AP Statistics – 11.2b	Name:
Goal: Chi-Square (χ^2) Test for Independence and Review the 3 Chi-Square Tests	Date:

P (A) = P(A | B)

$P(CHD) = \frac{190}{8474} = .02$

$P(CHD | \text{LOW ANGER}) = \frac{53}{3110} = .017$

Remember the probability definition -
**IN PROBABILITY… these must be EQUAL.
**FOR CHI-SQUARE… we are testing whether they are significantly different.

I. Setting up Hypothesis for Test of Independence/ Association:

Do Angry People Have More Heart Disease?

A study followed a random sample of 8474 people with normal blood pressure for about four years.[11] All the individuals were free of heart disease at the beginning of the study. Each person took the Spielberger Trait Anger Scale test, which measures how prone a person is to sudden anger. Researchers also recorded whether each individual developed coronary heart disease (CHD). This includes people who had heart attacks and those who needed medical treatment for heart disease. Here is a two-way table that summarizes the data:

Remember: EC=(RT*CT)/TT

	Low anger	Moderate anger	High anger	Total
CHD	53 (69.7)	110	27	190
No CHD	3057	4621	606	8284
Total	3110	4731	633	8474

MATRIX A

Clearly label "EXPECTED COUNTS" in ()'s if you want to put in the table.

EXPECTED COUNTS

	LOW ANGER	MED ANGER	HIGH ANGER
CHD	69.7	106.1	14.2
NO CHD	3040.3	4624.9	618.8
	3,110	4,731	633

MATRIX B

1. Why is this a $\chi 2$ Test for Independence?

① THIS IS 1 RANDOM SAMPLE FROM A POPULATION OF INTEREST (PEOPLE WITH NORMAL BLOOD PRESSURE)

② EACH OBSERVATION IS CLASSIFIED BY 2 CATEGORICAL VARIABLES (ANGER LEVEL - LOW, MED, HIGH AND HEART DISEASE - YES, NO)

2. There are 2 ways to write the hypothesis for this test. State BOTH:

OPTION 1

H_0: ANGER AND HEART DISEASE ARE INDEPENDENT IN POPULATION WITH NORMAL BLOOD PRESSURE.

H_A: ANGER + HEART DISEASE ARE NOT INDEPENDENT

OPTION 2:

H_0: THERE IS NO ASSOCIATION BETWEEN ANGER LEVEL AND HEART DISEASE IN POPULATION WITH NORMAL BLOOD PRESSURE.

H_A: THERE IS AN ASSOCIATION BETWEEN ANGER + HEART DISEASE

II. Does the data provide coinvincing evidence of an association between anger level and heart disease in the population of interest? Conduct an appropriate chi-square test to find out.

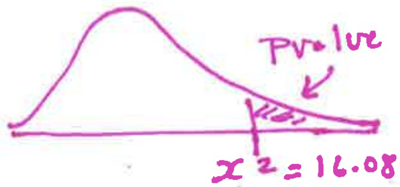
3. Conditions

- RANDOM SAMPLE OF 8,474 PEOPLE WITH NORMAL BLOOD PRESSURE
- INDEPENDENT - SAMPLING WITHOUT REPLACEMENT. CHECK 10% CONDITION. THERE ARE MORE THAN 10(8474) = 84,740 people with normal B.P.
- LARGE SAMPLE - ALL EXPECTED COUNTS ARE GREATER THAN 5. SEE THE TABLE ABOVE

On FRQ's write the Chi-Square Formula $\chi^2 = \Sigma (OBS - EXP)/EXP$

4. Mechanics

- Name the test: χ^2 TEST OF INDEPENDENCE (OR ASSOCIATION)
- Significance level: $\alpha = .05$
- Sketch the graph (χ^2 graph is optional)



- Degrees of freedom $df = (3-1)(2-1) = 2$
- Test Statistic

$$\chi^2 = \frac{(53-69.7)^2}{69.7} + \dots = 16.08$$

- P-value

Pvalue $= P(\chi^2 > 16.08) = .00032$

STAT TESTS χ^2-TEST

5. Conclusion in context

and less than alpha equal .05

SINCE THE PVALUE IS VERY SMALL and less than alpha equal .05, WE REJECT H_0. AND CONCLUDE THAT ANGER LEVEL AND HEART DISEASE ARE NOT INDEPENDENT. THERE IS SUFFICIENT EVIDENCE TO CONCLUDE THAT ANGER LEVEL AND HEART DISEASE ARE ASSOCIATED IN THE POPULATION OF PEOPLE WITH NORMAL BLOOD PRESSURE

III. χ^2 Test for Independence from start to finish – CYU page 718

H_0: THERE IS NO ASSOCIATION BETWEEN EXCLUSIVE TERRITORY AND FRANCHISES SUCCESS IN NEW FRANCHISE FIRMS (VARIABLES ARE INDEPENDENT)

H_A: THERE IS AN ASSOCIATION BETWEEN THE 2 VARIABLES

CONDITIONS

Random – random sample of new franchises

Independent – there are more than 10(170) new franchises in the U.S.

