

Chương 6 Kiểm định giả thuyết thống kê

Nguyễn Thị Mộng Ngọc
University of Science, VNU - HCM
ngtmmngoc@hcmus.edu.vn

Các khái niệm
trong kiểm
định giả
thuyết thống
kê

Giả thuyết H_0 và đối
thuyết H_1
Các loại sai lầm
Giá trị p -value

Kiểm định giả
thuyết -
Trường hợp
một mẫu

Kiểm định giả thuyết
cho kỳ vọng
Kiểm định giả thuyết
cho tỷ lệ

Kiểm định giả
thuyết -
Trường hợp
hai mẫu

Kiểm định giả thuyết
so sánh hai trung
bình - Trường hợp hai
mẫu độc lập

Kiểm định giả thuyết
so sánh hai trung
bình - Trường hợp hai
mẫu không độc lập

Kiểm định giả thuyết
so sánh hai tỷ lệ

Kiểm định giả thuyết thống kê

Định nghĩa 1

Giả thuyết thống kê là những giả thuyết nói về các tham số, dạng quy luật phân phối, hoặc tính độc lập của các biến ngẫu nhiên.

Việc tìm ra kết luận bác bỏ hay không bác bỏ một giả thuyết gọi là **kiểm định giả thuyết thống kê**.

Ví dụ: Một nhà sản xuất cho rằng khối lượng trung bình của 1 gói mì Omachi là 75 gam. Để kiểm tra ý kiến này là đúng hay sai, chọn ngẫu nhiên một số gói mì Omachi để tiến hành kiểm tra, tính toán, ...

Các khái niệm
trong kiểm
định giả
thuyết thống
kê

Giả thuyết H_0 và đối
thuyết H_1
Các loại sai lầm
Giá trị p -value

Kiểm định giả
thuyết -
Trường hợp
một mẫu

Kiểm định giả thuyết
cho kỳ vọng
Kiểm định giả thuyết
cho tỷ lệ

Kiểm định giả
thuyết -
Trường hợp
hai mẫu

Kiểm định giả thuyết
so sánh hai trung
bình - Trường hợp hai
mẫu độc lập

Kiểm định giả thuyết
so sánh hai trung
bình - Trường hợp hai
mẫu không độc lập

Kiểm định giả thuyết
so sánh hai tỷ lệ

Kiểm định giả thuyết thống kê

Ví dụ 1

Trong một báo cáo nói rằng: thu nhập bình quân của những người làm trong ngành thư viện ở Việt Nam là 7 triệu đồng một tháng thì ta có thể coi đó là một giả thuyết thống kê, giả thuyết này nói về một tham số (kỳ vọng) của biến ngẫu nhiên X biểu thị mức lương của những người làm trong ngành thư viện. Dựa vào số liệu của một mẫu điều tra về thu nhập và quy tắc kiểm định để đưa một kết luận là bác bỏ hay chấp nhận giả thuyết nói trên.

Các khái niệm
trong kiểm
định giả
thuyết thống
kê

Giả thuyết H_0 và đối
thuyết H_1
Các loại sai lầm
Giá trị p -value

Kiểm định giả
thuyết -
Trường hợp
một mẫu

Kiểm định giả thuyết
cho kỳ vọng
Kiểm định giả thuyết
cho tỷ lệ

Kiểm định giả
thuyết -
Trường hợp
hai mẫu

Kiểm định giả thuyết
so sánh hai trung
bình - Trường hợp hai
mẫu độc lập

Kiểm định giả thuyết
so sánh hai trung
bình - Trường hợp hai
mẫu không độc lập

Kiểm định giả thuyết
so sánh hai tỷ lệ

Giả thuyết H_0 và đối thuyết H_1

Trong bài toán kiểm định giả thuyết,

- giả thuyết cần được kiểm định gọi là **giả thuyết không (null hypothesis)**, kí hiệu H_0 ;
- mệnh đề đối lập với H_0 gọi là **đối thuyết (alternative hypothesis)**, kí hiệu H_1 .

Khi xây dựng giả thuyết,

- Khi xây dựng H_0 , trong cấu trúc của H_0 **luôn luôn có dấu "="**, có thể là dấu "=" hoặc " \leq " hoặc " \geq ".
- Khi xây dựng H_1 , trong cấu trúc của H_1 **không được có dấu "="**, có thể là dấu " \neq " hoặc "<" hoặc ">" tùy bài toán.

$$\begin{cases} H_0 : p \leq 0.05 & \text{Tỷ lệ sản phẩm kém ở mức chấp nhận được} \\ H_1 : p > 0.05 & \text{Tỷ lệ sản phẩm kém cao hơn mức cho phép} \end{cases}$$

$$\begin{cases} H_0 : \theta \leq \theta_0 \\ H_1 : \theta > \theta_0 \end{cases}$$

Kiểm định hai phía	Kiểm định 1 phía bên trái	Kiểm định 1 phía bên phải
$H_0: \theta = \theta_0$ $H_1: \theta \neq \theta_0$	$H_0: \theta \geq \theta_0$ $H_1: \theta < \theta_0$	$H_0: \theta \leq \theta_0$ $H_1: \theta > \theta_0$

Ví dụ 4: Một nhà sản xuất nước giải khát tuyên bố rằng chai chứa nước loại 2 lít chứa trung bình ít nhất là 67,6 ounce nước giải khát. Một mẫu các chai nước chứa loại 2 lít sẽ được chọn ra, và lượng chứa bên trong sẽ được đo lường để kiểm định phát biểu của nhà sản xuất. Trong trường hợp này, ta có thể đặt giả thuyết: $H_0: \mu \geq 67,6$ vs $H_1: \mu < 67,6$.

Sai lầm loại I và sai lầm loại II

Khi kiểm định giả thuyết thống kê, vì chỉ dựa trên những thông tin từ mẫu đang xét để đưa kết luận nên chúng ta có thể mắc phải một trong 2 loại sai lầm sau:

- **Sai lầm loại I:** là sai lầm mắc phải khi ta bác bỏ giả thuyết H_0 trong khi thực tế thì giả thuyết H_0 đúng.
- **Sai lầm loại II:** là sai lầm mắc phải khi ta không bác bỏ giả thuyết H_0 trong khi thực tế thì giả thuyết sai.

Sai lầm loại I và sai lầm loại II

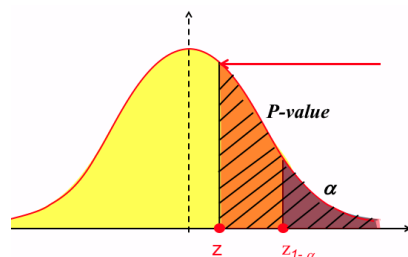
Thực tế \ Quyết định	H_0 đúng	H_0 sai
Không bác bỏ H_0	Không có sai lầm ($1 - \alpha$)	Sai lầm loại II β
Bác bỏ H_0	Sai lầm loại I α	Không có sai lầm ($1 - \beta$)

Trong đó,

- $P(\text{sai lầm loại I xảy ra}) = \alpha$, α chính là mức ý nghĩa của kiểm định và α thường được chọn trong khoảng từ 1% đến 10%.
- $P(\text{sai lầm loại II xảy ra}) = \beta$.
- $(1 - \beta)$ là độ mạnh của kiểm định.

Giá trị p_{value} hay $p_{\text{giá trị}}$

- $p_{\text{giá trị}}$ là giá trị xác suất dùng làm thước đo cho bằng chứng thu được từ mẫu chống lại giả thuyết H_0 .
- Giá trị $p_{\text{giá trị}}$ càng nhỏ càng cho thấy bằng chứng chống lại H_0 .
- $p_{\text{giá trị}}$ là mức ý nghĩa nhỏ nhất dùng để bác bỏ giả thuyết H_0 .



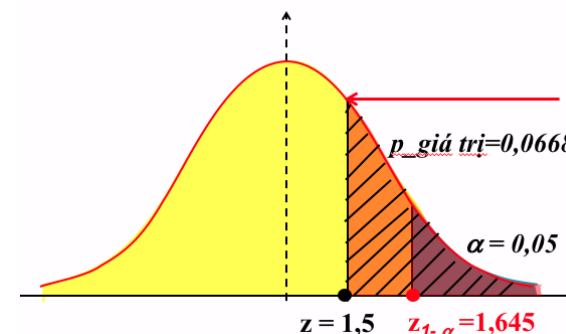
Quy tắc bác bỏ giả thuyết H_0 khi sử dụng $p_{\text{giá trị}}$:

- bác bỏ giả thuyết H_0 khi $p_{\text{giá trị}} < \alpha$;
- không đủ cơ sở bác bỏ giả thuyết H_0 khi $p_{\text{giá trị}} \geq \alpha$;

Ví dụ tính $p_{\text{giá trị}}$

Phương pháp $p_{\text{giá trị}}$ sử dụng giá trị của đại lượng thống kê kiểm định (z hoặc t) để tính một giá trị xác suất gọi là $p_{\text{giá trị}}$ (hay giá trị p).

