Economics 300-01: Quantitative Methods in Economics
Wesleyan University, Fall 2007

Selected Answers to Problem Set #4

4-26: Larry wishes to bid on 6 independent jobs, each with a constant probability of success of 40%. Thus this bidding process can be understood as a series of Bernoulli trials; the probably of success can be found with the binomial distribution.

1. Intuitively, each of the six bids that Larry submits has a 40% chance of acceptance, so the expected number of bids that actually would be accepted is $6 \times 0.4 = 2.4$. In general, if X is the random variable that represents number of successes in n trials, and each (independent) trial has (constant) probability π, then $\mathrm{E}[X] = n \cdot \pi$.

2. For every job he gets, Larry earns \$200. The fixed cost of preparing all six bids is \$300. Thus, Larry's profit function can be written as
$$P = \$200X - \$300\,.$$
To find the expected profit, we must find the expected value of the function above. Since it is a linear function, we know that
$$\mathrm{E}[P] = \$200\mathrm{E}[X] - \$300\,.$$
That is, the expectation of a sum (or difference) is the sum (difference) of the expectations. And the expectation of a constant is simply a constant. Since $\mathrm{E}[X] = 2.4$ from part (a) above, $\mathrm{E}[P] = \$480 - \$300 = \$180$.

3. When is profit positive? When $\$200X - \$300 > 0$, i.e. when $X > 1.5$. What is the chance (that is, the probability) that this will occur? Since only an integer number of bids actually will be accepted, $\Pr(X > 1.5) = \Pr(X \geq 2)$. Using the cumulative binomial table (Table III.c), we find that for $n = 6$ and $\pi = 0.4$, $\Pr(X \geq 2) = 0.767$, or roughly 77%. Thus, the probability that he does not make a profit (i.e. makes a loss) is $1 - 0.767 = 0.233$, or about 23%.

4-27: Now we suppose that the profit function for Larry is
$$P = \$200X - \$300 - \$20X^2\,.$$

1. The expected number of contracts he will be awarded is unrelated to the form of the profit function. Hence, he still is expected to win 2.4, as before.

2. To find the expected profit, notice that
$$\mathrm{E}[P] = \$200\mathrm{E}[X] - \$300 - \$20\mathrm{E}[X^2]\,.$$
Importantly, $\mathrm{E}[X^2] \neq (\mathrm{E}[X])^2$! Try an example for yourself: the expectation of a nonlinear function is never equal to the non-linear function of the expectation.

   So how can we determine the value of $\mathrm{E}[X^2]$ in order to solve this problem? Recall the following definition of the variance:
$$\sigma^2 = \mathrm{E}[X^2] - \mathrm{E}[X]^2\,.$$

We can rearrange this equation to find

$$\mathrm{E}[X^2] = \sigma^2 + \mathrm{E}[X]^2 .$$

We have already solved for $\mathrm{E}[X] = 2.4$ above; thus $(\mathrm{E}[X])^2 = 5.76$. How do we find the variance? Earlier in the text it was noted that for a binomially distributed random variable, $\sigma^2 = n \cdot \pi \cdot (1-\pi)$. So in this case, $\sigma^2 = 6 \times 0.4 \times 0.6 = 1.44$. Thus, $\mathrm{E}[X^2] = 5.76 + 1.44 = 7.2$.

Substituting this value into the expected profit equation reveals an expected profit of \$36.

There is a longer — yet still correct — way to solve this problem as well. Recognizing that $\mathrm{E}[P] = \mathrm{E}[g(X)] = \sum_{x \in X} g(x)\, p(x)$, where $g(X) = \$200X - \$300 - \$20X^2$, we can construct the following table, where the probabilities $p(x)$ come from Table III.b of the text:

x	$g(x)$	$p(x)$	$g(x)p(x)$
0	−300	0.047	−14.10
1	−120	0.187	−22.44
2	20	0.311	6.22
3	120	0.276	33.12
4	180	0.138	24.84
5	200	0.037	7.40
6	180	0.004	0.72
		1.000	36.06

Thus, $\mathrm{E}[g(X)] \approx 36$, which matches the above calculation. (The rounding error is due to the values of $p(x)$ being truncated at three decimal places.) Notice how much *more* work this is!

