



Facultade de Ciencias Económicas e Empresariais

Traballo de
fin de grao

Análise mostral:
estudo do deseño
sistemático

André López Pérez

Xullo 2020

Resumo

As enquisas son un instrumento cun papel chave na nosa sociedade, pois constitúen unha base para a toma de decisións, tanto no ámbito privado coma no público. Para a súa elaboración será necesario en primeiro lugar a selección dunha mostra, tendo en conta que o tratamento de datos de toda a poboación non vai resultar posible por razóns financeiras e de tempo. Neste traballo profundizamos nun tipo concreto de mostraxe probabilística, coma é a mostraxe sistemática. Malia non ser o método máis axeitado para determinados casos, pretendemos examinar cales son os seus pros e contras, enfatizando os primeiros a través dunha aplicación empírica. Por outra banda, pertencendo ao ámbito estatístico, será necesaria a introdución dunha serie de fórmulas que nos van permitir calcular os estimadores do total, da media e da proporción, así coma da súa varianza. Deste xeito, o traballo permítenos obter unha análise desta modalidade de mostraxe, tanto dende un punto de vista teórico coma práctico. Mais non se trata dunha visión estática, senón que podemos apreciar coma os investigadores están introducindo cambios constantemente, aparecendo así novos casos particulares da mostraxe sistemática, coma son a replicada ou a especial.

A extensión do traballo é de 8.561 palabras

Índice

Resumo.....	2
Índice.....	3
Índice de abreviaturas	5
Índice de táboas e gráficos.....	6
Índice de táboas	6
Índice de gráficos.....	7
Introdución.....	8
Planificación.....	11
Desenvolvemento do traballo.....	12
1. A mostraxe sistemática.....	12
1.1 Procedemento para seleccionar unha mostra sistemática	12
1.2 Vantaxes e desvantaxes da mostraxe sistemática	15
1.3 Relación da mostraxe sistemática coa mostraxe estratificada e por conglomerados	16
2. Estimadores.....	19
2.1 Varianza dos estimadores	20
2.1.1 Comparación entre a Mostraxe Aleatoria Simple e a Mostraxe Sistemática a partir das súas varianzas.....	22
3. Casos particulares.....	23
3.1 Casos particulares na poboación	23
3.1.1 Poboación aleatoria	23
3.1.2 Poboación periódica	23
3.1.3 Poboación con tendencia lineal.....	24
3.2 Casos particulares na mostraxe sistemática	25
3.2.1 Mostraxe sistemática replicada	25
3.2.2 Mostraxe sistemática especial	25
4. Mostras interpenetrantes.....	26

5. Comparación entre unha mostra sistemática e unha mostra aleatoria simple: caso particular do Grao en Economía na Universidade de Santiago de Compostela	27
Conclusións e ampliación.....	37
Bibliografía.....	39

Índice de abreviaturas

G.E: Grupo expositivo

G.I: Grupo interactivo

K: Coeficiente de elevación

M: Mostra

m.a.s.: Mostraxe Aleatoria Simple

N: Número total de unidades da poboación

n: Tamaño mostral

Pr.: Probabilidade

S₂ : Cuasivarianza poboacional

sis.: Sistemático

s.r.: Sen reposición

U_j : Punto de arranque

V: Varianza poboacional

\hat{Y} : Estimador do total

$\hat{\theta}$: Estimador

π_j : Probabilidade

Índice de táboas e gráficos

Índice de táboas

TÁBOA 1. DIFERENTES MOSTRAS SISTEMÁTICAS NO ESPAZO DE K MOSTRAS POSIBLES	13
TÁBOA 2. DIFERENTES MOSTRAS SISTEMÁTICAS POSIBLES CON N=20 E N=200	14
TÁBOA 3. DIFERENTES MOSTRAS SISTEMÁTICAS POSIBLES CON N=20 E N=200 RESALTANDO UN ESTRATO E UN CONGLOMERADO	17
TÁBOA 4. DISTRIBUCIÓN DO ALUMNADO DO GRAO EN ECONOMÍA EN GRUPOS EXPOSITIVOS E INTERACTIVOS SEGUNDO CURSOS	27
TÁBOA 5. DISTRIBUCIÓN DO ALUMNADO DO 1º CURSO DO GRAO EN ECONOMÍA NOS DIFERENTES GRUPOS EXPOSITIVOS E INTERACTIVOS.....	28
TÁBOA 6. DISTRIBUCIÓN DO ALUMNADO DO 2º CURSO DO GRAO EN ECONOMÍA NOS DIFERENTES GRUPOS EXPOSITIVOS E INTERACTIVOS.....	28
TÁBOA 7. DISTRIBUCIÓN DO ALUMNADO DO 3º CURSO DO GRAO EN ECONOMÍA NOS DIFERENTES GRUPOS EXPOSITIVOS E INTERACTIVOS.....	29
TÁBOA 8. DISTRIBUCIÓN DO ALUMNADO DO 4º CURSO DO GRAO EN ECONOMÍA NOS DIFERENTES GRUPOS EXPOSITIVOS E INTERACTIVOS.....	29
TÁBOA 9. DIFERENTES MOSTRAS SISTEMÁTICAS POSIBLES PARA O GRAO EN ECONOMÍA CON N=500 E N=50	30
TÁBOA 10. MOSTRA ALEATORIA SIMPLE DO GRAO EN ECONOMÍA CON N=500 E N=50	33
TÁBOA 11. MOSTRA ALEATORIA SIMPLE DO GRAO EN ECONOMÍA CON N=500 E N=50 SEN REPRESENTACIÓN DO 4º CURSO...	35

Índice de gráficos

GRÁFICO 1. MOSTRAXE SISTEMÁTICA DUNHA POBOACIÓN PERIÓDICA.....	24
GRÁFICO 2. MOSTRAXE SISTEMÁTICA DUNHA POBOACIÓN CON TENDENCIA LINEAL.....	24

Introdución

As técnicas de mostraxe son unha ferramenta fundamental empregada en moitos ámbitos, os cales van dende a economía á medicina ou á política. Os métodos cuantitativos teñen a súa utilidade coma instrumentos de apoio para a toma de decisións (Cid Cid, Delgado Manríquez, & Leguey Galán, 1999). Polo tanto, será necesario que a precisión dos datos sexa elevada, tendo en consideración que vai repercutir directamente nas decisións adoptadas. O obxectivo final da mostraxe vai ser obter a maior cantidade de información posible co menor custo, para proceder así á inferencia estatística (Cid Cid, Delgado Manríquez, & Leguey Galán, 1999). É dicir, o principio de economicidade vai estar presente en toda investigación, tanto dende o punto de vista financeiro coma temporal. Mais esta cuestión está directamente ligada coa precisión que se busque no estudo, pois é evidente que canta máis exactitude se busque neste, ou mellor dito, menor erro mostral se espere, o custo de dita investigación vai ser maior.

Podemos apreciar como esta cuestión está presente a diario nos medios de comunicación, nas institucións ou nas empresas. Todos os partidos políticos desexan coñecer cal é a intención de voto dos cidadáns, para así poder modificar ou continuar coas súas actuacións durante unha campaña electoral, o cal se deriva a través das enquisas, e polo tanto, das técnicas de mostraxe. O concello de Madrid, para saber a aceptación que podía ter Madrid Central no conxunto de individuos que abrangue o territorio, tamén levou a cabo unha serie de enquisas relacionadas con esta cuestión. Polo tanto, con estes dous exemplos pretendo amosar como se trata dun aspecto que ten unha gran relevancia nos nosos días e dunha ferramenta máis que útil para organismos, administracións ou partidos políticos.

