

Properties of Expectation

Assigned reading: *Ross* Chapter 7, Sections 1-5 (not starred sections).

Noncredit Exercises: pp. 379-389: 1-8, 10-12, 26, 38-40, 45, 50-51.

1. Let X and Y be jointly uniformly distributed on the unit square (with vertices $(0, 0)$, $(0, 1)$, $(1, 0)$, $(1, 1)$).

- (a) Are X and Y independent? Are they uncorrelated?

As we learned in class, when two random variables are jointly uniform on an axis-aligned rectangle, then they are independent, and the marginals are uniformly distributed on their corresponding intervals. Thus, in this case, X is uniform on $(0, 1)$, and Y is uniform on $(0, 1)$. Since they are independent, they are necessarily uncorrelated.

- (b) Let $W = X + Y$ and $Z = X - Y$. Are W and Z independent? Are they uncorrelated?

First we'll deal with independence. To do so, we'll need the joint density f_{WZ} . We use the usual approach. First, we solve the equations

$$a = u_1 + v_1, \quad b = u_1 - v_1 \quad \Rightarrow \quad u_1 = \frac{a+b}{2}, \quad v_1 = \frac{a-b}{2}$$

Next, we compute the determinant of the Jacobian

$$J = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad \Rightarrow \quad \det\{J(u, v)\} = -2$$

Thus,

$$f_{ZW}(a, b) = \frac{f_{XY}(u_1, v_1)}{|\det\{J(u, v)\}|} = \frac{1}{2}, \quad 0 \leq \frac{a+b}{2} \leq 1, \quad 0 \leq \frac{a-b}{2} \leq 1$$

The region in which f_{WZ} is non-zero (also called the support of f_{WZ}) is *not* an axis-aligned rectangle, so W and Z cannot be independent.

Now for the question of correlation... First, with X and Y uniform on the unit interval, we have

$$E[X] = E[Y] = \frac{1}{2}, \quad E[X^2] = E[Y^2] = \frac{1}{3}, \quad \text{Var}(X) = \text{Var}(Y) = \frac{1}{12}.$$

Thus,

$$E[W] = E[X] + E[Y] = 1, \quad E[Z] = E[X] - E[Y] = 0, \quad \text{Var}(W) = \text{Var}(Z) = \text{Var}(X) + \text{Var}(Y) = \frac{1}{6}.$$

Finally,

$$\text{Cov}(W, Z) = E[ZW] - E[Z]E[W] = E[ZW] = E[(X+Y)(X-Y)] = E[X^2 - Y^2] = 0.$$

Therefore, W and Z are uncorrelated, but they are *not* independent.

2. Let X and Y be jointly uniformly distributed on the square with vertices $(2, 0)$, $(1, 1)$, $(0, 0)$, $(1, -1)$.

- (a) Are X and Y independent? Are they uncorrelated?

Since the support is not an axis-aligned rectangle, X and Y cannot be independent. To determine if they are uncorrelated, we compute the covariance. First, $E[Y] = 0$ (note: $E[X] \neq 0$), so $\text{Cov}(X, Y) = E[XY] - E[X]E[Y] = E[XY]$.

$$E[XY] = \int_{u=0}^1 \int_{v=2-u}^u \frac{1}{2} dv du + \int_{u=1}^2 \int_{v=u-2}^u \frac{1}{2} dv du = 0$$

Therefore, X and Y are uncorrelated, but they are *not* independent.

- (b) Let $W = X + Y$ and $Z = X - Y$. Are W and Z independent? Are they uncorrelated?

First we'll deal with the question of independence by computing the joint density. As with question 2, we have

$$a = u_1 + v_1, \quad b = u_1 - v_1 \quad \Rightarrow \quad u_1 = \frac{a+b}{2}, \quad v_1 = \frac{a-b}{2}$$

and

$$J = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad \Rightarrow \quad \det\{J(u, v)\} = -2.$$

Thus,

$$f_{ZW}(a, b) = \frac{f_{XY}(u_1, v_1)}{|\det\{J(u, v)\}|} = \frac{1}{4}, \quad 0 \leq a \leq 2, \quad 0 \leq b \leq 2.$$

Since W and Z are jointly uniform on an axis-aligned rectangle, they are independent, and the marginals are uniformly distributed on their corresponding intervals. Thus, in this case, W is uniform on $(0, 2)$, and Z is uniform on $(0, 2)$. Since they are independent, they are necessarily uncorrelated.

