Testing of Hypothesis

A statistical hypothesis is an assertion about the probability distribution of a population.

Suppose we consider marks of students in two tests that have been conducted

Let $f(x)$ be the distⁿ. Then we want to test hypothesis

Null hypothesis $\leftarrow H_0$: $f(x)$ is $N(\mu, \sigma^2)$.

Alternative hyp $\leftarrow H_1$: $f(x)$ is not $N(\mu, \sigma^2)$

Another example: The average marks of students in PS class. Consider two groups. Group I: the students who regularly follow the class. Group II : the students who are irregular.

$\mu_1 \rightarrow$ average marks of the first group.

$\mu_2 \rightarrow$ average marks of the second group

Then we would like to test

$$H_0: \mu_1 = \mu_2 \rightarrow \text{Null hyp}$$

$$H_1: \mu_1 > \mu_2 \rightarrow \text{Alt. Hyp.}$$

Consider a drug being used for a certain disease. Let p be the proportion of patients who recover using the drug. Suppose it is known that $\boxed{p = 0.8}$ $\rightarrow$

A new drug is discovered and let P is the proportion of recovery using the new drug. Then we want to test ~~of~~

$H_0: P \leq 0.8$

$H_1: P > 0.8$

null hyp

alt. hyp

$$\boxed{\begin{aligned} H_0 &: P > 0.8 \\ H_1 &: P \leq 0.8 \end{aligned}}$$

Simple Hypothesis $\longrightarrow$ The prob. distn is completely specified.

Composite Hypothesis → The prob. distn is not completely specified.

Test of Statistical Hypothesis → is a procedure based on a random sample from the given population to check if the null or alternative are more likely under the sample.

So we reject H_0 on the basis of sample (accept H_1)

or we do not reject H_0 (accept H_0)

Acceptance and Rejection Regions.

Let S be the sample space of the random sample.

A rejection region (critical region)

of the test is that part of the sample space which corresponds to the rejection of the null hypothesis

$R \subset S$

$X \in R$

$X \in R^c \to$ acceptance region $R^c \subset S$

<u>Example</u>: A coin is tossed thrice independently and $p = P(\text{Head})$

$H_0: p = \frac{1}{4}$

$H_1: p = 3/4$

$S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$

We may choose

$A = R^c = \{HTT, THT, TTH, TTT\}$

$R = \{HHH, HHT, HTH, THH\}$

That is, if one or more heads are observed, we feel H_0 is true and if two or three heads are observed then we feel H_1 is true.

<u>Two Types of Errors</u>: We are likely to make two types of errors when we conduct a test of hypothesis:

Type I Error: Rejecting H_0 when it is tone

→ Error of the first kind

Type II Error: Accepting H_0 when it is false

→ Errr of the second kind

$$\alpha = P(\text{type I error})$$

$$\beta = P(\text{type II error})$$

In the above example, we can find

$$\alpha = P(\underline{X} \in R \mid H_0 \text{ is true})$$

$X \to$ number of heads

$X \sim \text{Bin}(3, p)$

$$\alpha = P_{p=1/4}(X \geq 2) = \binom{3}{2}\left(\frac{1}{4}\right)^2 \frac{3}{4} + \binom{3}{3}\left(\frac{1}{4}\right)^3$$

$$= \frac{10}{64} = 5/32$$

$$\beta = P\underset{p=3/4}{(}X \le 1) = \binom{3}{0}\left(\frac{1}{4}\right)^3 + \binom{3}{1}\frac{3}{4}\left(\frac{1}{4}\right)^2$$

$$= \frac{10}{64} = \frac{5}{32}.$$

$1 - \beta = \gamma =$ Power of the test

$= P($ Accepting H_0 when it is false$)$

Most Powerful Test: It is not possible to minimize both α & β

simultaneously. So generally we fix α (up to a certain level) and then consider test procedure which minimizes β (or maximizes $\gamma = 1-\beta$). This is called most powerful test (MP test).

Neyman & Pearson proved that

for simple null hyp. against simple alternative hypothesis the MP test exists and can be found.

Testing for Parameters of a Normal

Population: Let $X_1, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$ population.

Testing for mean μ.

