第12章　數學附錄

一　設 $X_1 \sim (\mu_1, \sigma_1^2)$， $X_2 \sim (\mu_2, \sigma_2^2)$，且 X_1 與 X_2 獨立，自 X_1 與 X_2 中獨立隨機抽取二個大樣本（ $n_1 \ge 30, n_2 \ge 30$)O，則 $\mu_1 - \mu_2$ 的信賴區間為 $(\overline{X}_1 - \overline{X}_2) \pm Z_{\alpha/2} \cdot \sigma_{\overline{X}_1 - \overline{X}_2}$，其中

$$\sigma_{\overline{X}_1 - \overline{X}_2} = \sqrt{(\sigma_1^2 / n_1) + (\sigma_2^2 / n_2)}$$

證明①估計 $\mu_1 - \mu_2$ 時，選擇 $\overline{X}_1 - \overline{X}_2$ 為點估計式

②依據中央極限定理可得

$$\overline{X}_1 \sim N(\mu_1, \sigma_1^2 / n_1), \quad \overline{X}_2 \sim N(\mu_2, \sigma_2^2 / n_2)$$

再依據常態分配加法定理可得 $\overline{X}_1 - \overline{X}_2$ 的抽樣分配：

$$\overline{X}_1 - \overline{X}_2 \sim N(\mu_1 - \mu_2, \sigma_1^2 / n_1 + \sigma_2^2 / n_2)$$

③由於 $\overline{X}_1 - \overline{X}_2$ 為常態分配，因此可得 $\overline{X}_1 - \overline{X}_2$ 的 $(1-\alpha)$ 的機率區間為：

$$P(\left|(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)\right| \le Z_{\alpha/2} \cdot \sigma_{\overline{X}_1 - \overline{X}_2}) = 1 - \alpha$$

其中 $\sigma_{\overline{X}_1 - \overline{X}_2} = \sqrt{(\sigma_1^2 / n_1) + (\sigma_2^2 / n_2)}$

④根據上面的機率區間可得 $(\mu_1 - \mu_2)$ 的 $(1-\alpha)$ 的信賴區間為：

$$P((\overline{X}_1 - \overline{X}_2) - Z_{\alpha/2} \cdot \sigma_{\overline{X}_1 - \overline{X}_2} \le \mu_1 - \mu_2 \le (\overline{X}_1 - \overline{X}_2) + Z_{\alpha/2} \cdot \sigma_{\overline{X}_1 - \overline{X}_2})$$
$$= 1 - \alpha$$

二　證明 $E(S_p^2) = \sigma^2$，即 S_p^2 為 σ^2 的不偏估計式。

證明

$$E\left(\frac{\sum (X_1 - \overline{X}_1)^2 + \sum (X_2 - \overline{X}_2)^2}{n_1 + n_2 - 2}\right)$$
$$= \frac{1}{n_1 + n_2 - 2} E[\sum (X_1 - \overline{X}_1)^2 + \sum (X_2 - \overline{X}_2)^2]$$
$$= \frac{1}{n_1 + n_2 - 2} [E(\sum (X_1 - \overline{X}_1)^2) + E(\sum (X_2 - \overline{X}_2)^2)]$$
$$= \frac{1}{n_1 + n_2 - 2} [(n_1 - 1) E(\frac{\sum (X_1 - \overline{X}_1)^2}{n_1 - 1}) + (n_2 - 1) E(\frac{\sum (X_2 - \overline{X}_2)^2}{n_2 - 1})]$$
$$= \frac{1}{n_1 + n_2 - 2} [(n_1 - 1)\sigma^2 + (n_2 - 1)\sigma^2]$$
$$= \frac{(n_1 - 1) + (n_2 - 1)}{n_1 + n_2 - 2} \sigma^2 = \sigma^2$$

因只要 X_1， X_2 為常態分配，且其變異數均為 σ^2

$$E(\frac{\sum(X_1-\overline{X}_1)^2}{n_1-1})=\sigma^2,\quad E(\frac{\sum(X_2-\overline{X}_2)^2}{n_2-1})=\sigma^2$$

三 設 $X_1\sim(\mu_1,\sigma_1^2)$， $X_2\sim(\mu_2,\sigma_2^2)$，且 X_1 與 X_2 獨立，自 X_1 與 X_2 中獨立隨機抽取二個大樣本($n_1\ge 30, n_2\ge 30$)則 $\mu_1-\mu_2$ 的信賴區間為 $(\overline{X}_1-\overline{X}_2)\pm Z_{\alpha/2}\cdot S_{\overline{X}_1-\overline{X}_2}$，

其中

$$S_{\overline{X}_1-\overline{X}_2}=\sqrt{(S_1^2/n_1)+(S_2^2/n_2)}$$

證明 根據中央極限定理可得

$$\overline{X}_1\sim N(\mu_1,\sigma_1^2/n_1),\quad \overline{X}_2\sim N(\mu_2,\sigma_2^2/n_2)$$

再依據常態分配加法定理可得

$$\overline{X}_1-\overline{X}_2\sim N(\mu_1-\mu_2,\sigma_1^2/n_1+\sigma_2^2/n_2)$$

