آمار و احتمال مهندسی

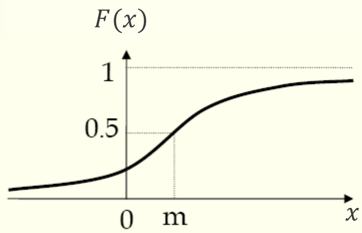
تابعی از یک متغیر تصادفی

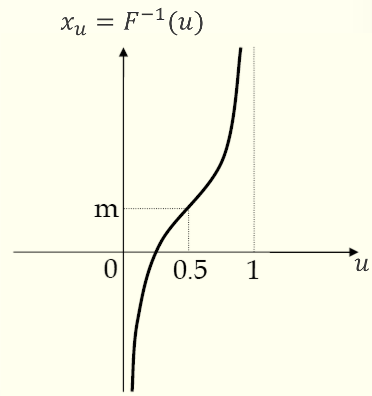
(Ross 5.7)

صدک‌ها (نقاط درصد یا percentile)

- وقتی می‌گوییم $F_X(x) = u$ یعنی به احتمال u متغیر تصادفی X از مقدار x بزرگتر نمی‌شود (کوچکتر یا مساوی است).
- فرض کنید می‌خواهیم ببینیم متغیر تصادفی X به احتمال 0.95 کوچکتر یا مساوی کدام x است؟ $F_X(x) = P(X \leq x) = 0.95$
- u در فاصله $[0,1]$ داده شده است و x در فاصله $(-\infty, +\infty)$ مطلوب است: $x_u = F^{-1}(u)$
- x_u را صدک $(100u)$-ام متغیر تصادفی X می‌گوییم. مثلاً $x_{0.95}$ صدک نود و پنجم است.

صدک

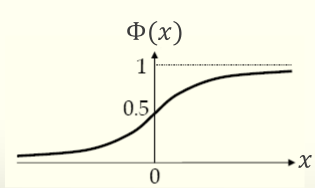




دهک، چارک، و میانه

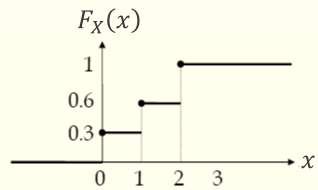
- صدک $10k$-ام را دهک (decile) k-ام می‌نامیم.
  - برای مثال صدک دهم را دهک اول، و صدک بیستم را دهک دوم گوییم.
- صدک ۲۵-ام را چارک (quartile) اول یا پایین و صدک ۷۵-ام را چارک سوم یا بالا گویند.
- صدک ۵۰-ام را چارک دوم یا میانه (median) می‌نامیم.
- پس میانه متغیر تصادفی X عددی مانند m است که: $F_X(m) = 0.5$
- مثال. برای متغیر تصادفی گوسی داریم:

$$\Phi(0) = 0.5 \Rightarrow m = 0$$



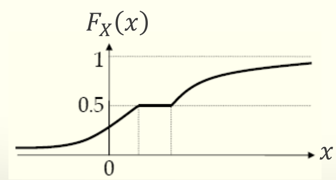
میانه

- ولی ممکن است F پرش داشته باشد و در هیچ نقطه‌ای $F_X(x) = 0.5$ نشود.



$$\longrightarrow \quad m = 1$$

- یا ممکن است یک بازه میانه باشد:



تعریف دقیق میانه

- برای متغیر تصادفی X، میانه m مقداری است که برای آن هر دو شرط زیر برقرار باشند:

$$1)\ P\{X \geq m\} \geq \frac{1}{2}$$

و

$$2)\ P\{X \leq m\} \geq \frac{1}{2}$$

- این تعریف حتی در صورتی که میانه به صورت یک بازه باشد و یا تابع F پرش داشته باشد نیز میانه را به درستی مشخص می‌کند.

مثال

- متغیر تصادفی Y با تابع چگالی احتمال زیر مفروض است:

$$f_Y(y) = 3y^2 : \quad 0 < y < 1$$

- میانه Y را محاسبه کنید.

