Begin with X, transform to $Y = g(X)$.
How to get dist. of Y from dist. of X.

STAT 712 fa 2022 Lec 7 slides

Transformations of a random variable

Karl B. Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

Let X be a rv with support $\boxed{\mathcal{X}}$ and let $Y = g(X)$ for some $g : \mathcal{X} \to \mathcal{Y}$.

support of $\mathcal{Y}$

Examples:

- X = diameter of sand dollar. $Y = \pi \cdot (X/2)^2$, surface area of underside.
- X = daily rate of return. $Y = A(1 + X)$, value of investment A after a day.
- X = white blood cells per μl. $Y = \mathbf{1}(X > 11 \times 10^3)$, whether above normal.

How to find the distribution of Y from the distribution of X?

Inverse mapping of a function

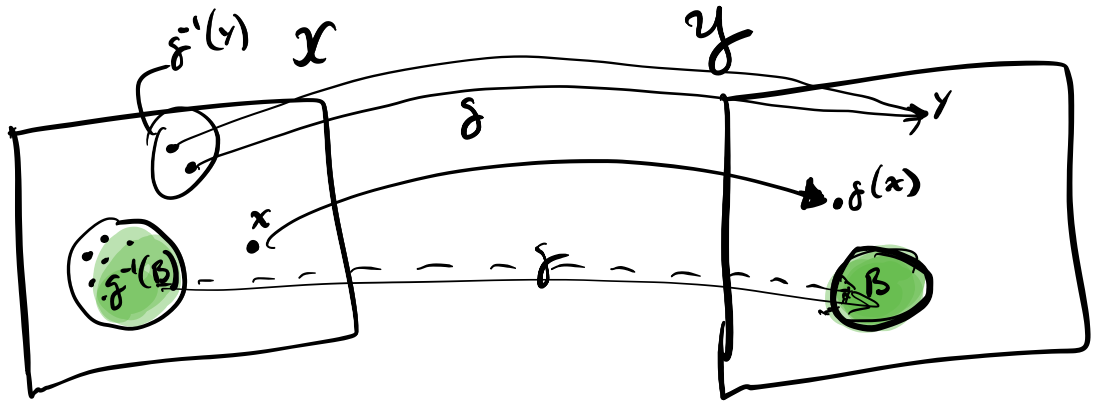
For any function $g : \mathcal{X} \to \mathcal{Y}$, define the mapping

$$\boxed{g^{-1}(B)} = \{x \in \mathcal{X} : g(x) \in B\} \quad \text{for all } B \subset \mathcal{Y}.$$

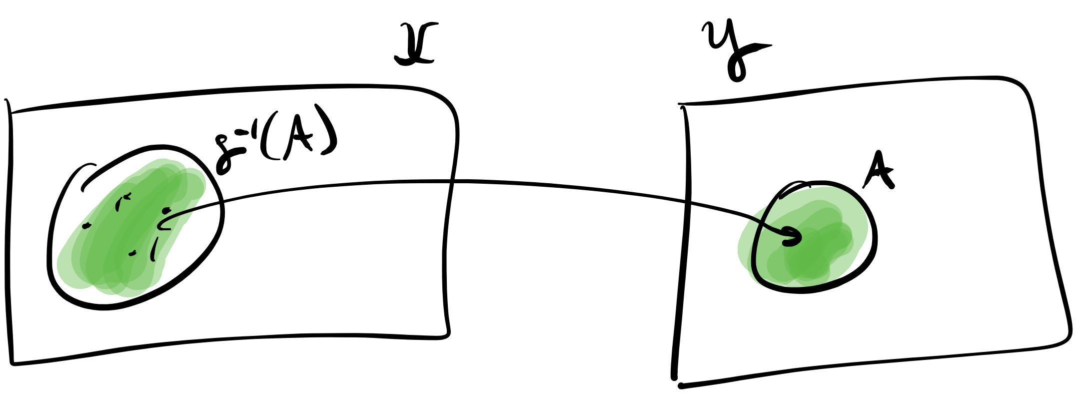
We call $g^{-1}(B)$ the *inverse image of B under g*.

