# Chi Square Practice Problems

Solve all problems using a chi square analysis. You must use statistics to support your answers.

1. A zookeeper hypothesizes that changing the intensity of the light in the primate exhibits will reduce the amount of aggression between the baboons. In exhibit A, with a lower light intensity, he observes 36 incidences of aggression over a one month period. In exhibit B, with normal lights, he observes 42 incidences of aggression. Should he support or reject his hypothesis?

The null hypothesis is that we would expect no difference in number of acts of aggression when comparing low light to normal light. There were 78 total acts of aggression. So we would expect 39 acts to occur in each of the 2 rooms (78/2).

|  |  |  |
| --- | --- | --- |
| Observed | Expected | (o-e)2/e |
| 36 | 39 | .23 |
| 42 | 39 | .23 |

X2 = .46

Degrees of freedom = 1

Critical value = 3.84

Because .46 is below 3.84, we can accept the null hypothesis. (Which proves the zookeeper’s hypothesis as incorrect.)

2. A high school, students can choose to enter one of three doors. Custodians noticed that door #3 was always getting broken and suggested that more students use that door because it has a hands-free opener. Science minded students counted the number of students entering each door to see if the custodians were right.

Door #1 had 60 students enter | Door #2 had 66 students enter | Door #3 had 80 students enter.

Were the custodians right?

The null hypothesis is that we would expect no difference in which door is preferred by students. The doors were opened 206 times total. 206/3 = 68.67.

|  |  |  |
| --- | --- | --- |
| Observed | Expected | (o-e)2/e |
| 60 | 69 | 1.17 |
| 66 | 69 | 0.13 |
| 80 | 69 | 1.75 |

X2 = 3.05

Degrees of freedom = 2

Critical value = 5.99

Because 3.05 is below 5.99, we can accept the null hypothesis. The students do not prefer any door.

3. A scientist predicts that the kittens born with a congenital birth defect will be 25% based on the hypothesis that it is caused be a recessive gene in that breed of cat. After surveying several litters, he found that 44 out of 125 kittens had the defect. Is his hypothesis correct?

The null hypothesis is that we would expect the congenital birth defect not to be recessive. Recessive traits show up 25% of the time in heterozygous communities. 25% of 125 is 31.25.

|  |  |  |
| --- | --- | --- |
| Observed | Expected | (o-e)2/e |
| 44 | 31.25 | 5.202 |

X2 = 5.202

Degrees of freedom = 1

Critical value = 3.84

Because 5.202 is above 3.84, so we reject the null hypothesis. (Which supports the scientist’s hypothesis.)

4. Suppose you take a random sample of 30 students who are using a new math text and a second sample of 30 students who are using a more traditional text. You compare student achievement on the state test given to all students at the end of the course. Based on state test performance, would you recommend the new math book?

|  |  |  |
| --- | --- | --- |
|  | Passed State Test | Failed State Test |
| New Textbook | 26 | 4 |
| Old Textbook | 22 | 8 |

The null hypothesis is that there is no significant difference between the number of students who passed that used the new and old books. 48 tests were passed. 48/2 = 24

|  |  |  |
| --- | --- | --- |
| Observed | Expected | (o-e)2/e |
| 26 | 24 | 0.167 |
| 22 | 24 | 0.167 |

X2 = 0.334

Degrees of freedom = 1

Critical value = 3.84

Because 0.334 is below 3.84 we accept the null hypothesis. (The new math books have no affect on students pass rate – so likely shouldn’t recommend buying the books.)

Practice Problems: Standard Deviations and Variance Answers

**6.** Calculate the standard deviation for the following three sample data sets:

a. 13 21 27 31 35 24 28 32 17 20 *sd = 7.05*

b. 100 115 112 113 95 87 90 104 107 98 *sd = 9.76*

c. 55 54 59 55 52 51 57 49 61 57 *sd = 3.68*

Consider the following three data sets A, B and C.

A = {9,10,11,7,13}

B = {10,10,10,10,10}

C = {1,1,10,19,19}

a) Calculate the mean of each data set.

b) Calculate the standard deviation of each data set.

c) Which set has the largest standard deviation?

d) Is it possible to answer question c) without calculations of the standard deviation?

A.B. Data set A

mean = (9+10+11+7+13)/5 = 10

Standard Deviation=((9-10)2+(10-10)2+(11-10)2+(7-10)2+(13-10)2 )/5 = 4 = √4 = 2

Data set B

mean = (10+10+10+10+10)/5 = 10
Standard Deviation =((10-10)2+(10-10)2+(10-10)2+(10-10)2+(10-10)2 )/5=0 =√0=0

 Data set C

mean = (1+1+10+19+19)/5 = 10

Standard Deviation =((1-10)2+(1-10)2+(10-10)2+(19-10)2+(19-10)2 )/5 = 64.8
=√64.8= 8.05

C. Data Set C

D. Yes, just by looking at the range of the numbers. Standard deviation is a measure of how spread out the numbers are – so we can see right off that B is going to have the smallest deviation and A has the largest difference between the numbers so it will have the largest standard deviation.