Math 170E: Winter 2023

Lecture 10, Wed 1st Feb

The binomial distribution

Last time:

- Let X be a discrete random variable taking values on a countable set $S \subseteq \mathbb{R}$
- We define its expected value to be
$$\mu_X = \mathbb{E}[X] = \sum_{x \in S} x p_X(x)$$
- If $g : S \to \mathbb{R}$ is a function, the expected value of $g(X)$ is
$$\mathbb{E}[g(X)] = \sum_{x \in S} g(x) p_X(x)$$
- Given $p \in (0, 1)$, $X \sim$ Bernoulli(p) if it has PMF
$$p_X(x) = \begin{cases} p & \text{if } x = 1 \\ 1 - p & \text{if } x = 0 \end{cases}$$
- We define the MGF $M_X(t)$, for $t \in \mathbb{R}$ by
$$M_X(t) = \mathbb{E}[e^{tX}].$$
- For any $r = 1, 2, 3, \ldots$, if M_X is well-defined and smooth around $t = 0$, then
$$\left.\frac{d^r}{dt^r} M_X\right|_{t=0} = \mathbb{E}[X^r].$$
- we defined the variance by var$X = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$.

Today:

We'll discuss today:

- the Binomial distribution and how to compute with it

Flip n fair coins. $1 \leq k \leq n$

$$\mathbb{P}(\text{exactly } k \text{ heads}) = \frac{|\{\text{choosing } k \text{ objects out of } n\}|}{2^n} = \frac{\binom{n}{k}}{2^n}$$

$X = \#\text{heads}$

$$\mathbb{P}(\text{exactly } k \text{ heads}) = \mathbb{P}(X=k) = P_X(k).$$

Definition 2.17:

- A Bernoulli trial is an experiment that has a probability $p \in (0, 1)$ of success and probability $(1 - p)$ of failure.
- Suppose we run $n \geq 1$ *independent, identical* Bernoulli trials.
- Let X be the number of successes. random variable.
- X takes values in $S = \{0, 1, \ldots, n\}$. discrete
- We say that X is a Binomial random variable with parameters n, p and write $X \sim \text{Binomial}(n, p)$

Total # of trials

prob of each success.

Proposition 2.18:

If $X \sim \text{Binomial}(n, p)$, then it has PMF

$$p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad \text{if } x \in \{0, 1, \ldots, n\}.$$

$p = 1/2$: $p_X(x) = \binom{n}{x}\left(\frac{1}{2}\right)^x\left(1-\frac{1}{2}\right)^{n-x}$

$X \sim Bin(n, 1/2)$ $= \binom{n}{x} \underbrace{\frac{1}{2^x}\,\frac{1}{2^{n-x}}}_{\frac{1}{2^n}}$.

~~Proof:~~ Fix $x \in \{0, 1, \ldots, n\}$. $p_X(x) = \mathbb{P}(X = x)$.

$\underbrace{\{1, 1, 1, 0, 0, 1, \ldots, 0\}}$.

$\hookrightarrow$ x-successes.

$\leftarrow$ "= probab. of exactly x successes out of n".

- $\binom{n}{x}$ ways of choosing the x-successes out of n trials
- x successes occur with prop. $\underbrace{p \times \cdots \times p}_{x} = p^x$
- $n-x$ failures ——— $\underbrace{(1-p) \times \cdots \times (1-p)}_{n-x} = (1-p)^{n-x}$

$$\sum_{x=0}^{n} P_X(x) = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x}.$$

$$= (p + 1 - p)^n = \underline{1}.$$

So P_X is a genuine PMF.

<u>Binomial Th$^{\underline{m}}$</u>

$$(y+z)^n = \sum_{x=0}^{n} \binom{n}{x} \times y^x z^{n-x}.$$

Example 11: Let's play Chuck-a-Luck!

- I roll 3 fair dies
- You pick a number in $\{1, 2, 3, 4, 5, 6\}$
- If your number comes up:
  - once, you win \$10
  - twice, you win \$20
  - thrice, you win \$30
- If your number doesn't come up, you owe me \$10
- Do you play?
- i.e. What is ~~my~~ your expected win per wager? $\to \mathbb{E}[Y]$?

