[5.4] Conditional Expectation

DEF: Conditional Exp.

$$E[Y|x] = \sum_{\text{all } y} y\, f(y|x) \quad (X \,\&\, Y \text{ are discrete})$$

$$= \int_{-\infty}^{\infty} y\, f(y|x)\,dy \quad (X \,\&\, y \text{ are cont.})$$

other Notations

$E[Y|X=x]$ $\qquad E_{Y|X}(Y)$

Note $E[Y|x]$ is a function of x!

$$g(x) = E[Y|x]$$

Thm 5.4.1 If $X \,\&\, Y$ are jointly dist R.V.'s

then
$$E_X[E(Y|X)] = E(Y)$$
or $E(Y) = E_X[E(Y|X)]$

In the continuous case,

$$E(E(Y|X)) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} y f(y|x) dy \right] f(x) dx$$

and

$$E(Y) = \int_{-\infty}^{\infty} y f_2(y) dy$$

proof:

$$E_X[E(Y|X)] = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} y f(y|x)dy\right] f_X(x)dx$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\, \underbrace{f(y|x)f_X(x)}_{f(x,y)}\, dydx$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f(x,y)dxdy$$

$$= \int_{-\infty}^{\infty} y \underbrace{\left[\int_{-\infty}^{\infty} f(x,y)dx\right]}_{f_Y(y)} dy$$

$$= \int_{-\infty}^{\infty} y f_Y(y)dy \quad = E(Y)$$

Ex: 5.4.2

Suppose the number of misspelled words in a student's term paper is distributed as a Poisson Distribution with mean 20. Your roommate, on average, finds 85% of such errors. We have a hierarchical structure like:

$$Y|X \sim \text{BIN}(X, 0.85)$$
$$X \sim \text{POI}(20)$$

Two ways to find E(Y):

1) Find pdf of Y and then compute the mean of Y.
2) E(Y) = E(E(Y|X))

To do the first, we need to find the joint pdf and integrate out the X's (this is not needed if you just want the mean)

Look up the mean and variance for the Poisson and the Binomial, then use it (and properties of Expectations and variances) to get the answer:

E(X) = 20 Var(X) = 20
E(Y|X) = np = X(0.85) Var(Y|X) = npq = X(0.85)(1-0.85)

Then:

$$E(Y) = E(E(Y|X)) = E[X(0.85)] = 0.85E[X]$$
$$= (0.85)(20) = 17$$

Thm 5.4.2 If X&Y are independent,

$$E(Y|x) = E(Y)$$
$$E(X|y) = E(X)$$

Def: The conditional variance of Y given X=x is

$$\operatorname{Var}(Y|x) = E\{[Y - E(Y|x)]^2|x\}$$
$$= E\{Y^2|x\} - \{[E(Y|x)]^2\}$$

Thm 5.4.3 If X&Y are jointly dist, then

$$\operatorname{Var}(Y) = E_X[\operatorname{Var}(Y|X)] + \operatorname{Var}_X[E(Y|X)]$$

EX:

$$\operatorname{Var}(Y) = E_X[\operatorname{Var}(Y|X)] + \operatorname{Var}(E_X(Y|X))$$
$$= E_X[X(0.85)(0.15)] + \operatorname{Var}(X(0.85))$$
$$= (0.85)(0.15)E(X) + (0.85)^2\operatorname{Var}(X)$$
$$= (0.85)(0.15)(20) + (0.85)^2(20)$$
$$= 2.55 + 14.45$$
$$= 17$$

5.4 cont

extension of thm 5.4.1

Thm 5.4.4 If X and Y are jointly dist R.V. If X and Y are jointly dist R.V. $h(x,y)$ is a function, then

$$E[h(X,Y)] = E_X\left[E[h(X,Y)|X]\right]$$

This says we can find the expectation of $h(X,Y)$ by first finding

$$E\left[h(x,Y)|x\right]$$

and then taking the expectation wrt X.

