



گروه ۲

ارزیابی کارایی سیستم‌های کامپیوتری



تمرین امتیازی

متغیرهای تصادفی $X_1, X_2, \dots, X_n$ را توأم گوسی گویند اگر pdf توأم آن‌ها به صورت زیر باشد:

$$f_{\underline{X}}(\underline{x}) \triangleq f_{X_1, \dots, X_n}(x_1, \dots, x_n) = \frac{\exp \left\{ -\frac{1}{2}(\underline{x} - \underline{m})^T K^{-1}(\underline{x} - \underline{m}) \right\}}{(2\pi)^{n/2} |K|^{1/2}} \quad (1)$$

که در آن $\underline{x}$ و $\underline{m}$ بردارهای ستونی تعریف شده در زیر هستند:

$$\underline{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \underline{m} = \begin{bmatrix} m_1 \\ \vdots \\ m_n \end{bmatrix} = \begin{bmatrix} E[X_1] \\ \vdots \\ E[X_n] \end{bmatrix}$$

و K ماتریس کوواریانس تعریف شده در زیر می‌باشد:

$$K = \begin{bmatrix} \text{Cov}(X_1, X_1) & \text{Cov}(X_1, X_2) & \cdots & \text{Cov}(X_1, X_n) \\ \text{Cov}(X_2, X_1) & \text{Cov}(X_2, X_2) & \cdots & \text{Cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_n, X_1) & \text{Cov}(X_n, X_2) & \cdots & \text{Cov}(X_n, X_n) \end{bmatrix} \quad (1)$$

از آنجا که $\text{Cov}(X_i, X_j) = \text{Cov}(X_j, X_i)$ است، پس K ماتریسی متقارن است.

* تابع توزیع گوسی ۲ متغیره را با توجه به تعریف $\rho_{X,Y} = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}$ بدست آوریم:
بنابراین برای ماتریس کواریانس داریم (K ماتریس کوواریانس):

$$K = \begin{bmatrix} \sigma_1^2 & \rho_{X,Y} \sigma_1 \sigma_2 \\ \rho_{X,Y} \sigma_1 \sigma_2 & \sigma_2^2 \end{bmatrix}, \quad |K| = \sigma_1^2 \sigma_2^2 (1 - \rho_{X,Y}^2) \quad (2)$$

$$K^{-1} = \frac{1}{\sigma_1^2 \sigma_2^2 (1 - \rho_{X,Y}^2)} \begin{bmatrix} \sigma_2^2 & -\rho_{X,Y} \sigma_1 \sigma_2 \\ -\rho_{X,Y} \sigma_1 \sigma_2 & \sigma_1^2 \end{bmatrix} \quad (3)$$

در معادله (۱) جمله توان عبارت (جمله توان) متناسب است با:

$$\begin{aligned} & \frac{1}{\sigma_1^2 \sigma_2^2 (1 - \rho_{X,Y}^2)} \begin{bmatrix} x - m_1 & y - m_2 \end{bmatrix} \begin{bmatrix} \sigma_2^2 & -\rho_{X,Y} \sigma_1 \sigma_2 \\ -\rho_{X,Y} \sigma_1 \sigma_2 & \sigma_1^2 \end{bmatrix} \begin{bmatrix} x - m_1 \\ y - m_2 \end{bmatrix} \\ &= \frac{1}{\sigma_1^2 \sigma_2^2 (1 - \rho_{X,Y}^2)} \begin{bmatrix} x - m_1 & y - m_2 \end{bmatrix} \begin{bmatrix} \sigma_2^2 (x - m_1) - \rho_{X,Y} \sigma_1 \sigma_2 (y - m_2) \\ -\rho_{X,Y} \sigma_1 \sigma_2 (x - m_1) + \sigma_1^2 (y - m_2) \end{bmatrix} \quad (۴) \\ &= \frac{\left(\frac{x-m_1}{\sigma_1}\right)^2 - 2\rho_{X,Y} \left(\frac{x-m_1}{\sigma_1}\right) \left(\frac{y-m_2}{\sigma_2}\right) + \left(\frac{y-m_2}{\sigma_2}\right)^2}{(1 - \rho_{X,Y}^2)} \end{aligned}$$

مثال: بردار متغیرهای تصادفی (X, Y, Z) دارای توزیع توأم گوسی با میانگین صفر و ماتریس کوواریانس زیر را در نظر بگیرید.

$$K = \begin{bmatrix} \text{Cov}(X, X) & \text{Cov}(X, Y) & \text{Cov}(X, Z) \\ \text{Cov}(Y, X) & \text{Cov}(Y, Y) & \text{Cov}(Y, Z) \\ \text{Cov}(Z, X) & \text{Cov}(Z, Y) & \text{Cov}(Z, Z) \end{bmatrix} = \begin{bmatrix} 1 & 0.2 & 0.3 \\ 0.2 & 1 & 0.4 \\ 0.3 & 0.4 & 1 \end{bmatrix}$$

مطلوبست توزیع حاشیه‌ای X, Z را بدست آورید.

جواب: ۲ روش حل داریم:

۱- انتگرال گیری مستقیم pdf برای یافتن pdf حاشیه‌ای.

۲- استفاده از این واقعیت که pdf حاشیه‌ای X, Z نیز یک گوسی است که دارای همان مجموعه میانگین‌ها، واریانس‌ها و کوواریانس مشترک است.

بنابراین طبق روش ۲، pdf توأم (X, Z) دارای بردار میانگین صفر و ماتریس کوواریانس به شکل زیر می‌باشد:

$$K = \begin{bmatrix} \text{Cov}(X, X) & \text{Cov}(X, Z) \\ \text{Cov}(Z, X) & \text{Cov}(Z, Z) \end{bmatrix} = \begin{bmatrix} 1 & 0.3 \\ 0.3 & 1 \end{bmatrix}$$

مثال: اگر $X_1, \dots, X_n$ متغیرهای تصادفی با $\text{Cov}(X_i, X_j) = 0 \quad (i \neq j)$ باشند و دارای توزیع توأم گوسی باشند، نشان دهید متغیرهای تصادفی مستقل هستند.

ماتریس کوواریانس قطری است و داریم:

$$K = \text{diag}[\text{VAR}(X_i)] = \text{diag}[\sigma_i^2] \quad , \quad K^{-1} = \text{diag}\left[\frac{1}{\sigma_i^2}\right]$$

$$(\underline{x} - \underline{m})^T K^{-1} (\underline{x} - \underline{m}) = \sum_{i=1}^n \frac{(x_i - m_i)^2}{\sigma_i^2}$$

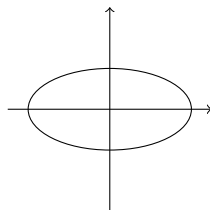
$$\begin{aligned} f_{\underline{X}}(\underline{x}) &= \frac{\exp \left\{ -1/2 \sum_{i=1}^n [(x_i - m_i)/\sigma_i]^2 \right\}}{(2\pi)^{n/2} |K|^{1/2}} \\ &= \prod_{i=1}^n \frac{\exp \left\{ -1/2 [(x_i - m_i)/\sigma_i]^2 \right\}}{\sqrt{2\pi\sigma_i^2}} \\ &= \prod_{i=1}^n f_{X_i}(x_i) \quad \blacksquare \end{aligned}$$

رابطه (۴) جمله نما در گوسی ۲ متغیره را نشان می‌دهد.

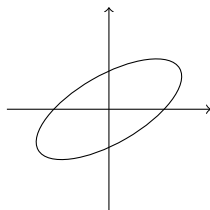
شکل تابع توزیع گوسی ۲ متغیره سطحی زنگوله‌ای است که مرکز آن (m_1, m_2) می‌باشد و مقادیر σ_1, σ_2 و ρ_{XY} شکل زنگوله را مشخص می‌کنند.

همه مقادیر x, y که به ازای آن‌ها مقدار نما، عدد ثابتی باشد، مجموعه سطح (level set) و یا کانتور تابع توزیع را بیان می‌کنند.

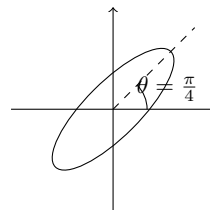
شکل‌های زیر به ازای مقادیر مختلف $\sigma_1, \sigma_2, \rho_{XY}$ کانتورهای pdf ثابت را نشان می‌دهند:



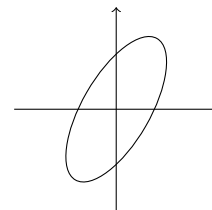
$$\rho_{XY} = 0$$



$$\begin{aligned} \sigma_1 &> \sigma_2 \\ 0 &< \theta < \pi/4 \\ \rho_{XY} &\neq 0 \end{aligned}$$



$$\begin{aligned} \sigma_1 &= \sigma_2 \\ \theta &= \pi/4 \\ \rho_{XY} &\neq 0 \end{aligned}$$



$$\begin{aligned} \sigma_1 &< \sigma_2 \\ \pi/4 &< \theta < \pi/2 \\ \rho_{XY} &\neq 0 \end{aligned}$$

مقدار زاویه θ عبارت است از:

$$\theta = \frac{1}{2} \tan^{-1} \left(\frac{2\rho_{XY}\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2} \right)$$

که در آن برای $\sigma_1^2 = \sigma_2^2$ ، $\tan^{-1} \infty$ ، همان $\pi/2$ است و $\theta = \pi/4$

مثال: در حالت خاص رابطه (۴) وقتی $m_1 = m_2 = 0$ و $\sigma_1^2 = \sigma_2^2 = 1$ است، از روی توزیع توأم که به صورت زیر در

می‌آید، توزیع حاشیه‌ای Y, X را با روش انتگرال‌گیری بدست آورید.

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{\frac{-(x^2-2\rho xy+y^2)}{2(1-\rho^2)}} \quad -\infty < x, y < \infty$$

توزیع حاشیه‌ای X با انتگرال‌گیری $f_{X,Y}(x, y)$ روی y بدست می‌آید:

$$f_X(x) = \frac{e^{-x^2/2(1-\rho^2)}}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-(y^2-2\rho xy)/2(1-\rho^2)} dy$$

با اضافه و کم کردن عبارت درون انتگرال را به صورت مربع کامل در می‌آوریم. یعنی:

$$y^2 - 2\rho xy + \rho^2 x^2 - \rho^2 x^2 = (y - \rho x)^2 - \rho^2 x^2$$

بنابراین:

$$\begin{aligned} f_X(x) &= \frac{e^{-x^2/2(1-\rho^2)}}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\{(y-\rho x)^2 - \rho^2 x^2\}/2(1-\rho^2)} dy \\ &= \frac{e^{-x^2/2(1-\rho^2)} \times e^{\rho^2 x^2/2(1-\rho^2)}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{e^{-[(y-\rho x)^2]/2(1-\rho^2)}}{\sqrt{2\pi(1-\rho^2)}} dy \\ &= \frac{e^{-x^2/2}}{\sqrt{2\pi}} \end{aligned}$$

مقدار انتگرال 1 است، چون درون انتگرال، تابع توزیع گاوسی با میانگین ρx و واریانس $1 - \rho^2$ است. مشاهده می‌شود، توزیع حاشیه‌ای X یک pdf گاوسی با میانگین 0 و واریانس 1 است. به همین ترتیب توزیع حاشیه‌ای Y بدست می‌آید.

تمرین: از روی توزیع توأم دو متغیره در فرمول زیر، pdf شرطی $f_{X|Y}(x|y)$ و $f_{Y|X}(y|x)$ را بدست آورید و نشان دهید مثلاً pdf X به شرط $Y = y$ یک گاوسی است که دارای میانگین شرطی $(\frac{\sigma_1}{\sigma_2})(y - m_2)$ و $m_1 + \rho_{X,Y}(\frac{\sigma_1}{\sigma_2})(y - m_2)$ واریانس شرطی $\sigma_1^2(1 - \rho_{X,Y}^2)$ است.

توزیع توأم نرمال دومتغیره :

$$\begin{aligned} f_{X,Y}(x, y) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho_{X,Y}^2}} \\ &\times \exp\left(-\frac{1}{2(1-\rho_{X,Y}^2)}\left[\left(\frac{x-m_1}{\sigma_1}\right)^2 - 2\rho_{X,Y}\left(\frac{x-m_1}{\sigma_1}\right)\left(\frac{y-m_2}{\sigma_2}\right) + \left(\frac{y-m_2}{\sigma_2}\right)^2\right]\right) \end{aligned}$$

سپس نشان دهید وقتی $\rho_{XY} = 0$ است، توزیع شرطی X به شرط $Y = y$ همان توزیع حاشیه‌ای X می‌باشد. یعنی X و Y مستقل هستند.

نکته مهم: اگر X و Y متغیر تصادفی مستقل باشند، $\rho_{XY} = 0$ است ولی اگر $\rho_{XY} = 0$ باشد، لزوماً X و Y مستقل نیستند.

تنها موردی که $\rho_{XY} = 0$ منجر به استقلال X و Y می‌شود، دقیقاً موقعی است که X و Y متغیر تصادفی گاوسی باشند.

امید متغیر تصادفی:

سؤال: توزیع دو جمله‌ای منفی را در نظر بگیرید

$$Y \sim \text{NegBin}(r, p) \quad E[Y] = \frac{r}{p}$$

در مورد Y_i کدام گزینه صحیح است؟

$$Y = \sum_{i=1}^? Y_i$$

(A) $Y_i : i$ امین امتحان شیر باشد، یعنی $Y_i \sim \text{Ber}(p)$ ، $i = 1, \dots, r$

(B) Y_i : تعداد امتحان برای بدست آوردن i امین موفقیت (پس از مشاهده $(i - 1)$ امین موفقیت). یعنی $Y_i \sim \text{Geo}(p)$ ، $i = 1, \dots, r$

(C) Y_i : تعداد موفقیت در n امتحان، یعنی که در مقابل به دنبال $P(Y_i = 1)$ هستیم و $Y_i \sim \text{Bin}(n, p)$ ، $i = 1, \dots, r$

جواب: B

مشخص کنید که هر مورد زیر عدد و یا تابع است.

