Math 170E: Winter 2023

Lecture 9, Mon 30th Jan

Special mathematical expectation

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

$Y(\omega) = g(X(\omega))$.

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

- Given $p \in (0, 1)$, $X \sim \text{Bernoulli}(p)$ if it has PMF

$$p_X(x) = \begin{cases} p & \text{if } x = 1 \\ 1 - p & \text{if } x = 0 \end{cases}$$

Today:

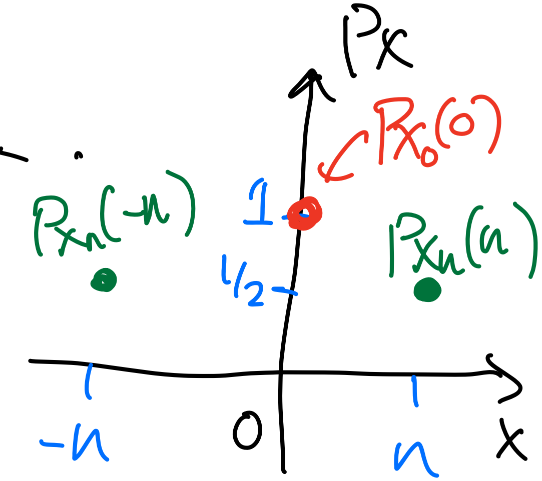
We'll discuss today:

- (centered) rth moments of random variables
- how to compute the variance and standard deviation of a random variable
- how to compute the moment generating function (MGF) of a discrete random variable
- how the MGF can be used to compute the mean, variance and moments of a discrete r.v.

$$\sum_{x \in \{-a, a\}} x p_X(x) = -a \tfrac{1}{2} + a \tfrac{1}{2}$$

An example: Consider the discrete r.v.s:

$$X_0 = 0, \text{ with probab. } 1.$$



Given $n \in \mathbb{N}$,

$$X_n = \begin{cases} n & \text{with prob. } 1/2 \\ -n & \text{with prob. } 1/2 \end{cases}$$

$\hookrightarrow$ $\mathbb{E}[X_0] = 0$, $\mathbb{E}[X_n] = 0$

$\hookrightarrow$ X_n feels more "spread out" than X_0 & even X_m for $m < n$.

Q: Can we devise some measure of spread for a r.v.?

$\hookrightarrow$ $X - \mathbb{E}[X]$ $\hookrightarrow$ $\mathbb{E}[|X - \mathbb{E}[X]|^2] = \sum_{x \in S} |x - \mathbb{E}[X]|^2 \mathbb{P}_X(x)$,

$|x|^2 = x^2$ $\quad = \sum_{x \in S} (x - \mathbb{E}[X])^2 \mathbb{P}_X(x)$

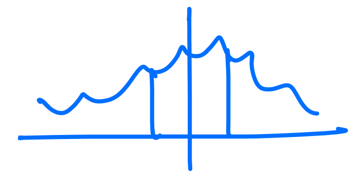
$\hookrightarrow$ $\mathbb{E}[(X_0 - \mathbb{E}[X_0])^2] = \mathbb{E}[X_0^2] = 0$

$\mathbb{E}[(X_n - \mathbb{E}[X_n])^2] = \mathbb{E}[X_n^2] = n^2 \to +\infty$ as $n \to +\infty$.

Definition 2.11:

If X is a discrete random variable taking values in a countable set $S \subseteq \mathbb{R}$ and $b \in \mathbb{R}$. We define the r^{th} moment of X about b is defined to be

$$\mathbb{E}[(X-b)^r] \quad , r \in \mathbb{N}.$$



b

If $b = 0$, we call this the r^{th} moment of X

$$\mathbb{E}[(X-b)^r] = \sum_{x \in S} (x-b)^r P_X(x),$$

$X = 2b$, prob $1/2$

$b^r \cdot \frac{1}{2}$

$$\mathbb{E}[X^r].$$

$$\mathbb{E}[X^2], \mathbb{E}[X],$$

Example 7:

$r=3$ $\qquad b=-1/2$

Let $X \sim \text{Bernoulli}(\frac{1}{3})$. What is the 3^{rd} moment of X about $-\frac{1}{2}$?

$$\begin{cases} 1 & \text{w/ prob } 1/3 \\ 0 & \text{———— } 2/3 \end{cases}$$

$$\mathbb{E}[(X-b)^r] = \mathbb{E}[(X+1/2)^3]$$

$$= \sum_{x\in\{0,1\}} (x+1/2)^3 P_X(x).$$

$$= (0+1/2)^3 P_X(0) + (1+1/2)^3 P_X(1).$$

$$= \frac{1}{8} \times \frac{2}{3} + \frac{3^3}{2^3}\cdot\frac{1}{3}$$

$$= \frac{1}{8}\times\frac{2}{3} + \frac{9}{8}$$

$$= \frac{1}{8}\times\frac{2}{3} + \frac{9\times 3}{8\times 3} = \frac{2+27}{24} = \frac{29}{24} \;//.$$

Definition 2.12:

Let X be a discrete random variable. We define the variance of X to be

$$\mathrm{var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] \rightsquigarrow \text{units}^2.$$

whenever it converges.

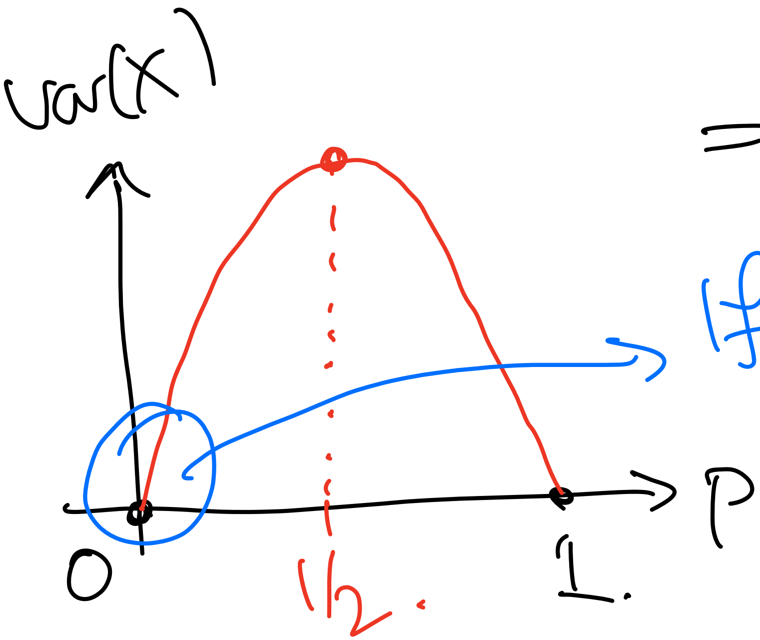
- We use the notation $\sigma_X^2 = \mathrm{var}(X)$ $\hookrightarrow \mu_X = \mathbb{E}[X].$

The standard deviation of X is $\sigma_X = \sqrt{\mathrm{var} X}$ $\rightsquigarrow$ units

Example 8: Given $p \in (0,1)$, let $X \sim \text{Bernoulli}(p)$. $\mathrm{var}(X) = \mathbb{E}[(X-\mu_X)^2]$.
What is $\mathrm{var}(X)$?

Last time: $\mu_X = p$.

$$\begin{aligned}\mathrm{var}(X) = \mathbb{E}[(X-p)^2] &= (0-p)^2 p_X(0) + (1-p)^2 p_X(1)\\ &= p^2(1-p) + (1-p)^2 p\\ &= p(1-p)\{\cancel{p} + 1 - \cancel{p}\}\\ &= p(1-p).\end{aligned}$$



If p small, $1-p \approx 1$,

$\mathbb{E}(X) = p \approx 0$

$\mathrm{var}(X) = p(1-p) \approx p$.

