**SUY DIỄN THỐNG KÊ CHO HAI MẪU**

**Ước lượng hiệu hai kì vọng toán của hai biến ngẫu nhiên phân phối chuẩn**

Giả sử có hai tổng thể nghiên cứu, trong đó các biến ngẫu nhiên *X*1 và *X*2 cùng phân phối chuẩn với các tham số đặc trưng tương ứng là  và  với  chưa biết. Để ước lượng sự chênh lệch  giữa hai kỳ vọng toán.

Từ hai tổng thể lập hai mẫu ngẫu nhiên độc lập kích thước tương ứng là *n1* và *n*2­



Từ đó ta tìm được các thống kê đặc trưng mẫu tương ứng là và . Để chọn thống kê *G* thích hợp xét các trường hợp sau

***Trường hợp 1. Nếu đã biết phương sai  của các tổng thể***. Khi đó

 có phân phối chuẩn tắc *N*(0,1). Ta có



**Khoảng tin cậy đối xứng** của  với độ tin cậy  là



**Khoảng tin cậy bên phải**: 

**Khoảng tin cậy bên trái**: 

***Trường hợp 2. Nếu* *chưa biết các phương sai  của các tổng thể song giả thiết rằng ***

Khi đó  trong đó .

**Khoảng tin cậy đối xứng** của  với độ tin cậy  là



**Khoảng tin cậy bên phải**: 

**Khoảng tin cậy bên trái** : 

***Trường hợp 3. Nếu chưa biết các phương sai  của các tổng thể và không có căn cứ để cho rằng chúng bằng nhau***

Khi đó  có phân phối Student với số bậc tự do là 

Với 

Khoảng ước lượng của  với độ tin cậy  là



**Ví dụ**

Từ một chuồng nuôi lợn lấy ngẫu nhiên bốn con đem cân và thu được trọng lượng tương ứng của chúng là 64, 66, 89 và 77kg. Từ một chuồng khác lấy ra ba con đem cân thu được trọng lượng là 56, 71kg và 53kg. 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 có phân phối chuẩn.

**Giải**

Gọi  tương ứng là trọng lượng của lợn ở hai chuồng nói trên. Theo giả thiết  phân phối chuẩn. Gọi trọng lượng trung bình lần lượt là  Đây là bài toán ước lượng hiệu số  khi chưa biết phương sai của tổng thể.

Nếu có thể cho rằng phương sai của chúng bằng nhau (chẳng hạn cả hai chuồng cùng nuôi một giống lợn và được chăm sóc như nhau), từ đó suy ra khoảng tin cậy đối xứng như sau:



Từ hai mẫu cụ thể ta có



**Từ đó** 

Ta có 

Do đó 

hay 

**Ước lượng hiệu tỉ lệ của 2 tổng thể**

Giả sử có hai tổng thể nghiên cứu, trong đó các biến ngẫu nhiên *X*1 và *X*2 cùng phân phối. Giả sử các tỷ lệ  của 2 tổng thể chưa biết, với độ tin cậy  tìm khoảng tin cậy cho .

Từ hai tổng thể trên rút ra hai mẫu ngẫu nhiên độc lập kích thước *n*1 và *n*2 và. Xét thống kê

 với 

Khi thì biến ngẫu nhiên *G* sẽ phân phối xấp xỉ *N*(0,1).

Khi *n*1, *n*2 đủ lớn đủ lớn ta xấp xỉ

 (\*)

**Khoảng tin cậy đối xứng** mức  của hiệu  như sau



**Khoảng tin cậy bên phải**: 

**Khoảng tin cậy bên trái** : 

**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 thứ nhất thì có 50 người mua. Còn với thị trường thứ hai, khi bán thử sản phẩm cho 50 khách hàng tiềm năng 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

**Giải**

Gọi  tương ứng là thị phần mà doanh nghiệp có thể đạt được ở hai thị trường. Vậy mức độ chênh lệch về thị phần là .

Với hai mẫu cụ thể



Với 

Vậy 

hay  với độ tin cậy 0,95.