Dito o anterior, compre explicar as razóns polas que a mostraxe resulta necesaria no tratamento e estudo de cuestións de diversa índole. En primeiro lugar, tal e como se mencionou anteriormente, é preciso comentar que o investigador ten unha serie de limitacións, tanto temporais coma financeiras. É dicir, se se quere coñecer a porcentaxe de españois que merca o xornal a diario, non é factible entrevistar aos 46 millóns de habitantes, pois conlevaría uns elevadísimos custos monetarios e unha cantidade de tempo non asumible polo investigador. Isto é aínda máis obvio se se fai o emprego da mostraxe para cuestións nas que se considere a poboación dun continente ou do mundo, resultando imposible a observación censal polas causas xa mencionadas. Polo tanto, o investigador seleccionará unha mostra de toda esta poboación, podendo así levar a cabo o seu estudo, cumprindo así co principio de economicidade. A mostraxe tamén será preferible ao censo no caso das poboacións homoxéneas. É dicir, cando a varianza da característica que estemos a estudar sexa mínima nos elementos da poboación,

será preferible o emprego da técnica da mostraxe, pois levará consigo un menor custo, á vez que non implicará un maior erro.

Tamayo (2001) afirma que a mostraxe é necesaria naqueles casos en que “a medición do elemento trae consigo a destrución ou deterioro do mesmo”. Isto é propio do ámbito da medicina, química ou física. Por exemplo, na procura dunha vacina contra unha enfermidade hai elementos da poboación que van sufrir unha reacción non desexada ou negativa, polo que o uso da mostraxe permite facer máis pequeno o número dos afectados.

Dado que ao longo deste traballo de investigación imos introducir e empregar unha serie de elementos relativos a esta técnica estatística, considero necesario coma paso previo definilos, de forma que se facilite a comprensión posterior do traballo en conxunto. Consideraremos as definicións que realiza López (2004) para os tres primeiros termos que se amosan a continuación, así coma a definición que realiza Pérez López (1999) de censo:

- Poboación: Conxunto de persoas ou obxectos dos que se desexa coñecer algo nunha investigación
- Mostraxe: Proceso mediante o cal se seleccionan as compoñentes da mostra do total da poboación.
- Mostra: Subconxunto ou parte do universo ou poboación en que se levará a cabo a investigación.
- Censo: Recollida ou toma de información de todos os elementos que forman a poboación

Cabe destacar, e seguindo o comentado no primeiro parágrafo, que a mostraxe é a fase previa á inferencia estatística. Este último termo podémolo definir coma a aplicación, a partir da información proporcionada por unha mostra seleccionada ao azar da poboación, de instrumentos estatísticos para obter conclusións e facer xeralizacións válidas da poboación baixo estudo (Quintana, 1996). Polo tanto é esencial que se minimicen os erros na mostraxe, tendo que cumprirse unha serie de condicións (Rodríguez Osuna, 2001):

- Innesgadez: En termos puramente estatísticos podemos afirmar que un estimador é innesgado se o valor do estatístico coincide cos valores do parámetro.
- Precisión: Vai estar estreitamente relacionada co tamaño da mostra. É dicir, a precisión é directamente proporcional ao tamaño da mostra.
- Para que a mostraxe sexa probabilística tense que cumprir a aleatoriedade e a independencia.

Chegado a este punto, e comentadas unhas breves nocións teóricas, é necesario precisar as diferentes fases que teñen lugar en toda investigación por mostraxe, as cales só me vou limitar a mencionar: establecemento do obxectivo, delimitación da poboación obxectivo e a poboación investigada, establecemento do marco, deseño da mostra, traballo de campo, procesamento dos datos, avaliación dos resultados, presentación dos resultados e difusión dos resultados. Ao longo deste proceso van ir xurdindo unha serie de erros, os cales podemos clasificar en dous grupos: erros de mostraxe e erros alleos á mostraxe (Tamayo, 2001). O erro mostral é aquel que se produce coma resultado de observar tan só unha fracción de todas as ocasións posibles (Blanco Villaseñor, Castellano Paulis, & Hernández Mendo, 2000). Por outra banda, o erro non

mostral ou alleo á mostraxe está asociado ao propio proceso de investigación. Dentro deste último podemos atopar erros na identificación do problema, erros de interpretación dos datos, nesgos introducidos polo entrevistador, etc.

Co obxectivo de minimizar os erros alleos á mostraxe é moi útil a elaboración e posta en funcionamento das enquisas piloto (Pérez Suárez, 1986). Unha enquisa piloto ten coma finalidade avaliar diferentes aspectos dun cuestionario, para ver se pode resultar efectivo, levándoo a cabo sobre unha mostra máis pequena, de forma que tamén se cumpra o principio de economicidade. A través desta o investigador pode comprobar a boa redacción das preguntas, os nesgos que poden estar introducidos ou axudar á redacción de novas preguntas.

Introducidos e definidos os termos que considero chave neste ámbito da estatística, é preciso diferenciar os dous tipos de mostraxe cos que nos podemos atopar: mostraxe probabilístico e mostraxe non probabilístico. Dentro do primeiro grupo podemos diferenciar entre as 5 técnicas seguintes: mostraxe aleatoria simple, mostraxe sistemática, mostraxe estratificada, mostraxe por conglomerados e mostraxe bietápica. Así, neste traballo ímonos centrar na mostraxe sistemática.

No primeiro apartado explícase como se leva a cabo a selección dunha mostra sistemática, as vantaxes e desvantaxes que posúe fronte ás outras modalidades, así como a consideración deste tipo de mostraxe coma caso particular da estratificada e por conglomerados. En segundo lugar procedemos a analizar os estimadores do total, da media e da proporción, ademais das súas varianzas, realizando unha comparación entre a mostraxe sistemática e a mostraxe aleatoria simple a partir deste último parámetro. Unido a isto desenvólvese un terceiro apartado, no que podemos observar os casos particulares ante os que nos podemos atopar. Neste punto consideramos práctico a división do epígrafe en dous subapartados. Por unha banda, os tipos de poboación cos que nos podemos atopar, observando que modalidade de mostraxe é máis apropiada para cada un destes. Por outra banda, un subapartado no que se fai referencia aos casos particulares no proceso de extracción da mostra. En cuarto lugar introducimos o concepto de mostras interpenetrantes, así coma a súa importancia para a obtención dun estimador innesgado da varianza. Por último, nun quinto apartado efectuamos unha comparación entre a mostraxe sistemática e a mostraxe aleatoria simple a partir dunha aplicación dos datos do Grao en Economía da Universidade de Santiago de Compostela.

Planificación

Neste epígrafe amósase o cronograma do proceso de elaboración do traballo, diferenciando as tarefas desenvolvidas para levalo a cabo e indicando a súa ubicación temporal. Do mesmo xeito, quero aproveitar este espazo para agradecer á profesora Mariña Lois o tempo dedicado e o apoio prestado para a realización deste traballo.

Novembro 2019						
L	M	X	J	V	S	D
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	

Decembro 2019						
L	M	X	J	V	S	D
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

Xaneiro 2020						
L	M	X	J	V	S	D
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

Febreiro 2020						
L	M	X	J	V	S	D
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	

Marzo 2020						
L	M	X	J	V	S	D
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

Abril 2020						
L	M	X	J	V	S	D
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30			

Maio 2020						
L	M	X	J	V	S	D
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

Xuño 2020						
L	M	X	J	V	S	D
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30					

Xullo 2020						
L	M	X	J	V	S	D
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

As diferentes tarefas aparecen marcadas no calendario nas seguintes cores:

Verde: asignación do tema e a titora do TFG.

Vermello: revisión bibliográfica.

Azul: redacción do texto inicial e corrección dos borradores.

Laranxa: entrega de borradores parciais á titora.

Amarelo: reunións coa titora.

Desenvolvemento do traballo

1. A mostraxe sistemática

1.1 Procedemento para seleccionar unha mostra sistemática

Este método está caracterizado por unha fácil execución, requirindo menos medios económicos ca outros coma a mostraxe aleatoria simple, e, á vez, sendo máis representativo (Tamayo, 2001). Con esa consideración previa, o seu procedemento lévase a cabo do modo que explicamos a continuación.

