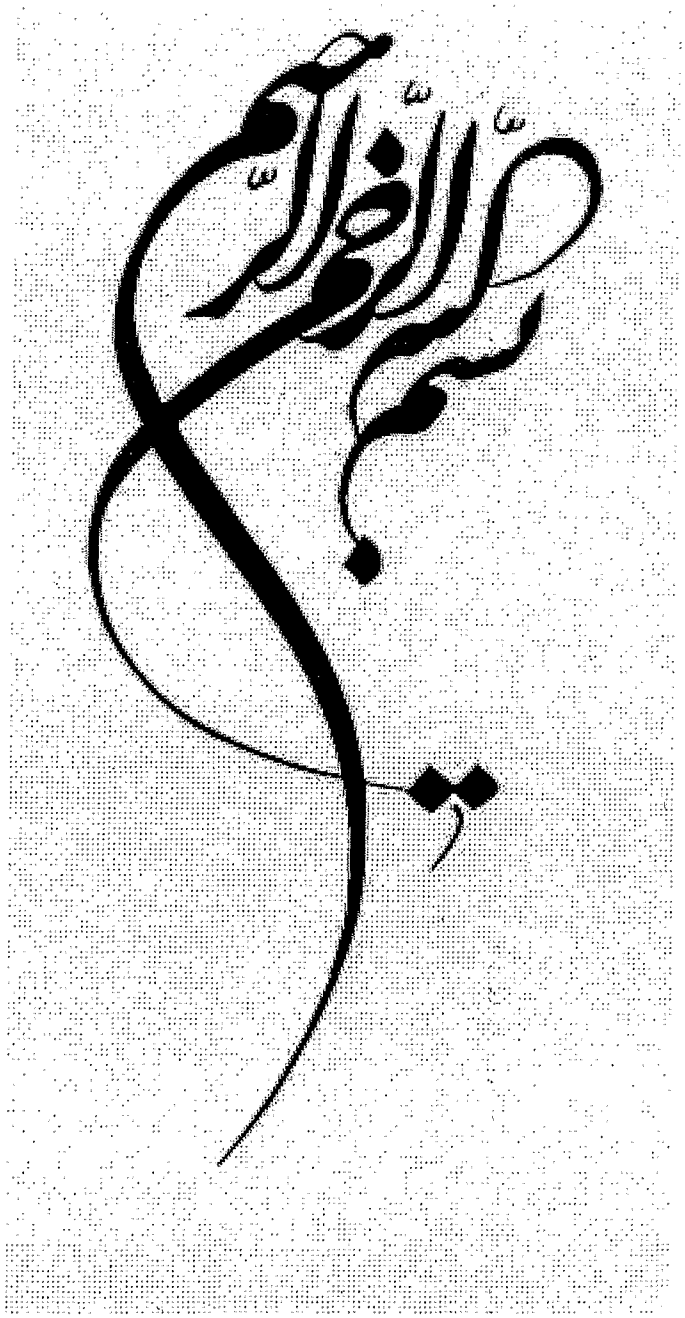

دانشگاه شهید بهشتی

دانشکده علوم ریاضی

گروه آمار

پایان‌نامه کارشناسی ارشد آمار بیمه

تقریب پواسون مرکب در مدل مخاطرهٔ فردی و کران‌های تقریب

توسط

مریم تیموریان سفیده‌خوان

استاد راهنما

دکتر محمد قاسم وحیدی‌اصل

استاد مشاور

دکتر محمد ذکایی

شهریور ۱۳۸۷

تقدیم به

پیشگاه لایزال یگانه‌اش

مادر و پدر بزرگوارم

و نثارم بر پایشان همه شرم است و شرم

خواهر و برادر مهربانم

و همه کسانی که همچون ستاره‌ای در آسمان زندگی‌ام درخشیده‌اند.