Thm 5.4.5 If X & Y are jointly dist R.V.'s, and $g(x)$ is a function, then

$$E[g(X)Y|x] = g(x)E[Y|x]$$

g(x) is "constant" with respect to Y|x

Corollary:

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

"partial" pullout of inner E, but not outside E

Ex: 5.4.3

$$(X,Y) \sim \text{MULT}(n, p_1, p_2)$$

$$X \sim \text{BIN}(n, p_1) \qquad E(X) = np_1$$

$$Y \sim \text{BIN}(n, p_2) \qquad E(Y) = np_2$$

$$Y|x \sim \text{BIN}(n-x, p), p = \frac{p_2}{1-p_1} \qquad E[Y|x] = \frac{(n-x)p_2}{1-p_1}$$

Find the covariance of X & Y

$$E[XY] = E\Big[E[XY|X]\Big] \quad \text{Thm 5.4.1}$$

$$= E\Big[XE[Y|X]\Big] \quad \text{Thm 5.4.5 (Corollary)}$$

$$= E\left[\frac{X(n-X)p_2}{1-p_1}\right]$$

$$= \frac{p_2}{1-p_1}E[X(n-X)]$$

$$= \frac{p_2}{1-p_1}E[nX - X^2]$$

$$= \frac{p_2}{1-p_1}\Big[nE[X] - E[X^2]\Big]$$

Since $\operatorname{Var}(X) = E(X^2) - \mu^2$, then

$$= \frac{p_2}{1-p_1}\Big[nE(X) - \operatorname{Var}(X) - \mu^2\Big]$$

Since $E(X) = np_1$ and $\operatorname{Var}(X) = np_1(1-p_1)$, then

$$= \frac{p_2}{1-p_1}\Big[n^2p_1 - np_1(1-p_1) - n^2p_1^2\Big] \quad \text{factor } np_1$$

$$= \frac{np_1p_2}{1-p_1}\Big[n - 1 + p_1 - np_1\Big]$$

$$= \frac{np_1p_2}{1-p_1}\Big[(n-1) - (n-1)p_1\Big] \quad \text{factor } (n-1)$$

$$= \frac{np_1p_2(n-1)}{1-p_1}[1-p_1]$$

Thus, $E[XY] = n(n-1)p_1p_2$

It follows that the covariance is

$$\operatorname{Cov}(X,Y) = E(XY) - E(X)E(Y)$$

$$= n(n-1)p_1p_2 - (np_1)(np_2)$$

$$= n^2p_1p_2 - np_1p_2 - n^2p_1p_2$$

$$= \boxed{-np_1p_2}$$

Bivariate Normal Distribution

$(X, Y) \sim \mathrm{BVN}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

$$-\infty < \mu_1 < \infty \qquad \sigma_1 > 0 \qquad -1 \leqslant \rho \leqslant 1$$
$$-\infty < \mu_2 < \infty \qquad \sigma_2 > 0$$

μ_1 is the mean of X

μ_2 is the mean of Y

σ_1 is the standard deviation of X

σ_2 is the standard deviation of Y

ρ is the correlation coefficient

Thm 5.4.6

If $E(Y|x)$ is a linear function of x, then

$$E(Y|x) = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1)$$

and

$$E_X\left[V(Y|X)\right] = \sigma_2^2(1 - \rho^2)$$

A pair of continuous R.V.'s X & Y is said to have a bivariate normal distribution if it has the joint pdf

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]\right\}$$

Thm 5.4.7

If $(X, Y) \sim \text{BVN}(\mu_1, \mu_2, \sigma_1^2 \sigma_2^2, \rho)$, then

$$X \sim \text{N}(\mu_1, \sigma_1^2)$$
$$Y \sim \text{N}(\mu_2, \sigma_2^2)$$

and ρ is the correlation coef

Thm 5.4.8

If $(X, Y) \sim \text{BVN}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

1. $Y|x \sim N\left[\mu_2 + \rho \frac{\sigma_2}{\sigma_1}(x - \mu_1), \sigma_2^2(1 - \rho^2)\right]$
2. $X|y \sim N\left[\mu_1 + \rho \frac{\sigma_1}{\sigma_2}(y - \mu_2), \sigma_1^2(1 - \rho^2)\right]$

Note: We learned previously that

if X is independent of Y, then $\text{Cov}(X, Y) = 0$ (or $\rho = 0$)

But, here in Thm 5.4.7, we learn that if $\rho = 0$, then

the joint pdf $f(x, y)$ can be factored into a product of marginal pdfs Hence, X and Y are independent!!!!

pdf of the Normal distribution!

We thus learn that for the NORMAL distribution, independence and uncorrelated are the same!!!

This isn't true in general!!!!