Name: _____ Hour: _____ Date: _____

Lesson 11.1: Day 2: Which color M&M is the most common part two?


The company that makes milk chocolate M&Ms claims the following distribution:
13% Brown, 14% Yellow, 20% Orange, 16% Green, 24% Blue, and 13% Red. Is this true?

1. Record the information from yesterday.

Observed values: Brown:_____ Yellow:_____ Orange:_____ Green:_____ Blue:_____ Red:_____

Expected values: Brown:_____ Yellow:_____ Orange:_____ Green:_____ Blue:_____ Red:_____

Test statistic: χ^2= _____

2. Check conditions:

Random: M n M's were randomly sampled.

Large counts: Which expected count is the lowest? Are all of the expected counts greater than 5?

Lowest Expected: Brown: ≥5 All expected counts ≥5 ✓

3. Calculate the P-value.

For this test *df* = *n* – 1, but n represents the number of 6 colors categories (colors).

What is the *df* for this test? 5

What is the test statistic for this test? χ^2 =

Use Table C to find the P-value:_____ Go to row df and move until you find the closest values to your χ^2, match to tail prob.

4. Make a conclusion. Use $\alpha = 0.05$.

Assuming the company's claimed color distribution is true, there is a ____ probability of getting a χ^2 value of ___ or greater purely by chance. This is/is not statistically significant. We fail to reject/reject the null hypothesis.

Follow up

5. Which color M&M had an observed value the farthest from the expected?

Color was _____ higher/lower than expected.

Name: _____ Hour: _____ Date: _____

Do the data provide significant evidence that the company was lying about the distribution of colors of M&Ms? Use $\alpha = 0.05$

STATE: Hypotheses: Significance level: .05

H_0: The claimed color distribution is true.

H_a: The claimed color distribution is not true.

PLAN: Name of procedure: chi-square test for goodness of fit χ^2 G.O.F.

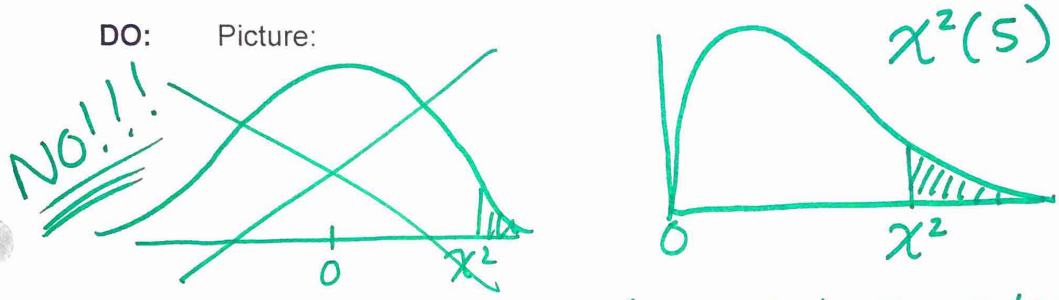
Check conditions:

Random: The MnM's were randomly sampled

Large Counts: All expected counts ≥ 5 ✓

Lowest: ___ ≥ 5

DO: Picture:



Specific Formula: $\chi^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{expected}}$

Work:

$$\chi^2 = \underset{\text{Brown}}{\frac{(O-E)^2}{E}} + \ldots \underset{\text{Red}}{\frac{(O-E)^2}{E}} =$$

Test statistic: $\chi^2 =$

P-value:

CONCLUDE: Assuming the claimed color distribution is true there is a ___ probability of getting a $\chi^2 =$ ___ or more purely by chance. This is / is not statistically significant. We reject/fail to reject the null and can/cannot conclude the alternative.

What parts of the usual 4-step process are missing in this test?

Parameter, statistic, General formula

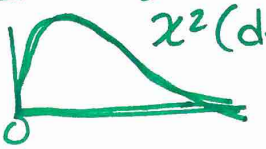
Name: _____ Hour: _____ Date: _____

Lesson 11.1: Day 2: Chi-Square Test for Goodness of Fit: 4 Steps

Important ideas:

LT#1 Conditions
① Random
② Large Counts:
All expected counts ≥ 5
Show lowest count.

LT#2 χ^2 Distribution



$\chi^2(df)$ - Always Right skewed
- Starts at 0.

LT#3 4 steps & Follow Up
State, plan, do, conclude.
If significant, discuss which values were largest contributor to χ^2.

Check Your Understanding

Does the warm, sunny weather in Arizona affect a driver's choice of car color? Cass thinks that Arizona drivers might opt for a lighter color with the hope that it will reflect some of the heat from the sun. To see if the distribution of car colors in Oro Valley, near Tucson, is different from the distribution of car colors across North America, she selected a random sample of 300 cars in Oro Valley. The table shows the distribution of car color for Cass's sample in Oro Valley and the distribution of car color in North America, according to www.ppg.com.

Color	White	Black	Gray	Silver	Red	Blue	Green	Other	Total
Oro Valley sample	84	38	31	46	27	29	6	39	300
North America	23%	18%	16%	15%	10%	9%	2%	7%	100%

1. Do these data provide convincing evidence that the distribution of car color in Oro Valley differs from the North American distribution?

STATE: Hypotheses: Significance level: $\alpha = .05$

① The distribution in Oro Valley is the same as N.A. distribution.
② Distribution in Oro Valley is not the same as N.A.

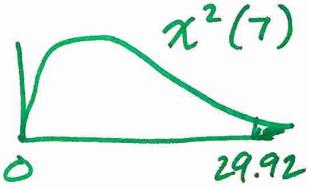
PLAN: Name of procedure: chi-square test for goodness of fit χ^2 GOF

Check conditions: Random: "Random sample of 300"

Large Counts: Lowest expected count = $300 \times .07 = 21 \geq 5$ ✓
All expected counts ≥ 5 ✓

DO: Specific Formula: $\chi^2 = \frac{(Obs. - Exp.)^2}{Exp}$

Picture:



Work: $\frac{(84-69)^2}{69} + \ldots + \frac{(39-21)^2}{21} = 29.92$

Test statistic: $\chi^2 = 29.92$

P-value: Less than .0005

CONCLUDE: Since the p-value is < .05, we have convincing evidence against the null. We reject the null and conclude the distribution is not the same in Oro Valley as it is in N.A.

2. If there is convincing evidence of a difference in the distribution of car color, perform a follow-up analysis. The largest components of the χ^2 test statistic come from "other" which was 18 more than expected, and "Gray" which was 17 below expected.