

Chương 4

Thống kê. Ước lượng tham số

Mục lục

1	Các đặc trưng của mẫu ngẫu nhiên	1
1.1	Trung bình mẫu ngẫu nhiên	1
1.2	Phương sai mẫu ngẫu nhiên	1
1.3	Tần suất mẫu ngẫu nhiên	1
1.4	Tính tham số đặc trưng mẫu trên máy tính bỏ túi	2
2	Phương pháp ước lượng bằng khoảng tin cậy cho kỳ vọng	2
2.1	Trường hợp đã biết phương sai: $V(X) = \sigma^2$	2
2.2	Trường hợp chưa biết phương sai, cỡ mẫu $n < 30$	3
2.3	Trường hợp chưa biết phương sai, cỡ mẫu $n > 30$	3
3	Phương pháp ước lượng cho tỷ lệ	5

1 Các đặc trưng của mẫu ngẫu nhiên

1.1 Trung bình mẫu ngẫu nhiên

Cho mẫu ngẫu nhiên $W_X = (X_1, X_2, \dots, X_n)$. Trung bình mẫu của mẫu ngẫu nhiên W_X :

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

Nếu biến ngẫu nhiên gốc có $E(X) = \mu; V(X) = \sigma^2 \Rightarrow E(\bar{X}) = \mu; V(\bar{X}) = \frac{\sigma^2}{n}$

1.2 Phương sai mẫu ngẫu nhiên

Phương sai mẫu của mẫu ngẫu nhiên W_X của biến ngẫu nhiên gốc X được ký hiệu là S và $E(S^2) = \frac{n-1}{n} \sigma^2$. Tuy nhiên để kỳ vọng của phương sai mẫu ngẫu nhiên trùng với phương sai của biến ngẫu nhiên gốc ta cần một sự hiệu chỉnh. Đó là phương sai hiệu chỉnh mẫu ngẫu nhiên. Phương sai hiệu chỉnh mẫu ngẫu nhiên: được ký hiệu là s và định nghĩa:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{n}{n-1} S^2$$

Khi đó: $E(s^2) = \sigma^2$

1.3 Tần suất mẫu ngẫu nhiên

Tần xuất mẫu là một thống kê ký hiệu và xác định bởi: $f = \frac{m}{n}$

1.4 Tính tham số đặc trưng mẫu trên máy tính bỏ túi

Tham khảo trên link:

- FX-570VN PLUS: <https://bitex.com.vn/vn/kho-ung-dung/1783/tinh-gia-tri-trung-binh,-phuong-sai-va-do-lech-chuan-bang-may-casio-fx570vn-plus.html>

- FX-580 VNX: <https://bitexedu.com/thcs/tai-lieu-thcs/toan-thcs/tinh-toan-thon-g-ke-tren-may-tinh-casio-fx-580vnx/>

Lưu ý: giá trị σ trong máy là độ lệch chuẩn mẫu chưa hiệu chỉnh. Trong tính toán, chúng ta cần dùng giá trị độ lệch chuẩn mẫu hiệu chỉnh là (s) trên máy tính.

Bài tập: Tính trung bình mẫu (\bar{x}), độ lệch chuẩn mẫu hiệu chỉnh (s) cho các bài sau:

Bài 1: Ở một địa điểm thu mua vải, kiểm tra một số vải thấy kết quả sau:

Số khuyết tật ở mỗi đơn vị	0	1	2	3	4	5	6
Số đơn vị kiểm tra (10m)	8	20	12	40	30	25	15

Bài 2: Cân 150 con vịt người ta thu được bộ số liệu sau:

Khối lượng	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00
Số lượng	2	6	24	35	39	24	14	6

Bài 3: Số liệu dưới đây cho tỷ lệ phần trăm một hóa chất trong 11 mẫu một loại xi măng:

Tỷ lệ (%)	6	15	8	8	6	9	17	18	4	8	10
-----------	---	----	---	---	---	---	----	----	---	---	----

Bài 4: Khảo sát trọng lượng X(kg) của 200 con lợn xuất chuồng ta được bảng số liệu sau:

X (kg)	[85-95)	[95-105)	[105-115)	[115-125)	[125-135)	[135-145)
Số lợn	10	30	45	80	30	5

Bài 5: Để kiểm tra trọng lượng loại sản phẩm S do nhà máy Q sản xuất trong quý IV năm 2020, người ta cần thử 100 sản phẩm loại này thu được bảng số liệu:

Trọng lượng (gam)	5-6	6-7	7-8	8-9	9-10	10-11	11-12	12-13
Số sản phẩm	5	10	15	20	29	10	6	5

