Department of Economics   **University of Victoria**

ECON 545: Econometric Analysis

Spherical Distributions

In our list of assumptions about the error term in our linear multiple regression model, we included one that incorporated both homoskedasticity and the absence of autocorrelation. That is, the individual values of the errors are assumed to be generated by a random process whose variance (σ^2) is constant, and all possible distinct pairs of these values are uncorrelated. We saw that this implied that the full error vector, ε, had a scalar covariance matrix, $\sigma^2 I_n$. We referred to this overall situation as one in which the values of the error term follow a **"Spherical Distribution"**. The purpose of this handout is to offer some background information about the origin of this terminology.

The discussion is quite general, so that you will realize that it applies to any random variables, not just the error term in our regression model. Further, so that we can look at some diagrams, let's consider the special case of two dimensions, rather than three, so that what would be a (3-dimensional) *sphere* becomes a (2-dimensional) *circle*. So, consider the pair of random values ε_i and ε_j, which we'll generically denote x and y. (This latter terminology has nothing to do with X and y in the regression model.) The values of these two random variables are plotted in the directions of the x and y axes in the graphs which follow.

In the three-dimensional plots we will see the joint probability density function, $p(x, y)$ in the direction of the z axis. All of these three-dimensional plots are for values in the range $-3 \leq x, y \leq 3$. Scales are given on the associated two-dimensional "contour" plots. The latter plots show "isolines" – that is, lines that join up (x, y) points that yield the same value (height) for $p(x, y)$. These contours are exactly analogous to the contour lines that you see on a topographic map to depict the nature of the terrain. They reflect what you see when you look down vertically on to the three-dimensional (bivariate) density plots.

The appearance of the density plots, and the shape of the associated contour plots, depend upon the variances of x and y, and the covariance (and hence correlation) between these two random variables.

If x and y have the same variance (*i.e*, if ε_i and ε_j are homoskedastic), and if they are uncorrelated (*i.e*, if ε_i and ε_j are **not** autocorrelated), then the contours will form circles. If there were three random variables we would need a four-dimensional density graph, and the contours would form a sphere. Hence the term **"Spherical Distribution"**. If there were four or more random variables the sphere would become a "hyper-sphere".

If x and y have different variances, the joint density surface is no longer symmetrical in the x and y directions, and then the coutour plot takes the form of an *ellipse*, rather than a circle. The same thing happens if x and y are correlated, even if they have the same variance. In this case, the slope of the primary axis of the ellipse is determined by the sign of the correlation between x and y.

