دانشکده‌ی علوم ریاضی

احتمال و کاربرد آن	۹ اردیبهشت ۹۳

جلسه‌ی ۱۸: روش‌های غیرمستقیم محاسبه‌ی میانگین و واریانس

مدرّس: دکتر شهرام خزائی *نگارنده: مجتبی تفاق و امیرمسعود گیوه‌چی*

۱ هدف

در این جلسه می‌خواهیم در قالب حل چند مسئله میانگین و واریانس متغیرهای تصادفی را بدون استفاده از تابع چگالی احتمال و به صورت غیر مستقیم محاسبه کنیم.

۲ مجموع متغیرهای تصادفی مستقل

یادآوری

در جلسه‌ی پیش دیدیم که اگر X و Y دو متغیر تصادفی پیوسته و مستقل از هم باشند و تعریف کنیم:

$$Z \triangleq X + Y$$

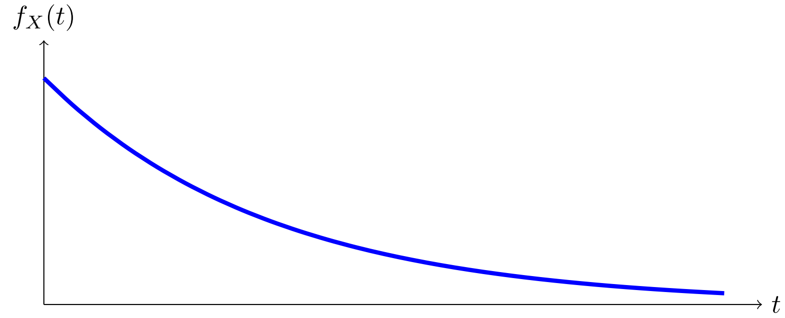
آنگاه تابع چگالی احتمال Z با توجه به رابطه زیر بدست می‌آید:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(t) f_Y(z-t) dt$$

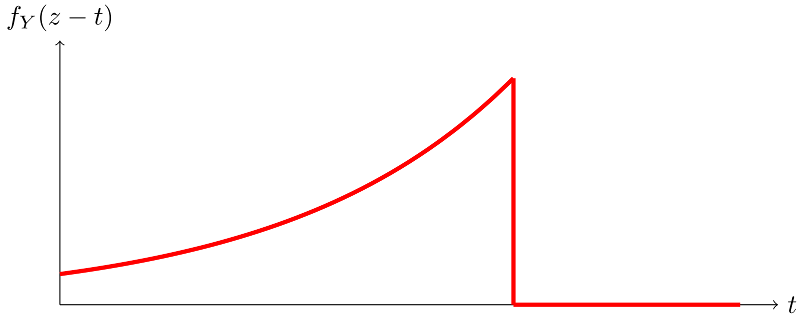
۱.۲ مجموع دو متغیر تصادفی نمایی

فرض کنید X و Y دو متغیر نمایی مستقل با پارامتر λ باشند و تعریف کنید:

$$Z \triangleq X + Y$$



(آ) $f_X(t) = \lambda e^{-\lambda t}$



(ب) $f_Y(z-t)$

شکل ۱: برای محاسبه‌ی f_Z دو نمودار بالا باید در هم ضرب شده و سپس انتگرال گرفته شود.

حال می‌خواهیم تابع چگالی احتمال متغیر تصادفی Z را محاسبه کنیم. به شکل بالا توجه کنید.

$$z < 0 \Longrightarrow f_X(t) f_Y(z-t) = 0 \Longrightarrow \int_{-\infty}^{\infty} f_X(t) f_Y(z-t) dt = 0 \Longrightarrow f_Z(z) = 0$$

$$z \geq 0 \Longrightarrow f_Z(z) = \int_{-\infty}^{\infty} f_X(t) f_Y(z-t) dt = \int_{0}^{z} \lambda e^{-\lambda t} \lambda e^{-\lambda(z-t)} dt = \lambda^2 \int_{0}^{z} e^{-\lambda z} dt = \lambda^2 z e^{-\lambda z}$$

بنابراین $f_Z(z)$ به صورت زیر محاسبه می‌شود.

$$f_Z(z) = \begin{cases} \lambda^2 z e^{-\lambda z} & z \geq 0 \\ 0 & z < 0 \end{cases}$$

