Chapter 3. Expectations

3.1. Introduction

In this chapter, we define five mathematical expectations—the mean, variance, standard deviation, covariance and correlation coefficient. We apply these general formula to an array of situations involving discrete and continuous random variables obeying single and joint probability distribution functions to evaluate expectations of both random variables and functions of random variables. Ideally, the reader observes the common analogy in the application of the five concepts expressed in a variety of different ways.

3.2. Mean of a Random Variable

The "mean" is another name for the "average". A third synonym for mean is the "expected value". Let x be a random variable with probability distribution $f(x)$. The mean of x is

$$\mu_x = E(x) = \sum_x x f(x) \tag{3.1.a}$$

if x is a discrete random variable, and

$$\mu_x = E(x) = \int_{-\infty}^{\infty} x f(x) dx \tag{3.1.b}$$

if x is a continuous random variable. We observe that this expectation weights each value of x by its corresponding probability.

Example 3.1: You all may be pretty upset that I suggest that the complicated formulae above give the average when you have been taught since elementary school that the average is given by:

$$\mu_x = \frac{1}{n} \sum_{i=1}^{n} x_i$$

where you just sum up all the elements in the set and divide by the number of elements. Well, let me reassure you that what you learned in elementary school is not wrong. This formula above is one case of equation (3.1) where the probability distribution, $f(x) = 1/n$. (This is called the discrete uniform probability distribution in the following chapter.) Since, in this case, $f(x)$ is not a function of x, it can be pulled out of the summation, giving the familiar result for the mean. However, the uniform distribution is just one of an infinite number of probability distributions. The general formula will apply for any probability distribution.

3.3. Mean of a Function of a Random Variable

Equation 3.1 gives the expected value of the random variable. In general, however, we need the expected value of a function of that random variable. In the general case, our equations which define the mean become:

$$\mu_{h(x)} = E(h(x)) = \sum_x h(x) f(x) \tag{3.2.a}$$

if x is discrete, and

$$\mu_{h(x)} = E(h(x)) = \int_{-\infty}^{\infty} h(x) f(x) dx \tag{3.2.b}$$

if x is continuous, where $h(x)$ is some arbitrary function of x. You should see that equation (3.1) is an example of the case of equation (3.2) where $h(x)=x$. This is one kind of function. However, there is no point in learning a formula for one function, when the formula for all functions is at hand. So, equation (3.2) is the equation to remember.

Example 3.2: In a gambling game, three coins are tossed. A man is paid \$5 when all three coins turn up the same, and he will lose \$3 otherwise. What is the expected gain?

In this problem, the random variable, x, is the number of heads. The distribution function, $f(x)$ is a uniform distribution for the 8 possible outcomes of the gambling game. The function $h(x)$ is the payout or forfeit for outcome x.

outcome	x	$h(x)$	$f(x)$
HHH	3	+\$5	1/8
HHT	2	-\$3	1/8
HTH	2	-\$3	1/8
HTT	1	-\$3	1/8
THH	2	-\$3	1/8
THT	1	-\$3	1/8
TTH	1	-\$3	1/8
TTT	0	+\$5	1/8

We know that the probability distribution $f(x)$ = 1/8, since there are 8 random, equally probable outcomes. Using equation (3.2), we find:

$$\mu_{h(x)} = \sum_{x} h(x) f(x) = \sum_{i=1}^{8} h(x)\left(\frac{1}{8}\right) = \frac{1}{8}\sum_{i=1}^{8} h(x) = \frac{1}{8}(5-3-3-3-3-3-3+5) = -1$$

The average outcome is that the gambler loses a dollar.

We can also work this same problem another way and make the distribution over the number of heads rather than all possible outcomes. In this case, the table looks like:

outcome	x	$h(x)$	$f(x)$
HHH	3	+$5	1/8
HHT,HTH,THH	2	-$3	3/8
TTH, THT, TTH	1	-$3	3/8
TTT	0	+$5	1/8

Then using equation (3.2) again, we find:

$$\mu_{h(x)} = \sum_{x} h(x) f(x) = \sum_{i=1}^{4} h(x) f(x) = \left[(5)\frac{1}{8} + (-3)\frac{3}{8} + (-3)\frac{3}{8} + (5)\frac{1}{8}\right] = -1$$

In fact there are other ways to define the problem. Choose a distribution that makes sense to you. They will all give the same answer so long as your distribution agrees with the physical reality of the problem.

3.4. Mean of a Function of Two Random Variables

Equation (3.2) gives the expected value of a function of one random variable. This equation can be simply extended to a function of two random variables.

$$\mu_{h(x,y)} = E(h(x,y)) = \sum_{x}\sum_{y} h(x,y) f(x,y) \tag{3.3.a}$$

if x and y are discrete random variables, and

$$\mu_{h(x,y)} = E(h(x,y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y) f(x,y)\,dx\,dy \tag{3.3.b}$$

if x and y are continuous random variables. You should see that equation (3.3) is entirely analogous to equation (3.2).

Example 3.3: Given the joint probability density function

$$f(x,y) = \begin{cases} \frac{2}{5}(2x+3y) & \text{for } 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

find the mean of $h(x,y) = y$.

Using equation (3.3.b) we have

$$\mu_{h(x,y)} = E(h(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y) f(x,y)dxdy = \int_0^1\int_0^1 y\left(\frac{2(2x+3y)}{5}\right)dxdy$$

$$\mu_{h(x,y)} = \int_0^1\int_0^1 y\left(\frac{2(2x+3y)}{5}\right)dxdy = \int_0^1 \frac{2}{5}\left(yx^2 + 3y^2x\right)\Big|_{x=0}^{x=1} dy = \int_0^1 \frac{2}{5}\left(y + 3y^2\right)dy$$

$$\mu_{h(x,y)} = \frac{2}{5}\left(\frac{y^2}{2} + y^3\right)\Bigg|_{y=0}^{y=1} = \frac{2}{5}\left(\frac{1}{2}+1\right) = \frac{6}{10}$$

What this says is that the average value of y for this joint probability distribution function is 0.6.

3.5. Variance of a Random Variable

The mean is one parameter of a distribution of data. It gives us some indication of the location of the random variable. It does not however give us any information about the distribution of the random variable. For example, in Figure 3.1., we observe the visually three PDFs with different values of the mean but the same value of the variance. The location of each PDF is different but the spread of the PDF remains the same. In Figure 3.2., we observe three PDFs with different values of the variance but the same value of the mean. Here, the location of the PDF remains constant but the spread of the PDF increases with increasing variance. The **variance** is a statistical measure of this spread.

