Lecture 10 11/24/2021

Statistical Issues related to Confidence Intervals

Statistical Inference – Overview

- Major Interest: μ (when $Var(X) = \sigma^2$ is known or unknown)

1. Point estimation (點估計): use $\bar{X}$ to estimate μ
2. Interval estimation (區間估計):

   Set $(1-\alpha)100\%$ as the confidence level

   use $\bar{X} \pm a$ or $(\bar{X} - a, \bar{X} + a)$ to estimate μ

where a is the "margin of error" (誤差範圍)

The margin of error is determined by the distribution of $\bar{X}$ and $(1-\alpha)100\%$

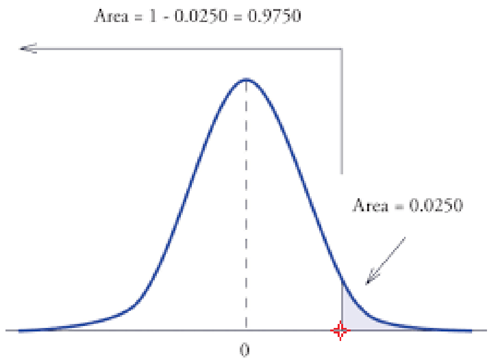
Formula:

- σ^2 is known → margin of error = $a = z_{\alpha/2} \dfrac{\sigma}{\sqrt{n}}$
- σ^2 is unknown → margin of error = $a = t_{n-1,\alpha/2} \dfrac{S}{\sqrt{n}}$ (Later)

where $z_{\alpha/2}$ or $t_{n-1,\alpha/2}$ are table values such that

$$\Pr(Z > z_{\alpha/2}) = \alpha/2 \quad \& \quad \Pr(T_{n-1} > t_{n-1,\alpha/2}) = \alpha/2 \quad \text{(Later)}$$

For 95% confidence interval, $\alpha = 0.05$, we should find $z_{\alpha/2} = z_{0.025}$



3. Hypothesis testing (假設檢定)

   Test whether $H_0 : \mu = \mu_0$ or not

   Note: Usually the researcher wants to reject H_0.

Examples of hypothesis testing (見下頁)

https://examples.yourdictionary.com/examples-of-hypothesis-testing.html

重點

- 兩個假說
    - Null hypothesis (虛無假說) vs. Alternative hypothesis (對立假說)
- 顯著水準

Example: 薄荷精油是否可緩解焦慮

Peppermint Essential Oil

Essential oils are becoming more and more popular. Chamomile, lavender, and ylang-ylang are commonly touted as anxiety remedies. Perhaps you'd like to test the healing powers of peppermint essential oil. Your hypothesis might go something like this:

- **Null hypothesis** - Peppermint essential oil has no effect on the pangs of anxiety.
- **Alternative hypothesis** - Peppermint essential oil alleviates the pangs of anxiety.
- **Significance level** - The significance level is 0.25 (allowing for a better shot at proving your alternative hypothesis).
- **P-value** - The p-value is calculated as 0.05.
- **Conclusion** - After providing one group with peppermint oil and the other with a placebo, you gauge the difference between the two based on self-reported levels of anxiety. Based on your calculations, the difference between the two groups is statistically significant with a p-value of 0.05, well below the defined alpha of 0.25. You conclude that your study supports the alternative hypothesis that peppermint essential oil can alleviate the pangs of anxiety.

Example: 維他命 C 是否可預防感冒

Vitamin C

Is it true that vitamin C has the ability to cure or prevent the common cold? Or is it just a myth? There's nothing like an in-depth experiment to get to the bottom of it all. A potential hypothesis test could look something like this:

- **Null hypothesis** - Children who take vitamin C are no less likely to become ill during flu season.
- **Alternative hypothesis** - Children who take vitamin C are less likely to become ill during flu season.
- **Significance level** - The significance level is 0.05.
- **P-value** - The p-value is calculated to be 0.20.
- **Conclusion** - After providing one group with vitamin C during flu season and the other with a placebo, you record whether or not participants got sick by the end of flu season. After conducting your statistical analysis on the results, you determine a p-value of 0.20. That is above the desired significance level of 0.05, and thus you fail to reject the null hypothesis. Based on your experiment, there is no support for the (alternative) hypothesis that vitamin C can prevent colds.

