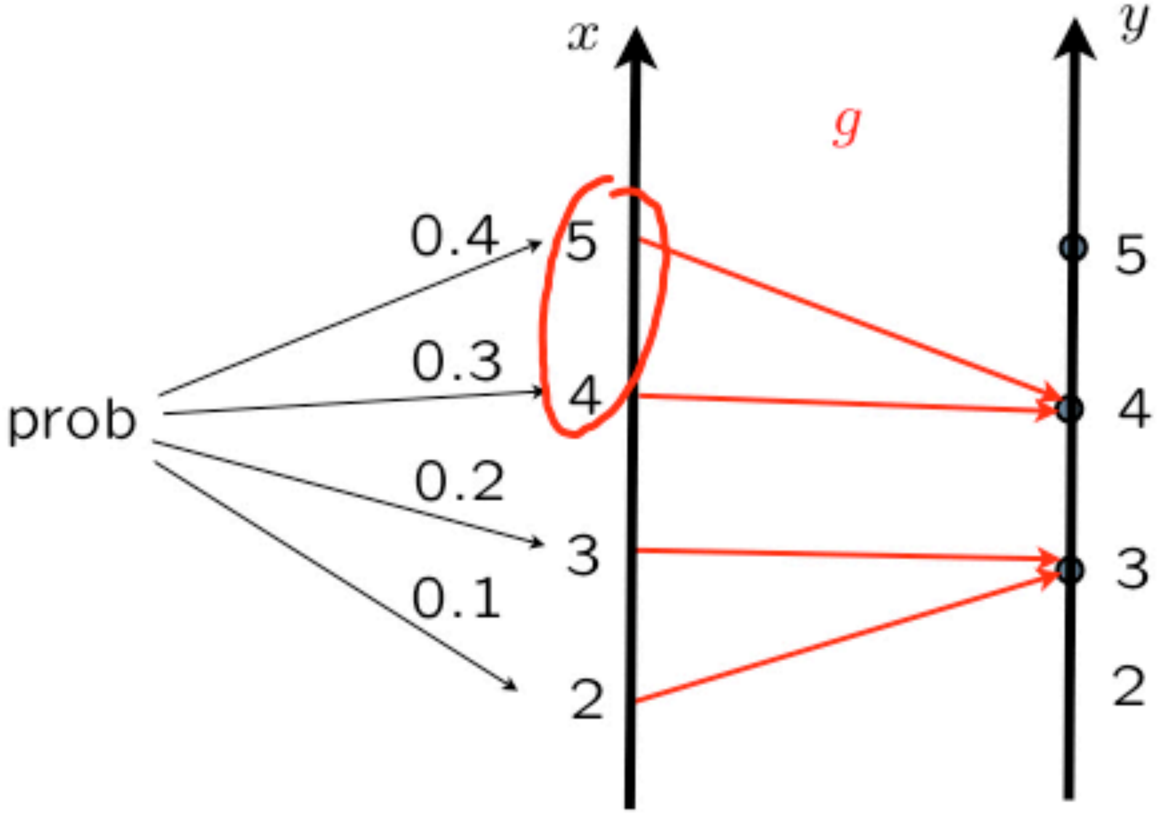
LECTURE 11: Derived distributions

- Given the distribution of X,
  find the distribution of $Y = g(X)$
  - the discrete case
  - the continuous case
  - general approach, using CDFs
  - the linear case: $Y = aX + b$
  - general formula when g is monotonic

- Given the (joint) distribution of X and Y,
  find the distribution of $Z = g(X, Y)$