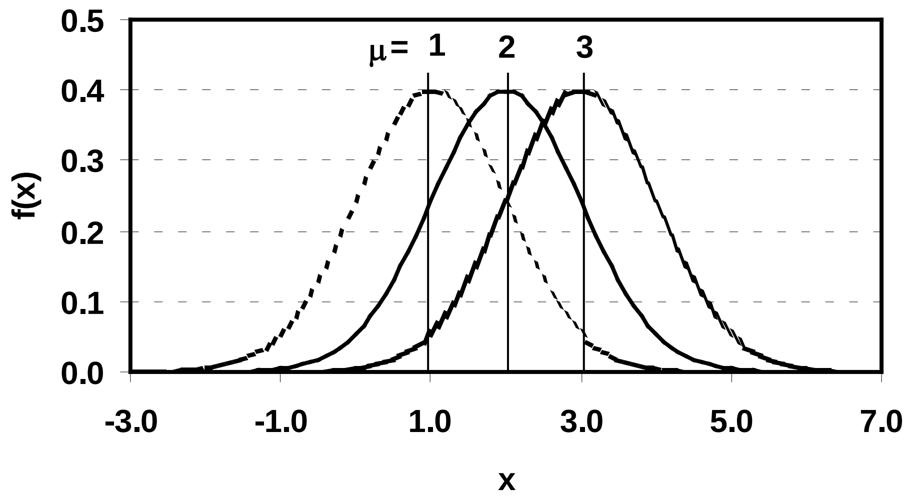

Figure 3.1. Three continuous probability density functions with common variance but different values of the mean.

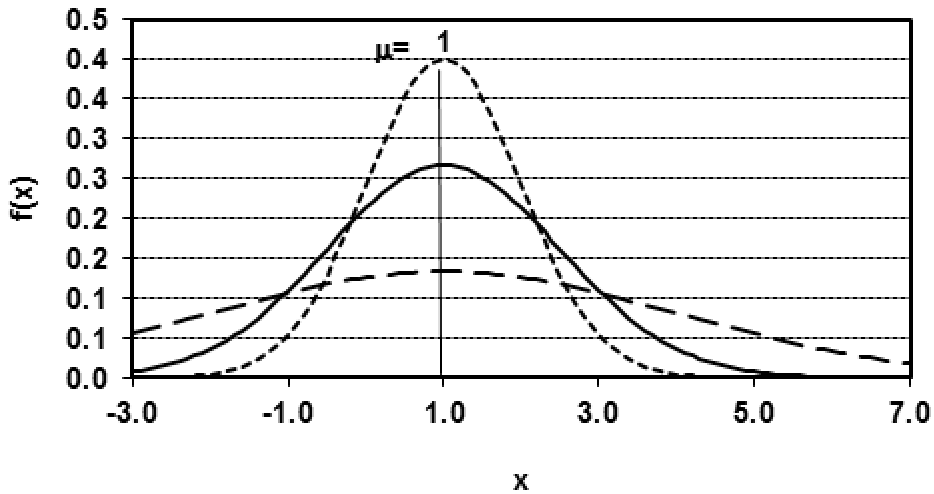

Figure 3.2. Three continuous probability density functions with common mean but different values of the variance.

We define the variance as follows. Let x be a random variable with PDF $f(x)$ and mean μ_x. The **variance** of x, σ_x^2, is defined as

$$\sigma_x^2 = E\left[(x-\mu_x)^2\right] \equiv \sum_x (x-\mu_x)^2 f(x) \tag{3.4.a}$$

if x is a discrete random variable, and

$$\sigma_x^2 = E\left[(x-\mu_x)^2\right] = \int_{-\infty}^{\infty} (x-\mu_x)^2 \, f(x)dx \qquad (3.4.b)$$

if x is a continuous random variable. You should see that equation (3.4) is just another case of equation (3.2) where $h(x) = (x-\mu_x)^2$. What the variance gives is "the average of the square of the deviation from the mean". The square is in there so that all the deviations are positive and the variance is a positive number.

Some tricks with the variance:

The definition for the variance given above, if evaluated properly, will always give the correct value of the variance. However, there is another shortcut formula that is often used. We derive the shortcut here.

$$\begin{aligned}
\sigma_x^2 &= E\left[(x-\mu_x)^2\right] = \int_{-\infty}^{\infty} (x-\mu_x)^2 \, f(x)dx \\
&= \int_{-\infty}^{\infty} x^2 f(x)dx - \int_{-\infty}^{\infty} 2x\mu_x f(x)dx + \int_{-\infty}^{\infty} \mu_x^2 f(x)dx \\
&= \int_{-\infty}^{\infty} x^2 f(x)dx - 2\mu_x \int_{-\infty}^{\infty} xf(x)dx + \mu_x^2 \int_{-\infty}^{\infty} f(x)dx \\
&= E\left[x^2\right] - 2\mu_x E[x] + \mu_x^2 \\
&= E\left[x^2\right] - 2\mu_x^2 + \mu_x^2 \\
&= E\left[x^2\right] - \mu_x^2 = E\left[x^2\right] - E[x]^2
\end{aligned}$$

Thus we have,

$$\sigma_x^2 = E\left[x^2\right] - E[x]^2 \qquad (3.5)$$

People frequently express the variance as "the difference between the mean of the squares and the square of the mean" of the random variable x. They do this because sometimes, you have $E\left[x^2\right]$ and $E[x]$ so the variance is often easier to calculate from equation (3.5) than it is from equation (3.4).

Example 3.4.: Calculate the mean and variance of the discrete data set of 10 numbers containing {1,2,3,4,5,6,7,8,9,10}.

The mean is calculated from equation (3.3.a), where the probability distribution is uniform, i.e., f(x) = 1/n. So μ_x = 5.5. The variance is calculated by squaring each number in the set so that you have a new set of x^2 containing {1,4,9,16,25,36,49,64,81,100}. Then the mean of this set of numbers (using equation (3.3.a)) is μ_{x^2} = 38.5. Now using equation (3.5), we have:

$$\sigma_x^2 = E\left[x^2\right] - E[x]^2 = 38.5 - (5.5)^2 = 8.25$$

The variance is always positive. If you don't get a positive answer using this formula then you have most certainly done something wrong.

Warning on using equation (3.5), the shortcut for the variance

The equation 3.5 may look very friendly but it comes with dangers. You often use this equation to obtain a small variance from the difference of two large numbers. Therefore, the answer you obtain may contain round-off errors. You need to keep all your insignificant figures in the averages in order to obtain the variance to the same number of significant figures.