۱- $P(Y = 2 | Y = 5)$ [عدد] ۳- $P(X = 2 | Y = y)$ [تابع y]

۲- $P(X = x | Y = 5)$ [تابع x] ۴- $P(X = x | Y = y)$ [تابع x, y]

از موارد زیر کدام صحیح و کدام غلط است

[صحیح]

$$\sum_x p(X = x | Y = 5) = 1 - 1$$

[غلط]

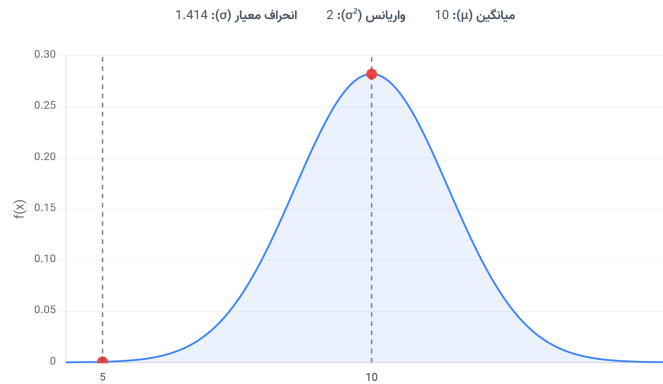
$$\sum_y p(X = 2 | Y = y) = 1 - 2$$

[صحیح]

$$\sum_x \underbrace{\left(\sum_y \underbrace{p(X = x | Y = y) p(Y = y)}_{P(X=x, Y=y)} \right)}_{P(X=x)} = 1 - 3$$

احتمال‌های نسبی در متغیرهای تصادفی پیوسته:

نکته: $f(x) > 1$ صحیح است، چرا؟ چون $f(x)$ احتمال نیست، بلکه بیانگر میزان چگالی توزیع در نقطه x است.



شکل ۱: نمودار تابع چگالی احتمال توزیع نرمال

فرض کنید X = زمان لازم برای حل یک مسأله باشد که دارای توزیع نرمال است

$$X \sim \mathcal{N}(10, 2) \quad (\sigma^2 = 2, \mu = 10)$$

چقدر محتمل‌تر است که مسأله را در مدت زمان ۱۰ دقیقه نسبت به مدت زمان ۵ دقیقه حل کنند؟

$$\frac{P(X = 10)}{P(X = 5)} = \begin{cases} A = \frac{0}{0} = \text{تعریف} \\ B = \frac{f(10)}{f(5)} \end{cases}$$

می‌دانیم در RV پیوسته، احتمال هر نقطه صفر است یعنی $P(X = 10) = P(X = 5) = 0$ ، برخلاف استدلال صحیح فوق، جواب B است.

زیرا در پیوسته:

$$P(X = a) = \lim_{\epsilon \rightarrow 0} P\left(a - \frac{\epsilon}{2} \leq X \leq a + \frac{\epsilon}{2}\right) = \lim_{\epsilon \rightarrow 0} \int_{a-\frac{\epsilon}{2}}^{a+\frac{\epsilon}{2}} f(x) dx = \epsilon f(a)$$

$$\Rightarrow \frac{P(X = a)}{P(X = b)} = \frac{\epsilon f(a)}{\epsilon f(b)} = \frac{f(a)}{f(b)}$$

$$\frac{f(10)}{f(5)} = \frac{\frac{1}{\sigma\sqrt{2\pi}} e^{-(10-\mu)^2/2\sigma^2}}{\frac{1}{\sigma\sqrt{2\pi}} e^{-(5-\mu)^2/2\sigma^2}}$$

$$= \frac{e^{-\frac{(10-10)^2}{2 \times 2}}}{e^{-\frac{(5-10)^2}{2 \times 2}}}$$

$$= \frac{e^0}{e^{-25/4}}$$

توزیع شرطی پیوسته: برای متغیرهای تصادفی پیوسته X و Y ، pdf شرطی X به شرط Y عبارت است از:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

به طور شهودی:

$$P(X = x|Y = y) = \frac{P(X=x, Y=y)}{P(Y=y)} \quad \longleftrightarrow \quad f_{X|Y}(x|y)\epsilon_X = \frac{f_{X,Y}(x,y)\epsilon_X\epsilon_Y}{f_Y(y)\epsilon_Y}$$

گسسته

پیوسته

اثبات می‌شود pdf شرطی $f_{X|Y}$ یک توزیع احتمال است. زیرا:

$$\int_{-\infty}^{\infty} f_{X|Y}(x|y) dx = \int_{-\infty}^{\infty} \frac{f_{X,Y}(x,y)}{f_Y(y)} dx = \frac{\int_{-\infty}^{\infty} f_{X,Y}(x,y) dx}{f_Y(y)} = \frac{f_Y(y)}{f_Y(y)} = 1$$

سؤال:

عفونت ویروسی: دو گروه مبتلا به بیماری ویروسی داریم

□ $G1$: 200 نفر که هر نفر مستقلاً با احتمال $P_1 = 0.1$ آلوده است

□ $G2$: 100 نفر که هر نفر مستقلاً با احتمال $P_2 = 0.4$ آلوده است

می‌خواهیم ($55 > \text{تعداد نفرات آلوده}$) P را بدست آوریم. که این یکی از استراتژی‌های زیر برای یافتن جواب بهتر است. ابتدا صورت سؤال را به صورت رسمی بیان می‌کنیم. یعنی RV ها و هدف را می‌نویسیم:
فرض کنید: $A = \text{تعداد نفرات آلوده در } G1$ باشد

$$A \sim \text{Bin}(200, 0.1)$$

و $B = \text{تعداد نفرات آلوده در } G2$

$$B \sim \text{Bin}(100, 0.4)$$

می‌خواهیم ($55 < A + B$) را بدست آوریم.

استراتژی ها :

A - کانولوشن : کار می‌کند ولی آزار دهنده است.

B - جمع متغیرهای ۲ جمله‌ای مستقل

C - جواب تقریبی به کمک جمع متغیرهای تصادفی پواسن مستقل (روش غلط، چون $np = 20$ است، و این فاصله به حد کمال از تقریب پواسن بزرگ نیست).

D - جواب تقریبی به کمک جمع متغیر تصادفی نرمال مستقل

E - هیچکدام از روش‌ها جواب دقیق نیست (به روشی دیگر نیاز داریم) (غلط است، چون D جواب مسأله است)

طبق روش D داریم $\leftarrow$ به تقریب نرمال

$$A \simeq X \sim \mathcal{N}(20, 18) \quad \mu_1 \simeq np = 200 \times 0.1 = 20$$

$$\sigma_1^2 \simeq np(1-p) = 20(1-0.1) = 18$$

$$B \simeq Y \sim \mathcal{N}(40, 24) \quad \mu_2 \simeq np = 100 \times 0.4 = 40$$

$$\sigma_2^2 \simeq np(1-p) = 40(1-0.4) = 24$$

حال: $W = X + Y$ عبارت است از:

جمع ۲ نرمال مستقل

↓

$$W = X + Y \sim \mathcal{N}(20 + 40 = 60, 18 + 24 = 42)$$

$$P(A + B > 55) \simeq P(W \geq 54.5)$$

عبارت فوق طبق اصل مربوط به تصحیح پیوستگی در تقریب ۲ جمله‌ای به نرمال بدست آمده است.

$$P(W \geq 54.5) = 1 - \Phi\left(\frac{54.5 - 60}{\sqrt{42}}\right) = 1 - \Phi(-0.85) \simeq 0.8023$$

برنامه‌ای کامپیوتری بنویسید که جواب اصلی $(P(A + B > 55))$ را برای RVهای ۲ جمله‌ای محاسبه کند و تحقیق کنید که جواب تقریبی $(P(W \geq 54.5))$ بهتر از جواب $P(W \geq 55)$ می‌باشد.

اگر به جمع ۲ متغیر تصادفی مستقل که در بخش قبل ذکر شد دقت کنید، در واقع در حال محاسبه کانولوشن توابع چگالی احتمال هستیم. به طور خاص:

کانولوشن P_X و P_Y برای متغیر تصادفی گسسته مستقل X و Y ، توزیع احتمال مجموع $(X + Y)$ را بدست می‌دهد.

$$P(X + Y = n) = \sum_k P(X = k)P(Y = n - k)$$

کانولوشن f_X و f_Y برای متغیرهای تصادفی پیوسته مستقل X و Y ، چگالی احتمال مجموع X, Y را بدست می‌دهد.

$$f_{X+Y}\left(\underset{\substack{\downarrow \\ \text{در گسسته } n \text{ مشابه}}}{\alpha}\right) = \int_{-\infty}^{\infty} f_X\left(\underset{\substack{\downarrow \\ \text{در گسسته } k \text{ مشابه}}}{k}\right) f_Y\left(\underset{\substack{\downarrow \\ \text{در گسسته } n-k \text{ مشابه}}}{\alpha - k}\right) dk$$

جمع متغیرهای تصادفی و محاسبه توزیع احتمال آنها برای ۳ نوع توزیع دو جمله‌ای، پواسون و نرمال را بحث کردیم. در زیر جمع متغیر تصادفی مستقل با توزیع یکنواخت را با ذکر مثال بررسی می‌کنیم که در واقع یک محاسبه کانولوشن است.

مثال:

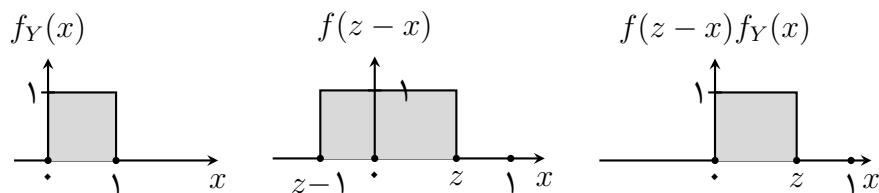
$$X \sim \text{Uni}(0, 1) \quad \text{و} \quad Y \sim \text{Uni}(0, 1)$$

متغیر تصادفی مستقل هستند. توزیع $X + Y$ یعنی f_{X+Y} را حساب کنید (مثال ۸.۶ پاپولیس چاپ ۴ صفحه ۱۸۳) فرض کنید $Z = X + Y$ باشد، می‌خواهیم $f_Z(z)$ را حساب کنیم. با توجه به آنکه حداقل و حداکثر X, Y ، ۰، ۱ هستند حداقل و حداکثر Z ، ۰ و ۲ می‌باشد.

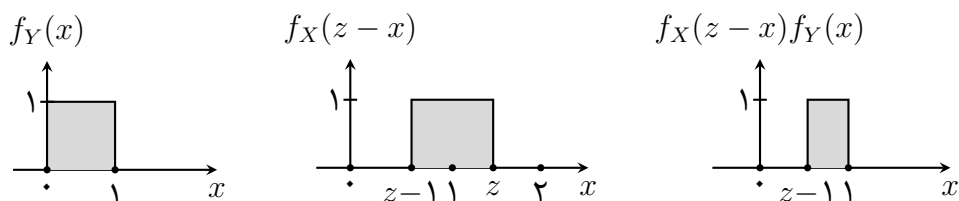
$$f_Z(z) = \int f_X(z - x)f_Y(x)dx$$

انتگرال فوق را برای ۲ حالت z :

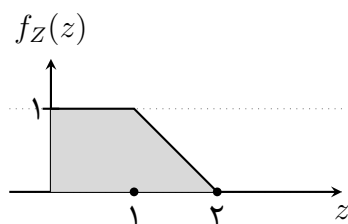
$$\begin{cases} 0 \leq z < 1 \\ 1 \leq z \leq 2 \end{cases} \quad \text{جداگانه بدست می‌آوریم.}$$

الف) $0 \leq z \leq 1$ 

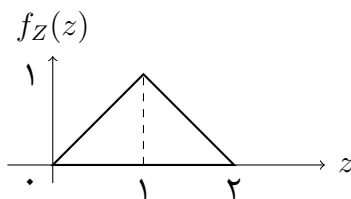
$$f_Z(z) = \int f_X(z-x)f_Y(x) dx = \int_z^0 1x dx = x \Big|_0^z = z - 0 = z$$

ب) $1 \leq z \leq 2$ 

$$f_Z(z) = \int f_X(z-x)f_Y(x) dx = \int_{z-1}^1 (1)(x) dx = x \Big|_{z-1}^1 = 1 - (z-1) = 2-z$$



با توجه به موجود بودن $f_Z(z)$ در قسمت الف و ب آن‌ها را با هم ترکیب می‌کنیم و چگالی $f_Z(z)$ به صورت زیر ترسیم می‌شود.



سوال:

اگر X و Y مستقل نباشند $f_Z(z)$ را حساب کنید (پاسخ صفحه ۱۸۱ کتاب پاپولیس).

$$f_Z(z) = \int_{-\infty}^{\infty} f_{XY}(z - y, y) dy = \int_{-\infty}^{\infty} f_{XY}(x, z - x) dx$$

کوواریانس و همبستگی

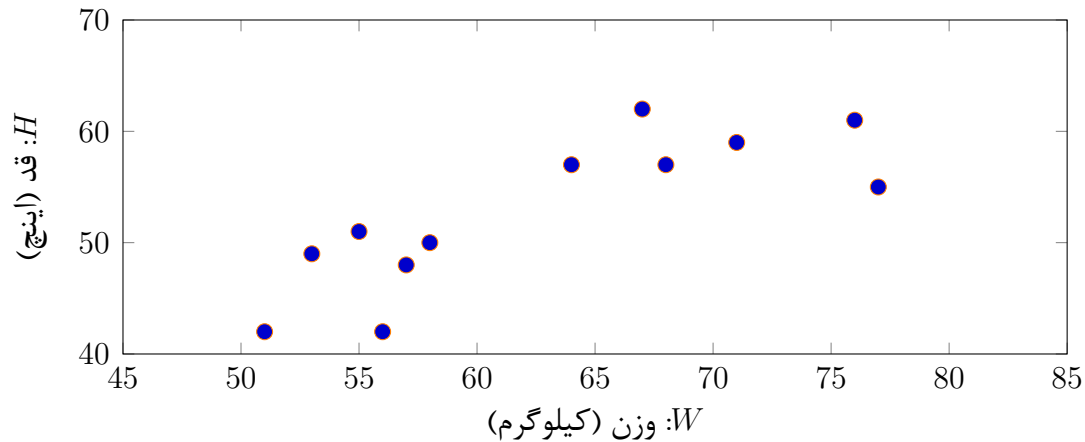
مثال: وزن و قد

جدول مقابل وزن بر حسب کیلوگرم و قد بر حسب اینچ ۱۲ نفر را نشان می‌دهد. در مورد کوواریانس این دو متغیر تصادفی بحث کنید.

$$E[W] = 62.75 \quad E[H] = 52.75 \quad E[WH] = 3355.83$$

$$\text{Cov}(W, H) = E[WH] - E[W]E[H] = 3355.83 - (62.75)(52.75) = 45.77 \quad (\text{positive correlation})$$

$W \cdot H$	قد (اینچ)	وزن (کیلوگرم)
۳۶۴۸	۵۷	۶۴
۴۱۸۹	۵۹	۷۱
۲۵۹۷	۴۹	۵۳
۴۱۵۴	۶۲	۶۷
۲۸۰۵	۵۱	۵۵
۲۹۰۰	۵۰	۵۸
۴۲۳۵	۵۵	۷۷
۲۷۳۶	۴۸	۵۷
۲۳۵۲	۴۲	۵۶
۲۱۴۲	۴۲	۵۱
۴۶۳۶	۶۱	۷۶
۳۸۷۶	۵۷	۶۸

