Lecture 12: 12/16/2020

One-sample test (Z test, T test)

Overview for the rest of semester

- One sample: $X_i \sim^{iid} E(X_i)=\mu, Var(X_i)=\sigma^2 \quad (i=1,\dots,n)$

    * Testing $\mu=\mu_0$

- Paired sample: (X_i, Y_i), $E(X_i)=\mu_1, E(Y_i)=\mu_2 \quad (i=1,\dots,n)$

    * Testing $\mu_1=\mu_2$?

- Two samples: $X_{1i} \sim^{iid} E(X_{1i})=\mu_1, Var(X_i)=\sigma_1^2 \quad (i=1,\dots,n_1)$

$$X_{2j} \sim^{iid} E(X_{2j})=\mu_2, Var(X_j)=\sigma_2^2 \quad (j=1,\dots,n_2)$$

    * Testing $\mu_1=\mu_2$?

Remark: all the procedures are studied under two scenarios with σ^2 (or σ_1^2 and σ_2^2) being "known" and "unknown".

假設檢定的兩種錯誤 (Two types of errors in testing)

		Truth	
		H_0 **is true**	H_0 **is false**
Decision	**Accept** H_0	OK	Type II error β
	Reject H_0	Type I error α	OK

α = Pr(reject $H_0 | H_0$ is true) → type I error rate =型一錯誤機率

β = Pr(accept $H_0 | H_0$ is false) → type I error rate =型二錯誤機率

Decision Rule:

- fix α = Pr(reject $H_0 | H_0$ is true) = level of significance (顯著水準)

Convention: fix $\alpha = 0.05$

*** One Sample Testing Problem**

Case 1: $H_0 : \mu = \mu_0$ versus $H_a : \mu > \mu_0$ (單邊檢驗, one-sided test)

Case 2: $H_0 : \mu = \mu_0$ versus $H_a : \mu < \mu_0$ (單邊檢驗, one-sided test)

Case 3: $H_0 : \mu = \mu_0$ versus $H_a : \mu \neq \mu_0$ (雙邊檢驗, two-sided test)

Steps:

- Set <u>the level of significance</u> (顯著水準)

  α = Pr(reject $H_0 | H_0$ is true) (usually $\alpha = 0.05$)

- Perform the testing procedure
  - Make one of the following decisions:
    - Reject H_0
    - Fail to reject H_0 → The evidence is NOT statistically significant (證據未達統計顯著水準)
  - We will introduce three methods to reach a decision.
    - Method 1: Check whether $\bar{x}_{obs}$ is located in the rejection region
    - Method 2: Check whether the p-value is smaller than α
    - Method 3: Check whether μ_0 is located in the $(1-\alpha)100\%$ confidence interval of μ (only suitable for Case 3)

*** Distribution Theory for the Normal Population**

- When σ^2 is known → $\dfrac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \sim^{H_0 \text{ is true}} N(0,1)$
- When σ^2 is unknown → $\dfrac{\bar{X} - \mu_0}{s / \sqrt{n}} \sim^{H_0 \text{ is true}} T_{n-1}$

Three methods of making the decision:

Based on the observed sample mean $\bar{x}_{obs}$, make the decision (reject H_0 or not)

Method 1: apply the decision rule → Lecture 11

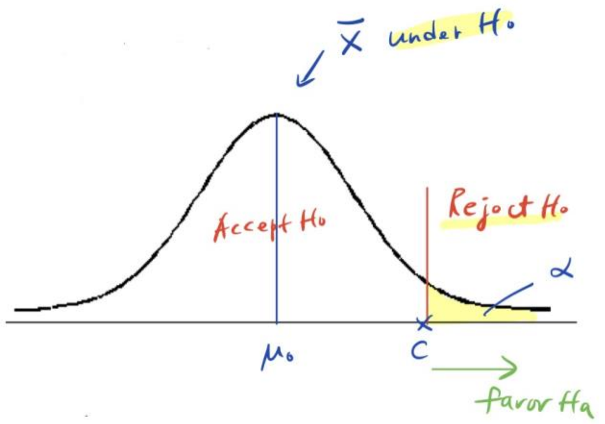
Method 2: use p-values

Method 3: use confidence intervals ← only for two-sided hypothesis

Situation: When $X_i \sim^{iid} N(\mu, \sigma^2)$ **and the value of** σ^2 **is known**

Case 1: $H_0 : \mu = \mu_0$ *versus* $H_a : \mu > \mu_0$

Method 1: Reject H_0 if $\bar{x}_{obs} > \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}$



Method 2: Reject H_0 if the p-value < α

p-value = $\Pr(\bar{X} > \bar{x}_{obs} \mid \mu = \mu_0)$

= the tail probability from $\bar{x}_{obs}$ to the tail in favor of H_a

$$= \Pr\left(\frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} > \frac{\bar{x}_{obs} - \mu_0}{\sigma / \sqrt{n}} \mid H_0\right)$$

$$= \Pr\left(Z > \frac{\bar{x}_{obs} - \mu_0}{\sigma / \sqrt{n}}\right)$$

$$= \Pr(Z > z_{obs})$$

H₀: μ=3.2 Ha: μ>3.2

14.8 **Arsenic contamination.** Arsenic is a compound naturally occurring in very low concentrations. Arsenic blood concentrations in healthy individuals are Normally distributed with mean $\mu = 3.2$ micrograms per deciliter (μg/dl) and standard deviation $\sigma = 1.5$ μg/dl.[5] Some areas are known to have naturally elevated concentrations of arsenic in the ground and water supplies. We take two SRSs of 25 adults residing in two different high-arsenic areas.

(a) We seek evidence *against* the claim that $\mu = 3.2$. What is the sampling distribution of the mean blood arsenic concentration $\bar{x}$ of a sample of 25 adults if the claim is true? Sketch the density curve of this distribution. (Sketch a Normal curve, and then mark the axis using what you know about locating the mean and standard deviation on a Normal curve.)

(b) Suppose that the data from the first sample give $\bar{x} = 3.35$. Mark this point on the axis of your sketch. Suppose that the data from the second sample give $\bar{x} = 3.75$ μg/dl. Mark this point on your sketch. Using your sketch, explain in simple language why one result is good evidence that the mean blood arsenic concentration of all adults in one high-arsenic area is greater than 3.2 μg/dl and why the outcome for the other high-arsenic area is not.