3. Let the random variables X and Y be such that $E[X] = 1$, $E[Y] = 4$, $Var(X) = 4$, $Var(Y) = 9$, and $\rho = 0.1$. Let $W = 3X + Y + 2$.

- (a) Find $E[W]$ and $Var(W)$.

Since expectation is linear,

$$E[3X + Y + 2] = 3E[X] + E[Y] + 2 = 9.$$

The variance is found with equal ease using the relation

$$Var(A + B) = Var(A) + Var(B) + 2Cov(A, B).$$

Since 2 is a constant, $Var(2) = 0$ and $Cov(A, 2) = 0$ for any A . Thus

$$Var(3X + Y + 2) = Var(3X + Y) = Var(3X) + Var(Y) + 2Cov(3X, Y) = 9Var(X) + Var(Y) + 6Cov(X, Y) = 48.6$$

- (b) If X and Y are jointly Gaussian random variables, what is $P\{W > 0\}$?

We know that a sum of Gaussian random variables is a Gaussian random variable. In this case, W is Gaussian with mean $\mu_W = E[3X + Y + 2] = 9$ and variance $\sigma_W^2 = Var(3X + Y + 2) = 48.6$. Thus,

$$P\{W > 0\} = 1 - \Phi\left(\frac{0 - \mu_W}{\sigma_W}\right) = 1 - \Phi\left(\frac{-9}{\sqrt{48.6}}\right) = \Phi\left(\frac{9}{\sqrt{48.6}}\right)$$

4. Let X and Y be jointly Gaussian zero-mean random variables, with variances σ_X^2 and σ_Y^2 (respectively) and with correlation coefficient ρ . Find the angle θ such that the random variables $Z = X \cos \theta + Y \sin \theta$ and $W = -X \sin \theta + Y \cos \theta$ are *independent* Gaussian random variables. Note that this coordinate transformation corresponds to a rotation of the axes by θ . Thus, in this problem you will show how to transform two dependent Gaussian random variables to independent Gaussian random variables by a simple rotation.

To simplify notation, define $a = \cos \theta$ and $b = \sin \theta$. Thus, $W = aX + bY$ and $Z = -bX + aY$. Since this is a linear transformation of jointly Gaussian random variables, we know that W and Z are jointly Gaussian. Thus, if W and Z are uncorrelated, they are independent. Thus, we need

$$\begin{aligned} 0 = Cov(Z, W) &= Cov(aX + bY, -bX + aY) \\ &= Cov(aX, -bX) + Cov(aX, aY) + Cov(bY, -bX) + Cov(bY, aY) \\ &= -abCov(X, X) + a^2Cov(X, Y) - b^2Cov(Y, X) + abCov(Y, Y) \\ &= ab(\sigma_Y^2 - \sigma_X^2) + (a^2 - b^2)Cov(X, Y) \\ &= \cos \theta \sin \theta (\sigma_Y^2 - \sigma_X^2) + (\cos^2 \theta - \sin^2 \theta) \rho \sigma_X \sigma_Y \\ &= \frac{1}{2} \sin 2\theta (\sigma_Y^2 - \sigma_X^2) + \rho \sigma_X \sigma_Y \cos 2\theta \end{aligned}$$

which leads to

$$\tan 2\theta = \frac{2\rho\sigma_X\sigma_Y}{(\sigma_Y^2 - \sigma_X^2)}.$$

For the special case of $\sigma_X = \sigma_Y$, we find $\theta = \pi/4$.

5. Let X and Y be jointly Gaussian zero-mean random variables, with variances σ_X^2 and σ_Y^2 (respectively) and with correlation coefficient ρ . Show that the marginal density f_X is the zero-mean Gaussian density with variance σ_X^2 . (Hint: First factor out the terms that do not depend on u . Then, complete the square in the exponent of the joint density function. Finally, integrate the joint density to obtain the marginal for X .)

