Example: Now, suppose there are 20 distinct items

$$\begin{tikzcd}[column sep=small] & \text{20 items} \arrow[dl, "\text{Group A}"'] \arrow[dr, "\text{Group B}"] & \\ N_A = 4 \text{ items} & & N - N_A = 16 \text{ items} \end{tikzcd}$$

Select the items at random.

(a) If I select the items for 9 times, one item each time, with replacement, define

X = # of time I select the items from Group A.

calculate $\mathbb{P}(X \geq 1)$?

Soln: $X \sim Bin(9, p)$ $\quad p = \frac{4}{20} = \frac{1}{5}$

$$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X=0) = 1 - \binom{9}{0}\left(\frac{1}{5}\right)^0\left(\frac{4}{5}\right)^9 = 1 - \left(\frac{4}{5}\right)^9 \quad //$$

(b) If I select 9 distinct items at random, calculate $\mathbb{P}(X \geq 1)$?

Soln: $\{X=0\} = \{$All 9 items come from Group B$\}$

$$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X=0) = 1 - \frac{\binom{4}{0}\binom{16}{9}}{\binom{20}{9}} = \text{simplify it} \quad //$$

Transformation of Random Variables

X : a R.V. $\quad g : \mathbb{R} \to \mathbb{R}$

Q: Given the distribution of X, how can I identify the distribution of $Y = g(X)$?

A: – If X is discrete, use P.M.F. Method.

– If X is continuous, use C.D.F. Method.

Example for Discrete R.V.: X has a p.m.f

$$\begin{cases} \mathbb{P}(X=-1) = 1/7 \\ \mathbb{P}(X=0) = 1/14 \\ \mathbb{P}(X=2) = 3/14 \\ \mathbb{P}(X=4) = 4/7 \end{cases}$$

Find the p.m.f of

$$Y = (X-1)^2$$

Soln: Draw a Table

X	-1	0	2	4
$Y=(X-1)^2$	4	1	1	9
$\mathbb{P}$	$1/7$	$1/14$	$3/14$	$4/7$

we have $\mathbb{P}(Y=1) = \mathbb{P}(X=0) + \mathbb{P}(X=2) = \frac{1}{14} + \frac{3}{14} = \frac{2}{7}$

$$\mathbb{P}(Y=4) = \mathbb{P}(X=-1) = \frac{1}{7}$$

$$\mathbb{P}(Y=9) = \mathbb{P}(X=4) = \frac{4}{7} \quad //$$

Example for Continuous R.V. : $X \sim N(0,1)$ w.

p.d.f : $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, c.d.f $\Phi(x) = \int_{-\infty}^{x} \phi(t)\,dt$

Identify the density of $Y = X^2$. (Hint : use $\Gamma(1/2) = \sqrt{\pi}$)

Soln : Use C.D.F MTD : for $y > 0$, C.D.F of Y

$$F_Y(y) = \mathbb{P}(Y \le y)$$
$$= \mathbb{P}(X^2 \le y)$$
$$= \mathbb{P}(-\sqrt{y} \le X \le \sqrt{y}) = \Phi(\sqrt{y}) - \Phi(-\sqrt{y})$$

Then, we have p.d.f :

$$f_Y(y) = \frac{d}{dy} F_Y(y)$$
$$= \frac{d}{dy}\{\Phi(\sqrt{y}) - \Phi(-\sqrt{y})\}$$

$\phi(\sqrt{y}) = \phi(-\sqrt{y})$

$$= \frac{1}{2\sqrt{y}}\phi(\sqrt{y}) + \frac{1}{2\sqrt{y}}\phi(-\sqrt{y}) = \frac{1}{\sqrt{y}}\phi(\sqrt{y})$$

$$= \frac{1}{\sqrt{2\pi}} y^{-\frac{1}{2}} e^{-\frac{y}{2}} = \frac{(1/2)^{1/2}}{\Gamma(1/2)} y^{\frac{1}{2}-1} e^{-y/2} \quad y > 0$$

$$\Rightarrow Y \sim \text{Gamma}\left(\tfrac{1}{2}, \tfrac{1}{2}\right)$$

$= \chi_1^2$ (Chi-Square distribution, 120B) //

More Examples:

① $X \sim \text{Unif}[-2, 3]$, $Y = |X-1|$, find density of Y.

