八、區間估計

Chapter 8 Interval Estimation

追尋那道光 李明聰

目錄

學習目標

知識(認知)

1.可以描述在推論統計中，區間估計的意涵。

2.可以說明各種情境下，信賴區間的意涵。

3.分辨信賴係數與顯著水準之間的差異性。

4.評價各種情境下，信賴區間的使用價值。

技能

1.能夠計算各種信賴水準下的信賴區間。

2.能夠計算各種情況下所需要的樣本數量。

3.綜合所學，能夠計算實務領域中，於特定情境下的信賴區間。

態度(情意)

1.意識到在日常生活或未來工作環境中，信賴區間資訊的重要性。

2.在各種信賴區間的資訊中，接受信賴區間所傳達的意涵。

教學使用時間：6 小時

推論是依據現有資訊、數值和資料提出研究調查結論的一種嚴謹程序。統計推論(statistical inference)是依據樣本資料(統計值)所提供的訊息，對母體參數提出結論的一種嚴謹程序。統計推論(結論)並非 100 %確定，因樣本無法完全代表整個母體，故運用統計推論時，除了提出統計結論外，尚須利用機率估算此統計結論的可靠程度或信賴程度。

母體平均值 μ 和母體比率 p 的區間估計，利用抽樣分布中的樣本平均值 $\bar{x}$ 和樣本比率 $\bar{p}$，分別進行母體平均值 μ 和母體比率 p 的區間估計。對未知的母體參數(parameter)標示出具有上下限之區間，並估算出在此區間中母體參數機率、可靠度和信賴度。

本章節是針對單一母體特定研究變數分布推論其母體參數(母數)的信賴區間。在經營一家餐廳或旅館時，經常面臨碰到抽樣調查產品品質、服務品質、消費者滿意度、消費者意見等，欲在特定信賴水準(推論成功機率)推論全部產品或消費者的相關參數的信賴區間，不確定性的問題與挑戰，就必須具體地善用區間估計的概念，清楚的呈現不確定性問題的信賴區間範圍，以便於管理者採取適當的方式因應。

8.區間估計

- 8.1母體平均值區間估計：大量樣本數
  - 8.1.1母體變異數已知
  - 8.1.2母體變異數未知
- 8.2母體平均值區間估計：小量樣本數
  - 8.2.1母體屬常態分布和母體變異數已知
  - 8.2.2母體屬常態分布和母體變異數未知
  - 8.2.3母體分布不確定和母體變異數已知
  - 8.2.4母體分布不確定和母體變異數未知
- 8.3母體比例區間估計
  - 8.3.1母體比例區間估計：大量樣本數
  - 8.3.2母體比例區間估計：小量樣本數
- 8.4決定樣本數量
  - 8.4.1估計母體平均值時，需要樣本數量
  - 8.4.2估計母體比例時，需要樣本數量
- 8.5母體變異數區間估計
  - 8.5.1卡方分布
  - 8.5.2卡方分布性質
  - 8.5.3利用卡方分布推估母體變異數的信賴區間
- 8.6單尾區間估計
  - 8.6.1母體平均值單尾區間估計
  - 8.6.2母體比例單尾區間估計

章節結構圖

8.1 母體平均值區間估計：大量樣本數

大量樣本數為 $n \geq 30$ 的情況，利用樣本統計值(statistic)—樣本平均值$\bar{x}$推估母體參數(母數)(parameter)—母體平均值 μ 的信賴區間(interval estimator or confidence interval)。

- 8.1母體平均值區間估計：大量樣本數
  - 8.1.1母體變異數已知
  - 8.1.2母體變異數未知
- 8.2母體平均值區間估計：小量樣本數
  - 8.2.1母體常態分布和母體變異數已知
  - 8.2.2母體常態分布和母體變異數未知
  - 8.2.3母體分布不確定和母體變異數已知
  - 8.2.4母體分布不確定和母體變異數未知

8.1.1 母體變異數已知

當母體變異數和標準(偏)差已知時，其母體平均值信賴區間的估算。

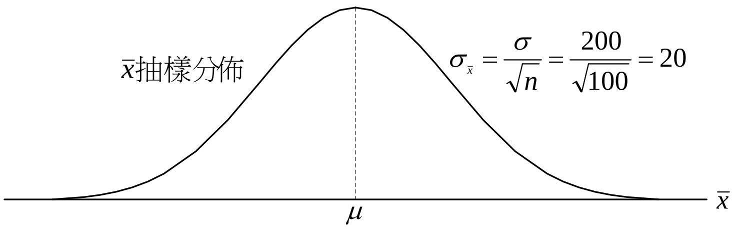
範例 8.1　台灣外食人口中，每人每週外食消費金額之分布屬於常態分布(Normal distribution)，平均值 NT$ 800 元，標準(偏)差 NT$ 200 元，在隨機抽樣中獲得樣本中每人每週消費金額之平均值$\bar{x}$。試計算樣本數量 n 為 100 時，每人每週消費金額之平均值$\bar{x}$抽樣分布的平均值$\mu_{\bar{x}}$和標準(偏)差$\sigma_{\bar{x}}$。

題解：隨機變數 X 代表每人每週外食消費金額，母體平均值 μ = NT$ 800 元，母體標準(偏)差 σ = NT$ 200 元，樣本數量 $n = 100$。

樣本平均值$\bar{x}$抽樣分布的母體平均值$\mu_{\bar{x}} = \mu$ = NT$ 800 元

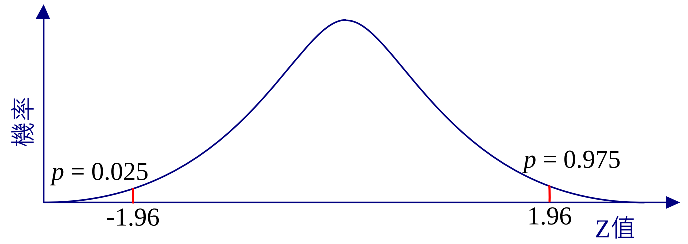
樣本平均值$\bar{x}$抽樣分布的母體標準(偏)差$\sigma_{\bar{x}} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}} = \frac{200}{\sqrt{100}} = \frac{200}{10}$ = NT$ 20.0 元

答案：樣本平均值$\bar{x}$抽樣分布的母體平均值：NT$ 800 元；母體標準(偏)差：NT$ 20.0 元



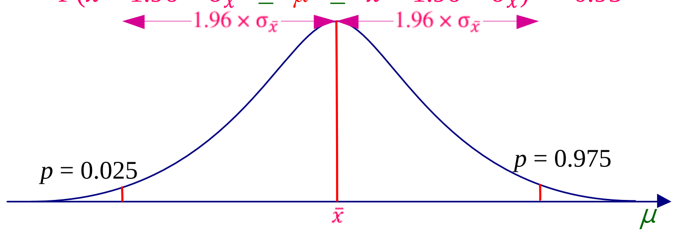
在實務領域中，母體平均值 μ 通常無法獲得，樣本平均值$\bar{x}$在各種隨機抽樣中的分布情況，會是以原始資料分布母體平均值 μ 為中心的常態分布，其母體標準(偏)差$\sigma_{\bar{x}}$ = NT\$20.0 元。

從標準常態分布累積機率表，可以發現 $P(Z \le -1.96) = 0.025$，$P(Z \le 1.96) = 0.975$，綜合前述內容 $P(-1.96 \le Z \le 1.96) = P(Z \le 1.96) - P(Z \le -1.96) = 0.975 - 0.025 = 0.950$。代表樣本平均值$\bar{x}$介在標準化 Z 值 ±1.96 範圍內機率是 0.950。



從母體特定觀測變數，其平均值 μ，標準(偏)差為 σ 的常態分布中，隨機抽取出 n 的觀測值為樣本，其樣本平均值$\bar{x}$可統計獲得，以推估一般情況下未知母體平均值 μ 的信賴區間。

$$P(-1.96 \le Z \le 1.96) = 0.95$$
$$P\left(-1.96 \le \frac{\bar{x}-\mu}{\sigma_{\bar{x}}} \le 1.96\right) = 0.95$$
$$P(-1.96 \times \sigma_{\bar{x}} \le \bar{x} - \mu \le 1.96 \times \sigma_{\bar{x}}) = 0.95$$
$$P(-1.96 \times \sigma_{\bar{x}} - \bar{x} \le -\mu \le 1.96 \times \sigma_{\bar{x}} - \bar{x}) = 0.95$$
$$P(1.96 \times \sigma_{\bar{x}} + \bar{x} \ge \mu \ge -1.96 \times \sigma_{\bar{x}} + \bar{x}) = 0.95$$
$$P(\bar{x} - 1.96 \times \sigma_{\bar{x}} \le \mu \le \bar{x} + 1.96 \times \sigma_{\bar{x}}) = 0.95$$



故在樣本平均值$\bar{x}$為中心點的區間$\bar{x} \pm 1.96 \times \sigma_{\bar{x}}$中，存在母體平均值 μ 機率為 0.95。以樣本平均值$\bar{x}$為中心點，$\pm 1.96 \times \sigma_{\bar{x}}$的區間範圍中，母體平均值 μ 數值落在此區間機率為 0.95。

利用隨機抽樣獲得樣本估計值，假設沒有非抽樣誤差時，抽樣誤差可以表示為：

抽樣誤差(margin of error) = $|\bar{x} - \mu|$

一般情況下，無法獲知母體平均值 μ 的真實數值，故無法利用上式直接估算抽樣誤差。

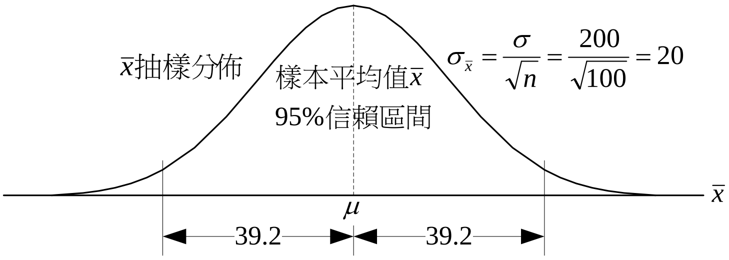
$$P(-1.96 \times \sigma_{\bar{x}} \le \bar{x} - \mu \le 1.96 \times \sigma_{\bar{x}}) = 0.95$$
$$\boxed{P(\mu - 1.96 \times \sigma_{\bar{x}} \le \bar{x} \le \mu + 1.96 \times \sigma_{\bar{x}}) = 0.95}$$

範例 8.2　台灣外食人口中，每人每週外食消費金額為常態分布(Normal distribution)，平均值 NT\$800 元，標準(偏)差 NT\$200 元，在隨機抽樣中獲得樣本中每人每週消費金額之平均值$\bar{x}$。試計算樣本數量 n 為 100 時，每人每週消費金額之平均值$\bar{x}$分布的 95 %信賴區間。

題解：隨機變數 X 代表每人每週外食消費金額，信賴水準 $1 - \alpha = 95\%$，顯著水準 $\alpha = 0.05$，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025} = 1.96$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。母體平均值 μ = NT\$800 元，標準(偏)差 σ = NT\$200 元，樣本數量 $n = 100$。樣品平均值分布之母體標準(偏)差$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{200}{\sqrt{100}} = \frac{200}{10}$ = NT\$20.0 元。

信賴區間 $\mu - z_{\frac{\alpha}{2}} \times \sigma_{\bar{x}} \le \bar{x} \le \mu + z_{\frac{\alpha}{2}} \times \sigma_{\bar{x}} \rightarrow 800 - 1.96 \times 20 \le \bar{x} \le 800 + 1.96 \times 20 \rightarrow 800 - 39.2 \le \bar{x} \le 800 + 39.2 \rightarrow 760.8 \le \bar{x} \le 839.2$

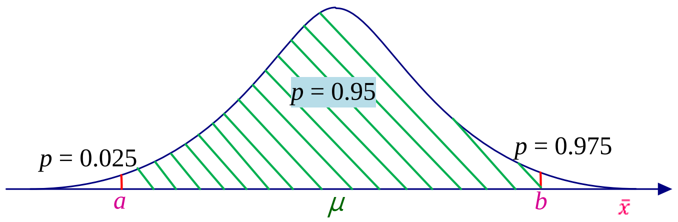
答案：每人每週消費金額之平均值$\bar{x}$分布的 95 %信賴區間 $\mu - 39.2 \le \bar{x} \le \mu + 39.2$ 即 $760.8 \le \bar{x} \le 839.2$



透過樣本獲得的樣本平均值$\bar{x}$，當作母體平均值 μ 的最佳點估計(point estimation)，惟通常母體平均值 μ 未知，如何得知此樣本平均值$\bar{x}$是否可靠與準確，在點估計的過程中，並無法回答此問題。區間估計(Interval estimation)為採用樣本資料計算出一個區間以估計母體參數(parameter)，並說明其可靠度的一種程序。

在母體平均值 μ 左右(上下)範圍內，必須包含 95 %樣本平均值$\bar{x}$的樣本才能代表母體。

若以 a 及 b 分別代表最小及最大樣本平均值$\bar{x}$，符合 $P(\bar{x}_{min} = a \le \mu \le \bar{x}_{max} = b) = P(a \le \mu \le b) = 0.95$。



區間(a, b)稱為信賴區間(confidence interval: CI)或信賴限界(confidence limit, C.L.)，a 代表下限(lower limit)或信賴下界(lower confidence limit)，b 代表上限(upper limit)或信賴上界(upper confidence limit)，而 $b - a$ 之區間範圍為樣本估值精密度(precision)或可靠度(reliability)之測定值。上列公式中機率數值 0.95 稱為信賴係數(confidence coefficient)。

信賴係數(confidence coefficient)是指在從一母體中透過隨機抽樣計算獲得的特定信賴區間(confidence interval)中，包含和出現母體參數(parameter)機率、信心、可靠度和信賴度。信賴水準通常以 $1 - \alpha$ 符號表示。其中 α 是可能發生錯誤機率，α 亦稱為信賴區間的顯著水準(level of significance)。

信賴係數(confidence coefficient) + 顯著水準(level of significance) = 1

95 %信賴區間或 0.95 信賴區間(95 % confidence interval or 0.95 confidence interval)是依據樣本資料估算出一個特定區間，確認在所有可能的樣本組合中，有 95 %的機會將母體參數估算(包含)在此區間中。

信賴係數(confidence coefficient) = 信賴水準(confidence level; level of confidence) = 信賴度(degree of confidence, β) = 推論成功機率 = $1 - \alpha$

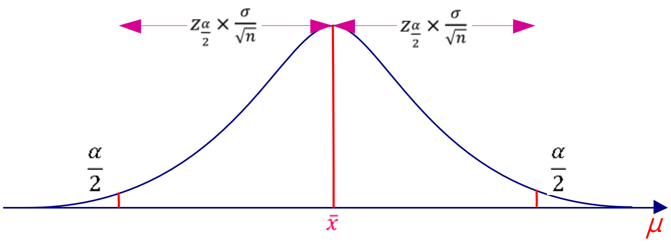
顯著水準(level of significance) = 推論失敗機率 = α

若有 n 個隨機抽樣的樣本，分別以 $x_1, x_2, x_3, \ldots, x_n$ 符號代表，欲利用樣本平均值$\bar{x}$來推算母體平均值 μ 的 95 %信賴區間，利用標準常態分布累積機率表在 0(標準常態分布平均值)左右範圍內佔有 95 %之標準常態變值 Z，$z_{0.025} = -1.96$ 及 $z_{0.975} = 1.96$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)，母體平均值 μ 的信賴區間為：

$$P(-z_{\frac{\alpha}{2}} \le \frac{\bar{x}-\mu}{\sigma_{\bar{x}}} \le z_{\frac{\alpha}{2}}) = 1 - \alpha$$

$$P(\bar{x} - z_{\frac{\alpha}{2}} \times \sigma_{\bar{x}} \le \mu \le \bar{x} + z_{\frac{\alpha}{2}} \times \sigma_{\bar{x}}) = 1 - \alpha$$

$$P(\bar{x} - z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}) = 1 - \alpha$$



其中$z_{\frac{\alpha}{2}}$：右尾機率為$\frac{\alpha}{2}$的標準常態化值；臨界值(critical value)：在特定信賴水準$(1-\alpha)$下的標準化值。

$\frac{\alpha}{2}$：為雙尾機率，雙尾機率之和$\frac{\alpha}{2}+\frac{\alpha}{2}=\alpha$。

$\sigma_{\bar{x}}=\frac{\sigma}{\sqrt{n}}$：標準誤(差)(standard error)，即為點估計值(樣本平均值)的標準(偏)差(standard deviation)。代表樣本平均值$\bar{x}$抽樣分布的分散(離散)程度，可以評量樣本平均值$\bar{x}$抽樣誤差的大小尺度。

$\bar{x}-z_{\frac{\alpha}{2}}\times\sigma_{\bar{x}}$：為信賴區間下限 a；$\bar{x}+z_{\frac{\alpha}{2}}\times\sigma_{\bar{x}}$：為信賴區間上限 b

$z_{\frac{\alpha}{2}}\times\frac{\sigma}{\sqrt{n}}$：為抽樣誤差(sampling error)、誤差界限、誤差範圍(margin of error)、最大誤差或可能機誤。臨界值 × 標準誤(差)。

$2\times z_{\frac{\alpha}{2}}\times\frac{\sigma}{\sqrt{n}}$：為信賴區間的寬度(width)或長度。

母體平均值 μ 在信賴水準為 95 %之信賴區間

$$P(\bar{x}-1.96\times\sigma_{\bar{x}} \le \mu \le \bar{x}+1.96\times\sigma_{\bar{x}}) = 0.95$$

同理，母體平均值 μ 在信賴水準為 99 %之信賴區間

$$P(\bar{x}-2.576\times\sigma_{\bar{x}} \le \mu \le \bar{x}+2.576\times\sigma_{\bar{x}}) = 0.99$$

同理，母體平均值 μ 在信賴水準為 90 %之信賴區間

$$P(\bar{x}-1.645\times\sigma_{\bar{x}} \le \mu \le \bar{x}+1.645\times\sigma_{\bar{x}}) = 0.90$$

在常態分布下，母體平均值 μ 的信賴區間通式

點估計值 ± 誤差範圍(margin of error)　　點估計值 ± 臨界值 × 標準誤(差)

$\bar{x}\pm z_{\frac{\alpha}{2}}\times\sigma_{\bar{x}}$　　$\bar{x}\pm z_{\frac{\alpha}{2}}\times\frac{\sigma}{\sqrt{n}}$

理想的信賴區間需要具備之條件：

A.信賴區間必須具有高的信賴水準(confidence level)。

B.信賴區間需要具有窄的寬度(範圍)。

範例 8.3　美味海產店過去一年每日營業額符合常態分布，上個月 31 天營業日平均日營業額$\bar{x}$ = 12000 元，過去經驗獲得營業額的標準(偏)差 σ = 1000 元，試求每日營業額平均值 μ 之 95 %信賴區間？

題解：隨機變數 X 代表每日營業額，信賴水準 95 %，顯著水準 $\alpha = 0.05$，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025} = 1.96$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。母體平均值信賴區間：$\bar{x}-z_{\frac{\alpha}{2}}\times\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x}+z_{\frac{\alpha}{2}}\times\frac{\sigma}{\sqrt{n}}$

下限：$\bar{x}-z_{\frac{\alpha}{2}}\times\frac{\sigma}{\sqrt{n}} = 12000-1.96\times\frac{1000}{\sqrt{31}} = 12000-352 = 11648$ 元

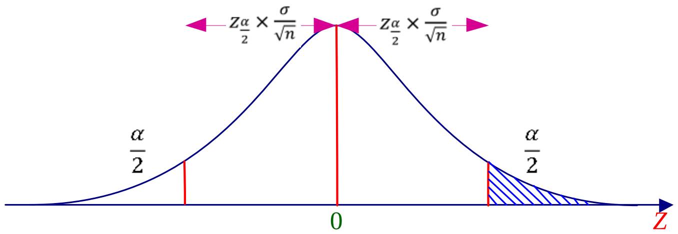
上限：$\bar{x}+z_{\frac{\alpha}{2}}\times\frac{\sigma}{\sqrt{n}} = 12000+1.96\times\frac{1000}{\sqrt{31}} = 12000+352 = 12352$ 元

故每日營業額平均值 μ 之 95 %信賴區間(CI)為 $11648 \le \mu \le 12352$ 元。表示上個月 31 天營業日營業額，所有可能樣本平均值在(11648, 12352)區間有 95 %的機會，將每日營業額平均值 μ 包含在內。

答案：每日營業額平均值 μ 之 95 %信賴區間為 11648 到 12352 元

練習 8.1　請依據下列的顯著水準 α 數值，查表獲得$z_{\frac{\alpha}{2}}$數值(z 值有效位數四捨五入取到小數點後第 4 位；z 值取正值)

a. $\alpha = 0.10$　　b. $\alpha = 0.01$　　c. $\alpha = 0.05$
d. $\alpha = 0.20$　　e. $\alpha = 0.02$　　f. $\alpha = 0.04$



題解：

a.$\alpha = 0.10$ → $z_{\frac{\alpha}{2}} = z_{\frac{0.10}{2}} = z_{0.05} = 1.6450$(代表在標準化 Z 分布中，雙尾機率 $\alpha = 0.10$，右尾機率 0.05，左尾機率 = 1 – 右尾機率 = 1 – 0.05 = 0.95。使用 Excel 軟體 NORM.S.INV 函數查詢獲得。使用 NORM.S.INV 函數需要輸入右尾機率)

b.$\alpha = 0.01$ → $z_{\frac{\alpha}{2}} = z_{\frac{0.01}{2}} = z_{0.005} = 2.5750$

c.$\alpha = 0.05$ → $z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025} = 1.9600$(代表在標準化 Z 分布中，雙尾機率 $\alpha = 0.05$，右尾機率 0.025，左尾機率 = 1 – 右尾機率 = 1 – 0.025 = 0.975。使用 Excel 軟體 NORM.S.INV 函數查詢獲得。使用 NORM.S.INV 函數需要輸入右尾機率)

d.$\alpha = 0.20$ → $z_{\frac{\alpha}{2}} = z_{\frac{0.20}{2}} = z_{0.10} = 1.2817$

e.$\alpha = 0.02$ → $z_{\frac{\alpha}{2}} = z_{\frac{0.02}{2}} = z_{0.01} = 2.3267$

f.$\alpha = 0.04$ → $z_{\frac{\alpha}{2}} = z_{\frac{0.04}{2}} = z_{0.02} = 2.0540$

「$z_{\frac{\alpha = 0.20}{2}} = z_{0.10} = 1.2817$」＜「$z_{\frac{\alpha = 0.10}{2}} = z_{0.05} = 1.6450$」＜「$z_{\frac{\alpha = 0.05}{2}} = z_{0.025} = 1.9600$」＜「$z_{\frac{\alpha = 0.04}{2}} = z_{0.02} = 2.0540$」＜「$z_{\frac{\alpha = 0.02}{2}} = z_{0.01} = 2.3267$」＜「$z_{\frac{\alpha = 0.01}{2}} = z_{0.005} = 2.5750$」。由此練習顯示「**顯著水準 α 數值愈大，獲得的對應標準化 Z 值數值愈小**」。反之，「**顯著水準 α 數值愈小，獲得的對應標準化 Z 值數值愈大**」。

練習 8.2　請依據下列信賴水準(confidence level)，找出其對應的顯著水準 α 數值：

a. 85 %信賴水準　　b. 90 %信賴水準　　c. 95 %信賴水準　　d. 99 %信賴水準

題解：

a.信賴水準 = 1 – α = 85 % → α = 1 – 信賴水準 = 1 – (1 – α) = 1 – 0.85 = 0.15

b.信賴水準 = 1 – α = 90 % → α = 1 – 信賴水準 = 1 – (1 – α) = 1 – 0.90 = 0.10

c.信賴水準 = 1 – α = 95 % → α = 1 – 信賴水準 = 1 – (1 – α) = 1 – 0.95 = 0.05

d.信賴水準 = 1 – α = 99 % → α = 1 – 信賴水準 = 1 – (1 – α) = 1 – 0.99 = 0.01

答案：各信賴水準對應的顯著水準分別為 a. α = 0.15, b. α = 0.10, c. α = 0.05, d. α = 0.01

練習 8.3　從一個標準(偏)差為 σ 的母體中，隨機抽出樣本數量 n。請列出母體平均值 90、95 和 99 %的信賴區間。

練習 8.4 請敘述信賴區間想要達到的兩個特性。

A.