Estando numeradas as unidades que forman a poboación de 1 a N, construímos unha serie de intervalos que van ir de k en k unidades ($k=N/n$). Para comezar seleccionando a mostra escollemos unha cifra entre 1 e k, cuxa notación matemática vai ser j, e a partir de aí imos ir sumando de k en k unidades. Así, imos constituír unha mostra que vai ser da seguinte maneira: $U_j, U_{j+k}, U_{j+2k}, U_{j+3k}, U_{j+(n-1)k}$. Cabe destacar que as diferentes mostras posibles non van ter ningún elemento en común, pois serán escollidas as cifras que ocupen unha determinada posición en cada un dos intervalos mencionados (Cid Cid, Delgado Manríquez, & Leguey Galán, 1999).

Polo tanto, con este sistema as extraccións son sempre sen reposición e, ademáis, no deseño sistemático a selección dunha unidade mostral determina a composición da mostra, tal e como podemos observar na seguinte táboa.

Táboa 1. Diferentes mostras sistemáticas no espazo de k mostras posibles

MOSTRAS POSIBLES				
M₁	-	M_J	-	M_K
U₁	-	U_J	-	U_K
U_{1+K}	-	U_{J+K}	-	U_{2K}
U_{1+2K}	-	U_{J+2K}	-	U_{3K}
.	-	.	-	.
.	-	.	-	.
.	-	.	-	.
U_{1+(n-1)K}	-	U_{J+(n-1)k}	-	U_{nK}

Fonte: *Elaboración propia*

Por ser as K mostras equiprobables e formadas por n unidades cada unha, a probabilidade de que cada unidade apareza na mostra é:

$$\pi_j = \Pr(U_j) = n/N.$$

Para entender mellor este tipo de mostraxe, poñamos por exemplo unha poboación de N=200, seleccionando unha mostra de n=20. Unha vez numeradas todas as unidades que se poden seleccionar para a mostra, tomaremos ao chou unha unidade mostral que non sexa maior ao coeficiente “k”, denominado coeficiente de elevación, e cuxo cálculo é $k=N/n$. Neste caso $k=200/20=10$.

A continuación haberá que ir sumando o coeficiente de elevación á primeira unidade mostral seleccionada ata obter unha mostra de tamaño n=20. Finalmente, coas táboas de números aleatorios seleccionamos un número comprendido entre 1 e 10.

Supoñendo que a cifra escollida é 7, a partir desta tomaríamos os demais elementos pertencentes á mostra sumando de 10 en 10, rematando cando n=20. Neste caso a mostra sería a seguinte:

7, 17, 27, 37, 47, 57, 67, 77, 87, 97, 107, 117, 127, 137, 147, 157, 167, 177, 187, 197.

Isto correspóndese coa mostra número 7 das que figuran na táboa reproducida a continuación.

Táboa 2. Diferentes mostras sistemáticas posibles con n=20 e N=200

MOSTRAS POSIBLES									
M ₁	M ₂	M ₃	M ₄	M ₅	M ₆	M ₇	M ₈	M ₉	M ₁₀
1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100
101	102	103	104	105	106	107	108	109	110
111	112	113	114	115	116	117	118	119	120
121	122	123	124	125	126	127	128	129	130
131	132	133	134	135	136	137	138	139	140
141	142	143	144	145	146	147	148	149	150
151	152	153	154	155	156	157	158	159	160
161	162	163	164	165	166	167	168	169	170
171	172	173	174	175	176	177	178	179	180
181	182	183	184	185	186	187	188	189	190
191	192	193	194	195	196	197	198	199	200

Fonte: *Elaboración propia*

Sendo as 10 mostras posibles equiprobables e formadas por 20 unidades cada unha, a probabilidade de que cada unidade apareza na mostra é:

$$\pi_j = \Pr(U_i) = n/N = 20/200 = 0,1$$

1.2 Vantaxes e desvantaxes da mostraxe sistemática

De acordo con Cid Cid, Delgado Manríquez e Leguey Galán (1999) e Pérez López (1999), existe un consenso acerca das vantaxes que a mostraxe sistemática presenta fronte a outros tipos de mostraxe:

- Selección rápida da mostra.
- Pode concebirse a mostraxe sistemática coma un caso particular da mostraxe estratificada. É dicir, cada un dos estratos está constituído por un intervalo de k unidades, seleccionándose así unha unidade de cada un.
- Se existe unha certa aleatoriedade entre os elementos seleccionados, nunca haberá un grupo de elementos de gran tamaño que quede sen estar representado.
- Pode considerarse un caso particular da mostraxe por conglomerados.
- Cando haxa aleatoriedade, pode considerarse a mostraxe sistemática coma un caso particular da mostraxe aleatoria simple.
- Non é necesario distinguir entre mostraxe con reposición e sen reposición, a diferenza doutras coma a aleatoria simple.
- O erro de mostraxe é habitualmente inferior ao presente na mostraxe aleatoria simple.
- Trátase dunha modalidade da mostraxe que se pode aplicar sen coñecer o tamaño exacto da mostra. É dicir, non se necesita un listado exhaustivo de toda a poboación. Este pode ser o caso da selección dun determinado número de coches que pasan por unha peaxe.

Por outra banda, a principall desvantaxe que se sinala habitualmente para a mostraxe sistemática é a presenza de nesgos dependendo da orde en que estén expostos os elementos do universo, é dicir, nos casos en que haxa periodicidade. Para evitar este problema pódese optar por desordenar os elementos da poboación (Rodríguez Osuna, 2001). Polo tanto, caberá a posibilidade dun aumento na varianza no caso en que a poboación estea ordenada de forma periódica (Azorín & Sánchez-Crespo, 1986).

De acordo con Pérez López (1999), outros inconvintes que se poden sinalar para este tipo de mostraxe son:

- Inexistencia de independencia, pois as unidades seleccionadas dependen da primeira escolla..
- A aleatoriedade só se produce na selección da primeira unidade da mostra..
- Problema teórico no relativo á estimación das varianzas..

- Dificultade á hora de obter estimadores do erro de mostraxe cometido, tendo presente que non todo par de elementos está representado na mostra. (Hedayat & Sinha, 1991) citado en (Fernandez Álvaro, López Blázquez, & Salamanca Miño, 2001)

1.3 Relación da mostraxe sistemática coa mostraxe estratificada e por conglomerados

Unha vez que introducimos as vantaxes máis significativas da mostraxe sistemática fronte a outros tipos de mostraxe, hai dúas que imos explicar con maior profundidade neste apartado. Por unha banda, a posible consideración da mostraxe sistemática coma un caso particular da mostraxe estratificada e, por outro lado, a concepción da mostraxe sistemática coma un caso da mostraxe por conglomerados.

Podemos considerar cada fila da táboa reproducida a continuación coma un estrato. É dicir, cada unha das filas vai representar un intervalo de lonxitude “k”, neste caso $k=10$. Polo tanto, a poboación divídese en n estratos ($n=N/k = 200/10 = 20$). A partir da táboa anteriormente construída, que contén os elementos da poboación numerados consecutivamente, poderíamos representala tal e como se amosa a continuación. Supoñendo a poboación dividida en n estratos e k conglomerados temos:

Táboa 3. Diferentes mostras sistemáticas posibles con n=20 e N=200 resaltando un estrato e un conglomerado

i/j	1									k
1	M ₁	M ₂	M ₃	M ₄	M ₅	M ₆	M ₇	M ₈	M ₉	M ₁₀
	1	2	3	4	5	6	7	8	9	10
	11	12	13	14	15	16	17	18	19	20
	21	22	23	24	25	26	27	28	29	30
	31	32	33	34	35	36	37	38	39	40
	41	42	43	44	45	46	47	48	49	50
	51	52	53	54	55	56	57	58	59	60
	61	62	63	64	65	66	67	68	69	70
	71	72	73	74	75	76	77	78	79	80
	81	82	83	84	85	86	87	88	89	90
i	91	92	93	94	95	96	97	98	99	100
	101	102	103	104	105	106	107	108	109	110
	111	112	113	114	115	116	117	118	119	120
	121	122	123	124	125	126	127	128	129	130
	131	132	133	134	135	136	137	138	139	140
	141	142	143	144	145	146	147	148	149	150
	151	152	153	154	155	156	157	158	159	160
	161	162	163	164	165	166	167	168	169	170
	171	172	173	174	175	176	177	178	179	180
	181	182	183	184	185	186	187	188	189	190
n	191	192	193	194	195	196	197	198	199	200

Fonte: *Elaboración propia.*

A liña subliñada en cor amarela representa un estrato, neste caso dende a unidade 1 á 10 (ambas incluídas). Se se fixese igual con todas as unidades da poboación, obteríamos 20 estratos, cada un formado por 10 unidades, de xeito que a obtención da mostra sistemática resulta practicamente idéntica á da mostra estratificada cunha unidade por estrato.

