第十章 兩個母體比較的統計推論

10.1 兩個母體平均數差的統計推論：獨立隨機樣本

定義：

假如 $X_1, X_2, \dots, X_{n_1} \overset{iid}{\sim} N(\mu_1, {\sigma_1}^2)$

$Y_1, Y_2, \dots, Y_{n_2} \overset{iid}{\sim} N(\mu_2, {\sigma_2}^2)$

且 $\bar{X}=\frac{\sum_{i=1}^{n_1}X_i}{n_1}$，$S_1{}^2=\frac{\sum_{i=1}^{n_1}(X_i-\bar{X})^2}{n_1-1}$

$\bar{Y}=\frac{\sum_{i=1}^{n_2}Y_i}{n_2}$，$S_2{}^2=\frac{\sum_{i=1}^{n_2}(Y_i-\bar{Y})^2}{n_2-1}$

如果兩常態母體獨立，則兩組樣本亦互相獨立。

- 兩母體平均數差的信賴區間

假如 $X_1, X_2, \ldots, X_{n_1} \overset{iid}{\sim} N(\mu_1, \sigma_1{}^2)$

$Y_1, Y_2, \ldots, Y_{n_2} \overset{iid}{\sim} N(\mu_2, \sigma_2{}^2)$

且兩常態母體互相獨立

求 $\mu_1 - \mu_2$ 之 $100(1-\alpha)\%$ C.I.

Case1 $\sigma_1{}^2, \sigma_2{}^2$ 已知

Sol： $\bar{X} \sim N(\mu_1, \frac{\sigma_1{}^2}{n_1})$

$\bar{Y} \sim N(\mu_2, \frac{\sigma_2{}^2}{n_2})$

則 $\bar{X} - \bar{Y} \sim N(\mu_1 - \mu_2, \frac{\sigma_1{}^2}{n_1} + \frac{\sigma_2{}^2}{n_2})$

$$\therefore \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1{}^2}{n_1} + \frac{\sigma_2{}^2}{n_2}}} \sim \mathrm{N}(0,1)$$

$$1-\alpha = P\left(-Z_{\alpha/2} < \frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}} < Z_{\alpha/2}\right)$$

$$= P\left(-Z_{\alpha/2}\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}} < (\bar{X}-\bar{Y}) - (\mu_1-\mu_2) < Z_{\alpha/2}\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}\right)$$

$$= P\left(-(\bar{X}-\bar{Y}) - Z_{\alpha/2}\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}} < -(\mu_1-\mu_2) < -(\bar{X}-\bar{Y}) + Z_{\alpha/2}\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}\right)$$

$$=P\left((\bar{X}-\bar{Y})-Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}<\mu_1-\mu_2\right.$$

$$\left.<(\bar{X}-\bar{Y})+Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$$

$$\left[(\bar{x}-\bar{y})-Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}},\ (\bar{x}-\bar{y})+Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right]$$

為$\mu_1-\mu_2$之100(1−α)% C. I.

Case2 $\sigma_1{}^2, \sigma_2{}^2$ 未知, $n_1 \geq 30$, $n_2 \geq 30$(大樣本)

Sol：$\sigma_1{}^2$ 未知由 $S_1{}^2$ 估計之

$\sigma_2{}^2$ 未知由 $S_2{}^2$ 估計之

$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{S_1{}^2}{n_1}+\frac{S_2{}^2}{n_2}}} \underset{\text{近似}}{\sim} \mathrm{N}(0,1)$$

$$1-\alpha \approx P\left(-Z_{\alpha/2} < \frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{S_1{}^2}{n_1}+\frac{S_2{}^2}{n_2}}} < Z_{\alpha/2}\right)$$

$$=P\left((\bar{X}-\bar{Y}) - Z_{\alpha/2}\sqrt{\frac{S_1{}^2}{n_1}+\frac{S_2{}^2}{n_2}} < \mu_1-\mu_2\right.$$

$$\left. < (\bar{X}-\bar{Y}) + Z_{\alpha/2}\sqrt{\frac{S_1{}^2}{n_1}+\frac{S_2{}^2}{n_2}}\right)$$

$$\left[(\bar{x}-\bar{y}) - Z_{\alpha/2}\sqrt{\frac{S_1{}^2}{n_1}+\frac{S_2{}^2}{n_2}},\ (\bar{x}-\bar{y}) + Z_{\alpha/2}\sqrt{\frac{S_1{}^2}{n_1}+\frac{S_2{}^2}{n_2}}\right]$$

為$\mu_1-\mu_2$之近似100(1−α)% C. I.

