TRƯỜNG ĐẠI HỌC
KINH TẾ KỸ THUẬT CÔNG NGHIỆP
UNIVERSITY OF ECONOMIC AND TECHNICAL INDUSTRIES

*Giảng viên:*

*Chu Bình Minh*



*Bài giảng*

# *Xác suất thống kê*

*Nam Dinh,Februay, 2008*

# PHẦN 2
# THỐNG KÊ TOÁN

## CHƯƠNG 5:
## LÝ THUYẾT ƯỚC LƯỢNG

# BÀI TOÁN:

Cho biến ngẫu nhiên gốc X với quy luật phân phối xác suất đã biết song chưa biết tham số θ. Ta phải ước lượng cho tham số θ.

# 1 PHƯƠNG PHÁP ƯỚC ĐIỂM

## 1.1 KHÁI NIỆM

Từ biến ngẫu nhiên gốc X ta lập mẫu ngẫu nhiên

$$W_n = (X_1, X_2, \dots, X_n)$$

Và từ mẫu này ta xây dựng thống kê

$$\hat{\theta} = f(X_1, X_2, \dots, X_n)$$

Nếu ta sử dụng thống kê $\hat{\theta}$ để ước lượng cho $\theta$ thì ta gọi là ước lượng điểm.

Có rất nhiều cách để ước lượng cho θ nhưng trong phạm vi bài giảng này chỉ trình bày loại ước lượng đơn giản nhất và ước lượng hiệu quả nhất.

## 1.2 ƯỚC LƯỢNG KHÔNG CHỆCH

Khi dùng thống kê $\hat{\theta}$ để ước lượng cho θ ta không thể căn cứ vào một vài trường hợp cụ thể mà kết luật được mà cần phải dựa vào giá trị trung bình của nó, do vậy ta có định nghĩa:

Định nghĩa



Thống kê $\hat{\theta}$ gọi là ước lượng không chệch của θ nếu $E\hat{\theta} = \theta$. Ngược lại gọi là ước lượng chệch.

Ví dụ

Cho X là biến ngẫu nhiên gốc X có quy luật $N(\mu, \sigma^2)$, hãy tìm các ước lượng không chệch của μ và $\sigma^2$.

Giải

Từ X ra rút ra mẫu $W = (X_1, X_2, X_3, X_4, X_5)$, dựa vào thống kê này ta có thể đưa ra các ước lượng không chệch cho μ và $\sigma^2$ như sau:

1, Đối với μ ta có



$$\widehat{\theta_1} = \bar{X} = \frac{X_1 + X_2 + X_3 + X_4 + X_5}{5}$$



Do $E\widehat{\theta_1} = E\bar{X} = \mu$ nên $\widehat{\theta_1}$ là ước lượng không chệch của μ

Hoặc

$$\widehat{\theta_2} = \frac{X_1 + 3X_2 + 4X_3 + X_4 + X_5}{10}$$

Do $E\widehat{\theta_2} = \mu$ nên $\widehat{\theta_2}$ là ước lượng không chệch khác của μ

2, Đối với $\sigma^2$ ta có



$$\widehat{\theta_3} = S^2 = \frac{1}{4}\sum_{i=1}^{5}(X_i - \bar{X})^2$$

Do $E\widehat{\theta_3} = \sigma^2$ nên $\widehat{\theta_3}$ là ước lượng không chệch của $\sigma^2$.

## 1.3 ƯỚC LƯỢNG HIỆU QUẢ

Như ví dụ trên ta thấy rằng một tham số θ có thể có nhiều ước lượng không chệch như $\widehat{\theta_1}, \widehat{\theta_2}$ nên một câu hỏi tự nhiên đặt ra là trong hai ước lượng không chệch của θ thì ước lượng nào tốt hơn. Mặt khác ta chú ý rằng nếu $\hat{\theta}$ để ước lượng không chệch cho $\theta$ không có nghĩa là mọi giá trị của $\hat{\theta}$ đều trùng khít với θ mà chỉ có nghĩa rằng trung bình các giá trị của $\hat{\theta}$ bằng θ. Do đó ta có thể coi ước lương không chệch nào của θ mà có trung bình của bình phương độ lệch so với θ bé hơn thì tốt hơn, nghia là ta so sánh 2 giá trị $E(\widehat{\theta_1} - \theta)^2$ và $E(\widehat{\theta_2} - \theta)^2$ để tìm ước lượng tốt hơn.

Nhưng:

$$E(\widehat{\theta_1} - \theta)^2 = E(\widehat{\theta_1} - E\widehat{\theta_1})^2 = D\widehat{\theta_1}, E(\widehat{\theta_2} - \theta)^2 = D\widehat{\theta_2}$$

Nên ta có định nghĩa:

# 1.3 ƯỚC LƯỢNG HIỆU QUẢ

## Định nghĩa

Cho $\widehat{\theta_1}$ và $\widehat{\theta_2}$ là hai ước lượng không chệch của θ. Nếu $D\widehat{\theta_1} < D\widehat{\theta_2}$ thì ta nói $\widehat{\theta_1}$ là ước lượng hiệu quả hơn $\widehat{\theta_2}$.
Nếu thống kê $\widehat{\theta_1}$ là ước lượng không chệch của θ có $D\widehat{\theta_1}$ nhỏ nhất thì ta nói $\widehat{\theta_1}$ là ước lượng hiệu quả nhất của θ.

Ví dụ

Cho X là biến ngẫu nhiên gốc X có quy luật $N(\mu, \sigma^2)$, từ X ra rút ra mẫu $W = (X_1, X_2, X_3, X_4, X_5)$. Như trên ta có

$$\widehat{\theta_1} = \bar{X} = \frac{X_1 + X_2 + X_3 + X_4 + X_5}{5}, \widehat{\theta_2} = \frac{X_1 + 3X_2 + 4X_3 + X_4 + X_5}{10}$$

Là 2 ước lượng không chệch của μ nhưng do $D\widehat{\theta_1} = \frac{\sigma^2}{5} < D\widehat{\theta_2} = \frac{28\sigma^2}{100}$ nên $\widehat{\theta_1}$ là ước lượng hiệu quả hơn $\widehat{\theta_2}$.

Chú ý.

- Trong thực tế do tham số θ chưa biết nên nếu ta dùng một giá trị $\hat{\theta}$ để ước lượng cho θ thì ta không thể đánh giá được sai số $|\theta - \hat{\theta}|$ nên ước lượng điểm không được sử dụng nhiều trong thực tế. Người ta thường dùng một khoảng giá trị để ước lượng cho θ, ước lượng như thế gọi là ước lượng khoảng sẽ được trình bày sau đây.