3. In part (b) above, the profit is still positive for $X = 2.4$. You should be able to readily confirm that he can earn a profit if he wins 2 (or more) jobs, but will not make a profit ($P < 0$) if $X \leq 1$. So this is the same answer as above: he has a 77% chance of earning a profit.

4-40: Let X be the (continuous) random variable that measures the life of a muffler. Then, using the notation we introduced in class, $X \sim \mathcal{N}(4.2, (1.8)^2)$.

1. If Mercury Mufflers does not replace a given muffler, it makes \$15 pure profit. However, if they must replace the muffler within the warranty period, then they lose \$15 − \$50 = \$35. Let α represent the proportion of the time the firm must pay a warranty claim, on average. Thus $(1-\alpha)$ gives the proportion of muffler sales that do *not* result in paying on the guarantee. We can then write the average net profit per sale as

$$\Pi = (1-\alpha) \cdot \$15 - \alpha \cdot \$35 .$$

Notice that α is the probability that the guarantee payment needs to be made; that is, $\alpha = \Pr(X < 3)$, or the proportion (probability) of muffler "lives" that are less than three years. How do we find $\Pr(X < 3)$? Recall that X is normally distributed. Since we are told that $\mu = 4.2$ and $\sigma = 1.8$, we see that 3 is less than one standard deviation *below* the mean. (You should draw the corresponding diagram if this is unclear to you.)

To compute this probability, we need to standardize:

$$\Pr(X < 3) = \Pr\left(\frac{X-\mu}{\sigma} < \frac{3-4.2}{1.8}\right) = \Pr(Z < -\frac{2}{3}) = \Pr(Z > \frac{2}{3}) \approx 0.25 ,$$

Then, the average net profit per muffler for Mercury is

$$\Pi = 0.75 \cdot \$15 - 0.25 \cdot \$35 = \$11.25 - \$8.75 = \$2.50\,.$$

2. To solve this part, we basically need to repeat the above analysis in reverse. Let x^* be the length (in years) of the new guarantee. Then to increase the average profit to \$5 per muffler, Mercury must reduce the probability that a muffler will fail. To find the answer, we set Π in the average net profit equation equal to \$5, and solve for the value of α that gives that per muffler profit level. Doing so yields $\alpha = 0.2$.

   So we wish to find $\Pr(X < x^*) = 0.2$. From table IV of the text, we can immediately confirm that $\Pr(Z > 0.84) = 0.2$. Of course we are looking for a *left* tail, but by symmetry we know that $\Pr(Z < -0.84) = 0.2$ as well.

   How do we get from Z, a standard normal random variable, to X, a normal random variable with mean $\mu = 4.2$ and standard deviation $\sigma = 1.8$? Intuitively we know that they must reduce the length of the guarantee, so that means $x^* < 3$. Standardizing would imply

$$\frac{x^* - \mu}{\sigma} = -0.84\,,$$

   so it follows that $x^* = \mu - (0.84 \times \sigma) = 4.2 - (0.84 \times 1.8) = 2.69$, or about 32.3 months. Our intuition tells us that if we found an answer *greater* than 3, we must have made an error (likely forgetting the minus sign in the above equation).

   Now to finally answer the question asked: Since the original guarantee was for 36 months (3 years), Mercury Mufflers should reduce their guarantee period by 3.7 months to garner an average net profit of \$5 per muffler sold. (Given the number of significant digits in this question, we probably should round our answer to four months.)

5-17: 1. The bivariate probability distribution for husbands' (X) and wives' (Y) income can be expressed as

	Y		
X	15	25	35
20	0.1	0.2	0.0
30	0.1	0.2	0.1
40	0.0	0.2	0.1

   I leave the graph of this joint probability distribution up to you.

