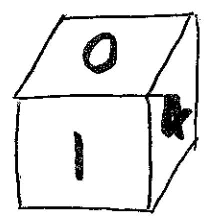
3 - 12 - 2008

- Expectation:

$$E(X) = \sum_{i=1}^{n} x_i \, p(X = x_i)$$

$\approx$ Average of X over many experiments

- Example 1



X: result of tossing fair die numbered 0, 1, 2, 3, 4, 5
Except 0, not possible

$$E(x) = 0 \cdot \cancel{\tfrac{1}{6}}_{\,0} + 1 \cdot \cancel{\tfrac{1}{6}}_{\,1/5} + 2 \cdot \cancel{\tfrac{1}{6}}_{\,1/5} + 3 \cancel{\tfrac{1}{6}}_{\,1/5} + 4 \cdot \cancel{\tfrac{1}{6}}\,1/5 + 5 \cdot \cancel{\tfrac{1}{6}}\,1/5$$

$= 3$

Example 2

Y: result of tossing magic die

numbered 0, 1, 2, 3, 4, 5

- almost always gives 3
- $\frac{1}{1000}$ prob to give 1, 2, 4, 5
- 0 prob to give 0

$$E(Y) = 3$$

- Same expectation, but different
- \$100 reward for guessing result
- which one to guess?

Example 2 !

- Reason:

  Less uncertainty

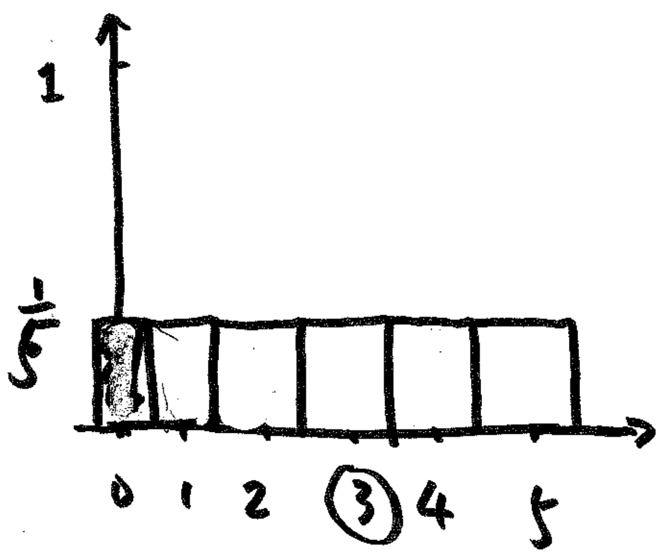
  Y differs less from $E(Y)$

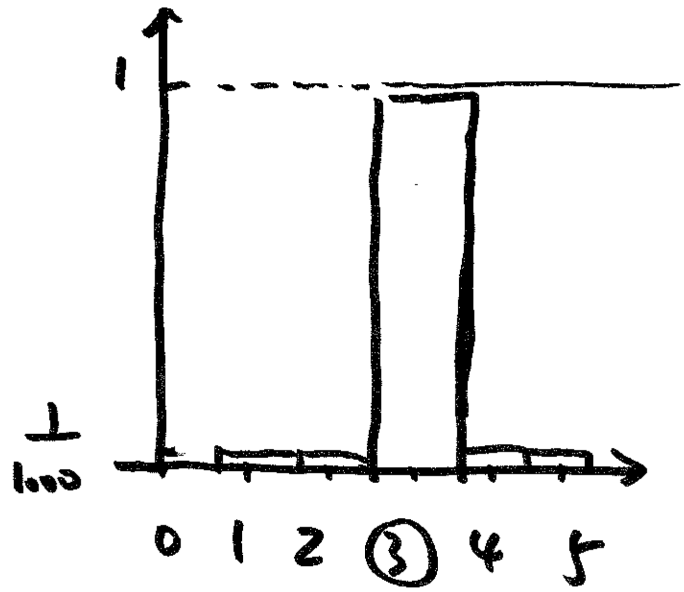
  than X from $E(X)$

- How to measure difference btw r.v. & its expectation?

  Variance

prob histogram of X in Example 1





prob histogram of Y in Example 2

X: # of heads in N coin toss

N	spread of values of X w/ > almost 0 prob
10	9/11 ≈ 0.82
25	15/26 ≈ 0.57
100	29/101 ≈ 0.29
400	58/400 ≈ 0.15

spread doubles as # of trials quadruples

Y : # of correct answers in N Qs

N	spread of values of Y w/ prob > almost 0	
10	5 / 11	≈ 0.45
25	11 / 26	≈ 0.42
100	22 / 101	≈ 0.21
400	44 / 401	≈ 0.11

Spread doubles as # of trials quadruples

$$V(x) = \sum_{i=1}^{k} (x_i - E(x))^2 \, p(X = x_i)$$

Example:

X: # of heads in 4 coin toss

$E(x) = 2$

x	0	1	2	3	4
$p(x)$	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{16}$

$$V(x) = (0-2)^2 \cdot \frac{1}{16} + (1-2)^2 \cdot \frac{1}{4}$$

$$+ (2-2)^2 \frac{1}{8} + (3-2)^2 \cdot \frac{1}{4}$$

$$+ (4-2)^2 \cdot \frac{1}{16} = 1$$

Analogy for proof of Lemma 5.28

$a_1 = 1, \quad a_2 = 2, \quad a_3 = \cancel{2}\,4$

$b_1 = 1, \quad b_2 = 2$

$$\sum_{i=1}^{3} a_i \sum_{j=1}^{2} b_i$$

$$(a_1 + a_2 + a_3)(b_1 + b_2)$$

$$= \sum_{i=1}^{3} \sum_{j=1}^{2} a_i b_j$$

$$\left(\underset{1}{a_1 b_1} + \underset{2}{a_1 b_2} + \underset{2}{a_2 b_1} + \underset{4}{a_2 b_2} + \underset{4}{a_3 b_1} + \underset{8}{a_3 b_2} \right)$$

$$= a_1 b_1 + (a_1 b_2 + a_2 b_1) + (a_2 b_2 + a_3 b_1) + a_3 b_2 \quad (*)$$

$$\sum_{(i,j):\, a_i b_j = 2} a_i b_j \qquad \sum_{(i,j):\, a_i b_j = 4} a_i b_j$$

(1 2)
(2 1)

$$a_1 b_1 = \sum_{\overline{(i,j):\ a_i b_j = 1}} a_i b_j$$

$$a_3 b_2 = \sum_{\overline{(i,j):\ a_i b_j = 8}} a_i b_j$$

$c_1 = 1, \quad c_2 = 2, \quad c_3 = 4, \quad c_4 = 8$

$$(*) = \sum_{k=1}^{4} \sum_{\overline{(i,j):\ a_i b_j = c_k}} a_i b_j$$

Analogy for proof of Lemma 5.28

X, Y: results of rolling ~~twice~~ two dice

x_1,	x_2,	x_3,	x_4,	x_5	x_6
y_1,	y_2,	y_3	y_4	y_5	y_6
1,	2	3	4	5	6

$$\sum_{(i,j)\,:\,x_i y_j = 4} P(X = x_i,\ Y = y_j)$$

$$= P(X=1,\ Y=4) + P(X=4,\ Y=1)$$

$$+ P(X=2,\ Y=2)$$

$$= P(XY = 4)$$