

CAPÍTULO V. NÚMERO DE ESTRATOS

En el capítulo II (II.2.8) se mencionó que el aumentar la cantidad de estratos al realizar muestreo estratificado, permite la obtención de una cierta ganancia; asimismo, mediante una pequeña demostración se expuso que cuando se realiza el muestreo utilizando una variable distinta a la de interés, existe un cierto número de estratos a partir del cual no se obtendrá una ganancia significativa y aumentarlos no tendrá utilidad alguna, además se podría incurrir en muchos más gastos, factor que no es deseable al realizar un diseño de muestra.

Con el fin de descubrir, para la población utilizada, los beneficios que se obtendrían al aumentar el número de estratos, se realizaron los cálculos de obtención de estratos, óptimos de tamaño de muestra y varianzas de la media estratificada utilizando la primera componente principal como base de muestreo, para 3, 4, 5 y 6 estratos.

La motivación para utilizar como base de muestreo únicamente a la primera componente principal, se debe a que en el capítulo anterior se descubrió que ésta por si sola acarrea los mayores beneficios que al utilizar cualquier otra de las variables de la población.

V.1 Cálculo de Límites y Óptimos por Número de Estratos

Las tablas correspondientes al proceso de búsqueda de estratos, sus límites correspondientes, y tamaños óptimos de muestra en cada uno de ellos para 3, 4, 5 y 6 estratos utilizando como base la primera componente, se siguen a continuación:

Considerando 3 estratos:

Tabla 5.1 Construcción de estratos (3)

Intervalo		Lim Inferior	Lim Superior	Frecuencia	$\sqrt{\text{Frec}}$	Acum.		
1	1	-20	-19	2	1.414214	1.41421		
	2	-19	-18	1	1	2.41421		
	3	-18	-17	5	2.236068	4.65028		
	4	-17	-16	3	1.732051	6.38233		
	5	-16	-15	2	1.414214	7.79655		
	6	-15	-14	13	3.605551	11.4021		
	7	-14	-13	16	4	15.4021		
	8	-13	-12	32	5.656854	21.059		
	9	-12	-11	52	7.211103	28.2701		
	10	-11	-10	88	9.380832	37.6509		
	11	-10	-9	111	10.53565	48.1865		
	12	-9	-8	156	12.49	60.6765		
	13	-8	-7	190	13.78405	74.4606		
	14	-7	-6	346	18.60108	93.0617		
	15	-6	-5	466	21.58703	114.649		
	16	-5	-4	660	25.69047	140.339		
	17	-4	-3	958	30.95158	171.291		X
	18	-3	-2	1305	36.12478	207.416		
	19	-2	-1	1685	41.04875	248.464		
	20	-1	0	1807	42.50882	290.973		
	21	0	1	1692	41.13393	332.107		X
	22	1	2	1511	38.87158	370.979		
	23	2	3	1328	36.44173	407.42		
	24	3	4	952	30.8545	438.275		
	25	4	5	646	25.41653	463.691		
	26	5	6	504	22.44994	486.141		
	27	6	7	459	21.42429	507.566		
	28	7	8	426	20.63977	528.205		
	29	8	9	259	16.09348	544.299		
	30	9	11	39	6.244998	550.544	183.515	X

Fuente: Elaboración propia

Tabla 5.2 Límites de estrato (3)

Rango		Límites de Estrato	
0	183.514612	-21	-3
183.5146117	367.029223	-3	1
367.0292234	550.543835	1	mas

Fuente: Elaboración Propia

Tabla 5.3 Cálculo de óptimos para el tamaño de muestra

n	No. Estrato	N_h	S_h	N_hS_h	n_h		W_h
200	1	3101	2.44967133	7596.4308	54.23891	54	0.19734
	2	6489	1.10873296	7194.5682	51.36959	51	0.412944
	3	6124	2.15872037	13220.004	94.39151	95	0.389716
	Σ	15714		28011.003		200	1

Fuente: Elaboración propia

Considerando 4 estratos:

Tabla 5.4 Construcción de estratos (4)

