

1. Uniform distribution on a rotated square

(a) X and Y are not independent since the footprint of their joint pdf is not rectangular.

The area of the square region is 2 so that the pdf of (X, Y) is $\frac{1}{2}$ within the region. Integrating out the y variable of the joint density function yields the triangle shaped pdf:

$$f_X(x) = \begin{cases} \int_{-1-x}^{1+x} \frac{1}{2} dy = 1+x, & -1 < x \leq 0 \\ \int_{x-1}^{1-x} \frac{1}{2} dy = 1-x, & 0 < x \leq 1 \\ 0 & \text{else} \end{cases}$$

Equivalently,

$$f_X(x) = \begin{cases} 1-|x| & -1 \leq x \leq 1 \\ 0 & \text{else} \end{cases}$$

Since the joint pdf is symmetric in x and y , it follows that the marginal density for Y is exactly the same.

(b)

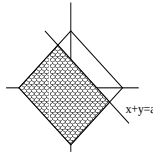
$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_{-1}^1 x(1-|x|) dx = 0$$

since an odd function integrates to zero.

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_{-1}^1 x^2(1-|x|) dx = 2 \int_0^1 x^2(1-x) dx = 2 \int_0^1 x^2 - x^3 dx = 2\left(\frac{1}{3} - \frac{1}{4}\right) = \frac{1}{6}$$

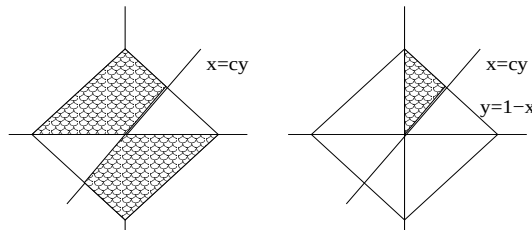
So $Var(X) = E[X^2] - E[X]^2 = \frac{1}{6}$.

(c) The range of A is the interval $[-1, 1]$. If a is in this interval, then the event $\{A \leq a\}$ is the same as (X, Y) falling within the halfspace $\{(x, y) : x + y \leq a\}$. The halfspace intersects the rotated square in a rectangle of dimensions $\frac{1+a}{\sqrt{2}} \times \sqrt{2}$. See the figure. The area of this rectangle is $1+a$. Since the pdf of (X, Y) is $\frac{1}{2}$ within



the rectangle, $P(A \leq a) = \frac{1+a}{2}$. Hence, $X + Y$ is uniformly distributed over the interval $[-1, 1]$.

(d) The range of C is the whole real line. Our approach will be to first find the CDF and pdf of C within the region $c \geq 0$. So let $c \geq 0$. Then $F_C(c)$ is one half the area of the shaded region in the diagram on the left below, which is 0.5 plus the area of the shaded region in the diagram on the right. View the area



as a triangle rotated clockwise a quarter revolution. The area of the triangle is one half the base times the height. The base is the unit length segment from $(0, 0)$ to $(0, 1)$. The height is the value of x at which the line $\{x = cy\}$ intersects the edge of the square, given by the equation $y = 1 - x$. Solving for x yields $x = \frac{c}{1+c}$. The area of the shaded triangle is thus $\frac{x}{2} = \frac{c}{2(1+c)}$ so that $F_C(c) = 0.5 + \frac{c}{2(1+c)}$ for $c \geq 0$. Differentiating this with respect to c yields that $f_C(c) = \frac{1}{2(1+c)^2}$ for $c \geq 0$.

We could find the pdf for negative c by a similar method. However, it is less work to notice that (X, Y) and $(-X, Y)$ have the same joint distribution, so that C has the same distribution as $-C$. Thus, f_C is an even function: $f_C(-c) = f_C(c)$. Therefore, $f_C(c) = \frac{1}{2(1+|c|)^2}$ for all real values of c .

2. Functions of independent exponential random variables

(a) Z takes values in the positive real line. So let $z \geq 0$.

$$\begin{aligned} P[Z \leq z] &= P[\min\{X_1, X_2\} \leq z] = P[X_1 \leq z \text{ or } X_2 \leq z] \\ &= 1 - P[X_1 > z \text{ and } X_2 > z] = 1 - P[X_1 > z]P[X_2 > z] = 1 - e^{-\lambda_1 z} e^{-\lambda_2 z} = 1 - e^{-(\lambda_1 + \lambda_2)z} \end{aligned}$$

Differentiating yields that

$$f_Z(z) = \begin{cases} (\lambda_1 + \lambda_2)e^{-(\lambda_1 + \lambda_2)z}, & z \geq 0 \\ 0, & z < 0 \end{cases}$$

That is, Z has the exponential distribution with parameter $\lambda_1 + \lambda_2$.