Example: Arsenic concentration

$H_0: \mu = 3.2$ *versus* $H_a: \mu > 3.2$

$\sigma = 1.5$, $n = 25$

Set $\alpha = 0.05 \Rightarrow z_\alpha = 1.645$

$$\frac{X - \mu_0}{\sigma/\sqrt{n}} = \frac{X - 3.2}{1.5/\sqrt{25}} \sim^{H_0 \text{ is true}} N(0,1)$$

<u>Method 1: Apply the rejection rule</u> (Lecture 11)

Reject H_0 if $\bar{x}_{obs} > 3.2 + 1.645\frac{1.5}{\sqrt{25}} \Leftrightarrow \bar{x}_{obs} > 3.6935$

When $\bar{x}_{obs}$ =3.35 → Fail to reject H_0

When $\bar{x}_{obs}$ =3.75 → Reject H_0

<u>Method 2: Use the p-value</u> (New & Important!!)

p-value = the probability of the tail area in favor of $H_a: \mu > 3.2$

When $\bar{x}_{obs}$ =3.35 → p-value = $\Pr(\bar{X} > \bar{x}_{obs} \mid H_0) = \Pr(\bar{X} > 3.35 \mid \mu = 3.2)$

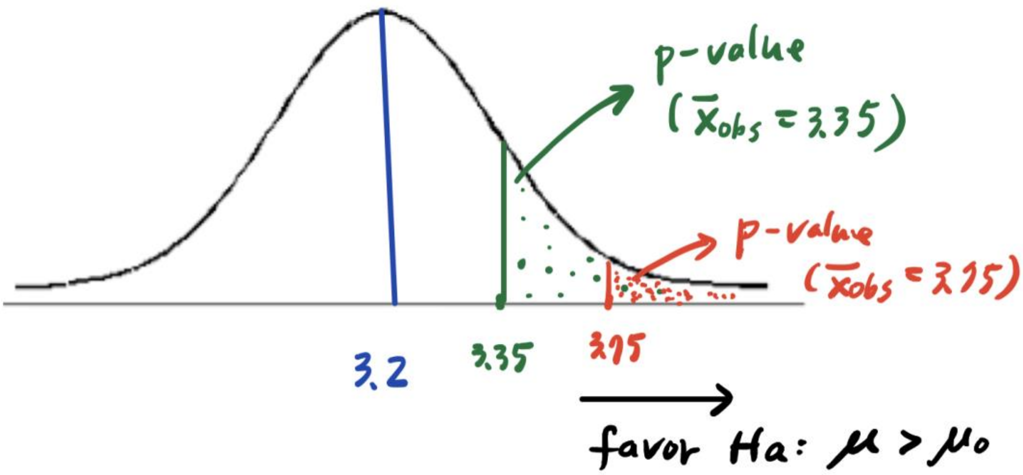
$$= \Pr\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > \frac{3.35 - 3.2}{1.5/\sqrt{25}}\right) = \Pr(Z > 0.5) = 0.3085$$

Since p-value = 0.3085 > 0.05 $= \alpha$ → Fail to reject H_0

When $\bar{x}_{obs}$ =3.75 → p-value = $\Pr(\bar{X} > \bar{x}_{obs} \mid H_0) = \Pr(\bar{X} > 3.75 \mid \mu = 3.2)$

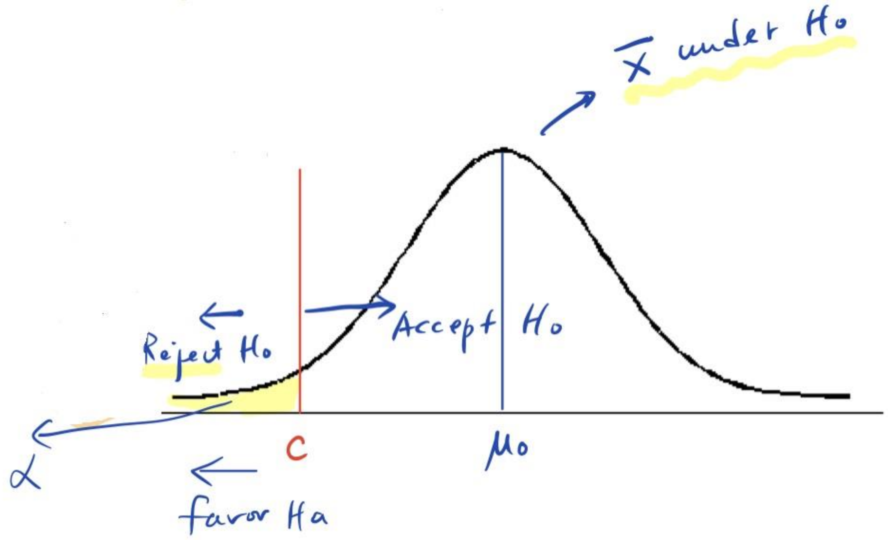
$$= \Pr\left(\frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} > \frac{3.75 - 3.2}{1.5 / \sqrt{25}}\right) = \Pr(Z > 1.833) = 0.0336$$

Since p-value = 0.0336 < 0.05 $= \alpha$ → Reject H_0



Case 2: $H_0 : \mu = \mu_0$ *versus* $H_a : \mu < \mu_0$

Method 1: Reject H_0 if $\bar{x}_{obs} < \mu_0 - z_{\alpha} \dfrac{\sigma}{\sqrt{n}}$



14.7 **Anemia.** Hemoglobin is a protein in red blood cells that carries oxygen from the lungs to body tissues. People with less than 12 grams of hemoglobin per deciliter of blood (g/dl) are anemic. A public health official in Jordan suspects that the mean μ for all children in Jordan is less than 12. He measures a sample of 50 children. Suppose that the "simple conditions" hold: The 50 children are an SRS from all Jordanian children and the hemoglobin level in this population follows a Normal distribution with standard deviation $\sigma = 1.6$ g/dl.