B.

練習 8.5 從一平均值 26.2 與標準(偏)差 4.1 的母體中，獲得 70 個隨機樣本。(信賴區間數值有效位數四捨五入取到小數點後第 2 位)

a.請計算母體平均值 μ 的 95 %信賴區間

b.請問信賴係數(confidence coefficient)等於 0.95 的意涵？

c.請計算母體平均值 μ 的 99 %信賴區間

d.在固定樣本數量的情況下，當信賴細數增加時，信賴區間會發生何種狀況？

題解：

a. 26.2 ± 0.96

b. 在重複性的抽樣過程中，有 95 %的機會信賴區間會包含母體平均值 μ

c. 26.2 ± 1.26 d. 增加

練習 8.6 由於全球景氣與經營環境的關係，造成全球性失業潮。學術機構為提供政府相關決策的參考，欲調查失業居民的年齡分布情況。現今針對失業居民進行隨機抽樣獲得 30 份樣本，失業居民的年齡資料(單位：歲)如下表所示。假設母體失業居民的年齡分布標準(偏)差為 15.2 歲。(a)試估算在 95 %信賴水準下，失業居民年齡平均值之信賴區間；(b)試估算在 90 %信賴水準下，失業居民年齡平均值之信賴區間。(答案有效位數四捨五入取到小數點後第 1 位)

25	65	54	25	35	45	46	48	52	53
56	51	28	35	45	42	41	25	29	28
29	26	33	44	54	56	48	47	45	23

題解：隨機變數 X 代表每位失業居民的年齡，樣本數量 $n = 30$，樣本平均值$\bar{x} = 41.1$ 歲，母體標準(偏)差 $\sigma = 15.2$ 歲。

(a) Step 1：信賴水準 95 %，顯著水準 $\alpha = 0.05$，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025} = 1.96$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。

Step 2：母體平均值信賴區間：$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} = 41.1 \pm 1.96 \times \frac{15.2}{\sqrt{30}} = 41.1 \pm 5.44$ 歲

(b) Step 1：信賴水準 90 %，顯著水準 $\alpha = 0.10$，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.10}{2}} = z_{0.05} = 1.645$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)

Step 2：母體平均值信賴區間：$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} = 41.1 \pm 1.645 \times \frac{15.2}{\sqrt{30}} = 41.1 \pm 4.75$ 歲

答案：(a)95 %信賴水準失業居民年齡平均值之信賴區間 35.7 到 46.5 歲；(b)90 %信賴水準失業居民年齡平均值之信賴區間 36.4 到 45.9 歲

練習 8.7 由於全球景氣與經營環境的關係，造成全球性失業潮。學術機構為提供政府相關決策的參考，欲調查失業居民的年齡分布情況。現今針對失業居民進行隨機抽樣獲得 50 份樣本，失業居民的年齡資料(單位：歲)如下表所示。假設母體失業居民的年齡分布標準(偏)差為 15.2 歲。(a)試估算在 95 %信賴水準下，失業居民年齡平均值之信賴區間；(b)試估算在 90 %信賴水準下，失業居民年齡平均值之信賴區間。(答案有效位數四捨五入取到小數點後第 1 位)

25	65	54	25	35	45	46	48	52	53
56	51	28	35	45	42	41	25	29	28
29	26	33	44	54	56	48	47	45	23

28	35	56	54	51	28	34	56	35	34
61	58	43	44	41	46	48	45	41	34

題解：

(a)Step 1：95 % confidence level，顯著水準 $\alpha = 0.05$，$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025} = 1.96$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。

Step 2　the confidence interval for μ is from　$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$

$n = 50$　　$\bar{x} = 42.1$ 歲　　$\sigma = 15.2$ 歲

$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} = 42.1 \pm 1.96 \times \frac{15.2}{\sqrt{50}} = 42.1 \pm 4.2$ 歲

(b)Step 1：90 % confidence level，顯著水準 $\alpha = 0.10$，$z_{\frac{\alpha}{2}} = z_{\frac{0.10}{2}} = z_{0.05} = 1.645$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。

Step 2　the confidence interval for μ is from　$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$

$n = 50$　　$\bar{x} = 42.1$ 歲　　$\sigma = 15.2$ 歲

$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} = 42.1 \pm 1.645 \times \frac{15.2}{\sqrt{50}} = 42.1 \pm 3.5$ 歲

答案：(a)95 %信賴水準失業居民年齡平均值之信賴區間 37.9 到 46.3 歲；(b)90 %信賴水準失業居民年齡平均值之信賴區間 38.6 到 45.6 歲

練習 8.8　從一常態分布 $N(5, 12)$中，隨機抽出 n 個基本單位為樣本，求符合 $P(\bar{x} - 0.98 \leq \mu \leq \bar{x} + 0.98) = 0.95$ 之樣本數量 n。(答案取到個位數)

練習 8.9　若全球 5 星級旅館客房面積趨近於常態分布，隨機抽出 36 間客房測量其實際面積，其平均面積為 15 坪，已知全體客房面積的標準(偏)差為 3 坪，在信賴度 $\beta = 0.95$，推測全體客房平均面積的信賴限界(confidence limit)。(答案有效位數四捨五入取到小數點後第 2 位)

練習 8.10　管理學院大一學生統計學的成績呈現 $N(\mu, 12.50)$，從學生中隨機抽取 65 位當作樣本，其平均分數為 65.25 分，試估算管理學院大一學生統計學平均分數之，(a)點估計值；(b)95 %信賴水準的最大誤差；(c)99 %信賴水準的最大誤差；(d)95 %信賴水準的信賴區間；(e)99 %信賴水準的信賴區間。(答案有效位數四捨五入取到小數點後第 2 位)

答案：(a)平均分數的點估計值 = 65.25 分；(b)信賴水準 95 %，顯著水準 $\alpha = 0.05$，右尾機率$\frac{\alpha}{2}$標準常態值 $z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025} = 1.96$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。最大誤差$z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} = 1.96 \times \frac{\sqrt{12.50}}{\sqrt{65}} = 0.86$ 分；(c)信賴水準 99 %，顯著水準 $\alpha = 0.01$，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.01}{2}} = z_{0.005} = 2.576$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。最大誤差$z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} = 2.576 \times \frac{\sqrt{12.50}}{\sqrt{65}} = 1.13$ 分；(d) $\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} = 65.25 \pm 1.96 \times \frac{\sqrt{12.50}}{\sqrt{65}} = 65.25 \pm 0.86$ 信賴區間(64.39~66.11 分)；(e) $\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} = 65.25 \pm 2.576 \times \frac{\sqrt{12.50}}{\sqrt{65}} = 65.25 \pm 1.13$ 信賴區間(64.12~66.38 分)

信賴區間 Excel 函數

利用 Excel 軟體中插入(I)→函數(F)…→在插入函數對話方塊中選取類別(C): **統計**，選取函數(N): **CONFIDENCE**→$\boxed{\text{確定}}$。在函數引數對話視窗中，Alpha 方塊輸入：欲推估信賴區間的顯著水準 α；Standard_dev 方塊輸入：母體標準(偏)差 σ；Size 方塊輸入：樣本數量 n。確定。即會在原先選定的儲存格中出現信賴區間的$z_{\frac{\alpha}{2}} \times \sigma_{\bar{x}}$數值。CONFIDENCE(alpha,standard_dev,size)。若將樣本平均值$\bar{x}$設為中心點左右分別$z_{\frac{\alpha}{2}} \times \sigma_{\bar{x}}$距離內區間為信賴區間。

當信賴區間的寬度 $2 \times z_{\frac{\alpha}{2}} \times \sigma_{\bar{x}} = 2 \times z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$愈小，表示區間估計的精準度(precision 準確度)愈高。故信賴區間寬度受其**顯著水準 α 的標準化 z 值**、**母體標準(偏)差 σ** 和**樣本數量 n** 三個數值的影響。因為「**顯著水準 α 數值愈大，獲得的對應標準化 Z 值數值愈小**」，在相同的樣本數量中，信賴水準(confidence level) $1-\alpha$ 下降(信賴區間的顯著水準 α 提高)會使精準度提高。母體標準差 σ 愈大，因母體標準差在分子，信賴區間寬度愈大，精準度愈低。樣本數量 n 愈大，因樣本數量在分母，信賴區間寬度愈小，精準度愈大。

8.1.2 母體變異數未知

在實際調查環境中，通常對於母體的參數未知，故，母體平均值 μ、變異數 σ^2 和標準(偏)差 σ 皆無法獲得明確的數值。在大量樣本數量($n \geq 30$)的情況下，使用樣本標準(偏)差 $S = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$作為母體標準(偏)差 σ 的點估計值。進行母體參數(parameter)區間估計時，可以利用樣本標準(偏)差 S 取代母體標準(偏)差 σ。大量樣本數量($n \geq 30$)，變異數 σ^2 和標準(偏)差 σ 未知時，母體平均值 μ 的信賴區間(**獲得近似數值，使用 t 分布進行運算比較精準**)【點估計值±誤差範圍(臨界值×標準誤)】為

$$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}} = \bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}}{\sqrt{n}} = \bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{\sqrt{\frac{\sum_{i=1}^{n}x_i^2-\frac{\left(\sum_{i=1}^{n}x_i\right)^2}{n}}{n-1}}}{\sqrt{n}}$$

範例 8.4 奇遇海產店每日營業額分布情況趨近於常態分布，上個月 31 天營業日，每日營業額列於下表，試求每日營業額平均值 μ 之 95 %信賴區間？(信賴區間有效位數取到小數 0 位)

題解：

營業日	營業額 x_i	$x_i-\bar{x}$	$(x_i-\bar{x})^2$	營業日	營業額 x_i	$x_i-\bar{x}$	$(x_i-\bar{x})^2$
1	16000	-500	250000	17	15000	-1500	2250000
2	15900	-600	360000	18	12000	-4500	20250000
3	17800	1300	1690000	19	15000	-1500	2250000
4	16600	100	10000	20	17000	500	250000
5	12000	-4500	20250000	21	16000	-500	250000
6	15000	-1500	2250000	22	15200	-1300	1690000
7	17000	500	250000	23	12000	-4500	20250000
8	16000	-500	250000	24	15000	-1500	2250000
9	15200	-1300	1690000	25	17000	500	250000
10	16500	0	0	26	16000	-500	250000
11	17500	1000	1000000	27	15200	-1300	1690000
12	18000	1500	2250000	28	16500	0	0
13	20000	3500	12250000	29	17500	1000	1000000
14	26000	9500	90250000	30	18000	1500	2250000
15	16000	-500	250000	31	18100	1600	2560000
16	20500	4000	16000000	合計	511500		206440000

樣本平均值$\bar{x} = \frac{\sum_{i=1}^{n}x_i}{n} = \frac{511500}{31} = 16500$。

使用 Z 分布運算**近似**信賴區間：信賴水準 95 %，顯著水準 $\alpha = 0.05$，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025}$ = 1.96(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。

$$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{s}{\sqrt{n}} = \bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}}{\sqrt{n}} = 16500 \pm 1.96 \times \frac{\sqrt{\frac{206440000}{31-1}}}{\sqrt{31}} = 16500 \pm 1.96 \times \frac{2623.2296}{5.5678} = 16500 \pm 923 \rightarrow$$

近似信賴區間 15577~17423 元

使用 t 分布運算**精準**信賴區間：自由度 $v = n - 1 = 31 - 1 = 30$，信賴水準 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，$t_{\frac{\alpha}{2},v} = t_{\frac{0.05}{2},30} = t_{0.025,30} = 2.0423$(使用 Excel 軟體 T.INV 函數查詢獲得)。

$$\bar{x} \pm t_{\frac{\alpha}{2},v} \times \frac{s}{\sqrt{n}} = 16500 \pm 2.0423 \times \frac{2623.2296}{\sqrt{31}} = 16500 \pm 962.2084 \rightarrow$$ 精準信賴區間 15537.79~17462.21 元

答案：每日營業額平均值 μ 之 95 %信賴區間(CI)近似值為 $15577 \le \mu \le 17423$ 元；精準值為 $15538 \le \mu \le 17462$ 元

練習 8.11　奇遇海產店每日販售清蒸鱸魚數量(尾)分布情況趨近於常態分布，上個月 31 天營業日，每日販售清蒸鱸魚數量(尾)x_i列於下表，試求每日販售清蒸鱸魚數量平均值 μ 之 95 %信賴區間？(答案有效位數四捨五入取到小數點後第 1 位)

題解：

營業日	數量 x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	22	-7.7742	60.4381
2	21	-8.7742	76.9865
3	16	-13.7742	189.7284
4	25	-4.7742	22.7929
5	18	-11.7742	138.6316
6	21	-8.7742	76.9865
7	25	-4.7742	22.7929
8	22	-7.7742	60.4381
9	22	-7.7742	60.4381
10	55	25.2258	636.3413
11	24	-5.7742	33.3413
12	26	-3.7742	14.2445
13	41	11.2258	126.0187
14	40	10.2258	104.5671
15	33	3.2258	10.4058
16	26	-3.7742	14.2445

營業日	數量 x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
17	35	5.2258	27.3091
18	38	8.2258	67.6639
19	39	9.2258	85.1155
20	42	12.2258	149.4703
21	26	-3.7742	14.2445
22	26	-3.7742	14.2445
23	28	-1.7742	3.1478
24	30	0.2258	0.0510
25	29	-0.7742	0.5994
26	34	4.2258	17.8574
27	35	5.2258	27.3091
28	32	2.2258	4.9542
29	31	1.2258	1.5026
30	36	6.2258	38.7607
31	25	-4.7742	22.7929
合計			2123.4194

樣本平均值$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{923}{31} = 29.7742$。

使用 Z 分布運算**近似**信賴區間：信賴水準 95 %，顯著水準 $\alpha = 0.05$，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025}$ = 1.96(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。

$$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{s}{\sqrt{n}} = \bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}}{\sqrt{n}} = 29.7742 \pm 1.96 \times \frac{\sqrt{\frac{2123.4194}{31-1}}}{\sqrt{31}} = 29.7742 \pm 1.96 \times \frac{8.4131}{5.5678} = 29.7742 \pm 2.9616$$

→　近似信賴區間 26.8126~32.7358 尾

使用 t 分布運算**精準**信賴區間：自由度 $v = n - 1 = 31 - 1 = 30$，信賴水準 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，$t_{\frac{\alpha}{2},v} = t_{\frac{0.05}{2},30} = t_{0.025,30} = 2.0423$(使用 Excel 軟體 T.INV 函數查詢獲得)。

$$\bar{x} \pm t_{\frac{\alpha}{2},v} \times \frac{s}{\sqrt{n}} = 29.7742 \pm 2.0423 \times \frac{8.4131}{\sqrt{31}} = 29.7742 \pm 3.0860 \rightarrow$$ 精準信賴區間 26.6883~32.8601 尾

答案：每日販售清蒸鱸魚數量平均值 μ 之 95 %近似信賴區間(CI)為 $26.8 \le \mu \le 32.7$ 尾；精準信賴區間(CI)為 $26.7 \le \mu \le 32.9$ 尾

練習 8.12　全球經歷金融海嘯，失業率攀升，依據「純純大學」調查 120 位失業勞工，結果顯示失業者平均花 30.5 週才找到下一份工作，樣本標準(偏)差 4.2 週。(A)請建構 95 %信賴水準下，失業者尋找下一份工作，需要之平均時間的信賴區間。(B)若說失業者尋找下一份工作，需要之平均時間為 27.8 週，是否合理？請論述。

題解：題目提供樣本數 n = 120。樣本平均值$\bar{x}$ = 30.5 週，樣本標準(偏)差 S = 4.2 週

使用 Z 分布運算**近似**信賴區間：信賴水準 95 %，顯著水準 α = 0.05，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025}$ = 1.96(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。

$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}} = 30.5 \pm 1.96 \times \frac{4.2}{\sqrt{120}} = 30.5 \pm 0.75$ → 近似信賴區間 29.7~31.3 週

使用 t 分布運算**精準**信賴區間：自由度 $v = n - 1 = 120 - 1 = 119$，信賴水準 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，

$t_{\frac{\alpha}{2},v} = t_{\frac{0.05}{2},119} = t_{0.025,119}$ = 1.9801(使用 Excel 軟體 T.INV 函數查詢獲得)。

$\bar{x} \pm t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} = 30.5 \pm 1.9801 \times \frac{4.2}{\sqrt{120}} = 30.5 \pm 0.7592$ → 精準信賴區間 29.7408~ 31.2592 元

答案：(A)近似和精準運算信賴區間 29.7~31.3 週；(B)若$\bar{x}$ = 27.8 週不屬於信賴區間中，明顯低於 95 %信賴水準之信賴區間下限值，其出現機率很低，若隨機抽取的樣本具有代表性，數值正確，可推測其原因可能是金融海嘯已經慢慢遠離，尋找下一份工作的間隔時間，才會有縮短的現象。

練習 8.13　從一個樣本平均值 60 和樣本標準(偏)差 10 的母體中獲得 50 個觀測值的隨機樣本。(A)請提供母體平均值 90 %的信賴區間；(B)請提供母體平均值 95 %的信賴區間；(C)請提供母體平均值 99 %的信賴區間。(答案有效位數四捨五入取到小數點後第 1 位)

題解：題目提供樣本數 n = 50。樣本平均值$\bar{x}$ = 60，樣本標準(偏)差 S = 10

(A)使用 Z 分布運算**近似**信賴區間：信賴水準 90 %，顯著水準 α = 0.10，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.10}{2}} = z_{0.05}$ = 1.6449(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。

$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}} = 60 \pm 1.6449 \times \frac{10}{\sqrt{50}} = 60 \pm 2.3262$ → 近似信賴區間 57.6738~62.3262

使用 t 分布運算**精準**信賴區間：自由度 $v = n - 1 = 50 - 1 = 49$，信賴水準 $1 - \alpha = 0.90$，顯著水準 $\alpha = 0.10$，

$t_{\frac{\alpha}{2},v=n-1} = t_{\frac{0.10}{2},50-1} = t_{0.05,49}$ = 1.6766(使用 Excel 軟體 T.INV 函數查詢獲得)。

$\bar{x} \pm t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} = 60 \pm 1.6766 \times \frac{10}{\sqrt{50}} = 60 \pm 2.3710$ → 精準信賴區間 57.6290~62.3710

(B)使用 Z 分布運算**近似**信賴區間：信賴水準 95 %，顯著水準 α = 0.05，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025}$ = 1.96(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。

$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}} = 60 \pm 1.96 \times \frac{10}{\sqrt{50}} = 60 \pm 2.7718$ → 近似信賴區間 57.2282~62.7718

使用 t 分布運算**精準**信賴區間：自由度 $v = n - 1 = 50 - 1 = 49$，信賴水準 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，

$t_{\frac{\alpha}{2},v=n-1} = t_{\frac{0.05}{2},50-1} = t_{0.025,49}$ = 2.0096(使用 Excel 軟體 T.INV 函數查詢獲得)。

$\bar{x} \pm t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} = 60 \pm 2.0096 \times \frac{10}{\sqrt{50}} = 60 \pm 2.8420$ → 精準信賴區間 57.1580~62.8420

(C)使用 Z 分布運算**近似**信賴區間：信賴水準 99 %，顯著水準 α = 0.01，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.01}{2}} = z_{0.005}$ = 2.5758(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。

$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}} = 60 \pm 2.5758 \times \frac{10}{\sqrt{50}} = 60 \pm 3.6428$ → 近似信賴區間 56.3572~63.6428

使用 t 分布運算**精準**信賴區間：自由度 $v = n - 1 = 50 - 1 = 49$，信賴水準 $1 - \alpha = 0.99$，顯著水準 $\alpha = 0.01$，

$t_{\frac{\alpha}{2},v=n-1} = t_{\frac{0.01}{2},50-1} = t_{0.005,49}$ = 2.6800(使用 Excel 軟體 T.INV 函數查詢獲得)。

