آمار و احتمال مهندسی

استقلال متغیرهای تصادفی

(Ross 6.2)

امید ریاضی مشترک

- برای یک متغیر تصادفی X، امید ریاضی برابر است با:

$$\mathrm{E}(X) = \int_{-\infty}^{+\infty} x\, f_X(x)\, dx$$

- برای دو متغیر تصادفی X و Y با تابع چگالی مشترک $f_{XY}(x,y)$ به دلیل مشخص نبودن نحوه ترکیب X و Y نمی‌توان امید ریاضی تعریف کرد.

- اما با تعمیم قضیه اساسی امید ریاضی، می‌توانیم برای تابع $g(X,Y)$ از دو متغیر تصادفی X و Y به صورت زیر بیان می‌شود:

$$\mathrm{E}\big(g(X,Y)\big) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y) f_{XY}(x,y)dx\, dy$$

امید ریاضی مشترک

- به طور مشابه برای متغیرهای تصادفی گسسته X و Y داریم:

$$E[g(X,Y)] = \sum_{x,y} g(x,y)P_{XY}(x,y)$$

- قبلا گفتیم که $E[X+Y] = E[X] + E[Y]$ و در حالت کلی:

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

- به عبارت دیگر امید ریاضی یک عملگر خطی بر روی متغیرهای تصادفی است.

خطی بودن امید ریاضی

اثبات:

$$\begin{aligned}
\mathrm{E}[X+Y] &= \sum_{x,y} (x+y)P_{XY}(x,y) \\
&= \sum_{x,y} xP_{XY}(x,y) + \sum_{x,y} yP_{XY}(x,y) \\
&= \sum_{x} x \sum_{y} P_{XY}(x,y) + \sum_{y} y \sum_{x} P_{XY}(x,y) \\
&= \sum_{x} xP_X(x) + \sum_{y} yP_Y(y) = E[X] + E[Y]
\end{aligned}$$

- به عنوان حالت کلی‌تر داریم:

$$\mathrm{E}(aX + bY) = a\mathrm{E}(X) + b\mathrm{E}(Y)$$

مثال

○ اگر تابع چگالی مشترک متغیرهای تصادفی X و Y به شکل زیر باشد:

$$f_{XY}(x,y) = \begin{cases} 2 & 0 < x < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

امید ریاضی XY چقدر است؟

$$E[XY] = \int_0^1 \int_x^1 xy \times 2 \; dy \, dx$$

$$\int_x^1 2xy \, dy = xy^2 \Big|_x^1 = x(1 - x^2) = x - x^3$$

$$\int_0^1 (x - x^3) dx = \frac{1}{2}x^2 - \frac{1}{4}x^4 \Big|_0^1 = \left(\frac{1}{2} - \frac{1}{4}\right) - 0 = \frac{1}{4} \;\rightarrow\; E[XY] = 0.25$$

متغیرهای تصادفی مستقل گسسته

○ دو متغیر تصادفی گسسته X و Y را مستقل گویند، هرگاه برای هر x_i و y_j،

$$\underbrace{P\{X = x_i, Y = y_j\}}_{P_{XY}(x_i, y_j)} = \underbrace{P\{X = x_i\}}_{P_X(x_i)} \underbrace{P\{Y = y_j\}}_{P_Y(y_j)}$$

○ استقلال دو متغیر تصادفی مشابه استقلال دو پیشامد تعریف می‌شود.

○ تعبیر شهودی: اگر دانستن مقدار متغیر تصادفی X اطلاعاتی راجع به مقدار متغیر تصادفی Y به ما ندهد، این دو متغیر مستقل از هم هستند.

○ استقلال دو متغیر تصادفی یک رابطه دو طرفه است.

مثال ۱

- احتمال شیر آمدن سکه‌ای برابر با p است:
  - سکه را $n+m$ بار پرتاب می‌کنیم.
  - متغیر تصادفی X را تعداد دفعاتی که سکه در n پرتاب اول شیر می‌آید، و متغیر تصادفی Y را تعداد دفعاتی که سکه در m پرتاب بعدی شیر می‌آید، تعریف می‌کنیم.

$$P(X=x,Y=y)=\binom{n}{x}p^x(1-p)^{n-x}\binom{m}{y}p^y(1-p)^{m-y}$$
$$=P(X=x)P(Y=y)$$

- بنابراین X و Y مستقل از هم هستند.

