The Bivariate Normal Distribution

This is Section 4.7 of the 1st edition (2002) of the book Introduction to Probability, by D. P. Bertsekas and J. N. Tsitsiklis. The material in this section was not included in the 2nd edition (2008).

Let U and V be two independent normal random variables, and consider two new random variables X and Y of the form

$$X = aU + bV,$$
$$Y = cU + dV,$$

where a, b, c, d, are some scalars. Each one of the random variables X and Y is normal, since it is a linear function of independent normal random variables.† Furthermore, because X and Y are linear functions of the same two independent normal random variables, their joint PDF takes a special form, known as the **bivariate normal** PDF. The bivariate normal PDF has several useful and elegant properties and, for this reason, it is a commonly employed model. In this section, we derive many such properties, both qualitative and analytical, culminating in a closed-form expression for the joint PDF. To keep the discussion simple, we restrict ourselves to the case where X and Y have zero mean.

Jointly Normal Random Variables

Two random variables X and Y are said to be **jointly normal** if they can be expressed in the form

$$X = aU + bV,$$
$$Y = cU + dV,$$

where U and V are independent normal random variables.

Note that if X and Y are jointly normal, then any linear combination

$$Z = s_1X + s_2Y$$

† For the purposes of this section, we adopt the following convention. A random variable which is always equal to a constant will also be called normal, with zero variance, even though it does not have a PDF. With this convention, the family of normal random variables is closed under linear operations. That is, if X is normal, then $aX + b$ is also normal, even if $a = 0$.

has a normal distribution. The reason is that if we have $X = aU + bV$ and $Y = cU + dV$ for some independent normal random variables U and V, then

$$Z = s_1(aU + bV) + s_2(cU + dV) = (as_1 + cs_2)U + (bs_1 + ds_2)V.$$

Thus, Z is the sum of the independent normal random variables $(as_1 + cs_2)U$ and $(bs_1 + ds_2)V$, and is therefore normal.

A very important property of jointly normal random variables, and which will be the starting point for our development, is that zero correlation implies independence.

Zero Correlation Implies Independence

If two random variables X and Y are jointly normal and are uncorrelated, then they are independent.

This property can be verified using multivariate transforms, as follows. Suppose that U and V are independent zero-mean normal random variables, and that $X = aU + bV$ and $Y = cU + dV$, so that X and Y are jointly normal. We assume that X and Y are uncorrelated, and we wish to show that they are independent. Our first step is to derive a formula for the multivariate transform $M_{X,Y}(s_1, s_2)$ associated with X and Y. Recall that if Z is a zero-mean normal random variable with variance σ_Z^2, the associated transform is

$$\mathbf{E}[e^{sZ}] = M_Z(s) = e^{\sigma_Z^2 s^2/2},$$

which implies that

$$\mathbf{E}[e^{Z}] = M_Z(1) = e^{\sigma_Z^2/2}.$$

Let us fix some scalars s_1, s_2, and let $Z = s_1X + s_2Y$. The random variable Z is normal, by our earlier discussion, with variance

$$\sigma_Z^2 = s_1^2\sigma_X^2 + s_2^2\sigma_Y^2.$$

This leads to the following formula for the multivariate transform associated with the uncorrelated pair X and Y:

$$\begin{aligned} M_{X,Y}(s_1, s_2) &= \mathbf{E}\left[e^{s_1X+s_2Y}\right] \\ &= \mathbf{E}[e^{Z}] \\ &= e^{(s_1^2\sigma_X^2+s_2^2\sigma_Y^2)/2}. \end{aligned}$$

Let now $\overline{X}$ and $\overline{Y}$ be *independent* zero-mean normal random variables with the same variances σ_X^2 and σ_Y^2 as X and Y, respectively. Since $\overline{X}$ and $\overline{Y}$ are independent, they are also uncorrelated, and the preceding argument yields

$$M_{\overline{X},\overline{Y}}(s_1, s_2) = e^{(s_1^2\sigma_X^2+s_2^2\sigma_Y^2)/2}.$$

Thus, the two pairs of random variables (X, Y) and $(\overline{X}, \overline{Y})$ are associated with the same multivariate transform. Since the multivariate transform completely determines the joint PDF, it follows that the pair (X, Y) has the same joint PDF as the pair $(\overline{X}, \overline{Y})$. Since $\overline{X}$ and $\overline{Y}$ are independent, X and Y must also be independent, which establishes our claim.

