STAT 1220
Common Final Exam

Spring 2012
May 3, 2012

PLEASE PRINT THE FOLLOWING INFORMATION:

Name: _____   Instructor: _____

Student ID #: _____   Section/Time: _____

THIS EXAM HAS TWO PARTS.

PART I.
Part I consists of 30 multiple choice questions. Each correct answer is scored 2 points; each incorrect (or blank) answer is scored 0, so there is no penalty for guessing. You may do calculations on the test paper, but your answers must be marked on the OPSCAN sheet with a soft lead pencil (HB or No. 2 lead). Any question with more than one choice marked will be counted as incorrect. If more than one choice seems correct, chose the one that is most complete or most accurate. Make sure that your name and ID number are written and correctly bubbled on the OPSCAN sheet.

PART II.
Part II consists of three free response questions, with values as indicated. You must show all work in the space provided or elsewhere on the exam paper in a place that you clearly indicate. Work on loose sheets will not be graded.

FOR DEPARTMENT USE ONLY:

Part II.

Question	1	2	3
Score			

Part I	Part II	TOTAL

Part I

1. The Department of Education wishes to estimate the proportion of all college students who have a job off-campus. It surveyed 1600 randomly selected students; 451 had such jobs. The *population* of interest to the Department of Education is:

   (a) All 1600 students surveyed.
   (b) The 451 students in the survey who had off-campus jobs.
   (c) All college students.
   (d) All college students who have off-campus jobs.
   (e) None of the above.

   Problems 2 through 4 pertain to the following sample data:

   $$3, \quad -2, \quad 1, \quad 0, \quad -5, \quad 3, \quad 2, \quad 0, \quad -1$$

2. The mean of this data set is about

   (a) 0.1   (b) 1.7   (c) 0.3   (d) -1.5 (e) 2

3. The median of this data set is

   (a) 0   (b) -2.5 (c) -1 (d) 0.5   (e) 0.1

4. The sample standard deviation of this data set is about

   (a) 2.4   (b) 6.6   (c) 3.2   (d) 9.5   (e) 2.6

5. The *standard deviation* of a numerical data set measures the _____ of the data.

   (a) variability
   (b) size
   (c) range
   (d) average
   (e) most frequent value

6. The distribution of the lengths of a commercially caught fish is bell-shaped with mean 26.8 cm and standard deviation 4.8 cm. Any fish measuring less than 22.0 cm must be released. The proportion of fish that must be released is about

   (a) .05   (b) .32   (c) .16   (d) .025   (e) .68

7. The score made by a particular student on a national standardized exam is the 65th percentile. This means that

   (a) About 65% of all scores on the exam were higher than his.
   (b) He got about 65% of the answers correct.
   (c) About 65% of all scores on the exam were lower than his.
   (d) His score is the 65th best on the exam.
   (e) His score is 65% of the average score.

Problems 8 through 10 pertain to the data in the following two-way classification table, summarizing a random sample of 1200 young adults who were questioned about their age and where they live.

Domicile	Age 18	19	20	21	22
With parents	41	65	86	113	156
Elsewhere	212	179	153	105	90

Assume the proportions hold for all young adults from 18 to 22 years of age. A young adult is selected at random.

8. The probability that he lives with his parents is about

   (a) 0.62   (b) 0.27   (c) 0.16   (d) 0.38   (e) 0.53

9. The probability that he is still a teenager (18 or 19) is about

   (a) 0.55   (b) 0.37   (c) 0.41   (d) 0.48   (e) 0.22

10. The probability that he lives with his parents given that he is still a teenager (18 or 19) is about

   (a) 0.21   (b) 0.19   (c) 0.31   (d) 7%   (e) 0.23

Problems 11 and 12 are based on the following information: 12% of all drivers do not have a valid driver's license, 6% of all drivers have no insurance, and 4% have neither.

11. The probability that a randomly selected driver either fails to have a valid license or fails to have insurance is about

   (a) 0.18   (b) 0.2   (c) 0.22   (d) 0.072   (e) 0.14

12. The probability that a randomly selected driver fails to have insurance, given that he fails to have a valid license, is about

   (a) 0.67   (b) 0.33   (c) 0.06   (d) 0.50   (e) .01

Problems 13 and 14 pertain to the following situation.

There are seven intersections with traffic signals between Tristan's and Isolde's homes. If x denotes the number of signals at which Tristan must stop because of a red light on a randomly selected trip to Isolde's place, the probability distribution of x is

x	0	1	2	3	4
P(x)	0.15	0.05	0.10		0.40

Tristan never has to wait at more than four of the signals.

