Some examples. One mother.

$Y \sim Po(\lambda)$, $X|Y \sim B(Y, p)$

$E[X|Y] = pY \Rightarrow E[X] = E[E[X|Y]] = pE[Y] = p\lambda$

$Var[X|Y] = Yp(1-p) \Rightarrow Var[X] = \lambda p(1-p) + p^2\lambda$

$= \lambda p(1-p+p) = \lambda p$

<u>Randomly chosen mother</u>

$$\left.\begin{aligned} X|Y &\sim B(Y,p) \\ Y|\Lambda &\sim Po(\Lambda) \end{aligned}\right\} \quad X|\Lambda \sim Po(p\Lambda)$$

$$\left.\begin{aligned} Y|\Lambda &\sim Po(\Lambda) \\ \Lambda &\sim exp\left(\tfrac{1}{\beta}\right) \end{aligned}\right\} \quad Y \sim \text{translated negative binomial with } n=1,\ p = \frac{1}{1+\beta}$$

$E[X] = E[E[X|Y]] = E[pY] = pE[Y] = pE[E(Y|\Lambda)] = pE[\Lambda] = p\beta$

$Var[X] = E[Var(X|Y)] + Var[E[X|Y]]$

$= E[Yp(1-p)] + Var[pY]$

$= p(1-p)E[Y] + p^2 Var[Y]$

$= p(1-p)E[E[Y|\Lambda]] + p^2[E[Var(Y|\Lambda)] + Var[E(Y|\Lambda)]$

$= p(1-p)E[\Lambda] + p^2[E[\Lambda] + Var[\Lambda]]$

$= p\beta - p^2\beta + p^2(\beta + \beta^2) = p\beta(1+p\beta)$

<u>Negative Binomial</u> with $n=1$

$Var[Y] = \mu(1+\mu)$

(In general $\mu(1+\frac{\mu}{n})$)

$p^* = \frac{1}{1+p\beta}$.

For covariance

$E[XY] = E[E[XY|X]] = E[X E[Y|X]]$.

4.5 Covariance and Correlation

Definition 4.5.1

The covariance of X and Y is defined by

$$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)] = E[XY] - E[X]\cdot E[Y]$$

Definition 4.5.2

The correlation (correlation coefficient) of X an Y is the number defined by $\rho_{X,Y} = \dfrac{Cov(X,Y)}{\sigma_X \sigma_Y}$

Consider the function $h(t) = E[((X-\mu_X)t + (Y-\mu_Y))^2]$

$$= t^2\sigma_X^2 + 2t\, Cov(X,Y) + \sigma_Y^2$$

$$h'(t) = 2t\sigma_X^2 + 2\, Cov(X,Y) = 0 \Rightarrow t = -\frac{Cov(X,Y)}{\sigma_X^2}$$

$$h''(t) = \sigma_X^2 > 0 \Rightarrow \text{minimum}$$

Substituted for t gives

$$h(t) = \frac{Cov(X,Y)^2\sigma_X^2}{\sigma_X^4} - \frac{2\, Cov(X,Y)^2}{\sigma_X^2} + \sigma_Y^2$$

$$= -\frac{Cov(X,Y)^2}{\sigma_X^2} + \sigma_Y^2 \geq 0$$

or

$$Cov(X,Y)^2 \leq \sigma_X^2\sigma_Y^2 \iff -\sigma_X\sigma_Y \leq Cov(X,Y) \leq \sigma_X\sigma_Y$$

$$\Rightarrow \quad |\rho_{X,Y}| \leq 1$$

Since $((X-\mu_X)t + (Y-\mu_Y))^2 \geq 0$ we have

$E[((X-\mu_X)t + (Y-\mu_Y))^2] = 0$

$h(t) = 0 \iff P[((X-\mu_X)t + (Y-\mu_Y))^2 = 0] = 1$.

$\Rightarrow P[(X-\mu_X)t + (Y-\mu_Y) = 0] = 1 \iff$

$P(Y = -tX + \mu_X t + \mu_Y) = 1$

Theorem 4.5.7

X, Y random variables

a) $-1 \leq \rho_{X,Y} \leq 1$

b) $|\rho_{X,Y}| = 1 \iff P(Y = aX + b) = 1$

$\rho_{X,Y} > 0 \Rightarrow t$ negative and a positive.

