**Section 11.2 (Part 1) - Inference for Relationships** (pp. 696-709)

**Comparing Distributions of a Categorical Variable** - In this section, we will compare more than two samples or groups. More generally, we are going to compare distributions of a single categorical variable across several populations or treatments.

**Example** - Market researchers suspect that background music may affect mood and buying behavior of customers. One study in a supermarket compared three randomly assigned treatments: no music, French accordion music, and Italian string music. Under each condition, the researchers recorded the numbers of bottles of French, Italian, and other wine purchased. Here is a table that summarizes the data

SORT OF LIKE TREATMENTS.

| | Music | | | |
|---|---|---|---|---|
| **Wine** | **None** | **French** | **Italian** | **Total** |
| French | 30 | 39 | 30 | **99** |
| Italian | 11 | 1 | 19 | **31** |
| Other | 43 | 35 | 35 | **113** |
| **Total** | **84** | **75** | **84** | **243** |

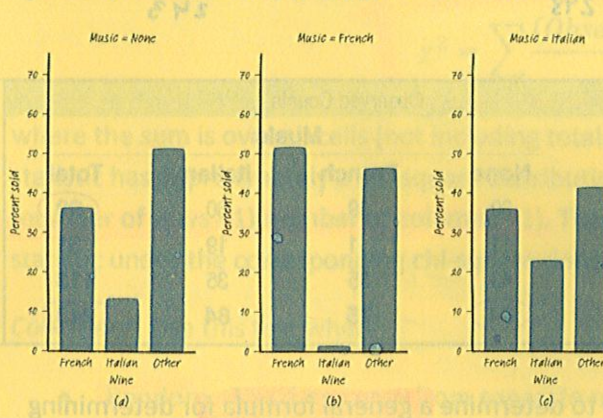
a) Calculate the conditional ~~probabilities~~ distribution of the type of wine sold for each treatment.

NO MUSIC:
$P(F) = \frac{30}{84} = 0.357$
$P(I) = 11/84 = 0.131$
$P(O) = 43/84 = 0.512$

FRENCH MUSIC:
$P(F) = 39/75 = 0.520$
$P(I) = 1/75 = 0.013$
$P(O) = 35/75 = 0.467$

ITALIAN MUSIC:
$P(F) = \frac{30}{84} = 0.357$
$P(I) = 19/84 = 0.226$
$P(O) = 35/84 = 0.417$

b) Given the graphs below, comparing the distributions of wine purchases under the different conditions, are the distributions similar or different?




ITALIAN SALES VERY LOW WHEN FRENCH MUSIC. HIGHER WHEN NO MUSIC OR ITALIAN MUSIC

THE %'S OF OTHER WERE CLOSE FOR ALL 3 TYPES OF MUSIC.

**Chi-Square Test for Homogeneity** - When we are trying to determine if the distribution of a categorical variable is the same for several populations or treatments, we conduct the chi-square test for homogeneity. For our example, this will involve testing the hypotheses:

*$H_0$: There is no difference in the distributions of wine purchases at this store when no music, French accordion music, or Italian string music is played.*

*$H_a$: There is a difference in the distributions of wine purchases at this store when no music, French accordion music, or Italian string music is played.*

**Computing Expected Counts** - Before we formally define the chi-square test for homogeneity, we will explore how to compute the expected counts.

| Observed Counts | | | | |
|---|---|---|---|---|
| | Music | | | |
| Wine | None | French | Italian | Total |
| French | 30 | 39 | 30 | 99 |
| Italian | 11 | 1 | 19 | 31 |
| Other | 43 | 35 | 35 | 113 |
| Total | 84 | 75 | 84 | 243 |

To find the expected counts, we start by assuming that $H_0$ is true. From the two way table, we can see that 99 of the 243 bottles of wine bought during the study were French wines.