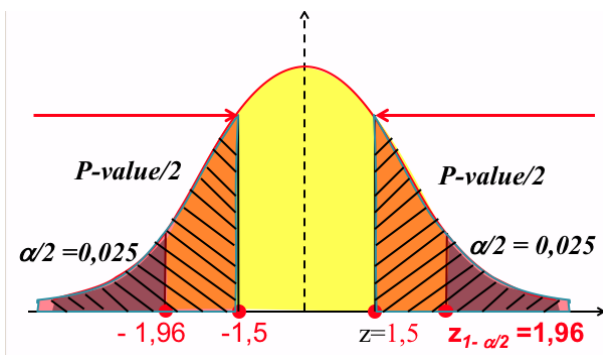
Trong kiểm định 1 phía bên phải, nếu giá trị thống kê của kiểm định $z = 1.5$ thì $p_{\text{giá trị}} = 1 - \Phi(z) = 1 - \Phi(1.5) = 1 - 0,9332 = 0,0668$.



Ví dụ tính p giá trị

Trong kiểm định hai phía, nếu giá trị thống kê của kiểm định $z = 1.5$ thì

$$p_{\text{giá trị}} = 2(1 - \Phi(|z|)) = 2(1 - \Phi(1.5)) = 2(1 - 0,9332) = 0,1336.$$



Các bước thực hiện trong một bài toán kiểm định giả thuyết

- B1: Thiết lập giả thuyết H_0 và đối thuyết H_1 .
- B2: Xác định tiêu chuẩn kiểm định: tính giá trị thống kê của kiểm định nhằm đánh giá dữ liệu mẫu có "thích hợp" với giả thuyết H_0 hay không.
- B3: Chọn mức ý nghĩa α , và xác định miền bác bỏ giả thuyết H_0 . Nếu giá trị kiểm định nằm trong miền này thì giả thuyết H_0 sẽ bị bác bỏ với mức ý nghĩa α .
- B4: Quyết định: rút ra kết luận về mặt thống kê: ở mức ý nghĩa α nào đó, bác bỏ hay không đủ cơ sở để bác bỏ giả thuyết H_0 .

Sau đó, rút kết luận cuối cùng về nội dung bài toán, nhằm trả lời một cách rõ ràng câu hỏi bài toán đặt ra.

Kiểm định giả thuyết về so sánh kỳ vọng với 1 số

- Trường hợp biết phương sai,
- Trường hợp không biết phương sai, mẫu lớn
- Trường hợp không biết phương sai, mẫu nhỏ.

Kiểm định giả thuyết về so sánh kỳ vọng với 1 số

TH 1: biết σ^2

• Các giả định:

- Mẫu ngẫu nhiên X_1, \dots, X_n được chọn từ tổng thể có phân phối chuẩn $\mathcal{N}(\mu, \sigma^2)$ với kỳ vọng μ chưa biết.
- Phương sai σ^2 đã biết.
- Cho trước giá trị μ_0 , cần so sánh kỳ vọng μ với μ_0 .

• Bài toán kiểm định có 3 dạng sau:

Hãy kiểm định một trong ba giả thuyết sau:

$$(a) \begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases} \quad (b) \begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases} \quad (c) \begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$$

với mức ý nghĩa α cho trước.

XSTK
N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê
Giả thuyết H_0 và đối thuyết H_1
Các loại sai lầm
Giá trị p_{value}
Kiểm định giả thuyết - Trường hợp một mẫu
Kiểm định giả thuyết cho kỳ vọng
Kiểm định giả thuyết cho tỷ lệ
Kiểm định giả thuyết - Trường hợp hai mẫu
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập
Kiểm định giả thuyết so sánh hai tỷ lệ

TH 1: biết σ^2

Các bước kiểm định:

- B1: Phát biểu giả thuyết H_0 và đối thuyết H_1
- B2: Xác định tiêu chuẩn kiểm định:
Chọn thống kê

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

làm tiêu chuẩn kiểm định.
Nếu giả thuyết H_0 đúng thì $Z \sim \mathcal{N}(0, 1)$.

Từ mẫu thực nghiệm, tính giá trị thống kê kiểm định:

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

XSTK
N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê
Giả thuyết H_0 và đối thuyết H_1
Các loại sai lầm
Giá trị p_{value}
Kiểm định giả thuyết - Trường hợp một mẫu
Kiểm định giả thuyết cho kỳ vọng
Kiểm định giả thuyết cho tỷ lệ
Kiểm định giả thuyết - Trường hợp hai mẫu
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập
Kiểm định giả thuyết so sánh hai tỷ lệ

TH 1: biết σ^2

Các bước kiểm định (tt):

- B3: Với mức ý nghĩa α , xác định miền bác bỏ: Với mức ý nghĩa α và dựa vào đối thuyết H_1 , xác định miền bác bỏ hoặc tính p -giá trị tương ứng:

Trường hợp	bác bỏ H_0 nếu	p -giá trị
(a) $H_1 : \mu \neq \mu_0$	$ z > z_{1-\alpha/2}$	$2(1 - \Phi(z))$
(b) $H_1 : \mu > \mu_0$	$z > z_{1-\alpha}$	$1 - \Phi(z)$
(c) $H_1 : \mu < \mu_0$	$z < -z_{1-\alpha}$	$\Phi(z)$

Trong đó, $z_{1-\alpha}$ được tìm từ bảng phân phối Gauss.

- B4 Kết luận:

- Nếu bác bỏ giả thuyết H_0 , ta kết luận H_1 đúng với $(1 - \alpha)100\%$ độ tin cậy.
- Ngược lại, ta kết luận chưa đủ cơ sở để bác bỏ giả thuyết H_0 với mức ý nghĩa α .

XSTK
N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê
Giả thuyết H_0 và đối thuyết H_1
Các loại sai lầm
Giá trị p_{value}
Kiểm định giả thuyết - Trường hợp một mẫu
Kiểm định giả thuyết cho kỳ vọng
Kiểm định giả thuyết cho tỷ lệ
Kiểm định giả thuyết - Trường hợp hai mẫu
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập
Kiểm định giả thuyết so sánh hai tỷ lệ

TH 1: biết σ^2

Ví dụ 1: Một hãng sản xuất vỏ xe quảng cáo rằng sản phẩm loại X của hãng có thể sử dụng không dưới 100 ngàn km, độ lệch chuẩn bằng 12 ngàn km. Một công ty vận tải mua 64 vỏ xe loại X, sau một thời gian sử dụng kết quả cho thấy độ bền trung bình là 98,5 ngàn km. Dựa vào thông tin này, hãy kết luận về lời quảng cáo của công ty, với mức ý nghĩa $\alpha = 5\%$.

XSTK
N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê
Giả thuyết H_0 và đối thuyết H_1
Các loại sai lầm
Giá trị p_{value}
Kiểm định giả thuyết - Trường hợp một mẫu
Kiểm định giả thuyết cho kỳ vọng
Kiểm định giả thuyết cho tỷ lệ
Kiểm định giả thuyết - Trường hợp hai mẫu
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập
Kiểm định giả thuyết so sánh hai tỷ lệ

TH 1: biết σ^2

Giải VD1: Theo đề ta có: $n = 64 > 30$, $\sigma = 12$, $\mu_0 = 100$, $\bar{x} = 98,5$, $\alpha = 0,05$. Gọi μ là độ bền trung bình của vỏ xe loại X ở hãng sản xuất này.

- GT:
$$\begin{cases} H_0 : \mu \geq 100 \\ H_1 : \mu < 100 \end{cases}$$
- Giá trị thống kê của kiểm định là :
$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{98,5 - 100}{12/\sqrt{64}} = -1.$$
- Với mức ý nghĩa $\alpha = 0,05$, tra bảng phân phối Gauss ta được $z_\alpha = -z_{1-\alpha} = z_{0,05} = -z_{0,95} = -1,645$.
- Ta thấy, $z = -1 > -1,645 = z_\alpha$ nên ta không đủ cơ sở bác bỏ giả thuyết H_0 với mức ý nghĩa $\alpha = 0,05$.
- Ta có thể kết luận rằng tuổi thọ trung bình của vỏ xe loại X này không thấp hơn 100 ngàn km với mức ý nghĩa 5%. Như vậy, lời quảng cáo của công ty là có thể tin.