For a single point $\{y\} \in \mathcal{Y}$, write $g^{-1}(y) := \boxed{g^{-1}(\{y\})} = \{x \in \mathcal{X} : g(x) = y\}$.

So g^{-1} takes subsets of $\mathcal{Y}$ and returns subsets of $\mathcal{X}$.



$f^{-1}(y)$
$\mathcal{X}$
$\mathcal{Y}$
y
f
x
$f(x)$
$f^{-1}(B)$
f
B



$\mathcal{X}$
$\mathcal{Y}$
$f^{-1}(A)$
A

$Y = g(X)$

Distribution of transformed random variable

To find the probability distribution of Y, we note that for any $A \in \mathcal{B}(\mathbb{R})$ we have

$$P(Y \in A) = P(g(X) \in A) = P(X \in g^{-1}(A))$$

Exercise: Let X take the values in $\mathcal{X} = \{-2, -1, 0, 1, 2\}$ with the probabilities

x	-2	-1	0	1	2
$P(X = x)$	$1/9$	$2/9$	$3/9$	$2/9$	$1/9$

Tabulate the probability distribution of $\boxed{Y = |X|}$.

$Y = g(X) = |X|$

① Identify $\mathcal{Y}$. $\mathcal{Y} = \{0, 1, 2\}$

y	$g^{-1}(y)$
0	$\{0\}$
1	$\{-1, 1\}$
2	$\{-2, 2\}$

$$P(Y=0) = P(g(X)=0) = P(X \in g^{-1}(0)) = P(X \in \{0\})$$
$$= 3/9$$

$$P(Y=1) = P(g(X)=1) = P(X \in g^{-1}(1)) = P(X \in \{-1,1\})$$
$$= \frac{4}{9}$$

$$P(Y=2) = P(g(X)=2) = P(X \in g^{-1}(2)) = P(X \in \{-2,2\})$$
$$= \frac{2}{9}$$

y	0	1	2
$P(Y=y)$	$3/9$	$\frac{4}{9}$	$\frac{2}{9}$

Finding the pmf of a transformed discrete rv

If X is discrete with pmf p_X then $Y = g(X)$ is discrete with pmf

$$p_Y(y) = P(X \in g^{-1}(y)) = \sum_{x \in g^{-1}(y)} p_X(x).$$

$= P(Y = y)$ "

Any transformation of a discrete rv will result in a discrete rv.

Strategy

1. Find the support $\mathcal{Y}$ of Y.
2. Identify the inverse function $g^{-1}(y)$ for each $y \in \mathcal{Y}$.

Exercise:

1. Let $X \sim \text{Poisson}(\lambda)$. Find the pmf of $Y = \mathbf{1}(X > 10)$. $= \begin{cases} 1 & \text{if } X > 10 \\ 0 & \text{if } X \le 10 \end{cases}$
2. Let $X \sim \text{Binomial}(n, p)$. Find the pmf of $Y = n - X$.

① $\mathcal{Y} = \{0, 1\}$

y	$g^{-1}(y)$
0	$\{x \in \mathcal{X} : g(x) = 0\} = \{0, 1, \ldots, 10\}$
1	$\{x \in \mathcal{X} : g(x) = 1\} = \underline{\{11, 12, \ldots\}}$

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

$$p_Y(y) = P(Y = y) = \begin{cases} P(X \in \{0, 1, \ldots, 10\}) & y = 0 \\ P(X \in \{11, 12, \ldots\}) & \underline{y = 1} \end{cases}$$

$$= \begin{cases} \sum_{x=0}^{10} \frac{e^{-\lambda} \lambda^x}{x!} & y = 0 \\ \sum_{x=11}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} & y = 1 \end{cases}$$

$$p_X(x) = \frac{e^{-\lambda} \lambda^x}{x!} \mathbb{1}(x \in \{0, 1, 2, \ldots\})$$

$$= \left(\sum_{x=11}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} \right)^{y} \left(1 - \sum_{x=11}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} \right)^{1-y}$$

② $X \sim p_X(x) = \binom{n}{x} p^x (1-p)^{n-x} \mathbb{1}(x \in \{0,1,\dots,n\})$

