

LECTURE 13: Conditional expectation and variance revisited;

Application: Sum of a random number of independent r.v.'s

- A more abstract version of the conditional expectation
 - view it as a random variable
 - the law of iterated expectations
- A more abstract version of the conditional variance
 - view it as a random variable
 - the law of total variance
- Sum of a random number
of independent r.v.'s
 - mean
 - variance

Conditional expectation as a random variable

- Function h
e.g., $h(x) = x^2$, for all x
 - Random variable X ; what is $h(X)$?
 - $h(X)$ is the r.v. that takes the value x^2 , if X happens to take the value x
- $g(y) = \mathbf{E}[X | Y = y] = \sum_x x p_{X|Y}(x | y)$
(integral in continuous case)
 - $g(Y)$: is the r.v. that takes the value $\mathbf{E}[X | Y = y]$, if Y happens to take the value y
- Remarks:
 - It is a function of Y
 - It is a random variable
 - Has a distribution, mean, variance, etc.

Definition: $\mathbf{E}[X|Y] = g(Y)$

The mean of $E[X | Y]$: Law of iterated expectations

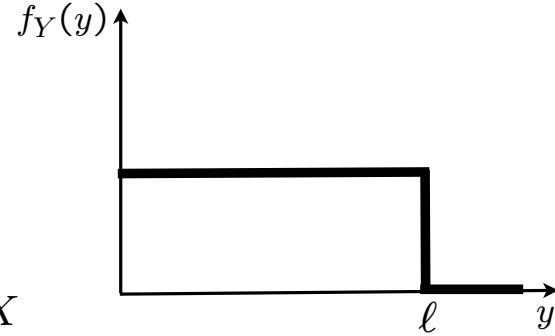
- $g(y) = E[X | Y = y]$

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

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

Stick-breaking example

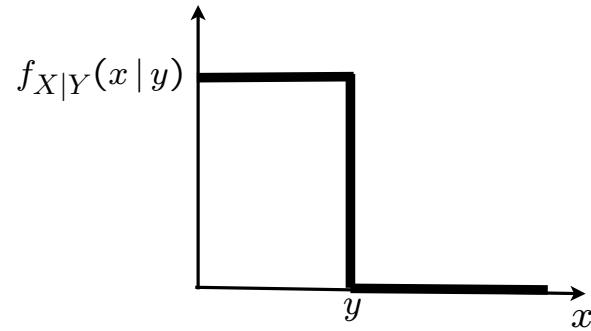
- Stick example: stick of length ℓ
break at uniformly chosen point Y
break what is left at uniformly chosen point X



$$\mathbf{E}[X \mid Y = y] =$$

$$\mathbf{E}[X \mid Y] =$$

$$\mathbf{E}[X] =$$



Forecast revisions

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

- Suppose forecasts are made by calculating expected value, given any available information
- X : February sales
- Forecast in the beginning of the year:
- End of January: will get new information, value y of Y

Revised forecast:

- Law of iterated expectations:

The conditional variance as a random variable

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

$$\text{var}(X | Y = y) = \mathbf{E}[(X - \mathbf{E}[X | Y = y])^2 | Y = y]$$

$\text{var}(X | Y)$ is the r.v. that takes the value $\text{var}(X | Y = y)$, when $Y = y$

- Example: X uniform on $[0, Y]$ $\text{var}(X | Y = y) =$

$$\text{var}(X | Y) =$$

Law of total variance: $\text{var}(X) = \mathbf{E}[\text{var}(X | Y)] + \text{var}(\mathbf{E}[X | Y])$

Derivation of the law of total variance

$$\text{var}(X) = \mathbf{E}[\text{var}(X | Y)] + \text{var}(\mathbf{E}[X | Y])$$