# 2 PHƯƠNG PHÁP ƯỚC LƯỢNG BẰNG KHOẢNG TIN CẬY

## 2.1 PHƯƠNG PHÁP CHUNG

### a, ĐỊNH NGHĨA.

*Khoảng* $(G_1, G_2)$ *của thống kê G gọi là khoảng tin cậy của tham số* $\theta$ *nếu với xác suất bằng* $(1-\alpha)$ *cho trước thoả mãn điều kiện:*

$$P(G_1 < \theta < G_2) = 1 - \alpha$$

*Xác suất* $(1-\alpha)$ *gọi là độ tin cậy của ước lượng, còn khoảng* $I = G_2 - G_1$ *gọi là độ dài khoảng tin cậy.*

# 2.1 PHƯƠNG PHÁP CHUNG

b, CÁC BƯỚC TIẾN HÀNH TÌM KHOẢNG $(G_1, G_2)$

Từ tổng thể lập mẫu ngẫu nhiên kích thước n:

$$W_n = (X_1, X_2, \dots, X_n)$$

Và từ đó xây dựng thống kê $G = f(X_1, X_2, \dots, X_n, \theta)$ sao cho quy luật phân phối của G không phụ thuộc vào các biến số và hoàn toàn xác định. Lúc đó với độ tin cậy $(1-\alpha)$ cho trước có thể tìm được cặp giá trị $\alpha_1, \alpha_2$ sao cho $\alpha_1 + \alpha_2 = \alpha$ và tương ứng ta tìm được cặp giá trị $g\alpha_1, g\alpha_2$ thoả mãn:

$$P(G < g\alpha_1) = \alpha_1$$

$$P(G > g\alpha_2) = \alpha_2$$

## 2.1 PHƯƠNG PHÁP CHUNG

Và từ đó suy ra $P(g\alpha_1 < G < g\alpha_2) = 1 - (\alpha_1 + \alpha_2) = 1 - \alpha$

Như vậy, với độ tin cậy $(1 - \alpha)$, ta đã tìn được khoảng $(g\alpha_1, g\alpha_2)$ cho G. Biến đổi tương đương biểu thức $g\alpha_1 < G < g\alpha_2$ ta có $G_1 < \theta < G_2$, suy ra

$$P(G_1 < \theta < G_2) = 1 - \alpha$$

hay $(G_1, G_2)$ chính là khoảng ước lượng cần tìm.

## 2.2 ƯỚC LƯỢNG KỲ VỌNG CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬN PHÂN PHỐI

Giả sử trong tổng thể biến ngẫu nhiên gốc X có phân phối chuẩn $N(\mu, \sigma^2)$ nhưng chưa biết tham số μ. Để ước lượng μ, từ tổng thể ta lập mẫu ngẫu nhiên kích thước n:

$$W_n = (X_1, X_2, \dots, X_n)$$

Để chọn thống kê G cho thích hợp ta xet hai trường hợp sau:



a, Đã biết phương sai $\sigma^2$ của biến ngẫu nhiên gốc X trong tổng thể.

Do trung bình mẫu:



$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

có $X_i$ có cùng quy luật phân phối với X và $E(\bar{X}) = \mu, D(\bar{X}) = \frac{\sigma^2}{n}$ nên $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

## 2.2 ƯỚC LƯỢNG KỲ VỌNG CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬN PHÂN PHỐI

Chọn thống kê

$$G = U = \frac{\bar{X} - \mu}{\sigma}\sqrt{n} \sim N(0,1)$$

với độ tin cậy $(1 - \alpha)$ cho trước có thể tìm được cặp giá trị $\alpha_1, \alpha_2$ sao cho $\alpha_1 + \alpha_2 = \alpha$ và tương ứng ta tìm được cặp giá trị $u_{1-\alpha_1}, u_{\alpha_2}$ thoả mãn:

$$P\left(G < u_{1-\alpha_1}\right) = \alpha_1$$

$$P\left(G > u_{\alpha_2}\right) = \alpha_2$$

suy ra

$$P\left(u_{1-\alpha_1} < U < u_{\alpha_2}\right) = 1 - (\alpha_1 + \alpha_2) = 1 - \alpha$$

do $u_{1-\alpha_1} = -u_{\alpha_1}$ nên ta có:

$$P\left(-u_{\alpha_1} < U < u_{\alpha_2}\right) = 1 - \alpha$$

## 2.2 ƯỚC LƯỢNG KỲ VỌNG CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬN PHÂN PHỐI

do $u_{1-\alpha_1} = -u_{\alpha_1}$ nên ta có:

$$P(-u_{\alpha_1} < U < u_{\alpha_2}) = 1 - \alpha$$

$$\Leftrightarrow P\left(-u_{\alpha_1} < \frac{\bar{X} - \mu}{\sigma}\sqrt{n} < u_{\alpha_2}\right) = 1 - \alpha$$



$$\Leftrightarrow P\left(\bar{X} - \frac{\sigma}{\sqrt{n}}u_{\alpha_2} < \mu < \bar{X} + \frac{\sigma}{\sqrt{n}}u_{\alpha_1}\right) = 1 - \alpha$$

Vậy với độ tin cậy $(1-\alpha)$ tham số μ của biến ngẫu nhiên gốc sẽ nằm trong khoảng



$$\left(\bar{X} - \frac{\sigma}{\sqrt{n}}u_{\alpha_2}, \bar{X} + \frac{\sigma}{\sqrt{n}}u_{\alpha_1}\right)$$

## 2.2 ƯỚC LƯỢNG KỲ VỌNG CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬN PHÂN PHỐI

- Khoảng tin cậy đối xứng: $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$ ta có khoảng tin cậy

$$(\bar{X} - \varepsilon, \bar{X} + \varepsilon)$$

với $\varepsilon = \frac{\sigma}{\sqrt{n}} u_{\alpha/2}$ được gọi là độ chính xác.

- Khoảng tin cậy bên phải: $\alpha_1 = 0, \alpha_2 = \alpha$ thì $u_{\alpha_1} = u_0 = +\infty$ và do đó khoảng tin cậy là

$$\left(\bar{X} - \frac{\sigma}{\sqrt{n}} u_\alpha, +\infty\right)$$

Biểu thức trên dùng để ước lượng giá trị tối thiểu của μ.

- Khoảng tin cậy bên trái: $\alpha_1 = \alpha, \alpha_2 = 0$ thì $u_{\alpha_2} = u_0 = +\infty$ và do đó khoảng tin cậy là

$$\left(-\infty, \bar{X} + \frac{\sigma}{\sqrt{n}} u_\alpha\right)$$

Biểu thức trên dùng để ước lượng giá trị tối đa của μ.

## 2.2 ƯỚC LƯỢNG KỲ VỌNG CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬN PHÂN PHỐI

Từ mẫu cụ thể

$$w = (x_1, x_2, \dots, x_n)$$

và từ đó ta tìm được giá trị cụ thể của trung bình mẫu $\bar{x}$.