Derived distributions — the discrete case

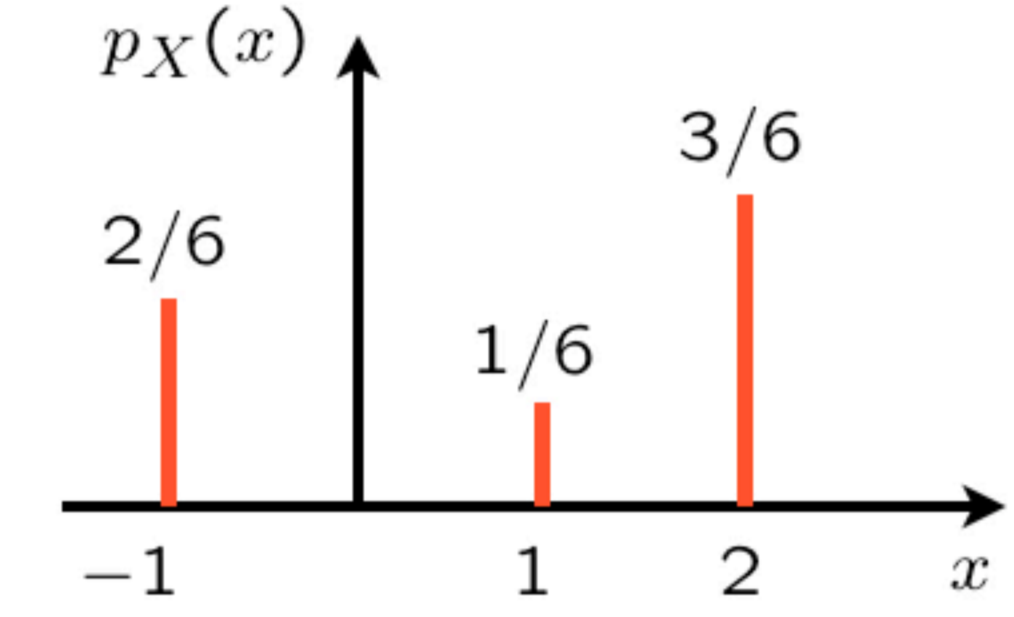
$Y = g(X)$



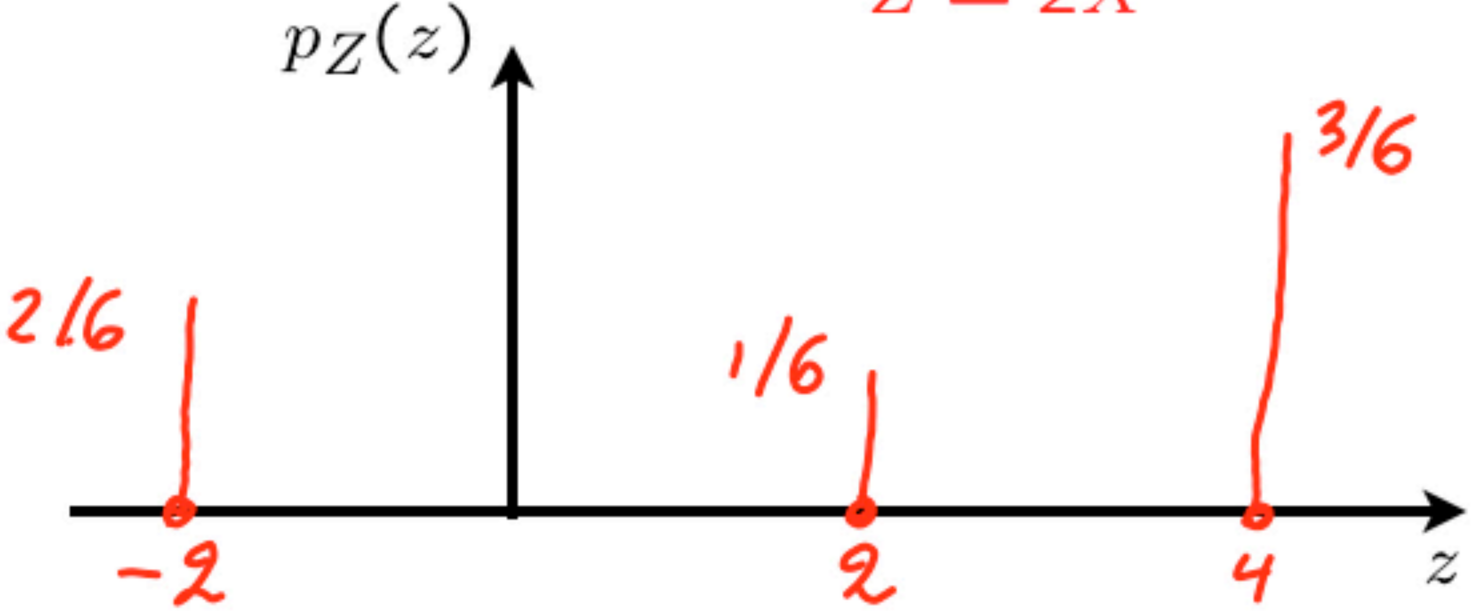
$$p_Y(4) = \mathbf{P}(Y = 4)$$
$$= \mathbf{P}(X = 4) + \mathbf{P}(X = 5)$$
$$= p_X(4) + p_X(5) = 0.3 + 0.4$$

$$p_Y(y) = \mathbf{P}(g(X) = y)$$
$$= \sum_{x:\, g(x) = y} p_X(x)$$

A linear function of a discrete r.v.

