Agribusiness Analysis and Forecasting

Multivariate Normal

Henry Bryant

Texas A&M University

Joint versus Marginal Distributions

Thus far, we have been thinking about each variable in a system individually. We have been thinking about each variable's univariate distribution, a.k.a, its *marginal distribution*. We have been implicitly assuming realizations of random variables occur *independently* of one another.

In this topic, we start thinking about the *joint distribution* (a.k.a., multivariate distribution) of a system of variables. This is a single probability distribution for all of a system's variables that accommodates some form of *dependence* among the variables.

Different Joint Distribution Approaches

* Multivariate Normal distribution (this topic). All variables have a normal marginal distribution.

* Mixed marginals where each variable has a different marginal distribution (next topic). For example:

- $Y_1 \sim$ Uniform
- $Y_2 \sim$ Normal
- $Y_3 \sim$ Empirical
- $Y_4 \sim$ Beta

Why use joint distributions?

Data are generated contemporaneously.

- Price and yield are observed each year for related commodities.
- Corn and sorghum used interchangeably for animal feed so prices are related.
- Steer and heifer prices are related.
- Yields of crops on the same farm have the same weather conditions.

Supply and demand forces affect prices similarly, bear market or bull market; prices move together.

- Prices for tech stocks move together.
- Prices for an industry or sector's stocks move together.

Why use joint distributions?

If correlation is ignored when random variables are correlated, results are biased: Suppose $\tilde{Y}_3 = \tilde{Y}_1 + \tilde{Y}_2$ and

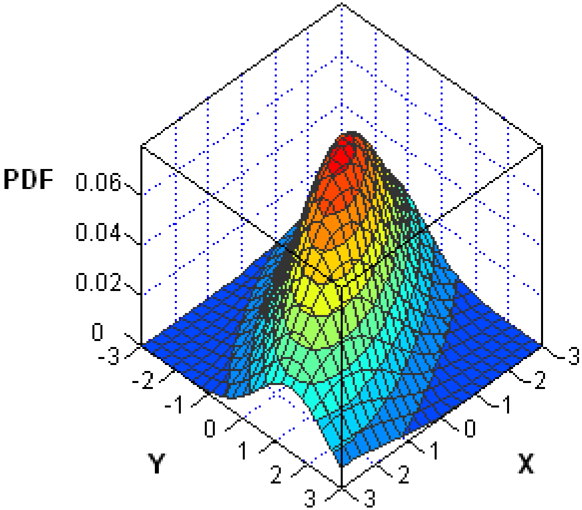
the model is simulated without correlation, assuming $\rho_{1,2}$ $=$0:

- If the true $\rho_{1,2} > 0$ then the model will understate the risk for Y_3.
- If the true $\rho_{1,2} < 0$ then the model will overstate the risk for Y_3.

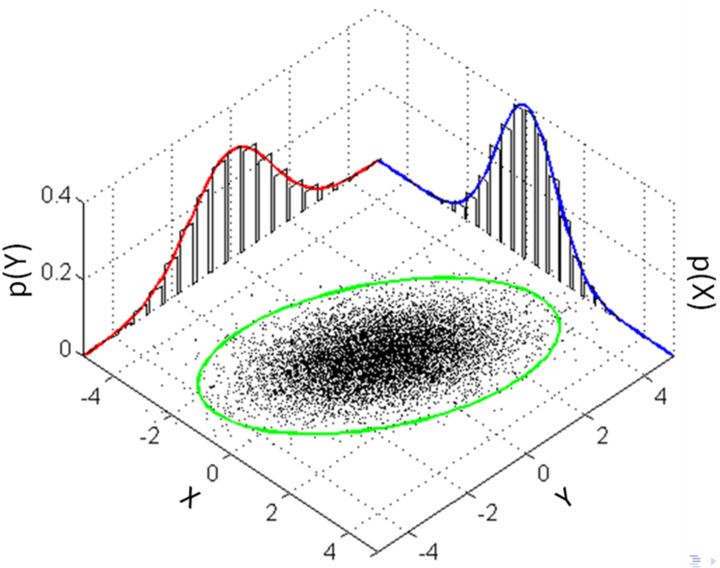
If $\tilde{Y}_3 = \tilde{Y}_1 * \tilde{Y}_2$:

- Not only the variance, but the mean of Y_3 is biased, as well.

MV Normal Joint PDF



MV Normal Marginal Distributions



Reminder...

You should only be simulating covariance stationary random variables.

- $E(Y_t) = E(Y_{t-1}) = \mu$
- $\sigma_t^2 = \sigma_{t-1}^2 = \sigma^2 < \infty$

These parameters are not a function of time.

Generating Draws for MV Normal

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} + \sqrt{\begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ & \sigma_{22} & \sigma_{23} \\ & & \sigma_{33} \end{bmatrix}} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$

- The z_i are standard normal draws
- μ_i is the mean for the i^{th} variable in the system
- σ_{ij} is the covariance between the i^{th} variable and the j^{th} variable
- $\sqrt{\ }$ denotes the *Cholesky decomposition* of the covariance matrix

Simulating MVN in Simetar

For 4 random variables...

- This is an array formula
- Start by highlighting 4 cells where the result will go
- Then
  =MVNORM(4x1 Means Vector, 4x4 Covariance Matrix)
  =MVNORM(A1:A4 , B1:E4)
  Control Shift Enter

Limitations to MV Normal

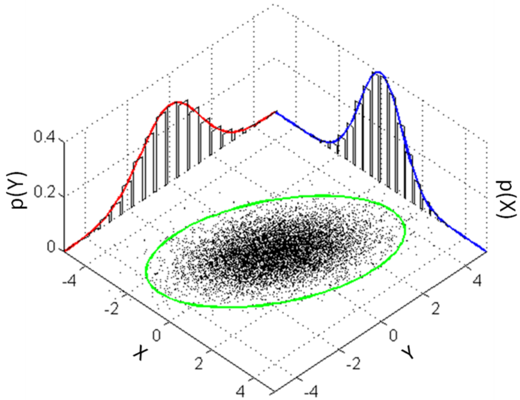
- Under MV Normal, all marginal distributions in the system are normal.

  But we won't generally have control over which marginal distributions are appropriate for the variables in our system.

- Under MV Normal, dependence among variables is strictly linear.

  But dependence among variables is often non-linear.

MV Normal Marginal Distributions



$$\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \Sigma = \begin{bmatrix} 1 & 0.6 \\ & 2 \end{bmatrix}$$