<div>XSTK</div> <div>N.T. M. Ngọc</div> <div>Các khái niệm trong kiểm định giả thuyết thông kê</div> <div>Giả thuyết H_0 và đối thuyết H_1</div> <div>Các loại sai lầm</div> <div>Giá trị p_{value}</div> <div>Kiểm định giả thuyết - Trường hợp một mẫu</div> <div>Kiểm định giả thuyết cho kỳ vọng</div> <div>Kiểm định giả thuyết cho tỷ lệ</div> <div>Kiểm định giả thuyết - Trường hợp hai mẫu</div> <div>Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập</div> <div>Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập</div> <div>Kiểm định giả thuyết so sánh hai tỷ lệ</div>	<div>TH 1: biết σ^2</div> <div>Ví dụ 2 : Dây chuyền sản xuất kem đánh răng P/S được thiết kế để đóng hộp những tuýt kem có trọng lượng trung bình là 6 oz (1 oz = 28g). Một mẫu gồm 30 tuýt kem được chọn ngẫu nhiên để kiểm tra định kỳ. Bộ phận điều khiển dây chuyền phải đảm bảo để trọng lượng trung bình mỗi tuýt kem là 6 oz; nếu nhiều hơn hoặc ít hơn, dây chuyền phải được điều chỉnh lại. Giả sử trung bình mẫu của 30 tuýt kem là 6.1 oz và độ lệch tiêu chuẩn của tổng thể $\sigma = 0.2$ oz. Thực hiện kiểm định giả thuyết với mức ý nghĩa 3% để xác định xem dây chuyền sản xuất có vận hành tốt hay không?</div>	<div>XSTK</div> <div>N.T. M. Ngọc</div> <div>Các khái niệm trong kiểm định giả thuyết thông kê</div> <div>Giả thuyết H_0 và đối thuyết H_1</div> <div>Các loại sai lầm</div> <div>Giá trị p_{value}</div> <div>Kiểm định giả thuyết - Trường hợp một mẫu</div> <div>Kiểm định giả thuyết cho kỳ vọng</div> <div>Kiểm định giả thuyết cho tỷ lệ</div> <div>Kiểm định giả thuyết - Trường hợp hai mẫu</div> <div>Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập</div> <div>Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập</div> <div>Kiểm định giả thuyết so sánh hai tỷ lệ</div>	<div>TH 1: biết σ^2</div>
<div>XSTK</div> <div>N.T. M. Ngọc</div> <div>Các khái niệm trong kiểm định giả thuyết thông kê</div> <div>Giả thuyết H_0 và đối thuyết H_1</div> <div>Các loại sai lầm</div> <div>Giá trị p_{value}</div> <div>Kiểm định giả thuyết - Trường hợp một mẫu</div> <div>Kiểm định giả thuyết cho kỳ vọng</div> <div>Kiểm định giả thuyết cho tỷ lệ</div> <div>Kiểm định giả thuyết - Trường hợp hai mẫu</div> <div>Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập</div> <div>Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập</div> <div>Kiểm định giả thuyết so sánh hai tỷ lệ</div>	<div>TH 1: biết σ^2</div> <div>Ví dụ 3 : Metro EMS: Một bệnh viện tại trung tâm thành phố cung cấp dịch vụ cấp cứu tại nhà. Với khoảng 20 xe cấp cứu, mục tiêu của trung tâm là cung cấp dịch vụ cấp cứu trong khoảng thời gian trung bình là 12 phút sau khi nhận được điện thoại yêu cầu. Một mẫu ngẫu nhiên gồm thời gian đáp ứng khi có yêu cầu của 40 ca cấp cứu được chọn. Trung bình mẫu là 13.25 phút. Biết rằng độ lệch tiêu chuẩn của tổng thể là $\sigma = 3.2$ phút. Giám đốc EMS muốn thực hiện một kiểm định, với mức ý nghĩa 5%, để xác định xem liệu thời gian một ca cấp cứu có bé hơn hoặc bằng 12 phút hay không?</div>	<div>XSTK</div> <div>N.T. M. Ngọc</div> <div>Các khái niệm trong kiểm định giả thuyết thông kê</div> <div>Giả thuyết H_0 và đối thuyết H_1</div> <div>Các loại sai lầm</div> <div>Giá trị p_{value}</div> <div>Kiểm định giả thuyết - Trường hợp một mẫu</div> <div>Kiểm định giả thuyết cho kỳ vọng</div> <div>Kiểm định giả thuyết cho tỷ lệ</div> <div>Kiểm định giả thuyết - Trường hợp hai mẫu</div> <div>Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập</div> <div>Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập</div> <div>Kiểm định giả thuyết so sánh hai tỷ lệ</div>	<div>TH 1: biết σ^2</div>

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

TH 1: biết σ^2

Ví dụ 4 :

Trong năm trước trọng lượng trung bình trước khi xuất chuồng của bò ở một trại chăn nuôi là 380 kg. Năm nay người ta áp dụng thử một chế độ chăn nuôi mới với hi vọng là bò sẽ tăng trọng nhanh hơn. Sau một thời gian áp dụng thử người ta lấy ngẫu nhiên 50 con bò trước khi xuất chuồng đem cân và tính được trọng lượng trung bình của chúng là 390 kg. Với mức ý nghĩa $\alpha = 0.01$ có thể cho rằng trọng lượng trung bình của bò trước khi xuất chuồng đã tăng lên hay không? Giả thiết trọng lượng của bò là BNN có phân phối chuẩn với độ lệch chuẩn là 35.2 kg.

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

TH 2: không biết σ^2 , $n \geq 30$

• Ví dụ:

Trạm cảnh sát giao thông trên đường cao tốc sẽ thực hiện việc bắn tốc độ định kỳ tại các địa điểm khác nhau để kiểm tra tốc độ của các phương tiện giao thông. Một mẫu về tốc độ của các loại xe được chọn để thực hiện kiểm định giả thuyết sau

$H_0 : \mu = 65$

$H_1 : \mu > 65$

Những vị trí mà bác bỏ H_0 là những vị trí tốt nhất được chọn để đặt radar kiểm soát tốc độ.
Tại địa điểm F, một mẫu gồm tốc độ của 64 phương tiện được bắn tốc độ ngẫu nhiên có trung bình là 66.2 mph và độ lệch tiêu chuẩn 4.2 mph. Sử dụng $\alpha = 5\%$ để kiểm định giả thuyết.

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

Kiểm định giả thuyết về so sánh kỳ vọng với 1 số

TH 2: không biết σ^2 , $n \geq 30$

• Các giả định:

Mẫu ngẫu nhiên X_1, \dots, X_n được chọn từ tổng thể có kỳ vọng μ và phương sai σ^2 không biết.

Sử dụng ước lượng không chệch S thay cho σ .

Cỡ mẫu lớn: $n \geq 30$.

• Khi cỡ mẫu lớn biến ngẫu nhiên

$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$

sẽ hội tụ về phân phối chuẩn hóa $Z \sim \mathcal{N}(0, 1)$. Khi đó, với mức ý nghĩa α , miền bác bỏ hoặc p -giá trị sẽ được tính tương tự như trường hợp biết phương sai, chỉ thay thế σ bằng s khi tính giá trị thống kê kiểm định.

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

TH 2: không biết σ^2 , $n \geq 30$

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

TH 2: không biết σ^2 , $n \geq 30$

Ví dụ: Một nghiên cứu được thực hiện để xác định mức độ hài lòng của khách hàng sau khi công ty điện thoại thay đổi, cải tiến 1 số dịch vụ khách hàng. Trước khi thay đổi, mức độ hài lòng của khách hàng tính trung bình là 77, theo thang điểm từ 0 đến 100. 350 khách hàng được chọn ngẫu nhiên để gửi bảng điều tra xin ý kiến sau khi các thay đổi được thực hiện, mức độ hài lòng trung bình tình được là 84, với độ lệch chuẩn là 28. Với mức ý nghĩa $\alpha = 5\%$, có thể kết luận khách hàng đã được làm hài lòng ở mức độ cao hơn được không?

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

TH 2: không biết σ^2 , $n \geq 30$

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

TH 2: không biết σ^2 , $n \geq 30$

Ví dụ khác:

Đo đường kính của 36 chi tiết máy ta được bảng số liệu sau:

Độ dài đường kính	10.10	10.12	10.20	10.25	10.30
Số chi tiết	3	15	14	2	2

Với mức ý nghĩa $\alpha = 0.05$ hãy cho kết luận về ý kiến: "Trung bình đường kính là 10.20"

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

Kiểm định giả thuyết về so sánh kỳ vọng với 1 số

TH 3: không biết σ^2 , $n < 30$

• Các giả định:

- Mẫu ngẫu nhiên X_1, \dots, X_n được chọn từ tổng thể có phân phối chuẩn $\mathcal{N}(\mu, \sigma^2)$ với kỳ vọng μ và phương sai σ^2 không biết.
- Sử dụng ước lượng S thay cho σ .
- Cỡ mẫu nhỏ: $n < 30$.

