

Name:

Stat 1040, Fall 2004  
Final Test, Wednesday December 15, 7:30-9:20 am

100 → **400**

Show your work. The test is out of 100 points and you have 110 minutes.

- 10 → **40** 1. Researchers think that eating "trans-fats" lowers the particle size of LDL molecules (so-called "bad cholesterol") in the body and hence increases the risk of heart disease. Stick margarine is high in "trans-fats" while butter is low in "trans-fats".

In a recent randomized, controlled, double-blind study, subjects were put on special diets for a period of 3 months. The *treatment group* received a special diet in which 20% of the calories came from stick margarine, while the *control group* received a special diet in which 20% of the calories came from butter. The food was prepared by taking the same low-fat diet and mixing either stick margarine or butter into foods such as muffins, casseroles, and hot cereals. Participants were required to eat all of the food provided in the special diets, and nothing else. At the end of the study, researchers measured the LDL particle size and found that the average LDL particle size for the treatment group was smaller than the average LDL particle size for the control group, and that the difference was "statistically significant".

- 8** (a) (2 points) Clearly explain what it means for the study to be *randomized*.

Participants were randomly assigned to the stick margarine (treatment) and butter (control) group.

- 8** (b) (2 points) Clearly explain what it means for the study to be *double-blind*.

Neither the participants eating the food nor the persons handing the food to the participants know who gets the food with stick margarine and who gets the food with butter.

- 8** (c) (2 points) Clearly explain what it means for the result to be *statistically significant*.

[Most likely, we would perform a 2-sample  $z$ -test, testing the null that both diets result in the same LDL particle size vs the alternative that the LDL particle sizes are different.] When the  $p$ -value of this test is between 1% and 5%, we would speak of a statistically significant result.

- 8** (d) (2 points) Why is it better to compare two groups like this instead of just putting all the people on the treatment diet (margarine) and comparing their LDL particle size at the beginning of the study to their LDL particle size at the end of the study?

There may be confounding factors that may have an effect on this longitudinal study, e.g., whether there is a seasonal effect (particle size may be different in Summer than in Spring/Fall), etc. Also, some people may have eaten margarine before, so there would be no noticeable difference for these people.

- 8** (e) (2 points) Why was the margarine or butter mixed into foods instead of being used as a spread?

The study should be double-blind, when used as a spread, some participants might be able to distinguish between margarine and butter - and, thus, the study wouldn't be blind any more.

10 → **40** 2. Six children attend a party. There are 7 party favors: 3 pink favors and 4 blue favors. The children are each given a party favor at random. [so this is without replacement!]

**8** (a) (2 points) What is the chance that the first child gets a pink favor?

1st pink:  $\frac{3}{7} = 0.4286 = \boxed{42.86\%}$

-2 for each calculation error

**8** (b) (2 points) What is the chance that the second child gets a pink favor?

2nd pink (and we don't know what the 1st was):  $\frac{3}{7} = 0.4286 = \boxed{42.86\%}$

[2nd pink (and 1st also pink):  $\frac{2}{6} = 0.3333 = 33.33\%$ ; but this wasn't the question!]

**8** (c) (2 points) If I see that the first 2 children received pink favors, what is the chance that the third child also gets a pink favor?

3rd pink (given 1st & 2nd pink):  $\frac{1}{5} = 0.2 = \boxed{20.0\%}$

**8** (d) (2 points) What is the chance that the left-over favor is blue?

7th blue (and we don't know what the 1st to 6th were):  $\frac{4}{7} = 0.5714 = \boxed{57.14\%}$

**8** (e) (2 points) What is the chance that the first two children get favors of different colors?

1st pink & 2nd blue (given 1st pink) or 1st blue & 2nd pink (given 1st blue):