Lúc đó với độ tin cậy $(1-\alpha)$, qua một mẫu cụ thể, khoảng tin cậy của μ là:



$$(g_1, g_2) = \left(\bar{x} - \frac{\sigma}{\sqrt{n}} u_{\alpha_2}, \bar{x} + \frac{\sigma}{\sqrt{n}} u_{\alpha_1}\right)$$

## 2.2 ƯỚC LƯỢNG KỲ VỌNG CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬN PHÂN PHỐI

Chú ý

Cùng với độ tin cậy $(1-\alpha)$ thì khoảng tin cậy nào ngắn hơn sẽ tốt hơn, trong trường hợp này khoảng tin cậy đối xứng sẽ ngắn nhất và có độ dài $I = 2\varepsilon$

Nếu bài toán cho trước độ dài đoạn tin cậy (độ chính xác $\varepsilon$) yêu cầu xác định kích thước tối thiểu của mẫu n với độ tin cậy $(1-\alpha)$ thì từ công thức

$$I = 2\varepsilon = \frac{2\sigma}{\sqrt{n}} u_{\alpha/2}$$

ta có:

$$n \geq \left[\frac{4\sigma^2}{I} u^2{}_{\alpha/2}\right] = \left[\frac{\sigma^2}{\varepsilon^2} u^2{}_{\alpha/2}\right]$$

## 2.2 ƯỚC LƯỢNG KỲ VỌNG CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬN PHÂN PHỐI

Ví dụ 1. Trọng lượng một loại sản phẩm là biến ngẫu nhiên phân phối chuẩn với độ lệch chuẩn là 1 gam. Cân thử 25 sản phẩm loại này ta thu được kết quả sau:

| Trọng lượng (gam) | 18 | 19 | 20 | 21 |
|---|---|---|---|---|
| Số SP tương ứng | 3 | 5 | 15 | 2 |

a, Với độ tin cậy 0,95 hãy tìm khoảng tin cậy đối xứng của trọng lượng trung bình loại sản phẩm nói trên.

b, Nếu yêu cầu độ chính xác của ước lượng là 0,1 và giữ nguyên độ tin cậy $1-\alpha$ thì phải điều tra một mẫu kích thước bằng bao nhiêu?

# 2.2 ƯỚC LƯỢNG KỲ VỌNG CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬN PHÂN PHỐI

Giải

a, X là trong lượng của sản phẩm

$$X \sim N(\mu, \sigma^2), \sigma = 1$$

$$1 - \alpha = 0{,}95 \Longrightarrow \alpha = 0{,}05$$

$$u_{\alpha/2} = u_{0{,}025} = 1{,}96$$

n =25, $\bar{x} = 19{,}64$



Khoảng tin cậy là

$$\left(19{,}64 - \frac{1}{5}1{,}96; 19{,}64 + \frac{1}{5}1{,}96\right)$$

# 2.2 ƯỚC LƯỢNG KỲ VỌNG CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬN PHÂN PHỐI

b, ta có $\varepsilon = 0{,}1$, theo công thức trên ta có

$$n \geq \left[\frac{\sigma^2}{\varepsilon^2} u^2{}_{\alpha/2}\right] = \left[\frac{1}{0{,}1^2}(1{,}96)^2\right] = 385$$

## 2.2 ƯỚC LƯỢNG KỲ VỌNG CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬN PHÂN PHỐI

b, Chưa biết phương sai $\sigma^2$ của biến ngẫu nhiên gốc X trong tổng thể và $n \leq 30$

Chọn thống kê

$$G = T = \frac{\bar{X} - \mu}{S}\sqrt{n} \sim T(n-1)$$



với độ tin cậy $(1-\alpha)$ cho trước có thể tìm được cặp giá trị $\alpha_1, \alpha_2$ sao cho $\alpha_1 + \alpha_2 = \alpha$ và tương ứng ta tìm được cặp giá trị $t_{1-\alpha_1}^{(n-1)}, t_{\alpha_2}^{(n-1)}$ thoả mãn:



$$P\left(G < t_{1-\alpha_1}^{(n-1)}\right) = \alpha_1$$

$$P\left(G > t_{\alpha_2}^{(n-1)}\right) = \alpha_2$$

## 2.2 ƯỚC LƯỢNG KỲ VỌNG CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬN PHÂN PHỐI

suy ra

$$P\left(t_{1-\alpha_1}^{(n-1)} < G < t_{\alpha_2}^{(n-1)}\right) = 1 - (\alpha_1 + \alpha_2) = 1 - \alpha$$

do $t_{1-\alpha_1}^{(n-1)} = -t_{\alpha_1}^{(n-1)}$ nên ta có:

$$P\left(-t_{\alpha_1}^{(n-1)} < G < t_{\alpha_2}^{(n-1)}\right) = 1 - \alpha$$

$$\Leftrightarrow P\left(-t_{\alpha_1}^{(n-1)} < \frac{\bar{X} - \mu}{S}\sqrt{n} < t_{\alpha_2}^{(n-1)}\right) = 1 - \alpha$$

$$\Leftrightarrow P\left(\bar{X} - \frac{S}{\sqrt{n}} t_{\alpha_2}^{(n-1)} < \mu < \bar{X} + \frac{S}{\sqrt{n}} t_{\alpha_1}^{(n-1)}\right) = 1 - \alpha$$

Vậy với độ tin cậy $(1-\alpha)$ tham số μ của biến ngẫu nhiên gốc sẽ nằm trong khoảng

$$\left(\bar{X} - \frac{S}{\sqrt{n}} t_{\alpha_2}^{(n-1)}, \bar{X} + \frac{S}{\sqrt{n}} t_{\alpha_1}^{(n-1)}\right)$$

## 2.2 ƯỚC LƯỢNG KỲ VỌNG CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬN PHÂN PHỐI

- Khoảng tin cậy đối xứng: $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$ ta có khoảng tin cậy

$$(\bar{X} - \varepsilon, \bar{X} + \varepsilon)$$

với $\varepsilon = \frac{S}{\sqrt{n}} t_{\alpha/2}^{(n-1)}$ được gọi là độ chính xác.

- Khoảng tin cậy bên phải: $\alpha_1 = 0, \alpha_2 = \alpha$ thì $u_{\alpha_1} = u_0 = +\infty$ và do đó khoảng tin cậy là



$$\left(\bar{X} - \frac{S}{\sqrt{n}} t_{\alpha}^{(n-1)}, +\infty\right)$$

Biểu thức trên dùng để ước lượng giá trị tối thiểu của μ.

- Khoảng tin cậy bên trái: $\alpha_1 = \alpha, \alpha_2 = 0$ thì $u_{\alpha_2} = u_0 = +\infty$ và do đó khoảng tin cậy là



$$\left(-\infty, \bar{X} + \frac{S}{\sqrt{n}} t_{\alpha}^{(n-1)}\right)$$

## 2.2 ƯỚC LƯỢNG KỲ VỌNG CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬN PHÂN PHỐI

Biểu thức trên dùng để ước lượng giá trị tối đa của μ.