• Bài toán kiểm định có 3 trường hợp:

(a) $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$

(b) $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu > \mu_0 \end{cases}$

(c) $\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu < \mu_0 \end{cases}$

với mức ý nghĩa α cho trước.

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

TH 3: không biết σ^2 , $n < 30$

Các bước kiểm định:

B1: Phát biểu giả thuyết H_0 và đối thuyết H_1

B2: Xác định tiêu chuẩn kiểm định:

Chọn thống kê

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

làm tiêu chuẩn kiểm định.

Nếu giả thuyết H_0 đúng thì T tuân theo phân phối Student với bậc tự do $n - 1$.

Từ mẫu thực nghiệm, tính giá trị thống kê kiểm định:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

TH 3: không biết σ^2 , $n < 30$

Các bước kiểm định (tt):

B3: Với mức ý nghĩa α , xác định miền bác bỏ:

Với mức ý nghĩa α và dựa vào đối thuyết H_1 , xác định miền bác bỏ hoặc tính p -giá trị tương ứng:

Trường hợp	bác bỏ H_0 nếu	p -giá trị
(a) $H_1 : \mu \neq \mu_0$	$ t > t_{1-\alpha/2}^{(n-1)}$	$2\mathbb{P}(T_{(n-1)} \geq t)$
(b) $H_1 : \mu > \mu_0$	$t > t_{1-\alpha}^{(n-1)}$	$\mathbb{P}(T_{(n-1)} \geq t)$
(c) $H_1 : \mu < \mu_0$	$t < -t_{1-\alpha}^{(n-1)}$	$\mathbb{P}(T_{(n-1)} \leq t)$

Trong đó,

$T_{(n-1)}$ là biến ngẫu nhiên tuân theo phân phối Student với bậc tự do $(n - 1)$;

$t_{1-\alpha}^{(n-1)}$ có được bằng cách tra bảng phân phối Student.

B4 Kết luận:

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

TH 3: không biết σ^2 , $n < 30$

Ví dụ:

Một loại đèn chiếu sáng được nhà sản xuất cho biết có tuổi thọ trung bình thấp nhất là 65 giờ. Kết quả kiểm tra từ mẫu ngẫu nhiên 21 bóng đèn cho thấy tuổi thọ trung bình là 62,5 giờ, với độ lệch chuẩn là 3. Với $\alpha = 0,01$, có thể kết luận gì về lời tuyên bố của nhà sản xuất?

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

TH 3: không biết σ^2 , $n < 30$

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

TH 3: không biết σ^2 , $n < 30$

Ví dụ

Cho 8 kết quả đo đặc về một đại lượng bởi cùng một máy đo không có sai lầm hệ thống:

369, 378, 315, 420, 385, 401, 372, 383

Với mức ý nghĩa $\alpha = 0.05$, hãy cho kết luận về ý kiến: "Giá trị trung bình là 380". Biết rằng đại lượng được đo có phân phối chuẩn.

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

Kiểm định giả thuyết cho tỷ lệ p

• Quan sát sự xuất hiện của biến cố "phần tử mang đặc tính A" trong n phép thử độc lập. Gọi Y là số lần xuất hiện biến cố trên thì $Y \sim B(n, p)$. Và

$$\hat{p} = \frac{Y}{n}$$

là một ước lượng không chệch cho p .

• Chọn thống kê

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

là tiêu chuẩn kiểm định. Nếu H_0 đúng, $Z \sim \mathcal{N}(0, 1)$ với điều kiện $n\hat{p} \geq 5$ và $n(1 - \hat{p}) \geq 5$.

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

Kiểm định giả thuyết về so sánh tỉ lệ tổng thể với 1 số

• Bài toán:

Cho tổng thể X , trong đó tỷ lệ phần tử mang đặc tính A nào đó trong tổng thể là p (p chưa biết). Từ mẫu ngẫu nhiên (X_1, X_2, \dots, X_n) hãy kiểm định

$$(a) \begin{cases} H_0 : p = p_0 \\ H_1 : p \neq p_0 \end{cases}$$

$$(b) \begin{cases} H_0 : p = p_0 \\ H_1 : p > p_0 \end{cases}$$

$$(c) \begin{cases} H_0 : p = p_0 \\ H_1 : p < p_0 \end{cases}$$

với mức ý nghĩa α ; với p_0 là giá trị cho trước.

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

Kiểm định giả thuyết cho tỷ lệ p

• B1: Phát biểu giả thuyết H_0 và đối thuyết H_1

• B2: Xác định tiêu chuẩn kiểm định:

Thống kê

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

là tiêu chuẩn kiểm định. Nếu H_0 đúng, $Z \sim \mathcal{N}(0, 1)$ với điều kiện $n\hat{p} \geq 5$ và $n(1 - \hat{p}) \geq 5$.

Từ mẫu thực nghiệm, tính giá trị thống kê kiểm định:

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}$$

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

Kiểm định giả thuyết cho tỷ lệ p

Các bước kiểm định (tt):

B3: Với mức ý nghĩa α , xác định miền bác bỏ: Với mức ý nghĩa α và dựa vào đối thuyết H_1 , xác định miền bác bỏ hoặc tính p -giá trị tương ứng:

Trường hợp	bác bỏ H_0 nếu	p -giá trị
(a) $H_1 : p \neq p_0$	$ z > z_{1-\alpha/2}$	$2(1 - \Phi(z))$
(b) $H_1 : p > p_0$	$z > z_{1-\alpha}$	$1 - \Phi(z)$
(c) $H_1 : p < p_0$	$z < -z_{1-\alpha}$	$\Phi(z)$

Trong đó, $z_{1-\alpha}$ được tìm từ bảng phân phối Gauss.

B4 Kết luận:

Nếu bác bỏ giả thuyết H_0 , ta kết luận H_1 đúng với $(1 - \alpha)100\%$ độ tin cậy.

Ngược lại, ta kết luận chưa đủ cơ sở để bác bỏ giả thuyết H_0 với mức ý nghĩa α .

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

Kiểm định giả thuyết về tỉ lệ p

Ví dụ: Giả sử sản phẩm của một công ty sản xuất vỏ xe ô tô đã chiếm 42% thị trường. Hiện tại, trước sự cạnh tranh của đối thủ và những điều kiện thay đổi của môi trường kinh doanh, ban lãnh đạo muốn kiểm tra lại xem thị phần công ty có còn là 42% hay không. Chọn ngẫu nhiên 550 ô tô trên đường, kết quả cho thấy có 219 xe sử dụng vỏ xe của công ty. Có thể kết luận gì với mức ý nghĩa $\alpha = 0,1$?

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

Tính $p_{\text{giá trị}}$

Ví dụ: Trong ví dụ về kiểm định giả thuyết tỉ lệ trong tổng thể trên, ta kết luận không thể bác bỏ giả thuyết H_0 với mức ý nghĩa $\alpha = 10\%$.
Tất nhiên, ta cũng không thể bác bỏ giả thuyết H_0 với mức ý nghĩa $\alpha < 10\%$.
Vậy, ở mức ý nghĩa $\alpha > 10\%$, liệu có thể bác bỏ giả thuyết H_0 ?
Nói cách khác, vấn đề là xác định mức ý nghĩa nhỏ nhất mà ở đó giả thuyết H_0 có thể bị bác bỏ. Mức ý nghĩa nhỏ nhất đó gọi là $p_{\text{giá trị}}$ (p_{value}).

Trong ví dụ trên, giá trị thống kê kiểm định tính được là $z = -1,037$ và kiểm định bên trái nên $p_{\text{giá trị}} = \phi(-1,037) = 1 - \phi(1,037) = 0,1499$.

XSTK
N.T. M. Ngọc
Các khái niệm trong kiểm định giả thuyết thông kê
Giả thuyết H_0 và đối thuyết H_1
Các loại sai lầm
Giá trị p_{value}
Kiểm định giả thuyết - Trường hợp một mẫu
Kiểm định giả thuyết cho kỳ vọng
Kiểm định giả thuyết cho tỷ lệ
Kiểm định giả thuyết - Trường hợp hai mẫu
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập
Kiểm định giả thuyết so sánh hai tỷ lệ

Kiểm định giả thuyết cho tỷ lệ p

Ví dụ khác: Trong kỳ nghỉ Giáng sinh và đầu năm mới, Cục An toàn giao thông đã thống kê được rằng có 500 người chết và 25000 người bị thương do các vụ tai nạn giao thông trên toàn quốc. Theo thông cáo của Cục ATGT thì khoảng 50% số vụ tai nạn có liên quan đến rượu bia. Khảo sát ngẫu nhiên 120 vụ tai nạn thấy có 67 vụ do ảnh hưởng của rượu bia. Sử dụng số liệu trên để kiểm định lời khẳng định của Cục An toàn giao thông với mức ý nghĩa $\alpha = 5\%$.