Example 3.5.: Use $\sigma_x^2 = E\left[x^2\right] - E[x]^2$ to obtain the variance of the following 10 numbers, using $f(x)$ = 1/10.

```
s = [9.92740197834152
  10.06375116530286
   9.98603320980938
  10.07806434475388
  10.04698164235319
  10.03746471832082
   9.96922239354823
   9.93320694775544
   9.93112251526341
   9.93822326228399]
```

Below we give a table that reports the means and variance when keeping different number of significant figures:

4 sig figs: μ_x = 9.990000000e+000 μ_{x^2} = 9.983000000e+001 σ_x^2 = 2.990000000e-002

5 sig figs: μ_x = 9.991000000e+000 μ_{x^2} = 9.982600000e+001 σ_x^2 = 5.919000000e-003

6 sig figs: μ_x = 9.991100000e+000 μ_{x^2} = 9.982630000e+001 σ_x^2 = 4.220790000e-003

7 sig figs: μ_x = 9.991150000e+000 μ_{x^2} = 9.982626000e+001 σ_x^2 = 3.181677500e-003

8 sig figs: μ_x = 9.991147000e+000 μ_{x^2} = 9.982626500e+001 σ_x^2 = 3.246624391e-003

9 sig figs: μ_x = 9.991147200e+000 μ_{x^2} = 9.982626470e+001 σ_x^2 = 3.242327932e-003

10 sig figs: μ_x = 9.991147220e+000 μ_{x^2} = 9.982626473e+001 σ_x^2 = 3.241958286e-003

all sig figs: μ_x = 9.991147218e+000 μ_{x^2} = 9.982626473e+001 σ_x^2 = 3.242000540e-003

You can see that when we only keep 4 significant figures, our calculated variance is off by 822%! You need to keep additional significant figures in the mean and the mean of the squares in order to get the variance with any accuracy.

For your information, the Matlab code that I used to generate the data is provided in the appendix of this chapter.

3.6. Standard Deviation

The standard deviation, σ_x, is the positive square root of the variance, σ_x^2.

Example 3.6.: Calculate the standard deviation of the discrete data set of 10 numbers containing {1,2,3,4,5,6,7,8,9,10}. We calculated the variance in the example 3.4. above. The standard deviation is the square root of the variance

$$\sigma_x = \sqrt{\sigma_x^2} = \sqrt{8.25} = 2.872281323$$

The standard deviation gives us a number *in the same units as the random variable x*, which describes the spread of the data.

3.7. Variance of a Function of a Random Variable

Let *x* be a random variable with PDF *f*(*x*). Let *g*(*x*) be an arbitrary function of *x*. We know that the mean of *g*(*x*) is, and mean $\mu_{g(x)}$ from equation (3.2). The **variance** of a function *g*(*x*), is

$$\sigma_{g(x)}^2 = E\left[\left(g(x) - \mu_{g(x)}\right)^2\right] = \sum_x \left(g(x) - \mu_{g(x)}\right)^2 f(x) \qquad (3.6.a)$$

if x is a discrete random variable and

$$\sigma_{g(x)}^2 = E\left[\left(g(x) - \mu_{g(x)}\right)^2\right] = \int_{-\infty}^{\infty} (g(x) - \mu_{g(x)})^2 f(x)dx \qquad (3.6.b)$$

if x is a continuous random variable. You should see that equation (3.6) is just another case of equation (3.2) where the function of the random variable is $h(x) = \left(g(x) - \mu_{g(x)}\right)^2$. As in the case where the function was $h(x) = (x - \mu_x)^2$, (in equation (3.4)), equation (3.6) can also be reduced to a second form:

$$\sigma^2_{g(x)} = E\left[g(x)^2\right] - E\left[g(x)\right]^2 \tag{3.7}$$

Beware: we have defined 3 functions, $f(x)$, $g(x)$, and $h(x)$. $f(x)$ is the probability distribution. $g(x)$ is the arbitrary function of the random variable that we would like to know about. $h(x)$ is the function with a mean that provides the variance of its argument. In other words, if $h(x) = \left(g(x) - \mu_{g(x)}\right)^2$ then $\sigma^2_{g(x)} = \mu_{h(x)}$.

Example 3.7.: Given the joint probability density function

$$f(x) = \begin{cases} \dfrac{x^2}{3} & \text{for -1} < \text{x} < 2 \\ 0 & \text{otherwise} \end{cases}$$

find the variance of $g(x) = x^{-1}$.

We will use equation (3.7). To do so we must find $E\left[g(x)^2\right]$ and $E\left[g(x)\right]$. $E\left[g(x)\right]$ is the mean of $g(x)$ and can be calculated from the formula for the mean, equation (3.2).

$$\mu_{g(x)} = E(g(x)) = \int_{-\infty}^{\infty} g(x) f(x) dx \tag{3.2.b}$$

$$E(x^{-1}) = \int_{-1}^{2} x^{-1}\left(\frac{x^2}{3}\right)dx = \left.\frac{x^2}{6}\right|_{x=-1}^{x=2} = \frac{2^2}{6} - \frac{(-1)^2}{6} = \frac{1}{2}$$

Now we repeat the calculation for the square of $g(x)$

$$E\left[\left(x^{-1}\right)^2\right] = E\left[x^{-2}\right] = \int_{-1}^{2} x^{-2}\left(\frac{x^2}{3}\right)dx = \left.\frac{x}{3}\right|_{x=-1}^{x=2} = \frac{2}{3} - \frac{(-1)}{3} = 1$$

Then we substitute into equation (3.7)

$$\sigma^2_{g(x)} = E\left[g(x)^2\right] - E\left[g(x)\right]^2 = 1 - \left(\frac{1}{2}\right)^2 = \frac{3}{4}$$

3.8. Variance & Covariance of a Function of Two Random Variables

By analogous methods, we can extend the variance definition to a function of two variables.

$$\sigma^2_{g(x,y)} = E\left[\left(g(x,y)-\mu_{g(x,y)}\right)^2\right] = \sum_x \sum_y \left(g(x,y)-\mu_{g(x,y)}\right)^2 f(x,y) \quad (3.8.\ a)$$

if *x* and *y* are discrete random variables and

$$\sigma^2_{g(x,y)} = E\left[\left(g(x,y)-\mu_{g(x,y)}\right)^2\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(g(x,y)-\mu_{g(x,y)}\right)^2 f(x,y)\,dx\,dy \quad (3.8.b)$$

if *x* and *y* are continuous random variables. You should see that equation (3.8) is just another case of equation (3.3) where the function of the random variable, let's call it $h(x,y) = \left(g(x,y)-\mu_{g(x,y)}\right)^2$. Again, equation (3.8) can be rewritten

$$\sigma^2_{g(x,y)} = E\left[g(x,y)^2\right] - E\left[g(x,y)\right]^2 = \mu_{h(x,y)} \quad (3.9)$$

Now, let's think about equation 3.8. If $g(x,y)=x$, then $h(x,y)=(x-\mu_x)^2$ and we calculate the variance of *x* from equation (3.8). If $g(x,y)=y$, then $h(x,y)=(y-\mu_y)^2$ and we have the variance of *y* from equation (3.8).

Now, if $h(x,y)=(x-\mu_x)(y-\mu_y)$, then we can use equation (3.8) to calculate the **COVARIANCE**, σ_{xy}. The covariance has the units of xy. There is no function $g(x)$ defined for the covariance, so equation (3.9) does not apply to the calculation of the covariance. But if you substitute $h(x,y)=(x-\mu_x)(y-\mu_y)$ into equation (3.2) and solve as we did to arrive with equation (3.5) you find:

$$\sigma_{XY} = E[xy] - E[x]E[y] = \mu_{XY} - \mu_X \mu_Y \quad (3.10)$$

The qualitative significance of the covariance is the dependency between variables x and y.