The Conditional Distribution of X Given Y

We now turn to the problem of estimating X given the value of Y. To avoid uninteresting degenerate cases, we assume that both X and Y have positive variance. Let us define†

$$\hat{X} = \rho \frac{\sigma_X}{\sigma_Y} Y, \qquad \tilde{X} = X - \hat{X},$$

where

$$\rho = \frac{\mathbf{E}[XY]}{\sigma_X \sigma_Y}$$

is the correlation coefficient of X and Y. Since X and Y are linear combinations of independent normal random variables U and V, it follows that Y and $\tilde{X}$ are also linear combinations of U and V. In particular, Y and $\tilde{X}$ are jointly normal. Furthermore,

$$\mathbf{E}[Y\tilde{X}] = \mathbf{E}[YX] - \mathbf{E}[Y\hat{X}] = \rho\sigma_X\sigma_Y - \rho\frac{\sigma_X}{\sigma_Y}\sigma_Y^2 = 0.$$

Thus, Y and $\tilde{X}$ are uncorrelated and, therefore, independent. Since $\hat{X}$ is a scalar multiple of Y, it follows that $\hat{X}$ and $\tilde{X}$ are independent.

We have so far decomposed X into a sum of two independent normal random variables, namely,

$$X = \hat{X} + \tilde{X} = \rho\frac{\sigma_X}{\sigma_Y}Y + \tilde{X}.$$

We take conditional expectations of both sides, given Y, to obtain

$$\mathbf{E}[X \,|\, Y] = \rho\frac{\sigma_X}{\sigma_Y}\mathbf{E}[Y \,|\, Y] + \mathbf{E}[\tilde{X} \,|\, Y] = \rho\frac{\sigma_X}{\sigma_Y}Y = \hat{X},$$

where we have made use of the independence of Y and $\tilde{X}$ to set $\mathbf{E}[\tilde{X} \,|\, Y] = 0$. We have therefore reached the important conclusion that the conditional expectation $\mathbf{E}[X \,|\, Y]$ is a linear function of the random variable Y.

Using the above decomposition, it is now easy to determine the conditional PDF of X. Given a value of Y, the random variable $\hat{X} = \rho\sigma_X Y/\sigma_Y$ becomes

† Comparing with the formulas in the preceding section, it is seen that $\hat{X}$ is defined to be the linear least squares estimator of X, and $\tilde{X}$ is the corresponding estimation error, although these facts are not needed for the argument that follows.

a known constant, but the normal distribution of the random variable $\tilde{X}$ is unaffected, since $\tilde{X}$ is independent of Y. Therefore, the conditional distribution of X given Y is the same as the unconditional distribution of $\tilde{X}$, shifted by $\hat{X}$. Since $\tilde{X}$ is normal with mean zero and some variance $\sigma^2_{\tilde{X}}$, we conclude that the conditional distribution of X is also normal with mean $\hat{X}$ and the same variance $\sigma^2_{\tilde{X}}$. The variance of $\tilde{X}$ can be found with the following calculation:

$$\begin{aligned}\sigma^2_{\tilde{X}} &= \mathbf{E}\left[\left(X - \rho\frac{\sigma_X}{\sigma_Y}Y\right)^2\right] \\ &= \sigma^2_X - 2\rho\frac{\sigma_X}{\sigma_Y}\rho\sigma_X\sigma_Y + \rho^2\frac{\sigma^2_X}{\sigma^2_Y}\sigma^2_Y \\ &= (1-\rho^2)\sigma^2_X,\end{aligned}$$

where we have made use of the property $\mathbf{E}[XY] = \rho\sigma_X\sigma_Y$.