Estando os estratos formados por unidades iguais ou moi parecidas entre si, imos encontrarnos cunha gran vantaxe como é a elevada representatividade da mostra. É dicir, se cada estrato é un intervalo de elementos moi similares, na selección da mostra imos obter unha unidade de cada un destes estratos ou intervalos, podendo ser esta moi representativa da poboación. No caso contrario, se nos atopamos ante unha gran periodicidade, sendo k igual ao periodo ou múltiplo de cada un destes, a representatividade vai ser moi baixa.

Neste punto resulta necesario definir un termo mencionado no parágrafo anterior, coma é o de representatividade. Tal e como sinalan Kruskal e Mosteller (1979), citados en Azorín e Sánchez-Crespo (1986), podemos falar de representatividade ou de mostra representativa nos seguintes casos:

- Que as unidades seleccionadas reflectan satisfactoriamente a variabilidade da poboación.
- Cando non se produzan tendencias ou preferencias na selección das unidades da mostra.
- Que podamos falar da mostra coma unha miniatura da poboación.

Resulta tamén pertinente mencionar unha das desvantaxes comentadas no apartado 1.2, coma é a ausencia da independencia na selección das unidades da mostra, pois depende en todo caso da primeira escolla. Sen embargo, no caso da mostraxe estratificada a independencia está presente en todo momento para a selección de ditas unidades.

Por outra banda, tal e como foi apuntado anteriormente, tamén se pode considerar a mostraxe sistemática coma caso particular da mostraxe por conglomerados. Cada unha das columnas da táboa representa un conglomerado, coma é o caso da subliñada de cor verde, que vai dende o número 6 ao 196, constituíndose o conglomerado de 10 en 10 unidades. Así, podemos apreciar como na táboa hai 10 conglomerados, cada un deles formado por 20 unidades. Desta forma, seleccionar unha mostra sistemática sería equivalente á selección dunha mostra por conglomerados de tamaño un, pois cada columna da táboa mencionada é representativa de todos os elementos da poboación numerados, sendo unha mostra sistemática posible. A pretensión é que cada columna forme un grupo de unidades do máis heteroxéneas posibles e con propiedades similares ás de toda a poboación, de tal forma que unha única columna, é dicir, un conglomerado, poda representar ben a toda a poboación. Polo tanto, son desexables columnas con heteroxeneidade dentro delas e homoxeneidade entre elas, a diferenza do estrato, que constitúe o caso oposto.

2. Estimadores

Un estimador é un estatístico (unha función da mostra) empregado para estimar un parámetro descoñecido da poboación (de la Fuente Fernández, 2016). Tendo en consideración que o obxectivo da mostraxe é poder levar a cabo a inferencia estatística, o proceso da estimación vai ser fundamental neste senso, poidendo ser este último puntual ou por intervalos. Para cada un dos tres parámetros que consideraremos neste epígrafe (media, total e proporción) teremos un estimador innesgado, tal e como se mostra a continuación.

- Estimador da media poboacional:

$$\hat{Y} = \frac{y_1 + y_2 + \dots + y_N}{N} = \frac{1}{N} \sum_{i=1}^N y_i$$

É a media da mostra sistemática:

$$\bar{y}_l = \frac{y_j + y_{j+k} + \dots + y_{(n-1)k}}{n}$$

- Estimador do total:

$$\hat{Y}_l = N \sum_{i=1}^N \frac{y_1 + y_2 + \dots + y_N}{N} = N \bar{y}_l$$

É o total da mostra sistemática:

$$\hat{y} = N \frac{y_j + y_{j+k} + \dots + y_{j+(n-1)k}}{n}$$

- Estimador da proporción poboacional:

$$P = \frac{A_1 + A_2 + \dots + A_N}{N} = \frac{1}{N} \sum_{j=1}^N A_j$$

É a proporción mostral:

$$P_j = \frac{A_j + A_{j+k} + \dots + A_{j+(n-1)k}}{n}$$

No último caso, tratándose de caracteres cualitativos, os A_j tomarán o valor 1 ou 0 dependendo de que posúan ou non unha característica concreta.

Neste apartado introducimos outro concepto, coma é o de estimador innesgado. Un estimador é innesgado cando a súa media ou esperanza matemática coincide co valor verdadeiro do parámetro descoñecido (Cid Cid, Delgado Manríquez, & Leguey Galán, 1999). Por exemplo, no caso particular do estimador da media afirmamos que é innesgado, pois a esperanza de tal estimador é igual ao parámetro descoñecido. Noutros termos, a media mostral coincide coa media poboacional.

2.1 Varianza dos estimadores

Á hora de explicar a varianza na mostraxe sistemática, vai ser preciso en primeiro lugar amosar as fórmulas da varianza e cuasivarianza poboacional:

- Varianza poboacional:

$$V(\hat{X}) = \frac{1}{N} \sum_{i=1}^N (X_i - \hat{X})^2$$

- Cuasivarianza poboacional:

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

A varianza dentro das diferentes mostras é:

$$V_{sis}(\hat{X}) = \frac{1}{kn} \sum_{i=1}^k \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2$$

Pola súa parte, a cuasivarianza dentro das diferentes mostras, tamén denominada cuasivarianza intramostral, é (Pérez López, 2005):

$$S_{sis}^2 = \frac{1}{k(n-1)} \sum_{i=1}^k \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2$$

Finalmente, a varianza estimada dos diferentes estimadores podemos escribila do seguinte xeito:

$$\hat{V}(\hat{\theta}) = \sum_{i=1}^n \frac{y_i^2}{\pi_i} (1 - \pi_i) + 2 \sum_{i=1}^n \sum_{j=1}^n \frac{y_i y_j (\pi_{ij} - \pi_i)}{\pi_i \pi_j}$$

- Varianza da media:

$$V(\hat{Y}) = \frac{1}{k} \sum_{j=1}^k (\hat{X}_j - \bar{X})^2$$

- Varianza do total:

$$V(\hat{Y}) = \frac{N^2}{k} \sum_{j=1}^k (\hat{X}_j - \bar{X})^2$$

- Varianza para a proporción:

$$V(\hat{P}) = \frac{1}{k} \sum_{j=1}^k (\hat{P}_j - P)^2$$

2.1.1 Comparación entre a Mostraxe Aleatoria Simple e a Mostraxe Sistemática a partir das súas varianzas

Neste senso cabe destacar outra forma de expresar a varianza da media mostral, que nos permitirá posteriormente levar a cabo unha comparación entre a m.a.s e a mostraxe sistemática (Cid Cid, Delgado Manríquez, & Leguey Galán, 1999):

$$V(\bar{y}) = \frac{N-1}{N} S^2 - \frac{k(n-1)}{N} S_{sis}^2$$

S^2 e S_{sis}^2 , que tamén se pode denotar coma S_{sis}^2 , son a cuasivarianza poboacional e a cuasivarianza dentro das mostras respectivamente, obténdose do seguinte xeito:

$$S^2 = \frac{1}{N-1} \sum_{i=1}^n \sum_{j=1}^k (y_{ij} - \bar{Y})^2$$

$$S_{sis}^2 = \frac{1}{k(n-1)} \sum_{i=1}^n \sum_{j=1}^k (y_{ij} - \bar{Y}_j)^2$$

Se $S^2 = S_{sis}^2$, a precisión da media dunha mostra sistemática é a mesma que a dunha mostra aleatoria simple. Para isto é preciso que exista aleatoriedade na disposición dos elementos da poboación. Por outra banda, se $S_{sis}^2 > S^2$ a variabilidade da media mostral para o caso da mostraxe sistemática será maior, polo que terá unha maior precisión dita mostra. Noutros termos, a precisión da media dunha mostra sistemática será maior canto maior sexa a heteroxeneidade das mostras, pois no caso de existir pouca heteroxeneidade as unidades sucesivas repetirían a información anterior, sendo as estimacións menos precisas (Cid Cid, Delgado Manríquez, & Leguey Galán, 1999).

