Session 22

Recall...

22.1

Defn: Given two j-dist RVs X and Y, the correlation between X and Y is defined as

$$\mathrm{corr}(X,Y) \triangleq E[XY];$$

the covariance between X and Y is defined as

$$\mathrm{cov}(X,Y) \triangleq E[(X-\overline{X})(Y-\overline{Y})];$$

The correlation coefficient between X and Y is defined as

$$r_{XY} \triangleq \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y}$$

Recall...

Defn: Two RVs X and Y are **uncorrelated** if their covariance is equal to zero.

This is true if any one of the following equivalent conditions is true:

1. $\operatorname{Cov}(X, Y) = 0$

2. $r_{XY} = 0$

3. $E[XY] = E[X] \cdot E[Y]$.

Defn: Two RVs X and Y are **orthogonal** if $E[XY] = 0$.

<u>Fact:</u> If $E[X^2] < \infty$ and $E[Y^2] < \infty$, then

$$|E[XY]| \le \sqrt{E[X^2] \cdot E[Y^2]},$$

with equality iff

$$Y = a_0 X, \qquad \text{(a.e.)}$$

almost everywhere.

for some constant a_0.

<u>Proof:</u> $E[(aX - Y)^2] \ge 0$

$$\Rightarrow E[a^2X^2 - 2aXY + Y^2] \ge 0$$

$$\Rightarrow E[X^2]a^2 + E[-2XY]a + E[Y^2] \ge 0$$

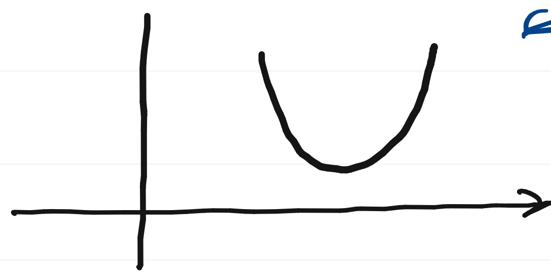
n.b. L.H.S. is a quadratic equation in a.

Let's look at two possible cases:

(i) $E[(aX-Y)^2] > 0$

(ii) $E[(aX-Y)^2] = 0$

(i): $0 < E[(aX-Y)^2] = E[X^2]a^2 - 2E[XY]a + E[Y^2]$



← there are no real roots.

$\Rightarrow$ Quadratic has no real roots in this case

$\Rightarrow$ "discriminant" of the quadratic is negative

$\Rightarrow$ "discriminant" of the quadratic is negative

Aside

$$az^2 + bz + c = 0 \Rightarrow z_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

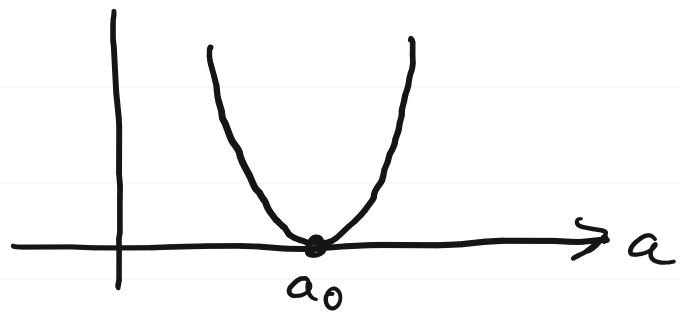
"Discriminant" $= b^2 - 4ac$

$$\Rightarrow \quad 4E^2[XY] - 4E[X^2] \cdot E[Y^2] < 0$$

$$\Rightarrow \quad (E[XY])^2 < E[X^2] \cdot E[Y^2]$$ ∎

(ii) If $E[(aX - Y)^2] = 0$

$\Rightarrow$ for some $a = a_0$, the quadratic has a single real root.



"discriminant" is equal to zero.

$$\Rightarrow \quad a = a_0$$

$$\Rightarrow \quad Y = a_0 X \quad \blacksquare$$

Joint Characteristic Functions

Defn: The joint characteristic function of two j-dist RVs X and Y is

$$\Phi_{XY}(\omega_1, \omega_2) \triangleq E\left[e^{i(\omega_1 X + \omega_2 Y)}\right]$$

$$= \underbrace{\iint_{\mathbb{R}^2} e^{+i(\omega_1 x + \omega_2 y)} f_{XY}(x,y)\, dx\, dy}_{\text{2-Dim. Fourier Transform}}$$

n.b. Inverse Fourier Transform Relationship:

$$f_{XY}(x,y) = \frac{1}{(2\pi)^2} \iint_{\mathbb{R}^2} \Phi_{XY}(\omega_1,\omega_2)\, e^{-i(\omega_1 x + \omega_2 y)}\, d\omega_1\, d\omega_2$$

Note:

1. $\Phi_X(\omega) = \Phi_{XY}(\omega, 0)$

   $\Phi_Y(\omega) = \Phi_{XY}(0, \omega).$

2. If $Z = aX + bY$

$$\Phi_Z(\omega) = E[e^{i\omega Z}] = E[e^{i\omega(aX+bY)}]$$

$$= E[e^{i[(a\omega)X + (b\omega)Y]}]$$

$$= \Phi_{XY}(a\omega, b\omega)$$

Theorem ("Convolution Theorem")

Let $\underline{X}$ and $\underline{Y}$ be two j-dist statistically independent random variables, and let $\underline{Z} = \underline{X} + \underline{Y}$. Then

$$\underline{\Phi}_{\underline{Z}}(\omega) = \underline{\Phi}_{\underline{X}}(\omega) \cdot \underline{\Phi}_{\underline{Y}}(\omega).$$

Proof:

$$\underline{\Phi}_{\underline{Z}}(\omega) = E\left[e^{i\omega \underline{Z}}\right] = E\left[e^{i\omega(\underline{X}+\underline{Y})}\right]$$

$$= E\left[e^{i\omega \underline{X}} \cdot e^{i\omega \underline{Y}}\right] \quad \text{---} \; (*)$$

$$= E\left[e^{i\omega \underline{X}}\right] \cdot E\left[e^{i\omega \underline{Y}}\right]$$

$$= \underline{\Phi}_{\underline{X}}(\omega) \cdot \underline{\Phi}_{\underline{Y}}(\omega) \quad \blacksquare$$

(*): <u>n.b.</u>

$$E[e^{i\omega X}\cdot e^{i\omega Y}] = \iint_{\mathbb{R}^2} e^{i\omega x}\cdot e^{i\omega y}\, f_{X,Y}(x,y)\,dx\,dy$$

$\Downarrow$ $X \perp\!\!\!\perp Y$

$$= \iint_{\mathbb{R}^2} e^{i\omega x}\cdot e^{i\omega y}\, f_X(x)\cdot f_Y(y)\,dx\,dy$$

$$= \int_{\mathbb{R}} e^{i\omega x} f_X(x)\,dx \cdot \int_{\mathbb{R}} e^{i\omega y} f_Y(y)\,dy$$

$$= E[e^{i\omega X}]\cdot E[e^{i\omega Y}]$$

<u>Fact</u>: The joint characteristic function of two jointly Gaussian random variables X and Y with j-pdf

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-r^2}} \exp\left\{\frac{-1}{2(1-r^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} - 2r\frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right]\right\}$$

is

$$\Phi_{XY}(\omega_1,\omega_2) = e^{i(\mu_X\omega_1+\mu_Y\omega_2)}\, e^{-\frac{1}{2}\left[\sigma_X^2\omega_1^2 + 2r\sigma_X\sigma_Y\omega_1\omega_2 + \sigma_Y^2\omega_2^2\right]}.$$

<u>Proof</u>: (You could prove this using the definition — 2D Fourier transform of a 2D Gaussian.)

The following approach is easier:

- It can be shown that if

$$\mathbb{Z} = w_1 \mathbb{X} + w_2 \mathbb{Y} \quad , \quad w_1, w_2 \in \mathbb{R}$$

and $\mathbb{X}$ and $\mathbb{Y}$ are jointly Gaussian, then $\mathbb{Z}$ is a Gaussian RV.

- If $\mathbb{Z}$ is a Gaussian RV, then it has characteristic function

$$\Phi_{\mathbb{Z}}(\omega) = e^{i \mu_Z \omega} \, e^{-\frac{1}{2} \sigma_Z^2 \omega^2}$$

where

$$\mu_Z = w_1 \mu_X + w_2 \mu_Y$$

and

$$\sigma_Z^2 = w_1^2 \sigma_X^2 + 2 r w_1 w_2 \sigma_X \sigma_Y + w_2^2 \sigma_Y^2 .$$

So we have

$$\Phi_Z(\omega) = e^{i\omega(\mu_X \omega_1 + \mu_Y \omega_2)} \cdot e^{-\frac{1}{2}\omega^2(\sigma_X^2 \omega_1^2 + 2r\omega_1\omega_2\sigma_X\sigma_Y + \sigma_Y^2\omega_2^2)}$$

$$= E\left[e^{i\omega(\omega_1 X + \omega_2 Y)}\right] \quad \text{---} \quad (*)$$

But

$$\Phi_Z(\omega)\Big|_{\omega=1} = E\left[e^{i(\omega_1 X + \omega_2 Y)}\right]$$

$$= e^{i(\omega_1\mu_X + \omega_2\mu_Y)} e^{-\frac{1}{2}\left[\sigma_X^2\omega_1^2 + 2r\,\sigma_X\sigma_Y\omega_1\omega_2 + \sigma_Y^2\omega_2^2\right]}$$

Defn: The joint moment generating function (mgf) of two j-dist RVs X and Y is

$$\phi_{XY}(s_1, s_2) \triangleq E\left[e^{s_1 X + s_2 Y}\right],$$

where $s_1, s_2 \in \mathbb{R}$ (or $s_1, s_2 \in \mathbb{C}$.)

