

PG 698

PENN STATE

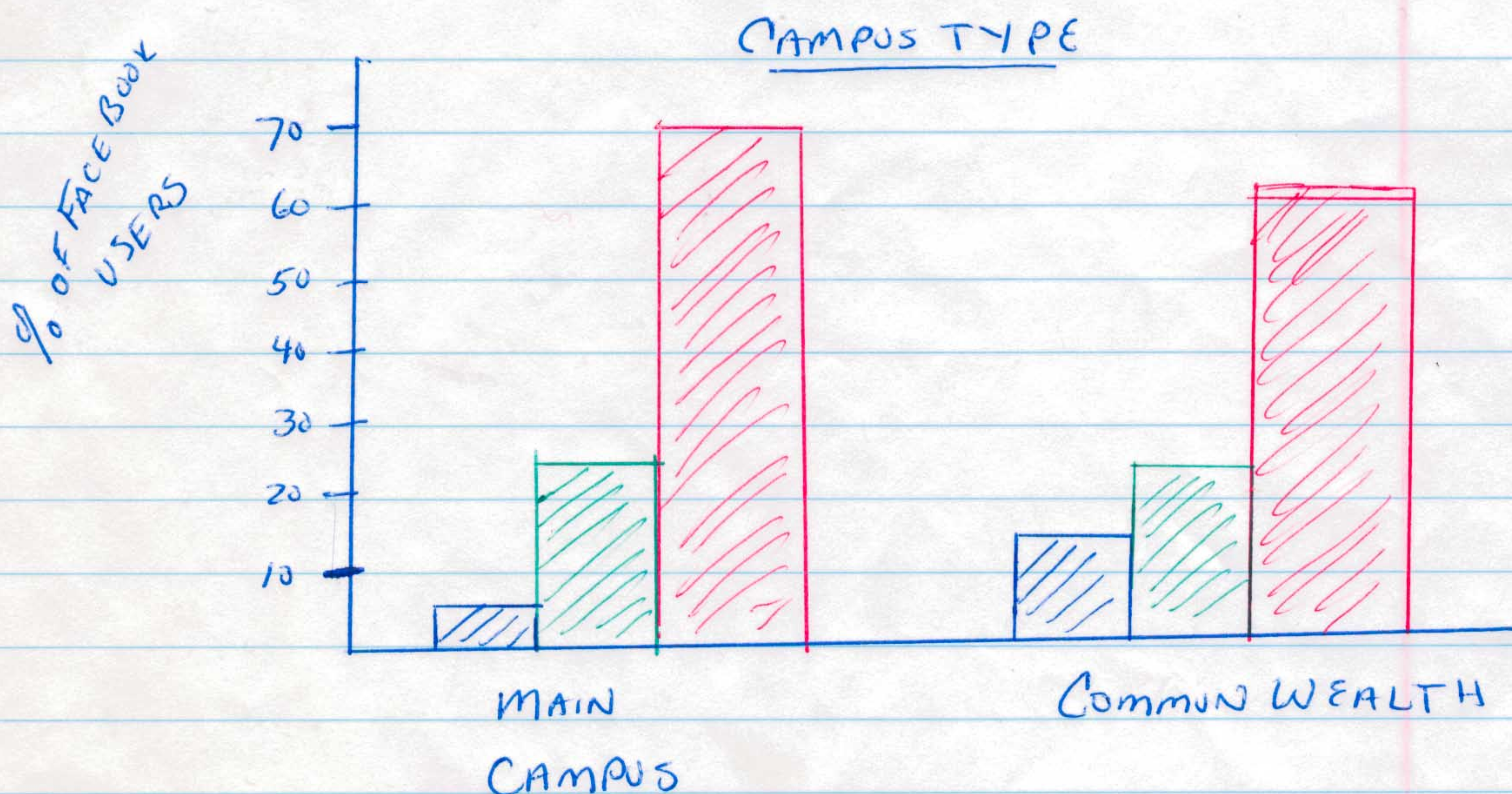
①

<u>USE FACEBOOK</u>	<u>MAIN CAMPUS</u>		<u>Common WEALTH</u>	
SEVERAL TIMES A MONTH	55	6.0%	76	12.1%
AT LEAST ONCE A WEEK	215	23.6%	157	25.0%
AT LEAST ONCE A DAY	640	70.3%	394	62.8%
TOTAL	910	99.9%	627	99.9%

