Joint Distribution

Covariance:

Let $X\ and\ Y$ be random variables with means $\mu_X\ and\ \mu_Y$ respectively. The covariance between $X\ and\ Y$, denoted by $Cov(X,Y)\ or\ \sigma_{XY}$ is given by $Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] = E[XY] - E[X]E[Y]$

Bilinearity of Covariance:

- $Cov(aX,Y) = aCov(X,Y) = Cov(X,aY)$
- $Cov(X_1 + X_2, Y) = Cov(X_1, Y\ ) + Cov(X_2, Y)$

Correlation Coefficient:
Let $X\ and\ Y$ be random variables with means $\mu_X\ and\ \mu_Y$ and variances ${\sigma_X}^2\ and\ {\sigma_Y}^2$respectively. The correlation coefficient between $X\ and\ Y$, denoted by ρ_{XY} is given by

$$\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{(VarX)(VarY)}};$$

Prove that, $|\rho_{XY}| \leq 1$

Proof: if $VarY \neq 0$, let $Z = X - \frac{Cov(X,Y)}{VarY}Y$. Then, we have,

$$0 \leq VarZ = Var\left(X - \frac{Cov(X,Y)}{VarY}Y\right) = VarX + Var\left(-\frac{Cov(X,Y)}{VarY}Y\right) + 2Cov\left(X, -\frac{Cov(X,Y)}{VarY}Y\right)$$

$$= VarX + \left(-\frac{Cov(X,Y)}{VarY}\right)^2 VarY + 2\left(-\frac{Cov(X,Y)}{VarY}\right)Cov(X,Y) = VarX - \frac{\left(Cov(X,Y)\right)^2}{VarY}$$

This is,

$$VarX - \frac{\left(Cov(X,Y)\right)^2}{VarY} \geq 0 \Rightarrow \left(Cov(X,Y)\right)^2 \leq VarX\ VarY$$

$$\Rightarrow |Cov(X,Y)| \leq \sqrt{VarX\ VarY} \Rightarrow |\rho_{XY}| = \left|\frac{Cov(X,Y)}{\sqrt{(VarX)(VarY)}}\right| \leq 1$$

Correlation Coefficient:

Let $X \text{ } and \text{ } Y$ be random variables with means $\mu_X \text{ } and \text{ } \mu_Y$ and variances ${\sigma_X}^2 \text{ } and \text{ } {\sigma_Y}^2$ respectively. The correlation coefficient between $X \text{ } and \text{ } Y$, denoted by ρ_{XY} is given by

$$\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{(VarX)(VarY)}}; \ |\rho_{XY}| \leq 1$$

Example: The joint density of a discrete 2D random variable (X, Y) is presented in the following table. Calculate ρ_{XY}

X\Y	0	1	2	3
0	0.840	0.030	0.020	0.010
1	0.060	0.010	0.008	0.002
2	0.010	0.005	0.004	0.001

$$E[XY] = 0.064; E[X] = 0.12; E[Y] = 0.148; Cov(X,Y) = 0.046;$$
$$E[X^2] = 0.16; E[Y^2] = 0.29; VarX = 0.146; VarY = 0.268; \ \rho_{XY} = 0.23$$

Theorem: Let $X \, and \, Y$ be random variables with correlation coefficient ρ_{XY}.
Then $|\rho_{XY}| = 1 \; if \; and \; only \; if \; Y = \beta_0 + \beta_1 X$ for some real numbers $\beta_0 \; and \; \beta_1 \neq 0$