$\bar{x} \pm t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} = 60 \pm 2.6800 \times \frac{10}{\sqrt{50}} = 60 \pm 3.7900$ → 精準信賴區間 56.2100~63.7900

答案：(A)近似信賴區間 57.67~62.33，精準信賴區間 57.63~62.37；(B)近似信賴區間 57.23~62.77，精準信賴區間 57.16~62.84；(C)近似信賴區間 56.36~63.64，精準信賴區間 56.21~63.79

練習 8.14 從一個樣本平均值 50 和樣本標準(偏)差 12 的母體中獲得 55 個觀測值的隨機樣本。(A)請計算樣本平均值的標準(偏)差$\sigma_{\bar{x}}$？(B)在 90 %信賴區間中，請計算誤差範圍(margin of error)？(答案有效位數四捨五入取到小數點後第 1 位)

題解：題目提供樣本數 $n = 55$。樣本平均值$\bar{x} = 50$，樣本標準(偏)差 S = 12

樣本平均值的標準(偏)差$\sigma_{\bar{x}} = \frac{S}{\sqrt{n}} = \frac{12}{\sqrt{55}} = 1.6181$

使用 Z 分布運算**近似**誤差範圍：信賴水準 90 %，顯著水準 $\alpha = 0.10$，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.10}{2}} = z_{0.05}$ = 1.6449(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。誤差範圍 = $z_{\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}} = 1.6449 \times \frac{12}{\sqrt{55}} = 2.6615$

使用 t 分布運算**精準**誤差範圍：自由度 $v = n - 1 = 55 - 1 = 54$，信賴水準 $1 - \alpha = 0.90$，顯著水準 $\alpha = 0.10$，$t_{\frac{\alpha}{2},v=n-1} = t_{\frac{0.10}{2},55-1} = t_{0.05,54} = 1.6737$(使用 Excel 軟體 T.INV 函數查詢獲得)。誤差範圍 = $t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} = 1.6737 \times \frac{12}{\sqrt{55}} = 2.7080$

答案：(A)樣本平均值的標準(偏)差$\sigma_{\bar{x}} = 1.6181$；(B)使用 Z 分布運算**近似**誤差範圍 = 2.6615，使用 t 分布運算**精準**誤差範圍 = 2.7080

練習 8.15 奇遇速食餐廳提供購餐車道服務，若服務人員訓練合格，對於購餐車道的消費者服務時間會趨近於常態分布。現今隨機抽取 30 輛進入購餐車道的車輛，測量服務人員的服務時間如下表所示(單位：秒)。(A)請計算母體平均值 90 %信賴區間；(B)請計算母體平均值 95 %信賴區間；(C)請計算母體平均值 99 %信賴區間；(D)在 95 %信賴區間中，請計算誤差範圍(margin of error)？(答案有效位數四捨五入取到小數點後第 1 位)

55	68	45	78	150	241	162	156	182	125
75	89	91	95	92	65	75	85	95	105
132	120	142	110	111	130	128	130	108	109

題解：樣本數量 $n = 30$，樣本平均值$\bar{x} = 111.6333$，樣本標準(偏)差 $S = 40.8575$

(A)使用 Z 分布運算**近似**信賴區間：信賴水準 90 %，顯著水準 $\alpha = 0.10$，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.10}{2}} = z_{0.05}$ = 1.6449(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。

$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}} = 111.6333 \pm 1.6449 \times \frac{40.8575}{\sqrt{30}} = 111.6333 \pm 12.2668$ → 近似信賴區間 99.3665~123.9001

使用 t 分布運算**精準**信賴區間：自由度 $v = n - 1 = 30 - 1 = 29$，信賴水準 $1 - \alpha = 0.90$，顯著水準 $\alpha = 0.10$，

$t_{\frac{\alpha}{2},v=n-1} = t_{\frac{0.10}{2},30-1} = t_{0.05,29} = 1.6991$(使用 Excel 軟體 T.INV 函數查詢獲得)。

$\bar{x} \pm t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} = 111.6333 \pm 1.6991 \times \frac{40.8575}{\sqrt{30}} = 111.6333 \pm 12.6716$ → 精準信賴區間 98.9618~124.3049

(B)使用 Z 分布運算**近似**信賴區間：信賴水準 95 %，顯著水準 $\alpha = 0.05$，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025}$ = 1.9600(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。

$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}} = 111.6333 \pm 1.9600 \times \frac{40.8575}{\sqrt{30}} = 111.6333 \pm 14.6168$ → 近似信賴區間 97.0165~126.2501

使用 t 分布運算**精準**信賴區間：自由度 $v = n - 1 = 30 - 1 = 29$，信賴水準 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，

$t_{\frac{\alpha}{2},v=n-1} = t_{\frac{0.05}{2},30-1} = t_{0.025,29} = 2.0452$(使用 Excel 軟體 T.INV 函數查詢獲得)。

$\bar{x} \pm t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} = 111.6333 \pm 2.0452 \times \frac{40.8575}{\sqrt{30}} = 111.6333 \pm 15.2527$ → 精準信賴區間 96.3806~126.8860

(C)使用 Z 分布運算**近似**信賴區間：信賴水準 99 %，顯著水準 $\alpha = 0.01$，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.01}{2}} = z_{0.005}$ = 2.5758(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。

$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}}$ = 111.6333 ± 2.5758 × $\frac{40.8575}{\sqrt{30}}$ = 111.6333 ± 19.2097 → 近似信賴區間 92.4236~130.8431

使用 t 分布運算**精準**信賴區間：自由度 $v = n - 1 = 30 - 1 = 29$，信賴水準 $1 - \alpha = 0.99$，顯著水準 $\alpha = 0.01$，

$t_{\frac{\alpha}{2},v=n-1} = t_{\frac{0.01}{2},30-1} = t_{0.005,29}$ = 2.7564(使用 Excel 軟體 T.INV 函數查詢獲得)。

$\bar{x} \pm t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}}$ = 111.6333 ± 2.7564 × $\frac{40.8575}{\sqrt{30}}$ = 111.6333 ± 20.5563 → 精準信賴區間 91.0771~132.1896

(D)使用 Z 分布運算**近似**誤差範圍：信賴水準 95 %，顯著水準 α = 0.05，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025}$ = 1.9600(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。近似誤差範圍 = $z_{\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}}$ = 1.9600 × $\frac{40.8575}{\sqrt{30}}$ = 14.6168

使用 t 分布運算**精準**誤差範圍：自由度 $v = n - 1 = 30 - 1 = 29$，信賴水準 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，

$t_{\frac{\alpha}{2},v=n-1} = t_{\frac{0.05}{2},30-1} = t_{0.025,29}$ = 2.0452(使用 Excel 軟體 T.INV 函數查詢獲得)。精準誤差範圍 = $t_{\frac{\alpha}{2},v=n-1} \times \frac{S}{\sqrt{n}}$ = 2.0452 × $\frac{40.8575}{\sqrt{30}}$ = 15.2527

答案：(A)近似信賴區間 99.37~123.90，精準信賴區間 98.96~124.30；(B)近似信賴區間 97.02~126.25，精準信賴區間 96.38~126.89；(C)近似信賴區間 92.42~130.84，精準信賴區間 91.08~132.19；(D)近似誤差範圍 = 14.62，精準誤差範圍 = 15.25

範例 8.5　奇遇速食餐廳提供購餐車道服務，從購餐車道隨機抽取 50 輛車(汽車、機車和腳踏車)，評量其服務時間為 X(單位：秒)，獲得$\sum_{i=1}^{50} x_i$ = 1320 與$\sum_{i=1}^{50} x_i^2$ = 35820。請估算服務購餐車道每一輛車平均時間的 95 %信賴區間。(答案有效位數四捨五入取到小數點後第 1 位)

題解：樣本數量 n = 50，樣本平均值$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{1320}{50}$ = 26.4 秒，樣本變異數 $S^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1} = \frac{\sum_{i=1}^{n} x_i^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}}{n-1} = \frac{35820-\frac{1320^2}{50}}{50-1} = \frac{35820-34848}{49} = \frac{972}{49}$ = 19.8367，樣本標準(偏)差 $S = \sqrt{S^2} = \sqrt{19.8367}$ = 4.4538 秒。

使用 Z 分布運算**近似**信賴區間：信賴水準 95 %，顯著水準 α = 0.05，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025}$ = 1.9600(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。

$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}}$ = 26.4 ± 1.9600 × $\frac{4.4538}{\sqrt{50}}$ = 26.4 ± 1.2345 → 近似信賴區間 25.1655~27.6345 秒

使用 t 分布運算**精準**信賴區間：自由度 $v = n - 1 = 50 - 1 = 49$，信賴水準 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，

$t_{\frac{\alpha}{2},v=n-1} = t_{\frac{0.05}{2},50-1} = t_{0.025,49}$ = 2.0096(使用 Excel 軟體 T.INV 函數查詢獲得)。

$\bar{x} \pm t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}}$ = 26.4 ± 2.0096 × $\frac{4.4538}{\sqrt{50}}$ = 26.4 ± 1.2658 → 精準信賴區間 25.1342~27.6658 秒

答案：服務購餐車道每一輛車平均時間的 95 5 %近似信賴區間 25.17~27.63 秒，精準信賴區間 25.13~27.67 秒

8.2 母體平均值區間估計：小量樣本數

小量樣本數為 n < 30 的情況，利用樣本統計值(statistic)—樣本平均值$\bar{x}$推估母體參數(母數)(parameter)—母體平均值 μ 的信賴區間(interval estimator or confidence interval)。

8.2.1 母體屬常態分布和母體變異數已知

母體屬於常態分布時，樣本平均值$\bar{x}$的抽樣分布亦是趨近於常態分布。故，$\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$。依據標準化常態分布 z 值，推估母體平均值 μ 在 $1 - \alpha$ 的信賴水準下之信賴區間：

$$P(-z_{\frac{\alpha}{2}} \leq \frac{\bar{x}-\mu}{\sigma_{\bar{x}}} \leq z_{\frac{\alpha}{2}}) = 1 - \alpha$$

$$P(\bar{x} - z_{\frac{\alpha}{2}} \times \sigma_{\bar{x}} \leq \mu \leq \bar{x} + z_{\frac{\alpha}{2}} \times \sigma_{\bar{x}}) = 1 - \alpha$$

$$P(\bar{x} - z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}) = 1 - \alpha$$

因此，在母體屬於常態分布，而樣本數量較少 $n < 30$，母體變異數 σ^2 和標準(偏)差 σ 已知的情況下，母體平均值 μ 的信賴區間【點估計值±誤差範圍(臨界值×標準誤)】為：

$$\bar{x} \pm z_{\frac{\alpha}{2}} \times \sigma_{\bar{x}} = \bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$$

範例 8.6 奇遇海產店過去一年每日營業額趨近於(符合)常態分布，上星期 7 天營業日平均日營業額 $\bar{x}$ = 12000 元，過去經驗獲得營業額的標準(偏)差 σ = 1000 元，試求每日營業額平均值 μ 之 95 %信賴區間？(答案有效位數四捨五入取到小數點後第 1 位)

題解：信賴水準 95 %，顯著水準 α = 0.05，右尾機率 $\frac{\alpha}{2}$ 標準常態值 $z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025}$ = 1.96(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。母體平均值信賴區間：$\bar{x} - z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$

下限：$\bar{x} - z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} = 12000 - 1.96 \times \frac{1000}{\sqrt{7}} = 12000 - 740.8 = 11259.2$ 元

上限：$\bar{x} + z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} = 12000 + 1.96 \times \frac{1000}{\sqrt{7}} = 12000 + 740.8 = 12740.8$ 元

每日營業額平均值 μ 之 95 %信賴區間(CI)為 $11259.2 \le \mu \le 12740.8$ 元。表示上週七天營業日營業額，所有可能樣本平均值在(11259.2, 12740.8)區間有 95 %的機會包含母體平均值 μ 在內。

範例 8.7 由一常態分布母體 $N(\mu_x, 7^2)$ 隨機獲得 $x_1, x_2, x_3, \ldots, x_{10}$ 樣本，再由另一常態分布母體 $N(\mu_y, 5^2)$ 隨機獲得 $y_1, y_2, y_3, \ldots, y_{12}$ 樣本，x_i 與 y_j 相互獨立。則 $\bar{x} = \frac{\sum_{i=1}^{10} x_i}{10}$ 與 $\bar{y} = \frac{\sum_{i=1}^{12} y_i}{12}$。(A)請寫出 $\bar{x} - \bar{y}$ 之抽樣分布；(B)若 $\bar{x}$ = 60 與 $\bar{y}$ = 50，計算 $\mu_x - \mu_y$ 之 90 %信賴區間，並解釋其意涵。(答案有效位數四捨五入取到小數點後第 2 位)

題解：$X \sim N(\mu_x, 7^2)$，$n_x = 10$，樣本平均值 $\bar{x} \sim N(\mu_x, \frac{7^2}{10})$，$Y \sim N(\mu_y, 5^2)$，$n_y = 12$，樣本平均值 $\bar{y} \sim N(\mu_y, \frac{5^2}{12})$。

(A) $\bar{x} - \bar{y} \sim N(\mu_x - \mu_y, \frac{7^2}{10} + \frac{5^2}{12})$

(B) 90 %信賴區間公式為 $(\bar{x} - \bar{y}) - 1.645 \times \sigma_{\bar{x}-\bar{y}} \le \mu_x - \mu_y \le (\bar{x} - \bar{y}) + 1.645 \times \sigma_{\bar{x}-\bar{y}}$

下限：$(\bar{x} - \bar{y}) - 1.645 \times \sigma_{\bar{x}-\bar{y}} = (60 - 50) - 1.645 \times \sqrt{\frac{7^2}{10} + \frac{5^2}{12}} = 10 - 4.3471 = 5.5629$

上限：$(\bar{x} - \bar{y}) + 1.645 \times \sigma_{\bar{x}-\bar{y}} = (60 - 50) + 1.645 \times \sqrt{\frac{7^2}{10} + \frac{5^2}{12}} = 10 + 4.3471 = 14.3471$

$\mu_x - \mu_y$ 之 90 %信賴區間(CI)為 $5.5629 \le \mu_x - \mu_y \le 14.3471$，表示有 90 %信心水準相信 $\mu_x - \mu_y$ 會位在 5.56 到 14.35 區間範圍內。

8.2.2 母體屬常態分布和母體變異數未知

母體屬於常態分布時，在母體變異數 σ^2 和標準(偏)差 σ 未知的情況下，樣本數量大時($n \ge 30$)，可以利用樣本標準(偏)差 S 取代母體標準(偏)差 σ，而估算母體平均值 μ 的信賴區間為 $\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}}$。在樣本數量較少($n < 30$)時，其標準化 z 值 $z = \frac{\bar{x}-\mu}{\frac{S}{\sqrt{n}}}$ 的分布並不近似標準常態分布(Standard normal distribution)，而是適用於自由度(degree of freedom, df) $n - 1$ 之 t 分布(t distribution)。母體屬於常態分布，而樣本數量較少 $n < 30$，母體變異數 σ^2 和標準(偏)差 σ 未知的情況下，應以 t 分布估算母體平均值 μ 的信賴區間。

8.2.2.1 *t* 分布

母體屬於常態分布(Normal distribution)時，其樣本平均值 $\bar{x}$ 的抽樣分布亦趨近於常態分布，$\bar{x} \sim N(\mu, \frac{\sigma^2}{n})$。將樣本平均值 $\bar{x}$ 標準化，標準常態分布為：

$$z = \frac{\bar{x}-\mu}{\sigma_{\bar{x}}} = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$$

在實際應用上母體變異數(變方)σ^2 或標準(偏)差 σ 經常未能獲知，利用樣本標準(偏)差 S 代替母體標準(偏)差 σ，可獲得自由度(degree of freedom, *df*) $n-1$ 之 t 分布(t distribution, t statistic or t score)。Gosset 稱為 t 值：t 值與 z 值同屬於無因次單位。

$$t = \frac{\bar{x}-\mu}{S_{\bar{x}}} = \frac{\bar{x}-\mu}{\frac{S}{\sqrt{n}}}$$

此分布又稱為學生氏 t 分布(Student's t-distribution)。t 分布之密度函數(density function)

$$f(t) = \frac{k}{\left(1+\frac{t^2}{v}\right)^{\frac{\times(v+1)}{2}}}，其中 -\infty < t < \infty$$

v(英文讀音 nu)為樣本變異數(S^2)的自由度 $v = n-1$，k 為一常數(constant)。

8.2.2.2 *t* 分布性質

$f(t)$式之 t 值分布為一理論抽樣分布(theoretical sampling distribution)。設常態分布之母體大小為 N，從此母體隨機抽取 n 個觀測值為樣本，所有可能樣本有 N^n 個(組合)，可求得 N^n 個(組合)t 值，即為 t 分布。

範例 8.8　一母體為 2、4 和 6 即基本單位數量 $N = 3$，從此母體隨機抽取 2 個觀測值($n = 2$)為一樣本，所有可能樣本組合有 $3^2 = 9$ 個，可計算獲得 9 個 t 值

樣本		平均值$\bar{x}$	S^2	樣本標準(偏)差 S	t 值
2	2	2	0	0.0000	$-\infty$
2	4	3	2	1.4142	–1
2	6	4	8	2.8284	0
4	2	3	2	1.4142	–1
4	4	4	0	0.0000	0
4	6	5	2	1.4142	1
6	2	4	8	2.8284	0
6	4	5	2	1.4142	1
6	6	6	0	0.0000	$+\infty$

母體平均值 $\mu = 4$

t 值次數分布表

t 值	$-\infty$	-1	0	1	$+\infty$
次數	1	2	3	2	1
機率	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{3}{9}$	$\frac{2}{9}$	$\frac{1}{9}$

t 值的期望值 $E(t)$為 0

$$E(t) = \frac{-\infty-1+0-1+0+1+0+1+\infty}{9} = 0$$

t 值分布之平均值為 0，以此為中心左右對稱，當母體基本單位數量 N 很大，樣本數量 n 也夠大時，t 值分布會非常接近標準常態分布之型態。

t 分布的變異數 $V(t)$

$$V(t) = \frac{v}{v-2}$$

t 分布的標準(偏)差 σ_t

$$\sigma_t = \sqrt{V(t)} = \sqrt{\frac{v}{v-2}}$$

當自由度 $v > 2$ 或樣本數量 $n > 3$ 時，$V(t) > 1$，故一般情況下，樣本數量 n 大於 3 時，t 分布的變異數皆會大於 1。當樣本數量 n 很大($n > 31$)時，自由度 $v = n-1$ 亦會變大，t 分布的變異數會接近 1，即 $V(t) \fallingdotseq 1$，與標準常態分布 $N(0,1)$的變異數 $V(z) = 1$ 接近。故當樣本數量大量增加($n > 31$)時，t 分布的平均值會趨

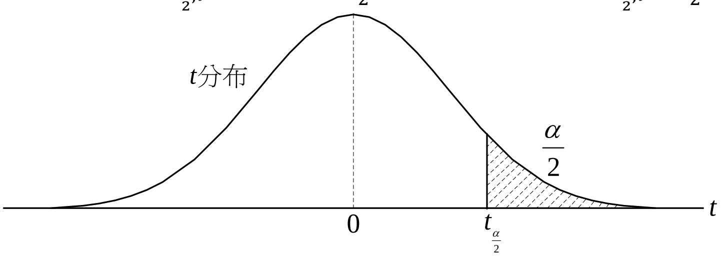
近於標準常態分布的平均值 0；*t* 分布的變異數會趨近於標準常態分布的變異數 1。因此，在樣本數量大量增加($n > 31$)時，*t* 分布會趨近於標準常態分布。因此，故當樣本數量高於 31($n > 31$)時，自由度 $v > 30$，可以利用標準常態分布取代 *t* 分布。

樣本數量 *n* 不同，因此 *t* 分布並非如標準化值 *Z* 為唯一曲線分布，而是一組曲線分布。樣本變異數 S^2 的自由度為 $v = n - 1$，所以 *t* 分布曲線隨自由度 *v* 數值而異，通常皆以 t_v 代表 *t* 分布曲線。

自由度

以樣本的統計量(值)來估計母體參數時，樣本中獨立或能自由變化的資料之個數，稱為該統計量(值)的自由度(degree of freedom, *df*)。例如：在估算樣本變異數 $S^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$ 或樣本標準(偏)差 $S = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$ 時，必須先計算樣本平均值 $\bar{x}$，在獲得樣本平均值之後，在所有 *n* 個觀測值中，只要確定 $n-1$ 個觀測值的數值，即可透過樣本平均值和 $n-1$ 個觀測值的數值，即可獲得剩下 1 個觀測值的數值。從此情況而論，可以自由調整的觀測值僅有 $n-1$ 個觀測值，故，樣本變異數 S^2 的自由度為 $v = n - 1$。

$t_{\frac{\alpha}{2}}$ 表示 *t* 分布**右**尾(標示文字在**右**下角)機率(面積)為 $\frac{\alpha}{2}$ 的 *t* 值。$t_{\frac{\alpha}{2},v}$ 代表在 *t* 分布中，自由度 $v = n - 1$ 時，右尾機率(面積)為 $\frac{\alpha}{2}$ 的 *t* 值。因此，從 $t_{\frac{\alpha}{2},v}$ 到 $+\infty$ 機率為 $\frac{\alpha}{2}$，亦可表示為 $P(t > t_{\frac{\alpha}{2},v}) = \frac{\alpha}{2}$。



