第 4 章　數學附錄

一 設 X 為代表母體資料特性的變數，其觀察值為：$x_1, x_{2,} \Lambda , x_N$，則

$$\sum_{i=1}^{N}(x_i - \mu)^2 \le \sum_{i=1}^{N}(x_i - A)^2$$

其中：A 為任意實數，μ 為 X 母體之均數。

證明 證明如下：

$$\begin{aligned}\sum_{i=1}^{N}(x_i - A)^2 &= \sum_{i=1}^{N}(x_i - \mu + \mu - A)^2 \\ &= \sum_{i=1}^{N}[(x_i - \mu)^2 + (\mu - A)^2 + 2(\mu - A)(x_i - \mu)] \\ &= \sum_{i=1}^{N}(x_i - \mu)^2 + \sum_{i=1}^{N}(\mu - A)^2 + 0\end{aligned}$$

故 $\sum_{i=1}^{N}(x_i - A)^2 \ge \sum_{i=1}^{N}(x_i - \mu)^2$

同理可得，若 X 為樣本資料之變數，其觀察值為 $x_1, x_{2,} \Lambda , x_n$，則

$$\sum_{i=1}^{n}(x_i - \overline{X})^2 \le \sum (x_i - A)^2$$

其中：A 為任意實數，$\overline{X}$ 為樣本平均數。

二 證明 $\sum_{i=1}^{N}(x_i - \mu) = 0$

證明 $\sum_{i=1}^{N}(x_i - \mu) = \sum_{i=1}^{N} x_i - \sum_{i=1}^{N} \mu = \sum_{i=1}^{N} x_i - N\mu = \sum_{i=1}^{N} x_i - \sum_{i=1}^{N} x_i = 0$

同理可得：$\sum_{i=1}^{n}(x_i - \overline{X}) = 0$

三 若 $Y = a + bX$，則 $\mu_Y = a + b\mu_X$

證明 $\mu_Y = \dfrac{\sum (a + bX)}{N} = \dfrac{\sum a + b(N \cdot \mu_X)}{N} = a + b\mu_X$

同理，若 $Y = a + bX$，則 $\overline{Y} = a + b\overline{X}$。

四 設有兩組資料，其分別以 X_1, X_2 表示，已知其平均數分別為 μ_1, μ_2，則 X_1 與 X_2 兩組資料之總平均 μ 為：

$$\mu = \frac{N_1\mu_1 + N_2\mu_2}{N_1 + N_2}$$

其中 N_1 與 N_2 分別為 X_1 與 X_2 之個數。

證明 根據定義

$$\mu = \frac{\sum X_1 + \sum X_2}{N_1 + N_2} = \frac{N_1 \cdot \mu_1 + N_2 \cdot \mu_2}{N_1 + N_2}$$

同理可得，X_1 與 X_2 樣本資料之總平均數($\overline{X}$)

$$\overline{X} = \frac{n_1 \overline{X}_1 + n_2 \overline{X}_2}{n_1 + n_2}$$

設以 X_1, X_2, Λ, X_k 表示，已知其平均數分別為 $\mu_1, \mu_2, \Lambda, \mu_k$，則 X_1, X_2, Λ, X_k，k 組資料之總平均 μ 為：

$$\mu = \frac{\sum_{i=1}^{k} N_i \mu_i}{\sum_{i=1}^{k} N_i}$$

其中 N_i 分別為 X_i 之觀察個數。

設 X_{ij} 代表 X_i 中第 j 個觀察值 $i = 1, \Lambda, k$，$j = 1, \Lambda, N_i$，則

$$\mu = \frac{\sum_{i=1}^{k} \sum_{j=1}^{N_i} X_{ij}}{\sum_{i=1}^{k} N_i} = \frac{\sum_{j=1}^{N_1} X_{1j} + \sum_{j=1}^{N_2} X_{2j} + \Lambda + \sum_{j=1}^{N_k} X_{kj}}{\sum N_i}$$

$$= \frac{N_1 \mu_1 + N_2 \mu_2 + \Lambda + N_k \mu_k}{\sum_{i=1}^{k} N_i} = \frac{\sum_{i=1}^{k} N_i \mu_i}{\sum_{i=1}^{k} N_i}$$