Intervalo		Lim Inferior	Lim Superior	Frecuencia	$\sqrt{\text{Frec.}}$	Acum.		
1	1	-20	-19	2	1.414214	1.41421		
	2	-19	-18	1	1	2.41421		
	3	-18	-17	5	2.236068	4.65028		
	4	-17	-16	3	1.732051	6.38233		
	5	-16	-15	2	1.414214	7.79655		
	6	-15	-14	13	3.605551	11.4021		
	7	-14	-13	16	4	15.4021		
	8	-13	-12	32	5.656854	21.059		
	9	-12	-11	52	7.211103	28.2701		
	10	-11	-10	88	9.380832	37.6509		
	11	-10	-9	111	10.53565	48.1865		
	12	-9	-8	156	12.49	60.6765		
	13	-8	-7	190	13.78405	74.4606		
	14	-7	-6	346	18.60108	93.0617		
	15	-6	-5	466	21.58703	114.649		X
	16	-5	-4	660	25.69047	140.339		
	17	-4	-3	958	30.95158	171.291		
	18	-3	-2	1305	36.12478	207.416		
	19	-2	-1	1685	41.04875	248.464		X
	20	-1	0	1807	42.50882	290.973		
	21	0	1	1692	41.13393	332.107		
	22	1	2	1511	38.87158	370.979		
	23	2	3	1328	36.44173	407.42		X
	24	3	4	952	30.8545	438.275		
	25	4	5	646	25.41653	463.691		
	26	5	6	504	22.44994	486.141		
	27	6	7	459	21.42429	507.566		
	28	7	8	426	20.63977	528.205		
	29	8	9	259	16.09348	544.299		
	30	9	11	39	6.244998	550.544	137.636	X

Fuente: Elaboración propia

Tabla 5.5 Límites de estrato (4)

Rango		Límites de Estrato	
0	137.635959	-21	-5
137.6359588	275.271918	-5	-1
275.2719176	412.907876	-1	3
412.9078763	550.543835	3	mas

Fuente: Elaboración propia

Tabla 5.6 Cálculo de óptimos para el tamaño de muestra

n	No. Estrato	N _h	S _h	N _h S _h	n _h		W _h
200	1	1483	2.30518913	3418.5955	32.1012	32	0.094374
	2	4608	1.09429556	5042.5139	47.35007	47	0.293242
	3	6338	1.13408126	7187.8071	67.49474	68	0.403335
	4	3285	1.71992354	5649.9488	53.05399	53	0.209049
	Σ	15714		21298.865		200	1

Fuente: Elaboración Propia

Considerando 5 estratos:

Tabla 5.7 Construcción de estratos (5)

Intervalo		Lim Inferior	Lim Superior	Frecuencia	√Frec	Acum.		
1	1	-21	-20	2	1.414214	1.41421		
	2	-20	-19	0	0	1.41421		
	3	-19	-18	1	1	2.41421		
	4	-18	-17	5	2.236068	4.65028		
	5	-17	-16	3	1.732051	6.38233		
	6	-16	-15	2	1.414214	7.79655		
	7	-15	-14	13	3.605551	11.4021		
	8	-14	-13	16	4	15.4021		
	9	-13	-12	32	5.656854	21.059		
	10	-12	-11	52	7.211103	28.2701		
	11	-11	-10	88	9.380832	37.6509		
	12	-10	-9	111	10.53565	48.1865		
	13	-9	-8	156	12.49	60.6765		
	14	-8	-7	190	13.78405	74.4606		
	15	-7	-6	346	18.60108	93.0617		X
	16	-6	-5	466	21.58703	114.649		

Fuente: Elaboración propia

Tabla 5.7 (Continuación)

Intervalo		Lim Inferior	Lim Superior	Frecuencia	$\sqrt{\text{Frec}}$	Acum.		
1	17	-5	-4	660	25.69047	140.339		
	18	-4	-3	958	30.95158	171.291		
	19	-3	-2	1305	36.12478	207.416		X
	20	-2	-1	1685	41.04875	248.464		
	21	-1	0	1807	42.50882	290.973		X
	22	0	1	1692	41.13393	332.107		
	23	1	2	1511	38.87158	370.979		
	24	2	3	1328	36.44173	407.42		
	25	3	4	952	30.8545	438.275		X
	26	4	5	646	25.41653	463.691		
	27	5	6	504	22.44994	486.141		
	28	6	7	459	21.42429	507.566		
	29	7	8	426	20.63977	528.205		
	30	8	9	259	16.09348	544.299		
	31	9	10	36	6	550.299		
	32	10	11	3	1.732051	552.031	110.406	X