標準常態分布與 *t* 分布圖形比較

t 分布圖形與標準常態分布圖形相似

都具有對稱於零、單峰及鐘形的特性

t 分布圖形散佈(spread)比標準常態分布圖形大，*t* 分布圖形尾端(tails)具有較高機率

以 $\frac{S}{\sqrt{n}}$ 替代 $\frac{\sigma}{\sqrt{n}}$ 標準化，使得 *t* 分布有較大的變異性。

t 分布自由度 *v* 愈大時，*t* 分布圖形愈趨近標準常態分布圖形

樣本數 *n* 越大，樣本標準(偏)差 s 估計母體標準(偏)差 σ 越準，估計值造成的額外變異性越少。

練習 8.16　*t* 分布出現極端值機率比常態分布出現極端值機率大還是小？為什麼？

題解：變異數 $V(t) = \frac{v}{v-2} > V(Z) = 1$，因為 *t* 分布的變異數比常態分布大，故 *t* 分布的分散程度比常態分布廣，因此，*t* 分布出現極端值機率比常態分布大。

t 分布 Excel 函數

利用 Excel 軟體中插入(I)→函數(F)…→在插入函數對話方塊中選取類別(C): **統計**，選取函數(N): **TDIST**→確定。在函數引數對話視窗中，x 方塊輸入：欲推估 *t* 分布的數值間 *x*；Deg_freedom 方塊輸入：自由度 *v*；Tails 方塊輸入：分布尾數個數，1 代表單尾分布；2 代表雙尾分布。確定。即會在原先選定的儲存格中出現 *t* 分布機率數值。TDIST(x,deg_freedom,tails)。

利用 Excel 軟體中插入(I)→函數(F)…→在插入函數對話方塊中選取類別(C): **統計**，選取函數(N): **TINV**→確定。在函數引數對話視窗中，Probability 方塊輸入：**雙尾** *t* 分布機率數值 *α*；Deg_freedom 方塊輸入：自由度 *v*。確定。即會在原先選定的儲存格中出現 *t* 分布的 *t* 數值。TINV(probability,deg_freedom)。

8.2.2.3 利用 t 分布推估母體平均值的信賴區間

母體資料屬於常態分布，母體變異數和標準差未知，都可以使用 t 分布進行信賴區間精準估算。樣本數高於 30，使用標準化 Z 值可以獲得信賴區間近似數值。樣本數量少於 30 時，若使用標準化 Z 值進行信賴區間估算，誤差量會比較大。

故，樣本數量少($n < 30$)，母體屬於常態分布，母體變異數 σ^2 和標準(偏)差 σ 未知時，母體平均值 μ 的信賴區間【點估計值±誤差範圍(臨界值×標準誤)】：

$$\bar{x} \pm t_{\frac{\alpha}{2},v=n-1} \times \frac{S}{\sqrt{n}}$$

其中 $1-\alpha$：信賴係數　　　α：信賴區間的顯著水準。

$t_{\frac{\alpha}{2},v=n-1}$：在自由度 $v = n-1$，**右**尾(文字標示於**右**下角)機率(面積)為$\frac{\alpha}{2}$的 t 值。

S：樣本標準(偏)差 $S = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$。

$\bar{x}$：樣本平均值。

v：自由度(degree of freedom, df)是在計算$\sum_{i=1}^{n}(x_i-\bar{x})^2$時，樣本所需提供的獨立資料的個數。

在計算$\sum_{i=1}^{n}(x_i-\bar{x})^2$時需要 n 個觀測值(資料)個數，包括 $x_1-\bar{x}, x_2-\bar{x}, x_3-\bar{x},\ldots, x_n-\bar{x}$。另外可知 $\sum_{i=1}^{n}(x_i-\bar{x}) = 0$，故只有 $n-1$ 個資料是屬於獨立性質。因此，只要預先知道 $n-1$ 個 $x_i-\bar{x}$ 的數值，最後一個 $x_i-\bar{x}$數值可以透過$\sum_{i=1}^{n}(x_i-\bar{x}) = 0$ 推估獲得。於是：樣本變異數 S^2 的自由度為 $v = n-1$。

$t_{\frac{\alpha}{2},v=n-1} \times \frac{S}{\sqrt{n}}$：為抽樣誤差(sampling error)、誤差界限、誤差範圍(margin of error)、最大誤差或可能機誤。臨界值 × 標準誤(差)。

$2 \times t_{\frac{\alpha}{2},v=n-1} \times \frac{S}{\sqrt{n}}$：為信賴區間的寬度(width)、長度。

範例 8.9　在自由度 10 的 t 分布中，請找出 $t_{0.05,10}$ 數值，該點在 t 分布右尾機率 $p = 0.05$。(答案有效位數四捨五入取到小數點後第 4 位)

題解：$P(t > 1.8125) = 0.05$(使用 Excel 軟體 T.INV 函數輸入左尾機率 0.95 自由度 10 查詢獲得；使用 TINV 函數輸入雙尾機率 0.10 自由度 10 查詢獲得)

答案：$t_{0.05,10} = 1.8125$

範例 8.10　在自由度 10 的 t 分布中，請找出 t 數值，該點在 t 分布左尾機率 $p = 0.05$。(答案有效位數四捨五入取到小數點後第 4 位)

題解：$P(t < -1.8125) = 0.05$(使用 Excel 軟體 T.INV 函數輸入左尾機率 0.05 自由度 10 查詢獲得；使用 TINV 函數輸入雙尾機率 0.10 自由度 10 查詢後再加上-號獲得)

答案：$t_{0.95,10} = -t_{0.05,10} = -1.8125$

練習 8.17　在自由度 10 的 t 分布中，請找出 $t_{0.975,10}$ 和 $t_{0.025,10}$ 數值，前述兩點屬於左右兩側機率皆為 $p = 0.025$。(答案有效位數四捨五入取到小數點後第 4 位)

題解：$P(t < -2.2281) = 0.025$，$P(t > 2.2281) = 0.025$(使用 Excel 軟體 T.INV 函數查詢獲得；或 T.INV.2T 函數輸入雙尾機率獲得正值 t 值)

答案：$t_{0.975,10} = -t_{0.025,10} = -2.2281$；$t_{0.025,10} = 2.2281$

範例 8.11　在自由度 26 的 t 分布中，請找出 $t_{0.01,26}$ 數值，該點在 t 分布右尾機率 0.01。

題解：$P(t > 2.4786) = 0.01$(使用 Excel 軟體 T.INV 函數輸入左尾機率 0.99 自由度 26 查詢獲得使用 TINV 函數輸入雙尾機率 0.02 自由度 26 查詢獲得)

答案：$t_{0.01,26} = 2.4786$

範例 8.12 奇遇海產店依據過去的資料分析顯示每日營業額(新台幣：元)符合常態分布，上星期 7 天營業日每日營業額如下表所示，試求每日營業額平均值 μ 之 95 %信賴區間？(答案有效位數四捨五入取到小數點後第 1 位)

題解：

營業日	營業額 x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	16000	0	0
2	18000	2000	4000000
3	17500	1500	2250000
4	16500	500	250000
5	12000	-4000	16000000
6	15000	-1000	1000000
7	17000	1000	1000000
合計	112000	0	24500000

樣本平均值$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{112000}{7} = 16000$，樣本標準(偏)差 $S = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}} = \sqrt{\frac{24500000}{7-1}} = 2020.7$。

自由度 $v = n - 1 = 7 - 1 = 6$，信賴水準 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，$t_{\frac{\alpha}{2},v=n-1} = t_{\frac{0.05}{2},7-1} = t_{0.025,6} = 2.4469$(使用 Excel 軟體 T.INV 函數查詢獲得)。

母體平均值信賴區間：$\bar{x} - t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} \le \mu \le \bar{x} + t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}}$

下限：$\bar{x} - t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} = 16000 - 2.4469 \times \frac{2020.7}{\sqrt{7}} = 16000 - 1868.8 = 14131.2$ 元

上限：$\bar{x} + t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} = 16000 + 2.4469 \times \frac{2020.7}{\sqrt{7}} = 16000 + 1868.8 = 17868.8$ 元

故每日營業額平均值 μ 之 95 %信賴區間(CI)為 $14131.2 \le \mu \le 17868.8$ 元。表示上週七天營業日營業額，所有可能樣本平均值在(14131.2, 17868.8)區間有 95 %的機會包含(母體)每日營業額平均值 μ 在內。

練習 8.18 阿文連鎖飲料店聲稱其珍珠奶茶每杯容量體積皆為 650 ml，容量分布屬於常態分布。現從其產品中隨機抽取 8 件樣本，測量其體積分別為 620、655、670、635、665、648、641 和 642 ml。請估算每杯珍珠奶茶平均容量的 95 %信賴區間。(答案有效位數四捨五入取到小數點後第 1 位)

題解：樣本平均值$\bar{x} = 647.00$ ml，樣本標準(偏)差 S = 16.27 ml。自由度 $v = n - 1 = 8 - 1 = 7$，信賴水準 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，$t_{\frac{\alpha}{2},v=n-1} = t_{\frac{0.05}{2},8-1} = t_{0.025,7} = 2.3646$(使用 Excel 軟體 T.INV 函數查詢獲得)。母體平均值信賴區間：$\bar{x} - t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} \le \mu \le \bar{x} + t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}}$

下限：$\bar{x} - t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} = 647 - 2.3646 \times \frac{16.27}{\sqrt{8}} = 647 - 13.6 = 633.4$ ml

上限：$\bar{x} + t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} = 647 + 2.3646 \times \frac{16.27}{\sqrt{8}} = 647 + 13.6 = 660.6$ ml

答案：每杯珍珠奶茶平均容量之 95 %信賴區間(CI)為 $633.4 \le \mu \le 660.6$ ml

練習 8.19 阿文連鎖飲料店聲稱其珍珠奶茶每杯容量體積皆為 600 ml，容量分布屬於常態分布。現從其產品中隨機抽取 25 件樣本，測量其平均體積為 586.6 ml，標準(偏)差為 35.7 ml。請估算每杯珍珠奶茶平均容量的 95 %信賴區間。(答案有效位數四捨五入取到小數點後第 1 位)

題解：樣本平均值$\bar{x}$ = 586.6 ml，樣本標準(偏)差 S = 35.7 ml，自由度 $v = n - 1 = 25 - 1 = 24$，信賴水準 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，$t_{\frac{\alpha}{2},v=n-1} = t_{\frac{0.05}{2},25-1} = t_{0.025,24} = 2.0639$(使用 Excel 軟體 T.INV 函數查詢獲得)。母體平均值信賴區間：$\bar{x} - t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} \le \mu \le \bar{x} + t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}}$

下限：$\bar{x} - t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} = 586.6 - 2.0639 \times \frac{35.7}{\sqrt{25}} = 586.6 - 14.74 = 571.9$ ml

上限：$\bar{x} + t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} = 586.6 + 2.0639 \times \frac{35.7}{\sqrt{25}} = 586.6 + 14.74 = 601.3$ ml

答案：每杯珍珠奶茶平均容量之 95 %信賴區間(CI)為 $571.9 \le \mu \le 601.3$ ml

練習 8.20　奇遇海產店依據過去的資料分析顯示每日營業額(新台幣：元)符合常態分布，前 14 天營業日每日營業額如下表所示，(a)試求每日營業額平均值 μ 之 95 %信賴區間？(b)試求每日營業額平均值 μ 之 90 %信賴區間？(答案有效位數四捨五入取到小數點後第 1 位)

題解：

營業日	營業額 x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	16000	250	62500
2	18000	2250	5062500
3	17500	1750	3062500
4	16500	750	562500
5	12000	-3750	14062500
6	14000	-1750	3062500
7	15200	-550	302500
8	15200	-550	302500
9	12000	-3750	14062500
10	18000	2250	5062500
11	15600	-150	22500
12	17000	1250	1562500
13	16500	750	562500
14	17000	1250	1562500
合計			

樣本平均值$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{220500}{14} = 15750$，樣本標準(偏)差 $S = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}} = \sqrt{\frac{49315000}{14-1}} = 1947.6811$，自由度 $v = n - 1 = 14 - 1 = 13$。

(a)信賴水準 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，$t_{\frac{\alpha}{2},v=n-1} = t_{\frac{0.05}{2},14-1} = t_{0.025,13} = 2.1604$(使用 Excel 軟體 T.INV 函數查詢獲得)。母體平均值信賴區間：$\bar{x} - t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} \le \mu \le \bar{x} + t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}}$

下限：$\bar{x} - t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} = 15750 - 2.1604 \times \frac{1947.6811}{\sqrt{14}} = 15750 - 1124.6 = 14625.4$ 元

上限：$\bar{x} + t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} = 15750 + 2.1604 \times \frac{1947.6811}{\sqrt{14}} = 15750 + 1124.6 = 16874.6$ 元

故每日營業額 μ 之 95 %信賴區間(CI)為 $14625.4 \le \mu \le 16874.6$ 元。表示上週七天營業日營業額，所有可能樣本平均值在(14625.4, 16874.6)區間有 95 %的機會包含(母體)每日營業額平均值 μ 在內。

(b)信賴水準 $1 - \alpha = 0.90$，顯著水準 $\alpha = 0.10$，$t_{\frac{\alpha}{2},v=n-1} = t_{\frac{0.10}{2},14-1} = t_{0.05,13} = 1.7709$(使用 Excel 軟體 T.INV 函數查詢獲得)。母體平均值信賴區間：$\bar{x} - t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} \le \mu \le \bar{x} + t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}}$

下限：$\bar{x} - t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} = 15750 - 1.7709 \times \frac{1947.6811}{\sqrt{14}} = 15750 - 921.8 = 14828.2$ 元

上限：$\bar{x} + t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} = 15750 + 1.7709 \times \frac{1947.6811}{\sqrt{14}} = 15750 + 921.8 = 16671.8$ 元

故每日營業額 μ 之 90 %信賴區間(CI)為 $14828.2 \le \mu \le 16671.8$ 元。表示上週七天營業日營業額，所有可能樣本平均值在(14828.2, 16671.8)區間有 90 %的機會包含(母體)每日營業額平均值 μ 在內。

練習 8.21　請依據下列自由度數值，找出 t 分布中右尾機率 0.05 的 t 值：(a) 5；(b) 10；(c) 15；(d) 20；(e) 25。

練習 8.22　在標準常態分布中有 95 %機率介於$-z_{0.025} = -1.96$ 和 $z_{0.025} = 1.96$ 之間。請下列自由度中，找出其對應的 t 值(即，找出$-t_{0.025}$ 和 $t_{0.025}$ 數值)：(a) 5；(b) 10；(c) 15；(d) 20；(e) 25。

練習 8.23　在自由度 $df = 10$ 的 t 分布曲線中，請使用 t 值表格找出下列的 t 值：(a) $t_{0.01}$；(b) $t_{0.025}$；(c) $t_{0.05}$；(d) $t_{0.10}$。

練習 8.24　深水大學對其學生每週外食次數(屬於常態分布)進行調查。透過隨機抽樣 120 位學生為樣本，其每週外食平均次數 8.50 次，標準(偏)差 2.50 次。請建立每週學生外食次數平均值 98 %的信賴區間。(答案有效位數四捨五入取到小數點後第 2 位)

題解：樣本數量 $n = 120$，信賴水準 $1 - \alpha = 0.98$，顯著水準 $\alpha = 0.02$，$z_{\frac{\alpha}{2}} = z_{0.01} = 2.3264$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)，$t_{\frac{\alpha}{2},v=n-1} = t_{\frac{0.02}{2},120-1} = t_{0.01,119} = 2.3580$(使用 Excel 軟體 T.INV 函數查詢獲得)。母體平均值信賴區間：$\bar{x} - t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} \le \mu \le \bar{x} + t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}}$

標準化 Z 值運算下限：$\bar{x} - z_{\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}} = 8.50 - 2.3264 \times \frac{2.5}{\sqrt{120}} = 8.50 - 0.53 = 7.97$

標準化 Z 值運算上限：$\bar{x} + z_{\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}} = 8.50 + 2.3264 \times \frac{2.5}{\sqrt{120}} = 8.50 + 0.53 = 9.03$

t 值運算下限：$\bar{x} - t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} = 8.50 - 2.3580 \times \frac{2.5}{\sqrt{120}} = 8.50 - 0.54 = 7.96$

t 值運算下限：$\bar{x} + t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} = 8.50 + 2.3580 \times \frac{2.5}{\sqrt{120}} = 8.50 + 0.54 = 9.04$

答案：每週學生外食次數平均值 98 %的信賴區間(CI)為 $7.97 \le \mu \le 9.03$ 次或 $7.96 \le \mu \le 9.04$ 次

練習 8.25　隨機抽取台灣國際觀光旅館經理級主管 12 位，調查其年薪(屬於常態分布)，結果顯示平均年薪為新台幣 1890000 元，標準(偏)差為新台幣 560000 元。試估算(A)台灣國際觀光旅館經理級主管平均年薪之 95 %信賴區間；(B)台灣國際觀光旅館經理級主管平均年薪之 99 %信賴區間；(C)在 99 %信賴水準下抽樣誤差；(D)台灣國際觀光旅館經理級主管平均年薪信賴水準 99 %的信賴區間寬度？(答案有效位數取到個位數)

題解：

(A)樣本數量 $n = 12$，信賴水準 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，$t_{\frac{\alpha}{2},v=n-1} = t_{\frac{0.05}{2},12-1} = t_{0.025,11} = 2.2010$(使用 Excel 軟體 T.INV 函數查詢獲得)。

$$\bar{x} - t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} \le \mu \le \bar{x} + t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}}$$

$$1890000 - 2.2010 \times \frac{560000}{\sqrt{12}} \le \mu \le 1890000 + 2.2010 \times \frac{560000}{\sqrt{12}}$$

$$1534191 \le \mu \le 2245809 \text{ 元}$$

(B)樣本數量 $n = 12$，信賴水準 $1 - \alpha = 0.99$，顯著水準 $\alpha = 0.01$，$t_{\frac{\alpha}{2},v=n-1} = t_{\frac{0.01}{2},12-1} = t_{0.005,11} = 3.1058$(使用 Excel 軟體 T.INV 函數查詢獲得)。

$$\bar{x} - t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} \le \mu \le \bar{x} + t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}}$$

$$1890000 - 3.1058 \times \frac{560000}{\sqrt{12}} \le \mu \le 1890000 + 3.1058 \times \frac{560000}{\sqrt{12}}$$

$$1387922 \le \mu \le 2392078 \text{ 元}$$

(C)99 %信賴水準下抽樣誤差$t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} = 3.1058 \times \frac{560000}{\sqrt{12}} = 502078$ 元

(D)信賴水準 99 %的信賴區間寬度 $2 \times t_{\frac{\alpha}{2},\nu} \times \frac{S}{\sqrt{n}} = 2 \times 3.1058 \times \frac{560000}{\sqrt{12}} = 1004155$ 元

答案：(A)95 %信賴區間 $1534191 \le \mu \le 2245809$ 元；(B)99 %信賴區間 $1387922 \le \mu \le 2392078$ 元；(C)99 %信賴水準下抽樣誤差 = 502078 元；(D)信賴水準 99 %的信賴區間寬度 1004155 元

範例 8.13 由常態分布 $N(\mu,\sigma^2)$母體中隨機抽出 x_1、x_2、x_3、x_4 和 x_5 五個變量，請計算母體平均值的 95 %信賴區間(A)母體變異數 σ^2 已知；(B)母體變異數 σ^2 未知。(信賴區間有效位數取到個位數)

題解：樣本平均值$\bar{x} = \frac{x_1+x_2+x_3+x_4+x_5}{5}$，樣本平均值亦趨近於常態分布$\bar{x} \sim N(\mu,\frac{\sigma^2}{5})$

樣本變異數 $S^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$，樣本標準(偏)差 $S = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$

(A)母體變異數 σ^2 已知

95 %信賴區間公式為$\bar{x} - 1.96 \times \sigma_{\bar{x}} \le \mu \le \bar{x} + 1.96 \times \sigma_{\bar{x}}$

$$\bar{x} - 1.96 \times \frac{\sigma}{\sqrt{5}} \le \mu \le \bar{x} + 1.96 \times \frac{\sigma}{\sqrt{5}}$$

$$\bar{x} - 0.8765 \times \sigma \le \mu \le \bar{x} + 0.8765 \times \sigma$$

(B)母體變異數 σ^2 未知，利用樣本變異數 S^2 取代母體變異數 σ^2 推估信賴區間。

95 %信賴區間公式為$\bar{x} - t_{\frac{\alpha}{2},\nu} \times \frac{S}{\sqrt{n}} \le \mu \le \bar{x} + t_{\frac{\alpha}{2},\nu} \times \frac{S}{\sqrt{n}}$，$t_{\frac{\alpha}{2},\nu} = t_{\frac{\alpha}{2},n-1} = t_{\frac{0.05}{2},5-1} = 2.7764$(使用 Excel 軟體 T.INV 函數查詢獲得)。

$$\bar{x} - 2.7764 \times \frac{S}{\sqrt{5}} \le \mu \le \bar{x} + 2.7764 \times \frac{S}{\sqrt{5}}$$

$$\bar{x} - 1.2416 \times S \le \mu \le \bar{x} + 1.2416 \times S$$

答案：(A) $\bar{x} - 0.8765 \times \sigma \le \mu \le \bar{x} + 0.8765 \times \sigma$；(B) $\bar{x} - 1.2416 \times S \le \mu \le \bar{x} + 1.2416 \times S$

練習 8.26 從一個樣本平均值 250.0 和樣本標準(偏)差 15.5 的母體中獲得 15 個隨機觀測值。(a)請計算母體平均值 90 %信賴區間；(b)母體平均值 95 %信賴區間；(c)母體平均值 99 %信賴區間；(d)假設樣本數量達 150，請重複前述 a、b 和 c 計算。

練習 8.27 一位台北市旅館總經理想要決定旅館中經理階層的平均年薪水準。隨機調查台北市 20 為旅館經理年薪獲得平均值$\bar{x}$ = NT\$75600 和標準(偏)差 S = NT\$15600。請計算台北市旅館經理階層平均年薪在下列信賴水準下的信賴區間：(a) 90 %；(b) 95 %；(c) 99 %

練習 8.28 假設有一隨機變數 X，可獲得其 26 個隨機樣本數，以估算母體平均值和標準(偏)差。下列三個母體平均值的信賴區間中，請分別列出信賴水準(confidence level)？(a) $\bar{x} \pm 2.0595\times\frac{S}{\sqrt{n}}$；(b) $\bar{x} \pm 2.4851\times\frac{S}{\sqrt{n}}$；(c) $\bar{x} \pm 1.7081\times\frac{S}{\sqrt{n}}$。

答案：(a)95 %；(b)98 %；(c)90 %

練習 8.29 假設從 101 個觀測值組成的隨機樣本中，獲得樣本平均值$\bar{x}$ = NT\$7000 和樣本標準(偏)差 S = NT\$600，假設原來母體屬於常態，使用 t 分布進行推估時，理論上正確，因為樣本數量大，故可以使用標準常態分布運算，即使母體變異數未知。請分別使用 t 分布和標準常態分布運算母體平均值 95 %信賴區間。(答案有效位數四捨五入取到小數點後第 2 位)

