Topic 2.7 Pdf of $Y = g(X)$

We do 4 things in this class

① Build models

* counting (all elements equally likely)
* conditional events with the Theorem of total prob.
* any $g(x) \geq 0$ that is piecewise continuous with finite integral.
* common pdfs and pmfs
* $\boxed{\text{given } f_X(x) \text{ and } Y = g(X), \text{ find } f_Y(y)}$

② Compute probabilities for a given model

* Axioms of probability
* pmf, pdf, cdf

③ Learn

* Bayes Rule
* Conditioning event C contains RV X (chop and scale)

④ compute summary statistics

- expected value, variance, moments
* conditional mean, conditional variance

PDF of a function of a RV (Chapter 4.5)

i.e., given X and its pdf $f_X(x)$,
and given $Y = g(X)$, what is $f_Y(y)$?

Recall: Y is a random variable

With respect to the 4 things we do in this class, this can be considered to be computing probabilities for the purpose of creating probability models

Why is this useful?

- characterize the output of a system if you know the functional form of the system, and know the PDF of the input
- have a good probability model for X, but want to know about Y

ex: measure V, I, compute power

We already know $E(g(X))$

Progressively harder problems

$P(g(X) \le y)$ for a specific value y

$F_Y(y)$ for all values of y

$f_Y(y)$ " " "

Some useful functions

$Y = aX + b$ (linear)

$Y = X^2$; $Y = +\sqrt{X}$

$Y = |X|$

$Y = 1/X$

$Y = \log(X)$

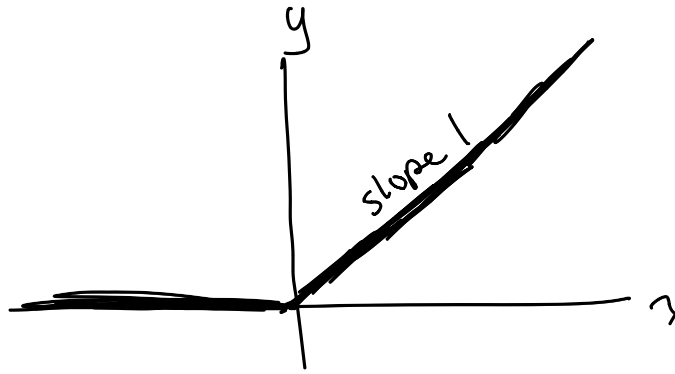
$Y = \sin(X)$

$Y = Q(X)$ (quantizer)

$Z = X - Q(X)$ (quantization error)

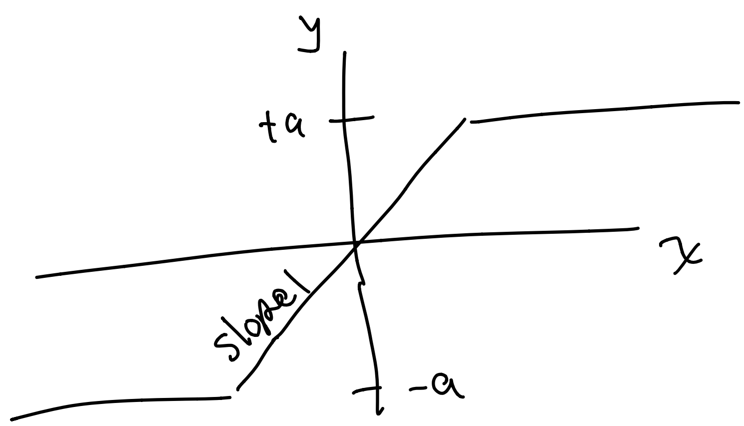
rectifier

$$Y = \begin{cases} X & x \geq 0 \\ 0 & x < 0 \end{cases}$$



limiter

$$Y = \begin{cases} -a & x < -a \\ X & -a \leq X \leq a \\ a & x > a \end{cases}$$



---

methods

1) General case (always works)

   a 2-step process to find

   a) $F_Y(y)$ and then

   b) $f_Y(y)$

2) specific cases ≡ short cuts – only relevant sometimes

   a) $g(X)$ is linear $Y = aX + b$

   b) $g(x)$ is monotonically increasing/decreasing and <u>differentiable</u>

   c) $g(X)$ has flat parts

Example (thru progressively harder problems)