Topic 1: Confidence interval for μ:

Case 1: $X_i \sim^{iid} N(\mu, \sigma^2)$ and σ is known

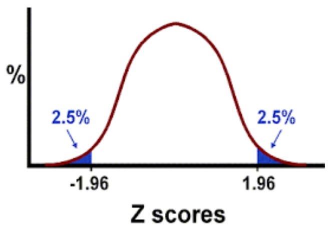
- Basis: $\bar{X} \sim N(\mu, Var(\bar{X}) = \frac{\sigma^2}{n}) \Leftrightarrow \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

$$\Pr(-z_{\alpha/2} \le \frac{|\bar{X}-\mu|}{\sigma/\sqrt{n}} \le z_{\alpha/2}) = \Pr(|Z| \le z_{\alpha/2}) = 1-\alpha$$

$(1-\alpha)100\%$ confidence interval satisfies

$$\Pr(\mu \in [\bar{X} - z_{\alpha/2} \cdot \sigma/\sqrt{n}, \bar{X} + z_{\alpha/2} \cdot \sigma/\sqrt{n}]) = 1-\alpha$$

* 95% confidence interval: use $z_{0.05/2} = z_{0.025} = 1.96$



* 90% confidence interval: use $z_{0.1/2} = z_{0.05} = 1.645$

Number skills of young men. The National Assessment of Educational Progress (NAEP) gave a test of basic arithmetic and the ability to apply it in everyday life to a sample of 840 men 21 to 25 years of age.[2] Scores range from 0 to 500; for example, someone with a score of 325 can determine the price of a meal from a menu. The mean score for these 840 young men was $\bar{x} = 272$. Consider the NAEP sample as an SRS from a Normal population with standard deviation $\sigma = 60$.

(a) What distribution describes how the sample mean $\bar{x}$ will vary if we take many such samples?

(b) What is the 95% confidence interval for the population mean score μ based on this one sample?

(c) What does it mean to say that we have "95% confidence" in this interval?

Sol: $n = 840$, $\sigma = 60$, $\bar{x}_{obs}$ = <u>observed value</u> of $\bar{X}$ = 272

a. The distribution of $\bar{X}$ follows $N(\mu, \text{var}(\bar{X}) = \frac{\sigma^2}{840} = \frac{3600}{840})$

b. $\bar{x}_{obs} \pm z_{0.025}\dfrac{\sigma}{\sqrt{n}} = 272 \pm 1.96\dfrac{60}{\sqrt{840}} = (267.94, 276.06)$

c. If we repeat the sampling procedure many times, there is 95% chance that the random interval will cover the true value of μ. (More later)

EXAMPLE 14.3 Healing of skin wounds

STATE: Biologists studying the healing of skin wounds measured the rate at which new cells closed a razor cut made in the skin of an anesthetized newt. Here are data from 18 newts, measured in micrometers (millionths of a meter) per hour:[3]

29	27	34	40	22	28	14	35	26
35	12	30	23	18	11	22	23	33

This is one of several sets of measurements made under different conditions. We want to estimate the mean healing rate for comparison with rates under other conditions.

FORMULATE: We will estimate the mean rate μ for all newts of this species by giving a 95% confidence interval.

SOLVE: We should start by checking the conditions for inference. For this first example, we will find the interval, then discuss how statistical practice deals with conditions that are never perfectly satisfied.

The mean of the sample is $\bar{x} = 25.67$. As part of the "simple conditions," suppose that from past experience with this species of newts we know that the standard deviation of healing rates is 8 micrometers per hour. For 95% confidence, the critical value is $z^* = 1.960$. A 95% confidence interval for μ is therefore

$$\begin{aligned}\bar{x} \pm z^*\frac{\sigma}{\sqrt{n}} &= 25.67 \pm 1.960\frac{8}{\sqrt{18}} \\ &= 25.67 \pm 3.70 \\ &= 21.97 \text{ to } 29.37\end{aligned}$$

newt: 蠑螈

14.6 **IQ test scores.** Here are the IQ test scores of 31 seventh-grade girls in a midwest school district:[4]

$\bar{x} = 105.84$

114	100	104	89	102	91	114	114	103	105	
108	130	120	132	111	128	118	119	86	72	
111	103	74	112	107	103	98	96	112	112	93

(a) These 31 girls are an SRS of all seventh-grade girls in the school district. Suppose that the standard deviation of IQ scores in this population is known to be $\sigma = 15$. We expect the distribution of IQ scores to be close to Normal. Make a stemplot of the distribution of these 31 scores (split the stems) to verify that there are no major departures from Normality. You have now checked the "simple conditions" to the extent possible.

(b) Estimate the mean IQ score for all seventh-grade girls in the school district using a 99% confidence interval. Follow the four-step process as illustrated in Example 14.3. (Note that IQ scores are typically rounded to the nearest integer when provided. For inference's sake, we treat IQ scores as a continuous random variable.)

