# Topic 11 Summary

2022 S2

11.1 Intro lecture

Our hypothesis tests so far:

H

$H_0$: proportion = val

$H_0$: mean = val

↓

Proportion test

known SD ↓

1 Sample Z test

↓

1 Sample t test

P

use Binomial or Normal (Z)

use Normal

use $t_{n-1}$

↓

$H_0$: $mean_{diff}$ = val

1 sample paired t test

↓

$H_0$: $mean_1 - mean_2$ = val

2 sample t test

Our final hypothesis tests:

Qualitative Data

$\chi^2$ tests

① Goodness of Fit

Ho: model fits data

- Homogeneity

② Independence

Ho: variable 1 is independent of variable 2

Form of TS:

$$= \sum_i \frac{(O_i - E_i)^2}{E_i}$$

A: Cochran's Rule

No more than 20% of categories have $E < 5$

Quantitative Data (Bivariate)

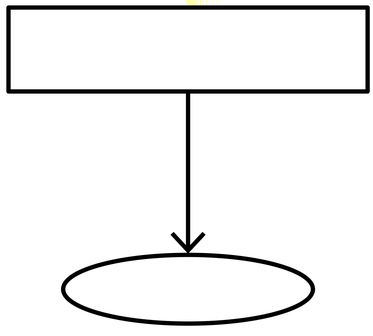
Regression Test

Ho: linear model does not fit data (slope = 0)

1 sample t test :)

L35, p13f

# Summary of Hypothesis Testing – Chi-square Test (Goodness of Fit)




| | O | E | $\frac{(O-E)^2}{E}$ |
|---|---|---|---|
| 1 | 7 | 10 | $\frac{(7-10)^2}{10} = 0.9$ |
| 2 | 7 | 10 | 0.9 |
| 3 | 9 | 10 | |
| 4 | 14 | 10 | $\frac{(14-10)^2}{10} = 1.6$ |
| 5 | 10 | 10 | |
| 6 | 13 | 10 | |
| | | | 4.4 |
| Totals | 60 | 60 | k = 6-1 = 5 |

$\alpha = 0.05$

**H**

$H_0$: die is fair → model fits $P(1) = P(2) \ldots = P(6) = \frac{1}{6}$

$H_1$: die is not fair

**A**

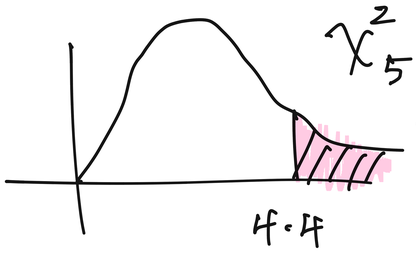
Cochran's Rule

No more than 20% of categories have $E < 5$

**T**

$T = 4.4$

**P**



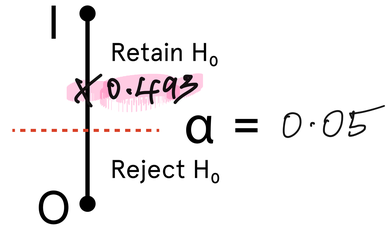

0.493

1-pchisq(4.4, 5)

**C**

Retain $H_0$

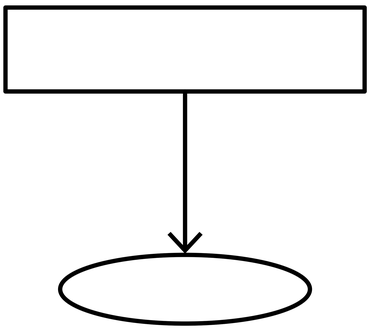
speedy way: chisq. test p23




**CI**

# Summary of Hypothesis Testing - Chi-square Test (~~Goodness of Fit~~ Independence, L35, p27f)

Assuming $H_0$ is true



| O | E | $\frac{(O-E)^2}{E}$ |
|---|---|---|
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

k =

$\alpha = 0.05$

**H**

$H_0$: Smoking & Background are independent

$H_1$: are not

**A**

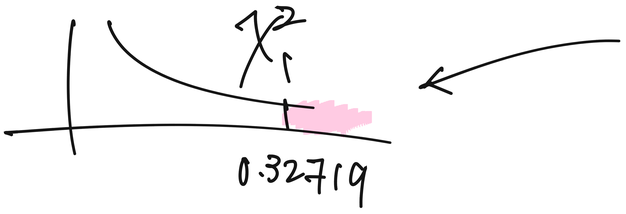
**T**

TS = 0.32719

$\chi^2_1 \leftarrow (2-1)\times(2-1)$

p79
L35

**P**

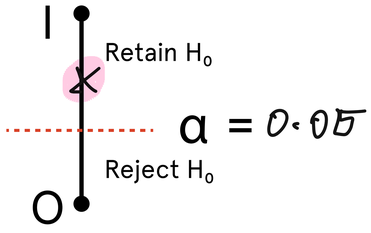



0.5673

**C**

Retain $H_0$

chi sq. test
p28, 29




**CI**

L30, p11, 24

# Summary of Hypothesis Testing – 1 Sample T Test

applied to Regression

Assuming $H_0$ is true

mean =

SD = ?

n =

OV =

SD =

| Sample Mean | EV = mean | |
|---|---|---|
| | SE = $\frac{SD}{\sqrt{n}}$ | X |

$\hat{y} = \alpha + \beta x$

$\alpha = 0.05$ ← slope parameter

## H

$H_0$ : ~~mean =~~ $\beta = 0$ ← No linear regression

$H_1$ : $\beta \neq 0$

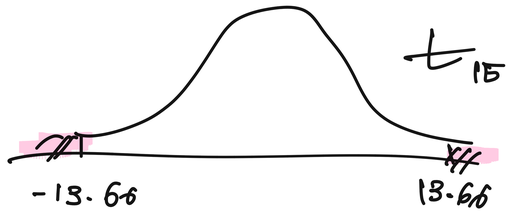
## A

## T

$$TS = \frac{OV - EV}{SE} = -13.66$$

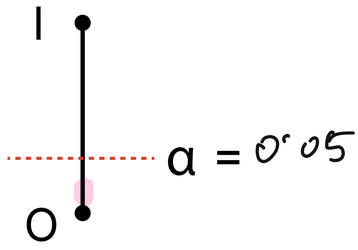
## P

PV = tiny




## C

Reject Ho → data is consistent with linear regression.




## CI