σ_{XY}	qualitative significance
$\sigma_{XY} > 0$	as x increases, y increases
$\sigma_{XY} = 0$	x and y are independent
$\sigma_{XY} < 0$	as x increases, y decreases

Example 3.8.: Given the joint probability density function

$$f(x,y)=\begin{cases}\frac{2}{5}(2x+3y) & \text{for } 0\le x\le 1, 0\le y\le 1\\ 0 & \text{otherwise}\end{cases}$$

find the covariance of x and y.

To find the covariance, we need: $E[xy], E[x]$ and $E[y]$. We already calculated $E[y]$ in example 3.3. and we found $E[y]$=0.6. Using a similar procedure, we calculate, the expected value of *x*, $E[x]$

$$E[x]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)dxdy=\int_0^1\int_0^1 x\left(\frac{2(2x+3y)}{5}\right)dxdy$$

$$E[x]=\int_0^1 \frac{2}{5}\left(\frac{2x^3}{3}+3y\frac{x^2}{2}\right)\Bigg|_{x=0}^{x=1} dy=\int_0^1 \frac{2}{5}\left(\frac{2}{3}+\frac{3y}{2}\right)dy$$

$$E[x]=\frac{2}{5}\left(\frac{2y}{3}+\frac{3y^2}{4}\right)\Bigg|_{y=0}^{y=1}=\frac{2}{5}\left(\frac{2}{3}+\frac{3}{4}\right)=\frac{17}{30}$$

In an analogous fashion, we calculate $E[xy]$

$$E[xy]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy=\int_0^1\int_0^1 xy\left(\frac{2(2x+3y)}{5}\right)dxdy$$

$$E[xy]=\int_0^1 \frac{2y}{5}\left(\frac{2x^3}{3}+3y\frac{x^2}{2}\right)\Bigg|_{x=0}^{x=1} dy=\int_0^1 \frac{2y}{5}\left(\frac{2}{3}+\frac{3y}{2}\right)dy$$

$$E[xy]=\frac{2}{5}\left(\frac{2y^2}{6}+\frac{3y^3}{6}\right)\Bigg|_{y=0}^{y=1}=\frac{2}{5}\left(\frac{2}{6}+\frac{3}{6}\right)=\frac{1}{3}$$

so using equation (3.10), we find:

$$\sigma_{XY} = E[xy] - E[x]E[y] = \frac{1}{3} - \frac{17}{30}\frac{6}{10} = \frac{100-102}{300} = -\frac{1}{150} = -0.006667$$

3.9. Correlation Coefficients

The magnitude of σ_{XY} does not say anything regarding the strength of the relationship between *x* and *y* because σ_{XY}. depends on the values taken by *x* and *y*. A scaled version of the covariance, called the correlation coefficient is much more useful. The correlation coefficient is defined as

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \tag{3.11}$$

This variable ranges from -1 to 1 and is 0 when σ_{XY} is zero. A negative correlation coefficient means that when *y* increases, *x* decreases and vice versa. A positive correlation coefficient means that when *x* increases, *y* also increases, and vice versa for decreasing.

ρ_{XY}	qualitative significance
$\rho_{XY} = 1$	*x* = *y*
$\rho_{XY} > 0$	as *x* increases, *y* increases
$\rho_{XY} = 0$	*x* and *y* are independent
$\rho_{XY} < 0$	as *x* increases, *y* decreases
$\rho_{XY} = -1$	*x* = -*y*

Example 3.9.: Given the joint PDF in example 3.8., find the correlation coefficient and make a statement about whether *x* is strongly or weakly correlated to *y*, relative to the variance of *x* and *y*.

To do this, we need the variance of *x* and *y*, which means we need $E[x^2]$ and $E[y^2]$

$$E[x^2] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2 f(x,y)dxdy = \int_0^1\int_0^1 x^2\left(\frac{2(2x+3y)}{5}\right)dxdy$$

$$E[x^2] = \int_0^1 \frac{2}{5}\left(\frac{x^4}{2} + yx^3\right)\Bigg|_{x=0}^{x=1} dy = \int_0^1 \frac{2}{5}\left(\frac{1}{2} + y\right)dy$$

$$E\left[x^2\right]=\frac{2}{5}\left(\frac{y}{2}+\frac{y^2}{2}\right)\Bigg|_{y=0}^{y=1}=\frac{2}{5}\left(\frac{1}{2}+\frac{1}{2}\right)=\frac{2}{5}$$

so

$$\sigma_X^2=E\left[x^2\right]-E[x]^2=\frac{2}{5}-\left(\frac{17}{30}\right)^2=0.0789$$

Now for *y*:

$$E\left[y^2\right]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(x,y)f(x,y)dxdy=\int_0^1\int_0^1 y^2\left(\frac{2(2x+3y)}{5}\right)dxdy$$

$$E\left[y^2\right]=\int_0^1\int_0^1 y^2\left(\frac{2(2x+3y)}{5}\right)dxdy=\int_0^1\frac{2}{5}\left(y^2x^2+3y^3x\right)\Bigg|_{x=0}^{x=1}dy=\int_0^1\frac{2}{5}\left(y^2+3y^3\right)dy$$

$$E\left[y^2\right]=\frac{2}{5}\left(\frac{y^3}{3}+\frac{3y^4}{4}\right)\Bigg|_{y=0}^{y=1}=\frac{2}{5}\left(\frac{1}{3}+\frac{3}{4}\right)=\frac{13}{30}$$

so

$$\sigma_Y^2=E\left[y^2\right]-E[y]^2=\frac{13}{30}-\left(\frac{6}{10}\right)^2=0.0733$$

so the correlation coefficient is

$$\rho_{XY}=\frac{\sigma_{XY}}{\sigma_X\sigma_Y}=\frac{-0.00667}{\sqrt{0.0789}\sqrt{0.0733}}=-0.0877$$

The small value of the correlation coefficient indicates that the random variables *x* and *y* are not strongly correlated, but they are weakly negatively correlated.

3.10. Means and Variances of linear combinations of Random Variables

These are several rules for means and variances. These rules have their basis in the theory of linear operators. A linear operator L[x] performs some operation on x, such that:

$$L[ax+by]=aL[x]+bL[y] \tag{3.12}$$

where *x* and *y* are variables and *a* and *b* are constants. This is the fundamental rule which all linear operators must follow.

