

So for α as in the diagram, considered a negative angle,

$$\begin{aligned} \tan \alpha &= \frac{-\rho}{\sqrt{1-\rho^2}} \\ &= \frac{-0.5}{\sqrt{0.75}} = -1/\sqrt{3} \end{aligned}$$

So $\alpha = -30^\circ$. Thus the angle at the corner of the shaded region is $-\alpha + 90^\circ = 120^\circ$. By rotational symmetry, the chance that (X, Z) lies in the shaded region is the ratio of angles $120^\circ/360^\circ = 1/3$. So

$$P(X \geq 0, Y \geq 0) = 1/3$$

In other words, about one-third of the father-son pairs had both father and son above average height.

Problem 2. Suppose you have data on two variables with a bivariate normal distribution, and $3/8$ of the data is above average in both variables. Estimate ρ .

Solution. Transform to standard units and use the same linear change of variable as in the solution of the previous problem. Now

$$\frac{3}{8} = \frac{135^\circ}{360^\circ}$$

so the angle of the corner at the origin is 135° . Thus α in the diagram is -45° , and by the previous solution

$$\frac{-\rho}{\sqrt{1-\rho^2}} = \tan \alpha = \tan(-45^\circ) = -1$$

So $\rho = 1/\sqrt{2}$.

Example 3. Conditional expectation of Y given X in an interval.

Suppose (X, Y) has standard bivariate normal density with correlation ρ .

Problem. For $a < b$, find $E(Y|a < X < b)$.

Solution. Given that X has a particular value $x \in (a, b)$, the expected value of Y is

$$E(Y|X = x) = \rho x.$$

Given just $(a < X < b)$ the precise value of X is unknown. But by the rule of average conditional expectations, $E(Y|a < X < b)$ can be found by integration of the conditional expectation $E(Y|X = x) = \rho x$ with respect to the conditional density of X given $a < X < b$. This gives

$$E(Y|a < X < b) = \int_a^b \rho x f_X(x|a < X < b) dx$$

where for $a < x < b$

$$\begin{aligned} f_X(x|a < X < b) dx &= P(X \in dx|a < X < b) \\ &= \frac{P(X \in dx, a < X < b)}{P(a < X < b)} \\ &= \frac{P(X \in dx)}{P(a < X < b)} \\ &= \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}}{\Phi(b) - \Phi(a)} dx \end{aligned}$$

Substituting this expression gives

$$E(Y|a < X < b) = \int_a^b \rho x \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}}{\Phi(b) - \Phi(a)} dx = \frac{\rho}{\sqrt{2\pi}} \frac{[e^{-\frac{1}{2}a^2} - e^{-\frac{1}{2}b^2}]}{\Phi(b) - \Phi(a)}$$

Example 4. Midterm and final.

Midterm and final scores in a large class have an approximately bivariate normal distribution, with parameters

midterm scores:	mean: 65	SD: 18
final scores:	mean: 60	SD: 20
correlation: 0.75		

Problem. Estimate the average final score of students who were above average on the midterm.

Solution. Let X and Y denote the midterm and final scores in standard units. The event "midterm score above average" is the same as the event $X > 0$. Take $a = 0$ and $b = \infty$ in the previous example to get

$$E(Y|X > 0) = \frac{\rho}{\sqrt{2\pi}} \left[\frac{1-0}{0.5} \right] = \frac{0.75 \times 2}{\sqrt{2\pi}} \approx 0.6$$

So the average final score of those who scored above average on the midterm is 0.6, in standard units. Thus the required score is

$$60 + 20 \times 0.6 = 72$$