

STAT220: WINTER 2008: QUIZ 1: JAN 24

Student name _____

Section _____

Please fill in your name and section. Answer both the questions, on the question paper.

It should take at most 20 minutes to complete this quiz.

Continue your answer on to the back of the piece of paper if necessary.

BOTH questions relate to the following (invented) story.

Pretend it is October, 1999. Suppose UW has exactly 40,000 undergraduate students registered. The Student Union wishes to undertake a short mail-in survey of student study habits. UW administrators will make available a computerized list of all 40,000 students, ordered by ID#. As you know, the first two digits of your ID# are the year of first registering at UW. For example, a student who first registered in 1996, has ID# **96.....**

1. (Explain your answers briefly: 2 points each part)

A group of Student Union representatives get together to discuss how to choose a sample of 400 students.

(a) Fred wants to take a simple random sample. Dave says this is impossible. Is Dave correct?

No, Dave is wrong. We have a list defining the population, so we can take a simple random sample.
(2 points) (*Depending how it is to be done, it might be impractical, because 40,000 is a large number.* 1 point only)

(b) Jane suggests they take the first 400 students on the list provided by UW. Is this a good idea?

No, the first 400 ID#s are likely to be seniors, and/or older students returning after a break from school. This is a confounding factor; they are likely not typical of the whole population in their study habits. *Do not have to say confounding factor so long as meaning is made clear.*

(c) Sue suggests taking a random number from 1 to 100, and then taking every 100th student on the list from that point: for example, if 31 is chosen, then numbers 31, 131, 231, 331, ... 39,931. Is this a good idea?

Yes, a much better idea. An easy way to get a good sample. Everyone has the same chance to be sampled, and there is likely no factor that places people exactly multiples of 100 apart on the list that could be associated with study habits. (*Saying only part of this is enough for 2 points.*)

(d) The Union representatives worry about **non-response**. Dave says: we'll go to the HUB cafeteria and fill up to the 400 responses we want by asking students there to fill out the survey. Is this a good idea?

No. The students in the HUB are likely not typical of the total population in their study habits.
Could be they study more (if they are studying there), or less, or at different hours, but not typical.

(e) Fred says: we expect about a 50% response rate, so to get 400 responses we should choose a sample of 800 to mail the survey to. Is this a good idea?

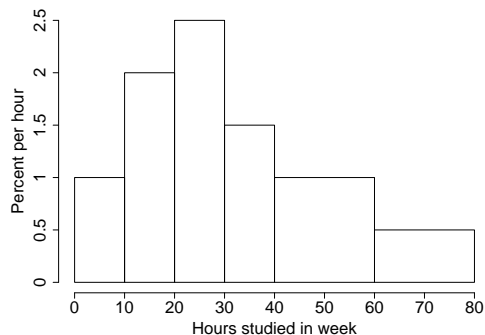
Yes, this is good. (*One additional comment for 2 points.*)

e.g. **He might be disappointed: he is very unlikely to get exactly 400 responses.**

or **It won't fix any of the sampling/(non)-response biases, but these are no worse than for an original sample of 400.** *Points such as these, with a "No", get 1 point only.*

2.(NO explanations necessary: circle your answer: 2 points each part)

Below is a histogram of the student responses to the number of hours spent studying in a particular week. The questions relate ONLY to the histogram summary of student responses, and NOT about any form of bias in the responses. Each box represents the responses in that interval. For example, students claiming to study at least 20 hours but less than 30 hours are said to study 20 to 30 hours, and scored in the 20 to 30 hours interval.



(a) The percentage of responses claiming 20 to 30 hours is approximately

2.5%, **25%**, 50%, "no way to know"

(b) More responses claim 30 to 40 hours than claim 40 to 60 hours: TRUE, **FALSE**, "about the same", "no way to know".

(c) More responses claim 18 to 19 hours than claim 21 to 22 hours: TRUE, FALSE, "about the same", **"no way to know"**.