Consider the differential operator: $L[x] = \frac{d}{dt}[x]$. Is it a linear operator? To prove or disprove the linearity of the differential operator, you must substitute it into equation (3.12) to verify it.

$$\frac{d}{dt}[ax+by] \stackrel{?}{=} a\frac{d}{dt}[x] + b\frac{d}{dt}[y]$$
$$\frac{d}{dt}[ax] + \frac{d}{dt}[by] \stackrel{?}{=} a\frac{d}{dt}[x] + b\frac{d}{dt}[y]$$
$$a\frac{d}{dt}[x] + b\frac{d}{dt}[y] = a\frac{d}{dt}[x] + b\frac{d}{dt}[y] \qquad \text{This is an identity.}$$

So, we have shown that the differential operator is a linear operator. What about the integral operator, $L[x] = \int xdt$?

$$\int [ax+by]dt \stackrel{?}{=} a\int xdt + b\int ydt$$
$$\int axdt + \int bydt \stackrel{?}{=} a\int xdt + b\int ydt$$
$$a\int xdt + b\int ydt \stackrel{?}{=} a\int xdt + b\int ydt \qquad \text{This is an identity.}$$

So, we have shown that the differential operator is a linear operator. What about the square operator, $L[x] = x^2$?

$$[ax+by]^2 \stackrel{?}{=} ax^2 + by^2$$
$$a^2x^2 + 2abxy + by^2 \stackrel{?}{=} ax^2 + by^2$$
$$a(1-a)x^2 + 2abxy + b(1-b)y^2 \stackrel{?}{=} 0$$

we can use the quadratic equation to solve for *x*:

$$x = \frac{-ab \pm \sqrt{a^2b^2 - a(1-a)b(1-b)y^2}}{a(1-a)}$$

For any given value of *y*, the solution to this quadratic formula are the only solutions which satisfy equation (3.12). In order for the operator to be linear, equation (3.12) must be satisfied for all *x*. Therefore, the square operator is not a linear operator.

Now, let's see if the mean is a linear operator (we will do this just for the continuous case, but the result could also be shown for the discrete case):

$$E[ax+by]\overset{?}{=}aE[x]+bE[y]$$

Substitute in the definition of the mean from equation (3.1)

$$\int_{-\infty}^{\infty}[ax+by]f(x)dx\overset{?}{=}a\int_{-\infty}^{\infty}xf(x)dx+b\int_{-\infty}^{\infty}yf(x)dx$$

The integral of a sum is the sum of the integrals. Constants can be pulled outside the integral, so

$$a\int_{-\infty}^{\infty}xf(x)dx+b\int_{-\infty}^{\infty}yf(x)dx=a\int_{-\infty}^{\infty}xf(x)dx+b\int_{-\infty}^{\infty}yf(x)dx$$

This is an identity. The mean is a linear operator. As a result, we have a few simplifications for the mean. In the equations below, we assume that a and b are constants.

The mean of a constant is the constant.

$$E(a)=a$$

Adding a constant to a random variable adds the same constant to the mean.

$$E(ax+b)=aE(x)+b$$

The mean of the sum is the sum of the means.

$$E(g(x)+h(x))=E(g(x))+E(h(x))$$

These rules also apply to joint PDFs.

$$E(g(x,y)\pm h(x,y))=E(g(x,y))+E(h(x,y))$$

If and only if x and y are **independent** random variables, then

$$E(xy)=E(x)E(y) \qquad \text{(only for independent } x \text{ and } y\text{)}$$

We can show that the variance is **not** a linear operator. However, by substituting in for the definition of the variance, equation (3.4), we can come up with several short-cuts for computing some variances of functions. Again, we assume that *a* and *b* are constants.

The variance of a constant is zero.

$$\sigma_b^2 = 0$$

Adding a constant to a random variable does not change the variance of the random variable.

$$\sigma_{ax+b}^2 = a^2\sigma_x^2$$

The variance of the sum is not the sum of the variances.

$$\sigma_{ax+by}^2 = a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\sigma_{xy}$$

If and only if *x* and *y* are **independent**, then

$$\sigma_{ax+by}^2 = a^2\sigma_x^2 + b^2\sigma_y^2 \qquad \text{(only for independent } x \text{ and } y\text{)}$$

We did not just make any of these theorems up. They can all be derived. As an example we now derive,

$$\sigma_{ax+by}^2 = a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\sigma_{xy}$$

We begin by direct substitution of (*ax*+*by*) into the definition of the variance:

$$\sigma_{g(x,y)}^2 \equiv E\left[\left(g(x,y) - \mu_{g(x,y)}\right)^2\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(g(x,y) - \mu_{g(x,y)})^2 f(x,y)dxdy$$

$$\sigma_{g(x,y)}^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(ax + by - \mu_{ax+by})^2 f(x,y)dxdy$$

$$\sigma_{g(x,y)}^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(a^2x^2 + b^2y^2 + 2abxy - 2ax\mu_{ax+by} - 2by\mu_{ax+by} + \mu_{ax+by}{}^2) f(x,y)dxdy$$

$$\sigma_{g(x,y)}^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}a^2x^2 f(x,y)dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}b^2y^2 f(x,y)dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}2abxyf(x,y)dxdy$$

$$-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}2ax\mu_{ax+by} f(x,y)dxdy - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}2by\mu_{ax+by} f(x,y)dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mu_{ax+by}{}^2 f(x,y)dxdy$$

$$\sigma^2_{g(x,y)} = a^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2 f(x,y)dxdy + b^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y^2 f(x,y)dxdy + 2ab\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy$$
$$-2a\mu_{ax+by}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)dxdy - 2b\mu_{ax+by}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x,y)dxdy + \mu_{ax+by}{}^2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy$$
$$\sigma^2_{g(x,y)} = a^2\mu_{x^2} + b^2\mu_{y^2} + 2ab\mu_{xy} - 2a\mu_{ax+by}\mu_x - 2b\mu_{ax+by}\mu_y + \mu_{ax+by}{}^2$$
$$\mu_{ax+by} = a\mu_x + b\mu_y$$
$$\sigma^2_{g(x,y)} = a^2\mu_{x^2} + b^2\mu_{y^2} + 2ab\mu_{xy} - 2a^2\mu_x{}^2 - 2ab\mu_y\mu_x - 2b^2\mu_y{}^2 - 2ab\mu_y\mu_x$$
$$+a^2\mu_x{}^2 + b^2\mu_y{}^2 + 2ab\mu_x\mu_y$$
$$\sigma^2_{g(x,y)} = a^2\left(\mu_{x^2} - \mu_x{}^2\right) + b^2\left(\mu_{y^2} - \mu_y{}^2\right) + 2ab\left(\mu_{xy} - \mu_y\mu_x\right)$$
$$\sigma^2_{g(x,y)} = a^2\sigma_x^2 + b^2\sigma_y^2 + 2ab\sigma_{xy}$$

Q.E.D.

3.11. An Extended Example for Discrete Random Variables

Consider the isomerization reaction:

$$\mathrm{A} \rightarrow \mathrm{B}$$

This reaction takes place in a plant which relies on raw material solution, which unfortunately, is supposed to have a concentration of reactant of 1.0 mol/liter but in reality varies +/- 20%. The reactor is jacketed and is supposed to be isothermal. Day to day observation of the thermocouples in the reactor indicates that temperature swings about 10% around its set point of 300 K.