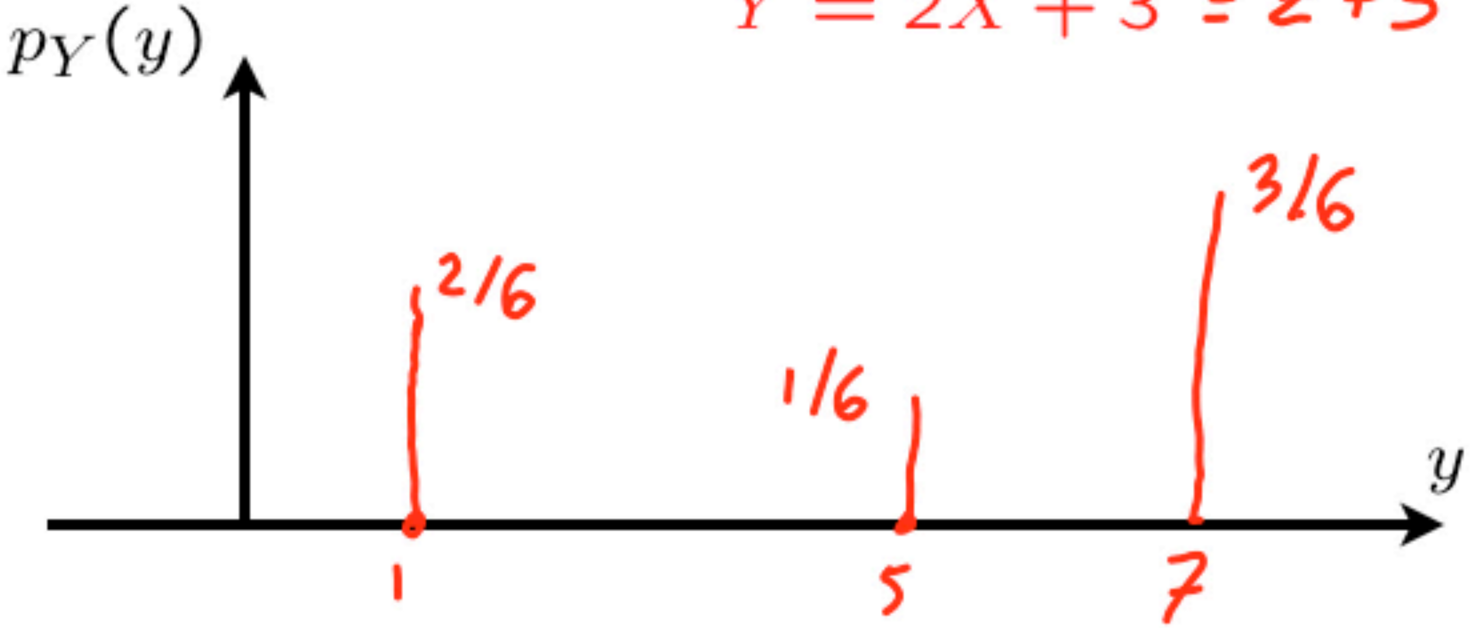


$Z = 2X$



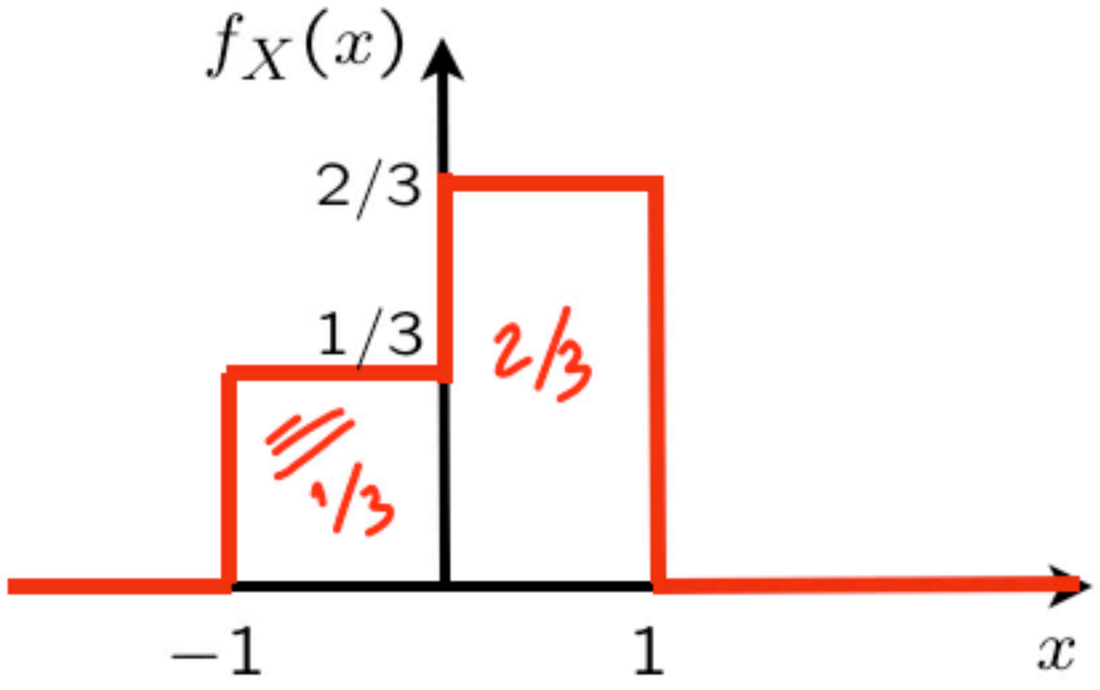
$$p_Y(y) = P(Y=y) = P(2X+3=y)$$
$$= P\left(X=\frac{y-3}{2}\right) = p_X\left(\frac{y-3}{2}\right)$$

