[5.3] Correlation

		Y			
		-1	0	1	$f_1(x)$
X	-1	0	$\frac{1}{4}$	0	$\frac{1}{4}$
	0	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{2}$
	1	0	$\frac{1}{4}$	0	$\frac{1}{4}$
	$f_2(y)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	

Marginals of X & Y

x	-1	0	1
$f_1(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

y	-1	0	1
$f_2(y)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Note that X and Y are <u>not</u> independent since $f(x,y)$ should be equal to $f_1(x)f_2(y)$ for all x, y

$$\text{but } 0 = f(1,1) \neq f_1(1)f_2(1) = \frac{1}{4}\cdot\frac{1}{4}$$

The mean and variance for X and Y is

$$E(X) = \sum_{\text{all } x} x f_1(x) = (-1)\left(\frac{1}{4}\right) + (0)\left(\frac{1}{2}\right) + (1)\left(\frac{1}{4}\right) = 0$$

$$V(X) = E(X^2) - \mu^2 = E(X^2) = (-1)^2\left(\frac{1}{4}\right) + (0)\left(\frac{1}{2}\right) + (1)^2\left(\frac{1}{4}\right) = \frac{1}{2}$$

$$\begin{aligned} E(XY) = \sum_X \sum_Y xyf(x,y) &= (-1)(-1)(0) + (-1)(0)(\tfrac{1}{4}) + (-1)(1)(0) \\ &+ (0)(-1)(\tfrac{1}{4}) + (0)(0)(0) + (0)(1)(\tfrac{1}{4}) \\ &+ (1)(-1)(0) + (1)(0)(\tfrac{1}{4}) + (1)(1)(0) = 0 \end{aligned}$$

Thus, the covariance is

$$\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y) = 0 - 0(0) = 0$$

Remember! If $X \perp\!\!\!\perp Y$, then $\text{Cov}(X,Y) = 0$

But it does not always work the other way around (meaning the converse is false). The distribution on the previous page is an example of this (The distribution was NOT independent, but the covariance was zero.) We call X & Y uncorrelated instead.

DEF: If X & Y are random variables with variances σ_X^2 & σ_Y^2 and cov σ_{XY} then the correlation coefficient of X & Y is

Greek letter "rho" (it's more like r than p)

$$\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

If $\rho = 0$, then X & Y are said to be uncorrelated.

Thm: If ρ is the correlation coeff of X and Y, then $-1 \le \rho \le 1$

and $\rho = \pm 1$ iff $Y = aX + b$ with prob 1.

for some $a \neq 0$ and b.

Thm 5.3.1

If ρ is the correlation coeff of $X \& Y$,

then $-1 \le \rho \le 1$

and $\rho = \pm 1$ iff $Y = aX + b$ with prob 1.

proof:

$$h(t) = E\left[(X-\mu_X)t + (Y-\mu_Y)\right]^2$$

$$= E\left[(X-\mu_X)^2 t^2 + 2t(X-\mu_X)(Y-\mu_Y) + (Y-\mu_Y)^2\right]$$

$$= t^2 E(X-\mu_X)^2 + 2tE\left[(X-\mu_X)(Y-\mu_Y)\right] + E(Y-\mu_Y)^2$$

Since $h(t) \ge 0$ for all values of t (why?)
$h(t)$ is a quadratic function, $h(t)$ has at most one real root and thus must have a non positive discriminant.

$$b^2 - 4ac \le 0$$

$$\left[2\underbrace{E\left[(X-\mu_X)(Y-\mu_Y)\right]}_{\sigma_{XY}}\right]^2 - 4\underbrace{E(X-\mu_X)^2}_{\substack{V(X)\\ \sigma_X^2}}\underbrace{E(Y-\mu_Y)^2}_{\substack{V(Y)\\ \sigma_Y^2}} \le 0$$

$$4\sigma_{XY}^2 - 4\sigma_X^2\sigma_Y^2 \le 0$$

$$4\sigma_{XY}^2 \le 4\sigma_X^2\sigma_Y^2$$

$$|\sigma_{XY}| \le \sigma_X\sigma_Y$$

$$-\sigma_X\sigma_Y \le \sigma_{XY} \le \sigma_X\sigma_Y$$

$$-1 \le \frac{\sigma_{XY}}{\sigma_X\sigma_Y} \le 1$$

$$-1 \le \rho \le 1$$