LỚP THẦY THI
TỔNG HỢP CÔNG THỨC THỐNG KÊ CUỐI KỲ HK202

Nội dung 1: Tìm khoảng ước lượng (khoảng tin cậy)

Dạng ước lượng	Độ chính xác	Khoảng ước lượng
1. Tỷ lệ:	$\varepsilon = Z_\alpha \dfrac{\sqrt{f.(1-f)}}{\sqrt{n}}$	$f - \varepsilon < p < f + \varepsilon$
2. Trung bình (kỳ vọng):		
❖ Tuân theo $N(a, \sigma^2)$, chưa biết σ^2 $(n < 30)$	$\varepsilon = t_{\alpha/2}^{(n-1)} \dfrac{s}{\sqrt{n}}$	$\bar{x} - \varepsilon < a < \bar{x} + \varepsilon$
❖ Mẫu lớn $(n \geq 30)$, chưa biết σ^2	$\varepsilon = Z_\alpha \dfrac{s}{\sqrt{n}}$	$\bar{x} - \varepsilon < a < \bar{x} + \varepsilon$
❖ Tuân theo $N(a, \sigma^2)$, đã biết σ^2 (ít thi)	$\varepsilon = Z_\alpha \dfrac{\sigma}{\sqrt{n}}$	$\bar{x} - \varepsilon < a < \bar{x} + \varepsilon$
3. Phương sai:		$\dfrac{(n-1).s^2}{\chi^2_{\frac{\alpha}{2}(n-1)}} < \sigma^2 < \dfrac{(n-1).s^2}{\chi^2_{1-\frac{\alpha}{2}(n-1)}}$

Xác định kích thước mẫu:

Dạng ước lượng	Kích thước mẫu
1. Tỷ lệ: (nếu không biết f dùng CT 2)	$n' = \left[\left(\dfrac{Z_\alpha\sqrt{f.(1-f)}}{\varepsilon'}\right)^2\right] + 1$ hoặc $n' = \left[\left(\dfrac{Z_\alpha}{\varepsilon'}\right)^2 .0{,}25\right] + 1$
2. Trung bình (kỳ vọng):	
❖ Chưa biết σ^2	$n' = \left[\left(Z_\alpha \dfrac{s}{\varepsilon'}\right)^2\right] + 1$
❖ Đã biết σ^2 (ít thi)	$n' = \left[\left(Z_\alpha \dfrac{\sigma}{\varepsilon'}\right)^2\right] + 1$

Nội dung 2: Kiểm định

Dạng 1: Kiểm định tỷ lệ 1 mẫu

1. Phát biểu giả thiết:

Giả thiết	Miền bác bỏ	Bác bỏ H, chấp nhận $\overline{H}$	P - value
$H: p = p_0$ $\overline{H}: p \neq p_0$	$W_\alpha = (-\infty, -Z_\alpha) \cup (Z_\alpha, +\infty)$	$\lvert U_{qs}\rvert > Z_\alpha \leftrightarrow U_{qs} \in W_\alpha$	$p = 2[1 - \Phi(\lvert U_{qs}\rvert)]$
$H: p = p_0$ $\overline{H}: p < p_0$	$W_\alpha = (-\infty, -Z_{2\alpha})$	$U_{qs} < -Z_{2\alpha} \leftrightarrow U_{qs} \in W_\alpha$	$p = \Phi(U_{qs})$
$H: p = p_0$ $\overline{H}: p > p_0$	$W_\alpha = (Z_{2\alpha}, +\infty)$	$U_{qs} > Z_{2\alpha} \leftrightarrow U_{qs} \in W_\alpha$	$p = 1 - \Phi(U_{qs})$

2. Tính giá trị quan sát:

$$U_{qs} = \frac{f - p_0}{\sqrt{p_0(1-p_0)}}\sqrt{n}$$

Dạng 2: Kiểm định tỷ lệ 2 mẫu:

1. Phát biểu giả thiết:

Giả thiết	Miền bác bỏ	Bác bỏ H, chấp nhận $\overline{H}$	P - value
H: $p_1 = p_2$ $\overline{H}$: $p_1 \neq p_2$	$W_\alpha = (-\infty, -Z_\alpha) \cup (Z_\alpha, +\infty)$	$\|U_{qs}\| > Z_\alpha \leftrightarrow U_{qs} \in W_\alpha$	$p = 2[1 - \Phi(\|U_{qs}\|)]$
H: $p_1 = p_2$ $\overline{H}$: $p_1 < p_2$	$W_\alpha = (-\infty, -Z_{2\alpha})$	$U_{qs} < -Z_{2\alpha} \leftrightarrow U_{qs} \in W_\alpha$	$p = \Phi(U_{qs})$
H: $p_1 = p_2$ $\overline{H}$: $p_1 > p_2$	$W_\alpha = (Z_{2\alpha}, +\infty)$	$U_{qs} > Z_{2\alpha} \leftrightarrow U_{qs} \in W_\alpha$	$p = 1 - \Phi(U_{qs})$