Case I : σ^2 is known

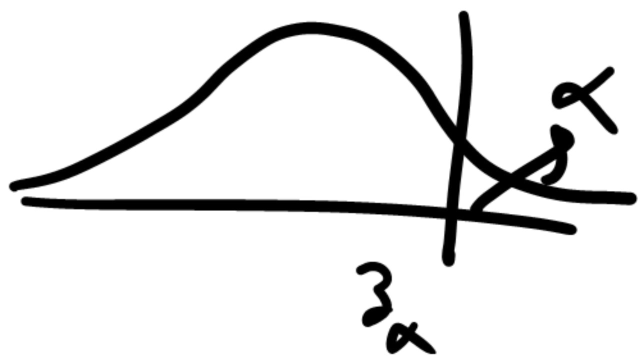
$$Z = \left.\frac{\sqrt{n}\,(\bar{X} - \mu_0)}{\sigma}\right\} \rightarrow \text{Test statistic}$$

(P1) $H_0 : \mu \leq \mu_0$ (fixed number)

$H_1 : \mu > \mu_0$

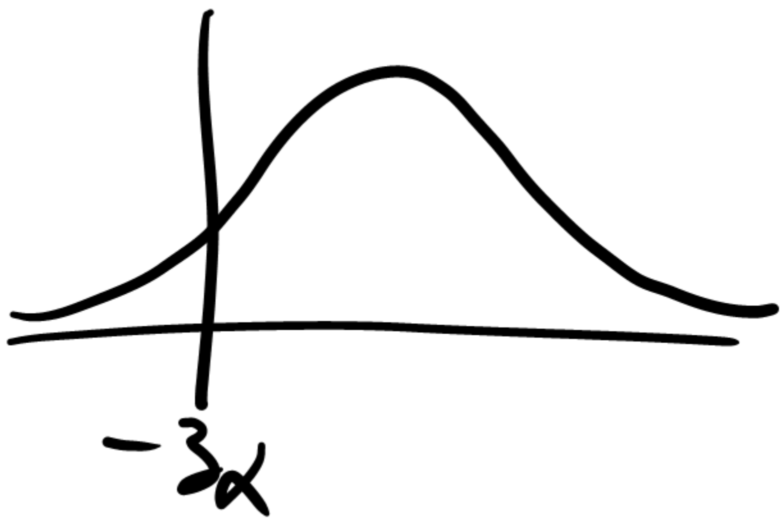
We reject H_0 if $Z > z_\alpha$

$z_\alpha \to$ upper $100(1-\alpha)\%$ point of standard normal distn



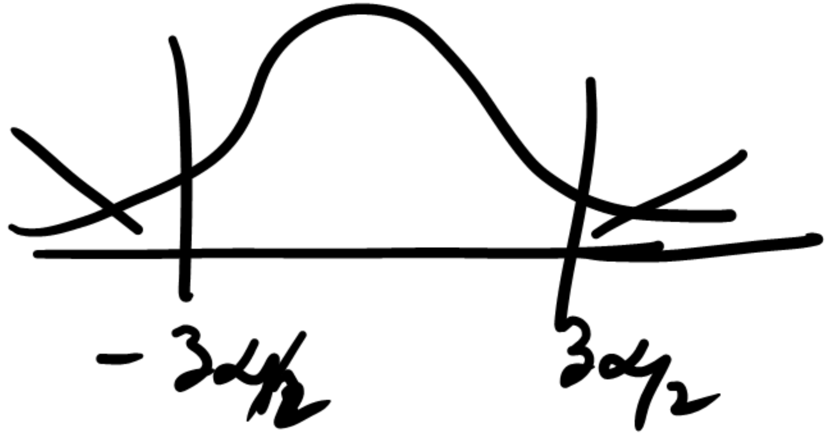
Accept H_0 otherwise.

P2
$$\left\{\begin{array}{l} H_0: \mu \geq \mu_0 \\ H_1: \mu < \mu_0 \end{array}\right\}$$



Reject H_0 if $Z < -z_\alpha$.