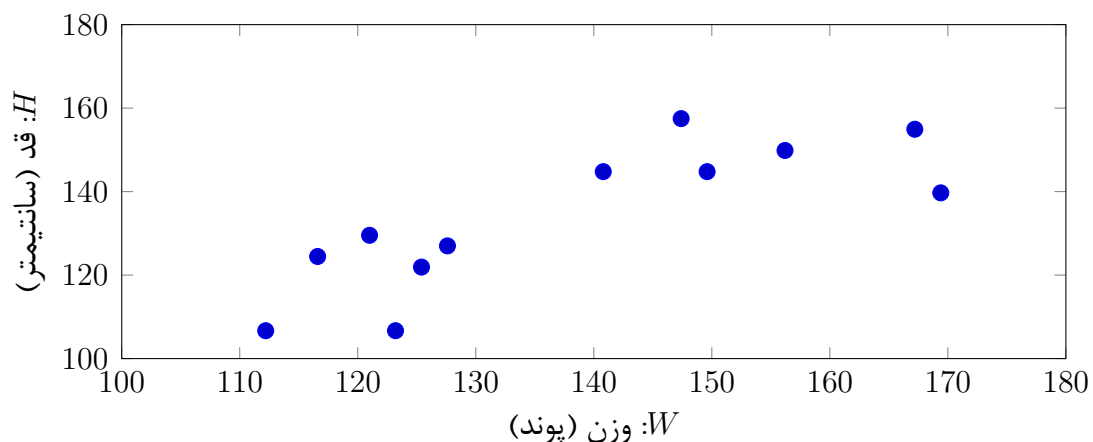


شکل ۲: نمودار پراکندگی وزن و قد (کیلوگرم و اینچ)

تغییر واحدها

اگر وزن بر حسب پوند و قد بر حسب سانتیمتر باشد، کوواریانس این دو متغیر تصادفی را به دست آورید. مشاهده می‌نمایید که همان داده منجر به نتایج متفاوتی می‌شود (هر اینچ معادل ۲.۵۴ سانتیمتر و هر کیلوگرم معادل ۲.۲ پوند می‌باشد). این مثال نشان می‌دهد: برای کوواریانس، **علامت** کوواریانس گویاتر از مقدار کوواریانس می‌باشد.

$$\begin{aligned} \text{Cov}(2.2W, 2.54H) &= E[2.2W \times 2.54H] - E[2.2W]E[2.54H] = \\ 18752.38 - (138.05)(133.99) &= 255.06 \quad (\text{positive correlation}) \end{aligned}$$



شکل ۳: نمودار پراکندگی وزن و قد (پوند و سانتیمتر)

نتیجه‌گیری: بنابراین نیازمند مفهومی موسوم به **همبستگی** داریم که وابستگی به واحد داده‌ها نداشته باشد.

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

که در آن:

$$\sigma_X^2 = \text{Var}(X),$$

$$\sigma_Y^2 = \text{Var}(Y),$$

و همواره:

$$-1 \leq \rho(X, Y) \leq 1$$

همبستگی، معیاری برای بیان رابطه خطی بین X و Y است. سه مورد زیر را در نظر بگیرید.
در دو مورد اول رابطه خطی بین X و Y وجود دارد و متغیرها همبسته شدید هستند:

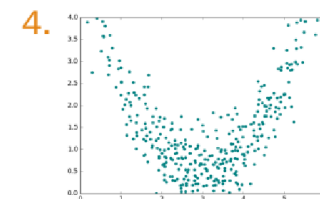
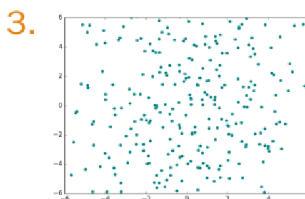
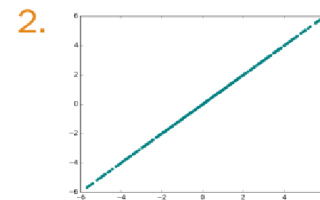
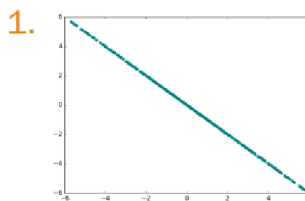
$$\square \rho(X, Y) = 1 \rightarrow Y = aX + b, \quad \text{where} \quad a = \frac{\sigma_Y}{\sigma_X} \quad (\text{همبستگی مثبت})$$

$$\square \rho(X, Y) = -1 \rightarrow Y = aX + b, \quad \text{where} \quad a = -\frac{\sigma_Y}{\sigma_X} \quad (\text{همبستگی منفی})$$

$$\square \rho(X, Y) = 0 \rightarrow X \text{ و } Y \text{ ناهمبسته هستند (در واقع رابطه خطی بین } X \text{ و } Y \text{ وجود ندارد)}$$

تمرین: ارتباط شکل‌ها با روابط

ارتباط شکل‌های ۱ تا ۴ را با روابط A, B, C, D مشخص نمایید.



A. $\rho(X, Y) = 1$

B. $\rho(X, Y) = -1$

C. $\rho(X, Y) = 0$

D. Other

پاسخ:

B. شکل ۱: $(Y = -\frac{\sigma_Y}{\sigma_X}X + b)$

خطی با ضریب زاویه منفی است. به مقدار ضریب زاویه هم توجه کنید و سعی کنید آن را متوجه شوید.

A. شکل ۲: $(Y = \frac{\sigma_Y}{\sigma_X}X + b)$

خطی با ضریب زاویه مثبت است. به مقدار ضریب زاویه هم توجه کنید و سعی کنید آن را متوجه شوید.

C. شکل ۳: $\rho(X, Y) = 0$ (ناهمبسته)

C. شکل ۴: $\rho(X, Y) = 0$ برای $Y = X^2$

همانطور که تاکید شد همبستگی "خطی بودن" را اندازه می‌گیرد و شکل ۴ نشان می‌دهد با آنکه کوواریانس X و Y صفر می‌باشد، این دو متغیر به طور "غیر خطی" با هم رابطه دارند.

مثال دیگر از اینکه کوواریانس صفر بیانگر استقلال نیست

فرض کنید متغیر تصادفی X هر یک از مقادیر قلمرو خود $\{-1, 0, 1\}$ را با احتمال $\frac{1}{3}$ اختیار می‌کند. متغیر تصادفی Y را به صورت زیر تعریف می‌کنیم:

$$Y = \begin{cases} 1 & \text{if } X = 0 \\ 0 & \text{otherwise} \end{cases}$$

کدام گزینه زیر در مورد تابع جرم احتمال توام X و Y صحیح است؟ و توزیع های حاشیه ای را بنویسید.

A.

		X		
		-1	0	1
Y	0	1/6	1/6	1/6
	1	1/6	1/6	1/6

B.

		X		
		-1	0	1
Y	0	1/3	0	1/3
	1	0	1/3	0

C.

		X		
		-1	0	1
Y	0	0	1/3	0
	1	1/3	0	1/3

جواب: B

		X			
		-1	0	1	
Y	0	1/3	0	1/3	2/3
	1	0	1/3	0	1/3
		1/3	1/3	1/3	

Marginal PMF of Y , $p_Y(y)$

Marginal PMF of X , $p_X(x)$

به سوال های زیر پاسخ دهید:

۱. $E[X]$ و $E[Y]$ را بیابید.
۲. $E[XY]$ را بیابید.
۳. $\text{Cov}(X, Y)$ را محاسبه کنید.
۴. آیا X و Y مستقل هستند؟

پاسخ

۱. $E[X] = 0$ $E[Y] = \frac{1}{3}$
۲. $E[XY] = 0$ XY همواره صفر است، بنابراین
۳. $\text{Cov}(X, Y) = E[XY] - E[X]E[Y] = 0$

۴. آیا مستقل هستند؟ خیر

$$P[Y = 0] = \frac{2}{3} \neq P(Y = 0|X = 0) = 0$$

واریانس مجموع متغیرهای تصادفی

برای دو متغیر تصادفی X و Y به طور کلی رابطه زیر برقرار است:

$$\text{Var}(X + Y) = \text{Var}(X) + 2\text{Cov}(X, Y) + \text{Var}(Y) = 0$$

اثبات:

$$\begin{aligned}\text{Var}(X + Y) &= \text{Cov}(X + Y, X + Y) \\ &= \text{Cov}(X, X) + \text{Cov}(X, Y) + \text{Cov}(Y, X) + \text{Cov}(Y, Y) \\ &= \text{Var}(X) + 2 \cdot \text{Cov}(X, Y) + \text{Var}(Y)\end{aligned}$$

توضیحات: $\text{Var}(X) = \text{Cov}(X, X)$ و تقارن کوواریانس $\text{Cov}(X, Y) = \text{Cov}(Y, X)$.

تعمیم فوق به صورت زیر است:

$$\text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i) + 2 \sum_{i=1}^n \sum_{j=i+1}^n \text{Cov}(X_i, X_j)$$

اثبات:

$$\begin{aligned}\text{Var}\left(\sum_{i=1}^n X_i\right) &= \text{Cov}\left(\sum_{i=1}^n X_i, \sum_{j=1}^n X_j\right) \\ &= \sum_{i=1}^n \sum_{j=1}^n \text{Cov}(X_i, X_j) \\ &= \sum_{i=1}^n \text{Var}(X_i) + \sum_{i=1}^n \sum_{j=1, j \neq i}^n \text{Cov}(X_i, X_j) \\ &= \sum_{i=1}^n \text{Var}(X_i) + 2 \sum_{i=1}^n \sum_{j=i+1}^n \text{Cov}(X_i, X_j)\end{aligned}$$

سوال:

روابط فوق در صورت استقلال متغیرها چگونه می‌باشد؟

۱ استقلال

استقلال یکی از مفاهیم بنیادی نظریه‌ی احتمال و آمار است. به‌طور کلی، استقلال را در دو سطح بررسی می‌کنیم:

۱. استقلال پیشامدها؛

۲. استقلال متغیرهای تصادفی.

در این فصل، ابتدا استقلال پیشامدها و سپس استقلال متغیرهای تصادفی گسسته و پیوسته را بررسی می‌کنیم.

۲ استقلال پیشامدها

۱.۲ تعریف استقلال دو پیشامد

فرض کنید E و F دو پیشامد در فضای نمونه‌ی Ω باشند. به‌صورت شهودی، E و F مستقل‌اند اگر وقوع یکی از آن‌ها، احتمال وقوع دیگری را تغییر ندهد. استقلال دو پیشامد را می‌توان با دو تعریف معادل بیان کرد.

تعریف اول: تعریف ضربی

دو پیشامد E و F مستقل‌اند اگر و تنها اگر

$$\Pr(E \cap F) = \Pr(E) \Pr(F).$$

این تعریف عمومی‌تر است، زیرا حتی هنگامی که یکی از پیشامدها احتمال صفر دارد نیز قابل استفاده است.

تعریف دوم: تعریف با احتمال شرطی

اگر $\Pr(F) > 0$ باشد، E و F مستقل‌اند اگر

$$\Pr(E | F) = \Pr(E).$$

به همین ترتیب، اگر $\Pr(E) > 0$ باشد، می‌توان نوشت

$$\Pr(F | E) = \Pr(F).$$

یعنی اطلاع از وقوع F احتمال E را تغییر نمی‌دهد و اطلاع از وقوع E نیز احتمال F را تغییر نمی‌دهد. معادل بودن دو تعریف از رابطه‌ی احتمال شرطی نتیجه می‌شود:

$$\Pr(E | F) = \frac{\Pr(E \cap F)}{\Pr(F)} = \Pr(E) \iff \Pr(E \cap F) = \Pr(E) \Pr(F).$$

۲.۲ مثال اول: پرتاب دو تاس

دو تاس سالم و مستقل را پرتاب می‌کنیم. نتیجه‌ی تاس اول و دوم را به ترتیب با D_1 و D_2 نشان می‌دهیم. فضای نمونه برابر است با

$$\Omega = \{(d_1, d_2) : d_1, d_2 \in \{1, 2, 3, 4, 5, 6\}\}.$$

این فضا 36 عضو هم‌احتمال دارد. پیشامدهای زیر را تعریف می‌کنیم:

$$E = \{D_1 = 1\}, \quad F = \{D_2 = 6\}, \quad G = \{D_1 + D_2 = 5\}.$$

اعضای این پیشامدها عبارت‌اند از

$$E = \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6)\},$$

$$F = \{(1, 6), (2, 6), (3, 6), (4, 6), (5, 6), (6, 6)\},$$

$$G = \{(1, 4), (2, 3), (3, 2), (4, 1)\}.$$

آیا E و F مستقل‌اند؟

احتمال‌های حاشیه‌ای برابرند با

$$\Pr(E) = \frac{6}{36} = \frac{1}{6}, \quad \Pr(F) = \frac{6}{36} = \frac{1}{6}.$$

اشتراک دو پیشامد فقط شامل حالت $(1, 6)$ است:

$$E \cap F = \{(1, 6)\}.$$

بنابراین

$$\Pr(E \cap F) = \frac{1}{36} = \frac{1}{6} \times \frac{1}{6} = \Pr(E) \Pr(F).$$

پس E و F مستقل‌اند. این نتیجه با شهود نیز سازگار است، زیرا نتیجه‌ی تاس اول هیچ اطلاعاتی درباره‌ی نتیجه‌ی تاس دوم نمی‌دهد.

آیا E و G مستقل‌اند؟

برای پیشامد G داریم

$$\Pr(G) = \frac{4}{36} = \frac{1}{9}.$$

اشتراک E و G فقط شامل حالت $(1, 4)$ است، پس

$$E \cap G = \{(1, 4)\}, \quad \Pr(E \cap G) = \frac{1}{36}.$$

اما

$$\Pr(E) \Pr(G) = \frac{1}{6} \times \frac{1}{9} = \frac{1}{54} \neq \frac{1}{36}.$$

بنابراین E و G مستقل نیستند. با استفاده از احتمال شرطی نیز داریم

$$\Pr(G | E) = \frac{1}{6} \neq \frac{1}{9} = \Pr(G).$$

پس دانستن اینکه تاس اول برابر یک شده است، احتمال رخ دادن مجموع پنج را تغییر می‌دهد.

۳.۲ تفاوت پیشامدهای ناسازگار و مستقل

دو پیشامد E و F را ناسازگار یا disjoint می‌نامیم اگر

$$E \cap F = \emptyset.$$

ناسازگار بودن یعنی دو پیشامد نمی‌توانند هم‌زمان رخ دهند. اگر هر دو پیشامد احتمال غیرصفر داشته باشند، داریم

$$\Pr(E \cap F) = 0,$$

در حالی که

$$\Pr(E) \Pr(F) > 0.$$

پس

$$\Pr(E \cap F) \neq \Pr(E) \Pr(F)$$

و دو پیشامد مستقل نیستند.