$n = 31,\ \bar{x}_{obs} = 105.84,\ \sigma = 15$

Stem-leaf plot (split the stem) → check whether the distribution is roughly symmetric

7	2 4
7	
8	
8	6 9
9	1 3
9	6 8
10	0 2 3 3 3 4
10	5 7 8
11	1 1 2 2 2 4 4 4
11	8 9
12	0
12	8
13	0 2

114	100	104	89	102	91	114	114	103	105	
108	130	120	132	111	128	118	119	86	72	
111	103	74	112	107	103	98	96	112	112	93

99% confidence interval → use $z_{0.01/2} = z_{0.005} = 2.575$

99% CI = $105.84 \pm 2.575 \frac{15}{\sqrt{31}} = [98.933, 112.767]$

Remarks on confidence intervals (important)

- Before sampling, $[\bar{X}-z_{\alpha/2}\cdot\sigma/\sqrt{n}, \bar{X}+z_{\alpha/2}\cdot\sigma/\sqrt{n}]$ is a random interval and the probability of this interval covers the true value of μ is $1-\alpha$.

- After sampling, the data $(x_1,\ldots,x_n)$ is a realization of the random sample and $\bar{x}_{obs}=\sum_{i=1}^{n}x_i/n$ becomes a number. Accordingly there are only two possibilities about whether the fixed interval $[\bar{x}_{obs}-z_{\alpha/2}\cdot\sigma/\sqrt{n}, \bar{x}_{obs}+z_{\alpha/2}\cdot\sigma/\sqrt{n}]$ contains the true μ: "yes" or "no". This is why $1-\alpha$ is called as a "confidence level" rather than a probability.

Case 2: $X_i \overset{iid}{\sim}$ an unknown distribution with $E(X_i)=\mu$ and $Var(X_i)=\sigma^2$ (known). Based on the Central Limit Theorem

$$\bar{X} \underset{n\to\infty}{\sim} N(\mu, Var(\bar{X})=\frac{\sigma^2}{n}) \quad \Leftrightarrow \quad \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \underset{n\to\infty}{\sim} N(0,1)$$

$$\Pr(\mu\in[\bar{x}_{obs}-z_{\alpha/2}\cdot\sigma/\sqrt{n}, \bar{x}_{obs}+z_{\alpha/2}\cdot\sigma/\sqrt{n}])\approx(1-\alpha)100\%$$

Note: The confidence level is an approximation when the population does not follow a normal distribution but the sample size is large.

More Philosophical Issues related to Confidence Intervals

1. Q: Why not set $\alpha = 0$? (confidence level = 100%)

   A: The interval will be too wide and becomes impractical.

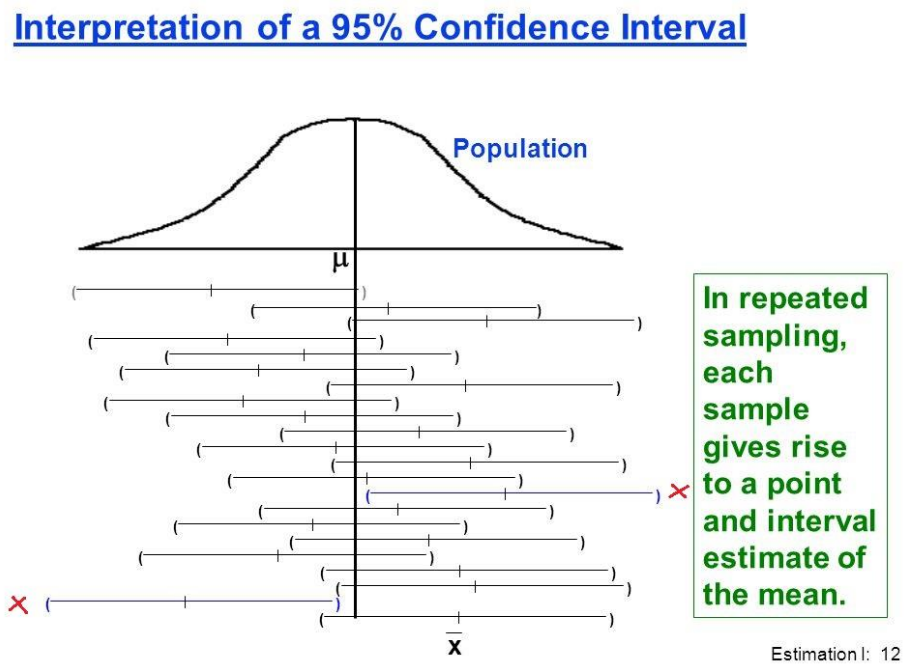
   統計學家告訴你: **"放棄完美主義 (不需要達到 100%), 你會獲得更多"**

2. $[\bar{X} - 1.96 \cdot \sigma / \sqrt{n}, \bar{X} + 1.96 \cdot \sigma / \sqrt{n}]$ is a <u>random interval</u>

   μ is a <u>fixed constant</u>.

   $\Pr(\mu \in [\bar{X} - 1.96 \cdot \sigma / \sqrt{n}, \bar{X} + 1.96 \cdot \sigma / \sqrt{n}]) = 0.95$

   Through <u>repeated sampling</u>, the probability that the random interval will cover the true value of μ is 0.95.