■

۲.۲ امید ریاضی مجموع متغیرهای تصادفی

فرض کنید X و Y دو متغیر تصادفی باشند و تعریف کنید:

$$Z \triangleq X + Y$$

در این قسمت می‌خواهیم $E[Z]$ را محاسبه کنیم. در حالت کلی اگر متغیر تصادفی Z تابعی از دو متغیر تصادفی X و Y مانند $g(X,Y)$ باشد، آن‌گاه داریم:

$$E[Z] = E[g(X,Y)] = \iint g(x,y) dx dy$$

اما در این حالت خاص $(Z = X + Y)$ داریم:

$$E[Z] = E[X] + E[Y]$$

به عنوان نمونه در مثال قبل داریم:

$$E[Z] = \frac{1}{\lambda} + \frac{1}{\lambda} = \frac{2}{\lambda}$$

حال می‌خواهیم در حالت گسسته رابطه بالا را ثابت کنیم.

$$E[X+Y] = \sum_x \sum_y (x+y) P_{X,Y}(x,y) = \sum_x \sum_y x P_{X,Y}(x,y) + \sum_y \sum_x y P_{X,Y}(x,y)$$

$$= \sum_x x\big(\sum_y P_{X,Y}(x,y)\big) + \sum_y y\big(\sum_x P_{X,Y}(x,y)\big) = \sum_x x P_X(x) + \sum_y y P_Y(y) = E[X] + E[Y]$$

در حالت پیوسته نیز رابطه‌ی بالا به طریق مشابه ثابت می‌شود.
در حالت کلی رابطه‌ی

$$E[X_1 + X_2 + \cdots + X_n] = E[X_1] + E[X_2] + \cdots + E[X_n]$$

نیز به استقرا ثابت می‌شود.
ویژگی فوق، ویژگی خطی امید ریاضی نامیده می‌شود. دقت کنید که این ویژگی مستقل از این که متغیرهای تصادفی X_i مستقل یا وابسته باشند، برقرار است. ■

۳.۲ واریانس مجموع متغیرهای تصادفی

رابطه‌ی

$$Var(X+Y) = Var(X) + Var(Y)$$

در حالت کلی برقرار نیست. مثلا اگر فرض کنیم $Y = -X$ در آن صورت $Var(X+Y) = 0$ زیرا $X+Y$ یک متغیر تک مقداری خواهد شد و رابطه‌ی بالا در این حالت صدق نمی‌کند. اما در حالتی که X و Y مستقل باشند، این رابطه برقرار است. به عنوان مثال فرض کنید:

$$X \sim N(\mu_X, \sigma_X)$$
$$Y \sim N(\mu_Y, \sigma_Y)$$
$$Z \triangleq X + Y$$

تابع چگالی احتمال Z از انتگرال پیچیده‌ی زیر قابل محاسبه است:

$$f_Z(z) = \int_{-\infty}^{\infty} \frac{1}{\sigma_X\sqrt{2\pi}} e^{-\frac{(t-\mu_X)^2}{\sigma_X^2}} \frac{1}{\sigma_Y\sqrt{2\pi}} e^{-\frac{(z-t-\mu_Y)^2}{\sigma_Y^2}} dt$$

با محاسبه‌ی انتگرال فوق می‌توان نشان داد متغیر تصادفی Z دارای توزیع گوسی با پارامترهای زیر خواهد بود:

$$E[Z] = \mu_X + \mu_Y$$
$$Var(Z) = \sigma_X^2 + \sigma_Y^2$$

البته بعدا با استفاده از تعریف تابع مولد گشتاور اثبات ساده‌تری برای این مسئله ارائه خواهیم کرد. بنابراین تابع چگالی احتمال Z به صورت زیر خواهد بود:

$$f_Z(z) = \frac{1}{\sqrt{2\pi(\sigma_X^2 + \sigma_Y^2)}} e^{-\frac{(z-(\mu_X+\mu_Y))^2}{\sigma_X^2+\sigma_Y^2}}$$

■

۳ چند مسئله

۱.۳ توپ طلا

مثال ۱ *دو آزمایش زیر را در نظر بگیرید:*

- **آزمایش اول:** *فرض کنید دو کیسه داریم. هر کدام از کیسه‌ها شامل N توپ است که دقیقاً یکی از آن‌ها توپ طلاست. از هر کدام از دو کیسه یک توپ برمی‌داریم.*
- **آزمایش دوم:** *در این قسمت یک کیسه شامل N توپ داریم که دقیقاً یکی از آن‌ها توپ طلاست و از آن بدون جای‌گذاری دو توپ برمی‌داریم.*

