Expected values of 2RVs and Joint moments

(Chapter 5, 6)

With 2 random variables, we care how they vary together.

Definition

Let $Z = g(X,Y)$ be any function of X and Y

Then $$E(Z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, f_{XY}(x,y)\, dy\, dx$$

or $$E(Z) = \sum_{\text{all } x}\sum_{\text{all } y} g(x,y)\, p_{XY}(x,y)$$

Some useful functions $Z = g(X,Y)$

$Z = X$

$Z = Y$

$Z = X + Y$ (sum of 2 RVs)

$Z = XY$ (product of 2 RVs)

$Z = g_1(X)\, g_2(Y)$ (product of 2 separable functions)

$Z = X^i Y^j$ (a joint moment of X and Y)

(example $Z = X^2 Y^2$)

$Z = (X - \mu_X)^i\, (Y - \mu_Y)^j$ (a joint centralized moment of X and Y)

Example: $Z = X + Y$

Easy to show that $\boxed{E(Z) = E(X) + E(Y)}$

proof: $E(Z) = E(X+Y) = \iint (x+y) f_{xy}(x,y)\,dx\,dy$

$= \iint x f_{xy}(x,y)\,dy\,dx + \iint y f_{xy}(x,y)\,dx\,dy$

$= \int x f_x(x)\,dx + \int y f_y(y)\,dy$

$= E(X) + E(Y)$

$\boxed{\text{Also, } E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n)}$

If X and Y are **independent**, and $Z = g(X, Y)$ can be decomposed into $Z = g_1(X) g_2(Y)$

then $E(Z) = E(g_1(X)) E(g_2(Y))$

In particular, $E(XY) = E(X) E(Y)$

proof $E(g_1(X) g_2(Y)) = \iint g_1(x) g_2(y) f_{xy}(x,y)\,dy\,dx$

$= \left[\int g_1(x) f_x(x)\,dx \right] \left[\int g_2(y) f_y(y)\,dy \right]$

$= E(g_1(X)) E(g_2(Y))$

Example: $Z = \exp(X) Y^5$

$E(Z) = E(\exp(X)) E(Y^5)$

if and only if X and Y are independent

Joint moments of X and Y

Recall : $E(X^n)$ is the n^{th} moment of X

$E((X-\mu_X)^n)$ is the n^{th} central moment of X

Joint moments $E(X^n Y^m)$

Joint central moments $E[(X-\mu_X)^n (Y-\mu_Y)^m]$

These summarize information about the joint behavior of X and Y

The more joint moments we know, the more we know about X and Y jointly

- Because while the joint PDF tells us everything about this, in practical situations it may be difficult to obtain or estimate.

Correlation $E(XY) = r_{XY}$ $n=1$ and $m=1$

If $E(XY)=0$, X and Y are said to be <u>orthogonal</u> but they may or may not be <u>uncorrelated</u>!

Covariance $E((X-\mu_X)(Y-\mu_Y)) = \mathrm{cov}(X,Y) = C_{XY}$

If $\mathrm{cov}(X,Y) = 0$, X and Y are said to be <u>uncorrelated</u>

Correlation coefficient

$$\rho_{xy} = \frac{\mathrm{cov}(X,Y)}{\sigma_x \sigma_y} = \frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{VAR}(X)\,\mathrm{VAR}(Y)}} = \frac{C_{xy}}{\sigma_x \sigma_y}$$

Note $-1 \le \rho_{xy} \le 1$

$\rho_{xy} = 0$ uncorrelated

$\rho_{xy} > 0$ positively correlated

$\rho_{xy} < 0$ negatively correlated

Be careful with the English terminology. These are mathematical definitions. When people, papers, and the press talk about the correlation, they often mean the correlation coefficient!

Q: Why covariance and not correlation?

A: Sometimes you don't care about the mean.

Ex: in a circuit, amplifiers are often given a large DC voltage (bias) so they'll operate as desired. The information signal is the variation about the bias. So covariance is more meaningful here.

