三、採樣方法

採樣最主要者為代表性，理論上可根據統計原理，推估需要之採樣數量，既可滿足樣品之代表性，又可以最省錢、省時、省人力的方法完成採樣分析工作，具有統計意義之採樣方法簡要介紹如下：

3.1 簡易隨機採樣（simple random sampling）

簡易隨機採樣為各種隨機採樣方法之基礎，針對母體中的全部個體，使之完全在均勻機率分布（uniform distribution）下抽取樣本，而每一個體被抽出之機率均為已知且相等，而簡易隨機採樣法的樣本取得，可利用隨機數量（亂數表）。但須先對母體編號，然後依照機率抽取，而當母體數大時，此種方法會比較麻煩。

在族群總樣品數為 N 的母體中任意取 n 個不同的樣品來（如果有重複相同的，則再取，直至不同為止），設採得為 X_1，X_2，…X_n，則此即為所得之簡單隨機樣本。此法雖為常用，但它必須先知道 N 值，而 N 往往不能預先知道。並且有時當 N 不太小時，一再證驗目前所抽到的樣本是否已抽過（因此必須重抽），也是較不方便的。以下為有關簡易隨機採樣方法中族群總樣品參數及平均值的採樣誤差：

$$\mathrm{Var}(\overline{X}) = \left[\frac{N-n}{N}\right]\frac{S^2}{n} = (1-f)\frac{S^2}{n}$$

$$\mathrm{Var}(N\overline{X}) = N^2\left[1-\frac{n}{N}\right]\frac{S^2}{n} = N^2(1-f)\frac{S^2}{n}$$

$\overline{X}$：為平均值（見表 2.1 中（2a）項）

f：$\frac{n}{N}$ 採樣因子

n：為採樣數

N：母體總數

S^2：樣品變異值（表 2.1（3a）項）

從採樣樣品中所得的數據估算母體真值時，常有人為因素或其他不確定因素所導致的誤差，而在其誤差中有部分是因隨機的選擇步驟而產生的變異數（Var），吾人稱之為隨機採樣誤差（random sampling error），由平均值 $\overline{X}$ 或族群母體總量本身分布的變異所產生的誤差為 $Var(\overline{X})$ 及 $Var(N\overline{X})$，而其中 S 為觀測樣品的標準偏差，其計算如下：

$$S^2 = \frac{\sum_{i=1}^{n} Xi^2 - (\sum_{i=1}^{n} Xi)^2 / n}{n-1}$$

因為在統計學上標準誤差 $\sigma(X) = \sqrt{Var(X)}$，所以關於樣品平均值及總量值 $N\overline{X}$ 的標準誤差為：

$$\sigma^2(\overline{X}) = \frac{1}{n}(1-f)S^2 \Rightarrow \sigma(\overline{X}) = \sqrt{\frac{(1-f)}{n}}S$$

$$\sigma^2(N\overline{X}) = \frac{1}{n}N^2(1-f)S^2 \Rightarrow \sigma(N\overline{X}) = \sqrt{\frac{(1-f)}{n}}NS$$

以上所討論的皆為抽出樣本後並不放回原母體的採樣情況。

另外有關於採樣點數目推估及信賴區間（CI）和常規極限值的關係，是試誤法（trial and error method）求得，由以下程序可得：

1. 先以任意的採樣點數(約 6～10)所算的 $\overline{X}$ 及 S^2 之估計值，而 $\overline{X}$ 和 S^2 計算方法（見表 2.1 第(2)、(3)項）。
2. 根據 1 步驟所得數據，去推算合理的採樣點 n_1（見表 2.1 第(8)項及表 2.2）。
3. 再根據合理的採樣點數 n_1 去採得 n_1 個樣品（可以順便多採少數幾點）以防估計或其他因素的誤差。
4. 分析每點樣品測值，是否偏離過大，若不符合標準，則予以剔除。

5. 根據合理數據來計算$\overline{X}$，S^2及標準偏差σ（即 S），與$\sigma_{\overline{x}}$（即$S_{\overline{x}}$）（見表 2.1 之第(2a)、(3a)、(4)、(5)項）。