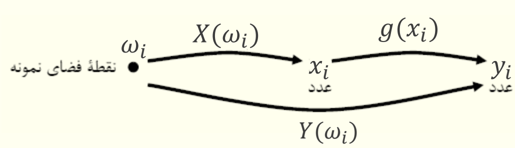
$$F_Y(m) = P(Y \leq m) = 0.5$$

$$\int_{-\infty}^{m} f_Y(y)dy = \frac{1}{2} \Rightarrow \int_{0}^{m} 3y^2 dy = \frac{1}{2}$$

$$\Rightarrow \ y^3 \Big|_0^m = \frac{1}{2} \Rightarrow m^3 - 0 = \frac{1}{2} \Rightarrow m = \frac{1}{\sqrt[3]{2}} \approx 0.7937$$

تابع یک متغیر تصادفی

- در بسیاری موارد با فرض دانستن تابع توزیع یک متغیر تصادفی، با توجه به اینکه در سیستم مورد نظر عملیاتی روی آن انجام میشود، علاقه‌مندیم که توزیع تابعی از یک متغیر تصادفی را به دست آوریم.
  - توزیع خروجی یک الگوریتم که مقادیر ورودی آن دارای توزیع یکنواخت هستند.
- اگر $g(x)$ یک تابع حقیقی باشد، می‌توانیم به هر $x_i = X(\omega_i)$، $y_i = g(x_i)$ را نسبت دهیم:



- پس با ترکیب دو تابع مواجه هستیم که یک متغیر تصادفی جدید به ما می‌دهد:

$$Y(\omega) = g(X(\omega))$$

به دست آوردن توزیع $Y = g(X)$ از توزیع X

- در حالت گسسته حل مساله ساده است:

$$P_Y(y) = P\{Y = y\} = P\{g(X) = y\} = \sum_{i:g(x_i)=y} P(x_i)$$

$$F_Y(y) = P\{Y \le y\} = P\{g(X) \le y\} = \sum_{i:g(x_i)\le y} P(x_i)$$

- مثلاً اگر متغیر تصادفی X دارای توزیع زیر باشد:

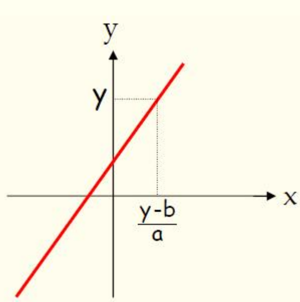
x_i	0	1	-1
P_i	1/3	1/3	1/3

برای $Y = X^2$ داریم:

$$P_Y(1) = P\{Y = 1\} = P\{X = 1\} + P\{X = -1\} = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$$

به دست آوردن توزیع $Y = g(X)$ از توزیع X

- اما برای حالت‌های غیرگسسته ابتدا به محاسبهٔ تابع CDF می‌پردازیم که ساده‌تر است.
- مثال ۱: $g(x) = ax + b$



$$Y = aX + b$$

$$F_Y(y) = P\{Y \le y\} = P\{aX + b \le y\}$$

$$= \begin{cases} P\left\{X \le \dfrac{y-b}{a}\right\} = F_X\left(\dfrac{y-b}{a}\right) & a > 0 \\ P\left\{X \ge \dfrac{y-b}{a}\right\} = 1 - F_X\left(\dfrac{y-b}{a}\right) & a < 0 \end{cases}$$

به دست آوردن توزیع $y = g(x)$ از توزیع x

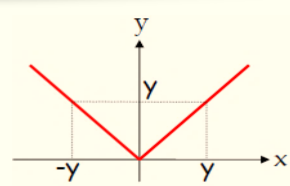
$$F_Y(y) = \begin{cases} P\left\{X \le \dfrac{y-b}{a}\right\} = F_X\left(\dfrac{y-b}{a}\right) & a > 0 \\ P\left\{X \ge \dfrac{y-b}{a}\right\} = 1 - F_X\left(\dfrac{y-b}{a}\right) & a < 0 \end{cases}$$

- با مشتق‌گیری نسبت به y داریم:

$$\Rightarrow f_Y(y) = \begin{cases} \dfrac{1}{a} f_X\left(\dfrac{y-b}{a}\right) & a > 0 \\ -\dfrac{1}{a} f_X\left(\dfrac{y-b}{a}\right) & a < 0 \end{cases} = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$