(a) We seek evidence *against* the claim that $\mu = 12$. What is the sampling distribution of $\bar{x}$ in many samples of size 50 if in fact $\mu = 12$? Make a sketch of the Normal curve for this distribution. (Sketch a Normal curve, and then mark the axis using what you know about locating the mean and standard deviation on a Normal curve.)

(b) The sample mean was $\bar{x} = 11.3$ g/dl. Mark this outcome on the sampling distribution. Also mark the outcome $\bar{x} = 11.8$ g/dl of a different study of 50 children in another country. Explain carefully from your sketch why one of these outcomes is good evidence that μ is lower than 12 and also why the other outcome is not good evidence for this conclusion.

Example: Anemia

$H_0 : \mu = 12$ *versus* $H_a : \mu < 12$

$\sigma = 1.6$, $n = 50$, Suppose $\alpha = 0.05 \Rightarrow z_\alpha = 1.645$

$$\frac{X - \mu_0}{\sigma / \sqrt{n}} = \frac{X - 12}{1.6 / \sqrt{50}} \sim^{H_0 \text{ is true}} N(0,1)$$

<u>Method 1: Apply the rejection rule</u>

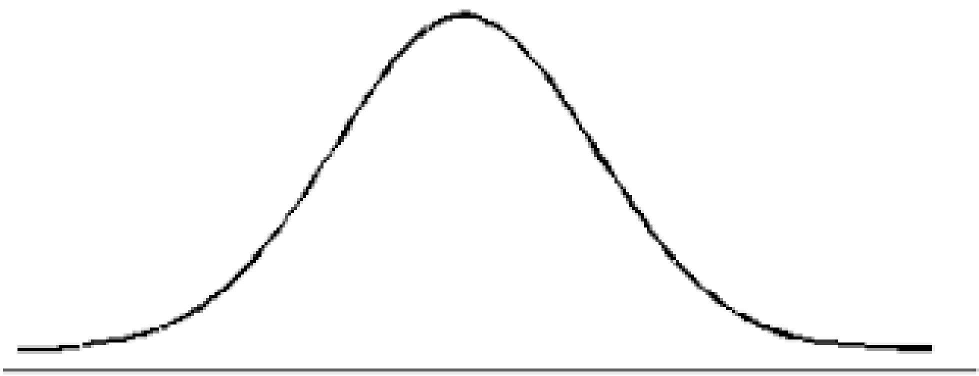
Reject H_0 if $\bar{x}_{obs} < 12 - 1.645 \frac{1.6}{\sqrt{50}} =$ 切點值 $\Leftrightarrow \bar{x}_{obs} < 11.628$

When $\bar{x}_{obs} = 11.3 \rightarrow$ Reject H_0

When $\bar{x}_{obs} = 11.8 \rightarrow$ Fail to reject H_0

<u>Method 2: Use the p-value</u>

p-value = the probability of the tail area from $\bar{x}_{obs}$ in favor of $H_a : \mu < 12$



When $\bar{x}_{obs}$ =11.3 → p-value = $\Pr(\bar{X} < \bar{x}_{obs} \mid H_0) = \Pr(\bar{X} < 11.3 \mid \mu = 12)$

$$= \Pr\left(\frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} < \frac{11.3 - 12}{1.6 / \sqrt{50}}\right)$$
$$= \Pr(Z < -3.09) \approx 0.0010$$

Since p-value = 0.0010 < 0.05 $= \alpha$ → Reject H_0

When $\bar{x}_{obs}$ =11.8 → p-value = $\Pr(\bar{X} < \bar{x}_{obs} \mid H_0) = \Pr(\bar{X} < 11.8 \mid \mu = 12)$

$$= \Pr\left(\frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} < \frac{11.8 - 12}{1.6 / \sqrt{50}}\right)$$
$$= \Pr(Z < -0.8839)$$
$$\approx 0.1867$$

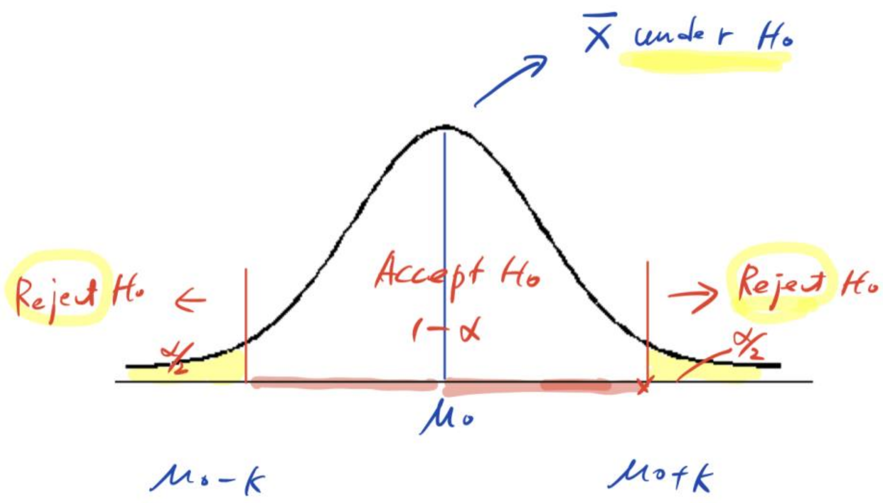
Since p-value = 0.1867 > 0.05 $= \alpha$ → Fail to reject H_0

Case 3: $H_0 : \mu = \mu_0$ *versus* $H_a : \mu \neq \mu_0$

<u>Method 1: Apply the rejection rule</u>

Reject H_0 if $\bar{x}_{obs} > \mu_0 + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ or $\bar{x}_{obs} < \mu_0 - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$

$$\Leftrightarrow \text{ Reject } H_0 \text{ if } z_{obs} = \left| \frac{\bar{x}_{obs} - \mu_0}{\sigma / \sqrt{n}} \right| > z_{\alpha/2}$$



EXAMPLE 14.9 Executives' blood pressures

STATE: The National Center for Health Statistics reports that the systolic blood pressure for males 35 to 44 years of age has mean 128 and standard deviation 15. The medical director of a large company looks at the medical records of 72 male executives in this age group and finds that the mean systolic blood pressure in this sample is $\bar{x} = 126.07$. Is this evidence that the company's executives have a different mean blood pressure from the general population?

