Expectation and Variance (Ch. 3.1 ER) Skim 3.2

Sample average $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$ where X_i iid

If $X_i \sim \text{Bernoulli}(q)$

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i \approx \frac{\#\{X_i = 0\}}{N} \times 0 + \frac{\#\{X_i = 1\}}{N}$$

$$\approx (1-q) \times 0 + q \times 1 = q$$

In general

$$\bar{X} = \sum_{x \in S_X} x \frac{\#\{X_i = x\}}{N} \approx \sum_{x \in S_X} x P(X = x)$$

This motivates the definition of expectation

Expected Value (Def 3.1.1/3.1.2)

$$\mathbb{E}[X] = \sum_{x \in S_X} x P(X = x)$$

Draft

Example

$$X \sim \text{Bernoulli}(q) \implies \mathbb{E}[X] = q$$

$$X \sim \text{Geometric}(q) \implies \mathbb{E}[X] = 1/q$$

does this make intuitive sense? Can you test with simulations?

[Note in 3.1.8 they use a slightly different definition of Geometric where $S = \{0, 1, \ldots\}$ not $S = \{1, 2, \ldots\}$ and get $\mathbb{E}[X] = (1-q)/q$]

Can skip Ex. 3.1.9 – 3.1.13

Properties of Expectation

1) Expectation of function (theorem 3.1.1)

If X is r.v. and $g : S_X \to \mathbb{R}$ is a function, then

$$\mathbb{E}[g(X)] = \sum_{x \in S_X} g(x) P(X = x)$$

Key point here is that we don't need to determine probabilities of $g(X)$ to determine $\mathbb{E}[g(X)]$

2) Expectation is linear (Theorem 3.1.2)

X, Y r.v. then $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

Example (Binomial Ex. 2.3.3)

watch 1Brown3Blue!

Let $X_i \sim \text{Bernoulli}(q)$ $i = 1, 2, 3, \ldots, N$

and assume independent!

note: two parameters

Let $S = \sum_{i=1}^{N} X_i$ then $S \sim \text{Binomial}(q, N)$

S is Model of many things

- Votes in election
- S/N is sample mean of Bernoulli trials
- responses on Y/N survey
- Positive tests in clinical trial
- Successful free throws

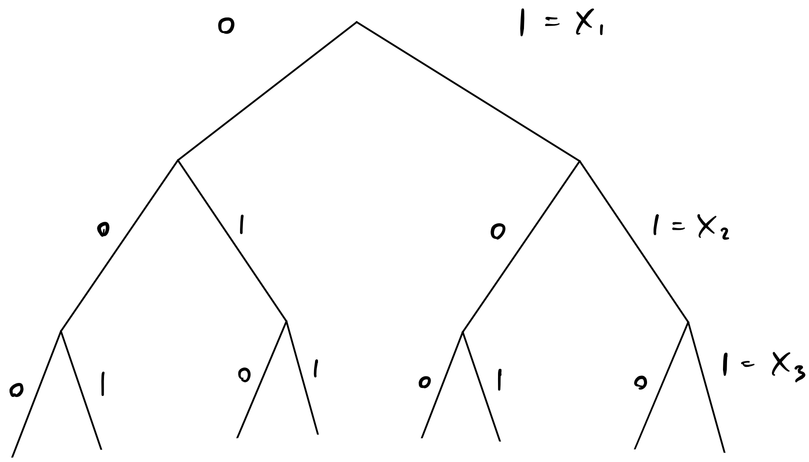
By linearity $\mathbb{E}[S] = \sum_{i=1}^{N} \mathbb{E}[X_i] = Nq$

I will also discuss conditional expectation

(Def 3.5.1)

skip sec 3.5.2

What about distribution? Take $N=3$



$S = \;\; 0 \quad 1 \quad 1 \quad 2 \quad 1 \quad 2 \quad 2 \quad 3$

Prob $= \;\; (1-q)^3 \quad (1-q)^2 q \quad (1-q)^2 q \quad (1-q)q^2 \quad (1-q)^2 q \quad (1-q)q^2 \quad (1-q)q^2 \quad q^3$

By theorem 1.5.1,

# of sequences w/ 2 1s

chance to get any particular sequence of flips w/ 2 1s

$$P(S=2) = 2 \times (1-q)\,q^2$$

In general

sequences w/ k heads (don't need to know formula)

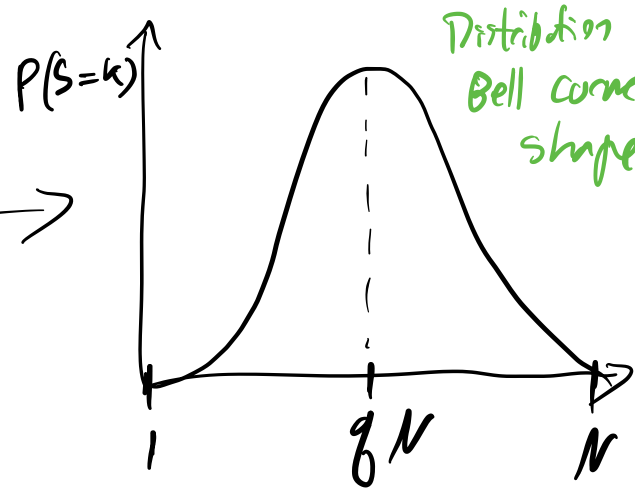
chance to get sequence w/ k heads

$$P(S=k) = \binom{N}{k}(1-q)^{N-k} q^k$$

(see Ch. 1 for details, don't need to know)

Key point:

many more ways to get $K \approx N/2$ heads then $K \approx 0$ or $K \approx N$ $\longrightarrow$



Distribution has Bell curve shape

Useful to quantify how "spread out" distribution is

Variance and Covariance (Ch 3.3 ER)

To measure spread of distribution we look at how much, on avg., it deviates from the mean

Variance (Def 3.1.1)

$$\mathrm{Var}(X) = \mathbb{E}\left[\left(X - \mathbb{E}[X]\right)^2\right]$$

Example $X \sim \mathrm{Bernoulli}(q)$

$$\Longrightarrow \mathbb{E}\left[(X-q)^2\right] = \mathbb{E}[X^2] - 2\mathbb{E}[X]q + q^2$$

$$= 0\cdot(1-q) + 1\cdot q^2 - 2q + q = q^2 - q = q(1-q)$$

$\longrightarrow$ X is least "variable" when $q \approx 0$ or $q \approx 1$

$S \sim \text{Binomial}(q) \implies \text{Var}(S) = Nq(1-q)$

I will explain more

<u>Problem</u> $\text{Var}(X)$ has different units than X

then $X \longrightarrow$ motivates Def 3.3.2

<u>Covariance (Def 3.3.3)</u>

$$\text{Cov}(X, Y) = \mathbb{E}\left[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\right]$$

I will also discuss thm 3.3.1, 3.3.3, 3.3.2