11.8 母體變異數的區間估計

用χ^2分配可
母體變異數
賴區間。

在許多的情況下，人們對於母體變異數或標準差可能比對母體平均數及母體比例來得關心。譬如汽車零件墊片製造商除要求墊片的平均直徑爲某一長度外，還規定直徑長度的標準差不得超過某一範圍。因爲若直徑超過規定的長度則不能使用，因此對於母體變異數的推估就顯得非常重要。母體變異數代表資料的分散程度，比如股票投資報酬率的變異數的大小代表獲利的風險程度；又如精密螺絲其精密度與直徑大小有關，因此其直徑的變異數代表精密度，而成爲品管的目標。母體變異數通常是不知道的，因此必須利用樣本變異數來估計它。在推估時必須求算樣本變異數的抽樣分配–卡方分配（χ^2分配 chi-square distribution）。

11.8.1 卡方分配

設從母體中隨機抽取樣本$(X_1, X_2, \cdots, X_n)$，則樣本變異數S^2爲：

樣本變異數

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1} \tag{11.20}$$

在一般情況下，S^2抽樣分配不易直接導出，但若我們假設母體爲$X \sim N(\mu, \sigma^2)$，而將S^2乘以$n-1$再除以σ^2（或稱爲卡方化）於是可得新統計量，該統計量爲一卡方分配（χ^2分配），稱爲卡方統計量。

卡方統計量

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{\sigma^2} \sim \chi^2_{n-1} \tag{11.21}$$

該統計量爲自由度$(n-1)$的卡方分配。

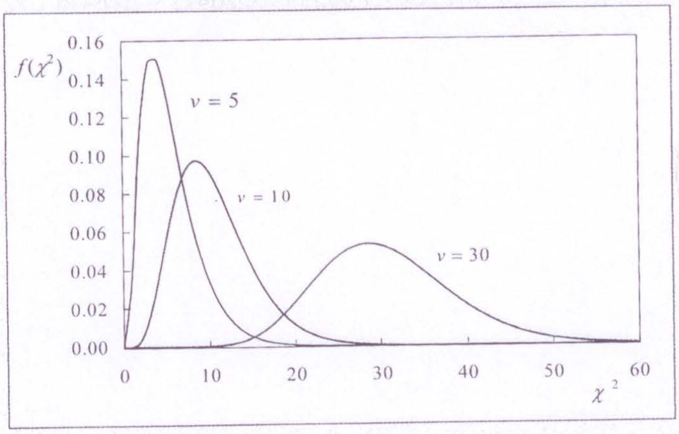
卡方分配可幫助我們推論母體變異數。卡方隨機變數的機率密度函數在初統中很少使用，因此我們只介紹卡方分配的重要定理與性質。

11.8.2 卡方分配的性質

卡方分配有幾個特性，說明如下：

①卡方分配爲一定義在大於等於 0（正數）範圍的右偏分配，不同的自由度決定不同的卡方分配，如圖 11.27 所示。

圖 11.27 卡方分配



②卡方分配只有一個參數即自由度，表爲 v 。卡方分配的平均數與變異數爲：

$$E(\chi_v^2) = v \quad , \quad V(\chi_v^2) = 2v$$

意即自由度爲 v 的卡方隨機變數，其平均數恰等於其自由度 v ，變異數爲 2 倍自由度。

③卡方分配隨自由度增加而逐漸對稱，當自由度趨近於無窮大時 $(v \to \infty)$ ，卡方分配會趨近常態分配，即：

$$\chi_v^2 \sim N(v, 2v)$$

④設 $X \sim N(\mu, \sigma^2)$ ，令 $Z^2 = (\dfrac{X-\mu}{\sigma})^2$ ，則 Z^2 爲自由度 1 的卡方分配。

卡方分配的臨界值

卡方分配常用的卡方值可查閱附表中的表六，部份卡方值如表 11.9 所示。卡方表中的值代表右尾機率爲 α 值的卡方值，表爲 χ_α^2 。表中第 1 行爲自由度 v ，第 1 列爲卡方右尾的機率。如 $\chi_{0.1}^2$ 表示卡方右尾的機率 $\alpha = 0.10$ 。我們可從卡方值表中找到常用的自由度及右尾機率的卡方值，例如若我們想求自

表 11.9　卡方值

$d.f$	$\chi^2_{0.995}$	$\chi^2_{0.95}$	$\chi^2_{0.90}$	$\chi^2_{0.10}$	$\chi^2_{0.05}$	$\chi^2_{0.005}$	$d.f$
1	.0000393	.0039321	.0157908	2.70554	3.84146	7.87944	1
2	.0100251	.102587	.210720	4.60517	5.99147	10.5966	2
3	.0717212	.351846	.584375	6.25139	7.81473	12.8381	3
4	.206990	.710721	1.063623	7.77944	9.48773	14.8602	4
5	.411740	1.145476	1.61031	9.23635	10.0705	16.7496	5
6	.675727	1.63539	2.20413	10.6446	12.5916	18.5476	6
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
17	5.69724	8.67176	10.0852	24.7690	27.5871	35.7185	17
18	6.26481	9.39046	10.8649	25.9894	28.8693	37.1564	18
19	6.84398	10.1170	10.6509	27.2036	30.1435	38.5822	19
20	7.43386	10.8508	12.4426	28.4120	31.4104	39.9968	20
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

