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

Economics 250 A: Introduction to Statistics

Midterm Exam I:

Time allotted: 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 100 possible marks. Answer all eight (8) questions (marks are indicated)

Part I (50 marks) This part tests your understanding of the concepts of population and sample.

Imagine the following situation where a discrete random variable X can take on 4 values with the following values and probability:

x	5	10	15	-5
$P(X = x)$	0.2	0.3	0.4	0.1

A researcher does not know this information and obtains a random sample of the following 5 observations from this population

1. **(10 marks)** What is the mean and variance of the population?

$$E[X] = .2 \times 5 + .3 \times 10 + .4 \times 15 - .1 \times .5 = 9.5$$
$$V[X] = .2 \times (5 - 9.5)^2 + .3 \times (10 - 9.5)^2 + .4 \times (15 - 9.5)^2 + .1 \times (-5 - 9.5)^2 = 37.25$$

2. **(10 marks)** What is the sample mean, median and variance of the sample?

$$\bar{X} = \frac{1}{5} \times (15 + 5 + 10 + 15 + 15) = 12$$
$$s_x^2 = \frac{1}{4} \times \left((15-12)^2 + (5-12)^2 + (10-12)^2 + (15-12)^2 + (15-12)^2\right) = 20$$

3. (**10 marks**) If another sample of 3 observations yielded a mean of 3 what is the sample mean for all 8 observations?

$$\bar{X}_8 = \frac{5}{8} \times 12 + \frac{3}{8} \times 3 = 8.625$$

4. **(20 marks)** If

$$Y = 10 - X^2$$

(a) Write out the probability distribution of Y?

Y_i	$P(Y = y_i)$
-15	0.3
-90	.3
-215	.4

(b) Calculate the $V[X]$ and show that this also proves that $E[X^2] \neq (E[X])^2$. Explain in words why this is not the case.

$$\begin{aligned} V[Y] &= E[X^2] - (E[X]^2) \\ &= 25 \times .3 + 100 \times .3 + 225 \times .6 - 12^2 = 28.5 \\ &\quad \text{Expectation is a linear operator !} \end{aligned}$$

Part II (50 marks) The section tests your understanding of conditional probabilities.

5. **(10 marks)** In a drug study people are divided evenly between those that get a drug and those that do not. The probability of getting better when participating in a drug test is 0.45. The probability of getting better when you get a drug is .9. If you observe someone who is better what is the probability that they had the drug? Calculate the probability of getting better when you do not get the drug?

$$\begin{aligned} P(D) &= .5 \quad P(B) = .45 \quad P(B \mid D) = .9 \\ P(D \mid B) &= \frac{P(B \cap D)}{P(B)} = \frac{P(D) \times P(B \mid D)}{P(B)} = \frac{.5 \times .9}{.45} = 1 \end{aligned}$$

:

$$P(B \mid \bar{D}) = \frac{P(B \cap \bar{D})}{P(\bar{D})} =$$
$$\begin{aligned} P(B) &= +P(B \cap D) + P(B \cap \bar{D}) \\ P(B \cap \bar{D}) &= 5. \times .9 - .45 = 0 \end{aligned}$$

: -0.065

6. (**20 marks**) Suppose there are three kinds of investors who make choices over cash and assets (stocks and bonds). Risk takers hold 10% in cash, risk neutral investors hold 50% in cash, and terrified folks hold 100% cash. Suppose there are 10 percent risk takers and twice as many terrified people as risk neutral people. Risk takers get a good night's sleep 70% of the time, whereas risk neutral people get a good night's sleep 50% of the time and sadly, terrified people never get a good night's sleep.

   (a) Write out all the information from this problem

$$\begin{aligned} P(RT) &= .10 \quad P(RN) = .3 \quad P(T) = .6 \\ P(S &\mid RT) = .7 \quad P(S \mid RN) = .5 \quad P(S \mid T) = 0 \end{aligned}$$

   (b) If you observe someone who just got a good night's sleep, what is the probability they were risk takers?

$$P(RT \mid S) = \frac{P(RT) \times P(S \mid RT)}{P(RT) \times P(S \mid RT) + P(RN) \times P(S \mid RN) + P(T) \times P(S \mid T)} = \frac{.1}{.1 \times .7 + .3}$$

   (c) What is the expected proportion of cash if you found someone who just had a good night's sleep?

$$\begin{aligned} E[C] &= C_{RT} \times P(RT \mid S) + C_{RT} \times P(RT \mid S) \\ &= .1 \times .32 + .5 \times .3 \times .68 = .422\% \text{ Cash} \end{aligned}$$

   (d) If the kind of investor you are is independent of education, and 70% of the population has a college degree. What is the probability of observing someone who had a good night's sleep but was a college graduate and a risk taker?

$$P(Coll \cap RT \mid S) = P(Coll) \times P(RT \mid S) = .7 \times .32 = .224$$

   :

