Final Exam

You have 3 hour to complete this open book exam of 160 points. Be concise and, above all, clear. Be sure that your name is printed on your book, and that you are neat. You must show all work to get full credit. Best of luck.

Question 1 (20 points)

Are the following statements true or false? Explain your reasoning.

1. The maximum likelihood estimator is efficient because its asymptotic variance is equal to the Cramer-Rao lower bound.

2. Because information matrix equality holds under correct model specification, one should always use the information matrix equality to simplify the asymptotic variance formula for the maximum likelihood estimator. This will ensure that the model is correctly specified.

3. Let \( X_i, i = 1, \ldots, n \) be i.i.d normal with mean \( \mu \) and variance \( \sigma^2 \). A researcher uses the statistic \( T = \sqrt{n} \left( \bar{X} - \mu \right) / \hat{\sigma} \), where \( \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2 \) is the sample variance. The researcher uses a \( t \)-table with \( n-1 \) degree of freedom to look up the critical values. This is wrong because the researcher does not correct for the degree of freedom in \( \hat{\sigma}^2 \).

4. For an i.i.d sample \( X_i \sim N(\mu, \sigma) \), the sample mean of the entire sample is a better estimator of \( \mu \) than the sample mean of the first half of the sample.

5. It is more common to use 95% confidence intervals than 99.9% confidence intervals because people are usual 95% confident instead of 99.9% confident.

Question 2 (20 points)

Suppose you decide to choose the size of your test as a function of the sample size, before knowing the sample size and the data. Consider three choices (1) \( \alpha = 0.05 \); (2) \( \alpha =
$1 - \Phi \left( \frac{n}{3} \right)$, where $\Phi (\cdot)$ is the standard normal CDF; (3) $\alpha = 1 - \Phi (n)$. Which of these three rules will you prefer? Please explain your reasoning. (Assume that you believe in asymptotic theory.)

Question 3 (40 points)
Let $X_i$ be a multinomial distribution such that $X_i = 1$ with probability $p_1$, $X_i = 2$ with probability $p_2$, $X_i = 3$ with probability $p_3$, $X_i = 4$ with probability $p_4$, and $X_i = 5$ with probability $1 - p_1 - p_2 - p_3 - p_4$, where $0 < p_1 + p_2 + p_3 + p_4 < 1$.

1. Define a parameter $\theta = p_1 + p_2^2$. Given a sample of size $n$ of i.i.d observations $X_i, i = 1, \ldots, n$, is there a uniformly minimum variance unbiased estimator for $\theta$? If there is derive it. If there isn’t explain your reasoning.

2. You are interested in testing $H_0 : \theta = 0.5, p_3 = p_4$, where against the alternative hypothesis that the null hypothesis is not true. Derive a Wald test statistic and its asymptotic distribution.

3. Describe a Likelihood Ratio test statistic and its asymptotic distribution.

4. For $n = 20$ and sample of $X_i$ as $1, 2, 3, 4, 5, 2, 3, 5, 2, 1, 4, 2, 3, 1, 2, 4, 2, 3, 1, 2, 4, 2, 3, 1, 2$, calculate the P-value of the Wald test.

5. Calculate the power of your Wald test when $p_1 = p_2 = p_3 = 0.2$ and $p_4 = 0.1$.

Question 4 (40 points)
Assume $Y_i, i = 1, \ldots, n$ be i.i.d. distributed according to the normal distribution $N (\mu, \sigma^2)$, where both $\mu$ and $\sigma^2$ are unknown parameters.

1. Explain how you can construct an exact 95% confidence interval for the variance parameter $\sigma^2$. 
2. For this question only, suppose the observations \( Y_i \) i.i.d with mean \( \mu \) and variance \( \sigma^2 \), but they are distributed as gamma random variables instead of as normal random variables, so that \( Y_i > 0 \) and \( f(y) = \frac{1}{\Gamma(k)\theta^k} y^{k-1} e^{-\frac{y}{\theta}} \) for \( \theta = \frac{\sigma^2}{\mu} \) and \( k = \frac{\mu^2}{\sigma^2} \). Is the exact finite sample confidence interval you constructed in (1) still valid in finite sample? Is it valid asymptotically? If they are, explain why. If they are not, explain whether you can come up with confidence intervals that are asymptotically valid, and whether you can come up with confidence intervals that are valid in finite samples.

3. If the observations \( Y_i \) no longer follow normal distributions, but they are still i.i.d with mean \( \mu \) and variance \( \sigma^2 \). Furthermore, you also know that \( E(Y_i - \mu)^4 = 3\sigma^4 \). Is the exact finite sample confidence interval you constructed in (1) still valid in finite sample? Is it valid asymptotically?

Question 5 (40 points)

1. Suppose \( X_i \sim N(\mu, 3), Y_i \sim N(2\mu, 2) \) and \( Z_i \sim N(4\mu, 1) \), independent of each other. Given a sample of i.i.d. observations \( (X_i, Y_i, Z_i) \). Obtain the maximum likelihood estimator of \( \mu \) and derive its asymptotic distribution. Is this the best unbiased estimator in the sense of having the smallest variance?

2. Now suppose \( W_i, i = 1, \ldots, n \) are i.i.d normally distributed with unknown mean and variance \( (\mu, \mu^2) \) for \( \mu > 0 \). Derive the maximum likelihood estimator of \( \mu \) and its asymptotic distribution.

3. The researcher is interested in testing the null hypothesis of \( H_0 : \mu^2 = 3 \) against the alternative hypothesis of \( H_1 : \mu^2 > 3 \). Formulate a test statistics and derive its asymptotic distribution.

4. Compute the power function of your test, when the size is \( \alpha = 0.05 \), under the alternative hypothesis that \( H_1 : \mu^2 = 3 + c/\sqrt{n} \), as a function of \( c \).
Question 6 (20 points)

Let $X$ be uniformly distributed between 5 and $5+\theta$, for $\theta > 0$. You have only one observation of $X$ and you want to test $H_0 : \theta = 3$ against $H_1 : \theta \neq 3$. Derive a likelihood ratio test of size 30%. Is this test uniformly most powerful?