6. 藉由 5 步驟所得之$\overline{X}$，$S_{\overline{x}}$，計算 CI（信賴區間）（見表 2.1 第(6)項）的信賴上、下限，與 RT（regulatory threshold）常規臨界值（見表 2.1 第(7)項）比較，是否超出依過去的統計數據所訂的 RT 值，若信賴上限超出 RT 值，而$\overline{X}$值又可視為真值μ時，則就不用再繼續採樣，而要探討此批廢棄物的其他背景特性。但若此次$\overline{X}$值不可代表真值，則就要根據新的$\overline{X}$等數據，再去算進一步的採點，一直重複，直至合理為止。或者是將步驟 3 的多餘樣品併入分析也是另一種方式。因為由表 2.1 第(8)項的採樣點數。

$$n=\frac{t_{0.1}{}^{2}S^{2}}{(RT-\overline{X})^{2}}$$

中可知當$\overline{X}$與 RT 值相近時，需採較多之樣品數，以避免判定過程中誤差的機會發生。

3.2 分層隨機採樣（stratified random sampling）

在分層隨機採樣過程中，將族群母體中各個體按不同目的有關之某種分類標準分為若干類或組，每類或組稱為層（stratum），並使得在此母體中的每一個元素皆屬於其中的一層且是唯一的一層，再由各層中以簡易隨機採樣法選取適量樣本的方法。

而在採樣人員認為採樣的對象其樣品可能有類別分布不同的趨向時，即可採用此法。而統計量的計算方法如下：

$m=m_1+\ldots\ldots+m_L$

$n=n_1+n_2+\ldots\ldots n_L$

m_L：為第 L 層內之母體數

m：為母體總數

$$\overline{X}_L=\frac{1}{n_L}\sum_{i=1}^{n_L}X_{Li}$$

$$S^2=\sum_{i=1}^{L}W_iS_i^2$$ 共 L 層

n_L：第 L 層所採的樣本數

n：為樣本總數

$\overline{X}_L$：為第 L 層樣本參數平均值

X_{Li}：為 L 層第 i 個樣品參數值

而依美國環保署採樣方法，分層隨機採樣樣品數：

$$\overline{X}=\sum_{k=1}^{r}W_k\overline{X}_L$$

$$S^2=\sum_{k=1}^{r}W_kS_k^2$$

$$n=\frac{t_{0.1}{}^2S^2}{\Delta^2}$$

$$\Delta=(RT-\overline{X})$$

W_k：k 層所佔比例

S_k^2：k 層之變異值

共 r 層

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{L}n_i\overline{X}_i$$

共 L 層

$\overline{X}$為所有樣品參數平均值

$$=\frac{1}{n}\sum_{i=1}^{L}\sum_{j=1}^{n_i}Xij$$

$V(\overline{X}i)$為第 i 層樣品數據推估得之隨機採樣誤差

$$V(\overline{X}i)=\frac{m_i-n_i}{m_i}\bullet\frac{S_i^2}{n_i}$$

其中 $S_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i} (X_{ij} - \overline{X}_i)^2 \quad 1 \le i \le L$

而 $\overline{X} = \frac{1}{n}\sum_{i=1}^{L} n_i \overline{X}_i = \sum_{i=1}^{L} W_i \overline{X}_i$ W_i 為各層之權數

$$V(\overline{X}) = \sum_{i=1}^{L} W_i V(\overline{X}_i) = \sum_{i=1}^{L} (\frac{n_i}{n})^2 (\frac{m_i - n_i}{m_i}) \frac{S_i^2}{n_i}$$

另外在樣本數之分配及決定有下述相同表示：

1. 等分配置 $n_i = \frac{n}{L}$

2. 比例配置 $n_i = \frac{m_i}{m} \bullet n$

3. 最佳配置 $n_i = \frac{m_i \quad S_i \quad C_i}{\sum m_i \quad S_i \quad C_i} \bullet n$

C_i 為各 i 層每一單位採樣之費用而在分層隨機採樣中的經費為 C 總成本：$C = C_o + \sum_{i=1}^{L} C_i n_i$ 其中 C_o 為基本費用

