Lecture 32, Part 1: April 12, 2021

- Finishing up chapter 9 (bootstrapped C.I)
- Prob. density functions.

Margin of error

$$\bar{X} \pm z_{\frac{\alpha}{2}} SD(\bar{X})$$

$$= \bar{X} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

- We have a confidence interval. Now we want to keep the **same confidence level**, but want to improve our accuracy. For example, say our *margin of error* is 4 percentage points, and we want it to be 1 percentage point. What should we do?

new $n = 16n$

$\sqrt{n_{new}} = 4\sqrt{n}$

A. increase width of CI 4 times by increasing SD

B. Decrease width of CI by increasing *n* by 4 times

C. Decrease width of CI by increasing *n* by 16 times $\rightarrow$ b/c of the square root law

$z_{\frac{\alpha}{2}}$ won't change $\rightarrow$ depends on the confidence level.

σ won't change, it is a constant associated with the population

Comparison with bootstrap CI

B= # of bootstrap samples =10,000

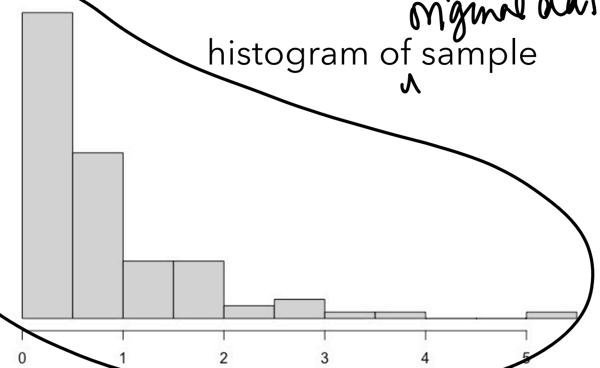
(sampling our sample w/repl., every sample size is the same as the original sample

- How do you create a bootstrap CI for the population mean?

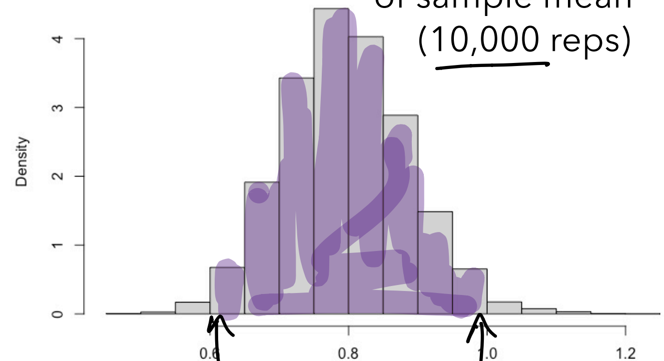
n=100

original data

histogram of sample



Bootstrapped histogram of sample mean (10,000 reps)



histogram of 10,000 sample means.

Took 10,000 samples & computed the samplemean each time.

~~take~~ take one resample of size 4 w/replacement

sample

(2, 4, 6, 8) → 2, 4, 2, 2 1st bootstrap sample

$\bar{x}$= 2.5

2, 6, 4, 2
$\bar{x} = 7/2 = 3.5$

~~Probability density functions~~

histogram represents dsn of sample mean $\bar{X}$
derived "empirically"

Want the values on the horizontal axis
that bound the middle 95%.

2.5th percentile & 97.5th percentile

250th value in sorted list → 0.628
9750th values → 0.976

95% C.I from bootstrapping (NO CLT) (0.628, 0.976)

Using CLT: From an original sample, ~~observed~~
value of $\bar{X}$ is 0.796, $SD(\bar{X}) \approx \frac{0.867}{10}$ (n=100)