題解：母體平均值信賴區間：$\bar{x} - t_{\frac{\alpha}{2},\nu} \times \frac{S}{\sqrt{n}} \le \mu \le \bar{x} + t_{\frac{\alpha}{2},\nu} \times \frac{S}{\sqrt{n}}$

$$7000 - 1.9840 \times \frac{600}{\sqrt{101}} \le \mu \le 7000 + 1.9840 \times \frac{600}{\sqrt{101}} \rightarrow 6881.55 \le \mu \le 7118.45$$

$$\bar{x} - z_{\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}} \le \mu \le \bar{x} + z_{\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}}$$

$$7000 - 1.960 \times \frac{600}{\sqrt{101}} \le \mu \le 7000 + 1.960 \times \frac{600}{\sqrt{101}} \rightarrow 6882.98 \le \mu \le 7117.02$$

答案：t 分布 $6881.55 \le \mu \le 7118.45$；標準常態分布 $6882.98 \le \mu \le 7117.02$

練習 8.30 假設從樣本數 $n = 81$ 的隨機樣本中獲得$\sum_{i=1}^{n} x_i = 3500$ 和$\sum_{i=1}^{n}(x_i - \bar{x})^2 = 8500$。(a)請計算樣本平均值和樣本標準(偏)差；(b)請利用 t 分布計算母體平均值 95 %信賴區間，運用表格查詢 t 值時，可以使用大約值即可；(c)利用標準常態分布計算母體平均值 95 %信賴區間。(答案有效位數四捨五入取到小數點後第 2 位)

題解：

練習 8.31 修讀本班統計學人數有 60 人，此次期中考試成績符合常態分布，透過隨機抽查 8 位學生成績分別為 65、75、88、95、68、80、85 和 92。請計算修讀本班統計學學生期中考試平均成績之 95 %信賴區間。(答案有效位數四捨五入取到小數點後第 2 位)

題解：母體總數 $N = 60$，樣本數量 $n = 8$，信賴水準 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，$t_{\frac{\alpha}{2},v=n-1} = t_{\frac{0.05}{2},8-1} = t_{0.025,7} = 2.3646$(使用 Excel 軟體 T.INV 函數查詢獲得)，樣本平均值$\bar{x} = 81.0$ 分，樣本標準(偏)差 $S = 10.9805$ 分。母體平均值信賴區間：$\bar{x} - t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} \le \mu \le \bar{x} + t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}}$

在有限母體中樣本平均值的抽樣分布，樣本平均值的變異數$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \times \frac{N-n}{N-1}$

母體變異數未知，使用樣本變異數取代 $V(\bar{x}) = \frac{S^2}{n} \times \frac{N-n}{N-1}$

下限：$\bar{x} - t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}} = 81.0 - 2.3646 \times \frac{10.9805}{\sqrt{8}} \times \sqrt{\frac{60-8}{60-1}} = 81.0 - 8.6182 = 72.38$ 分

上限：$\bar{x} + t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}} = 81.0 + 2.3646 \times \frac{10.9805}{\sqrt{8}} \times \sqrt{\frac{60-8}{60-1}} = 81.0 + 8.6182 = 89.62$ 分

答案：修讀本班統計學學生期中考試平均成績之 95 %信賴區間(CI)為 $72.38 \le \mu \le 89.62$ 分

8.2.3 母體分布不確定和母體變異數已知

母體的分布型態無法確定時，在母體變異數 σ^2 和標準(偏)差 σ 已知的情況下，樣本數量少($n < 30$)，可以利用柴比氏定理(Chebyshev's theorem, Bienayme-Chebyshev rule)進行母體平均值 μ 的信賴區間估算。在任何觀測值的分布資料中，至少有$(1 - \frac{1}{k^2})$比率或$(1 - \frac{1}{k^2}) \times 100$ %的資料，分布在算術平均值為中心，$\pm k$ 個標準(偏)差 S 的範圍內；樣本資料分布在$\bar{x} \pm k \times S$ 區間內，母體資料分布在 $\mu \pm k \times \sigma$ 區間內。

$$P(|\bar{x} - \mu| \le k \times \sigma_{\bar{x}}) \ge (1 - \frac{1}{k^2})$$
$$P(-k \times \sigma_{\bar{x}} \le \bar{x} - \mu \le k \times \sigma_{\bar{x}}) \ge (1 - \frac{1}{k^2})$$
$$P(-\bar{x} - k \times \sigma_{\bar{x}} \le -\mu \le -\bar{x} + k \times \sigma_{\bar{x}}) \ge (1 - \frac{1}{k^2})$$
$$P(\bar{x} + k \times \sigma_{\bar{x}} \ge \mu \ge \bar{x} - k \times \sigma_{\bar{x}}) \ge (1 - \frac{1}{k^2})$$
$$P(\bar{x} - k \times \sigma_{\bar{x}} \le \mu \le \bar{x} + k \times \sigma_{\bar{x}}) \ge (1 - \frac{1}{k^2})$$
$$P(\bar{x} - k \times \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + k \times \frac{\sigma}{\sqrt{n}}) \ge (1 - \frac{1}{k^2})$$

範例 8.14 奇遇海產店每日營業額分布情況不清楚，上星期 7 天營業日平均日營業額新台幣$\bar{x} = 12000$ 元，過去經驗獲得營業額的標準(偏)差為新台幣 $\sigma = 1000$ 元，試求每日營業額平均值 μ 之 95 %信賴區間？(答案有效位數四捨五入取到小數點後第 1 位)

題解：樣本數量 $n = 7$，樣本平均值$\bar{x} = 12000$ 元，母體標準(偏)差 $\sigma = 1000$ 元。

$$1 - \frac{1}{k^2} = 0.95 \rightarrow \frac{-1}{k^2} = 0.95 - 1 = -0.05 \rightarrow \frac{1}{k^2} = 0.05 \rightarrow \frac{1}{0.05} = 20 = k^2 \rightarrow k = \sqrt{k^2} = \sqrt{20} = 4.4721$$

母體平均值信賴區間：$\bar{x} - k \times \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + k \times \frac{\sigma}{\sqrt{n}}$

下限：$\bar{x} - k \times \frac{\sigma}{\sqrt{n}} = 12000 - 4.4721 \times \frac{1000}{\sqrt{7}} = 12000 - 1690.3 = 10309.7$ 元

上限：$\bar{x} + k \times \frac{\sigma}{\sqrt{n}} = 12000 + 4.4721 \times \frac{1000}{\sqrt{7}} = 12000 + 1690.3 = 13690.3$ 元

故每日營業額平均值 μ 之 95 %信賴區間(CI)為 $10309.7 \le \mu \le 13690.3$ 元。表示上週七天營業日營業額，所有可能樣本平均值在(10309.7, 13690.3)區間有 95 %的機會包含每日營業額平均值 μ 在內。

答案：每日營業額平均值 μ 之 95 %信賴區間(CI)為 $10309.7 \le \mu \le 13690.3$ 元

練習 8.32 觀光系產學攜手專班今年共有 345 人報名甄選，將錄取 25 名。若所有考生甄試成績非常態分布，其平均成績 76 分，標準(偏)差 6 分。若林小萱甄試成績 82 分，請估算其是否錄取？

題解：欲達錄取標準需在分布之達右側$\frac{25}{345}$ = 0.0725 機率以下，樣本數量少，在非常態分布中運用柴比氏定理推估母體平均值信賴區間

$$P(|x - \mu| \le k \times \sigma) \ge (1 - \frac{1}{k^2}) = 1 - 0.0725 \times 2 = 0.8551 \rightarrow k = 2.6268$$

$P(-k \times \sigma \le x - \mu \le k \times \sigma) \ge (1 - \frac{1}{k^2})$

$P(\mu - k \times \sigma \le x \le \mu + k \times \sigma) \ge (1 - \frac{1}{k^2})$

$\mu + k \times \sigma = 76 + 2.6268 \times 6 = 91.76 > 82$

答案：依據柴比氏定理推估林小萱成績在信賴區間內，故沒有錄取

8.2.4 母體分布不確定和母體變異數未知

母體的分布型態無法確定時，在母體變異數 σ^2 和標準(偏)差 σ 未知的情況下，樣本數量少($n < 30$)，無法利用柴比氏定理(Chebyshev's theorem, Bienayme-Chebyshev rule)進行母體平均值 μ 的信賴區間估算，必須使用無母數統計的方法進行估算。

表 母體平均值 μ 之信賴區間

母體屬性	條件	母體平均值信賴區間【點估計值±誤差範圍(臨界值×標準誤)】
常態分布或不確定 樣本數 $n \ge 30$	母體變異數 σ^2 已知	$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$
	母體變異數 σ^2 未知	$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{s}{\sqrt{n}}$ (近似) $\bar{x} \pm t_{\frac{\alpha}{2},v=n-1} \times \frac{s}{\sqrt{n}}$ (精準)
常態分布 樣本數 $n < 30$	母體變異數 σ^2 已知	$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$
	母體變異數 σ^2 未知	$\bar{x} \pm t_{\frac{\alpha}{2},v=n-1} \times \frac{s}{\sqrt{n}}$
分布不確定 樣本數 $n < 30$	母體變異數 σ^2 已知	$\bar{x} \pm k \times \frac{\sigma}{\sqrt{n}}$
	母體變異數 σ^2 未知	無母數統計的方法

在無法確定是否符合常態分布時，若樣本數量小於 30 個($n < 30$)，一般管理上都是期望以再增加樣本數量 n，來達到樣本數量多($n \ge 30$)時，利用標準常態分布的方式推估信賴區間。

8.3 母體比例區間估計

利用抽樣分布中的樣本比率$\bar{p}$，進行母體比率 p 的區間估計。在推估母體比率 p 的信賴區間時，必須先獲得樣本比率$\bar{p}$的抽樣分布。樣本比率$\bar{p}$的抽樣分布可以區分為大量樣本數($n \ge 30$)和小量樣本數($n < 30$)兩類。

8.3.1 母體比例區間估計：大量樣本數

樣本數量多($n \ge 30$)時，$n \times p \ge 5$ 同時必須 $n \times (1 - p) \ge 5$，樣本比率$\bar{p}$的抽樣分布會接近於常態分布，即為

$$樣本比率\bar{p} \sim N(p, \frac{p \times q}{n}) = N(p, \frac{p \times (1-p)}{n})$$

可以利用標準常態分布 z 值，估算抽樣誤差$|\bar{p} - p|$在$z_{\frac{\alpha}{2}} \times \sigma_{\bar{p}}$範圍內機率為 $1 - \alpha$，則

$$P(|\bar{p}-p| \le z_{\frac{\alpha}{2}} \times \sigma_{\bar{p}}) = P(|\bar{p}-p| \le z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}}) = P(|\bar{p}-p| \le z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times q}{n}}) = 1-\alpha$$

$$P(-z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}} \le \bar{p}-p \le z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}}) = 1-\alpha$$

$$P(-\bar{p}-z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}} \le -p \le -\bar{p}+z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}}) = 1-\alpha$$

$$P(\bar{p}+z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}} \ge p \ge \bar{p}-z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}}) = 1-\alpha$$

$$P(\bar{p}-z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}} \le p \le \bar{p}+z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}}) = 1-\alpha$$

在信賴係數和信賴水準(信賴度)為 $1-\alpha$，在信賴區間的顯著水準為 α，母體比例 p 的信賴區間為

$$\bar{p} \pm z_{\frac{\alpha}{2}} \times \sigma_{\bar{p}} \;=\; \bar{p} \pm z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}} \;=\; \bar{p} \pm z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times q}{n}}$$

其中 $1-\alpha$：信賴係數。　　　　　　　　α：信賴區間的顯著水準。

$z_{\frac{\alpha}{2}}$：標準常態分布右尾機率(面積)為$\frac{\alpha}{2}$的標準化 z 值。

$\sigma_{\bar{p}} = \sqrt{\frac{p\times(1-p)}{n}}$：為點估計值(樣本比率)抽樣分布的母體標準(偏)差 。

$z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times q}{n}}$：為抽樣誤差(sampling error)、誤差界限、誤差範圍(margin of error)或最大誤差。

$2 \times z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times q}{n}}$：為信賴區間的寬度或長度。

若母體比例 p 值未知時，可以利用樣本比例$\bar{p}$取代母體比例 p，以推估母體比例 p 的信賴區間【點估計值±誤差範圍(臨界值×標準誤)】：

$$\bar{p} \;\pm\; z_{\frac{\alpha}{2}} \times \sqrt{\frac{\bar{p}\times(1-\bar{p})}{n}}$$

若母體比例 p 值未知時，可以將母體比例 p 設為 0.50，以推估母體比例 p 的信賴區間，惟若設定 p = 0.50 所獲得的信賴區間較大，也較保守：

$$\bar{p} \;\pm\; z_{\frac{\alpha}{2}} \times \sqrt{\frac{0.50\times0.50}{n}}$$

範例 8.15　為調查高雄市消費者對有機飲食的支持比例，希望在信賴水準估計 95 %，從高雄市消費者中隨機抽取 300 位消費者為樣本，統計其對有機飲食的支持比例$\bar{p}$ = 0.56，請計算高雄市消費者對有機飲食支持比率的信賴區間？若過去曾有相關研究調查資料顯示有機飲食的支持比例 p = 0.4。(答案有效位數四捨五入取到小數點後第 4 位)

題解：樣本比例$\bar{p}$ = 0.56，母體比例 p = 0.4，樣本數量 n = 300，信賴水準 $1-\alpha$ = 0.95，顯著水準 α = 0.05，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025}$ = 1.96(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。母體比例

信賴區間：$\bar{p} - z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}} \le p \le \bar{p} + z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}}$

下限：$\bar{p} - z_{\frac{\alpha}{2}}\times\sqrt{\frac{p\times(1-p)}{n}} = 0.56 - 1.96\times\sqrt{\frac{0.4\times(1-0.4)}{300}} = 0.56 - 0.0554 = 0.5046$

上限：$\bar{p} + z_{\frac{\alpha}{2}}\times\sqrt{\frac{p\times(1-p)}{n}} = 0.56 + 1.96\times\sqrt{\frac{0.4\times(1-0.4)}{300}} = 0.56 + 0.0554 = 0.6154$

故支持有機飲食比率 p 之 95 %信賴區間(CI)為 $0.5046 \le p \le 0.6154$

若質疑以往對有機飲食的研究數值，不採用先前調查獲知的母體比例 p = 0.4，而以樣本比例$\bar{p}$ = 0.56 取代母體比例 p 以進行對母體比例 p 的信賴區間推估：

母體比例信賴區間：$\bar{p} - z_{\frac{\alpha}{2}} \times \sqrt{\frac{\bar{p}\times(1-\bar{p})}{n}} \le p \le \bar{p} + z_{\frac{\alpha}{2}} \times \sqrt{\frac{\bar{p}\times(1-\bar{p})}{n}}$

下限：$\bar{p} - z_{\frac{\alpha}{2}} \times \sqrt{\frac{\bar{p}\times(1-\bar{p})}{n}} = 0.56 - 1.96 \times \sqrt{\frac{0.56\times(1-0.56)}{300}} = 0.56 - 0.0562 = 0.5038$

上限：$\bar{p} + z_{\frac{\alpha}{2}} \times \sqrt{\frac{\bar{p}\times(1-\bar{p})}{n}} = 0.56 + 1.96 \times \sqrt{\frac{0.56\times(1-0.56)}{300}} = 0.56 + 0.0562 = 0.6162$

故支持有機飲食比率 p 之 95 %信賴區間(CI)為 $0.5038 \le p \le 0.6162$，信賴區間變大。

若質疑以往對有機飲食的研究數值，不值得信賴，不採用先前調查獲知的母體比例 $p = 0.4$，而以 0.50 取代母體比例 p 以進行對母體比例 p 的信賴區間推估：

母體比例信賴區間：$\bar{p} - z_{\frac{\alpha}{2}} \times \sqrt{\frac{0.50\times0.50}{n}} \le p \le \bar{p} + z_{\frac{\alpha}{2}} \times \sqrt{\frac{0.50\times0.50}{n}}$

下限：$\bar{p} - z_{\frac{\alpha}{2}} \times \sqrt{\frac{0.5\times0.5}{n}} = 0.56 - 1.96 \times \sqrt{\frac{0.5\times0.5}{300}} = 0.56 - 0.0566 = 0.5034$

上限：$\bar{p} + z_{\frac{\alpha}{2}} \times \sqrt{\frac{0.5\times0.5}{n}} = 0.56 + 1.96 \times \sqrt{\frac{0.5\times0.5}{300}} = 0.56 + 0.0566 = 0.6166$

故支持有機飲食比率 p 之 95 %信賴區間(CI)為 $0.5034 \le p \le 0.6166$，信賴區間再變大。

比較不同估算情況下的**信賴區間大小**。

練習 8.33　提供信賴區間，計算信賴水準

練習 8.34　假設高雄市有 1580000 人，透過隨機抽樣 560 人，發現其中有 263 人支持興建高雄港跨港纜車。請估算高雄市支持跨港纜車比率 p 的 99 %信賴區間。(答案有效位數四捨五入取到小數點後第 4 位)

題解：樣本支持比例$\bar{p} = \frac{263}{560} = 0.4696$，樣本數量 $n = 560$，信賴水準 $1 - \alpha = 0.99$，顯著水準 $\alpha = 0.01$，$z_{\frac{\alpha}{2}} = z_{\frac{0.01}{2}} = z_{0.005} = 2.5758$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。

母體比例信賴區間：$\bar{p} - z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}} \le p \le \bar{p} + z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}}$

下限：$\bar{p} - z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}} = 0.4696 - 2.5758 \times \sqrt{\frac{0.4696\times(1-0.4696)}{560}} = 0.4696 - 0.0543 = 0.4153$

上限：$\bar{p} + z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}} = 0.4696 + 2.5758 \times \sqrt{\frac{0.4696\times(1-0.4696)}{560}} = 0.4696 + 0.0543 = 0.5240$

答案：故支持跨港纜車比率 p 之 99 %信賴區間(CI)為 $0.4153 \le p \le 0.5240$

練習 8.35　阿文連鎖速食餐廳欲對其消費者進行服務滿意度調查，受訪消費者只有兩個選項分別為「滿意」和「不滿意」，隨機抽樣調查結果顯示 650 位消費者中有 210 位表示滿意其服務。試估算(A)滿意該連鎖速食餐廳消費者的比例之點估計值；(B)滿意該連鎖速食餐廳消費者的比例，在 95 %信賴水準下的最大誤差；(C)滿意該連鎖速食餐廳消費者的比例，在 95 %信賴水準下的信賴區間及區間長度。(答案有效位數四捨五入取到小數點後第 4 位)

題解：樣本數量 $n = 650$

(A)樣本滿意比例$\bar{p} = \frac{210}{650} = 0.3231$ = 滿意該連鎖速食餐廳消費者的比例之點估計值

(B)95 %信賴水準下的最大誤差 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025} = 1.96$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。

$$z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}} = 1.96 \times \sqrt{\frac{0.3231\times(1-0.3231)}{650}} = 1.96 \times 0.0183 = 0.0360$$

(C)信賴水準 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025} = 1.96$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。

母體比例信賴區間：$\bar{p} - z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}} \leq p \leq \bar{p} + z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}}$

下限：$\bar{p} - z_{\frac{\alpha}{2}}\times\sqrt{\frac{p\times(1-p)}{n}} = 0.3231 - 1.96\times\sqrt{\frac{0.3231\times(1-0.3231)}{650}} = 0.3231 - 0.0360 = 0.2871$

上限：$\bar{p} + z_{\frac{\alpha}{2}}\times\sqrt{\frac{p\times(1-p)}{n}} = 0.3231 + 1.96\times\sqrt{\frac{0.3231\times(1-0.3231)}{650}} = 0.3231 + 0.0360 = 0.3590$

滿意該連鎖速食餐廳消費者的比例，在 95 %信賴水準下的信賴區間(CI)為 $0.2871 \leq p \leq 0.3590$

滿意該連鎖速食餐廳消費者的比例，在 95 %信賴水準下的區間長度：$2 \times z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}} = 2 \times 1.96 \times \sqrt{\frac{0.3231\times(1-0.3231)}{650}} = 2 \times 1.96 \times 0.0183 = 0.0719$

答案：(A) 點估計值 = 0.3231；(B) 最大誤差 = 0.0360；(C) 信賴區間(CI)為 $0.2871 \leq p \leq 0.3590$；區間長度 = 0.0719

練習 8.36 依據以往的調查顯示，在五年內國內飯店經理級主管跳槽的比率為 25 %，期望估計誤差在 3 %以內，採用 95 %信賴水準。試估算(a)從新估算此比率時，需要隨機抽取多少經理級主管調查？(b)若沒有以前的調查資料，需要隨機抽取多少經理級主管調查？

練習 8.37 深水研究機構欲調查台灣家庭年所得低於 NT$ 400000 的比例。從 650 家庭的隨機樣本中估算，有 250 個家庭的年收入低於 NT$400000。請計算家庭年所得低於 NT$ 400000 比例的 95 %信賴區間。(答案有效位數四捨五入取到小數點後第 4 位)

練習 8.38 隨機抽樣高雄市大樓住戶 120 戶，發現有 100 戶裝設第四台，請估算高雄市大樓住戶第四台裝設率 95 %信賴區間。(答案有效位數四捨五入取到小數點後第 4 位)

題解：樣本之第四台裝設率$\frac{100}{120} = 0.8333$，信賴水準 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025} = 1.96$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。

母體比例信賴區間：$\bar{p} - z_{\frac{\alpha}{2}} \times \sqrt{\frac{\bar{p}\times(1-\bar{p})}{n}} \leq p \leq \bar{p} + z_{\frac{\alpha}{2}} \times \sqrt{\frac{\bar{p}\times(1-\bar{p})}{n}}$

下限$\bar{p} - z_{\frac{\alpha}{2}} \times \sqrt{\frac{\bar{p}\times(1-\bar{p})}{n}} = 0.8333 - 1.96\times\sqrt{\frac{0.8333\times(1-0.8333)}{120}} = 0.8333 - 0.0667 = 0.7667$

