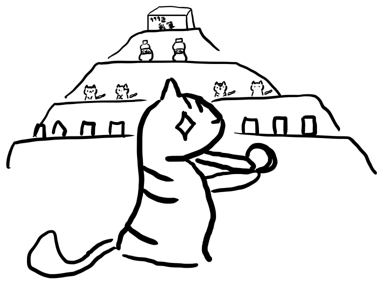
# 第 9 章
# 區間估計





本章將開始進入推論統計的應用。在點估計的章節中, 雖然可計算估計式的不偏性與有效性, 但我們僅能知道該估計式好不好, 卻無法得知該估計式的精確度。舉例來說, 比如球隊中的隊員是我們手邊有的未知參數估計式, 透過點估計理論可以幫我們找出隊上最棒的球員, 但到底該球員的表現有多好卻未知。因此本章將介紹可以直接看出精確度的參數估計方法: 區間估計。此外, 本章將介紹實際區間估計式與近似區間估計式的概念。



## —— 9.1 區間估計式的定義與建構 ——

**定義 9.1** 區間估計式 (*interval estimator*)
給定 $\{X_i\}_{i=1}^n \overset{i.i.d.}{\sim} f(x,\theta)$, 若統計量 $L(X_1, X_2, \ldots, X_n)$ 與 $U(X_1, X_2, \ldots, X_n)$ 滿足

1. $L(X_1, X_2, \ldots, X_n) < U(X_1, X_2, \ldots, X_n)$
2. $1-\alpha = P[L(X_1, X_2, \ldots, X_n) \le \theta \le U(X_1, X_2, \ldots, X_n)]$, 其中 $0 < \alpha < 1$。

在**抽樣之前**, 稱隨機區間 (*random interval*)

$$[L(X_1, X_2, \ldots, X_n), U(X_1, X_2, \ldots, X_n)]$$

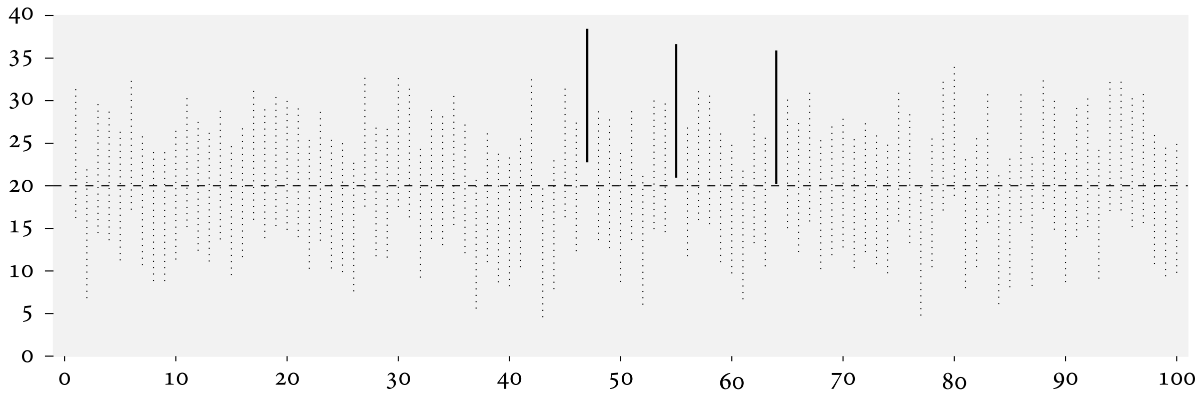
爲未知參數 $\theta$ 的 $100(1-\alpha)\%$ 區間估計式, 其中 $1-\alpha$ 爲區間估計式包含 $\theta$ 之「**事前機率**」。
在**抽樣之後**, 將樣本實現值 $\{x_i\}_{i=1}^n$ 代入, 則稱

$$[L(x_1, x_2, \ldots, x_n), U(x_1, x_2, \ldots, x_n)]$$

爲未知參數 $\theta$ 的 $100(1-\alpha)\%$ 區間估計值, 其中 $1-\alpha$ 爲區間估計值包含 $\theta$ 之「**事後信心**」。