If the specific type of music that is playing had no effect on wine purchase, the proportion of French wine sold under each condition should be 99/243 = 0.407. For instance, there were 84 bottles of wine bought when no music was playing. Then on average, how many bottles would we expect to be French wine?

$$\frac{99}{243} \cdot 84 = 34.22$$

The expected counts of the French wine bought under the other two conditions can be found in a similar way:

French music: $\frac{99}{243} \cdot \cancel{35}\,75 = 30.56$ Italian music: $\frac{99}{243} \cdot 84 = 34.22$

We repeat the procedure to find the expected counts for the other two types of wine.

Italian wine $31/243$

No music: $\frac{31}{243} \cdot 84 = 10.72$ French music: $\frac{31}{243} \cdot 75 = 9.57$ Italian music: $\frac{31}{243} \cdot 84 = 10.72$

Other wine $113/243$

No music: $\frac{113}{243} \cdot 84 = 39.06$ French music: $\frac{113}{243} \cdot 75 = 34.88$ Italian music: $\frac{113}{243} \cdot 84 = 39.06$

| Expected Counts | | | | |
|---|---|---|---|---|
| | Music | | | |
| Wine | None | French | Italian | Total |
| French | 34.22 | 30.56 | 34.22 | 99 |
| Italian | 10.72 | 9.57 | 10.72 | 31 |
| Other | 39.06 | 34.88 | 39.06 | 113 |
| Total | 84 | 75 | 84 | 243 |

| Observed Counts | | | | |
|---|---|---|---|---|
| | Music | | | |
| Wine | None | French | Italian | Total |
| French | 30 | 39 | 30 | 99 |
| Italian | 11 | 1 | 19 | 31 |
| Other | 43 | 35 | 35 | 113 |
| Total | 84 | 75 | 84 | $\boxed{243}$ |

Using the numbers marked in the table above, let's try to determine a general formula for determining the expected counts.

$$\frac{99}{243} \cdot 84 = \frac{99 \cdot 84}{243} = \frac{\text{(ROW TOTAL)(COLUMN TOTAL)}}{\text{TABLE TOTAL}}.$$

**Test Statistic** - The chi-square test statistic is computed the same way it was computed in the chi-square goodness-of-fit test.

$$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected}$$

Continuing with the music example, for French wine with no music, the observed count is 30 bottles and the expected count is 34.22. The contribution for the chi-square statistic for this cell is:

$$\frac{(30 - 34.22)^2}{34.22} = 0.52$$

This would be repeated for all nine cells and summed to determine the chi-square statistic.

$$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(30-34.22)^2}{34.22} + \frac{(39-30.56)^2}{30.56} + \cdots + \frac{(35-39.06)^2}{39.06} = \boxed{18.28}$$

---

**The Chi-Square Test for Homogeneity**

Suppose the Random, Large Sample Size, and Independent conditions are met. You can use the chi-square test for homogeneity to test

> *$H_0$: There is no difference in the distributions of a categorical variable for several populations or treatments.*
> *$H_a$: There is a difference in the distributions of a categorical variable for several populations or treatments.*

Start by finding the expected counts. Then calculate the chi-square statistic

$$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected}$$

where the sum is over all cells (not including totals) in the two-way table. If $H_0$ is true, the chi-square statistic has approximately a chi-square distribution with degrees of freedom = (number of rows - 1)(number of columns - 1). The *P*-value is the area to the right of the chi-square statistic under the corresponding chi-square density curve.

$(r-1)(c-1) = df$

*Conditions*: Use this test when:

- **Random** - The data come from separate random samples from each population of interest or from the groups in a randomized experiment.
- **Large Sample Size** - All expected counts are at least 5.
- **Independent** - Both the samples or groups themselves and the individual observations in each sample or group are independent. When sampling without replacement, check that the individual populations are at least 10 times as large as the corresponding samples (10% condition).

**Example (cont)** - In the wine example, the chi-square test statistic was 18.28. The degrees of freedom will be (3 - 1)(3 - 1) = 4. Using the calculator, the *P*-value is 0.0011.