Assume that $|\rho_{XY}| = 1$, then $\rho_{XY}^2 = 1$. i.e.,

$$\frac{(E[(X-\mu_X)(Y-\mu_Y)])^2}{(E[(X-\mu_X)^2])(E[(Y-\mu_Y)^2])} = 1$$

Let $W = X - \mu_X \; and \; Z = Y - \mu_Y$. Substitute this into the above equation, we have,

$$\frac{(E[WZ])^2}{E[W^2]E[Z^2]} = 1 \Rightarrow (E[WZ])^2 = E[W^2]E[Z^2]$$

Let $a = \frac{E[WZ]}{E[W^2]}$, then,

$$E[(aW-Z)^2] = a^2E[W^2] - 2aE[WZ] + E[Z^2] = \left(\frac{E[WZ]}{E[W^2]}\right)^2 E[W^2] - 2\frac{E[WZ]}{E[W^2]}E[WZ] + E[Z^2]$$

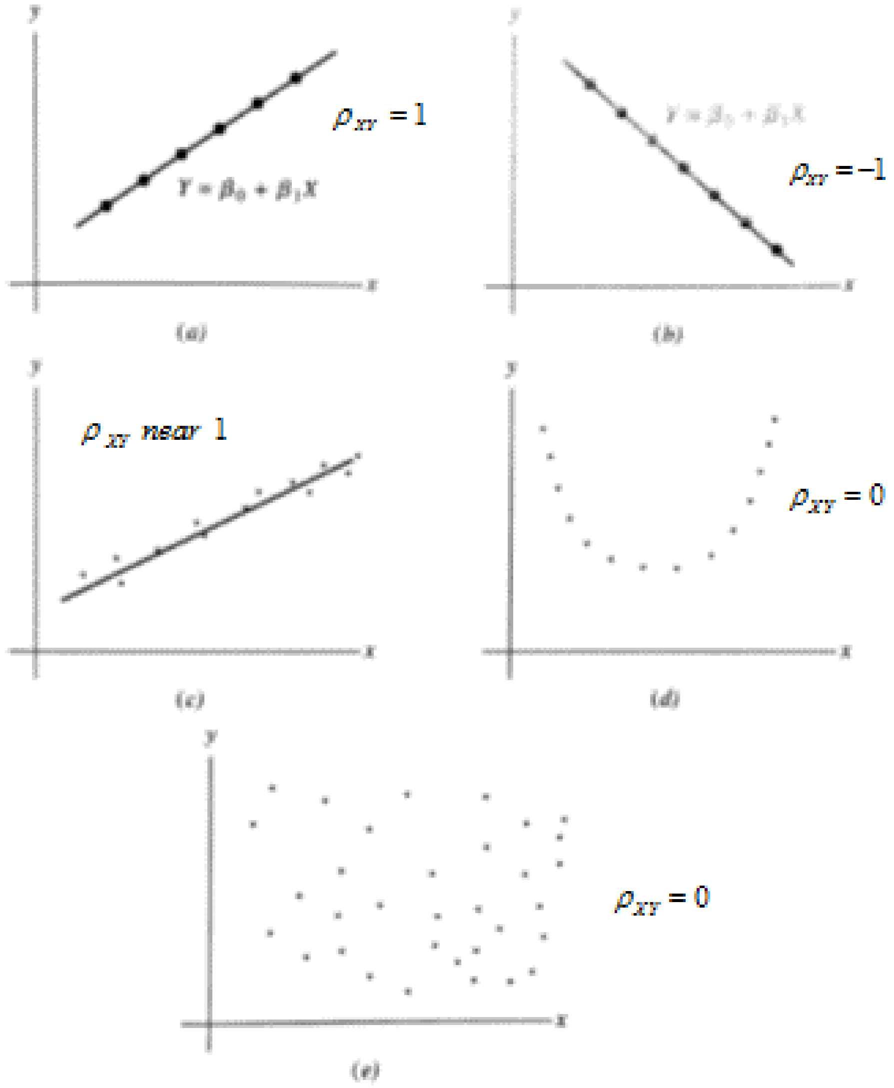
$$= -\frac{(E[WZ])^2}{E[W^2]} + E[Z^2] = \frac{E[W^2]E[Z^2] - (E[WZ])^2}{E[W^2]} = 0$$