2. Tính giá trị quan sát:

$$U_{qs} = \frac{f_1 - f_2}{\sqrt{\dfrac{\overline{p}(1-\overline{p})}{\overline{n}}}}$$

Trong đó: $f_1 = \dfrac{m_1}{n_1}$; $f_2 = \dfrac{m_2}{n_2}$; $\overline{n} = \dfrac{n_1.n_2}{n_1 + n_2}$; $\overline{p} = \dfrac{m_1 + m_2}{n_1 + n_2}$

Dạng 3: Kiểm định trung bình (kỳ vọng) 1 mẫu trường hợp biết phương sai σ^2

1. Phát biểu giả thiết:

Giả thiết	Miền bác bỏ	Bác bỏ H, chấp nhận $\overline{H}$	P - value
H: $a = a_0$ $\overline{H}$: $a \neq a_0$	$W_\alpha = (-\infty, -Z_\alpha) \cup (Z_\alpha, +\infty)$	$\|U_{qs}\| > Z_\alpha \leftrightarrow U_{qs} \in W_\alpha$	$p = 2[1 - \Phi(\|U_{qs}\|)]$
H: $a = a_0$ $\overline{H}$: $a < a_0$	$W_\alpha = (-\infty, -Z_{2\alpha})$	$U_{qs} < -Z_{2\alpha} \leftrightarrow U_{qs} \in W_\alpha$	$p = \Phi(U_{qs})$
H: $a = a_0$ $\overline{H}$: $a > a_0$	$W_\alpha = (Z_{2\alpha}, +\infty)$	$U_{qs} > Z_{2\alpha} \leftrightarrow U_{qs} \in W_\alpha$	$p = 1 - \Phi(U_{qs})$

2. Tính giá trị quan sát:

$$U_{qs} = \frac{\overline{x} - a_0}{\sigma}.\sqrt{n}$$

Dạng 4: Kiểm định trung bình (kỳ vọng) 1 mẫu trường hợp chưa biết phương sai σ^2, mẫu lớn ($n \geq 30$)

1. Phát biểu giả thiết:

Giả thiết	Miền bác bỏ	Bác bỏ H, chấp nhận $\overline{H}$	P - value
H: $a = a_0$ $\overline{H}$: $a \neq a_0$	$W_\alpha = (-\infty, -Z_\alpha) \cup (Z_\alpha, +\infty)$	$\|U_{qs}\| > Z_\alpha \leftrightarrow U_{qs} \in W_\alpha$	$p = 2[1 - \Phi(\|U_{qs}\|)]$
H: $a = a_0$ $\overline{H}$: $a < a_0$	$W_\alpha = (-\infty, -Z_{2\alpha})$	$U_{qs} < -Z_{2\alpha} \leftrightarrow U_{qs} \in W_\alpha$	$p = \Phi(U_{qs})$
H: $a = a_0$ $\overline{H}$: $a > a_0$	$W_\alpha = (Z_{2\alpha}, +\infty)$	$U_{qs} > Z_{2\alpha} \leftrightarrow U_{qs} \in W_\alpha$	$p = 1 - \Phi(U_{qs})$

2. Tính giá trị quan sát:

$$U_{qs} = \frac{\overline{x} - a_0}{s}.\sqrt{n}$$

Dạng 5: Kiểm định trung bình (kỳ vọng) 1 mẫu trường hợp chưa biết phương sai σ^2, mẫu bé ($n < 30$)

1. Phát biểu giả thiết:

Giả thiết	Miền bác bỏ	Bác bỏ H, chấp nhận $\overline{H}$	P - value
$H: a = a_0$ $\overline{H}: a \neq a_0$	$W_\alpha = (-\infty, -t_{\alpha/2}^{(n-1)}) \cup (t_{\alpha/2}^{(n-1)}, +\infty)$	$\lvert T_{qs}\rvert > t_{\alpha/2}^{(n-1)}$ $\leftrightarrow T_{qs} \in W_\alpha$	$p = 2P(T_{n-1} \geq \lvert T_{qs}\rvert)$
$H: a = a_0$ $\overline{H}: a < a_0$	$W_\alpha = (-\infty, -t_\alpha^{(n-1)})$	$T_{qs} < -t_\alpha^{(n-1)}$ $\leftrightarrow T_{qs} \in W_\alpha$	$p = P(T_{n-1} \leq -T_{qs})$
$H: a = a_0$ $\overline{H}: a > a_0$	$W_\alpha = (t_\alpha^{(n-1)}, +\infty)$	$T_{qs} > t_\alpha^{(n-1)}$ $\leftrightarrow T_{qs} \in W_\alpha$	$p = P(T_{n-1} \geq T_{qs})$

2. Tính giá trị quan sát:

$$T_{qs} = \frac{\overline{x} - a_0}{s}.\sqrt{n}$$

Dạng 6: Kiểm định trung bình (kỳ vọng) 2 mẫu độc lập, đã biết phương sai σ_1^2, σ_2^2

