

BHT Đoàn khoa MMT&TT – Training cuối kì II K16

# — TRAINING — XÁC SUẤT THỐNG KÊ



Lê Thị Ngọc Diễm  
MMCL2021

Today's  
trainer





Đặt câu hỏi tại đây

# Nội Dung



1. Vector ngẫu nhiên

2. Ước lượng

3. Kiểm định

4. Hồi quy & Tương quan





01

# Vector Ngẫu Nhiên

# Giới thiệu

Vector ngẫu nhiên rời rạc

1. Tính xác suất

2. Tính xác suất có điều kiện

3. Chứng minh độc lập

Vector ngẫu nhiên liên tục

1. Hàm mật độ đồng thời

2. Hàm mật độ thành phần

3. Tính xác suất

4. Tính xác suất có điều kiện

# Giới thiệu

- Nếu  $X$  và  $Y$  là hai biến ngẫu nhiên thì  $(X, Y)$  là một vector ngẫu nhiên.
- Phân phối của vector  $(X, Y)$  được gọi là **phân phối đồng thời**.
- Phân phối của từng biến  $X, Y$  được gọi là **phân phối thành phần**.

# Vector ngẫu nhiên rời rạc

Bảng phân phối xác suất đồng thời của X và Y

X \ Y	0	1	2
0	0.10	0.05	0.15
1	0.05	0.15	0.10
2	0.20	0.10	0.10

Tính chất:

$$\sum_{\substack{x \in \Omega(X) \\ y \in \Omega(Y)}} P_{(X,Y)}(x,y) = 1$$

Có những yêu cầu điển hình như sau:

1. Tính xác suất
2. Tính xác suất có điều kiện
3. Chứng minh độc lập

## 1. Tính xác suất

VD: Tính  $P(X + Y \geq 1)$

$$\begin{aligned}\text{Giải: } P(X + Y \geq 1) &= 1 - P(X + Y < 1) \\ &= 1 - P(X = 0, Y = 0) \\ &= 1 - 0.1 = 0.9\end{aligned}$$



# Vector ngẫu nhiên rời rạc

Bảng phân phối xác suất đồng thời của X và Y

X \ Y	Y		
	0	1	2
0	0.10	0.05	0.15
1	0.05	0.15	0.10
2	0.20	0.10	0.10

Tính chất:

$$\sum_{\substack{x \in \Omega(X) \\ y \in \Omega(Y)}} P_{(X,Y)}(x,y) = 1$$

Có những yêu cầu điển hình như sau:

2. Tính xác suất của X hoặc Y trong điều kiện của Y hoặc X.

VD: Tính  $P(Y \geq 1 | X > 1)$

$$P(X = x | Y = y) = \frac{P(X = x, Y = y)}{P_Y(y)}, \quad P_Y(y) > 0$$
$$P(Y = y | X = x) = \frac{P(X = x, Y = y)}{P_X(x)}, \quad P_X(x) > 0$$

$$\begin{aligned} P(Y \geq 1 | X > 1) &= \frac{P(Y \geq 1, X > 1)}{P(X > 1)} \\ &= \frac{P(Y = 1, X = 2) + P(Y = 2, X = 2)}{P(X = 2)} \\ &= \frac{0.1 + 0.1}{0.2 + 0.1 + 0.1} = 0.5 \end{aligned}$$



# Vector ngẫu nhiên rời rạc

Bảng phân phối xác suất đồng thời của X và Y

X \ Y	Y		
	0	1	2
0	0.10	0.05	0.15
1	0.05	0.15	0.10
2	0.20	0.10	0.10

Tính chất:

$$\sum_{\substack{x \in \Omega(X) \\ y \in \Omega(Y)}} P_{(X,Y)}(x,y) = 1$$

Có những yêu cầu điển hình như sau:

3. Hỏi X, Y có độc lập nhau không? Vì sao?

$$X, Y \text{ độc lập} \iff P_{(X,Y)}(x,y) = P_X(x) \cdot P_Y(y) \quad \forall x,y.$$