Proposition 2.13: If X is a discrete random variable and $a, b \in \mathbb{R}$, then:

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$$

$$\mathrm{var}(aX + b) = a^2 \mathrm{var}(X)$$

Proof:

$$\mathrm{var}(aX+b) = \mathbb{E}\left[\left(aX+b - \mathbb{E}[aX+b]\right)^2\right]$$

$$= \mathbb{E}\left[\left(aX+\cancel{b} - a\mathbb{E}[X] \cancel{- b}\right)^2\right].$$

$$= \mathbb{E}\left[a^2\left(X - \mathbb{E}[X]\right)^2\right] = a^2 \mathrm{var}(X). \quad //$$

$$\hookrightarrow \quad \sigma_{aX+b} = \sqrt{a^2 \mathrm{var}(X)} = |a|\, \sigma_X.$$

Proposition 2.14: If X is a discrete random variable then

$$\text{var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Proof:

$$\text{var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] \longrightarrow \sum_{x \in S} (x - \mu_X)^2 p_X(x)$$

$$\mathbb{E}[X^2] = \sum x^2 p_X(x)$$

$$= \mathbb{E}[X^2 - 2X\mathbb{E}[X] + \mathbb{E}[X]^2]$$

linearity

$$= \mathbb{E}[X^2] - 2\mathbb{E}[X\mathbb{E}[X]] + \mathbb{E}[X]^2$$

$$= \mathbb{E}[X^2] - 2\mathbb{E}[X]\mathbb{E}[X] + \mathbb{E}[X]^2$$

$$= \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

Example 9:

Let $m \geq 1$ and $X \sim \text{Uniform}(\{1, 2, \ldots, m\})$. Then

$\hookrightarrow \mu_X = \frac{m+1}{2}$.

$$\text{var}(X) = \frac{m^2 - 1}{12}$$

$\hookrightarrow \sum_{x=1}^{m} x$.

- e.g. $m = 6$, $\mu_X = \frac{7}{2} = 3.5$ and $\text{var}(X) = \frac{35}{12} \sim 2.92 \implies \sigma_X \sim 1.71$

$$\text{var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

$$= \mathbb{E}[X^2] - \frac{(m+1)^2}{4}.$$

$$\mathbb{E}[X^2] = \sum_{x=1}^{m} x^2 p_X(x) = \frac{1}{m} \sum_{x=1}^{m} x^2.$$

$$= \frac{1}{m} \cdot \frac{m(m+1)(2m+1)}{6}.$$

Evaluating $S_m = \sum_{x=1}^{m} x^2$

Look at $(x+1)^3 - x^3 = 3x^2+3x+1$ & sum both sides:

$$LHS \underset{\text{Telescoping}}{=} \sum_{x=1}^{m} \{(x+1)^3 - x^3\} = (\cancel{2^3} - 1^3) + (\cancel{3^3} - \cancel{2^3}) + \dots + ((m+1)^3 - \cancel{m^3})$$

$$= (m+1)^3 - 1.$$

$$RHS = \sum_{x=1}^{m} 3x^2+3x+1 = 3S_m + 3\sum_{x=1}^{m} x + m$$

$$= 3S_m + 3\frac{m(m+1)}{2} + m.$$

LHS=RHS

Thus, $3S_m = (m+1)^3 - 1 - \frac{3m(m+1)}{2} - m$

$$= \frac{m}{2}(2m^2+3m+1) = \frac{m(m+1)(2m+1)}{2}.$$

Definition 2.15: If X is a discrete random variable, we define the moment generating function (MGF) of X to be the function

$$M_X(t) = \mathbb{E}[e^{tX}], \quad t \in \mathbb{R}$$

whenever it exists.

