست‌شناسی، اقتصاد، مهندسی و کنترل کیفیت به‌کار می‌برند. فاصله تحمل، فاصله‌ای تصادفی است که نسبت مشخصی از جامعه را با سطح اطمینان مشخص پوشش می‌دهد. در این مقاله، به مقایسه فاصله‌های تحمل تقریبی برای متغیر تصادفی پواسون می‌پردازیم. فاصله‌های تحمل تقریبی براساس فاصله‌های اطمینان تقریبی برای پارامتر توزیع پواسون شامل فاصله والد، فاصله والد با تصحیح پیوستگی، فاصله امتیاز، فاصله با تثبیت واریانس، فاصله با تثبیت واریانس مرکزی-شده و فاصله فریمن و توکی ساخته می‌شوند. احتمال‌های پوشش و طول مورد انتظار فاصله‌های تحمل ارائه شده با انجام شبیه‌سازی مونت کارلو ارزیابی می‌شود. فاصله‌های تحمل ارائه شده در قالب یک مثال کاربردی استفاده می‌شوند.

واژه‌های کلیدی: توزیع پواسون، فاصله اطمینان، فاصله تحمل

مقدمه

یک فاصله اطمینان، براساس نمونه‌ای تصادفی برای برآورد پارامتر نامعلوم جامعه مانند میانگین یا واریانس به‌کار می‌رود. یک فاصله پیش‌گویی، براساس نمونه‌ای تصادفی برای پیش‌گویی مقداری از مشاهده آینده در جامعه به‌کار می‌رود. فاصله تحمل، فاصله‌ای تصادفی است که با ضریب اطمینان مشخصی، نسبتی از جامعه بررسی شده را در برمی‌گیرد. برای مثال، ۹۰ درصد اطمینان داریم که حداقل ۸۰ درصد واحدهای جامعه در فاصله (L, U) قرار دارند. فاصله (L, U) را یک فاصله تحمل $(0/9 \text{ و } 0/8)$ گویند.

همانند فاصله‌های اطمینان و پیش‌گویی، فاصله‌های تحمل را می‌توان برای توزیع‌های گسسته و پیوسته به‌کار برد. برای مثال از توزیع پیوسته، فرض کنید $x_1, \dots, x_n$ طول عمر n قطعه دارای توزیع نمایی با پارامتر θ باشند. اگر فاصله تحمل دوطرفه $(0/95 \text{ و } 0/9)$ را با (L, U) نشان دهیم این بدین معنی است که با اطمینان ۹۵ درصد حداقل ۹۰ درصد از قطعات دارای طول عمری بین (L, U) هستند. در حالت گسسته، برای مثال فرض کنید نمونه‌ای تصادفی از $n=1000$ لامپ برای بازرسی انتخاب شود که $x=10$ تای آن‌ها معیوب است. فرض کنید لامپ‌ها در بسته‌های ۵۰ تایی باشند و مهندس کارخانه تولیدکننده لامپ‌ها علاقمند است مشخص کند که با ۹۵ درصد اطمینان، حداقل ۹۰ درصد از چنین بسته‌هایی لامپ‌های معیوب بین L و U دارند. یا در مثالی دیگر، فرض کنید مهندسی نیاز دارد تا حد بالای تحمل U برای تعداد خاموش شدن ناگهانی سیستمی پیچیده

به‌دست آورد. برای مثال با ۹۵ درصد اطمینان، حداقل ۹۰ درصد از سیستم‌هایی که در طول یک سال کار می‌کنند دارای تعداد خاموش شدن ناگهانی حداکثر U هستند.

فاصله‌های تحمل ابزارهای مهمی در شاخه‌های مختلف از جمله کنترل کیفیت و مهندسی هستند. به‌عنوان مثال، در کنترل کیفیت، در فرایند تولید یک کالا، حدود مشخصه‌های فنی بالا^۱ (USL) و پایین^۲ (LSL) از قبل توسط مهندسان تعیین می‌شوند که این حدود باید مشخصه‌های محصولات تولیدی را در برگیرند. برای آن‌که متوجه شویم که چه نسبتی از این محصولات، در حدود مشخصه‌های فنی قرار می‌گیرند، یافتن یک فاصله‌ی تحمل (L, U) بر اساس نمونه‌ای از محصولات تولید شده می‌تواند مفید باشد. چنان‌چه فاصله‌ی تحمل درون حدود مشخصه‌های فنی قرار بگیرد یعنی $(L, U) \subset (LSL, USL)$ ، آن‌گاه کالای مورد نظر پذیرفته می‌شود و با اطمینانی مشخص می‌توان گفت که حداقل درصد مشخصی از محصولات، مطابق استانداردهای از پیش تعیین شده تولید شده‌اند. بنابراین، موضوع یافتن فواصل تحمل اهمیت زیادی دارند.

اولین تحقیقات در رابطه با فاصله‌های تحمل را ویلکس [۱] با تعیین حجم نمونه در مسئله فاصله‌ی تحمل آغاز کرد. پس از آن، تحقیقات گسترده‌ای درباره فواصل تحمل و کاربردهای آن انجام شده است. در ادبیات تحقیق، به مسئله یافتن فاصله‌های تحمل برای توزیع‌های پیوسته نسبت به توزیع‌های گسسته بیش‌تر پرداخته شده است. از جدیدترین کارهای انجام‌شده در زمینه فاصله‌های تحمل برای توزیع‌های پیوسته می‌توان به فرناندز [۲] و نقی‌زاده‌قمی و کیاپور [۳] اشاره کرد که به‌ترتیب فاصله‌های تحمل برای توزیع‌نمایی براساس داده‌های سانسور شده و کوتاه‌ترین فاصله‌های تحمل برای توزیع‌نمایی را براساس داده‌های رکوردی به‌دست آوردند. هم‌چنین میرمصطفائی و همکاران [۴] به محاسبه حدود تحمل برای زمان‌های تعمیر مینیمال سیستمی سری با مؤلفه‌های با طول عمر رایلی پرداختند. در توزیع نرمال چندمتغیره، کریشنامورسی و موندال [۵] روشی برای بهبود فاصله‌های تحمل ارائه کردند. بودج و متیو [۶] ناحیه‌های تحمل بیضوی و و دونگن و متیو [۷] ناحیه‌های تحمل مرجع و مرکزی را برای توزیع نرمال چندمتغیره بررسی کردند. در توزیع‌های گسسته، هان و چندرا [۸] مسئله یافتن فواصل تحمل را در توزیع‌های دو جمله‌ای و پواسن بررسی کردند. وانگ و تسونگ [۹] فاصله‌ی تحمل با بهبود احتمالات پوشش را در توزیع‌های دوجمله‌ای و پواسن بررسی کردند. کریشنامورسی و همکاران [۱۰] یک روش تقریبی ساده برای ساختن فواصل تحمل توزیع‌های دوجمله‌ای و پواسن ارائه دادند. یونگ [۱۱] به مقایسه فاصله‌های تحمل برای توزیع دوجمله‌ای منفی پرداخت. نقی‌زاده‌قمی و همکاران [۱۲] فاصله‌ی تحمل تقریبی برای توزیع پواسن-لیندلی ارائه کردند.

