Problem Set #2

Solutions

(1) Stock and Watson 2.2: (a) First,

$$E(W) = 3 + 6E(X) \text{ and } E(V) = 20 - 7E(Y).$$

From table 2.2,

$$E(X) = 0(.3) + 1(.7) = .7$$

and

$$E(Y) = 0(.22) + 1(.78) = .78.$$

Therrefore,

$$E(W) = 3 + 6(.7) = 7.2$$

and

$$E(V) = 20 - 7(.78) = 14.54.$$

(b) First,

$$\sigma_W^2 = \text{Var}(W) = 36\text{Var}(X) \text{ and } \sigma_V^2 = \text{Var}(V) = 49\text{Var}(Y).$$

Now, again from table 2.2,

$$E(X^2) = 0(.3) + 1(.7) = .7 \text{ and } E(Y^2) = 0(.22) + 1(.78) = .78.$$

Thus,

$$\text{Var}(X) = E(X^2) - E^2(X) = .7 - .7(.7) = .21.$$

Similarly,

$$\text{Var}(Y) = E(Y^2) - E^2(Y) = .78 - (.78)(.78) = .1716.$$

Thus,

$$\text{Var}(W) = 36(.21) = 7.56 \text{ and} \text{Var}(V) = 49(.1716) = 8.41.$$

(c)Note that

$$\begin{aligned}
\sigma_{WV} &= \text{Cov}(W, V) \\
&= E(WV) - E(W)E(V) \\
&= E[(3 + 6X)(20 - 7Y)] - E(W)E(V) \\
&= E(60 + 120X - 21Y - 42XY) - E(W)E(V) \\
&= 60 + 120E(X) - 21E(Y) - 42E(XY) - E(W)E(V)
\end{aligned}$$

From previous parts of this exercise, we know all of the above quantities, except for $E(XY)$. We can, however, calculate this from table 2.2:

$$E(XY) = 0(.15) + 0(.07) + 0(.15) + 1(.63) = .63.$$

Therefore,

$\text{Cov}(W, V) = 60 + 120(.7) - 21(.78) - 42(.63) - (7.2)(14.54) = 60 + 84 - 16.38 - 26.46 - 104.688 = -3.528$

It follows that $\text{Corr}(W, V) = \text{Cov}(W, V)/\sqrt{\text{Var}(W)\text{Var}(V)} = -3.528/\sqrt{7.56(8.41)} \approx -.44$. It is reasonable to expect a negative correlation between rain and the shorness of the commute.

(2) Stock and Watson 2.8.

(a) The damage random variable, Y takes on only 2 values: 0 and 20,000. The probability of each occurence is .95 and .05, respectively. Thus,

$$E(Y) = 0(.95) + 20,000(.05) = 1,000.$$

(b) Let y_j denote the amount of damage to the j^{th} house. Then,

$$\overline{Y} = \frac{1}{100}\sum_{i=1}^{100} y_j.$$

As we showed in class, under iid sampling:

$$E(\overline{Y}) = \mu_y = 1,000.$$

The second part of the question asks: what is $\Pr(\overline{Y} > 2,000)$?

To address this question, we must make use of the *central limit theorem* result which establishes that

$$\frac{\overline{Y} - \mu_y}{[\sigma_y/\sqrt{n}]} \to N(0, 1).$$

With $n = 100$, this approximation is typically quite accurate.

We have already showed that $\mu_y = 1,000$. The variance of Y, σ_y^2, is obtained as

$$\text{Var}(Y) = E(Y^2) - E^2(Y).$$

We need to calculate $E(Y^2)$ in order to calculate the variance. This is easily done:

$$E(Y^2) = 0(.95) + (20,000^2)(.05) = 20,000,000.$$

Therefore,

$$\text{Var}(Y) = 20,000,000 - (1,000)(1,000) = 19,000,000.$$

We can then calculate the standard deviation of the sample mean (the denominator of our expression in the central limit theorem formula) as the square root of 19,000,000/100, which is approximately 435.9.