$\mathcal{X} = \{0, 1, \dots, n\}$

$\boxed{Y = g(X) = n - X}$ $\quad \mathcal{Y} = \{0, 1, \dots, n\}.$

y	$g^{-1}(y)$
0	n
1	$n-1$
$\vdots$	
n	0

$y = g(x) = n - x$

$\iff x = n - y = g^{-1}(y)$

$$p_Y(y) = P(Y = y) = P(n - X = y)$$
$$= P(X = n - y)$$

$$= P(g(X) = y)$$
$$= P(X \in g^{-1}(y))$$
$$= P(X = n - y)$$
$$= p_X(n-y)$$
$$= \underbrace{\binom{n}{n-y}}_{\shortparallel} p^{n-y} (1-p)^{n-(n-y)}$$
$$= \binom{n}{y} (1-p)^y p^{n-y}$$

pmf of Binomial $(n, 1-p)$

Finding the pdf of a transformed continuous rv by the *cdf method*

Let X be a continuous rv with pdf f_X. Then the cdf of $Y = g(X)$ is given by

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = \int_{\{x:g(x)\leq y\}} f_X(x)dx$$

Then $f_Y(y) = \frac{d}{dy}F_Y(y)$, provided this is defined over the support of Y.

Exercises:

1. Let $X \sim f_X(x) = (3/2)x^2\mathbf{1}(-1 \leq x \leq 1)$. Find pdf of $Y = |X|$.
2. Let $X \sim \text{Uniform}(0,1)$. Find pdf of $Y = -\log X$.

① $Y = g(X) = |X|$. $\mathcal{X} = [-1, 1]$ $\mathcal{Y} = [0, 1]$

For $y \in [0,1]$

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

$$= \int_{-y}^{y} \tfrac{3}{2} x^2 \, dx$$

$$= \int_{\{x : |x| \le y\} = [-y, y]} \frac{3}{2} x^2 \, dx$$

$$= \int_{-y}^{y} \frac{3}{2} x^2 \, dx$$

$$= \left. \frac{x^3}{2} \right|_{-y}^{y}$$

$$= \frac{y^3}{2} - \frac{(-y)^3}{2}$$

$$= y^3$$

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ \boxed{y^3 \qquad 0 \le y \le 1} \\ 1 & 1 \le y \end{cases}$$

Then

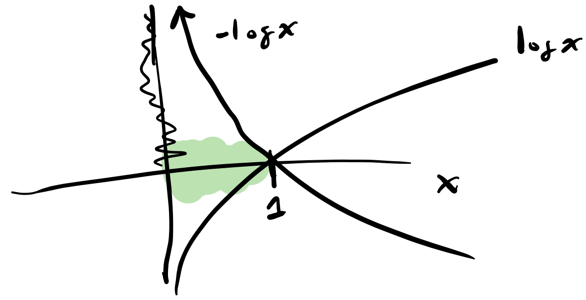
$$f_Y(y) = \frac{d}{dy} F_Y(y) = 3y^2 \qquad \text{for} \quad y \in [0,1]$$

② $X \sim \text{Unif}(0,1)$, $Y = -\log X$

Find f_Y.

$f_X(x) = \mathbb{1}(0 < x < 1)$

$\mathcal{X} = (0,1)$



Support of Y is $\mathcal{Y} = \underline{\underline{(0, \infty)}}$

For $y > 0$

$$F_Y(y) = P(Y \le y)$$

$$= P(-\log X \le y)$$

$$= P(X \ge e^{-y})$$

$$= \int_{e^{-y}}^{\infty} \mathbb{1}(0 < x < 1)\, dx$$

$$= \int_{e^{-y}}^{1} 1\, dx$$

$$= 1 - e^{-y}$$

For pdf of Y.

For $y > 0$,

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

$$= \frac{d}{dy}\left[1 - e^{-y}\right]$$

$$= e^{-y}.$$

$$\boxed{f_Y(y) = e^{-y}\, \mathbb{1}(y > 0)}$$

Theorem (*transformation method*)

Let X be a continuous rv with support $\mathcal{X}$ and let $Y = g(X)$ have support $\mathcal{Y}$.