مثال ۲

$$g(x) = |x| \quad \rightarrow \quad Y = |X|$$



$$F_Y(y) = P\{Y \le y\} = P\{|X| \le y\}$$

$$= \begin{cases} 0 & y < 0 \\ P\{-y \le X \le y\} = F_X(y) - F_X(-y) & y > 0 \end{cases}$$

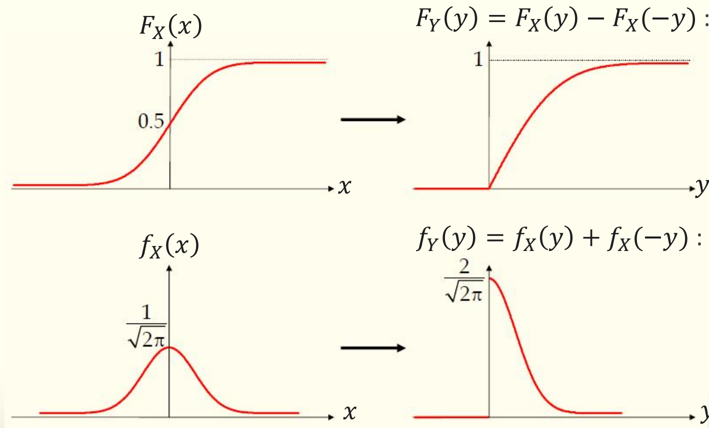
با مشتق‌گیری داریم:

$$f_Y(y) = \begin{cases} f_X(y) + f_X(-y) & y \ge 0 \\ 0 & y < 0 \end{cases}$$

ادامه مثال ۲ برای $X \sim N(0,1)$

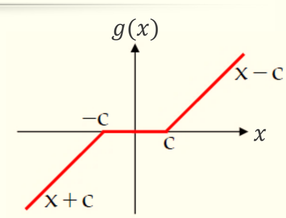
$F_Y(y) = F_X(y) - F_X(-y) : y > 0$

$f_Y(y) = f_X(y) + f_X(-y) : y > 0$



مثال ۳

$$g(x) = \begin{cases} x + c & x < -c \\ 0 & -c \le x \le c \\ x - c & x > c \end{cases} \quad \text{و} \quad Y = g(X)$$

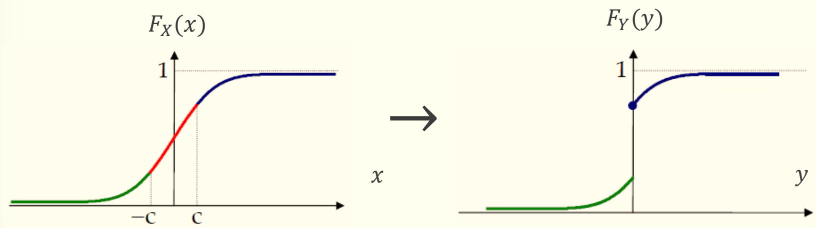


$$F_Y(y) = P\{Y \le y\} = \begin{cases} P\{X + c \le y\} & y < 0 \\ P\{X - c \le y\} & y > 0 \end{cases}$$

$$F_Y(y) = \begin{cases} F_X(y - c) & y < 0 \\ F_X(y + c) & y > 0 \end{cases} \quad \text{و} \quad P\{Y = 0\} = P\{-c < X < c\} = F_X(c) - F_X(-c)$$

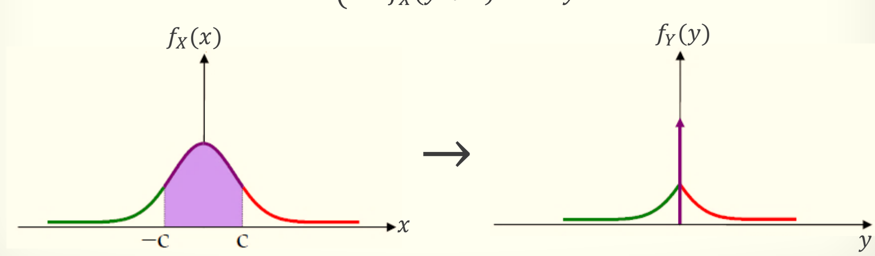
ادامه مثال ۳

$$F_Y(y) = \begin{cases} F_X(y-c) & y < 0 \\ F_X(y+c) & y \geq 0 \end{cases} \quad \text{و} \quad P\{Y=0\} = F_X(c) - F_X(-c)$$



