Queen's University
Faculty of Arts and Sciences
Department of Economics
Winter 2009

Economics 250 : Introduction to Statistics

Midterm Exam I : Solutions
Time allowed: 80 minutes.

Instructions:

READ CAREFULLY. Calculators are permitted. At the end of the exam are several formulae. Answers are to be written in the examination booklet. Do not hand in the question sheet. You are to answer **ALL** questions. **SHOW ALL YOUR WORK**. Most of the marks are awarded for showing how a calculation is done and not for the actual calculation itself!

There are a total of 60 possible marks to be obtained. Answer all 6 questions (marks are indicated).

Do all 6 questions

1. (10 Marks) Two number are chosen (with replacement) from the set {1, 2, 3, 4}. Specify,

   a. the sample space

   b. the event "the sum of the chosen numbers is 5"

   c. the event "the first number is greater than the second"

   Solution:

   a. The sample space consists of all possible pairs, (x, y), chosen from the set {1, 2, 3, 4}. That is,

   $$S = \{(1,1),(1,2),(1,3),(1,4),\ldots,(4,1),(4,2),(4,3),(4,4)\}$$

   b. A sum of 5 is possible by getting one of the following 4 pairs: $(1,4),(4,1),(2,3),(3,2)$. There are $2^4 = 16$ possible pairs in S and so the probability is $4/16 = 0.25$.

c. The first number is greater than the second in the following pairs: $(2, 1), (3, 1), (4, 1), (3, 2), (4, 2), (4, 3)$. There are 6 such pairs. Hence, the probability is $6/16 = 0.375$

2. (10 Marks) 6-49 is a lottery in which players pick 6 different integers (without replacement) between 1 and 49, with the order of selection being irrelevant. The lottery commission then selects six of these as the *winning* numbers. A player wins the grand prize if all six numbers that he or she has selected match the winning numbers. He or she wins the second prize if exactly five, and the third prize if exactly four of the six numbers chosen match the winning ones. Find the probability that a certain player's ticket wins

   a. the grand prize

   b. the second prize

   c. the third prize

   Solution:

   a. There are $\binom{49}{6}$ possible lottery tickets. To win the grand prize, your numbers need to match the winning numbers. Hence, the probability of winning the grand prize $\frac{1}{\binom{49}{6}}$.

   b. To win the second prize, we need to match 5 numbers out of 6 while the last one can be any of the remaining numbers. The probability of that is $\frac{\binom{6}{5}\binom{43}{1}}{\binom{49}{6}}$.

   c. Similarly, the probability of winning the third prize is $\frac{\binom{6}{4}\binom{43}{2}}{\binom{49}{6}}$.

3. (10 Marks) A corporation regularly takes deliveries of a particular sensitive part from three subcontractors. It found that the proportion of parts that are good or defective from the total received were as shown in the following table:

	Subcontractor		
Part	A	B	C
Good	0.27	0.30	0.33
Defective	0.02	0.05	0.03

a. If a part is chosen randomly from all those received, what is the probability that it is defective?

b. If a part is chosen randomly from all those received, what is the probability that it is from subcontractor B?

c. What is the probability that a part from subcontractor B is defective?

d. What is the probability that a randomly chosen defective part is from subcontractor B?

e. Is the quality of a part independent of the source of supply?

f. In terms of quality, which of the three subcontractors is most reliable?

Solution: The marginal probabilities can be easily added to the table:

		Subcontractor		
Part	A	B	C	
Good	0.27	0.30	0.33	P(Good) = 0.90
Defective	0.02	0.05	0.03	P(Defective) = 0.10
	P(A) = 0.29	P(B) = 0.35	P(C) = 0.36	

a. The answer is just the marginal probability of defective which is 0.10.

b. The answer is the marginal probability $P(A) = 0.35$.

c. We are looking for $P(\text{Defective}|B) = \frac{P(\text{Defective} \cap B)}{P(B)} = \frac{0.05}{0.35} \approx 0.14$.

d. Here we are instead looking for $P(B|\text{Defective}) = \frac{P(B \cap \text{Defective})}{P(\text{Defective})} = \frac{0.05}{0.1} = 0.50$.