We summarize our conclusions below. Although our discussion used the zero-mean assumption, these conclusions also hold for the non-zero mean case and we state them with this added generality; see the end-of-chapter problems.

Properties of Jointly Normal Random Variables

Let X and Y be jointly normal random variables.

- X and Y are independent if and only if they are uncorrelated.
- The conditional expectation of X given Y satisfies

  $$\mathbf{E}[X \mid Y] = \mathbf{E}[X] + \rho\frac{\sigma_X}{\sigma_Y}\big(Y - \mathbf{E}[Y]\big).$$

  It is a linear function of Y and has a normal PDF.
- The estimation error $\tilde{X} = X - \mathbf{E}[X \mid Y]$ is zero-mean, normal, and independent of Y, with variance

  $$\sigma^2_{\tilde{X}} = (1-\rho^2)\sigma^2_X.$$

- The conditional distribution of X given Y is normal with mean $\mathbf{E}[X \mid Y]$ and variance $\sigma^2_{\tilde{X}}$.

The Form of the Bivariate Normal PDF

Having determined the parameters of the PDF of $\tilde{X}$ and of the conditional PDF of X, we can give explicit formulas for these PDFs. We keep assuming that

X and Y have zero means and positive variances. Furthermore, to avoid the degenerate where $\tilde{X}$ is identically zero, we assume that $|\rho| < 1$. We have

$$f_{\tilde{X}}(\tilde{x}) = f_{\tilde{X}|Y}(\tilde{x}\,|\,y) = \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}\,\sigma_X} e^{-\tilde{x}^2/2\sigma_{\tilde{X}}^2},$$

and

$$f_{X|Y}(x\,|\,y) = \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}\,\sigma_X} e^{-\left(x - \rho\frac{\sigma_X}{\sigma_Y}y\right)^2/2\sigma_{\tilde{X}}^2},$$

where

$$\sigma_{\tilde{X}}^2 = (1-\rho^2)\sigma_X^2.$$

Using also the formula for the PDF of Y,

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\,\sigma_Y} e^{-y^2/2\sigma_Y^2},$$

and the multiplication rule $f_{X,Y}(x,y) = f_Y(y) f_{X|Y}(x\,|\,y)$, we can obtain the joint PDF of X and Y. This PDF is of the form

$$f_{X,Y}(x,y) = ce^{-q(x,y)},$$

where the normalizing constant is

$$c = \frac{1}{2\pi\sqrt{1-\rho^2}\,\sigma_X\sigma_Y}.$$

The exponent term $q(x,y)$ is a quadratic function of x and y,

$$q(x,y) = \frac{y^2}{2\sigma_Y^2} + \frac{\left(x - \rho\frac{\sigma_X}{\sigma_Y}y\right)^2}{2(1-\rho^2)\sigma_X^2},$$

which after some straightforward algebra simplifies to

$$q(x,y) = \frac{\dfrac{x^2}{\sigma_X^2} - 2\rho\dfrac{xy}{\sigma_X\sigma_Y} + \dfrac{y^2}{\sigma_Y^2}}{2(1-\rho^2)}.$$

An important observation here is that **the joint PDF is completely determined by σ_X, σ_Y, and ρ.**

In the special case where X and Y are uncorrelated ($\rho = 0$), the joint PDF takes the simple form

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y} e^{-\frac{x^2}{2\sigma_X^2} - \frac{y^2}{2\sigma_Y^2}},$$

which is just the product of two independent normal PDFs. We can get some insight into the form of this PDF by considering its contours, i.e., sets of points at which the PDF takes a constant value. These contours are described by an equation of the form

$$\frac{x^2}{\sigma_X^2} + \frac{y^2}{\sigma_Y^2} = \text{constant},$$

and are ellipses whose two axes are horizontal and vertical.

In the more general case where X and Y are dependent, a typical contour is described by

$$\frac{x^2}{\sigma_X^2} - 2\rho\frac{xy}{\sigma_X\sigma_Y} + \frac{y^2}{\sigma_Y^2} = \text{constant},$$

and is again an ellipse, but its axes are no longer horizontal and vertical. Figure 4.11 illustrates the contours for two cases, one in which ρ is positive and one in which ρ is negative.