Từ mẫu cụ thể

$$w = (x_1, x_2, \dots, x_n)$$

và từ đó ta tìm được giá trị cụ thể của trung bình mẫu $\bar{x}$.

Lúc đó với độ tin cậy $(1-\alpha)$, qua một mẫu cụ thể, khoảng tin cậy của μ là:

$$(g_1, g_2) = \left(\bar{x} - \frac{S}{\sqrt{n}} t_{\alpha_2}^{(n-1)}, \bar{x} + \frac{S}{\sqrt{n}} t_{\alpha_1}^{(n-1)}\right)$$

## 2.2 ƯỚC LƯỢNG KỲ VỌNG CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬN PHÂN PHỐI

Chú ý:

Do phân phối Student hội tụ nhanh về phân phối chuẩn khi $n \to +\infty$ nên khi n rất lớn (n > 30) thì xấp xỉ $t_\alpha^{(n-1)} \approx u_\alpha$.



Ví dụ 2. Với các yêu cầu của ví dụ 1 nhưng giả thiết ta chưa biết độ lệch chuẩn.

## 2.2 ƯỚC LƯỢNG KỲ VỌNG CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬN PHÂN PHỐI

Ví dụ 3. Để xác định kích thước trung bình của chi tiết do một máy tự động sản xuất, người ta lấy ngẫu nhiên 200 chi tiết đem đo và thu được bảng số liệu sau:

| Kích thước chi tiết(cm) | Số chi tiết tương ứng |
|---|---|
| 54,795-54,805 | 6 |
| 54,805-54,815 | 14 |
| 54,815-54,825 | 33 |
| 54,825-54,835 | 47 |
| 54,835-54,845 | 45 |
| 54,845-54,855 | 33 |
| 54,855-54,856 | 15 |
| 54,865-54,875 | 7 |

Với độ tin cậy 95% hãy ước lượng khoảng tin cậy đối xứng của kích thước trung bình của chi tiết do máy sản xuất. Giả thiết kích thước chi tiết tuân theo quy luật chuẩn.

## 2.3 ƯỚC LƯỢNG HIỆU CỦA HAI KỲ VỌNG TOÁN CỦA HAI BIẾN NGẪU NHIÊN PHÂN PHỐI CHUẨN

Giả sử ta xét một lúc hai tổng thể. Ở tổng thể thứ nhất ta xét biến ngẫu nhiên gốc $X \sim N(\mu_1, \sigma_1^2)$, ở tổng thể thứ hai ta xét biến ngẫu nhiên gốc $Y \sim N(\mu_2, \sigma_2^2)$ . Từ hai tổng thể nói trên rút ra hai mẫu ngẫu nhiên độc lập có kích thước tương ứng $n_1$ và $n_2$:



$$W_X = (X_1, X_2, \dots, X_{n_1})$$

$$W_Y = (Y_1, Y_2, \dots, Y_{n_2})$$

Để chọn thống kê G cho thích hợp ta xet hai trường hợp sau:

## 2.3 ƯỚC LƯỢNG HIỆU CỦA HAI KỲ VỌNG TOÁN CỦA HAI BIẾN NGẪU NHIÊN PHÂN PHỐI CHUẨN

a, Đã biết phương sai $\sigma_1^2$ và $\sigma_2^2$ của biến ngẫu nhiên gốc X, Y trong tổng thể.

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

với độ tin cậy $(1-\alpha)$ cho trước có thể tìm được cặp giá trị $\alpha_1, \alpha_2$ sao cho $\alpha_1 + \alpha_2 = \alpha$ và tương ứng ta tìm được cặp giá trị $u_{1-\alpha_1}, u_{\alpha_2}$ thoả mãn:

$$P\left(G < u_{1-\alpha_1}\right) = \alpha_1$$
$$P\left(G > u_{\alpha_2}\right) = \alpha_2$$

suy ra

$$P\left(u_{1-\alpha_1} < U < u_{\alpha_2}\right) = 1 - (\alpha_1 + \alpha_2) = 1 - \alpha$$

# 2.3 ƯỚC LƯỢNG HIỆU CỦA HAI KỲ VỌNG TOÁN CỦA HAI BIẾN NGẪU NHIÊN PHÂN PHỐI CHUẨN

do $u_{1-\alpha_1} = -u_{\alpha_1}$ nên ta có:

$$P\left(-u_{\alpha_1} < U < u_{\alpha_2}\right) = 1 - \alpha$$

$$\Leftrightarrow P\left(-u_{\alpha_1} < \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < u_{\alpha_2}\right) = 1 - \alpha$$

$$\Leftrightarrow P\left(\bar{X} - \bar{Y} - \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\,u_{\alpha_2} < \mu_1 - \mu_2 < \bar{X} - \bar{Y} + \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\,u_{\alpha_1}\right) = 1 - \alpha$$

## 2.3 ƯỚC LƯỢNG HIỆU CỦA HAI KỲ VỌNG TOÁN CỦA HAI BIẾN NGẪU NHIÊN PHÂN PHỐI CHUẨN

Vậy với độ tin cậy $(1-\alpha)$ tham số $\mu_1-\mu_2$ của biến ngẫu nhiên gốc sẽ nằm trong khoảng

$$\left(\bar{X}-\bar{Y}-\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}u_{\alpha_2}, \bar{X}-\bar{Y}+\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}u_{\alpha_1}\right)$$



- Khoảng tin cậy đối xứng: $\alpha_1=\alpha_2=\frac{\alpha}{2}$ ta có khoảng tin cậy

$$(\bar{X}-\bar{Y}-\varepsilon, \bar{X}-\bar{Y}+\varepsilon)$$

với $\varepsilon=\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}u_{\alpha/2}$ được gọi là độ chính xác.

## 2.3 ƯỚC LƯỢNG HIỆU CỦA HAI KỲ VỌNG TOÁN CỦA HAI BIẾN NGẪU NHIÊN PHÂN PHỐI CHUẨN

- Khoảng tin cậy bên phải: $\alpha_1 = 0, \alpha_2 = \alpha$ thì $u_{\alpha_1} = u_0 = +\infty$ và do đó khoảng tin cậy là

$$\left(\bar{X} - \bar{Y} - \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\, u_\alpha, +\infty\right)$$

Biểu thức trên dùng để ước lượng giá trị tối thiểu của $\mu_1 - \mu_2$.

- Khoảng tin cậy bên trái: $\alpha_1 = \alpha, \alpha_2 = 0$ thì $u_{\alpha_2} = u_0 = +\infty$ và do đó khoảng tin cậy là

$$\left(-\infty, \bar{X} - \bar{Y} + \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\, u_\alpha\right)$$

Biểu thức trên dùng để ước lượng giá trị tối đa của $\mu_1 - \mu_2$.

