GPA and Pronunciation

Eleanor Koster and Safiya Easthausen
AP Statistics, Selvaag
2017

Introduction

Is there an association between GPA and the pronunciation of certain words?

We wanted to find if there was an association between GPA and the pronunciation of certain words. We chose 8 words that we thought a lot of people pronounced differently and established two different pronunciations. The 8 words we chose were milk, bagel, february, comfortable, probably, espresso, caramel, and syrup.

Pronunciations

Milk-	P1= milk	P2= melk
Bagel-	P1= bae-gle	P2= bag-le
February-	P1= feb-yoo-air-ee	P2= feb-roo-air-ee
Comfortable-	P1= comf-der-ble	P2= com-fort-able
Probably-	P1= prah-ba-blee	P2= prob-lee
Espresso-	P1= ex-pres-o	P2= es-pres-o
Caramel-	P1= car-mel	P2= care-a-mel
Syrup-	P1= sir-up	P2= seer-up

Data Collection

We decided to collect our sample by taking an SRS of all classrooms during 3rd hour at Washburn (excluding classrooms without a class during this hour). We then used the random digit generator on our calculator to produce six numbers corresponding to six classrooms (on a master list of all classrooms at washburn within our constraints, we assigned each room a number).

We went to each classroom and had the teacher pick 10 students to be part of the sample. Then we separated them, so that we could reduce the possibility of influence one student's answers may have on another's. We recited the same script, "I have a series of cards. When I show you a card, please say the word written on it." Then they were shown the same cards in the same order and asked their GPA.

P1	P2	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2	P1	P2
Milk		Bagel		February		Comfortable		Probably		Espresso		Caramel		Syrup	
X		X			X	X		X			X	X		X	
	X		X	X		X		X			X	X			X
X		X		X			X	X			X	X			X
X		X			X		X	X		X		X			X
X		X		X		X		X			X		X	X	
X			X	X			X	X		X		X		X	
X		X		X			X	X			X	X			X
X		X			X		X	X			X		X		X
	X	X			X	X		X			X	X		X	
X		X		X		X			X		X	X		X	
	X	X		X		X		X		X			X	X	
X			X	X		X		X		X		X		X	
X			X		X	X		X		X			X	X	
X		X		X			X	X			X		X	X	
X		X			X	X		X		X		X		X	
X		X			X		X	X		X			X		X

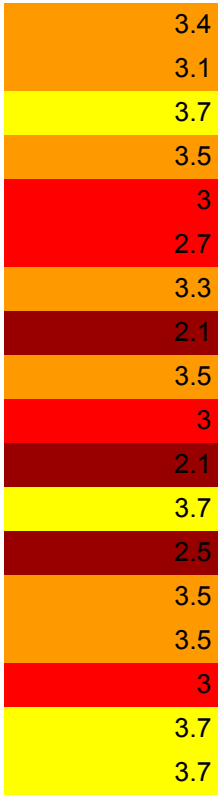
X			X			X			X			X			X			X			X		
X			X		X			X			X					X	X			X		
X			X			X			X				X		X			X			X		
X			X			X			X			X				X		X			X		
X			X			X			X			X				X		X			X		
X			X			X			X			X				X			X		X		
X			X			X			X			X				X		X			X		
X			X		X			X			X				X		X			X		
X			X			X			X			X			X			X				X	
X			X			X			X				X	X				X				X	
X			X			X			X			X				X		X			X		
	X		X			X			X				X			X		X				X	
	X		X			X			X			X				X		X			X		
X			X			X			X			X			X				X			X	
	X		X			X			X			X				X			X			X	
X			X			X			X		X					X						X	
X			X			X			X				X				X		X				
	X		X		X				X				X				X					X	