The reaction rate is given as

$$r_b = kC_A = k_o e^{-\frac{E_a}{RT}} C_A$$

where k is the rate constant, k_o is the pre-exponential factor of the rate constant, E_a is the activation energy, R is the gas constant, T is the temperature, and C_A is the concentration of the reactant. In one such reaction, $k_o = 20\frac{liters}{\min}$, $E_a = 10\frac{kJ}{mol}$, and $R = 8.314\frac{J}{mol \cdot K}$. Over a month, 20 spot measurements are made of the reactor, measuring the concentration of the reactant and the temperature.

Consider that the probability of obtaining any of the data points was uniform. Therefore, $f(x)=\frac{1}{n}$ where n is the number of measurements taken.

The tabulated data and the functions of that data are shown below:

runs	C_A	T	r_B	$C_A^{\ 2}$	T^2		$r_B^{\ 2}$	$C_A \cdot T$	$C_A \cdot r_B$	$T \cdot r_B$
	mol/liter	K	mol/min							
1	1.11	296.49	0.38	1.22	87904.81		0.15	327.92	0.42	113.49
2	1.01	272.80	0.25	1.02	74419.84		0.06	275.60	0.25	67.06
3	1.03	270.22	0.24	1.06	73020.02		0.06	278.42	0.25	64.96
4	0.82	324.55	0.40	0.67	105332.35		0.16	265.94	0.33	130.71
5	0.83	273.87	0.21	0.70	75006.19		0.04	228.67	0.17	56.61
6	1.11	274.20	0.28	1.23	75185.95		0.08	304.31	0.31	75.73
7	0.80	299.56	0.29	0.64	89733.67		0.08	239.93	0.23	86.56
8	0.84	325.13	0.42	0.71	105709.64		0.17	273.64	0.35	135.39
9	0.89	310.19	0.37	0.78	96220.00		0.13	274.75	0.32	113.75
10	1.16	271.78	0.28	1.35	73862.15		0.08	315.23	0.32	75.44
11	1.13	298.13	0.40	1.27	88878.56		0.16	336.10	0.45	118.94
12	1.14	304.56	0.44	1.30	92759.55		0.19	347.10	0.50	133.77
13	1.13	270.54	0.26	1.27	73189.31		0.07	305.20	0.30	71.58
14	1.04	280.21	0.29	1.09	78514.87		0.08	292.34	0.30	79.93
15	1.15	306.72	0.46	1.32	94077.94		0.21	352.47	0.52	139.67
16	0.87	319.16	0.40	0.76	101864.25		0.16	277.89	0.35	128.30
17	0.83	304.60	0.32	0.68	92778.97		0.10	251.76	0.26	97.07
18	1.06	303.42	0.40	1.11	92064.18		0.16	320.26	0.42	121.60
19	0.83	289.58	0.26	0.69	83856.40		0.07	241.11	0.22	75.75
20	0.89	301.14	0.33	0.79	90684.97		0.11	267.54	0.29	98.58
sum	19.66	5896.84	6.66	19.68	1745063.64		2.32	5776.21	6.57	1984.89
mean	0.98	294.84	0.33	0.98	87253.18		0.12	288.81	0.33	99.24
variance	0.02	321.47	0.01			covariance		-1.03	0.00	1.09
standard deviation	0.13	17.93	0.07			correlation		-0.43	0.14	0.83

We use the definition of the mean, $\mu = E(x) = \sum_x x f(x)$, to obtain expectation values for the following functions: C_A, T, r_B, $C_A^{\ 2}$, T^2, $r_B^{\ 2}$, $C_A \cdot T$, $C_A \cdot r_B$ and $T \cdot r_B$. The expectations are shown in the table above in the row marked mean. The variances of C_A, T, and r_B are calculated using the “difference between the mean of the square and the square of the mean” rule.

$$\sigma_{g(x)}^2 = E\left[g(x)^2\right] - E[g(x)]^2$$

Those variances are shown in the first three columns in the row marked variance. The covariances are obtained using the formula:

$$\sigma_{xy} = E[xy] - E[x]E[y]$$

and are shown in the last three columns for $C_A \cdot T$, $C_A \cdot r_B$ and $T \cdot r_B$. The standard deviations and correlation coefficients are given in the bottom row, obtained from:

$$\sigma_{g(x)} = \sqrt{\sigma^2_{g(x)}} \text{ and } \rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}.$$

Physical explanation of statistical results:

The mean and the standard deviation of the concentration show that statistically speaking:

$$C_A = 0.98 \pm 0.13 \frac{mol}{liter}$$

Similarly, $T = 294.5 \pm 17.9K$ and $r_B = 0.33 \pm 0.07 \frac{mol}{\min}$.

The physical meaning of the correlation coefficients are as follows. The concentration of A and the temperature (two independent random variables) should not be correlated. The correlation coefficient should be zero. It is -0.43. This non-zero value is a result of only having 20 data points. More data points would eventually average out to a correlation coefficient of zero.

The $C_A \cdot r_B$ correlation coefficient should be positive because as the concentration increases, the reaction rate increases. It is positive. The $C_A \cdot r_B$ correlation coefficient is small because the relationship is a linear (weak) relationship.

The $T \cdot r_B$ correlation coefficient should be positive because as the temperature increases, the reaction rate increases. It is positive. The $T \cdot r_B$ correlation coefficient is large because the relationship is an exponential (strong) relationship.

3.12. An Extended Example for Continuous Random Variables

A construction company has designed a distribution function which describes the area of their construction sites. The sites are all rectangular with dimensions *a* and *b*. The Joint PDF of the dimensions *a* and *b* are:

$$f(a,b) = \begin{cases} \frac{4}{21} \text{ab} & \text{for } 1 \le \text{a} < 2 \text{ and } 3 \le \text{b} \le 4 \\ 0 & \text{otherwise} \end{cases}$$

The company is interested in determining pre-construction site costs including fencing and clearing land. The amount of fencing gives rise to a perimeter cost. The Perimeter Costs, PC, are \$10 per meter of fencing required:

$$PC(a,b) = 10(2a + 2b)$$

The amount of land cleared is proportional to the area of the site and gives rise to an area cost. The Area Costs, AC, are \$20 per square meter of the lot:

$$AC(a,b) = 20ab$$

(a) Are a and b independent?
(b) Find the mean of a, b, PC, and AC.
(c) Find the variance of a, b, PC, and AC.
(d) Find the covariance of $a \cdot b$, $a \cdot PC$, $a \cdot AC$, $b \cdot PC$, $b \cdot AC$, and $PC \cdot AC$.
(e) Find the correlation coefficient of $a \cdot b$, $a \cdot PC$, $a \cdot AC$, $b \cdot PC$, $b \cdot AC$, and $PC \cdot AC$.

