4.4 Expected value of a function of a r.v.

Lec. 12 02/01

Ex A box contains 10 disks with radii $0, 1, 2, \ldots, 10$. inches
One disk is chosen at random.
Expected <u>area</u> of the disk = ?

$A = \pi R^2$, $\quad R$ = radius of the chosen disk, r.v. with $E[R] = 5$.

- But $E[A] \neq \pi \cdot 5^2 \approx \underline{78.5}$ in^2 (Warning!)

Instead, R takes values $0, 1, 2, \ldots, 10$ with equal probabilities $\frac{1}{11}$

A takes values $\pi \cdot 0^2, \pi \cdot 1^2, \ldots, \pi \cdot 10^2$ with same prob's $\frac{1}{11}$.

$$\Rightarrow \quad E[A] = \pi 0^2 \cdot \underbrace{\frac{1}{11}}_{p(x)} + \pi 1^2 \cdot \frac{1}{11} + \pi 2^2 \cdot \frac{1}{11} + \ldots + \pi 10^2 \cdot \frac{1}{11} \approx 110 \text{ in}^2$$

Prop. 4.1 Let X be a discrete r.v. with pmf $p(x)$.
Then for any function $g(\cdot)$,
$$E[g(X)] = \sum_x g(x) \, p(x)$$

Proof $g(x)$ is a r.v. with values $g(x_k)$

$$E[g(X)] = \sum_k g(x_k) \underbrace{P\{g(X) = g(x_k)\}}_{= P\{X = x_k\} = p(x_k)}$$

$P\{X = x_k\} = p(x_k)$, if g is one-to-one $\Rightarrow$ QED.

If g is not one-to-one, make a small perturbation of its values.

Properties of expectation:

1. $E[X+Y] = E[X] + E[Y]$ for any r.v. X, Y
2. $E[aX] = a \cdot E[X]$ for a r.v. X, $a \in \mathbb{R}$
3. $E[a] = a$ for $a \in \mathbb{R}$ (constant r.v.)

Proof of (2):

$$E[aX] = \sum_x ax \cdot p(x) \qquad \text{(by Prop. 4.1 with } g(x) = ax)$$
$$= a \cdot \sum_x x p(x) = a \cdot E[X]. \qquad \text{QED.}$$

4.5. Variance

Def X: r.v. with mean (= expected value) μ. <u>Variance</u> of X is

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

<u>Standard deviation</u> of X is

$$\sigma_X = \sqrt{\mathrm{Var}(X)}$$

More conveniently:

$$\begin{aligned}\mathrm{Var}(X) &= E[X^2 - 2\mu X + \mu^2] \\ &= E[X^2] - 2\mu \underbrace{E[X]}_{=\mu} + \mu^2 \qquad \text{(by the properties of expectation p.33)} \\ &= E[X^2] - \mu^2. \qquad \Rightarrow \text{ we proved:}\end{aligned}$$

Prop $\boxed{\mathrm{Var}(X) = E[X^2] - (E[X])^2}$

not the same!

Ex (refer to Ex. p.33): $E[R] = 5$, $E[R^2] = 0^2 \cdot \frac{1}{11} + 1^2 \cdot \frac{1}{11} + \dots + 10^2 \cdot \frac{1}{11} = 35$. $\mathrm{Var}(R) = 35 - 5^2 = 10$. $\sigma_R = \sqrt{10} \approx 3.2$ in

<u>Prediction.</u> Let X be a r.v.
What is the value t that <u>best predicts</u> X?

To answer this, we minimize the <u>risk</u>.

$$\ell(t) = E[(X-t)^2]$$

$$0 = \ell'(t) = \frac{d}{dt}\left(E[X^2] - 2t\,E[X] + t^2\right) = -2E[X] + 2t$$

$$\Rightarrow \quad t = E[X].$$

Thus: <u>the value that best predicts X is $E[X]$.</u>

(and the associated risk is $\mathrm{Var}(X)$).

