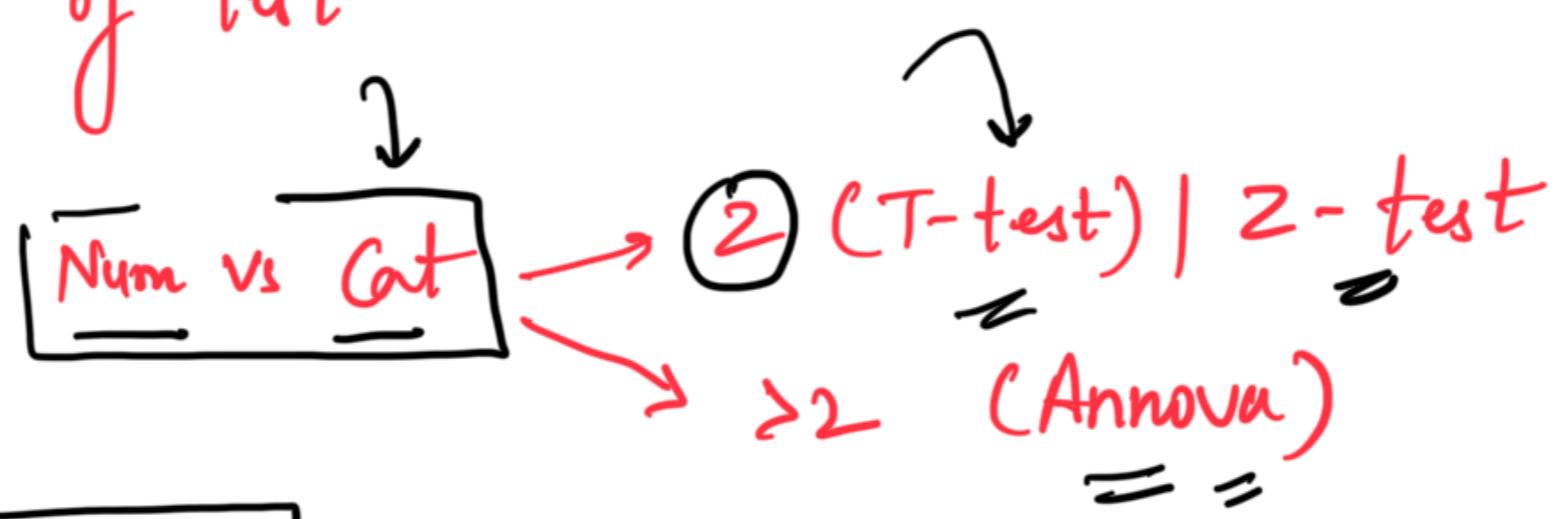


Class starts at 9:05 PM

Agenda

- ① Types of Test ✓
- ② Degree of Freedom ✓
- 3) Chi² goodness of fit test
- 4) Chi² Test for independence
- 5) Problem Solving. ✓

Types of test



Cat vs Cat → Chi square test
=

Num vs Num → Correlation

✓ gender product ✓ =
→ { M A } 2tut | t-test X
{ F B }
{ M C }
{ F D }

Degrees of Freedom → "flexibility"

Setup 1: Salary

✓ $P_1 - \boxed{35}$ ✓
✓ $P_2 - \boxed{n}$ |||
✓ $P_3 \rightarrow 24$ ✓

Avg → 25 L
=

$$35 + 36 + x = 35$$

$$\downarrow \downarrow \downarrow \quad 35 + \boxed{x} + 34 = 35$$

$\boxed{n-1}$

$$3 - 1 = 2$$

||| = =

$$y \\ n \\ \frac{34}{Avg \rightarrow 35}$$

Height & Weight

n_1	n_2	n_3
\bar{H}	\bar{w}	A
73	85 ♂	
68	73 ♂	
74	78 ♂	
x	82 ♂	
62	70 ♂	
Avg	71	81.2
	$n_1 - 1$	$n_2 - 1$

$$Dof = \frac{(n_1 - 1) + (n_2 - 1)}{n_1 + n_2 - 2}$$

$$n_1 + n_2 + n_3 - 3$$

$$= (n - 1) + (n - 1)$$

$$= 2(n - 1)$$

\equiv \equiv

$$(n_1 - 1) + (n_2 - 1) + (n_3 - 1) \\ = n_1 + n_2 + n_3 - 3.$$

Century vs Win

Win

what is the min value
that is req to
construct tri. table?

Century

	F	T	
F	50	-	314
T	-	-	46
176	184	260	

DOP = 1

$$(s-1) \times (c-1) = 2 \times 1$$

$$(2-1) \times (2-1) = 1 \times 1$$

$$= 1$$

$$\frac{176}{20} \\ 176$$

→

	A	B	C	D	
X	50	50	50	349	
Y	50	50	50	151	
Z	150	150	200	150	650

DOP = 6

$$Dof = (s-1) \times (c-1)$$

$$= (3-1) \times (4-1)$$

$$= 2 \times 3$$

$$= 6$$

$$Dof = 6$$

DOF ↑ Variability increases

$$\overline{x+y+1} = \overline{443}$$

$$\overline{x+y} = \overline{442}$$

(x,y)

$$x+y = 442$$

$$\rightarrow x + 32 + 1 = 443$$

$$x = 443 - 33$$

$$x = 410$$

$$\overline{x+y} = \overline{443}$$

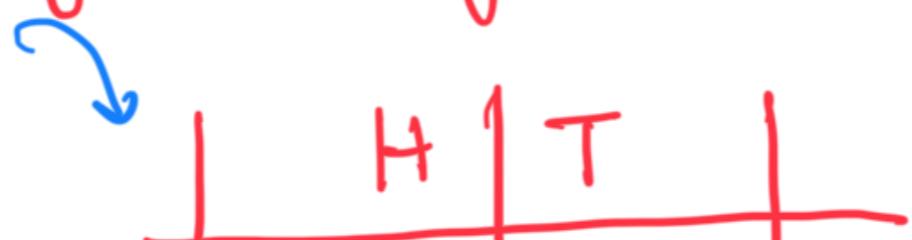
Coin Toss

To check if coin is fair

50 times

1000
=

H



(x-1) v (z-1)

