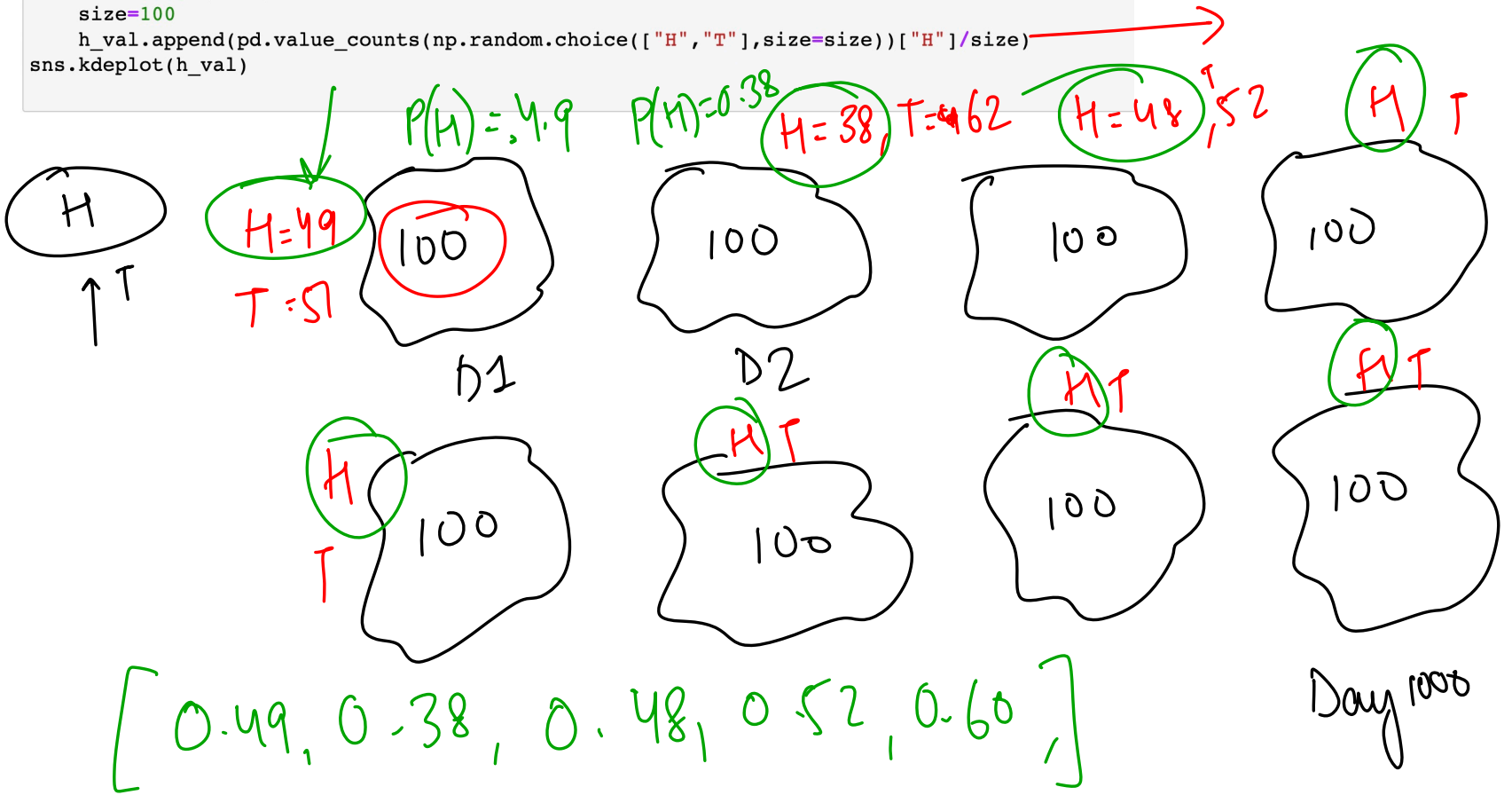
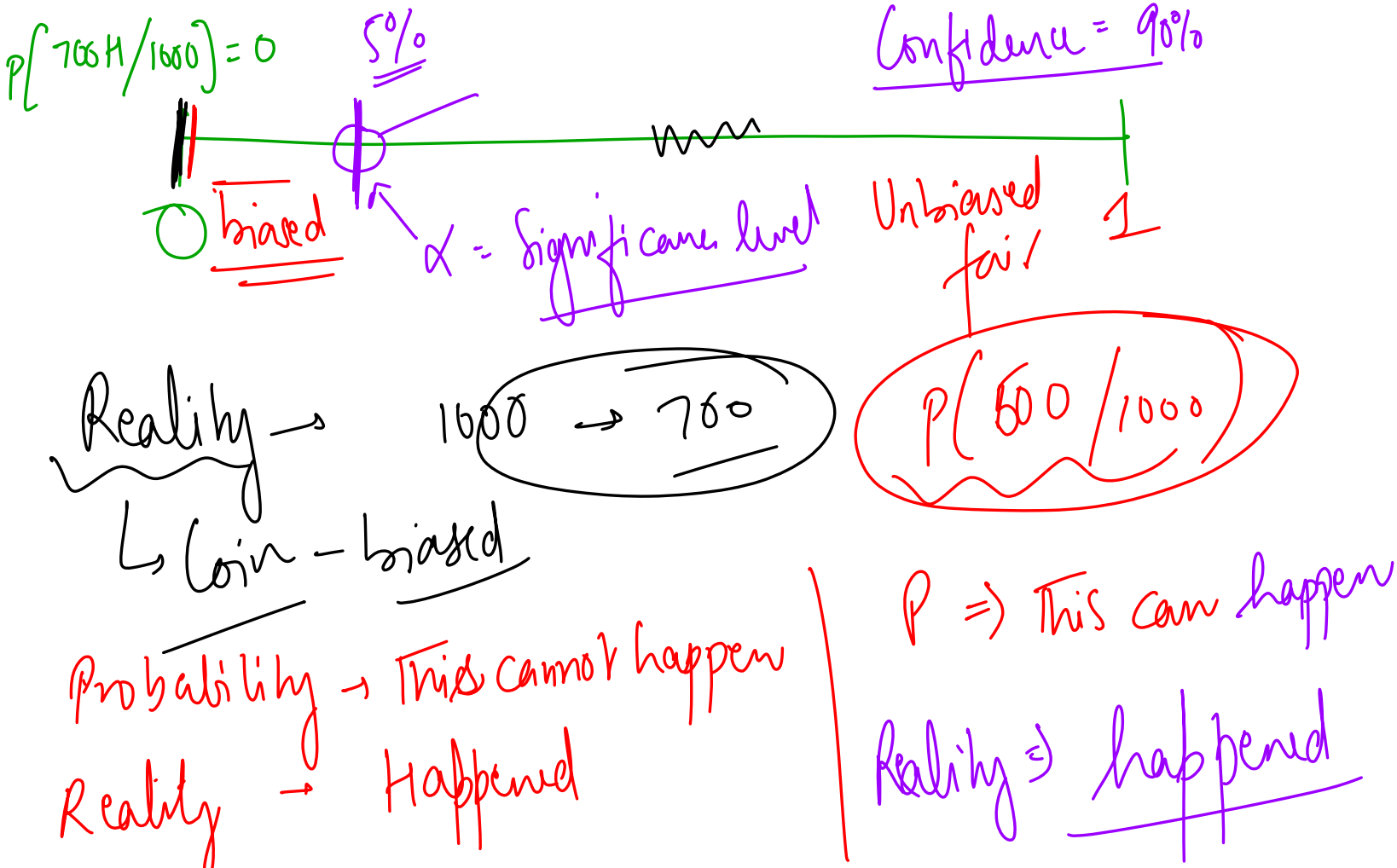
HYPOTHESIS TESTING - 2