Example 4.5.9

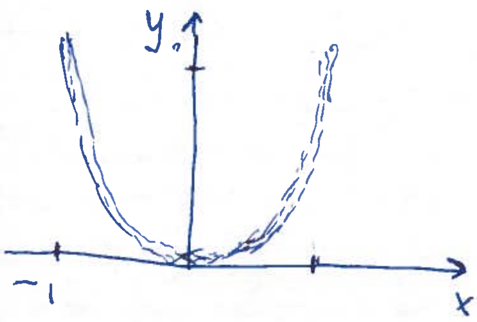
$X \sim$ uniform $[-1,1]$ — independent

$Z \sim$ uniform $[0, \frac{1}{10}]$

$Y = X^2 + Z$

$Cov(X,Y) = E[(X^2+Z)X] - E[X^2+Z] \cdot E[X]$

$= E[X^3] + E[Z] \cdot E[X] = E[X^3] = \frac{1}{2}\int_{-1}^{1} x^3\,dx = 0$



Correlation does not measure non linear relationship.

4.7. Inequalities.

Lemma 4.7.1 Let a, b be positive numbers and let p and q be any positive numbers (>1) satisfying

$$\frac{1}{p} + \frac{1}{q} = 1 \quad \Rightarrow \quad \left(q = \frac{p}{p-1} \right)$$

Then $\frac{1}{p} a^p + \frac{1}{q} b^q \geq ab$ with equality if and only if $a^p = b^q$

Proof. For given b let $g(a) = \frac{1}{p} a^p + \frac{1}{q} b^q - ab$

$$\frac{d}{da} g(a) = a^{p-1} - b = 0 \quad \Rightarrow \quad a = b^{\frac{1}{p-1}} \quad \text{or} \quad b = a^{p-1}$$

$$\frac{d^2}{da^2} g(a) > 0 \quad \Rightarrow \quad \text{minimum.}$$

$$g(b^{\frac{1}{p-1}}) = \frac{1}{p} b^{\frac{p}{p-1}} + \frac{1}{q} b^q - b^{\frac{1}{p-1}} \cdot b$$

$$= \frac{1}{p} b^q + \frac{1}{q} b^q - b^q = b^q \underbrace{\left(\frac{1}{p} + \frac{1}{q}\right)}_{1} - b^q = 0$$

$\Rightarrow$ $g(a) \geq 0$ and the ~~so~~ result follows.

$$g(a) = 0 \Rightarrow a^p = b^{\frac{p}{p-1}} = b^q.$$

Theorem 4.7.2

Hölder's inequality. X, Y random variables and $p > 0$, $q > 0$.

satisfies $\frac{1}{p} + \frac{1}{q} = 1$

Then
$$|E[XY]| \leq E|XY| \leq (E|X|^p)^{\frac{1}{p}} \cdot (E|Y|^q)^{\frac{1}{q}}$$

Proof. $-|XY| \leq XY \leq |XY| \Rightarrow -E(|XY|) \leq E[XY] \leq E[|XY|]$

$$\Leftrightarrow |E(XY)| \leq E(|XY|)$$

Let $a = \frac{|X|}{(E|X|^p)^{\frac{1}{p}}}$ and $b = \frac{|Y|}{(E|Y|^q)^{\frac{1}{q}}}$

From lemma 4.7.1

$$\underbrace{E\left[\frac{1}{p}\frac{|X|^p}{(E|X|)^p} + \frac{1}{q}\frac{|Y|^q}{E|Y|^q}\right]}_{1} \geq \frac{E(|XY|)}{(E|X|^p)^{\frac{1}{p}}(E|Y|^q)^{\frac{1}{q}}}$$

$$\Rightarrow E(|XY|) \leq (E|X|^p)^{\frac{1}{p}}(E|Y|^q)^{\frac{1}{q}}$$

In particular Theorem 4.7.3 Cauchy Schwartz

$$|E(XY)| \leq E(|XY|) \leq (E|X^2|)^{\frac{1}{2}}(E|Y|^2)^{\frac{1}{2}}$$

or $(Cov(X,Y))^2 \leq \sigma_X^2 \sigma_Y^2$.