Large Sample – the expected counts are all above 5. Expected Counts: 102.7, 20.3, 39.3, and 7.74.

III. χ^2 Test for Independence from start to finish – CYU page 718 (cont.)

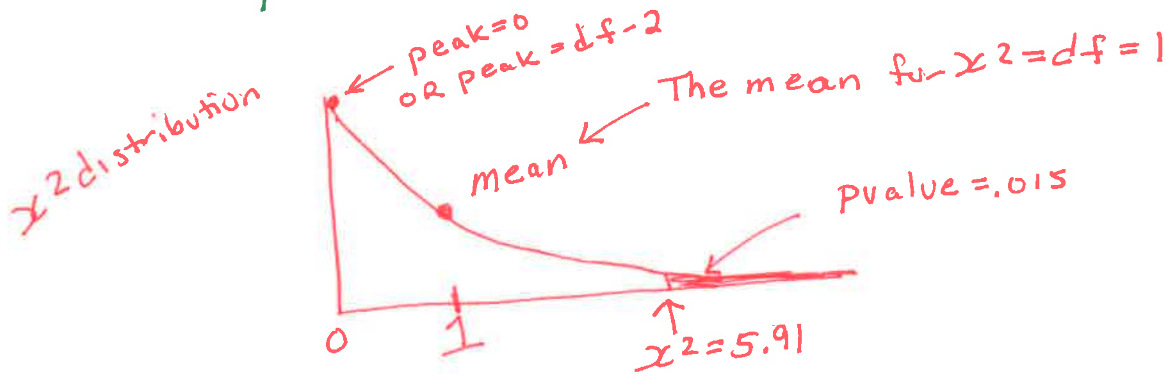
TEST: χ^2 TEST FOR INDEPENDENCE

$\alpha = .01$

$df = (2-1)(2-1) = 1$

$$\chi^2 = \frac{(108 - 102.7)^2}{102.7} + \ldots = 5.91$$

$$\text{pvalue} = P(\chi^2 > 5.91) = .015$$



CONCLUDE: Since the pvalue (.015) is greater than $\alpha = .01$, FAIL TO Reject Ho. We do not have enough evidence to conclude that there is an association between whether a franchise will be successful and whether the franchise has an exclusive territory. These 2 variables appear to be indendent.

$P(A) = P(A|B)$

$P(\text{success}) = \frac{123}{170} = .72$

$P(\text{success}|\text{exclusive})$

$108/123 = .76$

Referring back to the probability definition -
**IN PROBABILITY… these are NOT EQUAL; and conclude these variables are NOT independent.
**FOR CHI-SQUARE… the test determined that based on alpha=.01 we have enough evidence to state these variables are independent.

IV. Choose the Correct $\chi 2$ Test   Example: Online social networking

An article in the *Arizona Daily Star* (April 9, 2009) included the following table:

	18-24	25-34	35-44	45-54	55-64	65+	Total
Use Online Social Networks	137	126	61	38	15	9	386
Do Not Use Online Social Networks	46	95	143	160	130	124	698
Total	183	221	204	198	145	133	1084

Suppose that you decide to analyze this data using a chi-square test. However, without any additional information about how the data was collected, it isn't possible to know which chi-square test is appropriate.

Problem:

(a) Explain how you know that a goodness-of-fit test is not appropriate for analyzing these data.

- Since there are either two variables or two or more populations, a goodness-of-fit test is not appropriate.
- Goodness-of-fit tests are only appropriate when analyzing the distribution of **one variable** in **one population**.

(b) Describe how these data could have been collected so that a test for homogeneity is appropriate.

- To make a **test for homogeneity** appropriate, we would **need to take 6 independent random samples**, one from each age category, and then ask each person whether or not they use online social networks.
- Or to make a **test for homogeneity**, we could take **2 independent random samples**, one of online social network users and one of people that do not use online social networks, and ask each member of each sample how old they are.

(c) Describe how these data could have been collected so that a test for association/ independence is appropriate.

- To make a **test for association/independence appropriate:**
  - we would **take one random sample** from the population **and**
  - **ask each member** about **their age** **and** **whether or not they use online social networks**.

- This seems like the most reasonable method to collect the data, so a test of association/independence is probably the best choice. But, we can't know for sure unless we know how the data were collected.

V. Choose the Correct Inference Test

Example: Ibuprofen or acetaminophen?

EXAMPLE not included in the class handout. ***Review concepts on your own.

