Testing Hypothesis

Statistical Hypothesis —— A statement about population

Null hypothesis : H_0 Alternative hypothesis : H_1

$\begin{cases} H_0 : \mu = 0 \\ H_1 : \mu \neq 0 \end{cases}$ $\begin{cases} H_0 : \sigma^2 = 1 \\ H_1 : \sigma^2 \neq 1 \end{cases}$

$\begin{cases} H_0 : \sigma_x^2 = \sigma_y^2 \\ H_1 : \sigma_x^2 \neq \sigma_y^2 \end{cases}$ $\begin{cases} H_0 : \pi_1 = 2\pi_2 \\ H_1 : \pi_1 \neq 2\pi_2 \end{cases}$

$\begin{cases} H_0 : \text{Population is a normal distribution.} \\ H_1 : \text{Population is not a normal distribution.} \end{cases}$

Testing Hypothesis

Test —— Procedure (based on sample) leads to rejection or non-rejection of the hypothesis

Example : Test $H_0 : \mu = 1$ vs $H_1 : \mu \neq 1$.

"Reject H_0 if $\bar{X} > 2$ or $\bar{X} < 0$." is a reasonable test.

Draw a sample ⟶ $\bar{X} = 2.3$ ⟶ Reject H_0.

⟶ Conclude that $\mu \neq 1$.

But we may be wrong !

Two Types of Errors

	H_0 True	H_1 True
Do not reject H_0	Correct	Type II error
Reject H_0	Type I error	Correct

$\alpha = \Pr(\text{Type I error}) = \Pr(\text{Reject } H_0 \mid H_0 \text{ true})$

$\beta = \Pr(\text{Type II error}) = \Pr(\text{Do not reject } H_0 \mid H_1 \text{ true})$

Two Types of Errors

Example : life time of light bulbs with $\sigma = 300$ hours

$H_0 : \mu = 1200$ vs $H_1 : \mu = 1240$

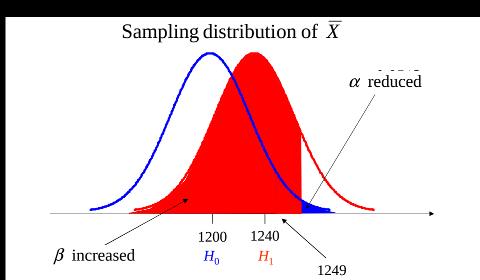
Draw a sample with size 100. $\bar{X} \sim N\left(\mu, \frac{300^2}{100}\right)$

Test —— "Reject H_0 if $\bar{X} > 1249$."

$\alpha = \Pr(\bar{X} > 1249 \mid \mu = 1200) = 0.0513 \left(\frac{1249 - 1200}{300/\sqrt{100}} \right)$

$\beta = \Pr(\bar{X} \leq 1249 \mid \mu = 1240) = 0.6179 \left(\frac{1249 - 1240}{300/\sqrt{100}} \right)$

Two Types of Errors



Choice of H_0 and H_1

α β

For fix n , we can only control one of the errors.

Convention : Control α .

Criterion 1 : Make Type I error a more serious error.

Example : Want to know if a man is guilty.

H_0 : He is not guilty. Type I error Jail an innocence. Type II error Release a criminal.

H_1 : He is guilty.

More serious error

Choice of H_0 and H_1

Convention : Control α to be small.

Pr (false rejection of H_0) = Pr (reject H_0 | H_0 true) = α

Pr (false acceptance of H_0) = Pr (accept H_0 | H_1 true) = β ← May be large

Reject H_0 ← Strong conclusion

Do not reject H_0 ← Weak conclusion

Criterion 2 : Put what you want to prove in H_1.

Choice of H_0 and H_1

Criterion 2 : Put what you want to prove in H_1.

Example : Want to show that Brand A is more popular than Brand B.

H_0 : B is more popular than A.
H_1 : A is more popular than B.

If reject H_0 → Data shows strong evidence to against H_0, i.e., high confidence that H_1 is true.