Tabla 5.8 Límites de estrato (5)

Rango		Límites de Estrato	
0	110.40618	-21	-6
110.4061776	220.81236	-6	-2
220.8123552	331.21853	-2	0
331.2185328	441.62471	0	4
441.6247104	552.03089	4	más

Fuente: Elaboración propia

Tabla 5.9 Cálculo de óptimos para el tamaño de muestra

n	No. Estrato	N_h	S_h	$N_h S_h$	n_h		W_h
200	1	1017	2.2226792	2260.4648	26.03549	26	0.064719
	2	3389	1.0932554	3705.0427	42.67379	43	0.215668
	3	3492	0.5747723	2007.105	23.11735	23	0.222222
	4	5483	1.1086581	6078.7726	70.01385	70	0.348925
	5	2333	1.4201037	3313.1019	38.15951	38	0.148466
	Σ	15714		17364.487		200	1

Fuente: Elaboración propia

Considerando 6 estratos:

Tabla 5.10 Construcción de estratos (6)

Intervalo		Lim Inferior	Lim Superior	Frecuencia	$\sqrt{\text{Frec.}}$	Acum.		
1	1	-20	-19	2	1.414214	1.41421		
	2	-19	-18	1	1	2.41421		
	3	-18	-17	5	2.236068	4.65028		
	4	-17	-16	3	1.732051	6.38233		
	5	-16	-15	2	1.414214	7.79655		
	6	-15	-14	13	3.605551	11.4021		
	7	-14	-13	16	4	15.4021		
	8	-13	-12	32	5.656854	21.059		
	9	-12	-11	52	7.211103	28.2701		
	10	-11	-10	88	9.380832	37.6509		
	11	-10	-9	111	10.53565	48.1865		
	12	-9	-8	156	12.49	60.6765		
	13	-8	-7	190	13.78405	74.4606		X
	14	-7	-6	346	18.60108	93.0617		
	15	-6	-5	466	21.58703	114.649		
	16	-5	-4	660	25.69047	140.339		
	17	-4	-3	958	30.95158	171.291		X
	18	-3	-2	1305	36.12478	207.416		
	19	-2	-1	1685	41.04875	248.464		X
	20	-1	0	1807	42.50882	290.973		
	21	0	1	1692	41.13393	332.107		X
	22	1	2	1511	38.87158	370.979		
	23	2	3	1328	36.44173	407.42		
	24	3	4	952	30.8545	438.275		X
	25	4	5	646	25.41653	463.691		
	26	5	6	504	22.44994	486.141		
	27	6	7	459	21.42429	507.566		
	28	7	8	426	20.63977	528.205		
	29	8	9	259	16.09348	544.299		
	30	9	11	39	6.244998	550.544	91.7573	X

Fuente: Elaboración propia

Tabla 5.11 Límites de estrato (6)

Rango		Límites de Estrato	
0	91.7573059	-21	-7
91.75730586	183.514612	-7	-3
183.5146117	275.271918	-3	-1
275.2719176	367.029223	-1	1
367.0292234	458.786529	1	4
458.7865293	550.543835	4	mas

Fuente: Elaboración propia

Tabla 5.12 Cálculo de óptimos para el tamaño de muestra

n	No. Estrato	N _h	S _h	N _h S _h	n _h		W _h
200	1	671	2.112652	1417.5895	19.81201	20	0.042701
	2	2430	1.11432445	2707.8084	37.84392	38	0.154639
	3	2990	0.56939755	1702.4987	23.79386	24	0.190276
	4	3499	0.57102802	1998.027	27.92412	28	0.222668
	5	3791	0.83655436	3171.3776	44.32269	44	0.24125
	6	2333	1.42010369	3313.1019	46.3034	46	0.148466
	Σ	15714		14310.403		200	1

