

Why is "Hypothesis Testing" important?

We want to weigh up evidence to make evidence based decisions.

Hypothesis Testing Framework

S State significance level

① Set up research question

H Hypotheses : H_0 vs H_1

② Weigh up evidence

A Assumptions

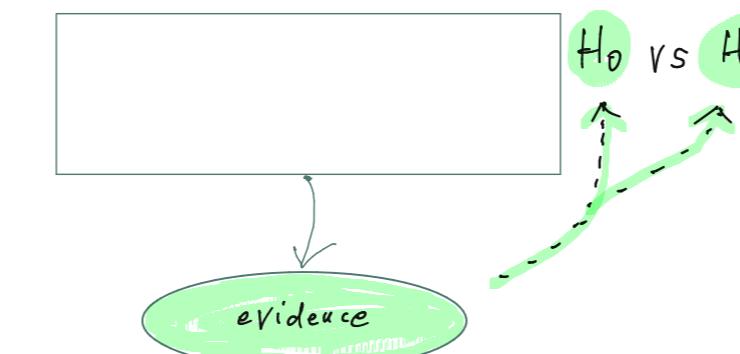
T Test Statistic

P P-value

③ Explain conclusion

C In terms of p-value and context

Module 4: Decisions with Data



what is the best decision?

⑨ Hypothesis Testing

1 Sample Z

10 Tests for Mean $\left\{ \begin{array}{l} 1 \text{ Sample T} \\ 2 \text{ Sample T} \end{array} \right.$

11 Tests for Relationship $\left\{ \begin{array}{l} p \text{ Chi-squared} \\ \text{Regression test} \end{array} \right.$

LO7 [capstone] Given real multivariate data and a problem, formulate an appropriate hypothesis and perform a range of hypothesis tests.

LO8 Interpret the p-value, conscious of the pitfalls associated with testing.

H_0 : parameter = value

H_1 : parameter ≠ value $\left\{ \begin{array}{l} \text{2 sided} \\ > \text{value} \\ < \text{value} \end{array} \right. \begin{array}{l} \\ \text{1 sided} \end{array}$

state & justify

TS - measures the difference between what is observed (data) and what is expected (H_0)

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

PV = chance of observing TS or something more extreme.

• Retain H_0 (large p-value)

• Reject H_0 (small p-value)

Assume H_0 is true



A large, empty oval shape outlined in brown, centered on the page. A small, dark gray arrow points downwards from the top center of the oval.

Box Model
Sample Sub
(or M.)

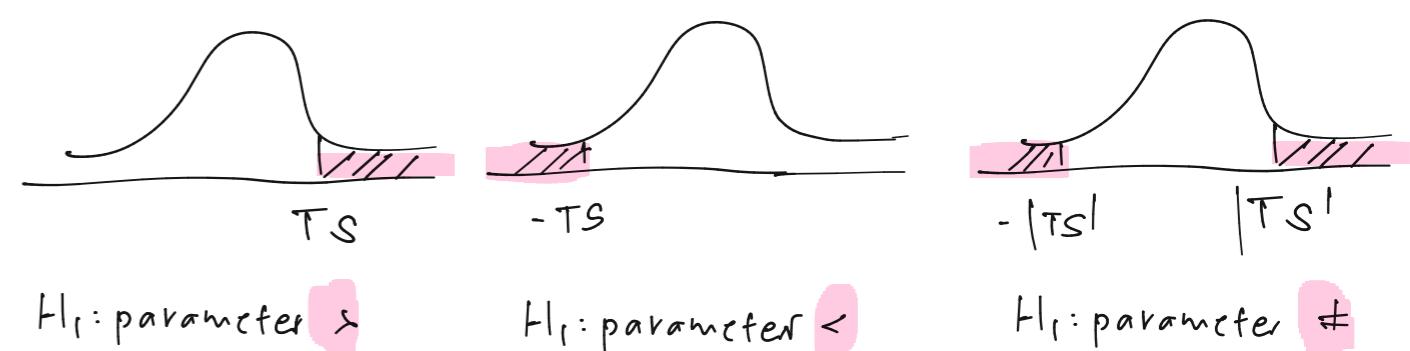
$$S = 0.05$$

A hand-drawn diagram illustrating the dissociation of water. On the left, there is a blue square containing the letters "H2O". To its right is a yellow blob containing the letters "H₃O". Below the yellow blob, the letter "H" is written above a vertical arrow pointing upwards, indicating the direction of equilibrium.

