Consider the following sample observations: 2781, 2900, 3013, 2856, and 2888. Suppose we want to do a two-sided test of the population mean, i.e.,
H0: \( \mu = 3000 \)
H1: \( \mu \neq 3000 \)

a) Compute the p-value, and state the conclusion "In English" (i.e., is there evidence that \( \mu \) is not 3000?) using alpha = 0.05.

b) Compute the appropriate confidence interval (CI). Is the conclusion the same as in part a? Explain.

One can also arrive at the same conclusion, without the p-value and CI, by what is called the rejection method. I'll walk you through it:

c) If H0 is true, compute the values of \( \bar{x} \) that have an area of 0.025 to the right and 0.025 to the left. (Together these areas add up to 0.05, i.e., alpha). These values of \( \bar{x} \) are called the critical values, and the regions beyond them (i.e., larger than the larger one, and smaller than the smaller one) are called the rejection region. So, in this part of the problem you are computing the rejection region.

d) Is the observed value of \( \bar{x} \) in the rejection region? If so, one can reject H0 in favor of H1; otherwise, one cannot say anything.

\[ \bar{x} = 2887.6, \quad s = 84.03, \quad n = 5 \]  

\[ t_{0.025} = \frac{2887.6 - 3000}{(84.03)/\sqrt{5}} = -2.78 \Rightarrow p\text{-value} = 2 \cdot P(t < -2.78) = 2(0.025) = 0.05 \]

At \( \alpha = 0.05 \), we have \( p\text{-value} < \alpha \Rightarrow \text{Reject H0 in favor of H1}. \)

In English: Data provide evidence that \( \mu \neq 3000 \).

\[ 2\text{-sided 95\% CI:} \quad \bar{x} \pm t_{0.025} \cdot \frac{s}{\sqrt{n}} = 2887.6 \pm 2.78 \cdot \frac{84.03}{\sqrt{5}} = [2783.2992] \]

In English: 3000 is not in the CI, and so, there is evidence that \( \mu \neq 3000 \).

\[ c) \quad \frac{\bar{x} - 3000}{s/\sqrt{n}} = 2.78 \Rightarrow \bar{x} = 3000 + 2.78 \cdot \frac{84.03}{\sqrt{5}} = 3104 \]

\[ \frac{\bar{x} - 3000}{s/\sqrt{n}} = -2.78 \Rightarrow \bar{x} = 3000 - 2.78 \cdot \frac{84.03}{\sqrt{5}} = 2895 \]

d) The observed \( \bar{x} \) (2887.6) is in the rejection region, and so there is evidence that \( \mu \neq 3000 \).
A sample of 210 Bell computers has 56 defectives. Theory suggests that a third of all Bell computers should be defective. Does this data contradict the theory (at alpha=0.05)? Specifically,

a) Do a z-test (p.2 above),

b) Do a chi-squared test with k=2 categories. Hint: The pi's (and pi_0's) of the k categories must sum to 1.

\[ \begin{align*}
\text{a) Let } \pi & = \text{ proportion of defectives in all Bell comps.} \\
H_0: \pi &= \frac{1}{3} \quad \text{Data says } \hat{p} = \frac{56}{210} = 0.267 \\
H_1: \pi &\neq \frac{1}{3} \quad \text{This is a two-sample, two-sided/tailed test of a proportion.} \\
\text{If } H_0 = \text{True, then } Z_{obs} &= \frac{(56/210) - (1/3)}{\sqrt{\frac{1}{210} \left( \frac{1}{3} \right) \left( 1 - \frac{1}{3} \right)}} = -2.05 \\
\therefore \ p\text{-value} &= 2 \ \text{prob}(Z < -2.05) = 2(0.0202) = 0.041 \\
\text{p-value < } \alpha \Rightarrow \text{Reject } H_0 \text{ in favor of } H_1, \ i.e. \ Data \ does \ not \ support \ the \ theory. \\
\end{align*} \]

\[ \begin{align*}
\text{b) This time let} \\
\pi_1 &= \text{ proportion of defectives in all Bell computers} \\
\pi_2 &= \text{ "non-defects" } \\
\end{align*} \]
\[ H_0: \pi_1 = \frac{1}{3}, \pi_2 = \frac{2}{3} \quad \Rightarrow \quad \pi_0 = \frac{1}{3}, \pi_2 = \frac{2}{3} \]

\[ H_1: \text{At least one is wrong} \]