Suppose

1. *g is monotone on $\mathcal{X}$ and*
2. *$\frac{d}{dy} g^{-1}(y)$ is continuous on $\mathcal{Y}$.*

Memorize

Then

$$f_Y(y) = f_X\left(g^{-1}(y)\right)\left|\frac{d}{dy} g^{-1}(y)\right| \quad \text{for } y \in \mathcal{Y}$$

Exercises: Prove the above and apply to the following:

1. Let $X \sim \text{Exponential}(\lambda)$. Find the pdf of $Y = \sqrt{X}$.
2. Let $X \sim \text{Gamma}(\alpha, \beta)$. Find the pdf of $Y = X^{-1}$.
3. Let $X \sim \text{Exponential}(1)$. Find the pdf of $Y = -\log X$.

① $X \sim f_X(x) = \frac{1}{\lambda} e^{-x/\lambda} \mathbb{1}(x>0)$, $\quad Y = \sqrt{X}$, $\quad$ so $\mathcal{Y} = (0, \infty)$

$$y = \sqrt{x} = g(x) \iff x = y^2 = g^{-1}(y), \qquad \frac{d}{dy} g^{-1}(y) = \frac{d}{dy} y^2 = \underline{2y}$$

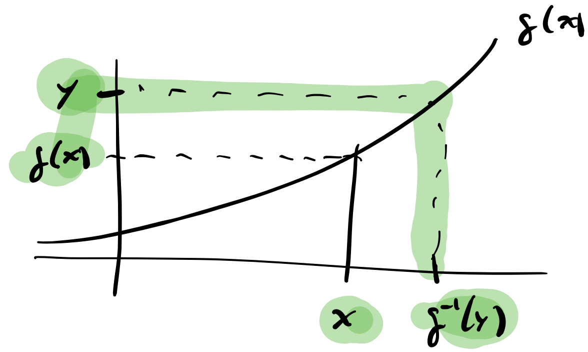
For $y \in (0, \infty)$,

$$f_Y(y) = \frac{1}{\lambda} e^{-y^2/\lambda} \left| 2y \right| = \frac{2y}{\lambda} e^{-y^2/\lambda} .$$

$$f_Y(y) = \frac{2y}{\lambda} e^{-y^2/\lambda} \mathbb{1}(y > 0)$$

Derive Transform method:

Let g monotone $\uparrow$. $\quad$ Use cdf method:



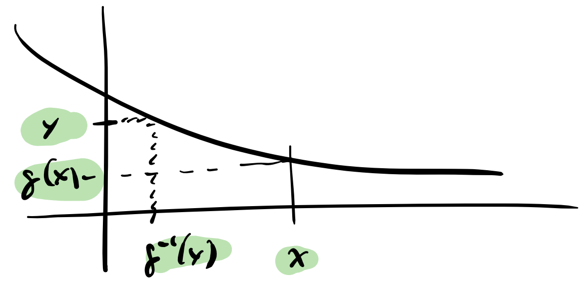
$$
\begin{aligned}
F_Y(y) &= P(Y \le y) \\
&= P(g(X) \le y) \\
&= P(X \le g^{-1}(y)) \\
&= F_X(g^{-1}(y))
\end{aligned}
$$

mono $\uparrow$

$$f_Y(y) = \frac{d}{dy} F_X(g^{-1}(y)) = f_X(g^{-1}(y)) \underbrace{\frac{d}{dy} g^{-1}(y)}_{>0} = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|$$

since g is monotone increasing.

Let g be monotone $\downarrow$



$$
\begin{aligned}
F_Y(y) &= P(Y \le y) \\
&= P(g(X) \le y) \\
&= P(X \ge g^{-1}(y)) \\
&= 1 - P(X < g^{-1}(y)) \\
&= 1 - F_X(g^{-1}(y))
\end{aligned}
$$

$$
\begin{aligned}
f_Y(y) = \frac{d}{dy} F_Y(y) &= \frac{d}{dy}\left[1 - F_X(g^{-1}(y))\right] \\
&= - f_X(g^{-1}(y)) \underbrace{\frac{d}{dy} g^{-1}(y)}_{<0} \\
&= f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| .
\end{aligned}
$$

Exercise: Let $Z \sim \text{Normal}(0,1)$. Find the pdf of $W = Z^2$.