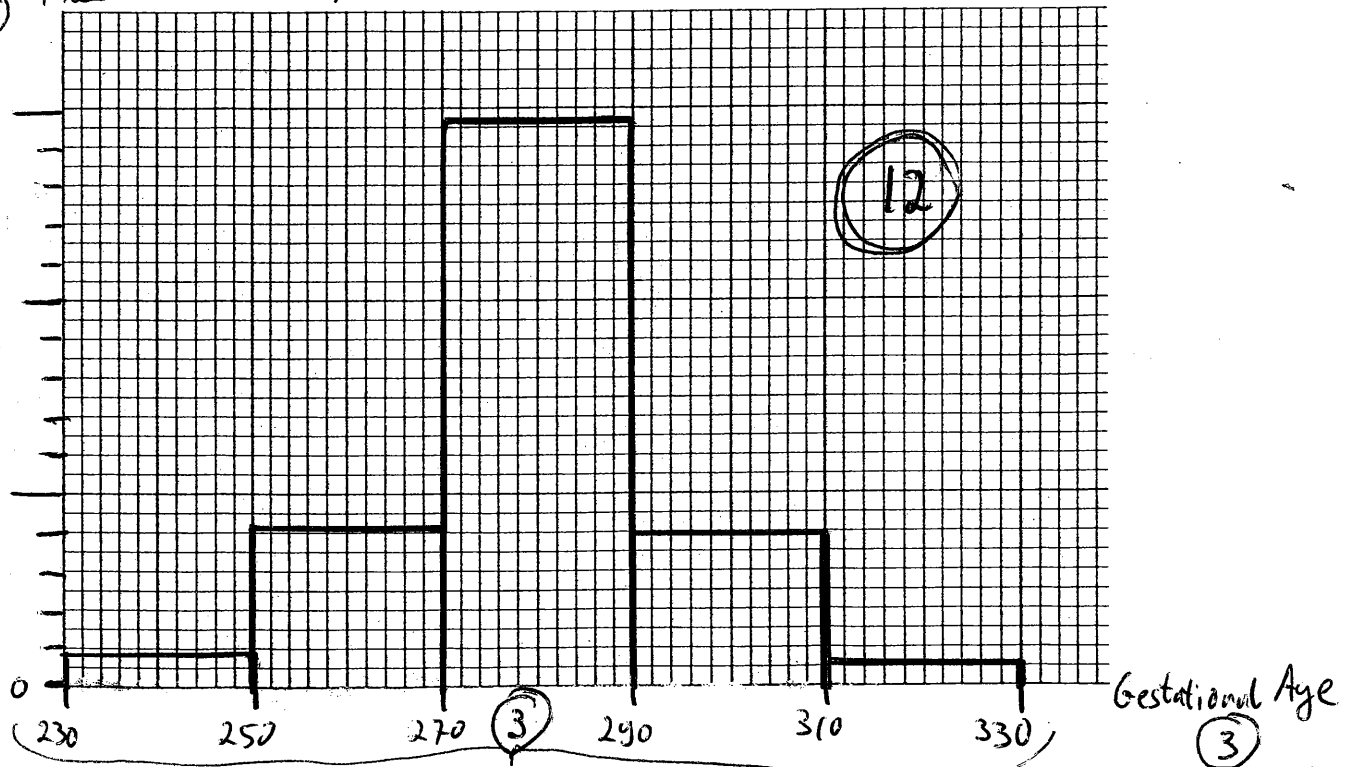
$$\frac{3}{7} \cdot \frac{4}{6} + \frac{4}{7} \cdot \frac{3}{6} = \frac{12}{42} + \frac{12}{42} = \frac{24}{42} = 0.5714 = \boxed{57.14\%}$$

6 → **24** 3. (6 points) The following table is for the gestational age of 1210 babies:

Gestational Age	Number
230-250	47
250-270	206
270-290	731
290-310	199
310-330	27

Draw a histogram for these data on the graph paper provided. Be sure to label the axes.

Frequency (or count or number)