(b) R takes values in the positive real line and by independence the joint pdf of X_1 and X_2 is the product of their individual densities. So for $r \geq 0$,

$$\begin{aligned} P[R \leq r] &= P\left[\frac{X_1}{X_2} \leq r\right] = P[X_1 \leq rX_2] \\ &= \int_0^\infty \int_0^{rx_2} \lambda_1 e^{-\lambda_1 x_1} \lambda_2 e^{-\lambda_2 x_2} dx_1 dx_2 \\ &= \int_0^\infty (1 - e^{-r\lambda_1 x_2}) \lambda_2 e^{-\lambda_2 x_2} dx_2 = 1 - \frac{\lambda_2}{r\lambda_1 + \lambda_2}. \end{aligned}$$

Differentiating yields that

$$f_R(r) = \begin{cases} \frac{\lambda_1 \lambda_2}{(\lambda_1 r + \lambda_2)^2} & r \geq 0 \\ 0, & r < 0 \end{cases}$$

3. Sums of independent normal random variables

Let X denote Jill's score and Y denote Jack's score. Then X and Y are independent normal random variables. The variable X has mean 170 and variance $20^2 = 400$, and Y has mean 160 and variance $15^2 = 225$.

(a) We want to compute $P[X \leq Y]$. This is the same as $P[Y - X \geq 0]$. The variable $-X$ is normal with mean -170 and variance 20^2 , so that $Y - X$ is a sum of independent normal random variables. The mean of $Y - X$ is $160 - 170 = -10$, and the variance is $15^2 + 20^2 = 625$. Therefore

$$P[X \leq Y] = P[Y - X \geq 0] = P\left[\frac{Y - X + 10}{\sqrt{625}} \geq \frac{10}{\sqrt{625}}\right] = Q\left(\frac{10}{\sqrt{625}}\right) = Q(0.4) = 1 - \Phi(0.4) = 0.3446.$$

(b) Similarly, $X + Y$ is normal with mean 330 and variance 625. Therefore

$$P[X + Y \geq 350] = P\left[\frac{X + Y - 330}{\sqrt{625}} \geq \frac{350 - 330}{\sqrt{625}}\right] = Q\left(\frac{20}{\sqrt{625}}\right) = Q(0.8) = 1 - \Phi(0.8) = 0.2119.$$

4. Sizing a Confidence Interval by the normal approximation to the binomial distribution

If n observations are made and p is given, then the probability that p lies in the random confidence interval $[\hat{p} - \delta, \hat{p} + \delta]$ satisfies

$$\begin{aligned} P[|\hat{p} - p| \leq \delta] &= P\left[\left|\frac{X}{n} - p\right| \leq \delta\right] = P\left[\left|\frac{X - np}{\sqrt{p(1-p)}}\right| \leq \delta \sqrt{\frac{n}{p(1-p)}}\right] \\ &\approx \Phi\left(\delta \sqrt{\frac{n}{p(1-p)}}\right) - \Phi\left(-\delta \sqrt{\frac{n}{p(1-p)}}\right) = 2\Phi\left(\delta \sqrt{\frac{n}{p(1-p)}}\right) - 1 \end{aligned}$$

(a) confidence $\approx 2\Phi\left(0.02 \sqrt{\frac{1000}{0.5(1-0.5)}}\right) - 1 = 2\Phi(1.265) - 1 = 0.794 = 79.4\%$

(b) confidence $\approx 2\Phi\left(0.02 \sqrt{\frac{1000}{0.1(1-0.9)}}\right) - 1 = 2\Phi(2.108) - 1 = 0.965 = 96.5\%$

(c) Select δ so that $2\Phi\left(\delta \sqrt{\frac{1000}{p(1-p)}}\right) - 1 = 0.99$. Observing from the table for standard normal that

$2\Phi(2.58) - 1 \approx 0.99$, we select δ so that $\delta \sqrt{\frac{1000}{p(1-p)}} = 2.58$, or

$$\delta = 2.58 \sqrt{\frac{p(1-p)}{1000}}$$

The 99% confidence interval for $p = 0.5$ requires $\delta = 2.58\sqrt{\frac{0.5(1-0.5)}{1000}} = 0.04$

(d) Similarly, the 99% confidence interval for $p = 0.1$ requires $\delta = 2.58\sqrt{\frac{0.1(0.9)}{1000}} = 0.025$

(e) The product $p(1-p)$, and hence the required δ , is maximized by $p = 0.5$. Thus, if $\delta = 0.04$ as found in part (c), then the confidence interval contains p with probability at least 0.99 (no matter what the value of p is), at least up to the accuracy of the normal approximation.

(f) The value of n that works for $p = 0.5$ works for any n (the situation is similar to that in part (e)) so n needs to be selected so that $0.01 = 2.58\sqrt{\frac{0.5(1-0.5)}{n}}$. This yields $n = (\frac{2.58}{0.01})^2(0.5)(1-0.5) \approx 16,641$.