## 2.3 ƯỚC LƯỢNG HIỆU CỦA HAI KỲ VỌNG TOÁN CỦA HAI BIẾN NGẪU NHIÊN PHÂN PHỐI CHUẨN

Từ các mẫu cụ thể

$$w_X = (x_1, x_2, \dots, x_{n_1})$$

$$w_Y = (y_1, y_2, \dots, y_{n_2})$$



và từ đó ta tìm được giá trị cụ thể của các trung bình mẫu $\bar{x}, \bar{y}$. Lúc đó với độ tin cậy $(1-\alpha)$, qua các mẫu cụ thể, khoảng tin cậy của $\mu_1 - \mu_2$ là:



$$(g_1, g_2) = \left(\bar{x} - \bar{y} - \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\,u_{\alpha_2}; \bar{x} - \bar{y} + \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\,u_{\alpha_1}\right)$$

## 2.3 ƯỚC LƯỢNG HIỆU CỦA HAI KỲ VỌNG TOÁN CỦA HAI BIẾN NGẪU NHIÊN PHÂN PHỐI CHUẨN

b, Chưa biết phương sai $\sigma_1^2$ và $\sigma_2^2$ của biến ngẫu nhiên gốc X, Y trong tổng thể nhưng giả thiết $\sigma_1^2 = \sigma_2^2 = \sigma^2$

Ta có:

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

và

$$\chi_X^2 = \frac{(n_1 - 1){S_X}^2}{\sigma^2} \sim \chi^2_{(n_1-1)}$$

$$\chi_Y^2 = \frac{(n_2 - 1){S_Y}^2}{\sigma^2} \sim \chi^2_{(n_2-1)}$$

# 2.3 ƯỚC LƯỢNG HIỆU CỦA HAI KỲ VỌNG TOÁN CỦA HAI BIẾN NGẪU NHIÊN PHÂN PHỐI CHUẨN

nên

$$\chi^2 = \chi_X^2 + \chi_Y^2 = \frac{1}{\sigma^2}\left((n_1 - 1){S_X}^2 + (n_2 - 1){S_Y}^2\right) \sim \chi^2_{n_1+n_2-2}$$

suy ra

$$T = \frac{U}{\sqrt{\frac{\chi^2}{n_1 + n_2 - 2}}}$$

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

Ta chọn thống kê

$$G = T = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{A} \sim T_{(n_1+n_2-2)}$$

# 2.3 ƯỚC LƯỢNG HIỆU CỦA HAI KỲ VỌNG TOÁN CỦA HAI BIẾN NGẪU NHIÊN PHÂN PHỐI CHUẨN

với

$$A = \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \sqrt{\frac{(n_1 - 1){S_X}^2 + (n_2 - 1){S_Y}^2}{n_1 + n_2 - 2}}$$

với độ tin cậy $(1 - \alpha)$ cho trước có thể tìm được cặp giá trị $\alpha_1, \alpha_2$ sao cho $\alpha_1 + \alpha_2 = \alpha$ và tương ứng ta tìm được cặp giá trị $t_{1-\alpha_1}^{(n_1+n_2-2)}, t_{\alpha_2}^{(n_1+n_2-2)}$ thoả mãn:

$$P\left(G < t_{1-\alpha_1}^{(n_1+n_2-2)}\right) = \alpha_1$$

$$P\left(G > t_{\alpha_2}^{(n_1+n_2-2)}\right) = \alpha_2$$

suy ra

$$P\left(t_{1-\alpha_1}^{(n_1+n_2-2)} < G < t_{\alpha_2}^{(n_1+n_2-2)}\right) = 1 - (\alpha_1 + \alpha_2) = 1 - \alpha$$

## 2.3 ƯỚC LƯỢNG HIỆU CỦA HAI KỲ VỌNG TOÁN CỦA HAI BIẾN NGẪU NHIÊN PHÂN PHỐI CHUẨN

do $t_{1-\alpha_1}^{(n_1+n_2-2)} = -t_{\alpha_1}^{(n_1+n_2-2)}$ nên ta có:

$$P\left(-t_{\alpha_1}^{(n_1+n_2-2)} < G < t_{\alpha_2}^{(n_1+n_2-2)}\right) = 1-\alpha$$

$$\Leftrightarrow P\left(-t_{\alpha_1}^{(n_1+n_2-2)} < \frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{A} < t_{\alpha_2}^{(n_1+n_2-2)}\right) = 1-\alpha$$

$$\Leftrightarrow P\left(\bar{X}-\bar{Y}-A.t_{\alpha_2}^{(n_1+n_2-2)} < \mu_1-\mu_2 < \bar{X}-\bar{Y}+A.t_{\alpha_1}^{(n_1+n_2-2)}\right) = 1-\alpha$$



Vậy với độ tin cậy $(1-\alpha)$ tham số $\mu_1-\mu_2$ của biến ngẫu nhiên gốc sẽ nằm trong khoảng



$$\left(\bar{X}-\bar{Y}-A.t_{\alpha_2}^{(n_1+n_2-2)}; \bar{X}-\bar{Y}+A.t_{\alpha_1}^{(n_1+n_2-2)}\right)$$

- Khoảng tin cậy đối xứng: $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$ ta có khoảng tin cậy

$$(\bar{X} - \bar{Y} - \varepsilon, \bar{X} - \bar{Y} + \varepsilon)$$

với $\varepsilon = A.t_{\alpha/2}^{(n_1+n_2-2)}$ được gọi là độ chính xác.

- Khoảng tin cậy bên phải: $\alpha_1 = 0, \alpha_2 = \alpha$ thì $u_{\alpha_1} = u_0 = +\infty$ và do đó khoảng tin cậy là

$$\left(\bar{X} - \bar{Y} - A.t_{\alpha}^{(n_1+n_2-2)}, +\infty\right)$$

Biểu thức trên dùng để ước lượng giá trị tối thiểu của $\mu_1 - \mu_2$.

- Khoảng tin cậy bên trái: $\alpha_1 = \alpha, \alpha_2 = 0$ thì $u_{\alpha_2} = u_0 = +\infty$ và do đó khoảng tin cậy là

$$\left(-\infty, \bar{X} - \bar{Y} + A.t_{\alpha}^{(n_1+n_2-2)}\right)$$

Biểu thức trên dùng để ước lượng giá trị tối đa của $\mu_1 - \mu_2$.