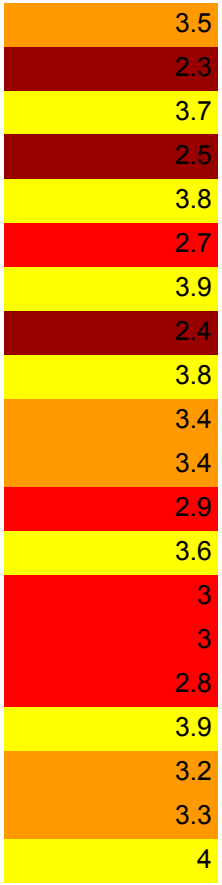
X		X		X		X		X			X	X			X
X		X			X		X	X			X		X		X
X		X		X		X		X			X	X			X
	X		X	X		X		X			X		X		X
X		X			X	X		X			X		X		X
X		X		X			X	X			X		X		X
	X	X		X		X		X		X		X			X
	X		X	X		X		X			X	X			X
	X	X		X			X	X		X		X		X	
	X	X		X			X	X		X		X			X
	X		X	X		X		X			X	X		X	
	X	X		X			X	X			X	X		X	
	X	X		X		X		X			X		X		X
	X	X		X		X		X			X		X		X
X		X			X	X		X			X		X		X
X		X			X		X	X			X		X		X
X		X		X		X		X			X	X			X
X		X			X		X	X			X		X		X
X		X			X	X		X			X	X			X

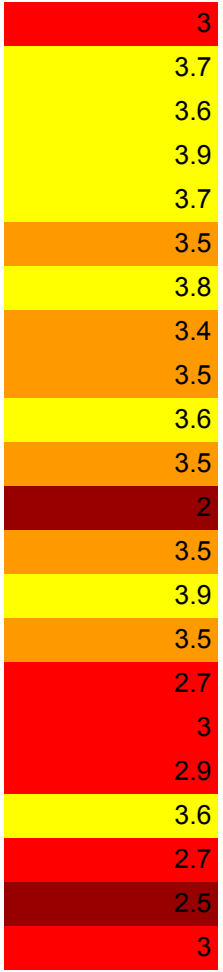
	X	X			X	X		X			X		X		X
	X	X			X	X		X		X		X			X
X		X			X		X	X			X	X			X
X		X		X		X		X		X		X			X
	X	X			X		X	X			X		X	X	
X		X			X	X		X			X	X		X	
X		X		X		X		X			X	X			X
41	19	53	7	34	26	39	21	55	5	20	40	38	22	26	34
3.202439024	3.331578947	3.224528302	3.385714286	3.379411765	3.065384615	3.276923077	3.180952381	3.272727273	2.92	3.21	3.26	3.373684211	3.018181818	3.196153846	3.279411765
68.33%	31.67%	88.33%	11.67%	56.67%	43.33%	65%	35%	91.67%	8.33%	33.33%	66.67%	63.33%	36.67%	43.33%	56.67%

	GPA				
#of times a person said P1	2.0-2.5	2.6-3.0	3.1-3.5	3.6-4.0	total
1	0	0	0	0	0
2	0	0	0	0	0
3	2	2	1	2	7
4	0	4	4	3	11
5	3	4	7	4	18
6	3	5	5	4	17
7	0	0	1	6	7
8	0	0	0	0	0
total	8	15	18	19	60