$Y = +\sqrt{X}$ $\qquad$ X is uniform between $[0,1]$

a) what's $E(Y)$?

$$E(Y) = \int_{-\infty}^{\infty} \sqrt{x}\, f_X(x)\, dx = \int_0^1 \sqrt{x}\, dx$$

(don't need to compute pdf of Y)

---

b) what's $P(Y \le 1/2)$? $\qquad y = 1/2$

$$P(Y \le 1/2) = P(\sqrt{X} \le 1/2) = P\left(X \le \left(\tfrac{1}{2}\right)^2\right)$$

$$= P(X \le 1/4) = 1/4$$

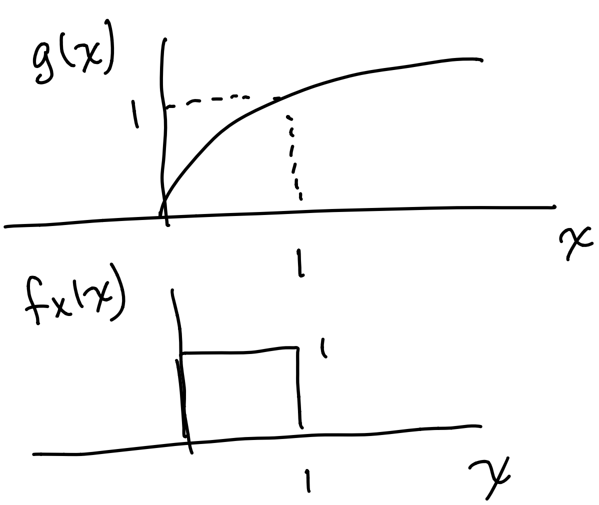
(transform event in Y into an event in X)

---

c) generalize: what's $P(Y \le y)$? (ie, what's $F_Y(y)$?)

what ranges of y are relevant?

Look at both $g(x)$ and $f_X(x)$ — any special cases?



YES. $\sqrt{x}$ is undefined for $x<0$

so $y<0$ is special

AND $f_X(x)$ is non zero for $0 \le x \le 1$ only, so

$y > 1$ is also special

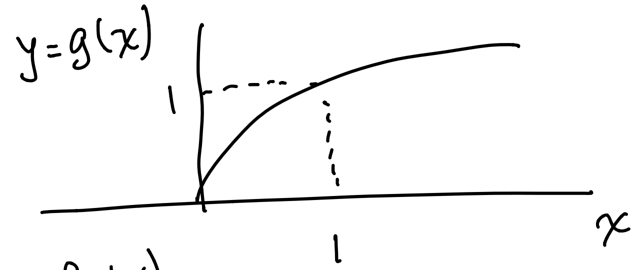
so to compute $P(Y \le y)$, we need to consider 3 regions:

$\{y < 0\}$, $\{0 \le y \le 1\}$, $\{1 < y\}$

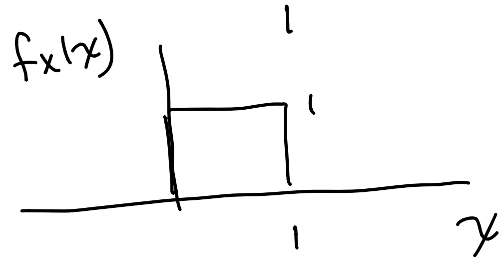
Case 1 : $y < 0$

$$P(Y \le y) = 0 \text{ if } y < 0 \quad \rightarrow \text{cannot happen}$$

Case 2: $y > 1$



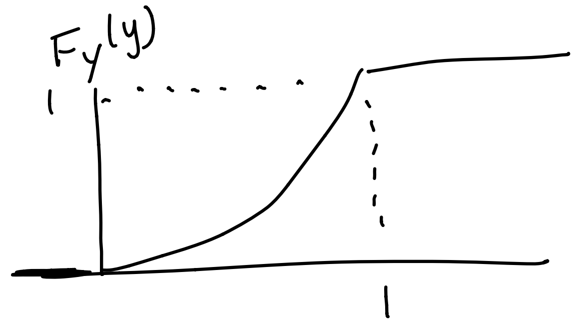
$$P(Y \le y) = P(\sqrt{X} \le y)$$
$$= P(X \le y^2)$$
$$= 1 \quad \text{if } y > 1$$