In a study reported by the *Annals of Emergency Medicine* (March 2009), researchers conducted a randomized, double-blind clinical trial to compare the effects of ibuprofen and acetaminophen plus codeine as a pain reliever for children recovering from arm fractures. There were many response variables recorded, including the presence of any adverse effect, such as nausea, dizziness, and drowsiness. Here are the results:

	Ibuprofen	Acetaminophen plus Codeine	Total
Adverse effects	36	57	93
No adverse effects	86	55	141
Total	122	112	234

Problem:

a) Explain why it was important to investigate this question with a randomized, double-blind clinical trial.

- IMPORTANCE OF RANDOMIZED EXPERIMENT
  - It is important that the treatments in an experiment be randomly assigned so that the two treatment groups are roughly equivalent at the beginning of the study.
  - Randomization reduced the effects of lurking (confounding) variables because these extraneous varibles should be balanced out among the two groups.
- IMPORTANCE OF DOUBLE-BLINDING
  - It is also important that both the patients and those administering the drugs and measuring the response do not know who is receiving which treatment.
  - This will keep the expectations the same for both groups of patients and not favor one treatment over the other.

Is the difference between the two groups statistically significant?

b) Conduct a Chi-square Test Homogeneity.
c) Conduct a 2-Sample Z-test for the difference of proportions.
d) Why do these 2 test provide the same results?

b) State: χ^2 Test for Homogeneity using $\alpha = 0.05$:

Ho: There is no difference in the proportions of patients like these who suffer adverse effects when taking ibuprofen or acetaminophen plus codeine.
Ha: There is a difference in the proportions…

Conditions:

- ***Random***: The treatments were assigned at random.
- ***Independent***: *Knowing if one subject had an adverse effect shouldn't give any additional information about the responses of other subjects, so the observations can be considered independent.*
- ***Large Sample Size*** The expected counts (listed below) are all at least 5.

Expected counts	Ibuprofen	Acetaminophen plus Codeine	
Adverse effects	48.5	44.5	93
No adverse effects	73.5	67.5	141
Total	122	112	234

Calculation:

- ***Test Statistic*** $\chi^2 = \dfrac{(36-48.5)^2}{48.5} + \cdots = 11.15$
- ***P-value*** df= (2 – 1)(2 – 1) = 1

  P-value = P($\chi^2 > 11.15$)= 0.0008 χ^2 cdf(11.15,e99,1)

Conclude: Because the P-value is less than α = 0.05, we reject H_0. We have convincing evidence that there is a difference in the proportions of patients like these who suffer adverse effects when taking ibuprofen or acetaminophen plus codeine.

c) <u>State</u>: 2-Sample Z-test for the difference of proportions using α= 0.05:

p_I = the true proportion of adverse effects for Ibup users

p_A = the true proportion of adverse effects for Acet. Users

$H_0: p_I - p_A = 0$

$H_a: p_I - p_A \neq 0$

Conditions:

- ***Random***: same
- ***Independent***: *same*
- ***Normal:*** Successes and failures are all greater than 10 - 36, 86, 57, 55

Calculation:

- ***Pooled Proportion***

$\hat{p}_I = \frac{36}{122} = .295$

$\hat{p}_A = \frac{57}{112} = .509$

$\hat{p}_C = \frac{36+57}{122+112} = .397$

- ***Test Statistic***

$$Z = \frac{.295 - .509}{\sqrt{(.397)(.603)}\sqrt{\frac{1}{122}+\frac{1}{112}}} = \frac{-.214}{.064} = -3.34$$

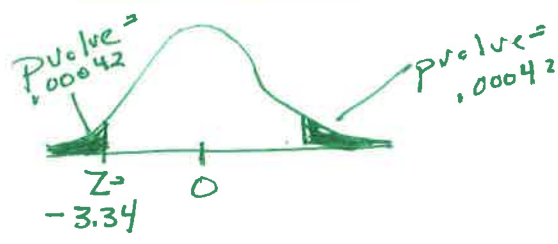
↑pooled $\hat{p}_C$

- ***P-value*** **P-value =** $P(Z < -3.34)$

$= 2(.0042)$

P-value $= .00084$

Normalcdf(−E99, −3.34, 0, 1) = .00042



Conclude: *same*

- Since we are comparing the proportion of subjects with adverse effects for just two treatments, we can also use a two-sample z test for the following hypotheses:
- Using technology, z = –3.339 and P-value = 0.0008.
- The P-value is exactly the same as the P-value from the chi-square test and $z^2 = (-3.339)^2 = 11.15 = \chi^2$.

CALC
Stat
TESTS
6: 2PropZTest
$x_1 = 36$
$n_1 = 122$
$x_2 = 57$
$n_2 = 112$
$p_1: \neq p_2$
⇓
$z = -3.34$
$p = 8.4E-4$
$\hat{p}_1 = .295$
$\hat{p}_2 = .509$
$\hat{p} = .397$
This is pooled $\hat{p}$.