P3 $\begin{cases} H_0: & \mu = \mu_0 \\ H_1: & \mu \neq \mu_0 \end{cases}$



Rejecting H_0 if $|Z| \geq z_{\alpha/2}$

Case II: σ^2 unknown

$$S^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2$$

$$T = \frac{\sqrt{n}\,(\bar{X} - \mu_0)}{S}$$

For (P1) $\rightarrow$ Reject H_0 if $T > t_{\alpha, n-1}$

$100(1-\alpha)\%$ point of t_{n-1} distn.

For (P2) $\rightarrow$ Rejecting H_0 if $T < -t_{\alpha, n-1}$

For (P3) $\rightarrow$ Reject H_0 if $|T| > t_{\alpha/2, n-1}$

Example : Suppose 12 electronic items

are tested and their lifetimes are (in months) recorded as (36.1, 40.2, 33.8, 38.5, 42, 35.8, 37, 41, 36.8, 37.2, 33, 36)

$H_0: \mu \geq 40$ vs $H_1: \mu < 40$

$\bar{x} = 37.2833$, $s = 2.7319$

$$T = \frac{\sqrt{n}(\bar{x} - 40)}{s} = \frac{\sqrt{12}(37.2833 - 40)}{2.7319}$$

$= -3.44$, $t_{0.05, 11} = 1.796$

We want to see if $T \leq -t_{\alpha, n-1}$

$$-3.44 < -1.796$$

is true

So H_0 is rejected.

Testing for Variance

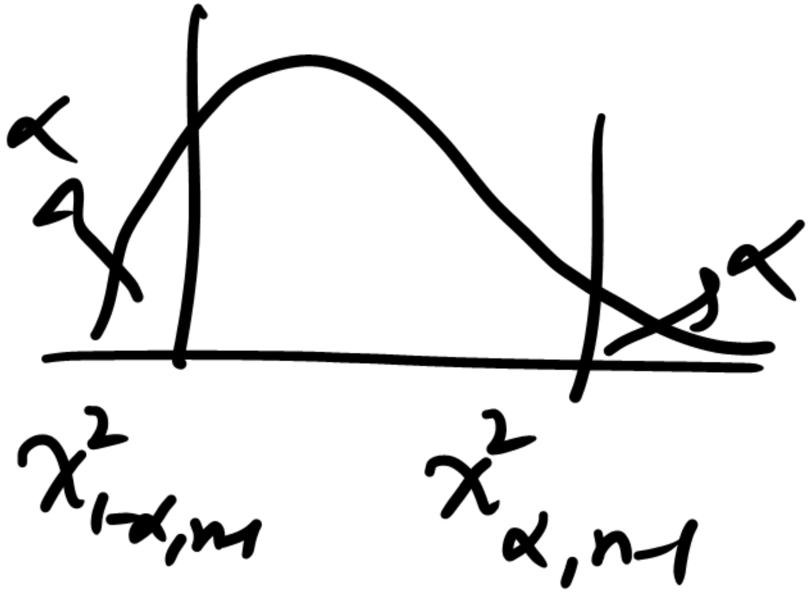
$$W = \frac{(n-1) S^2}{\sigma_0^2}$$

(P1) $H_0: \sigma^2 \le \sigma_0^2$, (P2): $H_0: \sigma^2 \ge \sigma_0^2$

$H_1: \sigma^2 > \sigma_1^2$ $\qquad$ $H_1: \sigma^2 < \sigma_0^2$

(P3) $H_0: \sigma^2 = \sigma_0^2$

$H_1: \sigma^2 \neq \sigma_0^2$



For (P1) : Reject H_0 if $W > \chi^2_{\alpha,n-1}$

for P2 : Reject H_0 if $W < \chi^2_{1-\alpha, n-1}$

for P3 : Reject H_0 if $W < \chi^2_{1-\frac{\alpha}{2}, n-1}$

or $W > \chi^2_{\alpha/2, n-1}$.