Since $(aW-Z)^2 \geq 0$, for the mean of a nonnegative random variable to be 0, the variable must equal 0. i.e.,

$$aW - Z = 0 \Rightarrow a(X-\mu_X) = Y - \mu_Y \Rightarrow Y = \mu_Y - a\mu_X + aX = \beta_0 + \beta_1 X$$

Correlation coefficient:

- If $\rho_{XY} = 1$, then $Y = \beta_0 + \beta_1 X \;\; with\, \beta_1 > 0.\, X \; and \; Y$ have perfectly positive correlation
- If $\rho_{XY} = -1$, then $Y = \beta_0 + \beta_1 X \;\; with\, \beta_1 < 0.\, X \; and \; Y$ have perfectly negative correlation
- If ρ_{XY} has value near 1 or -1, then $X \; and \; Y$ have linear trend
- If $\rho_{XY} = 0$, then $X \; and \; Y$ are not linearly related. $X \; and \; Y$ can be related. $X \; and \; Y$ can be unrelated.



Conditional density:

Consider a 2D random variable (X, Y). We are interested in “the random variable X given that $y = 30$”. $X|y = 30$, itself is a random variable.

Let (X, Y) be a 2D random variable with joint density f_{XY} and marginal density $f_X \ and \ f_Y$. Then:
The conditional density for $X \ given \ Y = y$, denoted by $f_{X|y}$ is given by

$$f_{X|y}(x) = \frac{f_{XY}(x, y)}{f_Y(y)}, \ f_Y(y) > 0$$

The conditional density for $Y \ given \ X = x$, denoted by $f_{Y|x}$ is given by

$$f_{Y|x}(y) = \frac{f_{XY}(x, y)}{f_X(x)}, \ f_X(x) > 0$$

Conditional density:

Example: Given (X,Y) with joint density f_{XY} in the table, calculate $f_{Y|x=1}(y)$ and $f_{X|y=2}(x)$

X\Y	0	1	2	3
0	0.840	0.030	0.020	0.010
1	0.060	0.010	0.008	0.002
2	0.010	0.005	0.004	0.001

$$f_{Y|x=1}(0) = \frac{f_{XY}(1,0)}{f_X(1)} = \frac{0.060}{0.08} = 6/8$$
$$f_{Y|x=1}(1) = \frac{f_{XY}(1,1)}{f_X(1)} = \frac{0.010}{0.08} = 1/8$$
$$f_{Y|x=1}(2) = \frac{f_{XY}(1,2)}{f_X(1)} = \frac{0.008}{0.08} = 1/10$$
$$f_{Y|x=1}(3) = \frac{f_{XY}(1,3)}{f_X(1)} = \frac{0.002}{0.08} = 1/40$$

$$f_{X|y=2}(0) = \frac{f_{XY}(0,2)}{f_Y(2)} = \frac{0.020}{0.032} = 5/8$$
$$f_{X|y=2}(1) = \frac{f_{XY}(1,2)}{f_Y(2)} = \frac{0.008}{0.032} = 1/4$$
$$f_{X|y=2}(2) = \frac{f_{XY}(2,2)}{f_Y(2)} = \frac{0.004}{0.032} = 1/8$$

Conditional density:

Example: Given a random variable (X, Y) with joint density f_{XY}, where,

$$f_{XY}(x,y) = \frac{c}{x}, \qquad 27 \le y \le x \le 33, \qquad c = \frac{1}{6 - 27\ln\left(\frac{33}{27}\right)} = 1.72$$

Find $f_{X|y}(x)\ and\ f_{Y|x}(y)$ and $\mu_{X|y=30} = E[X|y = 30]$

$$f_Y(y) = \int_y^{33} \frac{c}{x} dx = c \ln\frac{33}{y}\,;\ f_{X|y}(x) = \frac{f_{XY}(x,y)}{f_Y(y)} = \frac{c/x}{c\ln\frac{33}{y}} = \frac{1}{x\ln\frac{33}{y}};$$

