Lecture-27 P 1

Recap: Correlation, covariance
attendance / marks

Properties of Conditional
expectation

$$E[X] = E[E[X|Y]]$$

Discrete case:

$$E[X] = \sum_y E[X|Y=y]\, P(Y=y)$$

$Y = 1$ or 2 or 3
A B C

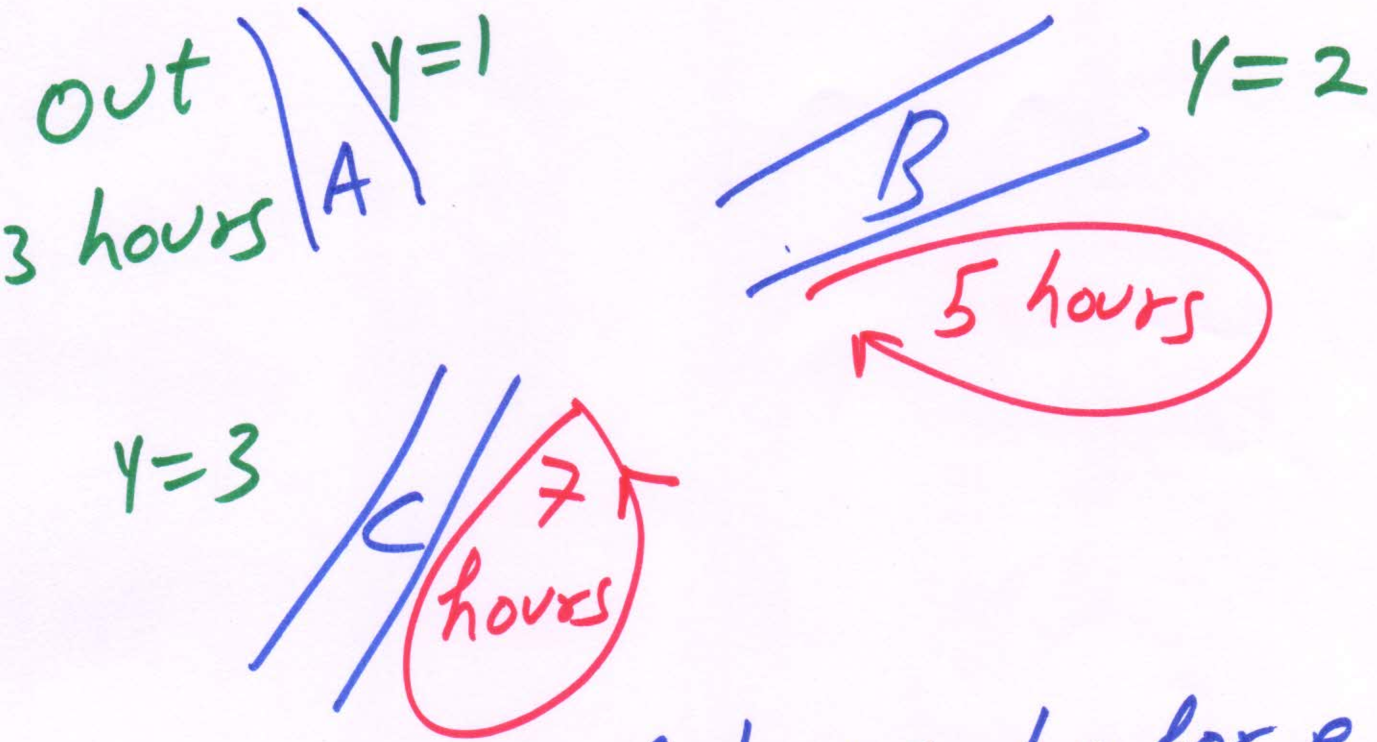
$$\boxed{= \sum_x x P(X=x)}$$

Continuous random variable

$$E[X] = \int_{-\infty}^{\infty} E[X|Y=y]\, f_Y(y)\, dy$$

E.g. Discrete random variable

A miner is trapped in a coal mine.



X: number of hours before he gets out of the tunnel.

$X \in \{ 3, 5+3, 7+3, 5+7+3, 5+5+3, 7+7+3, \ldots \}$

You can compute $E[X]$

$E[X] = E[E[X|Y]]$

$$E[X] = E[E[X|Y]]$$

$$= \sum_y E[X|Y=y]\, P[Y=y]$$

$$= \underline{E[X|Y=1]}\, P[Y=1] + \underline{E[X|Y=2]}\, P[Y=2] + E[X|Y=3]\, P[Y=3]$$

$$= \frac{1}{3}\left[\underline{3} + \underline{(\overset{B}{5} + \overset{\text{starting Point}}{E[X]})} + (7 + E[X])\right]$$

$$E[X] = 3 + 5 + 7 = 15 \text{ hours}$$

e.g.

How to choose a spouse??
There are 100 potential people that you will meet in your whole life. There is some ordering of these individuals. These is a ranking. You will meet these people in any of the 100! possible permutations.
What will be your strategy to get the best match?