XSTK
N.T. M. Ngọc
Các khái niệm trong kiểm định giả thuyết thông kê
Giả thuyết H_0 và đối thuyết H_1
Các loại sai lầm
Giá trị p_{value}
Kiểm định giả thuyết - Trường hợp một mẫu
Kiểm định giả thuyết cho kỳ vọng
Kiểm định giả thuyết cho tỷ lệ
Kiểm định giả thuyết - Trường hợp hai mẫu
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập
Kiểm định giả thuyết so sánh hai tỷ lệ

Kiểm định giả thuyết cho tỷ lệ p

Ví dụ khác:
Một nhà máy sản xuất sản phẩm với tỉ lệ loại một lúc đầu là 20%. Sau khi áp dụng phương pháp sản xuất mới, kiểm tra ngẫu nhiên 500 sản phẩm thấy có 150 sản phẩm loại một. Cho kết luận về tác dụng của phương pháp sản xuất với mức ý nghĩa 1%.

ĐS: phương pháp sản xuất mới làm tăng tỉ lệ sản phẩm loại một với mức ý nghĩa 1%

XSTK
N.T. M. Ngọc
Các khái niệm trong kiểm định giả thuyết thông kê
Giả thuyết H_0 và đối thuyết H_1
Các loại sai lầm
Giá trị p_{value}
Kiểm định giả thuyết - Trường hợp một mẫu
Kiểm định giả thuyết cho kỳ vọng
Kiểm định giả thuyết cho tỷ lệ
Kiểm định giả thuyết - Trường hợp hai mẫu
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập
Kiểm định giả thuyết so sánh hai tỷ lệ

Nội dung: Kiểm định giả thuyết cho trường hợp 2 mẫu

Kiểm định giả thuyết cho trường hợp hai mẫu độc lập

- So sánh hai kỳ vọng
 - Trường hợp biết phương sai
 - Trường hợp không biết phương sai, mẫu lớn
 - Trường hợp không biết phương sai, mẫu nhỏ
 - Trường hợp $\sigma_1^2 = \sigma_2^2 = \sigma^2$
 - Trường hợp $\sigma_1^2 \neq \sigma_2^2$
- So sánh hai tỉ lệ
- So sánh hai phương sai

Kiểm định giả thuyết cho trường hợp hai mẫu phụ thuộc

- So sánh hai kỳ vọng

XSTK
N.T. M. Ngọc
Các khái niệm trong kiểm định giả thuyết thông kê
Giả thuyết H_0 và đối thuyết H_1
Các loại sai lầm
Giá trị p_{value}
Kiểm định giả thuyết - Trường hợp một mẫu
Kiểm định giả thuyết cho kỳ vọng
Kiểm định giả thuyết cho tỷ lệ
Kiểm định giả thuyết - Trường hợp hai mẫu
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập
Kiểm định giả thuyết so sánh hai tỷ lệ

TH 1: đã biết σ_1^2, σ_2^2

Các giả định:

Quan sát X trên hai mẫu ngẫu nhiên lấy từ hai tổng thể 1 và 2 độc lập nhau.

- Trên tổng thể 1: $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$, lấy mẫu ngẫu nhiên cỡ n_1 có trung bình mẫu \bar{X}_1 .
- Trên tổng thể 2: $X \sim \mathcal{N}(\mu_2, \sigma_2^2)$, lấy mẫu ngẫu nhiên cỡ n_2 có trung bình mẫu \bar{X}_2 .
- Các phương sai σ_1^2 và σ_2^2 đã biết.

Bài toán kiểm định gồm các dạng sau:

Hãy kiểm định một trong những giả thuyết sau:

$$(a) \begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$$

$$(b) \begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 > \mu_2 \end{cases}$$

$$(c) \begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 < \mu_2 \end{cases}$$

với mức ý nghĩa α cho trước.

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

Các bước kiểm định:

B1: Phát biểu giả thuyết H_0 và đối thuyết H_1

B2: Xác định tiêu chuẩn kiểm định:

Chọn thống kê

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

làm tiêu chuẩn kiểm định. Nếu giả thuyết H_0 đúng thì $Z \sim \mathcal{N}(0, 1)$.

Từ mẫu thực nghiệm, tính giá trị thống kê kiểm định:

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

Ví dụ

Một công ty sản xuất sơn nghiên cứu về một loại phụ gia làm giảm thời gian khô của sơn. Thực hiện thí nghiệm trên 2 mẫu : mẫu 1 gồm 10 mẫu vật được sơn bằng loại sơn bình thường; mẫu 2 gồm 10 mẫu vật được sơn bằng loại sơn có chất phụ gia mới. Trong những nghiên cứu trước, biết rằng độ lệch chuẩn của thời gian khô sau khi quét sơn là 8 phút và không thay đổi khi thêm phụ gia vào. Thời gian khô trung bình của mẫu 1 và 2 lần lượt là $\bar{x}_1 = 121$ phút và $\bar{x}_2 = 112$ phút. Với mức ý nghĩa 5%, hãy cho kết luận về loại sơn với chất phụ gia mới.

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

Các bước kiểm định: (tt)

B3: Với mức ý nghĩa α , xác định miền bác bỏ:

Với mức ý nghĩa α và dựa vào đối thuyết H_1 , xác định miền bác bỏ hoặc tính p -giá trị tương ứng:

Trường hợp	bác bỏ H_0 nếu	p -giá trị
(a) $H_1 : \mu_1 \neq \mu_2$	$ z > z_{1-\alpha/2}$	$2(1 - \Phi(z))$
(b) $H_1 : \mu_1 > \mu_2$	$z > z_{1-\alpha}$	$1 - \Phi(z)$
(c) $H_1 : \mu_1 < \mu_2$	$z < -z_{1-\alpha}$	$\Phi(z)$

Trong đó, $z_{1-\alpha}$ được tìm từ bảng phân phối Gauss.

Lưu ý: Nếu p -giá trị $< \alpha$ thì ta bác bỏ giả thuyết H_0 với mức ý nghĩa α .

B4: Kết luận:

Nếu bác bỏ giả thuyết H_0 , ta kết luận H_1 đúng với $(1 - \alpha)100\%$ độ tin cậy.

Ngược lại, ta kết luận chưa đủ cơ sở để bác bỏ giả thuyết H_0 với mức ý nghĩa α .

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

Giải ví dụ:

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

TH 1: đã biết σ_1^2, σ_2^2

Ví dụ khác

Hai công thức khác nhau của nhiên liệu động cơ ôxy hóa đang được thử nghiệm để nghiên cứu chỉ số octane của chúng. Phương sai chỉ số octane của công thức thứ nhất là $\sigma_1^2 = 1,5$ và công thức thứ hai là $\sigma_2^2 = 1,2$. Hai mẫu ngẫu nhiên có cỡ mẫu $n_1 = 15$ và $n_2 = 20$ được nghiên cứu có chỉ số octane trung bình lần lượt là $\bar{x}_1 = 89,6$ và $\bar{x}_2 = 92,5$. Giả sử chỉ số octane có phân phối chuẩn. Nếu công thức 2 tạo ra chỉ số octane cao hơn so với công thức 1, thì nhà sản xuất muốn phát hiện nó. Hãy xây dựng và kiểm định giả thuyết thích hợp sử dụng $\alpha = 0,05$ và tính p -giá trị.

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

TH 2: chưa biết σ_1^2, σ_2^2 , mẫu lớn

Các giả định:

Quan sát X trên hai mẫu ngẫu nhiên lấy từ hai tổng thể 1 và 2 độc lập nhau.

- Trên tổng thể 1: $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$, lấy mẫu ngẫu nhiên cỡ n_1 có trung bình mẫu \bar{X}_1 .
- Trên tổng thể 2: $X \sim \mathcal{N}(\mu_2, \sigma_2^2)$, lấy mẫu ngẫu nhiên cỡ n_2 trung bình mẫu \bar{X}_2 .
- Các phương sai σ_1^2 và σ_2^2 chưa biết.
- Cỡ mẫu lớn: $n_1 \geq 30$ và $n_2 \geq 30$.