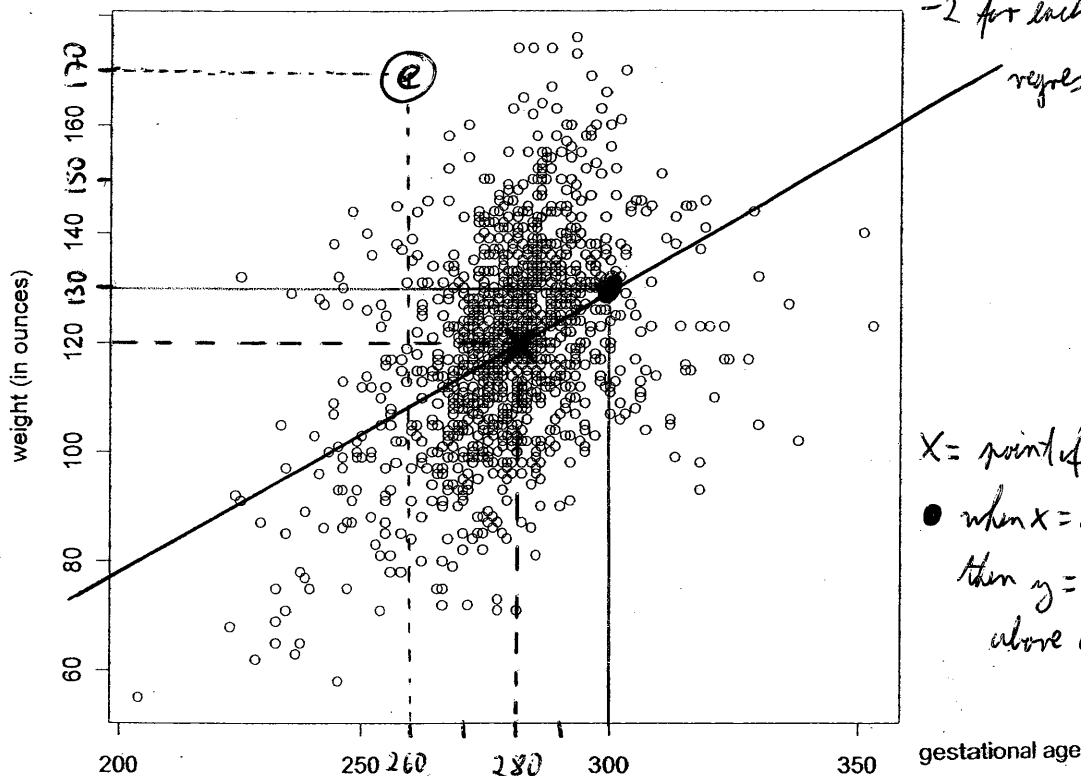
14 → 56 4. The following information was obtained from a group of 1221 babies:

X: Gestational age (in days): average = 280, SD = 15

Y: Birth-weight (in ounces): average = 120, SD = 18

The correlation coefficient was .42.

-2 if x, y not stated  
-4 if x, y swapped  
-2 for each calculation error



X = point of averages  
● when x = 20 days above avg x, then y = slope · 20 ≈ 10 ounces above avg y

20 (a) (5 points) Find the equation of the regression line for predicting birth-weight from gestational age.

$$\text{slope} = r \cdot \frac{SD_y}{SD_x} = .42 \cdot \frac{18}{15} = 0.504$$

$$\text{intercept} = \text{avg } y - \text{slope} \cdot \text{avg } x = 120 - 0.504 \cdot 280 = -21.12$$

$$\text{regression equation: } y = -21.12 + 0.504 \cdot x \quad \text{or} \quad \text{weight} = -21.12 + 0.504 \cdot \text{gest. age}$$

12 (b) (3 points) Sketch the regression line on the scatter diagram.

-6 if not through point of averages  
-6 if too steep / too flat

8 (c) (2 points) Predict the birth-weight for a baby with gestational age of 260 days.

for gest. age = 260:

$$\text{weight} = -21.12 + 0.504 \cdot 260 = 109.92 \text{ ounces}$$

8 (d) (2 points) Find the rms error of your prediction in part (c).

$$\text{r.m.s. error} = \sqrt{1 - r^2} \cdot SD_y = \sqrt{1 - 0.42^2} \cdot 18 = \sqrt{1 - 0.1764} \cdot 18$$

$$= \sqrt{0.8236} \cdot 18 = 0.9075 \cdot 18 = 16.34 \text{ ounces}$$

8 (e) (2 points) Would you be surprised to find that the baby in part (c) weighed 170 ounces? Explain.

$$\text{s.u.} = \frac{170 - 109.92}{16.34} = \frac{60.08}{16.34} = 3.68$$

3 yes, we would be quite surprised: A weight of 170 ounces is more than 3 rms errors above the predicted weight of 109.92 ounces. Only about 0.011% of babies with a gest. age of 260 have such a high (or even higher) weight.

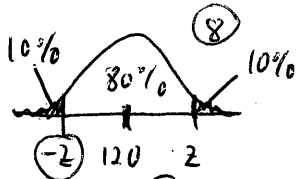
- 4 → **16** 5. (4 points) In an article in the May 2004 issue of the *Journal of occupational and Environmental Medicine*, authors describe an observational study of over 23,000 Detroit auto workers. They found that physically active employees had, on average, \$250 a year lower average health-care costs than sedentary employees. Their conclusion was that **to decrease health-care costs, sedentary employees should be encouraged to exercise**. Suggest one possible confounding factor that could account for the difference in health-care costs and explain why your confounding factor could make their conclusion incorrect.