3. In reality, the sampling procedure will only be performed <u>once</u>. There are only two possibilities: the interval either contains μ or the interval does not con*tain* μ.

動畫版: (code 貼在 R 檔案) https://youtu.be/2gl_TPMWTCI

Story: *Neyman and The confidence trick*

https://www150.statcan.gc.ca/n1/pub/12-001-x/2013002/article/11883/section3-eng.htm

Supplementary: *Frequentist vs. Bayesian*

Topic 2: Confidence Interval for μ when σ is unknown

- Use $S^2 = \dfrac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$ to estimate σ^2
- Replace $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ by $\dfrac{\bar{X}-\mu}{S/\sqrt{n}}$

Fact: $X_i \sim^{iid} N(\mu, \sigma^2)$, σ unknown,

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} = \frac{Z}{S/\sigma} \sim t_{(n-1)}$$ (T distribution with degree of freedom = n-1)

Formula: $(1-\alpha)\cdot 100\%$ confidence interval for μ with unknown σ^2

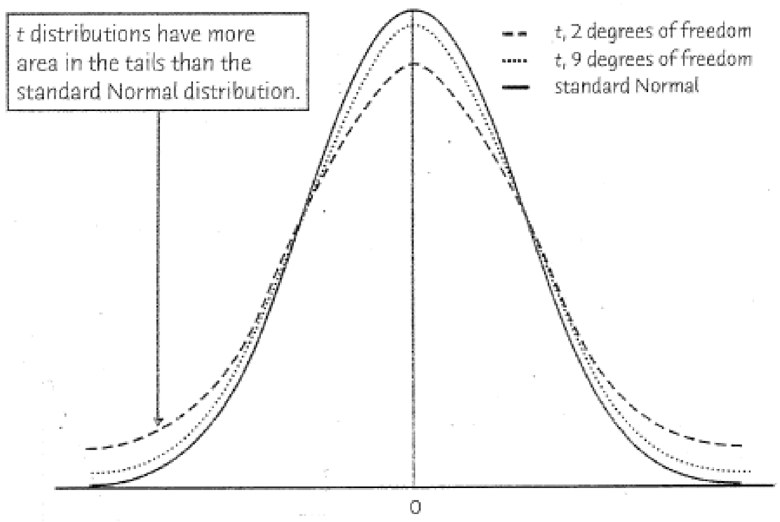
$$[\bar{X} - t_{\alpha/2,n-1}\cdot S/\sqrt{n}, \bar{X} + t_{\alpha/2,n-1}\cdot S/\sqrt{n}],$$

where $t_{\alpha/2,n-1}$ satisfies $\Pr(-t_{\alpha/2,n-1} \le T_{(n-1)} \le t_{\alpha/2,n-1}) = 1-\alpha$.

Note that when $X_i \sim^{iid} N(\mu, \sigma^2)$,

$$\Pr(-t_{\alpha/2,n-1} \le \frac{\bar{X}-\mu}{S/\sqrt{n}} \le t_{\alpha/2,n-1}) = 1-\alpha .$$

The density curves of the t distributions are similar in shape to the standard Normal curve. They are symmetric about 0, single-peaked, and bell-shaped.

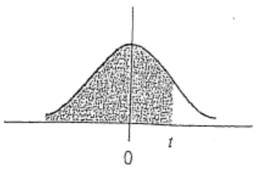


Remarks on T-table (compared with Z-table)

- The Z-table is made for N(0,1) and the T-table is made for T_v with selected values of v, the degree of freedoms.
- Since each row of the T-table represents a curve, the information has to be condensed. The columns are selected values of tail probability and the cell contains the cut-off value.

TABLE VI
Cumulative T distribution

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)
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	318.317	636.607
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	22.327	31.598
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055	3.930	4.318



$F(t) = P[T \le t]$

Examples

- $v = 5$: $\Pr(T_5 > 1.476) = 0.1$, $\Pr(T_5 > 2.015) = 0.05$, $\Pr(T_5 > 2.571) = 0.025$

- As v increases, $\Pr(T_5 > 2.015) = 0.05$, $\Pr(T_6 > 1.943) = 0.05$,

$$\Pr(T_{10} > 1.182) = 0.05, \quad \Pr(T_\infty > 1.645) = \Pr(Z > 1.645) = 0.05$$

- $\Pr(T_5 > 3) = p$, can we find p ?

The table tells us: $\Pr(T_5 > 2.571) = 0.025$ and $\Pr(T_5 > 3.365) = 0.01$

We only know that $0.01 < p < 0.025$ → T table does not contain enough information.

17.5 **Critical values.** What critical value t^* from Table C would you use for a confidence interval for the mean of the population in each of the following situations?

(a) A 95% confidence interval based on $n = 10$ observations.

(b) A 99% confidence interval from an SRS of 20 observations.

