STAT 345/445 Lecture 7

Section 2.1: Distributions of Functions of a Random Variable

Functions of random variables

Sometimes we want to transform a random variable.
For example:

- If X is the temperature in Fahrenheit, what is the distribution of the temperature in Celsius?

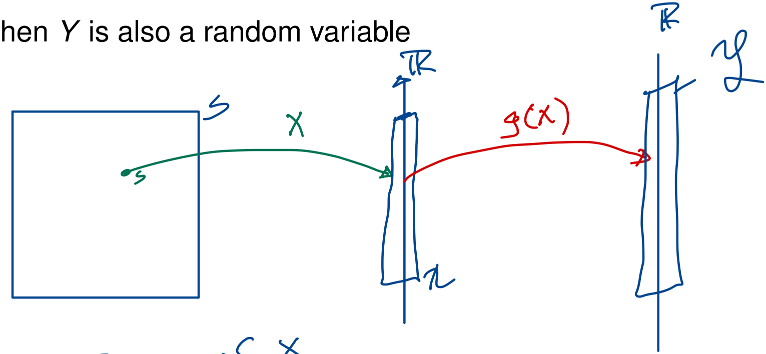
$$Y = (X - 32)\frac{5}{9}$$

- If X and Y denote height and weight, what is the distribution of the BMI?

$$B = \frac{X}{Y^2}$$

Functions of random variables

- Let X be a random variable and let $g(\cdot)$ be a function.
- Then Y is also a random variable



$\mathcal{X}$: Range of X

i.e. $\{x \in \mathbb{R} : \exists\, s \in S \text{ s.t. } X(s) = x\}$

$Y(s) = g(X(s))$

so Y is also a random variable

Functions of random variables

→ cdf or pdf/pmf of Y

- What is the distribution of the random variable $Y = g(X)$?
- Have the cdf $F_X(x)$ or pmf/pdf $f_X(x)$ of X
  - want to find the cdf $F_Y(y)$ or pmf/pdf $f_Y(y)$ of Y.

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

$$= P(X \leq \cancel{g^{-1}(y)})$$

notice that $g^{-1}(y)$ may not be just one value

Let's review *inverse mappings...*

Inverse mapping

- For a function $g(x) : \mathcal{X} \to \mathcal{Y}$ we define an **inverse mapping** as

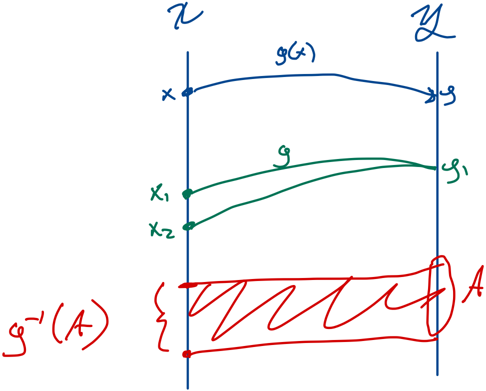
$$g^{-1}(A) = \{x \in \mathcal{X} : \ g(x) \in A\}$$

  for any set $A \subset \mathcal{Y}$
  - Note that $g^{-1}(A) \subset \mathcal{X}$

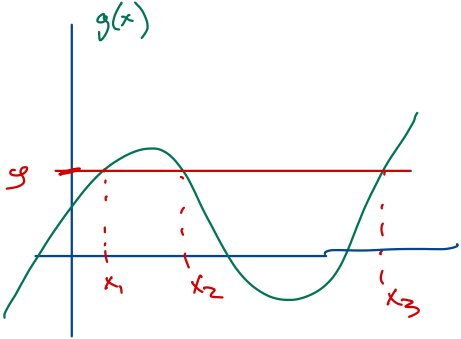
- In particular:

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

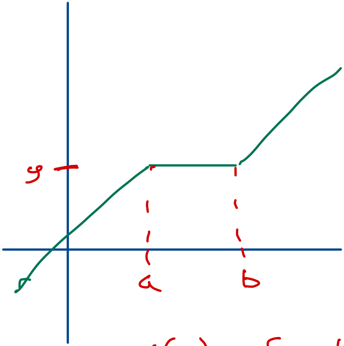
  - Can still be a set in $\mathcal{X}$ rather than just one number
  - Usually just write $g^{-1}(y)$



$$g^{-1}(\{y_1\}) = g^{-1}(y_1)$$
$$= \{x_1, x_2\}$$

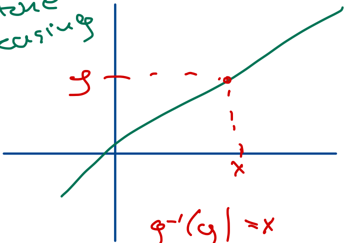


$g^{-1}(y) = \{x_1, x_2, x_3\}$



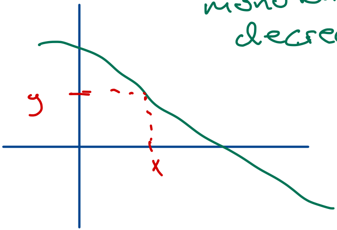
$g^{-1}(y) = [a, b]$

monotone increasing



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

monotone decreasing



More on inverse mapping

- A function $g(x) : \mathcal{X} \to \mathcal{Y}$ is a **one-to-one** function if and only if $\forall\, y \in \mathcal{Y}$ we have
$$g^{-1}(\{y\}) = \{x\}$$
  - Can write $g^{-1}(y) = x$
- *Strictly monotone* functions are one-to-one

Probability of a transformation

- Let X be a random variable in $(S, \mathcal{B}, P)$ and let $Y = g(X)$.
- Probabilities for Y can be obtained from probabilities of X and the inverse mapping $g^{-1}(\cdot)$
- In general $A \subset \mathcal{Y}$ inverse mapping

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

- Will look at discrete and continuous variables separately

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

In particular:

$$F_Y(y) = P(\underbrace{Y \le y}_{Y \in (-\infty, y]}) = P\left(X \in g^{-1}((-\infty, y])\right)$$

Discrete random variables

- Let X be a discrete random variable and let $Y = g(X)$ for some function $g(\cdot)$.
- Then Y is a discrete random variable
- Then

$$f_Y(y) = P(Y=y) = P(g(X)=y)$$
$$= P(X \in g^{-1}(y)) = P(X \in \{x : g(x)=y\})$$
$$= \sum_{x \in g^{-1}(y)} P(X=x) = \sum_{x \in g^{-1}(y)} f_X(x)$$

and

$$F_Y(y) = \sum_{u \le y} f_Y(y)$$

Discrete example 1

- Let $X \sim \text{Binom}(n, p)$, i.e. X has pmf

$p \in [0,1]$

$$f_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{for } x \in \{0, 1, 2, \ldots, n\} \\ 0 & \text{otherwise} \end{cases}$$

- X can be thought of as the number of successes in n independent Bernoulli trials
- What is the distribution of $Y = n - X$?

(number of failures instead number of successes)

whiteboard...

Discrete example 2

- Let $X \sim \text{Binom}(10, p)$. What is the distribution of $Y = |X - 5|$?

  on the whiteboard

~~Continuous random variables~~ Monotone transf.

- It's easiest to deal with *monotone* functions g:

$$\text{Increasing: } u > v \Rightarrow g(u) > g(v) \quad \forall u, v$$
$$\text{Decreasing: } u > v \Rightarrow g(u) < g(v)$$

- The **support** of a distribution (or random variable) is defined as

$$\mathcal{X} = \{x : f_X(x) > 0\} \tag{1}$$
$$\text{and let } \mathcal{Y} = \{y : \exists x \in \mathcal{X} \text{ such that } g(x) = y\} \tag{2}$$

support of Y support of X

- If g is monotone on $\mathcal{X}$ then it is *one-to-one* and *onto* from $\mathcal{X}$ to $\mathcal{Y}$.
  - Uniquely pairs an x to one y
  - Get an inverse function: $g^{-1}(y) = x$

cdf – method

Theorem ("cdf-method")