Example: Blood pressure

$H_0 : \mu = 128$

versus

$H_a : \mu \neq 128$

$\sigma = 15$, $n = 72$, Suppose $\alpha = 0.05 \Rightarrow z_{\alpha/2} = z_{0.025} = 1.96$ ← **(two-sided)**

Method 1: Check if $\left| \dfrac{\bar{x}_{obs} - 128}{15/\sqrt{72}} \right| > z_{\alpha/2}$

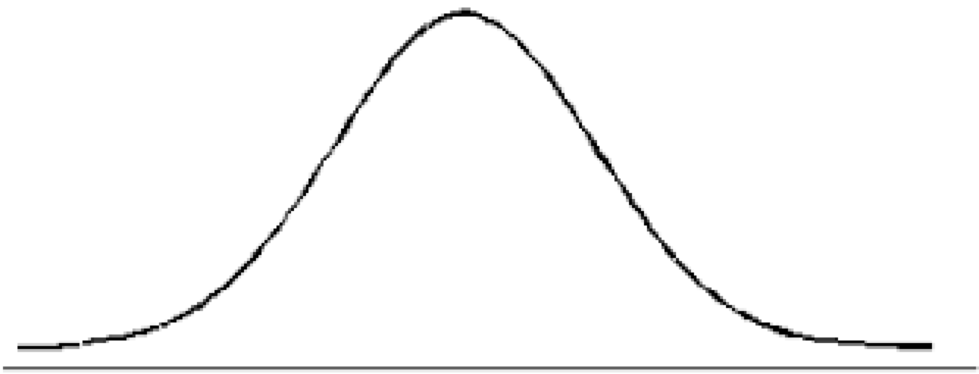
<u>Method 2: Use the p-value</u>

Tail area: the direction in favor of $H_a : \mu \neq 128$ (two sides: larger or smaller)

p-value = $2\Pr(Z > \left| \dfrac{\bar{x}_{obs} - \mu_0}{\sigma/\sqrt{n}} \right|) = 2\Pr(Z > |z_{obs}|)$ ← multiply by 2

$= 2\Pr(Z > 1.131) = 2 \times 0.0951 = 0.19$

Since p-value = 0.19 > 0.05 = α, we fail to reject H_0.



Remarks:

- 雙尾檢定的拒絕區包含兩個方向，每個方向在切點以外有 $\alpha/2$ 的機率，所以才會將 p-value 定義為 $2\Pr(Z > |z_{obs}|)$.

- 有可能發生單尾檢定會拒絕 (Z 切點 1.645)，但雙尾檢定不會拒絕 (Z 切點 1.96)，也就是雙尾檢定的標準較單尾更嚴格.

Method 3: Use confidence intervals ← **(New, only for two-sided test!!)**

Reject H_0 if $\mu_0 \notin [\overline{x}_{obs} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{x}_{obs} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}]$ (中點 = $\overline{x}_{obs}$!!)

If the $(1-\alpha)100\%$ confidence interval does not contain μ_0, then reject H_0.

$$126.07 \pm 1.96\frac{15}{\sqrt{72}} \Leftrightarrow (122.54, 129.46)$$

Since $\mu_0 = 128 \in (122.54, 129.46)$, we fail to reject H_0.

Remark:

The mid-point of a confidence interval is $\overline{x}_{obs}$ (the interval is a random interval),

NOT μ_0. → 同學常犯的錯誤!

Topic: Testing μ when σ^2 is estimated by S^2

$X_i \sim^{iid} N(\mu, \sigma^2)$, where σ^2 unknown but estimated by S^2

Distribution Theory:

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{(n-1)}$$

Case 1: $H_0: \mu = \mu_0$ *versus* $H_a: \mu > \mu_0$

Method 1: Reject H_0 if 切點值 $= \mu_0 + t_{n-1,\alpha}\dfrac{s}{\sqrt{n}} < \bar{x}_{obs}$

$$\Leftrightarrow \text{ Reject } H_0 \text{ if } t_{obs} = \frac{\bar{x}_{obs}-\mu_0}{s/\sqrt{n}} > t_{n-1,\alpha}$$

where $t_{n-1,\alpha}$ satisfies $\Pr(T_{n-1} > t_{n-1,\alpha}) = \alpha$

Method 2: Reject H_0 if p-value $< \alpha$

$$\begin{aligned}\text{p-value} = \Pr(\bar{X} > \bar{x}_{obs} \mid \mu = \mu_0) &= \Pr\left(\frac{\bar{X}-\mu_0}{s/\sqrt{n}} > \frac{\bar{x}_{obs}-\mu_0}{s/\sqrt{n}} \mid H_0\right) \\ &= \Pr\left(T_{n-1} > \frac{\bar{x}_{obs}-\mu_0}{s/\sqrt{n}}\right) \\ &= \Pr(T_{n-1} > t_{obs})\end{aligned}$$

Note: T table does not provide detailed information about the tail probability. So we can only find a region of the p-value.

Case 2: $H_0: \mu = \mu_0$ *versus* $H_a: \mu < \mu_0$

Method 1: Reject H_0 if $\bar{x}_{obs} < \mu_0 - t_{n-1,\alpha}\dfrac{s}{\sqrt{n}} =$ 切點值

$$\Leftrightarrow \text{ Reject } H_0 \text{ if } \frac{\bar{x}_{obs}-\mu_0}{s/\sqrt{n}} < -t_{n-1,\alpha}$$

Method 2: Reject H_0 if p-value $< \alpha$

$$\begin{aligned}\text{p-value} = \Pr(\bar{X} < \bar{x}_{obs} \mid \mu = \mu_0) &= \Pr\left(\frac{\bar{X}-\mu_0}{s/\sqrt{n}} < \frac{\bar{x}_{obs}-\mu_0}{s/\sqrt{n}} \mid H_0\right) \\ &= \Pr\left(T_{n-1} < \frac{\bar{x}_{obs}-\mu_0}{s/\sqrt{n}}\right) \\ &= \Pr(T_{n-1} < t_{obs}) = \Pr(T_{n-1} > |t_{obs}|)\end{aligned}$$

