# AP Statistics

## Chapter 11: Inference for Distribution of Categorical Data

## Day 2

**HW: Finish Lesson 2 Practice Worksheet**

$\chi^2 = \frac{(13-10)^2}{10} + \frac{(11-10)^2}{10} + \frac{(6-10)^2}{10} + \frac{(13-10)^2}{10} + \frac{(10-10)^2}{10} + \frac{(8-10)^2}{10}$

$\chi^2 = .9 + .1 + 1.6 + .4 + 0 + .4 = \boxed{3.4}$

2. The Mars Company reports that their M&M Peanuts Chocolate Candies are produced according to the following color distribution: 23% of each blue and orange, 15% each of green and yellow, and 12% each of red and brown. Joey bought a bag of Peanut Chocolate Candies and counted the colors of the candies in his sample: 12 blue, 7 orange, 13 green, 4 yellow, 8 red, and 2 brown.

- State the appropriate hypotheses for testing the company's claim about the color distribution of peanut M&M's.

  $H_0$: The company's stated color distribution for Peanut candies is correct.

  $H_a$: The company's stated color distribution of Peanut candies is incorrect.

- Calculate the expected count for each color, assuming that the company's claim is true.

  Candies: Blue Orange Green Yellow Red Brown

  Expected: 10.58 10.58 6.9 6.9 5.52 5.52

- Calculate the chi-square statistic for Joey's sample.

$\chi^2 = \frac{(13-10.58)^2}{10.58} + \frac{(7-10.58)^2}{10.58} + \frac{(13-6.9)^2}{6.9} + \frac{(4-6.9)^2}{6.9} + \frac{(8-5.52)^2}{5.52} + \frac{(2-5.52)^2}{5.52}$

$\chi^2 = .19059 + 1.2114 + 5.3938 + 1.2188 + 1.1147 + 2.2446 = \boxed{11.3734}$

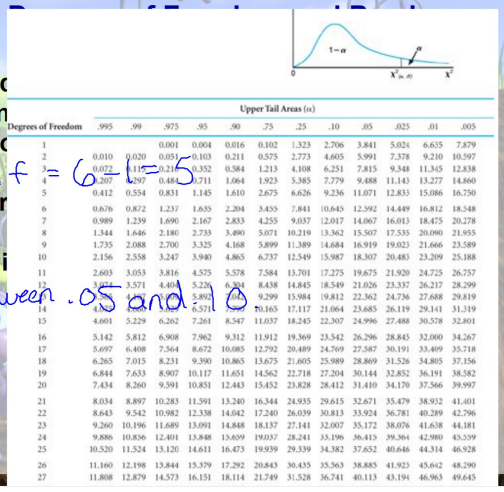
- To find
- goodn
- catego

$df = 6 - 1 = 5$

In yester

- What i

Between .05 and .10




Upper Tail Areas ($\alpha$)

