August 5, 2019

CORRELATION, COVARIANCE, INDEPENDENCE

If X and Y are random variables, then the covariance of X and Y is defined as

$$\begin{aligned}
\mathrm{cov}(X,Y) = \sigma_{XY} &= \mathbb{E}\left[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])\right] \\
&= \mathbb{E}\left[XY - Y\,\mathbb{E}[X] - X\,\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Y]\right] \\
&= \mathbb{E}[XY] - \mathbb{E}\left[Y\,\mathbb{E}[X]\right] - \mathbb{E}\left[X\,\mathbb{E}[Y]\right] + \mathbb{E}\big[\underbrace{\mathbb{E}[X]\mathbb{E}[Y]}_{\text{expected value of constant}}\big] \\
&= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \cancel{- \mathbb{E}[X]\mathbb{E}[Y]} \cancel{+ \mathbb{E}[X]\mathbb{E}[Y]} \\
&= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]
\end{aligned}$$

sometimes called the ~~product-moment~~ product-expectation

Qn: How do we calculate $\mathbb{E}[XY]$?

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x,y)\,dx\,dy$$

Qn: What happens when $X \perp Y$? Recall that $f(x,y) = f_X(x) f_Y(y)$.

$$\mathbb{E}[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x,y)\,dx\,dy = \int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} x\,y\, f_X(x) f_Y(y)\,dx\right\}dy$$

$$= \int_{-\infty}^{\infty} y f_Y(y) \left\{\int_{-\infty}^{\infty} x\, f_X(x)\,dx\right\} dy = \mathbb{E}[X]\,\mathbb{E}[Y]$$

NOTE: If $X \perp Y \Rightarrow \mathrm{cov}(X,Y) = 0$, i.e., $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$.

<u>Qn</u>: If X and Y are random variables such that $\text{cov}(X,Y)=0$, does this imply that $X \perp Y$?

<u>ANS</u>: No! (An exception: when X and Y are jointly Gaussian.)

<u>EX</u>: Let $X \sim N(0,1)$. Let $Y = X^2$.

But $\mathbb{E}[XY] = \mathbb{E}[X \cdot X^2] = \mathbb{E}[X^3] = 0$.

And, $\mathbb{E}[X]\mathbb{E}[Y] = 0 \cdot 1 = 0$.

But Y depends explicitly upon $X \Rightarrow X \not\perp Y$.

<u>WARNING</u>: (theoretical) "independence" $\longleftrightarrow$ joint density is factorizable <u>and</u> defined on a rectangle-like support set

(empirical) "statistical independence" $\longleftrightarrow$ covariance is zero or correlation is zero

RECALL: $\sigma_{XY} = \mathbb{E}\big[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])\big]$

- You get weird product units.
- Not invariant with respect to scalings.

PROOF: Suppose that X and Y are r.v. and define $\tilde{X} = \alpha X$ and $\tilde{Y} = \beta Y$.

$$\operatorname{cov}(\tilde{X},\tilde{Y}) = \operatorname{cov}(\alpha X, \beta Y) = \mathbb{E}\big[(\alpha X - \mathbb{E}[\alpha X])(\beta Y - \mathbb{E}[\beta Y])\big]$$

$$= \mathbb{E}\big[\alpha(X-\mathbb{E}[X])\,\beta\,(Y-\mathbb{E}[Y])\big]$$

$$= \alpha\beta \operatorname{cov}(X,Y)$$

ASIDE: what does covariance do with translation, i.e., if $\tilde{X} = \alpha X + a$ and $\tilde{Y} = \beta Y + b$? what would $\operatorname{cov}(\tilde{X},\tilde{Y})$ be? (ANS: No impact on covariance.)

GOAL: Generate an alternative to covariance that is unit-free?

CONCEPT OF CORRELATION: Standardize the random variables X and Y, i.e., to work with

$$Z_X = \frac{X-\mathbb{E}[X]}{\sigma_X} \quad \text{-and-} \quad Z_Y = \frac{Y-\mathbb{E}[Y]}{\sigma_Y}.$$

DEF: $\operatorname{CORR}(X,Y) = \operatorname{cov}(Z_X, Z_Y) = \rho_{XY} = \rho$

$$= \operatorname{cov}\left(\frac{X-\mathbb{E}[X]}{\sigma_X}, \frac{Y-\mathbb{E}[Y]}{\sigma_Y}\right) = \operatorname{cov}\left(\frac{X}{\sigma_X} - \frac{\mathbb{E}[X]}{\sigma_X}, \frac{Y}{\sigma_Y} - \frac{\mathbb{E}[Y]}{\sigma_Y}\right)$$

$$= \frac{1}{\sigma_X}\frac{1}{\sigma_Y}\operatorname{cov}(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_X} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y}$$

PROPERTIES OF COVARIANCE:

1. $\text{Cov}(X,Y) = \text{Cov}(Y,X)$
2. $\text{Cov}(X,X) = \text{Var}(X)$
3. $\text{Cov}(X,\alpha) = 0$

   $\alpha \in \mathbb{R}$
4. $\text{cov}(X+Y, Z)$
   $= \text{cov}(X,Z) + \text{cov}(Y,Z)$
5. $\text{cov}(\alpha X, \beta Y) = \alpha\beta\,\text{cov}(X,Y)$

PROPERTIES OF CORRELATION:

1. $\rho_{X,Y} = \rho_{aX+b,\,cY+d}$ where $a>0,\ c>0,\ b\in\mathbb{R},\ d\in\mathbb{R}$

   $a<0,\ c<0,\ b\in\mathbb{R},\ d\in\mathbb{R}$
2. If $X \perp Y$, $\rho_{XY} = 0$.