**Chứng minh X,Y phụ thuộc nhau:** Chỉ ra một cặp (x,y) sao cho

$$P_{(X,Y)}(x,y) \neq P_X(x) \cdot P_Y(y).$$

Chọn ngẫu nhiên 1 cặp  $(x, y)$

$\Rightarrow$  ví dụ chọn  $x = 1, y = 1$

Ta có:

$$P(X = 1, Y = 1) = 0.15 \quad (1)$$

$$\begin{aligned} P(X = 1).P(Y = 1) &= (0.5 + 0.15 + 0.10). (0.05 + 0.15 + 0.1) \\ &= 0.75 * 0.3 = 0.225 \quad (2) \end{aligned}$$

Từ (1), (2)  $\Rightarrow P(X = 1, Y = 1) \neq P(X = 1).P(Y = 1)$

$\Rightarrow X, Y$  không độc lập

# Vector ngẫu nhiên liên tục

Hàm mật độ đồng thời

$$f(x, y) = \begin{cases} C(x-2)e^{-y} & \text{nếu } 0 \leq x \leq 2, y \geq 0, \\ 0 & \text{nếu } (x, y) \text{ khác.} \end{cases}$$

Có những yêu cầu điển hình như sau:

1. Tìm C

2. Tìm hàm mật độ thành phần của X

3. Tính xác suất

VD: Tính  $P(X < 1)$

1. Tìm C

$$\iint_{\mathbb{R}^2} f(x, y) dx dy = 1$$

2. Tìm hàm mật độ thành phần của X

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

3. Tính xác suất

$$P((X, Y) \in D) = \iint_D f(x, y) dx dy$$



$$f(x,y) = \begin{cases} C(x-2)e^{-y} & \text{neu' } 0 \leq x \leq 2, y \geq 0 \\ 0 & \text{neu' } f(x,y) \text{ khac} \end{cases}$$

a) tìm C

$$\iint f(x,y) dx dy = 1 \Leftrightarrow \int_0^{+\infty} C(x-2)e^{-y} dy dx = 1$$

$$\Leftrightarrow C \int_0^{+\infty} e^{-y} dy \int_0^2 (x-2) dx = 1 \Leftrightarrow C \cdot 1 \cdot (-2) = 1$$

$$\Rightarrow C = -\frac{1}{2}$$

b) Tìm hàm mật độ thành phần của  $X$

$$f(x, y) = \begin{cases} -\frac{1}{2}(x-2)e^{-y} & \forall x \in [0, 2], y \geq 0 \\ 0 & \forall x \notin [0, 2], y < 0 \end{cases}$$

$$\begin{aligned} \text{TH: } x \in [0, 2], f_x(x) &= \int_{-\infty}^{+\infty} f(x, y) dy = \int_{-\infty}^0 f(x, y) dy + \int_0^{+\infty} f(x, y) dy \\ &= \int_{-\infty}^0 0 dy + \int_0^{+\infty} -\frac{1}{2}(x-2)e^{-y} dy \\ &= 0 - \frac{1}{2}(x-2) \int_0^{+\infty} e^{-y} dy = \\ &= -\frac{1}{2}(x-2) \end{aligned}$$

$$\text{TH: } x \notin [0, 2]: f(x, y) = 0 \forall y \Rightarrow f_x(x) = 0$$

$$f_x(x) = \begin{cases} -\frac{1}{2}(x-2) & \text{kh } x \in [0, 2] \\ 0 & \text{kh } x \notin [0, 2] \end{cases}$$

$$c) \text{ then } P(X < 1) = \int_{-\infty}^1 f_X(x) dx = \int_0^1 -\frac{1}{2}(x-2) dx = \frac{3}{4}$$



02

**Ước lượng**

# Nội dung Ước lượng:

○ Các bài toán ước lượng:

- Ước lượng **trung bình** tổng thể
- Ước lượng **tỷ lệ** tổng thể





## ƯỚC LƯỢNG TRUNG BÌNH TỔNG THỂ

TH1: Biết phương sai của tổng thể

$$\mu \in \left( \bar{x} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}; \bar{x} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right)$$



EX: Chiều cao của khu rừng ở độ tuổi 10 tuần có độ lệch chuẩn là:  $\sigma = 1.2cm$ . Đo ngẫu nhiên 20 cây, thấy chiều cao trung bình của 20 cây này là  $7.4cm$ . Giả sử chiều cao của khu rừng có phân phối chuẩn.

Yêu cầu: ước lượng chiều cao của cây trong rừng, với độ tin cậy 96%.