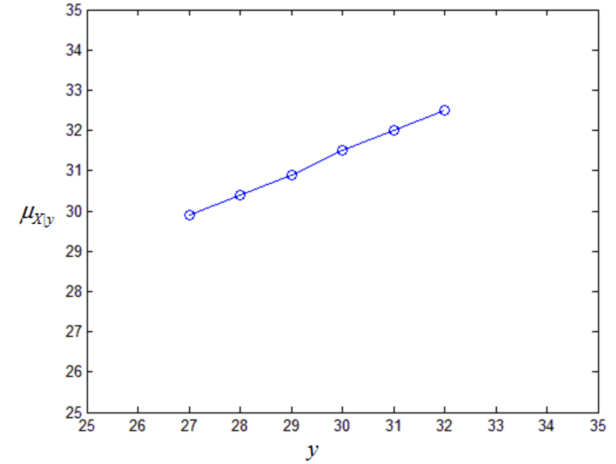
$$f_X(x) = \int_{27}^{x} \frac{c}{x} dy = \frac{c(x-27)}{x};\ f_{Y|x}(y) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{c/x}{c(x-27)/x} = \frac{1}{x-27}$$

$$f_{X|y=30}(x) = \frac{1}{0.0953x}$$

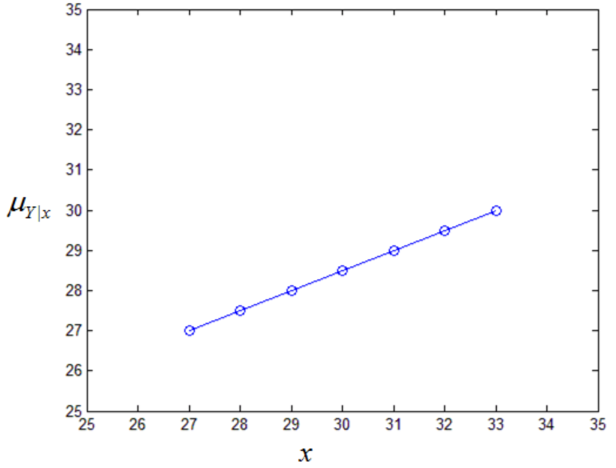
$$\mu_{X|y=30} = \int_{30}^{33} x\, f_{X|y=30}(x) dx = \int_{30}^{33} x \frac{1}{0.0953x} dx = 31.5$$

What is $E[X]$? $E[X] = \mu_{X|y}$?

Curves of Regression:
From the previous example, we can see that, in general, the mean of $X\ given\ Y = y\ or\ \mu_{X|y}$ is a function of y. When this function is graphed, we obtain what is called the curve of regression of $X\ on\ Y$.



Definition of curves of regression:
Let (X, Y) be a 2D random variable. Then,
The graph of the mean value of $X\ given\ Y = y$, denoted by $\mu_{X|y}$, is called the curve of regress of $X\ on\ Y$.
The graph of the mean value of $Y\ given\ X = x$, denoted by $\mu_{Y|x}$, is called the curve of regress of Y $on\ X$.



Example: Given a random variable (X, Y) with joint density f_{XY}, where,

$$f_{XY}(x, y) = \frac{c}{x}, \qquad 27 \le y \le x \le 33, c = \frac{1}{6 - 27\ln\left(\frac{33}{27}\right)} = 1.72$$

Find the curve of regression of $X\ on\ Y$ and the curve of regression of Y $on\ X$

We already found that $f_{X|y}(x) = \frac{1}{x\ln\frac{33}{y}}$ and $f_{Y|x}(y) = \frac{1}{x-27}$. Hence,

$$\mu_{X|y} = \int_y^{33} x\, f_{X|y}(x)dx = \int_y^{33} x \frac{1}{x\ln\frac{33}{y}} dx = \frac{33-y}{\ln\frac{33}{y}}\ ; \quad \mu_{Y|x} = \int_{27}^{x} y f_{Y|x}(y)dy = \int_{27}^{x} y \frac{1}{x-27} dy = \frac{x+27}{2}$$