7. (**10 marks**) Suppose fit people compose 30% of the population and those who do not watch the Olympics on TV are 75% of the population. Fit people watch the Olympics on TV 80% of the time. Write out the joint probability tablet.

$$\begin{aligned}
P(F) &= .3 \quad P(\bar{O}) = .75 \quad P(O \mid F) = .8 \\
P(F \cap O) &= P(F) \times P(O \mid F) = .3 \times .8 = .24 \\
P(F \cap \bar{O}) &= P(F) - P(F \cap O) = .3 - .24 = .06 \\
P(\bar{F} \cap O) &= P(O) - P(F \cap O) = .25 - .24 = .01 \\
P(\bar{F} \cap \bar{O}) &= P(\bar{O}) - P(F \cap \bar{O}) = .75 - .06 = .69
\end{aligned}$$

8. (**10 marks**) Suppose a variable X has a mean of 6 and a variance of 4. Another variable Y is constructed from it by adding 5 and multiplying it by 2. Transform the variable Y into a variable Z that has a mean of zero and a variance of 1. Show all work

$$\begin{aligned}
Y &= 5 + 2X \\
\bar{Y} &= 17 \quad s_y^2 = 16 \\
Z &= \frac{Y - 17}{4}
\end{aligned}$$

Formula Sheet

Statistics Formulas

Notation

- All summations are for $i = 1, \ldots, n$ unless otherwise stated.
- $\sim$ means 'distributed as'

Population Mean

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$$

Sample Mean

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Population Variance

$$\begin{aligned} \sigma^2 &= \frac{1}{N}\sum_{i=1}^{N}[x_i - \mu]^2 \\ &= \frac{1}{N}\sum_{i=1}^{N} {x_i}^2 - \mu^2 \end{aligned}$$

Sample Variance

$$s^2 = \tfrac{1}{n-1}\textstyle\sum_{i=1}^{n}[x_i - \bar{X}]^2$$

Alternatively

$$s^2 = \frac{1}{n-1}[\sum_{i=1}^{n} x_i^2 - n\bar{X}^2]$$

Grouped Data (with k classes)

$$\begin{aligned} \bar{X} &= \frac{1}{n}\sum_{j=1}^{k}\nu_j f_j \quad \text{where } \nu_j \text{ is the class mark for class } j \\ s^2 &= \frac{1}{n-1}\sum_{j=1}^{k} f_j(\nu_j - \bar{X})^2 \end{aligned}$$

Probability Theory

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \quad \text{(additive law)} \\ P(A \cap B) &= P(B)P(A \mid B) \quad \text{(multiplicative law)} \end{aligned}$$

If the E_i are mutually exclusive and exhaustive events for $i = 1, \ldots, n$, then

$$\begin{aligned} P(A) &= \sum_{i}^{n} P(A \cap E_i) = \sum_{i}^{n} P(E_i)P(A \mid E_i) \\ P(E_i \mid A) &= \frac{P(E_i)P(A \mid E_i)}{P(A)} \quad \text{(Bayes' Theorem)} \end{aligned}$$

Counting Formulae

$$P_R^N = \frac{N!}{(N-R)!}$$

$$C_R^N = \binom{N}{R} = \frac{N!}{(N-R)!R!}$$

Random Variables

Let X be a discrete random variable, then:

$$\begin{aligned}
\mu_x &= E[X] = \sum_x xP(X=x) \\
\sigma_x^2 &= V[X] = E[(x-\mu_x)^2] = \sum_x (x-\mu_x)^2 P(X=x) \\
\sigma_x^2 &= E[X^2] - (E[X])^2
\end{aligned}$$

The **covariance** of X and Y is

$$\begin{aligned}
Cov[X,Y] &= \sigma_{xy} = E\left[(X-\mu_x)(Y-\mu_y)\right] \\
&= E[XY] - E[X] \times E[Y]
\end{aligned}$$

The **correlation coefficient** of X and Y

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

If X and Y are independent random variables and a, b, and c are constants, then:

$$\begin{aligned}
E[a+bX+cY] &= a + b\mu_x + c\mu_y \\
V[a+bX+cY] &= b^2\sigma_x^2 + c^2\sigma_y^2
\end{aligned}$$

If X and Y are correlated then

$$V[a+bX+cY] = b^2\sigma_x^2 + c^2\sigma_y^2 + 2bc\sigma_{xy}$$

Coefficient of Variation (CV)

$$
\begin{aligned}
CV &= \frac{\sigma}{\mu} \times 100\% \text{ for population} \\
CV &= \frac{s}{\bar{X}} \times 100\% \text{ for sample}
\end{aligned}
$$

Univariate Probability Distributions

Binomial Distribution: For $x = 0, 1, 2, ..., n$ and :

$$Pr[X = x] = \binom{n}{x} \pi p^x (1 - p)^{n-x}$$

$$E[X] = np$$

$$V[X] = np(1 - p)$$