五 證明 $\sum (x_i - \mu)^2 = \sum x_i^2 - N\mu^2$

證明

$$\sum (x_i - \mu)^2 = \sum x_i^2 - 2\mu \sum x_i - N\mu^2 = \sum x_i^2 - 2N\mu^2 + N\mu^2$$

$$= \sum x_i^2 - N\mu^2$$

六 設 $Y = a + bX$，則 $\sigma_Y^2 = b^2 \sigma_X^2$，$\sigma_Y = |b| \sigma_X$

證明

$$\sigma_Y^2 = \frac{1}{N} \sum (a + bx - a - b\mu_x)^2 = \frac{\sum (bx - b\mu_x)^2}{N} = \frac{1}{N} b^2 \sum (x - \mu_x)^2 = b^2 \sigma_x^2 \text{ 故 } \sigma_Y = |b| \sigma_X$$

七 設 X_1 與 X_2 兩母體，其平均數與變異數分別為 μ_1, μ_2 及 σ_1^2, σ_2^2，則兩母體的全體變異數為：

$$\sigma^2 = \frac{1}{N_1 + N_2} \{ N_1 [\sigma_1^2 + (\mu_1 - \mu)^2] + N_2 [\sigma_2^2 + (\mu_2 - \mu)^2] \}$$

式中：$\mu = \dfrac{N_1 \mu_1 + N_2 \mu_1}{N_1 + N_2}$ (兩母體之全體平均數)

證明 $\sigma^2=\dfrac{\sum_{i=1}^{N_1}(x_{1i}-\mu)^2+\sum_{i=1}^{N_2}(x_{2i}-\mu)^2}{N_1+N_2}$

$$=\frac{1}{N_1+N_2}\left\{\sum_{i=1}^{N_1}\left[(x_{1i}-\mu_1)^2+(\mu_1-\mu)^2\right]+\sum_{i=1}^{N_2}\left[(x_{2i}-\mu_2)^2+(\mu_2-\mu)^2\right]\right\}$$

$$=\frac{1}{N_1+N_2}\left\{N_1\sigma_1^2+N_1(\mu_1-\mu)^2+N_2\sigma_2^2+N_2(\mu_2-\mu)^2\right\}$$

$$=\frac{1}{N_1+N_2}\left\{N_1[\sigma_1^2+(\mu_1-\mu)^2]+N_2[\sigma_2^2+(\mu_2-\mu)^2]\right\}$$

$$=\frac{1}{N_1+N_2}[N_1(\sigma_1^2+\mu_1)^2+N_2(\sigma_2^2+\mu_2)^2]-\mu^2$$

八 設有 k 個小母體，其個數分別為 N_1,Λ,N_K，平均數 μ_1,Λ,μ_K，變異數為 $\sigma_1^2,\Lambda,\sigma_K^2$，全體平均數($\mu$)與變異數 σ^2 為：

$$\mu=\frac{\sum_{i=1}^{K}N_i\mu_i}{\sum_{i=1}^{K}N_i},\quad \sigma^2=\frac{\sum_{i=1}^{K}N_i[\sigma_i^2+(\mu_i-\mu)^2]}{\sum N_i}$$

