

ỨNG DỤNG PHƯƠNG PHÁP CHỌN MẪU TRONG ĐIỀU TRA CHĂN NUÔI

Thạc Sĩ HÀ VĂN SƠN

Chi trong một thời gian ngắn, nhờ có chính sách đầu tư thích hợp của nhà nước vào trong nông nghiệp, nền kinh tế nông thôn đã đạt được những thành tựu đáng kể. Trong nông nghiệp nhà nước không chỉ đầu tư về cây lương thực mà còn chú trọng phát triển chăn nuôi. Mặc dù chỉ chiếm tỷ trọng nhỏ trong cơ cấu GDP nhưng sản phẩm chăn nuôi lại đóng vai trò quan trọng trong đời sống của nhân dân. Do đó việc thu thập thông tin về chăn nuôi là rất cần thiết để có các chính sách khuyến khích, đầu tư và phát triển chăn nuôi một cách hợp lý. Trong chăn nuôi tỷ lệ hộ chăn nuôi cá thể chiếm 80%, do vậy để thu thập số liệu về tình hình chăn nuôi trong điều kiện nguồn kinh phí hạn hẹp, thì việc tiến hành điều tra toàn bộ để nắm thông tin là một việc làm hết sức khó khăn. Trong điều kiện như vậy phương pháp điều tra chọn mẫu lại tỏ ra có nhiều ưu thế, nó phù hợp với xu thế của thống kê hiện đại.

1. Quy trình thực hiện cuộc điều tra chọn mẫu trong điều tra chăn nuôi của Cục Thống kê thành phố Hồ Chí Minh

Ta hãy xem xét phương án điều tra chăn nuôi 01/10/1995 của Cục Thống kê thành phố Hồ Chí Minh. Điều tra ở tất cả các huyện, quận ven về đàn gia súc, gia cầm và sản lượng sản phẩm chăn nuôi theo hai bước:

+ Bước 1:

Điều tra toàn diện số lượng heo, trâu, bò, bò sữa, ngựa, dê, cừu, thỏ, đàn gia cầm (gà, vịt,...) v.v... của các thành phần kinh tế tập thể, hộ gia đình và chăn nuôi tự túc của các cơ quan đoàn thể thuộc 6 huyện và 4 quận ven.

+ Bước 2:

Điều tra chọn mẫu ngẫu nhiên về kết quả chăn nuôi gia súc, gia cầm của các hộ chăn nuôi, đồng thời phúc tra lại kết quả điều tra toàn diện ở bước 1.

Đơn vị chọn mẫu là ấp (huyện) và tổ dân phố (quận). Dân chọn mẫu là danh sách các ấp và tổ dân phố của các quận, huyện được sắp xếp theo vị trí địa lý. Toàn thành

phố sẽ chọn ra 5% số ấp mẫu và tổ dân phố mẫu để điều tra bước 2 theo phương pháp chọn mẫu theo khối. Danh sách các ấp mẫu, tổ dân phố mẫu trong bước 2 do Cục Thống kê trực tiếp chọn theo phương pháp chọn mẫu theo xác suất tỷ lệ với quy mô. Khi chọn theo phương pháp này, các đơn vị nào có quy mô lớn, chiếm vị trí quan trọng sẽ có điều kiện được chọn vào mẫu nhiều hơn.

Mục đích điều tra chọn mẫu ở bước 2 là để phúc tra lại một số đơn vị ở bước 1, tìm tỷ lệ sai số giữa phúc tra và điều tra toàn diện, đồng thời để có số liệu chính xác của một số ấp mẫu được chọn làm tài liệu, tính toán phương sai và xác định lại quy mô mẫu cần điều tra kế tiếp. Cho nên kết quả phúc tra do các cán bộ ở phòng nông nghiệp Cục Thống kê thành phố kết hợp với các quận huyện làm.