Z Test

```
h_val=[]
for i in range(1000):
    size=100
    h_val.append(pd.value_counts(np.random.choice(["H","T"],size=size))["H"]/size)
sns.kdeplot(h_val)
```



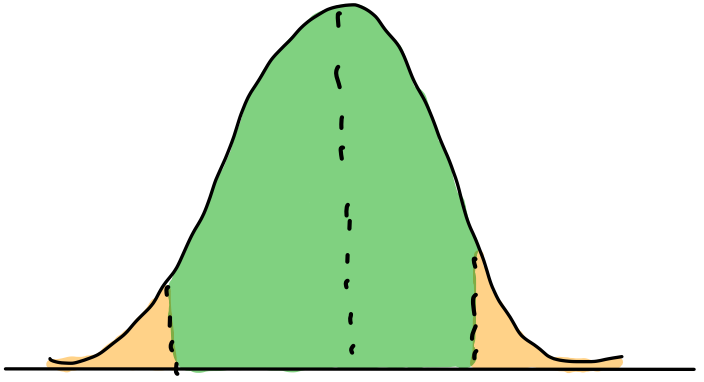

P[700 H/1000] = 0
5%
Confidence = 90%
biased
α = Significance level
Unbiased fair
1
Reality →
1000 → 700
P(500/1000
↳ Coin - biased
Probability → This cannot happen
Reality → Happened
P ⇒ This can happen
Reality ⇒ happened

