**1. CÁC DẠNG TOÁN**

1.1 Bảng số liệu của một sample gồm 100 ngày bất kì của 5 quốc gia về tỉ lệ chết do Covid19 và tổng số vaccine trên 100 người sử dụng cho các dạng toán.

1.2 Xây dựng khoảng tin cậy cho giá trị trung bình của tổng thể khi không biết phương sai.

1.3 Kiểm đinh giả thuyết với 1 tổng thể phương sai chưa biết.

1.4 Kiểm đinh giả thuyết với 1 tổng thể với propotion.

1.5 Kiểm định giả thuyết với 2 tổng thể với phương sai đã biết.

1.6 Kiểm định giả thuyết với 2 tổng thể với propotion.

1.7 Hồi quy tuyến tính

**2. Thống kê**

1, Việt Nam

|  |  |  |  |
| --- | --- | --- | --- |
| ***Death Rate*** |  | ***Vaccine per 100 people*** |  |
|  |  |  |  |
| Mean | 0.0162901 | Mean | 29.558263 |
| Standard Error | 0.0003682 | Standard Error | 2.4514151 |
| Median | 0.0170971 | Median | 3.59 |
| Mode | 0.0226098 | Mode | 0 |
| Standard Deviation | 0.0070344 | Standard Deviation | 46.318124 |
| Sample Variance | 4.948E-05 | Sample Variance | 2145.3686 |
| Kurtosis | -1.2451529 | Kurtosis | 0.7801284 |
| Skewness | -0.3440077 | Skewness | 1.4945964 |
| Range | 0.0214309 | Range | 154.86 |
| Minimum | 0.0036865 | Minimum | 0 |
| Maximum | 0.0251173 | Maximum | 154.86 |
| Sum | 5.9458686 | Sum | 10552.3 |
| Count | 365 | Count | 357 |

2,Italy

|  |  |  |  |
| --- | --- | --- | --- |
| ***Death rate*** |  | ***Vaccine per 100 people*** |  |
|  |  |  |  |
| Mean | 0.0298106 | Mean | 85.253753 |
| Standard Error | 0.0001393 | Standard Error | 3.2571345 |
| Median | 0.0298888 | Median | 90.26 |
| Mode | #N/A | Mode | #N/A |
| Standard Deviation | 0.002662 | Standard Deviation | 62.227467 |
| Sample Variance | 7.086E-06 | Sample Variance | 3872.2576 |
| Kurtosis | -0.116252 | Kurtosis | -1.5483604 |
| Skewness | 0.3267957 | Skewness | -0.0274143 |
| Range | 0.0126131 | Range | 188.53 |
| Minimum | 0.0224305 | Minimum | 0.09 |
| Maximum | 0.0350436 | Maximum | 188.62 |
| Sum | 10.880854 | Sum | 31117.62 |
| Count | 365 | Count | 365 |

3.Singapore

|  |  |  |  |
| --- | --- | --- | --- |
| ***Death rate*** |  | ***Vaccine per 100 people*** |  |
|  |  |  |  |
| Mean | 0.0010035 | Mean | 100.28416 |
| Standard Error | 4.279E-05 | Standard Error | 3.7721466 |
| Median | 0.0005694 | Median | 100.21 |
| Mode | 0.0006379 | Mode | 0.01 |
| Standard Deviation | 0.0008175 | Standard Deviation | 72.066759 |
| Sample Variance | 6.682E-07 | Sample Variance | 5193.6178 |
| Kurtosis | 0.5670981 | Kurtosis | -1.5525684 |
| Skewness | 1.4800113 | Skewness | -0.051979 |
| Range | 0.0024845 | Range | 209.24 |
| Minimum | 0.0004827 | Minimum | 0.01 |
| Maximum | 0.0029672 | Maximum | 209.25 |
| Sum | 0.3662784 | Sum | 36603.72 |
| Count | 365 | Count | 365 |

-Có thể thấy , tỉ lệ tử vong do Covid-19 của Singapore trong năm 2021 khác biệt do với các nước khác.

-Theo số liệu thống kê, tỉ lệ tử vong trung bình của Singapore là 0.0010035%, cao nhất là 0.0029672% và thấp nhất là 0.0004827% với Standard Deviation là 0.0008175.

- Tỉ lệ tử vong của Singapore ¾ giai đoạn đầu năm 2021 tỉ lệ tử vongthấp hơn 0.001, sau đó thì tăng nhanh đến mức cao nhất là 0.0029672. Vaccine rate: Có thể thấy , tỉ lệ tiêm vaccine Covid-19 của Singapore năm 2021 khá cao và tăng dần đều.

Theo số liệu thống kê, số liều vaccine trung bình của Singapore là 100.28416 liều trên mỗi 100 người, Standard Deviation là 72.066759. Mức thấp nhất là đầu năm 2021 là 0.01 liều trên mỗi 100 người và tăng mạnh đạt đến mức cao nhất là 209.25 liều trên mỗi 100 người vào cuối năm 2021. Từ hai biểu đồ trên ta có thể kết luận năm 2021 Singapore là quốc gia có số ca dương tính cũng như tử vong khá thấp. Cùng với đó là tỉ lệ tiêm vaccine khá cao.

