Risk & Asset Allocation
Case for Week 2

John A. Dodson

September 11, 2013

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

According to the exercise, we are interested in the event $X_1 = 1$ and what it tells us about X_2. Prior to the observation, we only have the marginal density for X_2,

$$\begin{aligned} f_{X_2}(x_2) &= \int_{-\infty}^{\infty} f_X(x_1, x_2)\, dx_1 \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_2^2} \end{aligned}$$

which has entropy

$$H_{X_2} = \mathrm{E}_{X_2}\left[-\log f_{X_2}(X_2)\right] = \log\sqrt{2\pi e}$$

Once we have the observation on X_1, we can work with the conditional density,

$$f_{X_2|X_1}(x_2) = \frac{f_X(1, x_2)}{\int_{-\infty}^{\infty} f_X(1, x)\, dx}$$
$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_2-\rho)^2}{2(1-\rho^2)}} \frac{1}{\sqrt{1-\rho^2}}$$

which has entropy

$$H_{X_2|X_1} = \mathrm{E}_{X_2|X_1}\left[-\log f_{X_2|X_1}(X_2)\right] = \log \sqrt{2\pi e\left(1-\rho^2\right)}$$

Clearly $H_{X_2|X_1} \leq H_{X_2}$. 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, $X_1 = x_1$.

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

$$f_{X_2|X_1}(x_2) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\left(x_2-\mu_2-\rho\frac{\sigma_2}{\sigma_1}(x_1-\mu_1)\right)^2}{2\sigma_2^2(1-\rho^2)}} \frac{1}{\sigma_2\sqrt{1-\rho^2}}$$

In other words,

$$X_2|X_1 \sim \mathcal{N}\left(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1-\mu_1), \sigma_2^2\left(1-\rho^2\right)\right)$$

suggesting the transformation

$$X_2|X_1 \triangleq \alpha + \beta_1 X_1 + \epsilon \tag{1}$$

where

$$\beta_1 = \frac{\operatorname{cov}(X_1, X_2)}{\operatorname{var} X_1} \tag{2a}$$
$$\alpha = \mathrm{E}\, X_2 - \beta_1\, \mathrm{E}\, X_1 \tag{2b}$$
$$\epsilon \sim \mathcal{N}\left(0, \operatorname{var} X_2 - \beta_1^2 \operatorname{var} X_1\right) \tag{2c}$$

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.