Agenda for today:

- Review of expectation
- Review of sample means and sample variability

} 20 min

- SQL (Allen) } 20-30 min

Expectation

$$E[X] = \sum_x x \cdot P(X = x)$$

value of the random variable

probability of that value

x	$P(X=x)$
2	1/4
3	1/2
4	1/4

$$E[X] = 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{2} + 4 \cdot \frac{1}{4}$$
$$= \frac{1}{2} + \frac{3}{2} + 1 = \boxed{3}$$

$$E[E[X]] = E[X]$$

Note: $E[X]$ is a constant!

$$E[c] = c$$

$E(X)$

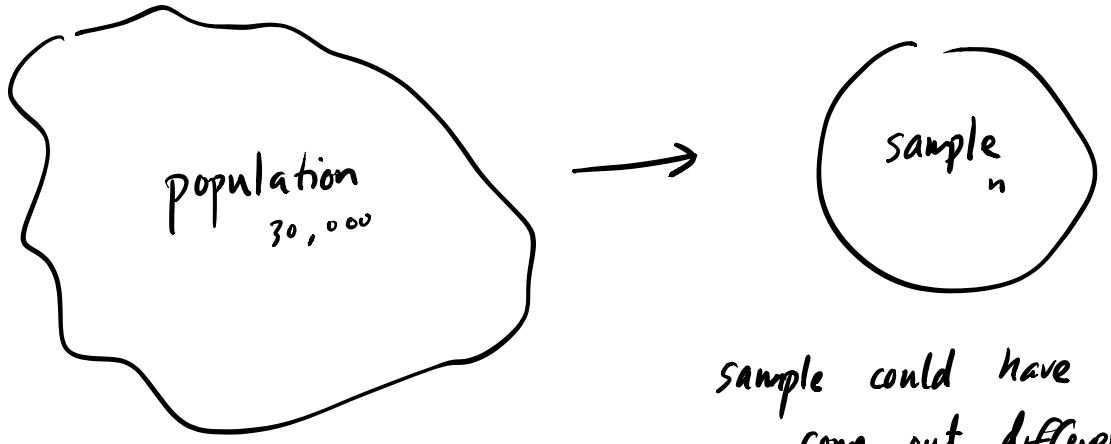
$E(X^2)$

$E(X)^2$

$Var(X)$

constants!

Sample Means



sample could have come out differently!

true average height: μ
SD of height: σ
$\geq$ population properties

→ The population could follow any distribution! We don't know its distribution (other than its mean and SD), but it doesn't matter.
→ μ is fixed! not random.

sample : {Person 1, Person 2, ..., Person n}

X_i = height of person i in our sample

X_i is a random variable! function of numerical sample.

⇒ $X_1, X_2, \ldots, X_n$ are i.i.d.

independent and identically distributed
→ because sampling with replacement

$E[X_i] = \mu$

$SD[X_i] = \sigma$

$$\bar{X}_n = \frac{X_1 + X_2 + \ldots + X_n}{n}$$ sample mean

$\bar{X}_n$ → what is this? **random variable** ← many students got this wrong on the Q.C.

$E[\bar{X}_n] = \mu$ expected value of sample mean is equal to the population mean

$SD[\bar{X}_n] = \frac{\sigma}{\sqrt{n}}$ square root law

Note (didn't cover live): Two random variables are identically distributed if they take on the exact same values with the same probabilities.

⇒ Just because two R.V.s have the same $E[X]$ and $SD[X]$, doesn't mean they are identically distributed!

⇒ e.g. $X \sim \text{Binomial}(10, \frac{1}{2})$, $y \sim \text{Normal}(5, \frac{5}{2})$

$E[X] = E[y] = 5$, $SD[X] = SD[y] = \sqrt{\frac{5}{2}}$ but X, y don't have the same dist!