<table>
<thead>
<tr>
<th>Cat. 1</th>
<th>Cat. 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected counts:</td>
<td>70/3 (210) = 70</td>
</tr>
<tr>
<td>Observed counts:</td>
<td>56</td>
</tr>
</tbody>
</table>

\[ \chi^2 = \frac{(70 - 56)^2}{70} + \frac{(140 - 154)^2}{140} = 2.8 + 1.4 = 4.20 \]

Table VII with df = 2 - 1 = 1 \( \Rightarrow \) 0.045 < p-value < 0.040

Infact, Table VII suggests that p-value \( \sim 0.04 \)

c) Yes! The 2-sided z-test gives p-value = 0.04

The chi-squared test gives p-value \( \sim 0.04 \)

Also note that \( \chi^2_{\text{obs}} = 4.20 \) turns out to be equal to \( z_{\text{obs}}^2 = (-2.05)^2 = 4.20 \).

This is not accidental: The chi-squared distribution is related to the square of a variable that is normally distributed.
Consider the data from an example in a past lecture where a survey of students in 370 yielded the following data:

17 students like Lab
48 " " Do not like Lab
15 " " have no opinion.

Suppose I believed that the proportion of students in each of the 3 categories (like, no-like, no-opinion) was equal. Does this data contradict that belief?

Let \( \theta_1 \): True proportion of students who like lab,
\( \theta_2 \): " " do not like Lab,
\( \theta_3 \): " " have no opinion.

Then \( H_0: \theta_1 = \frac{1}{3}, \theta_2 = \frac{1}{3}, \theta_3 = \frac{1}{3} \)

\( H_1: \) At least 1 of these is wrong.

\[
\chi^2 = \sum \frac{(\text{obs}_i - \text{exp}_i)^2}{\text{exp}_i} = \sum \left( \frac{n_i \theta_i - n_i}{n_i} \right)^2
\]

\[
n = \sum n_i = 17 + 48 + 15 = 80
\]

\[
= \frac{(80(\frac{1}{3}) - 17)^2}{80(\frac{1}{3})} + \frac{(80(\frac{1}{3}) - 48)^2}{80(\frac{1}{3})} + \frac{(80(\frac{1}{3}) - 15)^2}{80(\frac{1}{3})}
\]

\[
= \frac{1}{80(\frac{1}{3})} \left[ (26.7 - 17)^2 + (26.7 - 48)^2 + (26.7 - 15)^2 \right]
\]

\[
= \frac{1}{26.7} \left[ 94.09 + 453.69 + 136.89 \right] = 25.64
\]

\[
p\text{-value} = p(\chi^2 > 25.64) < 0.001 \quad \text{df} = 3-1 = 2 \quad \text{Table VII}
\]

So because \( p\text{-value} < \alpha \) (at \( \alpha = 0.05 \))

we can reject \( H_0 \) in favor of \( H_1 \) (at least one of the 3 proportions is not \( \frac{1}{3} \)), i.e. evidence contradicts the belief.

In English: At least 1 of the props is not \( \frac{1}{3} \).

FYI: Diagnosis: The "Categ. with the biggest deviation from \( \frac{1}{3} \)" is the "do not like lab" group.
The accompanying data resulted from an experiment in which seeds of five different types were planted and the number that germinated within 5 weeks of planting was observed for each seed type ("Nondestructive Optical Methods of Food Quality Evaluation," Food Science and Nutr., 1984: 232-279). Carry out a chi-squared test at level .01 to see whether the proportion of seeds that germinate in the specified period varies according to type of seed.