## 2.3 ƯỚC LƯỢNG HIỆU CỦA HAI KỲ VỌNG TOÁN CỦA HAI BIẾN NGẪU NHIÊN PHÂN PHỐI CHUẨN

Từ các mẫu cụ thể

$$w_X = (x_1, x_2, \dots, x_{n_1})$$
$$w_Y = (y_1, y_2, \dots, y_{n_2})$$

và từ đó ta tìm được giá trị cụ thể của các trung bình mẫu $\bar{x}, \bar{y}$. Lúc đó với độ tin cậy $(1-\alpha)$, qua các mẫu cụ thể, khoảng tin cậy của $\mu_1 - \mu_2$ là:

$$(g_1, g_2) = \left(\bar{x} - \bar{y} - A.t_{\alpha_2}^{(n_1+n_2-2)}; \bar{x} - \bar{y} + A.t_{\alpha_1}^{(n_1+n_2-2)}\right)$$



Chú ý:



Do phân phối Student hội tụ nhanh về phân phối chuẩn khi $n \to +\infty$ nên khi $n_1 + n_2$ rất lớn $(n_1 + n_2 > 30)$ thì xấp xỉ $t_\alpha^{(n_1+n_2-2)} \approx u_\alpha$.

## 2.3 ƯỚC LƯỢNG HIỆU CỦA HAI KỲ VỌNG TOÁN CỦA HAI BIẾN NGẪU NHIÊN PHÂN PHỐI CHUẨN

Ví dụ 4. Từ một chuồng nuôi lợn, người ta lấy ngẫu nhiên bốn con đem cân và thu được trọng lượng tương ứng của chúng là 64, 66, 89, 77. Từ chuồng khác lấy ra ba con để cân thì được trọng lượng là 56, 71, 53. Với độ tin cậy 95% hãy ước lượng sự khác biệt về trọng lượng trung bình của hai chuồng lợn đó. Giả thiết trọng lượng của lợn tuân theo quy luật phân phối chuẩn.

## 2.4 ƯỚC LƯỢNG XÁC SUẤT (TỈ LỆ) CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬT KHÔNG-MỘT

Giả sử trong tổng thể, biến ngẫu nhiên gốc X tuân theo quy luật không-một

$$P(X=1)=p, P(X=0)=1-p=q$$

Từ tổng thể lập mẫu ngẫu nhiên kích thước n: $W_n = (X_1, X_2, \dots, X_n)$,



$$\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i = f$$

đặt $m=\sum_{i=1}^{n}X_i \sim B(p;n)$ nên $E(m)=np, D(m)=npq$, suy ra $E\bar{X}=E\left(\frac{m}{n}\right)=p, D\bar{X}=D\left(\frac{m}{n}\right)=\frac{pq}{n}$



Khi n lớn và p không quá nhỏ thì

$$U=\frac{(\bar{X}-p)}{\sqrt{pq}}\sqrt{n}\sim N(0,1)$$

## 2.4 ƯỚC LƯỢNG XÁC SUẤT (TỈ LỆ) CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬT KHÔNG-MỘT

Chọn thống kê

$$G = U = \frac{(\bar{X} - p)}{\sqrt{pq}}\sqrt{n} \sim N(0,1)$$

tương tự ta có với độ tin cậy $(1-\alpha)$ cho trước có thể tìm được cặp giá trị $\alpha_1, \alpha_2$ sao cho $\alpha_1 + \alpha_2 = \alpha$ và tương ứng ta tìm được cặp giá trị $u_{1-\alpha_1}, u_{\alpha_2}$ thoả mãn:



$$P\left(G < u_{1-\alpha_1}\right) = \alpha_1$$



$$P\left(G > u_{\alpha_2}\right) = \alpha_2$$

suy ra

$$P\left(u_{1-\alpha_1} < U < u_{\alpha_2}\right) = 1 - (\alpha_1 + \alpha_2) = 1 - \alpha$$

## 2.4 ƯỚC LƯỢNG XÁC SUẤT (TỈ LỆ) CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬT KHÔNG-MỘT

do $u_{1-\alpha_1} = -u_{\alpha_1}$ nên ta có:

$$P\left(-u_{\alpha_1} < U < u_{\alpha_2}\right) = 1 - \alpha$$

$$\Leftrightarrow P\left(-u_{\alpha_1} < \frac{(\bar{X} - p)}{\sqrt{pq}}\sqrt{n} < u_{\alpha_2}\right) = 1 - \alpha$$



$$\Leftrightarrow P\left(\bar{X} - \frac{\sqrt{pq}}{\sqrt{n}}u_{\alpha_2} < p < \bar{X} + \frac{\sqrt{pq}}{\sqrt{n}}u_{\alpha_1}\right) = 1 - \alpha$$

Vậy với độ tin cậy $(1-\alpha)$ tham số p của biến ngẫu nhiên gốc sẽ nằm trong khoảng

$$\left(\bar{X} - \frac{\sqrt{pq}}{\sqrt{n}}u_{\alpha_2}, \bar{X} + \frac{\sqrt{pq}}{\sqrt{n}}u_{\alpha_1}\right)$$

# 2.4 ƯỚC LƯỢNG XÁC SUẤT (TỈ LỆ) CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬT KHÔNG-MỘT

nhưng do n lớn nên ta xấp xỉ p bởi: $p \approx f = \frac{m}{n}$ và do $\bar{X} = f$ nên ta có thể xấp xỉ khoảng ước lượng của p là

$$\left(f - \frac{\sqrt{f(1-f)}}{\sqrt{n}} u_{\alpha_2}, f + \frac{\sqrt{f(1-f)}}{\sqrt{n}} u_{\alpha_1}\right)$$



- Khoảng tin cậy đối xứng: $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$ ta có khoảng tin cậy

$$(f - \varepsilon, f + \varepsilon)$$



với $\varepsilon = \frac{\sqrt{f(1-f)}}{\sqrt{n}} u_{\alpha/2}$ được gọi là độ chính xác.

## 2.4 ƯỚC LƯỢNG XÁC SUẤT (TỈ LỆ) CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬT KHÔNG-MỘT

- Khoảng tin cậy bên phải: $\alpha_1 = 0, \alpha_2 = \alpha$ thì $u_{\alpha_1} = u_0 = +\infty$ và do đó khoảng tin cậy là

$$\left(f - \frac{\sqrt{f(1-f)}}{\sqrt{n}} u_\alpha, +\infty\right)$$



Biểu thức trên dùng để ước lượng giá trị tối thiểu của p.

- Khoảng tin cậy bên trái: $\alpha_1 = \alpha, \alpha_2 = 0$ thì $u_{\alpha_2} = u_0 = +\infty$ và do đó khoảng tin cậy là



$$\left(-\infty, f + \frac{\sqrt{f(1-f)}}{\sqrt{n}} u_\alpha\right)$$

Biểu thức trên dùng để ước lượng giá trị tối đa của p.

## 2.4 ƯỚC LƯỢNG XÁC SUẤT (TỈ LỆ) CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬT KHÔNG-MỘT

Ví dụ Hãy ước lượng tỉ lệ chính phẩm của một nhà máy bằng khoảng tin cậy đối xứng với độ tin cậy 0,95 nếu biết rằng kiểm tra 100 sản phẩm của nhà máy thì thấy có 10 phế phẩm.

