Class 9: Mean & Variance

Tuesday, October 1, 2019 11:38 AM

- Midterm Tues, Oct 8
- Next week: office hours Monday 10:30-11:30 10/7
  (still on Wednesday this week)

Expectation

Let X be a (discrete) random variable with PMF p.

(Recall $\mathbb{P}(X=k)=p(k)$ for all $k\in\mathbb{R}$.)

Then
$$\mathbb{E}[X]=\sum_k k\,p(k).$$

Examples:

① If $X\sim \mathrm{Ber}(p)$, meaning $X=\begin{cases}1 & \text{with prob } p\\ 0 & \text{with prob } 1-p,\end{cases}$

then $\mathbb{E}[X]=p$.

② If $X\sim \mathrm{Bin}(n,p)$, meaning $p(k)=\binom{n}{k}p^k(1-p)^{n-k}$,

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

= average # success out of n trials with success probability p.

③ Let $X\sim \mathrm{Geom}(p)$, meaning $p(k)=p(1-p)^{k-1}$

trials until first success

= prob. that k^{th} trial is first success, for $k=1,2,3,\ldots$

Then interpret $\mathbb{E}[X]$ as "average amount of time we have to wait until first success".

$$\mathbb{E}[X]=\sum_{k=1}^{\infty}k\,p(k)=\sum_{k=1}^{\infty}k\,p(1-p)^{k-1}$$

$$=p\sum_{k=1}^{\infty}k(1-p)^{k-1}$$

$$\frac{d}{dx}x^k=kx^{k-1}$$

$$=p\sum -\frac{d}{dp}\left[(1-p)^k\right]$$

$$= p \sum_{k=1}^{\infty} -\frac{d}{dp}\left[(1-p)^k\right]$$

on

$$= -p \frac{d}{dp}\left[\sum_{k=1}^{\infty} (1-p)^k\right]$$

$$\sum_{k=1}^{\infty} r^k = \frac{r}{1-r}$$

$$= -p \frac{d}{dp}\left[\frac{1-p}{1-(1-p)}\right]$$

$$= -p \frac{d}{dp}\left[\frac{1}{p} - 1\right]$$

$$= -p \left(-\frac{1}{p^2}\right)$$

$$= 1/p .$$

Inversely related to p.

Ex: $p = \frac{1}{2} \Rightarrow \mathbb{E}[X] = 2$ coin flips til first heads on average

---

Change of Variables / Law of the unconscious statistician

Let X be a r.v. with PMF p.

- Goal: Calculate $\mathbb{E}[f(X)]$ for functions $f: \mathbb{R} \to \mathbb{R}$.

  Ex: $\mathbb{E}[X^2] = ?$

- First way: Let $Y = f(X)$. Find PMF of Y and use it to compute $\mathbb{E}[f(X)] = \mathbb{E}[Y] = \sum_n n \mathbb{P}(Y=n)$.

- Direct way:

$$\mathbb{E}[f(X)] = \sum_k f(k) p(k)$$

Proof:

First way using $Y = f(X)$ must be correct.

Last time: $\mathbb{P}(Y=n) = \mathbb{P}(f(X)=n) = \sum_{k \text{ such that } f(k)=n} \mathbb{P}(X=k)$.

Thus

$$\mathbb{E}[f(X)] = \mathbb{E}[Y] = \sum n \mathbb{P}(Y=n)$$

Thus

$$\mathbb{E}[f(X)] = \mathbb{E}[Y] = \sum_n n\,\mathbb{P}(Y=n)$$

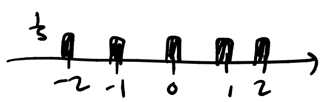
$$= \sum_n n \sum_{\substack{k \text{ s.t.} \\ f(k)=n}} p(k)$$

$$= \sum_n \sum_{\substack{k \text{ s.t.} \\ f(k)=n}} n\,p(k)$$

$$= \sum_k f(k)\,p(k).$$

Ex from last time:

Let X have PMF $p(k) = \frac{1}{5}$ for $k = -2, -1, 0, 1, 2$.