Margin of error $= 1.96 * \frac{0.887}{10}$

95% C.I. using CLT : $0.796 \pm 1.96 * \frac{0.887}{10}$

$(0.622, 0.97)$

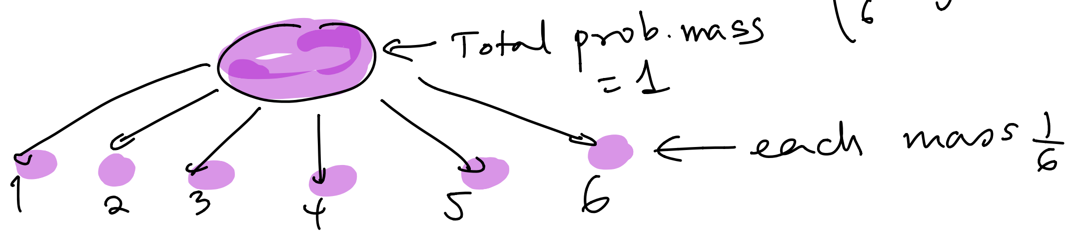
Bootstrap CI $(0.63, 0.98)$

Lecture 32 — Probability Density

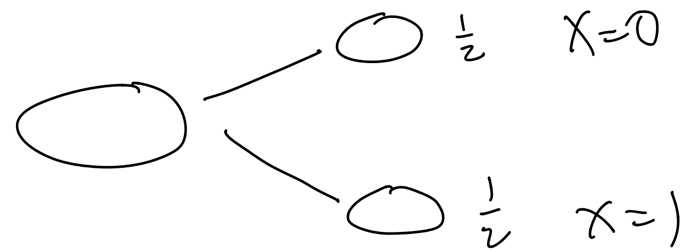
Discrete r.v.

Rolling a die, $X = \#$ of spots

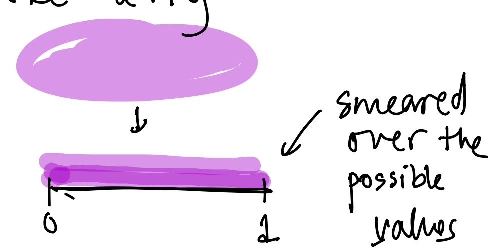
$$X = \left\{ \begin{matrix} 1 \\ 2 \\ 3 \\ \vdots \\ 6 \end{matrix} \right\} \text{ w.p. } \frac{1}{6}$$



Toss a coin :



Say instead X can take any value between 0 & 1



Instead of adding & sums we integrate .

Instead a prob. mass function (p.m.f)

you have a prob. density function

Such r.v. that can take <u>any</u> values in an interval on the real line are called

CONTINUOUS RANDOM VARIABLES.

X, $f(x) \leftarrow$ p.d.f. , $F(x) \leftarrow$ c.d.f

DENSITY — DISTRIBUTION

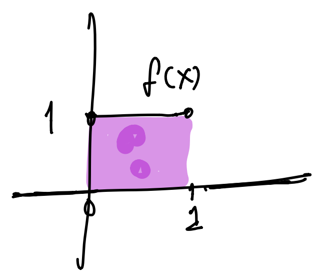
The pdf f has to satisfy:

1. $f(x) \geq 0$
2. $\displaystyle\int_{-\infty}^{\infty} f(x)\,dx = 1$ — Area under the curve $y = f(x)$

$F(x)$ = cumulative distribution function

$$= \int_{-\infty}^{x} f(t)\,dt.$$

Examples.
$$f(x) = \begin{cases} 0, & x \leq 0 \\ 1, & 0 < x < 1 \\ 0, & x \geq 1 \end{cases}$$



$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^1 1 \cdot dx = 1$$

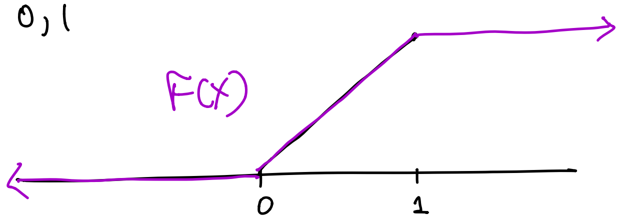
(or you could use the fact that it is a 1×1 square.)

The r.v. X that has this density function is called the Uniform r.v.

$X \sim U(0,1)$ ("X has the Uniform distribution on $(0,1)$")

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{0}^{x} 1\cdot dt = x$$

let x be b/w 0, 1



$$F(1) = \int_{-\infty}^{1} f(t)\,dt = \int_{0}^{1} f(t)\,dt = \int_{0}^{1} 1\cdot dt = 1$$

Examples Verify that f(x) are p.d.f. Find expressions for F(x)

① $$f(x) = \begin{cases} 0, & x \le 0 \\ 2x, & 0 < x < 1 \\ 0, & x \geq 1 \end{cases}$$

Also draw the graphs

② $$f(x) = \begin{cases} 0, & x < -1 \\ x+1, & -1 \le x \le 0 \\ 1-x, & 0 < x < 1 \\ 0, & x \geq 1 \end{cases}$$

③ $$f(x) = \begin{cases} e^{-x}, & x > 0 \\ 0, & \text{o/w.} \end{cases}$$ special fn.