CHI- SQUARE TEST ASSIGNMENT

1. SAMPLE SIZE OF 1600 CARDS HAVING 4 diffnt shapes.

so assuming they are equally present in the deck. = 1600/4 = 400 cards.

	Expected	Observed	$X^2 (O-E)^2/E$
Spades	400	404	16/400 = 1/25 = 0.04
Heart	400	420	400/400 = 1
Diamonds	400	400	0
Clubs	400	376	576/400 = 1.44
			2.48

[Degrees of freedom are how many bins we create or groups we create out of the data]

df = 4. for population

df = 4-1 for a sample size

= 3

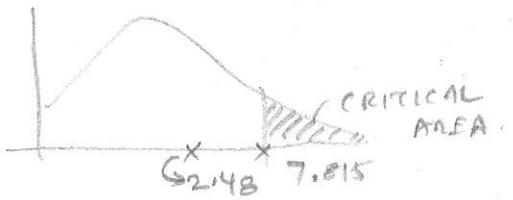
Lets assume CI for 95% CL

so lets get the critical value from the X^2 table.

df: 3 and critical area 5% (0.05)

Probability of exceeding the critical value is

7.815.



NULL HYPOTHESIS: SUITS ARE EQUALLY LIKELY

ALTERNATIE HYPO: HAVE DISCREPANCIES.

SO NULL HYPOTHESIS IS PROVED AS 2.48 is well within 7.815

2. SAMPLE SIZE OF 1662 cards with 5 diffnt types of cards. for a sample df = 5-1 = 4.

[Generally there are 2 jokers for every 52 CARDS]

	Expected	Observed	$X^2 (O-E)^2/E$
SPADES	400	404	0.04
HEARTS	400	420	1.0
DIAMONDS	400	400	0.0
CLUBS	400	356	4.84
JOKERS	62	82	6.45
			12.33

So for 1600 cards we can find

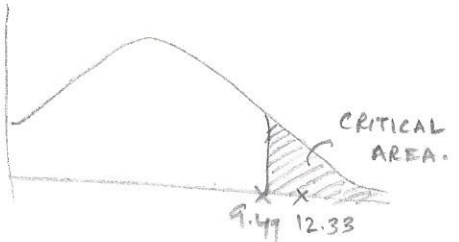
$\frac{1600 \times 2}{52}$ = 62 (approx) jokers

Degrees of freedom = 5-1 = 4.

Assume critical region: 5% (0.05) Area = 9.49

NULL HYPOTHESIS IS REJECTED

ALTERNATE HYPOTHESIS IS TRUE: HAVE DISCREPANCIES-



CHI-SQUARE TEST

5. Alpha level $\alpha = 0.05$, confidence level: 95%.

Store	# shoppers	Excepted	x^2		Sum = 14.6
A	262	220	220 − 262	$-42^2/220$	8.01
B	234	220	220 − 234	$-14^2/220 = 0.89$	
C	204	220	220 − 204	$16^2/220$	1.16
D	190	220	220 − 190	$30^2/220 = 4.09$	
E	210	220	220 − 210	$10^2/220 = 0.45$	

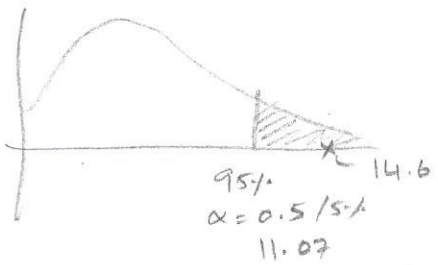
[Observed: Users do not have a preference and so the expectation is they are equally distributed

from the table we get df = 5, chi(0.05) = 11.07

So it is under the tail so reject the NULL HYPOTHESIS.

and ACCEPT ALTERNATE HYPOTHESIS

Therefore there is no preference for shoppers.



3. Tegin and cheetah cross. Results would be categorized as

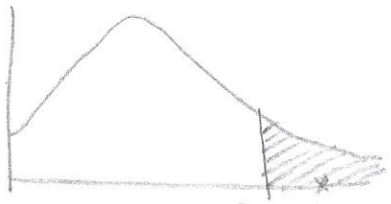
	Observed.	expected	x^2	
4 stripes only	50	60	$(50-60)^2/60$	1.66
3 spots only	41	58	$(41-58)^2/58$	4.98
9 both stripes and spots	85	58	$(85-58)^2/58$	12.56
				19.2

Assumption: Let us consider the critical region of 5%

degrees of freedom n = 3 (whole population) and not sample.

NULL HYPOTHESIS: DID SHE GET PREDICTED OUTCOME

= FALSE → It falls in the critical region



4.

COLOR \ POD SHAPE	INFLATED	CONSTRICTED	
YELLOW	556 (557.6)	184 (182.4)	740
GREEN	193 (191.4)	61 62.6	254
	749	245	994

Consider these categorical variables are independent [NULL HYPOTHESIS]

$P(A \cap B) = P(A) \cdot P(B)$

i.e. $Ye \cap Int = P(Ye) \cdot P(Int)$

$= \frac{740}{994} \cdot \frac{749}{994} = 0.561$ per instance

for 994 instances $= 994 \times 0.561$ PROPORTIONAL $= 557.6$

$$x^2 = \frac{(556-557.6)^2}{557.6} + \frac{(184-182.4)^2}{182.4} + \frac{(193-191.4)^2}{191.4} + \frac{(61-62.6)^2}{62.6}$$

$x^2 = 0.073$ degree of freedom $= (r-1)(c-1) = 1 \times 1 = 1$

SO NULL HYPOTHESIS IS TRUE (INDEPENDENT)

From Chi-SQUARE TABLE df = 1 and 5% STATISTICAL SIGNIFICANCE = 3.84 | 0.073 < 3.84