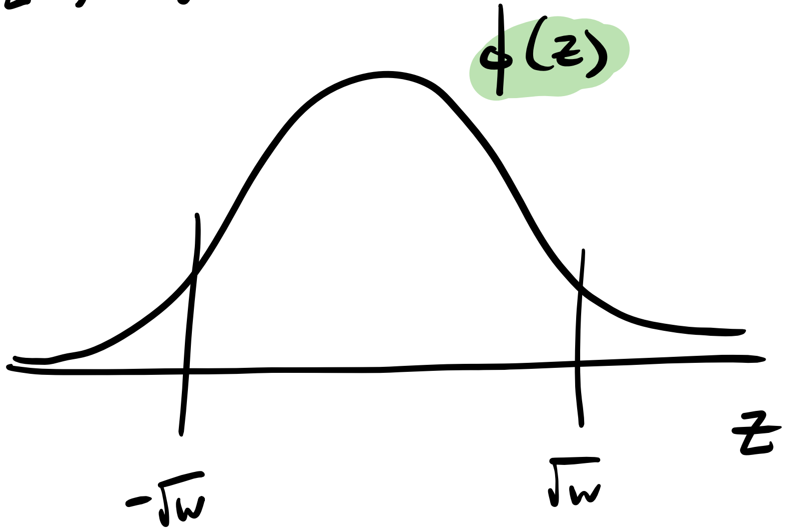
(i) Write down support of W. $\mathcal{W} = [0, \infty)$

For $w \in [0,\infty)$,

$$F_W(w) = P(W \le w)$$

$$= P(Z^2 \le w)$$

$$= P(-\sqrt{w} \le Z \le \sqrt{w})$$



$$= \Phi(\sqrt{w}) - \Phi(-\sqrt{w})$$

$$f_W(w) = \frac{d}{dw} F_W(w)$$

$$= \frac{d}{dw}\left[\Phi(\sqrt{w}) - \Phi(-\sqrt{w})\right]$$

$$= \phi(\sqrt{w}) \frac{d}{dw}\sqrt{w} - \phi(-\sqrt{w})\left(\frac{d}{dw} - \sqrt{w}\right)$$

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

$$= 2\ \phi(\sqrt{w}) \frac{1}{2\sqrt{w}}$$

$$= \frac{1}{\sqrt{w}} \phi(\sqrt{w})$$

$$= \frac{1}{\sqrt{w}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(\sqrt{w})^2}{2}}$$

$$= \frac{1}{\Gamma(\frac{1}{2}) 2^{1/2}} w^{1/2 - 1} e^{-w/2}$$

pdf of $\text{Gamma}(\frac{1}{2}, 2)$

a.k.a. Chi-square with df = 1.

$Y = g(X)$

Theorem (Probability integral transform)

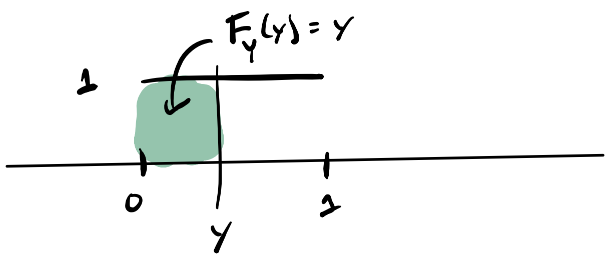
Let X be a rv with continuous cdf F_X. Then $Y = F_X(X) \sim$ Uniform$(0, 1)$.

Exercise: Prove the above (assume F_X is monotone). $\Rightarrow F_X^{-1}$ exists, is monotone increasing

Let $Y = F_X(X)$. Then cdf of Y is

$$
\begin{aligned}
F_Y(y) &= P(Y \le y) \\
&= P(F_X(X) \le y) \\
&= P\left(F_X^{-1}(F_X(X)) \le F_X^{-1}(y)\right) \\
&= P(\underline{X} \le F_X^{-1}(y)) \\
&= F_X\left(F_X^{-1}(y)\right) \\
&= y
\end{aligned}
$$

What is cdf of Unif(0,1)?



