Cubo Matemática Educacional
Vol. 4, Nº1, MAYO 2002

CONOCIENDO LA DISTRIBUCIÓN HIPERGEOMÉTRICA[1]

Eugenio Saavedra G.
Departamento de Matemática y C.C.,
Universidad de Santiago de Chile,
Casilla 307, Correo 2 Santiago, Chile,
E-mail: keno@fermat.usach.cl

Resumen: Este artículo muestra, a través de un ejemplo simple y cotidiano, la deducción del modelo matemático para la distribución hipergeométrica.

1 Introducción

Un curso de 2° Medio tiene 25 alumnos, de los cuales 10 son mujeres y 15 son hombres. Se realiza un experimento que consiste en elegir (al azar) 5 alumnos, para representar al curso en un acto del Colegio. Para efectuar la elección, se numeran 25 pelotitas desde 1 hasta 25, y se echan a una bolsa no transparente. Se escogen cinco pelotitas desde la bolsa (pueden sacarse las cinco al mismo tiempo o sacar una primero, luego la segunda, sin devolver la primera, después la tercera, sin devolver las anteriores, etc). Esta última modalidad es la que adoptamos. El número escogido, corresponde al alumno que lleva ese número en la lista del curso. Diremos que ocurrió un éxito si el número escogido corresponde a una alumna mujer,

[1] Parcialmente financiado por Cátedra Presidencial en Análisis Estocástico y por DICYT N° 049833SG .

en caso contrario, es decir, si el número escogido corresponde a un alumno hombre, diremos que ocurrió un fracaso. Se desea responder preguntas tales como: ¿cuál es la probabilidad de elegir tres alumnas mujeres y dos alumnos hombres?, o ¿cuál es la probabilidad de elegir sólo alumnas mujeres?, o ¿cuál es la probabilidad de elegir a lo menos un alumno hombre?.

2 Modelo Matemático Asociado al Experimento

En primera instancia, conozcamos la lista del curso.

Lista de Curso

1. Cisterna, Antonia	10. Lagos, Andrés	19. Toro, Francisco
2. Contreras, José	11. Meléndez, Gladys	20. Torres, Pamela
3. Diez, Mario	12. Méndez, Carlos	21. Valdés, Julio
4. Duarte, Juana	13. Morales, Andrea	22. Vallejos, Vanessa
5. Fuentes, Erwin	14. Muñoz, Carlos	23. Villarroel, Gonzalo
6. Gálvez, Claudio	15. Navarrete, Víctor	24. Zamora, Miguel
7. Gorostiza, María	16. Orellana, Natacha	25. Zamorano, Patricia
8. González, Rosario	17. Páez, Enzo	
9. Gutiérrez, Fabián	18. Romero, Daniel	

Según la lista de curso, las alumnas mujeres corresponden a los números 1, 4, 7, 8, 11, 13, 16, 20, 22, 25 y los alumnos hombres a los números 2, 3, 5, 6, 9, 10, 12, 14, 15, 17, 18, 19, 21, 23 y 24. Por ejemplo, si en la primera pelotita sale el 8, en la segunda 21, en la tercera 5, en la cuarta 24 y en la última pelotita sale 2, significa que hemos elegido a los alumnos Rosario, Julio, Erwin, Miguel y José. Si en cambio las pelotitas elegidas fuesen

primera pelotita: 17	segunda pelotita: 4	tercera pelotita: 2
cuarta pelotita: 10	quinta pelotita: 22	

entonces, los alumnos escogidos serían: Enzo, Juana, José, Andrés y Vanessa.

Notar que en el primer caso se obtuvieron cuatro hombres (4 fracasos) y una mujer (1 éxito), mientras que en el segundo caso fueron tres hombres (3 fracasos) y dos mujeres (2 éxitos).

Si repetimos 7 veces el experimento, es decir, sacar 5 pelotitas desde la bolsa con 25, entonces un posible resultado sería:
Una vez salieron cero mujeres y cinco hombres.
Dos veces salieron una mujer y cuatro hombres.
Dos veces salieron dos mujeres y tres hombres.
Una vez salieron tres mujeres y dos hombres.
Una vez salieron cuatro mujeres y un hombre.
Ninguna vez salieron cinco mujeres y cero hombres.

	Pelotita						
Repetición	1^a	2^a	3^a	4^a	5^a	total mujeres (n^o total de éxitos)	total hombres (n^o total de fracasos)
1^a	2	7	1	9	19	2	3
2^a	15	9	22	7	16	3	2
3^a	11	2	3	13	6	2	3
4^a	8	12	21	3	9	1	4
5^a	21	12	14	2	17	0	5
6^a	8	20	22	5	13	4	1
7^a	24	15	11	9	17	1	4

Resumiendo

nº de veces que se sacaron 5 pelotitas de las 25	nº de veces que salieron 0 mujer y 5 hombres	nº de veces que salieron 1 mujer y 4 hombres	nº de veces que salieron 2 mujeres y 3 hombres	nº de veces que salieron 3 mujeres y 2 hombres	nº de veces que salieron 4 mujeres y 1 hombre	nº de veces que salieron 5 mujeres y 0 hombres
7	1	2	2	1	1	0