به نظر شما این بار کدام آزمایش بهتر است؟

جواب *به عنوان معیار اول متوسط تعداد توپ‌های طلا را در این دو آزمایش مقایسه می‌کنیم. فرض کنید Y متغیر تصادفی باشد که نمایان‌گر تعداد توپ‌های طلای حاصل است. در آن صورت*

$$Y = X_۱ + X_۲ +$$

که $X_۱$ و $X_۲$ دو متغیر تصادفی هستند که به طریق روبه‌رو تعریف می‌شوند. هر کدام از X_i ها یک متغیر تصادفی برنولی هستند و در صورتی یک می‌شوند که توپ i ام طلا باشد. توجه کنید که در آزمایش اول $X_۱$ و $X_۲$ مستقل می‌باشند و در آزمایش دوم این‌گونه نیست اما می‌دانیم که در هر دو آزمایش $X_۱$ و $X_۲$ دارای توزیع برنولی با پارامتر $\frac{۱}{N}$ می‌باشند، بنابراین:

$$E[X_۱] = E[X_۲] = \frac{۱}{N} \Longrightarrow E[Y] = E[X_۱] + E[X_۲] = \frac{۲}{N}$$

پس متوسط تعداد توپ‌های طلا در هر دو آزمایش یکسان است. در این حالت به عنوان معیار دوم واریانس متغیر Y را در دو آزمایش مقایسه می‌کنیم. توزیع Y در آزمایش اول به صورت

$$p_Y(y) = \begin{cases} (\frac{N-۱}{N})^۲ & y = ۰ \\ ۲ \times \frac{۱}{N} \times \frac{N-۱}{N} & y = ۱ \\ \frac{۱}{N^۲} & y = ۲ \end{cases}$$

و در آزمایش دوم به صورت

$$p_Y(y) = \begin{cases} \frac{N-۲}{N} & y = ۰ \\ \frac{۲}{N} & y = ۱ \end{cases}$$

می‌باشد. با توجه به توزیع‌های بالا مشخص به طور شهودی مشخص است که متغیر تصادفی Y در آزمایش دوم واریانس کم‌تری نسبت به آزمایش اول دارد. به طور دقیق‌تر واریانس Y در دو حالت فوق به ترتیب برابر

$$\begin{aligned} Var(Y) &= E[Y^۲] - (E[Y])^۲ \\ &= \left[\left((\frac{N-۱}{N})^۲\right) \times ۰^۲ + \left(۲ \times \frac{۱}{N} \times \frac{N-۱}{N}\right) \times ۱^۲ + \left(\frac{۱}{N^۲}\right) \times ۲^۲\right] - (\frac{۲}{N})^۲ \\ &= \frac{۲(N-۱)}{N^۲} \end{aligned}$$

و

$$\begin{aligned} Var(Y) &= E[Y^2] - (E[Y])^2 \\ &= \left[\left(\frac{N-2}{N}\right) \times 0^2 + \left(\frac{2}{N}\right) \times 1^2\right] - \left(\frac{2}{N}\right)^2 \\ &= \frac{2(N-2)}{N^2} \end{aligned}$$

می‌باشد. بنابراین اگر شخصی بخواهد محتاطانه عمل کند از آزمایش دوم استفاده خواهد کرد و اگر ریسک‌پذیر باشد آزمایش اول را انتخاب خواهد نمود.

۲.۳ توپ طلا در حالتی دیگر

مثال ۲ *در ادامه آزمایش‌های مثال قبل را به گونه زیر تغییر می‌دهیم:*

- **آزمایش اول:** *فرض کنید N کیسه داریم. هر کدام از کیسه‌ها شامل N توپ است که دقیقاً یکی از آن‌ها توپ طلاست. از هر کدام از N کیسه یک توپ برمی‌داریم.*
- **آزمایش دوم:** *در این قسمت یک کیسه شامل N توپ داریم که دقیقاً یکی از آن‌ها توپ طلاست و از آن بدون جای‌گذاری N توپ (همه‌ی توپ‌ها!) را برمی‌داریم.*