Unif(0,1) cdf is $F_Y(y) = \begin{cases} 0 & y \le 0 \\ y & 0 < y < 1 \\ 1 & 1 \le y \end{cases}$

If F_X not continuous, or continuous but not monotone, F_X^{-1} is not well-defined...

Quantile function

The *quantile function* Q_X of a rv X with cdf F_X is the function

$$Q_X(\theta) = \inf\{x : F_X(x) \geq \theta\} \text{ for } \theta \in (0,1).$$

The *θ-quantile* of the rv with cdf F_X is defined as $Q_X(\theta)$.

- If F_X is continuous and strictly increasing on $\mathcal{X}$ then $\boxed{Q_X(\theta) = F_X^{-1}(\theta)}$.
- $Q_X : (0,1) \to \mathcal{X}$, where $\mathcal{X}$ the support of X.
- Q_X is always left-continuous.

Exercises: Find the quantile function of

1. $X \sim \text{Binomial}(2, 1/2)$ distribution. ✓
2. $X \sim \text{Exponential}(\lambda)$ distribution.
3. $X \sim \text{empDist}(x_1, \ldots, x_n)$. } hw

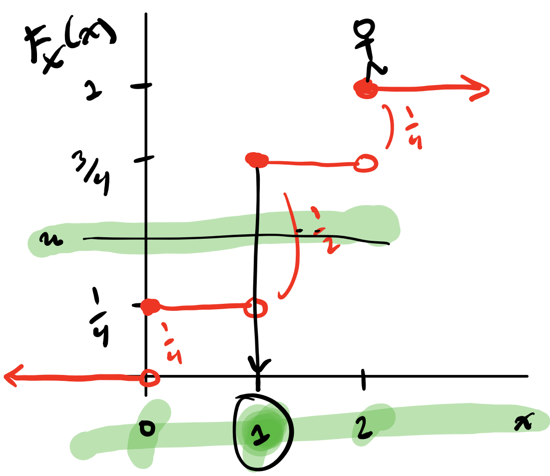
① $X \sim \text{Binom}(2, 1/2)$ dist (flip a coin twice)

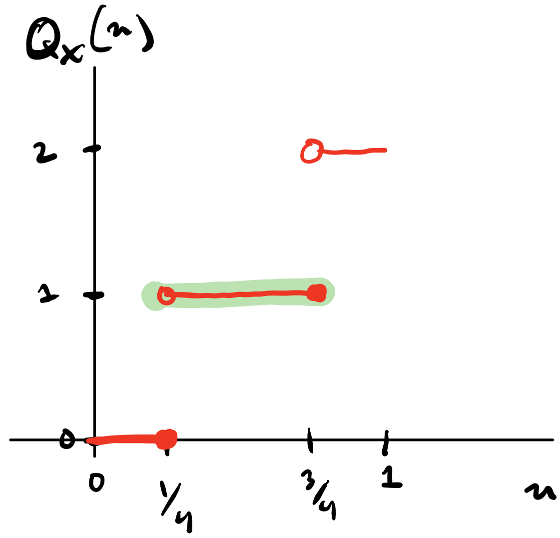
x	0	1	2
$p(X=x)$	$1/4$	$1/2$	$1/4$

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1/4 & 0 \le x < 1 \\ 3/4 & 1 \le x < 2 \\ 1 & 2 \le x \end{cases}$$

$Q_X(u) = \inf\{x \in \mathcal{X} : F_X(x) \ge u\}$ for $u \in (0,1)$.