資料來源：摘錄自附表表六。

由度 $v=5$，右尾機率 $\alpha=0.10$ 的卡方值，可在第 1 行自由度找到 5，第 1 列找到 $\chi^2_{0.10}$，兩者相交得到 $\chi^2_{5,0.10}=9.23635$，此即表示當自由度 $v=5$，$\alpha=0.10$ 時的卡方值爲 9.23635，如圖 11.28。

圖 11.28　卡方分配機率值

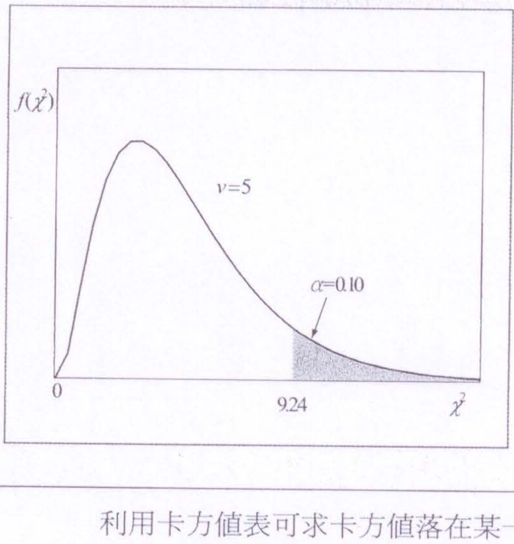
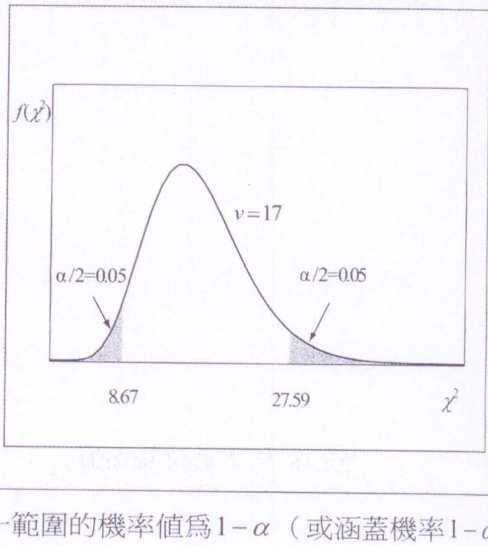


圖 11.29　卡方值的機率區間



利用卡方值表可求卡方值落在某一範圍的機率值爲 $1-\alpha$（或涵蓋機率 $1-\alpha$ 的卡方值的範圍）。表爲：

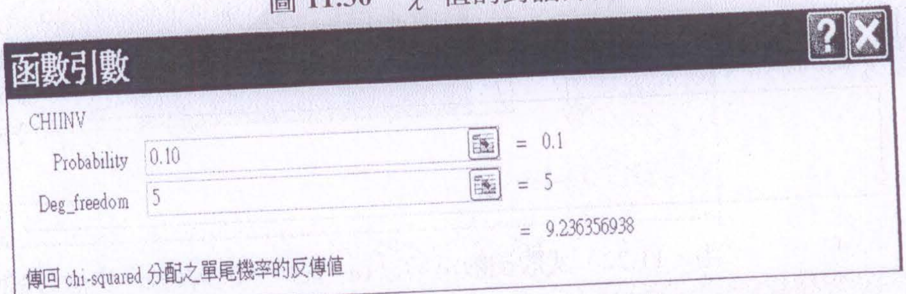
$$P(\chi^2_{1-(\alpha/2)} < \chi^2 < \chi^2_{\alpha/2}) = 1-\alpha$$

比如，若 $v=17$，$\alpha=0.10$ 則 $\chi^2_{1-(0.1/2)}=\chi^2_{0.95}=8.67176$，$\chi^2_{0.1/2}=\chi^2_{0.05}=27.5871$，意即 $v=17$ 時，$P(8.67176<\chi^2<27.5871)=0.9$，如圖 11.29。

Excel 的卡方值

Excel 的卡方值爲單尾機率 $P(\chi^2>\chi^2_\alpha)=\alpha$ 的卡方值。例如求自由度 $df=5$，右尾機率爲 0.1 的求法，如圖 11.30 所示。