Q: Why correlation coefficient and not covariance?

A: No units! The lack of units can make the correlation coefficient more interpretable.

Properties of correlation, covariance, and correlation coefficient for general RVs X and Y

Shortcut for computing covariance

recall $Var(X) = E((X-E(X))^2) = E(X^2) - E(X)^2$

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$

proof

$$= E((X-E(X))(Y-E(Y))) = E(XY - E(X)Y - XE(Y) + E(X)E(Y))$$

$$= E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y)$$

---

more properties

$$Cov(X,X) = Var(X)$$

$$Cov(X,Y) = Cov(Y,X)$$

$$Var(X+Y) = Var(X) + Var(Y) + \underline{2\,Cov(X,Y)}$$

cross-term disappears ONLY if X and Y are uncorrelated

proof

$$Var(X+Y) = E((X+Y)^2) - E(X+Y)^2$$

$$= E[X^2 + 2XY + Y^2] - (E(X) + E(Y))^2$$

$$= E(X^2) + 2E(XY) + E(Y^2) - E(X)^2 - 2E(X)E(Y) - E(Y)^2$$

$$= E(X^2) - E(X)^2 + 2[E(XY) - E(X)E(Y)] + E(Y^2) - E(Y)^2$$

$$= Var(X) + Var(Y) + 2\,Cov(X,Y)$$

$$Cov(aX+b, cY+d) = ac\,Cov(X,Y)$$

$$Cov(X+Y, U+V) = Cov(X,U) + Cov(X,V) + Cov(Y,U) + Cov(Y,V)$$

The most general. All relationships can be derived from this one

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

$$|Cov(X,Y)| \leq \sqrt{Var(X)Var(Y)}$$

Interpreting the correlation coefficient.

- indicates the strength of the linear relationship between X and Y

Suppose you want to estimate a RV Y from another RV X using a linear relationship.

Model : $Y = aX + b + N$

The noise RV N indicates it may not be an exact linear relationship

What's $E(Y)$, $Var(Y)$, ρ_{xy} ?

Define $E(X) = \mu$, $Var(X) = \sigma^2$

$E(Y) = E(aX + b + N) = a\mu + b + E(N)$

Assume $E(N) = 0$

$$Var(Y) = Var(aX + b + N)$$
$$= Var(aX) + var(N) + 2cov(aX, N)$$
$$= a^2\sigma^2 + \sigma_N^2$$

Assume noise uncorrelated with signal

$$E(XY) = E(aX^2 + bX + XN)$$
$$= aE(X^2) + b\mu$$

and since $\sigma^2 = E(X^2) - \mu^2$,
$$E(X^2) = \sigma^2 + \mu^2$$

so $E(XY) = a\sigma^2 + a\mu^2 + b\mu$

and, finally, $\rho_{xy} = \dfrac{cov(X,Y)}{\sqrt{var(X)Var(Y)}} = \dfrac{E(XY) - E(X)E(Y)}{\sigma\sqrt{a^2\sigma^2 + \sigma_N^2}}$

simplifying

$$\rho_{xy} = \frac{a\sigma^2 + a\mu^2 + b\mu - \mu(a\mu + b)}{\sigma\sqrt{a^2\sigma^2 + \sigma_N^2}}$$

$$= \frac{a\sigma}{\sqrt{a^2\sigma^2 + \sigma_N^2}}$$

depends on σ_N^2

<u>Interpretation</u>: ρ_{xy} describes how effective the linear relationship is

If $Y \approx aX + b$ is an accurate approximation, then σ_N^2 is small and

$$\rho_{xy} \to \frac{a\sigma}{|a|\sigma} = \pm 1 \qquad \text{as } \sigma_N^2 \to 0$$