上限$\bar{p} + z_{\frac{\alpha}{2}} \times \sqrt{\frac{\bar{p}\times(1-\bar{p})}{n}} = 0.8333 + 1.96\times\sqrt{\frac{0.8333\times(1-0.8333)}{120}} = 0.8333 + 0.0667 = 0.9000$

答案：高雄市大樓住戶第四台裝設率 95 %信賴區間(CI)為 $0.7667 \leq p \leq 0.9000$

8.3.2 母體比例區間估計：小量樣本數【選擇教材】

樣本數量少($n < 30$)時，於母體基本單位總數 N 為無窮大($N = \infty$)，樣本比率$\bar{p}$的抽樣分布會接近於二項分布，即為

$\bar{p}$~二項分布 B(n, p)

樣本數量少($n < 30$)時，於母體基本單位總數 N 為有限數值，樣本比率$\bar{p}$的抽樣分布會接近於超幾何分布，即為

$\bar{p}$~超幾何分布 Hypergeometric(N, r, n)

在樣本數量比較少的時候，無法使用簡單的公式推估(計算)信賴區間，可以使用查圖的方式粗略估計母體比例 p 信賴區間。

8.4 決定樣本數量

樣本數量 n 愈多，獲得有關母體參數(parameter)的資訊愈多，樣本的點估計值(point estimation)與原本母體參數(parameter)的數值愈接近，估計誤差愈小。惟研究調查實務上，樣本數量 n 愈多，抽樣(研究執行)成本愈高與時間耗費愈多。故在有限的經費和時效中，獲取誤差最小(可容忍誤差)的估計值，兩者折衷的平衡點，即必須考量抽樣樣本數量 n。

8.4.1 估計母體平均值時，需要樣本數量

當母體參數(parameter)分布趨近於常態分布。以樣本平均值$\bar{x}$估計母體平均值 μ，欲將估計誤差設定小於 A 值(與觀測值和平均值單位相同)，即估計誤差 = $|\bar{x}-\mu| \le A$。設定在 $100 \times (1-\alpha)$%的信賴水準下，母體變異數 σ^2 和標準(偏)差 σ 已知，母體平均值 μ 落於$\bar{x}-z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x}+z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$區間內，即$|\bar{x}-\mu| \le z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$。若$z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} \le A$ 成立，即$|\bar{x}-\mu| \le z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} \le A$ 亦可成立，隨機抽樣的樣本數量需達 $n \ge \frac{z_{\frac{\alpha}{2}}^2 \times \sigma^2}{A^2}$。

$$|\bar{x}-\mu| \le z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} \le A \;\rightarrow\; z_{\frac{\alpha}{2}} \times \frac{\sigma}{A} \le \sqrt{n} \xrightarrow{\text{等號兩邊皆開平方}} \frac{z_{\frac{\alpha}{2}}^2 \times \sigma^2}{A^2} \le n \;\rightarrow\; n \ge \frac{z_{\frac{\alpha}{2}}^2 \times \sigma^2}{A^2}$$

一般母體變異數 σ^2 未知，可以利用過去的樣本資料計算出樣本變異數 S^2 代替；或抽取少量樣本資料計算樣本變異數 S^2 代替。亦可利用$\frac{R}{6} < \sigma < \frac{R}{4}$ [R 為母體全距或母體全距的估計值]，若以$\frac{R}{4}$取代 σ，所獲得的樣本數量大小為最寬限量。

範例 8.16　為調查深水大學學生每週餐飲消費金額分布情況，希望估計誤差有 0.95 機率不超過新台幣 5 元，應從該大學學生中隨機抽取多少位學生為樣本？依據過去的研究調查資料顯示，大學生每週餐飲消費金額的標準(偏)差 σ 為新台幣 50 元。

題解：信賴水準 $1-\alpha = 0.95$，顯著水準 $\alpha = 0.05$，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025} = 1.96$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)，估計誤差 A = 5 元。

樣本數量 $n \ge \frac{z_{\frac{\alpha}{2}}^2 \times \sigma^2}{A^2} = \frac{1.96^2 \times 50^2}{5^2} = \frac{9604}{25} = 384.2$。非整數時，請採用無條件進位法，取到個位數 385，才能達到「估計誤差有 0.95 機率不超過新台幣 5 元」標準。

故應從該大學學生中抽取 385 位學生當成樣本，方能達成估計誤差 0.95 機率不超過新台幣 5 元的設定條件。

答案：需要隨機抽出 385 位學生當樣本

練習 8.39　一位總經理欲估算特定區域的家庭年收入。假設母體標準(偏)差為 NT$ 10000。此總經理想要樣本平均值與母體平均值的誤差在 NT$ 1200 有 0.95 機率。需要多少的樣本數量？

題解：信賴水準 $1-\alpha = 0.95$，顯著水準 $\alpha = 0.05$，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025} = 1.96$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。

樣本數量 $n \ge \frac{z_{\frac{\alpha}{2}}^2 \times \sigma^2}{A^2} = \frac{1.96^2 \times 10000^2}{1200^2} = \frac{384160000}{1440000} = 266.777$。非整數時，請採用無條件進位法，取到個位數 267，才能達到「誤差在 NT$ 1200 有 0.95 機率」標準。

答案：267 位

8.4.2 估計母體比例時，需要樣本數量

以樣本比例$\bar{p}$估計母體比例 p，欲將估計誤差設定小於 A 值(與觀測值和平均值單位相同，皆屬於無因次單位)，即估計誤差 = $|\bar{p}-p| \le A$。設定在 $100 \times (1-\alpha)$%的信賴水準下，母體的比例落於$\bar{p}-z_{\frac{\alpha}{2}} \times \sqrt{\frac{p \times (1-p)}{n}}$

$\leq p \leq \bar{p} + z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}}$區間內，即$|\bar{p} - p| \leq z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}}$。若$z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}} \leq A$ 成立，即$|\bar{p} - p| \leq z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}} \leq A$ 亦可成立，樣本數量需達 $n \geq \frac{z_{\frac{\alpha}{2}}^2 \times p\times(1-p)}{A^2}$。

$$|\bar{p} - p| \leq z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}} \leq A \rightarrow z_{\frac{\alpha}{2}} \times \frac{\sqrt{p\times(1-p)}}{\sqrt{n}} \leq A \rightarrow z_{\frac{\alpha}{2}} \times \frac{\sqrt{p\times(1-p)}}{A} \leq \sqrt{n} \rightarrow \frac{z_{\frac{\alpha}{2}}^2 \times p\times(1-p)}{A^2} \leq n \rightarrow n \geq \frac{z_{\frac{\alpha}{2}}^2 \times p\times(1-p)}{A^2}$$

一般母體比例 p 未知，可以利用過去的樣本資料計算出樣本比例$\bar{p}$代替；或抽取少量樣本資料計算樣本比例$\bar{p}$代替。亦可 0.5 代替 p，惟此可能造成所得樣本數量高於原始需求數量。

範例 8.17　為調查高雄市消費者對有機飲食的支持比例，希望估計誤差有 95 %機率小於 0.005，應從高雄市消費者中隨機抽取多少位消費者為樣本？若過去未有相關研究調查資料，故以保守估計 0.5 代替母體比例 p。

題解：信賴水準 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025} = 1.96$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)，估計誤差 A = 0.005。

樣本數量 $n \geq \frac{z_{\frac{\alpha}{2}}^2 \times p\times(1-p)}{A^2} = \frac{1.96^2\times0.5\times0.5}{0.005^2} = \frac{0.9604}{0.000025} = 38414.59$。非整數時，請採用無條件進位法，取到個位數 38415，才能達到「誤差有 95 %機率小於 0.005」標準。

故應從高雄市消費者中抽取 38415 位消費者當成樣本，方能達成估計誤差 95 %機率不超過 0.005 的設定條件。

答案：需要隨機抽出 38415 位消費者當樣本

練習 8.40　某民調機構希望預測特定候選人的得票比例。期望在 95 %的信賴水準下達到誤差在 0.05 以內，需要多少的樣本數量？

題解：信賴水準 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，右尾機率$\frac{\alpha}{2}$標準常態值$z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = z_{0.025} = 1.96$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)，誤差 $A = 0.05$。

樣本數量 $n \geq \frac{z_{\frac{\alpha}{2}}^2 \times p\times(1-p)}{A^2} = \frac{1.96^2\times0.5\times0.5}{0.05^2} = \frac{0.9604}{0.0025} = 384.16$。非整數時，請採用無條件進位法，取到個位數 385，才能達到「95 %的信賴水準下達到誤差在 0.05 以內」標準。

答案：385 位

練習 8.41　在某次總統大選前之民調抽樣數量 1000 人，在 95 %信心水準下，最大抽樣誤差 2 %，若抽樣數量增加到 2000 人，請計算最大抽樣誤差？

題解：抽樣數量 1000 人時的抽樣誤差 $A_{1000} = 0.02$

樣本數量 $n \geq \frac{z_{\frac{\alpha}{2}}^2 \times p\times(1-p)}{A_{1000}^2} = \frac{z_{\frac{\alpha}{2}}^2 \times p\times(1-p)}{0.02^2} = 1000 \rightarrow z_{\frac{\alpha}{2}}^2 \times p \times (1-p) = 1000\times0.02^2 = 0.4$

抽樣數量 2000 人時的抽樣誤差 A_{2000}

$$\frac{z_{\frac{\alpha}{2}}^2 \times p\times(1-p)}{A_{2000}^2} = \frac{0.4}{A_{2000}^2} = 2000 \rightarrow A_{2000}^2 = \frac{0.4}{2000} = 0.0002 \rightarrow A_{2000} = \sqrt{A_{2000}^2} = \sqrt{0.0002} = 0.01414$$

答案：最大抽樣誤差為 1.414 %

練習 8.42　在某次選舉前，抽樣調查某位候選人的支持率。在 95 %信心水準下，想要達到不超過 0.05 的估計誤差，至少需要多大的樣本？(A)385；(B)384；(C)271；(D)270。(99 年初等考試統計學大意)

8.5 母體變異數區間估計

餐廳、旅館或旅行社投資報酬率的變異數(variance)，亦代表獲利(經營)過程必須承擔的風險程度。餐廳每天營業額的變異數亦代表營運必須準備的備用食材之數量，然而備用食材準備太多，容易造成不新鮮及耗損；食材準備太少，無法滿足消費者的消費需求。因此，估算母體變異數 σ^2 之信賴區間就相當重要。

估算母體變異數 σ^2 的信賴區間，必須透過樣本變異數 S^2 的抽樣分布，即是屬於卡方分布(chi-square distribution)，利用 χ^2 分布標示。

8.5.1 卡方分布

卡方分布(Chi-square distribution)是將所有樣本之觀測值利用標準常態分布之標準化 z 值的平方和獲得。從母體中，有 N 個基本單位，分別為 $x_1, x_2, x_3, \ldots, x_N$，其母體平均值 $\mu = \frac{\sum_{i=1}^{N} x_i}{N}$，母體變異數 $\sigma^2 = \frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}$，標準化 z 值為

$$z_i = \frac{x_i-\mu}{\sigma} \rightarrow \text{等號兩邊皆取平方後為} z_i^2 = \frac{(x_i-\mu)^2}{\sigma^2}$$

若母體中，隨機抽出 n 個樣本，分別為 $x_1, x_2, x_3, \ldots, x_n$，其標準化 z 值的平方和為自由度 $v = n$ 的卡方值。此卡方分布，具有自由度 $v = n$，可以利用 $\chi^2 \sim \chi^2(n)$ 或 χ_n^2 符號表示。卡方值屬於無因次單位。

$$\sum_{i=1}^{n} z_i^2 = \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma^2} = \chi_n^2$$

若從 N 個基本單位的母體中，**隨機**抽出 n 個基本單位為樣本，可能的樣本組合為 N^n 種，即可獲得 N^n 個卡方值，統計各種卡方值的分布次數，即可獲得卡方分布曲線(x 軸為卡方值，y 軸為分布次數或頻率)。與 t 分布類同，當樣本數量 n 不同時，可獲得不同的卡方分布曲線。

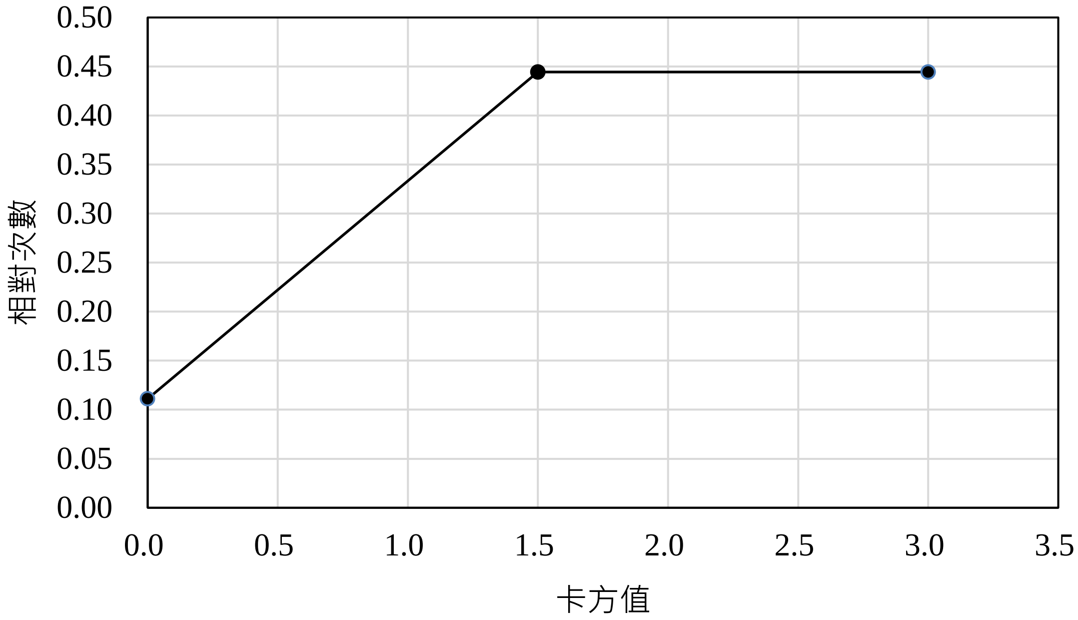
範例 8.18 從 3 個基本單位分別為 1、2 和 3 的母體中，隨機抽出 2 個基本單位為樣本，會產生 $N^n = 3^2 = 9$ 種不同的樣本組合，亦可獲得 9 個卡方值 χ^2

樣本	平均值 $\bar{x}$	$\sum_{i=1}^{n}(x_i-\mu)^2$	$\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma^2}$	$\sum_{i=1}^{n}(x_i-\bar{x})^2$	$\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sigma^2}$
1, 1	1.0	2	3.0	0.0	0.00
1, 2	1.5	1	1.5	0.5	0.75
1, 3	2.0	2	3.0	2.0	3.00
2, 1	1.5	1	1.5	0.5	0.75
2, 2	2.0	0	0.0	0.0	0.00
2, 3	2.5	1	1.5	0.5	0.75
3, 1	2.0	2	3.0	2.0	3.00
3, 2	2.5	1	1.5	0.5	0.75
3, 3	3.0	2	3.0	0.0	0.00

母體基本單位總數 $N = 3$，樣本數量 $n = 2$，母體平均值 $\mu = \frac{1+2+3}{3} = 2$。

母體變異數 $\sigma^2 = \frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N} = \frac{2}{3} = 0.6667$，卡方值 $\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma^2} = \chi_2^2$。

χ_2^2 卡方值	次數	相對次數
0.0	1	0.1111
1.5	4	0.4444
3.0	4	0.4444



當母體基本單位數量 N 很多，在不同的樣本數量 n 下，其卡方的分布情況如下圖所示，自由度大於 1 時，皆會產生波峰，波峰的最高點(中心點)即在卡方值小於其自由度的鄰近位置。

卡方分布機率密度函數 $f(x) = \frac{\frac{1}{2}^{\frac{v}{2}}}{\Gamma(\frac{v}{2})} \times x^{\frac{v}{2}-1} \times e^{\frac{-x}{2}}$，其中 $x \geq 0$，若 $x < 0$ 時，$f(x) = 0$。Γ 為 Gamma 函數 $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx$，$\alpha > 0$。

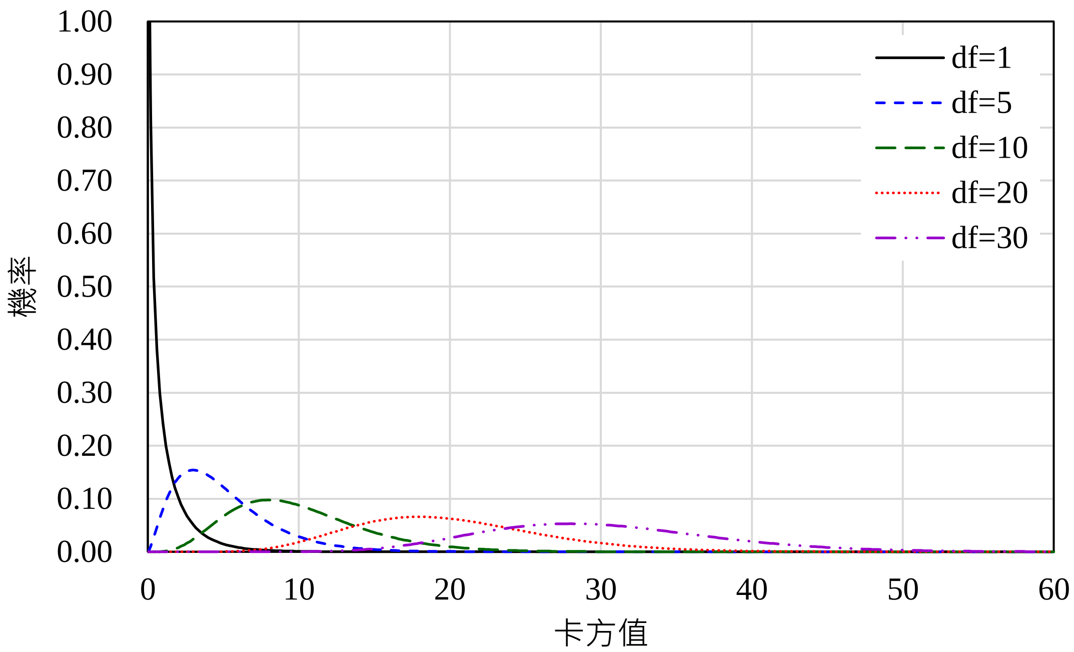


圖　卡方分布曲線

故卡方值χ_n^2的期望值為

$$E(\chi_n^2) = E\left[\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma^2}\right] = \frac{E\left[\sum_{i=1}^{n}(x_i-\mu)^2\right]}{\sigma^2} = \frac{n\times E\left[\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}\right]}{\sigma^2} = \frac{n\times\sigma^2}{\sigma^2} = n$$

$$Var(\chi_n^2) = V(\chi_n^2) = V\left[\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma^2}\right] = 2n$$

從母體所有基本單位中，隨機抽出 n 個基本單位為樣本，分別為 $x_1, x_2, x_3, \ldots, x_n$，樣本變異數 S^2 為：

$$S^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$$

若母體屬於(趨近於)常態分佈，其母體平均值 μ，變異數 σ^2，$x\sim N(\mu,\sigma^2)$，將樣本變異數 S^2 乘以$\frac{n-1}{\sigma^2}$，樣本變異數成為自由度 $v = n - 1$ 的卡方統計量或卡方分布，標示為χ_{n-1}^2或χ_v^2。

$$\chi_{n-1}^2 = \chi_{v=n-1}^2 = S^2 \times \frac{n-1}{\sigma^2} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1} \times \frac{n-1}{\sigma^2} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sigma^2}$$

若母體某特定變數之觀測值的平均值 μ 無法獲得時，利用樣本平均值$\bar{x}$取代，卡方值以χ_{n-1}^2標記，其自由度 $df = v = n - 1$。

$$\frac{(n-1)\times S^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sigma^2} = \chi_{n-1}^2 = \chi_{v=n-1}^2$$

$$E(\chi_{n-1}^2) = E\left[\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sigma^2}\right] = \frac{E[(n-1)\times S^2]}{\sigma^2} = \frac{(n-1)\times E(S^2)}{\sigma^2} = \frac{(n-1)\times \sigma^2}{\sigma^2} = n-1$$

$$Var(\chi_{n-1}^2) = V(\chi_{n-1}^2) = V\left[\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sigma^2}\right] = 2 \times (n-1)$$

卡方分布 Excel 函數

利用 Excel 軟體中插入(I)→函數(F)…→在插入函數對話方塊中選取類別(C): **統計**，選取函數(N): **CHIDIST**→確定。在函數引數對話視窗中，x 方塊輸入：欲推估卡方分布的數值間 x；Deg_freedom 方塊輸入：自由度 v。確定。即會在原先選定的儲存格中出現卡方分布機率數值。CHIDIST(x,deg_freedom)。

利用 Excel 軟體中插入(I)→函數(F)…→在插入函數對話方塊中選取類別(C): **統計**，選取函數(N): **CHIINV**→確定。在函數引數對話視窗中，Probability 方塊輸入：單尾分布機率數值 α；Deg_freedom 方塊輸入：自由度 v。確定。即會在原先選定的儲存格中出現卡方分布的卡方數值。CHIINV(probability,deg_freedom)。

8.5.2 卡方分布性質

在卡方分布中，所有卡方值皆為正值($\chi^2 > 0$)。

卡方分布屬於右偏分布。不同自由度有不同的卡方分布曲線。

當樣本數量 n 增加，致自由度 df 增加時，卡方分布曲線會趨近於左右對稱，當樣本數量 n 趨近於無窮大(∞)時，卡方分布會接近常態分布。$\chi_v^2 \sim N(v,2v)$。