به نظر شما کدام آزمایش بهتر است؟

جواب *همانند مثال قبل به عنوان معیار اول متوسط تعداد توپ‌های طلا را در این دو آزمایش مقایسه می‌کنیم. در این حالت به طور کاملا مشابه فرض کنید Y متغیر تصادفی باشد که نمایان‌گر تعداد توپ‌های طلای حاصل است. در آن صورت*

$$Y = X_1 + X_2 + \cdots + X_N$$

که X_i ها متغیرهای تصادفی هستند که به طریق روبه‌رو تعریف می‌شوند. هر کدام از X_i ها یک متغیر تصادفی برنولی است که در صورتی یک می‌شود که توپ i ام طلا باشد. توجه کنید که در آزمایش اول X_i ها مستقل می‌باشند و در آزمایش دوم این‌گونه نیست اما می‌دانیم که در هر دو آزمایش X_i ها دارای توزیع برنولی با پارامتر $\frac{1}{N}$ می‌باشند، بنابراین در هر دو آزمایش:

$$\forall i : E[X_i] = \frac{1}{N} \Longrightarrow E[Y] = E[X_1] + E[X_2] + \cdots + E[X_N] = \frac{N}{N} = 1$$

پس متوسط تعداد توپ‌های طلا در هر دو آزمایش یکسان است. هم‌اکنون به عنوان معیار دوم واریانس متغیر Y را در دو آزمایش مقایسه می‌کنیم. متغیر تصادفی Y در آزمایش اول یک متغیر دو جمله‌ای با پارامترهای N و $\frac{1}{N}$ و در آزمایش دوم یک متغیر تک مقداری با مقدار ثابت یک می‌باشد. واریانس Y در این دو آزمایش به ترتیب برابر $\frac{N-1}{N}$ و 0 است. بنابراین اگر شخصی بخواهد محتاطانه عمل کند از آزمایش دوم استفاده خواهد کرد و اگر ریسک‌پذیر باشد آزمایش اول را انتخاب خواهد نمود.

۳.۳ تشخیص استقلال متغیرهای تصادفی

در جلسه‌ی پیش دیدیم که اگر X و Y دو متغیر تصادفی مستقل باشند:

$$f_{X,Y}(x,y) = f_X(x) f_Y(y)$$

یا معادلاً:

$$f_{X|Y}(x|y) = f_X(x)$$

فرض کنید $f_{X,Y}$ را بدانیم. برای مثال دو توزیع زیر را در نظر بگیرید:

$$f_{X,Y}(x,y) = \begin{cases} ۶e^{-۲x-۳y} & x,y > ۰ \\ \\ ۰ & oth \end{cases}$$

$$f_{X,Y}(x,y) = \begin{cases} ۲۴xy & ۰ \leq x \leq ۱-y \leq ۱ \\ \\ ۰ & oth \end{cases}$$

حال سوال این است که در این دو حالت آیا X و Y از هم مستقل هستند؟
در حالت اول پاسخ مثبت است زیرا $f_{X,Y}$ به صورت ضرب دو توزیع نمایی می‌باشد ولی در حالت دوم نمی‌توان تابع توزیع مشترک را به صورت ضرب دو تابع توزیع نوشت و لذا پاسخ منفی است. ■

۴.۳ کلاه‌ها

مثال ۳ *N نفر هر کدام یک کلاه دارند و آن‌ها را در یک کیسه انداخته‌اند. سپس هر کدام یک کلاه را به تصادف از کیسه برمی‌دارند. به طور متوسط، چند نفر کلاه خود را به درستی از کیسه برمی‌دارند؟*

جواب *فرض کنید Y متغیر تصادفی نمایان‌گر تعداد افرادی باشد که کلاه خود را به درستی برداشته‌اند. در آن صورت داریم:*

$$Y = X_۱ + X_۲ + \cdots + X_N$$

که در عبارت بالا هر کدام از X_i ها یک متغیر تصادفی برنولی است که در صورتی یک می‌شود که نفر i ام کلاه خود را درست برداشته باشد. بنابراین میانگین متغیر تصادفی X_i برابر با $\frac{۱}{N}$ است، پس:

$$E[Y] = E[X_۱] + E[X_۲] + \cdots + E[X_N] = N \times \frac{۱}{N} = ۱$$

در ادامه واریانس Y را محاسبه می‌کنیم.

