Random Variables

Stat 510

DEFN

A RANDOM VARIABLE (RV) IS A MAPPING

$$\underbrace{X}_{\text{RV}} : \underbrace{\Omega}_{\text{SAMPLE SPACE}} \longrightarrow \underbrace{\mathbb{R}}_{\text{NUMBERS}}$$

ASSIGNS NUMBER $X(\omega)$

TO EACH OUTCOME $\omega \in \Omega$

FLIP TWO COINS   DEFINE X = # HEADS OBSERVED

ω	$X(\omega)$
HH	2
HT	1
TH	1
TT	0

SAMPLE SPACE

RV

PLACEHOLDER FOR POSSIBLE VALUE

x	$\mathbb{P}[X = x]$
0	0.25
1	0.50
2	0.25

NOTATION

$$\mathbb{P}(\{\omega \in \Omega ; X(\omega) = x\})$$

Defn

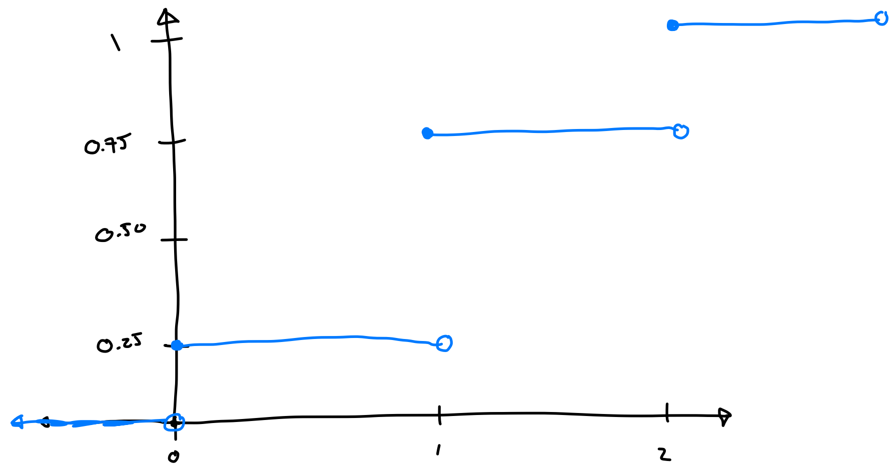
Cumulative Distribution Function CDF

$$\underset{\text{big}}{F_X}(\overset{\text{small}}{x}) : \underset{\#}{\mathbb{R}} \to \underset{\text{prob}}{[0,1]}$$

$$F_X(x) = \mathbb{P}(X \le x)$$

Example

x	$P[X=x]$
0	0.25
1	0.50
2	0.25



$$F_X(x) = P[X \le x]$$

$$F(0.5) = 0.25$$

$$F(-1) = 0$$

$$F(1) = 0.75$$

Thm — Equal in distribution

$X \sim F$ — CDF of F

$Y \sim G$ — CDF of G

IF $F(x) = G(x)$ for all x, THEN

$$P(X \in A) = P(Y \in A) \text{ for all } A.$$

WRITE $X \overset{d}{=} Y$

DOES NOT MEAN $X = Y$.

CONSIDER $P(X = 1) = P(X = -1) = 0.5$

DEFINE $Y = -X$.

$Y \overset{d}{=} X$

BUT $Y = 1$ WHEN $X = -1$

AND $Y = -1$ WHEN $X = 1$

THM

A FUNCTION $F : \mathbb{R} \to [0,1]$ IS A <u>CDF</u> IF

1. $x_1 < x_2 \Rightarrow F(x_1) \le F(x_2)$ NON-DECREASING

2. $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to \infty} F(x) = 1$ NORMALIZED

3. $F(x) = \lim_{\substack{y \to x \\ y > x}} F(y)$ RIGHT CONTINUOUS

DEFN

X IS DISCRETE IF IT TAKES COUNTABLY MANY VALUES.