2. Given the table of the joint distributions, it is relatively straight-forward to find the marginal distributions for X and Y, respectively. First, for X:

x	$p(x)$
20	0.3
30	0.4
40	0.3

   Notice that this is just the horizontal sum (over Y) for each row of the above joint distribution table. The mean of X is then

$$\begin{aligned} \mu_X &= \sum x \cdot p(x) \\ &= (20)(0.3) + (30)(0.4) + (40)(0.3) \\ &= 30\,. \end{aligned}$$

Similarly, the variance of X is

$$\begin{aligned}
\sigma_X^2 &= \sum (x-\mu_X)^2 \cdot p(x) \\
&= \sum x^2 \cdot p(x) - \mu_X^2 \\
&= (400)(0.3) + (900)(0.4) + (1600)(0.3) - 900 \\
&= 60\,.
\end{aligned}$$

For Y,the marginal distribution is given by summing down each column, i.e.

y	$p(y)$
15	0.2
25	0.6
35	0.2

The mean of Y can be found in exactly the same way as for X:

$$\mu_Y = \sum y \cdot p(y) = 25\,.$$

So too for the variance of Y:

$$\sigma_Y^2 = \sum y^2 \cdot p(y) - \mu_Y^2 = 40\,.$$

3. The covariance between X and Y can be found by using either of the following formulas:

$$\begin{aligned}
\sigma_{XY} &= \mathrm{E}[(X-\mu_X)(Y-\mu_Y)] &&= \sum_X \sum_Y (x-\mu_X)(y-\mu_Y)\cdot p(x,y) \\
&= \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] &&= \sum_X \sum_Y (xy)\cdot p(x,y) - \mu_X\mu_Y
\end{aligned}$$

Using the data of this problem, we employ the last of these formulas and compute the covariance as

$$\begin{aligned}
\sigma_{XY} &= (20)(15)(0.1) + (20)(25)(0.2) + (20)(35)(0.0) \\
&+ (30)(15)(0.1) + (30)(25)(0.2) + (30)(35)(0.1) \\
&+ (40)(15)(0.0) + (40)(25)(0.2) + (40)(35)(0.1) - (30)(25) \\
&= 770 - 750 = 20\,.
\end{aligned}$$

4. The question asks for the mean and variance of $S = X + Y$ to be calculated in two different ways. The "long" way (not recommended for exams) requires us to compute an S for each couple, then record the relative frequency (i.e.distribution) of these sums. This information is summarized below.

s	$p(s)$
35	0.1
45	0.3
55	0.2
65	0.3
75	0.1

Notice that the probabilities in this distribution sum up to one (as they must). Then one can show that

$$\mu_S = \sum s \cdot p(s) = 55\,,$$

and

$$\sigma_S^2 = \sum s^2 \cdot p(s) - \mu_S^2 = 140 \,.$$

The "short" way to solve this problem is to recognize that the random variable S is a linear function of X and Y, and so we can apply the appropriate formulas to this random variable, namely:

$$\mu_S = \text{E}[S] = \text{E}[X+Y] = \text{E}[X] + \text{E}[Y] = 30 + 25 = 55 \,.$$

Sure enough, this matches our above calculation. We also have a formula for the variance of a linear function of two random variables that we can apply here:

$$\text{var}(aX + bY) = a^2\text{var}(X) + b^2\text{var}(Y) + 2ab\,\text{cov}(X, Y) \,.$$

In this particular problem $a = 1$ and $b = 1$, so

$$\begin{aligned} \sigma_S^2 = \text{var}(X+Y) &= \text{var}(X) + \text{var}(Y) + 2\,\text{cov}(X, Y) \\ &= 60 + 40 + 2(20) = 140 \,. \end{aligned}$$

Again, this checks out with the solution to the "long" way.

5. Based on what we learned in the previous part, it should be relatively straight-forward to compute the mean and variance of the linear function $W = 0.6X + 0.8Y$. Notice that the mean is equal to

$$\text{E}[W] = 0.6\text{E}[X] + 0.8\text{E}[Y] = 18 + 20 = 38 \,,$$

and the variance equals

$$\begin{aligned} \sigma_W^2 = \text{var}(0.6X + 0.8Y) &= (0.6)^2\text{var}(X) + (0.8)^2\text{var}(Y) + 2(0.6)(0.8)\,\text{cov}(X, Y) \\ &= (0.36)(60) + (0.64)(40) + 2(0.48)(20) = 66.4 \,. \end{aligned}$$