Possible confounding factor(s): (just 1 needed)

⑫ Age: younger (and healthier) employees may be more physically active than older (less healthy) employees (4)

- 6 → **24** or: 6. (6 points) The birth-weight of babies follows the normal curve with an average of 120 ounces and an SD of 18 ounces. If a baby is at the 10th percentile for birth-weight, how much does the baby weigh?

avg = 120  
SD = 18



-2 for each calculation error

area between  $-1.30$  and  $1.30$ : 80.64%

weight =  $-1.30 \cdot 18 + 120 = 96.6$  ounces

↑ (2) ↑ (2) ↑ (2) ↑ (2)

- 12 → **48** 7. In a simple random sample of 4,700 U.S. adults, people were asked what foods and drinks they had consumed in the last 24 hours. It turned out that the average number of calories from sodas was 168 with an SD of 120.

- 32** (a) (8 points) Find an approximate 95% confidence interval for the average number of calories from sodas for all U.S. adults.

-2 for each calculation error

avg = 168

SD = 120

$SE_{sam} = \sqrt{4,700} \cdot 120 = 8,226.8$  (8)

$SE_{avg} = \frac{8,226.8}{4,700} = 1.75$  (8)

95% CI:  $168 \pm 2 \cdot 1.75 = 168 \pm 3.5 = 164.5$  to  $171.5$

↑ (2) ↑ (2) ↑ (2) ↑ (2) (4) (4)

- 8** (b) (2 points) Is your confidence interval still valid if you find out that the histogram of calories from sodas does not follow the normal curve? Explain briefly.

⑥ Yes, still valid: since the sample size (4,700) is very big, the prob. dist. for the average will follow the normal curve (and thus our interval is still valid). (2)

- 8** (c) (2 points) Is your confidence interval still valid if you find out that the sample is not a simple random sample? Explain briefly.

⑥ No, not valid: We need a SRS to correctly calculate a CI. (2)

12 → 48 8. (12 points) Researchers think that eating margarine lowers the particle size of LDL molecules (so-called "bad cholesterol") compared to eating butter. In a randomized, controlled, double-blind experiment, 105 people in the treatment group (margarine diet) had an average LDL M: margarine particle size of 252.6 Angstroms, with an SD of 4.6 Angstroms, while 110 people in the control group (butter diet) had an average LDL particle size of 254.8 Angstroms, with an SD of 4.1 B: butter Angstroms. Perform a test to determine whether the researchers' claim is correct. You must state a null and alternative hypotheses, compute a test statistic and a P-value, and clearly state your conclusions about the size of LDL molecules for people on margarine and butter diets such as those in this study.

2-sample z-test

1) null: margarine and butter diet result in same particle size, i.e.,  $\mu_{M} - \mu_{B} = 0$  (3)  
 alternative: margarine diet results in smaller particle size, i.e.,  $\mu_{M} - \mu_{B} < 0$  (1)

2)

<u>M</u>	<u>B</u>
sample size: 105	sample size: 110
sample $\mu_{M} = 252.6$	sample $\mu_{B} = 254.8$
sample $\sigma_{SD} = 4.6$	sample $\sigma_{SD} = 4.1$
$SE_{sam M} = \sqrt{105} \cdot 4.6 = 47.14$ (4)	$SE_{sam B} = \sqrt{110} \cdot 4.1 = 43.00$ (4)
$SE_{avg M} = \frac{47.14}{105} = 0.449$ (4)	$SE_{avg B} = \frac{43.00}{110} = 0.391$ (4)
$SE_{diff avg} = \sqrt{(0.449)^2 + (0.391)^2} = 0.595$ (4)	
$z = \frac{252.6 - 254.8}{0.595} = -3.70$ (4)	

3) area between -3.70 and 3.70: 99.978% (4)  
 → P-value:  $\frac{100\% - 99.978\%}{2} = 0.011\%$  (4)  
 4) reject the null (P-value < 5%) (4)  
 • result highly stat. sig. (P-value < 1%) (4)  
 • margarine diet results in smaller particle size (4)