4, KR

|  |  |  |  |
| --- | --- | --- | --- |
| ***Death rate*** |  | ***Vaccine per 100 people*** |  |
|  |  |  |  |
| Mean | 0.0233446 | Mean | 81.43526 |
| Standard Error | 0.0001735 | Standard Error | 3.0345771 |
| Median | 0.0240623 | Median | 94.16 |
| Mode | #N/A | Mode | #N/A |
| Standard Deviation | 0.0033139 | Standard Deviation | 57.975515 |
| Sample Variance | 1.098E-05 | Sample Variance | 3361.1603 |
| Kurtosis | 0.0039284 | Kurtosis | -1.4489766 |
| Skewness | -0.5730695 | Skewness | -0.0686878 |
| Range | 0.0131665 | Range | 184.31 |
| Minimum | 0.0156529 | Minimum | 0.28 |
| Maximum | 0.0288194 | Maximum | 184.59 |
| Sum | 8.5207712 | Sum | 29723.87 |
| Count | 365 | Count | 365 |

=> Theo số liệu thống kê, có thể thấy rằng tỉ lệ tử vong do COVID-19 của Hàn Quốc trong năm 2021 có xu hướng giảm đáng kể.

- Tỉ lệ tử vong trung bình của Hàn Quốc là 0.0233446%, đạt mức cao nhất là 0.0288194% và thấp nhất là 0.0156529%, có độ lệch chuẩn nhỏ là 0.0033139

- Đầu năm 2021 tỉ lệ tử vong của Hàn Quốc tăng cao đạt lên 0.018% vào ngày thứ 28-55 nhưng sau đó giảm dần xuống mức 0.008% đến cuối năm , tỉ lệ tiêm Vaccine thì ngược lại, có xu hướng thấp ở đầu năm và tăng cao về sau này.

- Theo thống kê, số liều vaccine trung bình của Hàn Quốc là 81.43526 liều trên mỗi 100 người, độ lệch chuẩn là 57.975515. Tỉ lệ thấp nhất là đầu năm 2021 là 0 liều trên mỗi 100 người và sau đó tăng mạnh đến mức cao nhất là 184.59 liều trên mỗi 100 người vào khoảng thời gian từ giữa năm đến cuối năm.

- Những số liệu này đã chứng minh rằng dịch bệnh COVID-19 đang phát triển theo hướng tích cực và khả quan hơn cùng với tỉ lệ giảm của các ca nhiễm bệnh là sự tăng cao về tỉ lệ tiêm vắc xin phòng bệnh.

5,Germany

|  |  |  |  |
| --- | --- | --- | --- |
| ***Death rate*** |  | ***Vaccine per 100 people*** |  |
|  |  |  |  |
| Mean | 0.0124545 | Mean | 61.741753 |
| Standard Error | 0.0002011 | Standard Error | 3.4178621 |
| Median | 0.0127085 | Median | 38.18 |
| Mode | #N/A | Mode | 0 |
| Standard Deviation | 0.0038422 | Standard Deviation | 65.298163 |
| Sample Variance | 1.476E-05 | Sample Variance | 4263.8501 |
| Kurtosis | -1.5585957 | Kurtosis | -1.0862884 |
| Skewness | 0.1325861 | Skewness | 0.6709344 |
| Range | 0.010484 | Range | 200.61 |
| Minimum | 0.0077381 | Minimum | 0 |
| Maximum | 0.0182221 | Maximum | 200.61 |
| Sum | 4.5458994 | Sum | 22535.74 |
| Count | 365 | Count | 365 |

DeadRate:

Có thể thấy được tỷ lệ tử vong do Covid-19 của Đức tăng cao và chạm đỉnh vào quý 1 của năm 2021 và có xu hướng giảm dần sau từng quý.

* Theo số liệu thống kê, tỉ lệ tử vong trung bình của Đức là 0.0124545%,  cao nhất là 0.0182221% và thấp nhất là 0.0077381% với Standard Deviation là 0.0038422.

Vaccine rate:

* Có thể thấy , tỉ lệ tiêm vaccine Covid-19 của Đức có chuyển biến tích cực trong suốt năm 2021  với số liều ngày càng tăng.
* Theo số liệu thống kê, số liều vaccine trung bình của Italy là 61.741753 liều trên mỗi 100 người, Standard Deviation là 65.298163. Mức thấp nhất là đầu năm 2021 là 0 liều trên mỗi 100 người  và tăng mạnh đạt đến mức cao nhất là 200.61 liều trên mỗi 100 người  vào cuối năm 2021

=> có thể thấy Đức là một nước có tỉ lệ tử vong do Covid 19 khá cao vào đầu năm 2021, khi mà người dân chưa tin tưởng vào vaccine và tỉ lệ những người đã được tiêm vaccine đầu năm của Đức là bằng 0. Nhưng sau đó, khi tỉ lệ người dân tiêm vaccine tăng dần thì tỉ lệ tử vong do Covid-19 cũng giảm 1 cách đáng kể.