در این مقاله، به مقایسه فاصله‌های تحمل تقریبی برای متغیر پواسن پرداخته می‌شود. ابتدا، تعریف فاصله‌های تحمل برای توزیع‌های گسسته و هم‌چنین روش عمومی ساختن فاصله‌های تحمل بیان می‌شود. سپس، فاصله‌های اطمینان تقریبی برای پارامتر توزیع پواسن ارائه می‌شوند. مقایسه فاصله‌های تحمل تقریبی با استفاده از معیارهای احتمال پوشش و طول مورد انتظار به‌کمک شبیه‌سازی مونت‌کارلو انجام می‌شود. در پایان از مثالی واقعی برای محاسبه فاصله‌های تحمل استفاده می‌شود.

1. Upper Specification Limit

2. Lower Specification Limit

نحوه ساختن فاصله‌های تحمل برای توزیع‌های گسسته

برای تعریف فاصله تحمل برای یک توزیع گسسته، فرض کنید X نمونه‌ای تصادفی از توزیعی گسسته و X دارای توزیع یک‌سان مستقل از X باشد. یک فاصله تحمل دوطرفه $[L(X), U(X)]$ با میزان پوشش p و سطح اطمینان $1-\alpha$ بدین صورت تعریف می‌شود

$$P_X\{P_X(L(X) \leq X \leq U(X) | X) \geq p\} \geq 1-\alpha, \quad (1)$$

که در آن $p, \alpha \in (0, 1)$ و معمولاً برای سادگی به آن یک فاصله تحمل دوطرفه $(p, 1-\alpha)$ می‌گویند. عبارت سمت چپ رابطه (۱) احتمال پوشش و p سطح پوشش نامیده می‌شود. در واقع احتمال پوشش این است که فاصله تحمل حداقل نسبت p از جامعه را در برگیرد یا پوشش دهد. در توزیع‌های پیوسته رابطه (۱) به صورت تساوی است، اما در توزیع‌های گسسته، احتمال پوشش، حداقل $1-\alpha$ است.

یک حد بالایی تحمل $(p, 1-\alpha)$ که با $U_1(X)$ نشان می‌دهیم بدین صورت تعریف می‌شود:

$$P_X\{P_X(X \leq U_1(X) | X) \geq p\} \geq 1-\alpha.$$

اگر k_p چندک مرتبه p از توزیع مورد نظر باشد، آن‌گاه حد بالایی تحمل یک‌طرفه $(p, 1-\alpha)$ در واقع همان حد بالایی اطمینان برای چندک k_p است، یعنی $P_X(k_p \leq U_1(X)) \geq 1-\alpha$. حد پایین تحمل $(p, 1-\alpha)$ نیز به‌طور مشابه تعریف می‌شود. فاصله تصادفی $[L_e(X), U_e(X)]$ که در رابطه (۲) صدق کند:

$$P_X(L_e(X) \leq k_{\frac{1-p}{2}}; k_{\frac{1+p}{2}} \leq U_e(X)) \geq 1-\alpha, \quad (2)$$

یک فاصله تحمل $(p, 1-\alpha)$ با کنترل چندک‌ها در هر دو دم توزیع نامیده می‌شود. اوون [۱۳] اولین بار چنین فواصلی را برای توزیع نرمال به کار برد. کریشنامورسی [۱۴] آن را یک فاصله تحمل با دم‌های برابر نامید.

فاصله $[L_e(X), U_e(X)]$ به گونه‌ای ساخته می‌شود که به‌طور هم‌زمان حداکثر نسبت $\frac{1-p}{2}$ از جامعه کم‌تر از

$L_e(X)$ باشد و حداکثر نسبت $\frac{1-p}{2}$ از جامعه بزرگ‌تر از $U_e(X)$ باشد.

فرض کنید $F_X(x | \theta)$ تابع توزیع متغیر تصادفی گسسته X با پارامتر θ و همچنین دارای خاصیت به‌طور تصادفی صعودی^۱ باشد، یعنی این که برای هر t داشته باشیم:

$$\Pr(X > t\hat{\theta}_1) \geq \Pr(X > t\hat{\theta}_2) \quad \forall \theta_1 > \theta_2.$$

در نتیجه چندک مرتبه p یعنی $k_p(\theta)$ که بدین صورت تعریف می‌شود:

$$k_p(\theta) = \inf\{x \in Z^+ : p \leq F_X(x | \theta)\},$$

یعنی کوچک‌ترین مقداری که برای آن $P(X \leq k_p(\theta) | \theta) \geq p$ ، تابعی صعودی نسبت به θ است. بنابراین اگر θ_u یک حد بالای اطمینان $100(1-\alpha)$ درصدی برای θ باشد، یعنی $P(\theta \leq \theta_u) \geq 1-\alpha$ آنگاه $k_p(\theta_u)$ که به‌عنوان کوچک‌ترین مقداری تعریف می‌شود که برای آن داشته باشیم $P(X \leq k_p(\theta_u) | \theta_u) \geq p$ ، یک حد بالای اطمینان $100(1-\alpha)$ درصدی برای $k_p(\theta)$ است. در نتیجه $k_p(\theta_u)$ یک حد بالای تحمل $(p, 1-\alpha)$ برای

1. Stochastically increasing

توزیع خواهد بود. به‌طور مشابه اگر θ_l یک کران پایین اطمینان $100(1-\alpha)$ درصدی برای θ باشد آن‌گاه $k_{1-p}(\theta_l)$ که به‌عنوان بزرگ‌ترین مقداری تعریف می‌شود که برای آن داشته باشیم:

$$P(X \geq k_p(\theta_l) | \theta_l) \geq p,$$

یک کران پایین اطمینان $100(1-\alpha)$ درصدی برای $k_p(\theta)$ و در نتیجه یک کران پایین تحمل $(p, 1-\alpha)$ برای توزیع خواهد بود. اگر (θ_l, θ_u) یک فاصله اطمینان $100(1-\alpha)$ درصدی برای θ باشد، آن‌گاه به‌طور هم‌زمان نامساوی‌های $k_{\frac{1-p}{2}}(\theta_l) \leq k_{\frac{1-p}{2}}(\theta_u)$ و $k_{\frac{1+p}{2}}(\theta) \leq k_{\frac{1+p}{2}}(\theta_u)$ با احتمال $1-\alpha$ برقرار هستند. بنابراین با توجه به رابطه (۲) فاصله $[k_{\frac{1-p}{2}}(\theta_l), k_{\frac{1+p}{2}}(\theta_u)]$ یک فاصله تحمل $(p, 1-\alpha)$ با دمه‌های برابر است.

توزیع پواسون

فرض کنید متغیر تصادفی X دارای توزیع پواسون با پارامتر λ با تابع جرم احتمال بدین صورت باشد:

$$f(x|\lambda) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x=0,1,2,\dots$$

براساس نمونه تصادفی $X_1, \dots, X_n$ از توزیع پواسون با پارامتر λ ، برآوردگر ماکسیمم درست‌نمایی $\hat{\lambda}$ برابر است با $\bar{X} = \hat{\lambda}$. در زیربخش بعدی، فاصله‌های اطمینان تقریبی برای پارامتر λ توزیع پواسون ارائه می‌شوند.