در واقع، پیشامدهای ناسازگار با احتمال غیرصفر به شدت به یکدیگر وابسته‌اند؛ زیرا وقوع یکی از آن‌ها نشان می‌دهد که دیگری حتماً رخ نداده است. تنها زمانی ممکن است دو پیشامد ناسازگار مستقل باشند که حداقل یکی از آن‌ها احتمال صفر داشته باشد.

۴.۲ استقلال یک پیشامد و متمم پیشامد دیگر

فرض کنید E و F مستقل باشند. می‌خواهیم ثابت کنیم E و متمم F ، یعنی F^c ، نیز مستقل‌اند. پیشامد E را می‌توان به دو قسمت ناسازگار تقسیم کرد:

$$E = (E \cap F) \cup (E \cap F^c).$$

چون دو قسمت سمت راست ناسازگارند،

$$\Pr(E) = \Pr(E \cap F) + \Pr(E \cap F^c).$$

در نتیجه،

$$\begin{aligned} \Pr(E \cap F^c) &= \Pr(E) - \Pr(E \cap F) \\ &= \Pr(E) - \Pr(E) \Pr(F) \\ &= \Pr(E)(1 - \Pr(F)) \\ &= \Pr(E) \Pr(F^c). \end{aligned}$$

در خط دوم از استقلال E و F و در خط آخر از رابطه‌ی $\Pr(F^c) = 1 - \Pr(F)$ استفاده کردیم. بنابراین

$$\Pr(E \cap F^c) = \Pr(E) \Pr(F^c),$$

پس E و F^c مستقل‌اند.

به همین روش می‌توان نشان داد که اگر E و F مستقل باشند، زوج‌های زیر نیز مستقل‌اند:

$$(E^c, F), \quad (E, F^c), \quad (E^c, F^c).$$

۵.۲ استقلال بیش از دو پیشامد

برای بیش از دو پیشامد، استقلال زوجی به‌تنهایی برای نتیجه‌گیری استقلال کامل کافی نیست. سه پیشامد E ، F و G مستقل زوجی‌اند اگر هر سه شرط زیر برقرار باشند:

$$\Pr(E \cap F) = \Pr(E) \Pr(F),$$

$$\Pr(E \cap G) = \Pr(E) \Pr(G),$$

$$\Pr(F \cap G) = \Pr(F) \Pr(G).$$

اما برای اینکه این سه پیشامد **مستقل متقابل** یا مستقل مشترک باشند، علاوه بر سه شرط زوجی، شرط سه‌تایی زیر نیز باید برقرار باشد:

$$\Pr(E \cap F \cap G) = \Pr(E) \Pr(F) \Pr(G).$$

بنابراین استقلال سه پیشامد در مجموع به چهار شرط نیاز دارد: سه شرط دوتایی و یک شرط سه‌تایی.

۶.۲ مثال دوم: استقلال زوجی بدون استقلال سه‌تایی

دوباره دو تاس سالم را پرتاب می‌کنیم و نتیجه‌های آن‌ها را با D_1 و D_2 نشان می‌دهیم. این بار پیشامدهای زیر را در نظر می‌گیریم:

$$E = \{D_1 = 1\}, \quad F = \{D_2 = 6\}, \quad G = \{D_1 + D_2 = 7\}.$$

پیشامد G شامل تمام حالت‌هایی است که مجموع دو تاس هفت می‌شود:

$$G = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}.$$

بنابراین

$$\Pr(E) = \Pr(F) = \Pr(G) = \frac{1}{6}.$$

بررسی استقلال زوجی

ابتدا E و F را بررسی می‌کنیم:

$$E \cap F = \{(1, 6)\},$$

پس

$$\Pr(E \cap F) = \frac{1}{36} = \frac{1}{6} \times \frac{1}{6} = \Pr(E) \Pr(F).$$

بنابراین E و F مستقل‌اند.

برای E و G نیز داریم

$$E \cap G = \{(1, 6)\},$$

پس

$$\Pr(E \cap G) = \frac{1}{36} = \frac{1}{6} \times \frac{1}{6} = \Pr(E) \Pr(G).$$

بنابراین E و G مستقل‌اند.

در نهایت، برای F و G داریم

$$F \cap G = \{(1, 6)\},$$

پس

$$\Pr(F \cap G) = \frac{1}{36} = \frac{1}{6} \times \frac{1}{6} = \Pr(F) \Pr(G).$$

بنابراین F و G نیز مستقل‌اند. هر سه شرط دوتایی برقرار است؛ در نتیجه E ، F و G استقلال زوجی دارند.

بررسی استقلال سه‌تایی

اشتراک سه پیشامد برابر است با

$$E \cap F \cap G = \{(1, 6)\}.$$

بنابراین

$$\Pr(E \cap F \cap G) = \frac{1}{36}.$$

از طرف دیگر،

$$\Pr(E) \Pr(F) \Pr(G) = \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} = \frac{1}{216}.$$

چون

$$\frac{1}{36} \neq \frac{1}{216},$$

شرط استقلال سه‌تایی برقرار نیست. پس این سه پیشامد با وجود استقلال زوجی، استقلال متقابل ندارند. این مثال نشان می‌دهد که برای بیش از دو پیشامد، بررسی استقلال زوجی کافی نیست.

۳ استقلال متغیرهای تصادفی

دو متغیر تصادفی زمانی مستقل‌اند که دانستن مقدار یکی، اطلاعاتی درباره‌ی مقدار دیگری به ما ندهد. اگر دو متغیر مستقل نباشند، آن‌ها را وابسته می‌نامیم. تعریف دقیق استقلال برای متغیرهای تصادفی گسسته و پیوسته متفاوت است.

۱.۳ استقلال متغیرهای تصادفی گسسته

فرض کنید X و Y دو متغیر تصادفی گسسته باشند. این دو متغیر مستقل‌اند اگر برای تمام مقادیر ممکن x و y داشته باشیم

$$\Pr(X = x, Y = y) = \Pr(X = x) \Pr(Y = y).$$

اگر تابع جرم احتمال توأم را با $p_{X,Y}$ و توابع جرم احتمال حاشیه‌ای را با p_X و p_Y نشان دهیم، همین تعریف با نمادگذاری دیگر به صورت

$$p_{X,Y}(x, y) = p_X(x)p_Y(y)$$

نوشته می‌شود.

نکته‌ی مهم این است که این تساوی باید برای تمام زوج‌های ممکن (x, y) برقرار باشد. بررسی یک زوج خاص برای اثبات استقلال کافی نیست؛ ولی یافتن حتی یک زوج که در آن تساوی برقرار نباشد، برای اثبات وابستگی کافی است.

۲.۳ مثال: مجموع دو تاس و نتیجه‌ی تاس اول

دو تاس سالم را پرتاب می‌کنیم و نتیجه‌های آن‌ها را با D_1 و D_2 نشان می‌دهیم. متغیر تصادفی مجموع را به صورت

$$S = D_1 + D_2$$

تعریف می‌کنیم. می‌خواهیم بررسی کنیم آیا S و D_1 مستقل‌اند یا خیر.

اگر فقط مقدار $S = 7$ را بررسی کنیم، ممکن است به نظر برسد این دو متغیر مستقل‌اند. زیرا

$$\Pr(S = 7) = \frac{6}{36} = \frac{1}{6}$$

و برای هر $d \in \{1, 2, 3, 4, 5, 6\}$ دقیقاً یک حالت وجود دارد که در آن $D_1 = d$ و $S = 7$ است. بنابراین

$$\Pr(S = 7, D_1 = d) = \frac{1}{36} = \Pr(S = 7) \Pr(D_1 = d).$$

اما این محاسبه فقط مقادیر متناظر با $S = 7$ را بررسی کرده است. برای استقلال، رابطه باید برای تمام مقادیر S و D_1 برقرار باشد.

برای مثال، $S = 5$ و $D_1 = 5$ را در نظر می‌گیریم. اگر $D_1 = 5$ و $S = 5$ باشد، باید $D_2 = 0$ شود که برای تاس ممکن نیست. پس

$$\Pr(S = 5, D_1 = 5) = 0.$$

ولی

$$\Pr(S = 5) = \frac{4}{36} = \frac{1}{9}, \quad \Pr(D_1 = 5) = \frac{1}{6},$$

و در نتیجه

$$\Pr(S = 5) \Pr(D_1 = 5) = \frac{1}{9} \times \frac{1}{6} = \frac{1}{54} > 0.$$

بنابراین

$$\Pr(S = 5, D_1 = 5) \neq \Pr(S = 5) \Pr(D_1 = 5),$$

پس S و D_1 مستقل نیستند و به یکدیگر وابسته‌اند. از نظر شهودی نیز دانستن مجموع دو تاس، مقادیر ممکن تاس اول را محدود می‌کند.

۳.۳ استقلال متغیرهای تصادفی پیوسته

فرض کنید X و Y دو متغیر تصادفی پیوسته باشند. این دو متغیر مستقل‌اند اگر برای تمام $x, y \in \mathbb{R}$ داشته باشیم

$$\Pr(X \leq x, Y \leq y) = \Pr(X \leq x) \Pr(Y \leq y).$$

با استفاده از تابع توزیع تجمعی، همین رابطه به صورت

$$F_{X,Y}(x, y) = F_X(x)F_Y(y)$$

نوشته می‌شود.

اگر چگالی‌های توأم و حاشیه‌ای وجود داشته باشند، شرط معادل استقلال آن است که

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

برای تمام مقادیر x و y برقرار باشد. چگالی‌های حاشیه‌ای نیز از انتگرال‌گیری روی متغیر دیگر به دست می‌آیند:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx.$$

۴.۳ نقش قلمرو در بررسی استقلال

برای بررسی استقلال، تنها جدا شدن عبارت جبری چگالی به دو عامل کافی نیست؛ این تساوی باید روی تمام قلمرو یا support توزیع برقرار باشد. به بیان دقیق‌تر، باید بتوان چگالی توأم را برای تمام (x, y) به صورت

$$f_{X,Y}(x, y) = g(x)h(y)$$

نوشت، به‌طوری که پس از نرمال‌سازی، $g(x) = f_X(x)$ و $h(y) = f_Y(y)$ باشند. همچنین قلمرو مشترک باید حاصل ضرب دو قلمرو یک‌بعدی باشد. اگر محدوده‌ی ممکن Y به مقدار X وابسته باشد، معمولاً استقلال برقرار نیست.

۵.۳ مثال: دو متغیر تصادفی گوسی

فرض کنید

$$X \sim \mathcal{N}(\mu_X, \sigma_X^2), \quad Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$$

و چگالی توأم آن‌ها برابر باشد با

$$\begin{aligned} f_{X,Y}(x,y) &= \frac{1}{2\pi\sigma_X\sigma_Y} \exp\left(-\frac{(x-\mu_X)^2}{2\sigma_X^2} - \frac{(y-\mu_Y)^2}{2\sigma_Y^2}\right) \\ &= \left[\frac{1}{\sqrt{2\pi}\sigma_X} \exp\left(-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right)\right] \left[\frac{1}{\sqrt{2\pi}\sigma_Y} \exp\left(-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}\right)\right]. \end{aligned}$$

عامل اول فقط تابع x و برابر $f_X(x)$ است و عامل دوم فقط به y وابسته و برابر $f_Y(y)$ است. بنابراین

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

و X و Y مستقل‌اند.

۶.۳ سؤال اول: چگالی نمایی

آیا متغیرهای تصادفی با چگالی توأم زیر مستقل‌اند؟

$$f_{X,Y}(x,y) = 6e^{-3x}e^{-2y}, \quad 0 < x < \infty, \quad 0 < y < \infty.$$

حل

ابتدا چگالی توأم را به صورت زیر می‌نویسیم:

$$f_{X,Y}(x,y) = (3e^{-3x})(2e^{-2y}).$$

چگالی‌های حاشیه‌ای نیز برابرند با

$$f_X(x) = 3e^{-3x}, \quad x > 0,$$

و

$$f_Y(y) = 2e^{-2y}, \quad y > 0.$$

پس

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

قلمرو توزیع نیز برابر حاصل ضرب $(0, \infty) \times (0, \infty)$ است. بنابراین X و Y مستقل‌اند.

۷.۳ سؤال دوم: چگالی روی مربع واحد

آیا متغیرهای تصادفی با چگالی توأم زیر مستقل‌اند؟

$$f_{X,Y}(x,y) = 4xy, \quad 0 < x < 1, \quad 0 < y < 1.$$

حل

چگالی حاشیه‌ای X برابر است با

$$f_X(x) = \int_0^1 4xy \, dy = 2x, \quad 0 < x < 1.$$

چگالی حاشیه‌ای Y نیز برابر است با

$$f_Y(y) = \int_0^1 4xy \, dx = 2y, \quad 0 < y < 1.$$

بنابراین

$$f_X(x)f_Y(y) = (2x)(2y) = 4xy = f_{X,Y}(x, y).$$

قلمرو نیز مربع $(0, 1) \times (0, 1)$ و حاصل ضرب دو بازه است. پس X و Y مستقل‌اند.

۸.۳ سؤال سوم: قلمرو مثلثی

تابع زیر را در نظر بگیرید:

$$f(x, y) = 8xy, \quad 0 < x < 1, \quad 0 < y < 1, \quad x + y < 1.$$

آیا متغیرهای تصادفی متناظر مستقل‌اند؟

توضیح قید و حل

اگر قید به صورت $x + y = 1$ نوشته شود، قلمرو فقط یک خط و دارای مساحت صفر است؛ بنابراین نمی‌تواند قلمرو یک چگالی دوبعدی معمولی باشد. برای ایجاد ناحیه‌ی مثلثی، قید مورد نظر باید $x + y < 1$ باشد.

همچنین تابع $8xy$ روی این مثلث نرمال نیست، زیرا

$$\int_0^1 \int_0^{1-x} 8xy \, dy \, dx = \frac{1}{3} \neq 1.$$

بنابراین $8xy$ با این قلمرو به تنهایی یک چگالی احتمال معتبر نیست. ثابت نرمال‌کننده‌ی صحیح برابر ۲۴ است و چگالی نرمال‌شده به صورت

$$f_{X,Y}(x, y) = 24xy, \quad x > 0, \quad y > 0, \quad x + y < 1$$

خواهد بود.

حتی پس از نرمال‌سازی، X و Y مستقل نیستند. دلیل اصلی، قلمرو مثلثی توزیع است. برای هر مقدار ثابت x ، محدوده‌ی ممکن Y برابر

$$0 < y < 1 - x$$

است؛ پس محدوده‌ی Y به مقدار X وابسته است. این قلمرو را نمی‌توان به صورت حاصل ضرب دو بازه‌ی مستقل نوشت.

