

Risk & Asset Allocation

Case for Week 2

John A. Dodson

September 10, 2014

1 Joint Density

The density for a multivariate normal random variable is

$$f_X(x) = (2\pi)^{-n/2} e^{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)} \sqrt{|\Sigma^{-1}|}$$

for positive-definite covariance matrix Σ of dimension n and mean vector μ .

For the in-class exercise this week, we have $n = 2$, $\mu = 0$, and

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

with correlation $-1 < \rho < 1$. Hence

$$\Sigma^{-1} = \frac{1}{1-\rho^2} \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}$$

and the density simplifies to

$$f_X(x) = \frac{1}{2\pi} e^{-\frac{x_1^2 - 2\rho x_1 x_2 + x_2^2}{2(1-\rho^2)}} \frac{1}{\sqrt{1-\rho^2}}$$

2 Conditional Density

It's conventional in the \mathbb{R}^2 case to use Descartes' (X, Y) instead of the more general (X_1, X_2) .

According to the exercise, we are interested in the event $X = 1$ and what it tells us about Y . Prior to the observation, we only have the marginal density for Y ,

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{(X,Y)}(x, y) dx \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} \end{aligned}$$

which has entropy

$$H_Y = E_Y[-\log f_Y(Y)] = \log \sqrt{2\pi e}$$

Once we have the observation on X , we can work with the conditional density,

$$\begin{aligned} f_{Y|X}(y) &= \frac{f_{(X,Y)}(1, y)}{\int_{-\infty}^{\infty} f_{(X,Y)}(1, y) dy} \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-\rho)^2}{2(1-\rho^2)}} \frac{1}{\sqrt{1-\rho^2}} \end{aligned}$$

which has entropy

$$H_{Y|X} = \mathbb{E}_{Y|X} [-\log f_{Y|X}(Y)] = \log \sqrt{2\pi e (1-\rho^2)}$$

Clearly $H_{Y|X} \leq H_Y$. To the extent that there is correlation, the observation acts to lower the entropy of the r.v. we are interested in describing.

3 Regression

We have demonstrated that (in the bivariate normal case at least) conditioning conveys information. We can see this in a more familiar light by considering the general bivariate result with four additional parameters and an arbitrary conditioning event, $X = x$.

You can confirm that the general conditional density works out to be

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

In other words,

$$Y|X \sim \mathcal{N}\left(\mu_Y + \rho \frac{\sigma_Y}{\sigma_X} (x - \mu_X), \sigma_Y^2 (1-\rho^2)\right)$$

suggesting the transformation

$$Y|X \triangleq \alpha + \beta X + \epsilon \tag{1}$$

where

$$\beta = \frac{\text{cov}(X, Y)}{\text{var } X} \tag{2a}$$

$$\alpha = \mathbb{E} Y - \beta \mathbb{E} X \tag{2b}$$

$$\epsilon \sim \mathcal{N}(0, \text{var } Y - \beta^2 \text{var } X) \tag{2c}$$

You may recognize this as the “population” result from classical statistics for ordinary least squares (OLS) regression.

This observation is useful because it gives us direction not only on grounding classical regression in modern estimation theory, but also on how we might adapt to situations where relationships are not well described by linear correlations.