فاصله‌های اطمینان تقریبی برای پارامتر λ توزیع پواسون

در این بخش فاصله‌های اطمینان تقریبی برای پارامتر λ توزیع پواسون را فهرست می‌کنیم که در بارکر [۱۵] موجود است.

الف. فاصله اطمینان با نمونه‌های بزرگ^۱ (LS)

فاصله اطمینان والد با تقریب توزیع نرمال برای λ بدین صورت است:

$$\left[\hat{\lambda} - c \sqrt{\frac{\hat{\lambda}}{n}} \hat{X} + c \sqrt{\frac{\hat{\lambda}}{n}} \right]$$

که در آن $c = z_{1-\frac{\alpha}{2}}$ چندک $(1-\frac{\alpha}{2})$ ام توزیع نرمال استاندارد است.

ب. فاصله اطمینان والد با تصحیح پیوستگی^۲ (CC)

با استفاده از تصحیح پیوستگی در رابطه (۳) فاصله اطمینان $100(1-\alpha)$ درصدی برای λ عبارت است از:

$$\left[\hat{\lambda} - c \sqrt{\frac{\hat{\lambda} + 0.5}{n}}, \hat{\lambda} + c \sqrt{\frac{\hat{\lambda} + 0.5}{n}} \right]$$

ج. فاصله اطمینان امتیاز^۳ (SC)

1. Large Sample confidence interval

2. Wald confidence interval with continuity correction

3. Score confidence interval

با در نظر گرفتن کمیت نرمال استاندارد مجانبی $\frac{\hat{\lambda} - \lambda}{\sqrt{\lambda/n}}$ فاصله اطمینان $100(1 - \alpha)$ درصدی برای λ

بدین صورت است:

$$\left[\hat{\lambda} + \frac{c^2}{2n} - c \frac{\sqrt{4\hat{\lambda} + \frac{c^2}{n}}}{\sqrt{4n}}, \hat{\lambda} + \frac{c^2}{2n} + c \frac{\sqrt{4\hat{\lambda} + \frac{c^2}{n}}}{\sqrt{4n}} \right]$$

د. فاصله اطمینان با تثبیت واریانس^۱ (VS)

این فاصله از کمیت نرمال استاندارد مجانبی $\frac{\sqrt{\hat{\lambda}} - \sqrt{\lambda}}{\sqrt{\frac{1}{4n}}}$ به دست می‌آید. با معکوس کردن این کمیت، این فاصله اطمینان به دست می‌آید:

$$\left[\hat{\lambda} + \frac{c^2}{4n} - c \sqrt{\frac{\hat{\lambda}}{n}}, \hat{\lambda} + \frac{c^2}{4n} + c \sqrt{\frac{\hat{\lambda}}{n}} \right]$$

ه. فاصله اطمینان با تثبیت واریانس مرکزی شده^۲ (RVS)

این فاصله از کمیت نرمال استاندارد مجانبی $\frac{\sqrt{\hat{\lambda} + a} - \sqrt{\lambda + a}}{\sqrt{\frac{1}{4n}}}$ برای مقدار ثابت مثبت a به دست می‌آید.

انسکوب [۱۶] نشان داد که فاصله زیر یک فاصله اطمینان تقریبی $100(1 - \alpha)$ درصدی برای λ است:

$$\left[\hat{\lambda} + \frac{c^2}{4n} - c \sqrt{\frac{\hat{\lambda} + \frac{3}{8}}{n}}, \hat{\lambda} + \frac{c^2}{4n} + c \sqrt{\frac{\hat{\lambda} + \frac{3}{8}}{n}} \right]$$

و. فاصله اطمینان فریمن و توکی^۳ (FT)

فریمن و توکی [۱۷] نشان دادند که این کمیت:

$$\sqrt{n} \left\{ [\sqrt{\hat{\lambda}} + \sqrt{\hat{\lambda} + 1}] - [\sqrt{\lambda} + \sqrt{\lambda + 1}] \right\}$$

دارای توزیع نرمال استاندارد مجانبی است. بنابراین یک فاصله اطمینان $100(1 - \alpha)$ درصدی برای

$\sqrt{\lambda} + \sqrt{\lambda + 1}$ بدین صورت است:

$$\left[\sqrt{\hat{\lambda}} + \sqrt{\hat{\lambda} + 1} - c \sqrt{\frac{1}{n}}, \sqrt{\hat{\lambda}} + \sqrt{\hat{\lambda} + 1} + c \sqrt{\frac{1}{n}} \right]$$

معکوس تابع $f(u) = \sqrt{u} + \sqrt{1+u}$ بدین صورت است:

1. Variance stabilizing confidence interval
2. Recentered variance stabilizing confidence interval
3. Freeman and Tukey confidence interval

$$g(v) = \left(\frac{v^2 - 1}{2v} \right)^2, \quad v \geq 1$$

بنابراین یک فاصل G اطمینان $100(1-\alpha)$ درصدی برای λ بدین صورت است:

$$\left[g \left(\sqrt{\hat{\lambda}} + \sqrt{\hat{\lambda} + 1} - c \sqrt{\frac{1}{n}} \right), g \left(\sqrt{\hat{\lambda}} + \sqrt{\hat{\lambda} + 1} + c \sqrt{\frac{1}{n}} \right) \right]$$

فاصله‌های تحمل تقریبی برای توزیع پواسون

فرض کنید $X_1, \dots, X_n$ یک نمونه تصادفی از توزیع پواسون با پارامتر λ باشد. نشان می‌دهیم که توزیع پواسون دارای خاصیت به‌طور تصادفی صعودی نسبت به پارامتر λ است. از آن‌جاکه اثبات مستقیم کار دشواری است، ابتدا خاصیت نسبت درست‌نمایی یکنوا^۱ را تعریف و با استفاده از یک لم در [۱۸] کافی است که خاصیت نسبت درست‌نمایی یکنوا را برای توزیع پواسون اثبات کنیم.

تعریف ۱. فرض کنید $\mathbf{X} = (X_1, \dots, X_n)$ یک نمونه تصادفی از یک توزیع احتمال $\theta \in \Theta \subset R$ باشد و برای هر $\theta_1, \theta_2 \in \Theta$ که $\theta_1 \neq \theta_2$ داشته باشیم $f_{\theta_1}(\mathbf{x}) \neq f_{\theta_2}(\mathbf{x})$ می‌گوییم خانواده $\{f_{\theta} : \theta \in \Theta \subset R\}$ دارای خاصیت نسبت درست‌نمایی یکنوا در آماره $T(\mathbf{X})$ است اگر برای $\theta_1 > \theta_2$ نسبت $\frac{f_{\theta_1}(\mathbf{x})}{f_{\theta_2}(\mathbf{x})}$ یک تابع غیرنزولی از $T(\mathbf{X})$ روی مجموعه‌ای که این نسبت وجود دارد، باشد.

قسمت ب لم زیرمنتسب به [۱۸] رابطه بین خاصیت نسبت درست‌نمایی یکنوا و خاصیت به‌طور تصادفی صعودی را نشان می‌دهد.

لم ۱. فرض کنید خانواده $\{f_{\theta} : \theta \in \Theta \subset R\}$ خانواده‌ای از توزیع‌ها باشند که دارای خاصیت نسبت درست‌نمایی یکنوا در $T(\mathbf{X}) = X$ باشد.

