

# 2009 FINAL: PRACTICE PROBLEMS

- Suppose that IQ's in the general population are normally distributed with mean 100 and SD = 10.
  - What is the probability that a randomly chosen person has IQ above 125?
  - What is the probability that a randomly chosen person has IQ between 85 and 120?
- A class had two tests, with the following summaries:  
Test 1: mean = 350, sd = 10      Test 2: mean = 23, sd = 1.5  
Mindy scored 400 on test 1, and 27 on Test 2. On which test did she do better relative to the class?
- Consider the following game. A player bets \$2. Two 6-sided dice are rolled. If the total is 2-4, the player wins \$3. If the total is 5-6, the player loses \$1 of the bet. If the total is 7, the player neither wins nor loses anything. If the total is 8-9, the player loses his entire bet. If 10-12 is rolled the player wins half the total of the dice.
  - What are the expected winnings of this game?
  - Does this game favor the player or the house?
- The weight of the eggs produced by a certain breed of hen is normally distributed with mean 62g and standard deviation  $\sigma = 4.5g$ .
  - What is the probability that a single egg randomly chosen weighs less than 60g?
  - What is the probability that the mean weight of 9 eggs randomly chosen is less than 60g?
- A sample of 100 teachers found that the average time that they had worked for the board of education is 12 years. Assume that we know the standard deviation of the population of teachers' service is 5 years.
  - Construct a 95% confidence interval for the mean terms of service for the teacher population in this school district.
  - The principal at Sorrentino Memorial High School states that the average teacher in his school has 15 years of service with a standard deviation of 7 years. The principal wishes to claim that he has a more-experienced staff (based on length of service not quality of teaching) than the other schools in the district. Can he make this claim? Why or why not? Assume that there are 30 teachers at SMHS.

6. A reporter wishes to conduct a survey to find the proportion of teenagers who are Chuck Norris fans. If the reporter wants to be 95% confident that his proportion is within 1% of the true proportion, how many teenagers must he interview?
7. A test was given to a random sample of 750 subjects. Their mean score was 318.5 and the standard deviation of the individual scores was 55.
- a) Compute the 95% confidence interval for the population mean score.
- b) Suppose  $H_0$  is that the population mean is 275, and  $H_1$  is that the population mean is not 275. What is the p-value based on the data? Do you reject  $H_0$ ?
8. A researcher wants to test the effectiveness of a new strength supplement. A random sample of 6 individuals was chosen and the amount of weight they could bench press was measured before and after they took the supplement for 6 weeks.

<u>Name</u>	<u>Before Using Supplement</u>	<u>After Using Supplement</u>
Mary	95	105
Margaret	65	70
Millie	70	75
Mollie	45	55
Maxi	90	90
Matilda	50	65

Test the claim that the supplement does not affect the weight that the individuals can bench press using an  $\alpha = 0.05$  test.

9. A manufacturer is testing the accuracy of its chlorine probes. The probes are placed in what is known to be water with a chlorine content of 2.00 ppm. The following are the readings that the probes it tested gave.

2.20 1.85 2.15 1.90 2.22 2.10 1.95 2.00 2.30 1.88  
2.15 2.22 1.87 1.92 1.95 2.20 2.05 2.12 1.95 2.15

Are these detectors accurate? Test an appropriate claim to answer this question.

10. In a study conducted to see if cancer incidence and geographic location are related, the following data was collected.

	<b>Leukemia</b>	<b>Melanoma</b>	<b>Other</b>
<b>Urban</b>	120	58	150
<b>Suburban</b>	85	100	60
<b>Rural</b>	65	105	85

- State the null hypothesis and the alternative hypothesis.
- Calculate all expected values.
- Calculate the chi square statistic.
- What are the appropriate degrees of freedom for the test?
- What is the p-value for this test?
- What do you conclude from the data about the hypothesis?

11. Identify the following type of data as *quantitative* or *categorical*.

- family income
- parent's employer
- college major
- grade point average
- types of sporting events attended
- number of sporting events attended

12. The number of hits a ball player got over the past 10 years is described below.

80    110    96    125    130    86    142    78    98    115

- (a) What is the mean and standard deviation of the data above.
- (b) Draw a boxplot of the data.
- (c) What is the probability that the player will have at least 70 hits the next season?

13. When is it better to use the median rather than the mean for a given set of data?

14. What are some things a statistician needs to do to ensure that he is obtaining a random sample for his study? Why is this important?