Á hora de escoller entre a m.a.s. e a mostraxe sistemática, é fundamental ter en conta a forma na que apareza ordenada a poboación. No seguinte apartado observaremos que modalidade de mostraxe resulta máis axeitada segundo o tipo de poboación que estemos a considerar.

3. Casos particulares

3.1 Casos particulares na poboación

No relativo á forma na que esté ordenada a poboación podemos establecer tres casos particulares: poboación aleatoria, poboación periódica e poboación con tendencia lineal (Echeverri Uribe, 2004).

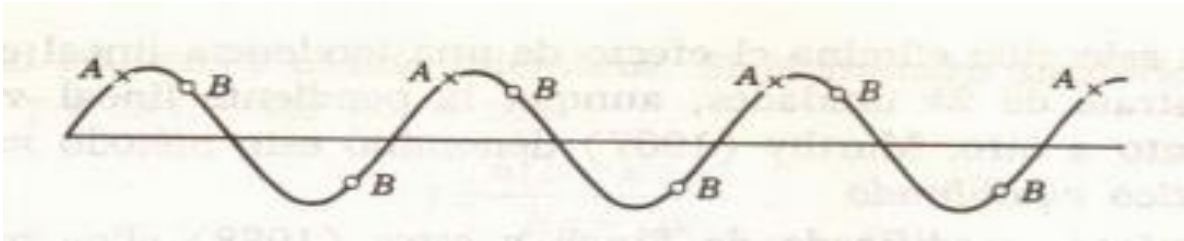
3.1.1 Poboación aleatoria

Os elementos están ordenados ao azar. Polo tanto, a variabilidade dentro da mostra sistemática será moi parecida á existente nunha mostra aleatoria simple, sendo estes dous tipos de mostraxe practicamente idénticos. Un caso que se pon en moitas ocasións coma exemplo de poboación aleatoria é unha lista na que aparecen as unidades poboacionais ordenadas alfabeticamente. Mais é preciso comprobar previamente que non exista unha relación entre as letras iniciais e as características que se estén a estudar. (Azorín & Sánchez-Crespo, 1986).

3.1.2 Poboación periódica

Os elementos están ordenados de forma que seguen unha tendencia cíclica. Neste caso é máis eficiente o uso da m.a.s. que o da mostraxe sistemática, pois se levamos a cabo este proceso co segundo método podemos chegar a atoparnos cunha mostra onde a variabilidade dentro desta é 0, non sendo así representativa. Co obxectivo de minimizar este problema, pódese optar por cambiar o punto de arranque ou inicio cada x elementos (Echeverri Uribe, 2004). Cando o investigador queira levar a cabo un estudo no que se teñan que seleccionar edificios ou bloques, co obxectivo de eliminar o efecto da periodicidade, este pode optar pola numeración “serpentina”, numerando así unha fila de edificios de esquerda a dereita, para continuar na seguinte fila numerando de dereita a esquerda (Azorín & Sánchez-Crespo, 1986). No gráfico 1 podemos observar o caso dunha poboación periódica.

Gráfico 1. Mostraxe sistemática dunha poboación periódica

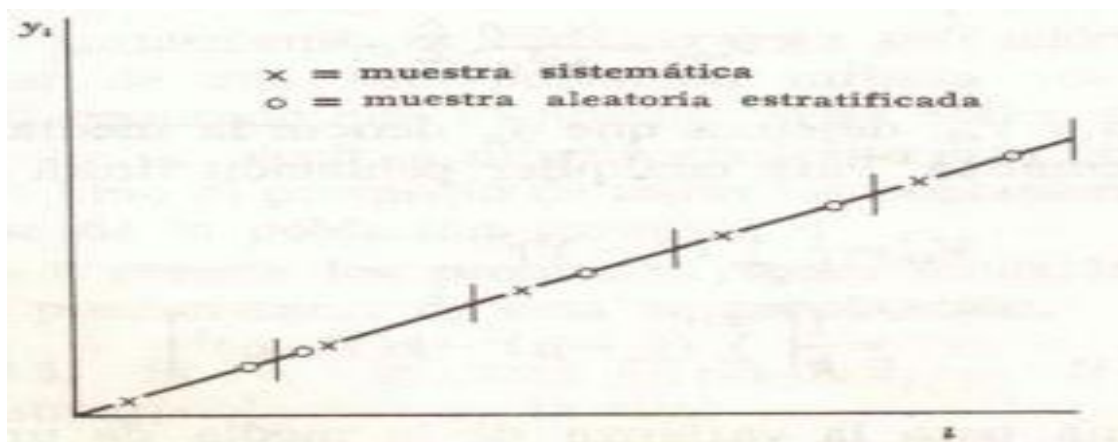


Fonte: Cochran (1987).

3.1.3 Poboación con tendencia lineal

Os elementos están ordenados seguindo un criterio ou esquema determinado. Neste caso podemos encontrar ante determinados estudos nos que se chegue a estimacións equivocadas. Isto pode darse cando estamos considerando ás persoas por orde de estatura (Instituto Interamericano de Ciencias Agrícolas de la O.E.A, 1958). É dicir, imaxinemos que necesitamos obter un número entre 1 e 10, os cales representan centímetros de estatura, para obter unha mostra sistemática. Se nos sae o 9, as mostras obtidas serán moi altas, polo que terá lugar unha sobreestimación da estatura media. Para que poidamos obter unha selección sistemática representativa, os valores das unidades seleccionadas deberán ter unha alta dispersión, formando así grupos heteroxéneos con respecto á característica que estemos a estudar (Gutiérrez Rojas, 2016). Neste tipo de poboacións poderanse mellorar os resultados da mostraxe empregando un deseño estratificado, pois leva consigo unha maior aleatoriedade na selección, a cal non se dá no caso da mostraxe sistemática (Mendenhall, Ott, & Scheaffer, 2006). Podemos observar no seguinte gráfico o caso dunha poboación con tendencia lineal, distinguindo segundo consideremos unha mostra sistemática ou unha mostra estratificada.

Gráfico 2. Mostraxe sistemática dunha poboación con tendencia lineal



Fonte: Cochran (1987).

3.2 Casos particulares na mostraxe sistemática

3.2.1 Mostraxe sistemática replicada

Este método consiste na selección de varias mostras sistemáticas, de xeito que se van obter un número de medias mostrais que vai permitir que nos acheguemos a unha estimación máis que aceptable dos parámetros poboacionais. A través desta mostraxe conséguese eliminar o efecto causado pola periodicidade, de aí gran parte do seu atractivo. O número de mostras que se deben seleccionar está en torno a 10. Polo tanto, para a súa aplicación opérase do seguinte xeito (Echeverri Uribe, 2004):

- Obter o valor de k , sendo $k=N/n$
- Obter o valor de k' , sendo $k'= 10*k$
- Xerar 10 números aleatorios entre 1 e k' . Deste xeito, cada un dos números xerados será o punto de arranque para cada unha das mostras sistemáticas
- Para calcular o segundo elemento de cada unha das mostras, só hai que sumarlle k' aos puntos de arranque ou inicio. E así sucesivamente, ata conseguir todo o tamaño da mostra que busquemos.