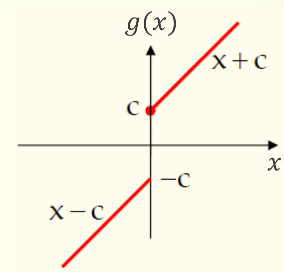
ادامه مثال ۳

$$f_Y(y) = \begin{cases} f_X(y-c) & y < 0 \\ \delta(y)\displaystyle\int_{-c}^{c} f_X(u)du & y = 0 \\ f_X(y+c) & y > 0 \end{cases}$$



مثال ۴

$$g(x) = \begin{cases} x + c & x \geq 0 \\ x - c & x < 0 \end{cases}$$



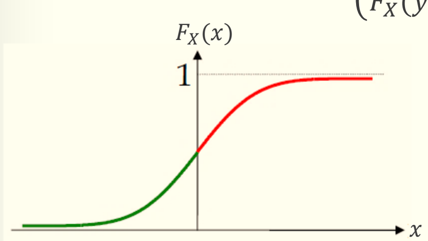
$$F_Y(y) = P\{Y \leq y\}$$

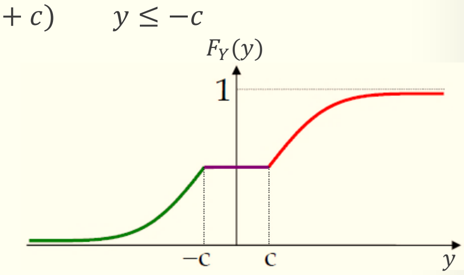
$$F_Y(y) = \begin{cases} P\{X + c \leq y\} & y \geq c \\ P\{X \leq 0\} & -c < y < c \\ P\{X - c \leq y\} & y \leq -c \end{cases}$$

$$F_Y(y) = \begin{cases} F_X(y - c) & y \geq c \\ F_X(0) & -c < y < c \\ F_X(y + c) & y \leq -c \end{cases}$$

ادامه مثال ۴

$$F_Y(y) = \begin{cases} F_X(y - c) & y \geq c \\ F_X(0) & -c < y < c \\ F_X(y + c) & y \leq -c \end{cases}$$

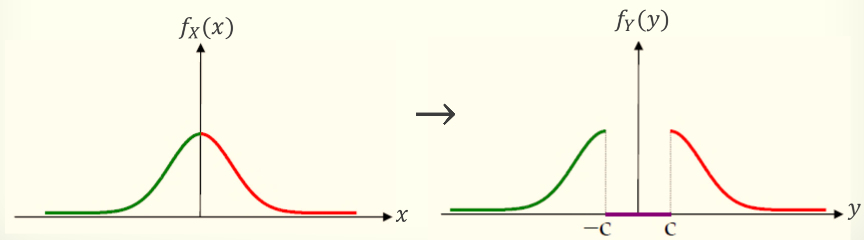




- به یاد داشته باشید که بی‌تغییر ماندن $g(x)$ در یک محدوده سبب پرش در F_Y می‌شود و به عکس پرش در $g(x)$ موجب بی‌تغییر ماندن F_Y در یک محدوده می‌شود.

ادامه مثال ۴

$$f_Y(y) = \begin{cases} f_X(y-c) & y \geq c \\ 0 & -c < y < c \\ f_X(y+c) & y \leq -c \end{cases}$$



محاسبه مستقیم f_Y بر حسب f_X

- قضیه: برای y داده شده، اگر معادله $g(x) = y$ دارای جواب‌های x_1، x_2، و ... باشد، خواهیم داشت:

$$f_Y(y) = \sum_i \frac{f_X(x_i)}{|g'(x_i)|}$$

که در آن:

$$f_X(x_i) = f_X(x)\Big|_{x=x_i(y)}$$

$$g'(x_i) = \frac{d}{dx} g(x)\Big|_{x=x_i(y)}$$

مشروط بر این که برای y داده شده، تعداد نقاط x_i قابل شمارش باشد و $g(x)$ در نقاط x_i مشتق‌پذیر باشد.