قدردانی

منت خدای را عزّوجل که طاعتش موجب قربت است و به شکر اندرش مزید نعمت.
با تشکر و سپاس فراوان از استاد ارجمند جناب آقای دکتر وحیدی که با حضور ارزشمند و راهنمائیهای مؤثرشان مرا در انجام این پروژه یاری نمودند و من همسو با علم و دانش، موفق زیستن را در مکتب ایشان آموختم. از استاد محترم جناب آقای دکتر ذکایی که مشاوره این رساله را بر عهده داشتند صمیمانه تشکر و قدردانی می‌نمایم.
از اساتید محترم جناب آقای دکتر فرید روحانی و جناب آقای دکتر امیر تیمور پاینده که زحمت مطالعه و داوری را تقبل فرمودند، کمال تشکر را دارم. از کلیهٔ عزیزان در پژوهشکدهٔ بیمه، به خصوص جناب آقای دکتر حسن‌زاده و همکاران محترم ایشان، و مدیریت محترم شرکت بیمهٔ آسیا که در به انجام رساندن این رساله زحمات فراوانی متحمل شدند، صمیمانه تشکر و قدردانی می‌کنم.

مریم تیموریان سفیده‌خوان

تهران — شهریور ماه ۱۳۸۷

پیشگفتار

داشتمان خاصی را در یک شرکت بیمه در نظر بگیرید. این شرکت برای برآورد میزان ادعای کل داشتمان و محاسبهٔ احتمال ورشکستگی و حق بیمهٔ مربوط به هر قرارداد، به تابع توزیع مجموع ادعای کل داشتمان نیاز دارد. بنابراین یافتن تابع توزیع مجموع متغیرهای تصادفی در صنعت بیمه امری بسیار با اهمیت است. فرض کنید X_i متغیر تصادفی حقیقی مقدار ونشانگر مقدار ادعای قرارداد iام باشد. بنابراین ما در پی یافتن توزیع متغیر تصادفی $S_n = \sum_{i=1}^{n} X_i$ می‌باشیم. حتی در حالتی که X_iها متغیرهای تصادفی برنولی باشند، یافتن تابع توزیع S_n برای داشتمان‌های بزرگ کار بسیار سختی است. لذا در اکثر موارد این توزیع با توزیع مناسبی تقریب زده می‌شود. توزیع‌های مناسب در طول این رساله، توزیع‌های پواسون و پواسون مرکب است. در بخش اول و دوم فصل دوم، دو مدل ممکن برای یک داشتمان توضیح داده‌شده است. در بخش سوم فصل دو روش استاین و کرستن برای یافتن کران بالای تقریب توزیع مجموع ادعای کل در شرایط خاص بیان شده است. در فصل سوم کران‌های خطا با مرتبهٔ بهتری برای تقریب مجموع ادعای کل با توزیع پواسون مرکب ارائه شده است. این کران‌ها بدون هیچ محدودیتی بر توزیع ادعاها، برقرار می‌باشند و به کران‌های با عامل جادویی معروفند. فصل چهارم اختصاص به یک مثال کاربردی دارد که در آن کران‌های خطا برای یک داشتمان خاص محاسبه شده است.

چکیده

تابع توزیع مجموع متغیرهای تصادفی نقشی مهم در مدیریت داشتمان‌ها و در نتیجه آمار بیمه دارد. به طور مثال اگر مدل یک داشتمان خاص مدل مخاطرهٔ فردی و هدف، یافتن مقدار ادعای کل این داشتمان در یک دورهٔ زمانی معین باشد، آنگاه یافتن تابع توزیع مجموع متغیرهای تصادفی امری مهم است. اما از طرفی یافتن تابع توزیع مجموع متغیرهای تصادفی در حالت کلی کاری بسیار مشکل و اغلب تقریباً ناممکن است. بنابراین تابع توزیع مجموع متغیرهای تصادفی را تقریب می‌زنیم. از جملهٔ توزیع‌های مناسب برای تقریب‌زدن، توزیع پواسون مرکب است. بدیهی است که هرتقریبی خطایی نیز دارد. بنابراین برای بهبود تقریب توزیع مجموع متغیرهای تصادفی، یافتن کران بالای تقریب اهمیت پیدا می‌کند. تاکنون روشهای متفاوتی برای یافتن این کران خطا ارائه شده است. در این رساله کران‌های بالای متفاوت و سه روش مهمی که از طریق آنها کران‌های بالایی با مرتبهٔ بهتر به‌دست می‌آیند، مورد بحث و بررسی قرار گرفته‌اند. توجه ما بیشتر به کران‌های بالایی است که هیچ محدودیت و شرطی برای توزیع متغیرهای تصادفی ندارند و از مرتبهٔ بهتری نسبت به کران‌های ارائه شده پیشین باشد. در انتها نیز کران‌های مزبور برای یک داشتمان واقعی محاسبه و مقایسه شده‌اند.