6. Notice that the mean difference is

$$\text{E}[D] = \text{E}[X] - \text{E}[Y] = 5 \,.$$

The variance of this measure is equal to

$$\begin{aligned} \sigma_D^2 = \text{var}(X - Y) &= \text{var}(X) + \text{var}(Y) - 2\,\text{cov}(X, Y) \\ &= 60 + 40 - 2(20) = 60 \,. \end{aligned}$$

7. E[D] is not an especially good measure of sex discrimination, as it ignores all confounding factors that might help explain why the husbands in our sample earn an average of \$5000 more than the wives. For example, data on education and experience for each couple is missing. (Not to mention that this is a *very* small sample with only 10 couples!)

   Moreover, how do we know if \$5000 is large or small? While it doesn't seem trivial, is this difference *statistically significant*? We will learn how to answer this kind of question in a few weeks. For the time being, notice that standard deviation of D is $\sqrt{60} = 7.75$ or \$7750, which is larger than the average difference. This finding suggests that sampling error could account for most of the observed difference between men's and women's salaries in our (very small) sample.

5-18: Consider the mean and variance of various tax schemes:

1. If S is taxed at a straight 20%, then $T = 0.2S = 0.2(X+Y)$, so

$$\mathrm{E}[T] = 0.2(\mathrm{E}[X] + \mathrm{E}[Y]) = 11\,,$$

and

$$\begin{aligned}\sigma_T^2 = \mathrm{var}(0.2X + 0.2Y) &= (0.04)\mathrm{var}(X) + (0.04)\mathrm{var}(Y) + 2(0.04)\,\mathrm{cov}(X,\,Y) \\ &= 2.4 + 1.6 + 1.6 = 5.6\,.\end{aligned}$$

Hence, $\sigma_T = \sqrt{\sigma_T^2} \approx 2.4$.

2. If S is taxed at 50% but with the first \$15,000 exempted, then $T = 0.5(S-15) = 0.5(X+Y) - 7.5$. Hence,

$$\mathrm{E}[T] = 0.5(\mathrm{E}[X] + \mathrm{E}[Y]) - 7.5 = 20\,,$$

and

$$\begin{aligned}\sigma_T^2 = \mathrm{var}(0.5X + 0.5Y - 7.5) &= (0.25)\mathrm{var}(X) + (0.25)\mathrm{var}(Y) + 2(0.25)\,\mathrm{cov}(X,\,Y) \\ &= 15 + 10 + 10 = 35\,.\end{aligned}$$

In this case, $\sigma_T = \sqrt{\sigma_T^2} \approx 5.9$.

3. If S is taxed according to the progressive schedule in the text, then notice that the probability distribution for T can be found directly from part (**d**) of (5-17):

s	t	$p(s) = p(t)$
35	5	0.1
45	7	0.3
55	11	0.2
65	16	0.3
75	22	0.1

Thus,

$$\mathrm{E}[T] = \sum_T t \cdot p(t) = 11.8\,,$$

and

$$\sigma_T^2 = \sum t^2 \cdot p(t) - \mu_T^2 = 27.36\,.$$

Therefore, $\sigma_T = \sqrt{\sigma_T^2} \approx 5.2$. Notice that in this final case, the relationship between T and S is not linear, so we must use the probability table.

5-38: 1. From the definition of the variance:

$$\sigma^2 = \mathrm{E}[X^2] - \mathrm{E}[X]^2\,,$$

so $\mathrm{E}[X^2] = [\mathrm{E}[X]]^2$ when the variance is zero — that is, when X is not a random variable but a constant. And this equation holds approximately when there is very little variation in X.

2. Now, recall the definition of the covariance:

$$\sigma_{XY} = \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y]\,.$$

Thus, $\mathrm{E}[XY] = \mathrm{E}[X]\mathrm{E}[Y]$ when the covariance is zero. In this case, X and Y could still be random variables — in fact, each could have a sizable variance — but they do not "co-vary" in any systematic (linear) way: they are uncorrelated. In class we noted that if two random variables are independent, they will also be uncorrelated. (Importantly, the converse does not hold.)

Revised: 24-Sep-2007