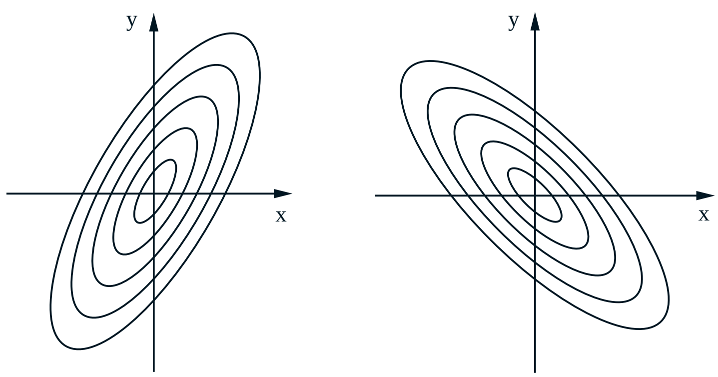

Figure 4.11: Contours of the bivariate normal PDF. The diagram on the left (respectively, right) corresponds to a case of positive (respectively, negative) correlation coefficient ρ.

Example 4.28. Suppose that X and Z are zero-mean jointly normal random variables, such that $\sigma_X^2 = 4$, $\sigma_Z^2 = 17/9$, and $\mathbf{E}[XZ] = 2$. We define a new random variable $Y = 2X - 3Z$. We wish to determine the PDF of Y, the conditional PDF of X given Y, and the joint PDF of X and Y.

As noted earlier, a linear function of two jointly normal random variables is also normal. Thus, Y is normal with variance

$$\sigma_Y^2 = \mathbf{E}\big[(2X - 3Z)^2\big] = 4\mathbf{E}[X^2] + 9\mathbf{E}[Z^2] - 12\mathbf{E}[XZ] = 4 \cdot 4 + 9 \cdot \frac{17}{9} - 12 \cdot 2 = 9.$$

Hence, Y has the normal PDF

$$f_Y(y) = \frac{1}{\sqrt{2\pi} \cdot 3} e^{-y^2/18}.$$

We next note that X and Y are jointly normal. The reason is that X and Z are linear functions of two independent normal random variables (by the definition of joint normality), so that X and Y are also linear functions of the same independent normal random variables. The covariance of X and Y is equal to

$$\begin{aligned}\mathbf{E}[XY] &= \mathbf{E}\big[X(2X-3Z)\big]\\ &= 2\mathbf{E}[X^2] - 3\mathbf{E}[XZ]\\ &= 2\cdot 4 - 3\cdot 2\\ &= 2.\end{aligned}$$

Hence, the correlation coefficient of X and Y, denoted by ρ, is equal to

$$\rho = \frac{\mathbf{E}[XY]}{\sigma_X\sigma_Y} = \frac{2}{2\cdot 3} = \frac{1}{3}.$$

The conditional expectation of X given Y is

$$\mathbf{E}[X\,|\,Y] = \rho\frac{\sigma_X}{\sigma_Y}Y = \frac{1}{3}\cdot\frac{2}{3}Y = \frac{2}{9}Y.$$

The conditional variance of X given Y (which is the same as the variance of $\tilde{X} = X - \mathbf{E}[X\,|\,Y]$) is

$$\sigma_{\tilde{X}}^2 = (1-\rho^2)\sigma_X^2 = \left(1-\frac{1}{9}\right)4 = \frac{32}{9},$$

so that $\sigma_{\tilde{X}} = \sqrt{32}/3$. Hence, the conditional PDF of X given Y is

$$f_{X|Y}(x\,|\,y) = \frac{3}{\sqrt{2\pi}\sqrt{32}}e^{-\frac{\left(x-(2y/9)\right)^2}{2\cdot 32/9}}.$$

Finally, the joint PDF of X and Y is obtained using either the multiplication rule $f_{X,Y}(x,y) = f_X(x)f_{X|Y}(x\,|\,y)$, or by using the earlier developed formula for the exponent $q(x,y)$, and is equal to

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{32}}e^{-\frac{\frac{y^2}{9}+\frac{x^2}{4}-\frac{2}{3}\cdot\frac{xy}{2\cdot 3}}{2(1-(1/9))}}.$$

We end with a cautionary note. If X and Y are jointly normal, then each random variable X and Y is normal. However, the converse is not true. Namely, if each of the random variables X and Y is normal, it does not follow that they are jointly normal, even if they are uncorrelated. This is illustrated in the following example.