Podríamos decir, entonces, que el número de éxitos que obtenemos al sacar 5 pelotitas desde la bolsa con 25 depende del azar. ¿Existirá algún patrón que siga la proporción de éxitos que resultan después de realizar el experimento 10 veces, 50 veces, 300 veces, o un "gran número"de veces?. Esta proporción (que dependerá del azar) se modelará matemáticamente, por este motivo relizaremos "muchas repeticiones del experimento", es decir, sacar 5 pelotitas desde la bolsa con 25. Como repetir el experimento , por ejemplo, 1000 veces es un poco lento y engorroso, simularemos en el computador la extracción de las 5 pelotitas desde la bolsa con 25. La tabla siguiente muestra los resultados obtenidos en una simulación.
Columna uno: indica el número de veces que se repite el experimento.

Columna dos: indica el número de veces que salieron cero mujeres (0 éxito) y cinco hombres (5 fracasos), al repetir el experimento la cantidad de veces que indica la columna uno.

Columna tres: indica el número de veces que salieron una mujer (1 éxito) y cuatro hombres (4 fracasos), al repetir el experimento la cantidad de veces que indica la columna uno.

Columna cuatro: indica el número de veces que salieron dos mujeres (2 éxitos) y tres hombres (3 fracasos), al repetir el experimento la cantidad de veces que indica la columna uno.

Columna cinco: indica el número de veces que salieron tres mujeres (3 éxitos) y dos hombres (2 fracasos), al repetir el experimento la cantidad de veces que indica la columna uno.

Columna seis: indica el número de veces que salieron cuatro mujeres (4 éxitos) y un hombre (1 fracaso), al repetir el experimento la cantidad de veces que indica la columna uno.

Columna siete: indica el número de veces que salieron cinco mujeres (5 éxitos) y cero hombres (0 fracaso), al repetir el experimento la cantidad de veces que indica la columna uno.

Columna ocho: indica la función, columna 2 dividida por columna 1, que anotamos f_0, y que llamamos frecuencia relativa de cero éxitos.

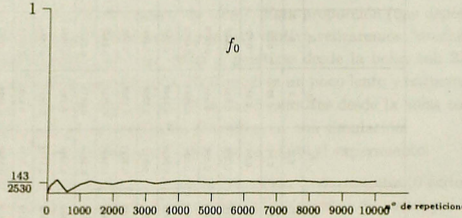
Columna nueve: indica la función, columna 3 dividida por columna 1, que anotamos f_1, se llama frecuencia relativa de un éxito.

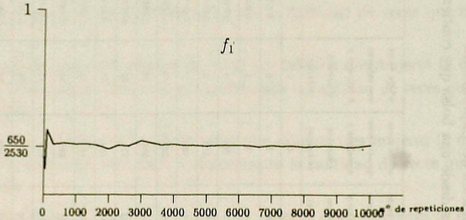
Análogamente, se definen las columnas diez, once y doce.

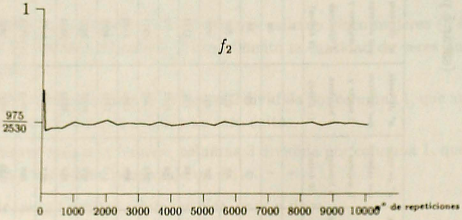
Simulación del Experimento
(sacar 5 pelotitas desde una bolsa que contiene 25)

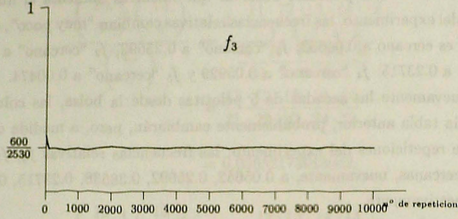
n° de veces que se repite el experimento	n° de veces que salieron 0 mujeres y 5 hombres (0 éxito)	n° de veces que salieron 1 mujer y 4 hombres (1 éxito)	n° de veces que salieron 2 mujeres y 3 hombres (2 éxitos)	n° de veces que salieron 3 mujeres y 2 hombres (3 éxitos)	n° de veces que salieron 4 mujeres y 1 hombre (4 éxitos)	n° de veces que salieron 5 mujeres y 0 hombre (5 éxitos)	f_0	f_1	f_2	f_3	f_4	f_5
10	0	2	5	3	0	0	0	0.2	0.5	0.3	0	0
50	1	7	28	13	1	0	0.02	0.14	0.56	0.26	0.02	0
100	4	35	34	23	3	1	0.04	0.35	0.34	0.23	0.03	0.01
1000	46	271	388	233	58	4	0.046	0.271	0.388	0.233	0.058	0.004
2000	94	491	796	496	115	8	0.047	0.2455	0.398	0.248	0.0575	0.004
3000	183	779	1143	704	182	9	0.061	0.2597	0.381	0.2347	0.0607	0.003
4000	258	1083	1479	938	232	10	0.0645	0.2708	0.3698	0.2345	0.058	0.0025
5000	283	1300	1914	1199	280	24	0.0566	0.26	0.328	0.2398	0.056	0.0048
6000	375	1589	2263	1421	330	22	0.0625	0.2648	0.3772	0.2368	0.055	0.0037
7000	413	1844	2689	1649	382	23	0.059	0.2634	0.3841	0.2356	0.0546	0.0033
8000	479	2042	3077	1898	461	43	0.0599	0.2553	0.3846	0.2373	0.0576	0.0054
9000	522	2339	3524	2097	481	37	0.058	0.2599	0.3916	0.233	0.0534	0.0041
10000	611	2668	3775	2341	554	51	0.0611	0.2668	0.3775	0.2341	0.0554	0.0051
15000	886	3868	5765	3495	938	48	0.0591	0.2579	0.3843	0.233	0.0625	0.0032
20000	1189	5237	7690	4596	1204	84	0.0595	0.2619	0.3845	0.2298	0.0602	0.0042
30000	1729	7918	11437	7089	1688	139	0.0576	0.2639	0.3812	0.2363	0.0563	0.0046