$$X+Y$$

$$M_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX}e^{tY}] = \mathbb{E}[e^{tX}]\,\mathbb{E}[e^{tY}]$$

$$M_X(t) = \mathbb{E}[g(X)] = \sum_{x \in S} e^{tx} p_X(x).$$

$$g(x) = e^{tx}$$

$$e^{tx} = \sum_{n=0}^{\infty} \frac{t^n x^n}{n!} \text{ true for any } t, x \in \mathbb{R}.$$

$$f(t) = \sum \frac{f^{(n)}(0)}{n!} t^n$$

Formally,

$$M_X(t) = \mathbb{E}\left[\sum_{n=0}^{\infty} \frac{t^n}{n!} X^n\right]$$

$$= \sum_{n=0}^{\infty} \mathbb{E}\left[\frac{t^n}{n!} X^n\right] = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathbb{E}[X^n]$$

Example 9:

Let $p \in (0,1)$ and $X \sim \text{Bernoulli}(p)$. What is the MGF of X?

$$M_X(t) = \sum_{x=0}^{1} e^{tx} p_X(x)$$

$$= e^{t\cdot 0} p_X(0) + e^{t\cdot 1} p_X(1)$$

$$= 1 - p + e^t p. \quad , \text{ for all } t \in \mathbb{R}.$$

$$\sum e^{tx}$$

Proposition 2.16: Let X is a discrete random variable with MGF $M_X(t)$ which is well-defined and smooth for $t \in (-\delta, \delta)$, for some $\delta > 0$. Then

$$\frac{d^r}{dt^r} M_X \bigg|_{t=0} = \mathbb{E}[X^r], \quad r \in \{1, 2, 3, \ldots\}$$

$-\delta \quad 0 \quad \delta \quad t$

$M_X(0) = 1$

Proof: $M_X(t) = \sum_{x \in S} e^{tx} p_X(x)$

$$\frac{d^r}{dt^r} M_X(t) = \frac{d^r}{dt^r}\left(\sum_{x \in S} e^{tx} p_X(x)\right)$$

$$= \sum_{x \in S} \frac{d^r}{dt^r}\left(e^{tx}\right) \cdot p_X(x).$$

$$\frac{d^r}{dt^r}\left(e^{tx}\right) = x^r e^{tx}.$$

$$= \sum_{x \in S} x^r e^{tx} p_X(x).$$

$$\frac{d^r}{dt^r} M_X(t) \bigg|_{t=0} = \sum_{x \in S} x^r p_X(x) = \mathbb{E}[X^r]$$

Proposition 2.16: Let X is a discrete random variable with MGF $M_X(t)$ which is well-defined and smooth for $t \in (-\delta, \delta)$, for some $\delta > 0$. Then

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

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

$$\frac{d}{dt} \log M_X$$
$$= \log \mathbb{E}[e^{tX}]$$
$$= \frac{d}{dt} \mathbb{E}[tX]$$
$$= \mathbb{E}[X]$$

see

Why? HW3

Example 10:

Let $p \in (0,1)$ and $X \sim \text{Bernoulli}(p)$. Use the MGF to compute $\text{var}(X)$.

$\hookrightarrow$ $M_X(t) = \underbrace{1-p+pe^t}_{\text{smooth \& well-defined}}$, for $t \in \mathbb{R}$.

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

$\hookrightarrow$ $\log M_X(t) = \log(1-p+pe^t)$.

$$\frac{d}{dt} \log M_X(t) = \frac{pe^t}{1-p+pe^t}$$

$\hookrightarrow$ $\frac{d^2}{dt^2} \log M_X(t) = \frac{d}{dt}\left(\frac{pe^t}{1-p+pe^t}\right)$

$$= \frac{(1-p+pe^t)(pe^t) - (pe^t)(pe^t)}{(1-p+pe^t)^2}.$$

$$\frac{d^2}{dt^2}\log M_X(t)\Big|_{t=0} = \frac{(1-p+p)\cdot p - p^2}{(1-p+p)^2}$$

$$= p - p^2 = p(1-p).$$

$$= var(X).$$