Statistics 1040, Section 009, Midterm 1 (200 Points)

Friday, February 17, 2006

Your Name: _____

from: FPP, p. 130, Exercise Set B, Question 8 (Answers: p. A-56)
& Ytat 1040, Fall 2004, Quiz 5, Question 2

Question 1: Correlation (20 Points)

Investigators take a sample of DINKS (dual–income families, where husband and wife both work and have no kids). The investigators have data on the husband's income and the wife's income. By definition,

$$\text{family income} = \text{husband's income} + \text{wife's income}.$$

The average family income was around \$50,000, and 10% of the couples had family income in the range \$45,000–\$55,000. Fill in the blanks, using the options given below, and **explain briefly**:

1. (**10 Points**) The correlation between wife's income and family income is
   (e) somewhat positive.

   Textbook: "Although wife's income must be less than family income, the two are positively associated."

2. (**10 Points**) Among couples whose family income is in the range \$45,000–\$55,000, the correlation between wife's income and husband's income is *(b) nearly −1* _____.

   Textbook: "If family income is practically constant, the more the wife makes, the less the husband can make."

Options: (a) -1   (b) nearly -1   (c) somewhat negative   (d) 0
(e) somewhat positive   (f) nearly 1   (g) 1   (h) -1.1   (i) 1.1

"slightly" wrong answers:
1: (d), (f)
2: (a), (c)

Grading: 0: "totally" wrong answer, no explanation
2: "totally" wrong answer, some explanation
5: "slightly" wrong answer, no explanation
6: "slightly" wrong answer, some explanation
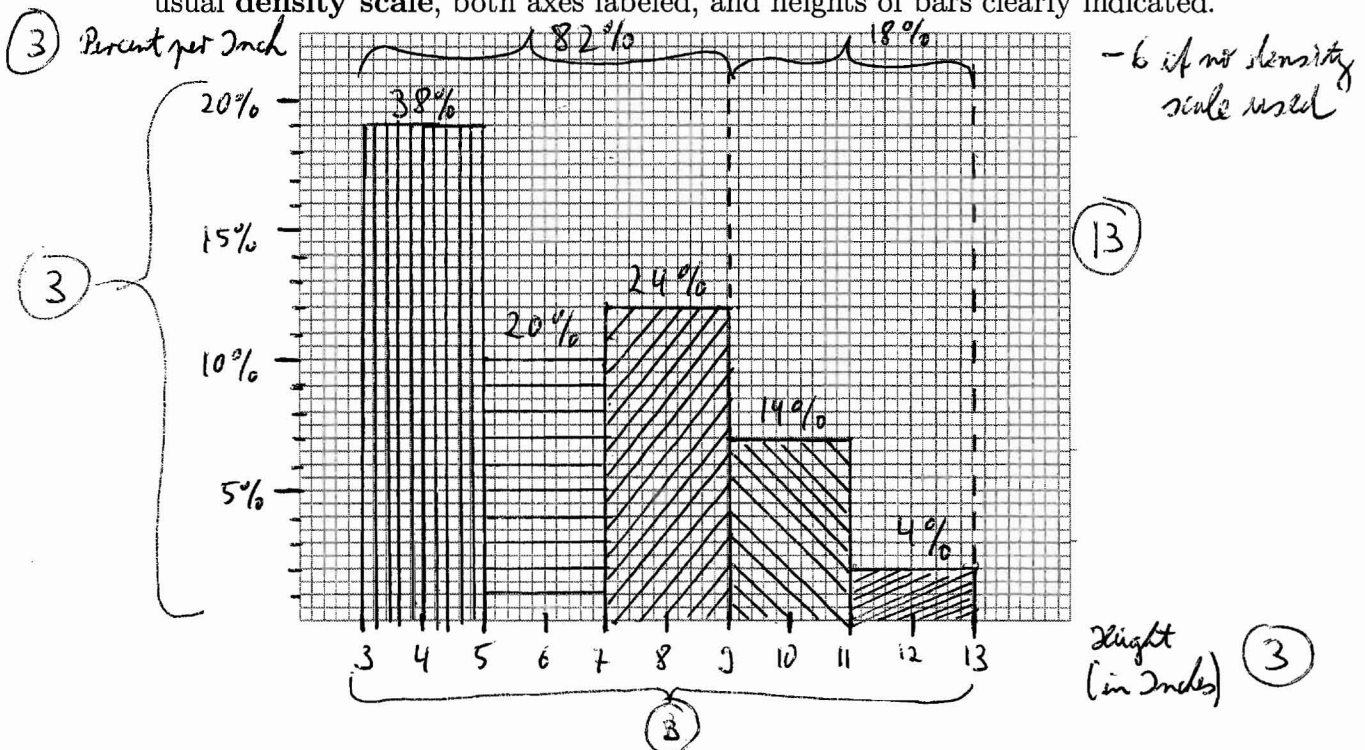*8: correct answer, no explanation*