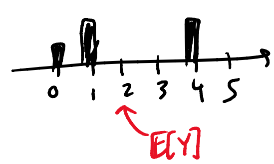
We saw $\mathbb{E}[X] = 0$.

Let $Y = X^2$. We saw $p_Y(n) = \begin{cases} \frac{1}{5} & \text{if } n=0 \\ \frac{2}{5} & \text{if } n = 1, 4 \end{cases}$.

Thus $\mathbb{E}[Y] = 0 \cdot p_Y(0) + 1 \cdot p_Y(1) + 4 \cdot p_Y(4)$

$$= 1 \cdot \tfrac{2}{5} + 4 \cdot \tfrac{2}{5}$$

$$= 2.$$



Direct way (change of variables):

$$\boxed{\mathbb{E}[f(X)] = \sum_k f(k)\,p(k) \quad \text{Use } f(x) = x^2.}$$

$$\mathbb{E}[Y] = \mathbb{E}[X^2] = \sum_k k^2 p(k)$$

$$= (-2)^2 p(-2) + (-1)^2 p(-1) + 0^2 p(0) + 1^2 p(1) + 2^2 p(2)$$

$$= 4 \cdot \tfrac{1}{5} + 1 \cdot \tfrac{1}{5} + 0 \cdot \tfrac{1}{5} + 1 \cdot \tfrac{1}{5} + 4 \cdot \tfrac{1}{5}$$

$$= 2.$$

Rule: (Linearity of Expectation)

If X is a r.v. and $a, b \in \mathbb{R}$ (constants, <u>not</u> random)

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

Proof:

Use change of variables with $f(x) = ax + b$ to get

$$\begin{aligned}
\mathbb{E}[aX+b] &= \sum_k (ak+b)\,p(k) \\
&= \left(\sum_k ak\,p(k)\right) + \left(\sum_k b\,p(k)\right) \\
&= a\underbrace{\left(\sum_k k\,p(k)\right)}_{\mathbb{E}[X]} + b\underbrace{\left(\sum_k p(k)\right)}_{1 \text{ because probabilities sum to } 1} \\
&= a\,\mathbb{E}[X] + b.
\end{aligned}$$

Variance

Variance is a measure of how "spread out" a distribution is around its mean.

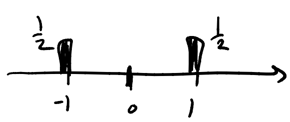
Ex:

$\mathbb{P}(X=0) = 1$

Zero variance

$\mathrm{Var}(X) = 0$

$\mathrm{Sd}(X) = 0$



$X = \begin{cases} 1 & \text{w.p. } \frac{1}{2} \\ -1 & \text{w.p. } \frac{1}{2} \end{cases}$

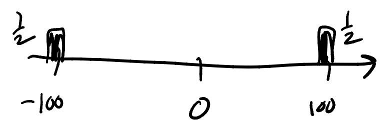
$\mathbb{E}[X] = 0$

Since $|X - \mathbb{E}[X]| = 1$ always,

$\mathrm{Var}(X) = 1$

$\mathrm{Sd}(X) = 1$

"w.p." means "with probability"



$X = \begin{cases} 100 & \text{w.p. } \frac{1}{2} \\ -100 & \text{w.p. } \frac{1}{2} \end{cases}$

$\mathbb{E}[X] = 0$

Now $|X - \mathbb{E}[X]| = 100$ always.

Higher variance,

$\mathrm{Var}(X) = 100^2$.

$\mathrm{Sd}(X) = 100$

Def: The variance of a r.v. X is defined by

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

- If X has PMF p and $\mathbb{E}[X] = m$, then

$$Var(X) = \sum_k (k-m)^2 p(k).$$

- Interpret $Var(X)$ as "mean squared error" or
  "avg squared distance from mean" or "deviation" ...

- $Var(X)$ is insensitive to which direction you deviate from mean.

- Why the square? Why $\mathbb{E}\big[\,|X - \mathbb{E}[X]|\,\big]$?

  No good reason, just usually more tractable, easier to compute.