或

$$\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}\sim N(0,1)$$

但因 σ_1^2， σ_2^2 未知，以 S_1^2， S_2^2 分別估計 σ_1^2 與 σ_2^2，在大樣本的情況下， S_1^2 與 S_2^2 均為 σ_1^2， σ_2^2 的一性估計式，因此下式趨近標準常態分配

$$\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}}\sim N(0,1)$$

因此可利用 Z 分配求得 $\mu_1-\mu_2$ 的信賴區間：

$(\overline{X}_1-\overline{X}_2)\pm Z_{\alpha/2}\cdot S_{\overline{X}_1-\overline{X}_2}$

<四> 設 $X\sim N(\mu_X,\sigma_X^2)$， $Y\sim N(\mu_Y,\sigma_Y^2)$， X 、 Y 獨立，且 $\sigma_X^2=\sigma_Y^2=\sigma^2$。自 X 、 Y 中隨機抽取二個獨立樣本，其樣本數分別為 n_X 與 n_Y。令 $\overline{X}$ 、 S_X^2， $\overline{Y}$ 、 S_Y^2 分別為二個樣本之平均數與變異數，則

$$\frac{(\overline{X}-\overline{Y})-(\mu_X-\mu_Y)}{S_P\sqrt{\frac{1}{n_X}+\frac{1}{n_Y}}}\sim t_{n_X+n_Y-2} \quad ①$$

其中 $S_p=(\frac{(n_x-1)S_X^2+(n_Y-1)S_Y^2}{n_X+n_Y-2})^{\frac{1}{2}}$

證明 ①式的左邊為：

$$\frac{(\overline{X}-\overline{Y})-(\mu_X-\mu_Y)}{S_p\sqrt{\frac{1}{n_X}+\frac{1}{n_Y}}}$$

$$=\frac{(\overline{X}-\overline{Y})-(\mu_X-\mu_Y)/\sigma\sqrt{\frac{1}{n_X}+\frac{1}{n_Y}}}{S_p/\sigma}\text{(分子分母同除以}\sigma^2\text{)}$$

(因 $\dfrac{(\overline{X}-\overline{Y})-(\mu_X-\mu_Y)}{\sigma\sqrt{\frac{1}{n_X}+\frac{1}{n_Y}}}\sim N(0,1)$)

(分母部份上下乘以(n_1+n_2-2)，並開根號)

$$=\frac{Z}{\sqrt{\frac{(n_X+n_Y-2)S_p^2}{\sigma^2(n_X+n_Y-2)}}}=\frac{Z}{\sqrt{\frac{\chi^2_{n_X+n_Y-2}}{n_X+n_Y-2}}}\sim t_{n_X+n_Y-2}$$

因為 $\dfrac{(n_X+n_Y-2)S_p^2}{\sigma_X^2}\sim\chi^2_{n_X+n_Y-2}$

証明請參閱下面

$$\frac{(n_X+n_Y-2)S_p^2}{\sigma^2}=\frac{(n_X-1)S_X^2+(n_Y-1)S_Y^2}{\sigma^2}$$

$$=\frac{(n_X-1)S_X^2}{\sigma^2}+\frac{(n_Y-1)S_Y^2}{\sigma^2}$$

$$=\chi^2_{n_X-1}+\chi^2_{n_Y-1}\text{（根據第11章附錄）}$$

$$=\chi^2_{n_X+n_Y-2}\text{（根據卡方分配的加法定理）}$$

五 設 $X\sim(\mu_X,\sigma_X^2)$，$Y\sim(\mu_Y,\sigma_Y^2)$，X、Y 獨立。自 X、Y 中隨機抽取 n_X 與 n_Y 為二個樣本，令 S_X^2 代表第一個樣本之變異數；S_Y^2 為第二個樣本之變異數，則

$$\frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2}\sim F_{n_X-1,n_Y-1}$$

證明　根據第10章數學附錄可知：

$$\frac{(n_X-1)S_X^2}{\sigma_X^2}\sim\chi^2_{n_X-1}\text{，}\quad\frac{(n_Y-1)S_Y^2}{\sigma_Y^2}\sim\chi^2_{n_Y-1}$$

上兩式相除可證得：

$$\frac{(n_X-1)S_X^2/\sigma_X^2(n_X-1)}{(n_Y-1)S_Y^2/\sigma_Y^2(n_Y-1)}=\frac{\chi^2_{n_x-1}/n_x-1}{\chi^2_{n_y-1}/n_y-1}=\frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2}\sim F_{n_X-1,n_Y-1}$$

六 根據 F 分配的倒數性質，求左尾 α 機率的 F 值為：

$$F_{v_1,v_2,1-\alpha}=\frac{1}{F_{v_2,v_1,\alpha}}$$

證明　左尾α機率的F值表為$F_{v_1,v_2,1-\alpha}$，可表為：

$$P(F_{v_1,v_2}\le F_{v_1,v_2,1-\alpha})=\alpha$$

取倒數可得$P(\frac{1}{F_{v_1,v_2}}\ge\frac{1}{F_{v_1,v_2,1-\alpha}})=\alpha$　①

因自由度v_2,v_1的右尾α機率的F值可表為$F_{v_2,v_1,\alpha}$，因此①式可表為

$P(F_{v_2,v_1}\ge F_{v_2,v_1,\alpha})=\alpha$　②

比較①②可知：

$$\frac{1}{F_{v_1,v_2,1-\alpha}}=F_{v_2,v_1,\alpha}，\quad F_{v_1,v_2,1-\alpha}=\frac{1}{F_{v_2,v_1,\alpha}}$$