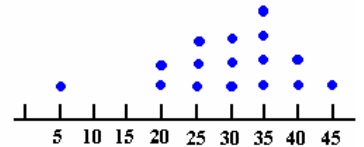
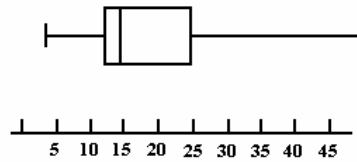
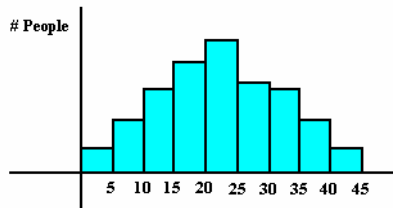
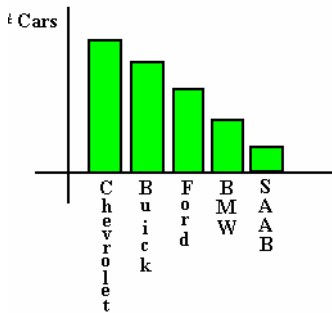
15. Match the two descriptions to the graphs below. (One from each category.)

(I)

(II)

(III)

(IV)



**Category #1:**  
**(distribution)**

- a) Skewed Left
- b) Skewed Right
- c) Approximately Symmetric
- d) No answer since the order of the  $x$ -axis data has no meaning.

**Category #2:**  
**(Type of chart)**

- a) Boxplot
- b) Pareto Chart
- c) Histogram
- d) Dotplot

16. The weights (in pounds) of 30 newborn babies are listed below.

5.5 5.7 5.8 5.9 6.1 6.1 6.4 6.4 6.5 6.6 6.7 6.7 6.7 6.9 7.0  
7.0 7.0 7.1 7.2 7.2 7.4 7.5 7.7 7.7 7.8 8.0 8.1 8.1 8.3 8.7

(a) Find  $D_4$  and  $Q_3$

(b) What percentile corresponds to a weight of 7.5 pounds?

17. Suppose that IQs in the general population are normally distributed with mean  $\mu = 110$  and standard deviation  $\sigma = 20$ . Suppose that a person is chosen at random from this population.

(a) What is the probability that that person has an IQ above 140?

(b) What is the probability that that person has an IQ between 60 and 105?

18. Given the frequency table below.

Miles Driven (per day)	Number of People
1-5	23
6-10	28
11-15	37
16-20	21
21-25	12
26-30	8
31-35	5

What is the mean and standard deviation for the number of miles driven per day?

19. You are trying to decide if you should suggest that your company take a certain project. The company's policy is to only take a project that can earn at least \$5000. From the research you have done you have concluded the following. If you take the project there is a 10% chance of earning a \$20000 profit. There is a 25% chance that the company can earn \$10000 profit and a 20% chance of earning \$5000. There is a 15% chance of breaking even with this project. On the down side, there is a 10% chance of losing \$3000 and a 20% chance of the company losing \$9000 in this venture. Based on all of this, will you suggest that the company should take this project? Explain your answer.
20. At Sorrentino Memorial High School, the principal is trying to convince more students taking Calculus to take the University course for credit. Forty-two students are currently enrolled in Calculus.. Of this group, 17 are taking it for credit. These students have an average grade of 90.5 with a standard deviation of 8.66. The other 25 students do not want to take the course for college credit. These students' average grade is 86.0 with a standard deviation of 4.85. The principal seeing this wishes to claim that students who take the course for credit also score better grades in the course. Is this claim valid?
21. The following table is of data collected in some unknown high school

Number of hours Studied (per night)	Score on SAT	Score on Calculus Final
2	1500	64
2.5	1600	70
3	1480	74
3.5	1710	78
4	1610	82
4.5	1520	88
5	1560	87

- (a) Determine if there is a strong correlation between the number of hours a student studies per night and their SAT score.
- (b) Based on the results of (a), what should a person who studied 3.25 hours score on the SAT?
- (c) Determine if there is a strong correlation between a person's SAT score and what they earn on the Calculus final.
- (d) Based on the results of (c), what should a person who scored a 1660 on the SAT score on the Calculus final?
- (e) Determine a 95% confidence interval for the slope of the regression line in part (c). Do the results of this confidence interval agree with the results you obtained in part (c)?

22. At Perrysville Prep School, 20 students were randomly chosen and their IQ scores measured. The average IQ score at Perrysville Prep was 110 with a standard deviation of 9.4. If the national average IQ score is 100, test the claim at the  $\alpha = 0.01$  level that Perrysville Prep students have higher IQ scores than the nation.