(a) a and b are independent if $f(x,y) = g(x)h(y)$ where the marginal distributions are defined in Chapter 2 as

$$g(x) = \int_{-\infty}^{\infty} f(x,y)dy \quad \text{and} \quad h(y) = \int_{-\infty}^{\infty} f(x,y)dx$$

We evaluate the marginal distributions.

$$g(a) = \int_{3}^{4} f(a,b)db = \int_{3}^{4} \frac{4}{21} abdb = \frac{4}{21} a \frac{b^2}{2}\Bigg|_3^4 = \frac{2}{3}a$$

$$h(b) = \int_{1}^{2} f(a,b)da = \int_{1}^{2} \frac{4}{21} abda = \frac{4}{21} b \frac{a^2}{2}\Bigg|_1^2 = \frac{2}{7}b$$

$$f(x,y) = \frac{4}{21}ab = g(x)h(y) = \left(\frac{2}{3}a\right)\left(\frac{2}{7}b\right) = \frac{4}{21}ab$$

Therefore, a and b are independent.

(b) Find the mean of a, b, PC, and AC.

The general formula for the mean is:

$$\mu_{h(x,y)} = E(h(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)f(x,y)dxdy$$

$$\mu_a = E(a) = \int_1^2\int_3^4 a\left(\frac{4}{21}ab\right)dbda = \int_1^2\left(\frac{4}{21}a^2\frac{b^2}{2}\right)_3^4 da = \left(\frac{4}{21}\frac{a^3}{3}\frac{7}{2}\right)_1^2 = \frac{98}{63} = 1.56\,\text{m}$$

$$\mu_b = E(b) = \int_1^2\int_3^4 b\left(\frac{4}{21}ab\right)dbda = \int_1^2\left(\frac{4}{21}a\frac{b^3}{3}\right)_3^4 da = \left(\frac{4}{21}\frac{a^2}{2}\frac{37}{3}\right)_1^2 = \frac{74}{21} = 3.52\,\text{m}$$

We can use the definition of the mean to compute the mean value of the perimeter costs.

$$\mu_{PC} = E(PC) = \int_1^2\int_3^4 20(a+b)\left(\frac{4}{21}ab\right)dbda = \frac{80}{21}\int_1^2\left(a^2\frac{b^2}{2} + a\frac{b^3}{3}\right)_3^4 da =$$

$$= \frac{80}{21}\left(\frac{a^3}{3}\frac{7}{2} + \frac{a^2}{2}\frac{37}{3}\right)_1^2 = \frac{80}{21}\left(\frac{49}{6} + \frac{111}{6}\right) = \frac{6400}{63} = \$101.59$$

OR remember that the mean is a linear operator, $E(ax+by) = aE(x) + bE(y)$

$$\mu_{PC} = E(PC) = E(20a) + E(20b) = 20E(a) + 20E(b) = 20(1.56) + 20(3.52) = \$101.6$$

For the area cost, we can use the definition of the mean.

$$\mu_{AC} = E(AC) = \int_1^2\int_3^4 20(ab)\left(\frac{4}{21}ab\right)dbda = \frac{80}{21}\int_1^2\left(a^2\frac{b^3}{3}\right)_3^4 da = \frac{80}{21}\left(\frac{a^3}{3}\frac{37}{3}\right)_1^2 = \frac{20720}{189} = \$109.63$$

OR remember $E(xy) = E(x)E(y)$ for independent variables.

$$\mu_{AC} = E(AC) = E(20ab) \neq 20E(a)E(b) = 20(1.56)(3.52) = \$109.63$$

because a and b are independent.

(c) Find the variance of a, b, PC, and AC.

The working equation to calculate the variance of a function is:

$$\sigma^2_{g(x,y)} = E\left[g(x,y)^2\right] - E\left[g(x,y)\right]^2$$

For these variables, we have calculated the mean (necessary to evaluate the function in the second term on the right hand side). We must next calculate the mean of the square (the first term on the right hand side) before we can calculate the variance.

$$\mu_{a^2} = E(a^2) = \int_1^2 \int_3^4 a^2 \left(\frac{4}{21} ab \right) dbda = \int_1^2 \left(\frac{4}{21} a^3 \frac{b^2}{2} \right)_3^4 da = \left(\frac{4}{21} \frac{a^4}{4} \frac{7}{2} \right)_1^2 = \frac{420}{168} = 2.50$$

$$\sigma_a^2 = E\left[a^2\right] - E[a]^2 = 2.50 - 1.5556^2 = 0.0802$$

$$\mu_{b^2} = E(b^2) = \int_1^2 \int_3^4 b^2 \left(\frac{4}{21} ab \right) dbda = \int_1^2 \left(\frac{4}{21} a \frac{b^4}{4} \right)_3^4 da = \left(\frac{4}{21} \frac{a^2}{2} \frac{175}{4} \right)_1^2 = \frac{525}{42} = 12.50$$

$$\sigma_b^2 = E\left[b^2\right] - E[b]^2 = 12.50 - 3.5238^2 = 0.0828$$

To calculate the variance of $PC(a,b) = 10(2a + 2b)$ remember

$$\sigma_{aX+bY}^2 = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\sigma_{XY}$$

Then we only need to calculate the covariance of *a* and *b*. The working formula for the covariance is:

$$\sigma_{XY} = E[xy] - E[x]E[y]$$

So we need the expectation value of $E(ab)$

$$\mu_{ab} = E(ab) = \int_1^2 \int_3^4 (ab) \left(\frac{4}{21} ab \right) dbda = \frac{4}{21} \int_1^2 \left(a^2 \frac{b^3}{3} \right)_3^4 da = \frac{4}{21} \left(\frac{a^3}{3} \frac{37}{3} \right)_1^2 = \frac{1036}{189} = 5.48$$

Then

$$\sigma_{ab} = 5.48 - 1.56(3.52) = 0.00$$

The covariance of *a* and *b* is zero. We should have known that because we showed in part (a) that *a* and *b* were statistically independent.

$$\sigma_{PC}^2 = \sigma_{20(a+b)=}^2 20^2\left(\sigma_a^2 + \sigma_b^2\right) + 2(20)(20)\sigma_{ab} = 65.2.$$

Lastly, for the area cost,

$$\mu_{AC^2} = E(AC^2) = \int_1^2\int_3^4 20^2(ab)^2\left(\frac{4}{21}ab\right)dbda = \frac{1600}{21}\int_1^2\left(a^3\frac{b^4}{4}\right)_3^4 da =$$

$$= \frac{1600}{21}\left(\frac{a^4}{4}\frac{175}{4}\right)_1^2 = \frac{4200000}{336} = 12500.0$$

$$\sigma_{AC}^2 = E\left[AC^2\right] - E\left[AC\right]^2 = 12500.0 - 109.63^2 = 481.$$

(d) Find the covariance of $a \cdot b$, $a \cdot PC$, $a \cdot AC$, $b \cdot PC$, $b \cdot AC$, and $PC \cdot AC$.

