Lecture 27: Conditional Expectation given a Random Variable

Example

Let $X \sim N(0,1)$, $Y = X^2$. Then, $E(Y|X) = ?$

Solu: $E(Y|X) = E(X^2|X) = X^2 = Y$

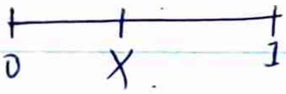
Example Let $X \sim N(0,1)$, $Y = X^2$. Then, find $E(X|Y) = ?$

solu: $E(X|Y) = E(X|X^2) = 0$ since if we observe $X^2 = a$, then $X = \pm\sqrt{a}$ equally likely, Average will be zero.

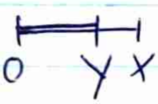
Example

stick of length 1, break off a random piece, break off another piece. find expected value or the conditional expectation, for the length of the second piece.

Solu:



$X \sim \text{Unif}(0,1)$



$Y|X \sim \text{Unif}(0,X)$ If we know $X = x$, then it's going to be $\text{Unif}(0,x)$

$E(Y|X=x) = x/2$ (on average)

So, $E(Y|X) = X/2$

$$E(E(Y|X)) = \frac{1}{2} E(X)$$
$$= 1/4$$
$$= E(Y) \text{ (Adam's Law)}$$

$E(X) = \frac{1}{2}$

Since $X \sim \text{Unif}(0,1)$

Useful Properties

① $E(h(X)\,Y|X) = h(X)\,E(Y|X)$ [X is Known so, $h(X)$ is a constant Taking out what's Known]

Example $Y = X^2$

$$E(Y|X) = E(X^2|X) = X^2 E(1|X)$$
$$= X^2$$

② $E(Y|X) = E(Y)$ if X, Y are independent.

③ $E(E(Y|X)) = E(Y)$ [Iterated Expectation/Adam's Law]

④ $E((Y - \underbrace{E(Y|X)}_{\text{Our prediction}})\,h(X)) = 0$, i.e., $\underbrace{Y - E(Y|X)}$ (Residual)

Actual value of Y

How far off is the prediction?

predicted value of Y

Any function of X

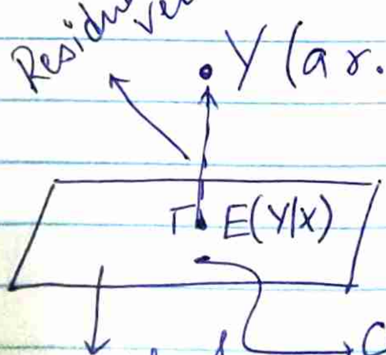
is uncorrelated with any function of X

What's left over after we try to predict Y

$$Cov(Y - E(Y|X), h(X)) = E((Y - E(Y|X))\,h(X)) - \underbrace{E(Y - E(Y|X))}_{0}\,E(h(X))$$

$$E(Y) - E(E(Y|X))$$
$$E(Y) - E(Y)$$

$$= E((Y - E(Y|X))\,h(X))$$



Y (a r.v., we are treating a r.v. (is a function) as a point or vector)

Conditional expectation is a projection

$E(Y|X) \rightarrow$ is a function of X, a point in the plane, that is closest to Y

Consists of all possible function of X (like X, X^2, e^X, etc.) (Collection of r.v.s.)

goes through origin since one function of X is 0.

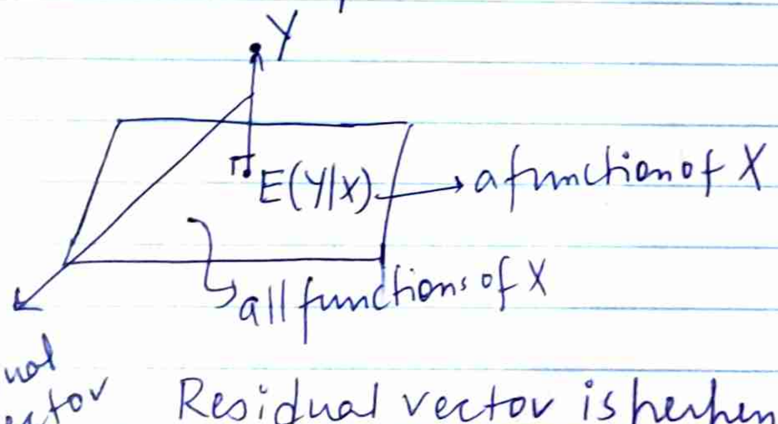
<u>Note</u> :- If Y is already a function of X, then $E(Y|X) = Y$, because that says if it's already in the plane, we don't need to project it anywhere.

But, if Y is not already a function of X, then we are projecting it down to whatever function of X is closest in a certain sense.

$$\langle X, Y \rangle = E(XY)$$

inner product
(fancy word for dot product)

The only assumption here is that we're working with function of X. All random variables we want to assume have finite variance.