مثال ۲

- فرض کنید Z تعداد کل شیرها در $n+m$ پرتاب باشد.
  - متغیر تصادفی X تعداد شیرها در n پرتاب اول

- آیا X و Z مستقل از هم هستند؟ خیر
  - برای نشان دادن وابستگی X و Z، کافی است مقداری برای Z پیدا کنیم که راجع به مقدار X به ما اطلاعات دهد و یا بالعکس.
  - اگر $Z=0$ باشد، آنگاه واضح است که $X=0$
  - بنابراین دانستن مقدار Z، مقدار دقیق X را مشخص کرد، در صورتی که به فرض استقلال این دو متغیر، آگاهی از مقدار Z نباید هیچ اطلاعاتی راجع به X به ما می‌داد.

مثال ۳

- فرض کنید متغیر تصادفی N تعداد درخواست‌هایی باشد که یک سرور در یک روز دریافت می‌کند.
  - دیدیم که N دارای توزیع پواسون است: $N\sim\mathrm{Poi}(\lambda)$
  - هر درخواست به طور مستقل از یک انسان (با احتمال p) و یا بات (با احتمال $1-p$) می‌آید.
  - X را تعداد درخواست‌های انسان‌ها در یک روز تعریف می‌کنیم:

$$X|N \sim Binomial(N,p)$$

  - Y را تعداد درخواست‌های بات‌ها در یک روز تعریف می‌کنیم:

$$Y|N \sim Binomial(N,1-p)$$

مثال ۳

- از قضیه احتمال کل داریم:

$$P(X=i,Y=j)=P(X=i,Y=j|X+Y=i+j)P(X+Y=i+j)$$
$$+P(X=i,Y=j|X+Y\neq i+j)P(X+Y\neq i+j)$$

- توجه کنید که:

$$P(X=i,Y=j|X+Y\neq i+j)=0$$

بنابراین:

$$P(X=i,Y=j)=P(X=i,Y=j|X+Y=i+j)P(X+Y=i+j)$$

مثال ۳

$$P(X=i,Y=j|X+Y=i+j)=\binom{i+j}{i}p^i(1-p)^j$$ ← توزیع دوجمله‌ای

$$P(X+Y=i+j)=e^{-\lambda}\frac{\lambda^{i+j}}{(i+j)!}$$ ← توزیع پواسون

$$P(X=i,Y=j)=\binom{i+j}{i}p^i(1-p)^j e^{-\lambda}\frac{\lambda^{i+j}}{(i+j)!}$$

$$=\frac{(i+j)!}{i!\,j!}(\lambda p)^i(\lambda(1-p))^j e^{-\lambda(p+1-p)}\frac{1}{(i+j)!}$$

$$=e^{-\lambda p}\frac{(\lambda p)^i}{i!}\times e^{-\lambda(1-p)}\frac{(\lambda(1-p))^j}{j!}$$

مثال ۳

$$P(X=i,Y=j)=e^{-\lambda p}\frac{(\lambda p)^i}{i!}\times e^{-\lambda(1-p)}\frac{(\lambda(1-p))^j}{j!}$$

$$=P(X=i)P(Y=j)$$

که در آن $X\sim Poi(\lambda p)$ و $Y\sim Poi(\lambda(1-p))$.

- بنابراین X و Y مستقل از هم هستند.
- توجه کنید که X و Y توزیع پواسون دارند، ولی $X|N$ و $Y|N$ دارای توزیع دوجمله‌ای هستند. به عبارت دیگر در اختیار داشتن اطلاعات اضافه نظر ما را راجع به شکل توزیع تغییر داد.