If $Y \approx aX + b$ is less accurate, σ_N^2 larger,

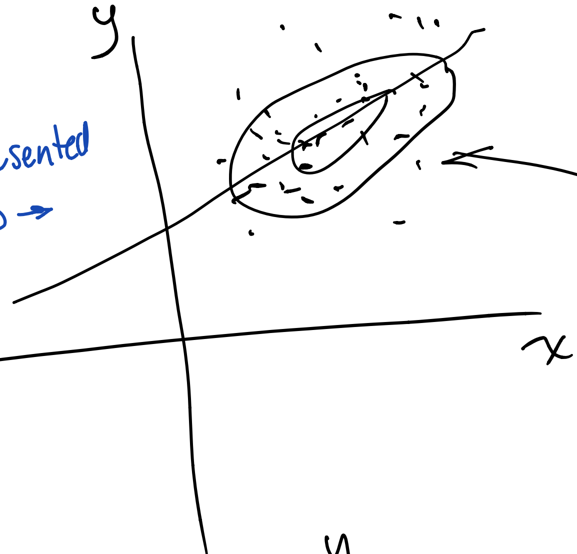
then $\rho_{xy} \approx \frac{a\sigma}{\sigma_N} \to 0$ as σ_N^2 grows

If $\rho = 0$, there is no <u>linear</u> relationship between $X + Y$.

if $\rho = \pm 1$ there is a direct linear relationship

Scatter plot examples

the joint pdf, represented by isocurves →

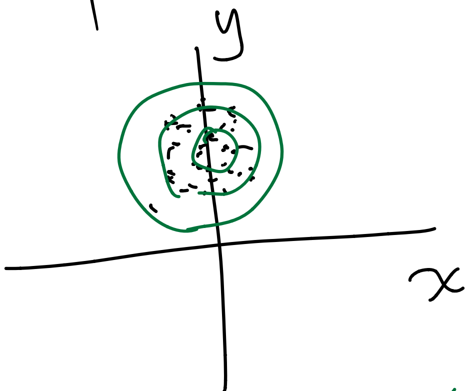


(these represent equal height contours for the joint pdf of X and Y)

(lines for which $f_{XY}(x,y) = c$ for different c.)

