

EX: Which of the following joint density functions have (correlation) $\rho_{XY} = 0$?

(Remember that $\rho_{XY} = 0$ is possible with symmetry even if X and Y are not independent.)

a)

$$f(x, y) = \begin{cases} \frac{1}{\pi} & (x-1)^2 + y^2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

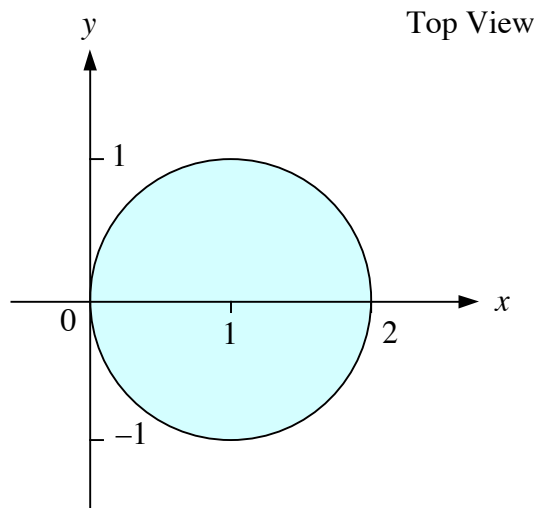
b)

$$f(x, y) = \begin{cases} \frac{1}{[\ln(2)]^2} \frac{1}{xy} & 1 \leq x \leq 2 \text{ and } 1 \leq y \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

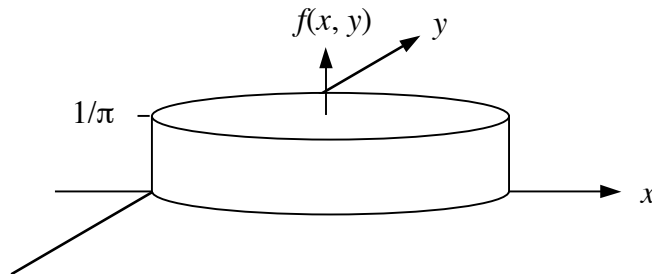
c) Estimate by eye the value of x for which $F(x) = 0.25$.

$$f(x, y) = \begin{cases} x + y & 0 \leq x \leq 1 \text{ and } 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

SOL'N: a) The region, $x^2 + y^2 \leq 1$, on which $f(x, y) \neq 0$ is called the support of $f(x, y)$. It is a circle of radius one, centered at $x = 1$ and $y = 0$, as shown below.



Since $f(x, y) = 1/\pi$ is constant on its support, $f(x, y)$ is a cylinder as shown below.



By symmetry, we can say that the mean values for x and y are $\mu_x = 1$ and $\mu_y = 0$. Since correlation is measured relative to mean values we have the following conceptual tool that allows us to simplify the calculation of ρ_{xy} :

TOOL: Variance and covariance terms are always expressed in terms of distances from means:

$$\sigma_X^2 = E((X - \mu_X)^2), \quad \sigma_{XY} = E((X - \mu_X)(Y - \mu_Y))$$

Thus, we get the same values for a probability density function (pdf) such as $f_X(x)$ or $f(x, y)$ shifted so that the mean values are at the origin. In other words, slide $f(x, y)$ around on the xy -plane so that $\mu_X = 0$ and $\mu_Y = 0$. For this shifted $f(x, y)$, the calculation of variances is simplified:

$$\sigma_X^2 = E(X^2), \quad \sigma_{XY} = E(XY)$$

Applying this tool to the present problem yields a new pdf that is a cylinder centered on the origin:

$$f(x, y) = \begin{cases} \frac{1}{\pi} & x^2 + y^2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Since correlation is defined in terms of variances and covariance, the correlation of this new $f(x, y)$ will be the same as for the original $f(x, y)$:

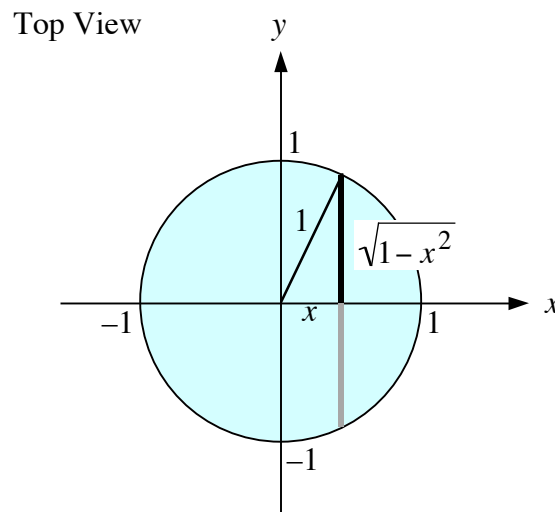
$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

The covariance, σ_{XY} , appears in the numerator of the definition of correlation, and we examine it first in case we can show it is zero. If we can do so, then it follows that the correlation is zero and we are done.

