1. In this problem, please save all the parameter values and expected values you estimate, not only to those digits displayed, but to all the digits your calculator computes, by storing them in your calculator memory. One reason is that there can be cancellation in subtractions resulting in loss of significant digits. Also, some estimates found in Problem 1 will be used again in Problem 2.

In a random sample of 10,000 people, the following probabilities were observed or would be predicted for blood types:

<table>
<thead>
<tr>
<th>Type</th>
<th>Observed</th>
<th>Theory I</th>
<th>Theory II</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.4541</td>
<td>$u(1 - v)$</td>
<td>$p^2 + 2pr$</td>
</tr>
<tr>
<td>B</td>
<td>0.0774</td>
<td>$(1 - u)v$</td>
<td>$q^2 + 2qr$</td>
</tr>
<tr>
<td>O</td>
<td>0.4384</td>
<td>$(1 - u)(1 - v)$</td>
<td>$r^2$</td>
</tr>
<tr>
<td>AB</td>
<td>0.0301</td>
<td>$uv$</td>
<td>$2pq$</td>
</tr>
</tbody>
</table>

According to Theory I, there are two forms (alleles) of one gene, C and c, and two forms of another gene, D and d. Here C has probability $u$, so c has probability $1 - u$, and D has probability $v$, so d has probability $1 - v$. Combinations of genes would give the following blood types: CD = AB, Cd = A, cD = B, cd = O. Having C is independent of having D. (A person gets one allele from each parent; we assume one is dominant over another, as in CC or Cc both giving “C,” and DD or Dd both giving “D.”)

Estimate the probability $u$ from the observed relative frequency of A’s plus AB’s, and $v$ from B’s plus AB’s.

In Theory II, there are three forms of one gene, C, c and t. The probabilities are $p$ for C, $q$ for c, and $r$ for t, so $p + q + r = 1$. Each person has two of the alleles, one from each parent. Since Cc = cC, Ct = tC and ct = tc (we assume here it doesn’t matter which parent each came from), there are 6 possible different combinations of two alleles a person can have. Blood groups would be determined as follows: CC or Ct is A, cc or ct is B, Cc is AB, and tt is O.

Estimate $(p + r)^2$ from the observed relative frequencies of A’s plus O’s, then take the square root to estimate $p + r$. Likewise estimate $(q + r)^2$ from the observed relative frequencies of B’s plus O’s, take the square root to estimate $q + r$ and then $p$ as $1 - q - r$. Then estimate $r$ as $1 - p - q$.

Test both theories by chi-squared. Which fits the data better? Is either theory rejected by the data, at the 0.05 level?

2. For the same data as in the previous problem, again test the same two hypotheses, but instead of a $\chi^2$ test, use the Wilks likelihood ratio test. Since the maximum likelihood is hard to compute for Theory II, don’t compute it, but again use the estimates as for the $\chi^2$ test. We know that the likelihood, maximized at the unknown MLEs, will be at least as large as it is at the estimated parameters used. By Wilks’s theorem, if a $d$-dimensional hypothesis $H_0$ is true, then for $n$ large (which it is), his likelihood ratio statistic has a distribution approximately $\chi^2_{k-1-d}$, since the dimension of the full multinomial model is $k - 1$.

Beside doing the tests, see if the values of the statistics are similar, for a non-rejected hypothesis.
3. Consider Problem 21 in Section 11.6, pp. 462-463 of Rice, but changed as follows. Omit the smallest and largest measurements for each of the two types of bearings, leaving 8 for each. Then, for the data on Type I, normality is not rejected, but for Type II it is. 
   (a) In what way(s) can one see that the Type II data may be non-normal? 
   (b) The non-normality indicates that the analysis in part (a) of Rice’s problem would be inappropriate, so we don’t need to do it, and this also answers his part (c). So, compare the two samples by a nonparametric method as in Rice’s part (b).

4. Problem 38 in §11.6 pp. 466-467 of Rice, but only for Pyrazolone-T and omitting the observations (0,0), leaving 9 pairs of observations.

5. An example on pp. 477-478 of Rice has a table with data from 7 different labs. The Shapiro-Wilk test was applied to the four separate data sets from Labs 1, 2, 3, and 7, and none of the four tests rejected normality, where the mean and variance could depend on the lab (even though as the box plot shows, the observation 3.81 from Lab 7 might be suspected of being an outlier).
   (a) Find the sample variance of the 10 observations from Lab 2. 
   (b) Find the sample variance of the 10 observations from Lab 3. 
   (c) Test whether these two variances are the same, using an F-test at level 0.05. It should be a two-sided test, but the tables of the F distribution in Rice are for one-sided tests. So, see if the ratio of the larger to the smaller of the two sample variances is larger than the 0.975 quantile of the F distribution with the appropriate degrees of freedom in numerator and denominator. 
   (d) Find also the p-value of the test, in other words the probability of observing a value of the F statistic as large or larger than the one $F_{obs}$ actually observed. You can do this in R by finding the probability that $F \leq F_{obs}$ as $pf(F_{obs}, n_1, n_2)$ where $n_1$ and $n_2$ are the degrees of freedom in the numerator and denominator respectively, then subtracting from 1 to get $P(F > F_{obs})$. (If R isn’t conveniently available to you there is a p-value calculator for F distributions provided on the Web at http://davidmlane.com/hyperstat/F_table.html. Or, when I Google “F distribution” I find the site as the third listed one after Wikipedia and Wolfram.) Then multiply by 2 to get a p-value for the 2-sided test. 
   (e) There are $\binom{7}{2} = 21$ possible comparisons of two different labs. The p-value from part (d) is not really valid because the specific comparison was chosen after noticing from the box plots that the Lab 2 seemed to have a high variance and the Lab 3 data a small one. So multiply the p-value from part (d) by 21 as a “Bonferroni” correction for multiple tests. Is the result still less than 0.05? If so, one can decide that there are significant differences in the variances between labs, and the basic assumption of the Analysis of Variance that the variances should be equal would be violated. If the result is larger than 0.05, one might conclude that a purportedly significant ratio found in parts (c) and (d) was not really significant.