$\hookrightarrow$ Let X be the number of times your # comes up

$$X \in \{0, 1, 2, 3\}.$$

Each die, comes up with your # with prob $1/6$.
does not come up —— $5/6$.

$\hookrightarrow X \sim \text{Binomial}(3, 1/6)$.

$$\mathbb{P}(X=x) = \binom{3}{x}\left(\frac{1}{6}\right)^x \left(1-1/6\right)^{3-x} = \binom{3}{x}\frac{5^{3-x}}{6^3}$$

for all $x \in \{0,1,2,3\}$.

Let Y be your winnings on a given game.

$Y \in \{-10, 10, 20, 30\}$.

If $Y = -10 \iff X = 0$. $\hookrightarrow p_Y(-10) = p_X(0)$.

$Y = 10 \iff X = 1$

$Y = 20 \iff X = 2$

$Y = 30 \iff X = 3$.

$$\mathbb{E}[Y] = \sum_{y \in \{-10, 10, 20, 30\}} y\, p_Y(y)$$

$-c, c, 2c, 3c$.

$$= -10P_Y(-10) + 10P_Y(10) + 20P_Y(20) + 30P_Y(30).$$

$$= -10P_X(0) + 10P_X(1) + 20P_X(2) + 30P_X(3).$$

$$= -10\binom{3}{0}\frac{5^3}{6^3} + 10\cdot\binom{3}{1}\frac{5^2}{6^3} + 20\cdot\binom{3}{2}\frac{5}{6^3} + 30\cdot\binom{3}{3}\frac{5^0}{6^3}.$$

$\vdots$

$$= -0.787.$$

On average, you lose ~\$0.79 per game.

Not a fair game.

Proposition 2.18: If $X \sim \text{Binomial}(n, p)$, then its MGF is

$$M_X(t) = (1 - p + pe^t)^n. \quad \text{for all } t \in \mathbb{R}.$$

Proof:

$$M_X(t) = \mathbb{E}[\underbrace{e^{tX}}_{g(X)}] = \sum_{x=0}^{n} e^{tx} p_X(x).$$

$$e^{ab} = (e^a)^b.$$

$$= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x (1-p)^{n-x}.$$

$$= \sum_{x=0}^{n} \binom{n}{x} \boxed{e^{tx} p^x} (1-p)^{n-x}.$$

$$= (e^t p)^x$$

$$= \sum_{x=0}^{n} \binom{n}{x} (e^t p)^x (1-p)^{n-x}.$$

Binomial Th =

$$= (1 - p + pe^t)^n.$$

$$\mathbb{E}[X] = \sum_{x=0}^{n} \binom{n}{x} x p^x (1-p)^{n-x}.$$

Proposition 2.19: If $X \sim \text{Binomial}(n, p)$, then

$$\mathbb{E}[X] = np.$$

$M_X(t) = (\text{—})^n$

$\log M_X(t) = n \log(\text{—})$.

Proof:

$$\hookrightarrow \mathbb{E}[X] = \frac{d}{dt} M_X(t) \Big|_{t=0}$$

$$= \frac{d}{dt} (1-p+pe^t)^n \Big|_{t=0}$$

$$= npe^t (1-p+pe^t)^{n-1} \Big|_{t=0}$$

$$= np(1-p+p)^{n-1} = np.$$

Also:

$$\mathbb{E}[X] = \frac{d}{dt} \log M_X(t) \Big|_{t=0}.$$

$$= \frac{d}{dt} n \log(1-p+pe^t) \Big|_{t=0}.$$

$$= n \frac{d}{dt} \log(1-p+pe^t) \Big|_{t=0} = np.$$

Proposition 2.20: If $X \sim \text{Binomial}(n, p)$, then

$$\text{var}(X) = np(1-p).$$

Proof:

$$\text{Var}(X) = \underbrace{\mathbb{E}[X^2]}_{\text{computed as}} - \mathbb{E}[X]^2 .$$

