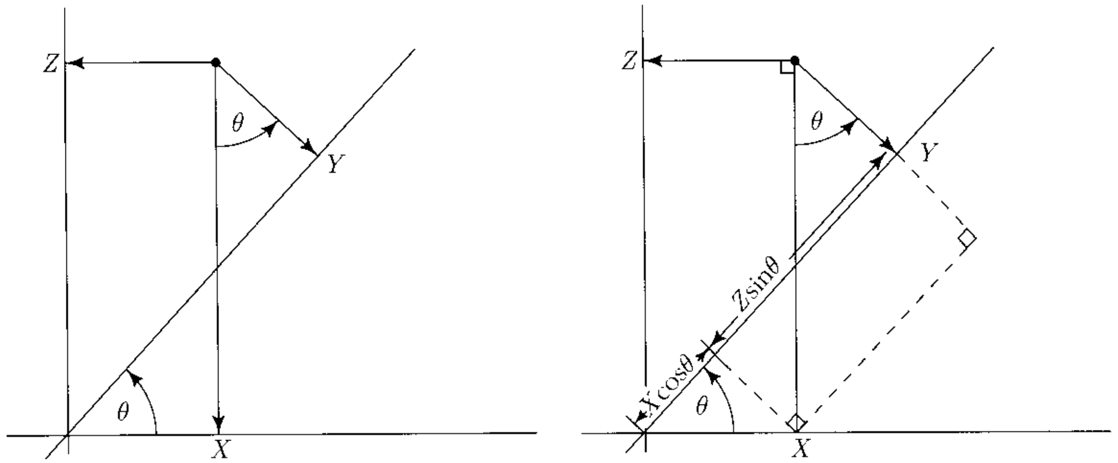
variables. A basic ingredient is the correlation coefficient, denoted here by $\rho$, often also by $r$:

$$\rho = Corr(X,Y) = E(X^*Y^*)$$

where $X^*$ is $X$ in standard units, and $Y^*$ is $Y$ in standard units. This correlation $\rho$ is a theoretical quantity, defined by expected values or integrals with respect to a bivariate distribution. In practice, such correlations are usually estimated by the corresponding empirical correlation obtained from data, with the empirical distribution of a data list $(x_1, y_1), \dots, (x_n, y_n)$ instead of the theoretical distribution, and averages instead of expectations.

## Constructing Correlated Normal Variables

To get a pair of correlated standard normal variables $X$ and $Y$, start with a pair of independent standard normal variables, say $X$ and $Z$. Let $Y$ be the projection of $(X, Z)$ onto an axis at an angle $\theta$ to the $X$-axis, as in the left-hand diagram:




By the geometry of the right-hand diagram

$$Y = X\cos\theta + Z\sin\theta$$

By rotational symmetry of the joint distribution of $X$ and $Z$, the distribution of $Y$ is standard normal. Thus

$$E(X) = E(Y) = E(Z) = 0$$
$$SD(X) = SD(Y) = SD(Z) = 1$$
$$\begin{aligned}\rho(X,Y) = E(XY) &= E[X(X\cos\theta + Z\sin\theta)] \\ &= E(X^2)\cos\theta + E(XZ)\sin\theta \\ &= \cos\theta\end{aligned}$$



since $E(X^2) = 1$, and $E(XZ) = E(X)E(Z) = 0$ by independence of $X$ and $Z$. To summarize, $X$ and $Y$ are standard normal variables with correlation $\rho = \cos\theta$. Note the special cases

| | | |
|---|---|---|
| $\theta = 0$ | when $\rho = 1$ | $Y = X$ |
| $\theta = \pi/2$ | when $\rho = 0$ | $Y = Z$ is independent of $X$ |
| $\theta = \pi$ | when $\rho = -1$ | $Y = -X$ |

For each $\rho$ between $-1$ and 1, there is an angle $\theta = \arccos\rho$, which makes $X$ and $Y$ have correlation $\rho$. Then $\cos\theta = \rho$, $\sin\theta = \sqrt{1-\rho^2}$, and

$$Y = \rho X + \sqrt{1-\rho^2}\,Z$$

where $X$ and $Z$ are independent normal (0, 1). The joint distribution of $X$ and $Y$ so defined is the *standard bivariate normal distribution with correlation $\rho$*.

### Standard Bivariate Normal Distribution

$X$ and $Y$ have standard bivariate normal distribution with correlation $\rho$ if and only if

$$Y = \rho X + \sqrt{1-\rho^2}\,Z$$

where $X$ and $Z$ are independent standard normal variables.

**Marginals.** Both $X$ and $Y$ have standard normal distribution.

**Conditionals.** Given $X = x$, $Y$ has normal $(\rho x, 1-\rho^2)$ distribution. Given $Y = y$, $X$ has normal $(\rho y, 1-\rho^2)$ distribution.

**Joint density.** The joint density of $X$ and $Y$ is

$$f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)\right\}$$

**Independence.** For $X$ and $Y$ with standard bivariate normal distribution, $X$ and $Y$ are independent if and only if $\rho = 0$.

The next two pages display the geomerty of linear transformation from $(X, Z)$ to $(X, Y)$. Following these pages is a discussion of the results presented in the above box.