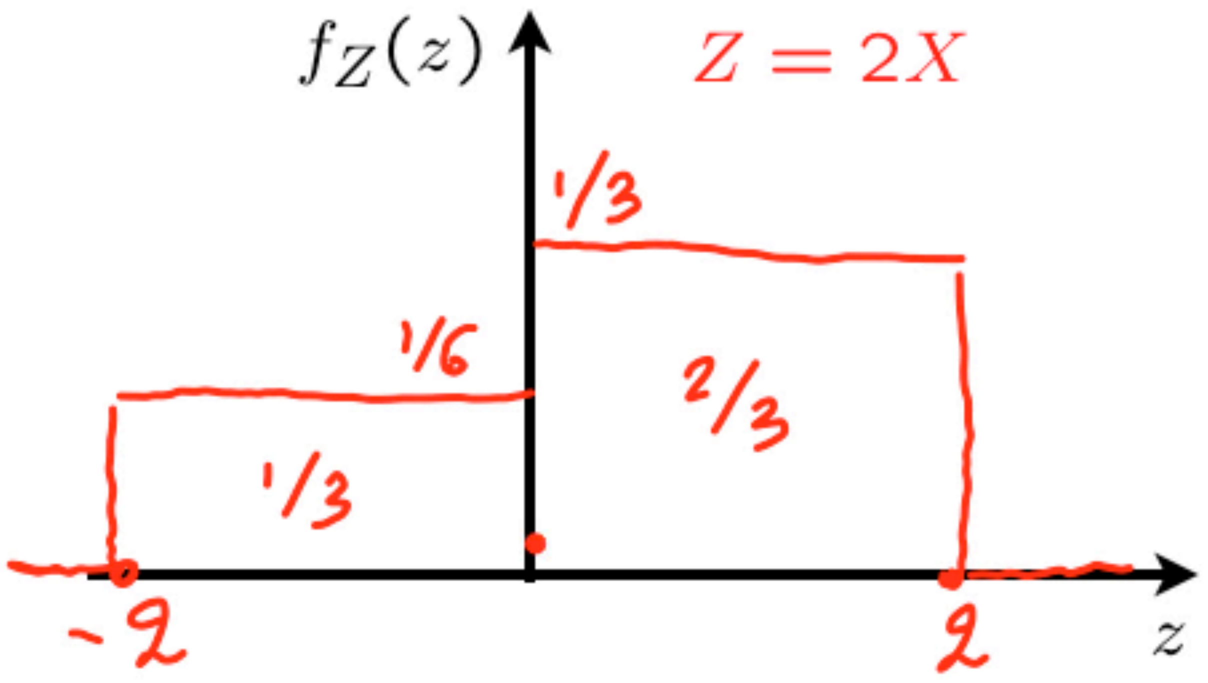
$Y = 2X + 3 = Z+3$

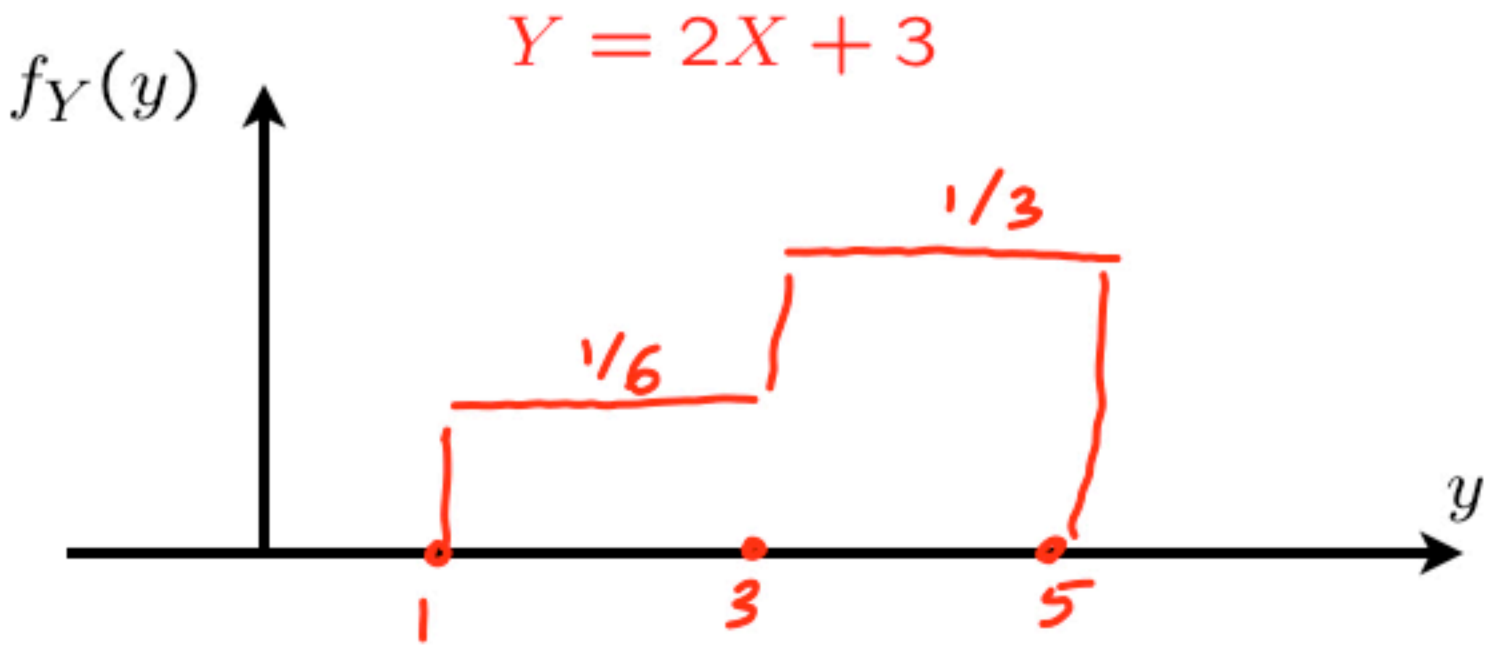


$$\boxed{Y = aX + b : \quad p_Y(y) = p_X\left(\frac{y-b}{a}\right)}$$

A linear function of a continuous r.v.







A linear function of a continuous r.v.

$Y = aX + b$

$a > 0$

$P(Y=y) = P(aX+b=y) = P\left(X = \frac{y-b}{a}\right)$

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

$= P\left(X \le \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right)$

$f_Y(y) = f_X\left(\frac{y-b}{a}\right) \cdot \frac{1}{a}$

$a < 0$

$= P\left(X \ge \frac{y-b}{a}\right)$

$= 1 - P\left(X \le \frac{y-b}{a}\right)$

$= 1 - F_X\left(\frac{y-b}{a}\right)$

$f_Y(y) = -f_X\left(\frac{y-b}{a}\right) \cdot \frac{1}{a}$

$$\boxed{f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)}$$

$$\boxed{p_Y(y) = p_X\left(\frac{y-b}{a}\right)}$$

A linear function of a normal r.v. is normal

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

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-(x-\mu)^2/2\sigma^2}$$

$Y = aX + b, \quad a \neq 0$

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$

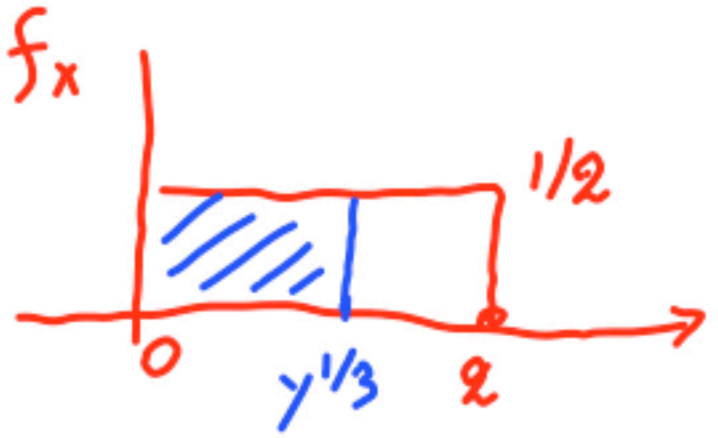
$$f_Y(y) = \frac{1}{|a|} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\left(\frac{y-b}{a}-\mu\right)^2/2\sigma^2}$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma |a|} e^{-(y-\underbrace{b-a\mu}{})^2/2\underbrace{\sigma^2 a^2}{}}$$

A general function $g(X)$ of a continuous r.v.

- **Two-step procedure:**
  - Find the CDF of Y: $F_Y(y) = \mathbf{P}(Y \le y)$ $= P(g(x) \le y)$
  - Differentiate: $f_Y(y) = \dfrac{dF_Y}{dy}(y)$

