CHI SQUARE TEST OF ASSOCIATION/INDEPENDENCE

*This test is used to determine whether there is a significant association between 2 categorical variables from the **same sample**.*

To determine of there was a relationship between smoking status and socioeconomic levels, researchers categorized 356 male employees:

ACTUAL COUNTS	Socioeconomic Level		
	High	Middle	Low
Current Smoker	51	22	43
Former Smoker	92	21	28
Never Smoked	68	9	22

DETERMINE EXPECTED COUNTS:

Expected Count = (Row Total)(Column Total)/ Sample Size

EXPECTED COUNTS	Socioeconomic Level		
	High	Middle	Low
Current Smoker	68.75	16.94	30.30
Former Smoker	83.57	20.60	36.83
Never Smoked	58.68	14.46	25.86

H STATE NULL AND ALTERNATIVE HYPOTHESES

H_o : There is no association between smoking status and SES in the population of federal male employees

H_a : There is an association between smoking status and SES in the population of federal male employees

A DETERMINE THAT CONDITIONS FOR TEST ARE ACCEPTABLE:

- SRS.. unknown but the sample was random
- Counts (not percents)… yes
- Every expected count ≥ 1 and $80\% \geq 5$… yes

T PERFORM TEST USING…

FORMULA/TABLE E:

a) Chi-Square Statistic: $X^2 = \Sigma (O_i - E_i)^2 / E_i$

$= (51 - 68.75)^2/68.75 + (22 - 16.94)^2/16.94 + \ldots + (22 - 25.86)^2 /25.86$

$= 18.51$

b) Degrees of Freedom = $(r - 1)(c - 1) = (3 - 1)(3 - 1) = 4$

Number of rows in table — Number of columns in table

c) P-Value

i) Table E:

Any X^2 statistic > 18.51 (df = 4) has P-value < .001

ii) Calculator:

DISTR → 7:X^2 cdf (18.51, 100, 4) → p = .00098

CALCULATOR:

a) Store observed counts in a [R,C] matrix:

MATRIX → EDIT → 1: [A] → 3 X 3 → Enter Counts → QUIT

b) Perform X^2 Test:

STAT → TESTS → C: X^2 – Test → X^2 = 18.51, P-value = .0098

NOTE:

If MATRIX [A] = Observed Counts, MATRIX [B] = Expected Counts

S STATE CONCLUSION IN CONTEXT:

There is very strong evidence ($p < .001$) to reject H_o and conclude an association exists between smoking status and SES in the population of federal male employees… to determine direction and nature of associations, use 2-way table techniques.