Case 3: $H_0 : \mu = \mu_0$ *versus* $H_a : \mu \neq \mu_0$

Method 1: Reject H_0 if $\bar{x}_{obs} > \mu_0 + t_{n-1,\alpha/2} \frac{s}{\sqrt{n}}$ or $\bar{x}_{obs} < \mu_0 - t_{n-1,\alpha/2} \frac{s}{\sqrt{n}}$

$\Leftrightarrow$ Reject H_0 if $\left| \frac{\bar{x}_{obs} - \mu_0}{s/\sqrt{n}} \right| > t_{n-1,\alpha/2}$

Method 2: Reject H_0 if p-value $< \alpha$

p-value = $2\Pr(T_{n-1} > \left| \frac{\bar{x}_{obs} - \mu_0}{s/\sqrt{n}} \right|) = 2\Pr(T_{n-1} > |t_{obs}|)$

Method 3: Use Confidence intervals

Rule: Reject H_0 if $\mu_0 \notin [\bar{x}_{obs} - t_{n-1,\alpha/2} \frac{s}{\sqrt{n}}, \bar{x}_{obs} + t_{n-1,\alpha/2} \frac{s}{\sqrt{n}}]$

Note: If the $(1-\alpha)100\%$ confidence interval does not contain μ_0, then reject H_0.

Exercises: T test

Assume $X_i \sim^{iid} N(\mu, \sigma^2)$ for $i = 1, \ldots, 25$.

However σ^2 is unknown but the sample variance equals $S^2 = 5^2$.

Let the level of significance = α = 0.05.

(a). $H_0 : \mu = 3$ vs. $H_a : \mu > 3$. The observed value of $\bar{X}$ is $\bar{x}_{obs} = 4$.

Method 1:

Reject H_0 if $\bar{x}_{obs} > 3 + t_{24,0.05} \frac{5}{\sqrt{25}}$

$\Leftrightarrow \bar{x}_{obs} > 4.711$ (i.e. $t_{24,0.05} = 1.711$)

Since $\bar{x}_{obs} = 4 < 4.711$, we fail to reject H_0

Method 2: p-value

$$\begin{aligned} p-value &= \Pr(\bar{X} > \bar{x}_{obs} \mid \mu = \mu_0) \\ &= \Pr(T_{24} > \frac{4-3}{5/\sqrt{25}}) \\ &= \Pr(T_{24} > 1) \in (0.1, 0.25) \end{aligned}$$

Since the p-value $\in (0.1, 0.25)$ $> \alpha = 0.05$, we fail to reject H_0

ν	.40 .60	.25 .75	.10 .90	.05 .95	.025 .975	.01 .99	.005 .995	.001 .999	.0005 (Area to right) .9995 (Area to left)
21	0.257	0.686	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.256	0.686	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.256	0.685	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.256	0.685	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787	3.450	3.725

Note: T-table 編表過於簡略 → 無法得到確切的 p-value

(b). $H_0: \mu = 3$ vs. $H_a: \mu < 3$. The observed value of $\bar{X}$ is $\bar{x}_{obs} = 2$.

Method 1: reject H_0

if $\bar{x}_{obs} < 3 - t_{24,0.05}\dfrac{5}{\sqrt{25}} \Leftrightarrow \bar{x}_{obs} < 1.289$

Since $\bar{x}_{obs} = 2 > 1.289$, we fail to reject H_0

Method 2:

p-value= $\Pr(\bar{X} < \bar{x}_{obs} \mid \mu = \mu_0)$

$= \Pr(T_{24} < \dfrac{2-3}{5/\sqrt{25}})$

$= \Pr(T_{24} < -1) = \Pr(T_{24} > 1)$

p-value $\in (0.1, 0.25)$

Since the p-value $> \alpha = 0.05$, we fail to reject H_0

Note: the T distribution is symmetric around zero.

(c). $H_0: \mu = 3$ vs. $H_a: \mu \neq 3$. The observed value of $\bar{X}$ is 2.

Method 1: Reject H_0 if $|t_{obs}| > t_{24,0.025}$

Since $|t_{obs}| = \left|\dfrac{2-3}{5/\sqrt{25}}\right| = 1 < t_{24,0.025} = 2.064$, we fail to reject H_0.

Method 2:

The p-value = $2 \cdot \Pr(T_{24} > |t_{obs}|) = 2 \cdot \Pr(T_{24} > \left|\dfrac{2-3}{5/\sqrt{100}}\right|) \in (0.2, 0.5)$

Since the p-value $> \alpha = 0.05$, we fail to reject H_0.

Method 3: Reject H_0 if $\mu_0 = 3$ falls outside $2 \pm t_{24,0.05} \dfrac{5}{\sqrt{25}} = (0.968, 3.711)$

Since $\mu_0 = 3 \in (0.289, 3.711)$, we fail to reject $H_0 : \mu = 3$

Comparison between T and Z confidence interval and tests

A. Confidence intervals – comparison

$(1-\alpha)100\%$ **Z- interval:** $[\bar{X} - z_{\alpha/2} \cdot \sigma / \sqrt{n}, \bar{X} + z_{\alpha/2} \cdot \sigma / \sqrt{n}]$

$(1-\alpha)100\%$ **T- interval:** $[\bar{X} - t_{(n-1),\alpha/2} \cdot S / \sqrt{n}, \bar{X} + t_{(n-1),\alpha/2} \cdot S / \sqrt{n}]$

Fact: Suppose $\sigma = S$. (估計值碰巧和真值相同).

Since $z_{\alpha/2} \le t_{(n-1),\alpha/2}$, so that T intervals are always <u>wider</u> (less precise) than Z intervals under the same confidence level. → 天下沒有白吃的午餐

B. Hypothesis testing (舉例)

$H_0 : \mu = \mu_0$ vs. $H_a : \mu > \mu_0$

p-value based on Z test: $\Pr(Z > \dfrac{\bar{x}_{obs} - \mu_0}{\sigma / \sqrt{n}})$

p-value based on T test: $\Pr(T_{(n-1)} > \dfrac{\bar{x}_{obs} - \mu_0}{S / \sqrt{n}})$

If $\sigma = S$, $\Pr(Z > \dfrac{\bar{x}_{obs} - \mu_0}{\sigma / \sqrt{n}}) \le \Pr(T_{(n-1)} > \dfrac{\bar{x}_{obs} - \mu_0}{s / \sqrt{n}})$

Fact: it is more difficult to reject H_0 for a T test.