If do not reject H_0 → Not enough evidence to against H_0, but may not have high confidence on H_0.

Choice of H_0 and H_1

Example : Clinical trials on drug design.

Standard medicine : cure rate $\pi_0 = 0.6$
New medicine : believe that cure rate $\pi > 0.6$

$H_0 : \pi \le 0.6$ (new medicine is not better)

$H_1 : \pi > 0.6$ (new medicine is better)

Criterion 1 : Type I error (switch to worse medicine) ← more serious
Type II error (abandon better medicine)

Criterion 2 : What we want to prove is H_1.

Construction of Test Procedure

Example : life time of light bulbs with $\sigma = 300$ hours

$H_0 : \mu = 1200$ vs $H_1 : \mu = 1240$ Preset α = 0.05.

Draw a sample with size 100. $\bar{X} \sim N\left(\mu, \frac{300^2}{100}\right)$

Reasonable test — "Reject H_0 if $\bar{X} > c$."

$$0.05 = \alpha = 1 - \Phi\left(\frac{c-1200}{300/\sqrt{100}}\right) \Rightarrow \frac{c-1200}{30} = z_{0.05} = 1.645 \quad c = 1249.35$$

"Reject H_0 if $\bar{X} > 1249.35$." ← Constructed before observing data.

Construction of Test Procedure

Example : life time of light bulbs with $\sigma = 300$ hours

$H_0 : \mu = 1200$ vs $H_1 : \mu = 1240$ Preset α = 0.05.

Draw a sample with size 100. $\bar{X} \sim N\left(\mu, \frac{300^2}{100}\right)$

"Reject H_0 if $\bar{X} > 1249.35$."

$$\beta = \Phi\left(\frac{1249.35-1240}{300/\sqrt{100}}\right) = 0.6225$$

$\Pr(\text{Reject } H_0 \mid H_1 \text{ true}) = 1 - \beta = 0.3775$ ← Power of the test.

Construction of Test Procedure

Example : life time of light bulbs with $\sigma = 300$ hours

$H_0 : \mu = 1200$ vs $H_1 : \mu = 1240$ Preset α = 0.05.

Draw a sample with size 400. $\bar{X} \sim N\left(\mu, \frac{300^2}{400}\right)$

"Reject H_0 if $\bar{X} > c$."

$$0.05 = \alpha = 1 - \Phi\left(\frac{c-1200}{300/\sqrt{400}}\right) \Rightarrow \frac{c-1200}{15} = z_{0.05} = 1.645 \quad c = 1224.675$$

$$\beta = \Phi\left(\frac{1224.675-1240}{300/\sqrt{400}}\right) = 0.1534$$

Power = 0.8466

Construction of Test Procedure

Example : life time of light bulbs with $\sigma = 300$ hours

$H_0 : \mu = 1200 \quad \text{vs} \quad H_1 : \mu = 1240$

Preset α = 0.05.

Preset power = 0.9 , i.e. β = 0.1

$$\frac{c - 1200}{300/\sqrt{n}} = 1.645$$

$$\frac{c - 1240}{300/\sqrt{n}} = -1.282$$

$\rightarrow$ $n = 482 \ , \ c = 1222.48$

Significance Probability

Example : life time of light bulbs with $\sigma = 300$ hours

$H_0 : \mu = 1200 \quad \text{vs} \quad H_1 : \mu = 1240$

$n = 482$

"Reject H_0 if $\bar{X} > 1222.48$."

$\bar{X} = 1223$ → Reject H_0 .

Stronger evidence → $\bar{X} = 1237$ → Reject H_0 .

Data strongly disagree with H_0 .