②

YOU NEED TO COMPARE PROPORTIONS SINCE THERE IS SUCH A LARGE DIFFERENCE IN SAMPLE SIZES (910 VS. 627) FROM THE 2 CAMPUSES.

③

KEY:

- AT LEAST ONCE A DAY
- AT LEAST ONCE A WEEK
- SEVERAL TIMES A MONTH

PG 702C

PENN STATE FACE BOOK USERS

- ① H_0 : THERE IS NO DIFFERENCE IN THE DISTRIBUTION OF FACEBOOK USE BETWEEN STUDENTS AT PENN STATE'S MAIN CAMPUS AND ITS COMMONWEALTH CAMPUSES

H_a : THERE IS A DIFFERENCE IN THE DISTRIBUTIONS OF FACEBOOK USE BETWEEN THE 2 CAMPUSES

- ② EXPECTED COUNTS = $\frac{\text{ROW TOTAL} \times \text{COLUMN TOTAL}}{\text{TABLE TOTAL}(n)}$

FACEBOOK USE	MAIN CAMPUS	COMMONWEALTH	TOTAL
SEVERAL TIMES A MONTH	$\frac{131 \times 910}{1537} = 77.56$	$\frac{131 \times 627}{1537} = 53.44$	131
ONCE A WEEK	220.25	151.75	372
ONCE A DAY	612.19	421.81	1034
TOTAL	910	627	1537

- * USE CALC
- ① ENTER MATRIX [A] WITH ~~EXPECTED~~ OBSERVED
 - ② STAT TESTS χ^2 -TEST
 - ③ MATRIX [B] NOW HAS EXPECTED

③ $\chi^2 = \sum \frac{(\text{OBSERVED} - \text{EXPECTED})^2}{\text{EXPECTED}}$

$$\chi^2 = \frac{(55 - 77.56)^2}{77.56} + \dots + \frac{(394 - 421.81)^2}{421.81} = 19.49$$