Definition: 檢定力

Power = $1-\beta$ **=** $\Pr(\text{Reject } H_0 \mid H_0 \text{ is false})$ = $\Pr(\text{Reject } H_0 \mid H_a \text{ is true})$

		Truth	
		H_0 **is true**	H_0 is false
Decision	Accept H_0	OK	Type II error β
	Reject H_0	Type I error α	OK Power = 1- β

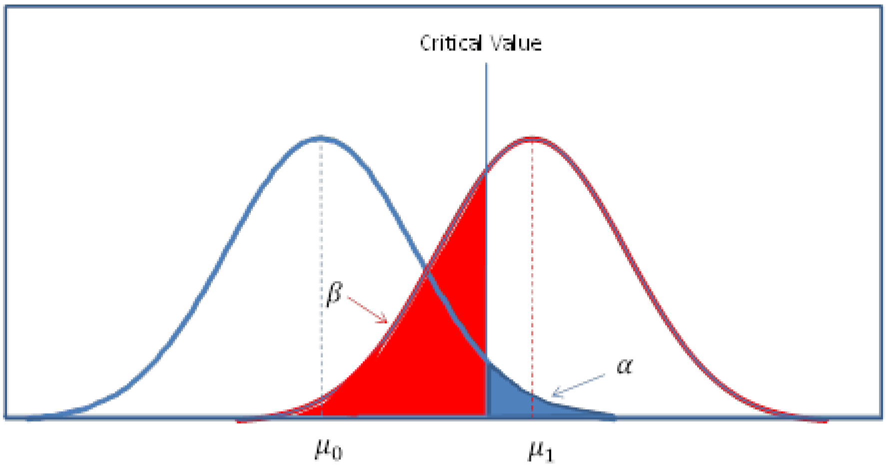
Power calculation:

Step 1: Need to specify $H_a : \mu = \mu_a \ (\mu_1)$

- Step 2: Use Method 1 to obtain the rule of rejecting H_0

- Step =3: use the fact $H_a : \mu = \mu_a$ to find the probability of reject H_0

Case 1: $H_a : \mu = \mu_a > \mu_0$



Exercises: $X_i \sim^{iid} N(\mu, \sigma^2)$ for $i = 1, \ldots, 100$ (Note: n = 100)

$\sigma^2 = 5^2$ is known and α = 0.05.

Case 1: $H_0 : \mu = 3$ vs. $H_a : \mu > 3$.

Step 1: Rejection rule

Reject if $\bar{X} > \mu_0 + z_\alpha \dfrac{\sigma}{\sqrt{n}}$ → $\bar{X} > 3 + 1.645\dfrac{5}{5}$ (保留此形式，不要化簡)

Step 2: Calculate the probability of rejecting H_0 under $H_a : \mu = \mu_a$

- When $\mu_a = 4$

  Power = 1- β = Pr(Reject $H_0 \mid \mu_a = 4$)

$$= \Pr(\bar{X} > 3 + 1.645\frac{5}{10} \mid \mu_a = 4)$$

$$= \Pr(\frac{\bar{X} - \mu_a}{\sigma / \sqrt{100}} > \frac{3 - 4 + 1.645\frac{5}{10}}{5/10} \mid \mu_a = 4)$$ (請保留住<u>算式細節</u>)

$$= \Pr(Z > \frac{3-4}{5/10} + 1.645)$$

$$= \Pr(Z > (-2) + 1.645) = \Pr(Z > -0.355) \ = 0.6387$$

When $\mu_a = 5$

Power = $1-\beta$ = Pr(Reject $H_0 |$ $\mu_a = 5$) = $\Pr(\bar{X} > 3 + 1.645\frac{5}{10} | \mu_a = 5)$

$$= \Pr(\frac{\bar{X}-\mu_a}{\sigma/\sqrt{100}} > \frac{3-5+1.645\frac{5}{10}}{5/10} | \mu_a = 5) = \Pr(Z > (-4)+1.645) = \Pr(Z > -2.355) \approx 0$$

Power $= \Pr(Z > -2.355) \approx 1$

Remark: From $\mu_a = 4$ to $\mu_a = 5$ when μ_a moves away from μ_0, the corresponding power increases from 0.6387 to 1.

Power Curve: X-axis: μ_a, Y-axis: Pr(Reject $H_0 |$ μ_a),

Case 2: $H_0 : \mu = 3$ vs. $H_a : \mu < 3$.

Step 1: Rejection rule

Reject if $\bar{X} < \mu_0 - z_\alpha \frac{\sigma}{\sqrt{n}}$ → $\bar{X} < 3 - 1.645\frac{5}{10}$

Step 2: Calculate the probability of rejecting H_0 under $H_a : \mu = \mu_a$

When $\mu_a = 2$

Power = $1-\beta$ = Pr(Reject $H_0 |$ $\mu_a = 2$) = $\Pr(\bar{X} < 3 - 1.645\frac{5}{10} | \mu_a = 2)$

$$= \Pr(\frac{\bar{X}-\mu_a}{\sigma/\sqrt{100}} < \frac{3-2-1.645\frac{5}{10}}{5/10} | \mu_a = 2)$$