کلمات کلیدی: تقریب پواسون مرکب، مدل مخاطرهٔ فردی، عامل جادویی، فاصلهٔ تغییرات کل.

فهرست مندرجات

فصل ۱

تعاریف و مفاهیم اولیه

۱.۱ مقدمه

این فصل شامل تعریف‌ها و مفاهیم اولیهٔ مورد نیاز می‌باشد. در بخش اول، تعریف‌ها و قضیه‌های مهم در مورد نمادهای لانداو تشریح شده است. این نمادها به خصوص در ساده‌نویسی گزاره‌هایی که شامل مفهوم حد هستند، به کار می‌روند. در بخش دوم این فصل قضیهٔ رادون—نیکودیم بیان شده است که در فصل‌های بعدی از آن استفاده خواهد شد. در بخش سوم نیز توضیح مختصری در بارهٔ فاصلهٔ تغییرات کل که به عنوان معیاری برای اندازه‌گیری خطای تقریب استفاده می‌شود و نیز لم‌هایی که برای اثبات قضیه‌ها در فصل‌های بعد به‌کار خواهد رفت، ارائه شده است.

۲.۱ نمادهای O و o

نمادهای O و o که به افتخار لانداو، ریاضیدان آلمانی، به نمادهای لانداو مشهور اند، در سال ۱۸۹۲ توسط پل‌گوستاوهانریش باخمن مطرح گردید.

تعریف ۱.۲.۱ فرض کنید $f, g : \mathcal{N} \longrightarrow \mathcal{R}$. در این صورت f از مرتبهٔ حداکثر g است هرگاه ثابت مثبتی مانند C و عدد صحیح مثبتی مانند n_0 موجود باشند به‌طوری که برای هر n که $n \geq n_0$،

$$|f(n)| \leq C|g(n)|. \qquad (۱.۲.۱)$$

این مفهوم با نماد $f = O(g)$ نشان داده می شود.
عدد C یکتا نیست؛ زیرا رابطهٔ (۱.۲.۱) برای هر عدد صحیح بزرگتر از C نیز برقرار است. تعریف کلی‌تری نیز برای O به صورت زیر ارائه شده است.

تعریف ۲.۲.۱ فرض کنید g و f دو تابع حقیقی مقدار باشند که در همسایگی نقطهٔ x_0 تعریف شده‌اند. گوییم f از مرتبهٔ O تابع g است، وقتی $x \longrightarrow x_0$، و می‌نویسیم $f = O(g)$ هرگاه یک همسایگی V از x_0 و ثابت مثبتی مانند M وجود داشته باشند به‌طوری که برای هر x که $x \in V$،

$$|f(x)| \leq M|g(x)| \qquad (۲.۲.۱)$$

اگر g در نزدیکی x_0 صفر نشود، آنگاه رابطهٔ بالا معادل است با شرط

$$\limsup_{x \longrightarrow x_0} |\frac{f(x)}{g(x)}| < \infty.$$

نماد o ی کوچک مفهوم مهم دیگری است که به صورت زیر تعریف می‌شود.

تعریف ۳.۲.۱ فرض کنید g و f دو تابع حقیقی مقدار باشند که در همسایگی نقطهٔ x_0 تعریف شده‌اند. $f = o(g)$، وقتی $x \longrightarrow x_0$، هرگاه برای هر $\epsilon > 0$ همسایگی $V = V(\epsilon)$ای از x_0 وجود داشته باشد به‌طوری که برای هر x که $x \in V$ ، $|f(x)| \leq \epsilon|g(x)|$. اگر g در همسایگی x_0 صفر نشود، در این صورت شرط $f = o(g)$ ، $x \longrightarrow x_0$، معادل است با

$$\lim_{x \longrightarrow x_0} \frac{f(x)}{g(x)} = 0.$$

بدیهی است که برای $x \longrightarrow x_0$ ، اگر $f = o(g)$ آنگاه $f = O(g)$ اماعکس این رابطه برقرار نیست.