Some examples follow, all for the case where x and y follow a bivariate normal distribution with zero means for x and y.
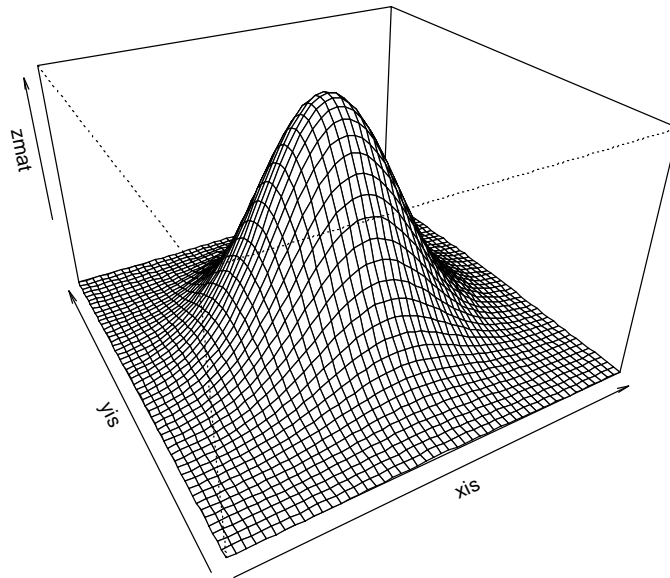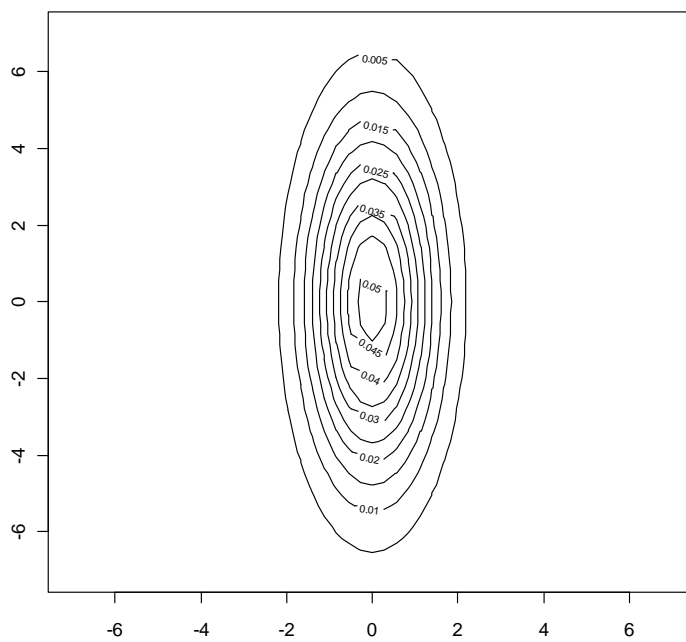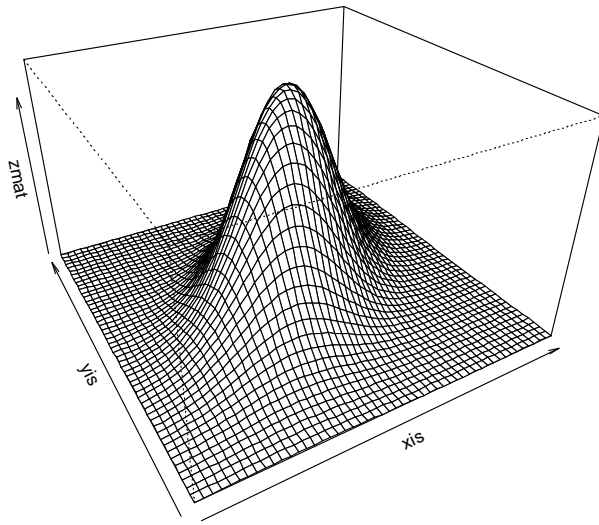
$$E\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \quad \text{var.}(x) = \text{var.}(y) = 1; \quad \text{cov.}(x, y) = 0 \; ; \text{ so } V\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