$$= \Pr(Z < \frac{3-2}{5/10} - 1.645) \quad = \Pr(Z < 0.355) \quad = 0.6387$$

When $\mu_a = 1$

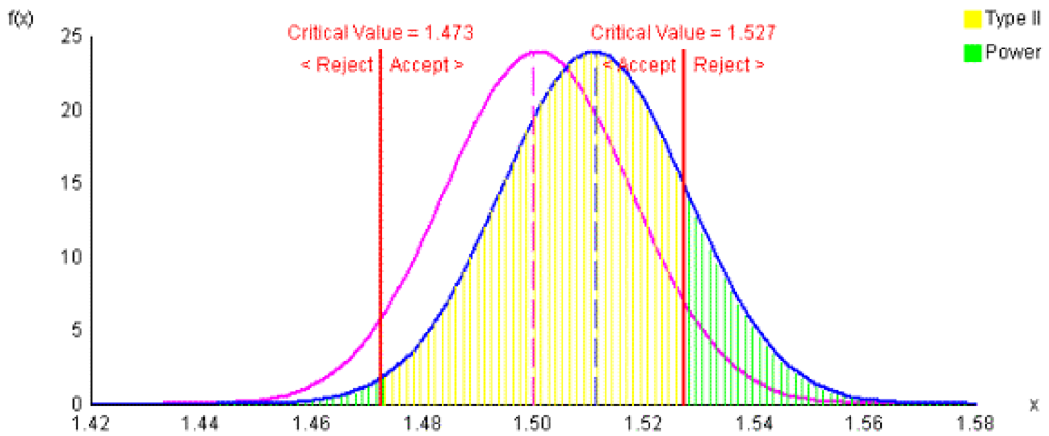
Power = $1-\beta$ = Pr(Reject $H_0 |$ $\mu_a = 1$) = $\Pr(\bar{X} < 3 - 1.645\frac{5}{10} | \mu_a = 1)$

$$= \Pr(Z < \frac{3-1}{5/10} - 1.645) \quad = \Pr(Z < 2.355) \approx 1$$

Remark: From $\mu_a = 2$ to $\mu_a = 1$ when μ_a moves away from μ_0, the corresponding power increases.

Case 3: test: $H_0 : \mu = 3$ vs. $H_a : \mu \neq 3$.

→ **the power probability contains two pieces**



Step 1: Rejection Rule: Reject if

- $\bar{X} < \mu_0 - z_{\alpha/2}\dfrac{\sigma}{\sqrt{100}}$ → $\bar{X} < 3 - 1.96\dfrac{5}{10}$

  or $\bar{X} > \mu_0 + z_{\alpha}\dfrac{\sigma}{\sqrt{100}}$ → $\bar{X} < 3 + 1.96\dfrac{5}{10}$

Step 2: 由這個公式重新依據 $H_a : \mu = \mu_a$ 求出(正確) reject H_0 的機率.

- When $\mu_a = 2$ (注意與上面的單尾情形不同，要小心)

$$\beta = \Pr(\text{Accept } H_0 \mid \mu_a = 2)$$

$$= \Pr(3 - 1.96\frac{5}{10} < \bar{X} < 3 + 1.96\frac{5}{10} \mid \mu_a = 2)$$

$$\text{Power} = \Pr(Z < \frac{(3-2)}{5/10} - 1.96) + \Pr(Z > \frac{(3-2)}{5/10} + 1.96)$$

$$= \Pr(Z < 0.04) + \Pr(Z > 3.96) \approx 0.516 + 0 \text{ (主要的值在第一項)}$$

When $\mu_a = 4$

$$\beta = \Pr(3 - 1.96\frac{5}{10} < \bar{X} < 3 + 1.96\frac{5}{10} \mid \mu_a = 4)$$

$$\text{Power} = \Pr(\bar{X} < 3 - 1.96\frac{5}{10} \mid \mu_a = 4) + \Pr(\bar{X} > 3 + 1.96\frac{5}{10} \mid \mu_a = 4)$$

$$= \Pr(Z < \frac{(3-4)}{5/10} - 1.96) + \Pr(Z > \frac{(3-4)}{5/10} + 1.96)$$

$$= \Pr(Z < -3.96) + \Pr(Z > -0.04) \approx 0 + 0.516 \text{ (主要值在第二項)}$$

Power as a function of μ (draw in class)

Case 1: $H_a : \mu > \mu_0$ Case 2: $H_a : \mu < \mu_0$ Case 3: $H_a : \mu \neq \mu_0$

Remarks:

- $\alpha = \Pr(\text{reject } H_0 \mid \mu = \mu_0)$ → false probability
- power = $\Pr(\text{reject } H_0 \mid \mu = \mu_a)$ → correct probability
- Both are the probability of rejecting H_0 but the underlying distributions are different such that α is based on $H_0 : \mu = \mu_0$ while the power is based on $H_a : \mu = \mu_a$.

Analysis of Power illustrated by Case 1

Case 1: $H_a : \mu = \mu_a > \mu_0$ → $\mu_0 - \mu_a < 0$

Power $= \Pr(Z > \dfrac{\mu_0 - \mu_a}{\sigma / \sqrt{n}} + z_\alpha)$ ($\mu_0 - \mu_a < 0$ 越負,切點越左邊,大於切點的機率越高)

The power value is determined by

- Distance between μ_0 and μ_a:
  - $|\mu_a - \mu_0|$ is larger, then the power is larger
- The value of σ:
  - σ is smaller, power is larger
- The sample size n:
  - n is larger, the power is larger

Find the smallest sample size to achieve the goal on the power value

- 當研究者心目中有個 μ_a (研究者認為真正的 μ 值), 希望檢定的 power 在此 μ_a 值時要達到某個標準 (如 power 至少九成). 可以求"最小樣本數".

Case 1: $H_0: \mu = 3$ vs. $H_a: \mu > 3$.

$$\Pr(\bar{X} > 3 + 1.645\frac{5}{\sqrt{n}} \mid \mu = 4) \geq 0.8$$

➔ $\Pr(Z > \frac{3-4}{5/\sqrt{n}} + 1.645) \geq 0.8$ (讓 $\sqrt{n}$ 成為變數)

$\frac{3-4}{5/\sqrt{n}} + 1.645 \approx -0.84$ (From Z-table: $\Pr(Z > -0.84) \approx 0.8$)

➔ $5(1.645 + 0.84) \approx \sqrt{n}$ ➔ $12.425 \approx \sqrt{n}$

Smallest sample size n= 155. (自行檢查老師有沒有算錯)

Case 2: $Power = \Pr(\bar{X} < \mu_0 - z_\alpha \frac{\sigma}{\sqrt{n}} \mid \mu = \mu_a) \geq K \quad (0 < K < 1)$

$\Pr(\frac{\bar{X} - \mu_a}{\sigma/\sqrt{n}} < \frac{\mu_0 - \mu_a}{\sigma/\sqrt{n}} - z_\alpha) = \Pr(Z < \frac{\mu_0 - \mu_a}{\sigma/\sqrt{n}} - z_\alpha) \geq K$ ➔ 畫圖