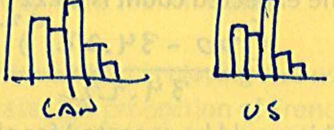
Conclusion: SINCE P-VALUE = 0.0011 < 0.05 = $\alpha$, REJECT $H_0$. THERE IS CONVINCING EVIDENCE THAT THERE IS A DIFF IN THE DISTS OF WINE SALES WHEN THE MUSIC IS DIFFERENT.

**Technology** - To conduct this test with a calculator, press [2nd] [$x^{-1}$] (MATRIX) <EDIT> and choose 1: A. Enter the dimensions of the contingency table. Enter the observed counts. Press [STAT] <TESTS> and choose C: $\chi^2$-Test.

NEED TO LOOK @ EXPECTED COUNTS IN [B]

**Application** - Complete CYU question 2 on ~~p. 708.~~ 711

① GRAPH




②

| Quality of life | Canada | United States | |
|---|---|---|---|
| Much better | 75 | 541 | 616 |
| Somewhat better | 71 | 498 | 569 |
| About the same | 96 | 779 | 875 |
| Somewhat worse | 50 | 282 | 332 |
| Much worse | 19 | 65 | 84 |
| Total | 311 | 2165 | 2476 |

| | | | |
|---|---|---|---|
| $\frac{(616)(311)}{2476}$ | $\frac{(616)(2165)}{2476}$ | 77.37 | 538.63 |
| $\frac{(569)(311)}{2476}$ | $\frac{(569)(2165)}{2476}$ | 71.47 | 497.53 |
| $\frac{(875)(311)}{2476}$ | $\frac{(875)(2165)}{2476}$ | 109.91 | 765.09 |
| $\frac{(332)(311)}{2476}$ | $\frac{(332)(2165)}{2476}$ | 41.70 | 290.30 |
| $\frac{(84)(311)}{2476}$ | $\frac{(84)(2165)}{2476}$ | 10.55 | 73.45 |

STATE: WE WANT TO TEST $H_0$: NO DIFF IN DIST. BETWEEN US + CANADA VS. $H_a$: THERE IS A DIFF.

PLAN: IF CONDITIONS MET, USE $\chi^2$ TEST FOR HOMOGENEITY. AT $\alpha = 0.01$

RANDOM: SEPARATE SRS'S ✓

LARGE SAMPLE SIZE: EXP COUNTS > 5 ✓

INDEPENDENT: LESS THAN 10% OF PEOPLE WHO HAVE HAD HEART ATTACKS. ✓

∴ COND'S MET.

DO:

$$\chi^2 = \sum \frac{(O-E)^2}{O} = \frac{(75-77.37)^2}{75} + \frac{(71-71.47)^2}{71} + \dots + \frac{(65-73.45)^2}{65} = \boxed{11.725}$$

P-VALUE = 0.0195 w/ df = 4

CONCLUDE: SINCE $P = 0.0195 > \alpha = 0.01$ WE FAIL TO REJECT $H_0$. THEREFORE ∄ ENOUGH EVIDENCE TO CONCLUDE THERE IS A DIFFERENCE IN DIST'S OF QUALITY OF LIFE IN CANADA + US.

**Follow-up Analysis**

```
Chi-Square Test: None, Franch, Italian
Expected counts are printed below observed counts
Chi-Square contributions are printed below expected counts

        None   French  Italian   Total
1         30       39       30      99
       34.22    30.56    34.22
       0.521    2.334    0.521

2         11        1       19      31
       10.72     9.57    10.72
       0.008    7.672    6.404

3         43       35       35     113
       39.06    34.88    39.06
       0.397    0.000    0.422

Total     84       75       84     243

Chi-Sq = 18.279, DF = 4, P-Value = 0.001
```

MUSIC

FR

IT

OT.

OBS

EXP

CONTRIB.

WAY BELOW

WAY ABOVE.

⇒ SALES OF ITALIAN WINE ARE STRONGLY AFFECTED BY MUSIC.