<u>Example</u>: A random sample of size 25 has $S^2 = 2$. We want to test

$H_0: \sigma^2 \leq 1$

$H_1: \sigma^2 > 1$

$$W = \frac{(n-1)S^2}{\sigma_0^2} = \frac{24 \times 2}{1} = 48$$

$$\chi^2_{0.05,24} = 37....$$

$W > \chi^2_{0.05,24}$. So H_0 is rejected

Testing for Parameters of Two Normal Populations :

$X_1, \ldots, X_m \sim N(\mu_1, \sigma_1^2)$

indept

$Y_1, \ldots, Y_n \sim N(\mu_2, \sigma_2^2)$

For comparing μ_1 & μ_2

<u>Case I</u>: σ_1^2 and σ_2^2 are known

P1 $H_0: \mu_1 \leq \mu_2$
$H_1: \mu_1 > \mu_2$

P2 $H_0: \mu_1 \geq \mu_2$
$H_1: \mu_1 < \mu_2$

P3 : $H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$.

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$$

For P1 : Reject H_0 if $Z > z_\alpha$

For P2 : Reject H_0 if $Z < -z_\alpha$

For (P3) : Reject H_0 if $|z| \geqslant z_{\alpha/2}$.

Case II : $\sigma_1^2 = \sigma_2^2 = \sigma^2$ unknown

$$T = \sqrt{\frac{mn}{m+n}} \; \frac{(\bar{X} - \bar{Y})}{S_p}$$

$$S_p^2 = \frac{(m-1)S_1^2 + (n-1)S_2^2}{(m+n-2)}$$

For P1 Reject H_0 if $T > t_{\alpha, m+n-2}$

For P2 Reject H_0 if $T < -t_{\alpha, m+n-2}$

For P3 Reject H_0 if $|T| \geq t_{\frac{\alpha}{2}, m+n-2}$

<u>Case III</u> : σ_1^2 & σ_2^2 are completely unknown

Welch - Smith - Satherwaite procedure

$$T^* = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}}$$

$$\nu = \frac{\left(\frac{S_1^2}{m} + \frac{S_2^2}{n}\right)^2}{\left[\frac{S_1^4}{m^2(m-1)} + \frac{S_2^4}{n^2(n-1)}\right]}$$

we take ν rounded to nearest integer.

(P1) Reject H_0 if $T^* > t_{\alpha,\nu}$

(P2) Reject H_0 if $T^* < -t_{\alpha,\nu}$

(P3) Reject H_0 if $|T^*| \geq t_{\frac{\alpha}{2},\nu}$.

Paired t- test: Here X_i & Y_i are not independent but they are paired.

$$(X_i, Y_i) \sim BVN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$$

$$d_i = X_i - Y_i, \quad d_i \sim N(\mu_1 - \mu_2, \sigma_d^2)$$

$$\bar{d} = \frac{1}{n}\sum d_i, \qquad S_d^2 = \frac{1}{n-1}\sum (d_i - \bar{d})^2$$

$$T_1 = \frac{\sqrt{n}\,\bar{d}}{S_d}$$

For P1 Reject H_0 if $T_1 > t_{\alpha, n-1}$

For P2 Reject H_0 if $T_1 < -t_{\alpha, n-1}$

For P3 Reject H_0 if $|T_1| \geq t_{\alpha/2, n-1}$

Testing for Variances

P1 $H_0: \sigma_1^2 \leq \sigma_2^2$

$H_1: \sigma_1^2 > \sigma_2^2$

P2 $H_0: \sigma_1^2 \geq \sigma_2^2$

$H_1: \sigma_1^2 < \sigma_2^2$

P3 $H_0: \sigma_1^2 = \sigma_2^2$ vs $H_1: \sigma_1^2 \neq \sigma_2^2$

$$F = \frac{S_1^2}{S_2^2}$$

For (P1) Reject H_0 if $F > f_{\alpha, m-1, n-1}$

(P2) Reject H_0 if $F < f_{1-\alpha, m-1, n-1}$

(P3) Reject H_0 if $F < f_{1-\frac{\alpha}{2}, m-1, n-1}$

or $f > f_{\alpha/2, m-1, n-1}$

Examples: Comparing average yields of a crop in two different states

State 1: $m = 10$ $\sigma_1^2 = 100$

$\bar{x} = 825$ mt

State 2: $n = 10$ $\sigma_2^2 = 60$

$\bar{y} = 815$. mt

$$Z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} = 2.5$$

$$z_{0.05} = 1.645$$

$Z > z_{0.05}$.

So for P1 : H_0 is rejected

2. weights of athletes in two teams

Team 1 : $m = 8$, $\bar{x} = 90$ (kg)

$s_1^2 = 3.9$

Team 2 : $n = 8$, $\bar{y} = 95$ (kg). $s_2^2 = 4$.