0.05

pvalue $<$ α $\rightarrow$ Reject H_0

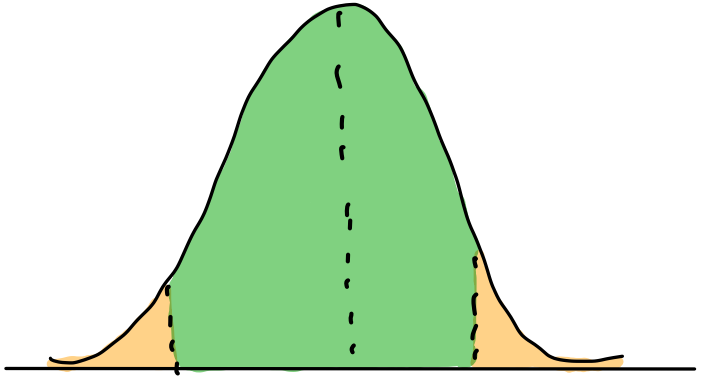
0.05

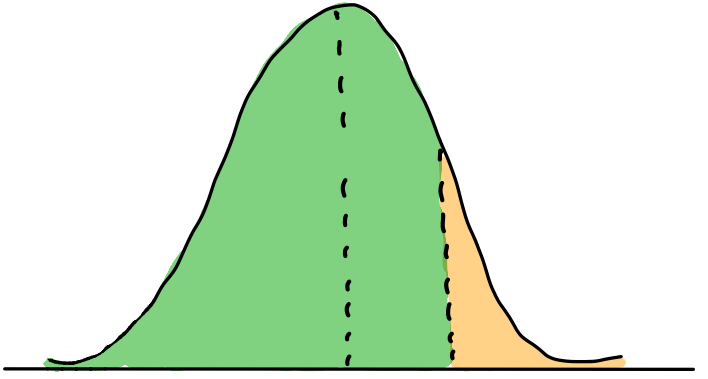
pvalue $>$ α $\rightarrow$ Fail to Reject H_0