متغیرهای تصادفی مستقل پیوسته

- دو متغیر تصادفی پیوسته X و Y را مستقل گویند، هرگاه برای هر x و y، پیشامدهای $\{X \leq x\}$ و $\{Y \leq y\}$ مستقل باشند، یعنی:

$$P\{X \leq x , Y \leq y\} = P\{X \leq x\}P\{Y \leq y\}$$

- به عبارت دیگر:

$$\mathrm{F}_{XY}(x,y) = \mathrm{F}_X(x)\mathrm{F}_Y(y)$$

- اگر از دو طرف مشتق بگیریم ($\frac{\partial^2}{\partial x \partial y}$)، خواهیم داشت:

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

مثال ۱

- تابع چگالی احتمال مشترک زیر، قابل نوشتن به صورت حاصلضرب تابعی از x در تابعی از y نیست، در نتیجه X و Y مستقل نیستند:

$$f_{XY}(x,y) = \begin{cases} \frac{2}{\pi}(1 - x^2 - y^2) & 0 < x^2 + y^2 < 1 \\ 0 & x^2 + y^2 > 1 \end{cases}$$

- داریم:

$$x^2 + y^2 < 1 \;\rightarrow\; y^2 < 1 - x^2 \;\rightarrow\; -\sqrt{1 - x^2} < y < \sqrt{1 - x^2}$$

$$f_X(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{2}{\pi}(1 - x^2 - y^2)\, dy = \frac{8}{3\pi}(1 - x^2)^{\frac{3}{2}} \;:\; -1 < x < 1$$

مثال ۱

- به طور مشابه و با توجه به تقارن خواهیم داشت:

$$f_Y(y) = \frac{8}{3\pi}(1-y^2)^{\frac{3}{2}} \ : \ \ -1 < y < 1$$

- واضح است که:

$$f_{XY}(x,y) \neq f_X(x) f_Y(y)$$

- بنابراین X و Y مستقل نیستند.

مثال ۲

- تابع چگالی احتمال مشترک نرمال را در نظر بگیرید:

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

- می‌توان این تابع را به صورت زیر نوشت:

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \times \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}}$$

- بنابراین: $f_{XY}(x,y) = f_X(x) f_Y(y)$
- در نتیجه X و Y مستقل بوده و هر دو دارای توزیع نرمال $N(0,\sigma)$ می‌باشند.

مثال ۳

○ متغیرهای تصادفی X و Y با تابع چگالی احتمال مشترک زیر را در نظر بگیرید:

$$f_{XY}(x,y) = 4xy \;:\; 0 < x,y < 1$$

این تابع قابل جداسازی به شکل زیر است:

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

$$f_X(x) = 2x : 0 < x < 1 \qquad f_Y(y) = 2y : 0 < y < 1$$

بنابراین X و Y مستقل از هم هستند.

○ اگر تابع چگالی مشترک به صورت $f_{XY}(x,y) = 4xy : 0 < x + y < 1$ باشد، با این که f_{XY} قابل جداسازی است، شرط $x + y < 1$ باعث وابستگی این دو متغیر تصادفی می‌شود.

مثال ۴

○ دو نفر قرار ملاقاتی برای ساعت ۱۲ تنظیم می‌کنند.
  ○ هر یک از آن‌ها به طور مستقل و با توزیع یکنواخت بین ساعت ۱۲ و ۱۲:۳۰ به محل قرار می‌رسد.
  ○ $X =$ میزان دقایقی که نفر اول بعد از ساعت ۱۲ می‌رسد: $X \sim U(0,30)$
  ○ $Y =$ میزان دقایقی که نفر دوم بعد از ساعت ۱۲ می‌رسد: $Y \sim U(0,30)$

○ احتمال این که اولین فردی که به محل قرار می‌رسد بیشتر از ۱۰ دقیقه منتظر دیگری شود چقدر است؟

$$P(X+10<Y) + P(Y+10<X) = 2P(X+10<Y) = ?$$ ← به دلیل تقارن

$$P(X+10<Y) = \iint_{x+10<y} f_{XY}(x,y)dx\,dy = \iint_{x+10<y} f_X(x)f_Y(y)\,dx\,dy$$

