2.1 Functions of random variables

Recall

- a Random variable X is defined as a $\boxed{\text{function}}$ from Sample space S to $\mathcal{X}$

$$X = X(s), \quad s \in S.$$

- information about the Random X is (completely) captured by its cdf

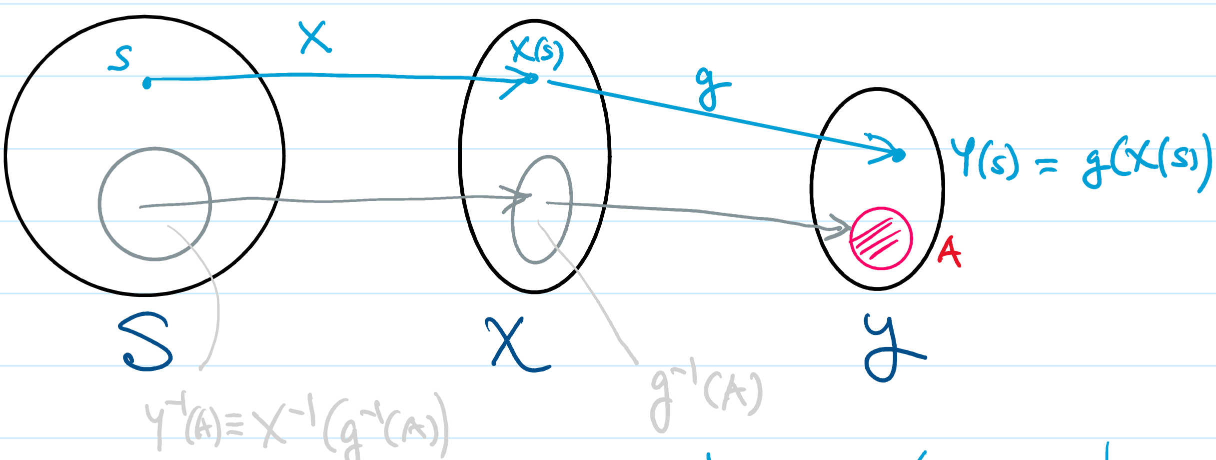
$$F_X(x) = P(X \leq x)$$
$$= P(s \in S \mid X(s) \leq x).$$

Take a function $g : \mathcal{X} \to \mathcal{Y}$, then

$$Y = g(X)$$

is Also a Random variable taking values in $\mathcal{Y}$, because Y is in fact a function on S

$$Y(s) = g(X(s)) \equiv g \circ X(s)$$
$$\not\equiv X \circ g$$



$$g^{-1}(A) := \{x \in \mathcal{X} \mid g(x) \in A\}$$

What is the probability distribution of $Y = g(X)$?

By def, $\forall A \subset \mathcal{Y}$

$$\boxed{\begin{aligned} P(Y \in A) &= P(g(X) \in A) \\ &:= P(X \in g^{-1}(A)) = P(X(s) \in g^{-1}(A)) \\ &= P(s \in X^{-1}(g^{-1}(A))) = P(s \in X^{-1} \circ g^{-1}(A)) \\ &= P(s \in (g \circ X)^{-1}(A)) \\ &= P(s \in Y^{-1}(A)) \\ &= P(Y(s) \in A) \end{aligned}}$$

What is nice is that we can "create" new random variables by applying a function to an existing random variable, instead of specifying the probability distribution from scratch, i.e., via a sample space and sigma algebra

EXAMPLE (BINOMIAL TRANSFORMATION)

- X is a binomial R.V. if its pmf

$$f_X(x) = P(X=x) = \binom{n}{x} p^x (1-p)^{n-x}$$

for $x = 0, 1, \ldots, n$.

We write

$$X \sim \text{Binomial}(n, p)$$

parameters

- Let $Y := g(X)$ where $g(x) = n - x$

i.e.

$Y = n - X$ is also a R.V.

But what is Y's distribution?

Clearly $Y \in \{0, 1, \ldots, n\}$

For $y \in \{0, 1, \ldots, n\}$

$$\begin{aligned} f_Y(y) &= P(Y = y) \qquad Y = n - X \\ &= P(X = n - y) \\ &= f_X(n-y) = \binom{n}{n-y} p^{n-y} (1-p)^y \\ &= \binom{n}{y} (1-p)^y p^{n-y} \end{aligned}$$

So $Y \sim \text{Binomial}(n, 1-p)$

CONTINUOUS RANDOM VARIABLES

if X a cont. RV, g a nice (cont.) function then $Y = g(X)$ is a cont. R.V.

$$
\begin{aligned}
F_Y(y) &= P(Y \leq y) \\
&= P(g(X) \leq y) \\
&= P(x \in \mathcal{X} : g(x) \leq y) \\
&= \int_{\{x:\ g(x) \leq y\}} f_X(x)\, dx
\end{aligned}
$$

The set $\{x : g(x) \leq y\}$ may be difficult to identify so F_y may be hard to derive in general.

if g is (strictly) Monotone, this gets simpler

either

- g is increasing i.e. $g(x) > g(y)$ if $x > y$
- g is decreasing i.e. $g(x) < g(y)$ if $x > y$.