GPA





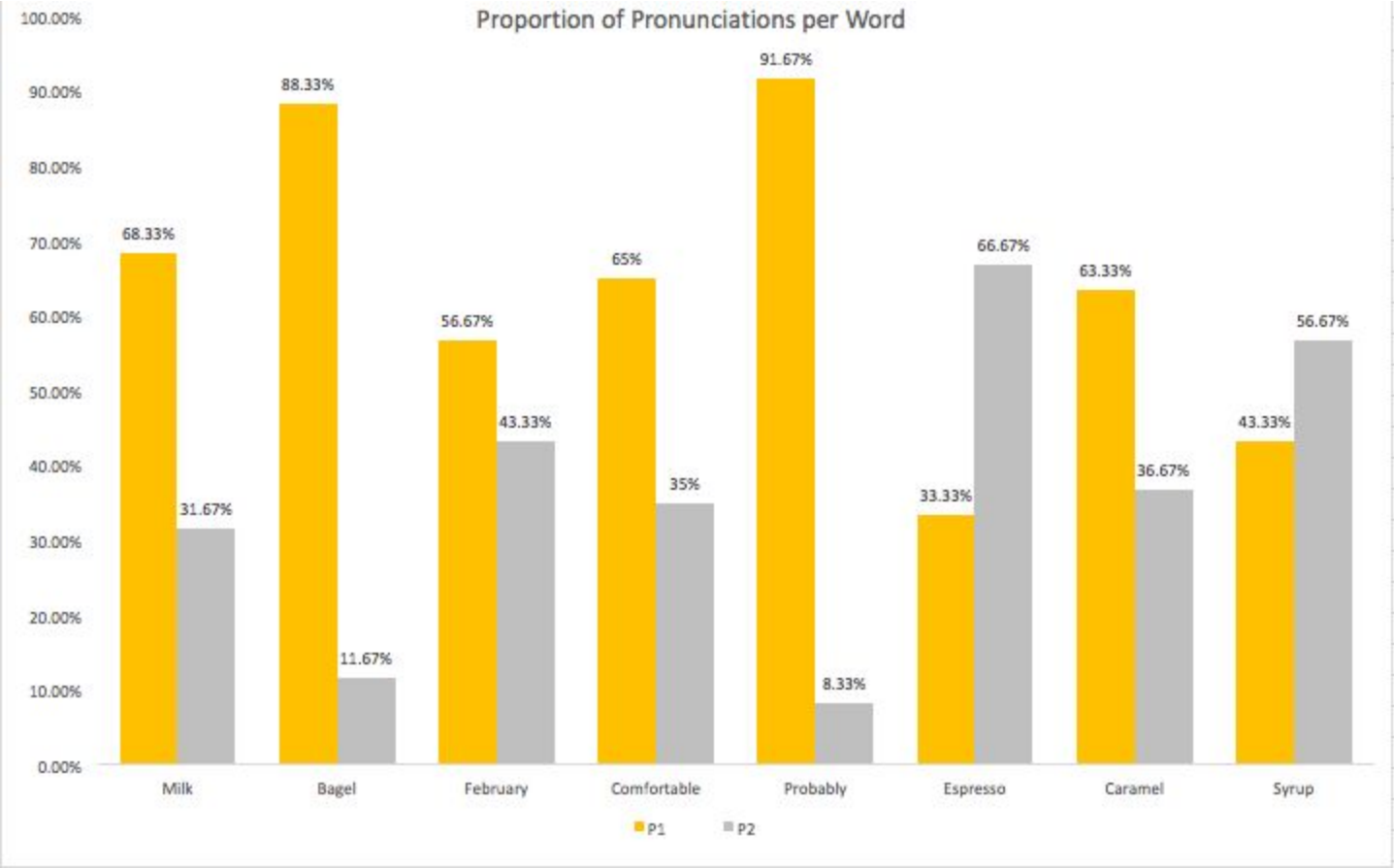


000 = 2.0-2.5

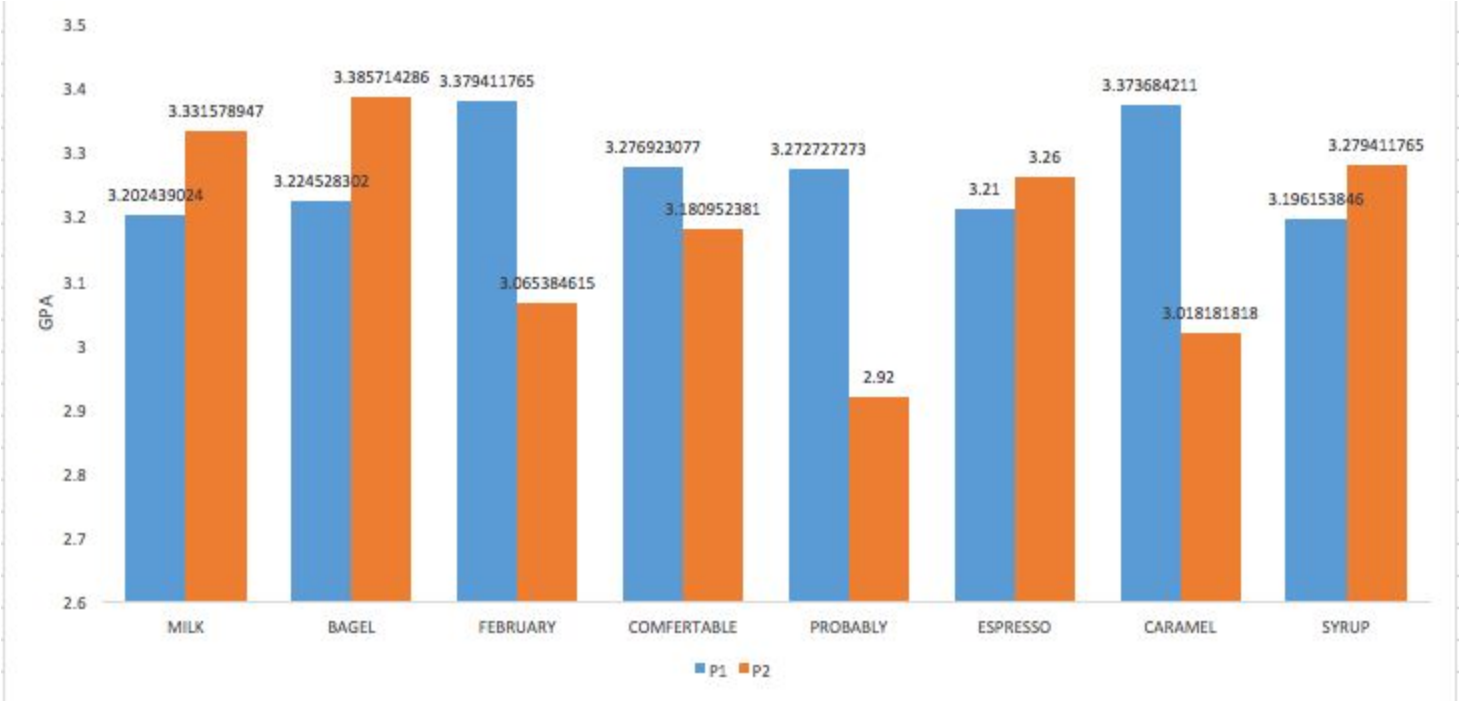
000 = 2.6-3.0

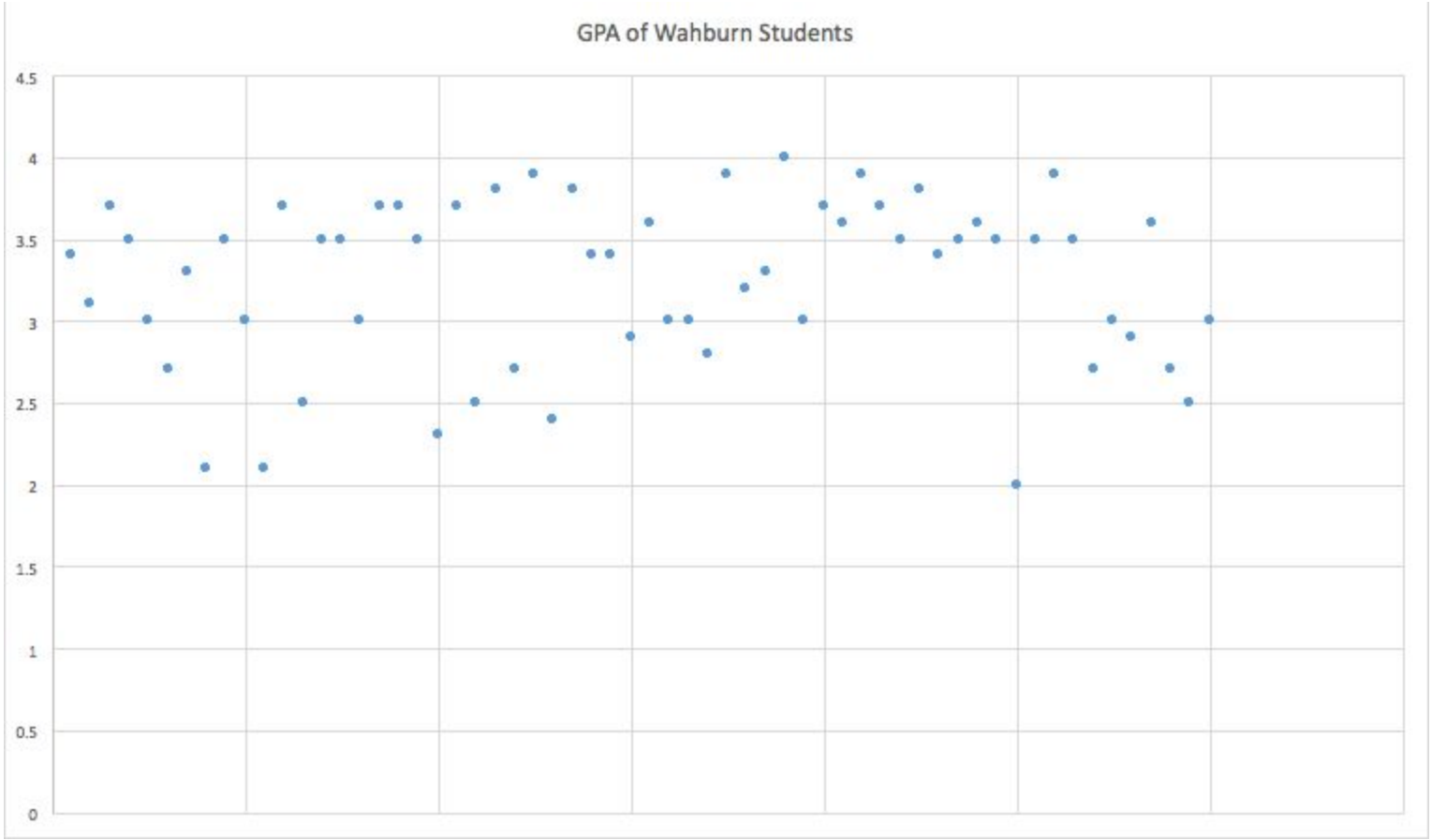
000 = 3.1-3.5

000 = 3.6-4/0


Proportion of Pronunciations per Word
100.00%
90.00%
80.00%
70.00%
60.00%
50.00%
40.00%
30.00%
20.00%
10.00%
0.00%
68.33%
31.67%
88.33%
11.67%
56.67%
43.33%
65%
35%
91.67%
8.33%
33.33%
66.67%
63.33%
36.67%
43.33%
56.67%
Milk
Bagel
February
Comfortable
Probably
Espresso
Caramel
Syrup
P1
P2

Average GPA per pronunciation




GPA of Wahburn Students
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0

Inference (One-Proportion Z Test)

Parameter- p: The proportion of all Washburn students who pronounce milk as P1.