Example: $Y = X^3$; X **uniform on** $[0, 2]$



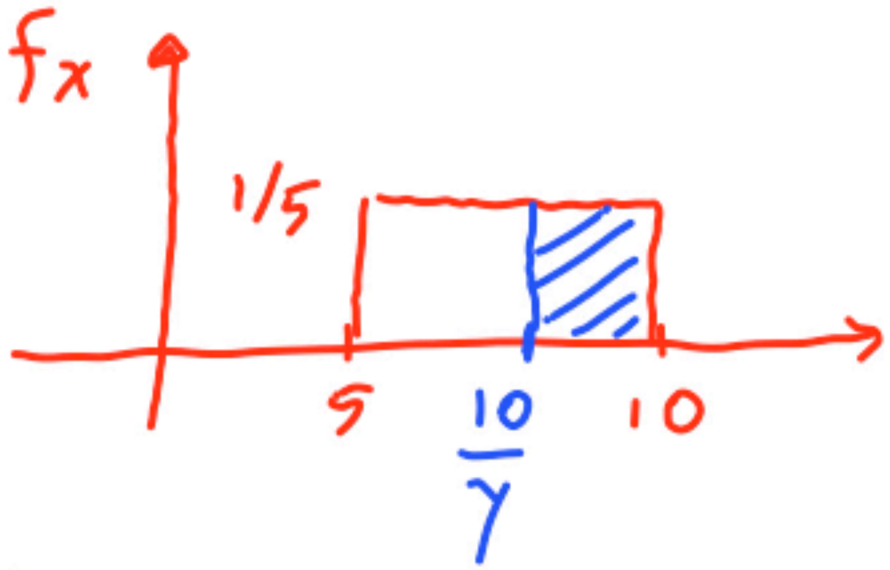
$0 \le y \le 8$

$$F_Y(y) = P(Y \le y) = P(X^3 \le y) = P(X \le y^{1/3}) = \frac{1}{2} y^{1/3}$$

$$f_Y(y) = \frac{dF_Y}{dy}(y) = \frac{1}{2} \cdot \frac{1}{3} y^{-2/3} = \frac{1}{6} \cdot \frac{1}{y^{2/3}}$$

Example: $Y = a/X$

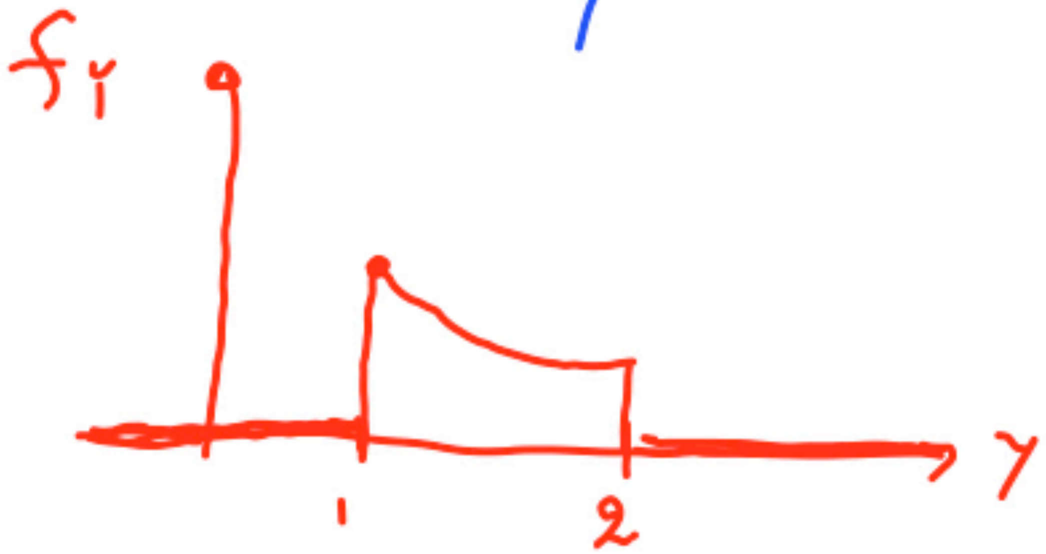
- You go to the gym and set the speed X of the treadmill to a number between 5 and 10 km/hr (with a uniform distribution). Find the PDF of the time it takes to run 10km.



$$\text{time} = Y = \frac{10}{X} \qquad 1 \le y \le 2$$

$$F_Y(y) = P(Y \le y) = P\left(\frac{10}{X} \le y\right)$$

$$= P\left(X \ge \frac{10}{y}\right) = \frac{1}{5}\left(10 - \frac{10}{y}\right)$$



$$f_Y(y) = \frac{1}{5}\,\frac{(-10)}{-y^2} = \frac{2}{y^2}, \quad 1 \le y \le 2$$

$$= 0, \quad \text{otherwise}$$

A general formula for the PDF of $Y = g(X)$ when g is monotonic

$x^3 \quad \frac{a}{x}$

decreasing $x < x' \Rightarrow g(x) < g(x')$

Assume g strictly ~~increasing~~

and differentiable

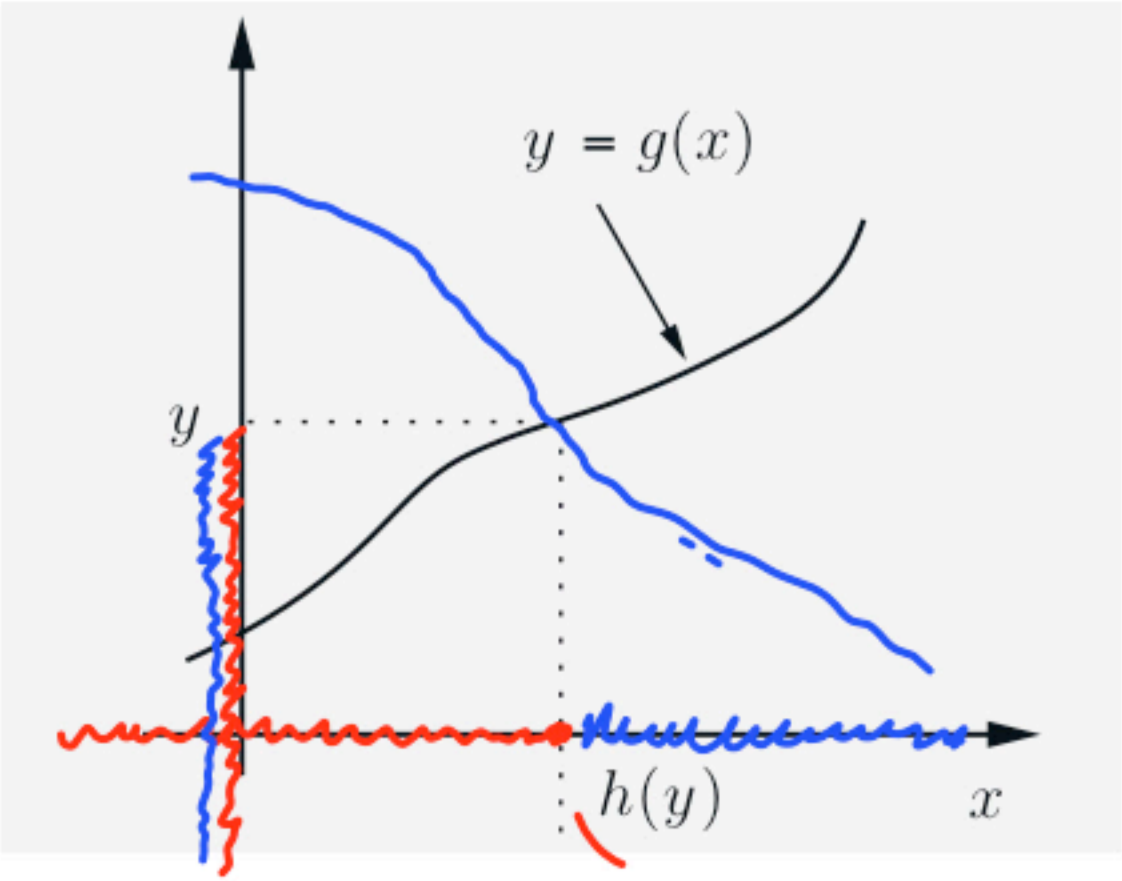
$$F_Y(y) = P(Y \le y) = P(X \le h(y)) = F_X(h(y))$$