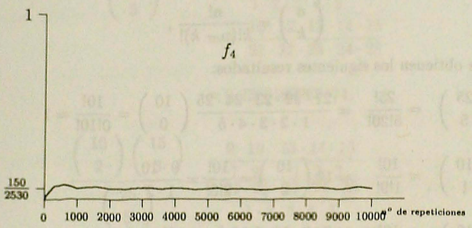
Los gráficos siguientes muestran el número de veces que se repite el experimento versus f_0 (respectivamente, f_1, f_2, f_3, f_4 y f_5).

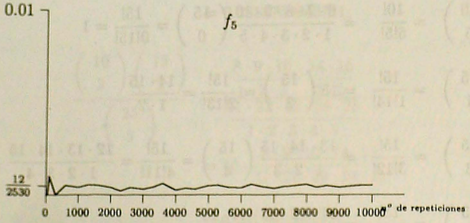







1
f_3
$\frac{600}{2530}$
0 1000 2000 3000 4000 5000 6000 7000 8000 9000 10000
n° de repeticiones


1
f_4
$\frac{150}{2530}$
0 1000 2000 3000 4000 5000 6000 7000 8000 9000 10000
n° de repeticiones


0.01
f_5
$\frac{12}{2530}$
0 1000 2000 3000 4000 5000 6000 7000 8000 9000 10000
n° de repeticiones

De los gráficos anteriores podemos observar que mientras aumenta el número de repeticiones del experimento, las frecuencias relativas cambian "muy poco", obteniéndose que: f_0 es cercano a 0.05652; f_1 "cercano" a 0.25692; f_2 "cercano" a 0.38538; f_3 "cercano" a 0.23715; f_4 "cercano" a 0.05929 y f_5 "cercano" a 0.00474. Además, al simular nuevamente las sacadas de 5 pelotitas desde la bolsa, las columnas 2, 3, 4 y 5, de la tabla anterior, probablemente cambiarán, pero, a medida que crece el número de repeticiones del experimento; las frecuencias relativas f_0, f_1, f_2, f_3, f_4 y f_5 son cercanas, nuevamente, a 0.05652, 0.25692, 0.38538, 0.23715, 0.05929 y 0.00474, respectivamente.

Por otra parte, el coeficiente binomial (el cuál se deduce intuitivamente en [3]) se define, para todo $n \geq 1$ y todo $0 \leq k \leq n$, por

$$\binom{n}{k} = \frac{n!}{k!(n-k)!},$$

de donde se obtienen los siguientes resultados:

$$\binom{25}{5} = \frac{25!}{5!20!} = \frac{21 \cdot 22 \cdot 23 \cdot 24 \cdot 25}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} \quad \binom{10}{0} = \frac{10!}{0!10!} = 1$$

$$\binom{10}{1} = \frac{10!}{1!9!} = 10 \quad \binom{10}{2} = \frac{10!}{2!8!} = \frac{9 \cdot 10}{1 \cdot 2}$$

$$\binom{10}{3} = \frac{10!}{3!7!} = \frac{8 \cdot 9 \cdot 10}{1 \cdot 2 \cdot 3} \quad \binom{10}{4} = \frac{10!}{4!6!} = \frac{7 \cdot 8 \cdot 9 \cdot 10}{1 \cdot 2 \cdot 3 \cdot 4}$$

$$\binom{10}{5} = \frac{10!}{5!5!} = \frac{6 \cdot 7 \cdot 8 \cdot 9 \cdot 10}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} \quad \binom{15}{0} = \frac{15!}{0!15!} = 1$$

$$\binom{15}{1} = \frac{15!}{1!14!} = 15 \quad \binom{15}{2} = \frac{15!}{2!13!} = \frac{14 \cdot 15}{1 \cdot 2}$$

$$\binom{15}{3} = \frac{15!}{3!12!} = \frac{13 \cdot 14 \cdot 15}{1 \cdot 2 \cdot 3} \quad \binom{15}{4} = \frac{15!}{4!11!} = \frac{12 \cdot 13 \cdot 14 \cdot 15}{1 \cdot 2 \cdot 3 \cdot 4}$$