الف. اگر f_{θ} تابعی غیر نزولی از X باشد آن‌گاه $E_{\theta}(\psi(X))$ تابعی غیرنزولی از θ است. تابع ψ تعریف نشده است

ب. خانواده توزیع‌های فوق دارای خاصیت به‌طور تصادفی صعودی است.

ج. اگر ψ تابعی باشد که تنها یک بار تغییر علامت دهد، یعنی یک x_0 وجود داشته باشد که به ازای

$$\psi(x) \leq 0, \quad x \leq x_0 \quad \text{و} \quad \psi(x) \geq 0, \quad x \geq x_0$$

در این صورت یک θ_0 وجود دارد که به ازای

$$E_{\theta}(\psi(x)) \leq 0, \quad \theta \leq \theta_0 \quad \text{و} \quad E_{\theta}(\psi(x)) \geq 0, \quad \theta \geq \theta_0$$

به جز این که $E(\psi(X))$ برای هر θ کاملاً

مثبت یا کاملاً منفی باشد.

قضیه ۱. فرض کنید X دارای توزیع پواسون با پارامتر λ باشد. در این صورت X خاصیت نسبت درست‌نمایی یکنوا دارد.

اثبات. به‌ازای $\lambda_1 > \lambda_2$ داریم:

$$\frac{f(x|\lambda_1)}{f(x|\lambda_2)} = e^{-(\lambda_1 - \lambda_2)} \left(\frac{\lambda_1}{\lambda_2} \right)^x$$

نسبت فوق تابعی غیرنزولی نسبت به X است و در نتیجه طبق تعریف ۱ دارای خاصیت نسبت درست‌نمایی یکنوا است.

با استفاده از قضیه ۱ و قسمت ب لم ۱ می‌توان نتیجه گرفت که توزیع پواسون خاصیت به‌طور تصادفی صعودی دارد. بنابراین یک حد بالایی تحمل $(p, 1-\alpha)$ برای توزیع پواسون با پارامتر λ ، کوچکترین مقدار صحیح $k_p(\lambda_u)$ است که برای آن $P(X \leq k_p(\lambda_u) | \lambda_u) \geq p$ ، λ_u یک حد بالایی اطمینان $100(1-\alpha)$ درصدی برای λ است. به‌طور مشابه $k_{1-p}(\lambda_l)$ که بزرگ‌ترین مقداری است که برای آن $P(X \geq k_{1-p}(\lambda_l) | \lambda_l) \geq p$ یک حد پایینی اطمینان $100(1-\alpha)$ درصدی برای λ است، یک حد پایینی تحمل $(p, 1-\alpha)$ برای توزیع پواسون با پارامتر λ است. همچنین فاصله تحمل با دمه‌های برابر $(p, 1-\alpha)$ بدین‌صورت است:

$$\left[\frac{k_{1-p}(\lambda_l) + k_{1+p}(\lambda_u)}{2}, \frac{k_{1-p}(\lambda_l) + k_{1+p}(\lambda_u)}{2} \right],$$

که در آن (λ_l, λ_u) یک فاصله اطمینان دو طرفه $100(1-\alpha)$ درصدی برای λ است.

بررسی احتمال پوشش و طول مورد انتظار فاصله‌های تحمل تقریبی

در این بخش به ارزیابی احتمال‌های پوشش و طول مورد انتظار فاصله‌های تحمل با دمه‌های برابر برای توزیع پواسون می‌پردازیم. فرض کنید (λ_l, λ_u) یک فاصله اطمینان $100(1-\alpha)$ درصدی برای λ و $[L(x), U(x)]$ یک فاصله تحمل $(p, 1-\alpha)$ باشد که در آن $L(x) = k_{1-p}(\lambda_l)$ و $U(x) = k_{1+p}(\lambda_u)$. احتمال پوشش دقیق فاصله تحمل $(p, 1-\alpha)$ با دمه‌های برابر با استفاده از این رابطه به‌دست می‌آید:

$$\sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} \mathbf{I}\{L(x) \leq k_{1-p}(\lambda) \text{ and } k_{1+p}(\lambda) \leq U(x)\},$$

که در آن $\mathbf{I}\{0\}$ تابع نشان‌گر است و $k_q(\lambda)$ چندک q ام توزیع پواسون با پارامتر λ است. طول مورد انتظار فاصله تحمل $(p, 1-\alpha)$ با دمه‌های برابر نیز با استفاده از این رابطه به‌دست می‌آید:

$$\sum_{x=0}^{\infty} [U(x) - L(x)] \frac{e^{-\lambda} \lambda^x}{x!}.$$

برای ارزیابی احتمال‌های پوشش و طول مورد انتظار فاصله‌های تحمل $n = 20$ را در نظر می‌گیریم و احتمال‌های پوشش و طول مورد انتظار را برای فاصله‌های تحمل $(0/90, 0/90)$ ، $(0/95, 0/95)$ و $(0/99, 0/99)$ مقایسه می‌کنیم. برای انجام این کار، مقادیر λ از بازه $(1, 15)$ انتخاب می‌شوند. شکل‌های ۱ تا ۳ نمودار احتمال پوشش فاصله‌های تحمل $(0/90, 0/90)$ ، $(0/95, 0/95)$ و $(0/99, 0/99)$ را نشان می‌دهند. همچنین آماره‌های خلاصه برای احتمال‌های پوشش فاصله‌های تحمل در جدول ۱ نمایش داده شده است. احتمال‌های پوشش ملاکی برای مطلوب بودن عملکرد فاصله‌های تحمل هستند به نحوی که اگر احتمال‌های پوشش فاصله‌های تحمل $(p, 1-\alpha)$ به سطح اطمینان اسمی $(1-\alpha)$ نزدیک باشند، این نشان دهنده عملکرد و رفتار مطلوب فاصله‌های تحمل است. از شکل‌های

۱ تا ۳ مشاهده می‌شود که فاصله‌های تحمل LS ، CC ، و VS رفتار مطلوبی دارند و احتمال‌های پوشش آن‌ها به سطح اطمینان اسمی نزدیک‌تر است. فاصله‌های تحمل RVS و FT عملکرد محافظه‌کارانه^۱ دارند، به این مفهوم که احتمال‌های پوشش آن‌ها بزرگ‌تر از سطح اطمینان اسمی است. با ملاحظه جدول ۱ و با مشاهده میانگین احتمال‌های پوشش نیز می‌توان موارد بیان شده را دریافت، به این ترتیب که فاصله‌های تحمل LS ، CC ، و VS دارای میانگین احتمال پوشش نزدیک به سطح اطمینان اسمی و فاصله‌های تحمل RVS و FT میانگین احتمال پوشش بزرگ‌تر از سطح اطمینان اسمی دارند. از جدول ۱ ملاحظه می‌شود که یک فاصله اطمینان $(1-2\alpha)$ ، فاصله تحمل SC با میانگین احتمال پوشش نزدیک به $(1-\alpha)$ تولید می‌کند. برای مثال با فاصله‌های اطمینان ۹۰،۹۵ و ۹۹ درصد به ترتیب می‌توان فاصله‌های تحمل SC با میانگین احتمال پوشش نزدیک به ۰،۹۵، ۰،۹۷۵ و ۰،۹۹۵ تولید کرد. ملاکی دیگری که برای ارزیابی فاصله‌های تحمل وجود دارد طول مورد انتظار فاصله‌های تحمل است که به‌طور حتم فاصله‌های تحمل با طول مورد انتظار کم‌تر دارای عملکرد بهتری هستند. شکل ۴ نمودار طول مورد انتظار فاصله‌های تحمل و جدول ۲ آماره‌های خلاصه مربوط به طول مورد انتظار آن‌ها را نشان می‌دهند. در مورد نمودارهای طول مورد انتظار فاصله‌های تحمل نیز واضح است که فاصله تحمل LS دارای کم‌ترین و فاصله تحمل RVS دارای بیش‌ترین طول مورد انتظار است.

