Chapter 9: Hypothesis Testing for Comparing Population Parameters

Hypothesis testing can address many different types of questions. We are not restricted to looking at the estimated value of a single population parameter. We can also test to determine if population parameters are the same for different populations. Consider the following grade distributions from Fall 2016 in MATH 1107. Do women and men perform the same in MATH 1107 or are there differences in their performances? We could compare final course averages between men and women.

Analysis Variable : course_average course_average						
Gender	N Obs	Mean	Median	Minimum	Maximum	Std Dev
F	78	83.83	88.90	19.20	110.20	19.21
M	42	79.01	81.70	0.00	111.20	19.01

Or we could consider the percentage of letter grades earned by gender.

Table of Gender by Letter_Grade								
Gender(Gender)	Letter_Grade(Letter_Grade)							
Frequency	A	B	C	D	F	W	WF	Total
F	39	20	9	2	3	4	1	78
M	11	16	7	3	2	2	1	42
Total	50	36	16	5	5	6	2	120

In baseball, the roles of pitchers and batters (position players) are different. Pitchers aren't required to be good hitters. However, strong defense in the field will not make up for weak hitting. Does this lead to any differences in these two types of players? Historically this has led to a difference in the average age of batters and pitchers. Since average player age has changed over the years, we need to pair together the average ages by year.

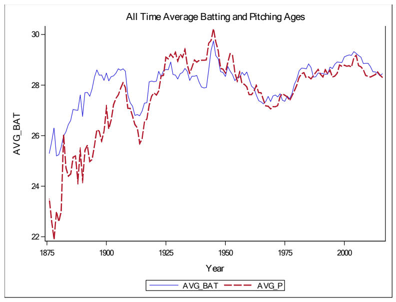

Figure 1: Average Batting and Pitching age by Year

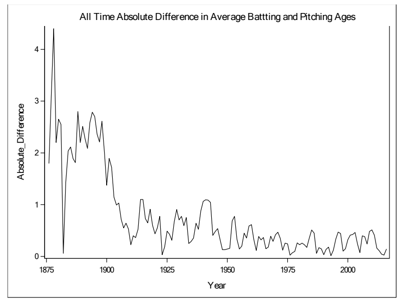

Figure 2: Difference of Batting and Pitching Age by year

1 Hypothesis Testing for Comparing Independent Population Means

Let's run a hypothesis test at 95% confidence to determine if there is a difference in the average score of men and women in MATH 1107. To perform the hypothesis test, we must first translate our question of interest into a statement about the difference of the population parameters. For the performance in MATH 1107 question here, we wonder if the average score of women is the same as the average score of men. Notationally speaking, is $\mu_W = \mu_M$? Since our null hypothesis must always specify a hypothesized value, we translate this to

$$H_0 : \mu_W - \mu_M = 0.$$

The alternative hypothesis, H_A, contains the values of the parameter we accept if we reject the null. In this case we wonder if the averages are different and are conducting a two-tailed test where

$$H_A : \mu_W - \mu_M \neq 0.$$

The general form of the test statistic for the difference of two means is

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

using the sample means, sample standard deviations and sample sizes from populations 1 and 2 in our hypothesis.

At 95% confidence and a two-tailed test, our critical value is ± 1.96. For our average score problem, we get a test statistic of

$$t = \frac{83.83 - 79.01}{\sqrt{\frac{19.21^2}{78} + \frac{19.01^2}{42}}} = 1.319.$$

Since the test statistic does not fall into the rejection region, we do not reject the null hypothesis. That is, we have no statistical evidence to suggest that the average score for women is different from the average score for men in MATH 1107. Of course, the TI 83/84 series calculators can do most of the work for us. To compare independent population means, us the **2-SampTTest** command.