Case3 ${\sigma_1}^2, {\sigma_2}^2$未知但相等, $n_1 < 30$, $n_2 < 30$(小樣本)

Sol：令$\sigma^2 = {\sigma_1}^2 = {\sigma_2}^2$

且$Sp^2 = \dfrac{(n_1-1){S_1}^2+(n_2-1){S_2}^2}{n_1+n_2-2}$

已知$\dfrac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma^2}{n_1}+\dfrac{\sigma^2}{n_2}}} \sim N(0,1)$

$$\frac{(n_1-1)S_1{}^2+(n_2-1)S_2{}^2}{\sigma^2} \sim x^2(n_1+n_2-2)$$

$$且\ T=\frac{\dfrac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\sigma^2(\frac{1}{n_1}+\frac{1}{n_2})}}}{\sqrt{\dfrac{(n_1-1)S_1{}^2+(n_2-1)S_2{}^2}{\sigma^2}\Big/_{(n_1+n_2-2)}}}$$

$$=\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{Sp^2(\frac{1}{n_1}+\frac{1}{n_2})}}$$

$$=\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{Sp\sqrt{(\frac{1}{n_1}+\frac{1}{n_2})}} \sim t(n_1+n_2-2)$$

$$1-\alpha = P\left(-t_{\alpha/2}(n_1+n_2-2) < \frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{Sp\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} < t_{\alpha/2}(n_1+n_2-2)\right)$$

$$= P\left((\bar{X}-\bar{Y}) - t_{\alpha/2}(n_1+n_2-2)\, Sp\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} < \mu_1-\mu_2 < (\bar{X}-\bar{Y}) + t_{\alpha/2}(n_1+n_2-2)\, Sp\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$$

$$\left[\left((\bar{X}-\bar{Y}) - t_{\alpha/2}(n_1+n_2-2)\, Sp\sqrt{\frac{1}{n_1}+\frac{1}{n_2}},\ (\bar{X}-\bar{Y}) + t_{\alpha/2}(n_1+n_2-2)\, Sp\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right]\right.$$

為$\mu_1-\mu_2$之100(1-α)% C. I.

上述所求之信賴區間可進一步獲得如下資訊：

如果[+,+]表示$\mu_1 > \mu_2$

[-,-]表示$\mu_1 < \mu_2$

[-,+]表示$\mu_1 = \mu_2$

- 兩個母體平均數差的假設檢定

假設$X_1, X_2, \ldots, X_{n_1} \overset{\text{iid}}{\sim} \mathrm{N}(\mu_1, \sigma_1{}^2)$

$$Y_1, Y_2, \ldots, Y_{n_2} \overset{\text{iid}}{\sim} \mathrm{N}(\mu_2, \sigma_2{}^2)$$

且兩常態母體獨立，給定顯著水準 α ，求下列檢定

H_0 ： $\mu_1 = \mu_2$

H_1 ： $\mu_1 \neq \mu_2$

Case1 $\sigma_1{}^2, \sigma_2{}^2$已知

Sol：上述假設檢定等同

H_0：$\mu_1 - \mu_2 = 0$

H_1：$\mu_1 - \mu_2 \neq 0$

$\overline{X}\text{-}\overline{Y} \xrightarrow[\text{推論}]{} \mu_1 - \mu_2$

已知$\overline{X}-\overline{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1{}^2}{n_1} + \frac{\sigma_2{}^2}{n_2}\right)$

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

在H_0為真之下

拒絕H_0 if $\bar{X}-\bar{Y} > C_2$ or $\bar{X}-\bar{Y} < C_1$

α=P(拒絕$H_0|H_0$為真)

$=P(\bar{X}-\bar{Y} > C_2 \text{ or } \bar{X}-\bar{Y} < C_1 | \mu_1 - \mu_2 = 0)$