Let X be a random variable with cdf $F_X(x)$ and let $Y = g(X)$. Then

(a) If g is an increasing function on $\mathcal{X}$ then

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

(b) If g is a decreasing function on $\mathcal{X}$ *and* X is continuous , then

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

In general: i.e. for both discrete and contiuous

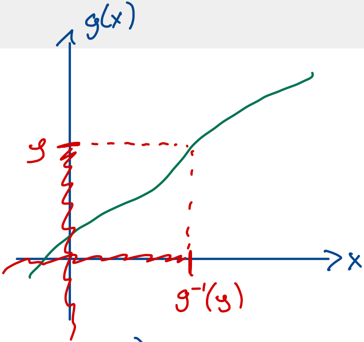
$$F_Y(y) = 1 - F_X(g^{-1}(y)) + P(X = g^{-1}(y))$$

cdf – method, proof

(a) If g is an increasing function on $\mathcal{X}$ then

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

$Y \le y \;\Rightarrow\; X \le g^{-1}(y)$



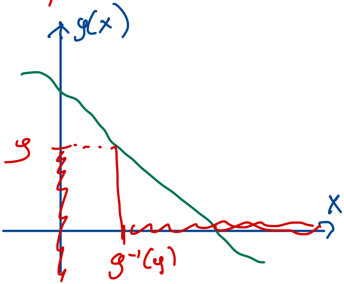
(b) If g is a decreasing function on $\mathcal{X}$ *and* X is continuous, then

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

In general: i.e. for both discrete and continuous

$$F_Y(y) = 1 - F_X(g^{-1}(y)) + P(X = g^{-1}(y))$$

$Y \le y \;\Rightarrow\; X \ge g^{-1}(y)$



Example: Exponential and Weibull

Let $X \sim \text{Expo}(1)$, i.e.

$$F_X(x) = \begin{cases} 0 & , \; x < 0 \\ 1 - e^{-x} & , \; x \geq 0 \end{cases}$$

Let $Y = X^\alpha$ for $\alpha > 0$. What is the distribution of Y?

on whiteboard

pdf – method

Theorem ("pdf method")

Let X be a continuous random variable with pdf $f_X(s)$ and let $Y = g(X)$ where g is a *monotone* function.

Suppose that $f_X(x)$ is continuous on $\mathcal{X}$ and that $g^{-1}(y)$ has a continuous derivative on $\mathcal{Y}$.

Then the pdf of Y is given by

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

$$0 \text{ otherwise}$$

proof on whiteboard

Example: Exponential and Uniform

Let $X \sim \text{Expo}(1)$, i.e.

$$f_X(x) = \begin{cases} e^{-x} & , \; x \geq 0 \\ 0 & , \; x < 0 \end{cases}$$

Let $Y = 1 - e^{-X}$. What is the distribution of Y? on whiteboard

Note:

$$F_X(x) = \begin{cases} 1 - e^{-x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

i.e. $g(x) = F_X(x)$

Probability integral transformation

Theorem

Let X have a continuous cdf $F_X(x)$ and let $Y = F_X(X)$. Then Y is uniformly distributed on $(0, 1)$, i.e.

$$F_Y(y) = \begin{cases} 0 & , \ y \leq 0 \\ y & , \ 0 < y < 1 \\ 1 & , \ y \geq 1 \end{cases}$$

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

and $Y \sim \text{Uniform} \quad \Rightarrow \quad X = F_X^{-1}(Y) \sim F_X$

When g is monotone only on certain intervals

- See Theorem 2.1.8 for more detail
- If $\mathcal{X}$ support of X can be split into sets $A_1, \ldots, A_k$ and g can be split into $g_1(x), \ldots, g_k(x)$ such that
  - $g(x) = g_i(x)$ for $x \in A_i$
  - g_i is a monotone function from A_i *onto* $\mathcal{Y}$

  then

$$f_Y(y) = \begin{cases} \sum_{i=1}^{k} f_X\left(g_i^{-1}(y)\right) \left| \frac{d}{dy} g_i^{-1}(y) \right| & , \ y \in \mathcal{Y} \\ 0 & , \ \text{otherwise} \end{cases}$$

Example: Standard normal and χ^2 distribution

Let $X \sim N(0,1)$, i.e.

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \qquad \text{for } x \in \mathbb{R}$$

Let $Y = X^2$. What is the distribution of Y?

$Y \sim \chi^2_1$ (Chi-square distr. with 1 d.f.)

on whiteboard