PROBABILITY MASS FUNCTION (PMF) OF X

$$f_X(x) = P[X = x]$$

DEFN

RV X IS CONTINUOUS IF THERE IS A FUNCTION f_X SUCH THAT

- $f_X(x) \geq 0$ FOR ALL x
- $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$
- $\mathbb{P}[a < X < b] = \int_a^b f_X(x)\,dx$ FOR $a \leq b$

WE CALL $f_X(x)$ THE PROBABILITY DENSITY FUNCTION

CDF

PDF

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx$$

PDF

$$f_X(x) = F_X'(x)$$

AT POINTS x WHERE $F(x)$ IS DIFFERENTIABLE

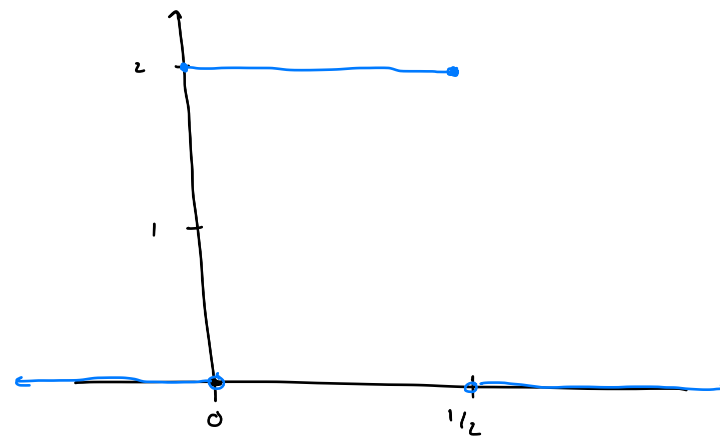
CONTINUOUS RV X

- $\mathbb{P}[X = x] = 0$
- $f(x) > 1$ IS ALLOWED

EX

$$f_X(x) = \begin{cases} 2 & \text{FOR } 0 \le x \le 1/2 \\ 0 & \text{OTHERWISE} \end{cases}$$

SOMETIMES IMPLIED



$$P[X = 0.25] = \int_{0.25}^{0.25} 2\,dx = 0 \checkmark$$

Lemma

F CDF OF X

- $\mathbb{P}[X = x] = F(x) - F(x^-)$ $\qquad F(x^-) = \lim_{y \uparrow x} F(y)$

- $\mathbb{P}[x < X \leq y] = F(y) - F(x)$

- $\mathbb{P}[X > x] = 1 - F(x)$

- IF X CONTINUOUS

$$\left.\begin{aligned} F(b) - F(a) &= \mathbb{P}[a < X < b] = \mathbb{P}[a \leq X < b] \\ &= \mathbb{P}[a < X \leq b] = \mathbb{P}[a \leq X \leq b] \end{aligned}\right\} \mathbb{P}[X = x] = 0$$

NAMED RANDOM VARIABLES

- TABLE IN BOOK
- WIKIPEDIA

$$X \sim F$$

"HAS DISTRIBUTION"

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

NORMAL DISTRIBUTION

PARAMETERS

DEFINES "FORM" OF DISTRIBUTION

DEFINES SPECIFIC DISTRIBUTION

DISCRETE	CONTINUOUS
POINT MASS	UNIFORM
BERNOULLI	NORMAL
BINOMIAL	EXPONENTIAL
GEOMETRIC	GAMMA
POISSON	BETA
	t
	χ^2
	F
	CAUCHY

Binomial Distribution

$$X \sim \text{Binom}(n, p)$$

pmf

$$f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \ldots, n, \quad 0 \le p \le 1$$

"sample space" — $x = 0, 1, 2, \ldots, n$

Parameter space — n, $0 \le p \le 1$

with $n = 20$, $p = 0.3$

$$f_X(x) = \binom{20}{x} 0.3^x \, 0.7^{20-x}$$

$$x = 0, 1, 2, \ldots, 20$$

- Additional info on Wikipedia

Normal Distribution

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

PARAMETERS

PDF

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \,\mathrm{Exp}\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \quad -\infty < x < \infty, \quad -\infty < \mu < \infty, \quad \sigma > 0$$

Probability in R

- PDF / PMF d*(· , parameters)
- CDF p*(· , parameters)
- Quantile q*(· , parameters)
- Random r*(· , parameters)

WARNING!

READ DOCS!

TRANSFORMATIONS

GIVEN RV X, WHAT IS DISTRIBUTION OF...

- $Y = X^2$?
- $Z = \frac{X-\mu}{\sigma}$?
- $Y = e^X$?
- $Y = r(X)$?

THREE METHODS

- DISTRIBUTION FUNCTION
- CHANGE-OF-VARIABLES
- MGF

KNOWN RELATIONSHIPS