Fuente: Elaboración propia

IV.2 Cálculo de Varianzas de la Media Estratificada

Con los límites de los estratos y el tamaño óptimo de muestra asignado a cada uno de ellos, posteriormente se calculan las varianzas de la media estratificada que son los valores que realmente interesa conocer ya que con ellos es posible realizar una comparación de ganancias conforme vayan aumentando los números de estratos. A manera de simplificación y para fines de una presentación visualmente mejor distribuida, se utilizarán las mismas 5 variables consideradas en el capítulo anterior ya que éstas podrían considerarse como representativas de la población de la que forman parte. Las tablas correspondientes a los resultados se presentan enseguida:

Tabla 5.13 Varianzas de la media estratificada para 3 estratos

CPI	ESCOLAR	S_MIN_PR	P_TODOS	P_5SMYM	POCU_CTO
Desv. Est.1	0.899511	0.523532	0.006641	0.027746	0.452461
Desv. Est.2	0.954773	0.364965	0.021223	0.036654	0.258618
Desv. Est.3	1.68672	0.751058	0.162119	0.148152	0.245537
$W_h^2(1-n_h/N_h)S_h^2/n_h$					
	0.000573	0.000194	3.13E-08	5.45E-07	0.000145
	0.003024	0.000442	1.49E-06	4.46E-06	0.000222
	0.004478	0.000888	4.14E-05	3.45E-05	9.49E-05
V(\bar{Y}_{est})	0.008075	0.001524	4.29E-05	3.95E-05	0.000462

Fuente: Elaboración propia

Tabla 5.14 Varianzas de la media estratificada para 4 estratos

CPI	ESCOLAR	S_MIN_PR	P_TODOS	P_5SMYM	POCU_CTO
Desv. Est.1	0.842082	0.556521	0.00442	0.03004	0.484203
Desv. Est.2	0.828866	0.389906	0.011319	0.027807	0.299384
Desv. Est.3	1.115263	0.396109	0.036866	0.051043	0.211868
Desv. Est.4	1.455454	0.64426	0.172588	0.137849	0.196542
$W_h^2(1-n_h/N_h)S_h^2/n_h$					
	0.000193	8.43E-05	5.32E-09	2.46E-07	6.38E-05
	0.001244	0.000275	2.32E-07	1.4E-06	0.000162
	0.002944	0.000371	3.22E-06	6.17E-06	0.000106
	0.001719	0.000337	2.42E-05	1.54E-05	3.13E-05
V(\tilde{Y}_{est})	0.006099	0.001068	2.76E-05	2.32E-05	0.000364

Fuente: Elaboración propia

Tabla 5.15 Varianzas de la media estratificada para 5 estratos

CPI	ESCOLAR	S_MIN_PR	P_TODOS	P_5SMYM	POCU_CTO
Desv. Est.1	0.851715	0.569863	0.003706	0.033063	0.487851
Desv. Est.2	0.775208	0.420301	0.008633	0.025136	0.328545
Desv. Est.3	0.835729	0.326984	0.01554	0.031293	0.206131
Desv. Est.4	1.109247	0.422875	0.0464	0.060725	0.187253
Desv. Est.5	1.262651	0.566364	0.167507	0.1229	0.163432
$W_h^2(1-n_h/N_h)S_h^2/n_h$					
	0.000114	5.1E-05	2.16E-09	1.72E-07	3.74E-05
	0.000642	0.000189	7.96E-08	6.75E-07	0.000115
	0.00149	0.000228	5.15E-07	2.09E-06	9.06E-05
	0.002113	0.000307	3.7E-06	6.33E-06	6.02E-05
	0.00091	0.000183	1.6E-05	8.62E-06	1.52E-05
V(\tilde{Y}_{est})	0.005268	0.000958	2.03E-05	1.79E-05	0.000319

