AP Statistics—*X2* Tests Name\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_KEY\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Pd\_\_\_\_

A radio station trying to determine what kind of music to play takes a simple random sample of 50 students at each of three locations: a local middle school, a high school, and a college. The students are asked to choose which of three different music genres they most enjoy hearing on the radio. Here are the results:



1. How many samples (and populations) and how many categorical variables are involved in this situation? Identify them.

3 POPULATIONS—MIDDLE SCHOOL, HIGH SCHOOL, COLLEGE

1 VARIABLE—MUSIC GENRE

1. Perform the appropriate statistical test to determine if there is a difference in the music preference of these three age groups.

CHI SQUARE TEST OF HOMOGENEITY AT ALPHA = 5% (PICK AN ALPHA LEVEL)

Ho: THE DISTRIBUTION OF FAVORITE MUSIC GENRE IS THE SAME ACROSS THE THREE AGE LEVELS.

Ha: THE DISTRIBUTION OF FAVORITE MUSIC GENRE IS NOT THE SAME ACROSS THE THREE AGE LEVELS.

|  |  |  |  |
| --- | --- | --- | --- |
|  | HIP HOP | ALT | POST ROCK |
| MIDDLE | 22 | 20 | 8 |
| HIGH | 22 | 20 | 8 |
| COLLEGE | 22 | 20 | 8 |

THE SAMPLES ARE ALL RANDOMLY SELECTED

EXPECTED COUNTS ARE ALL AT LEAST 5🡪

ASSUME THE POPULATIONS ARE ALL >= 500

X2 = (28-22)2 / 22 + … + (14-8)2 / 8 = 10.67

DF = 4

P(X2 ≥10.67) = .03 SKETCH THE X2 DISTRIBUTION W/ RIGHT TAIL SHADED

SINCE THE P-VALUE OF 3% IS LOWER THAN THE ALPHA LEVEL OF 5%, I REJECT Ho. THERE IS STATISTICALLY SIGNIFICANT EVIDENCE THAT THE THREE AGE LEVELS OF STUDENTS DIFFER IN THEIR MUSIC GENRE PREFERENCE.

1. If you chose a chi-square test for homogeneity in #1, explain how the data could have been obtained to make a chi-square test for independence appropriate. If you chose a test for independence, explain how the data could have been obtained to make a test for homogeneity appropriate.

IF THE SAMPLE WAS TAKEN AS 150 TOTAL STUDENTS WHO WERE THEN CATEGORIZED BY AGE LEVEL AND MUSIC GENRE PREFERENCE, THIS WOULD HAVE BEEN A CHI SQUARE TEST OF INDEPENDENCE, AND I COULD HAVE QUESTIONED WHETHER AGE LEVEL IS INDEPENDENT OF MUSIC GENRE.

1. For each situation, identify the number of populations and number of categorical variables, then state the appropriate test that would be used.
2. At a state university, the student population is approximately one-third male and two-thirds female. Over a two-day period the gender of each student entering the student union is recorded, with the following results: males, 452; females, 1548. Determine whether males and females are as likely to enter the union as would be predicted on the basis of their percentage in the student population.

1 POPULATION—STUDENTS AT THE STATE UNIVERSITY WHO WOULD ENTER THE UNION; SAMPLE CONSISTS OF 2000 STUDENTS WHO ENTERED THE UNION

1 CATEGORICAL VARIABLE—SEX

CHI SQUARE GOODNESS OF FIT TEST

(Ho: P(MALE) = 1/3 AND P(FEMALE) = 2/3, Ha: THE PROPORTIONS ARE DIFFERENT)

1. A researcher is interested in whether or not a significant trend exists regarding the popularity of certain work shifts among police officers. A random sample of 60 police officers is selected from a large metropolitan police force. The officers are asked to indicate which of three work shifts they preferred. The results show that 40 officers prefer the first shift, 10 prefer the second shift, and 10 prefer the third shift. Do the results deviate significantly from what would be expected due to chance?

1 POPULATION—POLICE OFFICERS; SAMPLE CONSISTS OF 60 OFFICERS FROM A LARGE METRO FORCE

1 CATEGORICAL VARIABLE—SHIFT PREFERENCE

CHI SQUARE GOODNESS OF FIT TEST

(Ho: P(FIRST) = P(SECOND) = P(THIRD) = 1/3; Ha: AT LEAST TWO PROPORTIONS ARE DIFFERENT)

1. A professor wants to determine whether her department should keep the requirement of college algebra as a prerequisite for an Introductory Statistics course. Accordingly, she allows some students to register for the course on a pass-fail basis regardless of whether or not they have had the prerequisite. Of the 70 students in the class, 40 have had algebra and 30 have not. At the end of the semester, the professor compares the number of students passing or failing the class with whether or not they had algebra. The results are in the table below. Are students more likely to pass the course if they have taken college algebra?

|  |  |  |
| --- | --- | --- |
|  | Pass | Fail |
| Algebra | 34 | 6 |
| No Algebra | 12 | 18 |

1 POPULATION—STUDENTS WHO WOULD TAKE INTRO STAT; THE SAMPLE IS 70 STUDENTS IN THE CLASS

2 VARIABLES—STATUS AT THE END OF THE COURSE (PASS/FAIL) AND PREREQUSITE (ALGEBRA OR NOT)

CHI SQUARE ASSOCIATION TEST

1. In a study of high-achieving high school graduates, the researchers surveyed 828 high school graduates who were considered “academic superstars” and 433 graduates who were considered “solid performers.” One question on the survey asked the distance from their home to the college they attended. Assuming it is reasonable to regard these two samples as random sample of academic superstars and solid performers nationwide, use the accompanying data to determine if it is reasonable to conclude that the distribution of responses over the distance from home categories is not the same for academic superstars and solid performers.

|  |  |
| --- | --- |
|  | Distance of College from Home (in miles) |
| Student Group | Less than 40 | 40 to 99 | 100 to 199 | 200 to 399 | 400 or more |
| Academic Superstars | 157 | 157 | 141 | 149 | 224 |
| Solid Performers | 104 | 95 | 82 | 65 | 87 |

2 POPULATIONS—ACADEMIC SUPERSTARS AND SOLID PERFORMERS NATIONWIDE; THERE IS A SAMPLE FROM EACH POPULATION

1 VARIABLE—DISTANCE OF COLLEGE FROM HOME

CHI SQUARE HOMOGENEITY TEST