- Variance squares the "units". To get to original units, use the standard deviation

$$Sd(X) = \sqrt{Var(X)}$$
$$= \sqrt{\mathbb{E}[(X-\mathbb{E}[X])^2]}.$$

Note:

$$\mathbb{E}[f(X)] \neq f(\mathbb{E}[X])$$
$$\sqrt{\mathbb{E}[(X-\mathbb{E}[X])^2]} \neq \mathbb{E}\big[\,|X-\mathbb{E}[X]|\,\big]$$

Ex:

Let $X \sim Ber(p)$. $\quad X = \begin{cases} 1 & \text{w.p. } p \\ 0 & \text{w.p. } 1-p \end{cases}$.

We saw that $\mathbb{E}[X] = p$.

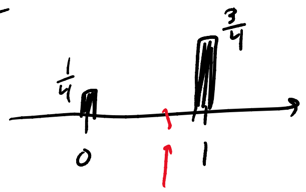
So $Var(X) = \mathbb{E}[(X-p)^2] = \sum_k (k-p)^2 \mathbb{P}(X=k)$

$$= (1-p)^2 \mathbb{P}(X=1) + (0-p)^2 \mathbb{P}(X=0)$$
$$= (1-p)^2 p + (0-p)^2 (1-p)$$

$$= (1-p)^2 p + (0-p)^2(1-p)$$
$$= p(1-p)^2 + p^2(1-p)$$
$$= p(1-p)(1-p+p)$$
$$= p(1-p).$$ = success probability times failure probability

So $Sd(x) = \sqrt{p(1-p)}$.

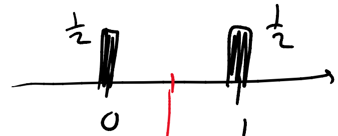
$p = \frac{3}{4}$



$E[X] = \frac{3}{4}$

$Var(x) = \frac{3}{4} \cdot \frac{1}{4} = \frac{3}{16}$

$p = \frac{1}{2}$

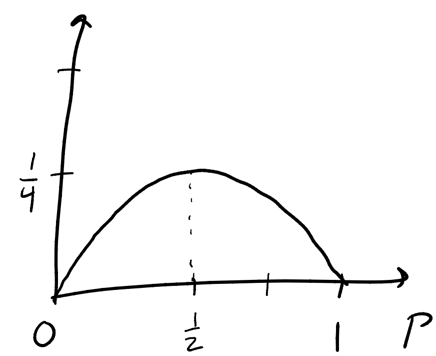


$E[X] = \frac{1}{2}$

$Var(x) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

Plot $Var(x)$ as a function of p:

$Var(x) = p(1-p)$



Ex: (Decision Theory)

↪ How do people choose between random outcomes?

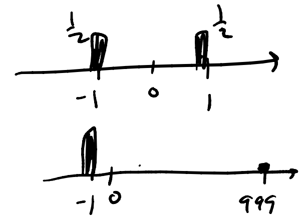
Two Options

① Win \$1 w.p. $\frac{1}{2}$ or lose \$1 w.p. $\frac{1}{2}$,

✓ ② Win \$999 w.p. $\frac{1}{1000}$ or lose \$1 w.p. $\frac{999}{1000}$

Mean/variance:

Mean/variance:



① $\text{Mean} = 1 \cdot \frac{1}{2} + (-1) \cdot \frac{1}{2} = 0$

$\text{Variance} = (1-0)^2 \cdot \frac{1}{2} + (-1-0)^2 \cdot \frac{1}{2} = 1.$

② $\text{Mean} = 999 \cdot \frac{1}{1000} + (-1) \cdot \frac{999}{1000} = 0$

$\text{Variance} = 999^2 \cdot \frac{1}{1000} + 1^2 \cdot \frac{999}{1000} = 999.$

③ Lose \$999 w.p. $\frac{1}{1000}$, win \$1 w.p. $\frac{999}{1000}$.

versus ①, most people choose ①.

Again option ③ has mean 0 and variance = 999.

This illustrates loss aversion.

↳ Risk aversion is closely related.

Ⓐ Win \$1 certainly.

Ⓑ Lose \$999 or win \$1001, each w.p. $\frac{1}{2}$.

Both cases mean \$1. Ⓐ is more popular because smaller variance.