Soln: Use C.D.F again, for $y > 0$

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

$$= \mathbb{P}(|X-1| \le y)$$

$$= \mathbb{P}(-y \le X-1 \le y)$$

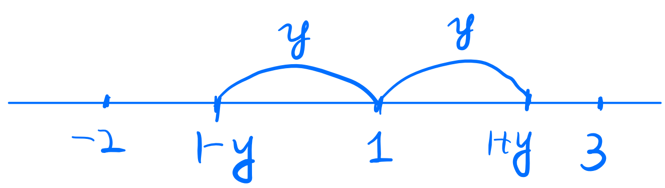
$$= \mathbb{P}(1-y \le X \le 1+y)$$

$$= F_X(1+y) - F_X(1-y)$$

p.d.f $f_Y(y) = \frac{d}{dy} F_Y(y) = f_X(1+y) + f_X(1-y)$

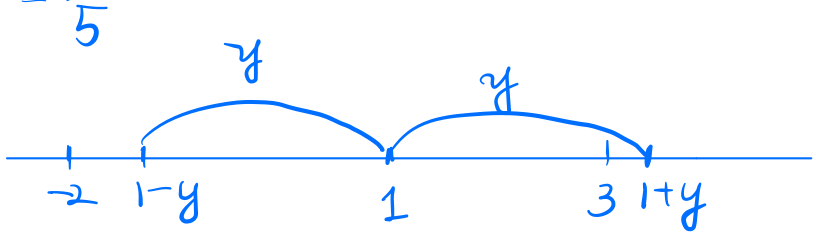
Note that $f_X(x) = \frac{1}{5}$ for $-2 < x < 3$, then

(1) $0 < y < 2$:



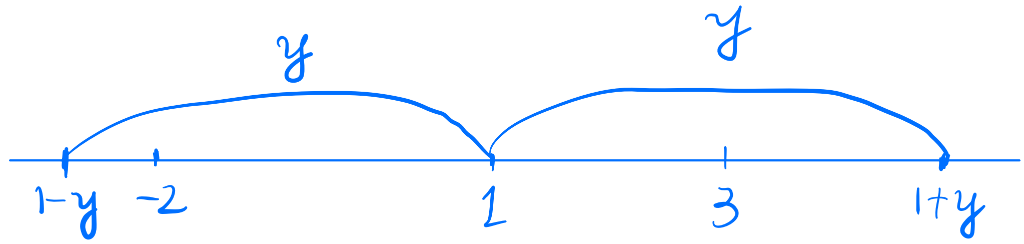
$$f_X(1+y) = f_X(1-y) = \frac{1}{5}$$

(2) $2 \le y < 3$:



$$f_X(1+y) = 0, \quad f_X(1-y) = \frac{1}{5}$$

(3) $y \geq 3$:



$$f_X(1+y) = f_X(1-y) = 0$$

Thus,

$$f_Y(y) = \begin{cases} 2/5 & 0 < y < 2 \\ 1/5 & 2 \leq y < 3 \\ 0 & \text{o.w.} \end{cases}$$

//

② $X \sim N(0,1)$, $Y = e^X$

(a) Calculate p.d.f of X.

<u>Soln</u>: $F_Y(t) = \mathbb{P}(Y \leq t) = \mathbb{P}(e^X \leq t)$

$$= \mathbb{P}(X \leq \log t) = F_X(\log t) \qquad t > 0$$

so $f_Y(t) = \frac{d}{dt} F_Y(t) = \frac{1}{t} f_X(\log t)$

where $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$ $(x \in \mathbb{R})$ is p.d.f of X.

Thus, $f_Y(t) = \begin{cases} \frac{1}{t\sqrt{2\pi}} \exp\{-\frac{(\log t)^2}{2}\} & t > 0 \\ 0 & t \leq 0 \end{cases}$ //

(b) Find the n-th moment $\mathbb{E}[Y^n]$ of Y.

Soln: $\mathbb{E}[Y^n] = \mathbb{E}[(e^X)^n] = \mathbb{E}[e^{nX}] = M_X(n)$

where $M_X(t) = e^{\frac{1}{2}t^2}$ is the MGF of X.

Thus $\mathbb{E}[Y^n] = e^{n^2/2}$. //

Belows are extra exercises :

Ex 1. Suppose a R.V. X has density

$$f(x) = \begin{cases} 6x(1-x) & \text{for } 0<x<1 \\ 0 & \text{elsewhere} \end{cases}$$

Calculate cumulative distribution function for X, $\mathbb{P}(X<1/4)$, and $\mathbb{P}(X>1/2)$.