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

TH 1: đã biết σ_1^2, σ_2^2

Ví dụ khác

Arnold Palmer và Tiger Woods là hai golf thủ giỏi nhất từ trước đến nay. Để so sánh hai người nếu cả hai đều đang chơi ở đỉnh cao, dữ liệu mẫu sau đây cho biết kết quả điểm thi đấu golf 18 lỗ trong giải PGA. Điểm của Palmer từ mùa giải 1960, trong khi của Woods từ mùa giải 1999 (theo Golf Magazine, tháng 2, 2020).

Arnold Palmer , 1960	Tiger Woods, 1999
$n_1 = 112$	$n_2 = 84$
$\bar{x}_1 = 69,95$	$\bar{x}_2 = 69,56$

Giả sử độ lệch chuẩn tổng thể là 2,5 cho cả hai golf thủ. Có thể cho rằng không có chênh lệch trung bình tổng thể về điểm thi đấu golf 18 lỗ của hai golf thủ với mức ý nghĩa 1% ?

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

TH 2: chưa biết σ_1^2, σ_2^2 , mẫu lớn

- Đối với trường hợp mẫu lớn, khi phương sai tổng thể σ_1^2 và σ_2^2 không biết, ta thay thế bằng các phương sai mẫu S_1^2 và S_2^2 mà không tạo ra nhiều khác biệt.
- Khi $n_1 \geq 30$ và $n_2 \geq 30$, dưới giả thuyết H_0

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

sẽ xấp xỉ phân phối chuẩn hóa $\mathcal{N}(0,1)$.

Từ mẫu thực nghiệm, tính giá trị thống kê kiểm định:

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

- Miền bác bỏ (hoặc p - giá trị) trong trường hợp này được tính tương tự như trường hợp biết phương sai.

TH 2: chưa biết σ_1^2, σ_2^2 , mẫu lớn

Ví dụ

Một bài báo về Kỹ thuật Radio và Vật lý điện tử (1984) đã nghiên cứu hoạt động của một máy phát ngẫu nhiên khi có tiếng ồn bên ngoài. Số chu kỳ được đo trong một mẫu là 100 lần với hai mức điện áp nhiễu khác nhau, 100 mV và 150 mV. Với mức điện áp nhiễu 100 mV, số chu kỳ trung bình là 7,9 và độ lệch chuẩn $s_1 = 2,6$. Với điện áp nhiễu 150 mV, số chu kỳ trung bình là 6,9 và độ lệch chuẩn $s_2 = 2,4$.

Ban đầu, người ta nghi ngờ rằng việc tăng điện áp, tiếng ồn sẽ làm giảm số chu kỳ trung bình. Dữ liệu có hỗ trợ xác nhận này không? Sử dụng $\alpha = 0,05$ và giả định rằng hai tổng thể có phân phối chuẩn. Tính p -giá trị của kiểm định?

TH 2: chưa biết σ_1^2, σ_2^2 , mẫu lớn

Ví dụ khác

Một trại chăn nuôi chọn một giống gà để tiến hành nghiên cứu hiệu quả của hai loại thức ăn A và B. Sau một thời gian nuôi thử nghiệm trong cùng điều kiện bằng hai loại thức ăn này, người ta chọn 50 con gà nuôi bằng thức ăn A thấy khối lượng trung bình là 2,2 kg, độ lệch chuẩn mẫu hiệu chỉnh là 1,25 kg. Chọn 40 con gà nuôi bằng thức ăn B thấy khối lượng trung bình là 1,2 kg, độ lệch chuẩn mẫu hiệu chỉnh là 1,02 kg. Hãy đánh giá hiệu quả của hai loại thức ăn trên với mức ý nghĩa 1%.

TH 3: chưa biết σ_1^2, σ_2^2 , mẫu nhỏ

Các giả định:

Quan sát X trên hai mẫu ngẫu nhiên lấy từ hai tổng thể 1 và 2 độc lập nhau.

- Trên tổng thể 1: $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$, lấy mẫu ngẫu nhiên cỡ n_1 có trung bình mẫu \bar{X}_1 .
- Trên tổng thể 2: $X \sim \mathcal{N}(\mu_2, \sigma_2^2)$, lấy mẫu ngẫu nhiên cỡ n_2 trung bình mẫu \bar{X}_2 .
- Các phương sai σ_1^2 và σ_2^2 chưa biết.
- Cỡ mẫu nhỏ: $n_1 < 30$ và $n_2 < 30$.

Ta xét hai trường hợp:

- 1 Trường hợp phương sai bằng nhau $\sigma_1^2 = \sigma_2^2$.
- 2 Trường hợp phương sai khác nhau $\sigma_1^2 \neq \sigma_2^2$.

Bài toán kiểm định gồm các dạng sau:

Hãy kiểm định một trong những giả thuyết sau:

$$(a) \begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases} \quad (b) \begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 > \mu_2 \end{cases} \quad (c) \begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 < \mu_2 \end{cases}$$

với mức ý nghĩa α cho trước.

TH 3a: $\sigma_1^2 = \sigma_2^2$ chưa biết, mẫu nhỏ

Các bước kiểm định:

- B1: Phát biểu giả thuyết H_0 và đối thuyết H_1
- B2: Xác định tiêu chuẩn kiểm định:
 - Trường hợp $\sigma_1^2 = \sigma_2^2 = \sigma^2$, ta sử dụng một ước lượng chung cho cả σ_1^2 và σ_2^2 là S_p^2 gọi là phương sai mẫu chung

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}.$$

- Dưới giả thuyết H_0 , thống kê

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

sẽ xấp xỉ phân phối Student với $n_1 + n_2 - 2$ bậc tự do.

- Từ mẫu thực nghiệm, tính giá trị thống kê kiểm định:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

Các bước kiểm định: (tt)

B3: Với mức ý nghĩa α , xác định miền bác bỏ:

Với mức ý nghĩa α và dựa vào đối thuyết H_1 , xác định miền bác bỏ hoặc tính p -giá trị tương ứng:

Đặt $df = n_1 + n_2 - 2$, miền bác bỏ và p - giá trị trong trường hợp này có dạng:

Trường hợp	bác bỏ H_0 nếu	p -giá trị
(a) $H_1 : \mu_1 \neq \mu_2$	$ t > t_{1-\alpha/2}^{df}$	$2\mathbb{P}(T_{(df)} \geq t)$
(b) $H_1 : \mu_1 > \mu_2$	$t > t_{1-\alpha}^{df}$	$\mathbb{P}(T_{(df)} \geq t)$
(c) $H_1 : \mu_1 < \mu_2$	$t < -t_{1-\alpha}^{df}$	$\mathbb{P}(T_{(df)} \leq t)$

Trong đó,

- $T_{(df)}$ là biến ngẫu nhiên tuân theo phân phối Student với bậc tự do df ;
- $t_{1-\alpha}^{df}$ có được bằng cách tra bảng phân phối Student.

Lưu ý: Nếu p -giá trị $\leq \alpha$ thì ta bác bỏ giả thuyết H_0 với mức ý nghĩa α .

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

TH 3b: $\sigma_1^2 \neq \sigma_2^2$ chưa biết, mẫu nhỏ

Các bước kiểm định: (tt)

B3: Với mức ý nghĩa α , xác định miền bác bỏ:

Miền bác bỏ (hay p -giá trị tương ứng) trong trường hợp này được xác định giống như trường hợp phương sai bằng nhau, chỉ thay bậc tự do df cho bởi phương trình (1).

B4: Kết luận:

- Nếu bác bỏ giả thuyết H_0 , ta kết luận H_1 đúng với $(1 - \alpha)100\%$ độ tin cậy.
- Ngược lại, ta kết luận chưa đủ cơ sở để bác bỏ giả thuyết H_0 với mức ý nghĩa α .

