Topic 3: Non-Spherical Disturbances

Our basic linear regression model is

$$\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad ; \qquad \boldsymbol{\varepsilon} \sim N[\mathbf{0}\,, \sigma^2 I_n]$$

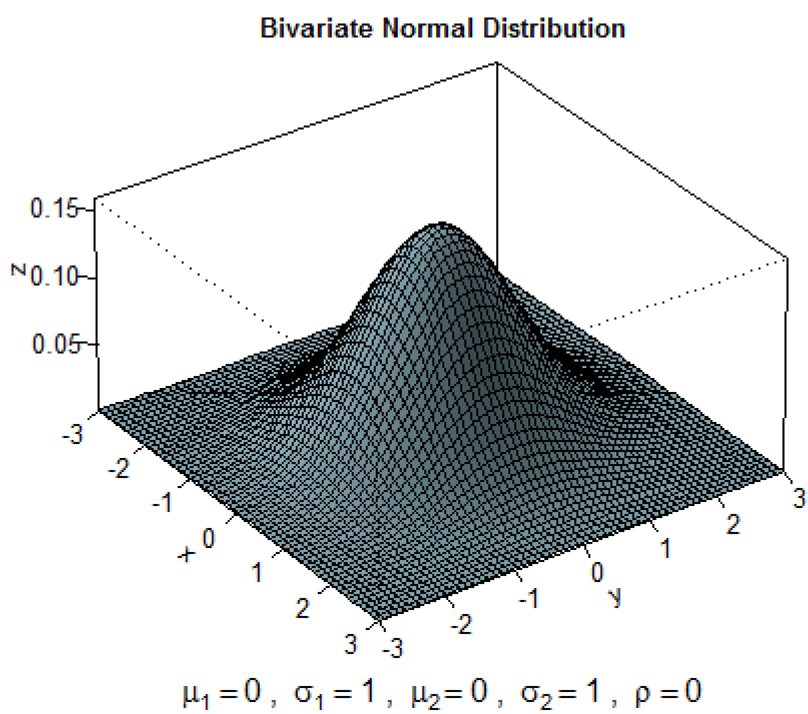
Now we'll generalize the specification of the error term in the model:

$$E[\boldsymbol{\varepsilon}] = \mathbf{0} \quad ; \quad E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'] = \Sigma = \sigma^2\Omega \quad ; \qquad \text{(\& Normal)}$$

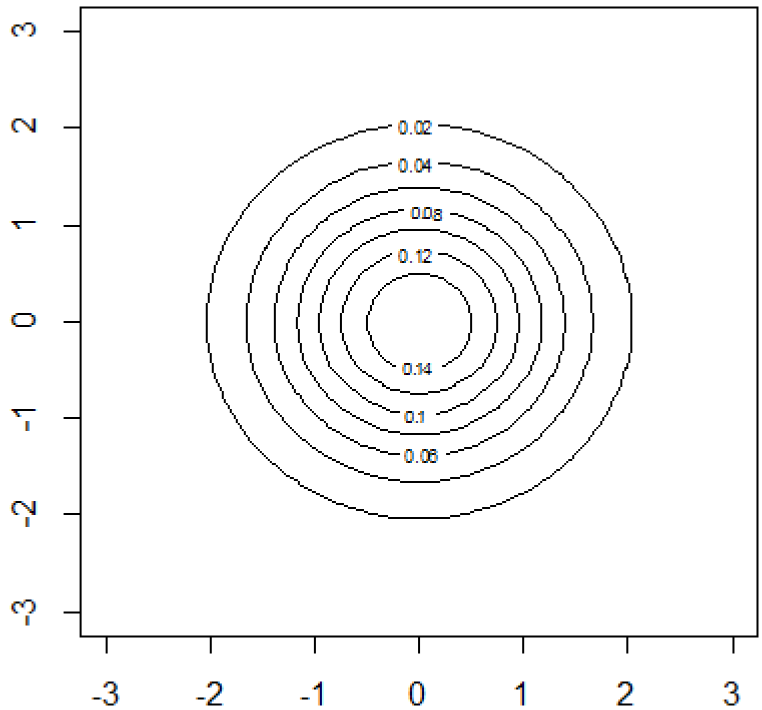
This allows for the possibility of one or both of

- Heteroskedasticity
- Autocorrelation (Cross-section; Time-series; Panel data)