$$\binom{15}{5} = \frac{15!}{5!10!} = \frac{11 \cdot 12 \cdot 13 \cdot 14 \cdot 15}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5}$$

y por tanto,

$$\frac{\binom{10}{0}\binom{15}{5}}{\binom{25}{5}} = \frac{1 \cdot \frac{11 \cdot 12 \cdot 13 \cdot 14 \cdot 15}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5}}{\frac{21 \cdot 22 \cdot 23 \cdot 24 \cdot 25}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5}} = \frac{11 \cdot 12 \cdot 13 \cdot 14 \cdot 15}{21 \cdot 22 \cdot 23 \cdot 24 \cdot 25} = 0.05652173913043,$$

$$\frac{\binom{10}{1}\binom{15}{4}}{\binom{25}{5}} = \frac{10 \cdot \frac{12 \cdot 13 \cdot 14 \cdot 15}{1 \cdot 2 \cdot 3 \cdot 4}}{\frac{21 \cdot 22 \cdot 23 \cdot 24 \cdot 25}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5}} = 5 \cdot \frac{10 \cdot 12 \cdot 13 \cdot 14 \cdot 15}{21 \cdot 22 \cdot 23 \cdot 24 \cdot 25} = 0.2569169960474,$$

$$\frac{\binom{10}{2}\binom{15}{3}}{\binom{25}{5}} = \frac{\frac{9 \cdot 10}{1 \cdot 2} \cdot \frac{13 \cdot 14 \cdot 15}{1 \cdot 2 \cdot 3}}{\frac{21 \cdot 22 \cdot 23 \cdot 24 \cdot 25}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5}} = 10 \cdot \frac{9 \cdot 10 \cdot 13 \cdot 14 \cdot 15}{21 \cdot 22 \cdot 23 \cdot 24 \cdot 25} = 0.3853754940711,$$

$$\frac{\binom{10}{3}\binom{15}{2}}{\binom{25}{5}} = \frac{\frac{8 \cdot 9 \cdot 10}{1 \cdot 2 \cdot 3} \cdot \frac{14 \cdot 15}{1 \cdot 2}}{\frac{21 \cdot 22 \cdot 23 \cdot 24 \cdot 25}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5}} = 10 \cdot \frac{8 \cdot 9 \cdot 10 \cdot 14 \cdot 15}{21 \cdot 22 \cdot 23 \cdot 24 \cdot 25} = 0.2371541501976,$$

$$\frac{\binom{10}{4}\binom{15}{1}}{\binom{25}{5}} = \frac{\frac{7\cdot 8\cdot 9\cdot 10}{1\cdot 2\cdot 3\cdot 4}\cdot 15}{\frac{21\cdot 22\cdot 23\cdot 24\cdot 25}{1\cdot 2\cdot 3\cdot 4\cdot 5}}$$

$$= 5\cdot\frac{7\cdot 8\cdot 9\cdot 10\cdot 15}{21\cdot 22\cdot 23\cdot 24\cdot 25}$$

$$= 0.05928853754941,$$

$$\frac{\binom{10}{5}\binom{15}{0}}{\binom{25}{5}} = \frac{\frac{6\cdot 7\cdot 8\cdot 9\cdot 10}{1\cdot 2\cdot 3\cdot 4\cdot 5}}{\frac{21\cdot 22\cdot 23\cdot 24\cdot 25}{1\cdot 2\cdot 3\cdot 4\cdot 5}}$$

$$= \frac{6\cdot 7\cdot 8\cdot 9\cdot 10}{21\cdot 22\cdot 23\cdot 24\cdot 25}$$

$$= 0.004743083003953.$$

Estos resultados motivan el siguiente modelo matemático.

Si denotamos por p_0 a la probabilidad de que ninguna alumna mujer salga al escoger a 5 alumnos desde los 25 del curso (probabilidad de cero éxito), entonces definimos:

$$p_0 = \frac{\binom{10}{0}\binom{15}{5}}{\binom{25}{5}}.$$

Se puede verificar, después de un poco de trabajo, que el numerador de p_0 corresponde al número de maneras en que pueden escogerse 5 alumnos hombres del total de 15 hombres del curso, y el denominador resulta ser el número de maneras en que 5 alumnos pueden escogerse del total de 25 del curso (sin distinguir sexo necesariamente).

Si ahora llamamos p_1 a la probabilidad de que salga una alumna mujer y 4 alumnos hombres, al escoger 5 alumnos de los 25 del curso (probabilidad de un éxito y cuatro fracasos), entonces definimos:

$$p_1 = \frac{\binom{10}{1}\binom{15}{4}}{\binom{25}{5}}.$$

En este caso, también se puede verificar que el numerador de $\mathbf{p_1}$ corresponde al número de maneras en que puede escogerse a una alumna mujer, del total de 10 mujeres del curso y 4 alumnos hombres del total de 15 alumnos de este sexo.