(c) An 80% confidence interval from a sample of size 7.

critical value = $t_{\alpha/2,n-1} = t^*$

a. $t_{0.025,9} = 2.262$; b. $t_{0.005,19} = 2.861$; **c.** $t_{0.1,6} = 1.440$

EXAMPLE 17.2 Healing of skin wounds

Let's look again at the biological study we saw in Example 14.3. We follow the four-step process for a confidence interval, outlined on page 361.

STATE: Biologists studying the healing of skin wounds measured the rate at which new cells closed a razor cut made in the skin of an anesthetized newt. Here are data from 18 newts, measured in micrometers (millionths of a meter) per hour (μm/h):[4]

29	27	34	40	22	28	14	35	26
35	12	30	23	18	11	22	23	33

This is one of several sets of measurements made under different conditions. We want to estimate the mean rate for comparison with rates under other conditions.

1	1 2 4
1	8
2	2 2 3 3
2	6 7 8 9
3	0 3 4
3	5 5
4	0

FIGURE 17.2 Stemp of the healing rates in Example 17.2.

To apply the T-interval, your need to check whether the normality assumption is plausible.

FORMULATE: We will estimate the mean rate μ for all newts of this species by giving a 95% confidence interval.

SOLVE: We must first check the conditions for inference.

- As in Chapter 14 (page 361), we are willing to regard these newts as an SRS from their species.
- The stemplot in Figure 17.2 does not suggest any strong departures from Normality.

We can proceed to calculation. For these data,

$$\bar{x} = 25.67 \text{ and } s = 8.324$$

The degrees of freedom are $n - 1 = 17$. From Table C we find that for 95% confidence, $t^* = 2.110$. The confidence interval is

$$\begin{aligned}\bar{x} \pm t^* \frac{s}{\sqrt{n}} &= 25.67 \pm 2.110 \frac{8.324}{\sqrt{18}} \\ &= 25.67 \pm 4.14 \\ &= 21.53 \text{ to } 29.81 \text{ micrometers per hour}\end{aligned}$$

Exercise: $n = 6$, $\bar{x}_{obs} = 89.01$, $s = 5.36$, $\alpha = 0.05$, $t_{0.025,5} = 2.571$

Rats eating oat bran. Exercise 17.2 gave the summary data (mean 89.01, standard error 5.36 mg/dl) for the cholesterol levels of 6 rats fed a diet enriched in fiber from oat bran. Cholesterol levels are usually approximately Normal, and we can regard these 6 rats as an SRS of the population of lab rats fed a diet enriched in oat fiber. Give the 95% confidence interval for the mean cholesterol level μ in this population.

$$[\bar{x}_{obs} - t_{\alpha/2,n-1} \cdot S/\sqrt{n}, \bar{x}_{obs} + t_{\alpha/2,n-1} \cdot S/\sqrt{n}] \quad \Rightarrow 89.01 \pm 2.571 \cdot 5.36/\sqrt{6}]$$

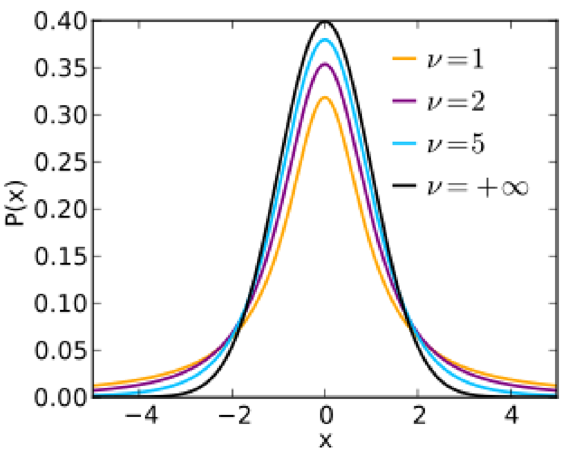
Summary of Confidence Intervals:

- Original population is normal
  - ■ σ^2 is known, use Z-interval (exact result)
  - ■ σ^2 is unknown and estimated by S^2, use T-interval (exact result)
- Original population is non-normal
  - ■ When $n \to \infty$, $S^2 \to \sigma^2$, $t_{\alpha/2,n} \approx z_{\alpha/2}$
  - ■ When n is not large and the population is skewed, using a T-interval to estimate μ is not accurate.

Comparison between Z-interval and T-interval given the same confidence level

Fix the confidence level $(1-\alpha)100\%$, $t_{\alpha/2,n-1} > z_{\alpha/2}$

Note: T_v has heavier tails than N(0,1)



If it happens that the value of σ and that of S are close, T-intervals will be <u>wider</u> than Z-intervals which is the price paid by the extra estimation of σ.

統計意義：區間寬度變長代表估計的不確定性增加，這是從“確知σ”到“估計σ”所付出的<u>代價</u>

Given the margin of error, find the smallest sample size (重要)

決定最小樣本數

EXAMPLE 15.6 How many observations?