A

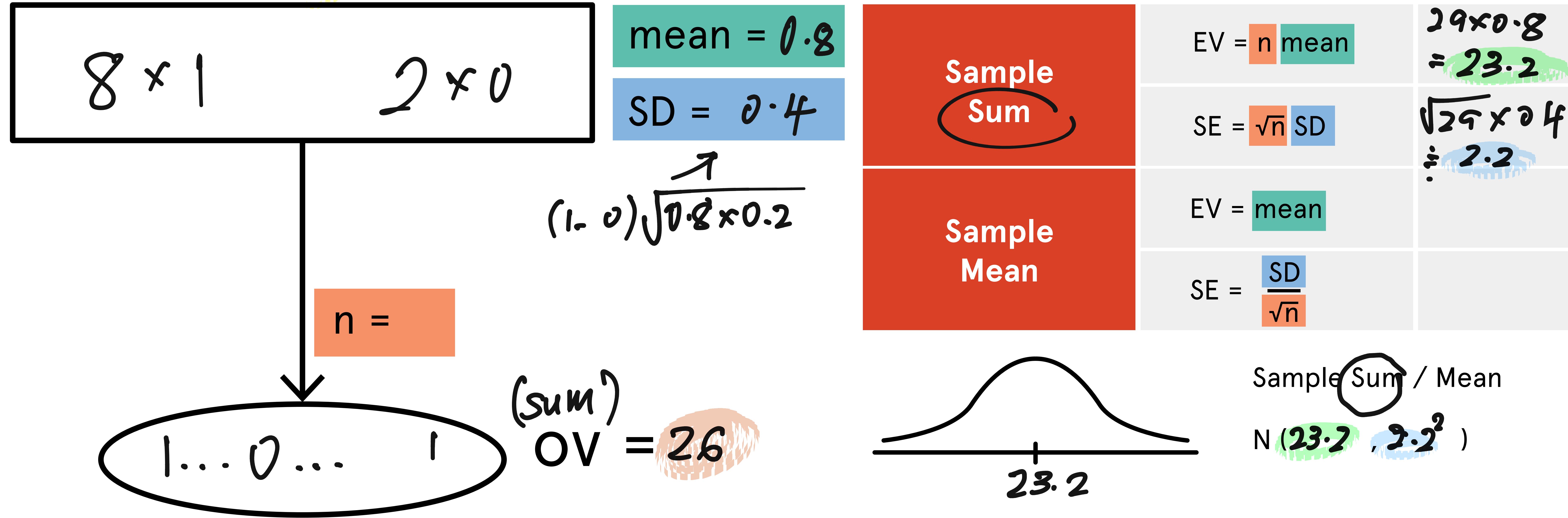
$$TS = \frac{OV - EV}{SE} = \langle \text{number} \rangle$$

 PV = P_{H_0} (getting TS or something
more extreme)



Summary of Hypothesis Testing - Proportion Test

Assuming H_0 is true

**S**

$$\alpha = 0.05$$

H

$$H_0: P(\text{peanut desensitisation}) = p = 0.8$$

$$H_1: p > 0.8$$

A

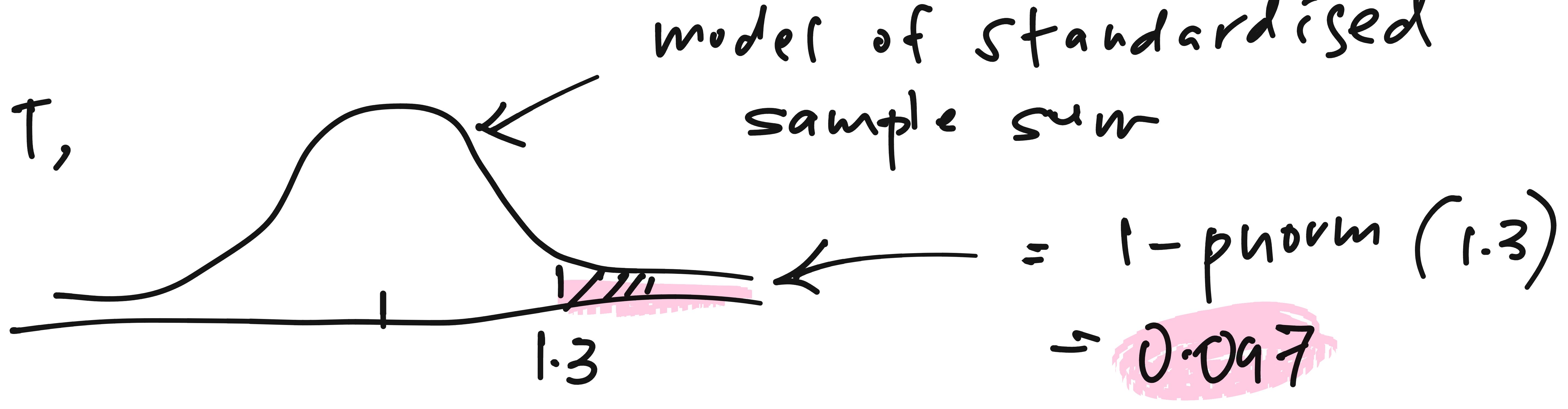
- The participants are independent of each other.
- $P(\text{desensitisation})$ = same for all participants.