10 → 40 9. (10 points) Researchers think anti-epileptic drugs accelerate bone loss in elderly women. To investigate, 12 women were randomly selected from all elderly women taking anti-epileptic drugs and they were monitored for a period of 5 years. At the end of the study, researchers measured their bone mineral density on a standardized scale. The average of the 12 measurements was -0.24 with an SD of 1.22. It is known that bone density measurements follow the normal curve. Test the hypothesis that the average for all such women is 0.0 against the alternative hypothesis that it is less than 0.0. State a null and an alternative hypothesis, find a test statistic and a P-value and clearly state your conclusions. (Note: negative values of bone mineral density correspond to accelerated bone loss.)

t-test: • sample size < 30 ✓  
 • SD for bone unknown ✓  
 • data follow normal curve ✓  
 1) null: drugs have no effect on bone loss, (3)  
 i.e.,  $\mu_{avg} = 0$  (1)  
 alternative: drugs accelerate bone loss, (3)  
 i.e.,  $\mu_{avg} < 0$  (1)

$t = \frac{-0.24 - 0}{0.368} = -0.65$  (4)  
 df = 12 - 1 = 11 (4)  
 3)  $t = -0.65$  above -0.70  
 ↓  
 25%  
 → P-value greater than 25% (4)  
 4) • do not reject the null (P-value > 25%) (4)  
 • drugs have no effect on bone loss (4)

2)  $SD^* = \sqrt{\frac{12}{11}} \cdot 1.22 = 1.044 \cdot 1.22 = 1.27$  (4)  
 $SE_{sam} = \sqrt{12} \cdot 1.27 = 3.46 \cdot 1.27 = 4.41$  (4)  
 $SE_{avg} = \frac{4.41}{12} = 0.368$  (4)

-2 for each calculation error  
 -36 for incorrect test  
 -4 if null alt swapped

-2 for each calculation error  
 -36 for incorrect test  
 -4 if null alt swapped

12 → **48** 10. (12 points) The following table represents an exit poll of 794 Utah voters from the 2004 election.

Age	Bush	Kerry	total
18-29	157	37	194
30-44	171	52	223
45-59	154	86	240
60 and over	97	40	137
total	579	215	794

expected

141.5	52.5
162.6	60.4
175.0	65.0
99.9	37.1

-2 for each calculation error  
-36 for incorrect test  
-4 if null, alt swapped

$$8 \times (2) = (16)$$

Treat this as though it is a simple random sample of Utah voters. Test to see whether reported voting behavior is independent of age for Utah voters for the 2004 Presidential election. You must clearly state a null and an alternative hypothesis, find a test statistic and a P-value, and clearly state your conclusions.

$\chi^2$  - Test for independence:

1, null: age and voting behavior are independent, i.e., boxes are identical (1)

alternative: age and voting behavior are not independent (voting behavior changes with age), (3)  
i.e., at least one box is different (1)

2, expected:  $\frac{194 \cdot 579}{794} = 141.5$  etc. (see table "expected" above)

$$\chi^2 = \text{sum of } \frac{(\text{obs} - \text{exp})^2}{\text{exp}} = \frac{(157 - 141.5)^2}{141.5} + \frac{(171 - 162.6)^2}{162.6} + \frac{(154 - 175.0)^2}{175.0} + \frac{(97 - 99.9)^2}{99.9} + \frac{(37 - 52.5)^2}{52.5} + \frac{(52 - 60.4)^2}{60.4} + \frac{(86 - 65.0)^2}{65.0} + \frac{(40 - 37.1)^2}{37.1}$$

$$df = (4-1) \cdot (2-1) = 3 \text{ (2)} = 17.49 \text{ (6)}$$

4 → **16** 11. (4 points) A sociologist is interested in whether men and women have different reactions to stress in the workplace. He gives a survey with 50 questions. For each question he does a chi-square test for independence between the participant's gender and their response to the question. He obtains 50 P-values (one for each of the questions), and finds that 3 of them are statistically significant. Then he starts writing his paper and describes how men and women differ with respect to stress in the workplace. What mistake is he making?

3,  $\chi^2 = 17.49$  greater than 11.34

→ P-value less than 1% (4)

4, • reject the null (P-value < 5%) (4)

• result highly stat. sig. (P-value < 1%) (4)

• voting behavior changes with age (they are not independent) (4)

Q11: When we run many tests, some of them will show a significant result just by chance. In fact, when we get a sig. result (P-value < 5%), we assume that the null is incorrect, but there remains a 5% chance that it is correct. So, out of 50 tests, we would expect that 5% (= 2.5 tests) show a sig. result. Observing 3 sig. tests is nothing unusual when we do 50 tests.