$$DF = (\text{ROW} - 1)(\text{COLUMN} - 1) = 2 \cdot 1 = 2$$

$$p \approx 0$$

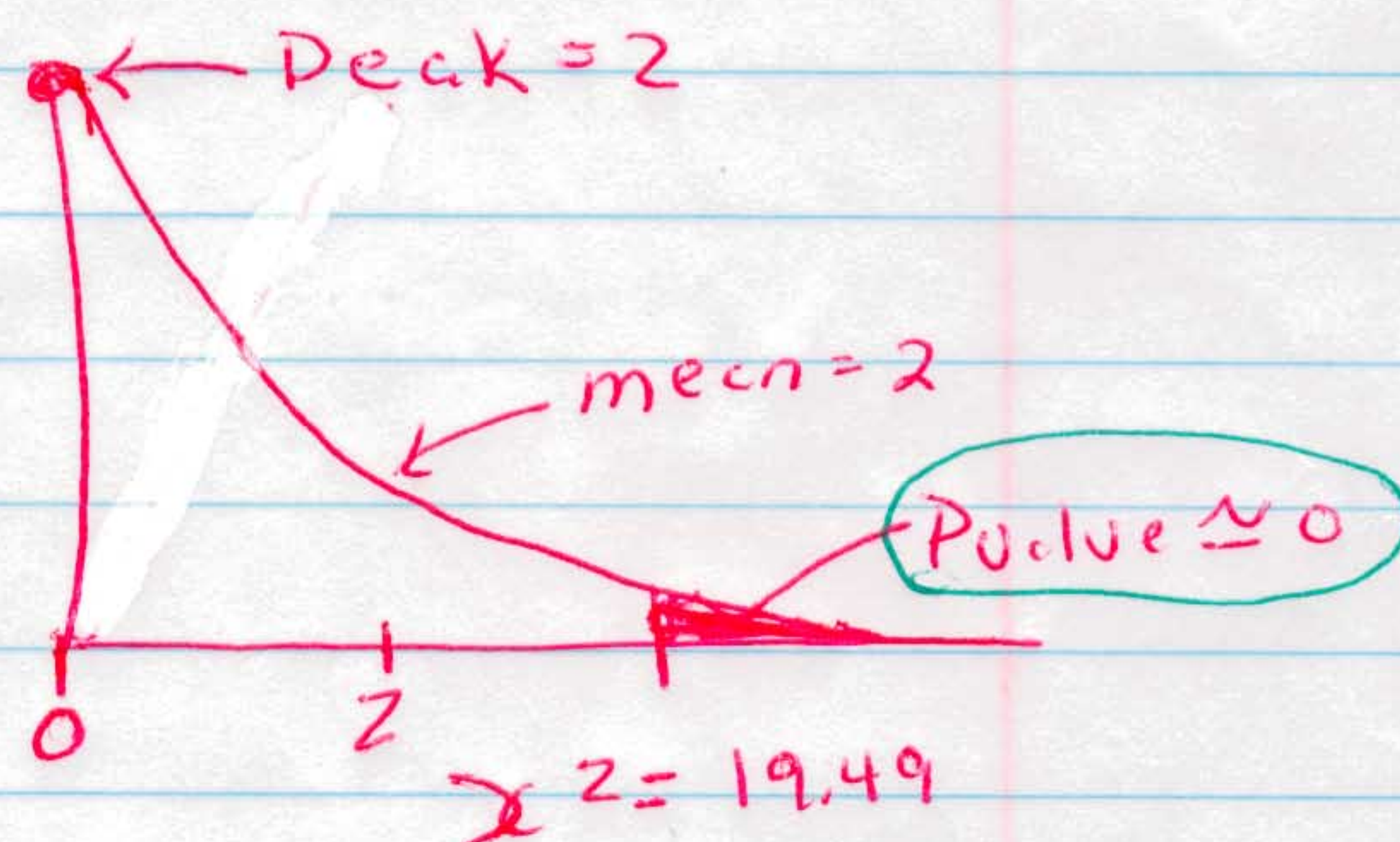
Pg 705

PENN STATE FACE BOOK USERS (continued from pg 702c)

① $\chi^2 = 19.49$ $df = 2$

$$P(\chi^2 > 19.49) \approx 0$$

$$\chi^2_{cdf}(19.49, 199, 2) = .000059$$



② ASSUMING THAT THERE IS NO DIFFERENCE IN THE DISTRIBUTION OF FACEBOOK USE BETWEEN STUDENTS ON PENN STATE'S MAIN CAMPUS AND STUDENTS AT PENN STATE'S COMMONWEALTH CAMPUS, THE PROBABILITY OF OBSERVING A DIFFERENCE IN THE DISTRIBUTIONS OF FACEBOOK USE FOR THE 2 SAMPLES AS LARGE OR LARGER THAN THE ONE FOUND IN THIS STUDY IS ABOUT 6 IN 100,000.

③ SINCE THE PVALUE IS SO SMALL, REJECT H_0 . It appears that the distribution of facebook use is different among students at PENN States main Campus and students at PENN State's common wealth campuses

PG 708

HEART ATTACK PATIENTS

①

QUALITY OF

LIFE

Canada

U.S.

Canada

U.S.