Example 4.29. Let X have a normal distribution with zero mean and unit variance. Let Z be independent of X, with $\mathbf{P}(Z=1) = \mathbf{P}(Z=-1) = 1/2$. Let

$Y = ZX$, which is also normal with zero mean. The reason is that conditioned on either value of Z, Y has the same normal distribution, hence its unconditional distribution is also normal. Furthermore,

$$\mathbf{E}[XY] = \mathbf{E}[ZX^2] = \mathbf{E}[Z]\,\mathbf{E}[X^2] = 0 \cdot 1 = 0,$$

so X and Y are uncorrelated. On the other hand X and Y are clearly dependent. (For example, if $X = 1$, then Y must be either -1 or 1.) If X and Y were jointly normal, we would have a contradiction to our earlier conclusion that zero correlation implies independence. It follows that X and Y are *not* jointly normal, even though both marginal distributions are normal.

The Multivariate Normal PDF

The development in this section generalizes to the case of more than two random variables. For example, we can say that the random variables $X_1, \ldots, X_n$ are jointly normal if all of them are linear functions of a set $U_1, \ldots, U_n$ of independent normal random variables. We can then establish the natural extensions of the results derived in this section. For example, it is still true that zero correlation implies independence, that the conditional expectation of one random variable given some of the others is a linear function of the conditioning random variables, and that the conditional PDF of $X_1, \ldots, X_k$ given $X_{k+1}, \ldots, X_n$ is multivariate normal. Finally, there is a closed-form expression for the joint PDF. Assuming that none of the random variables is a deterministic function of the others, we have

$$f_{X_1,\ldots,X_n} = ce^{-q(x_1,\ldots,x_n)},$$

where c is a normalizing constant and where $q(x_1, \ldots, x_n)$ is a quadratic function of $x_1, \ldots, x_n$ that increases to infinity as the magnitude of the vector $(x_1, \ldots, x_n)$ tends to infinity.

Multivariate normal models are very common in statistics, econometrics, signal processing, feedback control, and many other fields. However, a full development falls outside the scope of this text.

Solved Problems on The Bivariate Normal Distribution

Problem 1. Let X_1 and X_2 be independent standard normal random variables. Define the random variables Y_1 and Y_2 by

$$Y_1 = 2X_1 + X_2, \qquad Y_2 = X_1 - X_2.$$

Find $\mathbf{E}[Y_1]$, $\mathbf{E}[Y_2]$, $\mathrm{cov}(Y_1, Y_2)$, and the joint PDF f_{Y_1,Y_2}.

Solution. The means are given by

$$\mathbf{E}[Y_1] = \mathbf{E}[2X_1 + X_2] = \mathbf{E}[2X_1] + \mathbf{E}[X_2] = 0,$$
$$\mathbf{E}[Y_2] = \mathbf{E}[X_1 - X_2] = \mathbf{E}[X_1] - \mathbf{E}[X_2] = 0.$$

The covariance is obtained as follows:

$$\begin{aligned}\text{cov}(Y_1, Y_2) &= \mathbf{E}[Y_1 Y_1] - \mathbf{E}[Y_1]\mathbf{E}[Y_2] \\ &= \mathbf{E}\big[(2X_1 + X_2) \cdot (X_1 - X_2)\big] \\ &= \mathbf{E}\big[2X_1^2 - X_1 X_2 - X_2^2\big] \\ &= 1.\end{aligned}$$