**2. CÁC DẠNG TOÁN**

Bảng số liệu của một **sample** gồm 100 ngày bất kì của 5 quốc gia về tỉ lệ chết do Covid19 và tổng số vaccine trên 100 người , sử dụng chung cho các dạng toán ở dưới :

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | ***DEATH RATE*** |  |  |  |
|  | ***Vietnam*** | ***Italy*** | **Singapore** | ***Germany*** | ***Korea*** |
| Mean | 0.0160415 | 0.0296723 | 0.0009901 | 0.0122736 | 0.023137 |
| Standard Error | 0.0006609 | 0.0002781 | 8.584E-05 | 0.0003697 | 0.0003414 |
| Median | 0.0166002 | 0.0299015 | 0.0005389 | 0.0126604 | 0.0240972 |
| Mode | #N/A | #N/A | #N/A | #N/A | #N/A |
| Standard Deviation | 0.0066088 | 0.0027807 | 0.0008584 | 0.0036973 | 0.0034142 |
| Sample Variance | 4.368E-05 | 7.732E-06 | 7.368E-07 | 1.367E-05 | 1.166E-05 |
| Kurtosis | -1.1003723 | -0.0908589 | 0.56745 | -1.4163146 | -0.1157976 |
| Skewness | -0.3265185 | 0.0414298 | 1.5225288 | 0.2211367 | -0.623614 |
| Range | 0.0212352 | 0.0120041 | 0.0024808 | 0.0104384 | 0.0130833 |
| Minimum | 0.0038821 | 0.0229455 | 0.0004827 | 0.0077589 | 0.0156983 |
| Maximum | 0.0251173 | 0.0349496 | 0.0029635 | 0.0181973 | 0.0287816 |
| Sum | 1.6041547 | 2.9672314 | 0.099007 | 1.2273641 | 2.3137019 |
| Count | 100 | 100 | 100 | 100 | 100 |
| Confidence Level(95.0%) | 0.0013113 | 0.0005517 | 0.0001703 | 0.0007336 | 0.0006775 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | ***Vaccine per 100 people*** |  |  |  |
|  | ***Vietnam*** | ***Italy*** | ***Singapore*** | ***Germany*** | ***Korea*** |
|  |  |  |  |  |  |
| Mean | 31.451616 | 85.0254 | 91.9831 | 62.5993 | 79.6515 |
| Standard Error | 5.0627219 | 6.3795029 | 7.2255584 | 6.4485445 | 5.9760327 |
| Median | 1.69 | 85.56 | 77.75 | 38.2 | 90.78 |
| Mode | 0 | #N/A | #N/A | 0 | #N/A |
| Standard Deviation | 50.373447 | 63.795029 | 72.255584 | 64.485445 | 59.760327 |
| Sample Variance | 2537.4841 | 4069.8058 | 5220.8694 | 4158.3726 | 3571.2967 |
| Kurtosis | 0.4187682 | -1.4977887 | -1.4727009 | -0.9553865 | -1.4421024 |
| Skewness | 1.4374174 | 0.0876676 | 0.2279674 | 0.7236436 | 0.0354756 |
| Range | 152.05 | 187.93 | 208.74 | 200.61 | 184.18 |
| Minimum | 0 | 0.22 | 0.01 | 0 | 0.34 |
| Maximum | 152.05 | 188.15 | 208.75 | 200.61 | 184.52 |
| Sum | 3113.71 | 8502.54 | 9198.31 | 6259.93 | 7965.15 |
| Count | 99 | 100 | 100 | 100 | 100 |
| Confidence Level(95.0%) | 10.046807 | 12.658318 | 14.337075 | 12.795311 | 11.857745 |

2.1: Xây dựng khoảng tin cậy trên 1 **sample** **(x̄)** về **giá trị trung bình** với **phương sai(σ)** chưa biết

Bài Toán 1 : 5 tổng thể là số liệu thống kê về tỉ lệ chết do Covid19 của 5 nước Việt Nam, Italy, Singapore ,Korea và Đức năm 2021 theo từng ngày. Ta lấy ngẫu nhiên 1 sample gồm 100 phần tử của từng tổng thể của 5 nược và giá trị trung bình tỉ lệ chết do Covid19 của 5 quốc gia Việt Nam, Italy, Singapore ,Korea và Đức lần lượt là 0.0160415, 0.0296723, 0.0009901, 0.0122736, 0.023137 và độ lệch chuẩn của mẫu của 5 quốc gia trên lần lượt là 0.00793815, 0.001553463 , 0.00088524, 0.003057031, 0.000107124.

Hãy xây dựng khoảng tin cậy với độ tin cậy 95% cho giá trị trung bình của 5 tổng thể của 5 quốc gia.

Giải:

Sau khi Sử dụng data analysis trong excel ta tính được Confidence level(95%)

của 5 quốc gia Việt Nam, Italy, Singapore ,Korea và Đức lần lượt là 0.0013113, 0.0005517, 0.0001703, 0.0007336 và 0.0006775.

Hoặc Áp dụng công thức :

* Khoảng tin cây là ()

-Khoảng tin cậy 95% của giá trị trung bình của tổng thể tỉ lệ chết do Covid19 của Việt Nam là:

(0.0147302 ; 0.0173528 ).