## ƯỚC LƯỢNG TRUNG BÌNH TỔNG THỂ

### TH1: Biết phương sai của tổng thể

EX: Chiều cao của khu rừng ở độ tuổi 10 tuần có độ lệch chuẩn là:  $\sigma = 1.2cm$ . Đo ngẫu nhiên 20 cây, thấy chiều cao trung bình của 20 cây này là  $7.4cm$ . Giả sử chiều cao của khu rừng có phân phối chuẩn.

Yêu cầu: ước lượng chiều cao của cây trong rừng, với độ tin cậy 96%.

Giải: Ta có:  $\sigma = 1.2cm; n = 20; \bar{x} = 7.4cm; 1 - \alpha = 0.96$

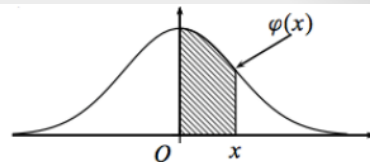
$$\text{Ta có: } \varphi\left(Z_{\frac{\alpha}{2}}\right) = \frac{0.96}{2} = 0.48 \Rightarrow Z_{\frac{\alpha}{2}} = \varphi^{-1}(0.48) = 2.06$$

Gọi  $\mu$  là chiều cao trung bình của cây trong rừng.

Khoảng ước lượng chiều cao của khu rừng là:

$$\mu \in \left( 7.4 - 2.06 \frac{1.2}{\sqrt{20}}; 7.4 + 2.06 \frac{1.2}{\sqrt{20}} \right) = (6.847; 7.953)$$

Bảng B. Giá trị  $\varphi(x) = \int_0^x f(z)dz = \frac{1-\alpha}{2}$



x	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817



## ƯỚC LƯỢNG TRUNG BÌNH TỔNG THỂ

TH2: Chưa biết phương sai của tổng thể

$$\mu \in \left( \bar{x} - t_{n-1; \frac{\alpha}{2}} \frac{s}{\sqrt{n}}; \bar{x} + t_{n-1; \frac{\alpha}{2}} \frac{s}{\sqrt{n}} \right)$$



EX: Theo dõi mức nguyên liệu hao phí để sản xuất 1 đơn vị sản phẩm, người ta thu được kết quả sau:

NLHP (gr)	19 – 19.5	19.5 – 20	20 – 20.5	20.5 – 21
Số SP	2	10	8	5

Hãy ước lượng mức nguyên liệu hao phí trung bình của cả nhà máy, với độ tin cậy 97%.

## ƯỚC LƯỢNG TRUNG BÌNH TỔNG THỂ

TH2: Chưa biết phương sai của tổng thể

EX: Theo dõi mức nguyên liệu hao phí để sản xuất 1 đơn vị sản phẩm, người ta thu được kết quả sau:

NLHP (gr)	19 – 19.5	19.5 – 20	20 – 20.5	20.5 – 21
Số SP	2	10	8	5

Hãy ước lượng mức nguyên liệu hao phí trung bình của cả nhà máy, với độ tin cậy 97%.

Giải: Từ mẫu ta có:  $n = 25$ ;  $\bar{x} = 20.07$ ;  $s = 0.454$

Độ tin cậy:  $1 - \alpha = 0.97 \Rightarrow \alpha = 0.03 \Rightarrow t_{n-1; \frac{\alpha}{2}} = t_{24; 0.015} = 2.3609$  (bảng tra Student).

Gọi  $\mu$  là mức hao phí trung bình của các sản phẩm trong nhà máy.

Khoảng ước lượng mức hao phí trung bình là:

$$\mu \in \left( 20.07 - 2.3609 \frac{0.454}{\sqrt{25}}; 20.07 + 2.3609 \frac{0.454}{\sqrt{25}} \right) = (19.856; 20.284)$$