HW: Read ~~pp. 696-709; do problems 27, 29, 31, 35, 37, 39 on pp. 724-726.~~

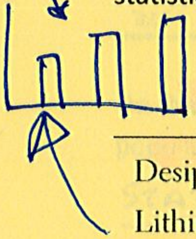
27, 29, 31, 37, 39

**Section 11.2 (Part 2) - Inferences for Relationships** (pp. 709-731)

**Comparing Several Proportions** - In chapter 10, we used the two-sample z-test to compare proportions from two samples from two populations. What if you want to compare proportions from more than two samples? The chi-square test for homogeneity can be used to do this.

The example on p. ~~710~~ illustrates this.

RELAPSE



Are the differences between the groups statistically significant at the 1% level?

| Group | Treatment | Cocaine Relapse? Yes | No |
|---|---|---|---|
| 1 | Desipramine | 10 | 14 |
| 2 | Lithium | 18 | 6 |
| 3 | Placebo $n=48$ | 20 | 4 $(n=24)$ |
| | | 48 | 24 |

| | | |
|---|---|---|
| Desipramine: | Yes: 10/24 = 0.417 | No: 14/24 = 0.583 |
| Lithium: | Yes: 18/24 = 0.75 | No: 6/24 = 0.25 |
| Placebo: | Yes: 20/24 = 0.833 | No: 4/24 = 0.167 |

**STATE**: WE WANT TO PERFORM A TEST OF

$H_0: p_1 = p_2 = p_3$ VS.

$H_A$: AT LEAST TWO OF THE $p_i$'S ARE DIFF WHERE $p_i$ = ACTUAL % OF CHRONIC COCAINE USERS LIKE THE ONES IN THIS EXPERIMENT WHO WOULD RELAPSE UNDER TREATMENT.

USE $\alpha = 0.01$

**PLAN**: IF COND'S ARE MET, WE WILL PERFORM A $\chi^2$ TEST FOR HOMOGENEITY.

RANDOM: RANDOM ASSIGNMENT ✓

LARGE SAMPLE SIZE: EXPECTED COUNTS ≥ 5 ✓

~~46.5 44.5 73.5 67.5~~

$$\begin{bmatrix} 16 & 8 \\ 16 & 8 \\ 16 & 8 \end{bmatrix}$$

INDEPENDENT: SUBJECTS IND ✓ ∴ CONDS MET.

~~DO: TEST STATISTIC: $\chi^2 = \frac{(36-46.5)^2}{36} + \dots + \boxed{11.15}$~~

~~P-VALUE (2-1)(2-1) = 1 df $\boxed{0.0008}$~~

~~CONCLUDE:~~

**DO**: TS: $\chi^2 = \frac{(10-16)^2}{16} + \frac{(18-16)^2}{16} + \dots = \boxed{10.5}$

P-VALUE: $df = (3-1)(2-1) = 2$ $\quad p = \boxed{0.0052}$

**CONCLUDE**: SINCE P-VALUE = 0.0052 < $\alpha$ = 0.01, REJECT $H_0$. WE HAVE SUFFICIENT EVIDENCE TO CONCLUDE RELAPSE RATES FOR THE 3 TREATMENTS ARE NOT ALL THE SAME.

## Relationships between Two Categorical Variables - Chi-Square Test for Association/Independence

Another common situation that leads to a two-way table is when a *single* random sample of individuals is chosen from a *single* population and then classified according to two categorical variables. The goal is to analyze the relationship between the variables. This leads to the chi-square test for association/independence.

### Chi-Square Test for Association/Independence

Suppose the Random, Large Sample Size, and Independent conditions are met. You can use the chi-square test for association/independence to test

> $H_0$: *There is no association between two categorical variables in the population of interest.*
> $H_a$: *There is an association between two categorical variables in the population of interest.*

Or, alternatively

> $H_0$: *Two categorical variables are independent in the population of interest.*
> $H_a$: *Two categorical variables are not independent in the population of interest.*

Start by finding the expected counts. Then calculate the chi-square statistic

$$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected}$$

df = (r-1)(c-1)

where the sum is over all cells (not including totals) in the two-way table. If $H_0$ is true, the chi-square statistic has approximately a chi-square distribution with degrees of freedom = (number of rows - 1)(number of columns - 1). The *P*-value is the area to the right of the chi-square statistic under the corresponding chi-square density curve.