1

Question 2: Histograms (**35 Points**)

When the Tribbles invaded the spaceship Enterprise, suppose that crew member Spock decided to take the logical step of seeing what the crew was up against, and he wanted to graphically represent the size of the Tribbles. Suppose that the table below summarizes the heights of the 50 Tribbles he found on the bridge. (Class intervals include the left but not the right endpoints.) [If you don't know what Tribbles are, take a look at http://www.startrek.com/startrek/view/series/TOS/episode/68744.html – there are several photos available at this Web site.]

Tribble Height (inches)	Number of Tribbles	*Percentage*	*Width*	*Height = $\frac{Percentage}{Width}$*
3–5	19	19/50 = 38%	2	38%/2 = 19%
5–7	10	10/50 = 20%	2	20%/2 = 10%
7–9	12	12/50 = 24%	2	24%/2 = 12%
9–11	7	7/50 = 14%	2	14%/2 = 7%
11–13	2	2/50 = 4%	2	4%/2 = 2%
	50	100%		

1. (**25 Points**) Draw a histogram of these height data, with the vertical axis on the usual **density scale**, both axes labeled, and heights of bars clearly indicated.

③ Percent per Inch
82%   18%
–6 if no density scale used
⑬
20%   38%
15%
10%   24%
20%
5%   14%
4%
3 4 5 6 7 8 9 10 11 12 13
Height (in Inches) ③
③ ③

2. (**10 Points**) If a Tribble is in the 82nd percentile for height, about how tall is it? (Note: Use the histogram, NOT the normal curve here.)

slightly wrong, using intervals "7-9" or "9-11" : ⑤

Its height is about: ___9 inches___ ⑩

82nd percentile means: 82% to the left of this value and 18% to the right of this value; this is at a height of approx. 9 inches

Question 3: Normal Curve (45 Points)

Spock did some additional height measurements of **all** Tribbles aboard Enterprise (and not only of those found on the bridge) and determined that their overall size closely follows the normal curve, with an average of 6.5 inches and an SD of 2.5 inches.

Fill the blanks in the statements below and **show all the work needed to obtain the answer:**

-2 for each calculation error

1. (**15 Points**) The percentage of Tribbles that are between 8.0 and 10.0 inches tall is __19.35%__.
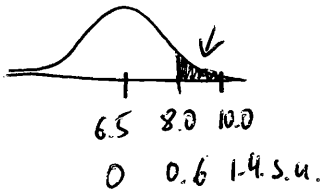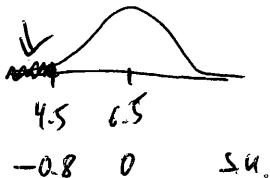
    S.u.: $\dfrac{8.0 - 6.5}{2.5} = 0.6$ ③   S.u.: $\dfrac{10.0 - 6.5}{2.5} = 1.4$ ③

    area between -0.6 and 0.6:
    45.15% ③

    area between -1.4 and 1.4:
    83.85% ③

    area between 0.6 and 1.4: $\dfrac{83.85\% - 45.15\%}{2} = 19.35\%$ ③

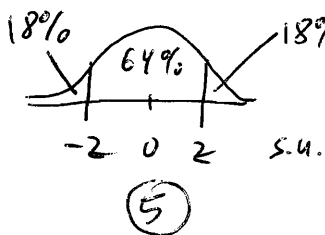2. (**15 Points**) The percentage of Tribbles that are less than 4.5 inches tall is __21.19%__.