Ví dụ Để ước lượng số cá có trong một hồ, người ta bắt 2000 con lên đánh dấu rồi thả trở lại. Vài hôm sau, bắt 400 con lên thì thấy có 80 con cá bị đánh dấu. Hãy ước lượng số cá có trong hồ với độ tin cậy 95%.

## 2.5 ƯỚC LƯỢNG HIỆU HAI THAM SỐ p CỦA HAI BIẾN NGẪU NHIÊN PHÂN PHỐI KHÔNG-MỘT

Giả sử ta xét một lúc hai tổng thể. Ở tổng thể thứ nhất ta xét biến ngẫu nhiên gốc $X$ tuân theo quy luật không-một với

$$P(X = 1) = p_1, P(X = 0) = 1 - p_1 = q_1$$

ở tổng thể thứ hai ta xét biến ngẫu nhiên gốc $Y$ tuân theo quy luật không-một với

$$P(Y = 1) = p_2, P(X = 0) = 1 - p_2 = q_2$$

Ta cần ước lượng hiệu $p_1 - p_2$ với độ tin cậy $1 - \alpha$

Từ hai tổng thể nói trên rút ra hai mẫu ngẫu nhiên độc lập có kích thước tương ứng $n_1$ và $n_2$:

$$W_X = (X_1, X_2, \dots, X_{n_1})$$

$$W_Y = (Y_1, Y_2, \dots, Y_{n_2})$$

## 2.5 ƯỚC LƯỢNG HIỆU HAI THAM SỐ p CỦA HAI BIẾN NGẪU NHIÊN PHÂN PHỐI KHÔNG-MỘT

Xét $\bar{X} - \bar{Y}$, với $\bar{X} = \frac{1}{n_1}\sum_{i=1}^{n_1} X_i = \frac{m_1}{n_1} = f_1$, $\bar{Y} = \frac{1}{n_2}\sum_{i=1}^{n_2} Y_i = \frac{m_2}{n_2} = f_2$. Do $E(\bar{X} - \bar{Y}) = E\bar{X} - E\bar{Y} = p_1 - p_2$, $D(\bar{X} - \bar{Y}) = D\bar{X} + D\bar{Y} = \frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}$ suy ra



$$U = \frac{\bar{X} - \bar{Y} - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} \sim N(0,1)$$

Ta chon thống kê



$$G = U = \frac{\bar{X} - \bar{Y} - (p_1 - p_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} \sim N(0,1)$$

## 2.5 ƯỚC LƯỢNG HIỆU HAI THAM SỐ p CỦA HAI BIẾN NGẪU NHIÊN PHÂN PHỐI KHÔNG-MỘT

Thực hiện tương tự ta có

$$P\left(\bar{X}-\bar{Y}-\sqrt{\frac{Cq_1}{n_1}+\frac{p_2q_2}{n_2}}u_{\alpha/2}<p_1-p_2\right.$$
$$\left.<\bar{X}-\bar{Y}+\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}u_{\alpha/2}\right)=1-\alpha$$

Khoảng tin cậy là: 

$$\left(\bar{X}-\bar{Y}-\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}u_{\alpha/2};\bar{X}-\bar{Y}+\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}u_{\alpha/2}\right)$$

Khi $n_1, n_2$ lớn ta xấp xỉ: $p_1 \approx f_1, p_2 \approx f_2$ ta có khoảng UL đối xứng:



$$\left(f_1-f_2-\sqrt{\frac{f_1(1-f_1)}{n_1}+\frac{f_2(1-f_2)}{n_2}}u_{\alpha/2};f_1-f_2\right.$$
$$\left.+\sqrt{\frac{f_1(1-f_1)}{n_1}+\frac{f_2(1-f_2)}{n_2}}u_{\alpha/2}\right)$$

## 2.5 ƯỚC LƯỢNG HIỆU HAI THAM SỐ p CỦA HAI BIẾN NGẪU NHIÊN PHÂN PHỐI KHÔNG-MỘT

ta có khoảng UL bên phải:

$$\left(f_1 - f_2 - \sqrt{\frac{f_1(1-f_1)}{n_1} + \frac{f_2(1-f_2)}{n_2}}\,u_\alpha ; +\infty\right)$$

khoảng UL bên trái:

$$\left(-\infty; f_1 - f_2 - \sqrt{\frac{f_1(1-f_1)}{n_1} + \frac{f_2(1-f_2)}{n_2}}\,u_\alpha\right)$$

## 2.5 ƯỚC LƯỢNG HIỆU HAI THAM SỐ p CỦA HAI BIẾN NGẪU NHIÊN PHÂN PHỐI KHÔNG-MỘT

Ví dụ Doanh nghiệp dự định đưa sản phẩm của mình vào hai thị trường khác nhau. Bán thử sản phẩm cho 100 khách hàng tiềm năng của thị trường thư nhất thì có 50 người mua. Còn với thị trường thứ hai bán thử sản phẩm cho 50 khách hàng thì có 20 người mua. Với độ tin cậy 95% hãy ước lượng mức độ chênh lệch về thị phần mà doanh nghiệp có thể đạt được tại hai thị trường đó.

## 2.6 ƯỚC LƯỢNG PHƯƠNG SAI CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬT PHÂN PHỐI CHUẨN

Giả sử trong tổng thể xét biến ngẫu nhiên gốc $X \sim N(\mu, \sigma^2)$ nhưng chưa biết phương sai $\sigma^2$ của nó. Để ước lượng $\sigma^2$ từ tổng thể ta lập mẫu ngẫu nhiên kích thước n

$$W = (X_1, X_2, \dots, X_n)$$

Để chọn thống kê G thích hợp ta xét hai trường hợp sau:



a, Đã biết kỳ vọng toán μ của biến ngẫu nhiên gốc.