We begin by doing a bit of work on the exponent of f_{XY} to make it a bit easier to factor the marginalizing integral.

$$\begin{aligned} 2(1 - \rho^2)Q(u, v) &= \frac{(u - \mu_x)^2}{\sigma_X^2} - 2\rho \frac{(u - \mu_x)(v - \mu_Y)}{\sigma_X \sigma_Y} + \frac{(v - \mu_Y)^2}{\sigma_Y^2} \\ &= \frac{(u - \mu_x)^2}{\sigma_X^2} - 2\rho \frac{(u - \mu_x)(v - \mu_Y)}{\sigma_X \sigma_Y} + \rho^2 \frac{(v - \mu_Y)^2}{\sigma_Y^2} + (1 - \rho^2) \frac{(v - \mu_Y)^2}{\sigma_Y^2} \\ &= \left[\frac{(u - \mu_x)}{\sigma_X} - \rho \frac{(v - \mu_Y)}{\sigma_Y} \right]^2 + (1 - \rho^2) \frac{(v - \mu_Y)^2}{\sigma_Y^2} \\ &= \frac{(u - \mu')^2}{\sigma_X^2} + (1 - \rho^2) \frac{(v - \mu_Y)^2}{\sigma_Y^2} \end{aligned}$$

in which $\mu' = \mu_x + \rho \frac{\sigma_X}{\sigma_Y}(v - \mu_Y)$. Now, we compute f_Y by integrating the joint density f_{XY} :

$$\begin{aligned} f_Y(v) &= \int_{-\infty}^{\infty} f_{XY}(u, v) du \\ &= \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp \left[-\frac{(u-\mu')^2}{2(1-\rho^2)\sigma_X^2} - \frac{(v-\mu_Y)^2}{2\sigma_Y^2} \right] du \\ &= \frac{1}{\sigma_Y\sqrt{2\pi}} \exp \left[-\frac{(v-\mu_Y)^2}{2\sigma_Y^2} \right] \int_{-\infty}^{\infty} \frac{1}{\sigma_X\sqrt{2\pi(1-\rho^2)}} \exp \left[-\frac{(u-\mu')^2}{2(1-\rho^2)\sigma_X^2} \right] du \end{aligned}$$

But the integral in the final expression is merely the integral of a Gaussian pdf (with mean μ' and variance $(1-\rho^2)\sigma_X^2$), and therefore, the integral evaluates to one, leaving us with

$$f_Y(v) = \frac{1}{\sigma_Y\sqrt{2\pi}} \exp \left[-\frac{(v-\mu_Y)^2}{2\sigma_Y^2} \right]$$

which is a Gaussian pdf with mean μ_Y and variance σ_Y^2 .

6. Suppose that the value of Y is to be estimated in terms of the random variables X_1 and X_2 by the linear predictor $g(X_1, X_2) = a + bX_1 + cX_2$. Further, suppose that $E[X_1] = E[X_2] = Cov(X_1, X_2) = 0$. Determine the values for a , b and c that minimize $E[(Y - (a + bX_1 + cX_2))^2]$.

We want to find parameters a , b and c to minimize $\mathcal{L} = E[(Y - a - bX_1 - cX_2)^2]$. To do this, we expand the square, use the linearity of expectation and then compute the partial derivatives with respect to the parameters and set these partials to zero.

$$\begin{aligned} \mathcal{L} &= E[Y^2] - 2aE[Y] - 2bE[X_1Y] - 2cE[X_2Y] + 2abE[X_1] + b^2E[X_1^2] + 2acE[X_2] + c^2E[X_2^2] + 2bcE[X_1X_2] + a^2 \\ &= E[Y^2] - 2aE[Y] - 2bE[X_1Y] - 2cE[X_2Y] + b^2E[X_1^2] + c^2E[X_2^2] + a^2 \end{aligned}$$

$$\frac{\partial}{\partial a} = -2\mu_Y + 2a = 0 \quad (1)$$

$$\frac{\partial}{\partial b} = -2E[YX_1] + 2bE[X_1^2] = 0 \quad (2)$$

$$\frac{\partial}{\partial c} = -2E[YX_2] + 2cE[X_2^2] = 0 \quad (3)$$

Now it's just algebra... First, solve for a using (1):

$$a = \mu_Y$$

Now solve for b using (2):

$$\begin{aligned} 0 &= -E[YX_1] + bE[X_1^2] \\ b &= \frac{E[YX_1]}{E[X_1^2]} = \frac{Cov(Y, X_1)}{\sigma_{X_1}} = \frac{\sigma_Y}{\sigma_{X_1}} \rho_{YX_1} \end{aligned}$$

Finally, solve for c using (3):

$$\begin{aligned} 0 &= -2E[YX_2] + 2cE[X_2^2] \\ c &= \frac{E[YX_2]}{E[X_2^2]} = \frac{Cov(Y, X_2)}{\sigma_{X_2}} = \frac{\sigma_Y}{\sigma_{X_2}} \rho_{YX_2} \end{aligned}$$