برخی از ویژگی‌های مفید نمادهای o و O در قضیهٔ زیر و بدون اثبات آمده است.

قضیه ۱.۲.۱ ویژگی‌های زیر برای نمادهای مرتبه برقرار است.

- اگر C_1 و C_2 مقادیر ثابتی باشند، آنگاه $C_1O(g) + C_2O(g) = O(g)$.
- اگر C_1 و C_2 مقادیر ثابتی باشند، آنگاه $C_1o(g) + C_2o(g) = o(g)$.
- $O(O(g)) = O(g)$.
- $O(o(g)) = o(O(g)) = o(o(g)) = o(g)$.
- $O(f)O(g) = O(fg)$.
- $O(f)o(g) = o(fg)$.

تابع g در رابطه‌های $f = O(g)$ یا $f = o(g)$، تابع پیمانه نامیده می‌شود. رابطهٔ $f = O(g)$ بیانگر این مطلب است که $\frac{f(x)}{g(x)}$ وقتی $x \longrightarrow x_0$، دارای کران متناهی است و رابطهٔ $f = o(g)$ نشانگر این است که وقتی $x \longrightarrow x$، $\frac{f(x)}{g(x)} \longrightarrow 0$؛ اما هیچ اطلاعی از سرعت همگرایی آن به‌دست نمی‌دهد. برای مثال $sinx = O(x)$ و $sinx = O(8x)$ وقتی $x \longrightarrow 0$، هر دو درست هستند. همچنین گزاره‌های $\frac{1}{1+x^2} = o(1)$ و $\frac{1}{1+x^2} = o(\frac{1}{x})$ وقتی $x \longrightarrow \infty$، نیز درست‌اند. بنابراین اطلاعی که نمادهای مرتبه به ما می‌دهند، کاملاً دقیق نبوده و از طرفی همین ویژگی مبهم بودن این امکان را به ما می‌دهد که آن‌ها را در شرایط مختلفی به‌کار ببریم. در پایان ذکر این نکته ضروری است که‌از نماد O به‌طور وسیعی در مطالعهٔ رشد توابع استفاده می‌شود. به این معنی که $f = O(g)$ فقط دلالت بر این دارد که تابع f رشدی سریعتر از g ندارد. در واقع این نماد، یک کران بالا برای f به ازای مقادیر بزرگ x ارائه می‌دهد.

۳.۱ قضیهٔ رادون—نیکودیم

در فصل سوم از قضیهٔ رادون—نیکودیم استفاده خواهیم کرد. لذا در این قسمت آن را بدون اثبات بیان می‌کنیم. برای اثبات این قضیه می‌توان به کتاب آش (۲۰۰۰) مراجعه کرد. قبل از بیان قضیه ابتدا تعریف استفاده شده در قضیه را ذکر می‌کنیم. فرض می‌کنیم $(\Omega, \mathcal{F})$ فضای اندازه‌پذیر بوده و μ اندازه‌ای سیگما—متناهی و λ اندازه‌ای علامتدار بر $\mathcal{F}$ باشند.

تعریف ۱.۳.۱ اندازهٔ λ را نسبت به اندازهٔ μ مطلقاً پیوسته گوییم اگر به ازای هر $A \in \mathcal{F}$ که $\mu(A) = 0$ ، آنگاه $\lambda(A) = 0$. این مفهوم با نماد $\lambda << \mu$ نمایش داده می‌شود.

قضیه ۱.۳.۱ فرض کنید اندازه μ نسبت به λ مطلقاً پیوسته باشد. بنابراین تابع اندازه‌پذیر بورل مانند $g : \Omega \longrightarrow \mathcal{R}$، وجود دارد به طوری که برای هر مجموعه $A \in \mathcal{F}$،

$$\lambda(A) = \int_A g d\mu.$$

اگر h تابع اندازه‌پذیر دیگری با همین خاصیت باشد، آنگاه تقریباً همه جا، نسبت به اندازهٔ μ، $g = h$.