**Bảng C. Giá trị phân phối Student  $n$  bậc tự do  $\mathbb{P}(-t_{\alpha/2}^n < T < t_{\alpha/2}^n) = \alpha$**

$\alpha / 2$ $n$	0,005	0,010	0,015	0,020	0,025	0,030	0,035	0,040
1	63,657	31,821	21,205	15,895	12,706	10,579	9,058	7,910
2	9,925	6,965	5,643	4,849	4,303	3,896	3,578	3,326
3	5,841	4,541	3,896	3,482	3,182	2,951	2,763	2,601
4	4,604	3,747	3,298	2,999	2,776	2,601	2,456	2,330
5	4,032	3,365	3,003	2,757	2,571	2,422	2,297	2,190
6	3,707	3,143	2,829	2,612	2,447	2,313	2,201	2,104
7	3,499	2,998	2,715	2,517	2,365	2,241	2,136	2,041
8	3,355	2,896	2,634	2,449	2,306	2,189	2,090	2,000
9	3,250	2,821	2,574	2,398	2,262	2,150	2,055	1,970
10	3,169	2,764	2,527	2,359	2,228	2,120	2,028	1,941
11	3,106	2,718	2,491	2,328	2,201	2,096	2,007	1,920
12	3,055	2,681	2,461	2,303	2,179	2,076	1,989	1,901
13	3,012	2,650	2,436	2,282	2,160	2,060	1,974	1,887
14	2,977	2,624	2,415	2,264	2,145	2,046	1,962	1,875
15	2,947	2,602	2,397	2,249	2,131	2,034	1,951	1,865
16	2,921	2,583	2,382	2,235	2,120	2,024	1,942	1,857
17	2,898	2,567	2,368	2,224	2,110	2,015	1,934	1,850
18	2,878	2,552	2,356	2,214	2,101	2,007	1,926	1,850
19	2,861	2,539	2,346	2,205	2,093	2,000	1,920	1,850
20	2,845	2,528	2,336	2,197	2,086	1,994	1,914	1,845
21	2,831	2,518	2,328	2,189	2,080	1,988	1,909	1,840
22	2,819	2,508	2,320	2,183	2,074	1,983	1,905	1,835
23	2,807	2,500	2,313	2,177	2,069	1,978	1,900	1,830
24	2,797	2,492	2,307	2,172	2,064	1,974	1,896	1,825
25	2,787	2,485	2,301	2,167	2,060	1,970	1,893	1,820





## ƯỚC LƯỢNG TỈ LỆ TỔNG THỂ

$$P \in \left( f - Z_{\frac{\alpha}{2}} \sqrt{\frac{f(1-f)}{n}}; f + Z_{\frac{\alpha}{2}} \sqrt{\frac{f(1-f)}{n}} \right)$$

EX: Bắt 300 con tôm trong ao nuôi 10.000 con thì có 60 con bị bệnh.

- a. Hãy ước lượng tỷ lệ tôm bị bệnh trong ao, với độ tin cậy 95%. Từ đó suy ra số con tôm bị bệnh trong ao là bao nhiêu?

## ƯỚC LƯỢNG TỈ LỆ TỔNG THỂ

EX: Bắt 300 con tôm trong ao nuôi 10.000 con thì có 60 con bị bệnh.

a. Hãy ước lượng tỷ lệ tôm bị bệnh trong ao, với độ tin cậy 95%. Từ đó suy ra số con tôm bị bệnh trong ao là bao nhiêu?

Giải: Ta có:  $N = 10.000, n = 300, m = 60$

Tỷ lệ tôm bị bệnh trong mẫu:  $f = \frac{60}{300} = 0.2$

Ta có: Độ tin cậy:  $1 - \alpha = 0.95 \Rightarrow \varphi\left(Z_{\frac{\alpha}{2}}\right) = \frac{1 - \alpha}{2} = \frac{0.95}{2} = 0.475 \Rightarrow Z_{\frac{\alpha}{2}} = \varphi^{-1}(0.475) = 1.96$

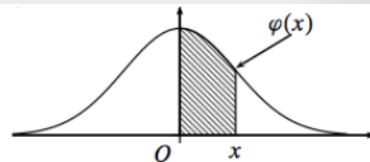
Gọi P là tỷ lệ tôm bị bệnh trong ao.

Khoảng ước lượng tỷ lệ tôm bị bệnh trong ao là:

$$P \in \left( 0.2 - 1.96 \sqrt{\frac{0.2(1-0.2)}{300}}; 0.2 + 1.96 \sqrt{\frac{0.2(1-0.2)}{300}} \right) = (0.1547; 0.2453)$$



Bảng B. Giá trị  $\varphi(x) = \int_0^x f(z)dz = \frac{1-\alpha}{2}$



x	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817





03

**Kiểm định**

# Nội dung Kiểm định:

○ Các bài toán kiểm định:

- **Kiểm định trung bình tổng thể**
- **Kiểm định tỷ lệ tổng thể**



# KIỂM ĐỊNH TRUNG BÌNH TỔNG THỂ

## TH1: Biết phương sai của tổng thể

Bước 1: Đặt giả thuyết

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

Bước 2: Tính giá trị kiểm định Z

$$Z = \frac{(\bar{x} - \mu_0) \sqrt{n}}{\sigma}$$

Bước 3: Tìm Z alpha/2

Bước 4: So sánh và kết luận

**EX:** Nếu máy móc làm việc bình thường thì trọng lượng của một loại SP là ĐLNN có phân phối chuẩn và có trọng lượng trung bình là 100gr, độ lệch chuẩn là 1gr.