<table>
<thead>
<tr>
<th>Seed type:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germinated</td>
<td>31</td>
<td>57</td>
<td>87</td>
<td>52</td>
<td>10</td>
</tr>
<tr>
<td>Failed to germinate</td>
<td>7</td>
<td>33</td>
<td>60</td>
<td>44</td>
<td>19</td>
</tr>
</tbody>
</table>

Specifically,

a) Does the statement of the problem require a test of homogeneity of 2 populations with respect to 5 categories, or vice versa?

b) Compute the p-value corresponding to your answer in part a.

c) State your conclusion "in English."

d) Diagnose the various terms in $X_{obs}^2$.

a) This question can be understood this way: what is the sample proportion of seeds that germinate, for a given seed type?

<table>
<thead>
<tr>
<th>Seed type</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample prop</td>
<td>31/37</td>
<td>57/57</td>
<td>87/97</td>
<td>52/52</td>
<td>10/10</td>
</tr>
<tr>
<td>= 0.82</td>
<td>0.63</td>
<td>0.59</td>
<td>0.54</td>
<td>0.34</td>
<td></td>
</tr>
</tbody>
</table>

These are sample props, $P_1, P_2, ..., P_5$.

The question is asking if the true/pop. props are equal, i.e., $P_1 = P_2 = ... = P_5$.

This condition is exactly what it means to say that 5 populations are homogeneous w.r.t. 2 categories. See the lecture notes.

So $r = 5$, $k = 2$

$H_0$: $P_1 = P_2 = ... = P_5$

$H_1$: At least 2 of the $P_i$'s are different.
b) Observed counts:

\[
\begin{pmatrix}
31 & 57 & 87 & 52 & 10 \\
7 & 33 & 60 & 44 & 19 \\
38 & 90 & 147 & 96 & 26 \\
\end{pmatrix}
\]

Sample size (N) = 1400

Expected counts (under $H_0$):

\[
\frac{1}{1400}
\begin{pmatrix}
237(38) & 237(90) & 237(147) & 237(96) & 237(26) \\
166(38) & 163(90) & 163(147) & 163(96) & 163(26) \\
\end{pmatrix}
\]

\[
= \begin{pmatrix}
22.52 & 53.32 & 87.10 & 56.88 & 17.18 \\
15.49 & 36.68 & 59.90 & 39.12 & 11.81 \\
\end{pmatrix}
\]

\[
= \frac{(31 - 22.52)^2}{22.52} + \frac{(57 - 53.32)^2}{53.32} + \frac{(87 - 87.10)^2}{87.10} + \frac{(52 - 56.88)^2}{56.88} + \frac{(10 - 17.18)^2}{17.18}
\]

\[
+ \frac{(7 - 15.49)^2}{15.49} + \cdots
\]

\[
= 3.20 + 0.25 + 0.0001 + 0.42 + 3.00 + 4.65 + 0.37 + 0.0002 + 0.61 + 4.37 = 16.86
\]

\[
\chi^2 = (5-1)(2-1) = 4
\]

\[
p-value = \Pr(\chi^2 > 16.86) = \frac{0.001}{0.005}
\]

At $\alpha = 0.01$, because $p-value < \alpha$, we reject $H_0$ ($\pi_1 = \pi_2 = \cdots = \pi_5$) in favor of $H_1$ (at least 2 $\pi_i$'s are different).

c) In English, and to answer the question asked, we say:

There is evidence from data that the proportion of seeds that germinate varies (or depends on) seed type.