For testing $H_0: \sigma_1^2 = \sigma_2^2$

$H_1: \sigma_1^2 \neq \sigma_2^2$

$$F = \frac{s_2^2}{s_1^2} = 1.026, \qquad f_{0.05,7,7} = 3.78$$

$$f_{0.95,7,7} = 0.3746$$

So H_0 cannot be rejected

ie the two variances are equal.

So for testing for equality of means

we can use pooled sample variance procedure.

$$T^* = \sqrt{\frac{mn}{m+n}}\ \frac{\overline{x}-\overline{y}}{S_p} = \sqrt{\frac{64}{16}}\ \frac{(-5)}{\sqrt{3.95}}$$

$$S_p^2 = \frac{7\times 3.9 + 7\times 4}{14} = 3.95$$

$$T^* \simeq -5.032,$$

$$t_{0.025,14} = 2.145, \quad , t_{0.05,14} = 1.761$$

For P2 $\quad T^{\alpha} < -t_{0.05,\,14}$

So H_0 is rejected

ie average weights of atheletes in Team 1 are less than average weights of atheletes in Team 2.

Example : Lives of some electronic items. (in hours)

Brand 1 : $m = 25$, $\bar{x} = 380$, $s_1^2 = 100$

Brand 2 : $n = 16$, $\bar{y} = 370$, $s_2^2 = 400$

$H_0: \sigma_1^2 \leq \sigma_2^2$

$H_1: \sigma_1^2 > \sigma_2^2$

$$F = \frac{s_2^2}{s_1^2} = 4$$

$$f_{0.05, 15, 24} = 2.1077$$

So H_0 is rejected.

We cannot pool sample variances

here as variances are not the same.

$$T_1 = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}} \cong 1.857$$

$$\nu = \frac{\left(S_1^2/m + S_2^2/n\right)^2}{\left[S_1^4/m^2(m-1) + S_2^4/n^2(n-1)\right]} = 19.86$$

$t_{0.05,19} = 1.729, \quad t_{0.025,19} = 2.093$

For P_3, H_0 cannot be rejected at 5% level of significance.

So the lives of electronic items in two pop^ns are similar.

Example (Paired-t-test)

15 persons underwent a diet program for 3 months and the weights x_i (before), y_i (after) are as follows.

Person	1	2	3	4	5	6	7
x_i	70	80	72	76	76	76	72
y_i	68	72	62	70	58	66	68
$d_i = x_i - y_i$	2	8	10	6	18	10	4

Person	8	9	10	11	12	13	14	15
x_i	78	82	64	74	92	74	68	84
y_i	52	64	72	74	60	74	72	74
d_i	26	18	-8	0	32	0	-4	10

$\bar{d} = 8.8$, $s_d = 10.98$, $n = 15$

$\dfrac{\sqrt{n}\,\bar{d}}{s_d} = 3.1$, $t_{0.025,14} = 2.145$

So at $\alpha = 0.025$, $H_0 : \mu_1 \leq \mu_2$

$H_1 : \mu_1 > \mu_2$

H_0 is rejected. So the training program is effective.

Law of Large Numbers (Weak)

Let $X_1, X_2, \ldots$ be independent and identically distributed random variables with mean μ and variance σ^2.

Let $Y_n = \frac{1}{n} \sum_{i=1}^{n} X_i$

Then for any $k > 0$,

$$P(|Y_n - \mu| > k) \longrightarrow 0$$

as $n \to \infty$

or $P(|Y_n - \mu| \leq k) \longrightarrow 1$ as $n \to \infty$

Or $\quad Y_n \xrightarrow{P} \mu$

Example: Let $X_1 \dots X_n, \dots \overset{i.i.d.}{\sim} P(\lambda)$

If $\lambda = 3$

$$Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} 3$$

2. Let $X_1, \dots X_2, \dots \quad N(2, 4)$.

$$E(X_i^2) = \mu^2 + \sigma^2 = 4 + 4 = 8$$

$$\frac{\sum X_i^2}{n} \xrightarrow{P} E(X_i^2) = 8$$

as $n \to \infty$