1. $\alpha$ 的常見設定值爲 $\alpha = 0.1, 0.05, 0.01$。
2. $1-\alpha$ 又稱作信心水準 (confidence level) 或信心參數 (confidence coefficient)。
3. 特別注意到, 由於 $L$ 與 $U$ 是隨機樣本的函數 (例如包含 $\bar{X}$), 因此給定的樣本不同, $L$ 與 $U$ 的實現值 (即 $\bar{X}$ 的實現值 $\bar{x}$) 也不同。是以 $L$ 與 $U$ 是隨機變數, 故稱作隨機區間。
4. $1-\alpha$ 指的是在抽取樣本前, 隨機區間包含未知參數之**事前機率**, 而非隨機區間的實現值包含未知參數的**事後機率**。將樣本實現值代入區間估計式後, 所得出的隨機區間實現值即爲**區間估計值** (interval estimates) 或稱**信賴區間** (confidence interval, CI)。由於區間估計值是區間估計式的實現值, 並不是隨機變數, 因此包含未知參數的機率不是 0 (區間中不包含未知參數) 就是 1 (區間中包含未知參數)。因此**在抽樣之後, 1−α 改稱為此區間估計值 (或信賴區間) 包含未知參數的信心**, 而非機率。
5. 有些教科書將信賴區間同時表示區間估計式與區間估計值。但此舉易造成觀念上的混淆, 因爲在區間估計式中 $1-\alpha$ 爲機率, 但在區間估計值中 $1-\alpha$ 爲信心, 不應混爲一談。
6. 信賴區間是由波蘭統計學家 Jerzy Neyman (1894-1981) 提出的概念。但有趣的是, 信心水準不是機率而是信心的說法, 讓當時學者 (包含 Ronald Fisher 在內) 都難以接受。

圖 9.1: 區間估計式圖示




7. 舉例來說, 給定 $\{X_i\}_{i=1}^n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$, 其中 $\sigma$ 爲已知常數。若以下成立:

$$\underbrace{1-\alpha}_{\text{事前涵蓋機率}} = P\left( \underbrace{\bar{X} - Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}}_{L(X_1, X_2, \ldots, X_n)} \leq \underbrace{\mu}_{\text{未知參數}} \leq \underbrace{\bar{X} + Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}}_{U(X_1, X_2, \ldots, X_n)} \right)$$

(a) $\left[\bar{X} \pm Z_{\alpha/2} \times \sigma/\sqrt{n}\right]$ 稱爲 $\mu$ 之 $100(1-\alpha)\%$ 區間估計式。其中因爲 $\bar{X}$ 是隨機變數, 故此爲隨機區間, 且其包含未知參數 $\mu$ 之事前機率爲 $1-\alpha$。意即若重覆抽樣分別計算其信賴區間, 則約有 $100(1-\alpha)\%$ 的信賴區間會包含未知參數 $\mu$。

(b) 若已知 $\sigma = 16$, 抽樣結果爲 $n = 16$, $\bar{X} = 20$, 且 $\alpha = 0.05$, $Z_{0.025} = 1.96$, 代入可得到 $\mu$ 之 95% 區間估計值爲

$$[L(x_1, x_2, \ldots, x_n), U(x_1, x_2, \ldots, x_n)] = \left[20 \pm 1.96\frac{16}{\sqrt{16}}\right] = [12.16, 27.84]$$

則稱 $[12.16, 27.84]$ 爲 $\mu$ 之 95% 區間估計值或信賴區間。**此區間包含 $\mu$ 之信心水準為 95%, 但信心不是機率, 不能解釋成 $P(\mu \in [12.16, 27.84])=0.95$**。若真實的 $\mu = 25$, 則此區間包含 $\mu$ 之機率爲 1; 若真實的 $\mu = 30$, 則此區間包含 $\mu$ 之機率爲 0。

(c) **信心水準並不是重覆建構的信賴區間包含未知參數的樣本比例。**例如硬幣出現正面的機率是 0.5, 但不保證投擲 2 次其中一定會有 1 次正面。圖 9.1 給予一個實際範例。給定 $\{X_i\}_{i=1}^{16} \sim N(20, 16^2)$, 重覆抽取 100 組隨機樣本, 分別建構 $\mu = 20$ 之 95% 區間估計值並將其繪於圖 9.1。令 $Y$ 爲這 100 組區間估計值中, **未包含**真實的 $\mu$ 之組數, 則 $Y \sim Binomial(100, 0.05)$, 且 $E(Y) = 5$。意即以平均來說, 會有 5 個區間估計式未包含真實的 $\mu$, 而此例中可看出有 3 組未包含 $\mu$。值得一提的是, 當重覆抽樣數越來越大, 根據弱大數法則, 區間估計值未包含 $\mu$ 的比例將越來越接近 $\alpha$。

(d) 由於未知參數是一固定常數, 因此解讀區間估計式之意涵時, 盡量避免用未知參數「落入」(fall) 區間估計式的字眼, 應該說區間估計式「包含」(contain) 未知參數。

**例題 9.1 區間估計式基本觀念**

A 95% confidence interval of a population mean is $[\underline{\mu}, \overline{\mu}]$. Which of the following statement is the most correct one?

