國立清華大學
NATIONAL TSING HUA UNIVERSITY

EE 306001
Probability

Lecture 17: further topics on random variable

李祈均

Outline

Reading 4.1

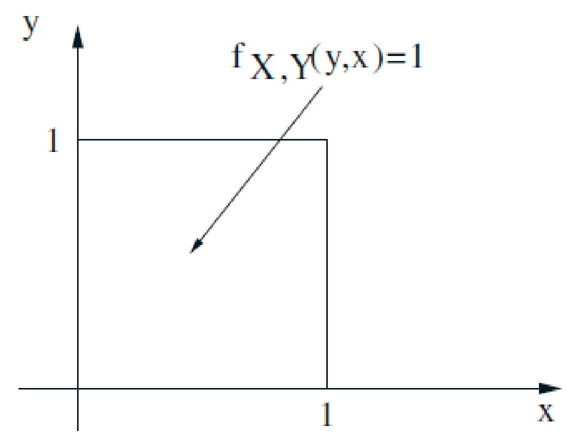
- Derived distributions

Derived distribution

Idea:

- Trying to find the distribution of a function of one or more known random variable with known probability law
- These types of distributions are called *derived distribution*

- Say for example:
  - You are told (X,Y) is uniform
  - You are interested in finding X/Y
    - Finding the ratio



Derived distribution

- Essentially, you are trying to obtain the pdf for:

$$g(X,Y) = \frac{Y}{X}$$

  – Involves deriving a distribution
  – Note that $g(X,Y)$ is also a random variable

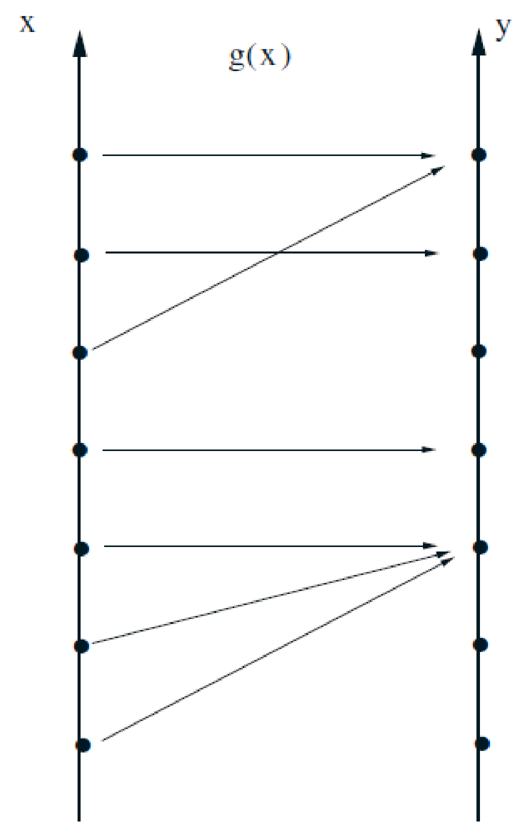
- *Important: When not to find them
  – If you don't need to find $g(X,Y)$ if not needed
  – For example:

$$E[g(X,Y)] = \iint g(x,y) f_{X,Y}(x,y)\,dx\,dy$$

How to find them?

- Discrete case
  - Obtain probability mass for each possible value of $Y = g(X)$

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



- Just sum up the probability in the X that leads to $Y = y$ to come up with t new PMF

Continuous case

- Can we do the same thing as in the discrete case?
  - Not really, we don't have probability that operates "points-to-points"
  - Because?
  - Probability in continuous case is zero at every individual point

- So what we want to find is not the probability,
  - We want to find the actual density for Y
  - In essence:
    - Find the probability of falling in a small interval around $Y = y$, and look back at the set of x's that leads to those y's
    - Then set those probability equal to each other to find density

Continuous case

- Instead of working with intervals, we work directly with CDF
  - What is CDFs? Cumulative distribution function
  - Essentially, a big intervals!