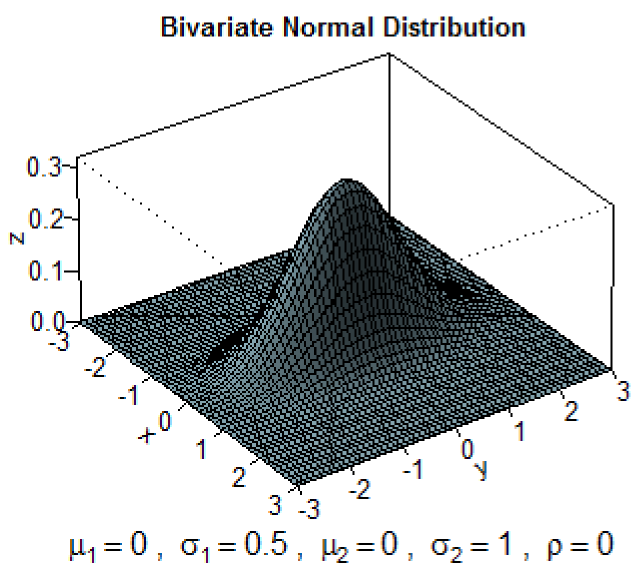
Spherical Disturbances – Homoskedasticity and Non-Autocorrelation

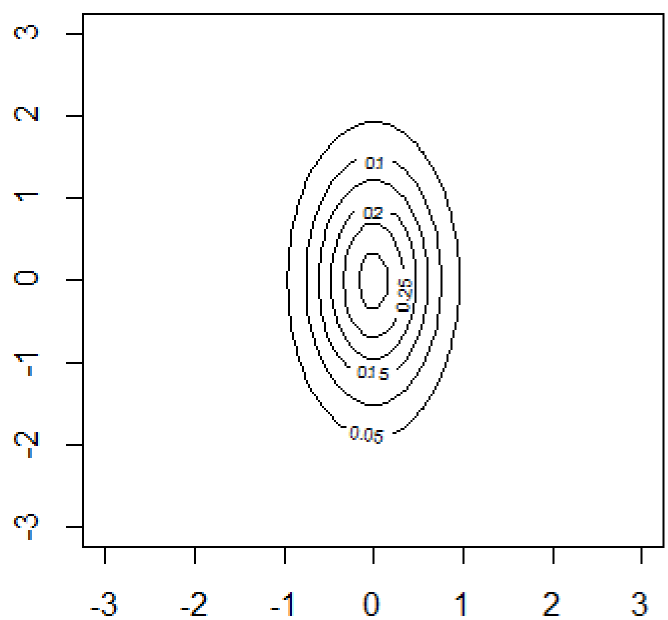


In the above, consider $x = \varepsilon_i$ and $y = \varepsilon_j$. The joint probability density function, $p(\varepsilon_i, \varepsilon_j)$, is in the direction of the *z* axis. Below is a contour of the above perspective. If we consider the joint distribution of *three* error terms instead of *two*, the circles below would become spheres, hence the terminology "spherical disturbances."