$\frac{\alpha}{2}=P(\bar{X}-\bar{Y} > C_2 | \mu_1 - \mu_2 = 0)$

$$=P\left(\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{{\sigma_1}^2}{n_1}+\frac{{\sigma_2}^2}{n_2}}} > \frac{C_2-0}{\sqrt{\frac{{\sigma_1}^2}{n_1}+\frac{{\sigma_2}^2}{n_2}}}\right)$$

$$= P(Z > Z_{\alpha/2})$$

$$\therefore \frac{C_2}{\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}} = Z_{\alpha/2} \Rightarrow C_2 = Z_{\alpha/2}\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}$$

$$\frac{\alpha}{2}=\mathrm{P}(\bar{X}-\bar{Y} < C_1 | \mu_1 - \mu_2=0)$$

$$=\mathrm{P}\left(\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}} < \frac{C_1-0}{\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}}\right)$$

$$= P(Z < -Z_{\alpha/2})$$

$$\therefore \frac{C_1}{\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}} = -Z_{\alpha/2} \Rightarrow C_1 = -Z_{\alpha/2}\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}$$

決策法則：

$$\text{拒絕}H_0 \text{ if } \bar{x}-\bar{y} > Z_{\alpha/2}\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}$$

$$\text{或}\ \bar{x}-\bar{y} < -Z_{\alpha/2}\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}$$

改寫成Z檢定統計量

$$拒絕H_0 \text{ if } \frac{\bar{x}-\bar{y}}{\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}} > Z_{\alpha/2}$$

$$或 \frac{\bar{x}-\bar{y}}{\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}} < -Z_{\alpha/2}$$

⇨拒絕H_0 if $z > Z_{\alpha/2}$ 或 $z < -Z_{\alpha/2}$

⇨拒絕H_0 if $|z| > Z_{\alpha/2}$ $z = \frac{\bar{x}-\bar{y}}{\sqrt{\frac{\sigma_1{}^2}{n_1}+\frac{\sigma_2{}^2}{n_2}}}$

檢定型態 \ 拒絕域 \ 分配條件	獨立常態母體			獨立一般母體
	①σ_1^2,σ_2^2已知	②σ_1^2,σ_2^2未知，$n_1\geq30,\ n_2\geq30$	③$\sigma_1^2=\sigma_2^2$但未知，$n_1<30,\ n_2<30$	④σ_1^2,σ_2^2未知，$n_1\geq30,\ n_2\geq30$
$H_0:\mu_1=\mu_2$ $H_1:\mu_1\neq\mu_2$ (雙尾)	$\lvert z\rvert > Z_{\alpha/2}$	$\lvert z\rvert > Z_{\alpha/2}$	$\lvert t\rvert > t_{\alpha/2}(n_1+n_2-2)$	同②
$H_0:\mu_1\leq\mu_2$ $H_1:\mu_1>\mu_2$ (右尾)	$z > Z_\alpha$	$z > Z_\alpha$	$t > t_\alpha(n_1+n_2-2)$	同②
$H_0:\mu_1\geq\mu_2$ $H_1:\mu_1<\mu_2$ (左尾)	$z < -Z_\alpha$ $Z=\dfrac{\bar{x}-\bar{y}}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$	$z < -Z_\alpha$ $Z=\dfrac{\bar{x}-\bar{y}}{\sqrt{\dfrac{S_1^2}{n_1}+\dfrac{S_2^2}{n_2}}}$	$t < -t_\alpha(n_1+n_2-2)$ $t=\dfrac{\bar{x}-\bar{y}}{Sp\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$ $Sp^2=\dfrac{(n_1-1)S_1^2+(n_2-1)S_1^2}{n_1+n_2-2}$	同②

10.2成對樣本之兩個母體平均數差的統計推論

- 成對樣本之平均數差的信賴區間

假設$X_{11}, X_{12}, \ldots, X_{1n}$為n人進行某一實驗前之數據；

$X_{21}, X_{22}, \ldots, X_{2n}$為n人進行某一實驗後之數據，

則$(X_{11}, X_{21}), (X_{12}, X_{22}), \ldots, (X_{1n}, X_{2n})$為n對相關測量值。

令D_i=$X_{1i}-X_{2i}$ i=1,2,…,n (n<30)

且假設$D_1, D_2, \ldots, D_n \overset{iid}{\sim} N(\mu_D, \sigma_D{}^2)$

求μ_D之$100(1-\alpha)\%$ C. I.