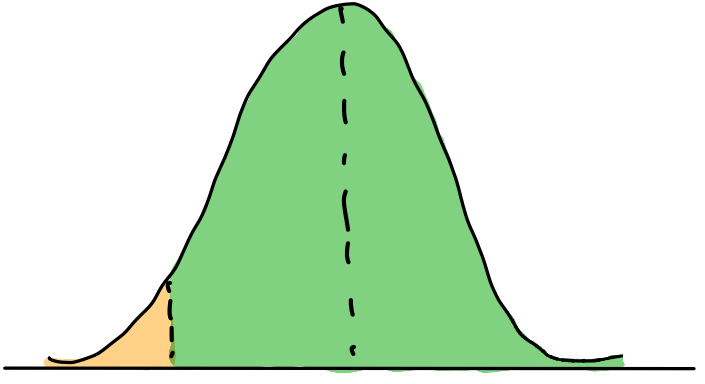


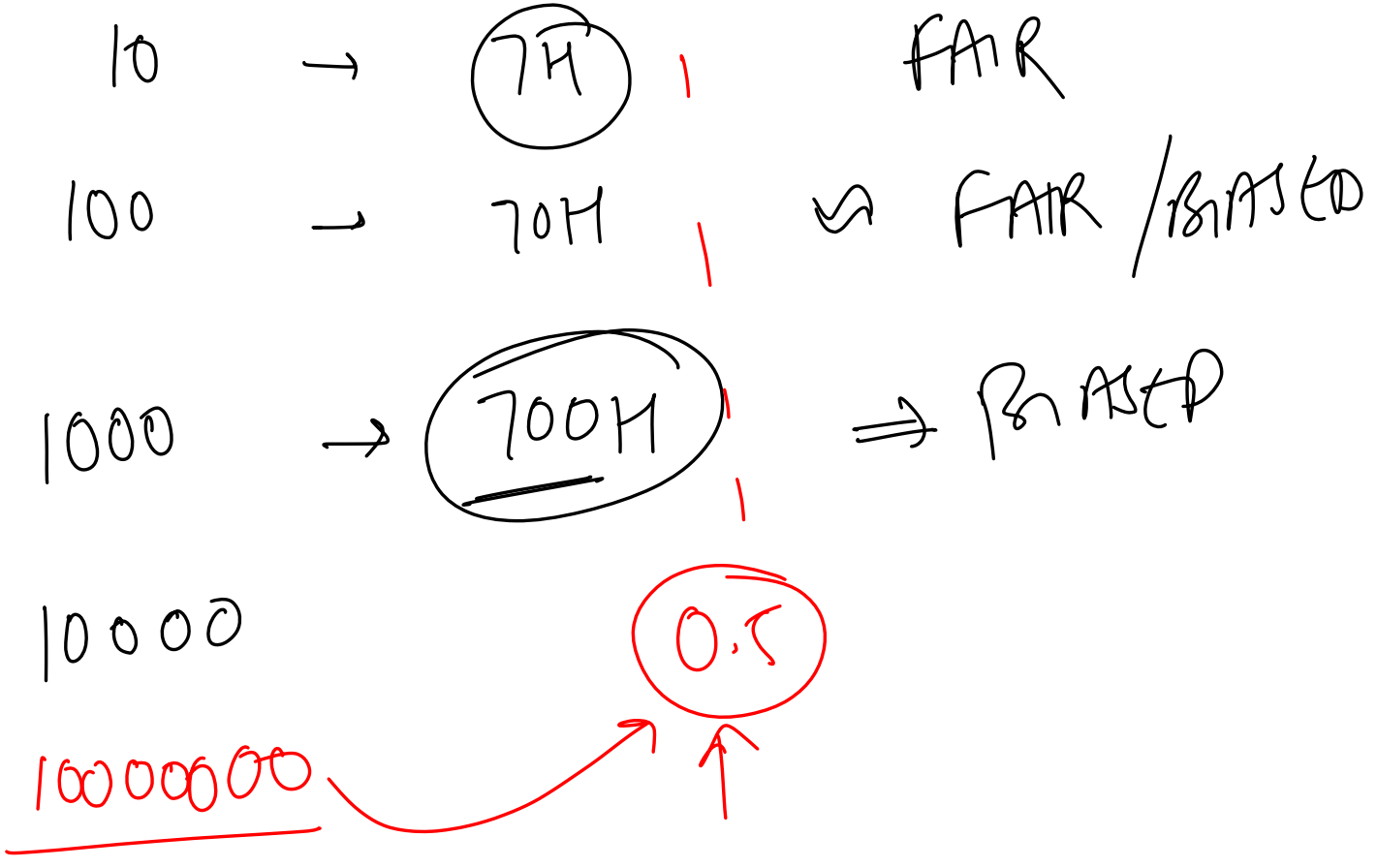
