1. Consider the simple discrete distribution with probability mass function

\[
  f(x|p) = \begin{cases} 
  (1 - p)p^{x-1} & \text{for } x = 1, 2 \text{ and } 3 \\
  p^3 & \text{for } x = 4.
  \end{cases}
\]

This problem concerns hypothesis testing based on a single observation from this distribution.

a) Consider the test \( \phi(x) = I[x = 1 \text{ or } x = 4] \).
   i) Give the power function for this test.
   ii) Is this test an "unbiased" test of \( H_0: p = .5 \) vs \( H_a: p \neq .5 \)? Explain.

b) Find a most powerful size \( \alpha = .15 \) test of \( H_0: p = .1 \) vs \( H_a: p = .2 \).

c) For what \( \alpha \)'s in \([0, 1]\) are there nonrandomized most powerful size \( \alpha \) tests of the hypotheses in part b)?

d) There is a UMP size \( \alpha = .15 \) test of \( H_0: p \leq .1 \) vs \( H_a: p > .1 \). Find it and argue very carefully that it is indeed UMP size \( \alpha = .15 \).

e) Make the Bayesian assumption that \( a \) priori, \( p \) is Uniform \((0, 1)\). The test \( \phi(x) = I[x = 3 \text{ or } x = 4] \) is a Bayes test for \( H_0: p \leq .6 \) vs \( H_a: p > .6 \). Show that this is correct for \( x = 3 \) (i.e. that a Bayes test will reject \( H_0 \) if \( x = 3 \)).

f) What is the size of the test in e)? (Argue carefully that your "size" is correct.)

2. Suppose that \( X_1, X_2, \ldots, X_n \) are iid with the marginal distribution specified in question 1.

a) Identify a two-dimensional sufficient statistic here and argue carefully that it is indeed sufficient.

b) Name any unbiased estimator of \( p \in [0, 1] \) in this context.

c) Find the maximum likelihood estimator of \( p \in [0, 1] \).

3. Consider the model where \( X_1, X_2, \ldots, X_n \) are iid Uniform \([0, \theta]\) for \( \theta > 0 \). In this model, the largest order statistic \( Y_n = \max_{1 \leq i \leq n} X_i \) is sufficient. So if we wish to do hypothesis testing in this model, we need only consider tests based on \( Y_n \).

a) Consider the set of hypotheses \( H_0: \theta = \theta_0 \) vs \( H_a: \theta > \theta_0 \) and tests of the form

\[
  \phi_k(y) = I[y > k].
\]

Find \( k_{\theta_0} \) so that \( \phi_{k_{\theta_0}} \) is of size \( \alpha = .05 \).
b) Invert the family of tests \( \{ \phi_{h_0} \} \) and produce the corresponding 95% confidence procedure for estimating \( \theta \).

4. Suppose that for parameters \( p_1 \geq 0, p_2 \geq 0 \) with \( p_1 + p_2 \leq 1 \), the variables \( X_1, X_2, \ldots, X_{40} \) are iid with the simple marginal discrete distribution specified in the following table.

| \( x \) | \( f(x|p_1, p_2) \) | 1 | 2 | 3 |
|-------|-----------------|---|---|---|
|       |                 | \( p_1 \) | \( p_2 \) | \( 1 - p_1 - p_2 \) |

However, we are not furnished with the full data set \( X_1, X_2, \ldots, X_{40} \), but rather only values of the statistics:

\[
N_1 = \sum_{i=1}^{10} I(X_i = 1) \\
N_2 = \sum_{i=1}^{10} I(X_i = 2) \\
N_3 = \sum_{i=1}^{10} I(X_i = 3)
\]

\[
N_{12} = \sum_{i=11}^{20} I(X_i = 1 \text{ or } X_i = 2) \\
N_{13} = \sum_{i=21}^{30} I(X_i = 1 \text{ or } X_i = 3) \\
N_{23} = \sum_{i=31}^{40} I(X_i = 2 \text{ or } X_i = 3)
\]

In fact, data in hand are \( N_1 = 1, N_2 = 4, N_3 = 5, N_{12} = 4, N_{13} = 5 \) and \( N_{23} = 7 \).

a) Write out a loglikelihood function, \( \ell(p_1, p_2) \), based on the \( N \)'s.

\( \ell(p_1, p_2) \) is complicated enough that pencil-and-paper analysis of it is not sensible. I used MathCad and found out the following about this function of 2 variables. In the first place, it is maximized at \( (p_1, p_2) = (.163, .372) \), where \( \ell(.163, .372) = -30.652 \). Further, for \( \ell_{ij}(p_1, p_2) \) the second partial of \( \ell \) with respect to \( p_i \) and \( p_j \),

\[
\ell_{11}(.163, .372) = -225.7, \quad \ell_{12}(.163, .372) = -64.9, \quad \text{and} \quad \ell_{22}(.163, .372) = -142.5.
\]

b) Give a large sample approximate 90% confidence interval for \( p_1 \) based on the above information.

Attached to this exam is a plot of \( \ell(p, p) \) versus \( p \). (This is the loglikelihood assuming that \( p_1 = p_2 = p \).) Use this plot and the information above to do the following.

c) Give a maximum likelihood estimate of \( p_1 \) if one supposes that \( p_1 = p_2 \).

d) Carry out a likelihood ratio test of \( H_0: p_1 = p_2 \) vs \( H_a: p_1 \neq p_2 \) at an approximately .05 level. (Since the null hypothesis imposes one restriction on the parameter vector, the appropriate limiting distribution will have 1 degree of freedom.)