STT 200 5:30 3-24-10

Note Title 3/24/2010

Ch 26 CHI SQUARE $\chi^2 = \sum_{\text{CELLS (CATEGORIES)}} \frac{(O-E)^2}{E}$

ON EXAMPLE 2 DELETE "400."

THERE ARE TWO #5. 5

HAND IN THE YELLOW ASSIGNMENT TODAY.
NAME, SECTION #, RECITATION #.

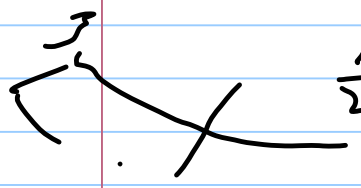
EXAMPLE 1. TOSS COIN 100 TIMES. SUPPOSE 60 H 40 T

ORIENTATION: WHY $\chi^2 = \sum \frac{(O-E)^2}{E}$

ANALOGOUS TO 17,4835 (50,50) COIN

COIN OBSERVED COUNTS H: 60, T: 40 — 2 CELLS.

MODEL OF FAIR COIN ⟹ EXPECTED COUNTS 50 50 (" ")



$$\chi^2 = \frac{(60-50)^2}{50} + \frac{(40-50)^2}{50} = \frac{100}{50} + \frac{100}{50} = 4$$

O for H, E for H

χ^2 TABLE

P-VALUE FOR TEST ??

H_0: $P(H) = .5$ $P(T) = .5$ vs H_A: NOT SO

THIS APPLICATION OF χ^2 DF = #CELLS − 1

DF (THIS APP) = #CELLS − 1 = 2 − 1 = 1

CELLS = 3

EXAMPLE 1.

	AA	Aa	aa	
MODEL (PAST)	.17	.48	.35	= 1
PROJECT TO 100	17	48	35	= 100
OBS	23	52	25	= 100

$$\chi^2 = \frac{(23-17)^2}{17} + \ldots =$$

χ^2 TABLE DF 0.05

2-1=1 ⇒ 4 ~ 3.84

P-(VALUE) ~ 0.05

BETWEEN 0.05 AND 0.025

ROW FOR DF 2 GET χ^2 OF 3 ← .1 .05

SAY P-VALUE > 0.1 ↪ X 4.605

IF GET χ^2 = 12 SAY P-VALUE < 0.005

BIN ≥ 50 S'S

#6. REQUIREMENT FOR χ^2 INDEP SAMPLES.

AND ALL E COUNTS AT LEAST 5.

NOTE: WHAT IS A χ^2 DISTRIBUTION?

$\chi^2_{df} \overset{\text{DIST OF}}{=} Z_1^2 + \cdots + Z_j^2$ INDEP $N[0,1]$.

NORMAL
BINOMIAL
POISSON
$t \sim Z$ DF∞
$\chi^2 \sim$ NORMAL

EXAMPLE #2. GATHER DATA & CROSS CLASSIFY

$n = 400$ 420

SEX	AA	Aa	aa	
M	30	80	40	150
F	40	140	90	270
	70			420

H_0: DIST OF GENE TYPE IS SAME FOR BOTH SEXES

DATA??

AA
AA
M 420 P(M) P(AA)
IF SEX, GENTYPE IND.

JUST DATA

	AA	Aa	aa
M	30/150	80/150	40/150
F	40/270	140/270	90/150

SO YOUR χ^2

	AA	Aa	aa	
M	70	150		150
F				270
	70	220	130	420

$\chi^2 = \frac{(30-25)^2}{25} + \ldots$ THROUGH 6 CELLS.

CELL M AA

LOOKS LIKE DATA IS TELLING US AA

E FOR CELL M

$15 \sim 420 \frac{70}{420} \frac{150}{420}$

$\Rightarrow E = 420 \hat{p}(AA) \hat{p}(M)$

$E = \left(\frac{\text{ROW}}{\text{TOT}}\right) \cdot \left(\frac{\text{COL}}{\text{TOT}}\right) / \left(\frac{\text{GRAND}}{\text{TOT}}\right)$

DF = #CELLS − PENALTIES. (REDUCES TO $(R-1)(C-1)$)

$(2-1)(3-1) = 2$

R. A. Fisher & R.

χ^2 WHAT DOES THIS HAVE TO DO WITH MENDEL'S DATA?

23rd MAY YR — W / WC

72	48
61	

= SAME χ^2

E E FROM "THEORY"

TOLD YOU $\chi^2_{dDF} \overset{DIST}{=} Z_1^2 + \cdots + Z_d^2$

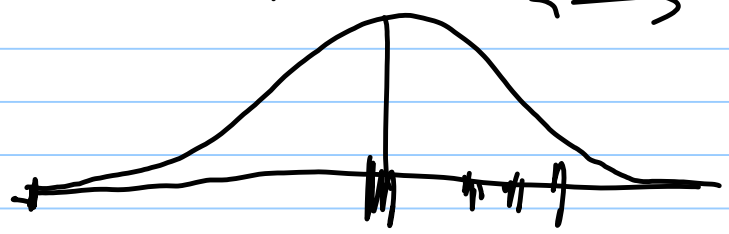
$\chi^2_{d_1} \oplus \chi^2_{d_2} \overset{DIST}{=} \chi^2_{d_1+d_2}$

d_1 INDEP d_2 EXPERIMENTS

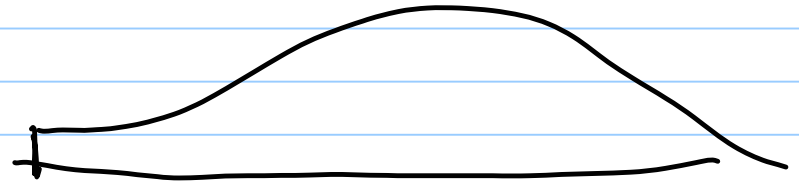
SUPER χ^2 DF LARGE

$\approx$ NORMAL

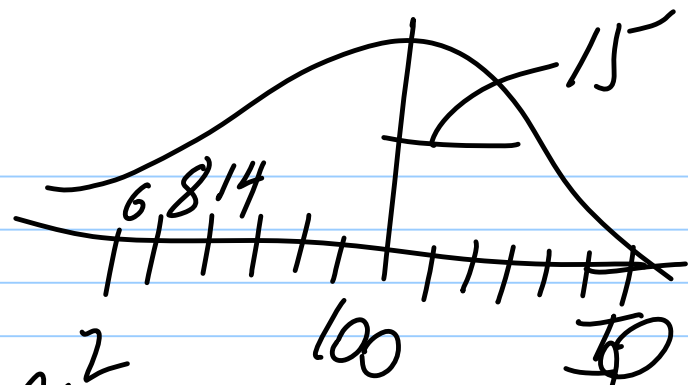
BROAD



~~CRYPTO~~ CYRIL BURT ?



AGGREGATED χ^2



χ^2 STAT LARGE — P-VAL SMALL