(A) If random samples were drawn again and again, with $\underline{\mu}$ and $\overline{\mu}$ computed each time, the population mean would lie in the interval $[\underline{\mu}, \overline{\mu}]$ 95% of the samples.

(B) If $\underline{\mu} = 0.1$ and $\overline{\mu} = 0.5$, the population mean would lie in the interval $[0.1, 0.5]$ with 95% probability.

(C) If $\underline{\mu} = 0.05$ and $\overline{\mu} = 0.55$, the population mean might still lie outside the interval $[0.05, 0.55]$.

(D) If $\underline{\mu} = 0.03$ and $\overline{\mu} = 0.56$, the population mean would lie in the interval $[0.03, 0.56]$ greater than 95% probability.

(E) (A) and (C). 《110 政大國貿》

《解》

選 (E)。

- (B) 與 (D) 錯在區間估計值中的信賴係數意涵是事前包含的機率, 不代表區間估計值包含未知參數的機率。
- 而只要信賴係數不是 100%, 真實母體參數就有可能在位於區間估計值之外。

**例題 9.2 區間估計式基本觀念之其二**

We are interested in the amount of time per week spending on Internet among college students. Let $\mu$ denote the population mean. Based on a randomly selected sample, the sample mean is 31 hours and the 95% confidence interval is $(28, 34)$. Which of the following statements is true?

(A) The probability that $\mu$ lies in $(28, 34)$ is 0.95.

(B) The probability that the sample mean lies in $(28, 34)$ is 0.95.

(C) About 95% of similarly constructed confidence interval will cover the sample mean.

(D) None of the above. 《105 政大財管》

《解》

選 (D)。

- (A) 與 (B) 錯在 0.95 應是區間估計值包含 $\mu$ 之信心而非機率。
- (C) 錯在應該將 sample mean 改為 population mean。

**定義 9.2** 誤差邊界 (*margin of error*)
若參數 $\theta$ 的點估計式爲 $\hat{\theta}$, 且其 $100(1-\alpha)\%$ 區間估計式爲 $\left[\hat{\theta} \pm \varepsilon\right]$。則 $\varepsilon$ 稱作誤差邊界。

1. 若欲估計的參數爲 $\theta = \mu$, 其點估計式爲 $\hat{\theta} = \bar{X}$。若欲估計的參數爲 $\theta = \sigma^2$, 其點估計式爲 $\hat{\theta} = S^2$。若欲估計的參數爲 $\theta = \mu_1 - \mu_2$, 則其點估計式爲 $\hat{\theta} = \bar{X}_1 - \bar{X}_2$。
2. 若區間估計式可寫成 $\left[\hat{\theta} \pm \varepsilon\right]$ 之形式, 則區間估計式的中點爲 $\hat{\theta}$, 且誤差邊界是區間估計式之寬度的一半。欲估計參數爲母體平均的區間估計式大多屬於這一類型。
3. 與母體變異數相關的未知參數, 例如 $\theta = \sigma^2, \sigma_1^2/\sigma_2^2$ 的區間估計式之形式爲 $\left[c_1\hat{\theta}, c_2\hat{\theta}\right]$, 其中 $c_1 < 1$, $c_2 > 1$。由於區間的中點並不是 $\hat{\theta}$, 因此在這個狀況下不定義誤差邊界。
4. 誤差邊界可以視作抽樣誤差 (sampling error) 的大小。誤差邊界越小 (或區間估計式的寬度越窄), 表示對於未知參數的區間估計越精確。隨著樣本數 $n$ 增加, 區間估計式的寬度會越來越窄, 估計會越精確。