Note

$g^{-1}(y) = \{x : g(x) = y\}$ is a singleton set

(a) if $g \uparrow$ then

taken to be its element

$$\{x : g(x) \leq y\} = \{x \leq g^{-1}(y)\}$$

So, $$F_Y(y) = \int_{\{x \leq g^{-1}(y)\}} f_X(x)\, dx$$

$$= F_X(g^{-1}(y))$$

(b) if $g \downarrow$ then

$$\{x : g(x) \leq y\} = \{x \geq g^{-1}(y)\}$$

So, $$F_Y(y) = \int_{\{x \geq g^{-1}(y)\}} f_X(x)\, dx$$

$$= 1 - \int_{\{x \leq g^{-1}(y)\}} f_X(x)\, dx$$

$$= 1 - F_X(g^{-1}(y))$$

We have proved

THM Let X have the cdf $F_X(x)$

$$\mathcal{X} = \{x : f_X(x) > 0\}$$

$$\mathcal{Y} = g(\mathcal{X}) := \{y : y = f(x) \text{ for some } x \in \mathcal{X}\}$$

(a) if $g : \mathcal{X} \to \mathcal{Y}$ is increasing, then $Y = g(X)$ is a R.v. taking values in $\mathcal{Y}$ with the cdf

$$F_Y(y) = F_X(g^{-1}(y)) \quad \forall y \in \mathcal{Y}$$

(b) if $g : \mathcal{X} \to \mathcal{Y}$ is decreasing, then $Y = g(X)$ is a R.V. in $\mathcal{Y}$ with cdf

$$F_Y(y) = 1 - F_X(g^{-1}(y)) \quad \forall y \in \mathcal{Y}$$

We can deduce the pdf from cdf

From part (a), $$f_Y(y) = \frac{d}{dy} F_Y(y)$$

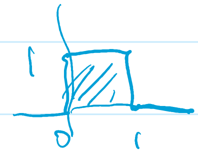
$$= f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y)$$

From (b), $$f_Y(y) = -f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y)$$

Combining the two cases to obtain

$$\boxed{f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|}$$

EXAMPLE : UNIFORM - EXPONENTIAL



Let $X \sim \text{Uniform}(0,1)$ i.e., $f_X(x) = \begin{cases} 1 & \text{if } x \in (0,1) \\ 0 & \text{otherwise} \end{cases}$

Take $Y = -\log X$.

what is the distribution of Y?

$\begin{cases} \mathcal{X} = (0,1) \\ \mathcal{Y} = (0,+\infty) \end{cases}$

$Y = g(X)$ where $g(x) = -\log x$

$g'(x) = -\frac{1}{x} < 0 \Rightarrow g \downarrow$ and $g^{-1}(y) = e^{-y} \quad \forall y > 0$

if $x \in (0,1)$ then $F_X(x) = \int_0^x f_X(x)\,dx$

$= x$

Thus, for $y > 0$

$$\begin{aligned} F_Y(y) &= 1 - F_X(g^{-1}(y)) \\ &= 1 - g^{-1}(y) \\ &= 1 - e^{-y} \end{aligned}$$

Hence $f_Y(y) = \frac{dF_Y(y)}{dy} = e^{-y}$ for $y > 0$

For $y \leq 0$, $F_Y(y) = 0$, so $f_Y(y) = 0$ as well.

We say $Y \sim \text{Exp}(1)$.

We say $Y \sim \text{Exp}(1)$.

EXAMPLE - INVERSE GAMMA

Let $X \sim \text{Gamma}(n, \beta)$ i.e.

$$f_X(x) = \frac{1}{(n-1)!\,\beta^n} x^{n-1} e^{-x/\beta} \quad \text{for } x > 0.$$

Let $Y = 1/X$, named inverse gamma R.V. what is the distribution of Y?

Let $g(x) = 1/x$ so $g^{-1}(y) = 1/y$ for $y \in (0,\infty)$
$g \downarrow$ on $\mathcal{X} = (0,\infty)$.

$\mathcal{Y}$

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|$$

$$= \frac{1}{(n-1)!\,\beta^n} (1/y)^{n-1} e^{-1/(\beta y)} \frac{1}{y^2}$$

$$= \frac{1}{(n-1)!\,\beta^n} y^{-(n+1)} e^{-1/(\beta y)}$$

EXAMPLE – SQUARE TRANSFORMATION

"Squaring up a standard Gaussian variable X gives the chi squared variable $Y = X^2$"

Let $g(x) = x^2$ for $x \in \mathbb{R}$

g is not a monotone function

$\mathcal{X} = \mathbb{R}, \quad \mathcal{Y} = \mathbb{R}_+ \cup \{0\}.$

$\forall y \geq 0$

$$
\begin{aligned}
F_Y(y) &= P(Y \leq y) \\
&= P(X^2 \leq y) \\
&= P(|X| \leq \sqrt{y})
\end{aligned}
$$

X is cont. R.V.