2 Phương pháp ước lượng bằng khoảng tin cậy cho kỳ vọng

Cách làm tổng quát:

B1: Chọn thống kê

B2: Áp dụng và tính khoảng tin cậy tùy theo yêu cầu bài toán

B3: Kết luận

2.1 Trường hợp đã biết phương sai: $V(X) = \sigma^2$

B1: Chọn thống kê: $U = \frac{\bar{X} - \mu}{\sigma} \sqrt{n}$

B2:

- Khoảng tin đối xứng (đối với yêu cầu bài toán ước lượng trung bình)

$$\left(\bar{x} - u_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}; \bar{x} + u_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}} \right)$$

Trong đó: \bar{x} , n đã biết, α là độ tin cậy. $u_{1-\frac{\alpha}{2}}$ được tra từ bảng phụ lục (ở cuối tài liệu).

$\varepsilon = u_{1-\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$ được gọi là sai số (độ chính xác) của ước lượng.

$l = 2 \cdot \varepsilon$ được gọi là độ dài của khoảng ước lượng.

- Khoảng tin cậy trái:

$$\left(-\infty; \bar{x} + u_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}} \right)$$

- Khoảng tin cậy phải:

$$\left(\bar{x} - u_{1-\alpha} \cdot \frac{\sigma}{\sqrt{n}}, +\infty \right)$$

B3: Kết luận

2.2 Trường hợp chưa biết phương sai, cỡ mẫu $n < 30$

B1: Chọn thống kê: $T = \frac{\bar{X} - \mu}{s} \sqrt{n}$

B2:

- Khoảng tin đối xứng (đối với yêu cầu bài toán ước lượng trung bình)

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

Trong đó: \bar{x} , n đã biết, α là độ tin cậy. $u_{1-\frac{\alpha}{2}}$ được tra từ bảng phụ lục (ở cuối tài liệu).

- Khoảng tin cậy trái:

$$\left(-\infty; \bar{x} + t_{1-\alpha}^{(n-1)} \cdot \frac{s}{\sqrt{n}} \right)$$

- Khoảng tin cậy phải:

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

2.3 Trường hợp chưa biết phương sai, cỡ mẫu $n > 30$

B1: Chọn thống kê: $U = \frac{\bar{X} - \mu}{s} \sqrt{n}$

B2:

- Khoảng tin đối xứng (đối với yêu cầu bài toán ước lượng trung bình)

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

Trong đó: \bar{x} , n đã biết, α là độ tin cậy. $u_{1-\frac{\alpha}{2}}$ được tra từ bảng phụ lục (ở cuối tài liệu).

- Khoảng tin cậy trái:

$$\left(-\infty; \bar{x} + u_{1-\alpha} \cdot \frac{s}{\sqrt{n}} \right)$$

- Khoảng tin cậy phải:

$$\left(\bar{x} - u_{1-\alpha} \cdot \frac{s}{\sqrt{n}}, +\infty \right)$$

B3: Kết luận

Bài tập:

Bài 1: Số liệu dưới đây cho tỷ lệ phần trăm một hóa chất trong 11 mẫu một loại xi măng:

Tỷ lệ (%)	6	15	8	8	6	9	17	18	4	8	10
-----------	---	----	---	---	---	---	----	----	---	---	----

Với độ tin cậy 95% hãy tìm ước lượng khoảng cho tỷ lệ phần trăm trung bình của loại hóa chất trên (giả sử tỷ lệ đó là biến ngẫu nhiên tuân theo luật phân phối chuẩn).

Bài 2: Khảo sát trọng lượng X (kg) của 200 con lợn xuất chuồng ta được bảng số liệu sau:

X (kg)	[85-95)	[95-105)	[105-115)	[115-125)	[125-135)	[135-145)
Số lợn	10	30	45	80	30	5

Với độ tin cậy 90% hãy ước lượng khoảng cho trọng lượng trung bình của đàn lợn xuất chuồng.

Bài 3: Để điều tra doanh thu của các gia đình kinh doanh loại mặt hàng A tại địa phương B, người ta khảo sát 100 gia đình kinh doanh loại mặt hàng này trong một tháng của năm 2019 thu được bảng số liệu:

Doanh thu (triệu VNĐ)	25	30	35	40	45	50	55	60	65
Số gia đình	4	9	17	25	20	10	8	4	3

Với độ tin cậy 95%, hãy ước lượng doanh thu trung bình/tháng của các gia đình kinh doanh loại mặt hàng A tại địa phương B.