مثال کاربردی

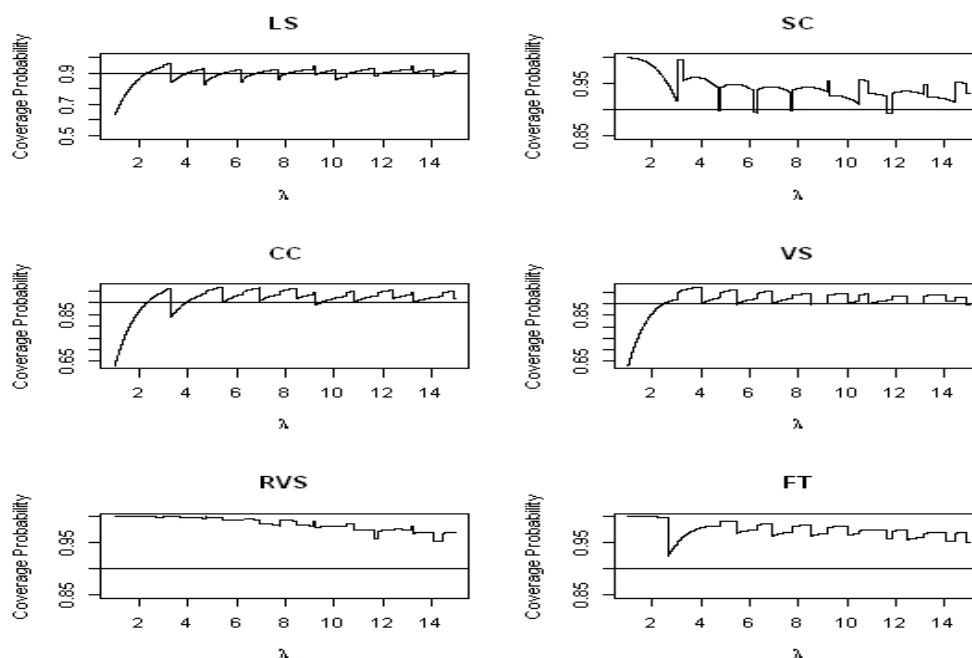
محصول فاقد انطباق یا ناقص، محصولی است که یک یا چند مشخصه آن با مشخصات محصول انطباق نداشته باشد. هریک از این مشخصات که با مشخصات محصول منطبق نباشد باعث ایجاد یک نقص^۲ یا عدم انطباق^۳ می‌شود. جدول ۳ داده‌های مربوط به تعداد نقص‌ها در ۲۶ نمونه ۱۰۰ تایی از صفحات مدار چاپ شده را نشان می‌دهد ([۱۹] صفحه ۲۹۰). معمولاً در فرایند تولید صفحات مدار چاپ شده ۱۶ نوع نقص متفاوت وجود دارد که بیش از ۶۰ درصد از تعداد کل نقص‌ها مربوط به دو نوع نقص می‌شود. عدم لحیم‌کاری کافی و لحیم‌کاری اتصالات سرد. بررسی این دو مورد منعکس‌کننده وجود مشکل در فرایند لحیم‌کاری است که اگر حذف شوند بازده فرایند به‌طور چشم‌گیری افزایش می‌یابد. به‌منظور سهولت، یک واحد بازرسی برابر ۱۰۰ صفحه از مدارهای چاپ شده تعریف شده است. چون در این ۲۶ نمونه تعداد ۵۱۶ نقص مشاهده شده است بنابراین برآورد ماکزیمم درست‌نمایی $\hat{\lambda}$ برابر است با $\hat{\lambda} = \frac{516}{26} = 19.85$. در نتیجه حدود کنترل آزمایشی شامل حد پایین کنترل^۴ (LCL) ، خط مرکز^۵ (CL) و حد بالای کنترل^۶ عبارتند از:

$$\begin{aligned} LCL &= \hat{\lambda} - 3\sqrt{\hat{\lambda}} = 19.85 - 3\sqrt{19.85} = 6.48 \\ CL &= \hat{\lambda} = 19.85 \\ UCL &= \hat{\lambda} + 3\sqrt{\hat{\lambda}} = 19.85 + 3\sqrt{19.85} = 33.22 \end{aligned}$$

1. Conservative
2. Defect
3. Nonconformity
4. Lower control limit
5. Center line
6. Upper control limit

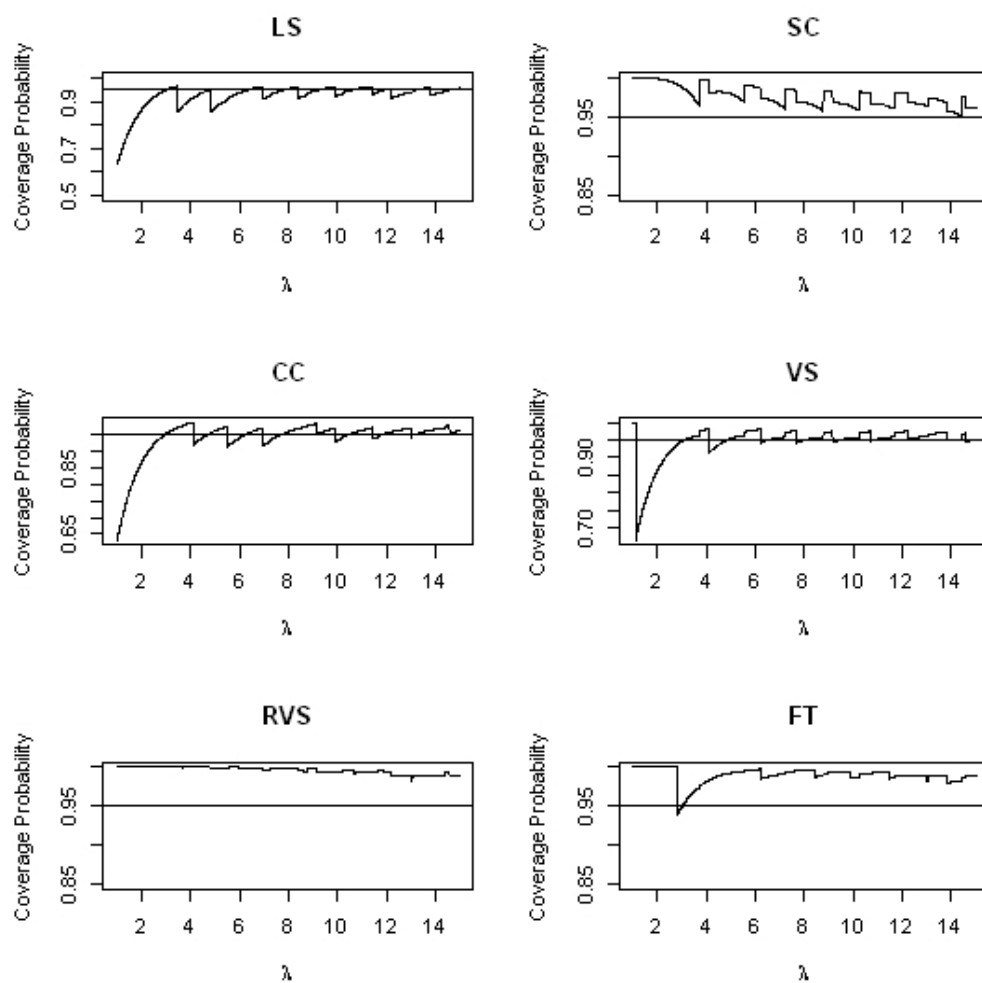
شکل ۵ نمودار کنترل مربوط به تعداد نقص‌ها را نشان می‌دهد که با استفاده از نرم‌افزار مینی‌تب^۱ نسخه ۱۷ رسم شده است. چنان‌که ملاحظه می‌شود نمونه‌های ۶ و ۲۰ خارج از کنترل رسم شده‌اند.

