TRƯỜNG ĐẠI HỌC TÀI CHÍNH-MARKETING
KHOA CƠ BẢN
BỘ MÔN TOÁN – THỐNG KÊ

BÀI GIẢNG

LÝ THUYẾT XÁC SUẤT VÀ THỐNG KÊ TOÁN

Giảng viên

ThS. Lê Trường Giang


PLEASE
GIVE ME
A JOB!
07889790209
© Mark Sutherland / SWNS.COM


500000
NĂM TRĂM NGHÌN ĐỒNG
VIỆT NAM
CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM


Reuters

Chương 3

MẪU NGẪU NHIÊN VÀ BÀI TOÁN ƯỚC LƯỢNG

Bài 3

ƯỚC LƯỢNG KHOẢNG CHO TỈ LỆ TỔNG THỂ

Bài 3. Ước lượng khoảng cho tỉ lệ tổng thể

- 3.1. Ước lượng khoảng cho tỷ lệ tổng thể
  - 3.1.1. Xây dựng khoảng ước lượng
  - 3.1.2. Ví dụ minh họa
  - 3.1.3. Bài tập nhóm
- 3.2.Khoảng ước lượng một phía
  - 3.2.1. Tối đa
  - 3.2.2.Tối thiểu

Tài liệu tham khảo

1. **Bài giảng Lý thuyết Xác suất và Thống kê toán,**

   Trường Đại học Tài Chính - Marketing.

2. Tập bài giảng Xác suất và Thống kê Toán – Lê Trường Giang.

3. Lê Sĩ Đồng (2013)- *Giáo trình Xác suất - Thống kê* –NXB GDVN.

4. Lê Khánh Luận, Nguyễn Thanh Sơn (2011)-*Lý thuyết xác suất và thống kê*-NXBĐHQG TpHCM.

5. Trần Lộc Hùng (2005)- *Giáo trình Xác suất Thống kê* –NXB GDVN.

6. Nguyễn Cao Văn, Trần Thái Ninh, Ngô Văn Thứ (2012) – Giáo trình Lý thuyết xác suất và Thống kê – NXB Đại học Kinh Tế Quốc Dân, HN.

Bài 3.
Ước lượng khoảng tham số tỉ lệ tổng thể

Giả sử trong tổng thể ta quan tâm những phần tử có tính chất A với tỷ lệ là p chưa biết. Từ tổng thể, ta chọn ra một mẫu gồm n phần tử, kiểm tra mẫu này ta có tỷ lệ phần tử có tính chất A là f. Với một mẫu chọn được, cùng với độ tin cậy $1-\alpha$ cho trước , nhiệm vụ của bài toán ƯLTL là cần xác định khoảng (p_1, p_2)

sao cho

$$\boxed{P\left(p_1 < p < p_2\right) = 1-\alpha}$$

3.1. Ước lượng khoảng của tỷ lệ tổng thể

Cho $\left(X_1, X_2, ..., X_n\right)$ là mẫu ngẫu nhiên của tổng thể X có tỉ lệ p,

F là tỉ lệ mẫu ngẫu nhiên,

f là tỉ lệ mẫu cụ thể,

n là kích thước mẫu,

$1-\alpha$ là độ tin cậy của ước lượng.

Ta xây dựng khoảng ước lượng (đối xứng) cho p:

XÂY DỰNG KHOẢNG ƯỚC LƯỢNG

Theo đlghtt, ta có $$G = \frac{F - p}{\sqrt{\frac{p(1-p)}{n}}} \xrightarrow{d} Z \sim N(0,1)$$

Với $1-\alpha$

ta cần xác định $\begin{cases} z_{\frac{1-\alpha}{2}} \\ -z_{\frac{1-\alpha}{2}} \end{cases}$ **thỏa mãn** $\begin{cases} P\left(G > z_{\frac{1-\alpha}{2}}\right) = \frac{\alpha}{2}; \\ P\left(G > -z_{\frac{1-\alpha}{2}}\right) = 1 - \frac{\alpha}{2}. \end{cases}$

