Note Title 3/24/2010

STT200 3pm 3-24-10

GIVE SECTION NUMBER AND NAME

SOLVE & RETURN THE HANDOUT FOR TODAY.

I WILL GO OVER SIMILAR EXERCISES.

YOU MAY CONSULT W/ OTHERS

EXAMPLE 2 SHOULD NOT SAY "400."

"CHI SQUARE" χ^2

$$\chi^2_{DF\ k\ (SAY)} \overset{DIST}{=} Z_1^2 + \cdots + Z_k^2$$

SEE χ^2 IN CONTEXTS OF SUM OF SQUARES.

DISTRIBUTIONS

- NORMAL
- POISSON
- BINOMIAL
- t
- χ^2

SO IF $K \to \infty$ χ^2 CAN BE GOTTEN FROM Z-TABLE!!

ANALOGUE OF EXAMPLE 1. "COUNTS DATA"

TOSS COIN 100 TIMES, CLASSIFY

H "EXPECTED"

	H	T
O	60	40
E	50	50

H_0: $P(H) = P(T) = \frac{1}{2}$

$E = 100\left(\frac{1}{2}\right), \ 100\left(\frac{1}{2}\right)$

$$\underset{\text{STATISTIC}}{\chi^2} = \sum_{\text{CELLS}} (O-E)^2/E$$

FOR THE PRESENT DATA

$$\underbrace{\chi^2}_{\text{STATISTIC}} = \underbrace{\frac{(60-50)^2}{50}}_{\text{CELL "H"}} + \underbrace{\frac{(40-50)^2}{50}}_{\text{CELL "T"}} = \frac{100}{50} + \frac{100}{50} = 4$$

ASIDE 70 30

50 50

FOR OUR 60, 40 DATA

$\chi^2 = 4$

$$\chi^2 = \frac{(70-50)^2}{50} + \frac{(30-50)^2}{50} = \frac{400}{50} + \frac{400}{50} = 16$$

∴ P-VALUE $P(\chi^2 > 4)$ IF H_0 TRUE

↑ DISTRIBUTION

← FAIR COIN

2 CELL - BINNING OF DATA

DEPENDS UPON DF WHICH IS FOR THIS APPLICATION

DF = # CELLS − 1
= 2 − 1 = 1

χ^2 TABLE

DF	0.05
1	3.841

CLOSEOD TO 4 = χ^2

SO P-VALUE IS ≈ 0.05.

PERTAINING TO ACTUAL EXAMPLE 1

		AA	Aa	aa
WILD	O	$\boxed{23}$	52	24
MODEL	E	$\boxed{17}$	48	35

MODEL (WILD)

AA	Aa	aa	
.17	.48	.35	(1)

IF MODEL IS CORRECT

E	17	48	35

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

DF = # CELL − 1

= 3 − 1

FOR THIS APP.

SAW 23 17

EXPECTED UNDER MODEL.

SO ASSOC WITH $\frac{(23-17)^2}{17}$

ASSOC WITH LARGE χ^2.

EXAMPLE 2.

~~"400"~~

H_0: MALE SALMON HAVE SAME GENE DISTRIBUTION AS FEMALE SALMON.

ONE DESIGN IS TO SAMPLE 420 FISH AND SORT BY GENE TYPE + SEX

	AA	Aa	aa	
M	30	80	40	150
F	40	140	90	270
	70	120	130	420 TOTAL

#CELLS R C

DF (R-1)(C-1)

~~NOTHING TO DO WITH n-1=419~~

ANOTHER DESIGN FIXES IN ADVANCE #M, #F IN SAMPLE.

THE ANALYSIS IS IDENTICAL IN BOTH DESIGNS.

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

ANALOGUE OF EXAMPLE 2.

— MERGE OF EXAMPLE 2 DATA. (PERHAPS IN CASES WHERE YOU HAVE TO DO SO TO ACHIEVE ALL $E \geq 5$).

H_0 DIST FOR M SAME AS FOR F.

	AA	(Aa or aa)	
M	30	120	150
F	40	230	~~275~~ 270
	70	350	420

AA

M ◯ — $\text{EXP} = 420\, P(M)\, P(AA)$

$$\approx \cancel{420}\,\frac{150}{\cancel{420}}\,\frac{70}{420}$$

$$= \frac{(\text{ROW TOT})(\text{COL TOT})}{N = \text{TOT}}$$

M → $30/150$ $120/150$

F → $40/270$ $230/270$

ARE THESE ≈ SAME?

SO "EXPECTED COUNTS"

O

	AA	NOT AA
m	30	120
F	40	230

E

AA	NOT AA	
$\frac{70\ 150}{420} =$	$\frac{350\ 150}{420} =$	150
$\frac{70\ 270}{420} =$	$\frac{350\ 270}{420} =$	270
70	350	420

OBS ↙ COUNTS

OBS COUNTS.

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

$$DF = (R-1)(C-1) = (2-1)(2-1) = 1$$

MENDEL