e. To assess if quality is independent of whether the product came from subcontractor A, we need to see if $P(\text{Good}|A) = P(\text{Good})$. The former is just $\frac{0.27}{0.29} \approx 93\%$. The latter is simply 0.90. Hence, clearly the two are not independent. Proceeding similarly for the subcontractors, for B we have $P(\text{Good}|B) = \frac{0.30}{0.35} \approx 85.71\%$ and $P(\text{Good}|C) = \frac{0.33}{0.36} \approx 91.67\%$. Neither of these equals 90%. Therefore, quality is not independent of source of supply.

f. We can answer this question by comparing how likely a given subcontractor is to produce a defective or good part. We have:

   1. P(A | Defective) = 0.02/0.29 ≈ 6.90%
   2. P(B | Defective) ≈ 14.29%
   3. P(C | Defective) = 0.03/0.36 ≈ 8.33%

   Hence, as A has the smallest defective rate (or the highest quality), A is the most reliable.

4. (10 Marks) Professor Mirza has made 30 exams of which eight are difficult, 12 are reasonable, and 10 are easy. The exams are mixed up, and he selects four of them at random to give to four sections of the course.

   a. Assuming that he can not give the same exam to different sections, what is the probability that all sections get a difficult exam?
   b. What is the probability that at least one section gets a difficult exam?
   c. How many sections would be expected to get a difficult test?
   d. How would your answers to a), b) and c) change if Professor Mirza *could* give the same exam to different sections?

   Solution:

   a. The chances of choosing a difficult exam on the first draw is just $\frac{8}{30}$. On the second, third and fourth draws the chances of picking a difficult exam are $\frac{7}{29}$, $\frac{6}{28}$, $\frac{5}{27}$, respectively. Hence, the probability of picking 4 difficult exams in a row is just $\frac{8}{20} \times \frac{7}{29} \times \frac{6}{28} \times \frac{5}{27} = \binom{8}{4}/\binom{30}{4} \approx 0.25\%$.
   b. We just apply the complement rule:

$$\begin{aligned} P(\text{at least one difficult exam}) &= 1 - P(\text{no difficult exams}) \\ &= 1 - \binom{22}{4}\Big/\binom{30}{4} \approx 73\% \end{aligned}$$

c. Let X be the number of difficult exams drawn. Then X is a discrete random variable that can take on the values $0, 1, 2, 3$ and 4. The expected value of X is

$$
\begin{aligned}
&E(X) \\
&= 0 \cdot P(X = 0) + 1 \cdot P(X = 1) + 2 \cdot P(X = 2) + 3 \cdot P(X = 3) + 4 \cdot P(X = 4) \\
&= 0 \cdot \frac{\binom{8}{0}\binom{22}{4}}{\binom{30}{4}} + 1 \cdot \frac{\binom{8}{1}\binom{22}{3}}{\binom{30}{4}} + 2 \cdot \frac{\binom{8}{2}\binom{22}{2}}{\binom{30}{4}} + 3 \cdot \frac{\binom{8}{3}\binom{22}{1}}{\binom{30}{4}} + 4 \cdot \frac{\binom{8}{4}\binom{22}{0}}{\binom{30}{4}} \\
&= 1.067
\end{aligned}
$$

d. When the same exam can be given to different section, we are choosing with replacement and the trials are independent. Moreover, the probability of choosing a difficult exam, $\frac{8}{30}$, is fixed across trials. Therefore, X as defined in b), is a binomial random variable with $n = 4, \pi = \frac{8}{30}$. Hence,

a.

$$
\begin{aligned}
P(\text{4 difficult exams}) &= P(X = 4) \\
&= \binom{4}{4}(\frac{8}{30})^4(1 - \frac{8}{30})^0 \\
&\approx 0.5\%.
\end{aligned}
$$

The ability to reuse exams makes it more likely that all exams will be difficult.

b.

$$
\begin{aligned}
P(\text{at least one difficult exam}) &= 1 - P(\text{no difficult exams}) \\
&= 1 - P(X = 0) \\
&= 1 - \binom{4}{0}(\frac{8}{30})^0(1 - \frac{22}{30})^4 \\
&\approx 71\%.
\end{aligned}
$$

The ability to reuse exams makes it *less* likely that at least one will be difficult as $8/30 < 7/29 < 6/28 < 5/25$.

c. The expectation of X is just $n \times \pi = 4 * \frac{8}{30} = 1.067$. Therefore, on average it doesn't matter if the exams can be reused across sections.