Sol：令$\mu_D=\mu_1-\mu_2$且n<30

則$\dfrac{\bar{D}-\mu_D}{S_D/\sqrt{n}}$~t(n-1)

$$\therefore 1-\alpha = P\left(-t_{\alpha/2}(n-1) < \frac{\bar{D}-\mu_D}{S_D/\sqrt{n}} < t_{\alpha/2}(n-1)\right)$$

$$= P\left(\bar{D} - t_{\alpha/2}(n-1)\frac{S_D}{\sqrt{n}} < \mu_D < \bar{D} + t_{\alpha/2}(n-1)\frac{S_D}{\sqrt{n}}\right)$$

$$\Rightarrow \left[\bar{d} - t_{\alpha/2}(n-1)\frac{s_d}{\sqrt{n}}, \bar{d} + t_{\alpha/2}(n-1)\frac{s_d}{\sqrt{n}}\right]$$

為 μ_D 之 $100(1-\alpha)\%$ C.I.

由於$\mu_D = \mu_1 - \mu_2$

上述所得信賴區間提供如下資訊

$[+,+]$表示$\mu_1 > \mu_2$

$[-,-]$表示$\mu_1 < \mu_2$

$[-,+]$表示$\mu_1 = \mu_2$

- 成對樣本之平均數的假設檢定

同前述討論信賴區間的條件，給定顯著水準 α，求下列假設檢定

H_0：$\mu_D = 0$

H_1：$\mu_D \neq 0$

Sol：$\bar{D} \xrightarrow[\text{推論}]{} \mu_D$

在H_0為真之下

拒絕H_0 if $\bar{D} > C_2$ or $\bar{D} < C_1$

α=P(拒絕$H_0|H_0$為真)

$=P(\bar{D} > C_2$ or $\bar{D} < C_1|\mu_D=0)$

$\frac{\alpha}{2}=P(\bar{D} > C_2|\mu_D=0)$

$=P(\frac{\bar{D}-\mu_D}{S_D/\sqrt{n}} > \frac{C_2-0}{S_D/\sqrt{n}})$

$= P(T > t_{\alpha/2}(n-1))$

$\therefore \frac{C_2-0}{S_D/\sqrt{n}} = t_{\alpha/2}(n-1) \Rightarrow C_2 = t_{\alpha/2}(n-1)\frac{S_D}{\sqrt{n}}$

$$\frac{\alpha}{2} = \mathrm{P}(\bar{D} < C_1 | \mu_D = 0)$$

$$= \mathrm{P}\left(\frac{\bar{D} - \mu_D}{S_D/\sqrt{n}} < \frac{C_1 - 0}{S_D/\sqrt{n}}\right)$$

$$= P(T < -t_{\alpha/2}(n-1))$$

$$\therefore \frac{C_1 - 0}{S_D/\sqrt{n}} = -t_{\alpha/2}(n-1)$$

$$\Rightarrow C_1 = -t_{\alpha/2}(n-1)\frac{S_D}{\sqrt{n}}$$

決策法則

拒絕 H_0 if $\bar{d} > t_{\alpha/2}(n-1)\frac{S_d}{\sqrt{n}}$

or $\bar{d} < -t_{\alpha/2}(n-1)\frac{S_d}{\sqrt{n}}$

改寫成t檢定統計量

拒絕 H_0 if $\frac{\bar{d}}{S_d/\sqrt{n}} > t_{\alpha/2}(n-1)$

or $\frac{\bar{d}}{S_d/\sqrt{n}} < -t_{\alpha/2}(n-1)$

⇨拒絕 H_0 if $|t| > t_{\alpha/2}(n-1)$ $\quad t = \frac{\bar{d}}{S_d/\sqrt{n}}$

分配條件 / 拒絕域 / 檢定型態	成對樣本
$H_0: \mu_0 = 0$ $H_1: \mu_0 \neq 0$ （雙尾）	$\lvert t \rvert > t_{\alpha/2}(n-1)$
$H_0: \mu_0 \leq 0$ $H_1: \mu_0 > 0$ （右尾）	$t > t_{\alpha}(n-1)$
$H_0: \mu_0 \geq 0$ $H_1: \mu_0 < 0$ （左尾）	$t < -t_{\alpha}(n-1)$ $t = \dfrac{\bar{d}}{s_d/\sqrt{n}}$