- Two step procedure
  - Get CDF of Y: $F_Y(y) = P(Y \leq y)$
  - Differentiate the CDF to obtain PDF

$$f_Y(y) = \frac{dF_Y}{dy}(y)$$

Example

- X: uniform on [0,2]
- $Y = X^3$
  - Now how to find the PDF for Y?

- Y can take value from [0,8]

Example

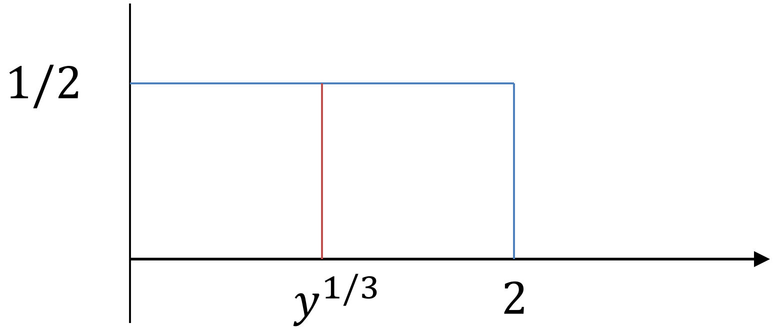
- Lets try to apply the two-step procedure

$$F_Y(y) = P(Y \le y) = P\left(X^3 \le y\right)$$

$$= P\left(X \le y^{1/3}\right)$$

What is this, X is uniform, and this is cumulative distribution in the interval of $\left[0, y^{1/3}\right]$

- $P\left(X \leq y^{1/3}\right)$
  - What is the probability (essentially the area)



  - Area: $\frac{1}{2}y^{1/3}$

$$F_Y(y) = P\left(X \leq y^{1/3}\right) = \frac{1}{2}y^{1/3}$$

$$f_Y(y) = \frac{dF_Y}{dy}(y) = \frac{1}{6y^{\frac{2}{3}}}, 0 \leq y \leq 8$$

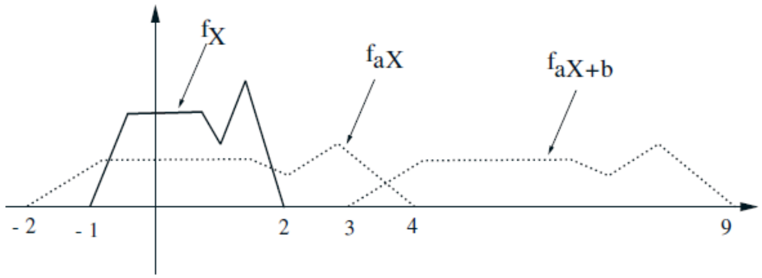
This is not a uniform distribution!

Linear case

- Lets say Y is linear function of X

$$Y = 2X + 5$$

- Imagine simple intuition to reason:
  - Stretch and rescale



From figure above,

Let’s write the formula first (in the case of Y is linear function of X: $Y = aX + b$)

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

- Multiplying means stretching
- Shifting just moving the pdfs

Let's derive this using two step approach

Assume $a > 0$

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

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

Now, let's differentiate this,

Using chain rule:

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

- With $a < 0$

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

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

Now, let's differentiate this,

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

Example on normal random variable

Suppose that X is normal random variable with mean μ, and variance σ^2

Now we say, $Y = aX + b$, a b are scalars, $a \neq 0$

Let's apply the formula

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

What is X?

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

Now what is Y?

