Statistical Analysis on Population Data

Kiki Mueller
Mentor: Prof. Marfani
Gender and Literacy

I researched the connection between literacy rates and gender to detect if male literacy tends to be more prevalent in various countries. There were two main reasons behind this statistic.

1. Patricia Kelley, an International Literacy Coordinator, focuses on the differences between the roles of a man and a women. It is much easier for a man to adjust his schedule to make time for education whereas a women is constantly depended on for family care.

2. Furthermore, Knezek discusses the fear of change; if a woman attains knowledge, she is able to fight for equal rights, leading to adjustments in society. However, there is statistical evidence claiming that equal literacy rates have an extremely positive effect on society.
Is literacy rate dependent on gender?

- Null hypothesis: The proportion of literate females is equivalent to the proportion of literate males
- Alternate hypothesis: The proportion of literate males is NOT equal to the proportion of literate females
- The P-value had to be below .05 in order to statistically prove a difference between literacy rates and gender.
Literacy is defined as “The percentage of people ages 15 and above who can, with understanding, read and write a short, simple statement on their everyday life.” The genders had a varying percentage within each country.

A 2-propZtest was conducted to determine the P-value.

If a statistical significance was shown, it would show that literacy rates are dependent on gender.
<table>
<thead>
<tr>
<th>Country</th>
<th>Adult (15+) literacy rate (%)</th>
<th>2011</th>
<th>Female</th>
<th>2011</th>
<th>Direction</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jamaica</td>
<td>82.06368</td>
<td>91.76553</td>
<td>p1&lt;p2</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mongolia</td>
<td>96.81157</td>
<td>97.87714</td>
<td>p1&lt;p2</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Puerto Rico</td>
<td>89.7129</td>
<td>90.91396</td>
<td>p1&lt;p2</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seychelles</td>
<td>91.40732304</td>
<td>92.25897952</td>
<td>p1&lt;p2</td>
<td>1.75E-12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Costa Rica</td>
<td>96.0423</td>
<td>96.47886</td>
<td>p1&lt;p2</td>
<td>1.30E-07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Botswana</td>
<td>84.62992</td>
<td>85.56156</td>
<td>p1&lt;p2</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estonia</td>
<td>99.76915</td>
<td>99.81971</td>
<td>p1&lt;p2</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antigua and Barbuda</td>
<td>98.4</td>
<td>99.42</td>
<td>p1&lt;p2</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dominican Rep.</td>
<td>89.9848332021186</td>
<td>90.2255199031432</td>
<td>p1&lt;p2</td>
<td>0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colombia</td>
<td>93.4882505559722</td>
<td>93.6671378327439</td>
<td>p1&lt;p2</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The 2-Prop Z tests results concluded the following:

1) In these following countries, we cannot say that literacy rates are dependent on gender:

   Columbia
   Lithuania
   Latvia
   Argentina
   Cuba
   Netherlands Antilles
2) Because $p_1$ was greater than $p_2$, the test proved that in these countries, the male literacy rate is higher:

- Albania
- Angola
- Bangladesh
- Bosnia and Herzegovina
- Cape Verde
- Chad
- Comoros
- Cote d'Ivoire
- Croatia
- Cyprus
- Ecuador
- Egypt
- Equatorial Guinea
- Eritrea
- Gabon
- Gambia
- Greece
- Guatemala
- Guinea-Bissau
- Indonesia
- Iraq
- Jordan
- Libya
- Macao, China
- Macedonia, FYR
- Mali
- Mauritania
- Mauritius
- Mexico
- Moldova
- Montenegro
- Morocco
- Myanmar
- Nepal
- Papua New Guinea
- Portugal
- Saudi Arabia
- Serbia
- Sierra Leone
the male literacy rate is higher:

- Sudan
- Syria
- Togo
- Turkey
- West Bank and Gaza
- Vietnam
- Yemen, Rep.
- Zimbabwe
- Vanuatu
- Romania
- Trinidad and Tobago
- South Africa
- Bulgaria
- Poland
- Uzbekistan
- Italy
- Brunei
- Armenia
- Samoa
- Turkmenistan
- Tajikistan
- Swaziland
- Ukraine
- Hungary
- Georgia
- Honduras
- Slovenia
3) Because $p_1$ was less than $p_2$, the test proved that in the following countries, the female literacy rate is higher:

- Jamaica
- Mongolia
- Puerto Rico
- Seychelles
- Costa Rica
- Botswana
- Estonia
- Antigua and Barbuda
- Dominican Rep.

Based on the results, it is apparent that male literacy rate is higher in a majority (67) of the countries tested using the 2-Prop Z test. In nine countries, the female literacy rate was proven to be higher. In six countries, the results were inconclusive.
Sugar Consumption and GDP

The second question I researched was the connection between sugar consumption and GDP.

According to the Food and Agriculture Organization of the UN, dietary habits and economic status are the main factors of sugar intake among countries.

The first factor is dietary habits. A “nutrition transition” which is the change in a country's dietary habits showed an increase in energy levels in low-income countries; this increase was attributed to rising levels of sugar intake. Trade determines a country's ability to receive sugar, as well as what valuable commodity a country has to offer. This would lead me to believe that countries with higher sugar intake have more income.
The second factor is agricultural policies; the food that is most desired will result in increased trade within the country. The equalization of low-income and high-income energy intake in 1993 was due to low-income countries obtaining more energy through sugar. Contrary to my data, this research shows that certain low-income countries are able to grow or obtain more sugar, resulting in the ability to consume more sugar daily. Furthermore, the span between 1993 and 2004 may have lead to further variation due to changes in these two factors.
**Question Two**

Is there a relationship between the amount of sugar consumed and GDP?

- Null hypothesis: average sugar intake in low GDP countries is equal to sugar intake in high GDP countries
- Alternative hypothesis: average sugar intake for low GDP countries is less than the sugar intake in high GDP countries
- There was an arbitrary GDP cut-off of $10,000, which was how I determined the difference between a low-GDP country and a high-GDP country. The data obtained was for 2004.
- Again, a P-value less than .05 would allow me to conclude the results are statistically significant and the alternative hypothesis is proven correct.
A 2-Samp-Ttest was conducted to discover whether there was any relation between the monetary value of a country and the amount of sugar consumed.

Amongst 136 low GDP countries, the mean sugar intake was 73.782 grams with a standard deviation of 43.813 grams.

Amongst 30 high GDP countries, the mean sugar intake was 118.91 grams with a standard deviation of 32.06 grams.
Using the values mentioned in the previous slide, the P-value was calculated to be 1.29E-8.

Because the calculation is less than .05, there is enough evidence to proof that low GDP countries consume less sugar than high GDP countries.