Khi đó

$$P\left(-z_{\frac{1-\alpha}{2}} < G < z_{\frac{1-\alpha}{2}}\right) = P\left(G < z_{\frac{1-\alpha}{2}}\right) - P\left(G < -z_{\frac{1-\alpha}{2}}\right) = 1 - \alpha$$

XÂY DỰNG KHOẢNG ƯỚC LƯỢNG

Suy ra

$$P\left(-z_{\frac{1-\alpha}{2}} < \frac{F-p}{\sqrt{\frac{p(1-p)}{n}}} < z_{\frac{1-\alpha}{2}}\right) = 1-\alpha \quad (*)$$

Khi n đủ lớn, theo đlghtt ta có thể thay

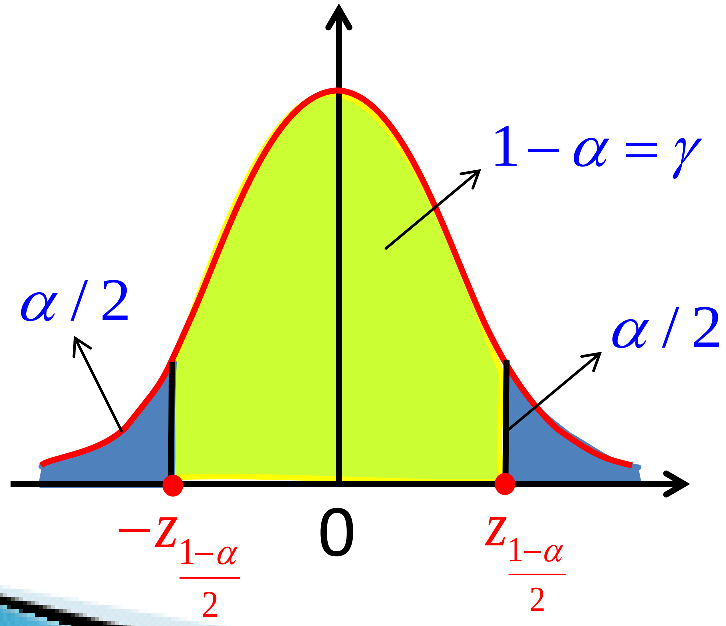
$$\sigma_X = \sqrt{p(1-p)} \to s_X = \sqrt{F(1-F)}$$

Khi đó, từ (*) ta suy ra

$$P\left(F - z_{\frac{1-\alpha}{2}}\sqrt{\frac{F(1-F)}{n}} < p < F + z_{\frac{1-\alpha}{2}}\sqrt{\frac{F(1-F)}{n}}\right) = 1-\alpha$$

Vậy , trên mẫu cụ thể ta thay F bởi *f*, ta được khoảng ước lượng của p với độ tin cậy $1-\alpha$

$$\left(f-\varepsilon, f+\varepsilon\right); \quad \varepsilon = z_{\frac{\gamma}{2}}.\sqrt{\frac{f\left(1-f\right)}{n}}$$



$$\begin{cases} n > 30 \\ nf > 5 \\ n\left(1-f\right) > 5 \end{cases}$$

MỘT SỐ VÍ DỤ

Ví dụ 1

Trước ngày bầu cử tổng thống, một cuộc thăm dò dư luận đã tiến hành. Người ta chọn ngẫu nhiên 100 người để hỏi ý kiến thì có 60 người nói rằng họ sẽ bỏ phiếu cho ông A. Hãy ước lượng (khoảng đối xứng) tỉ lệ cử tri bỏ phiếu cho ông A với độ tin cậy 95%.