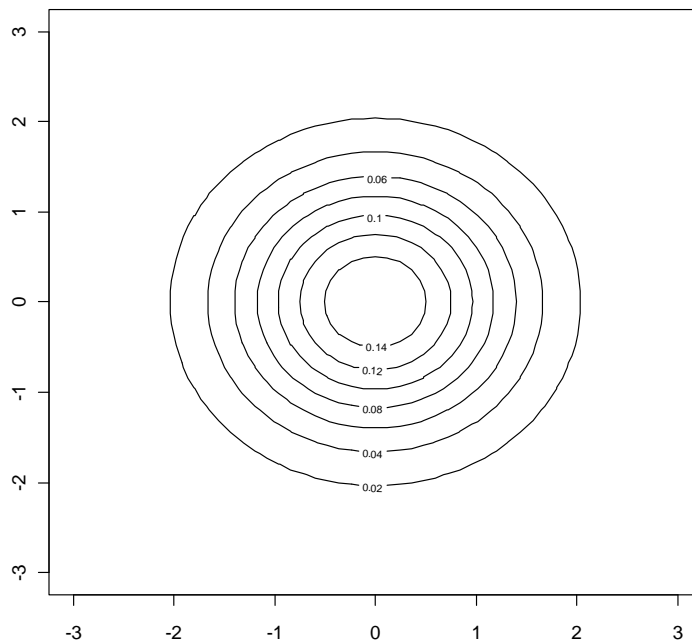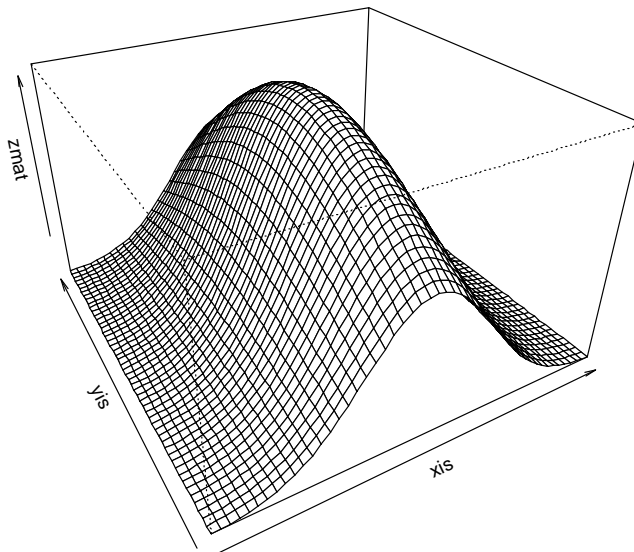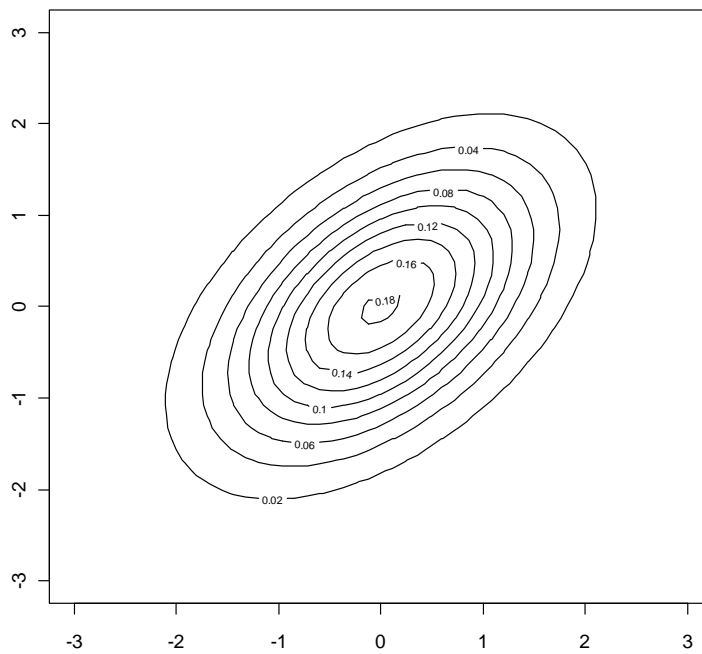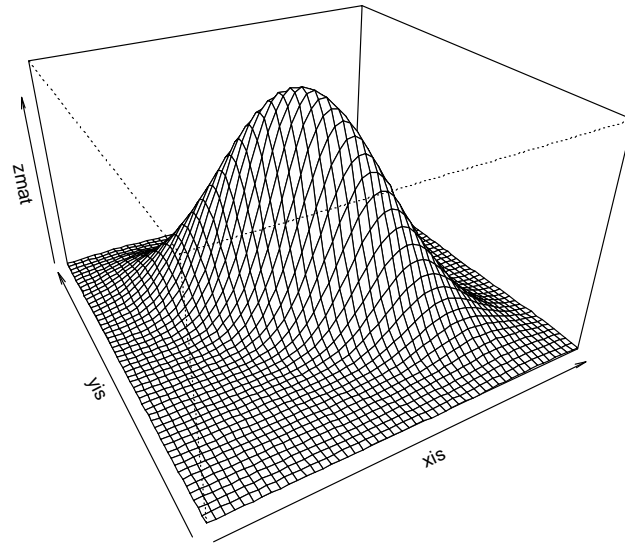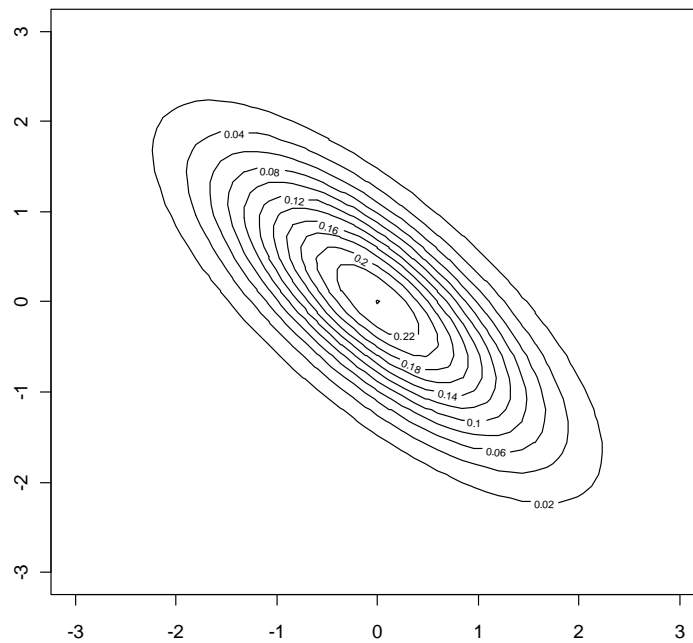
$$E\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \;\; \text{var.}(x) = 1 \; ; \; \text{var.}(y) = 9; \;\; \text{cov.}(x,y) = 0 \; ; \; \text{so} \; V\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 9 \end{pmatrix}$$
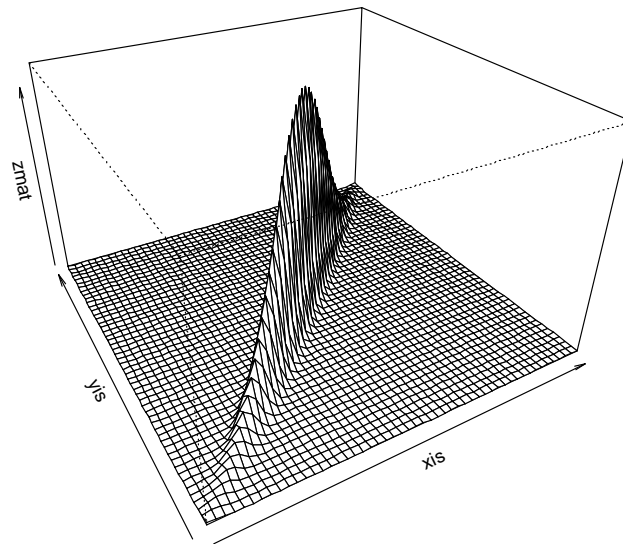
$$E\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \quad \text{var.}(x) = \text{var.}(y) = 1; \quad \text{cov.}(x, y) = 0.5 ; \text{ so } V\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$$