-Khoảng tin cậy 95% của giá trị trung bình của tổng thể tỉ lệ chết do Covid19 của Italy là:

(0.0291206; 0.030224).

-Khoảng tin cậy 95% của giá trị trung bình của tổng thể tỉ lệ chết do Covid19 của Singapore là:

(0.008198; 0.0011604).

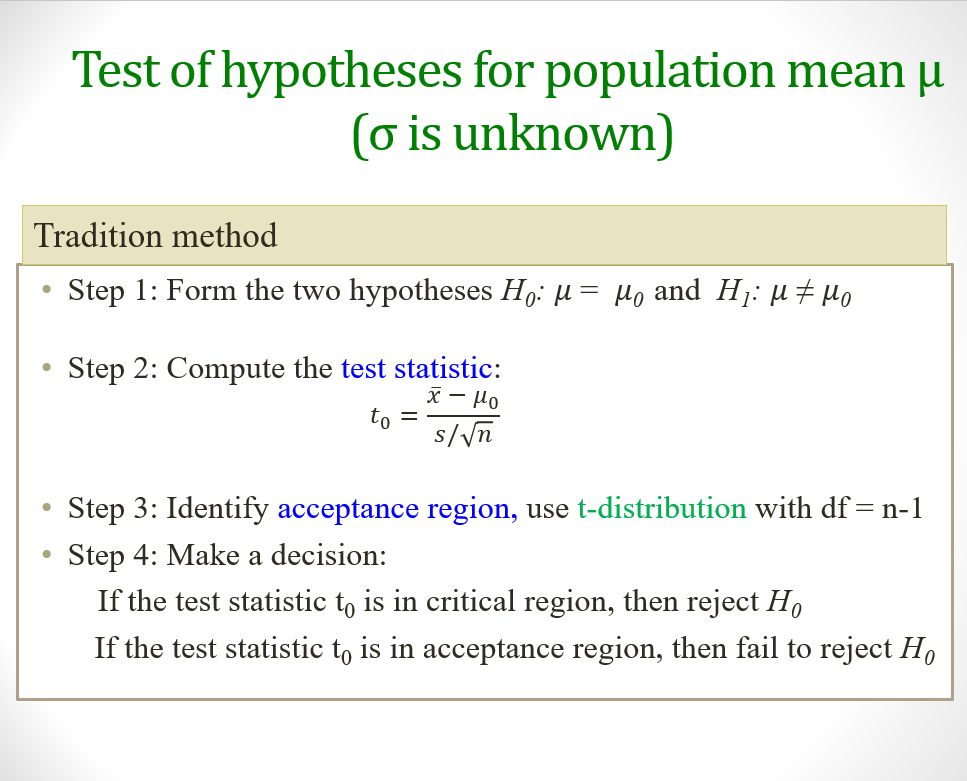
-Khoảng tin cậy 95% của giá trị trung bình của tổng thể tỉ lệ chết do Covid19 của Hàn Quốc là:

(0.01154; 0.0130072).

-Khoảng tin cậy 95% của giá trị trung bình của tổng thể tỉ lệ chết do Covid19 của Đức là:

(0.0224595;0.0238145).

2.2: Kiểm định giả thuyết trên 1 **sample** **(x̄)** về **giá trị trung bình** của tổng thể với **phương sai(σ)** chưa biết. ( so sánh với P-value)



**Xét mẫu thử về tỷ lệ tử vong của 100 ngày ngẫu nhiên từ Việt Nam**

- Kiểm định giả thuyết H0: μ = 0.01 so với H1: μ ≠ 0.01

Given:

T-Statistic:

* *Reject H0*
* Tỉ lệ tử vong của Việt Nam trong năm 2021 sẽ không thể bằng 0.01%

**Xét mẫu thử về vaccination của 100 ngày ngẫu nhiên từ Việt Nam**

- Kiểm định giả thuyết H0: μ = 25 so với H1: μ ≠ 25

Given:

T-Statistic:

* *Accept H0*
* Tỉ lệ tiêm Vaccine của Việt Nam trong năm 2021 là khoảng 25 liều/100 người

**Xét mẫu thử về tỷ lệ tử vong của 100 ngày ngẫu nhiên từ Italy**

- Kiểm định giả thuyết H0: μ = 0.030 so với H1: μ ≠ 0.030

Given:

T-Statistic:

* *Accept H0*
* Tỉ lệ tử vong của Italy trong năm 2021 là khoảng 0.030 %

**Xét mẫu thử về vaccination của 100 ngày ngẫu nhiên từ Italy**

- Kiểm định giả thuyết H0: μ = 86 so với H1: μ ≠ 86

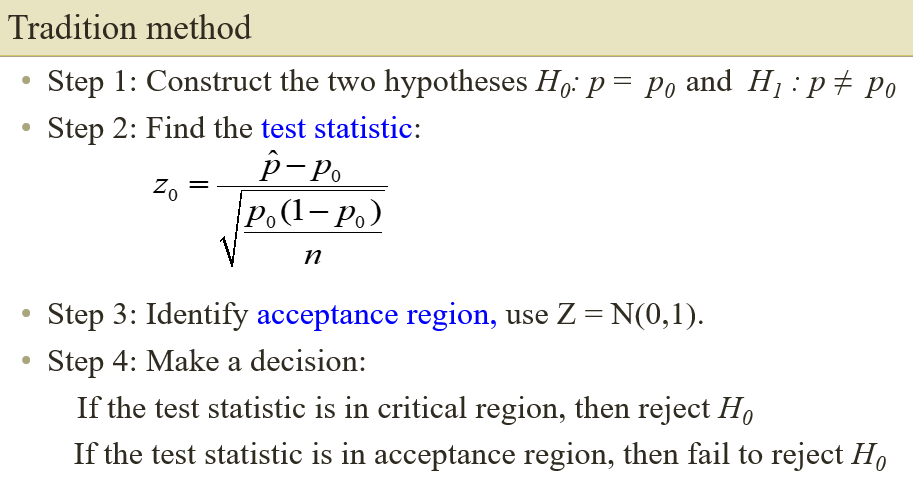
Given: 63.795029

T-Statistic:

* *Accept H0*
* Tỉ lệ tiêm Vaccine của Việt Nam trong năm 2021 là khoảng 86 liều/100 người

2.3: Kiểm định giả thuyết trên 1 **sample** **(x̄)** về **propotion(P)** của tổng thể

**Xét mẫu thử về tỷ lệ tử vong của 100 ngày ngẫu nhiên từ Vietnam**

****

Bài toán ví dụ : Tỉ lệ người đi tiêm vaccine tại Việt Nam đạt 40%. Có ý kiến cho rằng tỉ lệ này thấp do người dân lười đi ra ngoài. Theo dõi ngẫu nhiên 100 người thì có 30 người đi tiêm. Kết luận ý kiến trên với độ tin cậy 95%

- Kiểm định giả thuyết H0: p=0.4 so với H1: p ≠ 0.4

- 2 tail=> alpha/2 = 0.025= z=1.96

-p ̂=30/100=0.3

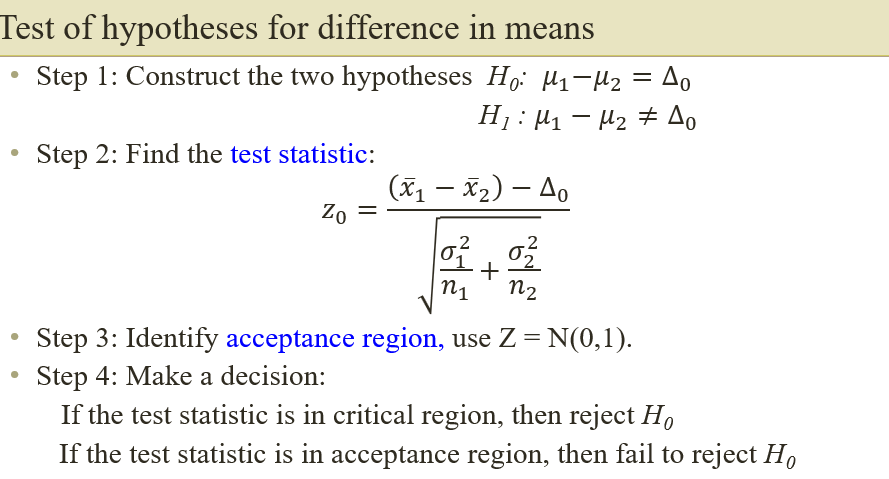
=> miền chấp nhận ( -∞,-1.96) or (1.96,+ ∞)

Z-Statistic: -2.041, nằm ngoài miền chấp nhận

* Reject *H0* , accept H1 : p ≠ 0.4
* Tỉ lệ người dân đi tiêp vaccine tại Việt Nam không đạt 40%

2.4: Kiểm định giả thuyết với 2 sample , **phương sai(σ)** đã biết.

***1. Sử dụng excel***



Excel: Data -> z-test : Two Sample for Means

Ý kiến 1: Tỷ lệ tiêm vaccine ở Việt Nam lớn hơn tỷ lệ tiêm vaccine ở Korea

alpha = 0.05

H0: μVN = μKR

H1: μVN > μKR ( right tail)

|  |  |  |
| --- | --- | --- |
| **z-Test: Two Sample for Means** |  |  |
|  |  |  |
|  | *Total vaccine VN* | *Total vaccine KR* |
| Mean | 31.451616 | 62.5993 |
| Known Variance | 46.31812 | 65.298163 |
| Observations | 99 | 100 |
| Hypothesized Mean Difference | 0 |  |
| z | -29.420745 |  |
| P(Z<=z) one-tail | 0 |  |
| z Critical one-tail | 1.6448536 |  |
| P(Z<=z) two-tail | 0 |  |
| z Critical two-tail | 1.959964 |  |

Ta có : z < z Critical one-tail

* Accept H0, reject H1

-Ở độ tin cậy =95%, không có đủ bằng chứng cho thấy tỉ lệ tiêm vaccine ở Việt Nam lớn hơn tỉ lệ tiêm vaccine ở KR