→ You reject the first k matches. maximize over k

→ You choose the one that's better than all these k rejects.

A: low, < 0.1

B: 0.1 to 0.3

C: 0.3 to 0.5

D: > 0.5

$P_k(\text{best})$ = probability of choosing the best prize, after rejecting the first k prizes, and then choosing the first prize that is better than all these k prizes.

$\cancel{E[X] = E[E[X|Y]]}$ 6

$$P_k(\text{best}) = \sum_{i=1}^{n} P_k(\text{best} \mid X=i)\, \underline{P(X=i)}$$

mutually exclusive

where X is the position of the best prize.

$P(X=i)$ = Probability that the best candidate is in the i^{th} position $= \frac{1}{n}$

$i = 1$ to n

$P_k(\text{best} \mid X=i) = ?$

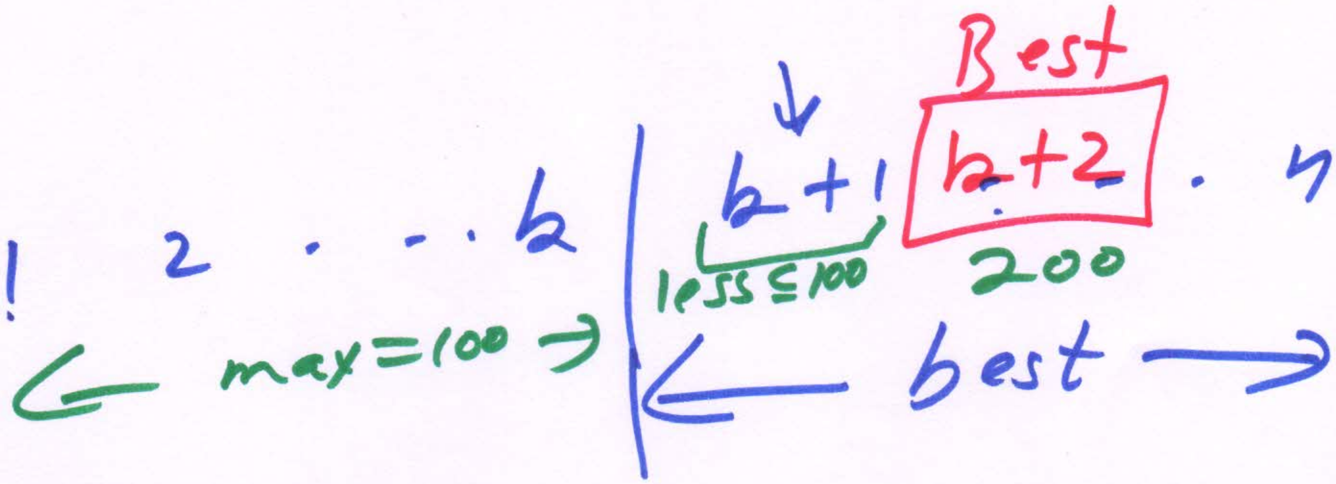
$1 \quad 2 \quad 3 \quad \ldots \quad k \mid k+1 \quad \ldots \quad n$

$\longleftarrow$ best $\longrightarrow$

$P_k(\text{best} \mid X=i) = 0$ for $i = 1$ to k

$P_k(\text{best} \mid X = i)$

for $i = k+1, k+2, \ldots, n$



let $i = k+1$

$P_k(\text{best} \mid X = k+1) = 1$

$P_k(\text{best} \mid X = k+2) = \dfrac{k}{k+1}$

P($k+1^{st}$ value is less than max (1 to k))

$P_k(\text{best} \mid X = i) = \dfrac{k}{i-1}$

$i \geq k+1$

The best of the first $i-1$ values is in the first k values.

$P_k(\text{best}) =$ 8

$$\sum_{i=1}^{k} P_k(\underset{\overset{\shortparallel}{0}}{\text{best}} / X = i)\, P(X=i) \quad +$$

$$\sum_{i=k+1}^{n} P_k(\text{best} \mid X=i)\, \underset{\overset{\shortparallel}{\frac{1}{n}}}{P(X=i)}$$

$$= \frac{1}{n} \sum_{i=k+1}^{n} \frac{k}{i-1}$$

$$= \frac{k}{n} \sum_{i=k+1}^{n} \frac{1}{i-1}$$

$$\simeq \frac{k}{n} \int_{k+1}^{n} \frac{1}{x-1}\, dx$$

$$= \frac{k}{n} \log\left(\frac{n-1}{k}\right) \simeq \boxed{\frac{k}{n} \log\left(\frac{n}{k}\right)}$$

$$\frac{k}{n} \log\left(\frac{n}{k}\right)$$

maximize this over k

$$k = \frac{n}{e}$$: reject the first 37% candidates.

$$P_{max} = \frac{n}{n \cdot e} \log\left(\frac{n \cdot e}{n}\right)$$

$$= \frac{1}{e} = 0.37$$

Zoom: chatwith me