$$[5.4] \text{ Conditional Expectation Note } \mathbb{E}[Y|x] \text{ is a function } yx!$$

$$E[Y|x] = \sum_{alley} y f(y|x) \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

$$= \int_{-\infty}^{\infty} y f(y|x)dy \quad (xdY \text{ are discribe})$$

proof:

$$E_X[E(Y|X)] = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} yf(y|x)dy \right] f_X(x)dx$$
$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y \underbrace{f(y|x)f_X(x)}_{f(x,y)} dydx$$
$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} yf(x,y)dxdy$$
$$= \int_{-\infty}^{\infty} y \left[\underbrace{\int_{-\infty}^{\infty} f(x,y)dx}_{f_Y(y)} \right] dy$$
$$= \int_{-\infty}^{\infty} yf_Y(y)dy = E(Y)$$

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:

 $Y|X \sim BIN(X, 0.85)$

 $X \sim \mathrm{POI}(20)$

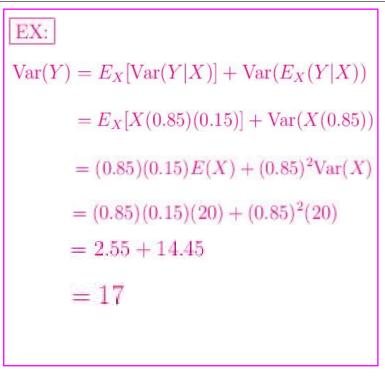
Two ways to find E(Y):
1) Find pdf of Y and then compute the mean of Y.
2) E(Y) = E(E(Y|X))
To do the first, we need to find the joint pdf and integrate out the X's (this is not needed if you just want the mean)

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:

 $\begin{array}{ll} {\sf E}({\sf X})=20 & {\sf Var}({\sf X})=20 \\ {\sf E}({\sf Y}|{\sf X})={\sf np}={\sf X}(0.85) & {\sf Var}({\sf Y}|{\sf X})={\sf npq}={\sf X}(0.85)(1\text{-}0.85) \end{array}$

Then:

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



5.4 cont
tx tension of thm 5.4.1
Thin 5.4.4

$$f = x and Y are jointly dist
tf = x and Y are jointly dist
 $f = x and y are jointly dist
f = x and y are jointly dist
 $f = x and y are jointly dist
f = x and y are jointly dist
 $f = x and y are jointly dist
f = x and y are jointly dist
 $f = x and y are jointly dist
f = x and y are jointly dist
f = x and y are jointly dist
 $f = x and y are jointly dist
f = x and y are jointly dist
 $f = x and y are jointly dist
f = x and y are jointly dist
f = x and y are jointly dist
 $f = x and y are jointly dist
f = y and y are jointly dist
f = y artial pullout of inner
f, but not outside E
f = but not outside E
f = x and y are jointly dist
f = y artial pullout of inner
f, but not outside E
f = y artial pullout of inner
f, but not outside E
f = y artial pullout of inner
f, but not outside E
f = y artial pullout of inner
f, but not outside E
f = y artial pullout of inner
f, but not outside E
f = y artial pullout of inner
f, but not outside E
f = y artial pullout of inner
f = y arti$$$$$$$$$

Find the covariance of $X \And Y$

Since $E(X) = np_1$ and $Var(X) = np_1(1 - p_1)$, then

$$\begin{split} E[XY] &= E\left[E[XY|X]\right] \quad \text{Thm 5.4.1} \\ &= E\left[XE[Y|X]\right] \quad \text{Thm 5.4.5 (Corollary)} \\ &= E\left[\frac{X(n-X)p_2}{1-p_1}\right] \\ &= E\left[\frac{X(n-X)p_2}{1-p_1}\right] \\ &= \frac{p_2}{1-p_1}E[X(n-X)] \\ &= \frac{p_2}{1-p_1}E[X(n-X)] \\ &= \frac{p_2}{1-p_1}E[nX-X^2] \\ &= \frac{p_2}{1-p_1}\left[nE[X] - E[X^2]\right] \\ &= \frac{p_2}{1-p_1}\left[nE[X] - E[X^2]\right] \\ &\text{Since Var}(X) &= E(X^2) - \mu^2, \text{ then} \\ &= \frac{p_2}{1-p_1}\left[nE(X) - \text{Var}(X) - \mu^2\right] \\ &= \frac{p_2}{1-p_1}\left[nE(X) - \text{Var}(X) - \mu^2\right] \\ &= \frac{n^2p_2}{1-p_1}\left[nE(X) - n^2p_1^2\right] \\ &= \frac{n^2p_1p_2(n-1)}{1-p_1}\left[1-p_1\right] \\ &\text{Thus, } E[XY] = n(n-1)p_1p_2 \\ &= n(n-1)p_1p_2 - (np_1)(np_2) \\ &= n(n-1)p_1p_2 - (np_1)(np_2) \\ &= n^2p_1p_2 - np_1p_2 - n^2p_1p_2 \\ &= \frac{-np_1p_2}{1-p_1} \right] \end{split}$$

Bivariate Normal Distribution

- $\begin{array}{l} (X,Y) \sim \mathrm{BVN}(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho) \\ & -\infty < \mu_1 < \infty \qquad \sigma_1 > 0 \\ & -\infty < \mu_2 < \infty \qquad \sigma_2 > 0 \end{array} \quad -1 \leqslant \rho \leqslant 1 \\ & \mu_1 \text{ is the mean of } X \\ & \mu_2 \text{ is the mean of } Y \\ & \sigma_1 \text{ is the standard deviation of } X \\ & \sigma_2 \text{ is the standard deviation of } Y \end{array}$
 - 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

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

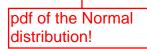
A pair of continuous R.V.'s X & Y is said to have a bivariate normal distribution if it has the joint pdf

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

Thm 5.4.7

If $(X, Y) \sim \text{BVN}(\mu_1, \mu_2, \sigma_1^2 \sigma_2^2, \rho)$, then $X \sim N(\mu_1, \sigma_1^2)$ $Y \sim N(\mu_2, \sigma_2^2)$ and ρ is the correlation coef Note: We learned previously that if X is independent of Y, then 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!!!!

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



This isn't true in general!!!!

Thm 5.4.8

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

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