13. The missing entry in the table is

    (a) 0.30   (b) 0.25   (c) 0.00   (d) 0.15   (e) 0.35

14. The average number of intersections at which Tristan must stop because of a red light on trips to Isolde's place is

    (a) 2.00   (b) 1.50   (c) 3.01   (d) 2.75   (e) 1.65

15. The probability that a person who has made a reservation for a trip on a twelve-vehicle ferry will actually arrive and make the trip is 0.85. If to account for "no-shows" the ferry company makes 13 reservations for a particular trip, the chance that all 13 vehicles will show is about

    (a) 0.15   (b) 0.06   (c) 0.02   (d) 0.12   (e) 0.08

16. A trailer manufacturing company buys screw fasteners in boxes of 5,000. Two percent of all fasteners are unusable. The mean and standard deviation of the number x of unusable fasteners in a randomly selected box are about

    (a) $(2, 0.98)$ (b) $(10, 1)$ (c) $(200, 9.9)$ (d) $(100, 98)$ (e) $(100, 9.9)$

17. There are two independent tests for a particular disease. The probability that Test I will fail to detect the disease in someone who has it is 0.30. The probability that Test II will fail to detect the disease in someone who has it is 0.20. The probability that the disease will go undetected in someone who has it when both tests are applied is about

    (a) 0.10   (b) 0.06   (c) 0.50   (d) 0.32   (e) 0.60

In Problems 18 and 19, z denotes the standard normal random variable.

18. $P(z \geq -1.26)$ is approximately

    (a) 0.2600   (b) 0.8962   (c) 0.1038   (d) 0.8849   (e) 0.1151

19. $P(0.13 \leq z \leq 1.58)$ is approximately

    (a) 0.2617   (b) 0.1587   (c) 0.4703   (d) 0.3912   (e) 0.0406

Problems 20 and 21 pertain to the following situation.
The level of cholesterol in the blood of women aged 20 to 55 in a particular country is normally distributed with mean 212 mg/dℓ and standard deviation 45.2 mg/dℓ.

20. The probability that a randomly selected such woman has cholesterol level between 200 and 240 mg/dℓ is about

    (a) 0.8802   (b) 0.8850   (c) 0.3936   (d) 0.3540   (e) 0.3388

21. The probability that the mean cholesterol level in a random sample of twenty such women is between 200 and 240 mg/dℓ is about

    (a) 0.3388   (b) 0.8802   (c) 0.2758   (d) 0.9402   (e) 0.3936

22. The hip width x of adult females is normally distributed with mean 38.3 cm and standard deviation 4.31 cm. The minimum width of an aircraft seat that will accommodate 98% of all adult women is about

    (a) 48.0 cm   (b) 42.6 cm   (c) 47.1 cm   (d) 40.5 cm   (e) 51.2 cm

23. An economist wishes to estimate the proportion of household income spent on energy (gas, electricity, and so on), at 95% confidence and to within five percentage points. Assuming no prior knowledge of the true proportion the minimum sample size needed is about

    (a) 271   (b) 542   (c) 664   (d) 385   (e) 165

24. Measurement of the waiting time of 45 randomly selected patients at a hospital emergency room gave mean and sample standard deviations 11.3 and 6.5 minutes. A 90% confidence interval for the mean waiting time of all patients is about

    (a) $11.3 \pm (1.645)\left(\dfrac{6.5}{\sqrt{45}}\right)$

    (b) $11.3 \pm (2.015)\left(\dfrac{6.5}{\sqrt{45}}\right)$

    (c) $11.3 \pm (1.680)\left(\dfrac{6.5}{\sqrt{45}}\right)$

    (d) $11.3 \pm (1.960)\left(\dfrac{6.5}{\sqrt{45}}\right)$

    (e) $11.3 \pm (1.679)\left(\dfrac{6.5}{\sqrt{45}}\right)$

25. In a random sample of 1,300 recent college graduates, 238 had student loans totalling at least $25,000 when they graduated. A 99% confidence interval for the proportion of all graduates with at least $25,000 in student loans when they graduated is about

    (a) 0.18 ± 0.05
    (b) 0.18 ± 0.005
    (c) 0.18 ± 0.07
    (d) 0.18 ± 0.03
    (e) 0.18 ± 0.01

    Problems 26 and 27 pertain to the following situation.
    In recent years the mean number of miles driven per year by residents of a particular state was 11,568. A researcher wishes to test, at the 5% level of significance, whether the mean is different now. In a random sample of 62 drivers, the mean and sample standard deviation of the number of miles driven last year were 10,822 and 1,741, respectively.