23. Ten people take a certain SAT prep course. The results of which are as follows...

Before	1000	980	860	890	1030	1100	810	980	960	1060
After	1080	1020	990	890	1010	1140	890	1060	1080	1070

Did the SAT prep course improve these people's scores?

24. Automobile manufacturers want to test for a difference in gas mileages among midsize cars, pickup trucks, and SUVs. The summarized data is below.

<b>MIDSIZE</b>	33.6	20.7	25.9	26.5	25.2	31.9	28.5	30.5	23.3	24.4
<b>PICK-UP</b>	20.3	19.7	27.8	21.5	20.6	25.9				
<b>SUV</b>	16.1	21.8	15.0	16.5	16.2	22.6	21.5	19.9		

(a) Would it be appropriate to run an ANOVA test on this data? Explain your answer.

(b) Fill in the ANOVA table below.

SOURCE	SS	DF	MS	F
Group				
Error				
TOTAL			XXX	

(c) Is there a difference in gas mileages among vehicle types? Test a claim to answer this question.

25. In a study of the effects of electromagnetic radiation, researchers visited the homes of children in the Denver area who had died of cancer (Leukemia, Lymphoma, Other), and classified the wiring outside the building as low current (LC) or high current (HC) to get:

	Leukemia	Lymphoma	Other
HC	52	10	17
LC	84	21	31

Test the claim that there is no relation between contracting some kind of cancer and the type of wiring that exists outside the buildings.

26. A pregnancy test is being tested before being put on the market. Researchers set up 3 trial samples as follows:

*Group 1:* Blood from 8 pregnant women and 2 non-pregnant women was drawn.

*Group 2:* Blood from 5 pregnant women and 5 non-pregnant women was drawn.

*Group 3:* Blood from 3 pregnant women and 7 non-pregnant women was drawn.

The test gave the following number of positive readings (indicating the blood was from a pregnant woman):

*Group 1:* 9

*Group 2:* 6

*Group 3:* 2

How accurate is this test? Show all work and give all explanations.

**For #27-34, determine which hypothesis test is the best to run for the given situation.**

27. We find the mean and standard deviation of 100 teenagers to see if they drink, on average, more coffee than teenagers did 5 years ago.

- (a) z test about a proportion
- (b) z test about a mean with one-sided alternative
- (c) z test about a mean with two-sided alternative
- (d) t test about a mean with one-sided alternative
- (e) t test about a mean with two-sided alternative
- (f) two-sample t test with one-sided alternative
- (g) two-sample t test with two-sided alternative
- (h) matched pairs, one-sided alternative
- (i) matched pairs, two-sided alternative
- (j) chi square test
- (k) ANOVA
- (l) inference for regression

**28.** We want to see if the mean cost of Sony TV sets is the same as the mean cost of Panasonic TV sets.

- (a) z test about a proportion
- (b) z test about a mean with one-sided alternative
- (c) z test about a mean with two-sided alternative
- (d) t test about a mean with one-sided alternative
- (e) t test about a mean with two-sided alternative
- (f) two-sample t test with one-sided alternative
- (g) two-sample t test with two-sided alternative
- (h) matched pairs, one-sided alternative
- (i) matched pairs, two-sided alternative
- (j) chi square test
- (k) ANOVA
- (l) inference for regression

**29.** We want to see if there is a relationship between people's average yearly salary (in dollars) and the amount of vacation time they take (in days).

- (a) z test about a proportion
- (b) z test about a mean with one-sided alternative
- (c) z test about a mean with two-sided alternative
- (d) t test about a mean with one-sided alternative
- (e) t test about a mean with two-sided alternative
- (f) two-sample t test with one-sided alternative
- (g) two-sample t test with two-sided alternative
- (h) matched pairs, one-sided alternative
- (i) matched pairs, two-sided alternative
- (j) chi square test
- (k) ANOVA
- (l) inference for regression

**30.** We want to see if there is a relationship between the day of the week (Sunday-Saturday) and the weather (sunny, rain, snow, etc.)

- (a) z test about a proportion
- (b) z test about a mean with one-sided alternative
- (c) z test about a mean with two-sided alternative
- (d) t test about a mean with one-sided alternative
- (e) t test about a mean with two-sided alternative
- (f) two-sample t test with one-sided alternative
- (g) two-sample t test with two-sided alternative
- (h) matched pairs, one-sided alternative
- (i) matched pairs, two-sided alternative
- (j) chi square test
- (k) ANOVA
- (l) inference for regression

**31.** We test to see if the average amount of money a teacher earns is less than the average salary of American professionals. The standard deviation for the mean salary of American professionals is \$2500.