Fuente: Elaboración propia

Tabla 5.16 Varianzas de la media estratificada para 6 estratos

CPI	ESCOLAR	S_MIN_PR	P_TODOS	P_5SMYM	POCU_CTO
Desv. Est.1	0.769725	0.562707	0.00329	0.023123	0.493735
Desv. Est.2	0.757515	0.467706	0.007167	0.027936	0.354136
Desv. Est.3	0.781056	0.338452	0.011951	0.027306	0.23734
Desv. Est.4	0.910274	0.323275	0.022331	0.036469	0.187912
Desv. Est.5	0.997697	0.390618	0.047838	0.059473	0.158758
Desv. Est.6	1.262651	0.566364	0.167507	0.1229	0.163432
$W_h^2(1-n_h/N_h)S_h^2/n_h$					
	5.24E-05	2.8E-05	9.57E-10	4.73E-08	2.16E-05
	0.000355	0.000136	3.18E-08	4.83E-07	7.77E-05
	0.000913	0.000171	2.14E-07	1.12E-06	8.43E-05
	0.001455	0.000184	8.76E-07	2.34E-06	6.2E-05
	0.001301	0.000199	2.99E-06	4.62E-06	3.3E-05
	0.000749	0.000151	1.32E-05	7.1E-06	1.25E-05
V(\tilde{Y}_{est})	0.004827	0.000869	1.73E-05	1.57E-05	0.000291

Fuente: Elaboración propia

En las tablas anteriores se muestra el proceso para el cálculo de las varianzas de manera individual para los distintos números de estrato, sin embargo, para una mejor comprensión y observación de ganancias por estrato, los datos a comparar se agrupan la tabla 5.17.

IV.3 Comparación de Varianzas

En la siguiente tabla el primer renglón de cifras corresponde la varianza de la media que se obtendría al realizar muestreo aleatorio simple a la población, y los renglones subsecuentes presentan la varianza de la media asociada al número de estratos manejado. En el apéndice D se encuentran los valores poblacionales de la media, varianza y desviación estándar para cada variable. Es posible observar como el utilizar muestreo aleatorio simple acarreará una varianza mayor que al utilizar muestreo estratificado, además conforme aumenta la cantidad de estratos, esta varianza sigue mejorándose aunque en menor medida cada vez.

Tabla 5.17 Varianzas de la media estratificada
por número de Estratos

ESTRATOS	ESCOLAR	S_MIN_PR	P_TODOS	P_5SMYM	POCU_CTO
V(\bar{Y})	0.023045	0.004106	8.26E-05	9.3E-05	0.001447
3	0.008075	0.001524	4.29E-05	3.95E-05	0.000462
4	0.006099	0.001068	2.76E-05	2.32E-05	0.000364
5	0.005268	0.000958	2.03E-05	1.79E-05	0.000319
6	0.004827	0.000869	1.73E-05	1.57E-05	0.000291

Fuente: Elaboración propia

De manera gráfica, para una mejor visualización de la tabla anterior:

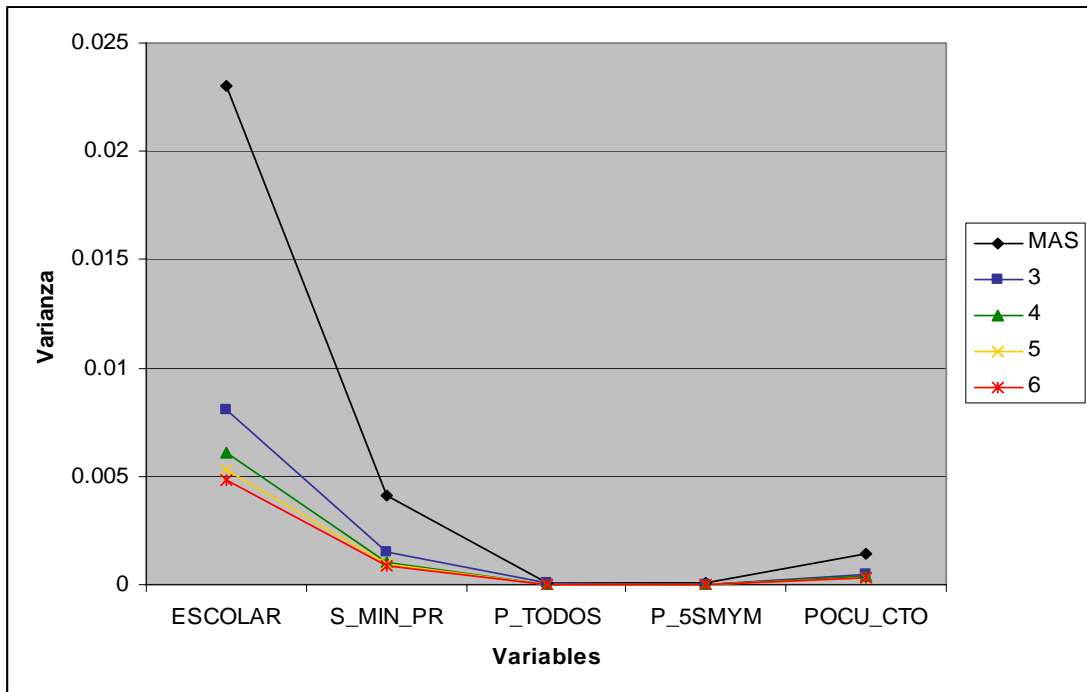


Figura 5.1 (a) Varianzas correspondientes al número de estratos
Fuente: Elaboración propia

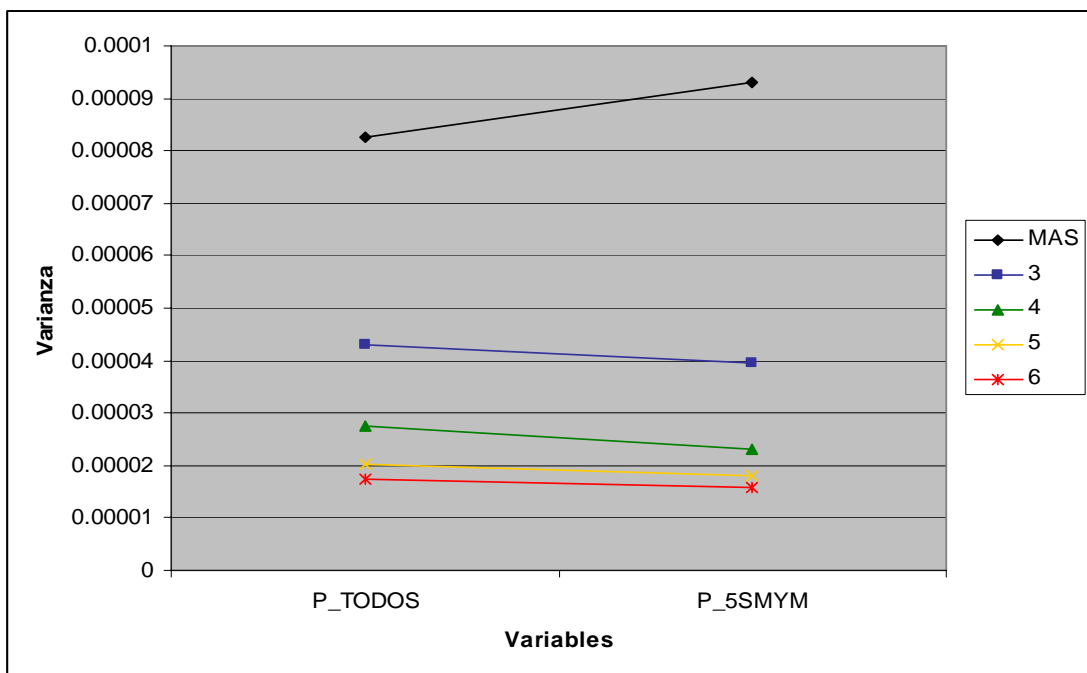


Figura 5.1 (b) Varianzas correspondientes al número de estratos
Fuente: Elaboración propia

La proporción de varianza cubierta por el muestreo estratificado con respecto al muestreo aleatorio simple, permite comprender en términos más generales la ganancia que se obtiene al preferir el primer método frente al segundo, ya que es más viable detectar un punto en el cual se esta llegando a cierta estabilidad.

