STOR 435.001: Lecture 17

Properties of Expectation - II
Limit Theorems

Jan Hannig

UNC Chapel Hill

Properties of expectation

Recall: For two random variables X and Y, conditional distribution of X given $Y = y$ is characterized by:

Discrete case: conditional p.m.f. $p_{X|Y}(x|y) = \frac{p(x,y)}{p_Y(y)}$

Continuous case: conditional density $f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$

Note: Conditional expectation of X given $Y = y$ is the function

$$g(y) := E(X|Y = y) = \begin{cases} \sum_x x p_{X|Y}(x|y), & \text{discrete case,} \\ \int x f_{X|Y}(x|y) dx, & \text{continuous case.} \end{cases}$$

Notation: $g(Y) = E(X|Y)$.

Computing expectations through conditioning

For random variables X and Y,

$$EX = E(E(X|Y))$$

Point of conditional expectation

1. Super important as a tool for calculating full expectations (previous slide). Can decompose a complicated problem into simpler parts.
2. Very important in it's own right for example for prediction. For example: predict earnings in a company (call this "Y") based on an explanatory variable like current demand, price of oil etc (call this "X"). One good predictor would be $\mathbb{E}(Y|X = x)$.

Example

Jan is planning to goto Durham for a party.

1. He has three options: he could take a bus. Assume that the amount of time it would take him he took a bus $X_{bus} = N(65, \sigma = 10)$.
2. He could take Uber. Assume that $X_{Uber} = N(35, \sigma = 3)$.
3. He could ask his wife. Assume that $X_{wife} = N(20, \sigma = 4)$.

Jan decides on one of these 3 choices uniformly at random. Let T be the amount of time it takes Jan to get to Durham. Find $\mathbb{E}(T)$.

Example: Davis library has the ground floor and 7 other floors which people typically take the elevator for. Assume that the number of people who enter an elevator on the ground floor is a Poisson random variable with mean 10. Suppose each person is equally likely to get off at any one of the 7 floors, independently of where the others get off. Compute the expected number of stops that the elevator will make before discharging all of its passengers.

Conditional expectation

A big supermarket store called “Mousy” (hatip to the movie Zootopia) is having a sale on Monday. Assume that the arrival of customers to Mousy can be modelled as a Poisson process with rate $\lambda = 50$ per hour. Also assume that the amount that each customer that enters the store can be modelled as

1. with probability .90 buy nothing, just browse.
2. with probability $.10$ the amount bought follows a Gamma distribution with parameter $\alpha = 40$ and $\lambda = 2$.

Let S denote the total amount spent in the store in the next 8 hours. Find $\mathbb{E}(S)$.

Properties of expectation

Prediction problem: X is unobservable random variable, Y is observable random variable (related to X). Provide the best guess (predictor) of X based on Y, that is, $g(Y)$ such that $E(X - g(Y))^2$ is smallest.

$E(Y|X)$ and prediction

$E(Y - g(X))^2 \geq E(Y - E(Y|X))^2$, that is, $E(Y|X)$ is the best predictor of Y given X.

Prediction example

In a large Math course with multiple sections, let X denote the score of a randomly selected student in Midterm 1 and Y be the score of the same student in Midterm 2. Suppose (X, Y) can be modelled using a Bivariate normal distribution with $\mu_X = 85, \sigma_X = 10$, $\mu_Y = 75, \sigma_Y = 16$. Further the correlation is $\rho = .8$. If the score on the first midterm namely X is 80 what would be your prediction of the score on the second midterm?

Prediction example

Fitness of individuals

Suppose you are modeling a population of microbes (e.g. virus). Suppose each microbe is born with an inherent random "death rate" Y which let us assume is Uniform $(0, 1)$. Further assume that conditional on the fitness $Y = y$, the lifetime of the microbe is an exponential random variable T with λ parameter $\lambda = y$.

1. Find $\mathbb{E}(T|Y)$ namely if you knew the "death rate" then what would be your best guess of the amount of time the microbe would survive?
2. Find $\mathbb{E}(Y|T)$ namely if you observed the microbe and figured out how long it survived then what would be your best guess of the fitness of the microbe?

Fitness continued

Other functionals of conditional distribution

1. For our class we are going to be happy just thinking about conditional expectation. However in the real world other concepts also very important.
2. Example: **Conditional variance**. If (X, Y) discrete then

$$\text{Var}(X|Y=y) = \sum_x (x - \mathbb{E}(X|Y=y))^2 p_{X|Y}(x|y)$$

3. Continuous:

$$\text{Var}(X|Y=y) = \int_{-\infty}^{\infty} (x - \mathbb{E}(X|Y=y))^2 f_{X|Y}(x|y) dx$$

4. Again both are functions of Y. Just as $\mathbb{E}(X|Y=y)$ is our "best guess" of X if you knew the value of Y, Var$(X|Y=y)$ measures the variability of the random variable X about this best guess. If for a particular value of Y, the above is small, then your best guess is going to be close to the true value of the random variable X. If large then less accurate predictions.