2. Hoàn thiện phương pháp điều tra chọn mẫu ngẫu nhiên vào trong chăn nuôi heo ở thành phố Hồ Chí Minh

Ta lấy số liệu điều tra chọn mẫu chăn nuôi heo 01/4/1995 của 6 huyện ngoại thành (Củ Chi, Hóc Môn, Thủ Đức, Bình Chánh, Nhà Bè, và Cần Giờ) để tính toán minh họa. (các số liệu minh họa ở đây được thực hiện trước khi có sự phân chia các quận huyện mới của thành phố) Các gia súc khác như trâu, bò..., cũng làm tương tự.

Chọn trong 519 ấp của 6 huyện ngoại thành lấy ra 29 ấp mẫu (chiếm 5,59% trên tổng số ấp của 6 huyện) theo xác suất tỷ lệ với quy mô, dựa vào bảng số ngẫu nhiên-Bảng Tippett - ta có danh sách các ấp mẫu.

Vì phương pháp chọn mẫu được áp dụng trong việc lựa chọn các ấp mẫu là phương pháp chọn mẫu chùm (cluster sampling) nên ta sử dụng các công thức trong chọn mẫu chùm để xác định sai số chọn mẫu.

Gọi:

Số ấp của 6 huyện là : $M = 519$ ấp

Số ấp mẫu : $m = 29$ ấp

Số hộ chăn nuôi heo trong ấp thứ i : N_i ($i=1,2,...,M$)

Số hộ chăn nuôi heo trong 519 ấp: $N = \sum_{i=1}^M N_i = 42.456$

Số hộ chăn nuôi heo của 29 ấp mẫu: $n = \sum_{i=1}^m N_i = 2.399$ hộ

Số heo của hộ thứ j trong ấp thứ i : y_{ij} Với $j=1,2,...,N_i$

Số heo bình quân của một hộ trong ấp i : y_i

Số heo bình quân của một hộ chăn nuôi heo của 29 ấp mẫu: y

Ta nhận thấy quy mô các khối mẫu trong tổng thể không bằng nhau và chọn không lặp nên ta có công thức tính phương sai chọn mẫu của y là σ_y^2 được ước lượng bởi s_y^2 theo công thức

$$s_y^2 = \frac{1}{m} \sum_{i=1}^m \left[\frac{N_i}{N} \right]^2 \frac{(\bar{y}_i - \bar{y})^2}{m-1} \left(1 - \frac{m}{M} \right)$$

Trong đó $\bar{N}_M = \frac{N}{M} = \frac{42.456}{519} = 81,80$

Dựa vào bảng tính số 1 ta tính được:

$$s_y^2 = \frac{1}{29} \frac{93,96737}{29-1} \left(1 - \frac{29}{519} \right) = 0,1092569$$

$$\Rightarrow s = \sqrt{0,1092569} = 0,33054$$

Như vậy sai số chọn mẫu của y là 0,33054 con.

Ngoài ra ta còn tính được:

$$s_y^2 = \frac{\sum_{i=1}^m (\bar{y}_i - \bar{y})^2}{m-1} = \frac{162,22}{29-1} = 5,7936$$

Số heo bình quân của 1 hộ trong 6 huyện Y được ước lượng bởi:

$$y \pm t_s \leftrightarrow 3,56 \pm t \cdot 0,33$$

Nếu chọn mức độ tin cậy là 95% thì $t=1,96$

$$\text{Hay } 3,56 - 1,96 \times 0,33 < Y < 3,56 + 1,96 \times 0,33$$

$$2,91 \text{ con} < Y < 4,21 \text{ con}$$

Bảng 1 : Bảng tính phương sai số lượng heo / hộ của 6 huyện
(Điều tra chọn mẫu 1/4/95)