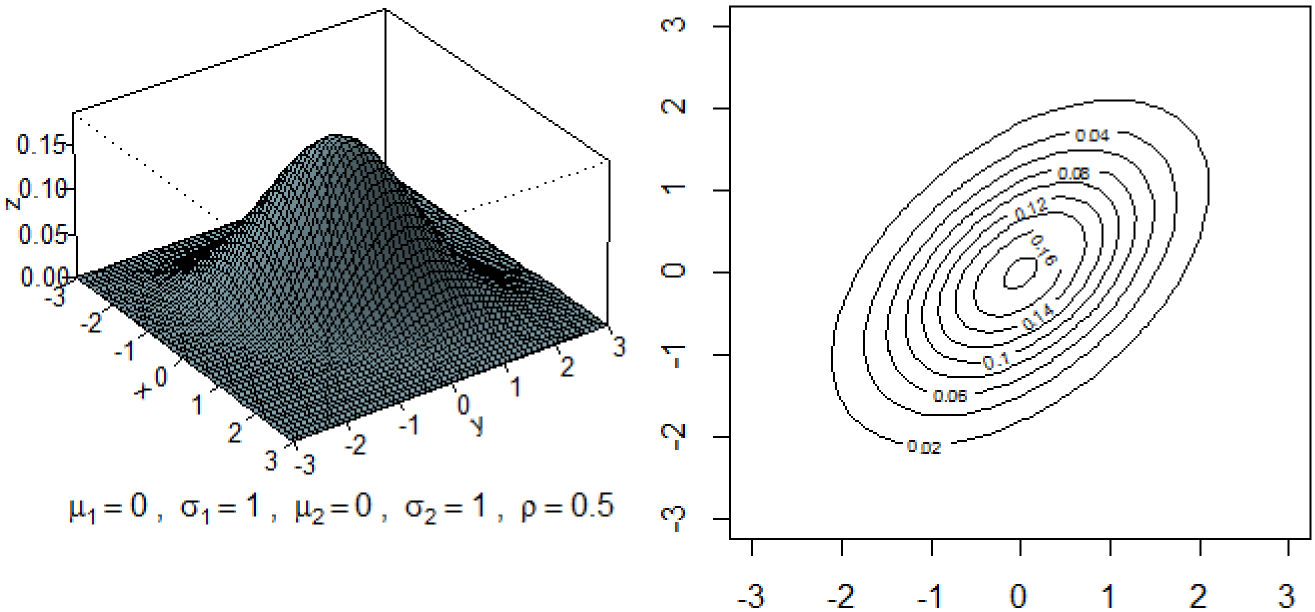


Non-Spherical Disturbances – Heteroskedasticity and Non-Autocorrelation

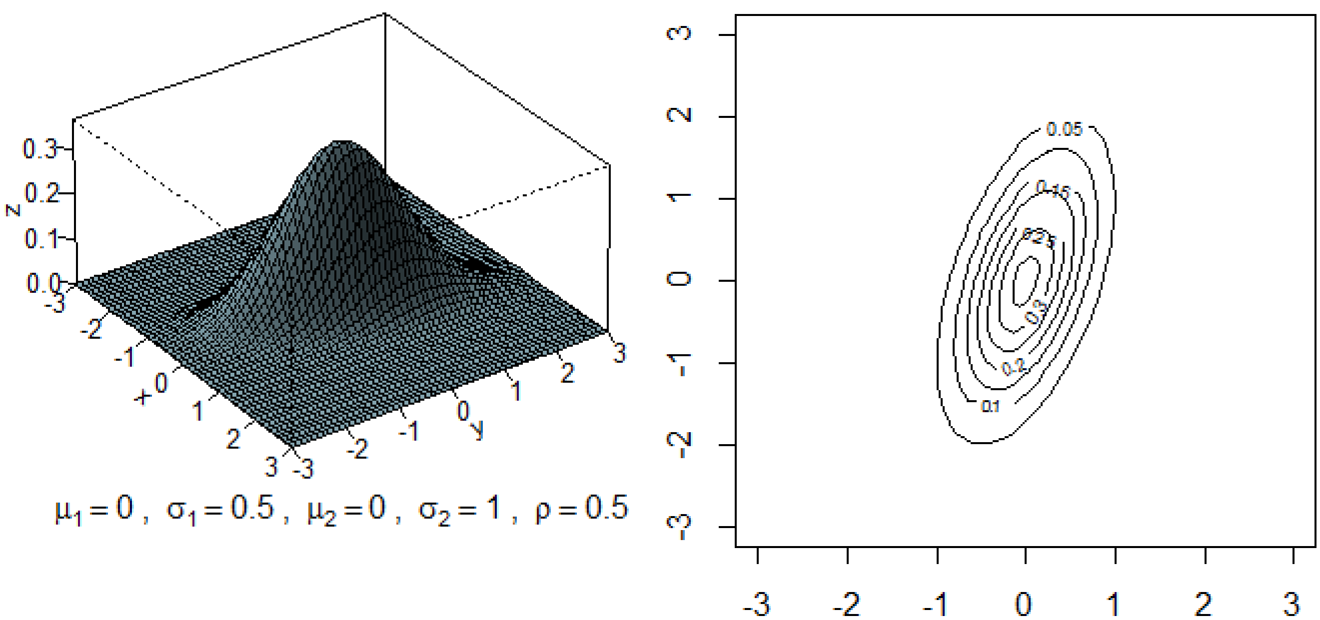




Non-Spherical Disturbances – Homoskedasticity and Autocorrelation



Non-Spherical Disturbances – Heteroskedasticity and Autocorrelation



How does the more general situation of non-spherical disturbances affect our (Ordinary) Least Squares estimator?

In particular, let's first look at the sampling distribution of $\boldsymbol{b}$:

$$\boldsymbol{b} = (X'X)^{-1}X'\boldsymbol{y} = (X'X)^{-1}X'(X\boldsymbol{\beta} + \boldsymbol{\varepsilon})$$

$$= \boldsymbol{\beta} + (X'X)^{-1}X'\boldsymbol{\varepsilon} \; .$$

So,

$$E(\boldsymbol{b}) = \boldsymbol{\beta} + (X'X)^{-1}X'E(\boldsymbol{\varepsilon}) = \boldsymbol{\beta} \;\; .$$

The more general form of the covariance matrix for the error term does not alter the fact that the OLS estimator is ***unbiased***.

