APMA 1650 - Spring 2021
Lecture 13

Friday, Feb 19, 2021

Summary of Continuous Random Variables

- Def: $F_X(x) = P(X \leq x)$ is absolutely continuous
- PDF: $f_X(x) = F'_X(x)$, $f_X(x) \geq 0$, $\int_{-\infty}^{\infty} f_X(x)\mathrm{d}x = 1$
- Area Rule:
  $P(a < X < b) = \int_a^b f_X(x)\mathrm{d}x = F_X(b) - F_X(a)$
- Expected Value $EX = \int_{-\infty}^{\infty} x f_X(x)\mathrm{d}x$
- LOTUS: $E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\mathrm{d}x$
- Variance: $\mathsf{Var}(X) = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x)\mathrm{d}x$

Functions of Continuous Random Variables

Suppose X is a continuous random variable and $g : \mathbb{R} \to \mathbb{R}$ is some function. Let

$$Y = g(X)$$

We know LOTUS

$$EY = \int_{-\infty}^{\infty} g(x) f_X(x) \mathrm{d}x.$$

Questions:

- What kind of random variable is Y?
- Is Y continuous?
- If so, what are $F_Y(y)$ and $f_Y(y)$?

Discrete Case

In the discrete case we could find the PMF of $Y = g(X)$ by

$$P_Y(y) = \sum_{\{x : g(x) = y\}} P_X(x).$$

This works for ANY $g : \mathbb{R} \to \mathbb{R}$!

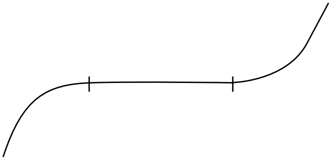
Simply sum the probabilities over all the $x \in R_X$ such that $g(x) = y$.

Continuous case

More complicated.

- $Y = g(X)$ may not be continuous anymore
- If g has flat parts then $g(X)$ will be a partially discrete (i.e $g(x) = \text{const}$)
- If g is not differentiable then $g(X)$ may not have a PDF.

"Ugly" cases:



Non continuous example (Flat parts)

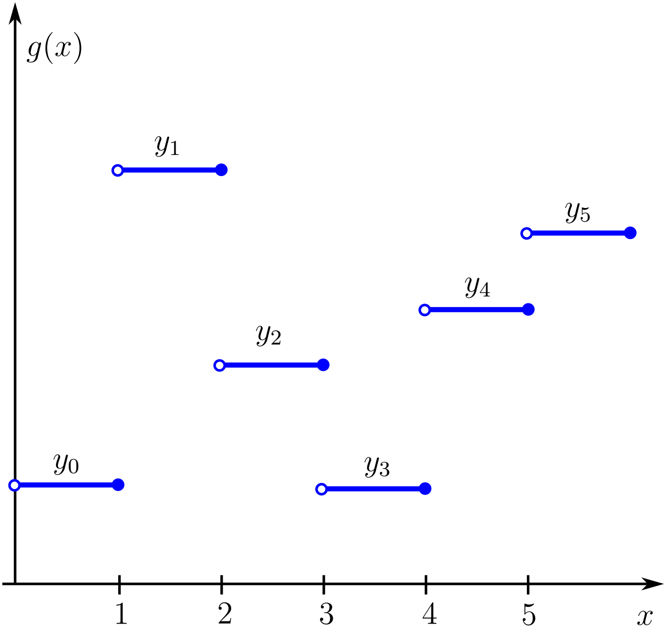
If $g(x)$ has flat parts, then $Y = g(X)$ is not continuous (it will end up mixed in general).

Example: If $\{y_1, y_2, \ldots\} \subset \mathbb{R}$ and

$$g(x) = \sum_{k=0}^{\infty} y_k I_{(k,k+1]}(x),$$

where

$$I_{(k,k+1]}(x) = \begin{cases} 1 & k < x \leq k+1 \\ 0 & \text{otherwise} \end{cases}$$



$g(x)$
y_1
y_5
y_4
y_2
y_0
y_3
1
2
3
4
5
x

$Y = g(X)$ is a discrete random variable with PMF

$$P_Y(y_k) = P(k < X \leq k+1)$$

LOTUS still works

$$\begin{aligned} EY &= \int_{-\infty}^{\infty} g(x) f_X(x) \mathrm{d}x \\ &= \sum_{k=0}^{\infty} y_k \int_k^{k+1} f_X(x) \mathrm{d}x \\ &= \sum_{k=0}^{\infty} y_k P(k < X \leq k+1) \\ &= \sum_{k=0}^{\infty} y_k P_Y(y_k) \end{aligned}$$

What about nice g? CDF Method

If g is differentiable and has no flat parts. Then $Y = g(X)$ is again a continuous RV.

CDF Method:

1. Find $R_Y = g(R_X)$
2. Find the CDF of Y

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

3. Find the density

$$f_Y(y) = F_Y'(y)$$

Example Suppose that $X \sim$ Uniform$(0, 1)$ and $Y = e^X$. What are $F_Y(y)$ and $f_Y(y)$?

Example What about $X \sim \text{Uniform}(-1, 1)$ and $Y = X^2$?