Bài 4: Đo độ xa X (đơn vị đo là mm) từ điểm trùng bia đến tâm bia của 16 lần bắn ta thu được số liệu sau:

2.10	1.95	2.07	2.03	1.91	2.08	1.98	2.10
2.06	1.92	1.95	2.11	2.00	1.96	2.08	1.91

Hãy ước lượng độ xa trung bình với độ tin cậy 95%. Giả sử độ xa X là biến ngẫu nhiên có phân phối chuẩn. Từ đó, tính xác suất để độ xa X lớn hơn 2mm.

Bài 5: Một nhà bán lẻ kiểm tra tuổi thọ (đơn vị tháng) của 10 ắc quy ô tô của công ty A và thu được bộ số liệu là:

27.6	28.7	34.7	29.0	22.9	29.6	29.4	30.3	36.5	34.7
------	------	------	------	------	------	------	------	------	------

Tìm ước lượng không chệch của độ lệch chuẩn của tuổi thọ trên. Với độ tin cậy 95%, xác định khoảng tin cậy cho tuổi thọ trung bình của loại ắc quy đó, biết tuổi thọ trên là biến ngẫu nhiên tuân theo luật phân phối chuẩn.

Bài 6: Để kiểm tra trọng lượng loại sản phẩm S do nhà máy Q sản xuất trong quý IV năm 2020, người ta cần thử 100 sản phẩm loại này thu được bảng số liệu:

Trọng lượng (gam)	5-6	6-7	7-8	8-9	9-10	10-11	11-12	12-13
Số sản phẩm	5	10	15	20	29	10	6	5

Với độ tin cậy 95% hãy ước lượng trọng lượng trung bình của loại sản phẩm S do nhà máy Q sản xuất tại thời điểm kiểm tra.

3 Phương pháp ước lượng cho tỷ lệ

B1: Chọn thống kê: $Z = \frac{f - p}{\sqrt{f(1-f)}} \cdot \sqrt{n}$

B2:

- Khoảng tin đối xứng (đối với yêu cầu bài toán ước lượng trung bình)

$$\left(f - u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{f(1-f)}{n}}; f + u_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{f(1-f)}{n}} \right)$$

Trong đó: $f = \frac{m}{n}$, n đã biết, α là độ tin cậy. $u_{1-\frac{\alpha}{2}}$ được tra từ bảng phụ lục (ở cuối tài liệu).

- Khoảng tin cậy trái:

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

- Khoảng tin cậy phải:

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

B3: Kết luận

Bài tập:

Bài 1: Thống kê ở một vùng trong 500 xe ô tô đăng ký có 68 xe thể thao. Với độ tin cậy 95% hãy xác định khoảng tin cậy đối xứng cho tỷ lệ xe thể thao ở vùng đó. Theo anh (chị) có cách nào để nâng cao độ chính xác của khoảng tin cậy cho tỷ lệ trên?

Bài 2: Tiến hành 120 phép đo như nhau, độc lập, thì thấy sự kiện A xuất hiện 42 lần. Xác định khoảng tin cậy đối xứng 99% cho tỷ lệ xuất hiện A.

Bài 3: Ở một trung tâm giống cây trồng, theo dõi 3070 cây cà phê thì có 1135 cây cho thu hoạch thấp. Với độ tin cậy 95% hãy xác định khoảng tin cậy cho tỷ lệ cây cà phê có thu hoạch thấp.

Bài 4: Một công ty dự định mở một siêu thị tại khu dân cư A. Để đánh giá khả năng mua hàng của khách hàng tại khu vực này, người ta đã điều tra ngẫu nhiên thu nhập trong một tháng của 100 gia đình và thu được bảng số liệu sau:

Thu nhập (triệu đồng)	34.0	34.5	35.0	35.5	36.0	36.5	37.0	37.5
Số gia đình	5	10	15	20	29	10	6	5

Tìm khoảng tin cậy cho tỷ lệ các gia đình có thu nhập $\geq 35,5$ triệu đồng/tháng với độ tin cậy 95%.