Chọn

$$G = \chi^2 = \frac{nS^{*2}}{\sigma^2}$$



với $S^{*2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$. Suy ra $\frac{nS^{*2}}{\sigma^2} = \sum_{i=1}^{n}(\frac{X_i - \mu}{\sigma})^2$, do $X_i \sim N(\mu, \sigma^2)$ nên $\frac{X_i - \mu}{\sigma} \sim N(0,1)$. Vậy

$$G = \chi^2 = \frac{nS^{*2}}{\sigma^2} \sim \chi^2(n)$$

## 2.6 ƯỚC LƯỢNG PHƯƠNG SAI CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬT PHÂN PHỐI CHUẨN

Do đó, với độ tin cậy $(1-\alpha)$ cho trước có thể tìm được cặp giá trị $\alpha_1, \alpha_2$ sao cho $\alpha_1 + \alpha_2 = \alpha$ và từ đó tìm được hai giá trị tới hạn khi bình phương tương ứng là $\chi^{2(n)}_{1-\alpha_1}, \chi^{2(n)}_{\alpha_2}$ thỏa mãn điều kiện

$$P\left(\chi^2 < \chi^{2(n)}_{1-\alpha_1}\right) = \alpha_1$$

$$P\left(\chi^2 > \chi^{2(n)}_{\alpha_2}\right) = \alpha_2$$

Do đó

$$P\left(\chi^{2(n)}_{1-\alpha_1} < \chi^2 < \chi^{2(n)}_{\alpha_2}\right) = 1 - (\alpha_1 + \alpha_2) = 1 - \alpha$$

Biến đổi tương đương ta có

$$P\left(\frac{nS^{*2}}{\chi^{2(n)}_{\alpha_2}} < \sigma^2 < \frac{nS^{*2}}{\chi^{2(n)}_{1-\alpha_1}}\right) = 1 - \alpha$$

## 2.6 ƯỚC LƯỢNG PHƯƠNG SAI CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬT PHÂN PHỐI CHUẨN

Như vậy, với độ tin cậy $(1-\alpha)$ khoảng tin cậy của $\sigma^2$ là

$$\left(\frac{nS^{*2}}{\chi^{2(n)}_{\alpha_2}}; \frac{nS^{*2}}{\chi^{2(n)}_{1-\alpha_1}}\right)$$

Từ khoảng tin cậy tổng quát ta có thể xây dựng được các khoảng tin cậy cụ thể sau:

- Nếu $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$ thì khoảng tin cậy có dạng:

$$\left(\frac{nS^{*2}}{\chi^{2(n)}_{\alpha/2}}; \frac{nS^{*2}}{\chi^{2(n)}_{1-\alpha/2}}\right)$$

## 2.6 ƯỚC LƯỢNG PHƯƠNG SAI CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬT PHÂN PHỐI CHUẨN

- Nếu $\alpha_1 = 0, \alpha_2 = \alpha$ thì khoảng tin cậy có dạng:

$$\left(\frac{nS^{*2}}{\chi_{\alpha}^{2(n)}};\ +\infty\right)$$

- Nếu $\alpha_1 = \alpha, \alpha_2 = 0$ thì khoảng tin cậy có dạng:



$$\left(0;\ \frac{nS^{*2}}{\chi_{1-\alpha}^{2(n)}}\right)$$

Với một mẫu cụ thể $w = (x_1, x_2, \ldots, x_n)$ có thể xác định được khoảng tin cậy cụ thể bằng số của $\sigma^2$ giống như đã làm ở các phần trên.

## 2.6 ƯỚC LƯỢNG PHƯƠNG SAI CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬT PHÂN PHỐI CHUẨN

Ví dụ. Mức hao phí nguyên liệu cho một đơn vị sản phẩm là biến ngẫu nhiên phân phối chuẩn với trung bình là 20 gam. Để ước lượng mức độ phân tán của mức hao phí này người ta cân thử 25 sản phẩm và thu được kết quả:

| | | | |
|---|---|---|---|
| *Hao phí nguyên liệu(gam)* | *19,5* | *20,0* | *20,5* |
| *Số sản phẩm tương ứng* | *5* | *18* | *2* |

Với độ tin cậy 95% hãy ước lượng $\sigma^2$ với $\alpha_1 = \alpha_2$.

## 2.6 ƯỚC LƯỢNG PHƯƠNG SAI CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬT PHÂN PHỐI CHUẨN

b, Chưa biết kỳ vọng toán μ của biến ngẫu nhiên gốc.

Chọn

$$G = \chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

Do đó, với độ tin cậy $(1-\alpha)$ cho trước có thể tìm được cặp giá trị $\alpha_1, \alpha_2$ sao cho $\alpha_1 + \alpha_2 = \alpha$ và từ đó tìm được hai giá trị tới hạn khi bình phương tương ứng là $\chi^{2(n-1)}_{1-\alpha_1}, \chi^{2(n-1)}_{\alpha_2}$ thỏa mãn điều kiện

$$P\left(\chi^2 < \chi^{2(n-1)}_{1-\alpha_1}\right) = \alpha_1$$

$$P\left(\chi^2 > \chi^{2(n-1)}_{\alpha_2}\right) = \alpha_2$$

# 2.6 ƯỚC LƯỢNG PHƯƠNG SAI CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬT PHÂN PHỐI CHUẨN

Do đó

$$P\left(\chi_{1-\alpha_1}^{2(n-1)} < \chi^2 < \chi_{\alpha_2}^{2(n-1)}\right) = 1 - (\alpha_1 + \alpha_2) = 1 - \alpha$$

Biến đổi tương đương ta có

$$P\left(\frac{nS^2}{\chi_{\alpha_2}^{2(n-1)}} < \sigma^2 < \frac{nS^2}{\chi_{1-\alpha_1}^{2(n-1)}}\right) = 1 - \alpha$$

Như vậy, với độ tin cậy $(1-\alpha)$ khoảng tin cậy của $\sigma^2$ là

$$\left(\frac{nS^2}{\chi_{\alpha_2}^{2(n-1)}}; \frac{nS^2}{\chi_{1-\alpha_1}^{2(n-1)}}\right)$$

## 2.6 ƯỚC LƯỢNG PHƯƠNG SAI CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬT PHÂN PHỐI CHUẨN

Từ khoảng tin cậy tổng quát ta có thể xây dựng được các khoảng tin cây cụ thể sau:

- Nếu $\alpha_1 = \alpha_2 = \alpha$ thì khoảng tin cậy có dạng:

$$\left(\frac{nS^2}{\chi^{2(n-1)}_{\alpha/2}}; \frac{nS^2}{\chi^{2(n-1)}_{1-\alpha/2}}\right)$$

- Nếu $\alpha_1 = 0, \alpha_2 = \alpha$ thì khoảng tin cậy có dạng:

$$\left(\frac{nS^2}{\chi^{2(n-1)}_{\alpha}}; +\infty\right)$$

- Nếu $\alpha_1 = 0, \alpha_2 = \alpha$ thì khoảng tin cậy có dạng:

$$\left(0; \frac{nS^2}{\chi^{2(n-1)}_{1-\alpha}}\right)$$

## 2.6 ƯỚC LƯỢNG PHƯƠNG SAI CỦA BIẾN NGẪU NHIÊN TUÂN THEO QUY LUẬT PHÂN PHỐI CHUẨN

Với một mẫu cụ thể $w = (x_1, x_2, \dots, x_n)$ có thể xác định được khoảng tin cậy cụ thể bằng số của $\sigma^2$ giống như đã làm ở các phần trên.

Ví dụ. Cùng giả thiết như ví dụ trên nhưng giả sử ta chưa biết giá trị trung bình.