Method of Transformations

Another approach that gives a direct way to compute the PDF is the Method of Transformations

First Assume:

- g is differentiable
- g is strictly increasing $x_1 < x_2 \Leftrightarrow g(x_1) < g(x_2)$

Let X be a continuous RV and $Y = g(X)$, then

$$f_Y(y) = \begin{cases} \frac{f_X(x)}{g'(x)} = f_X(x)\frac{\mathrm{d}x}{\mathrm{d}y} & \text{when } g(x) = y \\ 0 & \text{otherwise} \end{cases}$$

Proof

Since g is strictly increasing g^{-1} is well-defined. For each $y \in R_Y$ there is a unique $x = g^{-1}(y)$ such that $g(x) = y$

$$
\begin{aligned}
F_Y(y) &= P(Y \leq y) \\
&= P(g(X) \leq y) \\
&= P(X \leq g^{-1}(y)) \quad \text{strictly increasing} \\
&= P(X \leq x)
\end{aligned}
$$

Therefore for $y \in R_Y$ and $x = g^{-1}(y)$

$$
\begin{aligned}
f_Y(y) &= f_X(x) \frac{\mathrm{d}x}{\mathrm{d}x} \quad \text{chain rule} \\
&= \frac{f_X(x)}{g'(x)} \quad \text{using } \frac{\mathrm{d}x}{\mathrm{d}y} = \frac{1}{\frac{\mathrm{d}y}{\mathrm{d}x}} = \frac{1}{g'(x)}
\end{aligned}
$$

Change of variables interpretation

We can also see this as a change of variables formula on the "measures" $f_X(x)\mathrm{d}x$ and $f_Y(y)\mathrm{d}y$ via

$$y = g(x) \quad \Rightarrow \quad \mathrm{d}y = g'(x)\mathrm{d}x$$

or

$$f_Y(y)\mathrm{d}y = f_X(x)\mathrm{d}x \quad \Rightarrow \quad f_Y(y) = f_Y(x)\frac{\mathrm{d}x}{\mathrm{d}y}$$

This explains why LOTUS still holds

What about strictly decreasing?

Suppose that g is strictly decreasing

$$x_1 < x_2 \quad \Leftrightarrow \quad g(x_1) > g(x_2).$$

Then for $y = g(x)$

$$\begin{aligned} F_Y(y) &= P(g(X) \leq y) \\ &= P(X \geq x) \\ &= 1 - F_X(x) \end{aligned}$$

Therefore since $g'(x) < 0$

$$\begin{aligned} f_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y}(1 - F_X(x)) = -f_X(x)\frac{\mathrm{d}x}{\mathrm{d}y} \\ = f_X(x)\left|\frac{\mathrm{d}x}{\mathrm{d}y}\right| = \frac{f_X(x)}{|g'(x)|} \end{aligned}$$

Monotone case

Now Assume:

- g is differentiable
- g is strictly monotone (either increasing or decreasing)

Let X be a continuous RV and $Y = g(X)$, then

$$f_Y(y) = \begin{cases} \frac{f_X(x_*)}{|g'(x^*)|} = f_X(x_*) \left| \frac{\mathrm{d}x_*}{\mathrm{d}x} \right| & \text{when } g(x_*) = y \\ 0 & \text{otherwise} \end{cases}$$

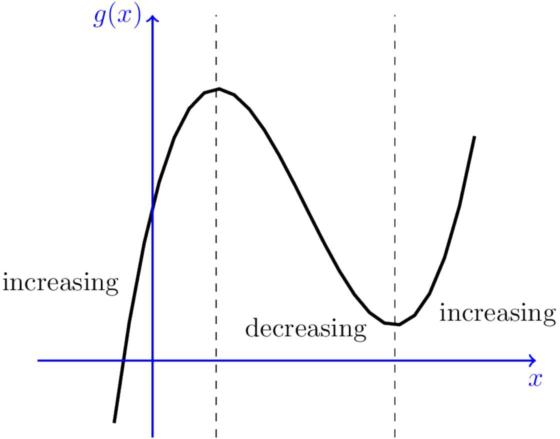
Example Consider X with PDF

$$f_X(x) = \begin{cases} 4x^3 & 0 < x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

What is $Y = 1/X$?

General Case

What if the function is not monotone (and therefore not invertible). Break it up into monotone pieces



Monotone case

Now Assume:

- g is differentiable
- R_X can be broken in to a finite number of intervals where $g(x)$ is strictly monotone.

The PDF of $Y = g(X)$ is given by

$$f_Y(y) = \sum_{k=1}^{n} \frac{f_X(x_k)}{|g'(x_k)|} = \sum_{k=1}^{n} f_X(x_k) \left| \frac{\mathrm{d}x_k}{\mathrm{d}y} \right|$$

where $x_1, x_2, \ldots x_n$ are all the solutions to $g(x) = y$.

Example Consider the PDF

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

Find the PDF of $Y = X^2$.

BONUS: How to simulate an RV

This is a very useful technique for numerically generating random variables from uniform ones

Let X be a random variable with invertible CDF $F_X(x)$, then

$$Y = F_X(X) \sim \text{Uniform}(0, 1).$$

This means that

$$F_X^{-1}(\text{Uniform}(0, 1)) \sim X$$

Proof: $y \in [0, 1]$

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

BONUS: How to simulate an RV

Steps (Inverse CDF method)

1. Find the inverse $F_X^{-1}(y)$ of the CDF $F_X(x)$ of the random variable X you want to generate
2. Generate $U \sim \text{Uniform}(0, 1)$ (MATLAB `rand`)
3. Calculate $F_X^{-1}(U)$
4. Profit

This actually works for ANY CDF (discrete or continuous) by defining

$$F_X^{-1}(y) = \inf\{x : F_X(x) = y\}$$