$$f_Y(y) = f_X(h(y)) \left| \frac{dh}{dy}(y) \right|$$

$$F_Y(y) = P(Y \le y) = P(X \ge h(y))$$

$$= 1 - P(X \le h(y)) = 1 - F_X(h(y))$$

$$f_Y(y) = -f_X(h(y)) \left| \frac{dh}{dy}(y) \right|$$



inverse function h → decreasing

$$\boxed{f_Y(y) = f_X\big(h(y)\big) \left| \frac{dh}{dy}(y) \right|}$$

Example: $Y = X^2$; X uniform on $[0,1]$

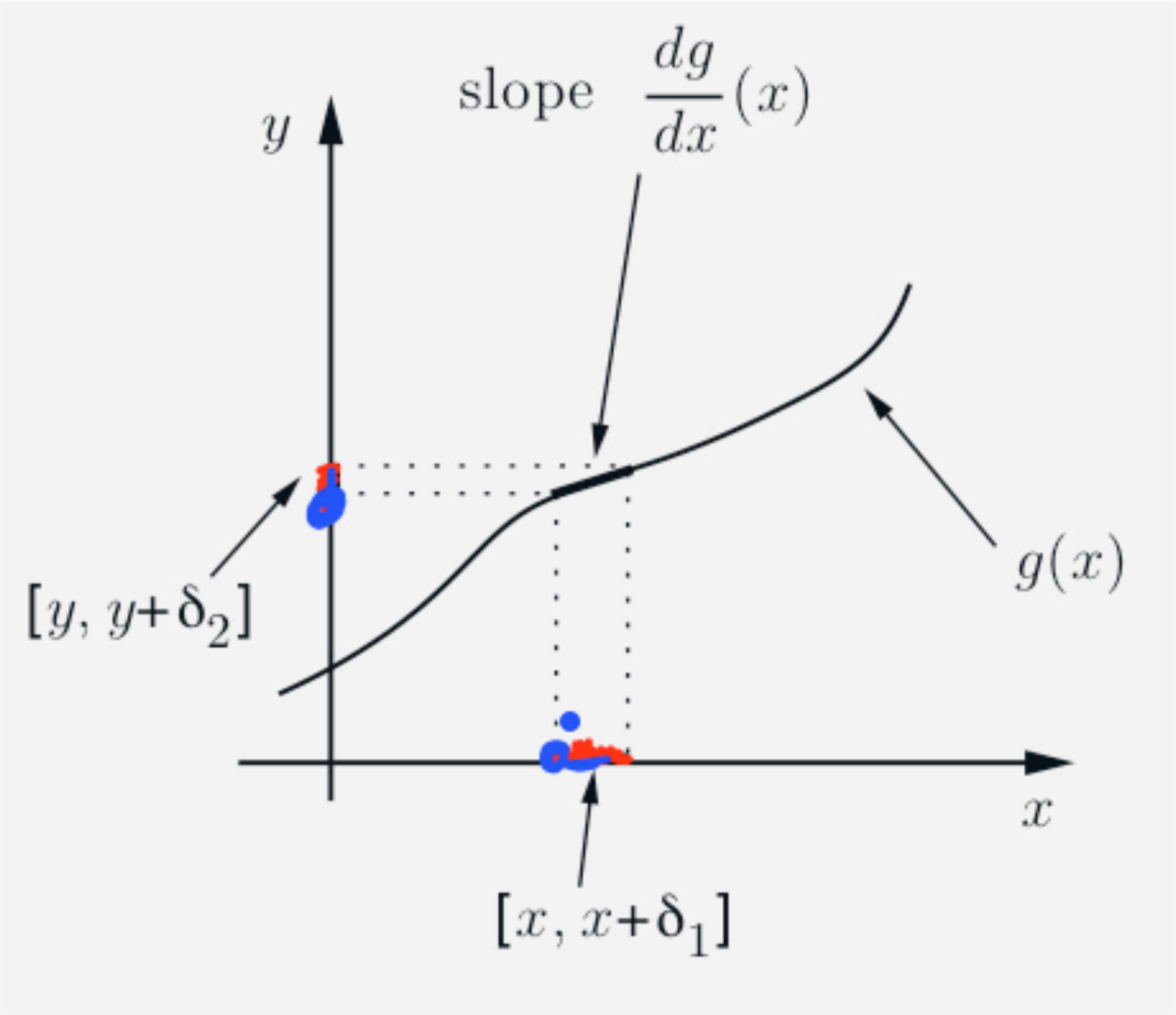
$$\boxed{f_Y(y) = f_X\big(h(y)\big)\left|\frac{dh}{dy}(y)\right|}$$