10.3兩個母體比例差 $P_1 - P_2$ 之統計推論

- 兩個樣本比例差 $\hat{P}_1 - \hat{P}_2$ 之抽樣分配

如同9.3節的條件及討論，可獲得

$$\hat{P}_1 = \frac{Y_1}{n_1} \overset{\text{近似}}{\sim} N\left(P_1, \frac{P_1(1-P_1)}{n_1}\right)$$

且 $$\hat{P}_2 = \frac{Y_2}{n_2} \overset{\text{近似}}{\sim} N\left(P_2, \frac{P_2(1-P_2)}{n_2}\right)$$

如果兩母體獨立，則 $\hat{P}_1$ 與 $\hat{P}_2$ 獨立

且 $$\hat{P}_1 - \hat{P}_2 \overset{\text{近似}}{\sim} N\left(P_1 - P_2, \frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}\right)$$

$$\therefore \frac{(\hat{P}_1 - \hat{P}_2) - (P_1 - P_2)}{\sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}} \overset{\text{近似}}{\sim} N(0,1)$$

- 兩個母體比例差 $P_1 - P_2$ 之區間估計

給定信賴水準 $1-\alpha$，求 $P_1 - P_2$ 之信賴區間

$$1-\alpha \approx P\left(-Z_{\alpha/2} < \frac{(\hat{P}_1-\hat{P}_2)-(P_1-P_2)}{\sqrt{\frac{P_1(1-P_1)}{n_1}+\frac{P_2(1-P_2)}{n_2}}} < Z_{\alpha/2}\right)$$

$$1-\alpha \approx P\left(-Z_{\alpha/2}\sqrt{\frac{P_1(1-P_1)}{n_1}+\frac{P_2(1-P_2)}{n_2}} < (\hat{P}_1-\hat{P}_2)-(P_1-P_2) < Z_{\alpha/2}\sqrt{\frac{P_1(1-P_1)}{n_1}+\frac{P_2(1-P_2)}{n_2}}\right)$$

$$1-\alpha \approx P\left(-(\hat{P}_1-\hat{P}_2)-Z_{\alpha/2}\sqrt{\frac{P_1(1-P_1)}{n_1}+\frac{P_2(1-P_2)}{n_2}} < -(P_1-P_2) < -(\hat{P}_1-\hat{P}_2)+Z_{\alpha/2}\sqrt{\frac{P_1(1-P_1)}{n_1}+\frac{P_2(1-P_2)}{n_2}}\right)$$

$$1-\alpha \approx P\left(\widehat{P_1}-\widehat{P_2}-Z_{\alpha/2}\sqrt{\frac{P_1(1-P_1)}{n_1}+\frac{P_2(1-P_2)}{n_2}} < P_1-P_2 < \widehat{P_1}-\widehat{P_2}+Z_{\alpha/2}\sqrt{\frac{P_1(1-P_1)}{n_1}+\frac{P_2(1-P_2)}{n_2}}\right)$$

$\widehat{P_1}$ 估計 P_1，$\widehat{P_2}$ 估計 P_2

則 $$\left[(\widehat{p_1}-\widehat{p_2})-Z_{\alpha/2}\sqrt{\frac{\widehat{p_1}(1-\widehat{p_1})}{n_1}+\frac{\widehat{p_2}(1-\widehat{p_2})}{n_2}},\ (\widehat{P_1}-\widehat{P_2})+Z_{\alpha/2}\sqrt{\frac{\widehat{p_1}(1-\widehat{p_1})}{n_1}+\frac{\widehat{p_2}(1-\widehat{p_2})}{n_2}}\right]$$