26. The rejection region for the relevant test is

    (a) $(-\infty, -2.000] \cup [2.000, \infty)$
    (b) $(-\infty, -2.000]$
    (c) $(-\infty, -1.999] \cup [1.999, \infty)$
    (d) $(-\infty, -1.999]$
    (e) $(-\infty, -1.670]$

27. The value of the test statistic for the test is about

    (a) -3.345 (b) 3.374 (c) 3.345 (d) -2.855 (e) -3.374

28. In a test of hypotheses of the form $H_0 : p = 0.69$ vs. $H_1 : p \neq 0.69$ a sample of size 1,250 produced the test statistic $z = -2.417$. The p–value (observed significance) of the test is about:

    (a) 0.008   (b) 0.016   (c) 0.032   (d) -0.032
    (e) impossible to tell from the information given

29. In a survey of several hundred college graduates with a degree in business, the linear regression equation relating overall GPA x and monthly starting salary y (in thousands of dollars) was $\widehat{y} = -0.572 + 1.501x$. Select the most accurate statement.

    (a) GPA and starting salary are slightly negatively correlated.
    (b) For each 1.5 point increase in GPA monthly salary increases by about \$572, on average.
    (c) For each 0.572 point increase in GPA monthly salary increases by about \$1500, on average.
    (d) For each 1 point increase in GPA monthly salary increases by about \$572, on average.
    (e) For each 1 point increase in GPA monthly salary increases by about \$1500, on average.

30. In order to compare the performance of students in large enrollment and small enrollment sections, thirty-five students from large sections and thirty-five students from small sections of a freshman mathematics course were randomly selected. The mean and sample standard deviation of grades on the common final exam for the students from large sections were 72.8 and 7.4; for the students from small sections the mean and sample standard deviation were 75.3 and 6.8. Assuming that the populations are normally distributed with equal standard deviations, a 90% confidence interval for the difference in the mean common final exam scores between students in the two types of sections is about

    (a) 2.5 ± 2.20
    (b) 2.5 ± 2.83
    (c) 2.5 ± 3.39
    (d) 2.5 ± 1.67
    (e) 2.5 ± 0.28

Part II

1. Over the past few years the proportion of American adults who use tobacco products has been 0.21 or 21%. A researcher wishes to test, at the 10% level of significance, whether the proportion is different now. He takes a random sample of 1,600 adults and finds that 304 use tobacco products.

   (a) State the null and alternative hypotheses for the test. [2 points]

   (b) Construct the rejection region. [2 points]

   (c) Compute the value of the test statistic, and make a decision. [4 points]

   (d) State a conclusion about the proportion of American adults who use tobacco products, based on the test you performed. [2 points]

   (e) Compute the p-value (observed significance) of the test and state its meaning in the context of the problem. [4 points]

2. A researcher believes that average dwelling house size in one region of the country has declined in the last five years. From property records he took a random sample of fifty houses from the population (Population 1) of all houses in service this year and a random sample of fifty houses from the population (Population 2) of all houses in service five years ago, with the following results:

$$
\begin{array}{llll}
\text{now} & n_1 = 50 & \bar{x}_1 = 2392 & s_1 = 325 \\
\text{past} & n_2 = 50 & \bar{x}_2 = 2434 & s_2 = 336
\end{array}
$$

The populations may be assumed to be normally distributed with equal standard deviations. Test whether the mean house size now is less than what it was five years ago, at the 10% level of significance, in the following series of steps.

(a) State the null and alternative hypotheses for the test. [2 points]

(b) Construct the rejection region. [2 points]

(c) Compute the value of the test statistic, and make a decision. [4 points]

(d) State a conclusion about the means, based on the test you performed. [2 points]

3. The circumference x in inches (measured four feet off the ground) and volume y in cubic feet for 37 pine trees ranging in circumference from 25.1 to 50.3 inches were measured. Summary statistics are:

$$n = 37 \quad 25.1 \le x \le 50.3 \quad \Sigma x = 1384 \quad \Sigma y = 1346 \quad SS_{xx} = 2365 \quad SS_{xy} = 5268 \quad SS_{yy} = 13{,}483$$

(a) Find the proportion of the variability in the volume of pine trees that is accounted for by size (circumference). [4 points]

(b) Find the regression line for predicting y from x. [4 points]

(c) If the circumference of a pine tree four feet off the ground is 36 inches, what volume is predicted by the regression equation found in part (b)? [2 points]

(d) For what sizes of pine tree (i.e., what values of the circumference) is a computation like that done in part (c) valid? Explain fully. [2 points]

(e) Using the value $s_e = 7.04$, construct a 95% confidence interval for the slope of the population regression line, and interpret its meaning in the context of the problem. [4 points]