3.2.2 Mostraxe sistemática especial

Hai moitos casos no día a día en que á hora de realizar un estudo de mostraxe non se pode coñecer con exactitude o valor de N . Este é o caso do número de coches que pasan por unha estrada, número de persoas que acuden a unha praia un día determinado... é dicir, estamos a falar de poboacións en continua formación. Non é posible aplicar a m.a.s., por non dispoñer da poboación enumerada, nin aplicar a mostraxe aleatoria sistemática, ante a imposibilidade de obter " k ". Polo tanto, a persoa encargada de levar a cabo o estudo estatístico, analizando o contexto da cuestión a tratar, escollerá " k ". A partir de aquí o procedemento exécutase do mesmo modo que no caso da mostraxe aleatoria sistemática. É dicir, selecciona un número entre 1 e k , que vai ser o punto de inicio, e posteriormente sumárase a " k " de forma sucesiva ata conseguir a mostra completa (Echeverri Uribe, 2004). É dicir, a principal diferenza nesta modalidade de mostraxe atópase na "subxectividade" para decidir cal é a " k ".

4. Mostras interpenetrantes

Tendo presente que cunha única mostra sistemática non vai ser posible a obtención dun estimador innesgado da varianza (Portela García-Miguez & Villeta López, 2007), o investigador poderá calculalo levando a cabo unha mostraxe sistemática replicada. Mediante este deseño, a mostra divídese en dúas ou máis submostras as cales son independentes entre sí. Cabe destacar que se pode levar a cabo nos principais métodos de mostraxe: aleatorio simple, estratificado, conglomerados e sistemático. (Vivanco, 2013)

Deste xeito, non tomaremos unha mostra sistemática de tamaño n , senón que obteremos m mostras, tendo cada unha destas un tamaño igual a $n' = n/m$. A lonxitude do intervalo para a selección das mostras será $k' = N/n'$. Sendo \hat{y}_i a media mostral da mostra sistemática, obteremos:

$$\widehat{\bar{y}}_m = \frac{1}{m} \sum_{i=1}^m \hat{y}_i, \text{ que é o estimador innesgado de } \bar{y}$$

É dicir, tendo en conta que na m.a.s. a media mostral é innesgada para a media poboacional, $\widehat{\bar{y}}_m$ tamén o será, pois é a media mostral dos m elementos \hat{y}_i , que se obteñen por m.a.s. Así, debemos ter presente que cada submostra é unha selección aleatoria da mostra maior.

Noutros termos, trátase de obter un conxunto de m mostras, obtendo destas o seu estimador, a través do seu valor medio. Polo tanto, súmanse os estimadores de cada mostra e dividídense entre o número de mostras que teñamos. Así, poderemos obter a varianza como:

$$V(\widehat{\bar{y}}_m) = \frac{1}{m(m-1)} \sum_{i=1}^m (\hat{y}_i - \bar{y})^2$$

Ademais, no relativo á validez do estudo, este método axuda a examinar a calidade da información, no senso de que estas submostras interpenetrantes teñen utilidade á hora de verificar a información acerca de erros non mostrais, coma poden ser aqueles nesgos introducidos polos entrevistador (Naciones Unidas, 2009).

5. Comparación entre unha mostra sistemática e unha mostra aleatoria simple: caso particular do Grao en Economía na Universidade de Santiago de Compostela

Supoñamos que queremos coñecer o nivel de satisfacción dos alumnos do Grao en Economía da Universidade de Santiago de Compostela co funcionamento do transporte público dende a facultade ás súas respectivas vivendas. Para iso levaremos a cabo dúas modalidades de mostraxe, coma son a mostraxe sistemática e a mostraxe aleatoria simple.

No Grao en Economía hai 500 alumnos, os cales se atopan repartidos entre os 4 cursos do seguinte xeito:

1º curso: 150 alumnos

2º curso: 150 alumnos

3º curso: 120 alumnos

4º curso: 80 alumnos

Á súa vez, estes alumnos estarán divididos en grupos expositivos e interactivos na forma en que se amosa na seguinte táboa:

Táboa 4. Distribución do alumnado do Grao en Economía en grupos expositivos e interactivos segundo cursos

Curso	G.E ₁	G.E ₂	G.E ₃	G.I ₁	G.I ₂	G.I ₃	G.I ₄	G.I ₅	G.I ₆
1º	50	50	50	25	25	25	25	25	25
2º	50	50	50	25	25	25	25	25	25
3º	40	40	40	20	20	20	20	20	20
4º	40	40		20	20	20	20		

Fonte: *Elaboración propia.*

Podemos apreciar na táboa que cada grupo expositivo se divide en dous grupos interactivos. Para os tres primeiros cursos atopámonos con 3 grupos expositivos e, polo tanto, 6 grupos interactivos por ano. A excepción neste senso constitúea o último curso do grao, no que se divide ao alumnado en 2 grupos expositivos e 4 interactivos, dado o menor número de alumnos para este caso.

Previamente á obtención dunha mostra sistemática e aleatoria simple consideramos axeitado distribuír o alumnado de cada un dos catro cursos do Grao en Economía nos correspondentes grupos de docencia expositiva e interactiva, de xeito que nos resulte sinxelo e intuitivo comprobar a que grupo pertence cada unidade seleccionada para a mostra. Esa distribución en grupos expositivos e interactivos é a que se recolle nas táboas que figuran a continuación.

Táboa 5. Distribución do alumnado do 1º curso do Grao en Economía nos diferentes grupos expositivos e interactivos

1º CURSO	
G.E ₁ (1-50)	G.I ₁ (1-25)
	G.I ₂ (25-50)
G.E ₂ (50-100)	G.I ₃ (50-75)
	G.I ₄ (75-100)
G.E ₃ (100-150)	G.I ₅ (100-125)
	G.I ₆ (125-150)

Fonte: *Elaboración propia.*

Táboa 6. Distribución do alumnado do 2º curso do Grao en Economía nos diferentes grupos expositivos e interactivos

2º CURSO	
G.E ₁ (150-200)	G.I ₁ (150-175)
	G.I ₂ (175-200)
G.E ₂ (200-250)	G.I ₃ (200-225)
	G.I ₄ (225-250)
G.E ₃ (250-300)	G.I ₅ (250-275)
	G.I ₆ (275-300)

Fonte: *Elaboración propia.*

Táboa 7. Distribución do alumnado do 3º curso do Grao en Economía nos diferentes grupos expositivos e interactivos

3º CURSO	
G.E ₁ (300-340)	G.I ₁ (300-320)
	G.I ₂ (320-340)
G.E ₂ (340-380)	G.I ₃ (340-360)
	G.I ₄ (360-380)
G.E ₃ (380-420)	G.I ₅ (380-400)
	G.I ₆ (400-420)

Fonte: *Elaboración propia.*

Táboa 8. Distribución do alumnado do 4º curso do Grao en Economía nos diferentes grupos expositivos e interactivos

4º CURSO	
G.E ₁ (420-460)	G.I ₁ (420-440)
	G.I ₂ (440-460)
G.E ₂ (460-500)	G.I ₃ (460-480)
	G.I ₄ (480-500)

Fonte: *Elaboración propia*

Así, cunha poboación de $N = 500$ queremos extraer unha mostra de $n = 50$ para poder levar a cabo a enquisa. Realizando unha mostraxe sistemática, procederemos do seguinte modo:

- O intervalo aleatorio de selección é $k = N/n$. Neste caso, $k = 500/50 = 10$.
- O punto de arranque, U_j , deberá estar comprendido entre 1 e k . Neste caso tal punto de arranque será $U_j = 6$.
- Deste xeito, a mostra que obteñamos será a que aparece destacada na seguinte táboa.

