What is Hypothesis Testing?

Worksheet

Hypothesis testing evaluates whether sample data provides enough evidence to reject a null hypothesis (H) in favor of an alternative hypothesis (H), based on a chosen significance level (commonly = 0.05).

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Questions

1. The null hypothesis (H) typically represents

A) a new effect
B) no effect or no difference
C) the researcher's desired outcome
D) a random guess

2. If the p-value < , you should

A) accept H
B) reject H
C) increase sample size
D) ignore the result

3. Rejecting a true H is called a

A) Type I error
B) Type II error
C) confidence interval
D) p-value error

4. A common significance level () used in research is

A) 0.5
B) 0.05
C) 5
D) 50

5. A company claims its light bulbs last = 1000 hours. A sample of n=36 bulbs has mean x=980 hours, =60 hours. Test at =0.05 (two-tailed).

6. A cereal box is labeled 500 g. A sample of n=25 boxes has mean x=495 g, =15 g. Test H: =500 at =0.05.

7. A new teaching method is tested: a sample of n=49 students scores x=78, historical =75, =14. Test at =0.05 one-tailed (H: >75).

8. Define: What is a null hypothesis (H)?

9. Define: What is an alternative hypothesis (H)?

10. Define: What is a p-value?

Answer Key

1. B) no effect or no difference - H is the default 'no effect' claim that testing tries to disprove.

2. B) reject H - A p-value below the significance level provides enough evidence to reject H.

3. A) Type I error - Type I error = false positive, rejecting a true null hypothesis.

4. B) 0.05 - = 0.05 (5%) is the most common threshold in practice.

5. z = (x)/(/n) = (980 1000)/(60/36) = 20/10 = 2.0 Critical z at =0.05 (two-tailed) = 1.96 |2.0| > 1.96 reject H: bulb lifespan is significantly different from 1000 hours

6. z = (495 500)/(15/25) = 5/3 1.67 Critical z (two-tailed, =0.05) = 1.96 |1.67| < 1.96 fail to reject H: not enough evidence boxes are underfilled

7. z = (78 75)/(14/49) = 3/2 = 1.5 Critical z (one-tailed, =0.05) = 1.645 1.5 < 1.645 fail to reject H: not enough evidence the method improves scores

8. The default claim of no effect or no difference, which testing tries to disprove.

9. The claim that there is an effect or difference, accepted only if H is rejected.

10. The probability of observing data as extreme as the sample, assuming H is true.