Statistics 2 Final Exam
Spring 2000

Printed Name _____
(Please print your name at the top of each page)

Signature _____

Student ID #_____

Circle your TA: *Susan Alber Nicola Armstrong Tony Hua Pawel Lasieki Michael Roberts*

You must show your work on all problems. No credit will be given for a correct answer with no justification.

1	2	3	4	5	6	7	8	9	10	11	12	Total

San Francisco Chronicle A5

Latinos Only Ethnic Group Not to Benefit From Boom

LOS ANGELES TIMES

WASHINGTON — Today's record-setting economic boom has enriched most members of every major ethnic group except one — Latinos.

The Federal Reserve's most recent survey of household finances shows that whites, blacks and Asians all have gained ground during the expansion.

By contrast, from 1995 to 1998 — the most recent period studied — the median Latino household net worth fell by a whopping 24 percent. For this group, median net worth (half of households have less, half have more) was $12,170 in 1995 but only $9,200 three years later.

Experts outside the Federal Reserve offered two principal explanations: an accelerating influx of poor immigrants, and a population singularly ill-equipped to profit from the Information Revolution.

Even as the wealth of the median Latino household was falling, the Latino community's total wealth was rising. That was reflected in data showing that average Latino household income (total wealth per household) rose from $61,000 in 1995 to $86,000 in 1998.

1. [4] This article claims that the median net worth (total wealth) per household decreased very substantially from 1995 to 1998. It also claims that the average net worth increased substantially over the same period. Is this possible, or is there something wrong with the numbers? Explain clearly and briefly.

2. [4] A red die and a blue die are rolled. You win $1 the outcome on the blue die is at least twice the outcome on the red die or if the outcome on the red die is at least twice the outcome on the blue. What is the chance you win $1?

3. [4] Five cards are dealt off the top of a shuffled deck. Find the chance that the first two cardsare diamonds and none of the others are.

4. [4] Observing between 45 and 55 heads on 100 tosses of a fair coin is as likely as observing between ____ and ____ heads on 10,000 tosses.

5. [4] A report by the Environmental Defense Association discusses the relationship between air pollution and annual death rates for a sample of 47 major cities in the US. The average death rate is reported as 9/1000 and the SD as 3/1000. The RMS error for predicting death rates from air pollution is reported as 4/1000. Is there anything wrong with these numbers? Or do you need more information to decide? Explain briefly.

6. [4] The May, 2000 issue of Consumer Reports presented the results of a large survey on the use and effectiveness of alternative medicines. Canadian and American subscribers to the magazine were sent a questionnaire and 46,806 responded. The article stated that a separate validation study demonstrated that the respondents were representative of the readership of the magazine; what precisely was meant by this is not clear, but for the purposes of this exam let us assume that they are representative with respect to age, gender, ethnicity, social economic status, and geography. Many common complaints were covered, including back pain. 20% of respondents reported that meditation helped them feel much better and 12% reported that over-the-counter drugs helped them feel much better.

   True or False and explain briefly: From these data one can conclude that meditation is more likely to help Consumer Reports readers suffering from back pain to feel much better than are over the counter drugs.

7. Statistics on the salary and education of supervisors in a large company were as follows:

Education:	Average = 14 years	SD = 3 years
Salary:	Average = $40,000	SD = $21,000
	Correlation = .40	

(a) [4] Find the regression equation for predicting salary from years of education.
(b) [2] Use the equation to predict the salary of a supervisor with 12 years of education.
(c) [2] A supervisor with 12 years of education decides to go back to school for 2 more years. Can you predict what his salary will be after that? If so, what is the prediction? If not, explain.

8. [4] A company that manufactures wire takes a simple random sample of its products and constructs a 95% confidence interval for the average tensile strength. A customer plans to purchase a roll of wire. You may assume that the roll to be purchased can be regarded as a random draw from the company's output. True or False and explain: the chance is 95% that the tensile strength of the wire on that roll will fall in the interval.

9. Two independent surveys plan to take simple random samples of a population in order to estimate the percentage of adults covered by medical insurance. They both plan to make 95% confidence intervals.

   (a) [2] What is the chance that the population percentage is contained in both intervals?

   (b) [2] True or False and explain briefly: If the surveys both have the same sample size, the intervals will have the same width.

   (c) [2] True or False and explain briefly: If both surveys have the same sample size combining their results will produce a confidence interval that is about half as wide as either of the confidence intervals based on the individual survey.

10. The Multiple Risk Intervention Trial tested the effects of intervention to reduce certain risk factors for coronary heart disease. The subjects were 12,866 men aged 35-57 at high risk for heart disease. 6428 were randomly assigned to be in the intervention group and 6438 were randomly assigned to be in the control group. The intervention included counseling on smoking.

(a) [4] On entry into the study, 59.3% of the intervention group were smoking and 59% of the control group were. If it is possible, find if this difference is statistically significant. If it is not possible, explain why not. Why might the investigators care?

(b) [4] After 6 years, 32.3% of the intervention group was smoking and 45.6% of the control group. If is possible, find if this difference is statistically significant. If it is not possible, explain why not.

11. [4] In the US there are two sources of statistics on national crime rates: (1) the FBI's Uniform Crime Reporting Program, which publishes summaries on all crimes reported to police agencies in jurisdictions covering virtually 100% of the population, and (2) the National Crime Survey, based on interviews with a nationwide sample of households. For purposes of this problem, you may assume that this is a simple random sample of 50,000 households.

    In 1992, 4.9% of the households in the sample reported at least one burglary during the last 12 months, and the FBI reported a rate of 32 per 1000 households. Can this difference be explained as chance error? If not, how would you explain it?

12. [6] In 1991 a study was done to assess possible effects of a new Virginia state law requiring the use of seat belts. The numbers below have been changed slightly for ease of calculation. Historical data for the treatment of the driver in a reported accident were as follows:

Treatment	None	Treated and released	Admitted to hospital	Died
percentage	50%	40%	8%	2%

A random sample of 500 accidents was taken the year after the seat belt law went into effect, with the following results:

Treatment	None	Treated and released	Admitted to hospital	Died
Number	300	165	30	5

Is there a statistically significant change relative to historical percentages?

Practice Final Solutions

1. It is possible. The histograms are skewed. The influx of poor immigrants lowered the median, while an increase in the net worth of the highest groups increased the average.

2. 1/2

3. (13/52)(12/51)(39/50)(38/49)(37/48)

4. 4950 and 5050

5. There is an error in the report. The RMS error can not be larger than the SD

6. This is not a random sample of readers. There is possible non-response bias. Cannot make this conclusion.

7. a. salary = 2800 x education + 800
b. $34,400
c. No, this is an observational study and it cannot be used to predict the effect of an intervention.

8. False - the interval is for the average, not for the population histogram.

9. a. .95 x .95
b. False. The estimated SEs will be different
c. False. The width of the confidence interval depends on the square root of the sample size.

10. a. It doesn't make sense to do a hypothesis test. Since the two groups were determined by randomization, any initial differences could only be due to chance.
b. z = 15.6. Statistically significant

11. z = 17. Could not be explained by chance. Not all burglaries were reported to police.

12. Chi-square = 21.1 with 3 degrees of freedom. P-value less than .01. There is a statistically significant change.