Finalmente, llamando $\mathbf{p_k}$, $k = 2, 3, 4, 5$, a la probabilidad de que salgan k alumnas mujeres y $5 - k$ alumnos hombres, al escoger 5 alumnos desde los 25 del curso (probabilidad de k éxitos y $5 - k$ fracasos), entonces definimos:

$$p_k = \frac{\binom{10}{k}\binom{15}{5-k}}{\binom{25}{5}},$$

para $k \in \{2, 3, 4, 5\}$, o sea

$$p_2 = \frac{\binom{10}{2}\binom{15}{3}}{\binom{25}{5}},$$

$$p_3 = \frac{\binom{10}{3}\binom{15}{2}}{\binom{25}{5}},$$

$$p_4 = \frac{\binom{10}{4}\binom{15}{1}}{\binom{25}{5}},$$

$$p_5 = \frac{\binom{10}{5}\binom{15}{0}}{\binom{25}{5}}.$$

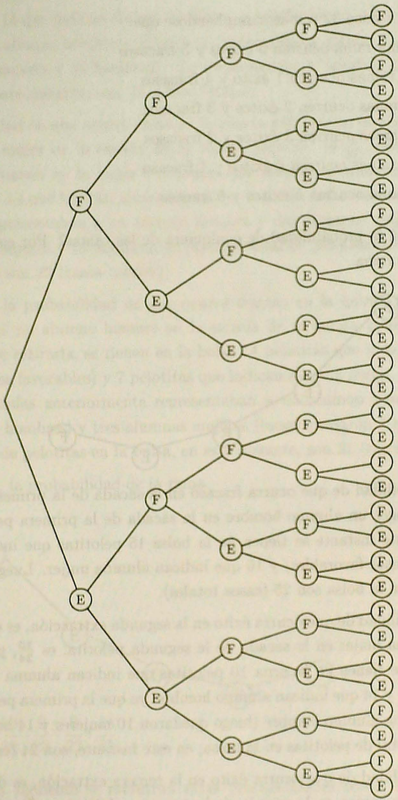
Notar que $p_0 + p_1 + p_2 + p_3 + p_4 + p_5 = 1$.

También aquí, es posible verificar que el numerador de $\mathbf{p_k}$ corresponde al número de maneras en que pueden escogerse k alumnas mujeres del total de 10 mujeres del curso y $5 - k$ alumnos hombres del total de 15 hombres del curso.

Las probabilidades antes definidas también podrían escribirse como:

$$p_k = \frac{\begin{pmatrix} n^o \text{ de mujeres del curso} \\ \\ n^o \text{ de mujeres escogidas} \end{pmatrix}\begin{pmatrix} n^o \text{ de hombres del curso} \\ \\ n^o \text{ de hombres escogidos} \end{pmatrix}}{\begin{pmatrix} n^o \text{ de alumnos del curso} \\ \\ n^o \text{ de alumnos escogidos} \end{pmatrix}}$$

El resultado de escoger 5 alumnos desde el total de 25 del curso, puede verse como en el siguiente árbol. El símbolo F significa que ocurrió fracaso y el símbolo E que ocurrió éxito.


F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E
F
E

Este árbol tiene 32 ramas, cumpliéndose que:

Sólo en una rama ocurren 0 éxitos y 5 fracasos

En cinco ramas ocurren 1 éxito y 4 fracasos

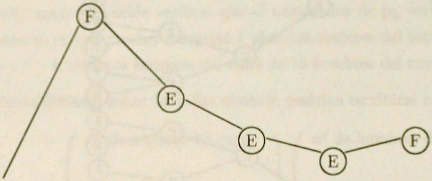
En diez ramas ocurren 2 éxitos y 3 fracasos

En diez ramas ocurren 3 éxitos y 2 fracasos

En cinco ramas ocurren 4 éxitos y 1 fracaso

En una rama ocurren 5 éxitos y 0 fracaso

¿Cuál será la probabilidad de cualquiera de las ramas? Por ejemplo, consideremos la rama



La probabilidad de que ocurra fracaso en la sacada de la primera pelotita, es decir, escoger un alumno hombre en la sacada de la primera pelotita, es $\frac{15}{25}$, pues en este instante se tienen en la bolsa 15 pelotitas que indican alumno hombre (casos favorables) y 10 que indican alumna mujer. Luego, el total de pelotitas en la bolsa son 25 (casos totales).

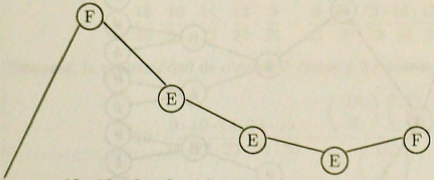
La probabilidad de que ocurra éxito en la segunda extracción, es decir, escoger una alumna mujer en la sacada de la segunda pelotita, es $\frac{10}{24}$, pues antes de retirarla, se tienen en la urna 10 pelotitas que indican alumna mujer (casos favorables) y 14 que indican alumno hombre, ya que la primera pelotita sacada representaba alumno hombre (luego quedaron 10 mujeres y 14 hombres). Por tanto, el total de pelotitas en la bolsa, en este instante, son 24 (casos totales).