H_0: Coin is fair

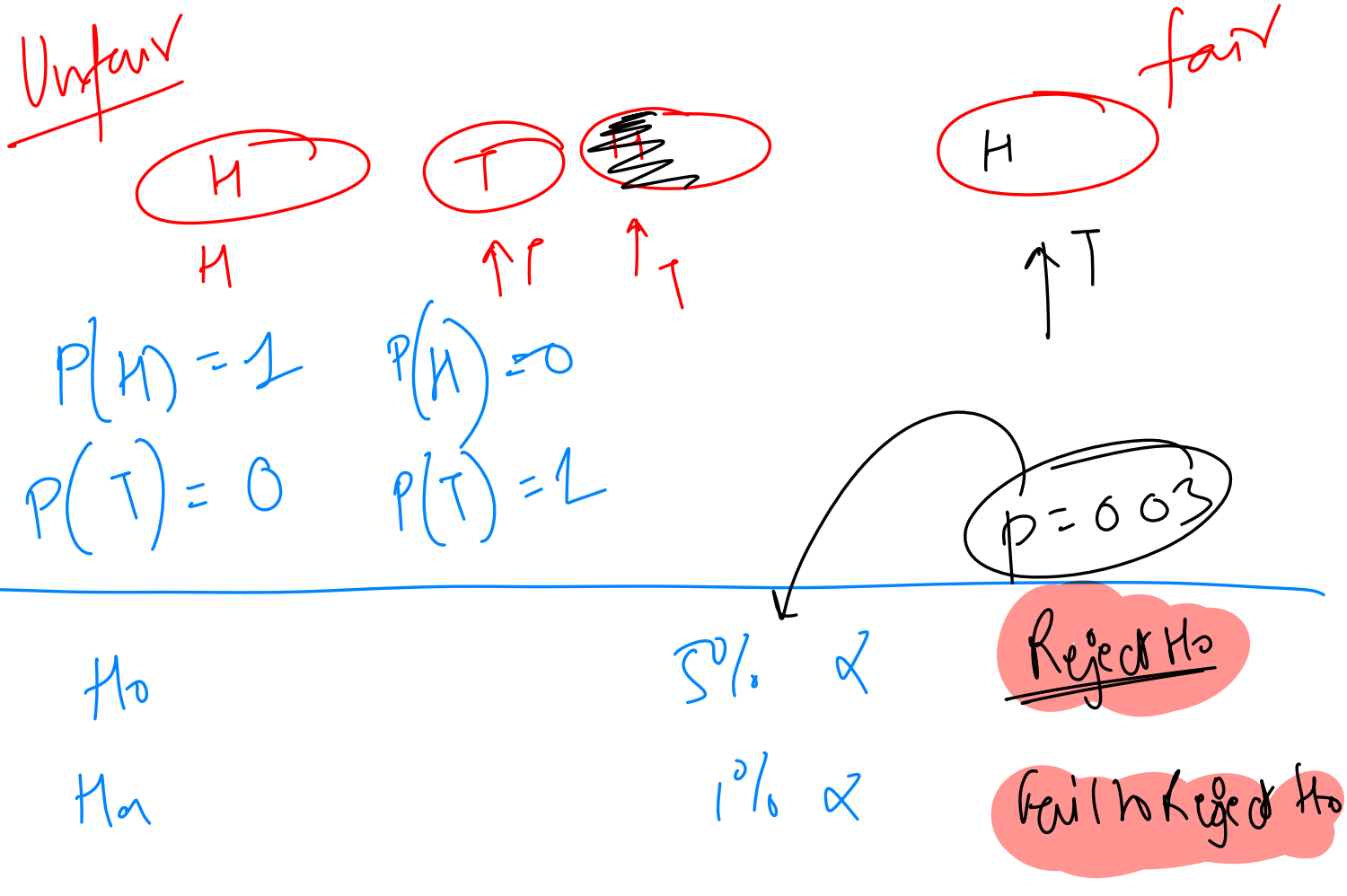
H_a: Coin is biased.