Much Better

75 24.1%

541 25.0%

77.37

538.63

Somewhat better

71 22.8%

498 23.0%

71.47

497.53

About the same

96 30.9%

779 36.0%

109.91

765.09

Somewhat worse

50 16.1%

282 13.0%

41.70

290.30

Much Worse

19 6.1%

65 3.0%

10.55

73.45

Total

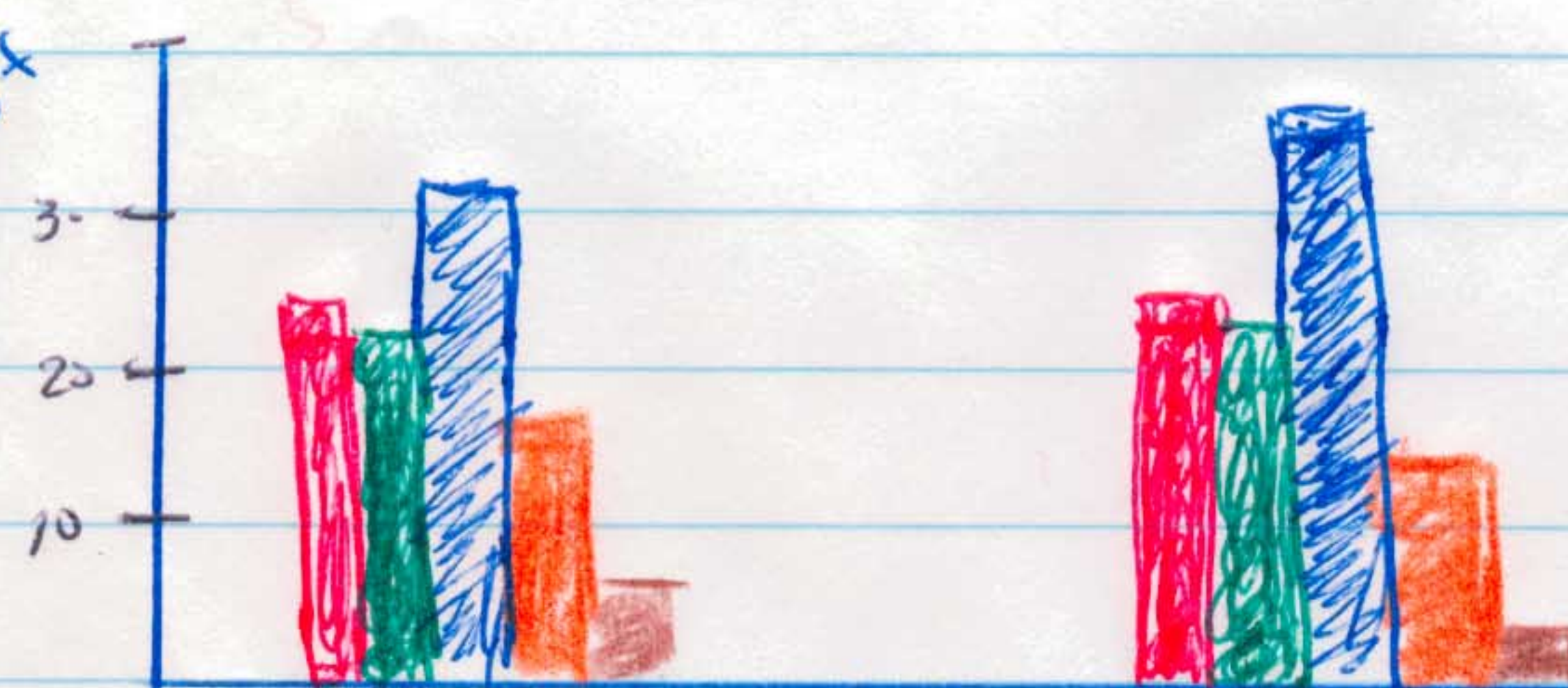
311 100%

2,165 100%

311.00

2165

Quality of Life



■ much better
 ■ somewhat better
 ■ About the same
 ■ somewhat worse
 ■ much worse

CANADA

US

② H_0 : There is no difference in the distribution of Quality of life in Canada and the U.S.