證明 $\sigma^2=\dfrac{\sum_{i=1}^{K}\sum_{j=1}^{N_i}(x_{ij}-\mu)^2}{\sum N_i}=\dfrac{\sum_{i=1}^{K}\sum_{j=1}^{N_i}[(x_{ij}-\mu_i)+(\mu_i-\mu)]^2}{\sum N_i}$

$$=\frac{1}{\sum N_i}\sum_{i=1}^{K}\sum_{j=1}^{N_i}[(x_{ij}-\mu_i)^2+2(x_{ij}-\mu_i)(\mu_i-\mu)+(\mu_i-\mu)^2]$$

$$=\frac{\sum_{i=1}^{K}\sum_{j=1}^{N_i}[(x_{ij}-\mu_i)^2+(\mu_i-\mu)^2]}{\sum N_i}=\frac{\sum_{i=1}^{K}[N_i\sigma_i^2+N_i(\mu_i-\mu)^2]}{\sum N_i}$$

九 幾何平均數的應用

設

$$P_n=P_0(1+r)^n$$

式中，r：年平均增加率(變動率)，P_n：第 n 期物價，P_0：第 0 期物價，則 P_0,P_1,Λ,P_n 為以 $(1+r)$ 為公比的等比數列。若欲求公比 $(1+r)$，即等於求 $\dfrac{P_1}{P_0},\dfrac{P_2}{P_1},\Lambda,\dfrac{P_n}{P_{n-1}}$ 的幾何平均數。因

$$\frac{P_n}{P_0}=(1+r)^n,\quad 1+r=\left(\frac{P_n}{P_0}\right)^{\frac{1}{n}}$$

又

$$\left(\frac{P_n}{P_0}\right)^{\frac{1}{n}}=\left(\frac{P_1}{P_0}\times\frac{P_2}{P_1}\times\frac{P_3}{P_2}\Lambda\frac{P_n}{P_{n-1}}\right)^{\frac{1}{n}}=\frac{P_1}{P_0},\frac{P_2}{P_1},\Lambda,\frac{P_n}{P_{n-1}}\text{ 的幾何平均數}=1+r$$

十 分組資料的動差

以平均數為中心的r級動差，分組資料的公式如下：

母體資料：

$$\frac{1}{N}\sum_{i=1}^{k}(x_i-\mu)^r f_i$$

式中：x_i為組中點，f_i為組次數，N為母體元素個數，k為組數。

樣本資料

$$\frac{1}{n}\sum_{i=1}^{k}(x_i-\overline{X})^r f_i$$

式中：x_i為組中點，f_i為組次數，n為樣本元素個數，k為組數。

以平均數為中心的2級、3級、4級動差的簡易計算法如下：

$M_2=U_2-(U_1)^2$

$M_3=U_3-3U_2U_1+2(U_1)^3$

$M_4=U_4-4U_3U_1+6U_2(U_1)^2-3(U_1)^4$

其中$U_r=\frac{\sum X^r}{N}$如$U_1=\frac{\sum X}{N}$，$U_2=\frac{\sum X^2}{N}$，$U_3=\frac{\sum X^3}{N}$，$U_4=\frac{\sum X^4}{N}$。U_r稱為以原點為中心的動差或簡稱為原動差。

十一 柴比氏定理

設X為一隨機數，其平均數為μ，變異數為σ^2，則對任何正數$K, K>0$。則

$$P(|X-\mu|<K\sigma)\geq 1-\frac{1}{K^2}，\quad P(|X-\mu|\geq K\sigma)\leq\frac{1}{K^2}$$

證明

$\sigma^2=\sum(x-\mu)^2 f(x)$

$=\sum_{S}(x-\mu)^2 f(x)+\sum_{\overline{S}}(x-\mu)^2 f(x)\geq\sum_{S}(x-\mu)^2 f(x)$

式中：$S=\left\{x\mid(x-\mu)^2\geq K^2\sigma^2\right\}$

故

$\sigma^2\geq K^2\sigma^2\sum_{S}f(x)$

$$\sigma^2\geq K^2\sigma^2 P(|x-\mu|\geq K\sigma)$$

即$P(|X-\mu|\geq K\sigma)\leq\frac{1}{K^2}$，$P(|X-\mu|<K\sigma)\geq 1-\frac{1}{K^2}$。