1. For three income groups (low, medium, high), let $\theta_1$, $\theta_2$, and $\theta_3$ be the respective fractions of individuals who favor a certain piece of legislation. Consider testing $H_0 : \theta_1 = \theta_2 = \theta_3$ against the alternative that these are not all equal. For samples of sizes 250, 200, and 150 taken in the three groups, respectively, the counts favoring or not favoring the legislation are as follows:

<table>
<thead>
<tr>
<th>Income Level</th>
<th>Favor</th>
<th>Do Not Favor</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>155</td>
<td>95</td>
<td>250</td>
</tr>
<tr>
<td>Medium</td>
<td>118</td>
<td>82</td>
<td>200</td>
</tr>
<tr>
<td>High</td>
<td>87</td>
<td>63</td>
<td>150</td>
</tr>
</tbody>
</table>

Is this a test of homogeneity or independence? Why?

This is a test whether the three populations are identical with respect the distribution of the variable Favor - Not Favor. Samples of specified sizes are taken from each population. Hence a test of homogeneity.

2. (Continuation) Show that the pooled estimate of the common value of $\theta$ assumed in $H_0$ is $\hat{\theta} = 0.6$.

This is the total number of "successes" divided by the total number of "trials":

$$\hat{\theta} = \frac{155 + 118 + 87}{250 + 200 + 150} = \frac{360}{600} = 0.6.$$ 

3. (Continuation) Suppose that the test statistic

$$\chi^2 = \frac{(155 - 150)^2}{150} + \frac{(95 - 100)^2}{100} + \frac{(118 - e_{21})^2}{e_{21}} + \frac{(82 - e_{22})^2}{e_{22}} + \frac{(87 - 90)^2}{90} + \frac{(63 - 66)^2}{66}$$

is to be used. Derive the appropriate $e_{21}$ and $e_{22}$, with details.

$$e_{21} = H_0\text{-fraction of 2nd population who "Favor" } = 0.6 \times 200 = 120,$$

$$e_{22} = 200 - e_{21} = 80.$$ 

4. (Continuation) The approximate $H_0$-distribution of this statistic is chi-square. Justify that the appropriate degrees of freedom is 2.

The degrees of freedom is (no. of rows - 1) x (no. of cols - 1) = (3 - 1) x (2 - 1) = 2 x 1 = 2.

5. (Continuation) Indicate the relevant probability model or models for the given data.

The relevant models are binomial(250, $\theta_1$), binomial(200, $\theta_2$), and binomial(150, $\theta_3$).