Soln: By definition of c.d.f / p.d.f, we have

$$\mathbb{P}(X \le x) = \int_{-\infty}^{x} f(t)\,dt$$

(i) $x \le 0$: $\mathbb{P}(X \le x) = \int_{-\infty}^{x} 0\,dt = 0$

(ii) $x \in (0,1)$: $\mathbb{P}(X \le x) = \int_{-\infty}^{x} f(t)\,dt$

$$= \int_{-\infty}^{0} 0\,dt + \int_{0}^{x} 6t(1-t)\,dt = 3t^2 - 2t^3 \Big|_0^x$$

$$= 3x^2 - 2x^3$$

(iii) $x \ge 1$: $F(x) = 1$

so $$F(x) = \begin{cases} 0 & x \le 0 \\ 3x^2 - 2x^3 & 0<x<1 \\ 1 & x \ge 1 \end{cases}$$

$$\mathbb{P}(X<1/4) = F(1/4) = \frac{3}{16} - \frac{2}{64} = \frac{5}{32}$$

$$\mathbb{P}(X>1/2) = 1 - F(1/2) = 1 - \left(\frac{3}{4} - \frac{2}{8}\right) = \frac{1}{2}$$

Ex 2: Suppose a R.V. X has density

$$f(x) = \begin{cases} \frac{x}{2} & 0 < x \leq 1 \\ \frac{1}{2} & 1 < x \leq 2 \\ \frac{3-x}{2} & 2 < x \leq 3 \\ 0 & \text{elsewhere} \end{cases}$$

Calculate the distribution function of X.

Soln: (i) $x \leq 0$, $F(x) = 0$

(ii) $0 < x \leq 1$: $F(x) = \int_{-\infty}^{0} 0\,dt + \int_0^x \frac{t}{2}\,dt = \left.\frac{t^2}{4}\right|_0^x = \frac{x^2}{4}$

(iii) $1 < x \leq 2$: $F(x) = \int_{-\infty}^{0} 0\,dt + \int_0^1 \frac{t}{2}\,dt + \int_1^x \frac{1}{2}\,dt$

$$= \frac{1}{4} + \frac{1}{2}(x-1)$$

(iv) $2 < x \leq 3$: $F(x) = \int_{-\infty}^{0} 0\,dt + \int_0^1 \frac{t}{2}\,dt + \int_1^2 \frac{1}{2}\,dt + \int_2^x \frac{3-t}{2}\,dt$

$$= \frac{3}{4} + \left.\frac{3t - t^2/2}{2}\right|_2^x$$

$$= \frac{3}{4} + \left(\frac{6x - x^2}{4} - \frac{6-2}{2}\right) = \frac{3x}{2} - \frac{x^2}{4} - \frac{5}{4}$$

(v) $x > 3$: $F(x) = 1$

Thus, $F(x) = \begin{cases} 0 & x \leq 0 \\ x^2/4 & 0 < x \leq 1 \\ x/2 - 1/4 & 1 < x \leq 2 \\ 3x/2 - x^2/4 - 5/4 & 2 < x \leq 3 \\ 1 & x > 3 \end{cases}$

Ex 3. Let X has distribution function (c.d.f.)

$$F(x) = \begin{cases} 0 & z < -2 \\ \frac{x+4}{8} & -2 \leq z < 2 \\ 1 & z \geq 2 \end{cases}$$

Calculate $\mathbb{P}(Z=-2)$, $\mathbb{P}(Z=2)$, $\mathbb{P}(-2<Z\leq 1)$, and $\mathbb{P}(2<Z\leq 3)$.

Soln: $\mathbb{P}(Z=-2) = F(-2) - \lim_{z \uparrow -2^-} F(z) = 0 - 0 = 0$

$\mathbb{P}(Z=2) = F(2) - \lim_{z \uparrow 2^-} F(z) = 1 - 1 = 0$

$\mathbb{P}(-2<Z\leq 1) = F(1) - F(-2) = \frac{1+4}{8} - 0 = \frac{5}{8}$

$\mathbb{P}(2<Z\leq 3) = F(3) - F(2) = 1 - 1 = 0.$

Ex 4. The c.d.f. of X is given by

$$F(x) = \begin{cases} 1-(1+x)e^{-x} & x>0 \\ 0 & x \leq 0 \end{cases}$$

then its density $f(x) = \begin{cases} xe^{-x} & x>0 \\ 0 & x<0 \end{cases}$ $\quad X \sim \text{Gamma}(2,1)$