- (a) z test about a proportion
- (b) z test about a mean with one-sided alternative
- (c) z test about a mean with two-sided alternative
- (d) t test about a mean with one-sided alternative
- (e) t test about a mean with two-sided alternative
- (f) two-sample t test with one-sided alternative
- (g) two-sample t test with two-sided alternative
- (h) matched pairs, one-sided alternative
- (i) matched pairs, two-sided alternative
- (j) chi square test
- (k) ANOVA
- (l) inference for regression

**32.** We take random samples of elderly adults, adults, teenagers, and children to determine if they have the same mean hours of sleep.

- (a) z test about a proportion
- (b) z test about a mean with one-sided alternative
- (c) z test about a mean with two-sided alternative
- (d) t test about a mean with one-sided alternative
- (e) t test about a mean with two-sided alternative
- (f) two-sample t test with one-sided alternative
- (g) two-sample t test with two-sided alternative
- (h) matched pairs, one-sided alternative
- (i) matched pairs, two-sided alternative
- (j) chi square test
- (k) ANOVA
- (l) inference for regression

**33.** A study is done to see if there is a relationship between a person's weight (in pounds) and blood pressure (in mmHg)

- (a) z test about a proportion
- (b) z test about a mean with one-sided alternative
- (c) z test about a mean with two-sided alternative
- (d) t test about a mean with one-sided alternative
- (e) t test about a mean with two-sided alternative
- (f) two-sample t test with one-sided alternative
- (g) two-sample t test with two-sided alternative
- (h) matched pairs, one-sided alternative
- (i) matched pairs, two-sided alternative
- (j) chi square test
- (k) ANOVA
- (l) inference for regression

34. We want to see if, on average, if people make more mistakes when driving after taking ibuprofen.

- (a) z test about a proportion
- (b) z test about a mean with one-sided alternative
- (c) z test about a mean with two-sided alternative
- (d) t test about a mean with one-sided alternative
- (e) t test about a mean with two-sided alternative
- (f) two-sample t test with one-sided alternative
- (g) two-sample t test with two-sided alternative
- (h) matched pairs, one-sided alternative
- (i) matched pairs, two-sided alternative
- (j) chi square test
- (k) ANOVA
- (l) inference for regression

35. At the 0.025 significance level, test the claim that the three brands have the same mean if the following sample results have been obtained.

<b>Brand A</b>	<b>Brand B</b>	<b>Brand C</b>
32	27	22
34	24	25
37	33	32
33	30	22
36		21
39		

36. A manager records the production output of three employees who each work on three different machines for three different days. The sample results are given below and he Minitab results follow.

Machine	Employee		
	<u>I</u>	<u>A</u>	<u>B</u>
<u>I</u>	16, 18, 19	15, 17, 20	14, 18, 16
<u>II</u>	20, 27, 29	25, 28, 27	29, 28, 26
<u>III</u>	15, 18, 17	16, 16, 19	13, 17, 16

### STATDISK ANOVA RESULTS

Source:	DF:	SS:	MS:	Test Stat, F:	Critical F:	P-Value:
Interaction:	4	15.4815	3.8704	0.71	2.9277	0.5981
Row Variable:	2	588.7407	294.3704	53.7027	3.5546	0.0000
Column Variable:	2	2.0741	1.037	0.1892	3.5546	0.8293

Does the machine type, the individual employee, and/or the interaction between the two have an effect on output? Explain.

37. Below are the penalties that occurred during the 2008 Stanley Cup finals last year.

	HOOKING	CROSS-CHECK	TRIPPING	INTERFERENCE
PENGUINS	5	3	5	9
REDWINGS	4	12	6	2

- Calculate the expected values for the penalties that each team incurred.
- Calculate the chi-square statistic for the data.
- Is there a relationship between the type of penalty incurred and the hockey team?
- Assume that you are testing the claim that the Redwings had a higher proportion of cross-check penalties than the Penguins with  $\alpha = 0.05$ , would the results of the chi-square test assist you in testing this claim?
- Assume that you are testing the claim that the Redwings had a lower proportion of interference penalties than the Penguins with  $\alpha = 0.01$ , would the results of the chi-square test assist you in testing this claim?