Ý kiến 2: Tỷ lệ tiêm vaccine ở Singapore nhỏ hơn tỷ lệ tiêm vaccine ở Đức

H0: μS = μGE

H1: μS < μGE (left- tail)

|  |  |  |
| --- | --- | --- |
| **z-Test: Two Sample for Means** |  |  |
|  |  |  |
|  | *Total vaccine Sing* | *Total vaccine Ger* |
| Mean | 91.9831 | 79.6515 |
| Known Variance | 72.06676 | 57.97551 |
| Observations | 100 | 100 |
| Hypothesized Mean Difference | 0 |  |
| z | 10.813771 |  |
| P(Z<=z) one-tail | 0 |  |
| z Critical one-tail | 1.6448536 |  |
| P(Z<=z) two-tail | 0 |  |
| z Critical two-tail | 1.959964 |  |

Ta có : z> z Criticle one-tail(-1.645)

=>Reject H0, accept H1

-Ở độ tin cậy =95%, không có đủ bằng chứng cho thấy tỉ lệ tiêm vaccine ở Singapore lớn hơn tỉ lệ tiêm vaccine ở Đức

Ý kiến 3: Tỷ lệ tiêm vaccine ở Italy bằng tỷ lệ tiêm vaccine ở Đức

H0: μItaly = μGE

H1: μItaly ≠ μGE (two-tail)

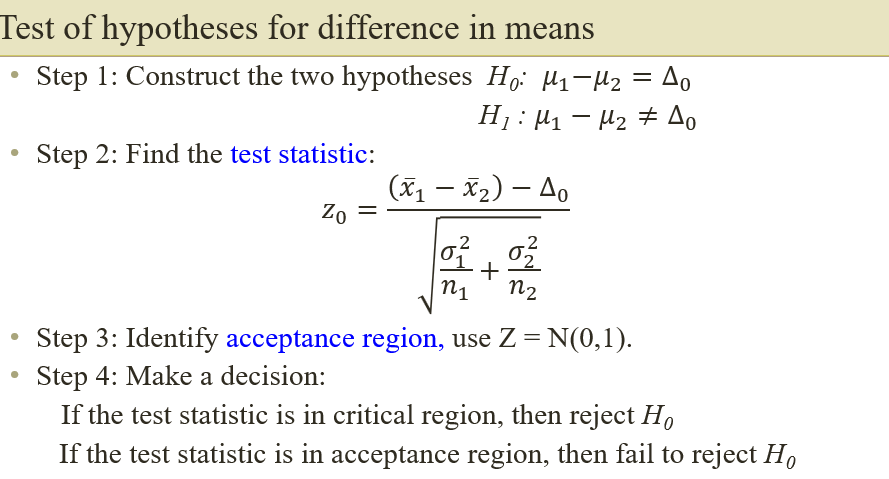
|  |  |  |
| --- | --- | --- |
| **z-Test: Two Sample for Means** |  |  |
|  |  |  |
|  | *Total vaccine Italy* | *Total vaccine Ger* |
| Mean | 85.0254 | 79.6515 |
| Known Variance | 62.22747 | 57.97551 |
| Observations | 100 | 100 |
| Hypothesized Mean Difference | 0 |  |
| z | 4.9015334 |  |
| P(Z<=z) one-tail | 4.755E-07 |  |
| z Critical one-tail | 1.6448536 |  |
| P(Z<=z) two-tail | 9.509E-07 |  |
| z Critical two-tail | 1.959964 |  |

Ta có : z> z Critical two tail

=>Reject H0, accept H1

-Ở độ tin cậy =95%, không có đủ bằng chứng cho thấy tỉ lệ tiêm vaccine ở Italy = tir lệ tiêm vaccine ở Đức

***2. Sử dụng tính toán thông thường***



Bài toán 1: Tỷ lệ tử vong do covid ở Việt Nam cao hơn Italy

H0: μVN = μItaly

H1: μVN >μItaly  (right)

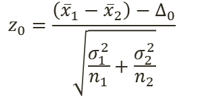
Lấy sample 100 ngày tương ứng có X̄VN = 0.016, X̄Italy = 0.03.

Biết σVN = 0.007, σItaly = 0.002662 với độ tin cậy là 95%.

Ta có :

X̄VN = 0.016, X̄Italy = 0.03, σVN = 0.007, σMA = 0.002662, α = 0.05

H0 : μVN – μItaly = 0; H1 : μVN – μItaly > 0

Áp dụng công thức  ta có : Z0 =-18.7 < Zα = 1.645

=>Accept H0,reject H1

=> với mực độ tin cậy là 95%, ko có đủ bằng chứng cho thấy tỉ lệ tử vong do covid ở Việt Nam cao hơn Italy

Bài toán 2: Tỷ lệ tử vong do covid ở Korea nhỏ hơn Việt Nam

H0: μKR = μVN

H1: μKR <μVN  (lèt tail)

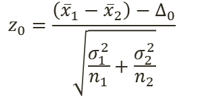
Lấy sample 100 ngày tương ứng có X̄VN = 0.016, X̄KR = 0.023.

