STAT 319/739 Bayesian statistics  
Project II  
Saad Mneimneh  
Computer Science  
Hunter College of CUNY

Part I  
Consider the following table.

<table>
<thead>
<tr>
<th>success</th>
<th>failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>group 1</td>
<td>5</td>
</tr>
<tr>
<td>group 2</td>
<td>11</td>
</tr>
</tbody>
</table>

(a) Perform a chi-squared test and a Fisher exact test (from homework 6) to decide if the successes in group 1 are too low. Is there a discrepancy in the results of the two tests? Explain.

(b) For this part and on, you will wear your Bayesian hat! Let \( p \) be the probability of success in group 1, and \( q \) the probability of success in group 2. Consider the following prior (\( p \) and \( q \) are independent):

\[
f(p, q) \propto p^{\alpha-1}(1-p)^{\beta-1}q^{\gamma-1}(1-q)^{\delta-1}
\]

and assume a binomial likelihood for the two groups that are independent when conditioned on \( p \) and \( q \), i.e.:

\[
P(\text{table}|p, q) = P(\text{row 1}|p)P(\text{row 2}|q)
\]

Find the posterior density for \( p \) and \( q \), \( f(p, q|\text{table}) \), when all parameters are zero, and compute \( P(p < q) \) using an approximation:

\[
P(p < q) \approx \frac{\sum_{p < q} f(p, q|\text{table})}{\sum_{p, q} f(p, q|\text{table})}
\]

for \( p, q \in \{0.01, 0.02, 0.03, \ldots, 0.99\} \). Use any programming language to compute your answer.

(c) Someone with two Bayesian hats suggests that if a value for \( p \) is revealed, say 0.8, then one would expect that \( q \) is also near 0.8. In other words, \( p \) and \( q \) should not be independent. Let

\[
\theta_1 = \ln \frac{p}{1-p} \quad \text{and} \quad \theta_2 = \ln \frac{q}{1-q}
\]
\[ \theta_2 | \theta_1 \sim N(\theta_1, \sigma^2) \]

Therefore, consider the following (improper) prior:

\[ f(\theta_1, \theta_2) \propto e^{-\frac{(\theta_1 - \theta_2)^2}{2\sigma^2}} \]

and find the corresponding prior \( f(p, q) \) (ask me if you don’t know how to make change of variable involving two variables), and show that the posterior \( f(p, q | \text{table}) \) is proper if the table has no zero entries (which is true for this example).

(d) Compute \( P(p < q) \) using the same technique in part (b) for \( \sigma = 1/2, \sigma = 1, \) and \( \sigma = 2, \) and compare your results to (a) and (b), and discuss.

Part II

A loaded die has probability 1/2 for 6. The probability of switching to a loaded die is 0.05. The probability of switching to a fair die is 0.1. We observe the following:

65116645313265124536664631636663162326455236266666625151631

Which parts of this sequence are generated by the loaded die? Use the Viterbi algorithm to figure this out. This needs a computer program.

Note: Use the log transformation described at the end of Note 11.