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

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

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

We can now recognize this PDF is also just a normal distribution:

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

It’s also a normal distribution

Linear scaling of normal distribution is still a normal distribution

Monotonic cases

- Generalizing the linear case to monotonic function
- Say let X be a random variable
  - Suppose that its range in contained in an interval I
- Now, consider Y is a function of X
  - $Y = g(X)$
  - Assume g is strictly monotonic over that interval I

- Monotonically increasing:
  - $g(x) < g(x')$ for all x, x' ∈ I, satisfying $x < x'$
- Monotonically decreasing:
  - $g(x) > g(x')$ for all x, x' ∈ I, satisfying $x < x'$

- Assume g is differentiable
- Also derivative will necessarily be non-negative in the increasing case, and non-positive in the decreasing case

- A strict monotonic function can be "inverted"

$$y = g(x), x = h(y)$$

- h is the inverted function

- Example

$$g(x) = ax + b, h(y) = \frac{y - b}{a}$$

$$g(x) = e^{ax}, h(y) = \frac{\ln y}{a}$$

- If strictly monotonic function is used, we can use the following formula directly

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

Redo the example

- You are driving from Boston to New York. You are on cruise control, hence your speed is uniformly distributed between 30 and 60 mph.

- You are going to drive a distance of 200 miles, and the time to take for the trip is 200/V
  - $T(V) = \frac{200}{V}$
    - h(T)=V=200/T
  - $V \sim UNIFORM(30,60)$
  - Now we have distribution of V, can we find the distribution of T?

$$f_T(t) = f_V(h(t)) \left|\frac{dh}{dt}(t)\right|$$

$$f_V\big(v = h(t)\big) = \frac{1}{30}$$

$$h(t) = 200/t$$

$$\left|\frac{dh}{dt}(t)\right| = 200/t^2$$

$$f_T(t) = \frac{1}{30}\frac{200}{t^2}$$

Another example

Let $Y = g(X) = X^2$, where X is a continuous uniform random variable on the interval $(0,1]$.

Within this interval, g is strictly monotonic,

First, find the inverse:

$$h(y) = \sqrt{y}$$

- Thus,

for any $y \in (0,1]$, we have

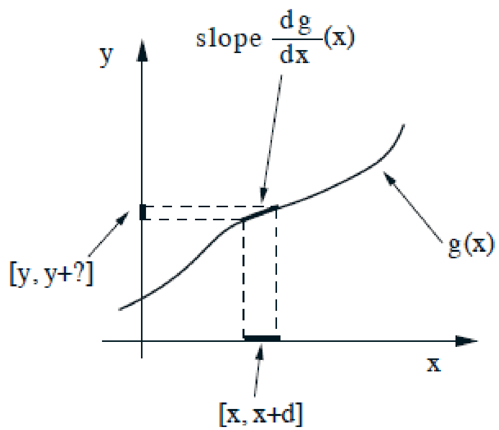
$$f_X(x = \sqrt{y}) = 1, \qquad \left|\frac{dh}{dy}(y)\right| = \frac{1}{2\sqrt{y}}$$

So finally,

$$f_Y(y) = \begin{cases} \dfrac{1}{2\sqrt{y}} & \text{, if } y \in (0,1] \\ 0, \text{otherwise} \end{cases}$$

Quick summarization

- Let $Y = g(X)$, g strictly monotonic



- What is happening?
  - Event $x \le X \le x + \delta$ is the same as $g(x) \le Y \le g(x + \delta)$
  - Or approximately, $g(x) \le Y \le g(x) + \delta \left| \left(\frac{dg}{dx} \right) (x) \right|$

- Hence, it explains the formula (it actually is a mass to mass transfer)

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

where $y = g(x)$

Example

Let X be a random variable with PDF f_X

Find the PDF of the random variable $Y = |X|$

a) When $f_X(x) = \begin{cases} \frac{1}{3}, \text{if } -2 < x \leq 1 \\ 0, \text{ otherwise} \end{cases}$

- Since $Y = |X|$, you can visualize the PDF for any given y as the following:

$$f_Y(y) = \begin{cases} f_X(y) + f_X(-y), \text{if y} \geq 0 \\ 0, \text{if y} < 0 \end{cases}$$

Since $Y = |X|, Y \geq 0$

So when $f_X(x) = \begin{cases} \frac{1}{3}, \text{if } -2 < x \leq 1 \\ 0, \text{otherwise} \end{cases}$

