STT 200 5:30 3-29-10

Note Title 3/29/2010

(1) KNOW FORMULAS - e.g. $\sum (O-E)^2/E$, $\frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = z$

ALL E ≥ 5 REQUIRED

$\bar{x} \pm t(\alpha/2)\, s/\sqrt{n}$

↑ DF $n-1$ REQUIRES NORMAL POPULATION DISTRIBUTION OR (FOR z) n LARGE (RELIES ON CENTRAL LIM. THEOREM)

2-SAMPLE COUNTERPART $(\bar{x}_1 - \bar{x}_2) \pm t(\alpha/2)\sqrt{s_1^2/n_1 \oplus s_2^2/n_2}$

$\hat{p} \pm z\sqrt{\hat{p}\hat{q}/n}$ OR 2-SAMPLE $\hat{p}_1 - \hat{p}_2 \pm z\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} \oplus \frac{\hat{p}_2\hat{q}_2}{n_2}}$

n_1, n_2 LARGE

FPC (WITHOUT REPL) $= \sqrt{\frac{N-n}{N-1}}$

* [CHI-SQ. DF = # CELLS − 1 FOR "GOODNESS OF FIT"]

E = n · (PROP OF CELL) FOR EACH CELL, DOES NOT DEPEND UPON DATA

EXAMPLE OF GOODNESS OF FIT: MODEL AA Aa aA aa
.1 .4 .3 .2

CRITERION ALL $E \geq 5$

$$\chi^2_{STAT} = \sum_{CELLS}^{4} \frac{(O-E)^2}{E}$$

(CATEGORIES)

O { CLASSIFY RANDOM SAMPLE $n = 60$ } 4 20 22 14

$E = np = 6\ 24\ 18\ 12$

$$\chi^2 = \frac{(4-6)^2}{6} + \frac{(20-24)^2}{24} + \frac{(22-18)^2}{18} + \frac{(14-12)^2}{12}$$

TAKE CARE

↖ CONTRIBUTION FROM CELL AA

DF = (GOODNESS OF FIT) = (#CELLS − 1)

DF = 4 − 1

SO IF (SUPPOSE) $\chi^2 = 2.2$ (??)

P-VALUE = 

DF = 3

.1 < P-VALUE < .9 UNINFORMATIVE

WANT TO KNOW THAT χ^2 BIG ≡ EVIDENCE AGAINST MODEL.

P-VALUE SMALL

χ^2 BIG $\sum \frac{(O-E)^2}{E}$

ABOUT $E \geq 5$ ALL CELLS

REWRITE (MERGE)

	AA or aa	Aa or aA
E	18	24 18

DF = 2

	AA	Aa or aA	aa
O	6	24	12
E	6	24	18

TO ACHIEVE ALL $E \geq 5$ OR EVEN GIVE CUSHION $E \gg 5$

χ^2 TEST OF INDEPENDENCE:

BATTERS:	LONG	MED	SHORT
WOOD BAT			
METAL BAT			

USE RANDOMM SAMPLE OF $n=100$ BATTERS - EACH GETS ONE HIT FOR THE RECORD.

SORT - BATTERS RANDOMLY ASSIGNED TO WOOD & METAL.

	L	M	S	
W		10		60
M				40
		20		100

INDEP LOOKS LIKE PROPORTIONALITY

20	15	40
40	30	80

REMAINDER.

CELL W M

RULE $\frac{(\text{ROWTOT})(\text{COLTOT})}{\text{TOT}}$

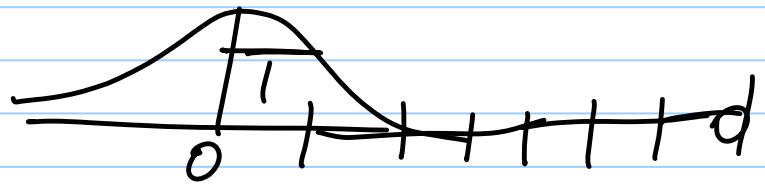
$E?$ $\quad E = \frac{60\ 20}{100} = 12$ WON'T ALWAYS BE AN INTEGER (USE DECIMAL)

a. CONTRIBUTION OF CELL WOOD MED $= \frac{(O-E)^2}{E} = \frac{(10-12)^2}{12}$

$\frac{(10-12)^2}{10} = .4$

6. STANDARDIZED RESIDUAL $\frac{O-E}{\sqrt{E}} \approx Z$ R.V.
IF INDEP IS SATISFIED.

SO SHOULD YOU SEE $\frac{O-E}{\sqrt{E}} = ?$



USE OF χ^2 TABLE

$\Rightarrow$ DF $= (R-1)(C-1)$ (JUST ABOVE)

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

(SPSS) χ^2 STATISTIC $= 4.23$

.99999 GOOD FIT!

.00001 BAAAD FIT! REJECT MODEL

DF	.1
2	4.60

So P-value would be ~ .1

So ~10% of time you'd see $\chi^2 \geq 4.6$ even if dist indep of bad type.

Merge of indep χ^2.

Study 1 χ^2 stat = 7.2 DF = 4

Study 2 χ^2 stat = 16.1 DF = 9

Can use these (provided they were indepent)

Super χ^2 (merge really) 16.1 + 7.2 DF = 9+4

DF 13

2236

P = .05

23.3 new χ^2 13 DF

OUR P-VALUE $< .05$

BACK TO CI

CALC α

EXAMPLE OF 0-1 DATA Z-BASED CI FOR p.

ASSIGN 100 BATTERS FINDING

TEST HYP FAIR χ^2 VS.

CI APPROACH (ESTIMATE $p = P(\text{WOOD})$

$\hat{p} = 60/100 = 0.6$ 95% Z-CI FOR p

$.6 \pm 1.96\sqrt{.6 \cdot .4}/\sqrt{100}$

χ^2 JUST DISCUSSED

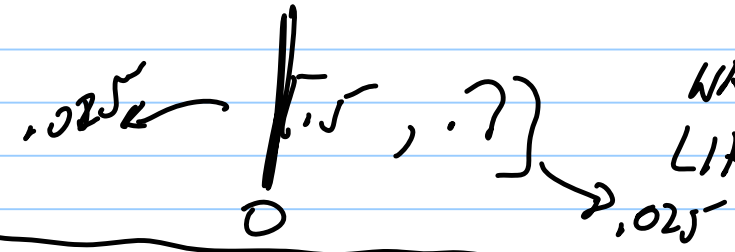
WOOD	METAL	
60	40	$\bar{x}$
50	50	$\bar{e}$

$$\chi^2_{DF=1} = \frac{(60-50)^2}{50} + \frac{(40-50)^2}{50}$$

$$= 2 + 2 = 4$$

P-VALUE .05

$.6 \pm 2\sqrt{.5/10} = .6 \pm .1$ or $(.5, .7)$.



WHAT CI IS DOING IS LIKE EVIL TWIN OF χ^2