為 P_1-P_2 之近似100(1-α)%信賴區間

上述的信賴區間可提供如下的資訊

$[+,+] \Rightarrow P_1 > P_2$

$[-,-] \Rightarrow P_1 < P_2$

$[+,+] \Rightarrow P_1 = P_2$

- 兩個母體比例差P_1-P_2之假設檢定

給定顯著水準α，求下列假設檢定

H_0：$P_1=P_2$

H_1：$P_1\neq P_2$

Sol：上述假設檢定可以寫為

H_0：$P_1-P_2=0$

H_1：$P_1-P_2\neq 0$

$$\widehat{P_1}-\widehat{P_2}\xrightarrow{\text{推論}}P_1-P_2$$

已知$\widehat{P_1}=\dfrac{Y_1}{n_1}$, $\widehat{P_2}=\dfrac{Y_2}{n_2}$

且$\widehat{P_1}-\widehat{P_2}\ \widetilde{\text{近似}}\ \mathrm{N}(\mathrm{P}_1-\mathrm{P}_2,\frac{P_1(1-P_1)}{n_1}+\frac{P_2(1-P_2)}{n_2})$

在H_0為真之下$(P_1=P_2)$

則$\widehat{P_1}-\widehat{P_2}\ \widetilde{\text{近似}}\ \mathrm{N}(0,\hat{P}(1-\hat{P})(\frac{1}{n_1}+\frac{1}{n_2}))$

其中$\hat{P}=\frac{Y_1+Y_2}{n_1+n_2}$

即$\frac{\widehat{P_1}-\widehat{P_2}}{\sqrt{\hat{P}(1-\hat{P})(\frac{1}{n_1}+\frac{1}{n_2})}}\ \widetilde{\text{近似}}\ \mathrm{N}(0,1)$

在H_0為真之下，拒絕H_0 if

$\widehat{P_1}-\widehat{P_2}>C_2$ or $\widehat{P_1}-\widehat{P_2}<C_1$

$\alpha = P($拒絕$H_0 \mid H_0$為真$)$

$= P(\widehat{P_1}-\widehat{P_2}>C_2 \text{ or } \widehat{P_1}-\widehat{P_2}<C_1 \mid P_1 = P_2)$

$$\frac{\alpha}{2} = P\left(\widehat{P_1}-\widehat{P_2}>C_2 \mid P_1 = P_2\right)$$

$$= P\left(\frac{\widehat{P_1}-\widehat{P_2}}{\sqrt{\hat{P}(1-\hat{P})(\frac{1}{n_1}+\frac{1}{n_2})}} > \frac{C_2}{\sqrt{\hat{P}(1-\hat{P})(\frac{1}{n_1}+\frac{1}{n_2})}}\right)$$

$$\approx P(Z > Z_{\alpha/2})$$

$$\therefore \frac{C_2}{\sqrt{\hat{P}\,(1-\hat{P})(\frac{1}{n_1}+\frac{1}{n_2})}} = Z_{\alpha/2}$$

$$\Rightarrow C_2 = Z_{\alpha/2}\sqrt{\hat{P}\,(1-\hat{P})(\frac{1}{n_1}+\frac{1}{n_2})}$$

$$\frac{\alpha}{2} = P\left(\widehat{P_1} - \widehat{P_2} < C_1 \mid P_1 = P_2\right)$$

$$= P\left(\frac{\widehat{P_1}-\widehat{P_2}}{\sqrt{\hat{P}\,(1-\hat{P})(\frac{1}{n_1}+\frac{1}{n_2})}} < \frac{C_1}{\sqrt{\hat{P}\,(1-\hat{P})(\frac{1}{n_1}+\frac{1}{n_2})}}\right)$$

$$\approx P\left(Z < -Z_{\alpha/2}\right)$$

$$C_1 = -Z_{\alpha/2}\sqrt{\hat{P}\,(1-\hat{P})(\frac{1}{n_1}+\frac{1}{n_2})}$$

結論：拒絕H_0 if

$$\widehat{P_1} - \widehat{P_2} > Z_{\alpha/2}\sqrt{\hat{P}\,(1-\hat{P})(\frac{1}{n_1}+\frac{1}{n_2})} \text{ or}$$

$$\widehat{P_1} - \widehat{P_2} < -Z_{\alpha/2}\sqrt{\hat{P}\,(1-\hat{P})(\frac{1}{n_1}+\frac{1}{n_2})}$$