تابع g تعریف شده در قضیه ۱.۳.۱ را مشتق رادون—نیکودیم یا چگالی λ نسبت به μ نامیده و با $\frac{d\lambda}{d\mu}$ نشان می‌دهیم. اگر μ اندازه لبگ باشد، آنگاه تابع g را چگالی λ می‌نامیم.

۴.۱ فاصلهٔ تغییرات کل

برای بررسی همگرایی در توزیع، از فاصلهٔ بین توزیع‌ها استفاده می‌شود. فاصلهٔ تغییرات کل d_{TV} برای دو متغیر تصادفی X و Y به صورت زیر تعریف می‌شود.

فرض کنید متغیرهای تصادفی X و Y بر فضای احتمال $(\Omega, \mathcal{F}, P)$ تعریف شده باشند و $\mathcal{B}$، سیگما میدان بورل (سیگما میدان روی خط حقیقی) باشد. در این صورت فاصلهٔ تغییرات کل بین توزیع‌های این دو متغیر تصادفی برابر است با

$$d_{TV}(X, Y) = \sup_{A \in \mathcal{B}} |P(X \in A) - P(Y \in A)| \qquad (۳.۴.۱)$$

اگر متغیرهای تصادفی X و Y به‌ترتیب دارای توابع توزیع F_X و F_Y باشند، آنگاه فاصلهٔ تغییرات کل بین توزیع‌های این دو متغیر تصادفی اغلب به صورت $d_{TV}(F_X, F_Y)$ نیز نشان داده می‌شود.

برای فاصلهٔ تغییرات کل تعریف‌های متفاوتی بیان شده است. این تعریف‌ها معادل باهم بوده و هر یک در شرایط مناسب استفاده می‌شود. در ادامه به یکی از این تعریف‌ها که در فصل سوم استفاده خواهد شد، اشاره می‌کنیم.

فرض کنید P_1 و P_2 دو اندازهٔ احتمال بر فضای اندازه‌پذیر $(\Omega, \mathcal{F})$ و نسبت به اندازهٔ μ مطلقاً پیوسته باشند. بنابراین طبق قضیه رادون—نیکودیم به‌ترتیب دارای توابع چگالی f_1 و f_2

نسبت به اندازهٔ μ خواهند بود. فاصلهٔ تغییرات کل میان این دو اندازه برابر است با

$$d_{TV}(P_1, P_2) = \frac{1}{2}\int |f_1 - f_2| d\mu.$$

در ادامه ثابت می‌کنیم d_{TV} یک متر است.

لم ۱.۴.۱ اگر d_{TV} فاصلهٔ تغییرات کل و X، Y و Z متغیرهای تصادفی باشند، آنگاه

$i)$ $d_{TV} \geq 0$.

$ii)$ $d_{TV}(X,Y) = d_{TV}(Y,X)$.

$iii)$ $d_{TV}(X,Y) \leq d_{TV}(X,Z) + d_{TV}(Z,Y)$.

$iv)$ $d_{TV}(X,Y) = 0 \iff P_X(A) = P_Y(A)$ ، $\forall A \in \mathcal{B}$.

برهان : با توجه به تعریف فاصلهٔ تغییرات کل و خواص قدرمطلق، روابط (i) و (ii) بدیهی‌اند. برای اثبات (iii) با استفاده از نابرابری مثلثی برای متغیرهای تصادفی X، Y، Z داریم :

$$\begin{aligned} |P(X \in A) - P(Y \in A)| &\leq |P(X \in A) - P(Z \in A)| \\ &+ |P(Z \in A) - P(Y \in A)| \end{aligned}$$

