

Name: _____ Section: (day/time) _____

AMS5 - MIDTERM
Thursday, February 4th, 2010

A Normal Table is on the last page of this exam.

You must explain all answers and/or show working for full credit.

There are some **bonus questions**. You can get full marks on this exam without answering these questions. If you do give correct answers, you will get extra marks. You cannot score more than 100% on the exam, however.

You are reminded of the University's policy on Academic Misconduct.

1. (2 marks) A large, representative sample of Americans was studied by the Public Health Service. The percentage of respondents who were left-handed decreased steadily with age, from 10% at 20 years to 4% at 70. "The data show that many people change from left-handed to right-handed as they get older." True or false? Why? If false, how do you explain the pattern in the data?

① + ①

False - the older people are not the same people as the younger ones

Older people were more likely to have been forced to use their right hands when learning to write.

2. (4 marks) A certain town has 25,000 families. These families own 1.6 cars, on average; the SD is 0.9. And 10% of the families have no cars at all. As part of an opinion survey, a simple random sample of 1,500 families is chosen. What is the chance that between 130 and 170 of the sample families will not own cars. Show your work.

①

families with no car = sum of draws.

Expected value $1500 \times (0 \times 0.9 + 1 \times 0.1) = 150$

Standard Error $\sqrt{1500 \times (1-0) \times 0.1 \times 0.9} = 11.619$

in standard units $130 = \frac{130 - 150}{11.619} = -1.72$

$170 = \frac{170 - 150}{11.619} = 1.72$

1

①



from table, chance is a 91.5%

①

3. (9 marks) We recorded the heights of 64 of you during class. The table below gives, for each height (in inches), the number of times that height was recorded.

height	(5	6)	43	59	60	62	63	64	65	66	67
frequency	5	1)	1	1	1	4	1	8	4	3	6
height	68	69	70	71	72	73	74	78	80	(87	99)
frequency	6	2	3	5	5	1	3	1	1	(1	1)

- (a) Which data values should be considered as errors? Why?

① heights of 5 and 6 inches - no one in class is that small tall.
87 and 99
(also 43" and 80" if justified in the explanation)

- (b) The median height for all the data is 67 inches. If the errors are excluded, what is the median height?

① These are 56 measurements remaining.
Median is average of 28th and 27th. These are both 67 inches.
The median is still 67 inches.

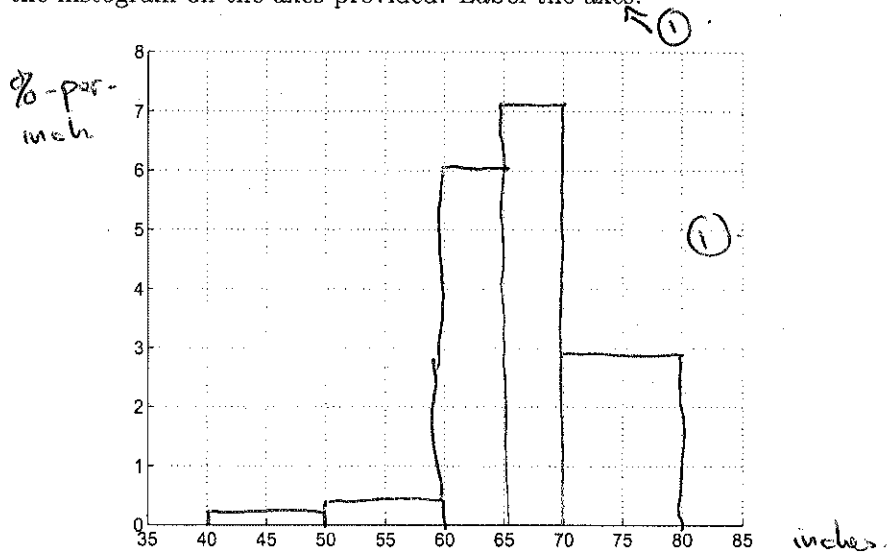
- (c) For the data excluding the errors, complete the table below, where the class interval includes the upper limit, but not the lower limit.

②

Class interval (inches)	frequency	percentage	width of class interval	percent-per-inch
40 - 50	1	1.8	10	0.18
50-60	2	3.6	10	0.36
60-65	18	30.1	5	6.1
65-70	20	35.7	5	7.1
70-80	16	28.6	10	2.9

[CONTINUED]

(d) Sketch the histogram on the axes provided. Label the axes.



(e) Does the histogram follow the Normal curve? Explain briefly.

① approximately - unimodal
vaguely symmetric.

(f) Excluding the outliers, the mean height is 67.2 inches and the SD is 5.3 inches. Using the Normal approximation, what proportion of students are 72 inches tall or taller?