Next, consider the covariance matrix of our OLS estimator in this more general situation:

$$V(\boldsymbol{b}) = V[\boldsymbol{\beta} + (X'X)^{-1}X'\boldsymbol{\varepsilon}] = V[(X'X)^{-1}X'\boldsymbol{\varepsilon}]$$

$$= [(X'X)^{-1}X'V(\boldsymbol{\varepsilon})X(X'X)^{-1}]$$

$$= [(X'X)^{-1}X'\sigma^2\Omega X(X'X)^{-1}]$$

$$\neq [\sigma^2(X'X)^{-1}] \;\; .$$

So, under our full set of modified assumptions about the error term:

$$\boldsymbol{b} \sim N[\boldsymbol{\beta}\, , V^*]$$

where

$$V^* = \sigma^2[(X'X)^{-1}X'\Omega X(X'X)^{-1}] \; .$$

- So, the usual computer output will be misleading, *numerically*, as it will be based on using the wrong formula, namely $s^2(X'X)^{-1}$.
- The standard errors, t-statistics, *etc*. will all be incorrect.
- As well as being *unbiased*, the OLS point estimator of $\boldsymbol{\beta}$ will still be ***weakly consistent***.
- The I.V. estimator of $\boldsymbol{\beta}$ will still be ***weakly consistent***.

- **However**, the usual estimator for the covariance matrix of $\boldsymbol{b}$, namely $s^2(X'X)^{-1}$, will be an *inconsistent estimator* of the true covariance matrix of $\boldsymbol{b}$!
- This has serious implications for inferences based on confidence intervals, tests of significance, *etc*.
- So, we need to know how to deal with these issues.
- This will lead us to some *generalized estimators*.
- First, let's deal with the most pressing issue – the inconsistency of the estimator for the covariance matrix of $\boldsymbol{b}$.

White's Heteroskedasticity-Consistent Covariance Matrix Estimator

- If we knew $\sigma^2\Omega$, then the "estimator" of the covariance matrix for $\boldsymbol{b}$ would just be:

$$V^* = [(X'X)^{-1}X'\sigma^2\Omega X(X'X)^{-1}]$$

$$= \frac{1}{n}\left[\left(\frac{1}{n}X'X\right)^{-1}\left(\frac{1}{n}X'\sigma^2\Omega X\right)\left(\frac{1}{n}X'X\right)^{-1}\right]$$

- Let $Q^* = \left(\frac{1}{n}X'\Sigma X\right)$ $(k \times k)$

$$= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij}\boldsymbol{x}_i\boldsymbol{x}_j'$$

$(k \times 1)$ $(1 \times k)$

- In the case of *heteroskedastic errors*, things simplify, because $\sigma_{ij} = 0$, for $i \neq j$.

Then, we have

$$Q^* = \frac{1}{n}\sum_{i=1}^{n} \sigma_i^2 \boldsymbol{x}_i \boldsymbol{x}_i'$$

- White (1980) showed that if we define

$$S_0 = \frac{1}{n}\sum_{i=1}^{n} e_i^2 \boldsymbol{x}_i \boldsymbol{x}_i'$$

  Then , $plim(S_0) = Q^*$.

- This means that we can estimate the model by OLS; get the associated residual vector, ***e*** ; and then a consistent estimator of V^*, the covariance matrix of ***b***, will be:

$$\hat{V}^* = \frac{1}{n}\left[\left(\frac{1}{n}X'X\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} e_i^2 \boldsymbol{x}_i \boldsymbol{x}_i'\right)\left(\frac{1}{n}X'X\right)^{-1}\right]$$

  or,

$$\hat{V}^* = n[(X'X)^{-1}S_0(X'X)^{-1}] .$$

- $\hat{V}^*$ is a **consistent estimator** of V^*, *regardless of the* (*unknown*) *form of the heteroskedasticity*.
- This includes no heteroskedasticity (*i.e*., homoscedastic errors).
- Newey & West produced a corresponding consistent estimator of V^* for when the errors possibly exhibit autocorrelation (of some unknown form).
- Note that the White and the Newey-West estimators modify just the estimated covariance matrix of ***b*** – not ***b*** itself.
- As a result, the *t*-statistics, *F*-statistic, *etc*., will be modified, but only in a manner that is appropriate *asymptotically*.
- So, if we have heteroskedasticity (or autocorrelation), whether we modify the covariance estimator or not, the usual test statistics will be unreliable **in finite samples**.
- Now let's turn to the estimation of $\boldsymbol{\beta}$, taking account of the fact that the error term has a non-scalar covariance matrix.
- Using this information should enable us to improve the *efficiency* of the LS estimator of the coefficient vector.