ECO251 QBA1
FIRST EXAM
February 16 and 17, 2006
ECO251 QBA1

Name: ____Key____________
Student Number : ____________________
Class Hour: ____________________

Remember – Neatness, or at least legibility, counts. In most non-multiple-choice questions an answer needs a calculation or short explanation to count.

Part I. (7 points)
(Source: Harvey J. Brightman) The following numbers are a sample and represent the pulse rates of 10 well-conditioned athletes.
31, 33, 36, 37, 37, 47, 44, 41, 38, 39.

Compute the following: **Show your work!**

a) The Median (1)
b) The Standard Deviation (3)
c) The 31st percentile (2)
d) The Coefficient of variation (1)

Solution:

Index	*x*	x^2	*x in order*
1	31	961	31
2	33	1089	33
3	36	1296	36
4	37	1369	37
5	37	1369	37
6	47	2209	38
7	44	1936	39
8	41	1681	41
9	38	1444	44
10	39	1521	47
	383	14875	

a) Median: $p(n+1) = .5(11) = 5.5 = a.b$ The middle numbers are the 5th and 6th number, which are 37 and 38. The median is the average of two middle numbers. $x_{.50} = \frac{x_5 + x_6}{2} = 37.5$.

Or $x_{1-p} = x_a + .b(x_{a+1} - x_a)$. So $x_{.50} = x_5 + .5(x_6 - x_5) = 37 + .5(38 - 37) = 37 + 0.5 = 387.5$

b) Standard Deviation: $\bar{x} = \frac{\sum x}{n} = \frac{383}{10} = 38.3$, $s^2 = \frac{\sum x^2 - n\bar{x}^2}{n-1} = \frac{14875 - 10(38.3)^2}{10-1}$

$= \frac{206.1}{9} = 22.90$. So $s = \sqrt{22.9} = 4.7854$. Note that saying a standard deviation is zero is equivalent to saying that all values of X are the same.

c) 31st percentile: $p(n+1) = .31(11) = 3.41$. So $a = 3$ and $.b = 0.41$

$x_{1-p} = x_a + .b(x_{a+1} - x_a)$ so $x_{1-.31} = x_{.69} = x_3 + 0.41(x_4 - x_3)$

$= 36 + 0.41(37 - 36) = 36.41$

d) $C = \frac{s}{\bar{x}} = \frac{4.7854}{38.3} = 0.1249$ or 12.49%

How this was done on Minitab – Version 1

Data (x) was placed in c2 (column 2).

```
MTB > describe c2
```

Descriptive Statistics: C2

```
Variable   N  N*   Mean  SE Mean  StDev  Minimum     Q1  Median     Q3
Maximum
C2        10   0  38.30     1.51   4.79    31.00  35.25   37.50  41.75
47.00
```

```
MTB > let c4 = c2*c2
```
Compute x^2 in column 4.

```
MTB > sort c2 c6;
SUBC> by c2.
```
x *in order* is put in column 6.

```
MTB > sum c2
```

Sum of C2

```
Sum of C2 = 383
```

```
MTB > sum c4
```

Sum of C4

```
Sum of C4 = 14875
```

```
MTB > print c2 c4 c6
```

Data Display

Row	C2	C4	C6
1	31	961	31
2	33	1089	33
3	36	1296	36
4	37	1369	37
5	37	1369	37
6	47	2209	38
7	44	1936	39
8	41	1681	41
9	38	1444	44
10	39	1521	47

Part II. (At least 35 points – 2 points each unless marked - Parentheses give points on individual questions. Brackets give cumulative point total.) Exam is normed on 50 points.

1. (Brightman) At an urban university, there are 7000 undergraduates whose ages are between 18 and 23, 2000 undergraduates between 24 and 29 years old, 1000 undergraduates between 30 1nd 35 years old and 1000 who are older than 35.
   a) Without doing any math, explain in plain English whether the mean will be below, the same as or above the median and why. (2)
   Answer: The 1000 people above 35 will pull the mean way up above the median. This distribution is skewed to the right.

   b) Where will the mode be relative to the mean and median? (1) [3]
   Answer: Since the median is usually between the mean and the mode, the mode must lie to the left of the mean and mode. If you made a diagram, it would show, left to right, mean median mode.

2. I have the average time of 10 randomly picked runners in the Boston Marathon.
   a) Is this a parameter or a statistic?
   Answer: This describes a sample and must be a statistic. A parameter describes a population.

   b) What symbol should you use to indicate this mean? [5]
   Answer: The symbol for a sample mean is $\bar{x}$.

3. For a rather shapeless distribution with one mode, a mean of 100 and a standard deviation of 2, we can say that the percent of data falling between 80 and 120 is
   a) At least 90%
   b) At most 90%
   c) 100%
   d) *At least 99%
   e) At most 99%
   f) None of the above. [7]
   Explanation: 120 is 100 + 10(2). 80 is 100 - 10(2). The Tchebyschev inequality says that the fraction of measurements more than $k = 10$ standard deviations from the mean is, at most,
   $\frac{1}{k^2} = \frac{1}{10^2} = \frac{1}{100}$. So the fraction between 80 and 120 is at least
   $1 - \frac{1}{100} = \frac{99}{100} = 99\%$.

4. For a mound-shaped (symmetrical) distribution with one mode, a mean of 100 and a standard deviation of 2, we can say that the percent of data falling above 96 is
   a) *About 97.5%
   b) About 95%
   c) Almost 100%
   d) About 68%
   e) None of the above
   Explanation: 96 is 100 - 2(2), i.e. 2 standard deviations below 100. The Empirical Rule says that 95% of data points will fall within 2 standard deviations of the mean. 5% will be in the tails of the distribution, because of symmetry, half (2.5%) will be above 100 + 2(2) = 104. The total is thus 95% + 2.5% = 97.5%.

5. The drawing of inferences about an unknown whole from a known part is
   a) Deductive reasoning
   b) *Inductive reasoning
   c) Census taking
   d) Sampling