$$\mathbb{E}[X^2] = \frac{d^2}{dt^2} M_X(t) \Big|_{t=0}.$$

$$= \frac{d}{dt}\left[npe^t(1-p+pe^t)^{n-1} \right]\Big|_{t=0}.$$

$$= np \frac{d}{dt}\left[e^t(1-p+pe^t)^{n-1} \right]\Big|_{t=0}$$

$$= np\left\{ e^t(1-p+pe^t)^{n-1} + e^t(n-1)\cdot pe^t(1-p+pe^t)^{n-2} \right\}\Big|_{t=0}$$

$$= np\left\{ (1-p+p)^{n-1} + (n-1)p(1-p+p)^{n-2} \right\}$$

$$= np\left\{ 1 + p(n-1) \right\}.$$

$$\mathrm{Var}(X) = np(1+p(n-1)) - n^2p^2$$

$$= np + np^2(n-1) - n^2p^2 \qquad \sum x^2 \binom{n}{x} p^x (1-p)^{n-x}$$

$$= np + \cancel{n^2p^2} - np^2 \cancel{- n^2p^2}$$

$$= np(1-p).$$

$$\hookrightarrow \mathrm{Var}(X) = \frac{d^2}{dt^2} \log M_X(t) \Big|_{t=0}.$$

Example 12: The power of a voter

- In a U.S. election, the candidate with the most votes in a state wins all the electoral college votes
- Let's suppose the number of electoral college votes is proportional to the population of the state
  - i.e. a state with population n has roughly cn electoral college votes
- We define the *average power of a citizen in a close election* in a state of $n = 2k + 1$ as the average number of electoral college votes a voters vote impacts if each of the other $n - 1 = 2k$ voters split their votes evenly between the two candidates
- the election is close: the other $2k$ voters vote independently and equally likely for either candidate
- What is the average power of a citizen in a close election?

The probability that a voter will make a difference
(i.e. choose the winning candidate)
is the same as the probability that if we flip

2k fair coins we have exactly k heads & k tails.

$\hookrightarrow \mathrm{Bin}(2k, 1/2)$.

$\hookrightarrow \mathbb{P}\left(\begin{array}{c}\text{your vote makes}\\ \text{a diff. in a state with}\\ n=2k+1 \text{ voters}\end{array}\right) = \binom{2k}{k}\frac{1}{2^k}\frac{1}{2^{2k-k}} = \binom{2k}{k}\frac{1}{2^{2k}}$

$$= \frac{(2k)!}{(k!)^2\, 2^{2k}}.$$

Approximation: Suppose k is very large

Stirling's Formula: $k! \sim k^{k+1/2}e^{-k}\sqrt{2\pi}$.

for $k \to +\infty$.

$$\frac{(2k)!}{(k!)^2\, 2^{2k}} \sim \frac{(2k)^{2k+1/2}e^{-2k}\sqrt{2\pi}}{\left(k^{k+1/2}e^{-k}\sqrt{2\pi}\right)^2 2^{2k}}$$

$$\sim \frac{(2k)^{2k+1/2}\cancel{e^{-2k}}\sqrt{2\pi}}{k^{2k+1}\cancel{e^{-2k}}\,2\pi\,2^{2k}}.$$

$$n = 2k+1$$

$$\frac{n}{2} \sim \frac{n-1}{2} = k$$

$$\sim \frac{2^{\cancel{2k}+1/2}\,k^{\cancel{2k}+1/2}\sqrt{2\pi}}{k^{\cancel{2k}+1}\,2\pi\cdot\cancel{2^{2k}}}$$

$$= \frac{c}{\sqrt{k}}.$$

$$\sim \frac{\sqrt{2}\cdot\sqrt{2\pi}}{2\pi\,k^{1/2}} \sim \frac{1}{\sqrt{k}}. \qquad k \sim \frac{n}{2}.$$

So

average power of a voter in a close election $\approx \left(\#\begin{smallmatrix}\text{electoral}\\\text{college votes}\end{smallmatrix}\right) \times \mathbb{P}\left(\begin{smallmatrix}\text{make}\\\text{a}\\\text{difference}\end{smallmatrix}\right)$

$$\sim cn\cdot\frac{1}{\sqrt{k}} \qquad k \sim \frac{n}{2}$$

$$\sim \frac{cn}{\sqrt{n}} \sim c\sqrt{n}.$$