- Essentially $f_X(x)$ for $-1 \leq x \leq 0$ get added to $f_Y(y)$ for $0 \leq y \leq 1$

Therefore,

$$f_Y(y) = \begin{cases} \frac{2}{3}, \text{if } 0 \leq y \leq 1 \\ \frac{1}{3}, \text{ if } 1 < y \leq 2 \\ 0, \text{otherwise} \end{cases}$$

b)

- What if

$$f_X(x) = \begin{cases} 2e^{-2x}, \text{if } x > 0 \\ 0, \text{otherwise} \end{cases}$$

- So we are told that X > 0; hence there is no negative values of X that need to be considered, thus:

$$f_Y(y) = f_X(y) = \begin{cases} 2e^{-2x}, \text{if } x > 0 \\ 0, \text{otherwise} \end{cases}$$

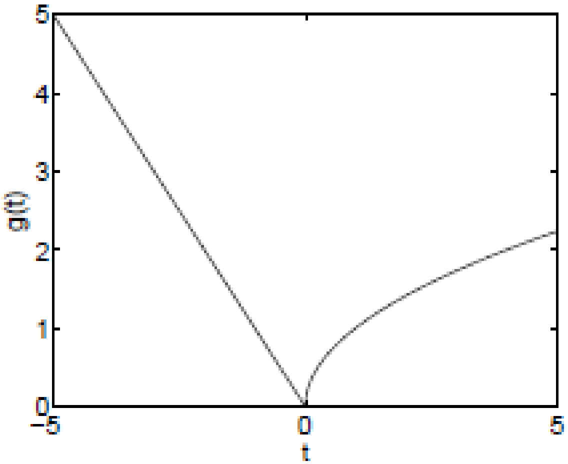
Example

- Let X be a normal random variable with mean 0, variance 1

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

Let $Y = g(X)$

$$g(t) = \begin{cases} -t, \text{for } t \leq 0 \\ \sqrt{t}, \text{for } t > 0 \end{cases}$$



Find PDF of Y

- Because of the definition on g, Y can only take non-negative values

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

$$= P(X \in [-y, 0]) + P(X \in (0, y^2))$$

$$= (F_X(0) - F_X(-y)) + \left(F_X(y^2) - F_X(0)\right)$$

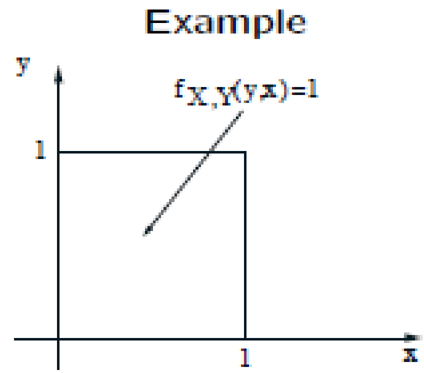
$$= F_X(y^2) - F_X(-y)$$

Think back in the original X space, in what region has mass been moved

- Taking the derivative:

$$f_Y(y) = 2yf_X\left(y^2\right) + f_X(-y)$$

$$= \frac{1}{\sqrt{2\pi}}\left(2ye^{-y^4/2} + e^{-y^2/2}\right)$$



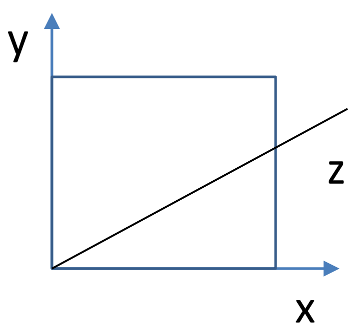
Back to the beginning, Find the PDF of $Z = g(X,Y) = \frac{Y}{X}$

- What is X?
  - Uniform distribution over [0,1]