This is a 2-dimensional bilateral Laplace transform.

Moment Theorem: $E[X^j \cdot Y^k]$ can be computed as

$$E[X^j \cdot Y^k] = \frac{\partial^j \partial^k}{\partial s_1^j \, \partial s_2^k} \left(\phi_{XY}(s_1, s_2) \right) \Bigg|_{\substack{s_1=0 \\ s_2=0}}$$

$$= \phi_{XY}^{(j,k)}(0,0).$$

Proof: Straightforward extension of the one-dimensional case. (exercise.)

Conditional Distributions

Defn: Let X and Y be two j-dist. RVs on $(\mathcal{S}, \mathcal{F}, P)$. The **joint conditional cdf of X and Y conditioned on $M \in \mathcal{F}$** is

$$F_{XY}(x,y|M) \triangleq P(\{X \le x\} \cap \{Y \le y\} | M)$$

$$= \frac{P(\{X \le x\} \cap \{Y \le y\} \cap M)}{P(M)}.$$

Sometimes, we can describe the event M in terms of X and/or Y.

Example: Let $M \triangleq \{x_1 < X \le x_2\}$.

Find $F_Y(y|M)$.

$$F_Y(y|M) = F_Y(y \mid \{x_1 < X \le x_2\})$$

$$= P(\{Y \le y\} \mid \{x_1 < X \le x_2\})$$

$$= \frac{P(\{Y \le y\} \cap \{x_1 < X_1 \le x_2\})}{P(\{x_1 < X \le x_2\})}$$

$$= \frac{F_{XY}(x_2, y) - F_{XY}(x_1, y)}{F_X(x_2) - F_X(x_1)}$$

$$\therefore F_Y(y \mid \{x_1 < X \le x_2\}) = \frac{F_{XY}(x_2, y) - F_{XY}(x_1, y)}{F_X(x_2) - F_X(x_1)}.$$

If we differentiate this w.r.t. y, this becomes

$$f_Y(y \mid \{x_1 < X \le x_2\}) = \frac{\partial}{\partial y} F_Y(y \mid \{x_1 < X \le x_2\})$$

$$= \cdots = \frac{\int_{x_1}^{x_2} f_{XY}(x,y)\,dx}{F_X(x_2) - F_X(x_1)} \quad --- \quad (*)$$

Of particular interest is the case of

$$f_Y(y \mid \{X = x\}) = \lim_{\Delta x \to 0} f_Y(y \mid \{x < X \le x + \Delta x\})$$

So we use $(*)$ and take

$$x_1 = x$$
$$x_2 = x + \Delta x$$

Then

$$f_{Y}(y \mid \{x < X \le x+\Delta x\}) \overset{(*)}{=} \frac{\int_{x}^{x+\Delta x} f_{X,Y}(\alpha, y)\, d\alpha}{F_X(x+\Delta x) - F_X(x)}$$

and

$$f_{Y}(y \mid \{X = x\}) = \lim_{\Delta x \to 0} f_{Y}(y \mid \{x < X \le x+\Delta x\})$$

$$= \lim_{\Delta x \to 0} \frac{\frac{1}{\Delta x} \int_{x}^{x+\Delta x} f_{X,Y}(\alpha, y)\, d\alpha}{\frac{1}{\Delta x}\left(F_X(x+\Delta x) - F_X(x)\right)}$$

$$= \frac{\frac{1}{\Delta x}\left(f(x+\Delta x, y) - f(x, y)\right)}{\frac{1}{\Delta x}\left(F_X(x+\Delta x) - F_X(x)\right)}$$

$$= \frac{f_{X,Y}(x, y)}{f_X(x)}$$

where $\frac{\partial}{\partial x} f(x, y) = f_{X,Y}(x, y)$.

Thus

$$\boxed{f_Y(y \mid \{X = x\}) = \frac{f_{XY}(x,y)}{f_X(x)}}$$

Similarly, by symmetry

$$\boxed{f_X(x \mid \{Y = y\}) = \frac{f_{XY}(x,y)}{f_Y(y)}}$$

We will sometimes use the notation

$$f_X(x \mid \{Y = y\}) = f_X(x \mid y) = f(x \mid y)$$

$$f_Y(y \mid \{X = x\}) = f_Y(y \mid x) = f(y \mid x)$$