فاصله‌های اطمینان ۹۵ درصد و فاصله‌های تحمل (۰/۹۰,۰/۹۵) برای هر یک از ۶ روش مطرح شده در جدول ۴ ارائه شده است. فاصله تحمل (۰/۹۰,۰/۹۵) برای روش‌های LS و CC برابر (۲۹, ۱۱) و برای روش‌های SC, VS, RVS و FT برابر (۳۰, ۱۱) است. شکل ۶ نمودار هیستوگرام و نمودار کنترل داده‌ها را با نمایش فاصله‌های تحمل (۰/۹۰,۰/۹۵) نشان می‌دهد. فاصله تحمل (۳۰, ۱۱) به این مفهوم است که با اطمینان ۹۵ درصد حداقل ۹۰ درصد از نمونه‌های ۱۰۰ تایی انتخاب شده دارای حداقل تعداد نقص ۱۱ و حداکثر تعداد نقص ۳۰ هستند. می‌توان این فاصله تحمل را برای پیش‌بینی نمونه‌های ۱۰۰ تایی بعدی نیز به کار برد به این ترتیب که ۹۵ درصد اطمینان وجود دارد که حداقل ۹۰ درصد از نمونه‌های ۱۰۰ تایی بعدی دارای تعداد نقص بین ۱۱ و ۳۰ هستند. فرض کنید مهندس طرح، حدود مشخصه‌های فنی را (۳۵, ۷) در نظر بگیرند. چون فاصله تحمل (۰/۹۰,۰/۹۵) یعنی (۳۰, ۱۱) داخل حدود مشخصه‌های فنی قرار دارد بنابراین با ۹۵ درصد اطمینان می‌توان تضمین کرد که حداقل ۹۰ درصد از محصولات مطابق با استانداردهای تعیین شده از سوی مهندسین کارخانه تولید شده‌اند.

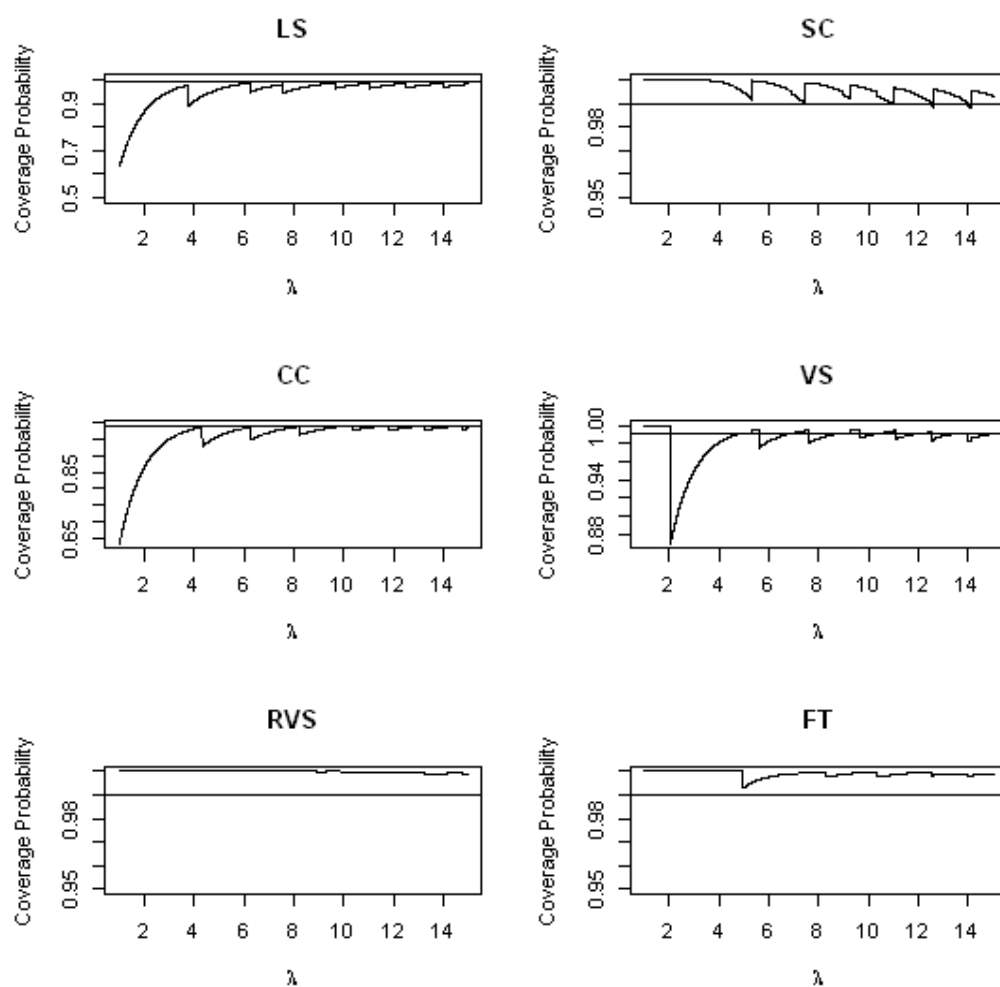


شکل ۱. نمودار احتمال پوشش فاصله‌های تحمل (۰/۹۰,۰/۹۰)

