Bivariate Normal Distrib

We have seen an example of an important DISCRETE Multiv. distrib.
— Now we look at an important continuous Multiv. distrib.

① DEFN: The 2 dim'l r. var $(\underline{X}, Y)$ is said to have the Bivariate Normal Prob. Distrib if its joint PDF is as follows

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\, e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]}$$

for $-\infty < x < \infty$
and $-\infty < y < \infty$

Parameters: $\mu_x, \mu_y, \sigma_x, \sigma_y, \rho$

where $-\infty < \mu_x, \mu_y < \infty$

$\sigma_x, \sigma_y > 0$

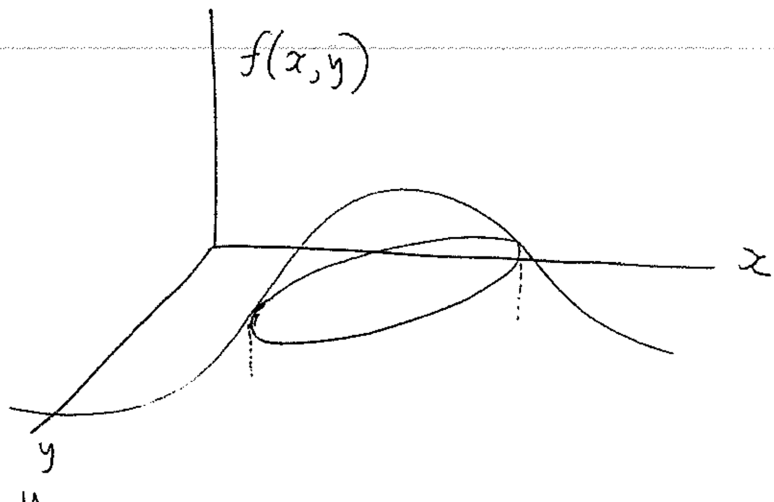
$|\rho| < 1$

SHAPE

Notice that the level surfaces of the $f(x,y)$ are ellipses.

$$\left[\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\right]$$

— and ρ controls the orientation of the ellipse



1) Marginals:

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma_x} e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2} \cdot \frac{1}{\sqrt{2\pi}\,\sigma_y\sqrt{1-\rho^2}}\, e^{\frac{-1}{2\sigma_y^2(1-\rho^2)}\left[y-\mu_y-\rho\frac{\sigma_y}{\sigma_x}(x-\mu_x)\right]^2} dy$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_x} e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2} \int \Big[\qquad\qquad\Big]\,dy$$

PDF for variable Normally distrib with mean $= \mu_y + \rho\frac{\sigma_y}{\sigma_x}(x-\mu_x)$

Var $= \sigma_y^2(1-\rho^2)$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_x} e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2}$$

(Similar result for $f_Y(y)$)

2) Conditional PDF's $\quad f(y|x) = \dfrac{f(x,y)}{f_X(x)}$

By the above it is clear that

$$f(y|x) = \frac{1}{\sqrt{2\pi}\,\sigma_y\sqrt{1-\rho^2}}\, e^{-\frac{1}{2\sigma_y^2(1-\rho^2)}\left[y-\mu_y-\rho\frac{\sigma_y}{\sigma_x}(x-\mu_x)\right]^2}$$

$$\sim N\left(\mu_y + \rho\frac{\sigma_y}{\sigma_x}(x-\mu_x),\ \sigma_y^2(1-\rho^2)\right)$$

[Similar result for $f(x|y)$]

$$E(Y|x) = \mu_y + \rho\frac{\sigma_y}{\sigma_x}(x-\mu_x)$$

$$Var(Y|x) = \sigma_y^2(1-\rho^2)$$

The Regression curve ~~for~~ of Y on X is therefore LINEAR.
ρ^2 may be interpreted as the Prop. of Var. of Y explained by X