H_A : There is a difference in the distribution of Quality of life in Canada and the U.S.

TEST: χ^2 TEST OF HOMOGENEITY $\alpha = .01$ $df = 4 - 1 = 4$

Conditions

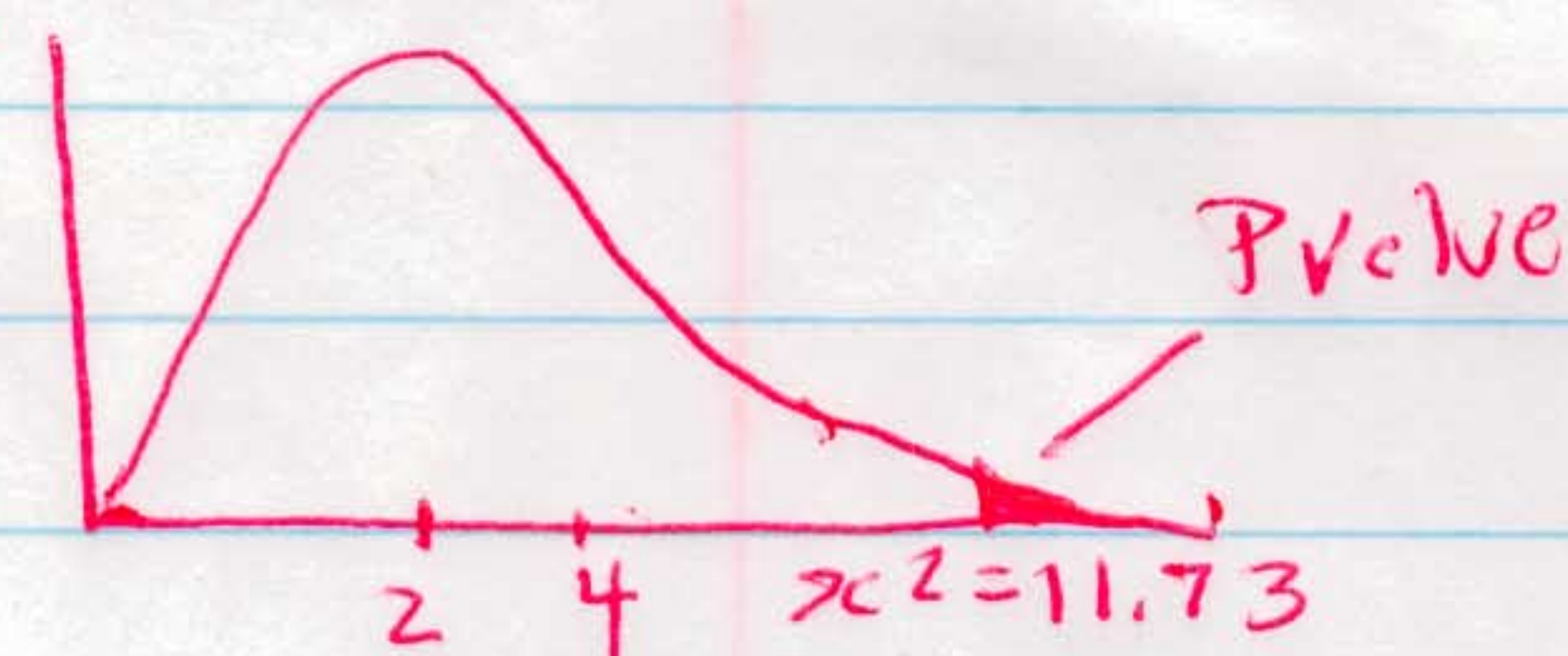
Random - Separate random samples

Independent - We clearly have less than 10% heart attack patients

Large Sample Size - All expected counts are at least 5 (see table above)

$$\chi^2 = \frac{(O-E)^2}{E} = 11.73$$

$$P(\chi^2 > 11.73) = .0195$$



Since the $p\text{-value}(.0195)$ is greater than $\alpha = .01$, We fail to reject H_0 . There is not enough evidence to conclude that there is a difference of quality of life for heart attack patients in the U.S. and Canada.