T

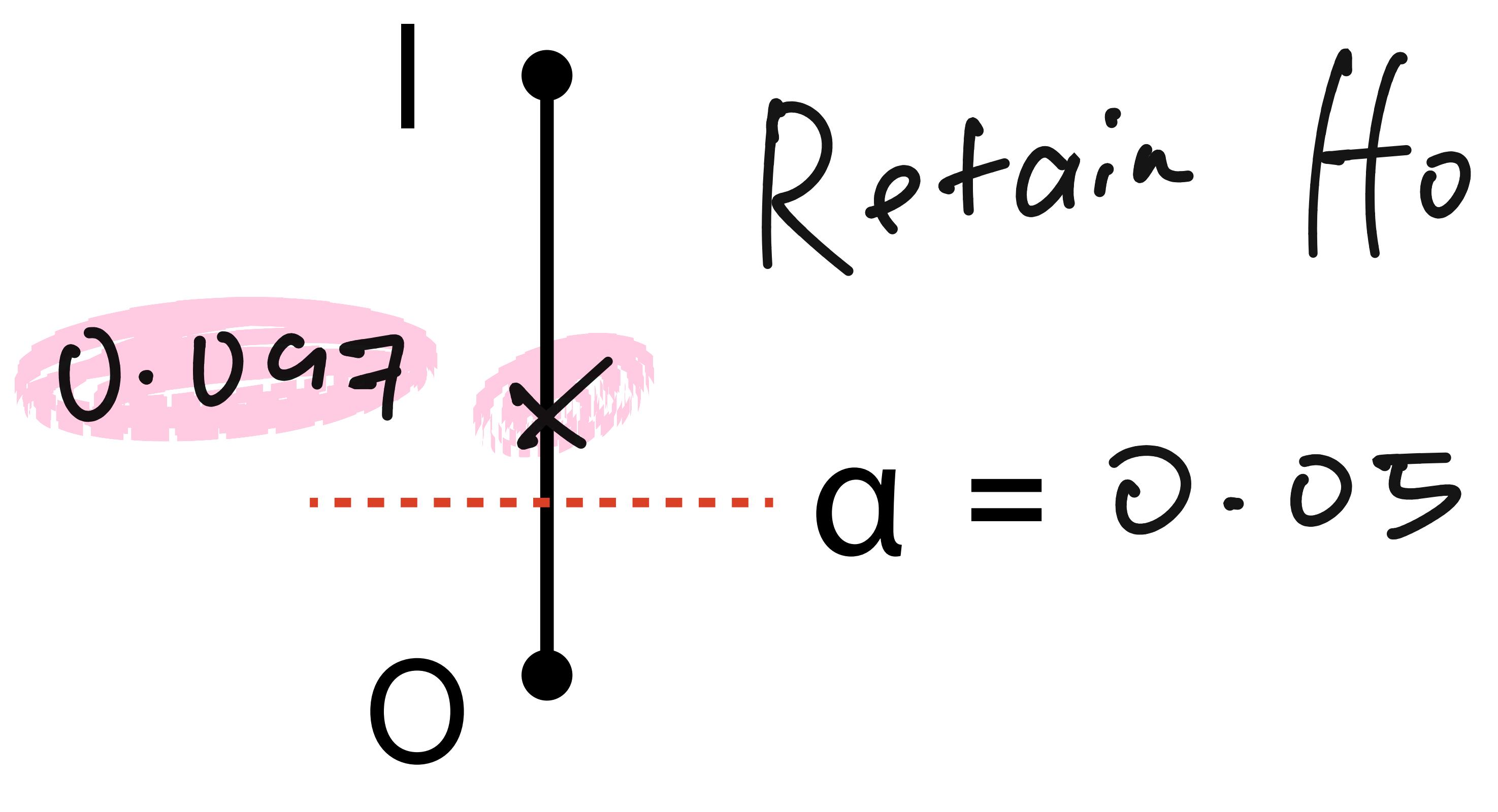
$$TS = \frac{OV - EV}{SE} = \frac{26 - 23.2}{2.2} = 1.3$$

P

Using CLT,

**C**

Data is consistent with a desensitisation rate of 0.8

**CI**

95% CI for sample sum:

$$\begin{aligned} OV &\pm 2 \cdot SE \\ 26 &\pm 2 \cdot 2.2 \\ (21.6, 30.4) &\quad \text{NB} \end{aligned}$$

How to simulate p-values

set.seed(1)

$$\text{box} = c(1, 1, 1, 1, 0) \quad \leftarrow \text{models } H_0$$

totals = replicate(100, sum(sample(box, 29, rep=TRUE)))

table(totals)

16	17	...	25	26	27	28
----	----	-----	----	----	----	----

1	2	16	6	4	2
---	---	----	---	---	---

p-value

$$= \frac{6 + 4 + 2}{100}$$

$$= \underline{\underline{0.12}}$$