**Kiểm định so sánh hai trung bình**

**Bài toán:** Giả sử có hai tổng thể  có có  chưa biết, . Với mức ý nghĩa , kiểm định giả thuyết .

Từ hai tổng thể trên rút ra hai mẫu độc lập



**Trường hợp 1. ** đã biết. Khi đó nếu  đúng thì

nếucó phân phối chuẩn hoặc  nếu  không có giả thiết này.

Do đó miền bác bỏ cho các dạng bài toán tương ứng như sau



**Trường hợp 2. ** chưa biết. Khi đó nếu  đúng thì thống kê .

Do đó miền bác bỏ cho các dạng bài toán tương ứng như sau



**Trường hợp 3. , ** có phân phối chuẩn,  đã biết. Khi đó nếu  đúng thì .

Do đó miền bác bỏ cho các dạng bài toán tương ứng như sau



**Trường hợp 4. , ** có phân phối chuẩn,  chưa biết nhưng giả định . Khi đó nếu  đúng thì thống kê .

Với 

Do đó miền bác bỏ cho các dạng bài toán tương ứng như sau



***Chú ý.*** *Phương pháp kiểm định trình bày ở trên còn có thể áp dụng khi  và  khác nhau. Chẳng hạn cần kiểm định giả thuyết  với  là giá trị định trước cũng được tiến hành tương tự như trên.*





**Ví dụ.** Tại một xí nghiệp người ta xây dựng hai phương án gia công cùng một loại chi tiết. Để đánh giá xem chi phí trung bình về nguyên liệu theo hai phương án ấy có khác nhau hay không người ta tiến hành sản xuất thử và thu được kết quả như sau:

Phương án 1: 2,5 3,2 3,5 3,8 3,5

Phương án 2: 2,0 2,7 2,5 2,9 2,3 2,6

Với mức ý nghĩa  hãy kết luận về vấn đề trên biết rằng chi phí nguyên liệu theo cả hai phương án gia công đều là các biến ngẫu nhiên phân phối chuẩn với 

**Giải**

Gọi tương ứng là chi phí nguyên liệu theo hai phương án gia công trên . Theo giả thiết  phân phối chuẩn.

Gọi chi phí nguyên liệu trung bình theo các phương án đó là .

Ta có bài toán kiểm định 

Do 

Từ mẫu cụ thể ta tính được





Do đó ta bác bỏ , thừa nhận  tức là chi phí nguyên liệu theo hai phương án đó thực sự khác nhau.

**Kiểm định so sánh hai tỉ lệ**

**Bài toán:** Gỉa sử từ hai tổng thể  có cùng đặc trưng định tính A.  có tỷ lệ  về đặc tính A,  có tỷ lệ  về đặc tính A. Hãy kiểm định giả thiết sau  với mức ý nghĩa .

Từ hai tổng thể rút ra hai mẫu ngẫu nhiên độc lập khích thước tương ứng là *n*1 và *n*2:



Khi và nếu *H*0 đúng 

 có phân phối xấp xỉ *N*(0,1).

Thông thường *p* chưa biết nên thay bằng ước lượng của nó là  với  đủ lớn.

Do đó ta có  phân phối xấp xỉ *N*(0,1) nếu . Do đó với mức ý nghĩa  cho trước và tùy thuộc dạng của đối thuyết  miền bác bỏ  được xây dựng theo các trường hợp sau:



**Ví dụ**

Kiểm tra ngẫu nhiên các sản phẩm cùng loại do hai nhà máy sản xuất được các số liệu sau:

Nhà máy Số sản phẩm được kiểm tra Số phế phẩm

**A**  

**B**  

Với mức ý nghĩa  có thể coi tỷ lệ phế phẩm của hai nhà máy như nhau được không?

**Giải**

Gọi  tương ứng là tỷ lệ phế phẩm của hai nhà máy A và B.

Cặp giả thuyết





Với hai mẫu cụ thể ta có





Vậy chưa có cơ sở để bác bỏ , tức là có thể coi tỷ lệ phế phẩm ở hai nhà máy là như nhau.