حال اگر از طرفین نابرابری بالا نسبت به $A \in \mathcal{B}$ سوپرمم بگیریم، خواهیم داشت:

$$d_{TV}(X,Y) \leq d_{TV}(X,Z) + d_{TV}(Z,Y).$$

برای اثبات حکم (iv)،

$$d_{TV}(X,Y) = 0 \iff \sup_{A \in \mathcal{B}} |P(X \in A) - P(Y \in A)| = 0$$

اگر و تنها اگر به ازای هر $A \in \mathcal{B}$، $P_X(A) = P_Y(A)$. ■

لم ۲.۴.۱ اگر d_{TV} فاصلهٔ تغییرات کل و X و Y متغیرهای تصادفی گسسته باشند، آنگاه

$$\begin{aligned} i) d_{TV}(X,Y) &\leq P(X \neq Y). \\ ii) d_{TV}(X,Y) &= \frac{1}{2}\sum_{k=0}^{\infty} |P(X=k) - P(Y=k)|. \end{aligned}$$

برهان : برای اثبات اولین نابرابری مجموعه دلخواه A، $A \in \mathcal{B}$، را در نظر می‌گیریم.

$$\begin{aligned} P(X \neq Y) &\geq P(X \in A, Y \in A^c) \\ &= P(X \in A) - P(X \in A, Y \in A) \geq P(X \in A) - P(Y \in A) \end{aligned}$$

همچنین با تعویض نقش‌های X و Y،

$$P(X \neq Y) \geq P(Y \in A) - P(X \in A)$$

بنابراین

$$P(X \neq Y) \geq |P(X \in A) - P(Y \in A)|.$$

حال اگر از طرفین رابطهٔ بالا نسبت به $A \in \mathcal{B}$ سوپرمم بگیریم، حکم حاصل می‌شود.
برای اثبات تساوی دوم فرض کنید که $A \subseteq \{0, 1, 2, ...\}$. در این صورت

$$\begin{aligned} \sum_{k=0}^{\infty} |P(X=k) - P(Y=k)| &= \\ &= \sum_{k \in A} |P(X=k) - P(Y=k)| + \sum_{k \in A^c} |P(X=k) - P(Y=k)| \\ &\geq |\sum_{k \in A} P(X=k) - P(Y=k)| + |\sum_{k \in A^c} P(X=k) - P(Y=k)| \\ &= |P(X \in A) - P(Y \in A)| + |P(X \in A^c) - P(Y \in A^c)| \\ &= 2|P(X \in A) - P(Y \in A)|. \end{aligned}$$

بنابراین

$$\sup_A |P_X(A) - P_Y(A)| \leq \frac{1}{2}\sum_{k=0}^{\infty} |P(X=k) - P(Y=k)|. \qquad (۴.۴.۱)$$

حال فرض کنید $A = \{k : P(X = k) \geq P(Y = k)\}$. در این صورت

$$\begin{aligned} ۲|P(X \in A) - P(Y \in A)| &= ۲|\sum_{k\in A} P(X = k) - P(Y = k)| \\ &= ۲\sum_{k\in A}[P(X = k) - P(Y = k)]. \end{aligned}$$

از طرفی

$$\begin{aligned} \sum_{k=۰}^{\infty} |P(X = k) - P(Y = k)| &= \sum_{k\in A} P(X = k) - P(Y = k) + \sum_{k\in A^c} P(Y = k) - P(X = k) \\ &= ۲\sum_{k\in A}\{P(X = k) - P(Y = k)\}. \end{aligned}$$

بنابراین

$$۲|P(X \in A) - P(Y \in A)| = \sum_{k=۰}^{\infty} |P(X = k) - P(Y = k)|$$

و

$$\sup_A |P_X(A) - P_Y(A)| \geq \frac{۱}{۲}\sum_{k=۰}^{\infty} |P(X = k) - P(Y = k)|. \qquad (۵.۴.۱)$$

از روابط (۴.۴۱) و (۵.۴۱) اثبات کامل می‌شود. ■

لم زیر برای یافتن کران بالای فاصلهٔ تغییرات کل مورد استفاده قرار خواهد گرفت.