$y = x^2 \Leftrightarrow x = \sqrt{y}$ $\quad h(y) = \sqrt{y}$

$f_Y(y) = \dfrac{1}{2\sqrt{y}}$

$0 \leq y \leq 1$

An intuitive explanation for the monotonic case



$$y = g(x) \qquad \delta_2 \approx \delta_1 \frac{dg}{dx}(x)$$

$$x = h(y) \qquad \delta_1 \approx \delta_2 \cdot \frac{dh}{dy}(y) \quad (*)$$

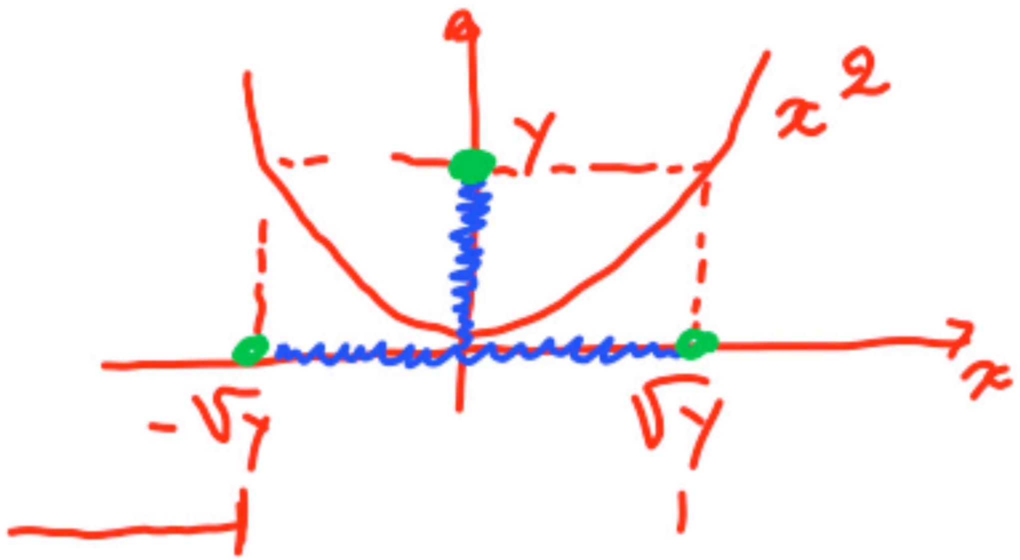
$$f_Y(y)\cancel{\delta_2} \approx P(y \le Y \le y+\delta_2) = P(x \le X + x+\delta_1)$$

$$\approx f_X(x)\,\delta_1 \approx f_X(x)\,\cancel{\delta_2}\,\frac{dh}{dy}(y)$$

$$f_Y(y) = f_X(x)\,\frac{dh}{dy}(y)$$

$$= f_X(h(y))\,\frac{dh}{dy}(y)$$

A nonmonotonic example: $Y = X^2$



- The discrete case:

$$p_Y(9) = \mathbf{P}(X=3) + \mathbf{P}(X=-3)$$

$$p_Y(y) = p_X(\sqrt{y}) + p_X(-\sqrt{y})$$

- The continuous case: $y \geq 0$

$$F_Y(y) = \mathbf{P}(Y \leq y) = \mathbf{P}(X^2 \leq y) = \mathbf{P}(|X| \leq \sqrt{y}) = \mathbf{P}(-\sqrt{y} \leq X \leq \sqrt{y})$$

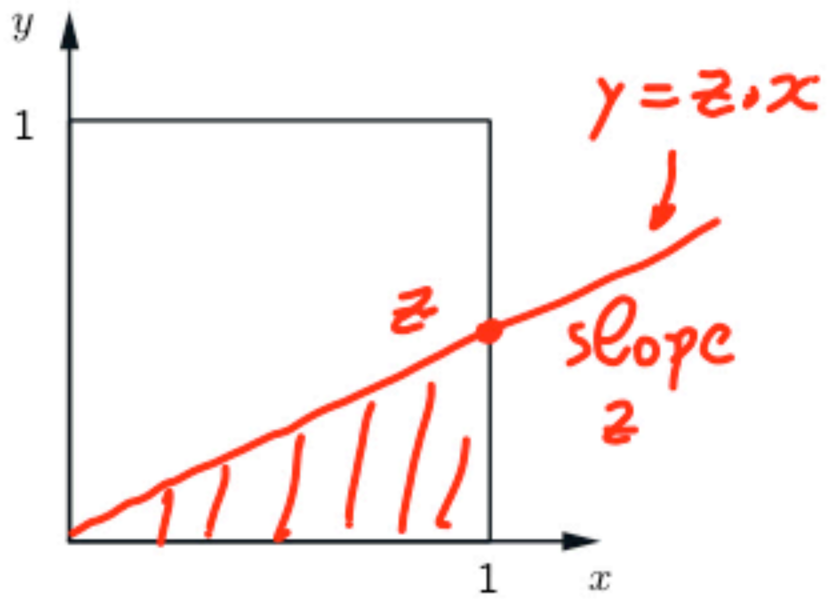
$$= F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

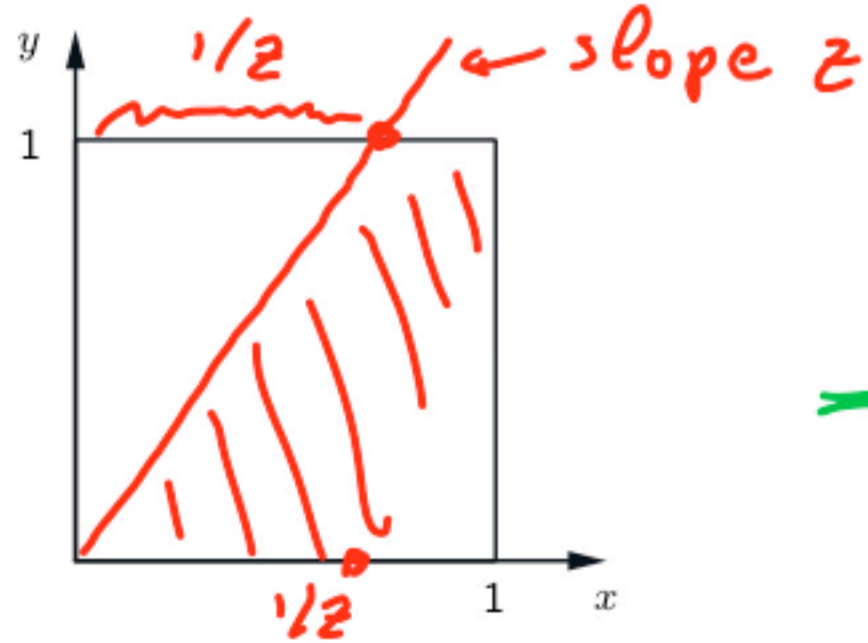
$$f_Y(y) = f_X(\sqrt{y})\frac{1}{2\sqrt{y}} + f_X(-\sqrt{y})\frac{\cancel{(-1)}}{2\sqrt{y}}$$