**定義 9.3** 樞紐量 (*pivotal quantity*)
建構 $\theta$ 之區間估計式時, 若統計量 $\varphi$ 滿足以下 2 個條件, 則 $\varphi$ 可作爲未知參數 $\theta$ 之樞紐量:

1. $\varphi$ 包含未知參數 $\theta$,
2. $\varphi$ 的抽樣分配已知, 且其分配之參數與 $\theta$ 無關。

1. 爲了強調 $\varphi$ 爲包含未知參數 $\theta$ 的統計量, 亦可用 $\varphi(\theta, X_1, X_2, \ldots, X_n)$ 表示之。
2. 舉例來說, 給定 $\{X_i\}_{i=1}^n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$,
   (a) 若 $\sigma^2$ 已知, $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ 可作爲 $\mu$ 的樞紐量;
   (b) 若 $\sigma^2$ 未知, $\dfrac{\bar{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$ 可作爲 $\mu$ 的樞紐量。
   (c) 若 $\mu$ 已知, $\dfrac{\sum_{i=1}^n (X_i-\mu)^2}{\sigma^2} \sim \chi^2(n)$ 可作爲 $\sigma^2$ 的樞紐量。
   (d) 若 $\mu$ 未知, $\dfrac{\sum_{i=1}^n (X_i-\bar{X})^2}{\sigma^2} \sim \chi^2(n-1)$ 可作爲 $\sigma^2$ 的樞紐量。
3. 樞紐量的分配形態將會影響區間估計式的形式。若未知參數 $\theta$ 的樞紐量爲對稱分配 (例如標準常態分配或 $t$ 分配), 則區間估計式可整理成 $\left[\hat{\theta} \pm \varepsilon\right]$ (對稱於 $\hat{\theta}$ 的區間) 的形式; 若未知參數 $\theta$ 的樞紐量不是對稱分配 (例如卡方分配或 $F$ 分配), 則區間估計式可整理成 $\left[c_1\hat{\theta}, c_2\hat{\theta}\right]$ (不對稱於 $\hat{\theta}$ 的區間)。但無論區間估計式是否爲對稱於 $\hat{\theta}$ 的區間, 由於 $\hat{\theta}$ 是對於 $\theta$ 的最佳預測, 因此 $\hat{\theta}$ 必然會在區間估計式之中。

**例題 9.3 區間估計式基本概念之其三**

(3%) A random sample $(x_i)$ is taken from the population whose mean is $\mu$. Which quantity is guaranteed to be in the 96% confidence interval? 《105 中興企研》

(A) 0 (B) $\bar{x}$ (C) $\mu$ (D) 0.96

《解》

選 (B)。點估計式會包含在區間估計式中。

**例題 9.4 區間估計式基本觀念之其四**

A confidence interval is defined as:

(A) a point estimate plus or minus a specific confidence level.

(B) a lower and upper confidence limit associated with a specific level of confidence.

(C) an interval that has a 95% probability of containing the population parameter.

(D) a lower and upper confidence limit that has a 95% probability of containing the population parameter. 《108 成大經研》

《解》

選 (B)。信賴區間是區間估計值 (將樣本結果代入區間估計式中, 爲實現值), 包含未知參數的機率不是 1 就是 0。

**例題 9.5 區間估計與點估計的比較**

The chief distinction between a point estimate and an interval estimate is that:

(A) an interval estimate is always correct (though it may not be precise) but a point estimate is usually incorrect

(B) a point estimate is always more useful than an interval estimate

(C) an interval estimate indicates the precision of the estimate while a point estimate does not

(D) a point estimate does not differ from an interval estimate except in minor respects

《106 成大財金》

《解》

選 (C)。

- (A) 錯在區間估計式也可能未包含未知參數, 因此不是永遠都是正確的。
- (B) 錯在一般而言, 區間估計相較於點估計多了抽樣誤差的資訊, 因此會比較有用。
- (D) 錯在區間估計式的概念與點估計式的概念明顯不同。