*Conditions*: Use this test when:

- **Random** - The data come from separate random samples from each population of interest or from the groups in a randomized experiment.
- **Large Sample Size** - All expected counts are at least 5.
- **Independent** - Both the samples or groups themselves and the individual observations in each sample or group are independent. When sampling without replacement, check that the individual populations are at least 10 times as large as the corresponding samples (10% condition).

**Example** - A study followed a random sample of 8474 people with normal blood pressure for about four years. All the individuals were free of heart disease at the beginning of the study. Each person took and anger scale test which measures how prone a person is to sudden anger. Researchers also recorded whether each individual developed coronary heart disease (CHD). This includes people who had heart attacks and those who needed medical treatment for heart disease.

| | Low anger | Moderate anger | High anger | Total |
|---|---|---|---|---|
| CHD | 53 | 110 | 27 | 190 |
| No CHD | 3057 | 4621 | 606 | 8284 |
| Total | 3110 | 4731 | 633 | 8474 |

EXPECTED COUNTS

| | | |
|---|---|---|
| 69.73 | 106.08 | 14.19 |
| 3040.27 | 4624.92 | 618.81 |

Do the data provide convincing evidence of an association between anger level and heart disease in the population of interest?

STATE: WE WANT TO TEST $H_0$: THERE IS NO ASSOC. BETWEEN ANGER LEVEL AND CHD IN POP OF PEOPLE W/NORMAL BLOOD PRESSURE. VS. $H_A$: THERE IS AN ASSOC. BETWEEN ANGER LEVEL AND CHD IN PEOPLE W/NORMAL B.P. WE WILL USE $\alpha = 0.05$.

PLAN: IF CONDITIONS ARE MET, USE $\chi^2$ TEST FOR ASSOC/INDEP.

RANDOM: SRS ✓

LARGE SAMPLE SIZE: EXPECTED COUNTS > 5 ✓

INDEPENDENT: S.T.A. MORE THAN 10 (8474) = 84740 PEOPLE W/NORMAL BLOOD PRESS. ✓ ∴ CONDS MET.

DO: TS: $\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(53-69.73)^2}{69.73} + \frac{(110-106.08)^2}{106.08} + \dots = \boxed{16.077}$

P-VALUE: $df = (2-1)(3-1) = 2$ $\boxed{P = 0.00032}$

CONCLUDE: BECAUSE $P = .00032 < 0.05 = \alpha$, REJECT $H_0$. ∴ WE HAVE SUFFICIENT EVIDENCE TO CONCLUDE THAT ANGER LEVEL AND H.D. ARE ASSOCIATED IN POP OF PEOPLE W/NORMAL BLOOD PRESSURE.

✓

## Using Chi-Square Tests Wisely

Both the chi-square test for homogeneity and the chi-square test for association/independence start with a two-way table of observed counts. They even calculate the test statistic, degrees of freedom, and *P*-value in the same way. However, the questions that these two tests answer are different.

- A chi-square test for homogeneity tests whether the distribution of a categorical variable is the same for each of several populations or treatments.
- A chi-square test for association/independence tests whether ~~tow~~ two categorical variables are associated in some population of interest.

*Trick*: Look at how the data were produced to decide which test is appropriate.

- If the data come from two or more independent random samples or treatment groups in a randomized experiment, then do a test for *homogeneity*.
- If the data come from a single random sample, with the individuals classified according to two categorical variables, use a test for *association/independence*.

HANDOUT ON $\chi^2$ TESTS.

DISCUSS VIOLATION OF COUNTS pp. 720-721

HW: Read pp. 709-724; do problems 43, 45, 49, 51, 53-58 on pp. 726-730.

41, 43, 45, 47, 51-56