Tabla 5.18 Proporción Varianza (Estr.) vs Varianza (MAS)

$V(\tilde{Y}_{est})/V(\tilde{Y})$	ESCOLAR	S_MIN_PR	P_TODOS	P_5SMYM	POCU_CTO
3	0.350406	0.371143	0.519189	0.425246	0.319254
4	0.264672	0.260038	0.334309	0.24976	0.251443
5	0.228585	0.233261	0.245771	0.192314	0.220322
6	0.209437	0.211559	0.209341	0.168837	0.201211

Fuente: Elaboración propia

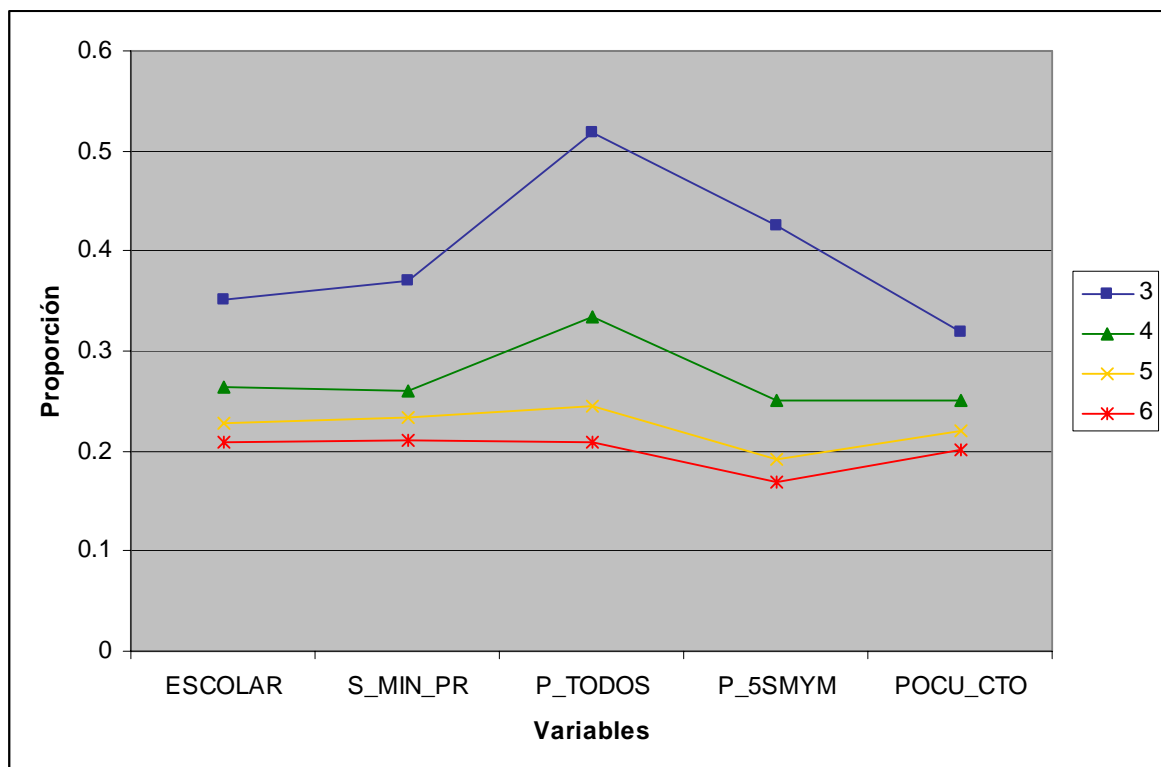


Figura 5.2 Proporción Varianza (Estr.) vs. Varianza (MAS)

Fuente: Elaboración propia

Al examinar la tabla y grafica anteriores, se observa como en un principio se tienen logros considerables cuando se incrementan los estratos, sin embargo, estos logros comienzan a ser cada vez menos evidentes, ya que en la gráfica la diferencia entre las líneas es cada vez menor, lo que permite intuir que el aumentar a 7 estratos las varianzas no estarían tan alejadas entre ellas y las correspondientes a la cantidad de estratos inmediata anterior. Este punto es de importancia ya que al realizar un aumento de estratos se puede incurrir en gastos de planeación, de preparación de encuestadores, etc.; es necesario llegar a un acuerdo entre un número de estratos conveniente que genere un costo aceptable que no rebasé un presupuesto inicialmente considerado.