$$\begin{aligned} E[Y^۲] &= E[(\sum_i X_i)^۲] \\ &= E[\sum X_i^۲ + ۲ \sum\sum_{۱\leq i<j\leq N} X_i X_j] \\ &= \sum_i E[X_i^۲] + ۲ \sum\sum_{۱\leq i<j\leq N} E[X_i X_j] \end{aligned}$$

بنا بر نحوه‌ی تعریف X_i ها، متغیرهای $X_i^۲$ و X_iX_j خود متغیرهای تصادفی برنولی با پارامترهای $\frac{۱}{N}$ و $\frac{۱}{N(N-۱)}$ می‌باشند.
پس عبارت بالا برابر است با:

$$E[Y^۲] = N \times \frac{۱}{N} + \binom{N}{۲} \times ۲ \frac{۱}{N(N-۱)} = ۲$$

بنابراین:

$$Var(Y) = E[Y^۲] - (E[Y])^۲ = ۲ - ۱ = ۱$$

۵.۳ شکستن میله

مثال ۴ *یک میله با طول l داریم. آن را از نقطه‌ای تصادفی (با توزیع یکنواخت) می‌شکنیم. فاصله‌ی این نقطه را از سر میله با متغیر تصادفی Y نشان می‌دهیم. قطعه‌ی به‌دست آمده را از نقطه‌ی تصادفی دیگری (با توزیع یکنواخت) می‌شکنیم. فاصله‌ی این نقطه را از همان سر قبلی میله با متغیر تصادفی X نشان می‌دهیم. توزیع و میانگین متغیر X را بیابید.*

جواب *چون توزیع متغیر تصادفی Y یکنواخت است، پس داریم:*

$$f_Y(y) = \begin{cases} ۱ & ۰ \leq y \leq ۱ \\ \\ ۰ & oth \end{cases}$$

و بنا بر فرض مسئله $f_{X|Y}(x|y)$ نیز یکنواخت است، بنابراین:

$$f_{X|Y}(x|y) = \begin{cases} \frac{۱}{y} & ۰ \leq x \leq y \\ \\ ۰ & oth \end{cases}$$

در نتیجه توزیع مشترک X و Y از رابطه‌ی زیر بدست می‌آید:

$$f_{X,Y}(x,y) = f_Y(y) f_{X|Y}(x|y) = \begin{cases} \frac{۱}{y} & ۰ \leq x \leq y \leq ۱ \\ \\ ۰ & oth \end{cases}$$

حال برای محاسبه‌ی تابع توزیع X به طریق زیر عمل می‌کنیم:

$$f_X(x) = \int_x^۱ \frac{۱}{y} dy = -\ln(x)$$

همچنین میانگین متغیر تصادفی X به صورت زیر قابل محاسبه است:

$$E[X] = \int_۰^۱ -x\ln(x)\ dx = \frac{۱}{۴}$$

۴ محاسبه میانگین متغیرهای تصادفی به کمک میانگین شرطی

فرض کنید $f_{X,Y}$ تابع توزیع مشترک دو متغیر تصادفی X و Y باشد. در آن صورت برای y داده شده، $f_{X|Y}(x|y)$ خود یک متغیر تصادفی است. از طرفی اگر g تابعی دلخواه باشد، آن‌گاه داریم:

$$E[g(X)|y] = \int g(x) f_{X|Y}(x|y) dx$$

به طور خاص $E[X|y]$ تابعی از y مانند $h(y)$ است. بنابراین می‌توانیم در مورد $h(Y) = E[X|Y]$ به عنوان یک متغیر تصادفی و متوسط آن که به صورت زیر نشان داده می‌شود، صحبت کنیم.

$$E[h(Y)] = E[E[X|Y]]$$

اما می‌توان ثابت نمود که $E[E[X|Y]] = E[X]$. در نتیجه در مثال پیشین، می‌توانیم بدین‌گونه عمل کنیم:

$$E[X] = E[E[X|Y]] = E[\frac{Y}{۲}] = \frac{۱}{۲} \times \frac{۱}{۲} = \frac{۱}{۴}$$