The bivariate normal is determined by the means, the variances, and the correlation coefficient, so we need to calculate the variances. We have

$$\sigma_{Y_1}^2 = \mathbf{E}[Y_1^2] - \mu_{Y_1}^2 = \mathbf{E}[4X_1^2 + 4X_1X_2 + X_2^2] = 5.$$

Similarly,

$$\sigma_{Y_2}^2 = \mathbf{E}[Y_2^2] - \mu_{Y_2}^2 = 5.$$

Thus

$$\rho(Y_1, Y_2) = \frac{\text{cov}(Y_1, Y_2)}{\sigma_{Y_1}\sigma_{Y_2}} = \frac{1}{5}.$$

To write the joint PDF of Y_1 and Y_2, we substitute the above values into the formula for the bivariate normal density function.

Problem 2. The random variables X and Y are described by a joint PDF of the form

$$f_{X,Y}(x, y) = ce^{-8x^2 - 6xy - 18y^2}.$$

Find the means, variances, and the correlation coefficient of X and Y. Also, find the value of the constant c.

Solution. We recognize this as a bivariate normal PDF, with zero means. By comparing $8x^2 + 6xy + 18y^2$ with the exponent

$$q(x, y) = \frac{\dfrac{x^2}{\sigma_X^2} - 2\rho\dfrac{xy}{\sigma_X\sigma_Y} + \dfrac{y^2}{\sigma_Y^2}}{2(1 - \rho^2)}$$

of the bivariate normal, we obtain

$$\sigma_X^2(1 - \rho^2) = 1/4, \qquad \sigma_Y^2(1 - \rho^2) = 1/9, \qquad (1 - \rho^2)\sigma_X\sigma_Y = -\rho/3.$$

From the first two equations, we have

$$(1 - \rho^2)\sigma_X\sigma_Y = 1/6,$$

which implies that $\rho = -1/2$. Thus, $\sigma_X^2 = 1/3$, and $\sigma_Y^2 = 4/27$. Finally,

$$c = \frac{1}{2\pi\sqrt{1 - \rho^2}\,\sigma_X\sigma_Y} = \frac{\sqrt{27}}{\pi}.$$

Problem 3. Suppose that X and Y are independent normal random variables with the same variance. Show that $X - Y$ and $X + Y$ are independent.

Solution. It suffices to show that the zero-mean jointly normal random variables $X - Y - \mathbf{E}[X - Y]$ and $X + Y - \mathbf{E}[X + Y]$ are independent. We can therefore, without loss of generality, assume that X and Y have zero mean. To prove independence, under the zero-mean assumption, it suffices to show that the covariance of $X - Y$ and $X + Y$ is zero. Indeed,

$$\text{cov}(X - Y, X + Y) = \mathbf{E}\big[(X - Y)(X + Y)\big] = \mathbf{E}[X^2] - \mathbf{E}[Y^2] = 0,$$

since X and Y were assumed to have the same variance.

Problem 4. The coordinates X and Y of a point are independent zero-mean normal random variables with common variance σ^2. Given that the point is at a distance of at least c from the origin, find the conditional joint PDF of X and Y.

Solution. Let C denote the event that $X^2 + Y^2 > c^2$. The probability $\mathbf{P}(C)$ can be calculated using polar coordinates, as follows:

$$\begin{aligned}\mathbf{P}(C) &= \frac{1}{2\pi\sigma^2}\int_0^{2\pi}\int_c^{\infty} re^{-r^2/2\sigma^2}\,dr\,d\theta \\ &= \frac{1}{\sigma^2}\int_c^{\infty} re^{-r^2/2\sigma^2}\,dr \\ &= e^{-c^2/2\sigma^2}.\end{aligned}$$

Thus, for $(x, y) \in C$,

$$f_{X,Y|C}(x, y) = \frac{f_{X,Y}(x, y)}{\mathbf{P}(C)} = \frac{1}{2\pi\sigma^2}e^{-\frac{1}{2\sigma^2}(x^2 + y^2 - c^2)}.$$

Problem 5.* Suppose that X and Y are jointly normal random variables. Show that

$$\mathbf{E}[X \mid Y] = \mathbf{E}[X] + \rho\frac{\sigma_X}{\sigma_Y}\big(Y - \mathbf{E}[Y]\big).$$

Hint: Consider the random variables $X - \mathbf{E}[X]$ and $Y - \mathbf{E}[Y]$ and use the result established in the text for the zero-mean case.