    S.u.: $\dfrac{4.5 - 6.5}{2.5} = -0.8$ ⑤

    area between -0.8 and 0.8: 57.63% ⑤

    area below -0.8: $\dfrac{100\% - 57.63\%}{2} = 21.19\%$ ⑤

3. (**15 Points**) When using the normal curve, a Tribble that is in the 82nd percentile for height is about __8.75__ inches tall.

    18% 64% 18%

    -2 0 2 s.u.   ⑤

    area between -0.90 and 0.90: 63.19% (closest to 64%) ⑤
    ↑

    original units: $0.90 \cdot 2.5 + 6.5 = 8.75$ inches ⑤

3

Question 4: Controlled Experiment/Observational Study (**60 Points**)

A recent study in Europe looked at a large group of women of childbearing age. The researchers asked each woman how much alcohol they had consumed over the past 12 months. The researchers found that women who drank moderate amount of alcohol were somewhat less likely to have infertillity problems than women who did not drink alcohol at all (November, 2001). The study said it "controlled for age, income, and religion".

1. (**15 Points**) Based on the information above, was this a controlled experiment or an (observational study? Circle your answer and explain briefly.

    ⑩

    no intervention was used:- nobody was told to drink / not to drink

    ⑤ correct explanation

    ① some explanation

2. (**15 Points**) Why did they "control for" age, income, and religion?

    These may be confounding factors ⑮

    -5 for missing keyword
    (but otherwise correct
    explanation)

3. (**15 Points**) Is this convincing evidence that infertility would decrease if women with infertility problems started to drink moderate amounts of alcohol? (Note: we are only asking about infertility. There may be other problems introduced by such behavior, but ignore them for answering this question).

    No! ⑩ we only know that there is association between drinking and fertility; drinking does not cause fertility

    [association is not causation !]

    ⑤ correct explanation

    ① some explanation

4. (**15 Points**) Suggest a possible confounding factor (other than age, income, or religion) and **clearly explain** why you think it might be a confounding factor.

    general health (condition):
    someone who has some other medical problem may not drink and also be less fertile

    ⑩ for correct confounding factor
    ⑤ for correct explanation
    ① for some explanation

4

Question 5: Regression (**40 Points**)

In one study, the correlation between the educational level of husbands and wives in a certain town was about 0.50; both averaged 12 years of schooling completed, with an SD of 3 years.

1) husband, 2) wife : $avg_x = 12$ $SD_x = 3$

Show your work.

1) wife, 2) husband : $avg_y = 12$ $SD_y = 3$ $r = 0.50$

1. (**15 Points**) Predict the educational level of a woman whose husband has completed 18 years of schooling.

   −2 for each calculation error

   The answer is: _____15_____ years

   $$su_x = \frac{x - avg_x}{SD_x} = \frac{18 - 12}{3} = 2$$ (5)

   $$su_y = su_x \cdot r = 2 \cdot 0.50 = 1$$ (5)

   $$y = su_y \cdot SD_y + avg_y = 1 \cdot 3 + 12 = 15$$ (5)

2. (**15 Points**) Predict the educational level of a man whose wife has completed 15 years of schooling.

   The answer is: _____13.5_____ years

   $$su_x = \frac{15 - 12}{3} = 1$$ (5)

   $$su_y = 1 \cdot 0.50 = 0.50$$ (5)

   $$y = 0.50 \cdot 3 + 12 = 13.5$$ (5)

3. (**10 Points**) Apparently, well–educated men marry women who are less well–educated than themselves. But the women marry men with even less education. How is this possible?

   Nothing unexpected – this is just the <u>regression effect</u> !

   The explanation given is an example of the <u>regression fallacy</u>

   (10) for correct keyword
   (5) for reasonable explanation (without keyword)
   (1) for some explanation