1. Phát biểu giả thiết:

Giả thiết	Miền bác bỏ	Bác bỏ H, chấp nhận $\overline{H}$	P - value
$H: a_1 = a_2$ $\overline{H}: a_1 \neq a_2$	$W_\alpha = (-\infty, -Z_\alpha) \cup (Z_\alpha, +\infty)$	$\lvert U_{qs}\rvert > Z_\alpha \leftrightarrow U_{qs} \in W_\alpha$	$p = 2[1 - \Phi(\lvert U_{qs}\rvert)]$
$H: a_1 = a_2$ $\overline{H}: a_1 < a_2$	$W_\alpha = (-\infty, -Z_{2\alpha})$	$U_{qs} < -Z_{2\alpha} \leftrightarrow U_{qs} \in W_\alpha$	$p = \Phi(U_{qs})$
$H: a_1 = a_2$ $\overline{H}: a_1 > a_2$	$W_\alpha = (Z_{2\alpha}, +\infty)$	$U_{qs} > Z_{2\alpha} \leftrightarrow U_{qs} \in W_\alpha$	$p = 1 - \Phi(U_{qs})$

2. Tính giá trị quan sát:

$$U_{qs} = \frac{\overline{x} - \overline{y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Dạng 7: Kiểm định trung bình (kỳ vọng) 2 mẫu lớn độc lập, chưa biết phương sai σ_1^2, σ_2^2, $\sigma_1^2 \neq \sigma_2^2$.

1. Phát biểu giả thiết:

Giả thiết	Miền bác bỏ	Bác bỏ H, chấp nhận $\overline{H}$	P - value
$H: a_1 = a_2$ $\overline{H}: a_1 \neq a_2$	$W_\alpha = (-\infty, -Z_\alpha) \cup (Z_\alpha, +\infty)$	$\lvert U_{qs}\rvert > Z_\alpha \leftrightarrow U_{qs} \in W_\alpha$	$p = 2[1 - \Phi(\lvert U_{qs}\rvert)]$
$H: a_1 = a_2$ $\overline{H}: a_1 < a_2$	$W_\alpha = (-\infty, -Z_{2\alpha})$	$U_{qs} < -Z_{2\alpha} \leftrightarrow U_{qs} \in W_\alpha$	$p = \Phi(U_{qs})$
$H: a_1 = a_2$ $\overline{H}: a_1 > a_2$	$W_\alpha = (Z_{2\alpha}, +\infty)$	$U_{qs} > Z_{2\alpha} \leftrightarrow U_{qs} \in W_\alpha$	$p = 1 - \Phi(U_{qs})$

2. Tính giá trị quan sát:

$$U_{qs} = \frac{\overline{x} - \overline{y}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Dạng 8: Kiểm định phương sai 1 mẫu

1. Phát biểu giả thiết:

Giả thiết	Miền bác bỏ
$H: \sigma^2 = \sigma_0^2$ $\overline{H}: \sigma^2 \neq \sigma_0^2$	$W_\alpha = (0, \chi^2_{1-\alpha/2}{}^{(n-1)}) \cup (\chi^2_{\alpha/2}{}^{(n-1)}, +\infty)$
$H: \sigma^2 = \sigma_0^2$ $\overline{H}: \sigma^2 < \sigma_0^2$	$W_\alpha = (0, \chi^2_{1-\alpha}{}^{(n-1)})$
$H: \sigma^2 = \sigma_0^2$ $\overline{H}: \sigma^2 > \sigma_0^2$	$W_\alpha = (\chi^2_{\alpha}{}^{(n-1)}, +\infty)$

2. Tính giá trị quan sát:

$$\chi^2_{qs} = \frac{(n-1)s^2}{\sigma_0^2}$$

Dạng 9: Phân tích phương sai một nhân tố (cỡ mẫu bằng nhau)

1. Phát biểu giả thiết:

$$H: \mu_1 = \mu_2 = \mu_3 = \ldots = \mu_I$$

Trung bình của tất cả các phương thức xử lý bằng nhau.

$$\overline{H}: \exists \mu_i \neq \mu_j$$

Có ít nhất 2 giá trị trung bình ở các phương thức xử lý khác nhau.

2. Tính các trung bình:

$$\overline{x_1} = \ldots,\ \overline{x_2} = \ldots,\ \overline{x_3} = \ldots,\ \overline{x_I} = \ldots,\ \overline{x} =$$

3. Tính các tổng bình phương (sai số, nghiệm thức, toàn thể):

$$SSE = SS_1 + SS_2 + SS_3 + \ldots SS_I = \sum_{j=1}^{J}(x_{1j} - \overline{x_1})^2 + \sum_{j=1}^{J}(x_{2j} - \overline{x_2})^2 + \ldots + \sum_{j=1}^{J}(x_{Ij} - \overline{x_I})^2$$

$$SSTr = J\sum_{i=1}^{I}(\overline{x_i} - \bar{x})^2$$

$$SST = SSTr + SSE$$

(Có thể tính SST bằng cách nhập toàn bộ bảng vào máy tính bỏ túi, $SST = n.\hat{s}^2$; sau đó tính SSTr, rồi tính $SSE = SST - SSTr$)

