MATH 210 - Final Exam Review
(updated 5/11/03)

Observing and Describing Data (Chapter 1)

1. Construct and interpret a relative frequency histogram. Use such a histogram to answer questions about percentiles, quartiles, median.
2. Compare/contrast the meaning of mean and median. When is it more appropriate to use the median?
3. Given the formula for standard deviation, explain how it captures the amount of variation within a set of data.
4. What does "resistant measure" mean? Which of the following are resistant measures and why: mean, standard deviation, median, IQR?

5. Normal Distributions (Section 1.3)

    a. What are the specific characteristics of any normal distribution?
    b. What types of data are often approximated by a normal distribution?
    c. Given the mean and standard deviation, compute the z-score for a specific data value. Describe in words what the z-score represents.
    d. Compute normal probabilities and percentiles.

Sampling Methods (Chapter 3)

1. Understand the following terms and be able to relate them to one another: population, sampling frame, sample.
2. Identify sources of possible bias in a given sampling scenario. Also explain why the bias is likely to exist and in what direction (either high or low) the results will likely be biased.
3. What is a "simple random sample?" How does one go about obtaining such a sample from a specified sampling frame? Why is it important to choose a random sample?
4. Describe how you might set up an experiment to determine if taking a daily dose of aspirin will reduce the incidence of heart attacks in a given population.

Probability (Chapters 4 and 5)

1. What does "probability" mean?
2. What is the "Law of Large Numbers?"
3. Define "independence" in general and also in the context of a two-way classification table.

4. Binomial Experiments (Section 5.1)

    a. Recognize a binomial experiment and know how to determine corresponding probabilities using the exact method (Table C) and normal approximation (Table A).
    b. Describe how a YES/NO polling situation can be viewed as a binomial experiment.

5. Central Limit Theorem (Section 5.2)

   a. Under what conditions can this theorem be applied?
   b. What is the conclusion of the theorem?
   c. Apply the Central Limit Theorem. For example: Suppose random variable X has $\mu=100$ and $\sigma=10$. For a sample of size 50 find:

$$P(\bar{x} > 102) \quad \text{and} \quad P(99 < \bar{x} < 101)$$

Inferential Statistics (Chapters 6, 7 and 8)

1. Perform hypothesis tests and construct confidence intervals for the following. **Formulas will be provided.** You should also know the assumptions that must be met for using each of these techniques.

a.	Single Population Mean (σ unknown)	t-test	(7.1)
b.	Difference in Two Population Means	t-test	(7.2)
c.	Single Population Proportion	z-test	(8.1)
d.	Difference in Two Population Proportions	z-test	(8.2)
e.	Multinomial Test (for a probability distribution) (χ^2-test for independence is *not* included!)	χ^2-test	(HB pp. 8.7–8.9)

   Also know how to set up and analyze matched pairs and randomized-comparative experiments, using the appropriate statistical techniques, either (a) or (b).

2. Explain what a Type I error is and how your choice of α relates to this error.
3. For any hypothesis test, explain in generic terms how you should determine what will be the null hypothesis and what will be the alternative hypothesis.
4. For any hypothesis test, describe the circumstances in which you reject the null hypothesis. (Explain both in terms of the sample data and in terms of the P-value.)
5. If there is not sufficient evidence to accept the alternative hypothesis what should we conclude about the null hypothesis?

Linear Regression (Sections 2.1–2.4, Handbook Ch. 2)

1. In a scatterplot describe any apparent association (positive vs. negative, strong vs. weak, linear vs. non-linear).
2. What additional computer output should be checked to support an apparently linear association you observe from the scatterplot?
3. Conduct the correct *hypothesis test* to determine if the LSR equation calculated from your sample data provides evidence of a relationship in the broader population of interest.
4. Use the LSR equation to predict a y value for a given value of x. Also interpret *prediction intervals* and *confidence intervals* for individual and mean outcomes.
5. The accuracy of prediction and confidence intervals depends on what assumptions about the residuals? What plot should be checked to verify this?
6. Does a strong correlation between x and y necessarily imply there is a cause-and-effect relationship between these variables? Explain.
7. In least-squares regression, what are the interpretations of b, r and r^2?