Táboa 9. Diferentes mostras sistemáticas posibles para o Grao en Economía con N=500 e n=50

M ₁	M ₂	M ₃	M ₄	M ₅	M ₆	M ₇	M ₈	M ₉	M ₁₀
1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100
101	102	103	104	105	106	107	108	109	110
111	112	113	114	115	116	117	118	119	120
121	122	123	124	125	126	127	128	129	130
131	132	133	134	135	136	137	138	139	140
141	142	143	144	145	146	147	148	149	150
151	152	153	154	155	156	157	158	159	160
161	162	163	164	165	166	167	168	169	170
171	172	173	174	175	176	177	178	179	180
181	182	183	184	185	186	187	188	189	190
191	192	193	194	195	196	197	198	199	200
201	202	203	204	205	206	207	208	209	210
211	212	213	214	215	216	217	218	219	220
221	222	223	224	225	226	227	228	229	230
231	232	233	234	235	236	237	238	239	240
241	242	243	244	245	246	247	248	249	250
251	252	253	254	255	256	257	258	259	260
261	262	263	264	265	266	267	268	269	270
271	272	273	274	275	276	277	278	279	280
281	282	283	284	285	286	287	288	289	290
291	292	293	294	295	296	297	298	299	300
301	302	303	304	305	306	307	308	309	310

311	312	313	314	315	316	317	318	319	320
321	322	323	324	325	326	327	328	329	330
331	332	333	334	335	336	337	338	339	340
341	342	343	344	345	346	347	348	349	350
351	352	353	354	355	356	357	358	359	360
361	362	363	364	365	366	367	368	369	370
371	372	373	374	375	376	377	378	379	380
381	382	383	384	385	386	387	388	389	390
391	392	393	394	395	396	397	398	399	400
401	402	403	404	405	406	407	408	409	410
411	412	413	414	415	416	417	418	419	420
421	422	423	424	425	426	427	428	429	430
431	432	433	434	435	436	437	438	439	440
441	442	443	444	445	446	447	448	449	450
451	452	453	454	455	456	457	458	459	460
461	462	463	464	465	466	467	468	469	470
471	472	473	474	475	476	477	478	479	480
481	482	483	484	485	486	487	488	489	490
491	492	493	494	495	496	497	498	499	500

Fonte: *Elaboración propia*

Podemos así observar como todos os grupos dos 4 cursos, tanto expositivos coma interactivos, quedan representados, poñendo así de manifesto unha das grandes vantaxes que presenta este tipo de mostraxe. E isto sucedería en calquera das 10 mostras que se poden dar para este caso. Sen embargo, isto non ocorre así no caso da mostraxe aleatoria simple.

Paralelamente, e tal e como se amosa no apartado 1.3., podemos considerar a mostraxe sistemática coma un caso particular da mostraxe estratificada e mostraxe por conglomerados. Desta maneira, cada unha das columnas (conglomerado) vai constituír unha mostra posible, incluíndo a todos os grupos expositivos e interactivos dos 4 cursos. Por outra banda, cada unha das filas da táboa, que vai incluír un intervalo de 10 unidades, poderá ser considerada coma un estrato. Polo tanto, no caso do Grao en Economía podemos diferenciar 10 conglomerados diferentes, así coma 50 estratos, escollendo unha unidade de cada un destes últimos para a constitución da mostra.

Comentado o anterior, será interesante obter unha mostra a través do método da mostraxe aleatoria simple, concretamente sen reposición. Sendo $N = 500$ e $n = 50$, o número de mostras que se poden obter é: $\binom{N}{n} = \binom{500}{50} = 2.314e + 69$. Sinalado isto, unha mostra aleatoria simple podería ser a seguinte:

M.A.S: 4, 6, 13, 17, 21, 58, 63, 64, 116, 123, 129, 141, 149, 154, 179, 184, 185, 226, 227, 245, 248, 271, 277, 283, 297, 299, 310, 341, 343, 349, 355, 362, 368, 381, 390, 397, 404, 407, 441, 443, 444, 452, 458, 461, 467, 468, 481, 485, 489, 498.

Así, podemos ver a mostra representada na seguinte táboa.

Táboa 10. Mostra aleatoria simple do Grao en Economía con N=500 e n=50

CURSO	GRUPOS		MOSTRAXE ALEATORIA SIMPLE
1º CURSO	G.E ₁	G.l ₁	4, 6, 13, 17, 21
		G.l ₂	
	G.E ₂	G.l ₃	58, 63, 64
		G.l ₄	
	G.E ₃	G.l ₅	116, 123
		G.l ₆	129, 141, 149
2º CURSO	G.E ₁	G.l ₁	154
		G.l ₂	179, 184, 185
	G.E ₂	G.l ₃	
		G.l ₄	226, 227, 245, 248
	G.E ₃	G.l ₅	271
		G.l ₆	277, 283, 297, 299
3º CURSO	G.E ₁	G.l ₁	310
		G.l ₂	
	G.E ₂	G.l ₃	341, 343, 349, 355
		G.l ₄	362, 368
	G.E ₃	G.l ₅	381, 390, 397
		G.l ₆	404, 407
4º CURSO	G.E ₁	G.l ₁	
		G.l ₂	441, 443, 444, 452, 458
	G.E ₂	G.l ₃	461, 467, 468,
		G.l ₄	481, 485, 489, 498

Fonte: *Elaboración propia.*

Observando esta mostra resáltase unha gran diferenza con respecto á outra modalidade de mostraxe empregada, coma é que non todos os grupos (expositivos e interactivos) quedan representados. Neste caso particular sí que se recolle algunha unidade de cada grupo expositivo, mais non se pode afirmar o mesmo para os grupos interactivos. Así, non se amosa ningunha unidade dos grupos interactivos 2 e 4 do 1º curso, do grupo interactivo 3 do 2º curso, do grupo interactivo 2 do 3º curso nin do grupo interactivo 1 do 4º curso.

Incluso podería darse o caso no que non quedase representada ningunha unidade dun curso enteiro, como pode ser na mostra seguinte:

M.A.S: 2, 7, 15, 18, 34, 36, 49, 50, 53, 55, 74, 79, 81, 86, 88, 96, 99, 128, 129, 134, 138, 146, 149, 153, 155, 158, 163, 173, 191, 204, 210, 221, 230, 245, 247, 252, 268, 284, 299, 317, 329, 335, 356, 359, 382, 393, 398, 407, 413, 417.

Que nunha táboa coma a anterior quedaría reproducida do seguinte xeito:

Táboa 11. Mostra aleatoria simple do Grao en Economía con N=500 e n=50 sen representación do 4º curso

CURSO	GRUPOS		MOSTRAXE ALEATORIA SIMPLE
1º CURSO	G.E ₁	G.l ₁	2, 7, 15, 18
		G.l ₂	34, 36, 49, 50
	G.E ₂	G.l ₃	53, 55, 74
		G.l ₄	79, 81, 86, 88, 96, 99
	G.E ₃	G.l ₅	
		G.l ₆	128, 129, 134, 138, 146, 149
2º CURSO	G.E ₁	G.l ₁	153, 155, 158, 163, 173
		G.l ₂	191
	G.E ₂	G.l ₃	204, 210, 221
		G.l ₄	230, 245, 247
	G.E ₃	G.l ₅	252, 268
		G.l ₆	284, 299
3º CURSO	G.E ₁	G.l ₁	317
		G.l ₂	329, 335
	G.E ₂	G.l ₃	356, 359
		G.l ₄	
	G.E ₃	G.l ₅	382, 393, 398
		G.l ₆	407, 413, 417
4º CURSO	G.E ₁	G.l ₁	
		G.l ₂	
	G.E ₂	G.l ₃	
		G.l ₄	

Fonte: *Elaboración propia.*

Nesta mostra dáse o caso no que ningún alumno do 4º curso aparece representado, os cales se corresponden coas unidades comprendidas entre 420 e 500. Así, as dúas táboas anteriores poñen de manifesto coma neste caso vai ser máis recomendable o emprego da mostraxe sistemática, incluíndose alumnos de todos os grupos expositivos e interactivos dos 4 cursos que comprende o Grao en Economía da Universidade de Santiago de Compostela.