上述可改寫成標準統計量檢定

拒絕H_0 if $z > Z_{\alpha/2}$ or $z < -Z_{\alpha/2}$

即拒絕H_0 if $|z| > Z_{\alpha/2}$, $z = \frac{\widehat{P_1} - \widehat{P_2}}{\sqrt{\hat{P}\,(1-\hat{P})(\frac{1}{n_1}+\frac{1}{n_2})}}$

檢定型態 \ 拒絕域 \ 檢定法	標準統計量檢定法
$H_0: P_1 = P_2$ $H_1: P_1 \neq P_2$ （雙尾）	$\lvert z \rvert \geq Z_{\alpha/2}$
$H_0: P_1 \leq P_2$ $H_1: P_1 > P_2$ （右尾）	$z \geq Z_{\alpha}$
$H_0: P_1 \geq P_2$ $H_1: P_1 < P_2$ （左尾）	$z \leq -Z_{\alpha}$ $\quad \widehat{P_1}=\frac{Y_1}{n_1},\ \widehat{P_2}=\frac{Y_2}{n_2}$ $z=\frac{\widehat{P_1}-\widehat{P_2}}{\sqrt{\hat{P}(1-\hat{P})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}},\ \hat{P}=\frac{Y_1+Y_2}{n_1+n_2}$

10.4兩個獨立常態母體變異數的統計推論

- 兩個獨立常態樣本變異數之抽樣分配假設

給定如下基本假設

1.$X_1, X_2, \ldots, X_{n_1} \overset{iid}{\sim} N(\mu_1, {\sigma_1}^2)$

2.$Y_1, Y_2, \ldots, Y_{n_2} \overset{iid}{\sim} N(\mu_2, {\sigma_2}^2)$

3.兩個常態母體獨立

由第7章的討論我們獲得如下抽樣分配

a. $\frac{(n_1-1){s_1}^2}{{\sigma_1}^2} \sim x^2(n_1-1), {s_1}^2 = \frac{\sum_{i=1}^{n_1}(X_i-\bar{X})^2}{n_1-1}$

b. $\frac{(n_2-1){s_2}^2}{{\sigma_2}^2} \sim x^2(n_2-1), {s_2}^2 = \frac{\sum_{i=1}^{n_2}(Y_i-\bar{Y})^2}{n_2-1}$

c. $F = \frac{\frac{(n_1-1){s_1}^2}{{\sigma_1}^2} \Big/ (n_1-1)}{\frac{(n_2-1){s_2}^2}{{\sigma_2}^2} \Big/ n_2-1} = \frac{{s_1}^2 / {\sigma_1}^2}{{s_2}^2 / {\sigma_2}^2}$

$= \frac{{s_1}^2}{{s_2}^2} \frac{{\sigma_2}^2}{{\sigma_1}^2} \sim F(n_1-1, n_2-1)$

- 兩個常態母體變異數比σ_1^2/σ_2^2之區間估計

給定信賴水準$1-\alpha$，求σ_1^2/σ_2^2之信賴區間

$$1-\alpha = P\left(F_{1-\alpha/2}(n_1-1, n_2-1) < \frac{S_1^2}{S_2^2}\frac{\sigma_2^2}{\sigma_1^2} < F_{\alpha/2}(n_1-1, n_2-1)\right)$$

$$= P\left(\frac{S_2^2}{S_1^2}F_{1-\alpha/2}(n_1-1, n_2-1) < \frac{\sigma_2^2}{\sigma_1^2} < \frac{S_2^2}{S_1^2}F_{\alpha/2}(n_1-1, n_2-1)\right)$$

$$= P\left(\frac{S_1^2}{S_2^2F_{\alpha/2}(n_1-1, n_2-1)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2F_{1-\alpha/2}(n_1-1, n_2-1)}\right)$$

$$= P\left(\frac{S_1^2}{S_2^2F_{\alpha/2}(n_1-1, n_2-1)} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2}F_{\alpha/2}(n_2-1, n_1-1)\right)$$

$\Rightarrow \left[\frac{s_1^2}{s_2^2 F_{\alpha/2}(n_1-1, n_2-1)}, \frac{s_1^2}{s_2^2} F_{\alpha/2}(n_2-1, n_1-1)\right]$

為 σ_1^2/σ_2^2 之100(1-α)% C. I.