مثال

- اگر $Y = aX^2$ و $a > 0$، تابع چگالی احتمال f_Y را بر حسب f_X بیابید.

برای $y > 0$، معادله $y = ax^2$ دو ریشه دارد: $x_1 = \sqrt{\frac{y}{a}}$ و $x_2 = -\sqrt{\frac{y}{a}}$

از سوی دیگر $g'(x) = 2ax$، بنابراین:

$$g'(x_1) = 2a\sqrt{\frac{y}{a}} = 2\sqrt{ay} \quad و \quad g'(x_2) = -2a\sqrt{\frac{y}{a}} = -2\sqrt{ay}$$

بنابراین داریم:

$$f_Y(y) = \frac{f_X(x_1)}{g'(x_1)} + \frac{f_X(x_2)}{|g'(x_2)|} = \frac{f_X\left(\sqrt{\frac{y}{a}}\right)}{2\sqrt{ay}} + \frac{f_X\left(-\sqrt{\frac{y}{a}}\right)}{2\sqrt{ay}}$$

اگر $y < 0$، معادله $g(x) = y$ ، جواب ندارد و $f_Y(y) = 0$.

راه حل دوم

- با استفاده از CDF داریم:

$$F_Y(y) = P\{Y \le y\} = P\{-\sqrt{\frac{y}{a}} \le X \le \sqrt{\frac{y}{a}}\}$$

$$F_Y(y) = F_X\left(\sqrt{\frac{y}{a}}\right) - F_X\left(-\sqrt{\frac{y}{a}}\right) \ : \ y > 0$$

با مشتق‌گیری داریم:

$$f_Y(y) = \frac{f_X\left(\sqrt{\frac{y}{a}}\right)}{2\sqrt{ay}} + \frac{f_X\left(-\sqrt{\frac{y}{a}}\right)}{2\sqrt{ay}}$$

مثال

- فرض کنید X یک متغیر تصادفی دارای توزیع نمایی با پارامتر λ باشد. متغیر تصادفی Z به صورت $Z = X^3$ تعریف می‌شود. تابع چگالی احتمال Z را به دست آورید.

$$f_X(x) = \lambda e^{-\lambda x} u(x)$$

$$z = x^3 \ \Rightarrow x_1 = z^{\frac{1}{3}}$$

$$g(x) = x^3 \ \Rightarrow g'(x) = 3x^2 \ \ \Rightarrow \ \ g'(x_1) = 3z^{\frac{2}{3}}$$

$$f_Z(z) = \frac{f_X(x_1)}{g'(x_1)} = \frac{f_X\left(z^{\frac{1}{3}}\right)}{3z^{\frac{2}{3}}} = \frac{\lambda}{3} z^{-\frac{2}{3}} e^{-\lambda z^{\frac{1}{3}}} u(z)$$

مثال

- فرض کنید تابع randU() با توزیع یکنواخت یک عدد تصادفی حقیقی در بازه (0,1) برمی‌گرداند.
- توزیع خروجی تابع unknownRand() چیست؟

```
double unknownRand( ) {
        double x, y ;
        x = randU( ) ;
        y = x++ ;
        y *= ++x ;
        return y ;
}
```

$$X \sim U(0,1) \ \rightarrow \ \ f_X(x) = 1 : 0 < x < 1$$

$$Y = X(X+2) = X^2 + 2X$$

$$y = x^2 + 2x \ \ \rightarrow y + 1 = (x+1)^2$$

$$\rightarrow \ \ x_1 = -1 + \sqrt{y+1} \ , x_2 = -1 - \sqrt{y+1}$$

ادامه مثال

- از آنجا که $0 < x < 1$ ، جواب $x_2 = -1 - \sqrt{y+1}$ قابل قبول نیست.

$$0 < x_1 < 1 \;\rightarrow\; 0 < -1 + \sqrt{y+1} < 1 \;\rightarrow\; 0 < y < 3$$

$$g(x) = x^2 + 2x \;\Rightarrow g'(x) = 2x + 2 \;\Rightarrow\; g'(x_1) = 2\sqrt{y+1}$$

$$f_Y(y) = \frac{f_X(x_1)}{g'(x_1)} = \frac{1}{2\sqrt{y+1}} \;:\; 0 < y < 3$$