   WARNING: $\rho_{XY} = 0 \not\Rightarrow$ independence.
3. $-1 \leq \rho_{XY} \leq 1$
4. $\rho_{XY} = +1 \iff Y = mX + b$ for some $m>0$ and $b\in\mathbb{R}$

   $\rho_{XY} = -1 \iff Y = mX+b$ for some $m<0$ and $b\in\mathbb{R}$
5. ρ measures only the linear dependence between X and Y. If the dependence is nonlinear, then ρ is an inapplicable or inappropriate measure of dependence.

BIVARIATE NORMALITY

Let $f(x,y)$ be a joint pdf of continuous random variables X and Y. We say that X and Y are bivariate normal iff the following four conditions hold.

(a) The pdf of X is $f_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}} e^{-\frac{(x-\mu_X)^2}{2\sigma_X^2}}$

The pdf of Y is $f_Y(y) = \frac{1}{\sigma_Y\sqrt{2\pi}} e^{-\frac{(y-\mu_Y)^2}{2\sigma_Y^2}}$.

(The marginals must be normal.)

(b) The conditional distribution of $Y|X=x$ is also normal for every $x \in (-\infty, \infty)$. That is, $Y|X=x$ is a r.v. that is normal for every x.

(Errors ~~xxxxx~~ for the regression are normally distributed.)

(c) The conditional expectation of $Y|X=x$, i.e., $\mu_{Y|X=x}$, is a linear function of x; that is,

$\mu_{Y|X=x} = \mathbb{E}[Y|X=x] = a + bx$ for some a and some b.

(Linearity)

(d) The conditional variance of Y, given $X=x$, is constant. We will denote it $\sigma^2_{Y|X=x}$. Because it is a constant, it is independent (in the functional sense) of x. (Homoscedasticity)

FACT: Suppose that (a) – (d) are true, i.e., X and Y are bivariate normal.

Then one can show that

quadratic form associated to a bivariate normal

$$f(x,y) = \frac{1}{\sigma_X \sigma_Y \sqrt{1-\rho^2}\cdot 2\pi} \exp\left\{-\frac{1}{2}\,\frac{1}{1-\rho^2}\,Q(x,y)\right\}$$

ρ_{XY} ← the actual number

$$\text{where } Q(x,y) = \left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\,\frac{x-\mu_X}{\sigma_X}\,\frac{y-\mu_Y}{\sigma_Y} + \left(\frac{y-\mu_Y}{\sigma_Y}\right)^2$$

where $\rho = \text{CORR}(X,Y)$.

Qn: what is $f_{Y|X}(y|x)$? well, it's $\underbrace{f(x,y) / f_X(x)}$.

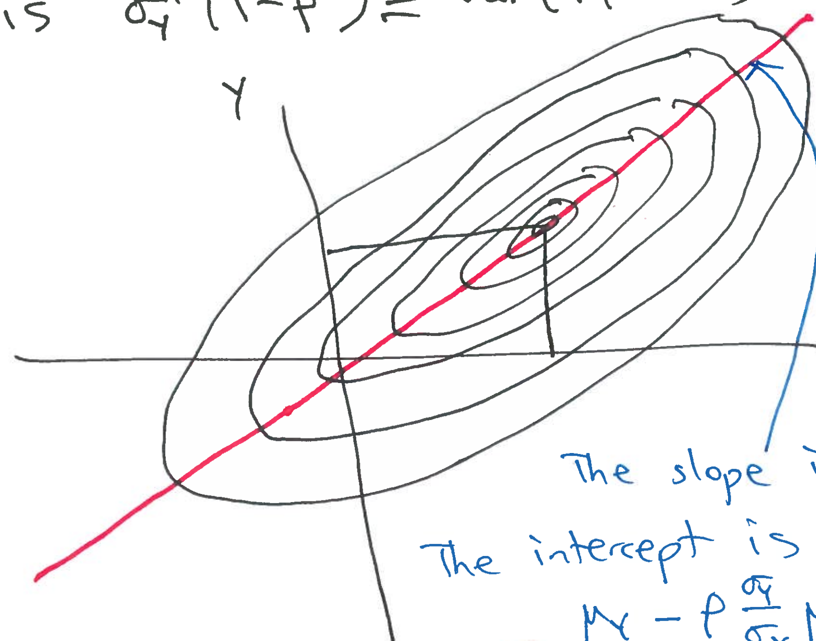
"usual" exercise for you

$$f_{Y|X}(y|x) = \frac{1}{\sigma_Y\sqrt{1-\rho^2}\,\sqrt{2\pi}} \exp\left\{-\frac{\left(y-\mu_Y-\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)\right)^2}{2\sigma_Y^2(1-\rho^2)}\right\}$$

This is a Gaussian density in slight disguise.

Its mean is $\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X) = \mathbb{E}[Y|X=x] = \mu_{Y|X=x}$.

Its variance is $\sigma_Y^2(1-\rho^2) = \text{Var}(Y|X=x) = \sigma^2_{Y|X=x}$.



$\hat{\beta}_1 = \dfrac{\text{Cov}(X,Y)}{\text{VAR}(X)} =$

The slope is $\rho\frac{\sigma_Y}{\sigma_X}$.

The intercept is $\mu_Y - \rho\frac{\sigma_Y}{\sigma_X}\mu_X$.

↳ $\bar{Y} - \hat{\beta}_1\bar{X} = \hat{\beta}_0$