上述信賴區間可提供如下訊息

[大於1, 大於1] $\Rightarrow$ $\sigma_1^2 > \sigma_2^2$

[小於1, 小於1] $\Rightarrow$ $\sigma_1^2 < \sigma_2^2$

[小於1, 大於1] $\Rightarrow$ $\sigma_1^2 = \sigma_2^2$

• 兩個常態母體變異數比σ_1^2/σ_2^2之假設檢定

給定信賴水準α，求下列假設檢定

H_0：$\sigma_1^2 = \sigma_2^2$

H_1：$\sigma_1^2 \neq \sigma_2^2$

Sol：上述假設檢定可改寫為

H_0：$\sigma_1^2/\sigma_2^2 = 1$

H_1：$\sigma_1^2/\sigma_2^2 \neq 1$

$$\frac{s_1^2}{s_2^2} \xrightarrow{\text{推論}} \frac{\sigma_1^2}{\sigma_2^2}$$

已知 $\frac{s_1^2}{s_2^2}\frac{\sigma_2^2}{\sigma_1^2} \sim F(n_1-1, n_2-1)$

在 H_0 為真之下 $\left(\sigma_1^2 / \sigma_2^2 = 1\right)$

$$\frac{s_1^2}{s_2^2} \sim F(n_1-1, n_2-1)$$

拒絕 H_0 if $\frac{s_1^2}{s_2^2} > c_2$ or $\frac{s_1^2}{s_2^2} < c_1$

$\alpha = P($拒絕 $H_0 \mid H_0$ 為真$)$

$$= P\left(\frac{s_1^2}{s_2^2} > c_2 \text{ or } \frac{s_1^2}{s_2^2} < c_1 \,\middle|\, \sigma_1^2 = \sigma_2^2\right)$$

$$\alpha = P\left(\frac{s_1^2}{s_2^2} > c_2 \;\middle|\; \sigma_1^2 = \sigma_2^2 \right)$$

$$= P\left(F > F_{\alpha/2}(n_1 - 1, n_2 - 1)\right)$$

$$\therefore \; c_2 = F_{\alpha/2}(n_1 - 1, n_2 - 1)$$

$$\alpha = P\left(\frac{s_1^2}{s_2^2} < c_1 \;\middle|\; \sigma_1^2 = \sigma_2^2 \right)$$

$$= P\left(F < F_{1-\alpha/2}(n_1 - 1, n_2 - 1)\right)$$

$$\therefore \; c_1 = F_{1-\alpha/2}(n_1 - 1, n_2 - 1)$$

拒絕 H_0 if $\dfrac{s_1^{\,2}}{s_2^{\,2}} > F_{\alpha/2}(n_1-1, n_2-1)$ or

$$\dfrac{s_1^{\,2}}{s_2^{\,2}} < F_{1-\alpha/2}(n_1-1, n_2-1)$$

拒絕 H_0 if $\dfrac{s_1^{\,2}}{s_2^{\,2}} > F_{\alpha/2}(n_1-1, n_2-1)$ or

$$\dfrac{s_2^{\,2}}{s_1^{\,2}} > F_{\alpha/2}(n_2-1, n_1-1)$$

檢定型態 \ 拒絕域 \ 檢定法	標準統計量檢定法
$H_0：\sigma_1^2 = \sigma_2^2$ $H_1：\sigma_1^2 \neq \sigma_2^2$ （雙尾）	$\frac{s_1^2}{s_2^2} \ge F_{\alpha/2}(n_1 - 1, n_2 - 1)$ or $\frac{s_2^2}{s_1^2} \ge F_{\alpha/2}(n_2 - 1, n_1 - 1)$
$H_0：\sigma_1^2 \le \sigma_2^2$ $H_1：\sigma_1^2 > \sigma_2^2$ （右尾）	$\frac{s_1^2}{s_2^2} \ge F_{\alpha}(n_1 - 1, n_2 - 1)$
$H_0：\sigma_1^2 \ge \sigma_2^2$ $H_1：\sigma_1^2 < \sigma_2^2$ （左尾）	$\frac{S_2^2}{S_1^2} \ge F_{\alpha}(n_2 - 1, n_1 - 1)$