Hướng dẫn tra bảng

Bảng giá trị tích phân Laplace (hàm phân phối xs Gauss)

$$\Phi_0(x)=\frac{1}{\sqrt{2\pi}}\int_0^x \exp\left(-\frac{t^2}{2}\right)dt$$

$$\gamma = 0.95;\ \Phi_0\left(z_{\frac{\gamma}{2}}\right)=\frac{\gamma}{2}=0,475 \qquad z_{\frac{\gamma}{2}} = z_{0,475} = 1,96$$

X	.00	.01	.02	.03	.04	.05	.06	.07	.08	.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	0389	0438	0478	0517	0557	0396	0636	0675	0714	0753
0.2	0793	0832	0871	0910	0948	0987	1026	1064	1103	1141
0.3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517
0.4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879
0.5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224
0.6	2257	2291	2324	2357	2389	2422	2454	2486	2517	2549
0.7	2580	2611	2642	2673	2703	2734	2764	2794	2823	2852
0.8	2881	2910	2939	2967	2995	3023	3051	3078	3106	3133
0.9	3159	3186	3212	3238	3264	3289	3315	3340	3365	3389
1.0	3413	3438	3461	3485	3508	3531	3554	3577	3599	3621
1.1	3643	3665	3686	3708	3729	3749	3770	3790	3810	3830
1.2	3849	3869	3888	3907	3925	3944	3962	3980	3997	4015
1.3	4032	4049	4066	4082	4099	4115	4131	4147	4162	4177
1.4	1492	4207	4222	4236	4251	4265	4279	4292	4306	4319
1.5	4332	4345	4357	4370	4382	4394	4406	4418	4429	4441
1.6	4452	4463	4474	4484	4495	4505	4515	4525	4535	4545
1.7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633
1.8	4641	4649	4656	4664	4671	4678	4686	4693	4699	4706
1.9	4713	4719	4726	4732	4738	4744	4750	4756	4761	4767
2.0	4772	4778	4783	4788	4793	4793	4803	4808	4812	4817
2,1	4821	4826	4830	4834	4838	4838	4846	4850	4854	4857

MỘT SỐ VÍ DỤ

Ví dụ 1

Hướng dẫn

+ Ta nhận thấy $\begin{cases} n = 100 > 30 \\ nf = 60 > 5 \\ n(1-f) = 40 > 5 \end{cases}$

+ Sai số (độ chính xác) của ước lượng

$$\varepsilon = z_{\frac{\gamma}{2}} \sqrt{\frac{f(1-f)}{n}} = 1{,}96.\sqrt{\frac{0{,}6.(1-0{,}6)}{100}} = 0{,}096$$

+ Khoảng ước lượng tỉ lệ $(f-\varepsilon; f+\varepsilon) = (0{,}504; 0{,}696)$.

MỘT SỐ VÍ DỤ

Ví dụ 2

Trước ngày bầu cử tổng thống, một cuộc thăm dò dư luận đã tiến hành. Người ta chọn ngẫu nhiên 100 người để hỏi ý kiến thì có 60 người nói rằng họ sẽ bỏ phiếu cho ông A. Để ước lượng tỷ lệ người dân bỏ phiếu cho ông A với độ tin cậy 90% và sai số không vượt quá 2% thì cần phải điều tra thêm ít nhất bao nhiêu người nữa.

Ví dụ 2

Hướng dẫn

Độ chính xác của ước lượng được xác định $\varepsilon = z_{\frac{\gamma}{2}}\sqrt{\dfrac{f(1-f)}{n}}$

Theo giả thiết ta có

$$\varepsilon \le 0,02 \Leftrightarrow z_{\frac{\gamma}{2}}\sqrt{\frac{f(1-f)}{n}} \le 0,02 \Leftrightarrow n \ge z_{0,45}^2\frac{f(1-f)}{0,02^2}$$

$$\Leftrightarrow n \ge (1,645)^2\frac{0,6.0,4}{(0,02)^2} \Leftrightarrow n \ge 1623,615$$

Vậy cần phải điều tra thêm ít nhất là 1524 người.