10
7H
FAIR
100
70H
FAIR / BIASED
1000
700H
BIASED
10000
0.5
10000000


Unfair
H
T
H
H
T
fair
H
T
P(H) = 1
P(T) = 0
P(H) = 0
P(T) = 1
p = 0.03
H0
Ha
5% α
1% α
Reject H0
fail to reject H0

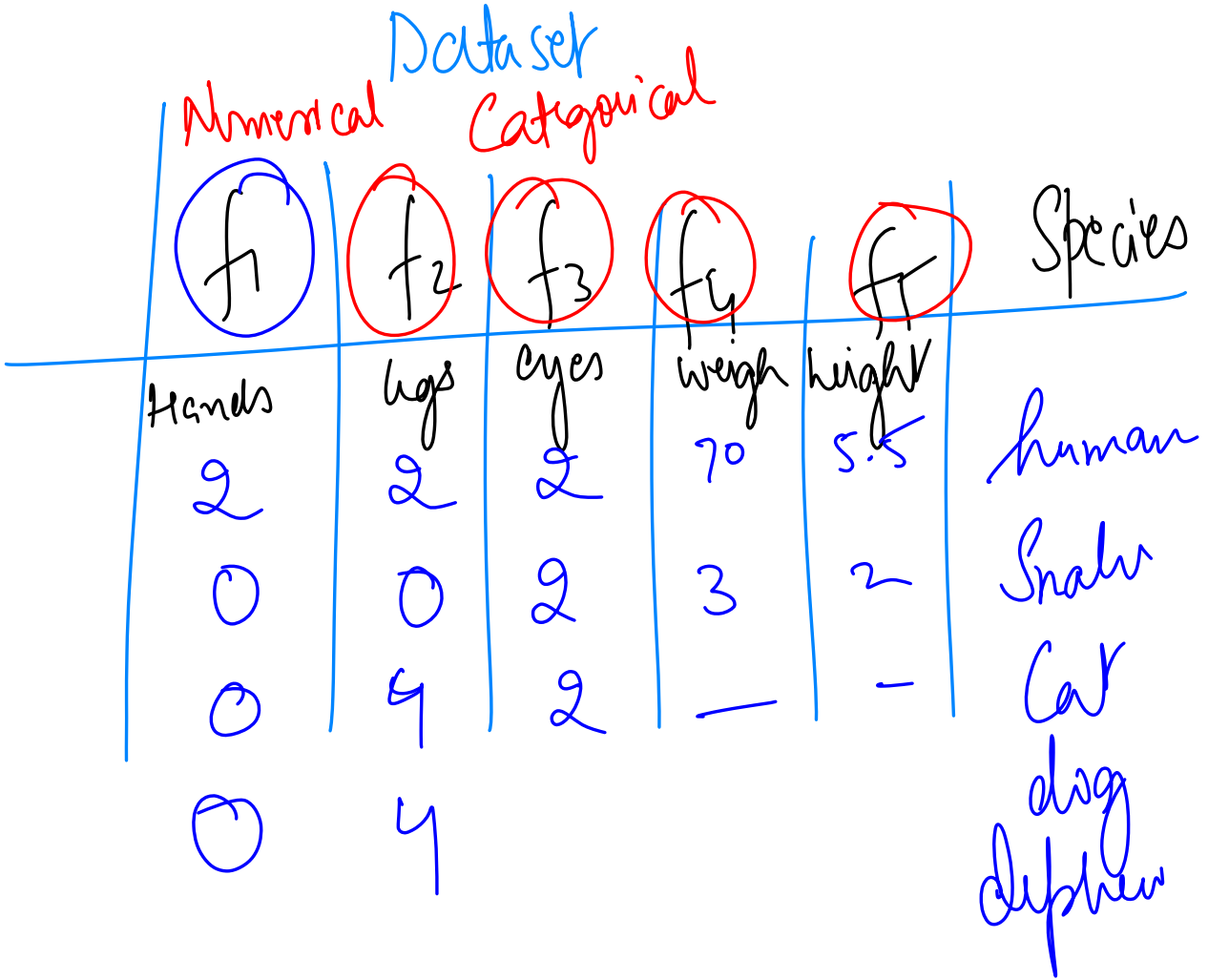
$1 - \text{Confidence} = \alpha$

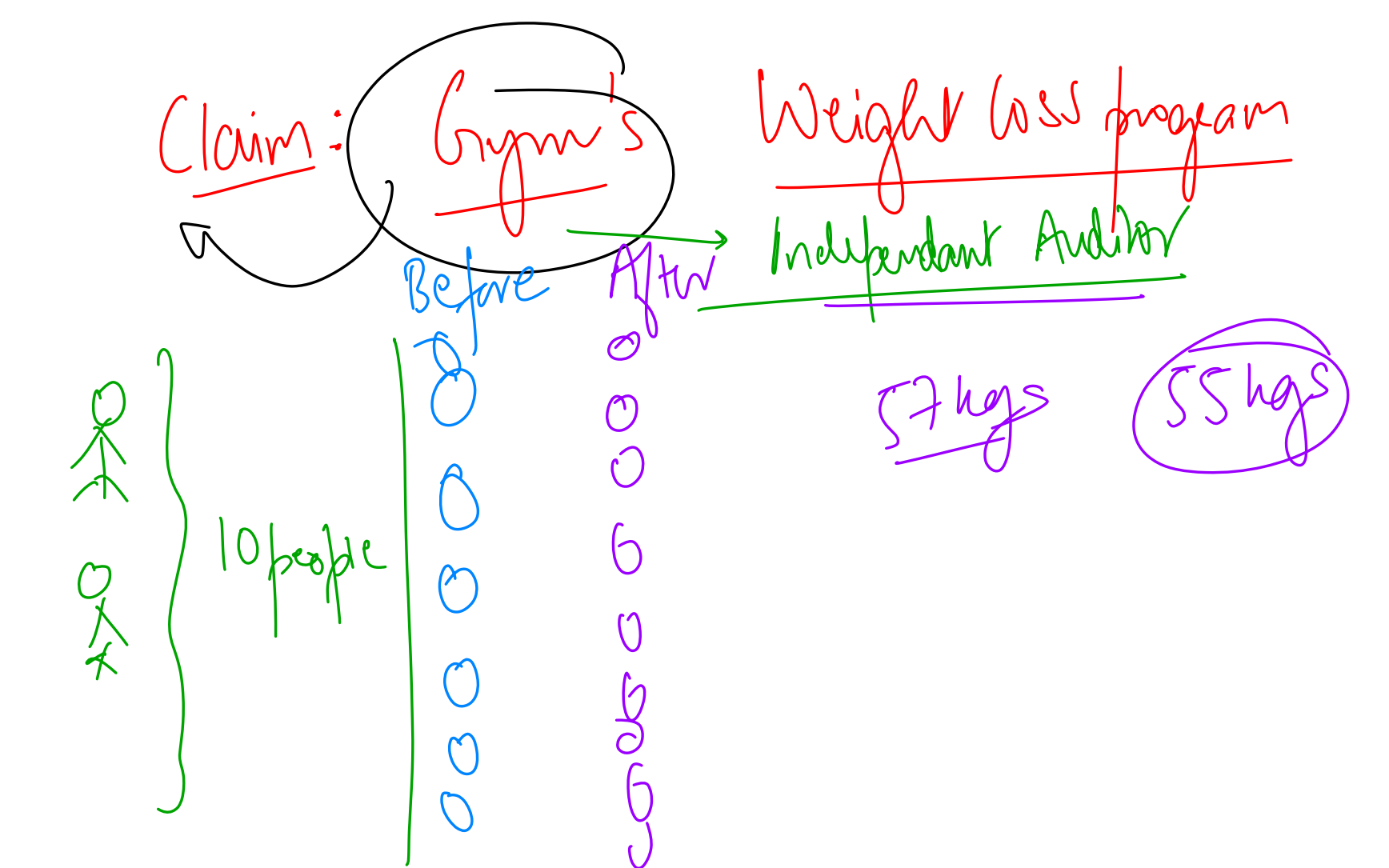
$1 - 90\% = \alpha \Rightarrow \alpha = 0.1$