case 3: $0 \le y \le 1$

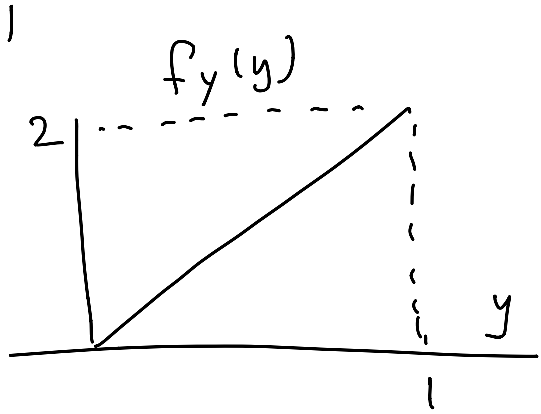
$$P(Y \le y) = P(\sqrt{X} \le y) = P(X \le y^2)$$
$$= \int_0^{y^2} 1 \, dx = x \Big|_0^{y^2} = y^2$$

$$\boxed{F_Y(y) = \begin{cases} 0 & y < 0 \\ y^2 & 0 \le y \le 1 \\ 1 & y > 1 \end{cases}}$$



d) what's $f_Y(y)$? check: is $F_Y(y)$ continuous?

$$f_Y(y) = \frac{d F_Y(y)}{dy} = \begin{cases} 0 & y < 0 \\ 2y & 0 \le y \le 1 \\ 0 & y > 1 \end{cases}$$



Recall: X was uniform

check: does pdf have area 1?

How do you know if your answer is correct?

1. $\int_{-\infty}^{\infty} f_Y(y)\, dy = 1$

2. It is a function of y, not x

3. You have found the answer for <u>every</u> value of y

4. You have substituted the relevant $f_X(x)$ or $F_X(x)$

5. See if your answer makes sense in a small region

Common mistakes to avoid

if $Y = X^2$ $\quad F_Y(y) = F_X(x)^2$

if $Y = X^2$ $\quad f_Y(y) = f_X(x)^2$

Neither one correct

The method of finding $F_Y(y)$ and differentiating to get $f_Y(y)$ always works.

We can use it to compute a shortcut for when $g(X)$ is linear

Example 4.31 Y is a linear function of X

$$Y = aX + b \quad \text{when } a \neq 0$$

whats $F_Y(y)$ and $f_Y(y)$?

2step process - find $F_Y(y)$ and differentiate to get $f_Y(y)$

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

2 cases from here: $a > 0$ and $a < 0$

case 1: $a > 0$

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

case 2: $a < 0$

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

(for a continuous RV)

next...

what's $f_Y(y) = \dfrac{dF_Y(y)}{dy}$?

case 1: $a > 0$ $\quad f_Y(y) = \dfrac{d}{dy} F_X\left(\dfrac{y-b}{a}\right)$

use the chain rule (see below)

$$= f_X\left(\frac{y-b}{a}\right) \frac{d}{dy}\left(\frac{y-b}{a}\right)$$

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

case 2: $a < 0$ $\quad f_Y(y) = \dfrac{d}{dy}\left[1 - F_X\left(\dfrac{y-b}{a}\right)\right]$

$$= -f_X\left(\frac{y-b}{a}\right) \frac{d}{dy}\left[\frac{y-b}{a}\right]$$

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

combine.

The linear shortcut

$$\boxed{f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right) \quad \text{for all } a \neq 0 \quad \text{when } Y = aX + b}$$

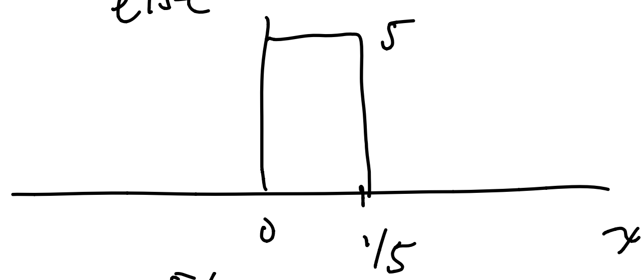
Recall the chain rule:

$$\frac{d}{dx} g_1(g_2(x)) = g_1'(g_2(x))\, g_2'(x)$$

So $\quad \dfrac{d}{dx} F_X(g_2(x)) = f_X(g_2(x)) \dfrac{d}{dx} g_2(x)$

Example of the linear shortcut

$$f_X(x) = \begin{cases} 5 & 0 \le x \le 1/5 \\ 0 & \text{else} \end{cases}$$



$$Y = g(X) = \frac{5X}{2} - 1$$

$$= aX + b \quad \text{where} \quad a = 5/2, \; b = -1$$

then $$f_Y(y) = \frac{1}{5/2} f_X\left(\frac{y+1}{5/2}\right)$$

Substitute: when $0 \le \frac{y+1}{5/2} \le 1/5$

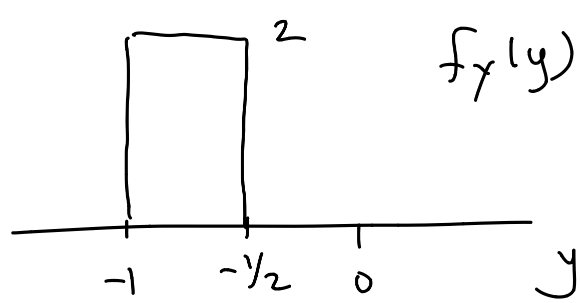
then $$f_Y(y) = \frac{1}{5/2}(5) = 2$$

and zero otherwise.

Simplify the condition... $0 \le y+1 \le 1/2$

$-1 \le y \le -1/2$

so $$f_Y(y) = \begin{cases} 2 & -1 \le y \le -1/2 \\ 0 & \text{else} \end{cases}$$



$\boxed{\text{stretched and shifted}}$

A linear function of a Gaussian RV is also a Gaussian RV

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

$Y = aX + b$

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

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

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

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

$Y \sim N(a\mu+b,\ a^2\sigma^2)$

This makes sense: $E(aX+b) = aE(X)+b = a\mu + b$

$Var(aX+b) = a^2 Var(A) = a^2\sigma^2$

Ex: $X \sim N(\mu, \sigma^2)$

$$Z = \frac{X-\mu}{\sigma}$$

$Z \sim N(0,1)$ standard normal

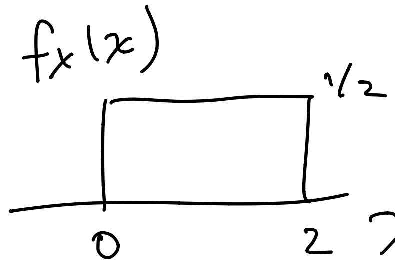
Derived RVs ($Y = g(X)$) are often mixed RVs

Example: X is a uniform RV between $[0, 2]$

and $Y = \min(X, 1) = \begin{cases} X & X < 1 \\ 1 & X \geq 1 \end{cases}$ (a clipper)

what is the pdf of Y?

Answer:

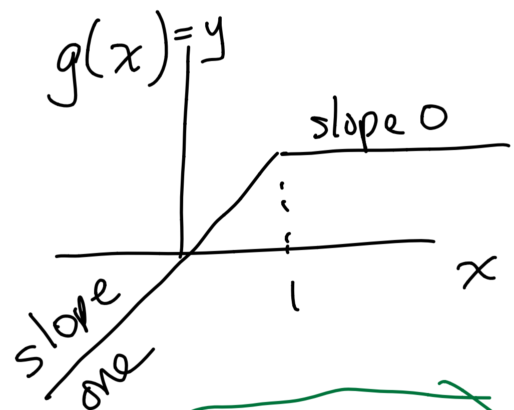


$$f_X(x) = \begin{cases} \frac{1}{2} & 0 \leq x \leq 2 \\ 0 & \text{else} \end{cases}$$

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt = \begin{cases} 0 & x < 0 \\ x/2 & 0 \leq x \leq 2 \\ 1 & x > 2 \end{cases}$$

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

what intervals of y are "interesting"?



case ① $y < 1$

case ② $y = 1$

case 3 $y > 1$

if $g(x)$ has a region with slope zero, then $Y = g(X)$ is likely a mixed RV.

Case 1: $y < 1$

We can apply the linear short cut

$$f_Y(y) = f_X(y) = \begin{cases} 1/2 & 0 < y < 1 \\ 0 & y < 0 \end{cases}$$

or do it out...