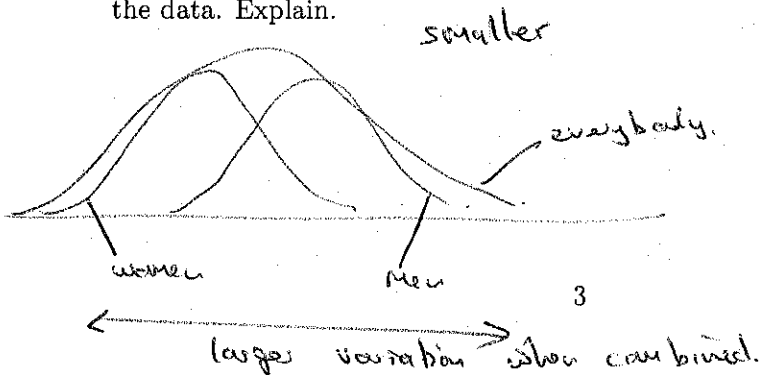
in standard units 72 is $\frac{72 - 67.2}{5.3} = 0.906$.

proportion is $\frac{1}{2}(100 - 64) = 18\%$.

(g) Excluding the outliers, what percentage of the students were actually 72 inches tall or taller?

① 11 out of 56 $\approx 19.6\%$

(h) [BONUS] (2 marks) The data includes heights for both men and women. If we were to look at the data for the men and women separately, would the standard deviations for the men and women be larger, smaller or about the same as the 5.3 inches found for all the data. Explain.



[TURN OVER]

4. (5 marks) Read the abstract of the article "Cannabis Use and Sexual Health" printed at the end of this exam paper.

(a) Was this a controlled experiment or an observational study? Why?

①

observational study.

participants self-assigned to use/not use.

(b) How were the subjects contacted? Give one potential problem with this method.

①

Random digit telephone survey

Non response bias.

(c) The investigators report that "after adjusting for demographic factors, daily cannabis use compared with no use was associated with an increased likelihood of reporting two or more sexual partners in the previous year".

①

Does using cannabis make you have more sexual partners? Explain.

No. There is an association between cannabis use and # sexual partners, but the survey does not show causation.

(d) What does "after adjusting for demographic factors" mean?

①

The demographics of the respondents (age, sex, etc) may not match the population (from census data). The weightings for the respondents were changed to give the demographic groups importance in relation to their proportion of the population.

(e) In the full article, the investigators state that the overall response rate was 56%. What effect might this response rate have on the results?

①

Introduce bias. People with particular behaviours may/may not want to participate.

[CONTINUED]

5. (7 marks) The article discussed in the previous question performed a telephone survey. Especially with potentially embarrassing questions or illegal behaviors, subjects may be unwilling to tell the truth.

(a) Would this bias the percentage of cannabis users reported in the study up or down? Explain briefly.

① Down - people who do use cannabis are likely to say they do not.

For a question with a yes/no response, one method to overcome this is as follows. You ask the subject to follow this procedure, in private, and tell you their answer at the end.

Toss a fair coin. If it comes up heads, answer honestly. If it comes up tails, toss the coin again, and answer "yes" if it comes up heads, and "no" if it comes up tails.

(b) Explain why this procedure means that, even if someone responds "yes", the investigators do not actually know whether the person does or does not use cannabis.

① Because they might be replying yes after tossing a tail followed by a head - which has no relationship with whether they do/do not use cannabis.

There are two different ways that result in someone answering "yes".

- (i) The first coin toss is heads, and they use cannabis.
- (ii) The first coin toss is tails, and the second coin toss is heads.

Are these events

(c) independent? mutually exclusive? Explain briefly.

① mutually exclusive, if (i) occurs, (ii) cannot.

[TURN OVER]

Consider a population where an unknown proportion, p of the population use cannabis.

- (d) What is the chance that a randomly selected person answers "yes" when asked if they use cannabis?

①

(i) $\frac{1}{2} \times p$ (H + all both)

(ii) $\frac{1}{2} \times \frac{1}{2}$ (H + T).

chance they answer "yes" = $0.25 + 0.5p$.

- (e) 1000 people were asked, and 325 replied "yes". What proportion of the population use cannabis?

①

$$\frac{325}{1000} = 0.25 + 0.5p \quad p = 0.15$$

- (f) Using the proportion of the population you calculated in part 5e, if you were to chose 10 people randomly from the population, what's the chance that 3 of them would be cannabis users?

①

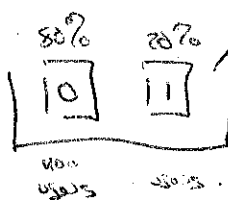
$n = 10$
 $k = 3$
 $p = 0.15$

$$\frac{10!}{3! 7!} (0.15)^3 (0.85)^7 = \underline{0.13}$$

- (g) If the population proportion is 20%, in a sample of 1000 you would expect there to be

①

200 users, plus or minus 12.6



expected value = $1000 \times (0.8 \times 0 + 0.2 \times 1) = 200$

standard error = $\sqrt{1000 \times (1-0) \times 0.2 \times 0.8} = 12.6$

[CONTINUED]