**性質 9.1** 建構區間估計式的步驟

1. 決定信心水準 $1-\alpha$,

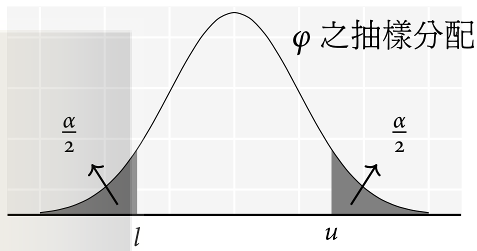
2. 找尋未知參數 $\theta$ 的樞紐量 $\varphi$,

3. 利用 $\varphi$ 的抽樣分配決定臨界值 $l, u$ 使得

$$1-\alpha = P[l \le \varphi(\theta, X_1, X_2, \ldots, X_n) \le u]$$

4. 移項整理使得 $\begin{aligned} 1-\alpha &= P[l \le \varphi(\theta, X_1, X_2, \ldots, X_n) \le u] \\ &= P[L(X_1, X_2, \ldots, X_n) \le \theta \le U(X_1, X_2, \ldots, X_n)] \end{aligned}$

1. 一般來說, 選取的 $l$ 與 $u$ 滿足




   (a) $P[\varphi(\theta, X_1, X_2, \ldots, X_n) \le l] = \alpha/2$

   (b) $P[\varphi(\theta, X_1, X_2, \ldots, X_n) \ge u] = \alpha/2$

   其值通常透過查臨界值表而來, 如右圖所示。

2. 多數情況下 $\varphi$ 不需自行推導, 但需知道分配參數在各種情境下對應使用的 $\varphi$ 為何。

3. 移項整理過程通常都有遞減轉換 (倒數或者同乘負號), 因此在移項整理過後:

   (a) $L(X_1, X_2, \ldots, X_n)$ 會包含 $u$, 即原先抽樣分配的右尾 (上界) 臨界值。

   (b) $U(X_1, X_2, \ldots, X_n)$ 會包含 $l$, 即原先抽樣分配的左尾 (下界) 臨界值。

4. 若欲建構單邊區間估計式, 即找尋 $U(X_1, X_2, \ldots, X_n)$ 或 $L(X_1, X_2, \ldots, X_n)$ 使得

   $1-\alpha = P\left[\underline{L} \le \theta \le U(X_1, X_2, \ldots, X_n)\right]$ 或 $1-\alpha = P\left[L(X_1, X_2, \ldots, X_n) \le \theta \le \overline{U}\right]$

   其中 $\underline{L}$ 為 $\theta$ 之參數空間的下界, $\overline{U}$ 為 $\theta$ 之參數空間的上界。舉例來說,

   (a) 母體均數 $\mu$ 的參數空間為 $(-\infty, \infty)$, 母體變異數 $\sigma^2$ 的參數空間為 $[\,0, \infty\,)$。

   (b) 母體比例 $p$ 的參數空間為 $[0, 1]$。

5. 未知參數 $\theta$ 之點估計式 $\hat{\theta}$ 是對於未知參數最佳之估計, 故應包含在區間估計式中。舉例來說, 若利用樣本平均 $\bar{X}$ 對於母體均數 $\mu$ 進行區間估計:

   (a) 區間估計值包含 $\mu$ 的機率不是 0 就是 1。

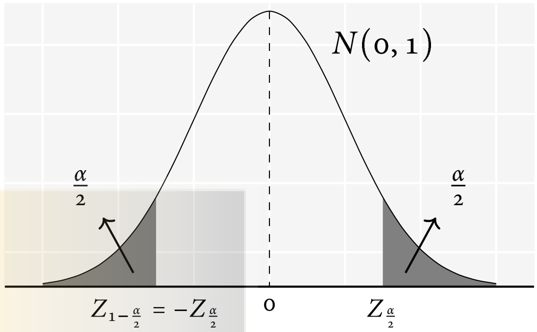
   (b) $\bar{X}$ 之實現值 $\bar{x}$ 必定在區間估計值中, 即區間估計值包含 $\bar{x}$ 的機率為 1。

**性質 9.2** 常態母體均數 $\mu$ 的區間估計式

給定 $\{X_i\}_{i=1}^n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$, 則

1. 若 $\sigma^2$ 已知, $\mu$ 的 $100(1-\alpha)\%$ 區間估計式為 $\left[\bar{X} \pm Z_{\frac{\alpha}{2}} \dfrac{\sigma}{\sqrt{n}}\right]$
2. 若 $\sigma^2$ 未知, $\mu$ 的 $100(1-\alpha)\%$ 區間估計式為 $\left[\bar{X} \pm t_{\frac{\alpha}{2}}(n-1) \dfrac{S}{\sqrt{n}}\right]$