Example 14.3 (page 361) reports a study of the healing rate of cuts in the skin of newts. We know that the population standard deviation is $\sigma = 8$ micrometers per hour. We want to estimate the mean healing rate μ for this species of newts within ± 3 micrometers per hour with 90% confidence. How many newts must we measure?

The desired margin of error is $m = 3$. For 90% confidence, Table C gives $z^* = 1.645$. We know that $\sigma = 8$. Therefore,

$$n = \left(\frac{z^*\sigma}{m}\right)^2 = \left(\frac{1.645 \times 8}{3}\right)^2 = 19.2$$

Because 19 newts will give a slightly larger margin of error than desired and 20 newts a slightly smaller margin of error, we must measure 20 newts. *Always round up to the next higher whole number when finding* n.

樣本數與誤差範圍

Sample size and margin of error. In the previous exercise, you calculated a 95% confidence interval for the population mean aspirin content μ based on this sample.

(a) Suppose that the SRS had only 4 tablets. Obviously the sample average would be specific to that particular sample. What would be the margin of error of a 95% confidence interval for the population mean μ in this case?

(b) Suppose that the SRS had 25 tablets. What would be the margin of error of a 95% confidence interval for μ?

(c) Compare the margins of error for samples of size 4, 10, and 25. How does increasing the sample size change the margin of error of a confidence interval?

$$\text{Margin of errors} = \begin{cases} z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}} & \text{if } \sigma \text{ is known} \\ t_{\alpha/2,n-1}\dfrac{s}{\sqrt{n}} & \text{if } \sigma \text{ is unknown} \end{cases}$$

a. $n = 4, z_{0.025} = 1.96$

b. $n = 25, z_{0.025} = 1.96$

c. As the sample size increases, the margin of error (the length of the confidence interval) decreases.

Summary of the Distribution Properties of $\bar{X}$

Case 1: σ is known

Fact 1a: $X_i \sim^{iid} N(\mu,\sigma^2)$,

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = Z \underset{\text{exactly for any } n}{\sim} N(0,1)$$

Fact 1b: $X_i \sim^{iid}$ non-normal with $E(X_i)=\mu$ and $Var(X_i)=\sigma^2,\ n\to\infty$,

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \underset{\substack{n\to\infty \\ \text{approximately}}}{\sim} N(0,1) \quad \text{(by the \textit{Central Limit Theorem})}$$

Case 2: σ is unknown and estimated by

$$S=\sqrt{\frac{\sum_{i=1}^{n}(X_i-X)^2}{n-1}} \quad (分母一定為\ n-1)$$

Fact 2a: $X_i \sim^{iid} N(\mu,\sigma^2)$

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} \underset{\substack{\forall n \\ \text{exactly}}}{\sim} T_{(n-1)} \quad \text{(T distribution with } n-1 \text{ degree of freedom)}$$

Fact 2b: $X_i \sim^{iid}$ non-normal with $E(X_i)=\mu$ and $Var(X_i)=\sigma^2,\ n\to\infty$

$$S\to\sigma \quad \& \quad \frac{\bar{X}-\mu}{S/\sqrt{n}} \underset{\substack{n\to\infty \\ \text{approximately}}}{\sim} T_{\infty}=Z$$

Case 3: σ is unknown and n is small

Fact 5a: If the data $(X_1,\ldots,X_n)$ are *roughly symmetric*, then

the distribution of $\frac{\bar{X}-\mu}{S/\sqrt{n}}$ may be approximated by $T_{(n-1)}$ （近似不會太離譜）

Fact 5b: If the population distribution is skew, we can NOT use T intervals.

Story: Neyman and the Confidence Trick

Neyman (1894-1981)

- a Polish mathematician and statistician who spent the first part of his professional career in Europe and the second part at the University of California, Berkeley. (In 1938, at the age of 44, Jerzy Neyman left the Department of Applied Statistics at University College, London and arrived in Berkeley to take up a position as Professor of Mathematics.)
- Neyman first introduced the modern concept of a confidence interval into statistical hypothesis testing and co-revised Ronald Fisher's null hypothesis testing (in collaboration with Egon Pearson).
- The Department of Statistics at UC Berkeley was founded by Jerzy Neyman. *Neyman is considered to be, with R.A Fisher, one of the two founders of modern statistics.*

"統計改變了世界": 信賴或是詐騙?