4. Trung bình bình phương sai số:

$$MSE = \frac{SSE}{I(J-1)}$$

5. Trung bình bình phương nghiệm thức:

$$MSTr = \frac{SSTr}{I-1}$$

6. Trung bình bình phương toàn thể:

$$MST = \frac{SST}{IJ-1}$$

7. Tính giá trị giá trị quan sát:

$$F = \frac{MSTr}{MSE}$$

8. Miền bác bỏ: $F > F_{\alpha;I-1;I(J-1)}$

(J: tổng số quan sát ở 1 nhóm, I: số nhóm so sánh)

Dạng 10: Kiểm định trung bình (kỳ vọng) 2 mẫu nhỏ độc lập, chưa biết phương sai σ_1^2, σ_2^2

1. Phát biểu giả thiết

Giả thiết	Miền bác bỏ	Bác bỏ H, chấp nhận $\overline{H}$	P - value
H: $a_1 = a_2$ $\overline{H}$: $a_1 \neq a_2$	$W_\alpha = (-\infty, -t_{\alpha/2^{(df)}}) \cup (t_{\alpha/2^{(df)}}, +\infty)$	$\lvert T_{qs}\rvert > t_{\alpha/2^{(df)}}$ $\leftrightarrow T_{qs} \in W_\alpha$	$p = 2P(T_{df} \geq \lvert T_{qs}\rvert)$
H: $a_1 = a_2$ $\overline{H}$: $a_1 < a_2$	$W_\alpha = (-\infty, -t_{\alpha^{(df)}})$	$T_{qs} < -t_{\alpha^{(df)}}$ $\leftrightarrow T_{qs} \in W_\alpha$	$p = P(T_{df} \leq -T_{qs})$
H: $a_1 = a_2$ $\overline{H}$: $a_1 > a_2$	$W_\alpha = (t_{\alpha^{(df)}}, +\infty)$	$T_{qs} > t_{\alpha^{(df)}}$ $\leftrightarrow T_{qs} \in W_\alpha$	$p = P(T_{df} \geq T_{qs})$

2. Tính giá trị quan sát:

Trường hợp $\sigma_1^2 = \sigma_2^2 = \sigma^2$:

$$T_{qs} = \frac{\bar{x} - \bar{y}}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \text{ với } {s_p}^2 = \frac{(n_1 - 1).s_1^2 + (n_2 - 1).s_2^2}{n_1 + n_2 - 2}, df = n_1 + n_2 - 2$$

Trường hợp $\sigma_1^2 \neq \sigma_2^2$:

$$T_{qs} = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \text{ với } df = \frac{\left[(s_1^2/n_1) + (s_2^2/n_2)\right]^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}}$$

Dạng 11: Kiểm định trung bình (kỳ vọng) 2 mẫu nhỏ phụ thuộc (không độc lập)

1. Phát biểu giả thiết

Giả thiết	Miền bác bỏ	Bác bỏ H, chấp nhận $\overline{H}$	P - value
H: $a_1 = a_2$ $\overline{H}$: $a_1 \neq a_2$	$W_\alpha = (-\infty, -t_{\alpha/2^{(n-1)}}) \cup (t_{\alpha/2^{(n-1)}}, +\infty)$	$\lvert T_{qs}\rvert > t_{\alpha/2^{(n-1)}}$ $\leftrightarrow T_{qs} \in W_\alpha$	$p = 2P(T_{n-1} \geq \lvert T_{qs}\rvert)$
H: $a_1 = a_2$ $\overline{H}$: $a_1 < a_2$	$W_\alpha = (-\infty, -t_{\alpha^{(n-1)}})$	$T_{qs} < -t_{\alpha^{(n-1)}}$ $\leftrightarrow T_{qs} \in W_\alpha$	$p = P(T_{n-1} \leq -T_{qs})$
H: $a_1 = a_2$ $\overline{H}$: $a_1 > a_2$	$W_\alpha = (t_{\alpha^{(n-1)}}, +\infty)$	$T_{qs} > t_{\alpha^{(n-1)}}$ $\leftrightarrow T_{qs} \in W_\alpha$	$p = P(T_{n-1} \geq T_{qs})$

2. Tính giá trị quan sát:

$$T_{qs} = \frac{\bar{D}}{S_D/\sqrt{n}} \text{ với } D_i = X_i - Y_i\text{ ; } \overline{D} = \frac{\sum_{i=1}^{n} D_i}{n}\text{ ; } S_D = \sqrt{\frac{\sum_{i=1}^{n}(D_i - \overline{D})^2}{n - 1}}$$

Dạng 12: Kiểm định trung bình (kỳ vọng) 2 mẫu lớn phụ thuộc (không độc lập)