Exercise 1 *Elementary Statistics is a required course for admittance to KSU's highly competitive nursing program. In fact, earning an A in MATH 1107 is practically a requirement for entry into the program. Due to this, do nursing-interest students earn higher average scores than other majors in the course? In a class of 135 students, there were 41 nursing-interest students. The average final score for nursing-interest students was 86.7 with a standard deviation of 14.56. The remaining 94 students averaged 74.8 points with a standard deviation of 19.36. Test at 95% confidence.*

Exercise 2 *The Philadelphia Phillies have a long a storied history in Major League Baseball. Is the average attendance at their games different since the Free Agency era began in 1977 than in prior years? Test at 99% confidence.*

Analysis Variable : Attendance Attendance							
N	**Mean**	**Std Dev**	**Minimum**	**Lower Quartile**	**Median**	**Upper Quartile**	**Maximum**
40	2378399.18	653989.62	1490638.00	1846532.50	2114224.50	2738413.00	3777322.00

Attendance 1977 to 2016

Analysis Variable : Attendance Attendance							
N	**Mean**	**Std Dev**	**Minimum**	**Lower Quartile**	**Median**	**Upper Quartile**	**Maximum**
87	557087.24	459250.59	112066.00	240600.00	341216.00	819698.00	2480150.00

Attendance 1890 to 1976

Is the result necessarily due to Free Agency?

2 Homework

1. Navidi/Monk Section 9.1: 15, 16, 21, 22, 23, 24, 33, 34 (ignore degrees of freedom)

3 Hypothesis Testing for Comparing Population Proportions

Let's return to the distribution of letter grades by gender from Fall 2016 and focus on the letter grade A.

Table of Gender by Letter_Grade								
Gender(Gender)	**Letter_Grade(Letter_Grade)**							
Frequency	**A**	**B**	**C**	**D**	**F**	**W**	**WF**	**Total**
F	39	20	9	2	3	4	1	78
M	11	16	7	3	2	2	1	42
Total	50	36	16	5	5	6	2	120

More women earned a grade of A than men. Is that due to more women taking the class? Or is a women more likely to earn an A than a man in MATH 1107? To put this in terms of hypothesis testing, is the percentage of women who earn the letter grade A the same as the percentage of men? The percentage of women who earned an A is $p_1 = \frac{39}{78} = \frac{1}{2}$. The percentage of men who earned an A is $p_2 = \frac{11}{42} = 0.261\,9$. Test at 99% confidence.

Example 3 *Our null hypothesis takes the form $H_0 : \mu_{PW} - \mu_{PM} = 0$. Since we asked if a women is more likely to earn an A than a man in MATH 1107, we are conducting a one-tailed test and $H_A : \mu_{PW} - \mu_{PM} > 0$. The general form of the test statistic for the difference of two proportions is*

$$z = \frac{p_1 - p_2}{\sqrt{\frac{x_1+x_2}{n_1+n_2} * (1 - \frac{x_1+x_2}{n_1+n_2}) * \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

using the sample proportions and sample sizes from populations 1 and 2 in our hypothesis. For our gender difference problem we get a test statistic of

$$z = \frac{\frac{39}{78} - \frac{11}{42}}{\sqrt{\frac{39+11}{78+42} * (1 - \frac{39+11}{78+42}) * \left(\frac{1}{78} + \frac{1}{42}\right)}} = 2.523$$

Comparing 2.523 to the critical value 2.33 leads us to reject the null hypothesis in favor of the alternative. Thus there is evidence to claim that a women is more likely to earn an A in MATH 1107 than a man. Of course, the TI 83/84 series calculators can do most of the work for us. To compare two population proportions, use the **2-PropZTest** command.

Exercise 4 *The "DFW" rate is a measure of the percentage of students who have not sufficiently mastered the material of a course in order to proceed to the next course in sequence. Using the data above, is the "DFW" rate different for men and women in MATH 1107? Test at 99% confidence.*

Exercise 5 *From our health study data, is the death rate different for men and women? Test at 95% confidence.*

Table of Sex by Status			
Sex	**Status**		
Frequency	**Alive**	**Dead**	**Total**
Female	1977	896	2873
Male	1241	1095	2336
Total	3218	1991	5209

4 Homework

1. Navidi/Monk Section 9.2: 13, 14, 19-22, 29, 30