In part (c) we found the covariance of $a \cdot b$ to be 0.0 because they were statistically independent. For the rest of these quantities, we use the rule

$$\sigma_{XY} = E[xy] - E[x]E[y]$$

where we already have the expectation values of the two factors in the second term on the r.h.s. We only need to find the first term on the r.h.s. to find the covariance. For the covariance between *a* and the perimeter cost we have

$$\mu_{aPC} = E(aPC) = \int_1^2\int_3^4 20(a^2+ab)\left(\frac{4}{21}ab\right)dbda = \frac{80}{21}\int_1^2\left(a^3\frac{b^2}{2} + a^2\frac{b^3}{3}\right)_3^4 da =$$

$$= \frac{80}{21}\left(\frac{a^4}{4}\frac{7}{2} + \frac{a^3}{3}\frac{37}{3}\right)_1^2 = \frac{80}{21}\left(\frac{105}{8} + \frac{259}{9}\right) = \frac{241360}{1512} = 159.63$$

$$\sigma_{aPC} = E[aPC] - E[a]E[PC] = 159.63 - (1.556)(101.59) = 1.56$$

For the covariance between *b* and the perimeter cost we have

$$\mu_{bPC} = E(bPC) = \int_1^2\int_3^4 20(ab+b^2)\left(\frac{4}{21}ab\right)dbda = \frac{80}{21}\int_1^2\left(a^2\frac{b^3}{3} + a\frac{b^4}{4}\right)_3^4 da =$$

$$= \frac{80}{21}\left(\frac{a^3}{3}\frac{37}{3} + \frac{a^2}{2}\frac{175}{4}\right)_1^2 = \frac{80}{21}\left(\frac{259}{9} + \frac{525}{8}\right) = \frac{543760}{1512} = 359.63$$

$$\sigma_{bPC} = E[bPC] - E[b]E[PC] = 359.63 - (3.5238)(101.59) = 1.65$$

For the covariance between *a* and the area cost we have

$$\mu_{aAC} = E(aAC) = \int_1^2\int_3^4 20(a^2b)\left(\frac{4}{21}ab\right)dbda = \frac{80}{21}\int_1^2\left(a^3\frac{b^3}{3}\right)_3^4 da =$$

$$= \frac{80}{21}\left(\frac{a^4}{4}\frac{37}{3}\right)_1^2 = \frac{44400}{252} = 176.19$$

$$\sigma_{aAC} = E[aAC] - E[a]E[AC] = 176.19 - (1.556)(109.63) = 5.65$$

For the covariance between *b* and the area cost we have

$$\mu_{bAC} = E(bAC) = \int_1^2\int_3^4 20(ab^2)\left(\frac{4}{21}ab\right)dbda = \frac{80}{21}\int_1^2\left(a^2\frac{b^4}{4}\right)_3^4 da =$$

$$= \frac{80}{21}\left(\frac{a^3}{3}\frac{175}{4}\right)_1^2 = \frac{98000}{252} = 388.89$$

$$\sigma_{bAC} = E[bAC] - E[b]E[AC] = 388.89 - (3.52)(109.63) = 3.00$$

For the covariance between the perimeter cost and the area cost we have

$$\mu_{ACPC} = E(ACPC) = \int_1^2\int_3^4 20^2(a^2b + ab^2)\left(\frac{4}{21}ab\right)dbda = \frac{1600}{21}\int_1^2\left(a^3\frac{b^3}{3} + a^2\frac{b^4}{4}\right)_3^4 da =$$

$$= \frac{1600}{21}\left(\frac{a^4}{4}\frac{37}{3} + \frac{a^3}{3}\frac{175}{4}\right)_1^2 = \frac{1600}{21}\left(\frac{555}{12} + \frac{1225}{12}\right) = \frac{2848000}{252} = 11301.63$$

$$\sigma_{ACPC} = E[ACPC] - E[AC]E[PC] = 11301.63 - (109.63)(101.59) = 164.3$$

(e) Find the correlation coefficient of $a \cdot b$, $a \cdot PC$, $a \cdot AC$, $b \cdot PC$, $b \cdot AC$, and $PC \cdot AC$.

The general formula for the correlation coefficient is:

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

$$\rho_{ab} = \frac{\sigma_{ab}}{\sigma_a \sigma_b} = \frac{0.0}{\sqrt{1.56}\sqrt{3.52}} = 0.0$$

$$\rho_{PCAC} = \frac{\sigma_{PCAC}}{\sigma_{PC}\sigma_{AC}} = \frac{164.3}{\sqrt{65.2}\sqrt{481}} = 0.93$$

$$\rho_{aPC} = \frac{\sigma_{aPC}}{\sigma_{a}\sigma_{PC}} = \frac{1.56}{\sqrt{1.56}\sqrt{65.2}} = 0.15$$

$$\rho_{bPC} = \frac{\sigma_{bPC}}{\sigma_{b}\sigma_{PC}} = \frac{1.65}{\sqrt{3.52}\sqrt{65.2}} = 0.11$$

$$\rho_{aAC} = \frac{\sigma_{aAC}}{\sigma_{a}\sigma_{AC}} = \frac{5.65}{\sqrt{1.56}\sqrt{481.}} = 0.21$$

$$\rho_{bAC} = \frac{\sigma_{bAC}}{\sigma_{b}\sigma_{AC}} = \frac{3.00}{\sqrt{3.52}\sqrt{481}} = 0.07$$

These correlations (with the exception of *a* and *b*) are all positive. They should be because as you increase one side of the lot (either *a* or *b*), you should increase both the perimeter and the area. Also, as you increase the perimeter, on average, you increase the area, given our distribution function.

3.13. Subroutines

Code 3.1. Variance as a function of truncation

This simple code illustrates the need to keep all significant figures in the intermediate calculations of averages before computing the variance using equation (3.5).

```
   n=10;
   r = rand(n,1);
   s = 10 + 0.1*(2*r - 1)
   s2 = s.^2;
   f = 1/n;
   format long
   mu_s = sum(f*s);
   mu_s2 = sum(f*s2);
   var_s = mu_s2 - mu_s^2;
   for i = 2:1:8
      mu_s_cut(i) = round(mu_s*(10^i))/(10^i);
      mu_s2_cut(i) = round(mu_s2*(10^i))/(10^i);
      var_s_cut(i) = mu_s2_cut(i) - mu_s_cut(i)^2;
      fprintf(1,'%i sig figs:  mu_s = %16.9e mu_s2 = %16.9e var_s = %16.9e\n',
i+2, mu_s_cut(i),mu_s2_cut(i),var_s_cut(i));
   end
```

```
   fprintf(1,'all sig figs: mu_s = %16.9e mu_s2 = %16.9e var_s = %16.9e\n',
mu_s,mu_s2,var_s);
```

3.14. Problems

Homework problems are posted on the course website.