Hypothesis- H_o: p = 0.5 The proportion of all washburn students who pronounce milk as P1 is 0.5 H_a: p > 0.5 The proportion of Washburn students who pronounce milk as P1 is greater than 0.5

Assumptions- Random- We took an SRS of 6 classrooms at Washburn.
Normal- LCC $np_o \geq 10$ $n(1-p_o) \geq 10$ $(60)(0.5) \geq 10 = 30 \geq 10$ $(60)(1-0.5) \geq 10 = (60)(0.5) \geq 10 = 30 \geq 10$ Independent- 10% Condition $n \leq 1/10N$ $60 \leq 1/10N = 600 \leq N$ There's more than 600 students at Washburn

Name- We used a one-proportion z test.

Test Statistic- $z=(\hat{p}-p_o)/\sqrt{(p_o(1-p_o)/n)}$ $z=(0.6833-0.5)/\sqrt{(0.5(0.5)/60)}=2.84$

P-value- From the table the p-value is 0.0023

Decision- Because the P-value is significant at the $\alpha=0.05$ level, we reject the null hypothesis

Conclusion- There is strong evidence that the proportion of Washburn students who pronounce milk as P1 is greater than 0.5.

Inference (One-Proportion Z Test)

Parameter- p: The proportion of all Washburn students who pronounce bagel as P1.

Hypothesis- H_o: p = 0.5 The proportion of all washburn students who pronounce bagel as P1 is 0.5 H_a: p > 0.5 The proportion of Washburn students who pronounce bagel as P1 is greater than 0.5

Assumptions- Random- We took an SRS of 6 classrooms at Washburn.
Normal- LCC $np_o \geq 10$ $n(1-p_o) \geq 10$ $(60)(0.5) \geq 10 = 30 \geq 10$ $(60)(1-0.5) \geq 10 = (60)(0.5) \geq 10 = 30 \geq 10$ Independent- 10% Condition $n \leq 1/10N$ $60 \leq 1/10N = 600 \leq N$ There's more than 600 students at Washburn

Name- We used a one-proportion z test.

