# 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.

**Statistical hypothesis** (or hypothesis) is a statement about parameter(s) $\theta$ 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, $H_0$.

**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.

**Decision:**

Reject $H_0$ and conclude that $H_1$ is substantiated

Not reject $H_0$

**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 when $H_0$ is false.

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

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

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

**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 $H_0$.

In statistics, we denote

$\alpha$=P(type I error)=P(reject $H_0 \mid H_0$ is true).

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

**Fire alarm example:**

Increase sensitivity of the detector:

There is a trade off between $\alpha$ and $\beta$. Given a hypothesis testing problem, we need to design a test such that $\alpha$ and $\beta$ are balanced.

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

**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 $\alpha$.