Story: When the idea of confidence intervals was first proposed in 1934

Bowley actually started the argument by wondering aloud whether confidence intervals were just *"a confidence trick"!*
He asked, *"Does [a confidence interval] really lead us to what we need—the chance that within the universe which we are sampling the proportion is within these certain limits? I think it does not.* I think we are in the position of knowing that either an improbable event had occurred or that the proportion in the population is within these limits… The statement of the theory is not convincing, and until I am convinced *I am doubtful of its validity.*

http://mathmuseum.tw/wp-content/pdf/034_1230088632.pdf （中文摘要）

第十一章談到尼曼—皮爾生假設檢定理論的發展，躋身統計學的巔峰地位，但遭遇到費雪一再攻擊。尼曼在加州柏克萊大學分校建立統計系，到 1981 年他去世時，已經把該系變成全世界最重要的統計學術殿堂之一。他常利用系上下午茶時間，催促學生或同事談談自己最新的研究與進展，又特別提攜女學生，鼓勵她們往學術生涯發展，而他的女弟子也在天文學、動物學、流行病學…等研究有傑出的貢獻。第十二章由愛滋病（AIDS）研究談起估計，區間估計（interval estimate）現在已經普及於幾乎所有統計分析，如民意調查…等。1934 年，尼曼的演講論文主題是抽樣調查分析，他在附錄中提出信賴區間的方法。但一開始就被大會主席批評，他覺得尼曼博士所謂的「信賴」，可能只是「一種獲取對方信賴之後的詐騙」。尼曼取巧避談機率，而將他創造出來的東西稱為「信賴區間」。

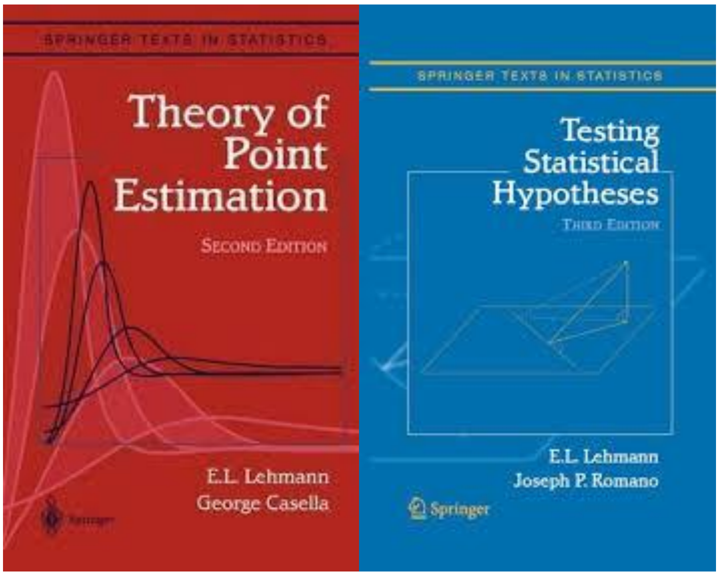
Story: RA Fisher and Neyman

We learned of the feud between Fisher and Pearson, and of Fisher's attacks on Neyman's work.
The latter was really brought to life in the story of Fisher attending a presentation by Neyman in France and how Neyman prepared himself for tough questions from Fisher only to find Fisher could not speak French.

https://errorstatistics.com/2018/02/24/deconstructing-the-fisher-neyman-conflict-wearing-fiducial-glasses-continued-2/

統計博士生最重要的兩本書

(Leymann 是 Neyman 的學生，任教於 UC Berkeley)



Two Philosophical Debate: "Frequentist" and Bayesian Approaches

https://www.quora.com/What-is-the-difference-between-frequentist-confidence-intervals-and-Bayesian-credible-intervals

Frequentist approach:

Frequentist confidence intervals treat the parameter as fixed and the sample as random. The bounds of the confidence interval are statistics of the dataset. The bounds have the property that 95% of the time under repeated redraws of the data, they will contain the true, fixed value.

Bayesian approach:

Bayesian credible intervals treat the parameter θ as random and the data as fixed and creates a distribution of reasonable values for the parameter, given the observed data.

先參考 ppt

Comparison between frequentist and Bayesian statistics (will not be on the exam)

Random sample $(X_1,...,X_n)$

Assume that $X_i \overset{iid}{\sim} f_\theta(x|\theta)$

However the value of θ is unknown in reality.

*** Frequentist approach:** view θ as a unknown constant

Likelihood function:

$$L(\theta) = \Pr(X_1 = x_1,...,X_n = x_n | \theta) = \prod_{i=1}^{n} f(x_i|\theta)$$

= the plausibility of a model parameter value θ, given specific observed data $(x_1,...,x_n)$.