لم ۳.۴.۱ برای دنباله‌های $\{X_i\}$، $i = ۱, ۲, ..., n$، و $\{Y_i\}$، $i = ۱, ۲, ..., n$، از متغیرهای تصادفی دلخواه داریم:

$$P(\sum_{i=۱}^{n} X_i \neq \sum_{i=۱}^{n} Y_i) \leq \sum_{i=۱}^{n} P(X_i \neq Y_i).$$

برهان : برای هر i، $i = ۱, ۲, ..., n$، تعریف می‌کنیم :

$$A_i = \{X_i \neq Y_i\}.$$

با استفاده از نابرابری بول،

$$\begin{aligned} P(\sum_{i=۱}^{n} X_i \neq \sum_{i=۱}^{n} Y_i) &\leq P(\bigcup_{i=۱}^{n} A_i) \\ &\leq \sum_{i=۱}^{n} P(A_i) = \sum_{i=۱}^{n} P(X_i \neq Y_i). \end{aligned}$$

لذا حکم ثابت می‌شود. ■

روش‌های بسیاری برای یافتن کوچکترین کران‌های بالا برای متر $d_{TV}(X,Y)$ موجود است. یکی از این روش‌ها، روش جفت‌سازی است . این روش را وانگ (۱۹۸٦) ارائه داد. فرض کنید X و Y متغیرهای تصادفی گسسته بر فضای احتمال $(\Omega,\mathcal{F},P)$ باشند. با توجه به لم (۲.۴.۱) برای یافتن کران بالای $d_{TV}(X,Y)$ می‌توان توزیع توأمی برای متغیرهای تصادفی X و Y یافت به‌طوری که $P(X=Y)$ ماکسیمم شود.

می‌دانیم که به ازای هر دو توزیع F و G توزیع‌های توأم بسیاری مانند H وجود دارند که دارای توزیع‌های حاشیه‌ای F و G باشند. به‌ویژه دو توزیع توأم H^* و H را می‌توان از توزیع‌های F و G تولید کرد به‌طوری که به‌ترتیب دارای بیشترین و کمترین همبستگی باشند. ویت(۱۹۷٦) ثابت می‌کند که اگر X دارای توزیع برنولی و Y دارای توزیع پواسون با $E(X)=E(Y)$ باشد، آنگاه روش استفاده شده در استخراج توزیع توأم H^* که دارای بیشترین همبستگی است، موجب ماکسیمم شدن $P(X=Y)$ نیز خواهد شد.اما در حالت کلی توزیع توأم H^* مقدار$P(X=Y)$ را ماکسیمم نخواهد کرد. می‌خواهیم بدانیم که آیا روشی وجود دارد که بتوان توزیع توأمی (جفت‌سازی) یافت به طوری که موجب برقراری تساوی $d_{TV}(X,Y)=P(X\neq Y)$ برای هر دو متغیر تصادفی گسستهٔ دلخواه X و Y شود؟ پاسخ مثبت است. برای این منظور ابتدا تعریف زیر را بیان می‌کنیم.

تعریف ۱.۴.۱ جفت‌ساز دو متغیر تصادفی گسستهٔ دلخواه X و Y را بیشین می‌نامند، هرگاه

$$d_{TV}(X,Y)=P(X\neq Y).$$

چنین جفت‌ساز بیشین را با $P^*(X\neq Y)$ نشان می‌دهیم.

برای یافتن توزیع توأمی که در تعریف (۱.۴.۱) صدق کند، فرض خواهیم کرد که X و Y متغیرهای تصادفی گسسته و حقیقی مقدار باشند که بر فضای احتمال $(\Omega,\mathcal{F},P)$ تعریف شده‌اند. مجموعهٔ B را به صورت زیر تعریف می‌کنیم:

$$B=\{x:0\leq P(X=x)<P(Y=x)\}.$$

برای هر $x\in B^c$ (B^c متمم مجموعه B است) تعریف می‌کنیم:

$$a_x = P(X=x)-P(Y=x)$$

$$\alpha_x = \frac{a_x}{\sum_{x\in B^c} a_x}$$

و برای هر $y \in B$ ،

$$b_y = P(Y = y) - P(X = y)$$
$$\beta_y = \frac{b_y}{\sum_{y \in B} b_y}$$

به راحتی ثابت می‌شود که برای هر x, y ،

$$a_x \beta_y = \alpha_x b_y. \quad ، \quad \sum_{x \in B^c} a_x = \sum_{y \in B} b_y$$

