CENTRAL MICHIGAN UNIVERSITY  
Department of Mathematics 

Ph.D. QUALIFYING EXAMINATION

STATISTICS  
Time: 180 minutes  
June 19, 1997

General Instructions:
There are two parts to the Exam: Part A consists of 10 questions from which you are to answer 8 and Part B consists of 8 questions from which you are to answer 6.

Part A
Please give complete solutions to eight of the following ten questions. Each question will receive equal weight when graded. Indicate which questions you wish to have graded on your answer sheet.

1. Sixty percent of new drivers have had driver education. During their first year, new drivers without driver education have a probability .08 of having an accident, but new drivers with driver education have only a .05 probability of having an accident. What is the probability a new driver has had driver education, given that the driver had no accidents during the first year?

2. An urn contains n white chips numbered 1 through n, n green chips numbered 1 through n, and n red chips numbered 1 through n. If two chips are to be drawn at random without replacement from the urn, what is the probability that both chips will be of the same color or bear the same number?

3. Let X be a random variable with cumulative distribution function

\[ F(x) = \begin{cases} 
0 & \text{for } x \leq 0 \\
1 - e^{-x} & \text{for } x > 0.
\end{cases} \]

(a) What is \( P \left( 0 \leq e^X \leq 4 \right) \)?
(b) Determine the moment-generating function of \( X \).

4. If the joint probability density function of \( X \) and \( Y \) is given by

\[ f(x, y) = \begin{cases} 
24xy & \text{for } 0 < x < 1, 0 < y < 1, x + y < 1 \\
0 & \text{elsewhere},
\end{cases} \]

find \( P \left( X < \frac{1}{2} \mid Y < \frac{1}{2} \right) \).

5. Let \( X \) and \( Y \) be independent random variables, each having an exponential distribution with parameter \( \theta = 1 \). Determine the probability that \( X + Y > 2 \).
6. Suppose the length of the western rattlesnake is normally distributed with a mean of 40 inches and a standard deviation of 2 inches. If five western rattlesnakes are collected, what is the probability that at least four are shorter than 41 inches? (Assume that their lengths are independent.)

7. Let \( X \) be the number of the trial on which the first success occurs when a Bernoulli trial with success probability \( \theta \) is repeated independently.
   (a) Derive the moment-generating function of \( X \).
   (b) Use the moment-generating function to determine the expected value of \( X \).

8. If the joint probability density of \( X \) and \( Y \) is given by
   
   \[
   f(x, y) = \begin{cases} 
   x + y & \text{for } 0 < x < 1, 0 < y < 1 \\
   0 & \text{elsewhere},
   \end{cases}
   \]
   
   determine the variance of \( W = 3X - 5Y + 10 \).

9. The maximum passenger load that a ferry can carry is 16,375 pounds. Suppose on any trip
   the ferry carries 100 passengers and the distribution of the weight of the passengers has a
   mean of 160 pounds and a standard deviation of 15 pounds. Assume that the weights of
   the passengers can be considered independent. Use Chebyshev's Theorem to obtain an upper
   bound for the probability that the ferry will exceed the limit on a trip.

10. A system has two components placed in series so that the system fails if either of the two
    components fails. The second component is twice as likely to fail as the first. If the two
    components operate independently, and if the probability that the entire system fails is 0.28,
    what is the probability that the first component fails?

Part B

Please give complete solutions to six [3 questions from 11 - 14 and 3 questions from 15 - 18] of
the following eight questions. Each question will receive equal weight when graded. Begin
each question on a separate sheet. Indicate which questions you wish to have graded on
your answer sheet.

11. \( X \) is a binomial random variable with parameters \( n \) and \( p \). Define \( Y = X/n \)
   (a) Show that \( P \left( \left| Y - p \right| \geq \delta \right) \leq \frac{1}{4n\delta^2} \).
   (b) Suppose we wanted to find a value of \( n \) such that
       \( P(Y - p < \delta) \geq .9 \), show that \( n \geq \frac{2.5}{\delta^2} \).
   (c) Show that when \( n \) is large, \( P(Y - p < \delta) = \Phi(2\delta\sqrt{n}) - 1 \),
       where \( \Phi \) is the cumulative distribution function of the standardized normal random variable.
12. \(X_1, X_2,\) and \(X_3\) are three independent \(\chi^2\) variables with degrees of freedom \(d_1, d_2, d_3,\) respectively.