برای مشاهده‌ی جبری وابستگی، چگالی‌های حاشیه‌ای نسخه‌ی نرمال‌شده را محاسبه می‌کنیم:

$$f_X(x) = \int_0^{1-x} 24xy \, dy = 12x(1-x)^2, \quad 0 < x < 1,$$

$$f_Y(y) = \int_0^{1-y} 24xy \, dx = 12y(1-y)^2, \quad 0 < y < 1.$$

حاصل ضرب آن‌ها برابر است با

$$f_X(x)f_Y(y) = 144xy(1-x)^2(1-y)^2,$$

که با $24xy$ برابر نیست. همچنین در نقاطی مانند $(0.75, 0.75)$ ، هر دو چگالی حاشیه‌ای مثبت‌اند ولی چون $x+y > 1$ است، چگالی توأم صفر است. پس شرط

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

برقرار نیست و X و Y وابسته‌اند.

۹.۳ جمع‌بندی استقلال متغیرهای تصادفی

برای متغیرهای گسسته، استقلال با تجزیه‌ی تابع جرم احتمال توأم و برای متغیرهای پیوسته با تجزیه‌ی تابع توزیع تجمعی یا چگالی توأم بررسی می‌شود. شرط استقلال باید برای تمام مقادیر ممکن برقرار باشد. علاوه بر عبارت جبری توزیع، شکل قلمرو نیز اهمیت دارد؛ قلمروهایی که در آن‌ها محدوده‌ی یک متغیر به مقدار متغیر دیگر وابسته است، معمولاً باعث وابستگی می‌شوند.

۴ تمرین: تبدیل خطی و استقلال

فرض کنید

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

و متغیر تصادفی Y به صورت

$$Y = 2X$$

تعریف شده باشد.

توزیع Y را به دو روش زیر به دست آورید:

۱. با استفاده از قاعده‌ی تبدیل خطی متغیر تصادفی؛

۲. با فرض استقلال X و Y .

سپس نتایج دو روش را مقایسه کنید و توضیح دهید کدام روش به پاسخ صحیح منجر می‌شود و چرا.

پاسخ

مقایسه دو روش برای متغیرهای نرمال

□ روش تبدیل خطی: اگر $X \sim N(\mu, \sigma^2)$ و $Y = aX + b$ آنگاه

$$Y \sim N(a\mu + b, a^2\sigma^2)$$

□ روش جمع متغیرهای تصادفی مستقل: اگر $X_1 \sim N(\mu_1, \sigma_1^2)$ و $X_2 \sim N(\mu_2, \sigma_2^2)$ مستقل باشند، آنگاه

$$Y = X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

بررسی مثال $Y = 2X$ با هر دو روش

روش اول: تبدیل خطی

$$Y = 2X \Rightarrow a = 2, b = 0$$

پس:

$$Y \sim N(2\mu, 4\sigma^2)$$

روش دوم: جمع متغیرها

$$Y = X + X$$

در نتیجه:

$$X + X \sim N(\mu + \mu, \sigma^2 + \sigma^2)$$

یعنی:

$$Y \sim N(2\mu, 2\sigma^2)$$

جواب غلط

سوال: از دو روش به دو جواب متفاوت رسیده‌ام. کدام روش به جواب غلط منجر شده است؟

جواب: روش جمع متغیرهای تصادفی مستقل جواب غلط تولید می‌کند، زیرا رابطه

$$Y = X_1 + X_2$$

فقط در حالتی درست است که X_1 و X_2 مستقل باشند؛ اما در

$$Y = X + X$$

استقلال برقرار نیست، چون X و X مستقل نیستند.

سوال: دو بازیکن C و B مشغول یک بازی تصادفی هستند که امتیازهای آنها طبق توزیع‌های نرمال زیر توزیع شده‌اند:

$$A_C \sim N(1657, 200^2)$$

$$A_B \sim N(1470, 200^2)$$

احتمال برنده شدن C را می‌خواهیم محاسبه کنیم:

$$P(C \text{ برنده شدن}) = P(A_C > A_B) = P(A_C - A_B > 0)$$

اگر A_C و A_B مستقل باشند، فرض کنید:

$$D = A_C - A_B = A_C + (-A_B)$$

کدام یک از گزینه‌های زیر توزیع D را، که برابر با $D = A_C - A_B$ است، به درستی بیان می‌کند؟

$$1. D \sim N(1657 - 1470, 200^2 - 200^2)$$

$$2. D \sim N(1657 - 1470, 200^2 + 200^2)$$

$$3. D \sim N(1657 + 1470, 200^2 + 200^2)$$

$$4. \text{هیچ کدام}$$

پاسخ

گزینه 2 صحیح است.

دلیل

با توجه به اینکه

$$A_B \sim N(\mu, \sigma^2),$$

اگر متغیر قرینه آن را در نظر بگیریم، خواهیم داشت:

$$-A_B \sim N(-\mu, (-1)^2 \sigma^2).$$

پس واریانس بدون تغییر باقی می‌ماند و فقط میانگین قرینه می‌شود. بنابراین برای

$$D = A_C + (-A_B)$$

داریم:

$$D \sim N(1657 - 1470, 200^2 + 200^2).$$

در نتیجه:

$$D \sim N(187, 2 \times 200^2).$$

محاسبه احتمال نهایی

از آنجا که واریانس D برابر است با

$$2 \times 200^2 = 80000,$$

انحراف معیار آن برابر می‌شود با:

$$\sigma_D = \sqrt{80000} \approx 283.$$

احتمال برنده شدن C برابر است با:

$$P(\square \text{ شدن}) = P(D > 0).$$

پس:

$$P(D > 0) = 1 - F_D(0) = 1 - \Phi\left(\frac{0 - 187}{283}\right) \approx 0.7454.$$

مجموع متغیرهای تصادفی مستقل

قوانین کلی

برای توزیع $X + Y$ وقتی X و Y مستقل هستند، روابط زیر برقرار است (و این نتایج به سادگی به بیش از دو متغیر نیز تعمیم می‌یابد):

۱. توزیع دوجمله‌ای: اگر $X \sim \text{Bin}(n_1, p)$ و $Y \sim \text{Bin}(n_2, p)$ آنگاه $X + Y \sim \text{Bin}(n_1 + n_2, p)$.

۲. توزیع پواسون: اگر $X \sim \text{Poi}(\lambda_1)$ و $Y \sim \text{Poi}(\lambda_2)$ آنگاه $X + Y \sim \text{Poi}(\lambda_1 + \lambda_2)$.

۳. توزیع نرمال: اگر $X \sim N(\mu_1, \sigma_1^2)$ و $Y \sim N(\mu_2, \sigma_2^2)$ آنگاه $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

مثال: کاربرد تقریب پواسون

مسئله: فرض کنید $A \sim \text{Bin}(30, 0.01)$ و $B \sim \text{Bin}(50, 0.02)$ و این دو متغیر مستقل باشند. آیا می‌توان $P(A + B = 2)$ را با روشی ساده محاسبه کرد؟

نکته کلیدی

هرگاه در توزیع دوجمله‌ای داشته باشیم $n \geq 20$ و $p \leq 0.05$ می‌توان از تقریب پواسون استفاده کرد. این شرط برای هر دو متغیر A و B برقرار است.

روش ۱: روش تقریبی ولی ساده

$$A \approx X \sim \text{Poi}(\lambda_1 = np = 30 \times 0.01 = 0.3)$$

و

$$B \approx Y \sim \text{Poi}(\lambda_2 = np = 50 \times 0.02 = 1).$$

پس:

$$X + Y \sim \text{Poi}(\lambda_1 + \lambda_2) = \text{Poi}(1.3).$$

بنابراین:

$$P(A + B = 2) \approx P(X + Y = 2) = \frac{\lambda^2}{2!} e^{-\lambda} = \frac{1.3^2}{2} e^{-1.3} \approx 0.2302.$$

روش ۲: روش دقیق ولی سخت‌تر

با استفاده از قانون احتمال کل برای متغیرهای گسسته:

$$P(A + B = 2) = \sum_{k=0}^2 P(A = k) P(B = 2 - k).$$

در نتیجه:

$$P(A + B = 2) = \sum_{k=0}^2 \binom{30}{k} (0.01)^k (0.99)^{30-k} \binom{50}{2-k} (0.02)^{2-k} (0.98)^{48+k} \approx 0.2327.$$

نتیجه

همان‌طور که مشاهده می‌شود، نتیجه روش تقریبی بسیار نزدیک به جواب دقیق است.

مثال

جدول زیر توزیع احتمال توأم ۳ متغیر تصادفی گسسته X, Y, Z را نشان میدهد.
به سوالات زیر پاسخ دهید؟

Z	Y	$X = 1$	$X = 2$	$X = 3$	$X = 4$	$X = 5$	$X = 6$
0	$Y = 1$	0.006	0.000	0.000	0.000	0.000	0.000
0	$Y = 2$	0.155	0.069	0.034	0.006	0.023	0.029
0	$Y = 3$	0.092	0.063	0.023	0.006	0.006	0.000
0	$Y = 4$	0.017	0.029	0.011	0.006	0.000	0.000
0	$Y \geq 5$	0.029	0.006	0.011	0.006	0.000	0.000
1	$Y = 1$	0.000	0.000	0.000	0.000	0.000	0.000
1	$Y = 2$	0.126	0.040	0.017	0.017	0.000	0.017
1	$Y = 3$	0.046	0.040	0.006	0.011	0.000	0.006
1	$Y = 4$	0.006	0.006	0.000	0.000	0.000	0.000
1	$Y \geq 5$	0.006	0.000	0.017	0.011	0.000	0.000

1- $P(Y \geq 5, X = 1, Z = 1)$

راهنمایی: دقت به جدول صورت سوال

2- $P(Z = z | X = x)$

راهنمایی: توزیع احتمال مربوطه را در قالب جدولی بدست آورید

3- $P(X = x | Z = z)$

راهنمایی: توزیع احتمال مربوطه را در قالب جدولی بدست آورید

جواب:

۱- جعبه نشان داده شده در جدول (۰.۰۶.۰)

۲- ابتدا باید متغیر تصادفی Y را به حاشیه بفرستیم و به توزیع $P(X = x, Z = z)$ دست پیدا کنیم.

به کمک توزیع فوق می‌توان به سوال‌هایی نظیر $P(X = 1, Z = 1)$ پاسخ داد که مقدار آن مطابق جدول 0.184 است.

سپس به کمک روابط زیر توزیع‌های $P(Z = z | X = x)$ و $P(X = x | Z = z)$ را به دست آورد:

$$P(Z = z | X = x) = \frac{P(Z = z, X = x)}{P(X = x)}$$

	$X = 1$	$X = 2$	$X = 3$	$X = 4$	$X = 5$	$X = 6$
$Z = 0$	0.299	0.167	0.080	0.023	0.029	0.029
$Z = 1$	0.184	0.086	0.040	0.040	0.000	0.023

$$P(X = 1, Z = 1) = 0.184$$

$$P(Z = 0 | X = 1) = \frac{P(Z = 0, X = 1)}{P(X = 1)} = \frac{0.299}{0.299 + 0.184} = \frac{0.299}{0.483} = 0.619$$

$$P(Z = 1 | X = 1) = \frac{P(Z = 1, X = 1)}{P(X = 1)} = \frac{0.184}{0.299 + 0.184} = \frac{0.184}{0.483} = 0.381$$

-۳

	$X = 1$	$X = 2$	$X = 3$	$X = 4$	$X = 5$	$X = 6$
$Z = 0$	0.619	0.659	0.667	0.364	1.000	0.556
$Z = 1$	0.381	0.341	0.333	0.636	0.000	0.444

$$P(X = x | Z = z) = \frac{P(X = x, Z = z)}{P(Z = z)}$$

$$P(X = 1 | Z = 0) = \frac{P(X = 1, Z = 0)}{P(Z = 0)} = \frac{0.299}{0.299 + 0.167 + 0.080 + 0.023 + 0.029 + 0.029} = \frac{0.299}{0.627} = 0.477$$

$$P(X = 1 | Z = 1) = \frac{P(X = 1, Z = 1)}{P(Z = 1)} = \frac{0.184}{0.184 + 0.086 + 0.040 + 0.040 + 0.000 + 0.023} = \frac{0.184}{0.373} = 0.493$$

	$X = 1$	$X = 2$	$X = 3$	$X = 4$	$X = 5$	$X = 6$
$Z = 0$	0.477	0.266	0.128	0.037	0.046	0.046
$Z = 1$	0.493	0.231	0.108	0.108	0.000	0.062

Law of the unconscious statistician : LOTUS

$$E[g(X)] = \sum_x g(x)p(x)$$

دو تاس ۶ وجهی را ۲ بار پرتاب می‌کنیم و آن‌ها را D_1 و D_2 می‌نامیم. مقدار $X = D_1 + D_2$ مقدار $Y = D_2$ امید $E[X | Y = 6]$ را حساب کنید.

جواب:

$$\begin{aligned} \sum_x xp(X = x | Y = 6) &= 2 \times 0 + 3 \times 0 + \dots + 6 \times 0 \\ &+ 7 \times \frac{1}{6} + 8 \times \frac{1}{6} + \dots + 12 \times \frac{1}{6} = 9.5 \end{aligned}$$

* به طور شهودی هم به جواب فوق می‌رسیم:

$$6 + E[\text{val of } D_1] = 6 + 3.5 = 9.5$$

سوال: ۱- $E[X]$ ۲- $E[X, Y]$ ۳- $E[X | Y]$ ۴- $E[X | Y = 6]$ ۵- $E[Y | X]$ ۶- $E[X = 1]$

هر کدام از اعداد فوق با کدام مورد زیر مطابق است.

A- یک عدد B- تابعی از Y C- تابعی از X ($g(X)$)

D- تابعی از X, Y ($g(X, Y)$) E- مفهومی ندارد

جواب: (1A) (2F) (3B) (4A) (5C)

(6E) $X = 1$ یک پیشامد است و امید یک پیشامد را نمی‌توان محاسبه کرد.

سوال: $E[X | Y]$ با کدام مورد مطابق است.