1. Phát biểu giả thiết

Giả thiết	Miền bác bỏ	Bác bỏ H, chấp nhận $\overline{H}$	P - value
H: $a_1 = a_2$ $\overline{H}$: $a_1 \neq a_2$	$W_\alpha = (-\infty, -Z_\alpha) \cup (Z_\alpha, +\infty)$	$\lvert U_{qs}\rvert > Z_\alpha \leftrightarrow U_{qs} \in W_\alpha$	$p = 2[1 - \Phi(\lvert U_{qs}\rvert)]$
H: $a_1 = a_2$ $\overline{H}$: $a_1 < a_2$	$W_\alpha = (-\infty, -Z_{2\alpha})$	$U_{qs} < -Z_{2\alpha} \leftrightarrow U_{qs} \in W_\alpha$	$p = \Phi(U_{qs})$
H: $a_1 = a_2$ $\overline{H}$: $a_1 > a_2$	$W_\alpha = (Z_{2\alpha}, +\infty)$	$U_{qs} > Z_{2\alpha} \leftrightarrow U_{qs} \in W_\alpha$	$p = 1 - \Phi(U_{qs})$

2. Tính giá trị quan sát: Tương tự như dạng 11 (thay kí hiệu T_{qs} thành U_{qs})

Nội dung 3: Hồi quy

Dạng 1: Viết phương trình đường hồi quy và tìm hệ số tương quan

Phương trình hồi quy tuyến tính mẫu của Y theo X: y = a + bx, với:

$$\begin{cases} a = \overline{y} - b\overline{x} \\ b = \dfrac{\overline{xy} - \overline{x}.\overline{y}}{\widehat{s_x}^2} \end{cases}$$

Phương trình hồi quy tuyến tính mẫu của X theo Y: x = c + dy, với:

$$\begin{cases} c = \overline{x} - d\overline{y} \\ d = \dfrac{\overline{xy} - \overline{x}.\overline{y}}{\widehat{s_y}^2} \end{cases}$$

Hệ số tương quan mẫu:

$$r_{XY} = \frac{\overline{xy} - \overline{x}.\overline{y}}{\widehat{s_x}.\widehat{s_y}}$$

Nhận xét mối quan hệ tuyến tính giữa X và Y:

$|r_{XY}| < 0{,}7$: nghèo nàn ; $0{,}7 \le |r_{XY}| < 0{,}8$: khá,

$0{,}8 \le |r_{XY}| < 0{,}9$: tốt; $|r_{XY}| \ge 0{,}9$: xuất sắc (xấp xỉ tuyến tính)

Hệ số xác định: r^2

Dạng 2: Ước lượng sai số chuẩn, phương sai của sai số

$$SSE = SST - SSR = n.\widehat{s_y}^2 - \frac{n(\overline{xy} - \overline{x}.\overline{y})^2}{\widehat{s_x}^2}$$

Sai số chuẩn đường hồi quy: $\sigma = \sqrt{\dfrac{SSE}{n-2}}$

Ước lượng phương sai: $\sigma^2 = \dfrac{SSE}{n-2}$

Dạng 3: Khoảng tin cậy cho các hệ số hồi quy

Khoảng tin cậy cho hệ số tự do A: $a - \varepsilon_a < A < a + \varepsilon_a$, trong đó:

$$\varepsilon_a = t_{\alpha/2^{(n-2)}} \frac{\sqrt{SSE.\overline{x^2}}}{\widehat{s_x}\sqrt{n(n-2)}}$$

Khoảng tin cậy cho hệ số góc B: $b - \varepsilon_b < B < b + \varepsilon_b$, trong đó:

$$\varepsilon_b = t_{\alpha/2^{(n-2)}} \frac{\sqrt{SSE}}{\widehat{s_x}\sqrt{n(n-2)}}$$

Dạng 4: Kiểm định cho các hệ số hồi quy

Kiểm định A:

1. Phát biểu giả thiết

Giả thiết	Miền bác bỏ	Bác bỏ H, chấp nhận $\overline{H}$	P - value
H: $A = a_0$ $\overline{H}$: $A \neq a_0$	$W_\alpha = (-\infty, -t_{\alpha/2^{(n-2)}}) \cup (t_{\alpha/2^{(n-2)}}, +\infty)$	$\|T_{qs}\| > t_{\alpha/2^{(n-2)}}$ $\leftrightarrow T_{qs} \in W_\alpha$	$p = 2P(T_{n-2} \ge \|T_{qs}\|)$
H: $A = a_0$ $\overline{H}$: $A < a_0$	$W_\alpha = (-\infty, -t_{\alpha^{(n-2)}})$	$T_{qs} < -t_{\alpha^{(n-2)}}$ $\leftrightarrow T_{qs} \in W_\alpha$	$p = P(T_{n-2} \le -T_{qs})$
H: $A = a_0$ $\overline{H}$: $A > a_0$	$W_\alpha = (t_{\alpha^{(n-2)}}, +\infty)$	$T_{qs} > t_{\alpha^{(n-2)}}$ $\leftrightarrow T_{qs} \in W_\alpha$	$p = P(T_{n-2} \ge T_{qs})$

Thông thường $a_0 = 0$

2. Tính giá trị quan sát:

$$T_{qs} = \frac{a - a_0}{\dfrac{\sqrt{SSE.\overline{x^2}}}{\widehat{s_x}\sqrt{n(n-2)}}}$$

Kiểm định B:

1. Phát biểu giả thiết

Giả thiết	Miền bác bỏ	Bác bỏ H, chấp nhận $\overline{H}$	P - value
H: $B = b_0$ $\overline{H}$: $B \neq b_0$	$W_\alpha = (-\infty, -t_{\alpha/2^{(n-2)}}) \cup (t_{\alpha/2^{(n-2)}}, +\infty)$	$\lvert T_{qs}\rvert > t_{\alpha/2^{(n-2)}}$ $\leftrightarrow T_{qs} \in W_\alpha$	$p = 2P(T_{n-2} \geq \lvert T_{qs}\rvert)$
H: $B = b_0$ $\overline{H}$: $B < b_0$	$W_\alpha = (-\infty, -t_{\alpha^{(n-2)}})$	$T_{qs} < -t_{\alpha^{(n-2)}}$ $\leftrightarrow T_{qs} \in W_\alpha$	$p = P(T_{n-2} \leq -T_{qs})$
H: $B = b_0$ $\overline{H}$: $B > b_0$	$W_\alpha = (t_{\alpha^{(n-2)}}, +\infty)$	$T_{qs} > t_{\alpha^{(n-2)}}$ $\leftrightarrow T_{qs} \in W_\alpha$	$p = P(T_{n-2} \geq T_{qs})$

Thông thường $b_0 = 0$

2. Tính giá trị quan sát:

$$T_{qs} = \frac{b - b_0}{\dfrac{\sqrt{SSE}}{\widehat{s_x}\sqrt{n(n-2)}}}$$

Dạng 5: Kiểm định sự phù hợp của hàm hồi quy tuyến tính đơn

1. Phát biểu giả thiết:

H: $R^2 = 0$ hoặc ($B = 0$): Phương trình đường hồi quy không thích hợp

$\overline{H}$: $R^2 \neq 0$ hoặc ($B \neq 0$): Phương trình đường hồi quy thích hợp

2. Tính giá trị quan sát:

$$F = \frac{r^2}{\dfrac{1 - r^2}{n - 2}}$$

3. Miền bác bỏ: $W_\alpha = (F_{\alpha^{(1, n-2)}}; +\infty)$
4. Bác bỏ H, chấp nhận $\overline{H}$ khi $F > F_{\alpha^{(1, n-2)}} \leftrightarrow F \in W_\alpha$

Dạng 6: Kiểm định mối tương quan tuyến tính X, Y

1. Phát biểu giả thiết:

H: X,Y không có tương quan tuyến tính ($R = 0$)

$\overline{H}$: X,Y có tương quan tuyến tính ($R \neq 0$)

2. Tính giá trị quan sát:

$$T_{qs} = r\sqrt{\frac{n - 2}{1 - r^2}}$$

3. Miền bác bỏ: $W_\alpha = (-\infty, -t_{\alpha/2^{(n-2)}}) \cup (t_{\alpha/2^{(n-2)}}, +\infty)$
4. Bác bỏ H, chấp nhận $\overline{H}$ khi $\lvert T_{qs}\rvert > t_{\alpha/2^{(n-2)}} \leftrightarrow T_{qs} \in W_\alpha$

Dạng 7: Khoảng tin cậy cho giá trị dự đoán

$$a + bx_0 \pm t_{\alpha/2^{(n-2)}}.\sigma.\sqrt{\left[\frac{1}{n} + \frac{(\overline{x} - x_0)^2}{n.\widehat{s_x}^2}\right]}$$

TỔNG HỢP CÔNG THỨC XÁC SUẤT CUỐI KỲ

Nội dung 1: Công thức cộng xác suất

$$P(A+B) = P(A) + P(B) - P(AB)$$
$$P(A+B+C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$$

Tổng quát:

$$P(A_1+A_2+A_3+\ldots A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_iA_j) + \sum_{i<j<k} P(A_iA_jA_k) - \ldots + (-1)^{n-1}.P(A_1A_2A_3\ldots A_n)$$

Nếu A, B xung khắc thì $P(A+B) = P(A) + P(B)$

Nếu $A_1; A_2; A_3;\ldots.A_n$ XK đôi một thì $P(A_1+A_2+A_3+..A_n) = P(A_1) + P(A_2) + \ldots + P(A_n)$

Nếu $\{A_1; A_2; A_3; \ldots, A_n\}$ là nhóm biến cố đầy đủ thì $P(A_1) + P(A_2) + \ldots + P(A_n) = 1$

Nội dung 2: Công thức xác suất đầy đủ - công thức Bayes

Định lý: Giả sử $\{H_1, H_2, \ldots, H_n\}$ là hệ biến cố đầy đủ và F là một biến cố bất kỳ.

Công thức 1:

$$P(F) = P(H_1).P(F/H_1) + P(H_2).P(F/H_2) + \ldots + P(H_n).P(F/H_n) = \sum_{i=1}^{n} P(H_i).P(F/H_i)$$

⇨ Đây là công thức xác suất toàn phần hay công thức thức xác suất đầy đủ

Công thức 2:

$$P(H_k/F) = \frac{P(H_k.F)}{P(F)} = \frac{P(H_k).P(F/H_k)}{\sum_{i=1}^{n} P(H_i).P(F/H_i)}; k = 1, 2, 3, \ldots n \text{ và } P(F) \neq 0$$

⇨ Đây là công thức Bayes.