7. Problem 73, page 388 of Ross.

- We are told that $f_{R|S}$ is a Gaussian pdf with mean S . Thus, $E[R|S] = S$. Taking the expectations of both sides of this equality we obtain $E[S] = E[E[R|S]]$. Finally, using the fact that $E[S] = \mu$ and the relationship $E[R] = E[E[R|S]]$ we obtain $E[R] = \mu$.
- Since $f_{R|S}$ is a Gaussian pdf with unit variance, we know that the conditional variance $Var(R|S) = 1$. Thus, we easily conclude that $E[Var(R|S)] = 1$. We know from part (a) that $E[R|S] = S$, and thus $Var(E[R|S]) = Var(S) = \sigma^2$. Now, we merely apply Proposition 4.2: $Var(R) = E[Var(R|S)] + Var(E[R|S]) = 1 + \sigma^2$.
- In the following sequence of equalities, equation (4) is just the theorem of total probability. Equation (5) is obtained by merely substituting the appropriate densities (note that we aren't really concerned with the value for the constant C). Equation (6) is obtained by expanding the squares and combining the exponentials. Equation (7) is obtained by factoring out of the integral those terms that do not depend on s , and doing a bit of algebra to isolate the s^2 term in the exponent. In equation (8) we merely add and subtract a term that allows us to complete

the square in (9). Finally, (10) is obtained once we realize that there is a Gaussian pdf hiding in the integrand (of course we need to use a bit of the constant C to make this true), and thus, the integral is equal to one.

$$f_R(r) = \int f_{R|S}(r|s)f_S(s)ds \quad (4)$$

$$= C \int e^{-\frac{(r-s)^2}{2}} e^{-\frac{(s-\mu)^2}{2\sigma^2}} ds \quad (5)$$

$$= C \int \exp\left[-\frac{\sigma^2 r^2 - 2\sigma^2 r s + \sigma^2 s^2 + s^2 - 2s\mu + \mu^2}{2\sigma^2}\right] ds \quad (6)$$

$$= C \exp\left[-\frac{\sigma^2 r^2 + \mu^2}{2\sigma^2}\right] \int \exp\left[-\frac{s^2 - 2\frac{\mu + \sigma^2 r}{1 + \sigma^2} s}{\frac{2\sigma^2}{1 + \sigma^2}}\right] ds \quad (7)$$

$$= C \exp\left[-\frac{\sigma^2 r^2 + \mu^2}{2\sigma^2}\right] \int \exp\left[-\frac{s^2 - 2\frac{\mu + \sigma^2 r}{1 + \sigma^2} s + \left(\frac{\mu + \sigma^2 r}{1 + \sigma^2}\right)^2 - \left(\frac{\mu + \sigma^2 r}{1 + \sigma^2}\right)^2}{\frac{2\sigma^2}{1 + \sigma^2}}\right] ds \quad (8)$$

$$= C \exp\left[-\frac{\sigma^2 r^2 + \mu^2}{2\sigma^2}\right] \exp\left[-\frac{\left(\frac{\mu + \sigma^2 r}{1 + \sigma^2}\right)^2}{\frac{2\sigma^2}{1 + \sigma^2}}\right] \int \exp\left[-\frac{\left(s^2 + \frac{\mu + \sigma^2 r}{1 + \sigma^2}\right)^2}{\frac{2\sigma^2}{1 + \sigma^2}}\right] ds \quad (9)$$

$$= K \exp\left[-\frac{\sigma^2 r^2 + \mu^2}{2\sigma^2} - \frac{(\mu + \sigma^2 r)^2}{2(1 + \sigma^2)\sigma^2}\right] \quad (10)$$

At this point, we see that $f_R(r) = K e^{-Q(r)}$ in which $Q(r)$ is quadratic in r . Thus, f_R is Gaussian.

(d) Since $Cov(R, S) = E[RS] - E[R]E[S]$, we first compute $E[RS]$

$$E[RS] = E[E[RS|S]] = E[S \times E[R|S]] = E[S \times S] = E[S^2] = \mu^2 + \sigma^2$$

In this sequence of equalities, the first equality holds because in general $E[X] = E[X|Y]$. The second equality holds because, given S , we can treat the S to the left of the conditioning bar as a constant in the expression $E[RS|S]$. The third equality holds because (as we saw above) $E[R|S] = S$. Finally, recall that in general $Var(X) = E[X^2] - (E[X])^2$.