Biết σVN = 0.007, σKR = 0.003842 với độ tin cậy là 95%.

Ta có :

X̄KR = 0.023, X̄VN = 0.016, σKR = 0.003842, σVN = 0.007, α = 0.05

H0 : μVN – μKR = 0; H1 : μVN – μKR < 0

Áp dụng công thức  ta có : Z0 =8.766386267 >-Zα = -1.645

=> Accept H0,reject H1

=> với mực độ tin cậy là 95%, ko có đủ bằng chứng cho thấy tỉ lệ tử vong do covid ở Korea thấp hơn VN

Bài toán 3: Có giả thiết cho rằng : tỉ lệ chết do covid ở Đức khác Singapore.

H0: μGE = μSing

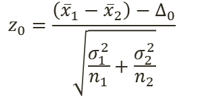
H1: μGE ≠ μSing

Lấy mẫu các mẫu ngẫu nhiên 100 ngày tương ứng

Có X̄GE = 0.0123, X̄Sing =0.001.

Ta có σGE = 0.0033, σSing = 0.0008 độ tin cậy là 95%.

H0 : μGE– μ Sing = 0; H1 : μGE – μ Sing ≠ 0

Áp dụng công thức  ta có : Z0 =33.28 >Zα = 1.96

=>Reject H0, Accept H1

với mức độ tin cậy là 95%, có chứng cứ ủng hộ tỷ lệ chết do covid ở Đức khác Singapore

2.4: : Kiểm định giả thuyết trên 2 **sample** **(x̄)** về **propotion(P)**

Bài toán: Các nhà nghiên cứu muốn kiểm tra tác dụng phụ của vắc xin COVID tại Việt nam. Trong nghiên cứu lâm sàng, 62 trong số 100 người dùng vắc xin A1 báo cáo các phản ứng phụ. Trong khi 48 người trong số 100 người dùng vắc xin A2 báo cáo các phản ứng phụ. Ở độ tin cậy 95%, vắc xin A1 có hoạt động giống với A2 không?

Claim:

H0: P1 = P2

H1: P1 ≠ P2

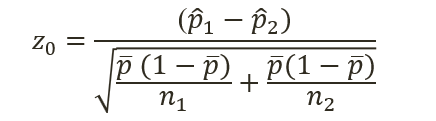
Ta có:

+p ̂1=62/100=0.62

+p ̂2=48/100=0.48

H0: P1 - P2=0 , H1: P1 - P2≠0

+P̄= 0.55

Áp dụng công thức ta có:

Z0=1.989 > Zα = 1.96

=>Reject H0, accept H1

=>Ở độ tin cậy là 95%, không có bằng chứng nào cho thấy vaccine A1 hoạt động giống với A2

2.5: Hồi quy tuyến tính

Bài toán : Một mẫu 100 ngày số liệu về số người chết và số người được tiêm vaccine từ 5 quốc gia: Việt Nam, Italy, Singapore ,Hàn Quốc và Đức**. Independent variable(X)** là tỉ lệ số người được tiêm vaccine trên 100 người, **Dependent variable(Y)** là tỉ lệ số người chết. Liệu có sự tương quan nào giữa tỉ lệ tiêm vaccine và tỉ lệ số người chết không?

Chỉ số correlation thể hiện mức độ hiệu quả của việc tiêm vaccine trên 100 người và tỉ lệ chết trong 1 khoảng thời gian. Tức là chỉ số càng đi về -1 thể hiện tỉ lệ tiêm vaccine của nước đó càng hiệu quả, ngược lại càng đi về 1 thể hiện mức độ không hiệu quả của quốc gia đó.

1.Việt Nam.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT | |  |  |  |  |
|  |  |  |  |  |  |
| *Regression Statistics* | |  |  |  |  |
| Multiple R | 0.4577276 |  |  |  |  |
| R Square | 0.2095146 |  |  |  |  |
| Adjusted R Square | 0.2014484 |  |  |  |  |
| Standard Error | 0.0059058 |  |  |  |  |
| Observations | 100 |  |  |  |  |
|  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |
| Regression | 1 | 0.0009059 | 0.000906 | 25.974 | 1.68391E-06 |
| Residual | 98 | 0.003418 | 3.49E-05 |  |  |
| Total | 99 | 0.004324 |  |  |  |
|  |  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* |  |
| Intercept | 0.0141205 | 0.0007006 | 20.15432 | 1E-36 |  |
| X Variable 1 | 5.991E-05 | 1.175E-05 | 5.096514 | 2E-06 |  |

* =0.0141205, = 5.991\*,  **=**0.0059058
* = 0.000905, = 0.003418, = 0.004324
* Trendline: y = 0.0141 + 6\*10-5\*x

Chỉ số correlation giữa tỉ lệ tiêm vaccine trên 100 người và tỉ lệ chết ở Việt Nam là 0.4577276

* Mức độ tiêm vaccine của Việt Nam chưa được hiệu quả.