$$f(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right]\right)$$

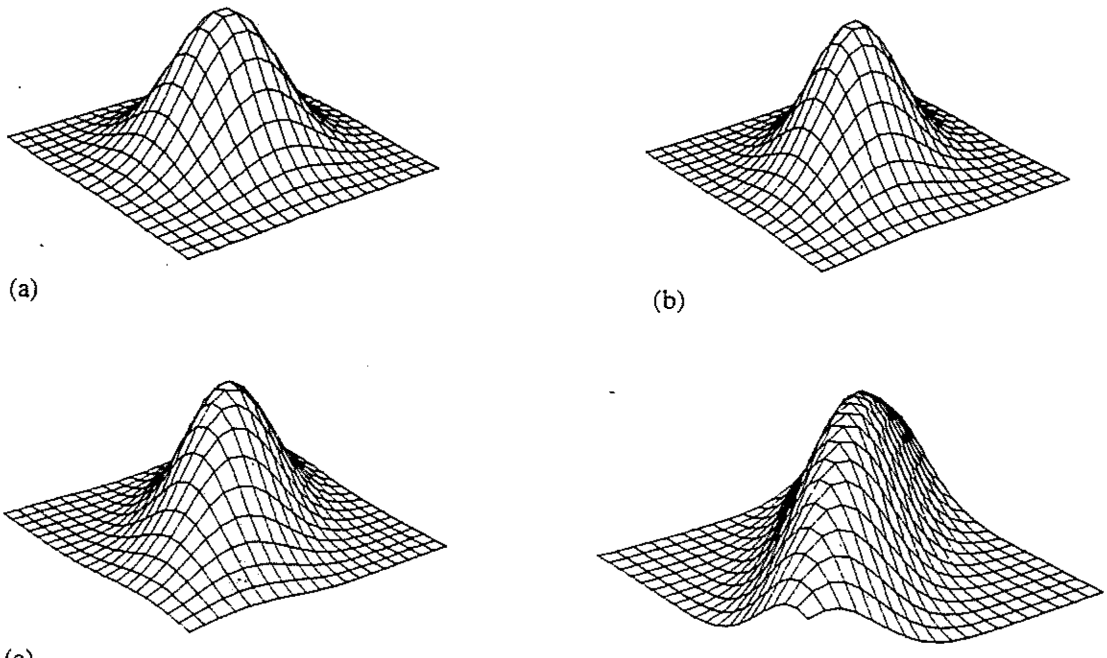
One of the earliest uses of this bivariate density was as a model for the joint distribution of the heights of fathers and sons. The density depends on five parameters:

$$-\infty < \mu_X < \infty \qquad -\infty < \mu_Y < \infty$$
$$\sigma_X > 0 \qquad \sigma_Y > 0$$
$$-1 < \rho < 1$$

The contour lines of the density are the lines in the xy plane on which the joint density is constant. From the equation above, we see that $f(x, y)$ is constant if

$$\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} = \text{constant}$$

The locus of such points is an ellipse centered at (μ_X, μ_Y). If $\rho = 0$, the axes of the ellipse are parallel to the x and y axes, and if $\rho \neq 0$, they are tilted. Figure 3-7 shows several bivariate normal densities, and Figure 3-8 shows the corresponding elliptical contours.



(a) (b) (c) (d)

Figure 3-7. Bivariate normal densities with $\mu_X = \mu_Y = 0$ and $\sigma_X = \sigma_Y = 1$ and (a) $\rho = 0$, (b) $\rho = .3$, (c) $\rho = .6$, (d) $\rho = .9$.

3) $\underline{Cov(X, Y)}$:

$$E\left[(X-\mu_x)(Y-\mu_y)\right] = \iint (x-\mu_x)(y-\mu_y) f(x,y)\, dx\, dy$$

$$= \int \frac{(x-\mu_x)}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2} \left\{ \underbrace{\int (y-\mu_y) \frac{1}{\sqrt{2\pi}\,\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2\sigma_y^2(1-\rho^2)}\left[y-\mu_y-\rho\frac{\sigma_y}{\sigma_x}(x-\mu_x)\right]^2} dy}_{\rho\frac{\sigma_y}{\sigma_x}(x-\mu_x)} \right\} dx$$

$$= \int (x-\mu_x) f_X(x)\left[\int (y-\mu_y) f(y|x)\, dy\right] dx$$

$$= \int \frac{x-\mu_x}{\sqrt{2\pi}\,\sigma_x}\left[\rho\frac{\sigma_y}{\sigma_x}(x-\mu_x)\right] e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2} dx$$

$$= \rho\frac{\sigma_y}{\sigma_x}\left[\sigma_x^2\right]$$

$$= \sigma_y \sigma_x \rho$$

Thus ρ is the Correlation Coefficient for X and Y

CASE $\rho = 0$

In this case, notice that

$$f(x,y) = f_X(x)\, f_Y(y)$$

I.E. X and Y are indep^t

Thus $\rho = 0 \implies X, Y$ indep^t

for Bivariate Normal ONLY