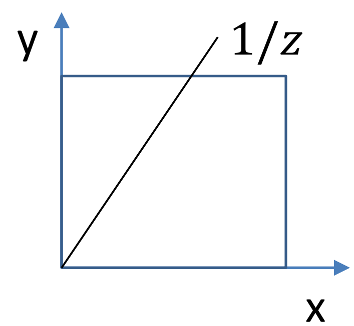
- What is Y?
  - Uniform distribution over [0,1]

- X,Y independent of one another

- So we need to find the CDF of Z, and Z is a ratio between X and Y
  - Can imagine z is essentially a slope term
  - So we need to find $P(Y/X < z)$
    - Split into two cases, and find the area under the curve



Case 1: $0 < z < 1$



Case 2: $z > 1$

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

Area for case 1: $\frac{z}{2}$

Area for case 2: $1 - \left(\frac{1}{(2z)}\right)$

Put it together:

$$= \begin{cases} \frac{z}{2}, \text{if } 0 \leq z \leq 1 \\ 1 - \frac{1}{2z}, \text{if } z > 1 \\ 0, \text{otherwise} \end{cases}$$

- In order to find the PDF, we then differentiate with respect to CDF:

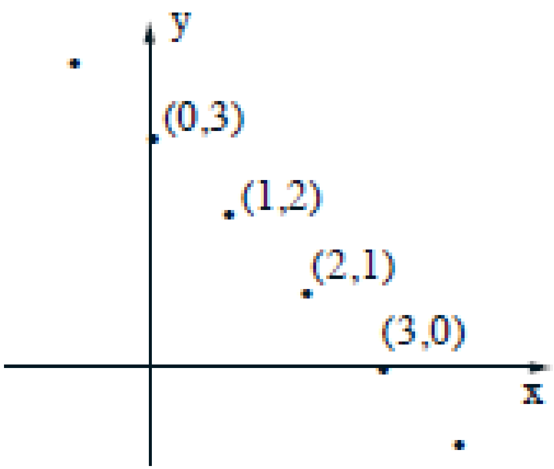
$$f_Z(z) = \begin{cases} \frac{1}{2}, \text{if } 0 \leq z \leq 1 \\ \frac{1}{(2z^2)}, \text{if z} > 1 \\ 0, \text{otherwise} \end{cases}$$

Okay,

So this is the general way of finding a derived distribution, and the steps involved when dealing with multiple random variables are exactly the same as single random variable

The distribution of X+Y

- Before jumping into continuous case, let's take a look at discrete case:
  - $W = X + Y$; X, Y, are independent



Say we want to compute $W = 3$

We look at each possible values of (X, Y) that would make the sum into 3

And add up the probabilities for each of these associations

$$p_W(w) = P(X + Y = w)$$

$$= \sum_x P(X = x)P(Y = w - x)$$

$$= \sum_x p_X(x)p_Y(w - x)$$

What is the mechanics of these?

- Put the two pmf's on top of each other
- Flip the pmf of Y
- Shift the flipped pmf by w (to right if $w > 0$)
- Cross-multiply and add

Continuous case?

- Extremely similar to discrete case, as always..

Say $\mathrm{Z} = X + Y$

First note the following:

$$P(Z \leq z|X = x) = P(X + Y \leq z|X = x)$$

$$= P(x + Y \leq z|X = x) = P(x + Y \leq z)$$

$$= P(Y \leq z - x)$$

If we take differentiation on both sides:

$$f_{Z|X}(z|x) = f_Y(z - x)$$

- Knowing that, then invoke multiplication rule for joint PDF:

$$f_{X,Z}(x,z) = f_X(x)f_{Z|X}(z|x) = f_X(x)f_Y(z-x)$$

Now to find the $f_Z(z)$ -> marginal PDF, integrate out x

$$f_Z(z) = \int_{-\infty}^{\infty} f_{X,Z}(x,z)dx = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx$$

$$p_W(w) = \sum_x p_X(x)p_Y(w-x)$$

Both cases are similar, and intuitions are exactly the same:

This operation is called **convolution!**