圖 11.30　χ^2 值的對話方塊



1.選取「公式」「插入函數 f_x」、「統計」、「CHIINV」、按確定。

2.在 Probability 輸入 0.10。

3.在 deg_freedom 輸入自由度 5。

11.8.3　由卡方分配導出母體變異數的信賴區間

母體變異數的區間估計及標準差的估計方法，可分爲兩個步驟：

①第 1 個步驟仍是選擇樣本變異數 S^2 做爲母體變異數 σ^2 的點估計式。即：

$$S^2=\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{n-1}$$

由於它是一個不偏估計式，所以一般選擇 S^2 來估計 σ^2。

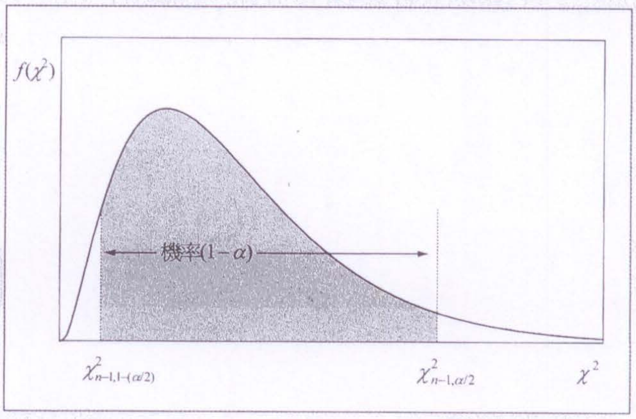
②第 2 個步驟是利用卡方分配導出 σ^2 的信賴區間

由於 $\frac{(n-1)S^2}{\sigma^2}$ 爲一自由度 $n-1$ 的卡方分配，因此利用 χ^2_{n-1} 的分配可求得統計量 $\frac{(n-1)S^2}{\sigma^2}$ 在 $1-\alpha$ 的機率區間，以公式表示如下：

$$P(\chi^2_{n-1,1-(\alpha/2)} \le \frac{(n-1)S^2}{\sigma^2} \le \chi^2_{n-1,\alpha/2}) = 1-\alpha \qquad (11.22)$$

以圖 11.31 表示。

圖 11.31 $(1-\alpha)$ χ^2 值的機率區間



由（11.22）式取倒數再乘以 $(n-1)S^2$，則可得 σ^2 的信賴區間為：

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

σ^2 的信賴區間並不對稱於 S^2，因此一般只能表為：

母體變異數 σ^2 的信賴區間

$$\frac{(n-1)S^2}{\chi^2_{n-1,\alpha/2}} \le \sigma^2 \le \frac{(n-1)S^2}{\chi^2_{n-1,1-(\alpha/2)}} \qquad (11.23)$$

母體標準差的區間估計

母體的變異數與標準差雖都是用來衡量資料的分散情形，但標準差因其單位與資料的單位相同，因此較為常用。在進行母體標準差 σ 的區間估計時，我們選擇樣本標準差 S 做為它的點估計式。雖然樣本標準差 S 不是 σ 的不偏估計式，但卻是一致性估計式，在目前尚無較佳的估計式可用來做為 σ 的點估計式的情況下，我們只好選擇樣本標準差 S 來估計 σ。而在做母體標準差的區間估計時，將 σ^2 的區間估計的公式開根號可得 σ 的區間估計公式如下：

$$\sqrt{\frac{(n-1)S^2}{\chi^2_{n-1,\alpha/2}}} \le \sigma \le \sqrt{\frac{(n-1)S^2}{\chi^2_{n-1,1-\alpha/2}}} \qquad (11.24)$$

例 11.13 貓食罐頭的容量

假設進口商抽驗 20 罐 catsin 貓食罐頭，發現每罐容量的樣本標準差為 25g，據此推測每罐容量的母體變異數及標準差 95%信賴區間為何？

解 利用（11.23）式求解如下：

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

$$\frac{19(25)^2}{32.85} \le \sigma^2 \le \frac{19(25)^2}{8.91}$$

其中 $\chi^2_{19,0.025} = 32.85$，$\chi^2_{19,0.975} = 8.91$。

因此可得 95%信賴水準下母體變異數的信賴區間為：

$$361.49 \le \sigma^2 \le 1332.77$$

依（11.24）式可得 95%信賴水準下標準差的信賴區間為：

$$\sqrt{361.49} \le \sigma \le \sqrt{1332.77}$$

$$19.01 \le \sigma \le 36.51$$

故可推論：「貓食罐頭容量的母體變異數，在 95%信賴水準下的信賴區間為 361.49~1332.77 之間。貓食罐頭容量的標準差 95%的信賴區間為 19.01g~36,51g」 ■■■

11.9 Excel 的使用

求算母體平均數的信賴區間時，可利用 Excel 的「公式」、「f_x 插入函數」來進行，其步驟如下：

1.選取「公式」、「f_x 插入函數」、「統計」、「CONFIDENCE」、按確定。

2.在 alpha 空白方格輸入顯著水準 α 值，在 standard_dev 輸入標準差的值在 size 輸入樣本數（sample size），按確定。產生的數值為 $Z_{\alpha/2}\sigma/\sqrt{n}$，此一數值為估計誤差。