Generally Acceptance

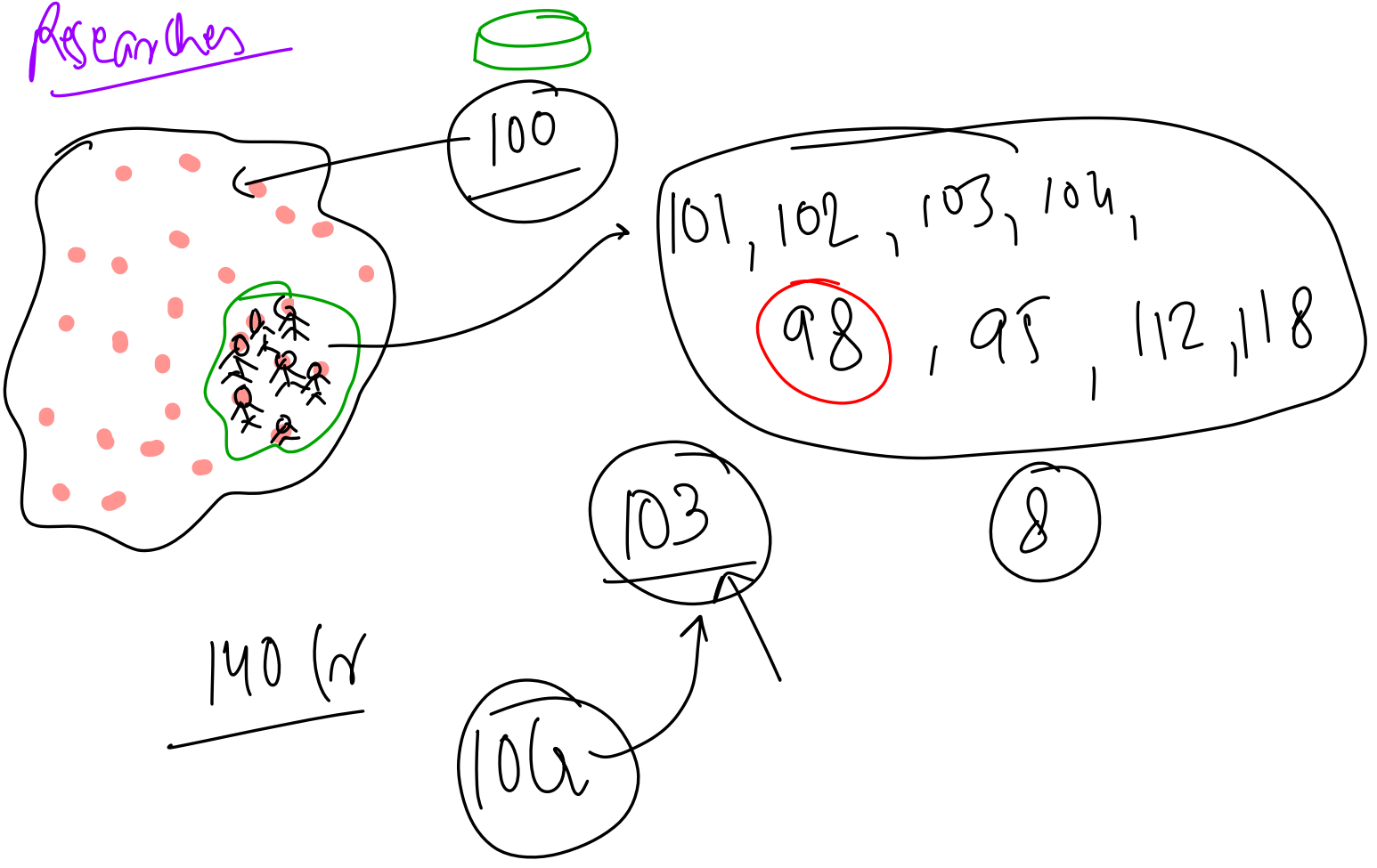
$\alpha = 0.05$ / Confidence = 95%

5%


Dataset
Numerical
Categorical
f1
f2
f3
f4
f5
Species
Hands
legs
eyes
weigh
height
2
2
2
70
5.5
human
0
0
2
3
2
Snake
0
4
2
—
—
Cat
0
4
dog
dolphin


Claim:
Gym's
Weight loss program
Independant Auditor
Before
After
10 people
57 kgs
55 kgs