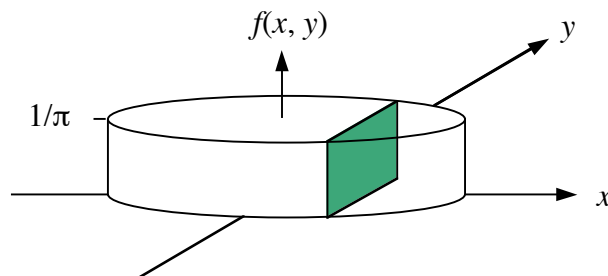
As noted above, $\sigma_{XY} = E(XY)$ for our new $f(x, y)$. Calculation of $E(x, y)$ involves a double integral:

$$E(X,Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x,y)dydx$$

To determine the correct substitutions for the limits of integration, we consider the support of a cross-sections of $f(x, y)$ in the y direction for one value of x , as shown below.



The illustration, below, shows the cross section itself at this same value of x .



From the diagrams we have the following integral:

$$\sigma_{XY} = E(X,Y) = \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} xy \frac{1}{\pi} dy dx$$

$$\sigma_{XY} = E(X,Y) = \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} xy \frac{1}{\pi} dy dx = \int_{-1}^1 x \frac{1}{\pi} \frac{y^2}{2} \Big|_{y=-\sqrt{1-x^2}}^{y=\sqrt{1-x^2}} dx = \int_{-1}^1 x \frac{1}{\pi} \cdot 0 dx$$

Rather than calculating the inner integral, we observe that it is the integral of an odd function, (i.e., $g(y) = -g(-y)$), with equal but opposite limits of integration. In other words, the value of the integrand at y is equal to minus one times the value of the integrand at $-y$. Thus, these two points cancel each other out in the integral. This happens at every value of y , and it follows that the integral is zero.

$$\sigma_{XY} = E(X,Y) = 0$$

We can verify this result by hand or by using the following Matlab[®] commands to compute the integral:

```
syms x y
int(x*y/pi, y, -sqrt(1-x^2), sqrt(1-x^2))
```

Since the covariance is zero, it follows that the correlation is zero:

$$\rho_{XY} \equiv \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = 0$$

- b) The correlation for this joint probability density function is zero since X and Y are independent.

$$\rho_{XY} = 0$$

To show independence of X and Y , we write $f(x, y)$ as the product of $f_X(x)$ and $f_Y(y)$.

$$f_X(x) = \begin{cases} \frac{1}{\ln(2)x} & 1 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{1}{\ln(2)y} & 1 \leq y \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

NOTE: We are able to separate the conditions on x and y for this $f(x, y)$. This is possible because the conditions on x (i.e., $1 \leq x \leq 2$) do not include y and vice versa. If the conditions for x depend on the value of y or vice versa, then X and Y are almost certainly dependent.

- c) We suspect the correlation for this joint probability density function is nonzero since X and Y are dependent. The dependence of X and Y is revealed by taking cross-sections and seeing that they have different shapes after being scaled vertically to make their area equal one. That is, the conditional probabilities change with the value of the variable being conditioned on.

Any hoped for symmetry is elusive, as the mean values in the x and y directions are not immediately apparent, and turn out to be awkward values:

$$\mu_X = \int_0^1 x f_X(x) dx = \int_{x=0}^{x=1} x \int_{y=0}^{y=1} (x+y) dy dx = \int_{x=0}^{x=1} x \left(x + \frac{1}{2} \right) dx = \frac{7}{12}$$

Note that the quantity in the parentheses in the last integral is $f_X(x)$, which we will use below.

By symmetry we have $\mu_Y = \mu_X = 7/12$. These values are awkward because they are not in the center of the figure. Thus, it seems unlikely that we will be able to pair up all points in such a way that they cancel out in calculation of correlation.

We proceed with the calculation of variances and covariance needed for the calculation of correlation.

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_0^1 x^2 \left(x + \frac{1}{2}\right) dx = \frac{5}{12}$$

Subtract the mean squared to find the variance:

$$\sigma_X^2 = E(X^2) - \mu_X^2 = \frac{5}{12} - \left(\frac{7}{12}\right)^2 = \frac{11}{144}$$

By symmetry the variance is the same for Y .

Now compute the covariance. The first step is to find $E(XY)$.

$$E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f(x, y) dy dx = \int_{x=0}^{x=1} x \int_{y=0}^{y=1} y(x + y) dy dx$$

or

$$E(XY) = \int_{x=0}^{x=1} x \left(\frac{1}{2}x + \frac{1}{3}\right) dx = \frac{1}{3}$$

Now subtract the product of the means.

$$\sigma_{XY} = E(XY) - \mu_X \mu_Y = \frac{1}{3} - \left(\frac{7}{12}\right)^2 = \frac{-1}{144}$$

Now we plug values into the formula for correlation:

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{-1/144}{\sqrt{11/144} \sqrt{11/144}} = \frac{-1}{11}$$

The correlation is nonzero but very small.