Statistics 1040, Section 008, Midterm 1 (200 Points)

Friday, February 15, 2008

Your Name: _____

from: Final, Fall 2007, Question 1 (Solutions: → Course Web Page)

Question 1: Controlled Experiments / Observational Studies (**35 Points**)

In the December 10 issue of NEWSWEEK, medical writer Jerry Adler says:

"It's not too soon to start thinking about New Year's resolutions, and here's mine, as a medical writer: I will not report on any amazing new treatments for anything, unless they were tested in large, randomized, placebo–controlled, double–blind clinical trials published in high–quality peer–reviewed medical journals. If that means not telling NEWSWEEK's readers about, say, a new magnetized–water cure for osteoporosis, cancer and autism — well, there are infomercials to fill that gap."

1. (**10 Points**) Explain what it means for a study to be double–blind.

The subjects do not know whether they are in the treatment or in the control group; (5) nor do the doctors and nurses know who work with these subjects. (5)

2. (**15 Points**) Give 3 different reasons *why* a medical study should be double–blind.

- it guards against bias in the subjects' responses (5)
- it guards against bias in the doctors' and nurses' behavior towards the subjects (5)
- it guards against bias in doctors' assessment of a disease, i.e., did the patient improve / fully recover from the disease or not? (5)

3. (**10 Points**) What is a placebo? Why is it used?

- a placebo is a drug or vaccination (e.g., a sugar pill or a salt water injection) that resembles the treatment, but has no medical effect (5)
- it is used such that the subjects' response will be related to the treatment itself and <u>not</u> to the <u>idea</u> of the treatment (5)

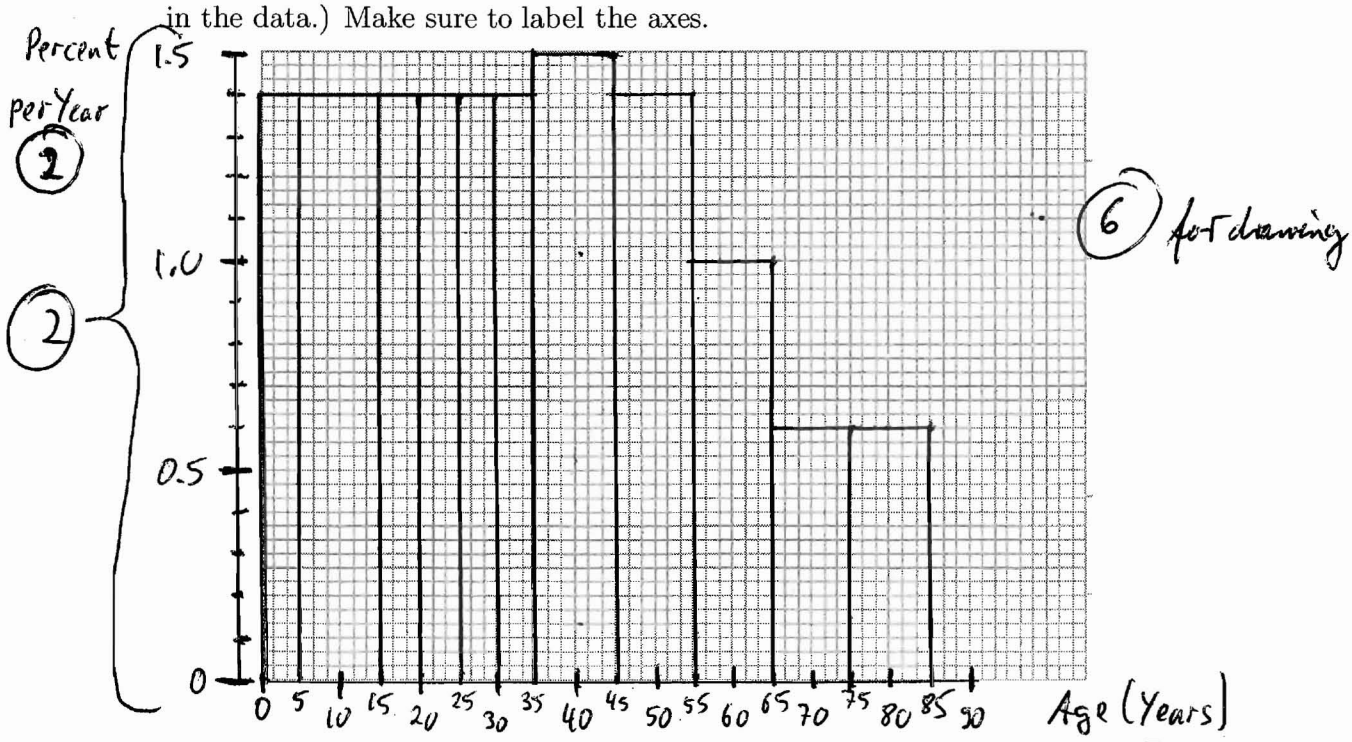
from: Quiz 2, Spring 2008 (!), Question 1 & FPP, Chapter 3, Review Exercise 2 [with one extra part] (Solutions: → Course Packet → Workbook)