Ex Compare the two investment options:

Stock A: 3% return with prob. 0.8; 2% loss with prob. 0.2

Stock B: 5% return with prob. 0.6; 2% loss with prob. 0.4

Expected profit (in %):

$$E[X_A] = 3\cdot 0.8 - 2\cdot 0.2 = 2\%$$

$$E[X_B] = 5\cdot 0.6 - 2\cdot 0.4 = 2.2\% \quad \leftarrow \text{slightly better}$$

However, compare the risks:

$$E[X_A^2] = 3^2\cdot 0.8 + 2^2\cdot 0.2 = 8$$

$$E[X_B^2] = 5^2\cdot 0.6 + 2^2\cdot 0.4 = 16.6$$

$$Var(X_A) = E[X_A^2] - (E[X_A])^2 = 8 - 2^2 = 4$$

$$Var(X_B) = E[X_B^2] - (E[X_B])^2 = 16.6 - 2.2^2 = 11.76 \quad \leftarrow \text{higher}$$

$$\sigma_{X_A} = \sqrt{2.4} \approx 2\%$$

$$\sigma_{X_B} = \sqrt{8.56} \approx 3.4\%$$

So, typical profit for A is $(2 \pm 2)\%$

typical profit for B is $(2.2 \pm 3.4)\%$.

Ans: A is better, less risk.

Properties of variance

1. It is <u>not</u> true in general that $Var(X+Y) = Var(X) + Var(Y)$.
2. $Var(aX) = a^2 \cdot Var(X)$, so $\sigma_{aX} = a \cdot \sigma_X$
3. $Var(X+b) = Var(X)$, so $\sigma_{X+b} = \sigma_X$.

Proof of (3):

$$Var(X+b) = E\left[\left((X+b) - E(X+b)\right)^2\right] = E\left[\left(X+b-E[X]-b\right)^2\right]$$ (by prop's of expectation p.33)

$$= E\left[\left(X - E[X]\right)^2\right] = Var(X).$$ QED.

Lec. 13 $\boxed{02/03/2012}$

Ex (The matching problem)

N homeworks are returned to N students at random.
How many students will get their own homework on average?

(Recall that we calculated before $P\{\text{at least 1 student gets his/her own HW}\} \approx 1 - \frac{1}{e}$, see p.17).

Method of indicators: $X = \#$ students getting their own HW.

Represent $X = \sum_{i=1}^{N} X_i$ where $\underset{\substack{\uparrow \\ \text{"indicators"}}}{X_i} = \begin{cases} 1, & \text{if student } i \text{ gets his/her own HW} \\ 0, & \text{otherwise} \end{cases}$

$$\Rightarrow E[X] = \sum_{i=1}^{N} \underbrace{E[X_i]}_{=?}$$

$E[X_i] = 1 \cdot P\{X_i = 1\} + 0 \cdot P\{X_i = 0\} = P\{X_i = 1\} = P\{\text{student } i \text{ gets own HW}\} = \frac{1}{n}$. (why?)

Therefore $E[X] = N \cdot \frac{1}{N} = 1$.

Ex [What is the standard deviation of the students who get their own hw?

$$E[X^2] = E\left[\left(\sum_{i=1}^{N} X_i\right)^2\right] = E\left[\sum_{i=1}^{N} X_i^2 + \sum_{i \neq j} X_i X_j\right] = \sum_{i=1}^{N} E[X_i^2] + \sum_{i \neq j} E[X_i X_j].$$

$E[X_i^2] = E(X_i) = \frac{1}{N}$, as before.

$E[X_i X_j] = 1 \cdot P\{X_i X_j = 1\} + 0 \cdot P\{X_i X_j = 0\} = P\{X_i = 1 \text{ and } X_j = 1\} = P\{\text{both students } i,j \text{ get their own HW}\}$

$= \frac{1}{N(N-1)}$. (Why? e.g. by conditioning on one of students)

Therefore

$$E[X^2] = N \cdot \frac{1}{N} + \underbrace{N(N-1)}_{\#\text{ pairs of } i \neq j} \cdot \frac{1}{N(N-1)} = 2.$$

Hence $Var(X) = E[X^2] - (E[X])^2 = 2 - 1 = 1$ $\qquad \sigma_X = \sqrt{1} = 1$