Các bước giải bài toán ước lượng tỷ lệ

Bước 1	**Xác định các tham số** **(n, f,** $1-\alpha$ **)**
Bước 2	**Tính độ chính xác (mức sai số)** $\varepsilon = z_{\frac{\gamma}{2}} \sqrt{\frac{f(1-f)}{n}} \Rightarrow n = z_{\frac{\gamma}{2}}^2 \frac{f(1-f)}{\varepsilon^2}$
Bước 3	**Kết luận** $p \in (f-\varepsilon; f+\varepsilon)$

3.2. Khoảng ước lượng một phía

Khoảng tin cậy tối đa của p với độ tin cậy $1-\alpha$

$$p \le f + z_{\gamma-\frac{1}{2}} \cdot \sqrt{\frac{f(1-f)}{n}}$$

Khoảng tin cậy tối thiểu của p với độ tin cậy $1-\alpha$

$$f - z_{\gamma-\frac{1}{2}} \cdot \sqrt{\frac{f(1-f)}{n}} \le p$$

Ví dụ 3

Kiểm tra ngẫu nhiên 400 sản phẩm do một máy sản xuất thấy có 20 phế phẩm. Với mức ý nghĩa 5%,

a) Hãy ước lượng tỷ lệ phế phẩm tối đa của máy đó.

b) Hãy ước lượng tỷ lệ phế phẩm tối thiểu của nhà máy đó.

Hướng dẫn tra bảng

Bảng giá trị tích phân Laplace (hàm phân phối xs Gauss)

$$\Phi_0(x) = \frac{1}{\sqrt{2\pi}} \int_0^x \exp\left(-\frac{t^2}{2}\right) dt$$

$$\alpha = 0,05 \Rightarrow \Phi_0\left(z_{0,5-\alpha}\right) = 0,5 - \alpha = 0,45 \Rightarrow z_{0,45} = 1,645$$

X	.00	.01	.02	.03	.04	.05	.06	.07	.08	.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	0389	0438	0478	0517	0557	0396	0636	0675	0714	0753
0.2	0793	0832	0871	0910	0948	0987	1026	1064	1103	1141
0.3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517
0.4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879
0.5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224
0.6	2257	2291	2324	2357	2389	2422	2454	2486	2517	2549
0.7	2580	2611	2642	2673	2703	2734	2764	2794	2823	2852
0.8	2881	2910	2939	2967	2995	3023	3051	3078	3106	3133
0.9	3159	3186	3212	3238	3264	3289	3315	3340	3365	3389
1.0	3413	3438	3461	3485	3508	3531	3554	3577	3599	3621
1.1	3643	3665	3686	3708	3729	3749	3770	3790	3810	3830
1.2	3849	3869	3888	3907	3925	3944	3962	3980	3997	4015
1.3	4032	4049	4066	4082	4099	4115	4131	4147	4162	4177
1.4	1492	4207	4222	4236	4251	4265	4279	4292	4306	4319
	4332	4345	4357	4370	4382		4406	4418	4429	4441
1.6	4452	4463	4474	4484	4495	4505	4515	4525	4535	4545
1.7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633
1.8	4641	4649	4656	4664	4671	4678	4686	4693	4699	4706
1.9	4713	4719	4726	4732	4738	4744	4750	4756	4761	4767
2.0	4772	4778	4783	4788	4793	4793	4803	4808	4812	4817
2,1	4821	4826	4830	4834	4838	4838	4846	4850	4854	4857

Ví dụ 3

Ta có $$f = \frac{20}{400} = 0,05$$

$$\varepsilon = z_{\gamma - \frac{1}{2}} \sqrt{\frac{f\left(1-f\right)}{n}} = 1,645.\sqrt{\frac{0,05.0,95}{400}} = 0,0179$$