Remarks:

- 機率: 給定 θ, 問 $\Pr(X = x|\theta)$
- 統計: 給一組資料 $(X_1,...,X_n) = (x_1,...,x_n)$, 看這組資料最支持哪個值的 θ
- Maximum likelihood estimator (MLE) of θ (最大概似估計量)

→ find $\hat{\theta}$ maximize $L(\theta)$

Bayesian approach: view θ as random

- $p(\theta)$: prior distribution (先驗分布)
  - Describe the probability distribution that would express one's beliefs about the parameter before data evidence is taken into account.
- Likelihood as the evidence supported by the data:

$$p(x_1,\ldots,x_n \mid \theta) \quad = \prod_{i=1}^{n} f(x_i \mid \theta)$$

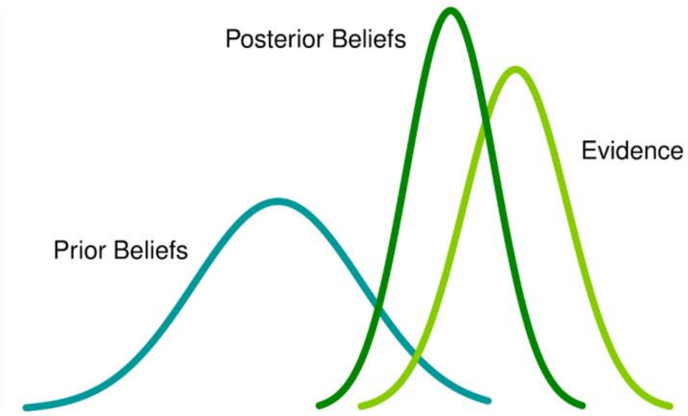
- Posterior distribution (後驗分布)

$$p(\theta \mid x_1,\ldots,x_n) = \frac{p(x_1,\ldots,x_n,\theta)}{p(x_1,\ldots,x_n)} = \frac{p(x_1,\ldots,x_n \mid \theta)P(\theta)}{\int_\theta p(x_1,\ldots,x_n \mid \theta)P(\theta)}$$

→ related to Bayes' theorem

Estimation of θ

- Posterior mean: $E(\theta \mid x_1,\ldots,x_n) = \int_\theta \theta p(\theta \mid x_1,\ldots,x_n)$
- Posterior mode: maximize $p(\theta \mid x_1,\ldots,x_n)$



橫軸: value of θ 縱軸: probability

Review: Bayes' theorem:

Given that $A_1,\ldots A_n$ "partition" (切割) the sample space,

$$\Pr(A_j \mid B) = \frac{\Pr(B \cap A_j)}{\mathrm{P(B)}} = \frac{\Pr(B \cap A_j)}{\sum_{k=1}^{K} \Pr(B \cap A_k)} = \frac{\Pr(B \mid A_j)\Pr(A_j)}{\sum_{k=1}^{K} \Pr(B \mid A_k)\Pr(A_k)}$$

Note: "partition" means that $A_i \cap A_j = \phi$ for $i \neq j$, and $A_1 \cup \ldots A_n = S$.

Criticisms on Bayesian inference by frequentists

1. The prior $p(\theta)$ is too subjective → change the prior, change the answer

Remedy: empirical Bayes, non-informative prior

- It is difficult to compute $p(\theta \mid X_1, \ldots, X_n)$ or posterior mean (mode)

Remedy: Bayesian statisticians developed useful computation methods such as

MCMC, metropolis, Gibbs sampling, ...)

貝氏定理淺談

壹、前言

貝氏定理(*Bayes' theorem*)是條件機率中的一個重要內容，在歷史上，貝氏定理被歸之於十八世紀英國長老派教會牧師貝氏(*Reverend Thomas Bayes*, 1702-1761)所提出。事實上，貝氏並未明確給出定理的公式，只是在一篇論文〈關於解決機遇論的一個問題 "*An Essay towards solving a Problem in the Doctrine of Chances.*"〉中，討論其相關問題，法國著名的數學家拉普拉斯(*Laplace* 1749-1827)，才是第一位給出明確公式的人。甚至於貝氏在生前並沒有公布他的發現，而是由他的一位牧師朋友，也是數學家的普萊斯 (*Richard Price*)，於 1764 年將貝氏的論文發表在英國皇家學會的會刊上。

http://b014.hchs.hc.edu.tw/ezfiles/14/1014/img/164/409339189.pdf

MODERN SCIENCE AND THE BAYESIAN-FREQUENTIST CONTROVERSY
Bradley Efron
The 250-year debate between Bayesians and frequentists is unusual among philosophical arguments in actually having important practical consequences.Whenever noisy data is a major concern, scientists depend on statistical inference to pursue nature's mysteries. **19th Century science was broadly Bayesian in its statistical methodology, while frequentism dominated 20th Century scientific practice.** This brings up a pointed question: which philosophy will predominate in the 21st Century? One thing is already clear – statistical inference will pay an increased role in scientific progress as scientists attack bigger, messier problems in biology, medicine, neuroscience, the environment, and other fields that have resisted traditional deterministic analyses. This talk argues that a combination of frequentist and Bayesian thinking will be needed to deal with the massive data sets scientists are now bringing us. Three examples are given to suggest how such combinations might look in practice. A large portion of the talk is based on my presidential address to the American Statistical Association and a related column in Amstat News.
https://pdfs.semanticscholar.org/575b/9274b3ec2b4db5e2baecfd1949be1d7c800b.pdf