Công thức 3:

$$P(F_2/F_1) = \frac{P(F_1F_2)}{P(F_1)} = \frac{\sum_{i=1}^{n} P(H_i).P(F_1F_2/H_i)}{\sum_{i=1}^{n} P(H_i).P(F_1/H_i)}$$

⇨ Đây là công thức mở rộng.

Nội dung 3: Các quy luật phân phối xác suất

1. **Phân phối Đều:** $X \sim U(a, b)$, nếu hàm mật độ của nó có dạng:

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a; b] \\ 0, & x \notin [a; b] \end{cases}$$

$$E(X) = med(X) = \frac{a+b}{2}; D(X) = \frac{(b-a)^2}{12}; Mod(X) = x, \text{ với } x \in [a; b]$$

2. **Phân phối Mũ:** $X \sim E(\lambda)$, nếu hàm mật độ của nó có dạng:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

$$E(X) = \frac{1}{\lambda}; D(X) = \frac{1}{\lambda^2}; Mod(X) = 0; Med(X) = \ln(2)/\lambda$$

- Công thức **mất trí nhớ**: $X \sim E(\lambda)$ thì với mọi $a,b \geq 0$, ta có:

$$P(X > a + b \mid X > a) = P(X > b)$$

- Nếu các biến ngẫu nhiên $X_i \sim E(\lambda_i)$, $i = 1, 2, 3, \ldots, n$ độc lập thì ta có:

$$Y = \min\{X_1, X_2, X_3, \ldots X_n\} \sim E(\lambda_1 + \lambda_2 + \lambda_3 + \ldots + \lambda_n)$$

3. Phân phối Poisson: $X \sim P(\lambda)$:

$E(X) = D(X) = \lambda;\ \lambda - 1 \le Mod(X) \le \lambda$ và $Mod(X) \in \mathbb{N}$

$$P(X = k) = \frac{e^{-\lambda}.\lambda^k}{k!}$$

$$P(k_1 \le X \le k_2) = \sum_{k_1}^{k_2} \frac{e^{-\lambda}.\lambda^k}{k!}$$

- Nếu các biến ngẫu nhiên $X_i \sim P(\lambda_i)$, $i = 1, 2, 3, \ldots, n$ độc lập thì ta có:

$$Y = X_1 + X_2 + X_3 + \ldots + X_n \sim P(\lambda_1 + \lambda_2 + \lambda_3 + \ldots + \lambda_n)$$

4. Phân phối Chuẩn: $X \sim N(a, \sigma^2)$, nếu hàm mật độ của nó có dạng:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}},\ x \in R$$

$E(X) = Mod(X) = Med(X) = a;\ D(X) = \sigma^2$

$$P(k_1 \le X \le k_2) = \int_{k_1}^{k_2} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \Phi\left(\frac{k_2 - a}{\sigma}\right) - \Phi\left(\frac{k_1 - a}{\sigma}\right) = \overline{\Phi}\left(\frac{k_2 - a}{\sigma}\right) - \overline{\Phi}\left(\frac{k_1 - a}{\sigma}\right) (*)$$

$$P(|X - a| < \varepsilon) = 2.\ \Phi\left(\frac{\varepsilon}{\sigma}\right), \varepsilon > 0$$

- Nếu các biến ngẫu nhiên $X_i \sim N(a_i, \sigma_i^2) = 1, 2, 3, \ldots, n$ độc lập thì ta có:

$$Y = X_1 + X_2 + X_3 + \ldots + X_n \sim N(a, \sigma^2) \text{ với } a = a_1 + a_2 + a_3 + \ldots + a_n,\ \sigma^2 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \ldots + \sigma_n^2$$

- Nếu các biến ngẫu nhiên $X \sim N(a_1, \sigma_1^2)$ và $Y \sim N(a_2, \sigma_2^2)$ độc lập thì ta có:

$$Z = c.X + d.Y \sim N(a, \sigma^2) \text{ với } a = c.a_1 + d.a_2;\ \sigma^2 = c^2.\sigma_1^2 + d^2.\sigma_2^2$$

5. Phân phối Siêu bội: $X \sim H\ (N, M, n)$:

$E(X) = np;\ D(X) = npq\frac{N - n}{N - 1}$ với $p = \frac{M}{N}$ và $q = 1 - p$

$$P(X = k) = \frac{C_M^k.C_{N-M}^{n-k}}{C_N^n} \text{ và } P(k_1 \le X \le k_2) = \sum_{k_1}^{k_2} \frac{C_M^k.C_{N-M}^{n-k}}{C_N^n}$$

Trong trường hợp $n << N$ (n rất nhỏ so với N lớn): $P(X = k) = \frac{C_M^k.C_{N-M}^{n-k}}{C_N^n} \approx C_n^k.p^k.q^{n-k};\ k = 0, 1, 2,, n$