Measure of agreement between data and H_0

significance probability / p-value

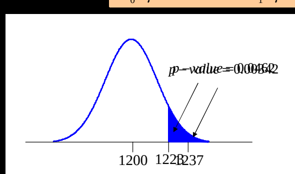
Significance Probability

significance probability of an observation

—— smallest α for which it leads to the rejection of H_0

Example : life time of light bulbs with $\sigma = 300$ hours

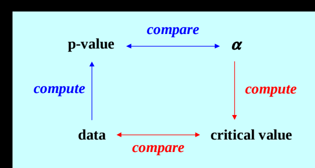
$H_0 : \mu = 1200 \quad \text{vs} \quad H_1 : \mu = 1240$ $\quad n = 482$



$p - value \text{ of } (\bar{X} = 1237)$

$$= 1 - \Phi\left(\frac{1237 - 1200}{300/\sqrt{482}}\right) = 0.00342$$

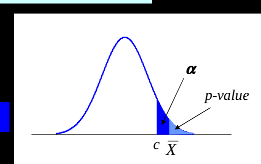
Significance Probability



$\bar{X} > c \Leftrightarrow p - value < \alpha$

Reject H_0 if data in region

Reject H_0 ⇔ p-value < α

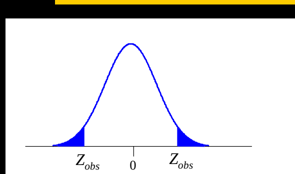


Large Sample, Known Variance

$H_0 : \mu = \mu_0 \quad \text{vs} \quad H_1 : \mu \neq \mu_0$ $\quad \sigma^2$ known

Test statistic $Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$

Reject H_0 at significance level α if $|Z_{obs}| > Z_{\alpha/2}$.



If $Z_{obs} < 0$,

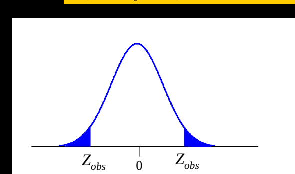
$p\text{-value} = 2\Phi(Z_{obs})$

Large Sample, Known Variance

$H_0 : \mu \geq \mu_0 \quad \text{vs} \quad H_1 : \mu < \mu_0$ $\quad \sigma^2$ known

Test statistic $Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$

Reject H_0 at significance level α if $Z_{obs} < -Z_{\alpha}$.



$p\text{-value} = \Phi(Z_{obs})$

Large Sample, Known Variance

Example : package delivery time with known $\sigma = 5$ mins

$H_0 : \mu \le 28$ vs $H_1 : \mu > 28$ $\quad \alpha = 0.05$

Sample with $n = 100$ $\quad \overline{X} = 31.5$, $S^2 = 28.9$

Test statistic $Z = \dfrac{\overline{X} - 28}{5/\sqrt{100}}$

Reject H_0 at significance level $\alpha = 0.05$ if $Z_{obs} > Z_{0.05} = 1.645$.

$Z_{obs} = \dfrac{31.5 - 28}{5/\sqrt{100}} = 7 > 1.645$ $\quad$ Reject H_0 at $\alpha = 0.05$.

$p - value = 1 - \Phi(7) \approx 0$

Small Sample, Normal Population, Unknown Variance

Normal population $\quad \sigma^2$ unknown

$H_0 : \mu \ge \mu_0$ vs $H_1 : \mu < \mu_0$

Test statistic $T = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$

Reject H_0 at significance level α if $T_{obs} < -t_{n-1,\alpha}$.

Small Sample, Unknown Variance

Example : package delivery time $\quad X \sim N(\mu, \sigma^2)$

$H_0 : \mu \le 28$ vs $H_1 : \mu > 28$ $\quad \alpha = 0.05$

Sample with $n = 10$ $\quad \overline{X} = 32.3$, $S^2 = 46.3$

Test statistic $T = \dfrac{\overline{X} - 28}{S/\sqrt{10}}$

Reject H_0 at significance level $\alpha = 0.05$ if $T_{obs} > t_{9,0.05} = 1.833$.

$T_{obs} = \dfrac{32.3 - 28}{\sqrt{46.3/10}} = 1.998 > 1.833$ $\quad$ Reject H_0 at $\alpha = 0.05$.