چگالی توأم متغیرهای X و Y را به صورت زیر تعریف می‌کنیم:
برای $x \in B$،

$$f(x,y) = \begin{cases} P(X = x) & y = x \\ ۰ & \text{در غیر این صورت} \end{cases}$$

و برای $y \in B^c$ ،

$$f(x,y) = \begin{cases} P(Y = y) & x = y \\ ۰ & \text{در غیر این صورت} \end{cases}$$

و برای $x \in B^c$ و $y \in B$،

$$f(x,y) = a_x \beta_y = \alpha_x b_y. \qquad (۶.۴.۱)$$

تابع f، تمام شرایط تابع چگالی را دارا است؛ زیرا اولاً مقادیر تابع f نامنفی است. ثانیاً

$$\begin{aligned} \sum f(x,y) &= \sum_{x \in B} P(X = x) + \sum_{y \in B^c} P(Y = y) + \sum_{x \in B^c} \sum_{y \in B} \alpha_x b_y \\ &= \sum_{x \in B} P(X = x) + ۱ - \sum_{y \in B} P(Y = y) + \sum_{x \in B^c} \sum_{y \in B} \frac{a_x}{\sum_{x \in B^c} a_x} b_y \\ &= \sum_{x \in B} P(X = x) + ۱ - \sum_{y \in B} P(Y = y) + \sum_{y \in B} b_y = ۱. \end{aligned}$$

قضیه ۱.۴.۱ اگر متغیرهای X و Y دارای تابع چگالی توأم f تعریف شده به صورت (۶.۴.۱) باشند، آنگاه

$$d_{TV}(X, Y) = P(X \neq Y).$$

برهان : با استفاده از لم ۲.۴.۱ ،

$$\begin{aligned} d_{TV} &= \frac{1}{2}|\sum_{x=0}^{\infty} P(Y=x) - P(X=x)| \\ &= \frac{1}{2}[\sum_{x\in B}\{P(Y=x)-P(X=x)\} + \sum_{x\in B^c}\{P(X=x)-P(Y=x)\}] \\ &= \frac{1}{2}[\sum_{x\in B} b_x + \sum_{x\in B^c} a_x] \end{aligned}$$

اما با توجه به اینکه $\sum_{x\in B^c} a_x = \sum_{x\in B} b_x$، بنابراین

$$d_{TV} = \sum_{x\in B^c} a_x. \qquad (۷.۴.۱)$$

از طرفی با استفاده از قانون احتمال کل و توزیع توأم f،

$$\begin{aligned} P(X\neq Y) &= \sum_{x=0}^{\infty} P(X=x, Y\neq x) \\ &= \sum_{x\in B} P(X=x, Y\neq x) + \sum_{x\in B^c} P(X=x, Y\neq x) \\ &= \sum_{x\in B} 0 + \sum_{x\in B^c}[\sum_{y\in B} a_x\beta_y + \sum_{y\in B^c, y\neq x} 0] \\ &= \sum_{x\in B^c}\sum_{y\in B}[P(X=x)-P(Y=x)][\frac{P(Y=y)-P(X=y)}{\sum_{y\in B}[P(Y=y)-P(X=y)]}] \\ &= \sum_{x\in B^c} P(X=x) - P(Y=x) = \sum_{x\in B^c} a_x. \end{aligned}$$

بنابراین $d_{TV}(X,Y) = P(X\neq Y)$. ■

همان طور که قبلاً اشاره شد، روش جفت‌سازی برای اثبات قضیهٔ زیر به‌کار می‌رود.

قضیه ۲.۴.۱ هرگاه $\{X_1, X_2, ..., X_n\}$ و $\{Y_1, Y_2, ..., Y_n\}$ دنباله‌هایی از متغیرهای تصادفی گسسته و مستقل باشند، در این صورت

$$d_{TV}(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} Y_i) \leq \sum_{i=1}^{n} d_{TV}(X_i, Y_i).$$