Put observed counts in Matrix [A] and Test (χ^2 Test) Matrix B has expected counts

12.2 NOTES

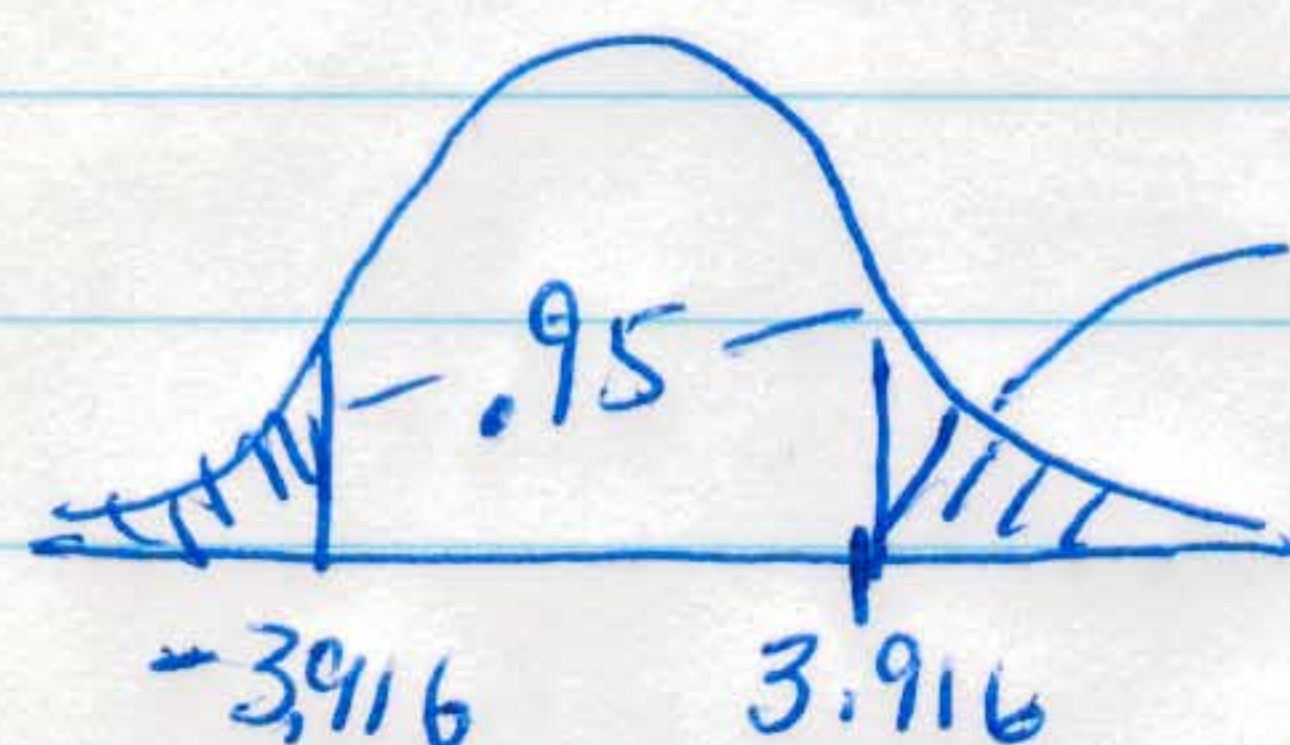
5

PG 712

DATA EXPLORATION Second hand stores

① $\chi^2 = 15.334$
 $z = 3.916 \longrightarrow z^2 = (3.916)^2 = 15.335$ (this matches χ^2)