4. 尼氏分配（neyman allocation）

$$n_i = \frac{m_i S_i}{\sum m_i S_i} \bullet n$$

而採樣數之推估(n)，參照其它採樣方法之採樣數計算公式而得。

3.3 系統隨機採樣（systematic random sampling）

將全體所有的個體予以依次編號、訂定固定間隔或時間。每隔若干號、固定間隔或固定時間抽取一個樣本，而樣本數與母體數關係，則端看間隔的劃分而定，間隔大時樣品數小，間隔小時樣本數大，例如每隔 20

個號碼或時間取一個樣品，則總樣品數佔母體族群為 5%，同理，每隔 100 個取得一個，則樣本數為全體的 1%。若全體族群母體個數為 N 時，所要採樣樣本總數為 n 時，吾人取 $I=\frac{N}{n}$ 稱為抽樣間隔（sampling interval），若 I 不為整數，則用四捨五入法取其整數。

至於樣本個體依次為 S，2S，3S...而這些在全體中的位次，則可用下列公式求之

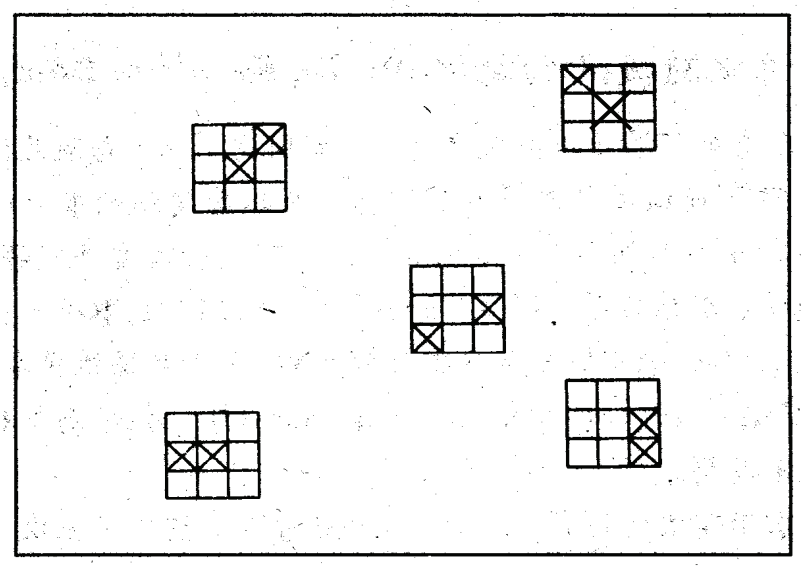
$O=(S-1)\times I+f$

O 為樣本個體在全體中的位置（在全部母體已依次編號下），S 為等間隔的樣品順序位次，I 為間隔大小，f 為第一個抽取樣品在全部母體所佔的位置，至於其他統計量之計算，因其中 f 是由第一區間(I)中隨機取得，故平均值等各參數均與簡單隨機採樣之計算相同。而若將樣品層化現象的母體分層，每一層區域內再以系統隨機採樣，則各參數之計算方式與分層隨機採樣相同。

3.4 階段式採樣（subsampling）

階段式抽樣是先由一個原始 N 個單位（一單位中含有多個樣品）中抽取 n 個單位的隨機樣本，稱為主要（或第一段）抽樣單位（primary sampling unit 簡稱 psu)而再從 n 個單位中的第 i 個被選的主要單位再選 m 個單位，稱為次要（或第二段）抽樣單位（secondary sampling unit 簡稱 ssu）而主要單位當中皆含 M 個單位，若就只進行至 ssu 中分析，則稱為二段式採樣，若繼續由 ssu 抽取更小單位進行採樣，則為三段式採樣，而三段以上之採樣，稱之為多段式採樣（multistage sampling)。而圖 3.1 為二段式採樣之一例子，而在此所討論的僅為在主要單位選出 M 個單位，其所含次要單位數皆相同的情形，以下為二段式採樣各統計量之計算（圖

3.1）：



⊠代表單元採樣樣本

圖 3.1　2 階段式採樣圖（N=81，n=5；M=9，m=2）

$$\overline{X}=\frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}X_{ij}=\frac{1}{n}\sum_{i=1}^{n}\overline{X}_{ij}$$ $\overline{X}$為平均值

$$S^2(\overline{X})=V(\overline{X})=\frac{N-n}{N}\bullet\frac{S_1^2}{n}+\frac{M-m}{M}\frac{S_2^2}{n\cdot m}$$ $V(\overline{X})$：平均值之變異數