A- یک تابع از Y B- یک تابع از X C- یک عدد

جواب: A

* مورد ستاره دار صفحه قبل را دقیق‌تر بررسی کنید:

$$E[X | Y = 6] = 9.5$$

Let $W = \text{value of } D_1$, $X = W + Y$

W, Y مستقل هستند.

$$\begin{aligned}
\therefore E[X | Y = y] &= E[W + Y | Y = y] = E[W + y | Y = y] = y + E[W | Y = y] \\
&\downarrow \\
&(E[y | Y = y] = y) \\
&= y + \sum_w wp(W = w | Y = y) \stackrel{\text{به علت استقلال}}{=} y + \sum_w wp(W = w) \\
&= y + E[W] = y + 3.5 \\
E[X | Y] &= 3.5 + Y \quad \text{پس:}
\end{aligned}$$

* تحقیق کنید:

$$\begin{aligned}
g(Y) &= E[X | Y] \quad E[E[X | Y]] = E[g(Y)] = \sum_y P(Y = y) E[X | Y = y] \\
&= \sum_y P(Y = y) \sum_x xp(X = x | Y = y) \\
\text{تعریف امید شرطی} \quad &= \sum_y (\sum_x xp(X = x | Y = y) P(Y = y)) = \sum_y \sum_x xp(X = x, Y = y) \\
\text{جابجا کردن مکان} \quad &= \sum_x \sum_y xp(X = x, Y = y) = \sum_x x \sum_y p(X = x, Y = y) \\
\text{به حاشیه راندن متغیر تصادفی } X \quad &= \sum_x xp(X = x) \\
&= E[X]
\end{aligned}$$

* برنامه بازگشتی زیر را در نظر بگیرید:

$f()$: uniform

X is randomly selected from $[1, 2, 3]$

if $(x == 1)$; ret 3

elseif $(x == 2)$; ret $(5 + f())$

else ret $(7 + f())$

فرض کنید Y مقدار بازگشت داده شده توسط $f()$ باشد، $E[Y]$ را حساب کنید.

$$E[Y] = E[Y | X = 1] \overbrace{p(X = 1)}^{\frac{1}{3}} + E[Y | X = 2] \overbrace{P(X = 2)}^{\frac{1}{3}} + E[Y | X = 3] \overbrace{p(X = 3)}^{\frac{1}{3}} \quad \text{جواب:}$$

وقتی $X = 1$ است، 3 را برمی‌گرداند. پس $E[Y | X = 1] = 3$

وقتی $X = 2$ است، 5 به علاوه مقدار بازگشتی $f()$ را برمی گرداند. پس:

$$E[Y | X = 2] = \ominus$$

$$E[5] + Y \quad \text{الف}$$

$$E[Y + 5] = 5 + E[Y] \quad \ominus$$

$$E[5] + E[Y | X = 2] \quad \text{ج}$$

وقتی $X = 3$ است، مقدار 7 به علاوه مقدار بازگشتی $f()$ را برمی گرداند. پس:

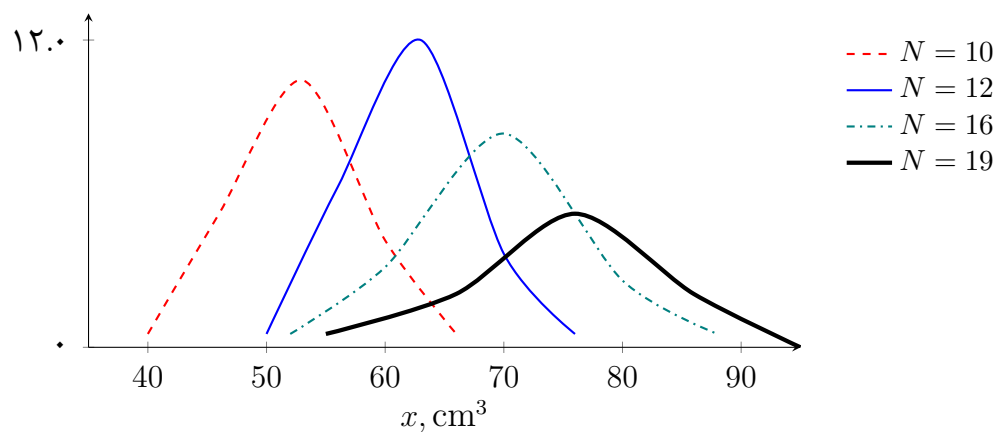
$$E[Y | X = 3] = 7 + E[Y]$$

$$E[Y] = 3 \times \frac{1}{3} + (5 + E[Y])\left(\frac{1}{3}\right) + (7 + E[Y])\left(\frac{1}{3}\right) \quad \text{برای جایگذاری:}$$

$$= \frac{1}{3}((3 + 5 + 7) + 2E[Y]) = 5 + \frac{2}{3}E[Y] \implies E[Y] = 15$$

* تلفیق متغیرهای تصادفی گسسته و پیوسته:

فرض کنید X متغیر پیوسته بیانگر حجم اشیاء و N متغیر گسسته بیانگر وزن همان اشیاء (10, 12, 16, 19) باشد. شکل زیر توزیع احتمال آن‌ها را نشان می‌دهد.



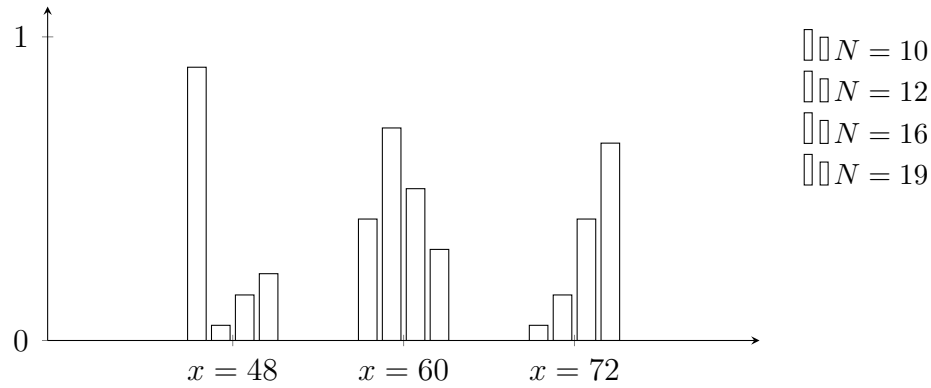
شکل ۴

PDF شرطی X به شرط N

A. $f_{X|N}(x | n)$

PMF شرطی N به شرط X

B. $p_{N|X}(n | x)$



شکل ۵

مثال : (از متغیر تصادفی تلفیقی)

یک سکه را $n + m$ بار پرتاب می‌کنیم و n شیر مشاهده می‌شود.

* قبل از آزمایش، X (احتمال شیر آمدن سکه) هر احتمالی می‌تواند باشد؛ $f_X(x)$

فرض شود تعداد شیرها $N =$

* با شرط $X = x$ پرتاب سکه مستقل هستند؛ $p_{N|X}(n|x)$

باور ما نسبت به X با مشاهده $N = n$ به‌روز شده و آن را با $f_{X|N}(x|n)$ بیان می‌کنیم.

توزیع‌های زیر را با توجه به A تا D متناظر کنید.

$X|N$ -۳ $N|X$ -۲ X -۱

Uni(0, 1) -A Bin($n + m, x$) -B C- یادگیری بیز

D- ایده ذهنی (subjective)

جواب:

X -۱ D ولی باور ما A می‌باشد.

$N|X$ -۲ B

$X|N$ -۳ C

شرح بیشتر: قبل از آزمایش، X می‌تواند هر احتمالی باشد (و در احتمال بیزی به آن توزیع prior یا پیشین گویند) و

معمولاً $X \sim \text{Uni}(0, 1)$

حال وقتی $N = n$ (تعداد شیر) باشد، با شرط $X = x$ ، پرتاب سکه‌ها مستقل است (و به آن likelihood گویند) و با نماد ریاضی زیر بیان می‌شود:

$$N|X \sim \text{Bin}(n + m, x)$$

یعنی توزیع شرطی $N|X$ دارای توزیع دوجمله‌ای است

در نهایت باور به روز شده X پس از مشاهده $N = n$ ، استنباط بهتری نسبت به متغیر تصادفی X به دست می‌دهد و در آمار بیزی به آن posterior (پسین) گوییم و با نماد $f_{X|N}(x|n)$ آن را نشان می‌دهیم. (به تلفیق دو توزیع گسسته و پیوسته توجه کنید)

$$\begin{aligned} f_{X|N}(x|n) &= \frac{\overbrace{P_{N|X}(n|x)}^{\text{یکنواخت}} \overbrace{f_X(x)}^{\text{جمله‌ای}}}{P_N(n)} = \frac{\binom{n+m}{m} x^n (1-x)^m \times 1}{P_N(n)} \\ &= \frac{\binom{n+m}{m} x^n (1-x)^m}{\underbrace{P_N(n)}_{\text{ثابت}}} = \frac{1}{c} x^n (1-x)^m \end{aligned}$$

که در آن $c = \int_0^1 x^n (1-x)^m dx$ است.

*** توزیع شرطی:**

فرض کنید X و Y متغیرهای تصادفی پیوسته باشند.

PDF شرطی X به شرط Y (که در آن فرض می‌کنیم $f_Y(y) > 0$ تا مخرج کسر صفر نشود) عبارت است از:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

$$\frac{\overbrace{f_{X|Y}(x|y)dx}^{x \text{ تابعی از } x}}{\underbrace{1}_{\text{احتمال}}} = \frac{f_{X,Y}(x, y)dxdy}{f_Y(y)dy}$$

$$P(x \leq X \leq x + dx \mid y \leq Y \leq y + dy) = \frac{P(x \leq X \leq x + dx, y \leq Y \leq y + dy)}{P(y \leq Y \leq y + dy)}$$

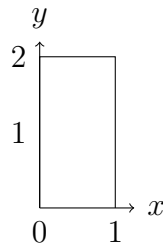
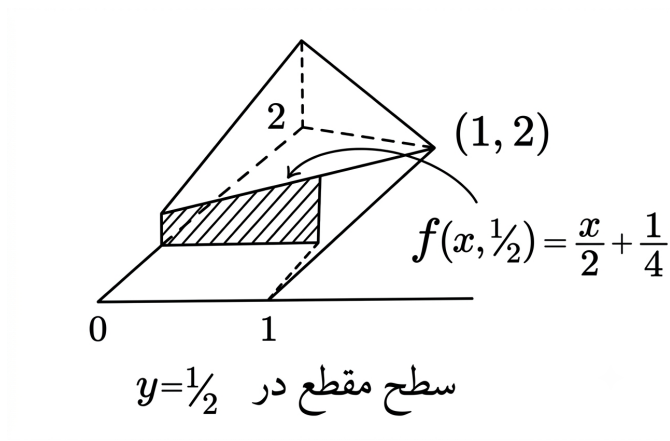
هر احتمال حاشیه‌ای را می‌توان به صورت مقدار امید نوشت.

$$P_X(x) = \int_y P_{X|Y}(x|y)P_Y(y) dy = E_y [P_{X|Y}(x|y)]$$

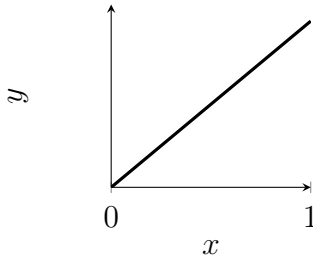
$$f(x, y) = \begin{cases} y(\frac{1}{2} - x) + x & 0 \leq x \leq 1, 0 \leq y \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

به سطح مقطع‌ها نگاه کنید. مثلاً سطح مقطع سطح با صفحه $y = \frac{1}{2}$ تابع زیر می‌باشد.

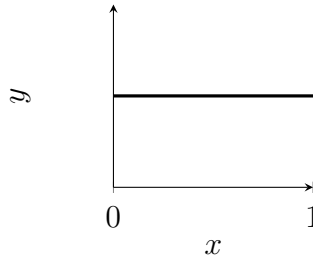
$$f(x, \frac{1}{2}) = \frac{1}{2} \left(\frac{1}{2} - x \right) + x = \frac{x}{2} + \frac{1}{4}$$



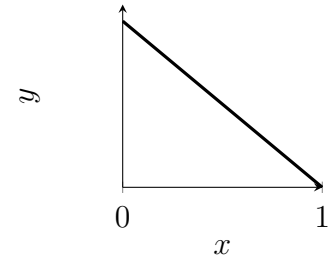
شکل ۶: نواحی در صفحه $x - y$ که $f(x, y) > 0$ است



$$f(x, 0) = x \text{ (ج)}$$



$$f(x, 1) = \frac{1}{2} \text{ (ب)}$$



$$f(x, 2) = 1 - x \text{ (ا)}$$

بررسی کنید که توزیع ۲ متغیره فوق، توزیع احتمال است.

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy &= \int_0^2 \int_0^1 \left[y \left(\frac{1}{2} - x \right) + x \right] dx dy \\ &= \int_0^2 \left\{ \int_0^1 \left[y \left(\frac{1}{2} - x \right) + x \right] dx \right\} dy \\ &= \int_0^2 \left[\int_0^1 y \left(\frac{1}{2} - x \right) dx + \int_0^1 x dx \right] dy \\ &= \int_0^2 \left[y \int_0^1 \left(\frac{1}{2} - x \right) dx + \int_0^1 x dx \right] dy \\ &= \int_0^2 \left[y \times 0 + \frac{1}{2} \right] dy \\ &= \int_0^2 \frac{1}{2} dy = \frac{1}{2} \cdot 2 = 1 \end{aligned}$$

در فرایند انتگرال گیری دوگانه فوق به انتگرال زیر رسیدیم.

$$\int_{-\infty}^{\infty} f(x, y) dx = \int_0^1 \left[y \left(\frac{1}{2} - x \right) + x \right] dx = \frac{1}{2} \quad 0 < y < 2$$

که توزیع حاشیه‌ای Y عبارت است از

$$f_Y(y) = \begin{cases} \frac{1}{2} & 0 < y < 2 \\ 0 & \text{otherwise} \end{cases}$$

که تشخیص می‌دهیم توزیع یکنواخت $U(0, 2)$ است.

$$\begin{aligned}
 f_X(x) &= \int_{-\infty}^{\infty} f(x, y) dy \\
 &= \int_0^2 \left[y \left(\frac{1}{2} - x \right) + x \right] dy && 0 < x < 1 \\
 &= \left(\frac{1}{2} - x \right) \int_0^2 y dy + 2x && 0 < x < 1 \\
 &= 2 \left(\frac{1}{2} - x \right) + 2x && 0 < x < 1 \\
 &= 1 && 0 < x < 1 \\
 &= 0 && \text{otherwise}
 \end{aligned}$$

که یک توزیع یکنواخت $U(0, 1)$ است.

در مثال فوق $f_X(x)f_Y(y) = \frac{1}{2} \neq f(x, y)$ که پیشامدها مستقل نیستند.

توزیع شرطی X به شرط $Y = y$ وقتی $0 < y < 2$ است، عبارت است از:

$$\begin{aligned}
 f(x|y) &= \frac{f(x, y)}{f_Y(y)} \\
 &= \frac{y \left(\frac{1}{2} - x \right) + x}{\frac{1}{2}} && 0 < x < 1 \\
 &= y(1 - 2x) + 2x && 0 < x < 1 \\
 &= 0 && \text{otherwise}
 \end{aligned}$$

که معادله فوق به ضریب مقیاس‌کننده $\frac{1}{2}$ همان معادله سطح مقطع در صفحه قبلی است. برای $y = \frac{1}{2}$ داریم:

$$f\left(x \middle| \frac{1}{2}\right) = \begin{cases} x + \frac{1}{2} & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

توجه داریم که این سطح مقطع‌ها که از آنها به عنوان چگالی‌های شرطی مقیاس نشده یاد می‌کنیم با تغییرات y ، از نظر شکل و مقیاس عوض می‌شوند (برای سه مقدار y شکل دو صفحه قبل نگاه شود)

به طور کلی اگر برای همه y ها و همه x هایی که $f_X(x) > 0$ است: $f(y|x) = f_Y(y)$
یا به طور معادل، برای همه x ها و همه y هایی که $f_Y(y) > 0$ است: $f(x|y) = f_X(x)$
یا به طور معادل، برای همه x ها و y ها: $f(x, y) = f_X(x) \cdot f_Y(y)$
می‌گوییم X و Y متغیرهای تصادفی مستقل هستند.