$p - value = \Pr(t_9 > 1.998) = 0.0384$

Two Samples Problem (Large Samples)

μ_x, σ_x^2 → Sample with size m

μ_y, σ_y^2 → Sample with size n

Independent

$H_0 : \mu_x \ge \mu_y$ vs $H_1 : \mu_x < \mu_y$

Test statistic $Z = \dfrac{\overline{X} - \overline{Y}}{\sqrt{S_x^2/m + S_y^2/n}}$

Reject H_0 at significance level α if $Z_{obs} < -Z_\alpha$.

Two Samples Problem (Small Samples)

$N(\mu_x, \sigma_x^2)$ → Sample with size m

$\sigma_x^2 = \sigma_y^2$

$N(\mu_y, \sigma_y^2)$ → Sample with size n

Independent

$H_0 : \mu_x \ge \mu_y$ vs $H_1 : \mu_x < \mu_y$

Test statistic $T = \dfrac{\overline{X} - \overline{Y}}{S_{pool}\sqrt{1/m + 1/n}}$

Reject H_0 at significance level α if $T_{obs} < -t_{m+n-2,\alpha}$.

Two Samples Problem (Small Samples)

Example : 4 scores from class A : 64, 66, 89, 77
3 scores from class B : 56, 71, 53

$H_0 : \mu_x = \mu_y$ vs $H_1 : \mu_x \ne \mu_y$ $\quad \alpha = 0.05$

Assumptions :
(i) Normal populations. (ii) Independent samples. (iii) Equal variances.

$T = \dfrac{\overline{X} - \overline{Y}}{S_{pool}\sqrt{1/4 + 1/3}}$ $\quad$ Reject H_0 at $\alpha = 0.05$ if $|T_{obs}| > t_{5,0.025} = 2.5706$.

$\overline{X} = 74, S_x^2 = 132.667, \overline{Y} = 60, S_y^2 = 93$ $\quad S_{pool}^2 = \dfrac{(3)(132.667) + (2)(93)}{3+2} = 117$

$|T_{obs}| = \left| \dfrac{74 - 60}{\sqrt{(117)(1/4 + 1/3)}} \right| = 1.695 < 2.5706$ $\quad$ Do not reject H_0 at $\alpha = 0.05$.

Two Samples Problem (Small Samples)

$N(\mu_x, \sigma_x^2)$ → Sample with size m

$\sigma_x^2 \neq \sigma_y^2$ **Independent**

$N(\mu_y, \sigma_y^2)$ → Sample with size n

$H_0: \mu_x = \mu_y$ vs $H_1: \mu_x \neq \mu_y$

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{S_x^2/m + S_y^2/n}} \qquad v = \frac{\left(S_x^2/m + S_y^2/n\right)^2}{S_x^4/\left(m^2(m-1)\right) + S_y^4/\left(n^2(n-1)\right)}$$

Reject H_0 at significance level α if $|T_{obs}| > t_{v,\alpha/2}$.

Paired Data

Example : effect of stimulus on blood pressure

Man	1	2	3	4	5	6	7	8	9	10	11	12
Before (X)	120	124	130	118	140	128	140	135	126	130	126	127
After (Y)	128	131	131	127	132	125	141	137	118	132	129	135
D = Y - X	8	7	1	9	-8	-3	1	2	-8	2	3	8

Test $H_0: \mu_D = 0$ vs $H_1: \mu_D \neq 0$ at $\alpha = 0.05$

X and Y are dependent → ~~Two sample T-Test~~

$D = Y - X$ $\qquad \mu_D = E(D) = \mu_y - \mu_x$

Paired Data

Example : effect of stimulus on blood pressure

Man	1	2	3	4	5	6	7	8	9	10	11	12
D = Y - X	8	7	1	9	-8	-3	1	2	-8	2	3	8

Test $H_0: \mu_D = 0$ vs $H_1: \mu_D \neq 0$ at $\alpha = 0.05$

Test statistic $T = \frac{\bar{D} - 0}{S_D/\sqrt{12}}$ $\qquad$ Reject H_0 if $|T_{obs}| > t_{11,0.025} = 1.796$

$\bar{D} = 1.833,\ S_D = 5.83$ $\qquad |T_{obs}| = \frac{1.833}{5.83/\sqrt{12}} = 1.09 < 1.796$

Do not reject H_0 at $\alpha = 0.05$.