5. Transformation of jointly continuous random variables, I

We have $X = g_1(Z, \Theta)$ and $Y = g_2(Z, \Theta)$, where $g_1(z, \theta) = \sqrt{2z}\cos(\theta)$ and $g_2(z, \theta) = \sqrt{2z}\sin(\theta)$, for (z, θ) ranging over the infinite half strip $z > 0$ and $0 \leq \theta \leq 2\pi$. We use coordinates (x, y) for the range of this mapping, so any point (z, θ) in the infinite half strip in the $z - \theta$ plane is mapped to the point (x, y) in the $x - y$ plane. The mapping can be written concisely as:

$$\begin{aligned}x &= \sqrt{2z}\cos(\theta) \\y &= \sqrt{2z}\sin(\theta)\end{aligned}$$

This mapping is applied to (Z, Θ) which has joint pdf $f_{Z, \Theta}(z, \theta) = \frac{1}{2\pi}e^{-z}$ within the infinite half strip. The range of the mapping is the entire plane. Indeed, for any $(x, y) \neq 0$ we can recover (z, θ) by

$$\begin{aligned}z &= \frac{x^2 + y^2}{2} \\ \theta &= \tan^{-1}\left(\frac{y}{x}\right)\end{aligned}$$

The Jacobian determinant of the transformation is

$$J(z, \theta) = \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \frac{\cos \theta}{\sqrt{2z}} & -\sqrt{2z} \sin \theta \\ \frac{\sin \theta}{\sqrt{2z}} & \sqrt{2z} \cos \theta \end{vmatrix} = 1 \neq 0, \quad \text{for all } z, \theta$$

Therefore the required pdf is

$$f_{X,Y}(x, y) = \frac{f_{Z, \Theta}(z, \theta)}{|J(z, \theta)|} = \frac{1}{2\pi}e^{-z} = \frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}, \quad -\infty < x < \infty, \quad -\infty < y < \infty$$

This joint density can be factored into the product of two standard normal density functions,

$$f_{X,Y}(x, y) = \left(\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\right) \cdot \left(\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\right) = f_X(x) \cdot f_Y(y)$$

so that indeed X and Y are independent, standard normal random variables.

(b) There are several ways to calculate the pdf of R . The easiest is to realize that $R = \sqrt{2Z}$. The range of R is the positive reals, and for $r > 0$ the CDF of R is given by

$$F_R(r) = P[R \leq r] = P[\sqrt{2Z} < r] = P[Z \leq \frac{r^2}{2}] = 1 - \exp(-\frac{r^2}{2})$$

Differentiating with respect to r yields

$$f_R(r) = \begin{cases} re^{-\frac{r^2}{2}}, & r > 0 \\ 0, & r \leq 0 \end{cases}$$

This is the Rayleigh probability density function. This problem shows that the Rayleigh pdf arises quite naturally.

6. Transformation of jointly continuous random variables, II

We are using the mapping, from the square region $\{(u, v) : 0 \leq u, v \leq 1\}$ in the $u - v$ plane to the triangle with corners $(0,0)$, $(3,0)$, and $(3,1)$ in the $x - y$ plane, given by

$$\begin{aligned}x &= 3u \\y &= uv.\end{aligned}$$

The mapping is one-to-one, meaning that for any (x, y) in the range we can recover (u, v) . Indeed, the inverse mapping is given by

$$\begin{aligned}u &= \frac{x}{3} \\v &= \frac{3y}{x}.\end{aligned}$$

The Jacobian determinant of the transformation is

$$J(u, v) = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 3 & 0 \\ v & u \end{vmatrix} = 3u \neq 0, \quad \text{for all } u, v \in (0, 1)^2.$$

Therefore the required pdf is

$$f_{X,Y}(x, y) = \frac{f_{U,V}(u, v)}{|J(u, v)|} = \frac{9u^2v^2}{|3u|} = 3uv^2 = \frac{9y^2}{x}$$

within the triangle with corners $(0,0)$, $(3,0)$, and $(3,1)$, and $f_{X,Y}(x, y) = 0$ elsewhere.

(b) By integrating out y from the joint pdf, we find

$$f_X(x) = \begin{cases} \int_0^{\frac{x}{3}} \frac{9y^2}{x} dy = \frac{x^2}{9} & \text{if } 0 \leq x \leq 3 \\ 0 & \text{else} \end{cases}$$

Therefore the conditional density $f_{Y|X}(y|x)$ is well defined only if $0 \leq x \leq 3$. For $0 \leq x \leq 3$ we have

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x, y)}{f_X(x)} = \begin{cases} \frac{81y^2}{x^3} & \text{if } 0 \leq y \leq \frac{x}{3} \\ 0 & \text{else} \end{cases}$$