مثال: pdf، X و Y به صورت زیر است:

$$f(x, y) = \frac{12}{7}(x^2 + xy) \quad 0 \leq x \leq 1, 0 \leq y \leq 1$$

a- $P(X > y)$

b- توزیع حاشیه‌ای (چگالی حاشیه‌ای) $f_X(x)$ را حساب کنید.

جواب:

بخش a:

$$P(X > y) = \iint_{X>y} f(x, y) dx dy = \int_0^1 \int_y^1 \frac{12}{7}(x^2 + xy) dx dy \quad x > y \rightarrow \begin{cases} y \in [0, 1] \\ x \in [y, 1] \end{cases}$$

$$= \int_0^1 \frac{12}{7} \left(\frac{x^3}{3} + \frac{x^2 y}{2} \right) \Big|_y^1 dy$$

یعنی x از y تا ۱ عوض شده

$$= \int_0^1 \left(\frac{12}{21} + \frac{12}{14}y - \frac{12}{21}y^3 - \frac{12}{14}y^3 \right) dy = \frac{9}{14}$$

بخش b:

برای $0 \leq x \leq 1$:

$$f_X(x) = \int_0^1 \frac{12}{7} f(x^2 + xy) dy = \frac{12}{7}x^2 + \frac{6}{7}x \quad \text{به حاشیه راندن } y$$

برای $x < 0$ و $x > 1$:

$$f_X(x) = 0$$

مثال: فرض کنید X تعداد افرادی باشد که در روز به نانوائی مراجعه می‌کنند و داریم $X \sim N(50, 25)$.

$Y_i =$ هر نفر i به مدت $Y_i \sim \text{Poisson}(8)$ منتظر می‌ماند.

X و Y_i مستقل هستند، زمان سپری شده برای همه مشتریان برابر با $W = \sum_{i=1}^X Y_i$ می‌باشد، $E[W]$ چیست؟

جواب:

$$E[W] = E\left(\sum_{i=1}^X Y_i\right) = E\left[E\left[\sum_{i=1}^X Y_i \mid X\right]\right] = E[X \cdot E[Y_i]] = E[X]E[Y_i] = 50 \cdot 8$$

وقتی تعداد مشخص باشد:

$$E\left[\sum_{i=1}^X Y_i \mid X = n\right] = \sum_{i=1}^n E[Y_i \mid X = n] = \sum_{i=1}^n E[Y_i] = nE[Y_i]$$

$$E\left[\sum_{i=1}^X Y_i \mid X\right] = X \cdot E[Y_i]$$

مثال: متغیر تصادفی X را مشاهده کرده‌ایم و می‌خواهیم در مورد Y پیشگویی کنیم. مثلاً $X =$ تعداد مشتری در صف در ساعت ۹ صبح و $Y =$ تعداد مشتری در ساعت ۱۰ صبح. فرض کنید $g(X)$ تابعی باشد که برای پیشگویی Y بکار می‌رود، یعنی $\hat{Y} = g(X)$. به دنبال آن $g(X)$ هستیم که $E[(Y - g(X))^2]$ را کمینه می‌کند.

جواب:

$$g(X) = E[X|Y]: \text{بهترین پیشگو:}$$

به طور شهودی $E[(Y - c)^2]$ وقتی که $c = E[Y]$ است، کمینه می‌شود. اکنون X را مشاهده می‌کنیم و Y به X وابسته است، بنابراین $c = E[Y|X]$ انتخاب می‌شود. مطلب فوق در واقع شروع یادگیری ماشین است.

متغیر تصادفی نشانگر (Indicator RV):

$X =$ متغیر تصادفی نشانگر برای پیشامد A است.

$$X = \begin{cases} 1 & \text{اگر } A \text{ رخ دهد} \\ 0 & \text{در غیر این صورت} \end{cases}$$

$$E[X] = P(A)$$

به طور مشابه برای هر Y ، $E[X|Y = y] = P(A|Y = y)$.

بنابراین $E[X] = E_Y[E_X[X|Y]] = E[E[X|Y]] = E[P(A|Y)]$

برای حالت گسسته: برای محاسبه این امید باید $E[X] = \sum_y P(A|Y=y)P(Y=y) = P(A)$ چگالی احتمال مربوطه را به طور صحیح بنویسید. (بسیار مهم).
برای حالت پیوسته: مشابه فوق می‌باشد.

قانون احتمال کل:

$$P(A) = \sum_y P(A|Y=y)P(Y=y)$$

مثال: استخدام منشی: با n نفر مصاحبه می‌کنیم. به $n!$ حالت منشیان می‌توانند وارد شوند، یعنی انواع ترتیب ورود n منشی، $n!$ است که همه را متساوی الاحتمال در نظر بگیرید، می‌خواهیم فقط ۱ نفر را استخدام کنیم.
ادعا: از نظر بهره‌وری بین «بهترین» و «متوسط» یک ضریب اختلاف α به 1 وجود دارد.
کمپانی A ضریب $\alpha = 25$ و کمپانی B ضریب $\alpha = 100$ را در نظر می‌گیرد.
فرض: بلافاصله پس از هر مصاحبه باید تصمیم بگیرید که استخدام کنید و یا خیر.
فیدبکی که از مصاحبه با نفر i بدست می‌آورد، یک رتبه‌بندی در مقایسه با $i-1$ کاندیدای قبلی به شما می‌دهد.
استراتژی استخدام: ابتدا از میان n نفر، با k نفر مصاحبه کنید و آنها را استخدام نکنید، سپس اولین کاندیدی که از k نفر اولیه بهتر باشد را استخدام کنید.

احتمال زیر را به این ترتیب تعریف می‌کنیم:

$P_k(best)$ = احتمال آنکه بهترین کاندیدا از میان n نفر انتخاب شود

X = موقعیت بهترین کاندیدا ($1, 2, \dots, n$)

$$P_k(Best) = \sum_{i=1}^n P_k(Best|X=i)P(X=i) \underbrace{P(X=i)}_{\frac{1}{n}} = \frac{1}{n} \sum_{i=1}^n P_k(Best|X=i)$$

هر کدام از n نفر می‌توانند بهترین باشند.

بدیهی است، اگر بهترین نفر جزو k نفر اولی باشد یعنی $i \leq k$ و چون ما او را استخدام نکرده‌ایم،

بنابراین: اگر $i \leq k$ ، $P_k(Best|X=i) = 0$

بنا بر استراتژی استخدام، ما شانس انتخاب بهترین کاندیدا را داریم. (فرض می‌کنیم در موقعیت i قرار دارد)، اگر بهترین کاندیدای تا نفر مصاحبه‌شونده $i-1$ جزو k نفر اول باشد. احتمال آن را به صورت $P_k(Best|X=i)$ نشان می‌دهیم.
تکرار می‌کنیم که، اگر بهترین نفر که در موقعیت i قرار دارد، جزو k نفر اول باشد. آنوقت، شانس انتخاب بهترین نفر

را از دست داده‌ایم (طبق استراتژی انتخاب ما) اگر $i \leq k$ ، $P_k(Best|X=i) = 0$

برای $i > k$: $P_k(Best|X=i) = P_k$ بهترین نفر در $i-1$ نفر قبلی جزو k نفر اول باشد) $P_k(Best|X=i) = P_k$ پس:

$$\begin{aligned} P_k(Best) &= \frac{1}{n} \sum_{i=1}^n P_k(Best|X=i) = \frac{1}{n} \sum_{i=k+1}^n \frac{k}{i-1} \\ &\approx \frac{k}{n} \int_{i=k+1}^n \frac{1}{i-1} di = \frac{k}{n} \ln(i-1) \Big|_{k+1}^n \\ &= \frac{k}{n} \ln \frac{n-1}{k} = \frac{k}{n} \ln \frac{n}{k} \end{aligned}$$

برای بیبشینه‌سازی، از $P_k(Best)$ نسبت به k مشتق می‌گیریم.

$$g(k) = \frac{k}{n} \ln \frac{n}{k}$$

$$g'(k) = \frac{1}{n} \ln \frac{n}{k} + \frac{k}{n} \left(\frac{-1}{k} \right) = \frac{1}{n} \ln \frac{n}{k} - \frac{1}{n}$$

قرار می‌دهیم $g'(k) = 0$ و بر حسب k حل می‌کنیم:

$$\frac{1}{n} \ln \frac{n}{k} - \frac{1}{n} = 0 \rightarrow \ln \frac{n}{k} = 1 \rightarrow \frac{k}{n} = e \rightarrow k = \frac{n}{e}$$

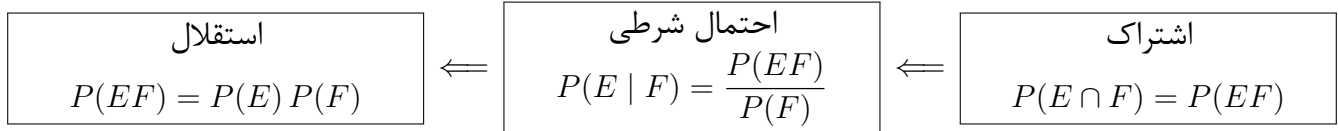
پس $\frac{n}{e}$ نفر را مصاحبه کنید و سپس بهترین را انتخاب کنید که احتمال آن می‌شود:

$$P_k(Best) \approx \frac{1}{e} = 0.368$$

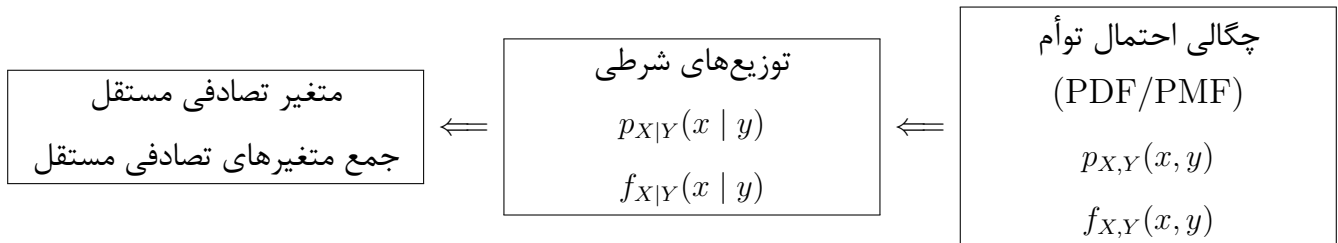
مرجع: کتاب CLRS

۵ توزیع‌های توأم و مثال ردگیری شیء در فضای دوبعدی

۱.۵ بحث چند پیشامد



توزیع‌های توأم چندمتغیره:



یک مثال: ردگیری شیء در فضای ۲بعدی (مقدمه‌ای بر یادگیری).

۱. **باور پیشین:** یک باور پیشین در مورد محل یک شیء در ۲ بعد داریم (X, Y) (مختصات شیء).
۲. **اندازه‌گیری (مشاهده):** یک مشاهده در مورد فاصله‌ی شیء که با نویز آغشته شده (یعنی طبق یک تابع توزیع می‌باشد) انجام می‌شود و این فاصله (مثلاً $D = 4$) به دست می‌آید.
۳. **باور پسین:** باور به‌روز شده (پسین = posterior) شما در مورد مکان شیء در فضای ۲بعدی پس از مشاهده اندازه‌گیری چه می‌باشد. در احتمال، در واقع در مورد احتمالات زیر بحث می‌شود.

$$\underbrace{f_{X,Y|D}(x, y | d)}_{\text{باور پسین (posterior belief)}} = \frac{\underbrace{f_{D|X,Y}(d | x, y)}_{\text{درست‌نمایی (گواه)}} \cdot \underbrace{f_{X,Y}(x, y)}_{\text{باور پیشین (prior belief)}}}{\underbrace{f_D(d)}_{\text{ثابت نرمال‌کننده}}}$$

(۱): فرض کنید $(0, 0)$ مبدأ؛ محل ۲بعدی شیء (X, Y) = فرض کنید توزیع پیشین یک نرمال ۲متغیره متقارن باشد.

۲.۵ باور پیشین و درست‌نمایی

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma^2} e^{-[(x-3)^2 + (y-3)^2]/2\sigma^2} = K_1 e^{-\frac{(x-3)^2 + (y-3)^2}{8}}$$

در آن فرض کنید $\sigma^2 = 2$ و $\mu_x = \mu_y = 3$ ؛ K_1 ثابت نرمال کننده.

۲- مشاهده $D = 4$ (قسمت درست‌نمایی مسئله). فرض کنید D فاصله‌ی شعاعی شیء از مبدأ باشد. فرض

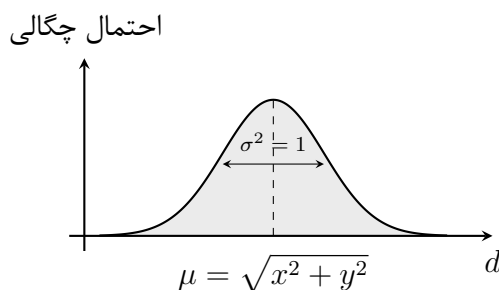
کنید محل واقعی شیء در مکان (x, y) را می‌دانیم. D هنوز به نویز آغشته شده که نویز دارای توزیع با واریانس

$\sigma^2 = 1$ فرض می‌شود. در بی‌نویزی، D مکان واقعی را نشان می‌دهد: $\mu = \sqrt{x^2 + y^2}$.

۳- اگر محل واقعی (x, y) را می‌دانیم، چقدر محتمل است (likelihood) که اندازه‌گیری ما منجر به $D = 4$ شود.

$$D | X, Y \sim N\left(\mu = \sqrt{x^2 + y^2}, \sigma^2 = 1\right)$$

در واقع اثر نویز را نرمال فرض کنیم.



هر اندازه‌گیری فاصله نسبت به محل واقعی شیء توزیع نرمال دارد.