1. 若 $\sigma^2$ 已知, 則




$$\varphi = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

給定

$$1-\alpha = P\left(Z_{1-\frac{\alpha}{2}} \le \varphi \le Z_{\frac{\alpha}{2}}\right)$$

其中標準常態分配對稱於 0, 如右圖所示。因此 $Z_{1-\alpha/2} = -Z_{\alpha/2}$, 將此結果代入可得到

$$\begin{aligned}
1-\alpha &= P\left(-Z_{\frac{\alpha}{2}} \le \varphi \le Z_{\frac{\alpha}{2}}\right) = P\left(-Z_{\frac{\alpha}{2}} \le \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le Z_{\frac{\alpha}{2}}\right) && \xrightarrow{\text{同乘 } \sigma/\sqrt{n}} \\
&= P\left(-Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le \bar{X}-\mu \le Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) && \xrightarrow{\text{同減 } \bar{X}} \\
&= P\left(-\bar{X}-Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le -\mu \le -\bar{X}+Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) && \xrightarrow{\text{同乘 } -1} \\
&= P\left(\bar{X}-Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X}+Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) \\
&= P\left(\mu \in \left[\bar{X} \pm Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]\right) \qquad (9.1)
\end{aligned}$$

2. 由式 (9.1) 可發現, 區間估計式的寬度與 $\bar{X}$ 無關, 僅與誤差邊界 $Z_{\alpha/2} \times \sigma/\sqrt{n}$ 有關, 且
   (a) 當 $n$ 越大, 則區間估計式的區間越窄。
   (b) 當 $\sigma$ 越大, 則區間估計式的區間越寬。
   (c) 當 $\alpha$ 越大, 則 $Z_{\alpha/2}$ 越小。例如 $Z_{0.05} = 1.645 < Z_{0.025} = 1.96$。此時區間估計式的區間越窄; 反之, 當 $1-\alpha$ 越大, 區間越寬。

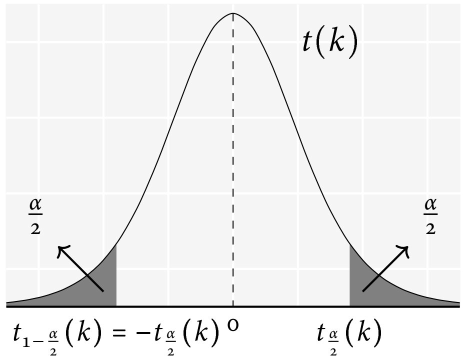
   $n$ 越大, 表示對於未知參數的估計越精準, 因此區間越窄。$1-\alpha$ 越大, 代表包含未知參數的信心越大, 因此在其他條件不變下, 應考慮較寬之區間。

3. 若 $\sigma^2$ 未知，則

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

給定

$$1-\alpha = P\left[t_{1-\frac{\alpha}{2}}(n-1) \le \varphi \le t_{\frac{\alpha}{2}}(n-1)\right]$$



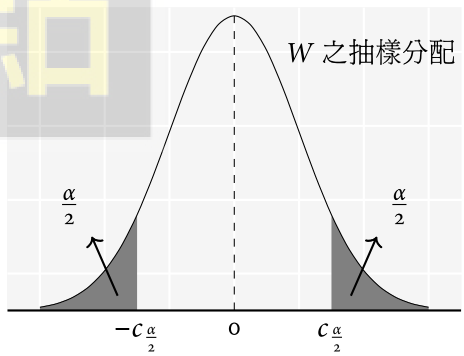

其中 $t$ 分配對稱於 0，如右圖所示。同理，可得知 $t_{1-\alpha/2}(n-1) = -t_{\alpha/2}(n-1)$，因此

$$\begin{aligned}
1-\alpha &= P\left(-t_{\frac{\alpha}{2}}(n-1) \le \varphi \le t_{\frac{\alpha}{2}}(n-1)\right) \\
&= P\left(-t_{\frac{\alpha}{2}}(n-1) \le \frac{\bar{X}-\mu}{S/\sqrt{n}} \le t_{\frac{\alpha}{2}}(n-1)\right) && \xrightarrow{\text{同乘 } S/\sqrt{n}} \\
&= P\left(-t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}} \le \bar{X}-\mu \le t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\right) && \xrightarrow{\text{同減 } \bar{X}} \\
&= P\left(-\bar{X}-t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}} \le -\mu \le -\bar{X}+t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\right) && \xrightarrow{\text{同乘 } -1} \\
&= P\left(\bar{X}-t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}} \le \mu \le \bar{X}+t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\right) \\
&= P\left(\mu \in \left[\bar{X} \pm t_{\frac{\alpha}{2}}(n-1)\frac{S}{\sqrt{n}}\right]\right)
\end{aligned}$$