Solution. Let $\tilde{X} = X - \mathbf{E}[X]$ and $\tilde{Y} = Y - \mathbf{E}[Y]$. The random variables $\tilde{X}$ and $\tilde{Y}$ are jointly normal. This is because if X and Y are linear functions of two independent normal random variables U and V, then $\tilde{X}$ and $\tilde{Y}$ are also linear functions of U and V. Therefore, as established in the text,

$$\mathbf{E}[\tilde{X} \mid \tilde{Y}] = \rho(\tilde{X}, \tilde{Y})\frac{\sigma_{\tilde{X}}}{\sigma_{\tilde{Y}}}\tilde{Y}.$$

Note that conditioning on $\tilde{Y}$ is the same as conditioning on Y. Therefore,

$$\mathbf{E}[\tilde{X} \mid \tilde{Y}] = \mathbf{E}[\tilde{X} \mid Y] = \mathbf{E}[X \mid Y] - \mathbf{E}[X].$$

Since X and $\tilde{X}$ only differ by a constant, we have $\sigma_{\tilde{X}} = \sigma_X$ and, similarly, $\sigma_{\tilde{Y}} = \sigma_Y$. Finally,

$$\mathrm{cov}(\tilde{X}, \tilde{Y}) = \mathbf{E}[\tilde{X}\tilde{Y}] = \mathbf{E}\Big[\big(X - \mathbf{E}[X]\big)\big(Y - \mathbf{E}[Y]\big)\Big] = \mathrm{cov}(X, Y),$$

from which it follows that $\rho(\tilde{X}, \tilde{Y}) = \rho(X, Y)$. The desired formula follows by substituting the above relations in the formula for $\mathbf{E}[\tilde{X} \,|\, \tilde{Y}]$.

Problem 6.*

(a) Let $X_1, X_2, \ldots, X_n$ be independent identically distributed random variables and let $Y = X_1 + X_2 + \cdots + X_n$. Show that

$$\mathbf{E}[X_1 \,|\, Y] = \frac{Y}{n}.$$

(b) Let X and W be independent zero-mean normal random variables, with positive integer variances k and m, respectively. Use the result of part (a) to find $\mathbf{E}[X \,|\, X + W]$, and verify that this agrees with the conditional expectation formula for jointly normal random variables given in the text. *Hint:* Think of X and W as sums of independent random variables.

Solution. (a) By symmetry, we see that $\mathbf{E}[X_i \,|\, Y]$ is the same for all i. Furthermore,

$$\mathbf{E}[X_1 + \cdots + X_n \,|\, Y] = \mathbf{E}[Y \,|\, Y] = Y.$$

Therefore, $\mathbf{E}[X_1 \,|\, Y] = Y/n$.

(b) We can think of X and W as sums of independent standard normal random variables:

$$X = X_1 + \cdots + X_k, \qquad W = W_1 + \cdots + W_m.$$

We identify Y with $X + W$ and use the result from part (a), to obtain

$$\mathbf{E}[X_i \,|\, X + W] = \frac{X + W}{k + m}.$$

Thus,

$$\mathbf{E}[X \,|\, X + W] = \mathbf{E}[X_1 + \cdots + X_k \,|\, X + W] = \frac{k}{k + m}(X + W).$$

This formula agrees with the formula derived in the text because

$$\rho(X, X + W)\frac{\sigma_X}{\sigma_{X+W}} = \frac{\mathrm{cov}(X, X + W)}{\sigma_{X+W}^2} = \frac{k}{k + m}.$$

We have used here the property

$$\mathrm{cov}(X, X + W) = \mathbf{E}\big[X(X + W)\big] = \mathbf{E}[X^2] = k.$$