② $\chi^2 = 15.334$ $df = 1$ $pvalue = \chi^2cdf(15.334, E99, 1) = .00009$
 $z = 3.916$



$normalcdf(3.916, E99, 0, 1) = .000045$
 $\times 2$

$pvalue = .00009$

- ③ Since the pvalue for the χ^2 Test is so small, we have convincing evidence that there is a difference in the proportion of second hand store shoppers that are women in the 2 cities. In fact we are 95% confident that the interval from .077 to .231 captures the true difference in the proportions of second hand store shoppers that are women in the 2 cities. This interval suggests that the proportion of second hand store shoppers who are women is between .08 and .23 higher in City 1 than in City 2.

PG 713

① This was an experiment. EACH INDIVIDUAL WAS exposed to a treatment that involved how he or she was contacted

②

CONTACT METHOD

Phone
Personal Interview
Written Response
Total

HAVE YOU USED COCAINE?

Answers "Yes"	Answers "No"	Total
168 197.33	632 602.67	800
200	600	800
224	576	800
592	1,808	2,400

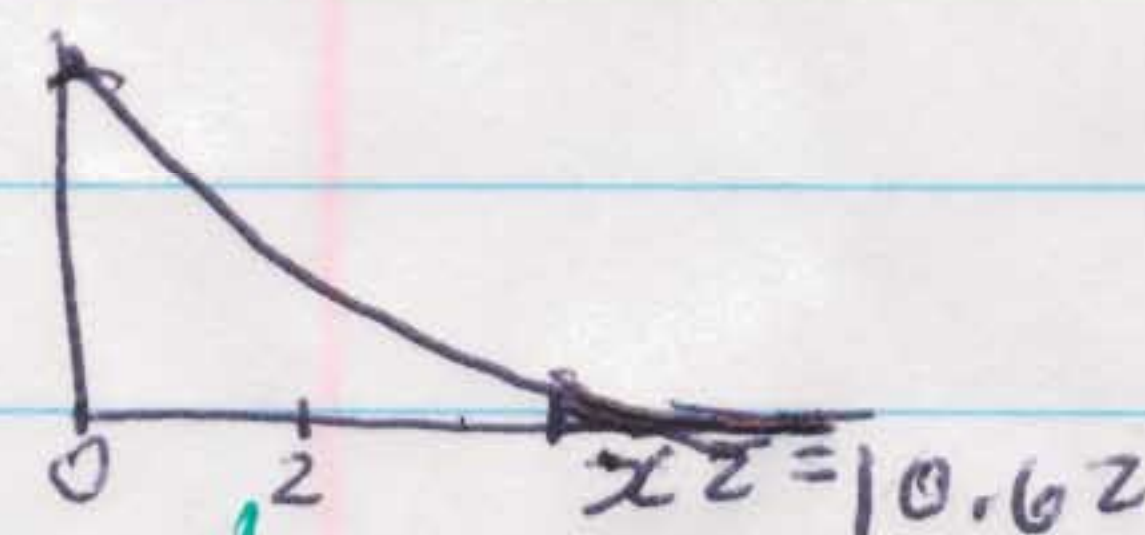
* EXPECTED COUNTS IN BLUE

③

H_0 : There is no difference in the actual proportion who answered "YES" based on contact method

$$\text{or } H_0: p_1 = p_2 = p_3$$

H_A : There is a difference in the actual proportion who answered "YES" based on contact method

Conditions

Random: the data comes from a randomized experiment

Independent: due to random assignment, groups can be viewed independent. Also individual responses are independent

large sample size: the expected counts are at least 5 (see table)

TEST: χ^2 TEST FOR HOMOGENEITY

$$\alpha = .05$$

$$\chi^2 = \frac{(O - E)^2}{E} = 10.62 \quad df = 2$$

$$p\text{-value} = P(\chi^2 > 10.62) = .0049$$

Since the p-value $< .05$, Reject H_0 . We have convincing evidence of a difference in the proportions of people who answered "Yes" based on how they were contacted.

