Mini Math
March 3

1. A company makes milk chocolate M&Ms claims the following distribution: 13% Brown, 14% Yellow, 20% Orange, 16% Green, 24% Blue, and 13% Red. A random sample of M&Ms are collected with the following counts: 55 Brown, 92 Yellow, 101 Orange, 73 Green, 74 Blue, 66 Red. Conduct a chi square test for goodness of fit to test the company's claim.

Solution: Full solution is not written (need to use state, plan, do, conclude framework): test-statistic: 25.69, p-value: less than 0.0005. Reject H_0.

2. Keri made a 6-sided die in her ceramics class and rolled it 90 times to test if each side was equally likely to show up. The table summarizes the outcomes of her 90 rolls.

Outcome of roll	1	2	3	4	5	6	Total
Frequency	12	28	12	13	10	15	90

(a) State the hypothesis that Keri should test.

> **Solution:** H_0 : The claimed distribution of $\frac{1}{6}$ probability for each roll is true (dice is fair).
> H_a : The claimed distribution is not true (dice is not fair).

(b) Calculate the expected count for each of the possible outcomes.

> **Solution:** $\frac{1}{6}(90 = 15$. Each expected count should be 15.

(c) Calculate the chi-square test statistic.

> **Solution:**
> $$\chi_5^2 = 14.4$$

(d) Use table C to find the p-value. What conclusion should you make?

> **Solution:** p-value is between 0.01 and 0.02. Because the p-value is less than 0.05, we reject the null hypothesis and have convincing evidence that the dice is not fair.

1. A social scientist selects a random sample of 25 freshmen, 25 sophomores, 25 juniors, and 25 seniors from various high schools across the state Kentucky. Each student was asked if they preferred in-person or remote learning. Here are the results:

	Freshman	Sophomore	Junior	Senior
Remote	3	12	14	15
In Person	22	13	11	10

(a) State the appropriate null and alternative hypotheses.

> **Solution:** H_o : There is no difference in the distribution of learning preference for all freshmen, sophmores, Juniors El Seniors in Kentucky.
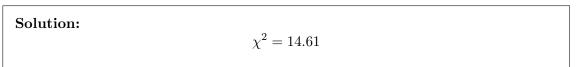>
> H_a :There is a difference in the distribution of learning preference all from all freshman sophmores, Juniors, and Seniors in Kentucky.

(b) Show the calculation for the expected count in the Remote/Senior cell. Then provide a complete table of expected counts.

> **Solution:**
>
> | | Freshman | Sophomore | Junior | Senior |
> |---|---|---|---|---|
> | Remote | 11 | 11 | 11 | 11 |
> | In Person | 14 | 14 | 14 | 14 |

(c) Calculate the value of the chi-square test statistic.

> **Solution:**
> $$\chi^2 = 14.61$$

(d) Interpret the P-value in plain language.

> **Solution:** p-value $= P(\chi^2_3 > 14.61) = 0.002182164$ Assuming the distribution for learning preferences are the same amongst freshman, sophmores, juniors, and seniors there is a roughly 0.2% chance of observing a χ^2 test statistic of 14..61 or greater in future samples. This is very condemning evidence against the null hypothesis and thus we reject H_0.

Mini Math

March 16

For each of the following situations decide what type of chi square test is appropriate. Explain.

1. A random sample of 200 students was asked to samples a new type of pizza that the school was considering using as a replacement for the current pizza. Each student stated if they were a freshman, sophomore, junior, or senior and asked if they liked the new pizza more than the current pizza (or not). The school would like the know if there is a relationship between grade level and pizza opinion.

   > **Solution:** 1 sample
   > 2 Variables (grade level, pizza opinion) Chi square test for independence

2. Another school is also considering changing their pizza vendor. This school selects a random sample of 50 freshman, 50 sophomores, 50 juniors, and 50 seniors. Each student tries the new pizza and states whether they like it more than the current pizza (or not). The school would like to know if the distribution of opinion differs across grade levels.

   > **Solution:** 4 samples
   > 1 variable
   > Chi square test for homogeneity .

3. A pizza shop claims that 30% of orders are placed on Fridays, 20% are placed on Saturdays, and 10% of orders are placed on other days of the week. A global pandemic may have change this distribution. The manger investigates so he knows how to staff the pizza shop appropriately. He selects a random sample of 300 orders and classifies each one according to the day of the week the order was placed. He wants to know if the distribution of orders is the same as it was before the global pandemic.

   > **Solution:** 1 sample
   > 1 variable (dar of order)
   > Chi square goodness of fit