8.5.3 利用卡方分布推估母體變異數的信賴區間

在卡方分布中，母體特定變數之觀測值的平均值 μ 無法獲得，僅以抽出 n 個基本單位的樣本平均值$\bar{x}$取代，在自由度 $df = v = n - 1$ 的環境，利用卡方值χ_{n-1}^2推估母體變異數的信賴區間：

$$P(\chi_{1-\frac{\alpha}{2},n-1}^2 \le \chi^2 \le \chi_{\frac{\alpha}{2},n-1}^2) = 1-\alpha$$

$$P(\chi_{1-\frac{\alpha}{2},n-1}^2 \le \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sigma^2} \le \chi_{\frac{\alpha}{2},n-1}^2) = 1-\alpha$$

其中$\chi_{1-\frac{\alpha}{2},n-1}^2$：代表在自由度 $df = v = n - 1$，右尾機率為 $1 - \frac{\alpha}{2}$所對應的卡方值。

$\chi_{\frac{\alpha}{2},n-1}^2$：代表在自由度 $df = v = n - 1$，右尾機率為$\frac{\alpha}{2}$所對應的卡方值。

$$P(\chi_{1-\frac{\alpha}{2},n-1}^2 \le \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sigma^2} \le \chi_{\frac{\alpha}{2},n-1}^2) = 1-\alpha$$

$$P(\frac{\chi_{1-\frac{\alpha}{2},n-1}^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \le \frac{1}{\sigma^2} \le \frac{\chi_{\frac{\alpha}{2},n-1}^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}) = P(\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\chi_{1-\frac{\alpha}{2},n-1}^2} \ge \sigma^2 \ge \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\chi_{\frac{\alpha}{2},n-1}^2}) = 1-\alpha$$

$$P(\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\chi_{\frac{\alpha}{2},n-1}^2} \le \sigma^2 \le \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\chi_{1-\frac{\alpha}{2},n-1}^2}) = 1-\alpha$$

$S^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}$等號兩邊皆乘以 $n - 1$ 為 $S^2 \times (n-1) = \sum_{i=1}^{n}(x_i-\bar{x})^2$

$$P(\frac{S^2\times(n-1)}{\chi_{\frac{\alpha}{2},n-1}^2} \le \sigma^2 \le \frac{S^2\times(n-1)}{\chi_{1-\frac{\alpha}{2},n-1}^2}) = 1-\alpha$$

母體變異數 σ^2 在信賴水準 $1-\alpha$ 的信賴區間為：

$$\frac{S^2\times(n-1)}{\chi^2_{\frac{\alpha}{2},n-1}} \le \sigma^2 \le \frac{S^2\times(n-1)}{\chi^2_{1-\frac{\alpha}{2},n-1}}$$

$$\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\chi^2_{\frac{\alpha}{2},n-1}} \le \sigma^2 \le \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}$$

母體標準(偏)差 σ 在信賴水準 $1-\alpha$ 的信賴區間為：

$$\sqrt{\frac{S^2\times(n-1)}{\chi^2_{\frac{\alpha}{2},n-1}}} \le \sigma \le \sqrt{\frac{S^2\times(n-1)}{\chi^2_{1-\frac{\alpha}{2},n-1}}}$$

$$\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\chi^2_{\frac{\alpha}{2},n-1}}} \le \sigma \le \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}}$$

範例 8.19　假設有一自由度 19 的卡方值。請找出下列右尾機率的卡方值 χ^2：(a) 0.005；(b)0.010；(c)0.050；(d)0.100。(答案有效數值取到小數點後四位)

題解：可以查詢卡方分布臨界值表、運用 Excel 軟體 CHISQ.INV.RT(右尾機率)或 CHIINV 函數第一個參數輸入右尾機率，第二個參數輸入自由度即可獲得或 CHISQ.INV 函數第一個參數輸入左尾機率，第二個參數輸入自由度亦可獲得。

答案：(a) $\chi^2_{0.005,19} = 38.5823$; (b) $\chi^2_{0.010,19} = 36.1909$; (c) $\chi^2_{0.050,19} = 30.1435$; (d) $\chi^2_{0.100,19} = 27.2036$

範例 8.20　假設有一自由度 19 的卡方值。請找出下列左尾機率的卡方值 χ^2：(a) 0.005；(b)0.010；(c)0.050；(d)0.100。(答案有效數值取到小數點後四位)

題解：可以查詢卡方分布臨界值表，或運用 Excel 軟體 CHISQ.INV.RT(右尾機率)函數第一個參數輸入右尾機率(此題目提供的是左尾機率，可以利用 1 – 左尾機率就可以獲得右尾機率，輸入前述函數)，第二個參數輸入自由度即可獲得。

答案：(a) $\chi^2_{1-0.005,19} = 6.8440$; (b) $\chi^2_{1-0.010,19} = 7.6327$; (c) $\chi^2_{1-0.050,19} = 10.1170$; (d) $\chi^2_{1-0.100,19} = 11.6509$

練習 8.43　假設χ^2_0為一特定的卡方值 χ^2。請找出下列條件下的卡方值χ^2_0：(a)$P(\chi^2 \ge \chi^2_0) = 0.01$，樣本數量 $n = 15$；(b)$P(\chi^2 \ge \chi^2_0) = 0.05$，樣本數量 $n = 20$；(c)$P(\chi^2 \ge \chi^2_0) = 0.10$，樣本數量 $n = 25$。(答案有效數值取到小數點後四位)

題解：可以查詢卡方分布臨界值表，或運用 Excel 軟體 CHISQ.INV.RT(右尾機率)函數第一個參數輸入右尾機率｛例如：$P(\chi^2 \ge \chi^2_0) = 0.01$，右尾機率就是 0.01｝，第二個參數輸入自由度(n - 1)即可獲得。

答案：(a) $\chi^2_0 = 29.1412$; (b) $\chi^2_0 = 30.1435$; (c) $\chi^2_0 = 33.1962$

範例 8.21　奇遇海產店每日營業額屬於常態分布，上星期 7 天營業日每日營業額之標準(偏)差 $S = 1200$ 元，試求每日營業額變異數 σ^2 在 95 %信賴水準下的信賴區間？(答案有效位數四捨五入取到小數點後第 1 位)

題解：樣本數量 $n = 7$，自由度 $df = v = n - 1 = 7 - 1 = 6$，信賴水準 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，樣本標準(偏)差 $S = 1200$ 元，$\chi^2_{1-\frac{\alpha}{2},n-1} = \chi^2_{1-\frac{0.05}{2},7-1} = \chi^2_{0.975,6} = 1.2373$，$\chi^2_{\frac{\alpha}{2},n-1} = \chi^2_{\frac{0.05}{2},7-1} = \chi^2_{0.025,6} = 14.4494$[運用 Excel 軟體 CHISQ.INV.RT(右尾機率)函數獲得]。

母體變異數信賴區間公式為$\frac{S^2\times(n-1)}{\chi^2_{\frac{\alpha}{2},n-1}} \le \sigma^2 \le \frac{S^2\times(n-1)}{\chi^2_{1-\frac{\alpha}{2},n-1}}$

母體變異數信賴區間下限：$\frac{S^2\times(n-1)}{\chi^2_{\frac{\alpha}{2},n-1}} = \frac{1200^2\times(7-1)}{14.4494} = \frac{8640000}{14.4494} = 597948.7$ 元 ^{2}

母體變異數信賴區間上限：$\frac{S^2\times(n-1)}{\chi^2_{1-\frac{\alpha}{2},n-1}}=\frac{1200^2\times(7-1)}{1.2373}=\frac{8640000}{1.2373}=6982946.7$ 元 ^{2}

每日營業額變異數 σ^2 之 95 %信賴區間(CI)為 $597948.7\le\sigma^2\le6982946.7$ 元 ^{2}。每日營業額標準(偏)差 σ 之 95 %信賴區間(CI)為 $773.3\le\sigma\le2642.5$ 元

練習 8.44　奇遇海產店每日營業額屬於常態分布，上星期 7 天營業日每日營業額如下表所示，試求每日營業額變異數 σ^2 在 95 %信賴水準下的信賴區間？(答案有效位數四捨五入取到小數點後第 1 位)

營業日	營業額 x_i	$x_i-\bar{x}$	$(x_i-\bar{x})^2$
1	16000	0	0
2	18000	2000	4000000
3	17500	1500	2250000
4	16500	500	250000
5	12000	-4000	16000000
6	15000	-1000	1000000
7	17000	1000	1000000
合計	112000		24500000

題解：樣本數量 $n=7$，自由度 $df=v=n-1=7-1=6$，信賴水準 $1-\alpha=0.95$，顯著水準 $\alpha=0.05$，樣本標準(偏)差 $S=\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}=\sqrt{\frac{24500000}{7-1}}=2020.726$，$\chi^2_{1-\frac{\alpha}{2},n-1}=\chi^2_{1-\frac{0.05}{2},7-1}=\chi^2_{0.975,6}=1.2373$，$\chi^2_{\frac{\alpha}{2},n-1}=\chi^2_{\frac{0.05}{2},7-1}=\chi^2_{0.025,7}=14.4494$[運用 Excel 軟體 CHISQ.INV.RT(右尾機率)函數獲得]。

母體變異數信賴區間公式為$\frac{S^2\times(n-1)}{\chi^2_{\frac{\alpha}{2},n-1}}\le\sigma^2\le\frac{S^2\times(n-1)}{\chi^2_{1-\frac{\alpha}{2},n-1}}$

下限：$\frac{S^2\times(n-1)}{\chi^2_{\frac{\alpha}{2},n-1}}=\frac{2020.7^2\times(7-1)}{14.4494}=\frac{24500000}{14.4494}=1695572.1$ 元 ^{2}

上限：$\frac{S^2\times(n-1)}{\chi^2_{1-\frac{\alpha}{2},n-1}}=\frac{2020.7^2\times(7-1)}{1.2373}=\frac{24500000}{1.2373}=19801180.0$ 元 ^{2}

故每日營業額變異數 σ^2 之 95 %信賴區間(CI)為 $1695572.1\le\sigma^2\le19801180.0$ 元 ^{2}。每日營業額標準(偏)差 σ 之 95 %信賴區間(CI)為 $1302.1\le\sigma\le4449.9$ 元

練習 8.45　阿文連鎖飲料店聲稱其珍珠奶茶每杯容量體積皆為 650 ml。現從其產品中隨機抽取 7 件樣本，測量其體積分別為 620、655、670、635、665、648 和 641 ml。若珍珠奶茶每杯容積呈現常態分布 $N(\mu,\sigma^2)$。請估算(A) σ^2 點估計值。(B)95 %信賴水準下，σ^2 的信賴區間。(答案有效位數四捨五入取到小數點後第 1 位)

題解：樣本數量 $n=7$，自由度 $df=v=n-1=7-1=6$，信賴水準 $1-\alpha=0.95$，顯著水準 $\alpha=0.05$，樣本標準(偏)差 $S=17.4329$，樣本變異數 $S^2=303.9048$，$\chi^2_{1-\frac{\alpha}{2},n-1}=\chi^2_{1-\frac{0.05}{2},7-1}=\chi^2_{0.975,6}=1.2373$，$\chi^2_{\frac{\alpha}{2},n-1}=\chi^2_{\frac{0.05}{2},7-1}=\chi^2_{0.025,6}=14.4494$[運用 Excel 軟體 CHISQ.INV.RT(右尾機率)函數獲得]。

母體變異數信賴區間公式為$\frac{S^2\times(n-1)}{\chi^2_{\frac{\alpha}{2},n-1}}\le\sigma^2\le\frac{S^2\times(n-1)}{\chi^2_{1-\frac{\alpha}{2},n-1}}$

下限：$\frac{S^2\times(n-1)}{\chi^2_{\frac{\alpha}{2},n-1}}=\frac{7.4329^2\times(7-1)}{14.4494}=\frac{1823.429}{14.4494}=126.19$ ml^2

上限：$\frac{S^2\times(n-1)}{\chi^2_{1-\frac{\alpha}{2},n-1}}=\frac{7.4329^2\times(7-1)}{1.2373}=\frac{1823.429}{1.2373}=1473.72$ ml^2

故每日營業額變異數 σ^2 之 95 %信賴區間(CI)為 $126.2\le\sigma^2\le1473.7$ ml^2

答案：(A) σ^2 點估計值 = 303.9048 ml^2；(B)信賴區間 $126.2\le\sigma^2\le1473.7$ ml^2

8.6 單尾區間估計

前述的區間估計皆是假設母體參數(parameter)皆有可能比樣本統計值(statistic)高或低，故在信賴水準為 $1-\alpha$，可能發生錯誤機率 α，平均分布於左右兩側。若欲將錯誤機率 α，全部放置在左側或右側，屬於單尾區間估計或單側區間估計。

8.6.1 母體平均值單尾區間估計

在雙尾區間估計時，母體變異數 σ^2 和標準(偏)差 σ 已知，抽樣樣本數量較多($n \geq 30$)，推估母體平均值 μ 在信賴水準 $1-\alpha$ 下的信賴區間為：

$$P(\bar{x} - z_{\frac{\alpha}{2}} \times \sigma_{\bar{x}} \leq \mu \leq \bar{x} + z_{\frac{\alpha}{2}} \times \sigma_{\bar{x}}) = 1-\alpha$$

其中 $z_{\frac{\alpha}{2}}$：為右尾機率為 $\frac{\alpha}{2}$ 的標準常態化值。

$\frac{\alpha}{2}$：為雙尾機率，雙尾機率和 $\frac{\alpha}{2} + \frac{\alpha}{2} = \alpha = 0.05$。

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$：為樣本平均值分布的標準(偏)差。

$\bar{x} - z_{\frac{\alpha}{2}} \times \sigma_{\bar{x}}$：為信賴區間下限 a；$\bar{x} + z_{\frac{\alpha}{2}} \times \sigma_{\bar{x}}$：為信賴區間上限 b。

在右單尾區間估計時，母體變異數 σ^2 和標準(偏)差 σ 已知，所有可能發生錯誤機率 α，皆在右側，推估母體平均值 μ 在信賴水準 $1-\alpha$ 下的信賴區間為

$$P(\mu \leq \bar{x} + z_\alpha \times \sigma_{\bar{x}}) = P(\mu \leq \bar{x} + z_\alpha \times \frac{\sigma}{\sqrt{n}}) = 1-\alpha$$

若母體變異數 σ^2 和標準(偏)差 σ 未知，利用樣本標準(偏)差 $S = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$ 取代母體標準(偏)差 σ，推估母體平均值 μ 在信賴水準 $1-\alpha$ 下的信賴區間為

$$P(\mu \leq \bar{x} + z_\alpha \times \frac{S}{\sqrt{n}}) = 1-\alpha$$

在左單尾區間估計時，母體變異數 σ^2 和標準(偏)差 σ 已知，所有可能發生錯誤機率 α，皆在左側，推估母體平均值 μ 在信賴水準 $1-\alpha$ 下的信賴區間為

$$P(\bar{x} - z_\alpha \times \sigma_{\bar{x}} \leq \mu) = P(\bar{x} - z_\alpha \times \frac{\sigma}{\sqrt{n}} \leq \mu) = 1-\alpha$$

若母體變異數 σ^2 和標準(偏)差 σ 未知，利用樣本標準(偏)差 $S = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$ 取代母體標準(偏)差 σ，推估母體平均值 μ 在信賴水準 $1-\alpha$ 下的信賴區間為

$$P(\bar{x} - z_\alpha \times \frac{S}{\sqrt{n}} \leq \mu) = 1-\alpha$$

在雙尾區間估計時，樣本數量少($n < 30$)，母體屬於常態分布，母體變異數 σ^2 和標準(偏)差 σ 未知時，母體平均值 μ 的信賴區間：

$$\bar{x} \pm t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}}$$

其中 $1-\alpha$：信賴係數　　　α：信賴區間的顯著水準。

$t_{\frac{\alpha}{2},v}$：在自由度 $v = n-1$，右尾機率(面積)為 $\frac{\alpha}{2}$ 的 t 值。

S：樣本標準(偏)差 $S = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$。

$\bar{x}$：樣本平均值。

v：自由度(degree of freedom, *df*)是在計算 $\sum_{i=1}^{n}(x_i-\bar{x})^2$ 時，樣本所需提供的獨立資料的個數。

在計算 $\sum_{i=1}^{n}(x_i-\bar{x})^2$ 時需要 n 個觀測值(資料)個數，包括 $x_1-\bar{x}, x_2-\bar{x}, x_3-\bar{x}, \ldots, x_n-\bar{x}$。另外可知 $\sum_{i=1}^{n}(x_i-\bar{x}) = 0$，故只有 $n-1$ 個資料是屬於獨立性質。因此，只要預先知道 $n-1$ 個 $x_i-\bar{x}$ 的數值，最後一個 $x_i-\bar{x}$ 數值可以透過 $\sum_{i=1}^{n}(x_i-\bar{x}) = 0$ 推估獲得。於是：樣本變異數 S^2 的自由度為 $v = n-1$。

$t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}}$：為抽樣誤差(sampling error)、誤差界限、誤差範圍(margin of error)、最大誤差或可能機誤。臨界值 × 標準誤(差)。

$2 \times t_{\frac{\alpha}{2},v} \times \frac{S}{\sqrt{n}}$：為信賴區間的寬度(width)、長度。

在右單尾區間估計時，母體變異數 σ^2 和標準(偏)差 σ 未知，利用樣本標準(偏)差 $S = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$ 取代母體標準(偏)差 σ，推估母體平均值 μ 在信賴水準 $1-\alpha$ 下的信賴區間為

$$P(\mu \le \bar{x} + t_{\alpha,v} \times \frac{S}{\sqrt{n}}) = 1-\alpha$$

在左單尾區間估計時，母體變異數 σ^2 和標準(偏)差 σ 未知，利用樣本標準(偏)差 $S = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}$ 取代母體標準(偏)差 σ，推估母體平均值 μ 在信賴水準 $1-\alpha$ 下的信賴區間為

$$P(\bar{x} - t_{\alpha,v} \times \frac{S}{\sqrt{n}} \le \mu) = 1-\alpha$$

表　單尾母體平均值 μ 之信賴區間

母體屬性	條件	單尾母體平均值信賴區間	
		母體變異數 σ^2 已知	母體變異數 σ^2 未知
常態分布或不確定 樣本數 $n \ge 30$	右單尾	$P(\mu \le \bar{x} + z_\alpha \times \frac{\sigma}{\sqrt{n}}) = 1-\alpha$	$P(\mu \le \bar{x} + z_\alpha \times \frac{S}{\sqrt{n}}) = 1-\alpha$(近似) $P(\mu \le \bar{x} + t_{\alpha,v} \times \frac{S}{\sqrt{n}}) = 1-\alpha$(精準)
	左單尾	$P(\bar{x} - z_\alpha \times \frac{\sigma}{\sqrt{n}} \le \mu) = 1-\alpha$	$P(\bar{x} - z_\alpha \times \frac{S}{\sqrt{n}} \le \mu) = 1-\alpha$(近似) $P(\bar{x} - t_{\alpha,v} \times \frac{S}{\sqrt{n}} \le \mu) = 1-\alpha$(精準)
常態分布 樣本數 $n < 30$	右單尾	$P(\mu \le \bar{x} + z_\alpha \times \frac{\sigma}{\sqrt{n}}) = 1-\alpha$	$P(\mu \le \bar{x} + t_{\alpha,v} \times \frac{S}{\sqrt{n}}) = 1-\alpha$
	左單尾	$P(\bar{x} - z_\alpha \times \frac{\sigma}{\sqrt{n}} \le \mu) = 1-\alpha$	$P(\bar{x} - t_{\alpha,v} \times \frac{S}{\sqrt{n}} \le \mu) = 1-\alpha$

範例 8.22　奇遇海產店販售澳洲生蠔數量屬於(趨近於)常態分布，上個月 31 天營業日每日販售澳洲生蠔平均值 $\bar{x}$ = 150 個，標準(偏)差 S = 20 個，試求在 95 %的信賴水準下，每日至少要準備多少澳洲生蠔？

題解：樣本數量 $n = 31$，信賴水準 $1-\alpha = 0.95$。樣本平均值 $\bar{x} = 150$ 個，樣本標準(偏)差 $S = 20$ 個。

使用標準化 Z 分布近似算法：顯著水準 $\alpha = 0.05$，$z_\alpha = z_{0.05} = 1.645$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。每日至少要準備的澳洲生蠔數量，將所有可能犯錯機率全放在右側，屬於右單尾區間估計，則 $P(\mu \le \bar{x} + z_\alpha \times \frac{S}{\sqrt{n}}) = 1-\alpha = 0.95$。

上限：$\bar{x} + z_\alpha \times \frac{S}{\sqrt{n}} = 150 + 1.645 \times \frac{20}{\sqrt{31}} = 150 + 5.9 = 155.9$ 個(需要無條件進位到個位數)。故每日至少準備生蠔數量為 156 個，才能應付 95 %的情況。因此，每日販售澳洲生蠔數量在 0~156 個，佔有 95 %機率。

使用 t 分布精準算法：自由度 $v = n-1 = 31-1 = 30$，顯著水準 $\alpha = 0.05$，$t_{\alpha,v} = t_{0.05,30} = 1.6973$(使用 Excel 軟體 T.INV 函數查詢獲得)。每日至少要準備的澳洲生蠔數量，將所有可能犯錯機率全放在右側，屬於右單尾區間估計，則 $P(\mu \le \bar{x} + t_{\alpha,v} \times \frac{S}{\sqrt{n}}) = 1-\alpha = 0.95$。

上限：$\bar{x} + t_{\alpha,v} \times \frac{S}{\sqrt{n}} = 150 + 1.6973 \times \frac{20}{\sqrt{31}} = 150 + 6.1 = 156.1$ 個(需要無條件進位到個位數)。故每日至少準備生蠔數量為 157 個，才能應付 95 %的情況。因此，每日販售澳洲生蠔數量在 0~157 個，佔有 95 %機率。

答案：至少要準備 156(近似)和 157(精準)個生蠔才能達到 95 %的信賴水準

範例 8.23 奇遇海產店販售澳洲生蠔數量屬於(趨近於)常態分布，上星期 7 天營業日每日販售澳洲生蠔平均值$\bar{x}$ = 150 個，標準(偏)差 S = 20 個，試求在 95 %的信賴水準下，每日至少要準備多少澳洲生蠔？