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

case 3: $y > 1$

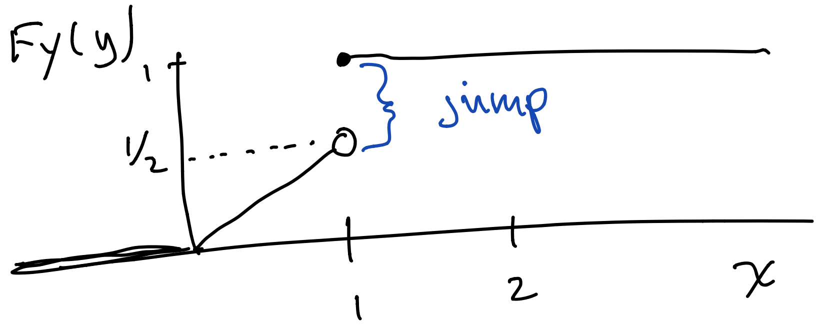
$$F_Y(y) = P(Y \le y) = 1$$

$$f_Y(y) = 0$$

case 2: $y = 1$

$$F_Y(y) = P(Y \le 1) = 1$$

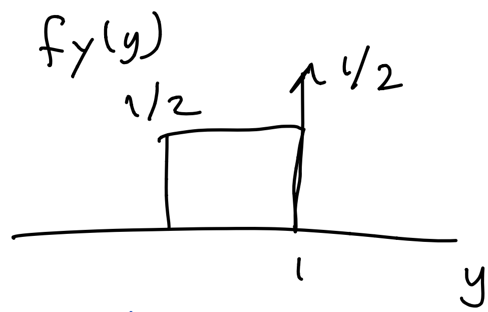
sketch



$$F_Y(y) = \begin{cases} F_X(y) & y < 1 \\ 1 & y \ge 1 \end{cases}$$

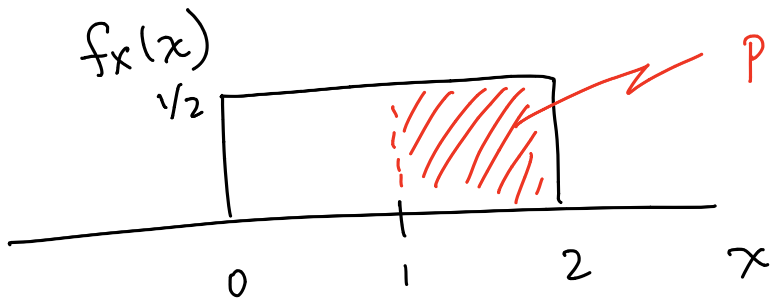
Differentiating,

$$f_Y(y) = \begin{cases} f_X(y) & y < 1 \\ 1/2\, \delta(y) & y = 1 \\ 0 & \text{else} \end{cases}$$

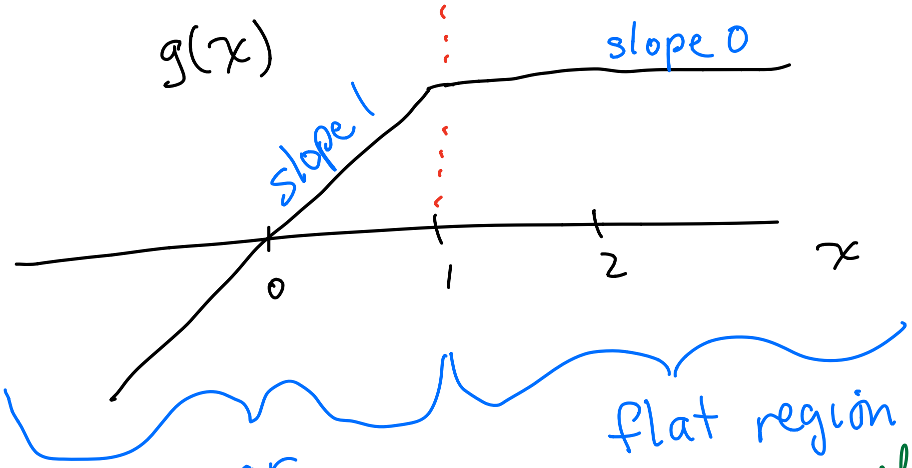


A mixed RV!