Now, getting back to our original question:

$$\begin{aligned}
\Pr(\overline{Y} > 2,000) &= \Pr([\overline{Y} - 1,000]/435.9 > [2,000 - 1,000]/435.9) \\
&= \Pr(Z > 2.29) \\
&= 1 - \Pr(Z \leq 2.29) \\
&= 1 - .9890 \\
&= .011
\end{aligned}$$

In the above, Z is a standard Normal variable, andthe number in the second to last line is simply read off the table in Stock and Watson. Now, consider what this really means. If one partipates in the insurance pool, then the expected (average) payment due to damage is the same as if you did not participate in the pool. However, the variance is significantly reduced. Without the pool, 5 percent of the time you will have to pay 20,000. With the pool, it is extremely unlikely (the probability is about 1 percent) that you will ever pay more than 2,000 in a given year!

3) Stock and Watson 2.10 (a)

$$\begin{aligned}
E(Y|X) &= E(X^2 + Z|X) \\
&= E(X^2|X) + E(Z|X) \\
&= X^2 + E(Z) \\
&= X^2.
\end{aligned}$$

The second line uses the fact that the expectation of a sum is the sum of the expectations, being careful to "distribute" the conditioning on X in each term. The third line notes that X and Z are independent [and therefore $E(Z|X) = E(Z)$]. In addition, if we know, or condition on X, then X and functions of X are no longer random - therefore $E(X^2|X) = X^2$. Finally, since Z is standard Normal, E(Z) = 0.

(b)

$$\begin{aligned}
\mu_y &= E(Y) \\
&= E(X^2 + Z) \\
&= E(X^2) + E(Z) \\
&= E(X^2) + 0 \\
&= 1
\end{aligned}$$

The only tricky part here is the last step. Since X is a standard Normal variable, $E(X) = 0$ and $\text{Var}(X) = 1$. Since $\text{Var}(X) = E(X^2) - E^2(X) = E(X^2)$, it follows that $E(X^2) = 1$.

(c)

$$\begin{aligned} E(XY) &= E(X[X^2+Z]) \\ &= E(X^3+XZ) \\ &= E(X^3)+E(XZ) \\ &= 0+E(X)E(Z) \\ &= 0 \end{aligned}$$

The second to last line used the fact that odd order moments of a standard normal variable are zero. The independence of X and Z was also used in this step.

(d)

$$\begin{aligned} \text{Cov}(X,Y) &= E(XY)-E(X)E(Y) \\ &= 0-E(X)E(Y) \\ &= 0-0(1) \\ &= 0 \end{aligned}$$

The second step used the result in part c. The final step subbed in the moments for $E(X)$ (as a std. Normal variable) and $E(Y)$ was calculated in (b).

(4)

$$\begin{aligned} \text{Cov}(aX+b,cY+d) &= E[(aX+b)(cY+d)]-E[(aX+b)]E[(cY+d)] \\ &= E[acXY+adX+bcY+bd]-[aE(X)+b][cE(Y)+d] \\ &= acE(XY)+adE(X)+bcE(Y)+bd-acE(X)E(Y)-adE(X)-bcE(Y)-bd \\ &= acE(XY)-acE(X)E(Y) \\ &= ac[E(XY)-E(X)E(Y)] \\ &= ac\text{Cov}(X,Y) \end{aligned}$$

(5)

$$\begin{aligned} \text{Var}(aX+Y) &= E[(aX+Y)^2]-E(aX+Y)E(aX+Y) \\ &= E(a^2X^2+2aXY+Y^2)-[aE(X)+E(Y)][aE(X)+E(Y)] \\ &= a^2E(X^2)+2aE(XY)+E(Y^2)-a^2E^2(X)-2aE(X)E(Y)-E^2(Y) \\ &= a^2[E(X^2)-E^2(X)]+2a[E(XY)-E(X)E(Y)]+[E(Y^2)-E^2(Y)] \\ &= a^2\text{Var}(X)+2a\text{Cov}(X,Y)+\text{Var}(Y). \end{aligned}$$

(6) This was discussed in class. (See the "tale of two estimators" powerpoint show for a complete answer). The first estimator is consistent, but biased. The second, in part b (which is the same as the powerpoint example, replacing n with 30) is unbiased and inconsistent, as its mass is placed farther and father away from θ as $n \to \infty$.