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

TH 3b: $\sigma_1^2 \neq \sigma_2^2$ chưa biết, mẫu nhỏ

Các bước kiểm định:

B1: Phát biểu giả thuyết H_0 và đối thuyết H_1

B2: Xác định tiêu chuẩn kiểm định:

Trường hợp $\sigma_1^2 \neq \sigma_2^2$, dưới giả thuyết H_0 , thống kê

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

sẽ xấp xỉ phân phối Student với bậc tự do df được xác định như sau

$$df = \frac{[(s_1^2/n_1) + (s_2^2/n_2)]^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}}$$

Từ mẫu thực nghiệm, tính giá trị thống kê kiểm định:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

Ví dụ:

Trong sản xuất chất bán dẫn, khắc hóa chất ướt thường được sử dụng để loại bỏ silic từ mặt sau của tấm wafer trước khi kim loại hóa. Tỷ lệ ăn mòn (etch) là một đặc tính quan trọng trong quá trình này và được biết là tuân theo phân phối chuẩn. Hai phương pháp khắc khác nhau đã được so sánh bằng cách sử dụng hai mẫu ngẫu nhiên gồm 10 tấm wafer cho mỗi dung dịch. Tỷ lệ ăn mòn quan sát được như sau:

Mẫu 1: 9,9 10,6 9,4 10,3 9,3 10,0 9,6 10,3 10,2 10,1

Mẫu 2: 10,2 10,0 10,6 10,2 10,7 10,7 10,4 10,4 10,5 10,3

a) Dữ liệu trên có hỗ trợ tuyên bố rằng tỷ lệ ăn mòn trung bình là giống nhau cho cả hai phương pháp không, giả sử hai phương sai tổng thể bằng nhau?

b) Thực hiện lại kiểm định trên câu a) với giả thiết hai phương sai tổng thể khác nhau?

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

Ví dụ khác

Ban lãnh đạo công ty cho rằng doanh số bán hàng tăng lên sau khi thực hiện các biện pháp khuyến mãi. Chọn ngẫu nhiên 13 tuần trước đợt khuyến mãi có được doanh số trung bình là 1234 triệu đồng và độ lệch chuẩn mẫu là 324 triệu đồng. Và chọn ngẫu nhiên 14 tuần sau đợt khuyến mãi có được doanh số trung bình là 1864 triệu đồng và độ lệch chuẩn mẫu là 289 triệu đồng. Hãy kiểm định ý kiến trên với $\alpha = 5\%$

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

Ví dụ khác

Xét hai chất xúc tác có thể được sử dụng trong một phản ứng hóa học hàng loạt. Mười hai lô được sử dụng chất xúc tác 1, dẫn đến năng suất trung bình là 86 và độ lệch chuẩn mẫu là 3. Mười năm lô được sử dụng chất xúc tác 2, và kết quả là năng suất trung bình 89 với độ lệch chuẩn là 2. Giả sử năng suất các phép đo xấp xỉ phân phối chuẩn với cùng độ lệch chuẩn. Với mức ý nghĩa 1%, có bằng chứng để khẳng định rằng chất xúc tác 2 tạo ra năng suất trung bình cao hơn chất xúc tác 1 không?

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

So sánh hai mẫu không độc lập, kiểm định cặp (paired t - test)

- Khi hai mẫu không độc lập thì mỗi giá trị quan trắc được trong một mẫu có mối liên hệ tương ứng với một giá trị quan trắc ở mẫu thứ hai. Như vậy, ta có thể ghép cặp từng giá trị trong hai mẫu với nhau.
- Việc ghép cặp là kết quả của việc
 - quan trắc giá trị trước và sau khi thực hiện 1 thí nghiệm. Chẳng hạn như đo trọng lượng trước và sau khi thực hiện một chế độ ăn kiêng.
 - so sánh cùng 1 đặc tính.
 - thí nghiệm trên cùng 1 địa điểm.
 - thí nghiệm với cùng thời gian.

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

So sánh hai mẫu không độc lập, kiểm định cặp (paired t - test)

- Xét (X_{1i}, X_{2i}) , với $i = 1, 2, \dots, n$, là tập gồm n cặp giá trị quan trắc với giả sử rằng kỳ vọng và phương sai của tổng thể đại diện bởi X_1 là μ_1 và σ_1^2 và kỳ vọng và phương sai của tổng thể đại diện bởi X_2 là μ_2 và σ_2^2 . X_{1i} và X_{2j} ($i \neq j$) độc lập.
- Định nghĩa độ sai khác giữa mỗi cặp trong tập hợp các giá trị quan trắc là
$$D_i = X_{1i} - X_{2i}, \quad i = 1, \dots, n \tag{2}$$
- Các $D_i, i = 1, \dots, n$ được giả sử có phân phối chuẩn.
- Goi $\mu_D = E(D_i)$, bởi vì D_1, \dots, D_n là những biến ngẫu nhiên độc lập và có cùng phân phối, nếu d_1, \dots, d_n là những giá trị của D_1, \dots, D_n , ta định nghĩa

XSTK
N.T. M. Ngọc
Các khái niệm trong kiểm định giả thuyết thông kê
Giả thuyết H_0 và đối thuyết H_1
Các loại sai lầm
Giá trị p_{value}
Kiểm định giả thuyết - Trường hợp một mẫu
Kiểm định giả thuyết cho kỳ vọng
Kiểm định giả thuyết cho tỷ lệ
Kiểm định giả thuyết - Trường hợp hai mẫu
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập
Kiểm định giả thuyết so sánh hai tỷ lệ

So sánh hai mẫu không độc lập
(paired t - test)

- Goi $\mu_D = E(D_i)$, bởi vì D_1, \dots, D_n là những biến ngẫu nhiên độc lập và có cùng phân phối, nếu d_1, \dots, d_n là những giá trị của D_1, \dots, D_n , ta định nghĩa
$$\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i \quad (3)$$
$$s_d^2 = \frac{1}{n-1} \sum_{i=1}^n (d_i - \bar{d})^2 = \frac{1}{n-1} \sum_{i=1}^n d_i^2 - \frac{n}{n-1} (\bar{d})^2 \quad (4)$$
- Ta cần kiểm định các giả thuyết và đối thuyết sau
$$(a) \begin{cases} H_0 : \mu_D = D_0 \\ H_1 : \mu_D \neq D_0 \end{cases} \quad (b) \begin{cases} H_0 : \mu_D = D_0 \\ H_1 : \mu_D < D_0 \end{cases} \quad (c) \begin{cases} H_0 : \mu_D = D_0 \\ H_1 : \mu_D > D_0 \end{cases}$$

XSTK
N.T. M. Ngọc
Các khái niệm trong kiểm định giả thuyết thông kê
Giả thuyết H_0 và đối thuyết H_1
Các loại sai lầm
Giá trị p_{value}
Kiểm định giả thuyết - Trường hợp một mẫu
Kiểm định giả thuyết cho kỳ vọng
Kiểm định giả thuyết cho tỷ lệ
Kiểm định giả thuyết - Trường hợp hai mẫu
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập
Kiểm định giả thuyết so sánh hai tỷ lệ

So sánh hai mẫu không độc lập
(paired t - test)

Các bước kiểm định

- Phát biểu giả thuyết H_0 và đối thuyết H_1
- Xác định mức ý nghĩa α
- Tính thống kê kiểm định

$$T = \frac{\bar{D} - D_0}{S_D / \sqrt{n}} \quad (5)$$

thống kê T có phân phối Student với $n - 1$ bậc tự do.

- Xác định miền bác bỏ với mức ý nghĩa α ,

XSTK
N.T. M. Ngọc
Các khái niệm trong kiểm định giả thuyết thông kê
Giả thuyết H_0 và đối thuyết H_1
Các loại sai lầm
Giá trị p_{value}
Kiểm định giả thuyết - Trường hợp một mẫu
Kiểm định giả thuyết cho kỳ vọng
Kiểm định giả thuyết cho tỷ lệ
Kiểm định giả thuyết - Trường hợp hai mẫu
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập
Kiểm định giả thuyết so sánh hai tỷ lệ

So sánh hai mẫu không độc lập
(paired t - test)

- Miền bác bỏ và p - giá trị trong trường hợp này có dạng

Đối thuyết	Miền bác bỏ	p - giá trị
$H_1 : \mu_D \neq D_0$	$ t > t_{1-\alpha/2}^{n-1}$	$p = 2\mathbb{P}(T_{n-1} \geq t)$
$H_1 : \mu_D < D_0$	$t < -t_{1-\alpha}^{n-1}$	$p = \mathbb{P}(T_{n-1} \leq t)$
$H_1 : \mu_D > D_0$	$t > t_{1-\alpha}^{n-1}$	$p = \mathbb{P}(T_{n-1} \geq t)$
- Kết luận: Nếu bác bỏ H_0 , ta kết luận H_1 đúng với $(1 - \alpha) * 100\%$ độ tin cậy. Ngược lại kết luận chưa đủ cơ sở để bác bỏ H_0 .
- Trường hợp cỡ mẫu $n > 30$, bài toán kiểm định hai mẫu phụ thuộc thực hiện tương tự như trường hợp một mẫu dựa trên mẫu ngẫu nhiên (D_1, \dots, D_n) .