$Q_X : (0,1) \to \mathcal{X}$

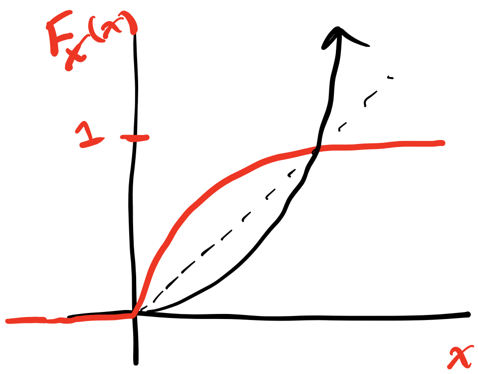


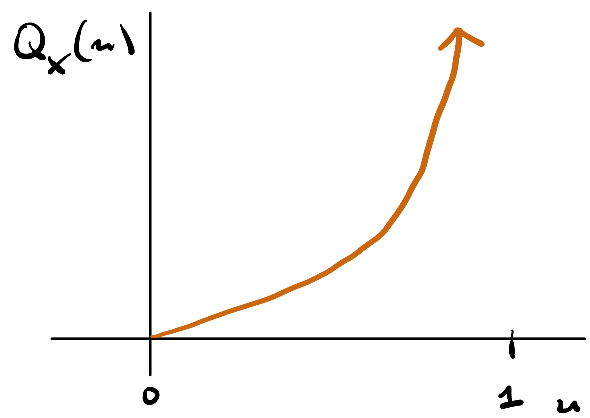


② $X \sim \text{Exponential}(\lambda)$. The cdf is

$$F_X(x) = \begin{cases} 0, & x \le 0 \\ 1 - e^{-x/\lambda}, & x > 0 \end{cases}$$

$$Q_X(u) = F_X^{-1}(u)$$ since F_X monotone on $\mathcal{X} = (0, \infty)$.





To find $Q_X(u)$ write

$$u = 1 - e^{-x/\lambda},$$

& solve for x. $\Rightarrow$

$$u = 1 - e^{-x/\lambda}$$

$$\iff 1 - u = e^{-x/\lambda}$$

$$\iff \log(1-u) = -x/\lambda$$

$$\iff -\lambda \log(1-u) = x$$

So

$$Q_X(u) = -\lambda \log(1-u)$$

Random number generation

Reverse the $\boxed{\text{probability integral transform}}$ to generate $X \sim F_X$ for any F_X:

1. Generate a Uniform(0, 1) realization.
2. Set $X = \boxed{Q_X(U)}$ where $Q_X(u) = \inf\{x : F_X(x) \geq u\}$ is the quantile function.

Exercise: Prove the above (assume F_X is monotone so that $Q_X = F_X^{-1}$).

Many careers spent getting computers to generate "random" Uniform(0, 1) values.

Let $U \sim \text{Unif}(0,1)$, and set $X = F_X^{-1}(U)$.

Then $P(X \leq x) = P(F_X^{-1}(U) \leq x) = P(F_X(F_X^{-1}(U)) \leq F_X(x))$
$= P(U \leq F_X(x)) = F_X(x)$,

so X has cdf F_X.

Maybe on exam.

Exercise: Get realization of $X \sim \text{Exponential}(\lambda)$ beginning with $U \sim \text{Unif}(0,1)$.

The quantile function is $Q_X(u) = -\lambda \log(1-u)$

Just set $X = -\lambda \log(1-U)$

For shift-and-scale transformations, remember that you can use mgfs:

Theorem (*mgf method*)

For any constants a and b, the mgf of $aX + b$ is

$$M_{aX+b}(t) = e^{bt} M_X(at).$$

Exercises:

1. Let $X \sim \text{Gamma}(\alpha, \beta)$. Find distribution of $Y = X/\beta$.
2. Let $X \sim \text{Normal}(\mu, \sigma)$. Find distribution of $Z = (X - \mu)/\sigma$.

$f_X(x; \underbrace{\mu, \sigma}_{\text{parameters}})$

Location-scale families

For any pdf f_Z the family of pdfs given by

$$\left\{ f_X(x; \mu, \sigma) = \frac{1}{\sigma} f_Z\left(\frac{x-\mu}{\sigma}\right) : \mu \in \mathbb{R}, \sigma > 0 \right\}$$

is the *location-scale family with standard pdf f_Z*. We call

- μ the *location parameter*
- σ the *scale parameter*

Example: A location-scale family with standard pdf $\overset{\phi}{f(z)} = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$ are

$$\left\{ f(x; \mu, \sigma^2) = \frac{1}{\sigma} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 \right], x \in \mathbb{R} : \mu \in \mathbb{R}, \sigma > 0 \right\}.$$

The location and scale parameters are μ and σ, respectively.

Just scale or just location families...