2.Italy.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT | |  |  |  |  |
|  |  |  |  |  |  |
| *Regression Statistics* | |  |  |  |  |
| Multiple R | 0.9108935 |  |  |  |  |
| R Square | 0.829727 |  |  |  |  |
| Adjusted R Square | 0.8279895 |  |  |  |  |
| Standard Error | 0.0011533 |  |  |  |  |
| Observations | 100 |  |  |  |  |
|  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |
| Regression | 1 | 0.0006351 | 0.000635 | 477.55 | 1.8664E-39 |
| Residual | 98 | 0.0001303 | 1.33E-06 |  |  |
| Total | 99 | 0.0007655 |  |  |  |
|  |  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* |  |
| Intercept | 0.0330481 | 0.0001928 | 171.429 | 3E-123 |  |
| X Variable 1 | -3.97E-05 | 1.817E-06 | -21.8528 | 2E-39 |  |

* =0.0330481, = -3.97\*,  **=**0.0011533
* = 0.0006351, = 0.0001303, = 0.0007655
* Trendline: y = 0.033 - 3.97E \*10-5\*x

Chỉ số correlation giữa tỉ lệ tiêm vaccine trên 100 người và tỉ lệ chết ở Italy là -0.9108935

* Mức độ tiêm vaccine của Italy rất hiệu quả.

3.Singapore.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT | |  |  |  |  |
|  |  |  |  |  |  |
| *Regression Statistics* | |  |  |  |  |
| Multiple R | 0.790718197 |  |  |  |  |
| R Square | 0.625235267 |  |  |  |  |
| Adjusted R Square | 0.621411137 |  |  |  |  |
| Standard Error | 0.000528141 |  |  |  |  |
| Observations | 100 |  |  |  |  |
|  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |
| Regression | 1 | 4.56048E-05 | 4.56E-05 | 163.49739 | 1.31464E-22 |
| Residual | 98 | 2.73354E-05 | 2.789E-07 |  |  |
| Total | 99 | 7.29402E-05 |  |  |  |
|  |  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* |  |
| Intercept | 0.000126049 | 8.57633E-05 | 1.4697357 | 0.1448365 |  |
| X Variable 1 | 9.39326E-06 | 7.34617E-07 | 12.78661 | 1.315E-22 |  |

* =0.000126049, = 9.39326E-06,  **=**0.000528141
* = 4.56048E-05, = 2.73354E-05, = 7.29402E-05
* Trendline: y = 0.00012 - 9.4E-06\*x

Chỉ số correlation giữa tỉ lệ tiêm vaccine trên 100 người và tỉ lệ chết ở Italy là 0.790718197

* Mức độ tiêm vaccine của Singapore kém hiệu quả.

4.Korea

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT | |  |  |  |  |
|  |  |  |  |  |  |
| *Regression Statistics* | |  |  |  |  |
| Multiple R | 0.8876912 |  |  |  |  |
| R Square | 0.7879957 |  |  |  |  |
| Adjusted R Square | 0.7858324 |  |  |  |  |
| Standard Error | 0.001711 |  |  |  |  |
| Observations | 100 |  |  |  |  |
|  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |
| Regression | 1 | 0.001066 | 0.0011 | 364.2547393 | 8.845E-35 |
| Residual | 98 | 0.000287 | 3E-06 |  |  |
| Total | 99 | 0.001353 |  |  |  |
|  |  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* |  |
| Intercept | 0.0154597 | 0.000239 | 64.672 | 3.44554E-82 |  |
| X Variable 1 | -5.09E-05 | 2.67E-06 | -19.085 | 8.84478E-35 |  |

* =0.0154597, = -5.09E-05,  **=**0.001711
* = 0.001066, = 0.000287, = 0.001353
* Trendline: y = 0.0154597 -5.09E-05\*x

Chỉ số correlation giữa tỉ lệ tiêm vaccine trên 100 người và tỉ lệ chết ở Italy là -0.8876912

* Mức độ tiêm vaccine của Korea rất hiệu quả.

5.Germany

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT | |  |  |  |  |
|  |  |  |  |  |  |
| *Regression Statistics* | |  |  |  |  |
| Multiple R | 0.7254903 |  |  |  |  |
| R Square | 0.5263362 |  |  |  |  |
| Adjusted R Square | 0.5215029 |  |  |  |  |
| Standard Error | 0.0023617 |  |  |  |  |
| Observations | 100 |  |  |  |  |
|  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |
| Regression | 1 | 0.000607 | 0.0006 | 108.8978234 | 1.377E-17 |
| Residual | 98 | 0.000547 | 6E-06 |  |  |
| Total | 99 | 0.001154 |  |  |  |
|  |  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* |  |
| Intercept | 0.0264384 | 0.000395 | 66.967 | 1.21747E-83 |  |
| X Variable 1 | -4.145E-05 | 3.97E-06 | -10.435 | 1.37657E-17 |  |

* =0.0264384, = -4.145E-05,  **=**0.0023617
* = 0.000607, = 0.000547, = 0.001154
* Trendline: y = 0.0264-4.145E-05\*x

Chỉ số correlation giữa tỉ lệ tiêm vaccine trên 100 người và tỉ lệ chết ở Italy là -0.7254903

* Mức độ tiêm vaccine của Germany rất hiệu quả.