a) Khoảng tin cậy tối đa

$$p \le f + \varepsilon = 0,0679$$

b) Khoảng tin cậy tối thiểu

$$p \ge f - \varepsilon = 0,0321$$

Các bước giải bài toán ước lượng tỷ lệ

Bước 1: Xác định

- Kích thước mẫu: n
- Tỉ lệ mẫu: f
- Độ tin cậy: $1-\alpha=\gamma$

Bước 2: Tính độ chính xác

I. Đối xứng $\varepsilon = z_{\frac{\gamma}{2}}\sqrt{\dfrac{f(1-f)}{n}}$

II. Một phía $\varepsilon = z_{\gamma-\frac{1}{2}}\sqrt{\dfrac{f(1-f)}{n}}$

Bước 3: Kết luận

i. Khoảng tin cậy đối xứng $p \in (f-\varepsilon; f+\varepsilon)$

ii. Khoảng tin cậy tối đa $p \le f+\varepsilon$

iii. Khoảng tin cậy tối thiểu $p \ge f-\varepsilon$

BÀI TẬP NHÓM

Bài 1. Một vùng có 3000 hộ gia đình. Để điều tra nhu cầu tiêu dùng một loại hàng hóa tại vùng đó người ta nghiên cứu ngẫu nhiên 100 gia đình và thấy có 74 gia đình có nhu cầu về loại hàng hóa trên. Với độ tin cậy 95% hãy ước lượng số gia đình trong vùng có nhu cầu về loại hàng hóa đó.

Bài 2. Để ước lượng tỷ lệ người dân có mức thu nhập trên 10 triệu đồng ở TP. HCM với độ tin cậy 95%, sai số không vượt quá 2% thì cần phải điều tra với số lượng bao nhiêu người, biết rằng tỉ lệ thực nghiệm là 0,8.

BÀI TẬP NHÓM

Bài 1

Hướng dẫn Bài 1

+ Gọi M ..., suy ra $p = \dfrac{M}{3000}$ + Ta nhận thấy $\begin{cases} n = 100 > 30 \\ nf = 74 > 5 \\ n(1-f) = 26 > 5 \end{cases}$

+ Độ chính xác của ước lượng

$$\varepsilon = z_{\frac{\gamma}{2}}\sqrt{\frac{f(1-f)}{n}} = 1,96.\sqrt{\frac{0,74.(1-0,74)}{100}} = 0,086$$

+ Khoảng ước lượng tỉ lệ $(0,654;0,826)$.

+ Kết luận (1962; 2478)

BÀI TẬP NHÓM

Bài 2

Hướng dẫn Bài 2

Độ chính xác của ước lượng được xác định $\varepsilon = z_{\frac{1-\alpha}{2}}\sqrt{\frac{f\left(1-f\right)}{n}}$

Theo giả thiết ta có

$$\varepsilon \le 0,02 \Leftrightarrow z_{\frac{\gamma}{2}}\sqrt{\frac{f\left(1-f\right)}{n}} \le 0,02 \Leftrightarrow n \ge z_{\frac{\gamma}{2}}^{2}\frac{f\left(1-f\right)}{0,02^{2}}$$

$$\Leftrightarrow n \ge \left(1,96\right)^{2}\frac{0,8.0,2}{\left(0,02\right)^{2}} \Leftrightarrow n \ge 1536,64$$

Vậy cần phải điều tra ít nhất là 1537 người

Bài 3. **Từ một lô hàng gồm 5000 sản phẩm, người ta chọn ngẫu nhiên ra 500 sản phẩm để kiểm tra thì thấy có 450 sản phẩm loại A.**

a) Hãy ước lượng số sản phẩm loại A có trong lô hàng với độ tin cậy 95%?

b) Nếu muốn ước lượng số sản phẩm loại A của lô hàng đạt độ chính xác như câu a) và độ tin cậy 99% thì cần kiểm tra thêm bao nhiêu sản phẩm nữa?

c) Nếu muốn ước lượng tỷ lệ sản phẩm loại A của lô hàng đạt độ chính xác $\varepsilon = 2,5\%$ thì độ tin cậy là bao nhiêu %?

Đáp số

a) (4369; 4632) sản phẩm.

b) cần phải điều tra 364 sản phẩm nữa.

c) độ tin cậy là 93.72%.

Đáp số

a) (4369; 4632) sản phẩm.

b) cần phải điều tra 364 sản phẩm nữa.

c) độ tin cậy là 93.72%.

XIN CHÂN THÀNH CẢM ƠN!