La probabilidad de que ocurra éxito en la tercera extración, es decir, escoger una alumna mujer en la sacada de la tercera pelotita, es $\frac{9}{23}$, pues antes de retirarla, se tiene en la bolsa 9 pelotitas que indican alumna mujer (casos favorables) y 14 que indican alumno hombre, ya que la primera pelotita sacada representaba alumno hombre y la segunda representaba alumna mujer (luego quedaron 9 mujeres y 14 hombres). En consecuencia, el total de pelotitas en la bolsa, en este instante, son 23 (casos totales).

La probabilidad de que ocurra éxito en la cuarta extracción, es decir, escoger una alumna mujer en la sacada de la cuarta pelotita, es $\frac{8}{22}$, pues antes de retirarla, se tienen en la bolsa 8 pelotitas que indican alumna mujer (casos favorables) y 14 que indican alumno hombre. Las tres pelotitas sacadas anteriormente representaban a un alumno hombre y dos alumnas mujeres, luego quedaron 8 mujeres y 14 hombres, es decir, el total de pelotitas en la urna en este instante son 22 (casos totales).

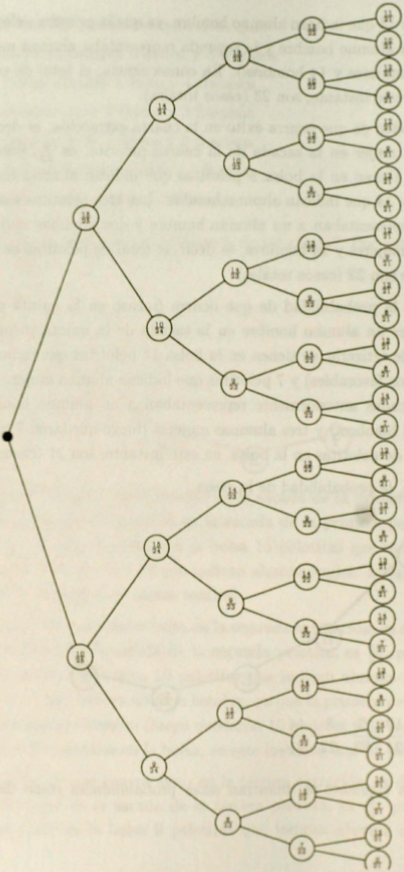
Finalmente, la probabilidad de que ocurra fracaso en la quinta pelotita, es decir, escoger un alumno hombre en la sacada de la quinta pelotita, es $\frac{14}{21}$, pues antes de retirarla, se tienen en la bolsa 14 pelotitas que indican alumno hombre (casos favorables) y 7 pelotitas que indican alumna mujer. Las cuatro pelotitas sacadas anteriormente representaban a un alumno hombre (luego quedaron 14 hombres) y tres alumnas mujeres (luego quedaron 7 mujeres). O sea, el total de pelotitas en la bolsa, en este instante, son 21 (casos totales).

Por lo tanto, la probabilidad de la rama.



es $\frac{15}{25}\cdot\frac{10}{24}\cdot\frac{9}{23}\cdot\frac{8}{22}\cdot\frac{14}{21}.$

En la figura siguiente se muestran estas probabilidades como diagrama de árbol.



Sólo en una rama ocurren 0 éxitos y 5 fracasos, y esta rama tiene probabilidad

$$\frac{15}{25}\cdot\frac{14}{24}\cdot\frac{13}{23}\cdot\frac{12}{22}\cdot\frac{11}{21} = \frac{11\cdot 12\cdot 13\cdot 14\cdot 15}{21\cdot 22\cdot 23\cdot 24\cdot 25}$$

$$= \frac{\binom{10}{0}\binom{15}{5}}{\binom{25}{5}}.$$

En cinco ramas ocurren 1 éxito y 4 fracasos y cada rama tiene la misma probabilidad, ésta es,

$$\frac{15}{25}\cdot\frac{14}{24}\cdot\frac{13}{23}\cdot\frac{12}{22}\cdot\frac{10}{21} = \frac{10\cdot 12\cdot 13\cdot 14\cdot 15}{21\cdot 22\cdot 23\cdot 24\cdot 25}.$$

Entonces, la probabilidad de obtener 1 éxito y 4 fracasos será

$$5\cdot\frac{10\cdot 12\cdot 13\cdot 14\cdot 15}{21\cdot 22\cdot 23\cdot 24\cdot 25} = \frac{\binom{10}{1}\binom{15}{4}}{\binom{25}{5}}.$$

En diez ramas ocurren 2 éxitos y 3 fracasos y cada rama tiene la misma probabilidad, ésta es,

$$\frac{15}{25}\cdot\frac{10}{24}\cdot\frac{14}{23}\cdot\frac{13}{22}\cdot\frac{9}{21} = \frac{9\cdot 10\cdot 13\cdot 14\cdot 15}{21\cdot 22\cdot 23\cdot 24\cdot 25}.$$

Entonces, la probabilidad de obtener 2 éxitos y 3 fracasos será

$$10\cdot\frac{9\cdot 10\cdot 13\cdot 14\cdot 15}{21\cdot 22\cdot 23\cdot 24\cdot 25} = \frac{\binom{10}{2}\binom{15}{3}}{\binom{25}{5}}.$$

