PROBABILITY PART.

1. For each of the following experiments and events determine the probability of the event. A deck of cards has 52 cards with numbers 1, 1, 2, 2, 3, 3, 4, 4.

(a) (4 pts.) Experiment: Three cards are drawn randomly without replacement. Event: At least two of the cards chosen have numbers larger than two.

(b) (5 pts.) Experiment: Cards are drawn randomly without replacement until a card with the number 4 is chosen. Event: The first 4 is chosen on the 5th draw.

(c) (6 pts.) Experiment: 100 cards are drawn randomly with replacement. Event: The sum of the numbers on the 100 cards is larger than 270. (Use an approximation).

(d) (7 pts.) Experiment: Same as in (c). Event: The number of 4's chosen is less than 20. (Use an approximation).

(e) (8 pts.) Experiment: Two cards are drawn randomly without replacement. This is repeated independently 84 times. Event: Two or more of the 84 trials result in 4's. (Both cards have 4's.) (Use an approximation)

2. Let $f(x) = 1.5x^2$ for $-1 \leq x \leq 1$. Let $X_1$ and $X_2$ be a random sample from this density.

(a) (7 pts.) Find the cdf and the density for $Y = X_1^2$.

(b) (5 pts.) Let $M = \max(X_1, X_2)$. Find the cdf for $M$.

(c) (6 pts.) Find the correlation coefficient $\rho(X_1, X_1 + X_2)$.

(d) (7 pts.) Let $W = X_1X_2$. Find $E(W)$ and $\text{Var}(W)$.

3. (a) (5 pts.) State the Weak Law of Large Numbers.

(b) (10 pts.) Use the Chebyshev's Inequality to prove the Weak Law of Large Numbers.

4. Let $X$ be a random variable with the density $f(x) = 2x$ for $0 \leq x \leq 1$. Conditionally on $X = x$, let $Y$ be uniformly distributed on $[0, x]$.

(a) (7 pts.) Find the marginal density of $Y$.

(b) (6 pts.) Given a random variable $U$, uniformly distributed on $[0, 1]$, how could a random variable $X$ be generated from $U$?

(c) (7 pts.) Find $Z = E(Y|X)$.

(d) (10 pts.) Find the joint cdf of $X$ and $Y$. 
5. Consider the reliability diagram below. The system has 5 components numbered 1, 2, 3, 4, 5. Let $A_k$ be the event that component $k$ works as it should. Suppose that $A_1, \ldots, A_5$ are independent. Denote $p_k = P(A_k)$ for each $k$.

![Reliability Diagram]

Figure 1: For problem 5.

(a) (5 pts.) Express the event $B$ that the entire system works successfully in terms of $A_k$.

(b) (5 pts.) Express $P(B)$ in terms of $p_k$.

(c) (5 pts.) Express $P(A_2|B)$ in terms of $p_k$.

STATISTICS PART.

6. Let $\lambda > 0$ be a unknown parameter. Let $X_1, X_2$ be a random sample from the distribution with density $f(x) = \lambda x e^{-\lambda x^2/2}$ for $x \geq 0$.

(a) (10 pts.) Find the maximum likelihood estimator of $\lambda$.

(b) (10 pts.) Find the method of moments estimator of $\lambda$.

7. The following data correspond to a weight loss for 7 people. Each person was put on a low-carb diet and a high-carb diet. The order of diets was determined randomly for each person.

| Low-carb diet | 1.3 | 1.0 | 2.1 | 2.0 | 1.4 | 2.6 | 2.7 |
| High-carb diet | 1.0 | 1.3 | 1.1 | 1.0 | 2.8 | 0.7 | 2.3 |

(a) (12 pts.) State a parametric model, define hypotheses, and carry out a test at level $\alpha = 0.05$ which will enable you to decide whether the low-carb diet is more effective than the high-carb diet for the weight loss.

(b) (8 pts.) Perform a nonparametric test of the same hypotheses. Find the $p$-value for this test.
8. In order to estimate the change $\Delta = p_1 - p_2$ in the proportion of a population of 45 thousand MSU students who wanted a decrease in tuition between time $t_1$ and time $t_2$, random samples of sizes $n_1 = 500$ and $n_2 = 400$ were taken independently at the two times. Let $X_1$ and $X_2$ be the numbers favoring a decrease in tuition for the two samples.

(a) (8 pts.) Define notation and give a formula for a 90% confidence interval on $\Delta$.

(b) (5 pts.) Apply the formula for the case $X_1 = 250, X_2 = 320$.

(c) (7 pts.) Explain the meaning of your interval in such a way that someone who had not ever studied statistics would understand.

(d) (8 pts.) If equal sample sizes, $n = n_1 = n_2$ were to be used how large must $n$ be in order that the estimator $\hat{\Delta}$ have probability at least 0.95 of being within 0.02 of $\Delta$?

9. Let $(X_1, X_2)$ be a random sample from the geometric distribution with unknown parameter $0 < p < 1$. Suppose that we wish to test $H_0 : p = 0.5$ vs $H_a : p < 0.5$.

(a) (5 pts.) Consider the test which rejects for $T = X_1 + X_2 > 12$. What is the level of significance $\alpha$?

(b) (10 pts.) Use a theorem to prove that this test is uniformly most powerful for this $\alpha$ level.

(c) (7 pts.) Find the power of this test for $p = 0.25$.

10. Let $(x_i, Y_i)$ be observed for $i = 1, \ldots, n$. Suppose that the $x_i$ are constants, and that $Y_i = \beta x_i + \epsilon_i$, where $\beta$ is an unknown parameter, and the $\epsilon_i$ are independent, each with the $N(0, \sigma^2)$ distribution.

(a) (8 pts.) Show that the least squares estimator of $\beta$ is $\hat{\beta} = (\sum_{i=1}^{n} x_i Y_i) / (\sum_{i=1}^{n} x_i^2)$. Do not use matrix or vector space methods.

(b) (5 pts.) Show that $\hat{\beta}$ is an unbiased estimator of $\beta$.

(c) (5 pts.) Find $\text{Var}(\hat{\beta})$.

(d) (8 pts.) For the following $(x_i, Y_i)$ pairs find a 95% confidence interval on $\beta$: $(1, 3), (2, 5), (3, 8), (4, 8)$.

11. (7 pts.) Let $X_1, \ldots, X_n$ be independent random variables, each with mean $\mu$ and variance $\sigma^2$. Consider the two estimators $\mu^* = \frac{1}{2^n-1} \sum_{i=1}^{n} 2^{n-i} X_i$ and $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$.

(a) Show that $\mu^*$ is an unbiased estimator of $\mu$.

(b) Determine the relative efficiency (relative size of variances) of $\mu^*$ to $\bar{X}$ and find its limiting value as $n \to \infty$. 