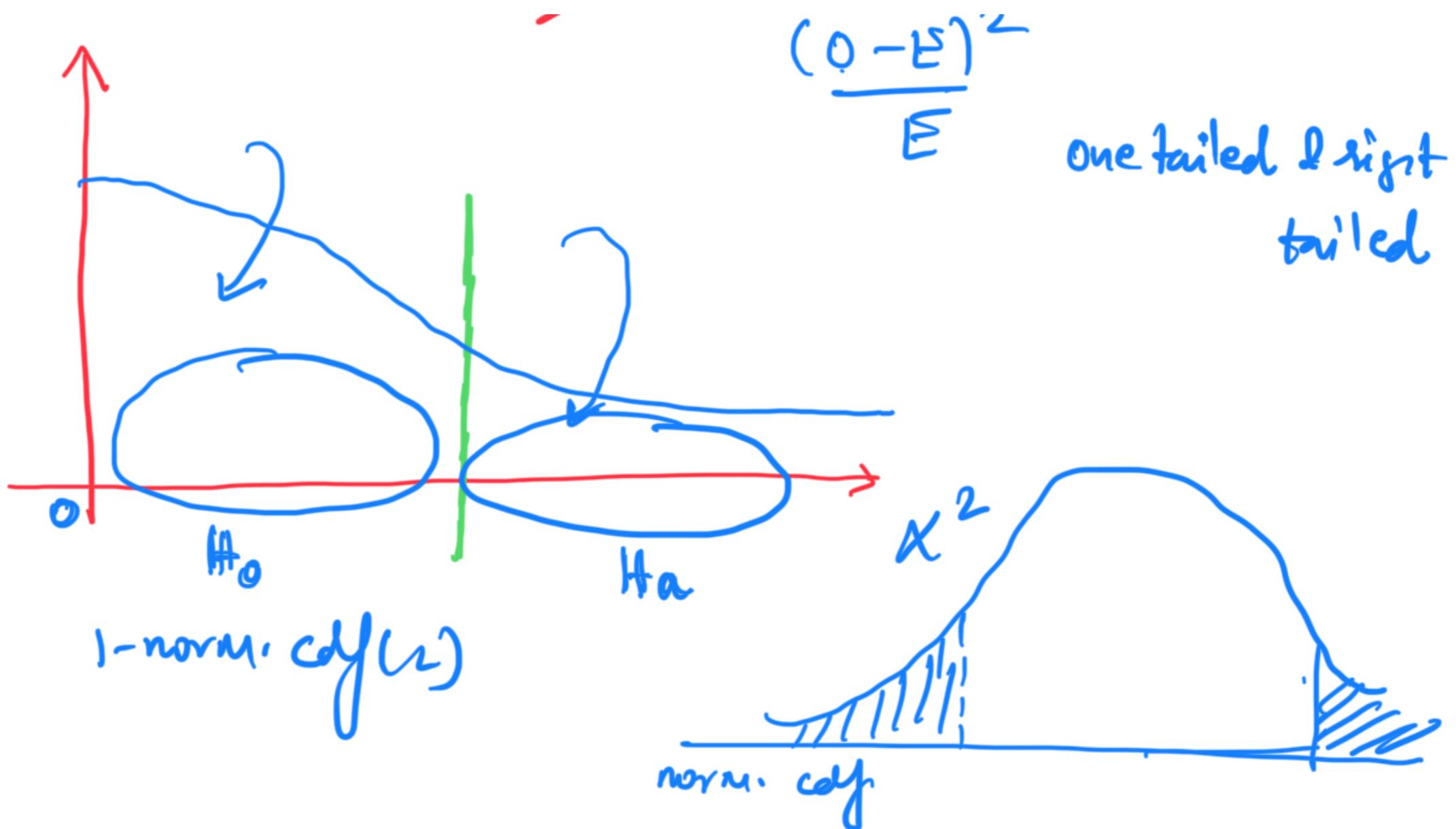
Expected	25	25	50
Actual	28	22	50			
	(X)	(X)				

H₀ coin is fair

$$\begin{aligned}
 &= (25 - 28) + (25 - 22) \\
 &= -3 + 3 \\
 &= (25 - 28)^2 + (25 - 22)^2 \\
 &= (3)^2 + (1)^2 \\
 &= 18
 \end{aligned}$$

$$\chi^2 = \sum_{k=1}^{25} \frac{(O-E)^2}{E}$$

H₀ coin is fair



Test for independence

Survey :- Gender impacts online | offline purchases

Preference vs Gender

(Online & offline)
2 catg

(M | F)
2 categories

H_0 :- Gender and preference are independent

H_a : Gender and preference are dependent

Obs	values		
	M	W	
offline	527	72	599
online	206	102	308
//	733	174	907

$$907 \times 2\% = 599$$

Expected Values

	M	W	
Online	485	115	599

66%

Offline	249	59	108	(24%)
	735	174	907	

66% of 735 \rightarrow 484

66% of 174 \rightarrow 115

34% of 735 \rightarrow 249

34% of 174 \rightarrow 59

Assumptions of Chi-square test

- ✓ 1) Cat vs Cat
- ✓ 2) Obs should be independent of one another
- ✓ 3) Emp frequency in each cell should be > 5

J
Z
N



rogerius
M