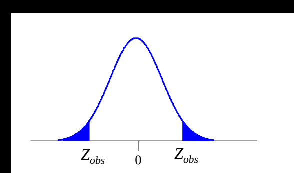
95% C.I. for $\mu_D = \mu_y - \mu_x$: $\quad 1.833 \pm 1.796 \frac{5.83}{\sqrt{12}} = 1.833 \pm 3.023$

Population Proportion

$H_0: \pi = \pi_0$ vs $H_1: \pi \neq \pi_0$

Test statistic $Z = \frac{\hat{\pi} - \pi_0}{\sqrt{\hat{\pi}(1-\hat{\pi})/n}}$ $\qquad \hat{\pi} = \frac{X}{n}$

Reject H_0 at significance level α if $|Z_{obs}| > Z_{\alpha/2}$.



If $Z_{obs} < 0$,

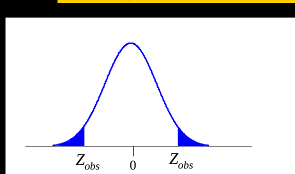
p- $p\text{-value} = 2\Phi(Z_{obs})$)

Population Proportion

$H_0: \pi \geq \pi_0$ vs $H_1: \pi < \pi_0$

Test statistic $Z = \frac{\hat{\pi} - \pi_0}{\sqrt{\hat{\pi}(1-\hat{\pi})/n}}$ $\qquad \hat{\pi} = \frac{X}{n}$

Reject H_0 at significance level α if $Z_{obs} < -Z_{\alpha}$.



$p\text{-value} = \Phi(Z_{obs})$

Population Proportion

Example : Crosses of peas $\qquad \pi = \Pr(\text{Yellow pea})$

Experiment : 176 Yellow, 48 Green

$H_0: \pi = 0.75$ vs $H_1: \pi \neq 0.75$ $\qquad \alpha = 0.05$

Test statistic $Z = \frac{\hat{\pi} - 0.75}{\sqrt{\hat{\pi}(1-\hat{\pi})/224}}$ $\qquad \hat{\pi} = \frac{176}{224} = 0.7857$

$$|Z_{obs}| = \frac{0.7857 - 0.75}{\sqrt{(0.7857)(0.2143)/224}} = 1.3021 < Z_{0.025} = 1.96$$

Do not reject H_0 at $\alpha = 0.05$.

$p - value = 2(1 - \Phi(1.0321)) = 0.1928$

Population Proportions

Independent binomial random variables

$X \sim b(n_1, \pi_1)$ $\quad$ $Y \sim b(n_2, \pi_2)$

Test $H_0 : \pi_1 \geq \pi_2$ vs $H_1 : \pi_1 < \pi_2$.

Test statistic $$Z = \frac{\hat{\pi}_1 - \hat{\pi}_2}{\sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_2}}}$$

$\hat{\pi}_1 = \frac{X}{n_1}$ $\quad$ $\hat{\pi}_2 = \frac{Y}{n_2}$ $\quad$ $\hat{\pi} = \frac{X+Y}{n_1+n_2}$

Reject H_0 at significance level α if $Z_{obs} < -Z_\alpha$.

Population Proportions

Example : Graduation rate

	ISP	Non-ISP
Graduated	189	158
Not Graduated	9	52
Total	198	210

$\hat{\pi}_{ISP} = \frac{189}{198} = 0.9545$

$\hat{\pi}_N = \frac{158}{210} = 0.7524$

$H_0 : \pi_{ISP} \leq \pi_N$ vs $H_1 : \pi_{ISP} > \pi_N$ $\quad$ $\alpha = 0.05$

$\hat{\pi} = \frac{189+158}{198+210} = 0.8505$

Reject H_0 at $\alpha = 0.05$.

$$Z_{obs} = \frac{0.9545 - 0.7524}{\sqrt{(0.9545)(0.0455)/198 + (0.7524)(0.2476)/210}} = 6.0757 > Z_{0.05} = 1.645$$