An intuitive approach to the previous problem

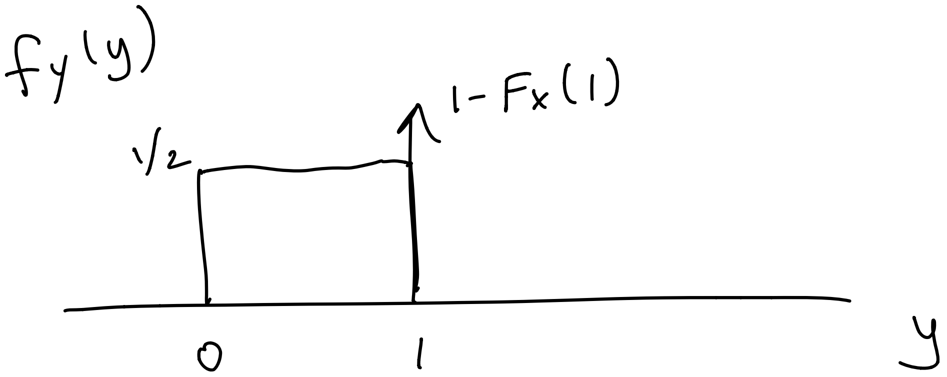


$P(X>1) = 1 - F_X(1)$



linear region doesn't change PDF

flat region forces all values of x in this range to 1 $\Rightarrow$ delta function with area $= P(X>1)$



Another special case (aka shortcut)

if $Y = g(X)$ $g(\cdot)$ is a differentiable and monotonic function

then we can directly compute $f_Y(y)$ from $f_X(x)$ without computing $F_Y(y)$

Monotonic:

increasing $g(x_1) < g(x_2)$ for all $x_1 < x_2$

decreasing $g(x_1) > g(x_2)$ for all $x_1 < x_2$

A monotonic function can be inverted

$$y = g(x) \Rightarrow x = h(y)$$

$$\Rightarrow g(h(y)) = y$$

ex: $y = g(x) = \exp(x) \Rightarrow h(y) = \frac{1}{a} \ln y$ for $a \neq 0$

ex: $y = ax + b \Rightarrow h(y) = \frac{y-b}{a}$ for $a \neq 0$

monotonic shortcut

then

$$\boxed{f_Y(y) = \frac{f_X(x)}{\left|\frac{dg(x)}{dx}\right|} \quad \text{if } g \text{ is a monotonic differentiable function}}$$

and recall $f_X(\underline{x}) = f_X(\underline{h(y)})$

Let's verify this is true (when $g(x)$ is monotonically increasing)

Use 2-step process

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

differentiate w.r.t y using the chain rule

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

To put this in the form above using $g(x)$ directly,

differentiate both sides of $\quad y = g(h(y))$

$$1 = \left[\frac{d}{dh} g(h(y))\right]\left[\frac{dh(y)}{dy}\right]$$

$$= \frac{dg(x)}{dx} \frac{dh(y)}{dy} \qquad \text{since } h(y) = x$$

so $$\frac{dh(y)}{dy} = 1 \Big/ \frac{dg(x)}{dx}$$

Add the absolute value, which will include the case where $g(x)$ is monotonically decreasing to get

$$\boxed{f_Y(y) = \frac{f_X(x)}{\left|\frac{dg(x)}{dx}\right|}}$$

when $g(x)$ monotonic + differentiable

but don't forget this needs to be expressed in terms of y!

Application of the monotonic shortcut

$$Y = g(X) = +\sqrt{X} \quad \text{and } X \text{ is uniform } [0,1]$$

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

(this is the same example as above. we should get the same answer)

$$\frac{dg}{dx} = \frac{1}{2\sqrt{x}}$$

$$f_Y(y) = \frac{f_X(x)}{\left|\frac{dg}{dx}\right|} = \begin{cases} \frac{f_X(x)}{1/2\sqrt{x}} & x \in [0,1] \\ 0 & \text{else} \end{cases}$$

But this is not in terms of y so we are not done!

$$f_Y(y) = \begin{cases} 2\sqrt{y^2}\, f_X(y^2) & y^2 \in [0,1] \\ 0 & \text{else} \end{cases}$$

$$= \begin{cases} 2y & \text{when } y \in [0,1] \\ 0 & \text{else} \end{cases}$$

(good. this is the same answer)

Simulating a RV on a computer

$X = \text{rand}(\cdots)$ $\quad X \in [0,1]$ uniformly distributed

X is actually discrete with, say 2^{31} values, generated pseudorandomly.