Researches



RV

probability mass function

Discrete

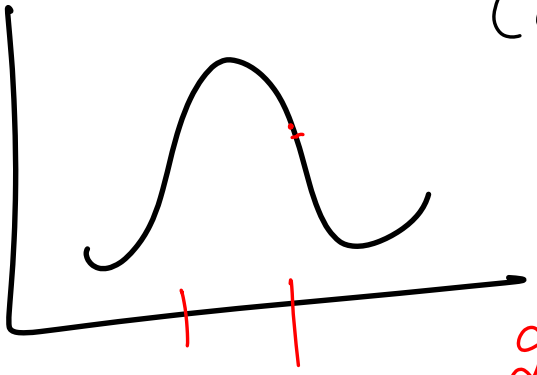
Binomial

pmf

cdf

Continuous

Normal



probability density fn:

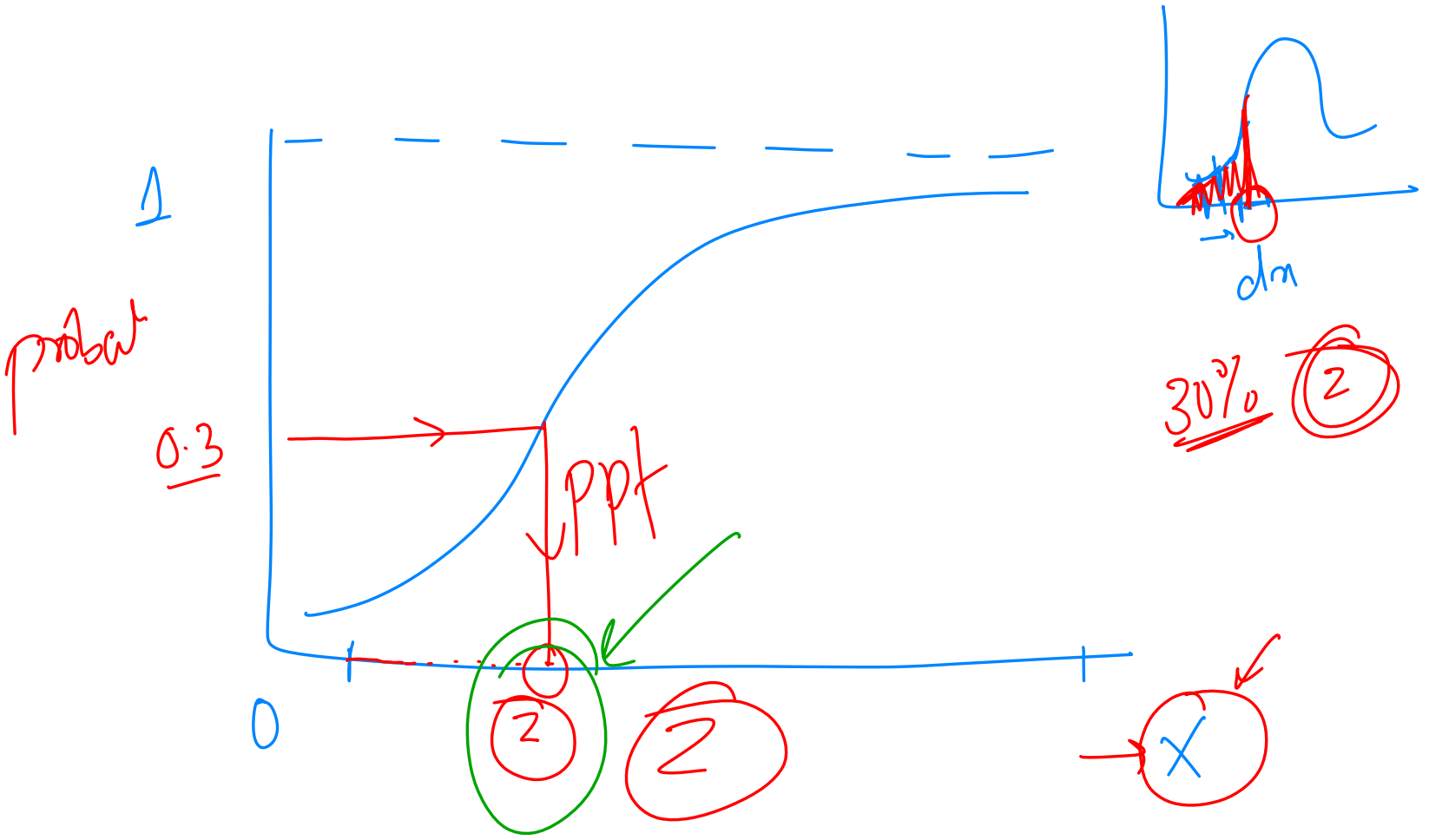
pdf

cdf

P[X ≤

2.12345678

2.12345678


1
product
0.3
ppt
0
z
z
dx
30%
z
x