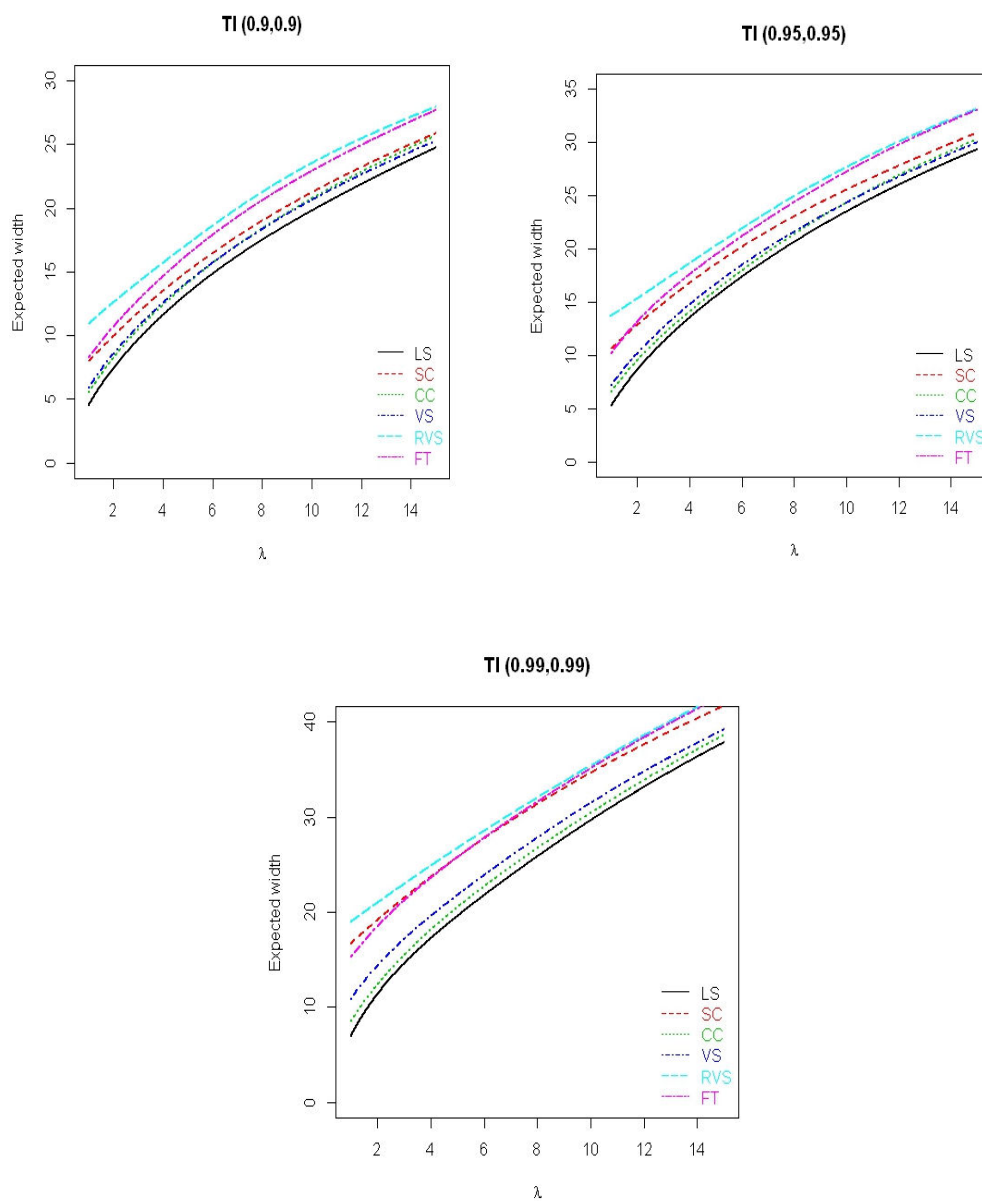
1. Minitab software



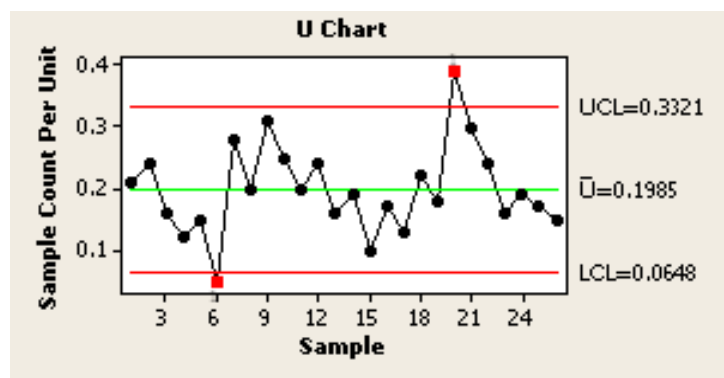
شکل ۲. نمودار احتمال پوشش فاصله‌های تحمل (۰/۹۵، ۰/۹۵)



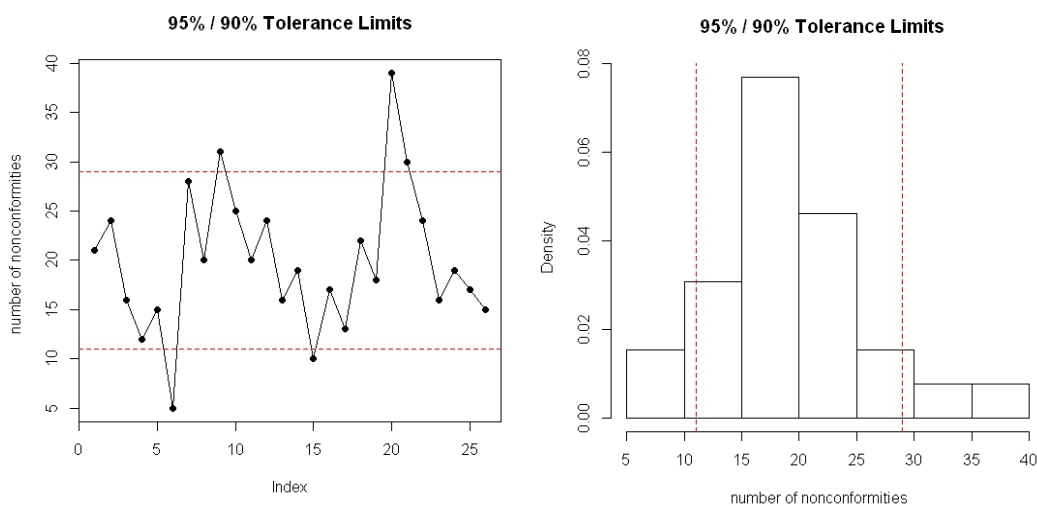
شکل ۳. نمودار احتمال پوشش فاصله‌های تحمل (۰/۹۹, ۰/۹۹)



شکل ۴. نمودار طول مورد انتظار فاصله‌های تحمل



شکل ۵. نمودار کنترل برای تعداد نقص‌ها در نمونه‌های ۱۰۰ تایی از صفحات مدار چاپ شده



شکل ۶. نمودار هیستوگرام و نمودار کنترل داده‌ها با نمایش فاصله‌های تحمل (۰/۹۵ و ۰/۹۰)