En diez ramas ocurren 3 éxitos y 2 fracasos y cada rama tiene la misma probabilidad, ésta es,

$$\frac{15}{25}\cdot\frac{10}{24}\cdot\frac{14}{23}\cdot\frac{9}{22}\cdot\frac{8}{21} = \frac{8\cdot 9\cdot 10\cdot 14\cdot 15}{21\cdot 22\cdot 23\cdot 24\cdot 25}.$$

Entonces, la probabilidad de obtener 3 éxitos y 2 fracasos será

$$10 \cdot \frac{8 \cdot 9 \cdot 10 \cdot 14 \cdot 15}{21 \cdot 22 \cdot 23 \cdot 24 \cdot 25} = \frac{\binom{10}{3}\binom{15}{2}}{\binom{25}{5}}.$$

En cinco ramas ocurren 4 éxitos y 1 fracaso y cada rama tiene la misma probabilidad, ésta es,

$$\frac{10}{25} \cdot \frac{9}{24} \cdot \frac{15}{23} \cdot \frac{8}{22} \cdot \frac{7}{21} = \frac{7 \cdot 8 \cdot 9 \cdot 10 \cdot 15}{21 \cdot 22 \cdot 23 \cdot 24 \cdot 25}.$$

Entonces, la probabilidad de obtener 4 éxitos y 1 fracaso será

$$5 \cdot \frac{7 \cdot 8 \cdot 9 \cdot 10 \cdot 15}{21 \cdot 22 \cdot 23 \cdot 24 \cdot 25} = \frac{\binom{10}{4}\binom{15}{1}}{\binom{25}{5}}$$

Finalmente, sólo en una rama ocurren 5 éxitos y 0 fracaso, y esta rama tiene probabilidad

$$\frac{10}{25} \cdot \frac{9}{24} \cdot \frac{8}{23} \cdot \frac{7}{22} \cdot \frac{6}{21} = \frac{6 \cdot 7 \cdot 8 \cdot 9 \cdot 10}{21 \cdot 22 \cdot 23 \cdot 24 \cdot 25} = \frac{\binom{10}{5}\binom{15}{0}}{\binom{25}{5}}.$$

Observación: Si modificásemos el experimento y permitiéramos que la pelotita extraida de la bolsa se repusiera en ella, entonces, la probabilidad de tener éxito en cualquiera de las 5 extracciones de la pelotita sería la misma, $\frac{10}{25} = \frac{2}{5}$, puesto que en todo instante la bolsa tendría 10 pelotitas que indican alumna mujer y 15 que indican alumno hombre. Por lo tanto, el experimento modificado sigue un esquema Bernoulli con parámetros $(5, \frac{2}{5})$. Así, la probabilidad de obtener k éxitos, con $k \in \{0, 1, 2, 3, 4, 5\}$, sería:

$$b\left(5,\frac{2}{5};0\right) = 0.07776 \qquad b\left(5,\frac{2}{5};3\right) = 0.2304$$

$$b\left(5,\frac{2}{5};1\right) = 0.2592 \qquad b\left(5,\frac{2}{5};4\right) = 0.0768$$

$$b\left(5,\frac{2}{5};2\right) = 0.3456 \qquad b\left(5,\frac{2}{5};5\right) = 0.01024$$

El Esquema Bernoulli y sus implicaciones pueden consultarse en [1], [2] o [3]. Como en el experimento primario la probabilidad de obtener k éxitos era:

$$p_0 = 0.05652173913043 \qquad p_3 = 0.2371541501976$$

$$p_1 = 0.2569169960474 \qquad p_4 = 0.05928853754941$$

$$p_2 = 0.3853754910711 \qquad p_5 = 0.004743083003953,$$

entonces, la probabilidad de obtener k éxitos cuando se extraen 5 pelotitas de una bolsa con 25, y la pelotita no se regresa a la bolsa, es "parecida", pero no igual, a la probabilidad que se obtiene cuando la pelotita se regresa a la bolsa.

3 El Modelo General

La elección, al azar, de 5 alumnos en un curso de 25 y el hecho de considerar como éxito la salida de una mujer, es una manera simple de introducir la llamada distribución hipergeométrica. Se tienen N objetos cualquiera, los cuales pueden clasificarse en dos tipos, digamos tipo I y tipo II. Se sabe que hay M objetos del tipo I y por tanto $N - M$ del tipo II. Se retiran, sin reposición, r objetos al azar (se dice que se toma una muestra de tamaño r). Suponiendo que $M < N$, $r \leq M$ y llamando p_k a la probabilidad de obtener exactamente k objetos del tipo I (éxitos) y $r - k$ del tipo II (fracasos) en la muestra, se define,

$$p_k = \frac{\binom{M}{k}\binom{N-M}{r-k}}{\binom{N}{r}}$$

$$= \frac{\begin{pmatrix} \text{n}^o \text{ de objetos tipo I} \\ \text{n}^o \text{ de objetos tipo I en la muestra} \end{pmatrix} \begin{pmatrix} \text{n}^o \text{ de objetos tipo II} \\ \text{n}^o \text{ de objetos tipo II en la muestra} \end{pmatrix}}{\begin{pmatrix} \text{n}^o \text{ total de objetos} \\ \text{tamaño de la muestra} \end{pmatrix}}$$

A este modelo se le conoce como distribución hipergeométrica de parámetros (N, M, r).