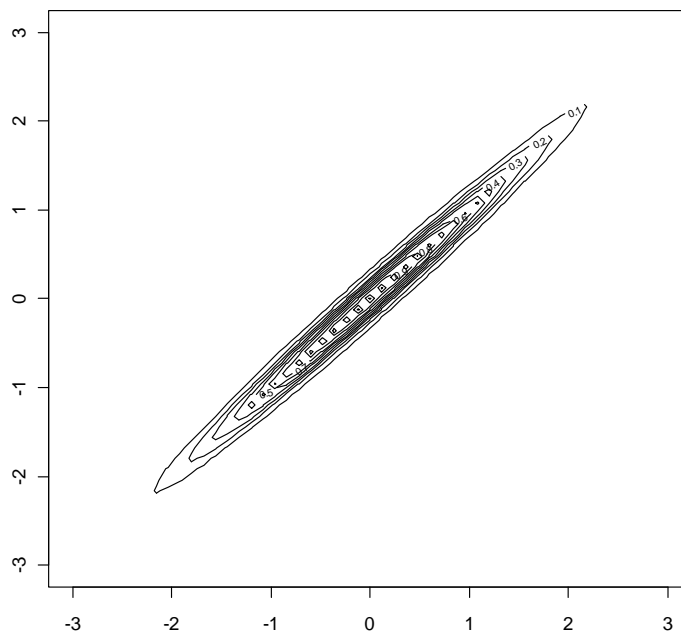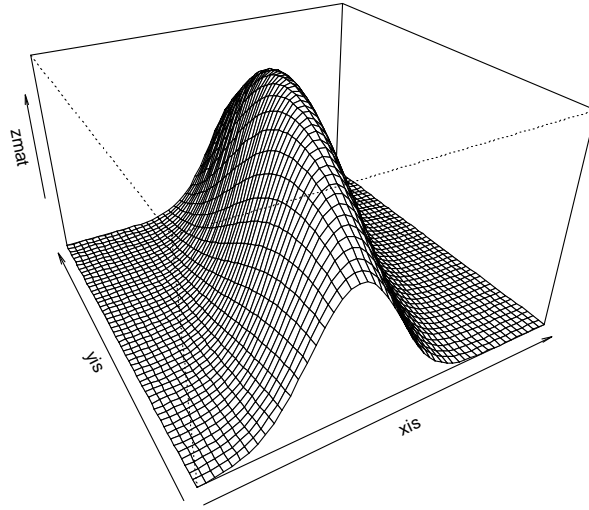
$$E\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \quad \text{var.}(x) = \text{var.}(y) = 1; \quad \text{cov.}(x, y) = -0.75 ; \text{ so } V\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & -0.75 \\ -0.75 & 1 \end{pmatrix}$$
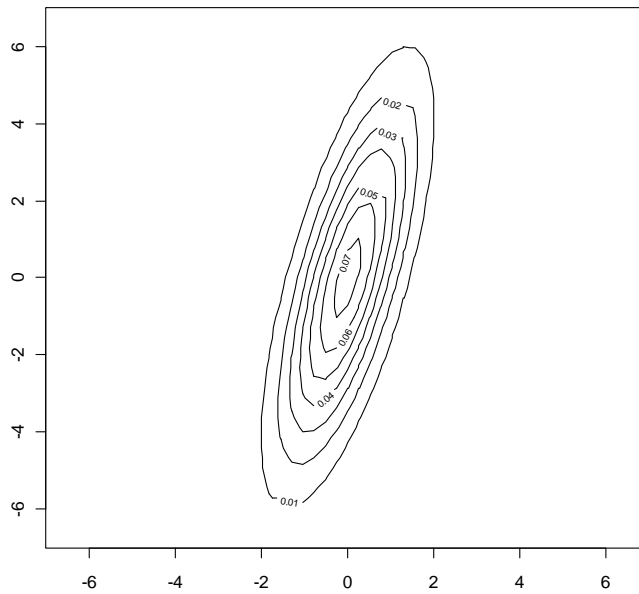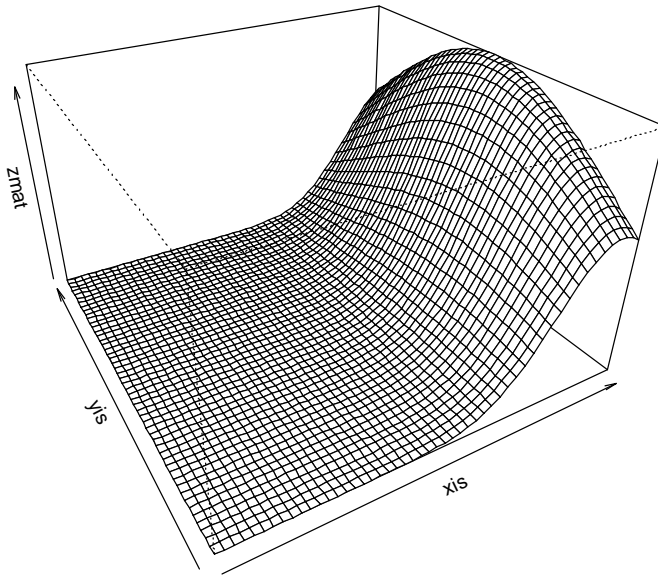
$$E\binom{x}{y}=\binom{0}{0}; \ \ \text{var.}(x)=\text{var.}(y)=1; \ \ \text{cov.}(x\,,\,y)=0.99 \ ; \ \text{so} \ V\binom{x}{y}=\begin{pmatrix} 1 & 0.99 \\ 0.99 & 1 \end{pmatrix}$$

$$E\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \quad \text{var.}(x) = 1 ; \text{var.}(y) = 9; \quad \text{cov.}(x , y) = 0.7 ; \text{ so } V\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0.7 \\ 0.7 & 9 \end{pmatrix}$$

In this final case: do *x* and *y* have the same means? Do they have the same variances? Are they correlated – if so, positively or negatively?