Qua một thời gian, người ta nghi ngờ trọng lượng trung bình của SP này thay đổi. Người ta cân thử 100 SP và tính được trọng lượng trung bình là 100,3gr. Hãy kết luận điều nghi ngờ này, với mức ý nghĩa 5%.



## KIỂM ĐỊNH TRUNG BÌNH TỔNG THỂ

### TH1: Biết phương sai của tổng thể

Giải: Ta có:  $\mu_0 = 100\text{gr}, \sigma = 1\text{gr}; n = 100; \bar{x} = 100,3; \alpha = 0.05$

Gọi  $\mu$  là trọng lượng trung bình của SP trong nhà máy hiện nay cần kiểm định.

$$\text{Đặt giả thuyết: } \begin{cases} H_0: \mu = \mu_0 = 100 \\ H_1: \mu \neq 100 \end{cases}$$

$$\text{Giá trị kiểm định: } Z = \frac{(100.3 - 100) \sqrt{100}}{1} = 3$$

$$\text{Ta có: Mức ý nghĩa } \alpha = 0.05 \quad \Rightarrow \varphi\left(Z_{\frac{\alpha}{2}}\right) = \frac{1 - 0.05}{2} = 0.475$$

$$\Rightarrow Z_{\alpha} = \varphi^{-1}(0.475) = 1.96$$

## KIỂM ĐỊNH TRUNG BÌNH TỔNG THỂ

TH1: Biết phương sai của tổng thể

Kết luận:

Do  $|Z| > Z_{\frac{\alpha}{2}}$  nên bác bỏ giả thuyết  $H_0$ , chấp nhận  $H_1$

Tức là:  $\mu \neq 100$

Vì  $\bar{x} > \mu_0 = 100$  nên trọng lượng trung bình của SP trong nhà máy hiện nay lớn hơn 100gr. Nghĩa là, máy móc là việc không bình thường.

## KIỂM ĐỊNH TRUNG BÌNH TỔNG THỂ



### TH2: Chưa biết phương sai của tổng thể

Bước 1: Đặt giả thuyết

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

Bước 2: Tính giá trị kiểm định T

$$T = \frac{(\bar{x} - \mu_0) \sqrt{n}}{s}$$

Bước 3: Tìm T<sub>n-1</sub>; alpha/2

Bước 4: So sánh và kết luận

**EX:** Trọng lượng trung bình của các bao gạo do một máy đóng bao sản xuất là ĐLNN có phân phối chuẩn, với trọng lượng trung bình theo quy định là 50kg/bao. Để xem máy đóng bao làm việc có bình thường không, người ta cân thử 25 bao, tính được trọng lượng trung bình 49.52kg/bao và độ lệch chuẩn hiệu chỉnh là 0.5kg/bao. Hãy cho biết kết luận về tình hình làm việc của máy đóng bao này, với mức ý nghĩa 5%.

## KIỂM ĐỊNH TRUNG BÌNH TỔNG THỂ

TH2: Chưa biết phương sai của tổng thể

Giải: Ta có:  $\mu_0 = 50\text{kg}$ ;  $n = 25$ ;  $\bar{x} = 49.52\text{kg}$ ;  $s = 0.5\text{kg}$

Gọi  $\mu$  là trọng lượng trung bình của SP trong nhà máy hiện nay cần kiểm định.