5. (10 Marks) On the basis of reconnaissance reports in the Kandahar area, Colonel Smith decides that the probability of an enemy attack against the left is 0.2, against the centre is 0.5, and against the right is 0.3. A flurry of enemy radio traffic occurs in preparation for the attack. Since deception is normal as a prelude to battle, Colonel Brown, having intercepted the radio traffic, tells General Quick that if the enemy wanted to attack on the left, the probability is 0.2 that we would have sent this particular radio traffic. He tells the general that the corresponding probabilities for an attack on the centre or the right are 0.7 and 0.1, respectively. How should General Quick use these two equally reliable staff members' views to get the best probability profile for the forthcoming attack?

   Solution: Let A be the event that the attack would be against the left, B be the event that the attack would be against the centre, and C be the event that the attack would be against the right. Let $\triangle$ be the event that this particular flurry of radio traffic occurs. The information Colonel Brown has provided are *conditional* probabilities of a particular flurry of radio traffic given that the enemy is preparing to attack against the left, the centre, and the right. However, Colonel Smith has presented *unconditional* probabilities, on the basis of reconnaissance reports, for the enemy attacking against the left, the centre, and the right. Because of these, the general should take the opinion of Colonel Smith as prior probabilities for A, B, and C. That is $P(A) = 0.2$, $P(B) = 0.5$, and $P(C) = 0.3$. Then he should calculate $P(A|\triangle)$, $P(B|\triangle)$ and $P(C|\triangle)$ based on Colonel Brown's view, using Bayes' theorem, as follows:

$$\begin{aligned}
P(A|\triangle) &= \frac{P(\triangle|A)P(A)}{P(\triangle|A)P(A) + P(\triangle|B)P(B) + P(\triangle|C)P(C)} \\
&= \frac{0.2 \times 0.2}{0.2 \times 0.2 + 0.7 \times 0.5 + 0.1 \times 0.3} \\
&= \frac{0.04}{0.42} \\
&\approx 0.095
\end{aligned}$$

   Similarly, $P(B|\triangle) = \frac{0.7 \times 0.5}{0.42} \approx 0.83$ and $P(C|\triangle) = \frac{0.1 \times 0.3}{0.42} \approx 0.071$.

6. (10 Marks) Edward's experience shows that 7% of the parcels he mails will not reach their destination. He has bought two books for \$20

each and wants to mail them to his brother. If he sends them in one parcel, the postage is \$5.20, but if he sends them in separate parcels, the postage is \$3.30 for each book. To minimize the expected value of his expenses (loss + postage), which way is preferable to send the books, as two separate parcels or as a single parcel?

Solution: Denote by p the probability that the parcel reaches it's destination. Of course, $p = 1 - 0.07 = 0.93$. We need to calculate the expected cost of sending one parcel and compare it with the expected cost of sending two parcels.

To compute the first, notice that if we call the parcel reaching its destination a "success" and not reaching its destination a "failure" then we can model delivery as a Bernoulli random variable. The "success" probability is p and payoffs (i.e. costs) are \$5.20 (only postage) upon "success" and $2 \times \$20 + \$5.20 = \$45.20$. The expected cost is:

$$\begin{aligned} E(1 \text{ parcel}) &= 5.20 \times p + (40 + 5.20) \times (1 - p) \\ &= 45.20 - 40p \\ &= \$8.00 \end{aligned}$$

When sending two parcels, we can model this as a binomial random variables with two trials, $\pi = 0.93$ and payoffs (i.e. costs) of \$3.30 when a single trial succeeds and $\$20 + \$6.60 = \$26.60$ when a single trial fails (we still have to pay postage for both parcels even though only one fails). The probability and payoffs for two trials are shown in the table below:

Case	Probability	Cost
2 successes (both parcels reach)	$\binom{2}{2}p^2(1-p)^0$	\$6.60
1 success (one of the two parcels reaches)	$\binom{2}{1}p^1(1-p)^1$	\$26.60
0 successes (neither parcel reaches)	$\binom{2}{0}p^0(1-p)^2$	\$46.60

The expected cost is:

$$\begin{aligned} E(2 \text{ parcels}) &= 6.60 \times p^2 + 26.60 \times 2p(1-p) + 46.60 \times (1-p)^2 \\ &= 46.60 - 40p \\ &= \$9.40 \end{aligned}$$

Given the expected costs, its better to send one parcel than two.