PG 718

Success	Exclusive Territory		Total
	Yes	No	
Yes	108 (102.74)	15 (20.26)	123
No	34 (39.26)	13 (7.74)	47
	142	28	170

H_0 : There is no association between territory type and success or not in the population of franchises

H_A : There is an association between territory type and success or not in the population of franchises

CONDITIONS

Random sample of franchises

Independent - there are more than 1,700 franchises in U.S.

Large sample size - All expected counts are at least 5 (see table)

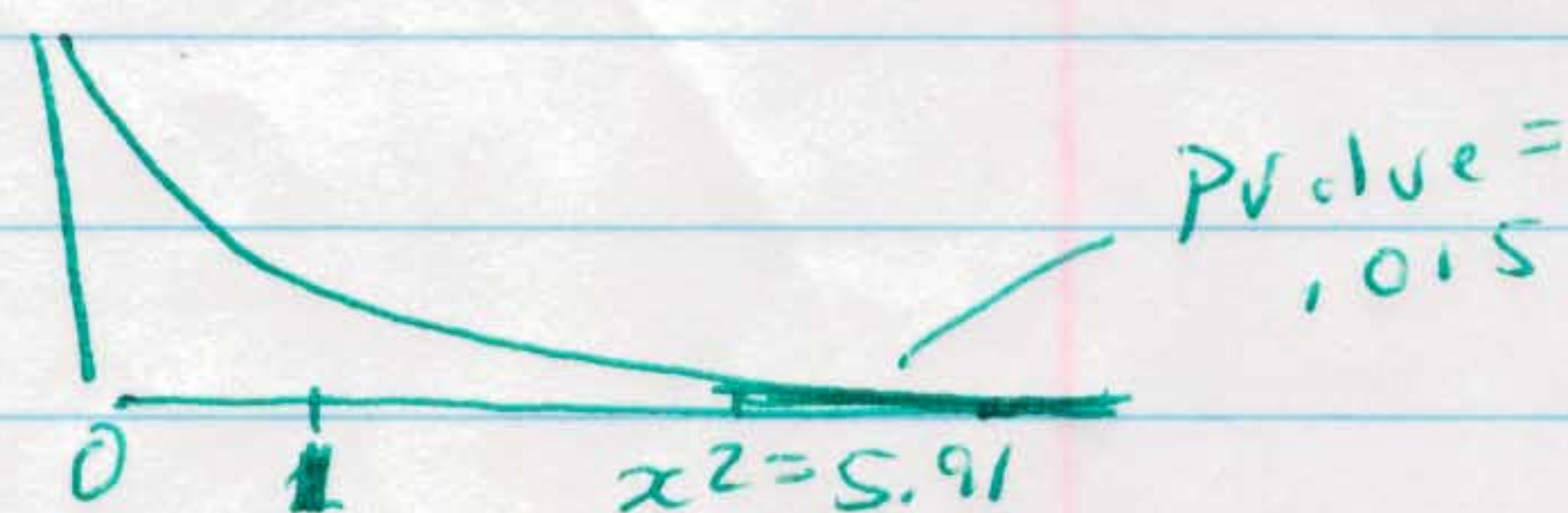
TEST - χ^2 Test of Association/Independence

$$\alpha = .01$$

$$df = (Rows - 1)(Columns - 1) = (1)(1) = 1$$

$$\chi^2 = \frac{(O - E)^2}{E} = 5.91$$

$$p\text{-value} = P(\chi^2 \geq 5.91) = .0150$$



CONCLUDE Since the p-value (.0150) is greater than $\alpha = .01$, We fail to reject H_0 . We do not have enough evidence to conclude that there is an association between whether franchises have an exclusive territory or not and whether they are successful or not.