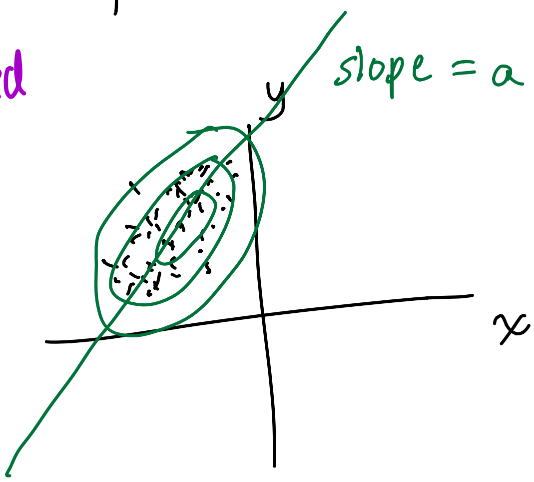
uncorrelated

$\rho \approx 0$



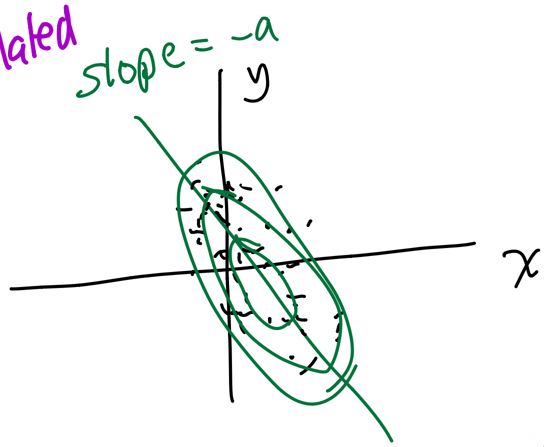
positively correlated

$\rho \approx 0.8$

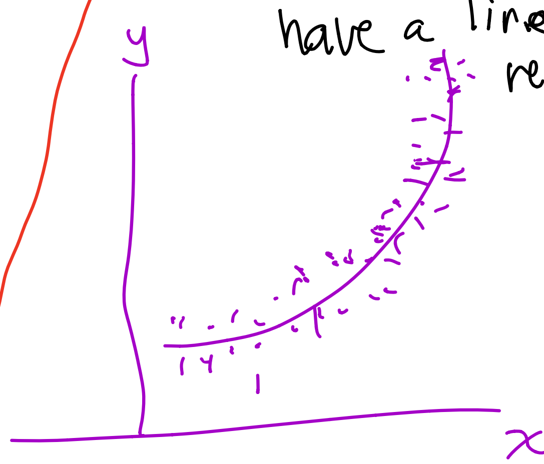


negatively correlated

$\rho \approx -0.8$



Not all RVs have a linear relationship



Covariance, correlation, and independence

Recall, X and Y are independent
if and only if $f_{XY}(x, y) = f_X(x) f_Y(y)$
for all x and all y

$\boxed{\text{if } X \text{ and } Y \text{ are independent, then } cov(X,Y) = 0}$

why? easy to prove...

Be careful, the reverse is not always true.

$\boxed{\text{if } cov(X,Y) = 0, \; X \text{ and } Y \text{ may not be independent!}}$

Example: uncorrelated but <u>NOT</u> independent.

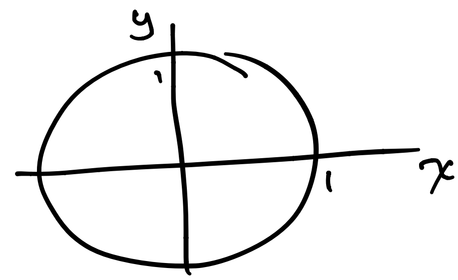
example 5.27

Θ is a uniformly distributed angle, $(0, 2\pi)$

$X = \cos\Theta$
$Y = \sin\Theta$

(x,y) are on the unit circle

Are X and Y independent? uncorrelated?



To show independence, you must be able to show

$f_{XY}(x,y) = f_X(x)\, f_Y(y)$ for all x and y.

But note that the region of support for $f_{XY}(x,y)$ is <u>on</u> the unit circle.

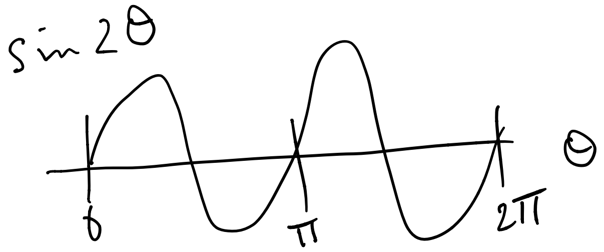
So we know immediately that they are NOT independent, because a circle is not in product form

---

What is $\mathrm{COV}(X,Y)$? $= E(XY) - E(X)E(Y)$
$= E(XY)$ because $E(X) = 0$

$$E(XY) = E(\cos\theta \sin\theta) = E\left(\frac{\sin 2\theta}{2}\right)$$

$$= \frac{1}{2}\,\frac{1}{2\pi}\int_0^{2\pi} \sin 2\theta \, d\theta = 0$$



$\Rightarrow$ X and Y are uncorrelated but <u>not</u> independent!

Summary for independent RVs

Independent RVs are always uncorrelated

Uncorrelated RVs may not be independent

Uncorrelated joint Gaussian RVs are independent

Independent RVs ONLY

$$E(g_1(X)\, g_2(Y)) = E(g_1(X))\, E(g_2(X))$$

$$r_{xy} = E(XY) = E(X)E(Y)$$

$$C_{xy} = \mathrm{Cov}(X,Y) = 0$$

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

no cross term

and of course

$$f_{xy}(x,y) = f_x(x)\, f_y(y)$$

$$f_x(x|y) = f_x(x)$$

(next topic..)