Residual vector is perpendicular to the plane
So, any function of X ~~vec~~ is perpendicular to the residual vector.

$$\therefore E\big(\underbrace{(Y - E(Y|X))}_{\text{Residual vector}}\, \underbrace{h(X)}_{\text{Any function of } X}\big) = 0$$

Proof of property (4)

$$E((Y - E(Y|X))h(X))$$
$$= E(Yh(X)) - E(E(Y|X)h(X)) \quad \text{(Using Linearity)}$$
$$= E(Yh(X)) - E(E(Yh(X)|X)) \quad \text{(putting back what's known. since } X \text{ is known, therefore } h(X) \text{ is known)}$$
$$= E(Yh(X)) - E(Yh(X))$$
[Adam's Law, $E(E(Y|X)) = E(Y)$] → Property (3)
$$= 0$$

Proof of property (3) [discrete case]

Let $E(Y|X) = g(X)$

$$E(g(X)) = \sum_x g(x) P(X=x) \quad \text{[LOTUS]}$$
$$= \sum_x E(Y|X=x) P(X=x)$$
$$= \sum_x \left(\sum_y y P(Y=y|X=x) \right) P(X=x)$$
$$= \sum_y \sum_x y P(Y=y, X=x)$$
(Joint PMF = Conditional PMF × Marginal PMF)
$$= \sum_y y \sum_x P(Y=y, X=x)$$
$$= \sum_y y P(Y=y) \quad \text{[Summing up over } X \text{ gives us the marginal distribution of } Y\text{]}$$
$$= E(Y)$$

Conditional Variance

$$Var(Y|X) = E(Y^2|X) - (E(Y|X))^2$$
$$= E((Y - E(Y|X))^2|X)$$

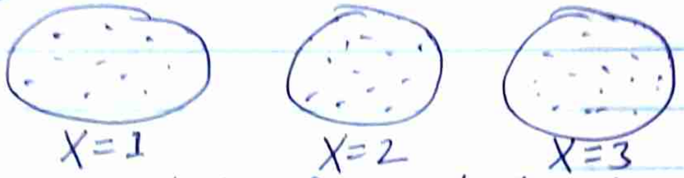
[Everything is based on the assumption that we know X]

Property (5)

$$Var(Y) = \underbrace{E(Var(Y|X))}_{\text{within}} + \underbrace{Var(E(Y|X))}_{\text{between}}$$

[EVE's Law] → EVVE

Example



3 groups, lots of people inside each group. Just imagine Y as height. We have some population of people which consists of 3 subpopulations. We want to know the mean and variance of the heights of people in this population. And, each ~~x~~ sub-population may have its own mean and variance. We want the mean & variance of overall population.

X {takes 3 values, namely 1, 2 or 3} ⟶ If we take a random person from this population, which subpopulation are they in.

$E(Y|X=1)$ → mean for subpopulation 1.

There are 2 types of variability going on in above example:

(i) Different subpopulations may have difference in height, so we have differences between populations.

(ii) Variability within each population. So, within each population, unless everyone in that subpopulation is the same height, we have variability within each of these subpopulations

$Var(E(Y|X))$ → saying look at the average within each population, and then take the average, take the variance of those numbers. So, that's looking between populations.

$E(Var(Y|X))$ → Look within each population. This says look within each population, take its variance and then average those numbers.

→ Replace each population by just its average height and take the variance of those.

Example pick a random city, pick random sample of n people in that city.

$X = \#$ people with disease

$Q =$ proportion of people in the random city with disease.

Different cities have different prevalences of the disease and we are picking a random city. So, Q is a random variable/probability.

Find $E(X)$, $Var(X)$, assuming $Q \sim Beta(a,b)$.
→ flexible

$X|Q \sim Binomial(n, Q)$
(Once we know what proportion of people in that city have the disease, then we are doing binomial)

Expected value of Bin(n,Q) = nQ

Solu: $E(X) = E(E(X|Q)) = E(\underset{\text{Beta}(a,b)}{nQ}) = \frac{na}{a+b}$

$$Var(X) = E(\underbrace{Var(X|Q)}) + Var(E(X|Q))$$

Bin(n,Q) if assuming Q as const. $var = nQ(1-Q)$

$$= E(nQ(1-Q)) + Var(nQ)$$

$$= E(nQ(1-Q)) + n^2 Var(Q)$$

$$= nE(Q(1-Q)) + n^2 Var(Q)$$

$$E(Q(1-Q)) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 q^a (1-q)^b \, dq$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a+1)(b+1)}{\Gamma(a+b+2)}$$

$\Gamma(x+1) = x\Gamma(x)$

$$= \frac{ab\cancel{\Gamma(a+b)}}{(a+b+1)(a+b)\cancel{\Gamma(a+b)}}$$

$$= \frac{ab}{(a+b)(a+b+1)}$$

$$Var(Q) = \frac{\mu(1-\mu)}{a+b+1}, \mu = a/_{a+b}.$$