Tên áp mẫu	Số hộ nuôi N_i	Tổng số heo	$\bar{y}_i$	$\bar{y}_i - \bar{y}$	$(\bar{y}_i - \bar{y})^2$	$\left(\frac{N_i}{N_u}\right)^2$	$\left(\frac{N_i}{N_u}\right)^2 (\bar{y}_i - \bar{y})^2$
1 Ấp An Hòa - An Phú	42	140	3.33	-0.22	0.05	0.26	0.01
2. Ấp Trảng Lắm - T.L. Hạ	101	300	2.97	-0.59	0.34	1.52	0.52
3. Lô 6- An Nhơn Tây	76	177	2.33	-1.23	1.51	0.86	1.30
4. Ấp chợ - Phú Hòa Đông	51	137	2.69	-0.87	0.76	0.39	0.29
5. Ấp Bàu Tròn - Nhuận Đức	121	229	1.89	-1.66	2.77	2.19	6.06
6. Ấp xóm Chùa - Tân An Hội	121	329	2.72	-0.84	0.70	2.19	1.54
7. Khu phố 6 - Thị Trấn	40	166	4.15	0.59	0.35	0.24	0.08
8. Ấp Thới Tứ - Thới Tam Thôn	193	618	3.20	-0.35	0.13	5.57	0.70
9. Ấp 2 - Tân Thới Hiệp	179	816	4.56	1.00	1.00	4.79	4.81
10. Ấp Hậu Lân - Bà Điểm	78	201	2.58	-0.98	0.96	0.91	0.87
11. Ấp Đồng Tiến - Đ. H. Thuận	52	359	6.90	3.35	11.20	0.40	4.53
12. Ấp 2 - Đông Thạnh	37	264	7.14	3.58	12.80	0.20	2.62
13. Ấp Trường Tre - Linh Xuân	61	821	13.46	9.90	98.05	0.56	54.53
14. Ấp Thảo Điền - An Phú	38	223	5.87	2.31	5.34	0.22	1.15
15. Ấp Gò Công - L.T. Mỹ	105	457	4.35	0.80	0.63	1.65	1.04
16. Ấp Vinh Thuận - Long Bình	85	274	3.22	-0.33	0.11	1.08	0.12
17. Ấp Tây A - Bình Trưng	40	154	3.85	0.29	0.09	0.24	0.02
18. Khu phố 9 - Thị Trấn	86	266	3.09	-0.46	0.22	1.11	0.24
19. Ấp 2 - Bình Hưng	34	203	5.97	2.41	5.83	0.17	1.01
20. Ấp 2 - Qui Đức	101	217	2.15	-1.41	1.98	1.52	3.02
21. Ấp 3 - Bình Chánh	124	307	2.48	-1.08	1.17	2.30	2.69
22. Ấp 2 - Tân Tạo	75	236	3.15	-0.41	0.17	0.84	0.14
23. Ấp 3 - Vĩnh Lộc A	178	469	2.63	-0.92	0.85	4.74	4.03
24. Ấp 2 - Lê Minh Xuân	68	198	2.91	-0.65	0.42	0.69	0.29
25. Khu phố 7 - Thị Trấn	54	156	2.89	-0.67	0.45	0.44	0.19
26. Ấp 6 - Phú Xuân	41	105	2.56	-1.00	0.99	0.25	0.25
27. Ấp 3 - Tân Qui Đông	9	64	7.11	3.55	12.63	0.01	0.15
28. Ấp Miếu Ba - Cần Thạnh	80	296	3.70	0.14	0.02	0.96	0.02
29. Ấp Bình Phước - Bình Khánh	129	351	2.72	-0.84	0.70	2.49	1.74
TỔNG CỘNG	2399	8533	3.56		162.22		93.97

Với tổng số hộ chăn nuôi của cả thành phố là 44.236 hộ thì tổng đàn heo của toàn thành phố được ước lượng bởi:

$$44.236 \cdot 2,91 < \hat{Y} < 44.236 \cdot 4,21$$

$$128.726 \text{ con} < \hat{Y} < 186.234 \text{ con}$$

Như vậy với mức độ tin cậy là 95% thì tổng đàn heo của thành phố có đến ngày 01.4.95 sẽ nằm trong khoảng (128.726 con, 186.234 con)

So với báo cáo chính thức về số lượng heo toàn thành phố có đến ngày 01.4.95 của Cục Thống kê thành phố Hồ Chí Minh là 180.720 con thì số liệu suy rộng trên hoàn toàn phù hợp với thực tế.