Case 3: $Power = \Pr(\bar{X} < \mu_0 - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \mid \mu = \mu_a) + \Pr(\bar{X} > \mu_0 + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \mid \mu = \mu_a) \geq K$

$$Power = \Pr(Z < \frac{\mu_0 - \mu_a}{\sigma/\sqrt{n}} - z_{\alpha/2}) + \Pr(\bar{X} > \frac{\mu_0 - \mu_a}{\sigma/\sqrt{n}} + z_{\alpha/2}) \geq K$$

➔ 不好算最小 n, 不會在考試為難你們☺

補充 假設檢定背後的故事

$H_0:\theta=\theta_0$ vs. $H_a:\theta=\theta_a$.（一般參數 θ，而不一定是 μ）

Neyman–Pearson lemma

From Wikipedia, the free encyclopedia

This article includes a list of references, but **its sources remain unclear** because it has **insufficient inline citations**. Please help to improve this article by introducing more precise citations. *(May 2018) (Learn how and when to remove this template message)*

In statistics, the **Neyman–Pearson lemma** was introduced by Jerzy Neyman and Egon Pearson in a paper in 1933.[1]

Suppose one is performing a hypothesis test between two simple hypotheses H_0: $\theta=\theta_0$ and H_1: $\theta=\theta_1$ using the likelihood-ratio test with threshold η, which rejects H_0 in favour of H_1 at a significance level of

$$\alpha = P(\Lambda(x)\le \eta \mid H_0),$$

where

$$\Lambda(x) := \frac{\mathcal{L}(\theta_0\mid x)}{\mathcal{L}(\theta_1\mid x)}$$

and $\mathcal{L}(\theta\mid x)$ is the likelihood function. Then, the Neyman–Pearson lemma states that the likelihood ratio, $\Lambda(x)$, is the **most powerful test** at significance level α.

If the test is most powerful for all $\theta_1\in\Theta_1$, it is said to be uniformly most powerful (UMP) for alternatives in the set Θ_1.

In practice, the likelihood ratio is often used directly to construct tests — see likelihood-ratio test. However it can also be used to suggest particular test-statistics that might be of interest or to suggest simplified tests — for this, one considers algebraic manipulation of the ratio to see if there are key statistics in it related to the size of the ratio (i.e. whether a large statistic corresponds to a small ratio or to a large one).

$$\text{Likelihood ratio statistics} = \frac{L(\theta_0\mid \text{data})}{L(\theta_a\mid \text{data})} \rightarrow \text{越小，代表越傾向拒絕}$$

Neyman　　　　*Egon Pearson*

http://www.agron.ntu.edu.tw/biostat/EgonPearson.html

Pearson 與 Neyman 第一個討論的問題就是 **Karl Pearson 的卡方適合度檢定(chi square goodness of fit test)**。這個檢定是被用來檢視觀察到的資料是否符合某特定的分配。但是這個檢定與其他一般檢定不同。**只要有一組資料統計分析家就有無限多種的方式能運用到這個檢定。**[1]當時還沒有一個準則來判定哪一種分析方式是最好的。每次使用這個檢定時，統計分析家就只能做任意的選擇。

Egon Pearson 將這個問題寫信給 Neyman，信件內容如下：

如果我將一筆資料套用卡方適合度檢定檢視是否符合常態分配，並且得到不顯著的 P-value，那我怎麼知道這筆資料是真的符合常態分配？**也就是說，我怎麼能夠確認其他版本的卡方適合度檢定或者是其他尚未發現的適合度檢定不會產生顯著的 P-value，**

[1] 這段意思是 Karl Pearson 的卡方檢定法不具有唯一性

讓這筆資料不符合常態分配。

Neyman 帶著這個問題回到華沙(Warsaw，波蘭首都)後，兩人的信件交流就開始了。**他們兩人對於 Fisher 使用概似函數來估計的想法念念不忘。於是兩人試圖找出概似函數與適合度檢定之間的關連性。**而他們兩人合作後第一篇的學術論文就因此而誕生了。這篇論文是他們合作所發表三篇經典論文中統計理論最艱深的，也帶動當時對於顯著性檢定觀念的變革。

但發表這篇經典論文卻有一個有趣的現象。在當時 Biometrika 算是頂尖一流的期刊，而 Egon Pearson 的父親 Karl Pearson 是負責該期刊的編輯。照道理說，該篇論文應該要在 Biometrika 上發表。但是這篇論文卻在 Egon Pearson 與 Neyman 合作所創立的新期刊 Statistical Research Memoirs 發表。原因出在於上一代兩位統計大師 Karl Pearson 與 Fisher 的交惡。Fisher 認為 Karl Pearson 所提倡的卡方適合度檢定是有瑕疵的，並於 1922 年發表論文第一次批評 Karl Pearson 的這個檢定。而 Karl Pearson 則認為 Fisher 的概似函數估計法完全是無稽之談。但是 Egon Pearson 與 Neyman 的研究成果偏偏又是以概似函數觀念為基礎所推得的。**在這種情況底下，Egon Pearson 為了避免觸怒他的父親 Karl Pearson，才會將這篇足以撼動統計學界的論文發表在一個名不見經傳的期刊上。**

http://mathshistory.st-andrews.ac.uk/Biographies/Pearson_Egon.html

Karl Pearson's work had been under attack from R A Fisher for a number of years and Egon later explained

I had to go through the painful stage of realising that K.P. [his father] could be wrong ... and I was torn between conflicting emotions:

a. finding it difficult to understand R.A.F. [Fisher],

b. hating him for his attacks on my paternal 'god',

c. realising that in some things at least he was right.

In [3] the friendship that developed between Neyman and Pearson during 1926 is described. It paints a picture of Pearson, and his difficulties, at this time:
Pearson was an introverted young man who felt inferior for a number of reasons. He had grown up in the great shadow of K.P. [Karl Pearson], "lovingly protected" in his childhood and kept out of the war in his youth. At Cambridge he had felt cut off from classmates of his own generation, all veterans of the conflict. He suspected that K.P. was disappointed in him, for he had not gone on to his second mathematics tripos but had taken his degree on the basis of work he had done during the war. After joining the staff of K.P.'s laboratory, he had continued to live at home and to have almost all his social contacts with relatives.