Test Statistic- $z=(\hat{p}-p_o)/\sqrt{(p_o(1-p_o)/n)}$ $z=(0.8833-0.5)/\sqrt{(0.5(0.5)/60)}=5.94$

P-value- From the table the p-value is less than 0.0002

Decision- Because the P-value is significant at the $\alpha=0.05$ level, we the null hypothesis.

Conclusion- There is strong evidence that the proportion of Washburn students who pronounce bagel as P1 is greater than 0.5.

Inference (One-Proportion Z Test)

Parameter- p: The proportion of all Washburn students who pronounce february as P1.

Hypothesis- H_o: p = 0.5 The proportion of all washburn students who pronounce february as P1 is 0.5 H_a: p > 0.5 The proportion of Washburn students who pronounce february as P1 is greater than 0.5

Assumptions- Random- We took an SRS of 6 classrooms at Washburn.
Normal- LCC $np_o \geq 10$ $n(1-p_o) \geq 10$ $(60)(0.5) \geq 10 = 30 \geq 10$ $(60)(1-0.5) \geq 10 = (60)(0.5) \geq 10 = 30 \geq 10$ Independent- 10% Condition $n \leq 1/10N$ $60 \leq 1/10N = 600 \leq N$ There's more than 600 students at Washburn

Name- We used a one-proportion z test.

Test Statistic- $z=(\hat{p}-p_o)/\sqrt{(p_o(1-p_o)/n)}$ $z=(0.5667-0.5)/\sqrt{(0.5(0.5)/60)}=1.03$

P-value- From the table the p-value is 0.1515

Decision- Because the P-value is not significant at the α=0.05 level, we fail to reject the null hypothesis

Conclusion- There is not strong evidence that the proportion of Washburn students who pronounce february as P1 is greater than 0.5

Inference (One-Proportion Z Test)

Parameter- p: The proportion of all Washburn students who pronounce comfortable as P1.

Hypothesis- H_o: p = 0.5 The proportion of all washburn students who pronounce comfortable as P1 is 0.5 H_a: p > 0.5 The proportion of Washburn students who pronounce comfortable as P1 is greater than 0.5

Assumptions- Random- We took an SRS of 6 classrooms at Washburn.
Normal- LCC $np_o \geq 10$ $n(1-p_o) \geq 10$ $(60)(0.5) \geq 10 = 30 \geq 10$ $(60)(1-0.5) \geq 10 = (60)(0.5) \geq 10 = 30 \geq 10$ Independent- 10% Condition $n \leq 1/10N$ $60 \leq 1/10N = 600 \leq N$ There's more than 600 students at Washburn

Name- We used a one-proportion z test.

Test Statistic- $z=(\hat{p}-p_o)/\sqrt{(p_o(1-p_o)/n)}$ $z=(0.65-0.5)/\sqrt{(0.5(0.5)/60)}=2.32$

P-value- From the table the p-value is 0.0102

Decision- Because the P-value is significant at the $\alpha=0.05$ level, we reject the null hypothesis

Conclusion- There is strong evidence that the proportion of Washburn students who pronounce comfortable as P1 is greater than 0.5

Inference (One-Proportion Z Test)

Parameter- p: The proportion of all Washburn students who pronounce probably as P1.

Hypothesis- H_o: p = 0.5 The proportion of all washburn students who pronounce probably as P1 is 0.5 H_a: p > 0.5 The proportion of Washburn students who pronounce probably as P1 is greater than 0.5

Assumptions- Random- We took an SRS of 6 classrooms at Washburn.
Normal- LCC $np_o \geq 10$ $n(1-p_o) \geq 10$ $(60)(0.5) \geq 10 = 30 \geq 10$ $(60)(1-0.5) \geq 10 = (60)(0.5) \geq 10 = 30 \geq 10$ Independent- 10% Condition $n \leq 1/10N$ $60 \leq 1/10N = 600 \leq N$ There's more than 600 students at Washburn

Name- We used a one-proportion z test.