Giá trị hàm phân phối chuẩn tắc

x	0	1	2	3	4	5	6	7	8	9
0,0	0,50000	50399	50798	51197	51595	51994	52392	52790	53188	53586
0,1	53983	54380	54776	55172	55567	55962	56356	56749	57142	57535
0,2	57926	58317	58706	59095	59483	59871	60257	60642	61026	61409
0,3	61791	62172	62556	62930	63307	63683	64058	64431	64803	65173
0,4	65542	65910	66276	66640	67003	67364	67724	68082	68439	68739
0,5	69146	69447	69847	70194	70544	70884	71226	71566	71904	72240
0,6	72575	72907	73237	73565	73891	74215	74537	74857	75175	75490
0,7	75804	76115	76424	76730	77035	77337	77637	77935	78230	78524
0,8	78814	79103	79389	79673	79955	80234	80511	80785	81057	81327
0,9	81594	81859	82121	82381	82639	82894	83147	83398	83646	83891
1,0	84134	84375	84614	84850	85083	85314	85543	85769	85993	86214
1,1	86433	86650	86864	87076	87286	87493	87698	87900	88100	88298
1,2	88493	88686	88877	89065	89251	89435	89617	89796	89973	90147
1,3	90320	90490	90658	90824	90988	91149	91309	91466	91621	91774
1,4	91924	92073	92220	92364	92507	92647	92786	92922	93056	93189
1,5	93319	93448	93574	93699	93822	93943	94062	94179	94295	94408
1,6	94520	94630	94738	94845	94950	95053	95154	95254	95352	95449
1,7	95543	95637	95728	95818	95907	95994	96080	96164	96246	96327
1,8	96407	96485	96562	96638	96712	96784	96856	96926	96995	97062
1,9	97128	97193	97257	97320	97381	97441	97500	97558	97615	97670
2,0	97725	97778	97831	97882	97932	97982	98030	98077	98124	98169
2,1	98214	98257	98300	98341	98382	98422	98461	98500	98537	98574
2,2	98610	98645	98679	98713	98745	98778	98809	98840	98870	98899
2,3	98928	98956	98983	99010	99036	99061	99086	99111	99134	99158
2,4	99180	99202	99224	99245	99266	99285	99305	99324	99343	99361
2,5	99379	99396	99413	99430	99446	99261	99477	99492	99506	99520
2,6	99534	99547	99560	99573	99585	99598	99609	99621	99632	99643
2,7	99653	99664	99674	99683	99693	99702	99711	99720	99728	99763
2,8	99744	99752	99760	99767	99774	99781	99788	99795	99801	99807
2,9	99813	99819	99825	99831	99836	99841	99846	99851	99856	99861
3,0	0,99865	3,1	99903	3,2	99931	3,3	99952	3,4	99966	
3,5	99977	3,6	99984	3,7	99989	3,8	99993	3,9	99995	
4,0	999968									
4,5	999997									
5,0	99999997									

Giá trị $t_{1-\alpha}^{(n)}$ của phân phối Student

$\begin{array}{c} 1 - \alpha \\ \text{Bậc tự do} \end{array}$	0,90	0,95	0,975	0,99	0,995	0,998	0,9995
1	3,078	6,314	12,706	31,821	63,526	318,309	363,6
2	1,886	2,920	4,303	6,965	9,925	22,327	31,600
3	1,638	2,353	3,128	4,541	5,841	10,215	12,922
4	1,533	2,132	2,776	3,747	4,604	7,173	8,610
5	1,476	2,015	2,571	3,365	4,032	5,893	6,869
6	1,440	1,943	2,447	3,143	3,707	5,208	5,959
7	1,415	1,895	2,365	2,998	3,499	4,705	5,408
8	1,397	1,860	2,306	2,896	3,355	4,501	5,041
9	1,383	1,833	2,262	2,821	3,250	4,297	4,781
10	1,372	1,812	2,228	2,764	3,169	4,144	4,587
11	1,363	1,796	2,201	2,718	3,106	4,025	4,437
12	1,356	1,782	2,179	2,681	3,055	3,930	4,318
13	1,350	1,771	2,160	2,650	3,012	3,852	4,221
14	1,345	1,761	2,145	2,624	2,977	3,787	4,140
15	1,341	1,753	2,131	2,606	2,947	3,733	4,073
16	1,337	1,746	2,120	2,583	2,921	3,686	4,015
17	1,333	1,740	2,110	2,567	2,898	3,646	3,965
18	1,330	1,734	2,101	2,552	2,878	3,610	3,922
19	1,328	1,729	2,093	2,539	2,861	3,579	3,883
20	1,325	1,725	2,086	2,528	2,845	3,552	3,850
21	1,323	1,721	2,080	2,518	2,831	3,527	3,819
22	1,321	1,717	2,074	2,508	2,819	3,505	3,792
23	1,319	1,714	2,069	2,500	2,807	3,485	3,767
24	1,318	1,711	2,064	2,492	2,797	3,467	3,745
25	1,316	1,708	2,060	2,485	2,787	3,450	3,725
26	1,315	1,796	2,056	2,479	2,779	3,435	3,707
27	1,314	1,703	2,052	2,473	2,771	3,421	3,690
28	1,313	1,701	2,048	2,467	2,763	3,408	3,674
29	1,311	1,699	2,045	2,462	2,756	3,396	3,659
$+\infty$	1,282	1,645	1,960	2,326	2,576	3,090	3,291