(a) Show \(Y = X_1 + X_2 + X_3 \sim \chi^2(d_1 + d_2 + d_3)\)

(b) Show that \(Y = \frac{X_1}{X_3}\) and \(W = X_1 + X_3\) are independent

(c) Show that \(\frac{X_1}{d_1} \quad \text{and} \quad \frac{X_2}{d_2} \quad \text{are two independent F variables.} \frac{(X_1 + X_3)/(d_1 + d_3)}{d_3}\)

13. A general class of distribution, called the exponential family of distributions, has the form:
\[
f(x) = \exp\{A(\theta)B(x) + C(x) + D(\theta)\},
\]
where \(A, B, C,\) and \(D\) are arbitrary functions of their arguments. Normal, Binomial, Poisson and many other distributions are in this family. Find the appropriate functions of \(A(\theta), B(x), C(x), D(\theta)\) for each of the following distributions:

(a) Normal Distribution
(b) Binomial Distribution
(c) Poisson Distribution.

14. Correlation coefficient \(\rho\), between two standardized random variables \(X\) and \(Y\) is defined:
\[
\rho = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)\text{var}(Y)}}
\]

(a) Show that \(|\rho| \leq 1\)

(b) Show that \(\rho = 1 - \frac{\sigma^2_{X-Y}}{2}\), where \(\sigma^2_{X-Y}\) is the variance of the difference, \(X - Y\).

(c) Show that \(\rho = \frac{\sigma^2_{X+Y}}{2} - 1\), where \(\sigma^2_{X+Y}\) is the variance of the sum, \(X + Y\).

15. A group claims that the mean length of time consumers keep their car is less than 7.5 years. In order to test the claim, a random sample of size 36 is taken to obtain a mean \(\bar{x} = 6.8\) years and a standard deviation \(s = 3.2\) years.

(a) State the null and the alternative hypotheses for this problem.
(b) State the test statistic.
(c) Find the rejection region for \(\alpha = 0.025\).
(d) Compute the observed value of the test statistic and state your decision.
(e) State an appropriate conclusion in terms of the original problem.
(f) What type of error are you liable to in (d) and why?

16. A random sample of size \(n\) is taken from a population with probability density function
\[
f(x, \theta) = \begin{cases} 
(\theta + 1)x^\theta, & 0 < x < 1 \\
0, & \text{elsewhere},
\end{cases}
\]
where parameter \(\theta > 0\).

(a) When is an estimator \(T\) consistent for a parameter \(\theta\) ?
(b) Find the moment estimator of \(\theta\) and show that this estimator is consistent.
(c) Find the maximum likelihood estimator of \(\theta\).
(d) Find a sufficient statistic for parameter \(\theta\).
17. Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution that has probability density function

$$f(x, \theta) = \begin{cases} \frac{xe^{-x/\theta}}{\theta^2}, & 0 < x < \infty \\ 0, & \text{otherwise,} \end{cases}$$

where $0 < \theta < \infty$. Let $y = \sum_{i=1}^{n} x_i$ and let the loss function be $L(\theta, \omega(y)) = (\theta - \omega(y))^2$.

The decision function $\omega(y) = ky$ is such that $k$ does not depend on $y$.

(a) Prove that the risk function is given by $R(\theta, \omega) = \theta^2[1 - 4nk + 2n(2n + 1)k^2]$.

(b) Find the value of $k$ for which the risk is minimum.

(c) Is the decision function $\omega(y) = ky$ biased or unbiased? If the decision function is biased for $\theta$, find a function of $y$ that is unbiased for $\theta$.

(d) On using the value of $k$ in (b), determine $\max \{ R(\theta, \omega) \}$ if it exists.

18. Suppose $X_1$ and $X_2$ constitute a random sample of size 2 from the population with probability density function

$$f(x, \theta) = \begin{cases} \theta x^{\theta-1}, & 0 < x < 1 \\ 0, & \text{elsewhere,} \end{cases}$$

Suppose the critical region $X_1X_2 \geq 3/4$ is used to test the null hypothesis $H_0$: $\theta = 2$ against the alternative hypothesis $H_1$: $\theta = 3$.

(a) What is the significance level of the test?

(b) What is the power of the test at $\theta = 3$?

(c) Compute the probability of type II error when $\theta = 3$. 