Hence
$$
\begin{aligned}
F_Y(y) &= P(X \leq \sqrt{y}) - P(X < -\sqrt{y}) \\
&= P(X \leq \sqrt{y}) - P(X \leq -\sqrt{y}) \\
&= F_X(\sqrt{y}) - F_X(-\sqrt{y})
\end{aligned}
$$

So
$$
\begin{aligned}
f_Y(y) &= \frac{d}{dy} F_Y(y) \\
&= \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y})
\end{aligned}
$$

For standard Gaussian $X \sim N(0,1)$,

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad x \in \mathbb{R}$$

Then

$$f_Y(y) = \frac{1}{2\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-y/2} + \frac{1}{2\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-y/2}$$

$$= \frac{1}{\sqrt{2\pi y}} e^{-y/2}$$

← pdf for Chi-squared R.V. with 1 degree of freedom

Probability Integral Transformation

THM 1 Let X be a continuous R.V. with cdf $F_X(x)$.
Let $Y = F_X(X)$.
Then
$$Y \sim \text{Uniform}(0,1)$$

Remarks

- THM 1 does NOT hold for discrete R.V. ⟸ since F_X would be step function.
- F_X might not be strictly increasing from 0 to 1.
- Define the inverse function of F_X as

$$F_X^{-1}(y) := \inf\{x : F_X(x) \geq y\} \quad \forall y \in (0,1)$$

in addition

$$F_X^{-1}(1) := +\infty, \quad F_X^{-1}(0) := -\infty.$$

THM 2 Let $F_X(x)$ be a cdf of a random variable.
Let $Y \sim \text{Uniform}(0,1)$
and set
$$Z = F_X^{-1}(Y)$$
Then Z has cdf F_X.

Remark
- Useful for generating R.V.'s from Uniform RV's.

Remark
- Useful for generating R.V.'s from Uniform RV's.
- In THM 2, there is NO restriction to continuity, it holds generally (for real-valued R.V.'s)

Proof. (THM 1)

For $y \in (0,1)$

$$\begin{aligned} P(Y \le y) &= P(F_X(X) \le y) \\ F_X^{-1} \uparrow \longrightarrow &= P(F_X^{-1}(F_X(X)) \le F_X^{-1}(y)) \\ &= P(X \le F_X^{-1}(y)) \leftarrow \text{see } (*) \\ &= F_X(F_X^{-1}(y)) \\ F_X \text{ cont} \longrightarrow &= y \end{aligned}$$

It is easy to check: $P(Y \le 1) = 1$ and $P(Y \le 0) = 0$.
Hence cdf of Y is that of Unif $(0,1)$ R.V. $\square$

(*):

- if F_X is strictly increasing at x, for $x \in \mathcal{X}$
  then let $y = F_X(x) \Rightarrow x = F_X^{-1}(y)$
  so $F_X^{-1}(F_X(x)) = F_X^{-1}(y) = x$.

- if F_X is "flat" at x, i.e. $\forall x \in$ some $[x_1, x_2]$
  such than $F_X(x') = F_X(x_1)$ $\forall x' \in [x_1, x]$
  and $F_X(x) < F_X(x_1)$ $\forall x < x_1$
  then
  $$F_X^{-1}(F_X(x)) = x_1 \quad (\leq x)$$
  Since
  $$P(X \in (x_1, x]) = 0$$
  $$P(X \in (F_X^{-1}(F_X(x)), x]) = 0 \qquad (**)$$

So in this scenario

$\forall a \in \mathbb{R}$

$$P(F_X^{-1}(F_X(X)) \leq a \mid F_X \text{ is flat at } X)$$
$$= P(F_X^{-1}(F_X(X)) \leq a, X \leq a \mid \dots) +$$
$$P(F_X^{-1}(F_X(X)) \leq a, X > a \mid \dots)$$

$$=: C + D \quad , \text{ where}$$

$$C = P(X \leq a \mid F_X \text{ is flat at } X)$$
because $X \leq a \Rightarrow F_X^{-1}(F_X(X)) \leq X \leq a$

$$0 \leq P(F_X^{-1}(F_X(X)) < X \mid F_X \text{ is flat at } X)$$

with prob. 1. $\longrightarrow$ $= 0$, due to $(**)$

- Combining the two scenarios

$$\forall a,\ P(F_X^{-1}(F_X(X)) \leq a \mid F_X \text{ flat at } X) = P(X \leq a \mid F_X \text{ flat at } X)$$

$$P(F_X^{-1}(F_X(X)) \leq a \mid F_X \uparrow \text{ at } X) = P(X \leq a \mid F_X \uparrow \text{ at } X)$$

with prob. 1.

to arrive at $(*)$ $\square$