Packet 6: Hypothesis Testing

Learning Objects:

- Understand basic concepts and formulating the hypothesis.
- Learn how to conduct and interpret classical hypothesis testing.

Hypothesis Testing: Simply put, to answer yes/no questions; E.g., whether people eating saturated fat are more likely to develop heart disease. θ: proportion of people with heart disease

Statistical hypothesis (or hypothesis) is a statement about parameter(s) θ of a population.

Null v.s. Alternative: The claim or the research hypothesis we want to establish is called alternative hypothesis, H_1, opposite of which is called the null hypothesis, H_0.

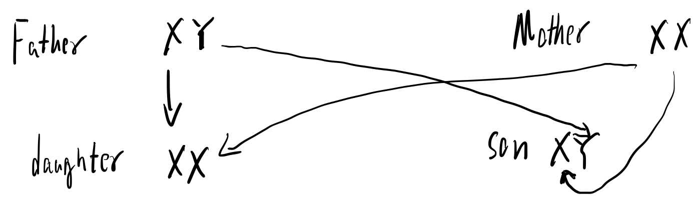
H_1 saturated fat $\rightarrow$ heart disease

H_0 saturated fat $\not\rightarrow$ heart disease

Rodolfo Gonzalez's hypothesis:

H_0: human intelligence genes are not only carried by X chromesome.

H_1: human intelligence genes are only carried by X chromesome.



Gonzalez' hypothesis has some interesting implications. If a male has some outstanding intellectual ability (associated with the X-chromosome) he is likely to be disappointed in the abilities of his sons because that ability can only be passed on to his daughters.

test whether Father IQ $\perp\!\!\!\perp$ son IQ

or whether female's $Var(IQ)$ < male's $Var(IQ)$

Decision:

Reject H_0 and conclude that H_1 is substantiated

Not reject H_0 We never say "accept H_0".

Type I and Type II Errors: Errors occur when decision is wrong

Type I error, H_0 is rejected when H_0 is true.

Type II error, H_0 is ~~accepted~~ not rejected when H_0 is false.

Court example: A suspect is not guilty v.s. A suspect is guilty.

H_0 v.s. H_1

Type I error: an innocent person is found guilty. (false rejection).

Type II error: a guilty person is found innocent. (false acceptance).

In the above criminal trail example, reject H_0 presumption of innocence.

legal right of accused

Presumption of Innocence One is considered innocent until proven guilty. v.s. No person shall be found guilty without being judged as such by a court.

Evidence is needed to reject H_0.

quantitative evidence → data

If there is no evidence to prove guilty, we do not reject H_0, but it does not mean we have proved H_0 is true.

In statistics, we denote

α=P(type I error)=P(reject H_0 | H_0 is true).

β=P(type II error)=P(fail to reject H_0 | H_0 is false). Power $= 1 - \beta$. of test

Fire alarm example:

truth: fire v.s. no fire

decision: alarm v.s. no alarm

H_0 no fire v.s. H_1 fire

which is H_0? yes: H_0: fire v.s. H_1 no fire

no: H_0: no fire v.s. H_1 fire

smoke detector will need a certain level of CO concentration to reject H_0: no fire

Increase sensitivity of the detector: more likely to ring the alarm

$\alpha = P(\text{alarm} \mid \text{no fire})$ false alarm ↑

$\beta = P(\text{no alarm} \mid \text{fire})$ ↓

never ring alarm $\alpha = 0$ $\beta = 1$

always alarm $\alpha = 1$ $\beta = 0$

There is a trade off between α and β. Given a hypothesis testing problem, we need to design a test such that α and β are balanced.

When type I error is more serious (like the criminal example), we design a test such that a preferred value of α is obtained (e.g., $\alpha = 0.05$). A good decision rule gives a small β.

Uniformly most powerful test (UMPT): a uniformly most powerful (UMP) test is a hypothesis test which has the greatest power, $1 - \beta$, among all possible tests of a given size α.