Test Statistic- $z=(\hat{p}-p_o)/\sqrt{(p_o(1-p_o)/n)}$ $z=(0.9167-0.5)/\sqrt{(0.5(0.5)/60)}=6.46$

P-value- From the table the p-value is less than 0.0002

Decision- Because the P-value is significant at the α=0.05 level, we reject the null hypothesis

Conclusion- There is strong evidence that the proportion of Washburn students who pronounce probably as P1 is greater than 0.5

Inference (One-Proportion Z Test)

Parameter- p: The proportion of all Washburn students who pronounce espresso as P1.

Hypothesis- H_o: p = 0.5 The proportion of all washburn students who pronounce espresso as P1 is 0.5 H_a: p > 0.5 The proportion of Washburn students who pronounce espresso as P1 is greater than 0.5

Assumptions- Random- We took an SRS of 6 classrooms at Washburn.
Normal- LCC $np_o \geq 10$ $n(1-p_o) \geq 10$ $(60)(0.5) \geq 10 = 30 \geq 10$ $(60)(1-0.5) \geq 10 = (60)(0.5) \geq 10 = 30 \geq 10$ Independent- 10% Condition $n \leq 1/10N$ $60 \leq 1/10N = 600 \leq N$ There's more than 600 students at Washburn

Name- We used a one-proportion z test.

Test Statistic- $z=(\hat{p}-p_o)/\sqrt{(p_o(1-p_o)/n)}$ $z=(0.3333-0.5)/\sqrt{(0.5(0.5)/60)}=-2.58$

P-value- From the table the p-value is 0.9951

Decision- Because the P-value is not significant at the α=0.05 level, we fail to reject the null hypothesis

Conclusion- There is not strong evidence that the proportion of Washburn students who pronounce espresso as P1 is greater than 0.5

Inference (One-Proportion Z Test)

Parameter- p: The proportion of all Washburn students who pronounce caramel as P1.

Hypothesis- H_o: p = 0.5 The proportion of all washburn students who pronounce caramel as P1 is 0.5 H_a: p > 0.5 The proportion of Washburn students who pronounce caramel as P1 is greater than 0.5

Assumptions- Random- We took an SRS of 6 classrooms at Washburn.
Normal- LCC $np_o \geq 10$ $n(1-p_o) \geq 10$ $(60)(0.5) \geq 10 = 30 \geq 10$ $(60)(1-0.5) \geq 10 = (60)(0.5) \geq 10 = 30 \geq 10$ Independent- 10% Condition $n \leq 1/10N$ $60 \leq 1/10N = 600 \leq N$ There's more than 600 students at Washburn

Name- We used a one-proportion z test.

Test Statistic- $z=(\hat{p}-p_o)/\sqrt{(p_o(1-p_o)/n)}$ $z=(0.6333-0.5)/\sqrt{(0.5(0.5)/60)}=2.07$

P-value- From the table the p-value is 0.0192

Decision- Because the P-value is significant at the $\alpha=0.05$ level, we reject the null hypothesis

Conclusion- There is strong evidence that the proportion of Washburn students who pronounce caramel as P1 is greater than 0.5

Inference (One-Proportion Z Test)

Parameter- p: The proportion of all Washburn students who pronounce syrup as P1.

Hypothesis- H_o: p = 0.5 The proportion of all washburn students who pronounce syrup as P1 is 0.5 H_a: p > 0.5 The proportion of Washburn students who say syrup pronounce syrup as P1 is greater than 0.5

Assumptions- Random- We took an SRS of 6 classrooms at Washburn.
Normal- LCC $np_o \geq 10$ $n(1-p_o) \geq 10$ $(60)(0.5) \geq 10 = 30 \geq 10$ $(60)(1-0.5) \geq 10 = (60)(0.5) \geq 10 = 30 \geq 10$ Independent- 10% Condition $n \leq 1/10N$ $60 \leq 1/10N = 600 \leq N$ There's more than 600 students at Washburn

Name- We used a one-proportion z test.

Test Statistic- $z=(\hat{p}-p_o)/\sqrt{(p_o(1-p_o)/n)}$ $z=(0.4333-0.5)/\sqrt{(0.5(0.5)/60)}=-1.03$

P-value- From the table the p-value is 0.8485

Decision- Because the P-value is not significant at the α=0.05 level, we fail to reject the null hypothesis

Conclusion- There is not strong evidence that the proportion of Washburn students who say syrup pronounce syrup as P1 than 0.5

Inference (One-Sample T Test)

Parameter- μ: The mean GPA of all Washburn students.