6. Phân phối nhị thức: $X \sim B(n, p)$:

$E(X) = np;\ D(X) = npq$ với $q = 1 - p$, $[np - q] \le Mod(X) \le [np - q]+1$ ([ ]: lấy phần nguyên) và $Mod(X) \in \mathbb{N}$

$$P(X = k) = C_n^k.p^k.q^{n-k}$$

$$P(k_1 \le X \le k_2) = \sum_{k_1}^{k_2} C_n^k.p^k.q^{n-k}$$

1. Trong trường hợp n lớn, $5\% \le p \le 95\%$. Khi đó biến ngẫu nhiên X có phân phối Nhị thức sẽ được xem như xấp xỉ phân phối Chuẩn $N(a = np, \sigma^2 = npq)$:

$$P(X = k) = \frac{1}{\sqrt{npq}\sqrt{2\pi}} e^{-\frac{(k - np)^2}{2npq}} \text{ và } P(k_1 \le X \le k_2) = \Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right)$$

2. Trong trường hợp n lớn, $p < 5\%$. Khi đó biến ngẫu nhiên X có phân phối Nhị thức sẽ được xem như xấp xỉ phân phối Poisson $P(\lambda = np)$

$$P(X = k) = \frac{e^{-(np)}.(np)^k}{k!} \text{ và } P(k_1 \le X \le k_2) = \sum_{k_1}^{k_2} \frac{e^{-(np)}.(np)^k}{k!}$$

- Nếu các biến ngẫu nhiên $X_i \sim B(n_i, p) = 1, 2, 3, \ldots, n$ độc lập thì ta có:

$$Y = X_1 + X_2 + X_3 + \ldots + X_n \sim B(n, p) \text{ với } n = n_1 + n_2 + n_3 + \ldots + n_n$$

Nội dung 4: Véc tơ ngẫu nhiên (X, Y) rời rạc

1. **Phân phối xác suất có điều kiện:**

Bảng phân phối xác suất của X với điều kiện $Y = y_j$ $(j = \overline{1,n})$ là:

X	x_1	x_2	...	x_m
P^{X/y_j}	$\frac{p_{1j}}{q_j}$	$\frac{p_{2j}}{q_j}$	...	$\frac{p_{mj}}{q_j}$

$$P(X = x_i \mid Y = y_j) = \frac{p_{ij}}{q_j}$$

Bảng phân phối xác suất của Y đối với điều kiện $X = x_i$ $(i = \overline{1, m})$ là:

X	y_1	y_2	...	y_n
P^{Y/x_i}	$\frac{p_{i1}}{p_i}$	$\frac{p_{i2}}{p_i}$	...	$\frac{p_{in}}{p_i}$

$$P(Y = y_j \mid X = x_i) = \frac{p_{ij}}{p_i}$$

2. **Điều kiện độc lập của X, Y:**

X và Y độc lập $\leftrightarrow P(X = x_i, Y = y_j) = P(X = x_i).P(Y = y_j)\ \forall i, j$ hay $p_{ij} = p_i q_j\ \forall i, j$.

$\leftrightarrow F(x, y) = F_X(x).F_Y(y)$

(F_X , F_Y là các hàm phân phối xác suất của X, Y, hay gọi la các hàm phân phối lề)

3. **Hàm phân phối xác suất đồng thời của (X, Y):**

$$F(x, y) = P(X < x, Y < y) = \sum_{x_i < x} \sum_{y_j < y} p_{ij}$$

Nội dung 5: Các đặc trưng của véc tơ ngẫu nhiên (X, Y)

1. **Kỳ vọng toán:** $E(X, Y) = (E(X), E(Y))$
2. **Hiệp phương sai (Coravian):** $Cov(X, Y) = E(XY) - E(X).E(Y)$
3. **Ma trận tương quan:**

$$D(X, Y) = \begin{bmatrix} cov(X, X) & cov(X, Y) \\ cov(Y, X) & cov(Y, Y) \end{bmatrix} = \begin{bmatrix} D(X) & cov(X, Y) \\ cov(Y, X) & D(Y) \end{bmatrix}$$

4. **Hệ số tương quan của X và Y:**

$$R_{XY} = \frac{cov(X,Y)}{\sqrt{D(X)}.\sqrt{D(Y)}} = \frac{E(XY) - E(X).E(Y)}{\sqrt{D(X)}.\sqrt{D(Y)}}$$

Nội dung 6: Định lý giới hạn trung tâm

Giả sử $X_1, X_2, X_3, \ldots X_n$ là các biến ngẫu nhiên độc lập cùng tuân theo một quy luật phân phối xác suất nào đó. Kí hiệu $E(X_i) = a$ và $D(X_i) = \sigma^2, \forall i$. Khi $n \to \infty$, chúng ta có sự hội tụ theo xác suất của các BNN sau:

Biến ngẫu nhiên $X = X_1 + X_2 + \ldots + X_n$ hội tụ về phân phối chuẩn $N(n.a, n.\sigma^2)$

Biến ngẫu nhiên $\overline{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}$ hội tụ về phân phối chuẩn $N(a, \frac{\sigma^2}{n})$

TÀI LIỆU THAM KHẢO:

BÀI GIẢNG XSTK – THẦY NGUYỄN BÁ THI