Question 2: Histograms (**45 Points**)

The age distribution of people in the U.S. in 2004 is shown below. (3) (3)

Age	Percent of population	Width	Height
0-5	7	5	7/5 = 1.4
5-15	14	10	14/10 = 1.4
15-20	7	5	7/5 = 1.4
20-25	7	5	7/5 = 1.4
25-30	7	5	7/5 = 1.4
30-35	7	5	7/5 = 1.4
35-45	15	10	15/10 = 1.5
45-55	14	10	14/10 = 1.4
55-65	10	10	10/10 = 1.0
65-75 ~~85~~	6	10	6/10 = 0.6
75 and ~~over~~	6	10	6/10 = 0.6

1. (**20 Points**) Draw a histogram for these data on the graph paper provided. (The class intervals include the left endpoint, not the right; for instance, on the second line of the table, 14% of the people were age 5 years or more but had not yet turned 15. The interval "75 and over" can be ended at 85. Men and women are combined in the data.) Make sure to label the axes.



Use your histogram to answer the following questions on the next page: (2)

(2)

2. (**5 Points**) Are there more children age 1, or elders age 71?
Circle your answer.

5

→ age 1: 1.4%

age 71: 0.6%

3. (**5 Points**) Are there more 21–year–olds, or 61–year–olds?
Circle your answer.

5

→ age 21: 1.4%

age 61: 1.0%

4. (**5 Points**) Are there more people age 0–4, or 55–59?
Circle your answer.

5

→ age 0–4: 4 · 1.4% = 5.6%

age 55–59: 4 · 1.0% = 4.0%

5. (**5 Points**) The percentage of people age 35 and over is around 25%, 50%, or 75%?
Circle your answer.

5

sum up percentages up to age 35:

7% + 14% + 7% + 7% + 7% + 7% = 49%

therefore, "age 35 and over" = 100% − 49% = 51% (closest to 50%)

6. (**5 Points**) To be at the 35^{th} percentile of the age distribution, one has to be about 15 years old, 20 years old, or 25 years old?
Circle your answer.

5

sum up percentages until we reach 35%:

7% + 14% + 7% + 7% = 35%

↑ 0–5 ↑ 5–15 ↑ 15–20 ↑ 20–25

therefore, 35% are reached for ages 0 to 25

[Note: explanations weren't required!]

from: Final, Fall 2007, Question 7 [with an extra part] (Solutions: → Course Web Page)

Question 3: Normal Curve (**50 Points**)

A grocery store carries a variety of "on the vine" tomatoes with an average weight of 5.0 ounces and an SD of 0.9 ounces. The weights of these tomatoes follow the normal curve.
Show your work!

-2 for each calculation error

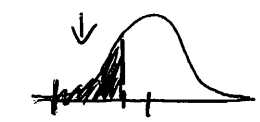
1. (**15 Points**) What percentage of them would weigh more than 6.0 ounces?
   The answer is: 13.57 %

s.u.: $\frac{6.0-5.0}{0.9} = \frac{1.0}{0.9} = 1.11$ (5)

area between −1.10 and 1.10: 72.87% (5)

area above 1.10: $\frac{100\% - 72.87\%}{2} = 13.57\%$ (5)

2. (**20 Points**) And what percentage would weigh between 3.7 ounces and 4.7 ounces?
   The answer is: 28.96 %



s.u.: $\frac{3.7-5.0}{0.9} = -1.44$ (4)