Hypotheses- H_o: μ = 3.0 The mean GPA of all Washburn students is 3.0 H_a: μ > 3.0 The mean GPA of all Washburn students is greater than 3.0

Assumptions- Random- We took an SRS of 6 classrooms at Washburn.
Normal- CLT n≥30 60≥30
Independent- 10% Condition n≤1/10N 60≤1/10N=600 There's more than 600 students at Washburn

Name- Since σ is unknown we will use a one-sample t test

Test statistic- $t= (\bar{x}-\mu_o) / (s/\sqrt{n})$ t=(3.24-3.0)/(0.508/√60)=3.66 df=n-1 df=60-1=59

P-value- From the table the p-value is less than 0.0005

Decision- Because the P-value is significant at the α=0.05 level, we reject the null hypothesis

Conclusion- There is strong evidence that the mean GPA of all Washburn students is greater than 3.0

Inference (Chi-Square Test of Association)

Parameter- We are interested in is there is an association between GPA and the pronunciation of certain words

Hypotheses- H_o: There is no association between GPA and the pronunciation of certain words

H_a: There is an association between GPA and the pronunciation of certain words

Assumptions- Random- We took an SRS of 6 classrooms at Washburn.
Normal- 10% Condition $n \leq 1/10N$ $60 \leq 1/10N=600$ There's more than 600 students at Washburn
Independent- Not all expected counts are at least 5

Name- We will use a chi-square test of association

Test statistic- $x^2=\sum (O-E)^2/E$ $x^2=\sum(0-0)^2/0+(0-0)^2/0+(0-0)^2/0+...=15.73300639$ $df=(n-1)(n-1)$ $df=(8-1)(4-1)=21$

P-value- From the table the p-value is greater than 0.25

Decision- Because the P-value is not significant at the $\alpha=0.05$ level, we fail to reject the null hypothesis

Conclusion- There is not strong evidence that there is an association between GPA and the pronunciation of certain words

Observed Counts

	GPA				
#of times a person said P1	2.0-2.5	2.6-3.0	3.1-3.5	3.6-4.0	total
1	0	0	0	0	0
2	0	0	0	0	0
3	2	2	1	2	7
4	0	4	4	3	11
5	3	4	7	4	18
6	3	5	5	4	17
7	0	0	1	6	7
8	0	0	0	0	0
total	8	15	18	19	60

Expected Counts

	GPA				
#of times a person said P1	2.0-2.5	2.6-3.0	3.1-3.5	3.6-4.0	total
1	0	0	0	0	0
2	0	0	0	0	0
3	0.93333	1.75	2.1	2.2167	7
4	1.4667	2.75	3.3	3.4833	11
5	2.4	4.5	5.4	5.7	18
6	2.2667	4.25	5.1	5.3833	17
7	0.93333	1.75	2.1	2.2167	7
8	0	0	0	0	0
total	8	15	18	19	60

Conclusion

- The proportion of Washburn students who pronounce milk as P1 is greater than 0.5
- The proportion of Washburn students who pronounce bagel as P1 is greater than 0.5
- The proportion of Washburn students who pronounce february as P1 is not greater than 0.5
- The proportion of Washburn students who pronounce comfortable as P1 is greater than 0.5
- The proportion of Washburn students who pronounce probably as P1 is greater than 0.5
- The proportion of Washburn students who pronounce espresso as P1 is not greater than 0.5
- The proportion of Washburn students who pronounce caramel as P1 is greater than 0.5
- The proportion of Washburn students who say syrup pronounce syrup as P1 is not greater than 0.5
- The mean GPA of all Washburn students is greater than 3.0
- There is no association between GPA and the pronunciation of certain words

Limitations/Error of Study

Some of the limitations we had were we had a pretty small sample size and we had a limited amount of time to collect data. Some causes of error within our study were that some people either didn't want to tell us their GPA or didn't know their GPA. We also had one class that didn't want to participate in the study. Another thing that could have affected our data is that the people talked to each other about what we were doing and figured it out and that might have changed their answers.