Conclusións e ampliación

Ao longo deste traballo poidemos comprobar coma a mostraxe sistemática se trata dun método moi vantaxoso, e mesmo única alternativa viable nalgúns casos determinados, como pode ser cando pretendamos obter unha mostra nunha estrada ou á saída dun teatro, a partir da selección de 1 de cada k elementos. Ademais das vantaxes e desvantaxes que sinalan as diferentes fontes bibliográficas consultadas, mediante a aplicación levada a cabo cos datos do Grao en Economía da Universidade de Santiago de Compostela confirmáronse algunhas destas, tales coma a representación de unidades de todos os grupos, neste caso expositivos e interactivos, así coma a doada obtención da mostra. E esta comparación entre tipos de mostraxe tamén se levou a cabo dende un ámbito máis matemático, desenvolvendo os estimadores do total, media e proporción, así coma as súas varianzas, o cal nos permitiu ver en que casos era preferible ou máis acertado o emprego dunha ou outra modalidade. Para isto consideramos os tres tipos de poboación ante os que nos podemos encontrar, e que desenvolvimos nun primeiro apartado de casos particulares, coma son a poboación periódica e con tendencia lineal. Ademais, tamén resultou relevante incluír neste traballo a consideración da mostraxe sistemática coma caso particular da mostraxe estratificada ou por conglomerados. Considerouse que foi necesario a inclusión de diferentes táboas para unha mellor visualización dos resultados, tendo presente que nalgún caso se empregou un marco de 200 alumnos.

Non obstante, ante os inconvintes que se foron descubrindo no relativo a determinados aspectos da mostraxe sistemática, houbo resposta por parte dos investigadores. Este é o caso da mostraxe sistemática replicada, que consegue eliminar o efecto da periodicidade na poboación. Esta desenvolveuse no apartado de mostraxas interpenetrantes, as cales nos axudan a obter un estimador innesgado da varianza. Ademais, poidemos apreciar coma este método tamén é aplicable para outras modalidades, coma é o caso da mostraxe aleatoria simple. É dicir, fóronse mellorando diferentes aspectos no relativo a este tipo de mostraxe, que, aínda así, dista de ser perfecto para o seu emprego en determinados ámbitos. Por exemplo, para o caso das enquisas dunhas eleccións, coma poden ser as autonómicas galegas do 12 de xullo, a mostraxe sistemática non levaría consigo unha gran representación da poboación. É dicir, podería producirse unha concentración das unidades seleccionadas nun determinado sexo ou idade, pois é posible que non se coñeza previamente a forma en que estea ordenada a poboación. Neste caso particular sería moito máis preciso o emprego da mostraxe polietápica, a cal

podemos considerar coma unha mostraxe por conglomerados, mais obtendo a mostra por etapas. Así, para os diferentes estudos que se poidan levar a cabo, haberá un método que se axuste mellor que o resto. Dito o anterior, poidemos observar en que casos concretos é máis recomendable a mostraxe sistemática, o cal se consegue grazas aos coñecementos teóricos e instrumentos estatísticos explicados ao longo deste traballo.

Por outra banda, no relativo ás posibles ampliacións, estas poderían vir do lado da aplicación empírica. Sen embargo, ante a situación extraordinaria na que nos atopamos inmersos, non foi posible levar a cabo esta tarefa. Tal e como tiñamos previsto ao comezo, a idea era realizar unha aplicación desta mostraxe, empregando unha serie de cuestionarios contestados por alumnos dos Graos en Economía e Administración e Dirección de Empresas da Universidade de Santiago de Compostela, relacionados estes co seu rendemento nas materias de tipo cuantitativo. Así, estes estaban formados por preguntas vinculadas co ámbito socioeconómico dos enquisados, os seus resultados en materias tales coma Matemáticas, Estatística e Econometría, así coma a especialidade de bacharelato de procedencia ou a asistencia a academias. Deste xeito, sería posible a posta en práctica das ferramentas estatísticas no ámbito da mostraxe que nos ocupa, coa axuda á vez dun software coma o SSPS. Ademais, a análise con datos actualizados ten unha dobre vantaxe. Dun lado, o coñecemento dos factores que están por detrás do baixo rendemento nas materias de tipo cuantitativo da nosa facultade, o cal é moi útil para a posible toma de medidas e aplicación de políticas nun futuro. Doutro lado, podemos confirmar as vantaxes que trae consigo o emprego desta mostraxe nun escenario real, o cal consolida o traballo realizado nesta investigación.

Bibliografía

- Azorín, F., & Sánchez-Crespo, J. L. (1986). *Métodos y aplicaciones del muestreo*. Madrid: Alianza Universidad Textos.
- Blanco Villaseñor, Á., Castellano Paulis, J., & Hernández Mendo, A. (2000). Generalizabilidad de las observaciones de la acción del juego en el fútbol. *Psicothema*, 12(2), 81-86.
- Cid Cid, A. I., Delgado Manríquez, C.A., & Leguey Galán, S. (1999). *Introducción al muestreo en poblaciones finitas*. Madrid: Nuevas Estructuras S.L.
- Cochran, W. G. (1987). *Técnicas de muestreo*. Ciudad de México: Compañía Editorial Continental.
- Echeverri Uribe, T. (2004). Estadísticas para las ciencias contables: extracción de la muestra en los diferentes diseños muestrales y cálculo de su tamaño. En T. Echeverri Uribe, *Estadísticas para las ciencias contables* (Vol. 45, págs. 19-88). Medellín: Universidad de Antioquia. Facultad de Ciencias Económicas.
- Fernandez Álvaro, C., López Blázquez, F., & Salamanca Miño, B. (2001). Aplicación del muestreo sistemático en el diseño de encuestas. *Metodología de Encuestas*, 3(1), 65-73.
- Gutiérrez Rojas, H. A. (2016). *Estrategias de Muestreo: Diseño de encuestas y estimación de parámetros*. Bogotá: Ediciones de la U.
- Hedayat, A.S. y Sinha, B. (1991). *Design and Inference in Finite Population Sampling*. New York: Wiley
- Instituto Interamericano de Ciencias Agrícolas de la O.E.A. (1958). *Curso Internacional de Economía Agrícola* (págs. 35-38). Lima.
- Kruskal, W. & Mosteller, F. (1979, 1980). *Representative sampling*. Int. Stat. Review, vols. 47 e 48.
- López, P. L. (2004). Población, muestra y muestreo. *Punto Cero*, 9(8), 69-74.
- Mendenhall, W., Ott, R. L., & Scheaffer, R. L. (2006). *Elementos de Muestreo*. Madrid: Thomson.
- Naciones Unidas (2009). *Diseño de muestras para encuestas de hogares: directrices prácticas*. Nueva York: Naciones Unidas.
- Pérez López, C. (1999). *Técnicas de muestreo estadístico: teoría, práctica y aplicaciones informáticas*. Madrid: Ra-Ma.
- Pérez López, C. (2005). *Muestreo estadístico: Conceptos y problemas resueltos*. Madrid: Pearson Educación, S.A.

- Pérez Suárez, R. (1986). *Nociones básicas de estadística*. Oviedo: Servicio de Publicaciones Universidad de Oviedo.
- Portela García-Miguez, J., & Villeta López, M. (2007). Muestreo Sistemático. En J. Portela García-Miguez e M. Villeta López, *Técnicas básicas de muestreo con SAS* (págs. 181-195). Madrid: Facultad de Estudios Estadísticos.
- Quintana, C. (1996). *Elementos de inferencia estadística*. San José: Editorial de la Universidad de Costa Rica.
- Rodríguez Osuna, J. (2001). *Métodos de muestreo*. Madrid: Centro de Investigaciones Sociológicas.
- Tamayo, G. (2001). Diseños muestrales en la investigación. *Semestre Económico*, 4(7), 1-14.
- Vivanco, M. (2013). *Muestreo estadístico. Diseños y aplicaciones*. Santiago de Chile: Editorial Universitaria .