۵ ارتباط میانگین و واریانس چند متغیر تصادفی مشهور با یک‌دیگر

۱.۵ متغیر تصادفی دوجمله‌ای

می‌دانیم متغیر تصادفی دوجمله‌ای با پارامترهای n و p جمع n متغیر تصادفی برنولی مستقل با پارامتر p می‌باشد. بنابراین با توجه به مطالب گفته شده، میانگین و واریانس آن n برابر میانگین و واریانس متغیر برنولی است. به طور دقیق‌تر فرض کنید $X_1, X_2, \cdots, X_n$ متغیرهای تصادفی برنولی مستقل با پارامتر p باشند و متغیر Y به صورت $Y = X_1 + X_2 + \cdots + X_n$ تعریف شود. در نتیجه Y یک متغیر تصادفی دوجمله‌ای با پارامترهای n و p است. از آن‌جا که متغیر تصادفی برنولی با پارامتر p دارای میانگین و واریانس به ترتیب p و $p(1-p)$ است، بنابراین داریم:

قضیه ۱ *اگر Y یک متغیر تصادفی دوجمله‌ای با پارامترهای n و p باشد، آن‌گاه:*

$$E[Y] = np$$
$$Var(Y) = np(1-p)$$

۲.۵ متغیر تصادفی فوق هندسی

می‌دانیم متغیر تصادفی فوق هندسی با پارامترهای (n, N, m) به صورت روبه‌رو تعریف می‌شود. فرض کنید N توپ داریم که m تای آن‌ها سفید و بقیه سیاه است. حال بدون جای‌گذاری n توپ به تصادف انتخاب می‌کنیم. در آن صورت اگر تعداد توپ‌های سفید انتخاب‌شده را با Y نشان دهیم، Y یک متغیر تصادفی فوق هندسی با پارامترهای (n, N, m) می‌باشد. در نتیجه، $Y = X_1 + X_2 + \cdots + X_n$ که در آن X_i یک متغیر تصادفی برنولی و نشان‌گر این است که توپ i ام سفید بوده است یا خیر، و از جلسات پیش می‌دانیم که میانگین هر کدام از X_i ها $\frac{m}{N}$ است. بر خلاف حالت قبل X_i ها از یک‌دیگر مستقل نیستند پس نمی‌توان گفت که

$$Var(Y) = Var(X_1) + Var(X_2) + \cdots + Var(X_n)$$

قضیه ۲ *اگر Y یک متغیر تصادفی فوق هندسی با پارامترهای (n, N, m) باشد، آن‌گاه:*

$$E[Y] = \frac{nm}{N}$$

۳.۵ متغیر تصادفی دوجمله‌ای منفی

متغیر تصادفی دو جمله‌ای منفی با پارامترهای r و p به صورت روبه‌رو تعریف می‌شود. فرض کنید یک سکه داریم که با احتمال p شیر و با احتمال $1-p$ خط می‌آید. حال تعداد پرتاب‌های لازم برای مشاهده‌ی r امین شیر را با متغیر تصادفی Y که یک متغیر دوجمله‌ای منفی با پارامترهای r و p است نمایش می‌دهیم. در آن صورت $Y = X_1 + X_2 + \cdots + X_r$ که X_i ها به صورت مقابل تعریف می‌شوند. X_i یک متغیر تصادفی هندسی با پارامتر p می‌باشد که نشان‌گر تعداد پرتاب‌ها بعد از مشاهده $i-1$ امین شیر تا ظهور i امین شیر است. با توجه به تعریف فوق X_i ها از یک‌دیگر مستقل می‌باشند و هر کدام دارای میانگین و واریانس به ترتیب $\frac{1}{p}$ و $\frac{1-p}{p^2}$ می‌باشند. پس قضیه زیر را داریم:

قضیه ۳ *اگر Y یک متغیر تصادفی دوجمله‌ای منفی با پارامترهای r و p باشد، آن‌گاه:*

$$E[Y] = \frac{r}{p}$$
$$Var(Y) = \frac{r(1-p)}{p^2}$$

۶ میانگین مجموع نامتناهی متغیر تصادفی

در این قسمت می‌خواهیم بررسی کنیم که آیا رابطه‌ی زیر همواره برای متغیرهای تصادفی دلخواه برقرار است؟

$$E[\sum_{i=1}^{\infty} X_i] = \sum_{i=1}^{\infty} E[X_i]$$

در حالت کلی پاسخ این سوال منفی است، اما تحت شرایط زیر این رابطه برقرار است:

۱. $\forall i : \Pr(X_i \geq 0) = 1$

۲. $\sum_{i=1}^{\infty} E[X_i] < \infty$