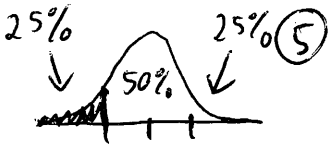
area between −1.45 and 1.45: 85.29% (3)

s.u.: $\frac{4.7-5.0}{0.9} = -0.33$ (4)

area between −0.35 and 0.35: 27.37% (3)

area between −1.45 and −0.35: $\frac{85.29\% - 27.37\%}{2} = 28.96\%$ (6)

3. (**15 Points**) Estimate the 25^{th} percentile of their weights.
   The answer is: 4.42 ounces



(5)

area between −0.65 and 0.65: 48.43% (closest to 50%)

original units: −0.65 · 0.9 + 5.0 = 4.42 ounces

(6) (1) (1) (1) (1)

from: Quiz 3, Fall 2007, Question 1
& FPP, Chapter 4, Review Exercise 3

(Solutions: → Course Web Page → Workbook)

Question 4: Average / SD (**40 Points**)

Part I:

Here is a list of numbers:

0.7 1.6 9.8 3.2 5.4 0.8 7.7 6.3 2.2 4.1
8.1 6.5 3.7 0.6 6.9 9.9 8.8 3.1 5.7 9.1

1. (**10 Points**) Without doing any arithmetic, guess whether the average is around (i) 1, (ii) 5, or (iii) 10. **Circle your answer and explain.**

(7)

"The average should be in the middle of the distribution: only three of the numbers are smaller than 1, and none are bigger than 10." (3)

2. (**10 Points**) Without doing any arithmetic, guess whether the SD is around (i) 1, (ii) 3, or (iii) 6. **Circle your answer and explain.**

(7)

"If the SD is 1, the entries 0.6 and 9.9 are much too far from the average. The SD can't be 6, because none of the numbers are more than 6 away from the average." (3)

from: Quiz 3, Fall 2007, Question 2
& FPP, Chapter 4, Review Exercise 7(b-c)

(Solutions: → Course Web Page → Workbook)

Part II:

A study on college students found that the men had an average weight of about 66 kg and an SD of about 9 kg. The women had an average weight of about 55kg and an SD of about 9 kg (Note that 1 kg = 2.2 lb).

1. (**10 Points**) Just roughly, what percentage of the men weighted between 57 kg and 75 kg?
Answer: 68 % (7)
Fill in your answer and explain.

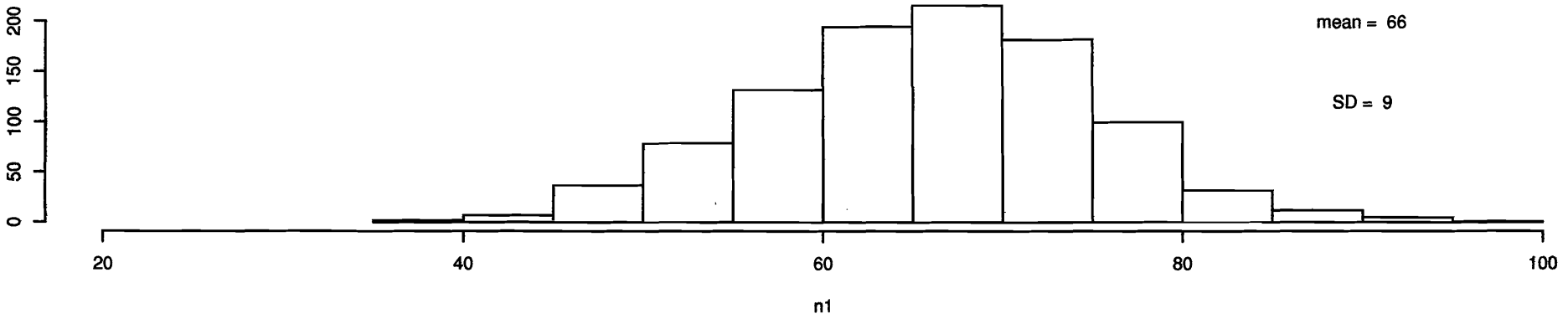
"The range is average ± 1 SD": 66 − 9 = 57, 66 + 9 = 75 (3)
our interval of interest

2. (**10 Points**) If you took the men and women together, would the SD of their weights be (i) smaller than 9 kg, (ii) just about 9 kg, or (iii) bigger than 9 kg? (7)
Circle your answer and explain.