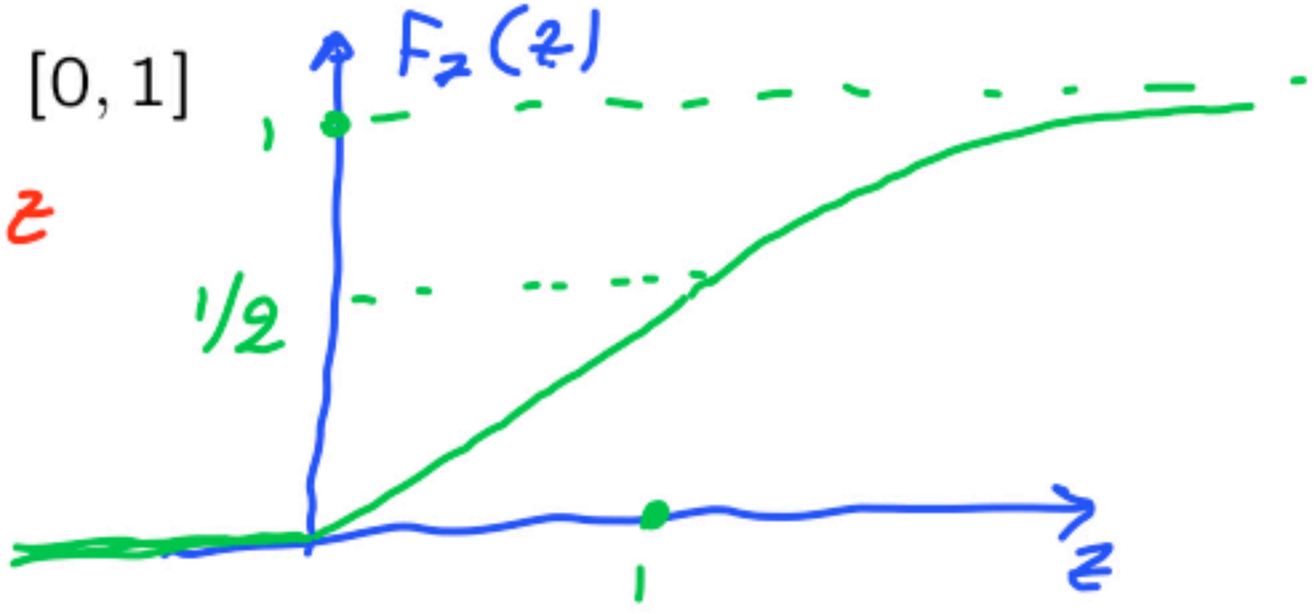
A function of multiple r.v.'s: $Z = g(X, Y)$

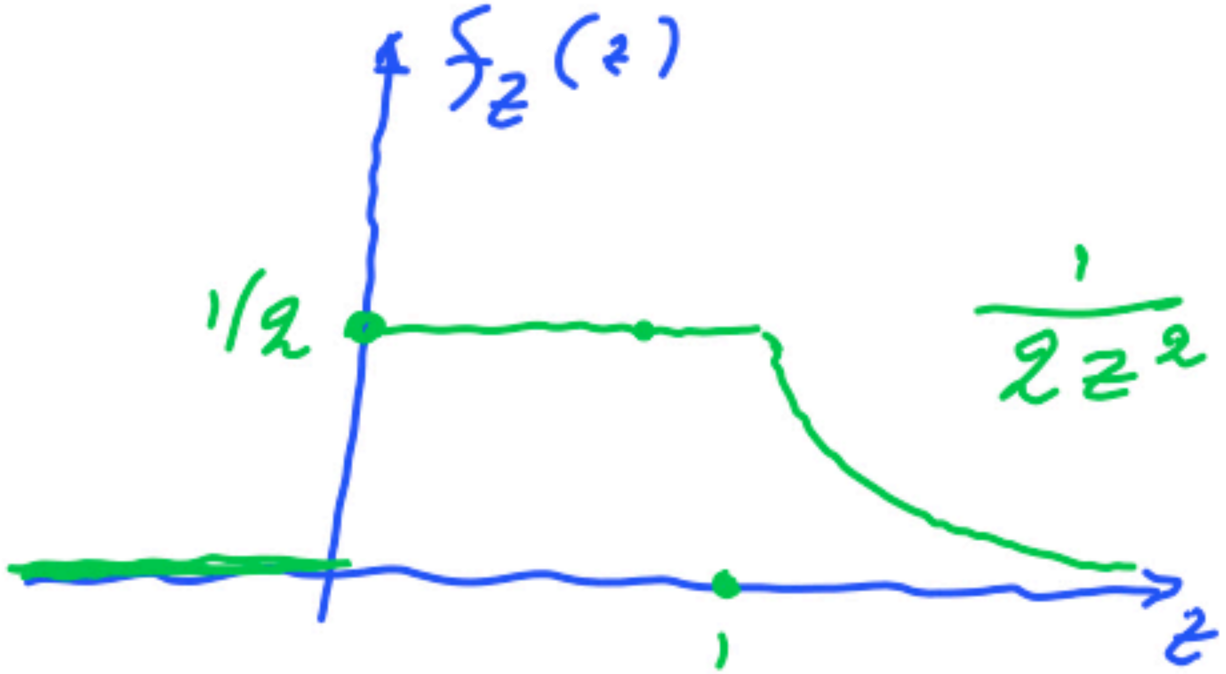
$Y = g(X)$

- Same methodology: find CDF of Z
- Let $Z = Y/X$; X, Y independent, uniform on $[0, 1]$









$$F_Z(z) = P\left(\frac{Y}{X} \le z\right) = 0, \quad z < 0$$

$$= \frac{1}{2} \cdot z, \quad 0 \le z \le 1$$

$$= 1 - \frac{1}{2z}, \quad z > 1$$