其中

$$S_1^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$$ 為母體族群 N 個主要單位之推估變異數

$$S_2^2=\frac{1}{n(m-1)}\sum_{i=1}^{n}\sum_{j=1}^{m}(X_{ij}-\overline{X})^2$$ 為母體族群 M 個次要單位之推估變異數

N：整個母體族群所含的 psu 總個數

n：隨機取樣抽出的 psu 個數

M：每一個從 psu 中抽取的次單位所含的樣本母體總數

m：次單位中隨機取出的次樣本數

而樣本數的決定必須考慮經費及精確度 2 個限制，由下列 2 個方程式

決定

$$V(\overline{X}) = \frac{N-n}{N}\frac{S_1^2}{n} + \frac{M-m}{M}\frac{S_2^2}{n \cdot m}$$

$C=C_1 \times n+C_2 \times n \times m$

在一定的經費之下 V(X)最小得最佳次樣品數(m_{opt})

$$m_{opt} = \frac{S_2}{S_1^2 - S_2^2/M} C_1/C_2$$

n 可用上式 $C=C_1 \times n+C_2 \times n \times m$ 求出

另外在一定的抽樣誤差下，而經費要最經濟時

$$m = \frac{S_2}{S_1^2 - \frac{S_2^2}{M}} \frac{C_1}{C_2}$$

同理 n 可由 $C=C_1 \times n+C_2 \times n \times m$ 求出

3.5 權威判斷方法（authoritative judgement sampling）

此抽樣方法為非機率抽樣，是由清楚了解所採集該廢棄物性質的人所採集，採出其所認為最能代表族群的元素樣本。而此種方法是要在掌握的採樣區間中，並無明顯分隔不平均的特徵下進行，採樣者根據現場貯存方式，以自己的主觀判斷來選擇，但是這些樣本結果的品質決定於取樣的人員不同而有離異。所以當族群分布均勻時同此法較為容易方便，但若以這些判斷樣品所得出的結果當做母體的有效推論時，需要謹慎考量。

3.6 混合採樣法（composite sampling）

前述五種方法皆為將個別單獨樣品測量分析後再予以平均，得出一個平均代表值，而本法則為將個別樣品先行取出後再予以充分攪拌混合均勻（有時按權重比例混合），再由此均勻混合部分抽出次樣品，再分析測量

此單一混合樣品的性質，雖然此種方法中一次採樣只產生一組較多數量的混合樣品，可使結果更具代表性。而混合個別樣品數量亦可由表 2.1，第(8)式算出。

至於其他統計量計算，方法與階段式類似，可參照其方法計算。

綜合上述六種採樣方法，吾人可將其中之各特點，列出其優缺點以供比較。(表 3.1)

表 3.1　各採樣理論之優缺點

採樣方法	優　點	缺　點
簡易隨機採樣	1. 方法簡單。 2. 因易估算族群總值及採樣誤差，準確度、精確度高。	因採樣樣品較為分散，所需採樣人力及經費較大。
分層隨機採樣	1. 若每層內之差異度愈小，可得更高精確度(比簡易隨機要高)。 2. 可求得各層之估算值。	1. 樣品數據資料之整理、推算工作，會比簡易隨機為繁。 2. 族群分布為未知傾向時會降低準確及精確度。
系統隨機採樣	1. 依隨機方式只須決定第一個，其餘依序，故較方便。 2. 污染物質分布均勻時，可得高精確度。	1. 族群分布為未知傾向時會降低準確及精確度。 2. 樣本個體成周期循環，而又與採樣間隔相近時誤差較大。
階段式採樣	採樣手續較方便。	1. 誤差較大。 2. 整理分析較繁雜。
權威判斷採樣	簡易、方便。	由於錯誤的判斷，誤差可能甚大，無法估算族群平均數及採樣標準偏差。
混合採樣	綜合簡易隨機採樣及階段式採樣優點。	為求更具代表性，需採集較多個別樣品，人力經費並沒有顯著節省。