Xác định cỡ mẫu:

Để việc tính toán đơn giản ta giả thiết trường hợp quy mô các khối trong tổng thể bằng nhau, khi đó phương sai chọn mẫu của ước lượng được tính theo công thức:

$$s_y^2 = \frac{1}{m} \sum_{i=1}^m \frac{(y_i - \bar{y})^2}{m-1}$$

Ta có $\varepsilon = t_{\alpha} \cdot s_y$ là phạm vi sai số chọn mẫu,

$$\varepsilon^2 = t_{\alpha}^2 s_y^2 = t_{\alpha}^2 \frac{1}{m} \sum_{i=1}^m \frac{(y_i - \bar{y})^2}{m-1}$$

$$\Rightarrow m = \frac{t_{\alpha}^2 \sum_{i=1}^m (y_i - \bar{y})^2}{\varepsilon^2 (m-1)} = \frac{t_{\alpha}^2 s_y^2}{\varepsilon^2}$$

$$\text{Đặt: } \varepsilon = E\% \times y \Rightarrow m = \frac{t_{\alpha}^2 s_y^2 \times 100^2}{E^2 \times y}$$

Giả sử độ tin cậy là 95% thì chúng ta cần phải điều tra với quy mô mẫu như sau:

$$m = \frac{1,96^2 \times 5,7936 \times 100^2}{E^2 \times 3,56^2}$$

Thông thường trong thực tế người ta chọn $\varepsilon = (10\%) \times y$

$$E = 10 \Rightarrow m = 176 \text{ ấp}$$

Như vậy theo tính toán, quy mô mẫu cho cuộc điều tra tiếp theo là 176 ấp, chiếm 34% tổng số ấp của 6 huyện.

3. Một số kiến nghị nhằm hoàn thiện việc ứng dụng phương pháp chọn mẫu vào trong điều tra chăn nuôi:

-Dàn chọn mẫu:

Dàn chọn mẫu sẽ là các ấp, tổ dân phố ở các huyện và quận yện. Tùy theo phương pháp chọn mẫu được sử dụng mà ta có các dàn chọn mẫu khác nhau.

Nếu phương pháp chọn mẫu là phương pháp ngẫu nhiên đơn thuần, thì dàn chọn mẫu sẽ là danh sách tất cả các hộ chăn nuôi ở tỉnh, thành phố.

Nếu phương pháp chọn mẫu là phương pháp chọn mẫu theo khối, thì dàn chọn mẫu là danh sách các khối và các hộ chăn nuôi trong các khối mẫu được chọn ra.

-Phương pháp chọn mẫu:

Qua thực tế điều tra chọn mẫu trong chăn nuôi ở thành phố Hồ Chí Minh, nên sử dụng phương pháp chọn mẫu phân tổ, kết hợp với chọn mẫu theo khối, chọn mẫu theo xác suất tỷ lệ với quy mô. Đơn vị chọn mẫu là ấp hoặc tổ dân phố. Đối với một số chỉ tiêu như năng suất, sản lượng chăn nuôi ta có thể áp dụng phương pháp chọn mẫu hai giai đoạn. Giai đoạn 1 đơn vị chọn mẫu là ấp hoặc tổ dân phố. Giai đoạn 2 đơn vị chọn mẫu là hộ chăn nuôi.

- Về công tác tổ chức thực hiện:

Đối với những địa phương chưa có kinh nghiệm trong việc tổ chức chọn mẫu trong điều tra chăn nuôi, cần phải tiến hành cuộc điều tra thử nghiệm trước khi tiến hành cuộc điều tra chính.

Công tác tập huấn cho cán bộ tham gia điều tra phải được thực hiện đầy đủ, có hiệu quả. Cán bộ tham gia điều tra phải hiểu đúng và đầy đủ mục đích, yêu cầu điều tra, nắm vững các khái niệm, các phương pháp đo lường trong cuộc điều tra.

Các biện pháp kiểm tra, kiểm soát phải được tiến hành thường xuyên để có thể sửa chữa những sai sót kịp thời.

Việc tuyên truyền rộng rãi về cuộc điều tra, tạo ấn tượng tốt cho các đối tượng điều tra cũng là việc làm hết sức cần thiết nhằm tạo ra tâm lý an tâm cho đối tượng điều tra để từ đó họ hợp tác tốt trong việc cung cấp số liệu.