| Degrees of Freedom | .995 | .99 | .975 | .95 | .90 | .75 | .25 | .10 | .05 | .025 | .01 | .005 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | 0.001 | 0.004 | 0.016 | 0.102 | 1.323 | 2.706 | 3.841 | 5.024 | 6.635 | 7.879 |
| 2 | 0.010 | 0.020 | 0.051 | 0.103 | 0.211 | 0.575 | 2.773 | 4.605 | 5.991 | 7.378 | 9.210 | 10.597 |
| 3 | 0.072 | 0.115 | 0.216 | 0.352 | 0.584 | 1.213 | 4.108 | 6.251 | 7.815 | 9.348 | 11.345 | 12.838 |
| 4 | 0.207 | 0.297 | 0.484 | 0.711 | 1.064 | 1.923 | 5.385 | 7.779 | 9.488 | 11.143 | 13.277 | 14.860 |
| 5 | 0.412 | 0.554 | 0.831 | 1.145 | 1.610 | 2.675 | 6.626 | 9.236 | 11.071 | 12.833 | 15.086 | 16.750 |
| 6 | 0.676 | 0.872 | 1.237 | 1.635 | 2.204 | 3.455 | 7.841 | 10.645 | 12.592 | 14.449 | 16.812 | 18.548 |
| 7 | 0.989 | 1.239 | 1.690 | 2.167 | 2.833 | 4.255 | 9.037 | 12.017 | 14.067 | 16.013 | 18.475 | 20.278 |
| 8 | 1.344 | 1.646 | 2.180 | 2.733 | 3.490 | 5.071 | 10.219 | 13.362 | 15.507 | 17.535 | 20.090 | 21.955 |
| 9 | 1.735 | 2.088 | 2.700 | 3.325 | 4.168 | 5.899 | 11.389 | 14.684 | 16.919 | 19.023 | 21.666 | 23.589 |
| 10 | 2.156 | 2.558 | 3.247 | 3.940 | 4.865 | 6.737 | 12.549 | 15.987 | 18.307 | 20.483 | 23.209 | 25.188 |
| 11 | 2.603 | 3.053 | 3.816 | 4.575 | 5.578 | 7.584 | 13.701 | 17.275 | 19.675 | 21.920 | 24.725 | 26.757 |
| 12 | 3.074 | 3.571 | 4.404 | 5.226 | 6.304 | 8.438 | 14.845 | 18.549 | 21.026 | 23.337 | 26.217 | 28.299 |
| 13 | 3.565 | 4.107 | 5.009 | 5.892 | 7.042 | 9.299 | 15.984 | 19.812 | 22.362 | 24.736 | 27.688 | 29.819 |
| 14 | 4.075 | 4.660 | 5.629 | 6.571 | 7.790 | 10.165 | 17.117 | 21.064 | 23.685 | 26.119 | 29.141 | 31.319 |
| 15 | 4.601 | 5.229 | 6.262 | 7.261 | 8.547 | 11.037 | 18.245 | 22.307 | 24.996 | 27.488 | 30.578 | 32.801 |
| 16 | 5.142 | 5.812 | 6.908 | 7.962 | 9.312 | 11.912 | 19.369 | 23.542 | 26.296 | 28.845 | 32.000 | 34.267 |
| 17 | 5.697 | 6.408 | 7.564 | 8.672 | 10.085 | 12.792 | 20.489 | 24.769 | 27.587 | 30.191 | 33.409 | 35.718 |
| 18 | 6.265 | 7.015 | 8.231 | 9.390 | 10.865 | 13.675 | 21.605 | 25.989 | 28.869 | 31.526 | 34.805 | 37.156 |
| 19 | 6.844 | 7.633 | 8.907 | 10.117 | 11.651 | 14.562 | 22.718 | 27.204 | 30.144 | 32.852 | 36.191 | 38.582 |
| 20 | 7.434 | 8.260 | 9.591 | 10.851 | 12.443 | 15.452 | 23.828 | 28.412 | 31.410 | 34.170 | 37.566 | 39.997 |
| 21 | 8.034 | 8.897 | 10.283 | 11.591 | 13.240 | 16.344 | 24.935 | 29.615 | 32.671 | 35.479 | 38.932 | 41.401 |
| 22 | 8.643 | 9.542 | 10.982 | 12.338 | 14.042 | 17.240 | 26.039 | 30.813 | 33.924 | 36.781 | 40.289 | 42.796 |
| 23 | 9.260 | 10.196 | 11.689 | 13.091 | 14.848 | 18.137 | 27.141 | 32.007 | 35.172 | 38.076 | 41.638 | 44.181 |
| 24 | 9.886 | 10.856 | 12.401 | 13.848 | 15.659 | 19.037 | 28.241 | 33.196 | 36.415 | 39.364 | 42.980 | 45.559 |
| 25 | 10.520 | 11.524 | 13.120 | 14.611 | 16.473 | 19.939 | 29.339 | 34.382 | 37.652 | 40.646 | 44.314 | 46.928 |
| 26 | 11.160 | 12.198 | 13.844 | 15.379 | 17.292 | 20.843 | 30.435 | 35.563 | 38.885 | 41.923 | 45.642 | 48.290 |
| 27 | 11.808 | 12.879 | 14.573 | 16.151 | 18.114 | 21.749 | 31.528 | 36.741 | 40.113 | 43.194 | 46.963 | 49.645 |

## Conditions for Chi-Square Goodness-of-Fit Tests

- The data comes from a random sample or a randomized experiment.
- **All expected counts are at least 5.**
- The population is at least 10 times as large as the sample (10% condition).

## When Were You Born?

**Are births evenly distributed across the days of the week? The one-way table below shows the distribution of births across the days of the week in a random sample of 140 births from local records in a large city:**

| Day: | Sun. | Mon. | Tues. | Wed. | Thur. | Fri. | Sat. |
|---|---|---|---|---|---|---|---|
| Birth: | 13 | 23 | 24 | 20 | 27 | 18 | 15 |

**Do the data give significant evidence that local births are not equally likely on all days of the week?**

$H_0$: Births are evenly distributed across the days of a week in this large city.

$H_a$: " not evenly distributed ".

1. As stated, this is an SRS.
2. All expected counts are at least 5.
3. We can assume there are at least 1400 births a week in this city.

## When Were You Born?

| Day: | Sun. | Mon. | Tues. | Wed. | Thur. | Fri. | Sat. |
|---|---|---|---|---|---|---|---|
| Birth: | 13 | 23 | 24 | 20 | 27 | 18 | 15 |
| Expected: | 20 | 20 | 20 | 20 | 20 | 20 | 20 |

$$\chi^2 = \frac{(13-20)^2}{20} + \frac{(23-20)^2}{20} \cdots + \frac{(15-20)^2}{20} = 7.6$$

$df = 7-1 = 6$ p-value = .2689

Using my calculator, I ran a $\chi^2$ - Goodness-of-fit test. Since the p-value is .2689 and is greater than the significance level of $\alpha = .05$, we fail to reject $H_0$. There is not sufficient evidence to conclude that the births are not evenly distributed in this city.

## Inhereted Traits

**Biologists wish to cross pairs of tobacco plants having genetic makeup Gg, indicating that each plant has one dominant gene (G) and one recessive gene (g) for color. In other words, the biologists predict that 25% of the offspring will be green, 50% will be yellow-green and 25% will be albino.**

$\chi^2 = 6.47$

$df = 3-1 = 2$

p-value = .0392

| Offspring color | Observed | Expected |
|---|---|---|
| Green | 23 | 21 |
| Yellow-green | 50 | 42 |
| Albino | 11 | 21 |

**Do these data differ significantly from what the biologists have predicted? Carry out an appropriate test at the α = 0.05 level to help answer this question.**

## Inhereted Traits

| Offspring color | Observed | Expected |
|---|---|---|
| Green | 23 | 21 |
| Yellow-green | 50 | 42 |
| Albino | 11 | 21 |

Since the p-value is .0392 and is less than the significance level of $\alpha = .05$, we reject $H_0$. There is sufficient evidence that the biologist prediction of offspring color is not correct.