Đặt giả thuyết: 
$$\begin{cases} H_0 : \mu = \mu_0 = 50 \\ H_1 : \mu \neq 50 \end{cases}$$

$$\text{Giá trị kiểm định: } T = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s} = \frac{(49.52 - 50)\sqrt{25}}{0.5} = -4.8$$

$$\text{Mức ý nghĩa: } \alpha = 0.05 \Rightarrow t_{n-1; \frac{\alpha}{2}} = t_{24; 0.025} = 2,0369$$

Ta có:  $|T| > t_{24; 0.025}$  nên bác bỏ giả thuyết  $H_0$ , chấp nhận  $H_1$ . Tức là,  $\mu \neq 50$

Do  $\bar{x} < \mu_0 = 50$  nên trọng lượng trung bình của bao gạo hiện nay nhỏ hơn 50kg. Máy làm việc không bình thường.



## KIỂM ĐỊNH TỈ LỆ TỔNG THỂ



Bước 1: Đặt giả thuyết

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

Bước 2: Tính giá trị kiểm định Z

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

Bước 3: Tìm Z alpha/2

Bước 4: So sánh và kết luận

**EX:** Một nhà máy sản xuất với tỷ lệ sản phẩm loại I lúc đầu là 20%. Sau khi áp dụng phương pháp sản xuất mới, kiểm tra ngẫu nhiên 500 sản phẩm thì thấy có 150 sản phẩm loại I. Hãy cho biết kết luận về phương pháp này, với mức ý nghĩa 1%.

## KIỂM ĐỊNH TỈ LỆ TỔNG THỂ

Giải: Ta có:  $p_0 = 0.2$ ;  $n = 500$ ;  $m = 150$ ;  $\alpha = 0.01$

Ta có: tỷ lệ sản phẩm loại I trong mẫu  $f = \frac{150}{500} = 0.3$

Gọi  $p$  là tỷ lệ sản phẩm loại I trong nhà máy hiện nay.

Đặt giả thuyết:  $\begin{cases} H_0 : p = 0.2 \\ H_1 : p \neq 0.2 \end{cases}$

Giá trị kiểm định:  $Z = \frac{(0.3 - 0.2)\sqrt{500}}{\sqrt{0.2(1 - 0.2)}} = 5.59$

Với mức ý nghĩa  $\alpha = 0.01$ , Ta có:  $Z_{\frac{\alpha}{2}} = \varphi^{-1}\left(\frac{1 - 0.01}{2}\right) = \varphi^{-1}(0.495) = 2.58$

Suy ra  $|Z| > Z_{\frac{\alpha}{2}}$  nên bác bỏ giả thuyết  $H_0$ , chấp nhận giả thuyết  $H_1$

A graphic of a spiral-bound notebook with a white page and a blue cover. The spiral binding is at the top. On the left side, there are two horizontal tabs, one yellow and one light blue. The page contains the number 04 in a large, bold, black font, centered within a light pink oval. Below the oval, the text 'Hồi quy và Tương quan' is written in a bold, blue font.

**04**

# **Hồi quy và Tương quan**

# HỒI QUY VÀ TƯƠNG QUAN

Tính hệ số  
tương quan  $r$



Viết phương trình  
hồi quy tuyến tính



Nhận xét tính tuyến  
tính của  $X$  và  $Y$



Tính giá trị của  $Y$   
theo  $X$  với  $X$  đã cho



## 1. Tính hệ số tương quan mẫu r

- r là một đại lượng đo tính tương quan tuyến tính của X và Y
- r là một con số, không có đơn vị, thuộc  $[-1, 1]$

$$r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sqrt{(\overline{x^2} - (\bar{x})^2)(\overline{y^2} - (\bar{y})^2)}}$$

## 2. Nhận xét tính tuyến tính của X và Y

- $|r| > 0,9$  : có sự tương quan tuyến tính mạnh giữa X, Y
  - $r > 0$  : tương quan tuyến tính **thuận** (X, Y đồng biến)
  - $r < 0$  : tương quan tuyến tính **nghịch** (X, Y nghịch biến)
- $|r| < 0,9$  : tương quan tuyến tính giữa X và Y yếu.
- $|r|$  càng gần 1 thì sự tương quan tuyến tính càng mạnh



### 3. Viết phương trình hồi quy tuyến tính

Phương trình hồi quy tuyến tính:

$$Y = A + B.X$$

Trong đó

$$B = \frac{\overline{xy} - \bar{x}.\bar{y}}{\overline{x^2} - (\bar{x})^2}$$

và

$$A = \bar{y} - B.\bar{x}$$

### 4. Tính giá trị của Y theo X

Dữ liệu đã có:

+ Có giá trị X từ đề bài

+ Có công thức tìm ra A và B

**Làm bài:** Thế A,B đã tìm và X từ đề bài vào phương trình hồi quy tuyến tính để tìm Y.



Điểm kiểm tra môn XSTK và môn Giải tích của 10 sinh viên được chọn ngẫu nhiên từ lớp có rất nhiều sinh viên:

XSTK	73	80	93	65	87	71	98	68	84	70
Giải tích	82	79	86	72	91	80	97	72	89	74

**1. Tính hệ số tương quan mẫu  $r$**

**2. Nhận xét tính tuyến tính của điểm XSTK và điểm Giải tích**

**3. Viết phương trình hồi quy tuyến tính của điểm Giải tích theo điểm XSTK**

**4. Tính giá trị của điểm theo điểm Giải tích theo điểm XSTK**

**Chú ý:** Xác định biến nào là biến phụ thuộc, biến nào là độc lập.

Trong trường hợp này, điểm Giải tích là biến phụ thuộc, điểm XSTK là biến độc lập.

Mode  $\Rightarrow$  Statistics  $\Rightarrow y = a + b.x \Rightarrow$  Nhập 2 hàng giá trị với biến độc lập là cột X, biến phụ thuộc là cột y  $\Rightarrow$  AC  $\Rightarrow$  OPTN  $\Rightarrow$  3. Tính hồi quy  $\Rightarrow$  Ra giá trị  $r$ ,  $A$ ,  $B$ .

$$a = 28.226$$

$$b = 0.684$$

$$r = 0.9095$$



Điểm kiểm tra môn XSTK và môn Giải tích của 10 sinh viên được chọn ngẫu nhiên từ lớp có rất nhiều sinh viên:

XSTK	73	80	93	65	87	71	98	68	84	70
Giải tích	82	79	86	72	91	80	97	72	89	74

**1. Tính hệ số tương quan mẫu  $r$**

**2. Nhận xét tính tuyến tính của điểm XSTK và điểm Giải tích**

**3. Viết phương trình hồi quy tuyến tính của điểm Giải tích theo điểm XSTK**

**4. Nếu giá trị của điểm XSTK = 100 thì giá trị của điểm Giải tích bằng bao nhiêu?**

$$a = 28.226$$

$$b = 0.684$$

$$r = 0.9095$$

1.  $r = 0.9095$

2.  $|r| > 0.9 \Rightarrow$  Điểm XSTK và điểm Giải tích có quan hệ tuyến tính mạnh.  $r > 0$  nên điểm XSTK và điểm Giải tích đồng biến.

3.  $Y = 28.226 + 0.684.X$

4. Với  $X = 100$ ,  $Y = 96.626$





# Nội Dung

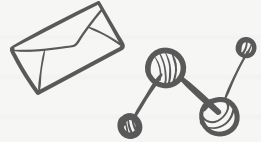


1. Vector ngẫu nhiên

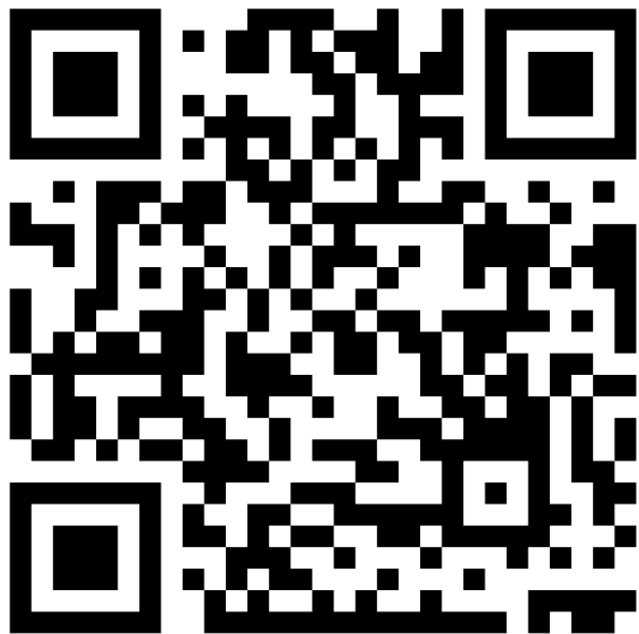
2. Ước lượng

3. Kiểm định

4. Hồi quy & Tương quan



# The End



B

H

T

Thank you for listening

**Form điểm danh**