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

$$\text{var}(X | Y = y) =$$

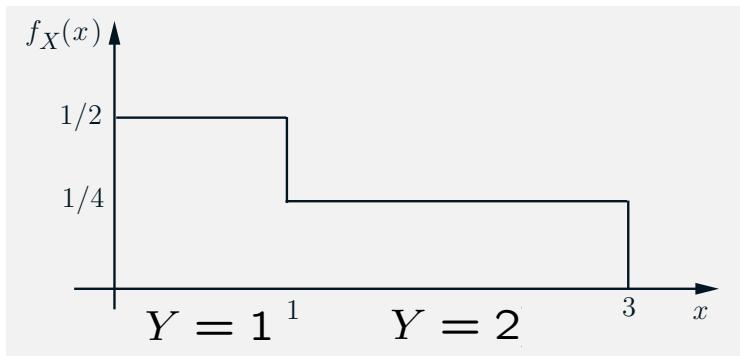
$$\text{var}(X | Y) =$$

$$\mathbf{E}[\text{var}(X | Y)] =$$

$$\text{var}(\mathbf{E}[X | Y]) =$$

A simple example

$$\text{var}(X) = \mathbb{E}[\text{var}(X | Y)] + \text{var}(\mathbb{E}[X | Y])$$



$$\text{var}(X | Y) = \begin{cases} \text{var}(X | Y = 1) = \\ \text{var}(X | Y = 2) = \end{cases}$$

$$\mathbb{E}[\text{var}(X | Y)] =$$

$$\mathbb{E}[X | Y] = \begin{cases} \mathbb{E}[X | Y = 1] = \\ \mathbb{E}[X | Y = 2] = \end{cases}$$

$$\mathbb{E}[\mathbb{E}[X | Y]] =$$

$$\text{var}(\mathbb{E}[X | Y]) =$$

Section means and variances

- Two sections of a class: $y = 1$ (10 students); $y = 2$ (20 students)
 x_i : score of student i
- Experiment: pick a student at random (uniformly)
random variables: X and Y
- Data:
 $y = 1 : \frac{1}{10} \sum_{i=1}^{10} x_i = 90$ $y = 2 : \frac{1}{20} \sum_{i=11}^{30} x_i = 60$

$$\mathbf{E}[X] =$$

$$\mathbf{E}[X \mid Y = 1] =$$

$$\mathbf{E}[X \mid Y] =$$

$$\mathbf{E}[X \mid Y = 2] =$$

$$\mathbf{E}[\mathbf{E}[X \mid Y]] =$$

Section means and variances (ctd.)

$$\mathbf{E}[X | Y] = \begin{cases} 90, & \text{w.p. } 1/3 \\ 60, & \text{w.p. } 2/3 \end{cases} \quad \mathbf{E}[\mathbf{E}[X | Y]] = 70 = \mathbf{E}[X]$$
$$\text{var}(\mathbf{E}[X | Y]) =$$

- More data: $\frac{1}{10} \sum_{i=1}^{10} (x_i - 90)^2 = 10$ $\frac{1}{20} \sum_{i=11}^{30} (x_i - 60)^2 = 20$

$$\text{var}(X | Y = 1) = \quad \text{var}(X | Y) =$$

$$\text{var}(X | Y = 2) = \quad \mathbf{E}[\text{var}(X | Y)] =$$

$$\text{var}(X) = \mathbf{E}[\text{var}(X | Y)] + \text{var}(\mathbf{E}[X | Y])$$

$\text{var}(X) = (\text{average variability within sections}) + (\text{variability between sections})$

Sum of a random number of independent r.v.'s

$$\mathbb{E}[Y] = \mathbb{E}[N] \cdot \mathbb{E}[X]$$

- N : number of stores visited
(N is a nonnegative integer r.v.)
- Let $Y = X_1 + \dots + X_N$

- X_i : money spent in store i
 - X_i independent, identically distributed
 - independent of N

$$\mathbb{E}[Y \mid N = n] =$$

- Total expectation theorem:

$$\mathbb{E}[Y] = \sum_n p_N(n) \mathbb{E}[Y \mid N = n]$$

- Law of iterated expectations:

$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y \mid N]]$$

Variance of sum of a random number of independent r.v.'s

$$Y = X_1 + \cdots + X_N$$

- $\mathbf{E}[Y | N] = N \mathbf{E}[X]$

$$\text{var}(\mathbf{E}[Y | N]) =$$

- $\text{var}(Y | N = n) =$

$$\text{var}(Y | N) =$$

$$\mathbf{E}[\text{var}(Y | N)] =$$

$$\text{var}(Y) = \mathbf{E}[\text{var}(Y | N)] + \text{var}(\mathbf{E}[Y | N])$$

$$\text{var}(Y) = \mathbf{E}[N] \text{var}(X) + (\mathbf{E}[X])^2 \text{var}(N)$$