4. 令 $W$ 為一對稱為 0 之分配 (標準常態或 $t$ 分配)，且臨界值之圖示如右。若 $\theta$ 之樞紐量滿足

$$\frac{\hat{\theta}-\theta}{SE(\hat{\theta})} \sim W$$

其中 $SE(\hat{\theta})$ 為 $\hat{\theta}$ 之標準誤 (standard error)，則




$$\begin{aligned}
1-\alpha &= P\left(-c_{\frac{\alpha}{2}} \le \frac{\hat{\theta}-\theta}{SE(\hat{\theta})} \le c_{\frac{\alpha}{2}}\right) = P\left(\hat{\theta}-c_{\frac{\alpha}{2}} \times SE(\hat{\theta}) \le \theta \le \hat{\theta}+c_{\frac{\alpha}{2}} \times SE(\hat{\theta})\right) \\
&= P\left(\theta \in \left[\hat{\theta} \pm c_{\frac{\alpha}{2}} \times SE(\hat{\theta})\right]\right)
\end{aligned}$$

意即 $\theta$ 之 $100(1-\alpha)\%$ 區間估計式為 $\left[\hat{\theta} \pm c_{\alpha/2} \times SE(\hat{\theta})\right]$。舉例來說，

(a) 給定 $\{X_i\}_{i=1}^n \overset{i.i.d.}{\sim} N(\mu,\sigma^2)$，其中 $\sigma^2$ 已知，則

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

其中 $\sigma^2$ 已知下, $SE(\bar{X}) = \sigma/\sqrt{n}$, 因此 $\mu$ 之 $100(1-\alpha)\%$ 區間估計式為

$$\left[\bar{X} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$$

(b) 給定 $\{X_i\}_{i=1}^{n} \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$, 其中 $\sigma^2$ 未知, 則

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

其中 $\sigma^2$ 未知下, $SE(\bar{X}) = S/\sqrt{n}$, 因此 $\mu$ 之 $100(1-\alpha)\%$ 區間估計式為

$$\left[\bar{X} \pm t_{\frac{\alpha}{2}}(n-1) \frac{S}{\sqrt{n}}\right]$$

5. 常態母體當 $\sigma$ 未知時, 若樣本數足夠大時 (自由度超過附表提供的最大值時), 由於 $t(n-1) \xrightarrow{d} N(0,1)$, 故給定 $n$ 足夠大下, $t_{\alpha/2}(n-1) \simeq Z_{\alpha/2}$。因此

$$\left[\bar{X} \pm t_{\frac{\alpha}{2}}(n-1) \frac{S}{\sqrt{n}}\right] \simeq \left[\bar{X} \pm Z_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}\right]$$

6. 若母體數 $N$ 有限 ($n/N \geq 0.05$), 則誤差邊界需考慮變異數的有限母體校正因子, 意即

(a) 若 $\sigma^2$ 已知, 有限母體下, $\mu$ 之 $100(1-\alpha)\%$ 區間估計式為

$$\left[\bar{X} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}\right]$$

(b) 若 $\sigma^2$ 未知, 有限母體下, $\mu$ 之 $100(1-\alpha)\%$ 區間估計式為

$$\left[\bar{X} \pm t_{\frac{\alpha}{2}}(n-1) \frac{S}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}\right]$$

**例題 9.6 誤差邊界之基本概念**

(5%) Assume $X_1, \ldots, X_n$ are iid random variables with distribution $N(\mu, \sigma^2)$, We construct a confidence interval for $\mu$, then the length of the confidence interval for $\mu$ would decrease if

(A) standard error for $\mu$ increases.

(B) $\sigma^2$ increases.

(C) $n$ Increases.

(D) sample mean $\bar{X}$ increases. 《104 中山企研》

《解》

選 (C)。$n$ 越大區間估計式之寬度越窄。

1. 當 $\sigma^2$ 越大, $SE(\hat{\mu})$ 越大, 區間估計式之寬度越寬。

2. $\bar{X}$ 之大小與區間估計式之寬度無關。