d) The large contributors to $\chi^2_{obs}$ are from the seed types 1 and 5, i.e., seed types 1 and 5 are the most different (in terms of germination rate, compared to what we would expect if germination rate did not depend on seed type).
Have you ever wondered whether soccer players suffer adverse effects from hitting "headers"? The authors of the article "No Evidence of Impaired Neurocognitive Performance in Collegiate Soccer Players" (The Amer. J. of Sports Medicine, 2002: 157-162) investigated this issue. The paper reported that 45 of the 91 soccer players in their sample had suffered concussion, 28 of 96 non-soccer athletes had suffered concussion, and only 8 of 53 student controls had suffered concussion. Denote

\[ p_{i1} = \text{pop. proportion of concussions among soccer players}, \]
\[ p_{i2} = \text{pop. proportion of concussions among non-soccer players}, \]
\[ p_{i3} = \text{pop. proportion of concussions among control group}. \]

Set up this problem as a test of homogeneity of three populations with respect to 2 categories. Specifically,

a) State the hypotheses in terms of \( p_{i1}, p_{i2}, p_{i3} \).

b) Write the data in the form of a contingency table.

c) Compute the expected counts.

d) Compute the p-value (or specify a range for it).

e) State the conclusion "in English."

f) Diagnose the various terms appearing in \( X_{\text{obs}}^2 \).

---

a) \( H_0: p_{i1} = p_{i2} = p_{i3} \)

\( H_1: \text{At least 2 of } p_i \text{'s are different} \)

b) Pop 1 (Soccer) \hspace{1cm} (45 \hspace{0.5cm} 46) \hspace{1cm} 91

\( = \hspace{0.5cm} 2 \hspace{0.5cm} (\text{non-soccer}) \hspace{1cm} (28 \hspace{0.5cm} 68) \hspace{1cm} 96 \)

\( = \hspace{0.5cm} 3 \hspace{0.5cm} (\text{control}) \hspace{1cm} (8 \hspace{0.5cm} 45) \hspace{1cm} 53 \)

\( = \hspace{0.5cm} 81 \hspace{0.5cm} 159 \hspace{0.5cm} 240 \)

e) \text{expected counts (under } H_0, \text{ i.e., when } H_0 = \text{True}): \]

\[
\begin{pmatrix}
\frac{91 \times 81}{240} = 30.7 & \frac{91 \times 159}{240} = 60.3 & 91 \\
\frac{96 \times 81}{240} = 32.4 & 63.6 & 96 \\
17.9 & 35.1 & 53
\end{pmatrix}
\]

\( \text{Just checking!} \)
d) \[ x^2 = \frac{(45-30.7)^2}{30.7} + \frac{(46-60.3)^2}{60.3} + \frac{(28-32.4)^2}{32.4} + \ldots + \frac{(\ldots)^2}{\ldots} = 6.66 + 3.39 + 0.59 + 0.30 + 5.48 + 2.79 = 19.22 \]

With \( df = (3-1)(2-1) \), Table VII \( \Rightarrow p\text{-value} < .001 \)

e) At \( \alpha = .05 \), or \( .01 \)

Reject \( H_0 \) in favor of \( H_1 \)

\[ \bar{Y}_1 = \bar{Y}_2 = \bar{Y}_3 \] At least 2 of the \( \bar{Y} \)'s are different.

In English: For at least 2 of the 3 populations (soccer, non-soccer, control), the proportion of concussions is different. This conclusion is true for any \( \alpha \geq .001 \).

f) Note that the conclusion in part e) is that the 3 populations are different, in terms of concussion rate in each pop. Given that each pop. has 2 categories, if the pops are different in terms of one of the categories, then they must also be different in terms of the other category. So, it is not possible to diagnose which category is the one in terms of which the pops are most different. However, FYI, the test of homogeneity is symmetric w.r.t. switching population and category. So, we can also conclude that the 2 populations of concussion and non-concussion are different w.r.t. at least one of the 3 categories (soccer, non-soccer, control). In that case, we can diagnose the \( x^2 \) values above, and conclude that the difference between the 2 pops is mostly in the soccer (and control) groups.