題解：樣本數量 $n = 7$，在樣本數量少和未知母體標準(偏)差的情況下，比照雙尾信賴區間的運算模式，運用 $t_{\alpha,v}$ 取代 z_α 運算信賴區間，信賴水準 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，$t_{\alpha,v} = t_{\alpha,n\text{-}1} = t_{0.05,7\text{-}1} = 1.9432$(使用 Excel 軟體 T.INV 函數查詢獲得)，樣本平均值$\bar{x} = 150$ 個，樣本標準(偏)差 $S = 20$ 個。

每日至少要準備的澳洲生蠔數量，將所有可能犯錯機率全放在右側，屬於右單尾區間估計，則

$$P(\mu \le \bar{x} + t_{\alpha,v} \times \frac{S}{\sqrt{n}}) = 1 - \alpha = 0.95$$

上限：$\bar{x} + t_{\alpha,v} \times \frac{S}{\sqrt{n}} = 150 + 1.9432 \times \frac{20}{\sqrt{7}} = 150 + 14.69 = 164.69$ 個

故每日至少準備生蠔數量為 165 個，才能應付 95 %的情況。因此，每日販售澳洲生蠔數量在 0~165 個，佔有 95 %機率。

練習 8.46 奇遇海產店販售澳洲生蠔數量屬於(趨近於)常態分布，上星期 7 天營業日每日販售澳洲生蠔數量如下表所示，試求在 95 %的信賴水準下，每日至少要準備多少澳洲生蠔？

營業日	販售生蠔數量 x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
1	25	-3	9
2	21	-7	49
3	27	-1	1
4	26	-2	4
5	32	4	16
6	28	0	0
7	37	9	81
合計	196		160

題解：樣本數 $n = 7$，自由度 $df = v = n - 1 = 7 - 1 = 6$，信賴水準 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，$t_{\alpha,v} = t_{\alpha,n\text{-}1} = t_{0.05,7\text{-}1} = 1.9432$(使用 Excel 軟體 T.INV 函數查詢獲得)，樣本平均值$\bar{x} = \frac{196}{7} = 28$ 個，樣本標準(偏)差 $S = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}} = \sqrt{\frac{160}{7-1}} = 5.1640$ 個。

每日至少要準備的澳洲生蠔數量，將所有可能犯錯機率全放在右側，屬於右單尾區間估計，則

$$P(\mu \le \bar{x} + t_{\alpha,v} \times \frac{S}{\sqrt{n}}) = 1 - \alpha = 0.95$$

上限：$\bar{x} + t_{\alpha,v} \times \frac{S}{\sqrt{n}} = 28 + 1.9432 \times \frac{5.1640}{\sqrt{7}} = 28 + 3.79 = 31.79$ 個

故每日至少準備生蠔數量為 32 個，才能應付 95 %的情況。因此，每日販售澳洲生蠔數量在 0~32 個，佔有 95 %機率。

練習 8.47 假設有一組樣本數量 $n = 15$ 觀測值的隨機樣本。母體標準(偏)差未知。假設樣本平均值 $\bar{x} = 15.0$ 和樣本標準(偏)差 $S = 5.0$。請計算母體平均值 μ 在 95 %左尾的信賴區間。(答案有效位數四捨五入取到小數點後第 2 位)

題解：信賴係數為 $1 - \alpha = 0.95$，顯著水準 $\alpha = 0.05$，$t_{\alpha,n\text{-}1} = t_{0.05,15\text{-}1} = 1.7613$(使用 Excel 軟體 T.INV 函數查詢獲得)。

$$\bar{x} - t_{\alpha,n\text{-}1} \times \frac{S}{\sqrt{n}} \le \mu \quad \rightarrow \quad 15.0 - 1.7613 \times \frac{5.0}{\sqrt{15}} \le \mu \quad \rightarrow \quad 12.73 \le \mu$$

答案：$12.73 \le \mu$

8.6.2 母體比例單尾區間估計

雙尾區間估計時，樣本數量多($n \geq 30$)，$n \times p \geq 5$ 同時必須 $n \times (1-p) \geq 5$，在信賴係數為 $1-\alpha$，信賴區間的顯著水準為 α，母體比例 p 的信賴區間為：

$$P(\bar{p} - z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}} \leq p \leq \bar{p} + z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}}) = 1-\alpha$$

$$\bar{p} \pm z_{\frac{\alpha}{2}} \times \sigma_{\bar{p}} = \bar{p} \pm z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}} = \bar{p} \pm z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times q}{n}}$$

其中 $1-\alpha$：信賴係數　　　　　　　α：信賴區間的顯著水準

$z_{\frac{\alpha}{2}}$：標準常態分布右尾機率(面積)為$\frac{\alpha}{2}$的標準化 z 值

在右單尾區間估計時，母體比例 p 已知，所有可能發生錯誤機率 α，皆在右側，推估(估算)母體比例 p 在信賴水準 $1-\alpha$ 下的信賴區間為：

$$P(p \leq \bar{p} + z_{\alpha} \times \sqrt{\frac{p\times(1-p)}{n}}) = 1-\alpha$$

若母體比例 p 值未知時，可以利用樣本比例$\bar{p}$取代母體比例 p，以推估(估算)母體比例的信賴區間：

$$P(p \leq \bar{p} + z_{\alpha} \times \sqrt{\frac{\bar{p}\times(1-\bar{p})}{n}}) = 1-\alpha$$

若母體比例 p 值未知時，可以將母體比例 p 設為 0.5，以推估(估算)母體比例的信賴區間，惟若設定 $p = 0.5$ 所獲得的信賴區間較大，也較保守。

在左單尾區間估計時，母體比例 p 已知，所有可能發生錯誤機率 α，皆在左側，推估(估算)母體比例 p 在信賴水準 $1-\alpha$ 下的信賴區間為：

$$P(\bar{p} - z_{\alpha} \times \sqrt{\frac{p\times(1-p)}{n}} \leq p) = 1-\alpha$$

若母體比例 p 值未知時，可以利用樣本比例$\bar{p}$取代母體比例 p，以推估(估算)母體比例的信賴區間：

$$P(\bar{p} - z_{\alpha} \times \sqrt{\frac{\bar{p}\times(1-\bar{p})}{n}} \leq p) = 1-\alpha$$

若母體比例 p 值未知時，可以將母體比例 p 設為 0.5，以推估母體比例的信賴區間，惟若設定 $p = 0.5$ 所獲得的信賴區間較大，也較保守。

範例 8.24　為調查高雄市消費者對有機飲食的支持比例，從高雄市消費者中隨機抽取 300 位消費者為樣本，統計其對有機飲食的支持比例$\bar{p} = 0.65$。為保守估計有機飲食的需求量，欲估計在信賴水準 95 %下支持比率信賴區間的最低值？(答案有效位數四捨五入取到小數點後第 4 位)

題解：樣本比例$\bar{p} = 0.65$，樣本數量 $n = 300$，信賴水準 $1-\alpha = 0.95$，顯著水準 $\alpha = 0.05$，$z_{\alpha} = z_{0.05} = 1.645$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。

題目需欲估計在信賴水準 95 %下支持比率信賴區間的最低值，因此，判斷屬於左尾區間估計，母體比例 p 值未知時，利用樣本比例$\bar{p}$取代母體比例 p，故下限值為$\bar{p} - z_{\alpha} \times \sqrt{\frac{\bar{p}\times(1-\bar{p})}{n}}$

$$\bar{p} - z_{\alpha}\times\sqrt{\frac{\bar{p}\times(1-\bar{p})}{n}} = 0.65 - 1.645 \times \sqrt{\frac{0.65\times0.35}{300}} = 0.65 - 0.0453 = 0.6047$$

答案：支持比率信賴區間的最低值 $p > 0.6047$

練習 8.48　為調查高雄市消費者對有機咖啡的支持比例，從高雄市消費者中隨機抽取 350 位消費者為樣本，統計其對有機咖啡的支持人數為 200 位。為保守估計有機咖啡的需求量，欲估計在信賴水準 99 %下接受率最低值的信賴區間？(答案有效位數四捨五入取到小數點後第 4 位)

題解：樣本比例 $\bar{p} = \frac{200}{350} = 0.5714$，樣本數 $n = 350$，$1-\alpha = 0.99$，$\alpha = 0.01$，$z_{\alpha} = z_{0.01} = 2.3267$(使用 Excel 軟體 NORM.S.INV 函數查詢獲得)。

題目需欲估計在信賴水準 99 %下接受率最低值的信賴區間，因此，判斷屬於左尾區間估計，母體比例 p 值未知時，利用樣本比例$\bar{p}$取代母體比例 p，故下限值為$\bar{p}-z_{\alpha}\times\sqrt{\frac{\bar{p}\times(1-\bar{p})}{n}}$。

$$\bar{p}-z_{\alpha}\times\sqrt{\frac{\bar{p}\times(1-\bar{p})}{n}} = 0.5714-2.3267\times\sqrt{\frac{0.5714\times0.4286}{350}} = 0.5714-0.0615 = 0.5099$$

答案：最低值的信賴區間 $p > 0.5099$

練習 8.49 欲估計常態母數均數的信賴區間，則下列敘述何者正確？(A)樣本數不變下，信賴水準愈高，信賴區間愈長；(B)樣本愈大，信賴區間愈長；(C)標準差愈小，信賴區間愈長；(D)以上皆正確。(99 年初等考試，統計學大意)A

討論議題

1.師生非同步討論議題：統計裡的信賴

各位同學請利用時間閱讀附加檔案中的【統計裡的信賴】pdf 檔案，此檔案是由高雄雄大學黃文璋老師所撰寫，嘗試從更完整的角度去理解信賴區間的意涵。第一回合請於 D+3 日中午 1200 點以前從「議題討論」區【張貼】標題：「信賴」，本文：請您詮釋一下【信賴】在統計學上的意涵(20 個字以上詮釋)。期望可以透過學習運用效益的交流，相互激勵，提升學習效益。

待有 35 篇第一回合【張貼】回應或第一回合【張貼】時間結束後，檢視其他同學的回應內容。第二回合【張貼】標題：「最佳詮釋」，本文：自己靜下心來，集思廣益，思考一下，哪一位同學詮釋得最好，其理由(10 個字以上)。第二回合【張貼】截止時間就是本議題在平台上的關閉時間。

2.師生非同步討論議題：信賴區間與信心水準的解讀

各位同學請利用時間閱讀附加檔案中的【信賴區間與信心水準】pdf 檔案，此檔案是由台灣大學研究生林柏佐同學所撰寫，特別要閱讀第三頁【信賴區間與信心水準】的單元內容，針對選舉的民調，推估母體比例的過程，詳細的閱讀瀏覽，期望同學可以更完整的理解母體比例信賴區間的推估。

第一回合請於 D+3 日中午 1200 點以前從「議題討論」區【張貼】標題：「信心水準」，本文：請您詮釋一下在選舉民調中【信心水準】的意涵(20 個字以上詮釋)。期望可以透過學習運用效益的交流，相互激勵，提升學習效益。

待有 35 篇第一回合【張貼】回應或第一回合【張貼】時間結束後，檢視其他同學的回應內容。第二回合【張貼】標題：「最佳詮釋」，本文：自己靜下心來，集思廣益，思考一下，哪一位同學詮釋得最好，其理由(10 個字以上詮釋)。第二回合【張貼】截止時間就是本議題在平台上的關閉時間。

3.學習者間同步教學討論議題：闖書平常考體驗

第一回合請於 D 日早上 0955 以前從「議題討論」區【張貼】標題：「考試心得」，本文：上大學第一學期最後一週期末考前進行闖書平常考，經過實際闖書平常考體驗後，請分享一下自己準備考試的得分關鍵(10 個字以上詮釋)。期望可以透過應考準備經驗的分享學習與相互交流，相互激勵，提升學習效益。

待有 35 篇第一回合【張貼】回應或第一回合【張貼】時間結束後，檢視其他同學的回應內容。即可開始張貼第二回合【張貼】標題：「領悟」，本文：自己靜下心來，集思廣益，思考一下，在

準備考試的技巧上，論述給自己最大領悟(10 個字以上)。透過同學間分享與討論，可以提升學習效益。加油！第二回合【張貼】截止時間就是本議題在平台上的關閉時間。

4.師生非同步討論議題：期末考後的檢討分析

第一回合請於 D+2 日中午 1200 點以前從「議題討論」區【張貼】標題：「檢討分析」，本文：上大學第一學期第一次挑戰鑑別度最高的期末考試題，經過一個學期的嚴謹教育洗禮後，寒假與未來精進學習的方向，請分享一下自己的檢討、分析與精進【最多 20 個字詮釋】。期望可以透過檢討、分析與精進的分享學習，相互激勵，提升學習效益。

待有 35 篇第一回合【張貼】回應或第一回合【張貼】時間結束後，檢視其他同學的回應內容。即可開始張貼第二回合【張貼】標題：「精進作為」，本文：自己靜下心來，集思廣益，思考一下，請論述在寒假和未來學習生涯上的精進具體作為【最多 20 個字詮釋】。第二回合【張貼】截止時間就是本議題在平台上的關閉時間。

5.師生非同步討論議題：期末精進學習

上大學第一學期已經到了第 17 週期末，統計學課程上經歷牛刀小試、來賓考試、上課練習、平常考、期中考試、議題討論等務實學習歷程，如何達到最佳的學習效益，已經有深入的體驗。第一回合請於 D+2 日中午 1200 點以前從「議題討論」區【張貼】標題：「最佳學習策略」，本文：詮釋一下自己認為最佳的學習策略【15 個字以上詮釋】。期望可以透過檢討、分析與精進的分享學習，相互激勵，提升學習效益。

待有 35 篇第一回合【張貼】回應或第一回合【張貼】時間結束後，檢視其他同學的回應內容。即可開始張貼第二回合【張貼】標題：「最後補強」，本文：自己靜下心來，集思廣益，思考一下，請論述接下來自己需要再精進補強之處【10 個字以上詮釋】。第二回合【張貼】截止時間就是本議題在平台上的關閉時間。

重點整理

Excel 函數彙整

Excel 函數	統計功能	輸入資料	輸出資料
T.INV	T分布	左尾機率，自由度	*t* 值
TINV	T分布	雙尾機率，自由度(舊版函數)	*t* 值(正值)
T.INV.2T	T分布	雙尾機率，自由度	*t* 值(正值)
T.DIST	T分布	X 值，自由度，累積運算	左尾機率(從-∞累積到 X 值機率)
TDIST	T分布	X 值，自由度，單尾/雙尾	右尾機率(從 X 值累積到∞機率)
T.DIST.2T	T分布	X 值，自由度	雙尾機率
T.DIST.RT	T分布	X 值，自由度	右尾機率(從 X 值累積到∞機率)
CHISQ.INV	卡方分布	左尾機率，自由度	卡方值
CHISQ.INV.RT	卡方分布	右尾機率，自由度	卡方值
CHIINV	卡方分布	右尾機率，自由度	卡方值

信賴係數(confidence coefficient) = 信賴水準(confidence level; level of confidence) = 信賴度(degree of confidence, β) = 推論成功機率 = $1-\alpha$

顯著水準(level of significance) = 推論失敗機率 = α

$$P(-z_{\frac{\alpha}{2}} \le \frac{\bar{x}-\mu}{\sigma_{\bar{x}}} \le z_{\frac{\alpha}{2}}) = 1-\alpha$$

$$P(\bar{x} - z_{\frac{\alpha}{2}} \times \sigma_{\bar{x}} \le \mu \le \bar{x} + z_{\frac{\alpha}{2}} \times \sigma_{\bar{x}}) = 1-\alpha$$

$$P(\bar{x} - z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}) = 1-\alpha$$

在常態分布下，母體平均值 μ 的信賴區間**通式**

點估計值 ± 誤差範圍(margin of error)

$$\bar{x} \pm z_{\frac{\alpha}{2}} \times \sigma_{\bar{x}} \qquad \bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$$

母體變異數 σ^2 和標準(偏)差 σ 未知

大量樣本數量($n \ge 30$)，變異數 σ^2 和標準(偏)差 σ 未知時，母體平均值 μ 的信賴區間為

$$\bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}} = \bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}}}{\sqrt{n}} = \bar{x} \pm z_{\frac{\alpha}{2}} \times \frac{\sqrt{\frac{\sum_{i=1}^{n}x_i^2 - \frac{\left(\sum_{i=1}^{n}x_i\right)^2}{n}}{n-1}}}{\sqrt{n}}$$

母體屬於常態分布，而樣本數量較少 $n < 30$，母體變異數 σ^2 和標準(偏)差 σ 未知的情況下，應以 t 分布估算母體平均值 μ 的信賴區間。

樣本數量少($n < 30$)，母體屬於常態分布，母體變異數 σ^2 和標準(偏)差 σ 未知時，母體平均值 μ 的信賴區間

$$\bar{x} \pm t_{\frac{\alpha}{2}} \times \frac{S}{\sqrt{n}}$$

母體的分布型態無法確定時，在母體變異數 σ^2 和標準(偏)差 σ 已知的情況下，樣本數量少($n < 30$)，可以利用柴比氏定理(Chebyshev’s theorem, Bienayme-Chebyshev rule)進行母體平均值 μ 的信賴區間估算。

$$P(|\bar{x}-\mu| \le k \times \sigma_{\bar{x}}) \ge (1-\frac{1}{k^2})$$

$$P(\bar{x} - k \times \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + k \times \frac{\sigma}{\sqrt{n}}) \ge (1-\frac{1}{k^2})$$

在信賴係數和信賴水準(信賴度)為 $1-\alpha$，在信賴區間的顯著水準為 α，母體比例 p 的信賴區間為

$$\bar{p} \pm z_{\frac{\alpha}{2}} \times \sigma_{\bar{p}} = \bar{p} \pm z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}} = \bar{p} \pm z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times q}{n}}$$

樣本數量少($n < 30$)時，於母體總數 N 為無窮大($N = \infty$)，樣本比率$\bar{p}$的抽樣分布會接近於二項分布，即為

$$\bar{p} \sim \text{二項分布}(p, \frac{p\times(1-p)}{n}) = \text{二項分布}(p, \frac{p\times q}{n})$$

樣本數量少($n < 30$)時，於母體總數 N 為有限數值，樣本比率$\bar{p}$的抽樣分布會接近於超幾何分布，即為

$$\bar{p} \sim \text{超幾何分布}(p, \frac{p\times(1-p)}{n} \times \frac{N-n}{N-1}) = \text{超幾何分布}(p, \frac{p\times q}{n} \times \frac{N-n}{N-1})$$

以樣本平均值$\bar{x}$估計母體平均值 μ，欲將估計誤差設定小於 A 值(與觀測值和平均值單位相同)，即$|\bar{x}-\mu| \le A$。設定在 $100 \times (1-\alpha)$%的信賴水準下，母體變異數 σ^2 和標準(偏)差 σ 已知，母體平均值 μ 落於$\bar{x} - z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{n}}$區間內，隨機抽樣的樣本數量需達 $n \ge \frac{z_{\frac{\alpha}{2}}^2 \times \sigma^2}{A^2}$。

以樣本比例$\bar{p}$估計母體比例 p，欲將估計誤差設定小於 A 值，即$|\bar{p}-p| \le A$。設定在 $100 \times (1-\alpha)$%的信賴水準下，母體的比例落於$\bar{p} - z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}} \le p \le \bar{p} + z_{\frac{\alpha}{2}} \times \sqrt{\frac{p\times(1-p)}{n}}$區間內，樣本數量需達 $n \ge \frac{z_{\frac{\alpha}{2}}^2 \times p \times (1-p)}{A^2}$。

利用卡方分布推估母體變異數的信賴區間

在卡方分布中，母體特定變數之觀測值的平均值 μ 無法獲得，僅以抽出 n 個基本單位的樣本平均值$\bar{x}$取代，在自由度 $df = v = n-1$ 的環境中，利用卡方值χ_{n-1}^2推估母體變異數的信賴區間：

$$P(\chi_{1-\frac{\alpha}{2},n-1}^2 \le \chi^2 \le \chi_{\frac{\alpha}{2},n-1}^2) = 1-\alpha$$

$$P(\chi^2_{1-\frac{\alpha}{2},n-1} \le \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sigma^2} \le \chi^2_{\frac{\alpha}{2},n-1}) = 1-\alpha$$

母體變異數 σ^2 在信賴水準 $1-\alpha$ 的信賴區間為：

$$\frac{S^2\times(n-1)}{\chi^2_{\frac{\alpha}{2},n-1}} \le \sigma^2 \le \frac{S^2\times(n-1)}{\chi^2_{1-\frac{\alpha}{2},n-1}}$$

$$\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\chi^2_{\frac{\alpha}{2},n-1}} \le \sigma^2 \le \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\chi^2_{1-\frac{\alpha}{2},n-1}}$$

在右單尾區間估計時，母體變異數 σ^2 和標準(偏)差 σ 已知，所有可能發生錯誤機率 α，皆在右側，推估母體平均值 μ 在信賴水準 $1-\alpha$ 下的信賴區間為：

$$P(\mu \le \bar{x} + z_\alpha \times \sigma_{\bar{x}}) = P(\mu \le \bar{x} + z_\alpha \times \frac{\sigma}{\sqrt{n}}) = 1-\alpha$$

在左單尾區間估計時，母體變異數 σ^2 和標準(偏)差 σ 已知，所有可能發生錯誤機率 α，皆在左側，推估母體平均值 μ 在信賴水準 $1-\alpha$ 下的信賴區間為：

$$P(\bar{x} - z_\alpha \times \sigma_{\bar{x}} \le \mu) = P(\bar{x} - z_\alpha \times \frac{\sigma}{\sqrt{n}} \le \mu) = 1-\alpha$$

在**右單尾區間估計**時，母體比例 p 已知，所有可能發生錯誤機率 α，皆在右側，推估(估算)母體比例 p 在信賴水準 $1-\alpha$ 下的信賴區間為：

$$P(p \le \bar{p} + z_\alpha \times \sqrt{\frac{p\times(1-p)}{n}}) = 1-\alpha$$

在**左單尾區間估計**時，母體比例 p 已知，所有可能發生錯誤機率 α，皆在左側，推估(估算)母體比例 p 在信賴水準 $1-\alpha$ 下的信賴區間為：

$$P(\bar{p} - z_\alpha \times \sqrt{\frac{p\times(1-p)}{n}} \le p) = 1-\alpha$$