For any pdf f_Z the families of pdfs given by

$$\{f_X(x;\sigma) = (1/\sigma)f_Z(x/\sigma) : \sigma > 0\} \quad \text{[loc. parameter = 0]}$$

$$\text{and} \quad \{f_X(x;\mu) = f_Z(x-\mu) : \mu \in \mathbb{R}\} \quad \text{[scl. parameter = 1]}$$

are the *scale family* and the *location family*, respectively, with standard pdf f_Z

Theorem (Standardizing a location-scale random variable)

For any pdf f and $\mu \in \mathbb{R}$, $\sigma > 0$, we have

1. *X has pdf given by $\boxed{\frac{1}{\sigma}f\left(\frac{x-\mu}{\sigma}\right)} \implies \boxed{Z = \frac{X-\mu}{\sigma}}$ has pdf given by $f(z)$.*
2. *Z has pdf given by $f(z) \implies X = \sigma Z + \mu$ has pdf given by $\frac{1}{\sigma}f\left(\frac{x-\mu}{\sigma}\right)$.*

Exercise: Prove the result.

① $X \sim f_X(x;\mu,\sigma) = \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right)$

Let $Z = \frac{X-\mu}{\sigma}$. Find pdf of Z:

$z = \frac{x-\mu}{\sigma} = g(x) \iff x = \sigma z + \mu = g^{-1}(z),$

$\frac{d}{dz} g^{-1}(z) = \sigma$

$$f_Z(z) = f_X(g^{-1}(z)) \left| \frac{d}{dz} g^{-1}(z) \right|$$

$$= \frac{1}{\sigma} f\left(\frac{(\sigma z + \mu) - \mu}{\sigma} \right) \left| \sigma \right|$$

$$= f(z).$$

Exercises:

1. Show that the collection of pdfs

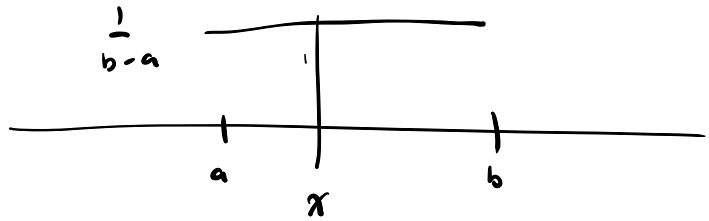
$$\left\{ f_X(x; a, b) = \frac{1}{b-a} \cdot \mathbf{1}(a < x < b) : -\infty < a < b < \infty \right\}$$

is a location scale family. Means identify standard pdf f_Z, also μ, σ.

2. Let $\boxed{Z \sim \text{Exponential}(1).}$ Find pdfs of $\boxed{X = \lambda Z + \mu}$ for $\lambda > 0$, $\mu \in \mathbb{R}$.

$f_Z(z) = e^{-z}$

$f_X(x) = \frac{1}{\lambda} e^{-\frac{(x-\mu)}{\lambda}} \mathbb{1}(x > \mu).$



$$z = \frac{x-a}{b-a} = g(x) \iff x = (b-a)z + a = g^{-1}(z)$$

$$\frac{d}{dz} g^{-1}(z) = b-a$$

$$f_Z(z) = f_X(g^{-1}(z)) \left|\frac{d}{dz} g^{-1}(z)\right| = \frac{1}{b-a} \mathbb{1}\left(a < g^{-1}(z) < b\right) \cdot |b-a|$$

$$= \cancel{\frac{1}{b-a}} \mathbb{1}\left(a < (b-a)z + a < b\right) \cancel{|b-a|}$$

$$= \mathbb{1}\left(0 < (b-a)z < b-a\right)$$

$$= \mathbb{1}(0 < z < 1)$$

So set $f_Z(z) = \mathbb{1}(0 < z < 1)$,

$\mu = a$, $\sigma = b - a$.

Then

$$f_X(x; a, b) = \frac{1}{b-a} \mathbb{1}(a < x < b) =$$

$$= \frac{1}{b-a} \mathbb{1}\left(0 < \frac{x-a}{b-a} < b\right)$$

$$= \frac{1}{b-a} \underbrace{f_Z}_{\text{Unif}(0,1) \text{ pdf.}} \left(\frac{x-a}{b-a}\right)$$