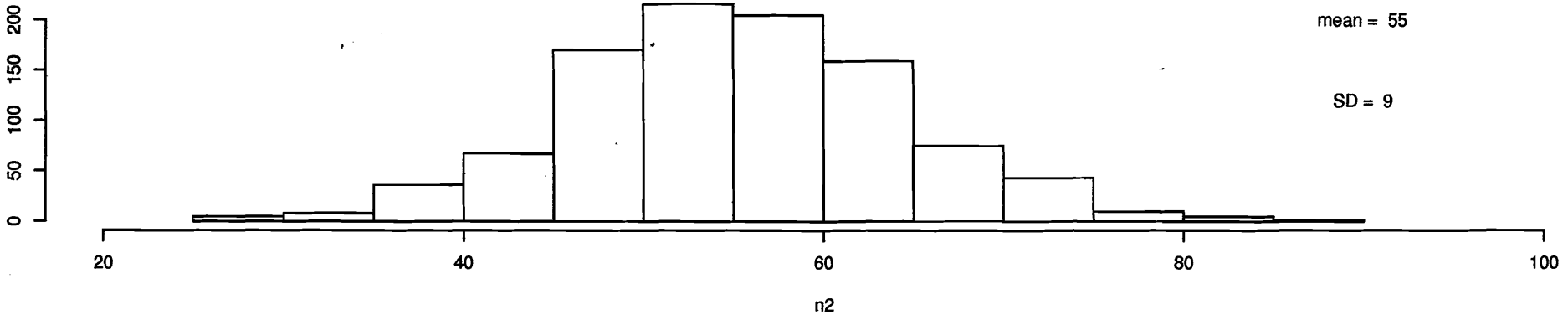
"If you take the men and women together, the spread in weights goes up." (3)

Question 4, Part II, Additional Explanation:

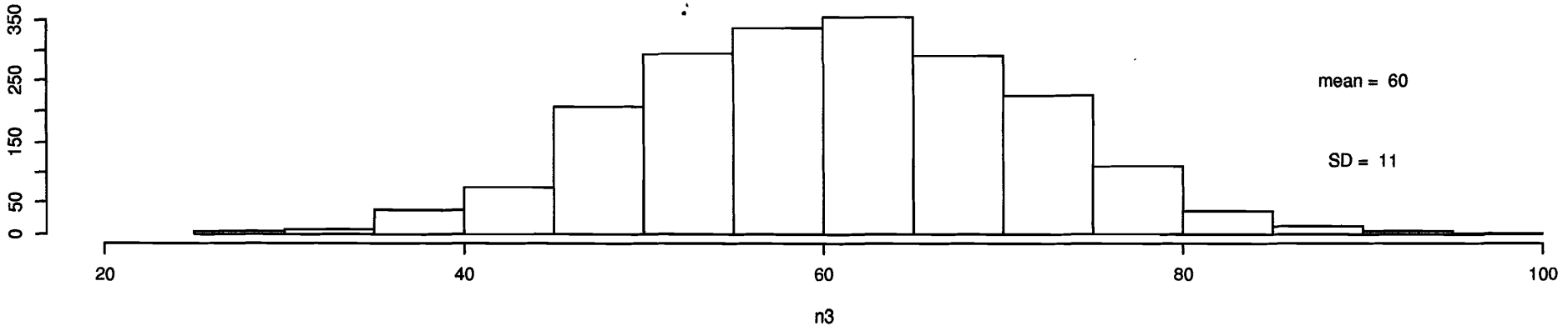
Men



Women



Both



From: Midterm 1, Spring 2006, Question 1
& FPP, Chapter 8, Exercise Set B, Question 8 (p. 130) [with different values]

(Solutions: → Course Webpage → Textbook p. A-56)

Question 5: Correlation (**30 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,

family income = husband's income + 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. (**15 Points**) The correlation between wife's income and family income is (e) somewhat positive.

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

2. (**15 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.

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

Options (for 1. and 2.): (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 (for each):
0: "totally" wrong, not explanation
2: "totally" wrong, some explanation
5: "slightly" wrong, not explanation
7: "slightly" wrong, some explanation
10: correct, not explanation
12: correct, some explanation
15: correct, correct explanation

Formulas:

$$\text{avg} = \frac{\text{sum of all numbers}}{\text{how many numbers}}$$

$$\text{SD} = \sqrt{\text{average of [(deviations from avg)}^2\text{]}}$$