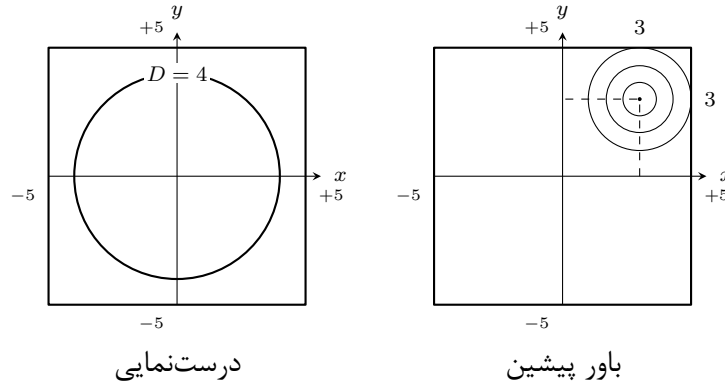
$$f_{D|X,Y}(D = d | X = x, Y = y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(d-\mu)^2/2\sigma^2} = \frac{1}{\sqrt{2\pi}} e^{-\frac{(d-\sqrt{x^2+y^2})^2}{2}} = K_2 e^{-\frac{(4-\sqrt{x^2+y^2})^2}{2}}$$

(با $\sigma^2 = 1$ و $\mu = \sqrt{x^2 + y^2}$ ؛ در مثال با 4)؛ K_2 ثابت نرمال کننده.

۳.۵ نمایش هندسی باور پیشین و درست‌نمایی

دو مؤلفه مسئله را می‌توان به صورت هندسی روی صفحه xy دید: باور پیشین یک تپه گاوسی حول $(3, 3)$ است و

درست‌نمایی $D = 4$ یک حلقه به شعاع 4 حول مبدأ می‌سازد.



جمع‌بندی بخش‌های ۱ و ۲:

(۱) باور پیشین:

$$f_{X,Y}(x, y) = K_1 \cdot e^{-\frac{[(x-3)^2 + (y-3)^2]}{8}}$$

(۲) درست‌نمایی (likelihood):

$$f_{D|X,Y}(d | x, y) = K_2 e^{-\frac{(d-\mu)^2}{2}} \quad \mu = \sqrt{x^2 + y^2}, \sigma^2 = 1$$

(۳) باور پسین (هدف):

$$f_{X,Y|D}(x, y | 4) = f_{X,Y|D}(X = x, Y = y | D = 4)$$

صورت بخش ۳. باور به‌روز شده (پسین) در مورد محل ۲ بعدی شیء پس از مشاهده اندازه‌گیری، طبق زیر و با توجه به بخش‌های ۱ و ۲ به‌دست می‌آید (هدف، چگالی پسین است):

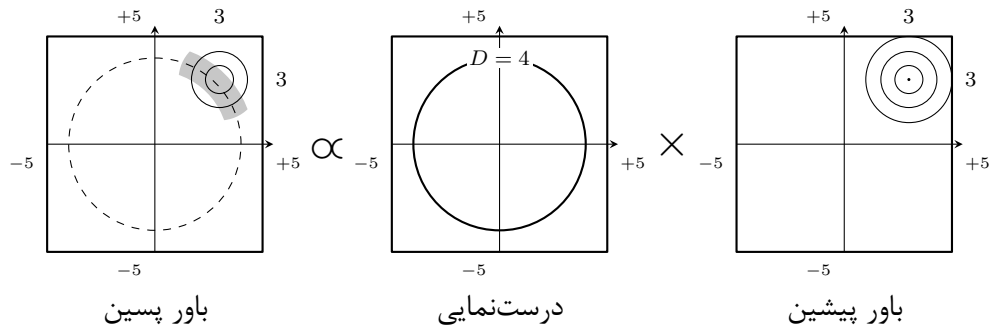
$$f_{X,Y|D}(X = x, Y = y | D = 4) \longleftarrow \text{(تئوری)}$$

۴.۵ محاسبه باور پسین (بخش ۳)

باور به‌روز شده (پسین) در مورد محل ۲ بعدی شیء با پردازش داده اندازه‌گیری، طبق زیر و با توجه به بخش‌های ۱ و ۲، از قضیه بیز محاسبه می‌شود:

$$\begin{aligned}
 f_{X,Y|D}(X=x, Y=y | D=4) &= \frac{\overbrace{f_{D|X,Y}(D=4 | X=x, Y=y)}^{D=4 \text{ درست‌نمایی}} \overbrace{f_{X,Y}(x, y)}^{\text{باور پیشین}}}{P(D=4)} \\
 &= \frac{K_2 \cdot e^{-\left(4-\sqrt{x^2+y^2}\right)^2/2} \cdot K_1 \cdot e^{-[(x-3)^2+(y-3)^2]/8}}{P(D=4)} \\
 &= \frac{K_3 e^{-\left[\frac{\left(4-\sqrt{x^2+y^2}\right)^2}{2} + \frac{(x-3)^2+(y-3)^2}{8}\right]}}{P(D=4)} = K_4 \cdot e^{-\left[\frac{\left(4-\sqrt{x^2+y^2}\right)^2}{2} + \frac{(x-3)^2+(y-3)^2}{8}\right]}
 \end{aligned}$$

(← $P(D=4)$ که ثابت است)



$$\begin{aligned}
 f_{X,Y}(x, y) &= K_1 e^{-\frac{[(x-3)^2+(y-3)^2]}{8}} && \text{باور پیشین} \leftarrow \\
 f_{X,Y|D}(x, y | 4) &= K_4 \cdot e^{-\left[\frac{\left(4-\sqrt{x^2+y^2}\right)^2}{2} + \frac{(x-3)^2+(y-3)^2}{8}\right]} && \text{باور پسین} \leftarrow
 \end{aligned}$$

۶ جمع دو متغیر تصادفی مستقل

فرض کنید D_1 و D_2 اعداد نمایش داده شده برای پرتاب ۲ تاس چهاروجهی باشد. بدیهی است D_1 و D_2 دارای توزیع مشابه بوده و روی ساپورت ۱ و ۲ و ۳ و ۴ و ۵ و ۶ به‌طور یکنواخت توزیع شده‌اند. $D_1 \sim D_2$ و

$$E(D_1) = E(D_2) = 3.5$$

۲ متغیر تصادفی مقابل را تعریف می‌کنیم:

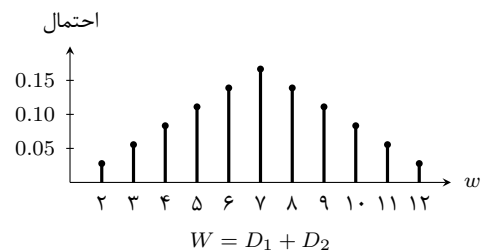
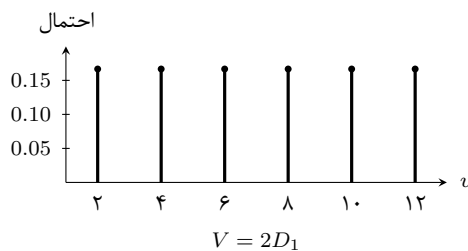
$$V = D_1 + D_1$$

$$W = D_1 + D_2$$

در هر ۲ مورد، ۲ متغیر تصادفی که روی اعداد ۱ تا ۶ دارای توزیع یکنواخت هستند را در حال جمع کردن هستیم. طبق قانون جمع‌پذیری:

$$E[V] = E[W] = 7$$

ولی توزیع‌های V و W کاملاً متفاوت است. W همه مقادیر ۲ تا ۱۲ را روی ساپورت خود دارد ولی V فقط مقادیر زوج ۲، ۴، ۶، ۸، ۱۰ و ۱۲ را به خود می‌گیرد. مشاهده می‌شود پراکندگی توزیع V از W بیشتر می‌باشد، با آنکه میانگین هر دو ۷ می‌باشد.



شکل: تابع جرم احتمال $V = 2D_1$ (سمت چپ؛ یکنواخت، احتمال هر مقدار برابر $\frac{1}{6}$) و $W = D_1 + D_2$ (سمت راست؛ مثلثی).

دلیل تفاوت در پراکندگی مربوط به آن است که در $W = D_1 + D_2$ پیشامدها مستقل ولی در $V = D_1 + D_1$ ، در واقع دو برابر مقدار انداختن یک تاس است. پس در میانگین وابستگی مهم نمی‌باشد ولی در واریانس وابستگی مهم است. چون در V ، وقتی D_1 بزرگ باشد، V هم بزرگ است ولی در W ، کمی ملایم‌تر است، چون D_2 ممکن است کوچک باشد. تغییرپذیری مقدار مجموع به رابطه‌ی ۲ متغیری که جمع می‌شوند وابسته است. در جمع n متغیر مستقل:

$$SD(S_n) = \sqrt{n} \sigma \iff \text{Var}(S_n) = \text{Var}(X_1) + \dots + \text{Var}(X_n) = n\sigma^2$$

σ^2 واریانس X_i می‌باشد. در مثال فوق:

$$SD(W) = \sqrt{2} \times 1,71 \iff SD(D_1) = 1,71$$

$$SD(V) = SD(D_1 + D_1) = SD(2D_1) = 2 SD(D_1) = 2 \times 1,71 \iff \text{ولی برای } V$$

$E(XY^2)$ را برای PDF توأم زیر به دست آورید:

$$f_{XY}(x, y) = \begin{cases} x + y & 0 < x \leq 1, 0 < y \leq 1 \\ 0 & \text{در غیر این صورت} \end{cases}$$

$$\begin{aligned}
 E[XY^2] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (xy^2) f_{XY}(x, y) dx dy = \int_0^1 \int_0^1 xy^2(x+y) dx dy \\
 &= \int_0^1 \int_0^1 (x^2y^2 + xy^3) dx dy = \int_0^1 \left(\frac{1}{3}y^2 + \frac{1}{2}y^3\right) dy = \frac{17}{72}
 \end{aligned}$$

۷ اطلاعات ارسالی در اسکایپ

درس ارزیابی کارایی

دکتر خونساری

۱۴۰۰

۱. یک مسئله. فرض کنید $U_1, U_2, \dots$ دنباله‌ای از متغیرهای تصادفی i.i.d با توزیع $\text{Uniform}[0,1]$ باشند. اگر N ایندکس کمینه‌ای باشد که $U_1 + U_2 + \dots + U_N \geq 1$ باشد، $E[N]$ را حساب کنید. جواب این سوال‌ها در اینترنت هست. منظور من از این سوال‌ها این است که ابهام‌هایی که دانشجویان معمولاً با آن مواجه هستند را به شما گوشزد کنم. اگر کسی راه حل خیلی جالب به همراه کد پایتون و یا آر پیدا می‌کند می‌تواند آن را به اشتراک بگذارد.

۲. جمع متغیرهای تصادفی بحث جالبی است؛ جمع تعدادی برنولی، دوجمله‌ای / جمع تعدادی نمایی، ارلانگ و یا گاما. اما در لینک بالا قضیه حد مرکزی بیان شده که نشان می‌دهد جمع متغیرهای تصادفی آ‌ی‌آ‌دی با آرایش ذکر شده نرمال می‌باشد. به تفاوت مفهومی مورد فوق‌الذکر کمال دقت را داشته باشید. سعی کنید برای درک بهتر کد پایتون و یا آر را بنویسید. مخصوصاً دقت کنید که در جمع متغیرهای تصادفی یکنواخت خیلی سریعتر از جمع مثلاً نمایی به نرمال همگرا می‌شوند. سعی کنید علت این تفاوت در نرخ همگرایی را متوجه شوید.

آدرس کد پایتون توزیع‌های نرمال، برنولی، دوتایی، نمایی و غیره :

<https://colab.research.google.com/drive/1o73Zrf5G9SVUFW2UY-qcGn3BhWCPktXr?usp=sharing>

۳ Tail of distribution is $P(X > x)$. For some distributions (light tailed) such as exponential it decays fast. For some distributions (long tailed, heavy tailed) such as Pareto it decays slowly. So, for heavy tailed you see big values in the support of the distribution many times and you couldn't ignore such values.

۴. نمایش تعاملی قضیه حد مرکزی با جمع توزیع‌های یکنواخت و نمایی:

<https://demonstrations.wolfram.com/IllustratingTheCentralLimitTheoremWithSumsOfUniformAndExponential/#popup1>

۵. برای دانشجویانی که ریاضی را فراموش کرده‌اند:

https://amsi.org.au/ESA_Senior_Years/PDF/

۶. توزیع بیشینه و کمینه متغیرهای مستقل.

Let $X_1, X_2, \dots, X_n$ be independent random variables (distribution functions $F_i(x)$ and tail distributions $G_i(x)$, $i = 1, \dots, n$).

Distribution of the maximum

$$\begin{aligned} P\{\max(X_1, X_2, \dots, X_n) \leq x\} &= P\{X_1 \leq x, \dots, X_n \leq x\} \\ &= P\{X_1 \leq x\} \cdots P\{X_n \leq x\} \quad (\text{independence!}) \\ &= F_1 \cdots F_n \end{aligned}$$

Distribution of the minimum

$$\begin{aligned} P\{\min(X_1, X_2, \dots, X_n) > x\} &= P\{X_1 > x, \dots, X_n > x\} \\ &= P\{X_1 > x\} \cdots P\{X_n > x\} \quad (\text{independence!}) \\ &= G_1(x) \cdots G_n(x) \end{aligned}$$

برای ماکزیمم به خود تابع توزیع فکر کنید: یعنی اگر بخواهیم ماکزیمم کمتر-مساوی ایکس باشد یعنی همه متغیرها هم کمتر-مساوی ایکس خواهند بود و به علت استقلال، احتمالاتِ توأم، مساوی حاصلضربِ احتمالات است؛ و برای مینیمم به مکملِ توزیع فکر کنید و اگر مینیمم بزرگتر از ایکس باشد یعنی همه بزرگتر از X خواهند بود و روابط فوق صحیح هستند.

سوال: چرا CDF احتمال است؟

لینک‌های ارسالی در گروه

<https://towardsdatascience.com/sum-of-exponential-random-variables-b023b61f0c0> •

- <https://homepage.divms.uiowa.edu/~mbognar/applets/gamma.html>
- look at other distributions and try to play with the parameters of each distribution and change the random variable in the support (range) of each distribution.
- <https://homepage.divms.uiowa.edu/~mbognar/>
- http://www.statisticalengineering.com/sums_of_random_variables.htm
- [https://stats.libretexts.org/Bookshelves/Probability_Theory/Book%3A_Introductory_Probability_\(Grinstead_and_Snell\)/07%3A_Sums_of_Random_Variables/7.02%3A_Sums_of_Continuous_Random_Variables](https://stats.libretexts.org/Bookshelves/Probability_Theory/Book%3A_Introductory_Probability_(Grinstead_and_Snell)/07%3A_Sums_of_Random_Variables/7.02%3A_Sums_of_Continuous_Random_Variables)
- https://www.probabilitycourse.com/chapter10/10_1_0_basic_concepts.php
- <https://www.geeksforgeeks.org/python-numpy-np-multinomial-method/>
- <https://www.geeksforgeeks.org/sympy-stats-multinomial-function-in-python/>
- https://www.probabilitycourse.com/chapter5/5_2_5_solved_prob.php

مراجع