

# STATISTICS FOR ECONOMISTS

## CLASS EXERCISE 2.

1. The data below give duration of unemployment by age, in July 1986.

Age	Duration of unemployment (weeks)				Total UE (000s)	Economically active (000s)
	<=8	8-26	26-52	>52		
	(percentage figures)					
16-19	27.2	29.8	24.0	19.0	273.4	1270
20-24	24.2	20.7	18.3	36.8	442.5	2000
25-34	14.8	18.8	17.2	49.2	531.4	3600
35-49	12.2	16.6	15.1	56.2	521.2	4900
50-59	8.9	14.4	15.6	61.2	388.1	2560
>=60	18.5	29.7	30.7	21.4	74.8	1110

(a) in what sense may these figures be regarded as probabilities? What does the figure 27.2 (top left cell) mean following this interpretation?

(b) which of the following statements are true:

- i. The probability of an economically active adult aged 25-34, drawn at random, being unemployed is 531.4/3600
- ii. If someone who has been unemployed for over one year is drawn at random, the probability that they are aged 16-19 is 19%.
- iii. For those aged 35-49 who became unemployed before July 1985, the probability of their still being unemployed is 56.2%.

(c) A person is drawn at random and found to have been unemployed for over one year. What is the probability that they are aged between 16 and 19?

2. "Odds" in horserace betting are defined as follows: 3/1 (three-to-one against) means a horse is expected to win once for every 3 times it loses; 3/2 means two wins out of 5 races; 4/5 (five to four *on*) means five wins out of 9 races, or for every four defeats, etc. Translate these odds into probabilities of victories. What makes the odds different from probabilities?

3. Historically the mean score in a statistics course is 63% with a standard deviation of 12, with scores following a normal distribution.

(i) If the department wants no more than 20% of students to fail the course, what should the pass mark be?

(ii) What is the probability that a randomly selected candidate scores between 60 and 70%?

4. In a normal distribution, we can estimate what percentage of values lie within a given number of standard deviations from the mean. What percentage of values lie within plus or minus: i) one; ii) two and iii) three st devs of the mean?

5. If you are told that the distribution is not normal, what rule would you use to estimate the percentage of values within a given number of standard deviations from the mean, and what percentage of values would lie within plus or minus two standard deviations of the mean?