Ejemplos:

1. Se sabe, por experiencia, que un paquete de semillas contiene 80% de semillas buenas y 20% de semillas malas. Si un paquete contiene 20 semillas y se eligen al azar 3 semillas del paquete.

   a) ¿Cuál es la probabilidad de que las 3 semillas sean malas?

   b) ¿Cuál es la probabilidad de que al menos una semilla sea mala?

   c) ¿Cuál es la probabilidad de que a lo más un tercio de las semillas escogidas sean malas?

Solución:

Número de objetos (semillas)	$N = 20$
Número de objetos tipo I (semillas malas)	$M = 4$
Número de objetos tipo II (semillas buenas)	$N - M = 16$
Tamaño de la muestra	$r = 3$

a) La probabilidad que desea calcular es p_3.

$$p_3 = \frac{\binom{4}{3}\binom{16}{0}}{\binom{20}{3}} = \frac{4}{\frac{18\cdot 19\cdot 20}{1\cdot 2\cdot 3}} = \frac{1}{285}.$$

b) La probabilidad buscada es $p_1 + p_2 + p_3$, como

$$p_0 + p_1 + p_2 + p_3 = 1,$$

entonces,

$$p_1 + p_2 + p_3 = 1 - p_0.$$

Pero,

$$p_0 = \frac{\binom{4}{0}\binom{16}{3}}{\binom{20}{3}} = \frac{\frac{14\cdot 15\cdot 16}{1\cdot 2\cdot 3}}{\frac{18\cdot 19\cdot 20}{1\cdot 2\cdot 3}} = \frac{28}{57},$$

o sea,

$$p_1 + p_2 + p_3 = 1 - \frac{28}{57} = \frac{29}{57}.$$

c) Como un tercio de 3 es 1, entonces se pide calcular $p_0 + p_1$. Pero,

$$p_1 = \frac{\binom{4}{1}\binom{16}{2}}{\binom{20}{3}} = \frac{4\cdot\frac{15\cdot 16}{1\cdot 2}}{\frac{18\cdot 19\cdot 20}{1\cdot 2\cdot 3}} = \frac{8}{19},$$

o sea,

$$p_0 + p_1 = \frac{28}{57} + \frac{8}{19} = \frac{52}{57}.$$

2. Consideremos un lote de artículos de alguna especie. Supongamos que un artículo cualquiera puede ser puesto a prueba y como resultado, clasificado como bien ajustado a las especificaciones que debe satisfacer el artículo (no defectuoso) o bien, que no se ajusta a éstas (defectuoso). Se toma una muestra

al azar, desde el lote, y se testea cada artículo en la muestra; entonces, para un cierto nivel de tolerancia, el lote es aceptado si el número de artículos defectuosos en la muestra es menor o igual a este nivel, y se rechaza el lote si el número de artículos defectuosos lo excede. Se tiene un lote de 100 artículos, de los cuales 15 son defectuosos (pero no se sabe cuales 15). Una muestra de tamaño 10 es seleccionada aleatoriamente del lote. Se desea encontrar el nivel de tolerancia, de modo que, con probabilidad superior a 0.8 aceptemos el lote.
Solución:

Número de objetos (artículos)	$N = 100$
Número de objetos tipo I (artículos defectuosos)	$M = 15$
Número de objetos tipo II (artículos no defectuosos)	$N - M = 85$
Tamaño de la muestra	$r = 10.$

Sea c el nivel de tolerancia para el lote; entonces, se pide encontrar el menor valor c de modo que la probabilidad de que el número de artículos en la muestra sea menor o igual a c, supere 0.8. Es decir, hallar c tal que $p_0+p_1+...+p_c \geq 0.8$.

Como,

$$p_0 = \frac{\binom{15}{0}\binom{85}{10}}{\binom{100}{10}} = 0.1808,$$

$$p_1 = \frac{\binom{15}{1}\binom{85}{9}}{\binom{100}{10}} = 0.3568,$$

$$p_2 = \frac{\binom{15}{2}\binom{85}{8}}{\binom{100}{10}} = 0.2919,$$

entonces,

$$p_0 + p_1 + p_2 = 0.8295.$$

Por lo tanto, el nivel de tolerancia es $c = 2$.

Referencias

[1] **González, J.**, *Bases Matemáticas de la Estadística* , Revista del Profesor de Matemáticas, Sociedad de Matemática de Chile , número especial, 45–61, 1999.

[2] **Iglesias, P. and Saavedra, E.**, *Probabilidad y Estadística Elementales,* Apunte Facultad de Matemáticas, Universidad Católica de Chile, (1997) 161 págs.

[3] **Saavedra, E.**, *Descubriendo Distribuciones de Probabilidades*, Monografía (por aparecer).