$\Rightarrow$ a good approximation to a continuous RV

Suppose we want Z to have a CDF $F_Z(z)$?

What $Z = g(X)$ will create a RV Z with the proper CDF

Answer: If $g(x) = F_Z^{-1}(z)$ then Z has the desired CDF

(as long as $g(X)$ is a monotonic function and $F_Z(z)$ is invertible)

Proof: $F_Z(z) = P(Z \le z) = P(g(X) \le z)$

$= P(X \le g^{-1}(z)) = F_X(g^{-1}(z))$

$= g^{-1}(z)$ in the desired range, since X is uniform so $F_X(x) = x$ in the desired range

$= F_Z(z)$

Some of the common continuous RVs are functions of other common RVs.

- Cauchy RV : θ is uniform $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$

  $Y = \tan\theta$ is Cauchy

- Chi-squared RV : X is Gaussian

  $Y = X^2$ is Chi-square

- Log normal RV : X is Gaussian

  $Y = \exp(X)$ is log-normal

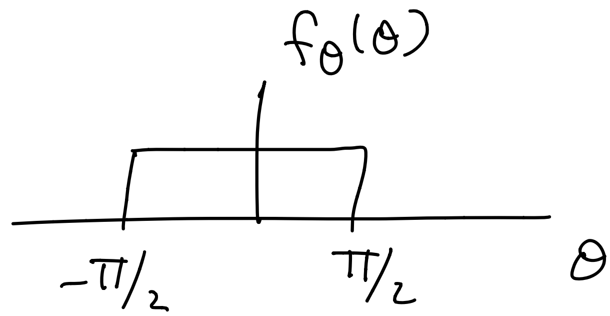
Deriving the Cauchy distribution

Given
$y = \tan\theta$ where θ is uniformly distributed on $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$

what's the pdf of y? (answer, the Cauchy distribution)

step 1: express $f_\theta(\theta)$

$$f_\theta(\theta) = \begin{cases} 1/\pi & -\frac{\pi}{2} < \theta < \frac{\pi}{2} \\ 0 & \text{else} \end{cases}$$



step 2: find $F_\theta(\theta)$

$$F_\theta(\theta) = \begin{cases} 0 & \theta \le -\pi/2 \\ (\theta + \pi/2)/\pi & -\pi/2 < \theta < \pi/2 \\ 1 & \theta \geq \pi/2 \end{cases}$$

step 3: $F_Y(y) = P(Y \le y) = P(\tan\theta \le y)$

$$= P(\theta \le \tan^{-1}(y)) = F_\theta(\tan^{-1}(y))$$

$$= \frac{\tan^{-1}(y) + \pi/2}{\pi} \quad \text{in region of interest}$$

(Note, by substituting here w/ $F_\theta(\theta)$ we can avoid chain rule)

Step 4: $f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{1}{\pi}\frac{1}{1+y^2}$ since $\frac{d(\tan^{-1}(y))}{dy} = \frac{1}{1+y^2}$

Review

- You can always apply the 2-step process.

$F_X(x)$ $\quad$ $F_Y(y)$ ⟵ find this 1st

$f_X(x)$ $\quad$ $f_Y(y)$ ⟵ goal

substitute either one of these, whichever makes the math easier

How to solve:

1) Can you apply one of the short cuts?
   - linear $g(X)$
   - any flat spots in $g(x)$
   - monotonic differentiable $g(X)$

2) Find $F_Y(y) = P(Y \le y)$
   - identify distinct regions for various values y.
     → these regions depend on $g(x)$ and $f_X(x)$.
   - manipulate $\{Y \le y\}$ into the same event but expressed using X, not Y.
   - substitute $F_X(x)$ here if it simplifies computation <u>and</u> substitute to eliminate x.

3) Differentiate $F_Y(y)$ to get $f_Y(y)$
   - apply chain rule if necessary
   - substitute $f_X(x)$ here if needed