جدول ۱. آماره‌های خلاصه احتمال‌های پوشش فاصله‌های تحمل

(۰/۹۹، ۰/۹۹)			(۰/۹۵، ۰/۹۵)			(۰/۹۰، ۰/۹۰)			
انحراف معیار	میانگین	میانگین	انحراف معیار	میانگین	میانگین	انحراف معیار	میانگین	میانگین	روش
۰/۰۵۸	۰/۹۷۰	۰/۹۴۹	۰/۰۵۲	۰/۹۳۸	۰/۹۲۳	۰/۰۴۴	۰/۹۰۵	۰/۸۹۳	LS
۰/۰۰۳	۰/۹۹۶	۰/۹۹۶	۰/۰۱۲	۰/۹۷۳	۰/۹۷۶	۰/۰۲۲	۰/۹۴۱	۰/۹۴۳	SC
۰/۰۵۹	۰/۹۷۹	۰/۹۵۷	۰/۰۵۳	۰/۹۵۳	۰/۹۳۷	۰/۰۵۰	۰/۹۲۵	۰/۹۱۶	CC
۰/۰۲۰	۰/۹۸۹	۰/۹۸۳	۰/۰۴۹	۰/۹۵۵	۰/۹۴۴	۰/۰۴۹	۰/۹۱۹	۰/۹۱۴	VS
۰/۰۰۰۳	۰/۹۹۹	۰/۹۹۹	۰/۰۴۴	۰/۹۹۷	۰/۹۹۵	۰/۰۱۲	۰/۹۸۷	۰/۹۸۶	RVS
۰/۰۰۰۱	۰/۹۹۹	۰/۹۹۸	۰/۰۰۹	۰/۹۹۰	۰/۹۸۹	۰/۰۱۴	۰/۹۷۴	۰/۹۷۵	FT

جدول ۲. طول مورد انتظار فاصله‌های تحمل

روش	(۰/۹۰,۰/۹۹)	(۰/۹۵,۰/۹۵)	(۰/۹۹,۰/۹۹)
LS	۱۶/۶۷۵	۱۹/۶۸۷	۲۵/۰۲۴
SC	۱۸/۳۴۰	۲۲/۲۹۴	۳۰/۶۵۷
CC	۱۷/۵۷۸	۲۰/۴۷۲	۲۵/۸۸۳
VS	۱۷/۵۴۴	۲۰/۷۰۲	۲۷/۰۸۲
RVS	۲۰/۶۲۱	۲۴/۳۹۲	۳۱/۷۱۳
FT	۱۹/۷۸۳	۲۳/۶۰۶	۳۰/۸۸۸

جدول ۳. داده‌های مربوط به تعداد نقص‌ها در نمونه‌های ۱۰۰ تا ۱۰۰۰ صفحات مدارهای چاپ شده

شماره نمونه	تعداد نقص‌ها	شماره نمونه	تعداد نقص‌ها
۱	۲۱	۱۴	۱۹
۲	۲۴	۱۵	۱۰
۳	۱۶	۱۶	۱۷
۴	۱۲	۱۷	۱۳
۵	۱۵	۱۸	۲۲
۶	۵	۱۹	۱۸
۷	۲۸	۲۰	۳۹
۸	۲۰	۲۱	۳۰
۹	۳۱	۲۲	۲۴
۱۰	۲۵	۲۳	۱۶
۱۱	۲۰	۲۴	۱۹
۱۲	۲۴	۲۵	۱۷
۱۳	۱۶	۲۶	۱۵

جدول ۴. فاصله‌های اطمینان ۹۵ درصد و فاصله‌های تحمل (۰/۹۰,۰/۹۵)

روش	فاصله اطمینان ۹۵ درصد	فاصله تحمل (۰/۹۰,۰/۹۵)
LS	[۱۷ / ۸۷۹, ۲۱ / ۲۸۱]	[۱۱, ۲۹]
SC	[۱۷ / ۹۵۱, ۲۱ / ۳۵۶]	[۱۱, ۳۰]
CC	[۱۷ / ۸۵۸, ۲۱ / ۳۰۲]	[۱۱, ۲۹]
VS	[۱۷ / ۹۱۶, ۲۱ / ۳۱۸]	[۱۱, ۳۰]
RVS	[۱۷ / ۸۹۹, ۲۱ / ۳۳۴]	[۱۱, ۳۰]
FT	[۱۷ / ۸۹۴, ۲۱ / ۳۳۹]	[۱۱, ۳۰]

منابع

1. Wilks S. S., "Determination of sample sizes for setting tolerance limits. Annals of Mathematical Statistics", 12 (1941) 91-96.
2. Fernandez A. J., "Two-sided tolerance intervals in the exponential case: Corrigenda and generalizations", Computational Statistics and Data Analysis, 54 (2010) 151-162.

3. Naghizadeh Qomi M., Kiapour A., "Shortest tolerance intervals controlling both tails of the exponential distribution based on record values", *Communications in Statistics – Simulation and Computation*, (2016), doi/full/10.1080/03610926.2014.990106.
4. MirMostafaei S.M.T.K., Naghizadeh Qomi M., Arturo J. Fernandez, "Tolerance limits for minimal repair times of a series system with Rayleigh distributed component lifetimes", *Applied Mathematical Modeling*, 40 (2016) 3153-3163.
5. Krishnamoorthy K., Mondal S., "Improved tolerance factors for multivariate normal distributions", *Communications in Statistics – Simulation and Computation*, 35 (2006) 461-478.
6. Mbodj M., Mathew T., "Approximate ellipsoidal tolerance regions for multivariate normal populations", *Statistics and Probability Letters*, 97 (2015) 41-45.
7. Dong X., Mathew T., "Central tolerance regions and reference regions for multivariate normal populations", *Journal of Multivariate Analysis* 134 (2015) 50-60.
8. Hahn G. J. , ChandraR., "Tolerance intervals for Poisson and binomial variables", *Journal of Quality Technology* 13(2) (1981) 100-110.
9. Wang H., Tsung F. , "Tolerance intervals with improved coverage probabilities for binomial and Poisson variables", *Technometrics*, 51 (2009) 25-33.
10. Krishnamoorthy K. , Xiaodong L., Sumona M., "Tolerance intervals for the distribution of the difference between two independent normal random variables", *Communications in Statistics-Theory and Methods*, 40 (1) (2011) 117-129.
11. Young D. S., "A procedure for approximate negative binomial tolerance intervals", *Journal of Statistical Computation and Simulation*, 84 (2) (2014) 438-450.
12. Naghizadeh Qomi M., Kiapour A., Young D. S., "Approximate tolerance intervals for the discrete Poisson-Lindley distribution, *Journal of Statistical Computation and Simulation*, 86 (2016) 841-854.
13. Owen D. B., "Control of percentages in both tails of the normal distribution", *Technometrics*, 6 (1964) 377-387.
14. Krishnamoorthy K. , "Handbook of Statistical Distributions with Applications", (2006), Boca Raton, FL: Chapman & Hall/CRC Press.
15. Barker L., "A Comparison of Nine Confidence Intervals for a Poisson Parameter When the Expected Number of Events is ≤ 5 ", *The American Statistician*, 56 (2002) 85-89.
16. Anscombe F. J., "The transformation of Poisson, binomial, and negative Binomial data", *Biometrika*, 35 (1948) 246-254.

17. Freeman M. F., Tukey J. W., "Transformations related to the angular and the square root, *Annals of Mathematical Statistics*, 21 (1950) 607-611.
18. Lehmann E. L., Romano J., "Testing statistical hypotheses" (2005) third edition, Springer.
19. Montgomery D. C., "Introduction to statistical quality control" (2005) New York: Wiley.