XSTK
N.T. M. Ngọc
Các khái niệm trong kiểm định giả thuyết thông kê
Giả thuyết H_0 và đối thuyết H_1
Các loại sai lầm
Giá trị p_{value}
Kiểm định giả thuyết - Trường hợp một mẫu
Kiểm định giả thuyết cho kỳ vọng
Kiểm định giả thuyết cho tỷ lệ
Kiểm định giả thuyết - Trường hợp hai mẫu
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập
Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập
Kiểm định giả thuyết so sánh hai tỷ lệ

Ví dụ:

Một bác sĩ dinh dưỡng nghiên cứu một chế độ ăn kiêng và tập thể dục mới để làm giảm lượng đường trong máu của các bệnh nhân bị bệnh tiểu đường. 10 bệnh nhân bị bệnh tiểu đường được chọn để thử nghiệm chương trình này, bảng kết quả bên dưới cho biết lượng đường trong máu trước và sau khi các bệnh nhân tham gia chương trình

Trước	268	225	252	192	307	228	246	298	231	185
Sau	106	186	223	110	203	101	211	176	194	203

Số liệu được cung cấp có đủ bằng chứng để kết luận rằng chế độ ăn kiêng và tập thể dục có tác dụng làm giảm lượng đường trong máu không? $\alpha = 0.05$.

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

Giải ví dụ

Gọi (X_{1i}, X_{2i}) với $i = 1, 2, \dots, n$ lần lượt là lượng đường trong máu trước và sau khi thực hiện chế độ ăn kiêng và tập thể dục của các bệnh nhân tiểu đường. Khi đó, $D_i = X_{1i} - X_{2i}$ là chênh lệch lượng đường trong máu trước và sau khi thực hiện chế độ ăn kiêng và tập thể dục của bệnh nhân tiểu đường thứ i . Trong mẫu thực nghiệm, ta tính được các giá trị d_i của D_i : 162, 39, 29, 82, 104, 127, 35, 122, 37, -18. Từ đó, tính

$$\bar{d} = \frac{1}{10} \sum_{i=1}^{10} d_i = 71,9$$

$$s_d^2 = \frac{1}{10 - 1} \sum_{i=1}^{10} (d_i - \bar{d})^2 = 3153,43 \quad \text{suy ra} \quad s_d = 56,1554.$$

Đặt $\mu_D = \mathbb{E}[D_i]$, khi đó ta tiến hành các bước kiểm định:

B1: Phát biểu giả thuyết H_0 và đối thuyết H_1

$$\begin{cases} H_0 : \mu_D = 0 \\ H_1 : \mu_D > 0 \end{cases}$$

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

Giải ví dụ (tt)

B2: Tìm thống kê kiểm định:

Khi giả thuyết H_0 đúng,

$$T = \frac{\bar{D} - D_0}{S_D/\sqrt{n}} \sim T(n - 1).$$

Từ mẫu thực nghiệm, tính giá trị thống kê kiểm định:

$$t = \frac{\bar{d} - D_0}{s_D/\sqrt{n}} = \frac{71,9 - 0}{56,1554/\sqrt{10}} = 4,0489$$

B3: Xác định miền bác bỏ:

Với mức ý nghĩa $\alpha = 5\%$, bác bỏ giả thuyết H_0 khi $t > t_{1-\alpha}^{n-1} = t_{0,95}^9 = 1,8331$.

Ta thấy, $t = 4,0489 > 1,8331 = t_{0,95}^9$ nên ta bác bỏ giả thuyết H_0 với mức ý nghĩa $\alpha = 5\%$.

B4: Kết luận:

Với 95% độ tin cậy, ta kết luận rằng chế độ ăn kiêng và tập thể dục có tác dụng làm giảm lượng đường trong máu của những bệnh nhân bị tiểu đường.

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

Ví dụ khác

Một quản lý công ty taxi đang cố gắng quyết định xem việc sử dụng vỏ xe radial thay vì vỏ có lớp bổ xiên thông thường có tiết kiệm nhiên liệu hơn không. 12 xe hơi được trang bị các vỏ xe radial và được lái theo một bài kiểm tra quy chuẩn. Không thay đổi tài xế, các xe trên sau đó được trang bị vỏ có lớp bổ xiên thông thường và được lái một lần nữa theo bài kiểm tra trên. Nhiên liệu tiêu thụ, theo kilomet mỗi lít, được ghi lại như sau:

Xe	Số kilomet mỗi Lít	
	Vỏ Radial	Vỏ bổ xiên
1	4,2	4,1
2	4,7	4,9
3	6,6	6,2
4	7,0	6,9
5	6,7	6,8
6	4,5	4,4
7	5,7	5,7
8	6,0	5,8
9	7,4	6,9
10	4,9	4,7
11	6,1	6,0
12	5,2	4,9

Ta có thể kết luận rằng các xe hơi được trang bị vỏ xe radial có mức tiêu hao nhiên liệu tốt hơn những xe được trang bị vỏ bổ xiên không? Giả sử các tổng thể có phân phối chuẩn.

XSTK

N.T. M. Ngọc

Các khái niệm trong kiểm định giả thuyết thông kê

Giả thuyết H_0 và đối thuyết H_1

Các loại sai lầm

Giá trị p_{value}

Kiểm định giả thuyết - Trường hợp một mẫu

Kiểm định giả thuyết cho kỳ vọng

Kiểm định giả thuyết cho tỷ lệ

Kiểm định giả thuyết - Trường hợp hai mẫu

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập

Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập

Kiểm định giả thuyết so sánh hai tỷ lệ

Ví dụ khác

5 nhân viên bán hàng được cho đi học lớp huấn luyện. Lớp huấn luyện có tác dụng không?

Nhân viên	Số lần bị khách hàng phàn nàn	
	Trước khi học	Sau khi học
A	6	4
B	20	6
C	3	2
D	0	0
E	4	0

- Các giả định
 - Hai mẫu độc lập,
 - Cỡ mẫu lớn và $np_1 > 5$; $n(1 - p_1) > 5$ và $mp_2 > 5$; $m(1 - p_2) > 5$.

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1 - \hat{p}) \left(\frac{1}{n} + \frac{1}{m} \right)}}$$

Để kiểm định hiệu quả của một loại thuốc ngừa bệnh, người ta tiêm thuốc này lên 150 con vật thí nghiệm, ngoài ra cũng có 150 con vật thuộc nhóm thứ hai không được tiêm thuốc ngừa. Sau đó, 300 con vật này được gieo bệnh để nghiên cứu, trong nhóm đã tiêm ngừa có 10 con chết và trong nhóm không tiêm ngừa có 30 con chết. Như vậy, tiêm ngừa có làm giảm tỷ lệ chết vì bệnh của các con vật không, với $\alpha = 1\%$?

<div>XSTK</div> <div>N.T. M. Ngọc</div> <div>Các khái niệm trong kiểm định giả thuyết thông kê</div> <div>Giả thuyết H_0 và đối thuyết H_1</div> <div>Các loại sai lầm</div> <div>Giá trị p_{value}</div> <div>Kiểm định giả thuyết - Trường hợp một mẫu</div> <div>Kiểm định giả thuyết cho kỳ vọng</div> <div>Kiểm định giả thuyết cho tỷ lệ</div> <div>Kiểm định giả thuyết - Trường hợp hai mẫu</div> <div>Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập</div> <div>Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập</div> <div>Kiểm định giả thuyết so sánh hai tỷ lệ</div>	<div>Giải ví dụ:</div>
--	------------------------

<div>XSTK</div> <div>N.T. M. Ngọc</div> <div>Các khái niệm trong kiểm định giả thuyết thông kê</div> <div>Giả thuyết H_0 và đối thuyết H_1</div> <div>Các loại sai lầm</div> <div>Giá trị p_{value}</div> <div>Kiểm định giả thuyết - Trường hợp một mẫu</div> <div>Kiểm định giả thuyết cho kỳ vọng</div> <div>Kiểm định giả thuyết cho tỷ lệ</div> <div>Kiểm định giả thuyết - Trường hợp hai mẫu</div> <div>Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu độc lập</div> <div>Kiểm định giả thuyết so sánh hai trung bình - Trường hợp hai mẫu không độc lập</div> <div>Kiểm định giả thuyết so sánh hai tỷ lệ</div>	<div>Ví dụ khác</div> <div>Một công ty sản xuất thuốc cần kiểm tra một loại thuốc có tác dụng là giảm việc xuất hiện cơn đau ngực ở các bệnh nhân. Công ty thực hiện thí nghiệm trên 400 người, chia làm hai nhóm: nhóm 1 gồm 200 được uống thuốc và nhóm 2 gồm 200 người được uống giả dược. Theo dõi thấy ở nhóm 1 có 8 người lên cơn đau ngực và nhóm 2 có 25 người lên cơn đau ngực. Với $\alpha = 0.05$, hay cho kết luận về hiệu quả của thuốc mới sản xuất.</div>
--	---