مثال ۴

$$f_X(x) = \frac{1}{30} : 0 < x < 30 \quad , \quad f_Y(y) = \frac{1}{30} : 0 < y < 30$$

$$P(X + 10 < Y) = \int_{10}^{30} \int_{0}^{y-10} \frac{1}{30} \times \frac{1}{30}\, dx\, dy$$

$$= \int_{10}^{30} \frac{1}{900}(y - 10)dy = \frac{1}{900}\left(\frac{1}{2}y^2 - 10y\right)\Big|_{10}^{30}$$

$$= \frac{1}{900}\left(\frac{900 - 100}{2} - 10(30 - 10)\right) = \frac{200}{900} = \frac{2}{9}$$

$$2P(X + 10 < Y) = 2 \times \frac{2}{9} = \frac{4}{9}$$ احتمال انتظار بیش از ده دقیقه ←

قضیه

قضیه. اگر X و Y مستقل باشند، آنگاه $Z = g(X)$ و $W = h(Y)$ نیز مستقل خواهند بود.

اثبات: در حالت خاصی که توابع g و h صعودی باشند، داریم:

$$P\{g(X) \le z , h(Y) \le w\} = P\{X \le g^{-1}(z), Y \le h^{-1}(w)\}$$

$$= P\{X \le g^{-1}(z)\}P\{Y \le h^{-1}(w)\}$$ (به دلیل استقلال دو متغیر تصادفی)

$$= P\{g(X) \le z\}P\{h(Y) \le w\}$$

مثال

○ فرض کنید X و Y دو متغیر تصادفی مستقل از هم باشند که هر دو دارای توزیع نمایی با پارامتر λ هستند. متغیر تصادفی W را به صورت $W = \mathrm{Min}\{X,Y\}$ تعریف می‌کنیم. تابع چگالی احتمال W را محاسبه کنید.

$$F_W(w) = P(W \le w) = 1 - P(W > w)$$

$$= 1 - P(X > w, Y > w) = 1 - P(X > w)P(Y > w) \quad \leftarrow \text{استقلال } X \text{ و } Y$$

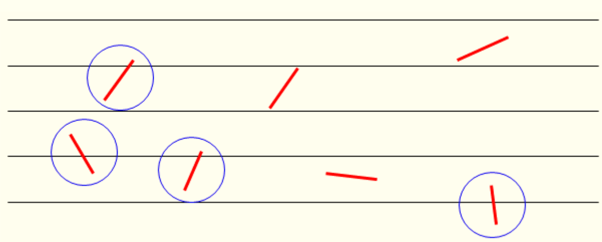
$$F_W(w) = 1 - \big(1 - F_X(w)\big)\big(1 - F_Y(w)\big)$$

$$= 1 - \left(e^{-\lambda w}\right)\left(e^{-\lambda w}\right) = 1 - e^{-2\lambda w}$$

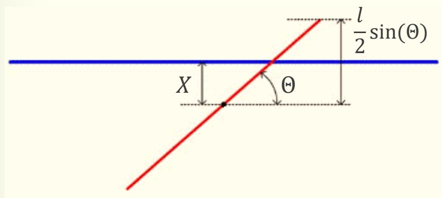
$$f_W(w) = \frac{dF_W(w)}{dw} = 2\lambda e^{-2\lambda w}$$

مثال: سوزن بوفون (Buffon’s Needle)

○ سوزنی به طول l را بر روی صفحه‌ای با خطوط موازی به فاصله t می‌اندازیم. احتمال این که سوزن یکی از این خطوط را قطع کند چقدر است؟



سوزن بوفون



- متغیر تصادفی X را فاصله مرکز سوزن از نزدیکترین خط می‌گیریم:

$X \sim U(0, t/2)$

- متغیر تصادفی Θ را زاویه حاده بین سوزن و خطوط می‌گیریم:

$\Theta \sim U(0, \pi/2)$

- متغیرهای تصادفی X و Θ مستقل از هم هستند، بنابراین تابع چگالی احتمال مشترک آن‌ها برابر است با:

$$P_{X,\Theta}(x,\theta) = P_X(x)P_\Theta(\theta) = \frac{1}{t/2} \times \frac{1}{\pi/2} = \frac{4}{t\pi}$$

$$0 < x < t/2 \;, \qquad 0 < \theta < \pi/2$$

سوزن بوفون

- سوزن در صورتی خط را قطع می‌کند که: $X \le \frac{l}{2}\sin(\Theta)$

$$P\left(X \le \frac{l}{2}\sin(\Theta)\right) = \int_0^{\pi/2} \int_0^{\frac{l}{2}\sin(\theta)} \frac{4}{t\pi}\, dx\; d\theta$$

$$\int_0^{\frac{l}{2}\sin(\theta)} \frac{4}{t\pi}\, dx = \frac{4}{t\pi} \times \left(\frac{l}{2}\sin(\theta) - 0\right) = \frac{2l}{t\pi}\sin(\theta)$$

$$\int_0^{\frac{\pi}{2}} \frac{2l}{t\pi}\sin(\theta)\; d\theta = -\frac{2l}{t\pi}\cos(\theta)\Big|_0^{\frac{\pi}{2}} = \frac{2l}{t\pi}$$

استقلال بیش از دو متغیر تصادفی

- n متغیر تصادفی گسسته $X_1, X_2, \ldots, X_n$ مستقل از هم هستند، اگر برای هر زیرمجموعه از X_iها داشته باشیم:

$$P\left(X_{i_1} = x_1, X_{i_2} = x_2, \ldots, X_{i_r} = x_r\right) = \prod_{k=1}^{r} P\left(X_{i_k} = x_k\right)$$

- و به طور خاص برای هر n متغیر:

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = \prod_{k=1}^{n} P(X_k = x_k)$$

استقلال بیش از دو متغیر تصادفی

- به همین ترتیب برای حالت پیوسته شرط استقلال به صورت زیر است:

$$P(X_1 \le a_1, X_2 \le a_2, \ldots, X_n \le a_r) = \prod_{k=1}^{n} P(X_k \le a_k)$$

- به عبارت دیگر:

$$F_{X_1 X_2 \ldots X_n}(a_1, a_2, \ldots, a_n) = \prod_{k=1}^{n} F_{X_k}(a_k)$$

و در نتیجه

$$f_{\vec{X}}(\vec{x}) = f_{X_1 X_2 \ldots X_n}(x_1, x_2, \ldots, x_n) = \prod_{k=1}^{n} f_{X_k}(x_k)$$

متغیرهای تصادفی i.i.d

- فرض کنید توزیع همه متغیرهای تصادفی X_i یکسان باشد، به عبارت دیگر داریم:

$$f_{X_i}(x) = f_X(x)$$

- همچنین فرض کنید X_iها مستقل هستند.
- به طور کلی، متغیرهای تصادفی که مستقل و دارای توزیع یکسان باشند را *i.i.d* می‌نامیم که مخفف عبارت Independent Identically Distributed است.
- اگر $\vec{X} = [X_1, X_2, \dots, X_n]$ و X_iها *i.i.d* باشند، داریم:

$$f_{\vec{X}}(\vec{x}) \underset{\downarrow}{=} f_{X_1}(x_1) f_{X_2}(x_2) \dots f_{X_n}(x_n) \underset{\downarrow}{=} f_X(x_1) f_X(x_2) \dots f_X(x_n)$$

استقلال

یکسان بودن توزیع

مثال

- فرض کنید یک سکه با احتمال رو آمدن p را n بار پرتاب می‌کنیم.
- متغیر تصادفی برنولی X_i را برابر خروجی پرتاب i-ام تعریف می‌کنیم:

$$X_i \sim Ber(p) \;\rightarrow\; P_{X_i}(x) = p^x (1-p)^{1-x}$$

- واضح است که توزیع همه X_iها یکسان است (از یک سکه ثابت در همه پرتاب‌ها استفاده شده) و X_i مستقل از هم هستند (هر پرتاب مستقل از سایر پرتاب‌هاست). بنابراین X_iها متغیرهای تصادفی i.i.d هستند.
- اگر $\vec{X} = [X_1, X_2, \dots, X_n]$ داریم:

$$P_{\vec{X}}(\vec{x}) = P_X(x_1) P_X(x_2) \dots P_X(x_n) = \prod_{i=1}^{n} p^{x_i} (1-p)^{1-x_i} = p^{\sum x_i} (1-p)^{n-\sum x_i}$$