01

LA ENSEÑANZA

DE LAS FUNCIONES DE DISTRIBUCIÓN DE PROBABILIDAD. EXPERIENCIAS PRÁCTICAS

ISSN: 2631-2662

Fecha de presentación: enero, 2018 **Fecha de aceptación: marzo, 2018** **Fecha de publicación: mayo, 2018**

LA ENSEÑANZA

DE LAS FUNCIONES DE DISTRIBUCIÓN DE PROBABILIDAD. EXPERIENCIAS PRÁCTICAS

THE TEACHING OF PROBABILITY DISTRIBUTION FUNCTIONS. PRACTICAL EXPERIENCES

Ana Lilia Castillo Coto[1]
E-mail: alcastillo@umet.edu.ec
Miguel Ángel Fernández Marín[2]
E-mail: mafernandez@umet.edu.ec
Fidel Ángel Gutiérrez Rodríguez[2]
E-mail: fagutierrez@umet.edu.ec
[1] Convenio Universidad de Cienfuegos-Universidad Metropolitana del Ecuador.
[2] Universidad Metropolitana del Ecuador

Cita sugerida (APA, sexta edición)

Castillo Coto, A. L., Fernández Marín, M. Á., & Gutiérrez Rodríguez, F. (2018). La enseñanza de las funciones de distribución de probabilidad. Experiencias prácticas. *Revista Metropolitana de Ciencias Aplicadas*, 1(2), 5-14. Recuperado de http://remca.umet.edu.ec/index.php/REMCA

RESUMEN

El presente artículo tiene como objetivo mostrar las posibles maneras de usar paquetes de programas como apoyo para el desarrollo de las habilidades de análisis estadístico. Aborda la experiencia de los autores en la enseñanza de la estadística a estudiantes de diversas carreras, con el apoyo de paquetes de programas estadísticos. Para el presente trabajo se usa la herramienta IBM SPSS Statistics Base. Con mayor o menor intensidad, el conocimiento de estadística es necesario en todas las profesiones, como sustento de la investigación científica y el desempeño empírico cotidiano. La enseñanza de la estadística está por ello presente en los planes de estudio de la mayoría de las carreras, sin embargo, no resulta ser de las asignaturas más gustadas por los estudiantes.

Palabras clave: Funciones, distribución de probabilidad, análisis estadístico, variables discretas, programas estadísticos.

ABSTRACT

In the paper the authors aim to show the possible ways to use some software packages as a support in the teaching processes for the development of statistical analysis skills in the students. It addresses the experience of the authors by teaching Statistics in several study programs, with the support of IBM SPSS Statistics Base. The knowledge of Statistics is necessary in all professions, as well as support of scientific research or for the daily empirical performance. The teaching of Statistics is therefore present in the curricula of all professions; however is not one of the most liked subjects by the students.

Keywords: Functions, probability distribution, statistical analysis, discrete variables, statistical programs.

INTRODUCCIÓN

Los autores han considerado necesario abordar un pequeño preámbulo sobre los principales términos más usados para explicar las variables aleatorias como función –específicamente las distribuciones binomial, binomial negativa, hipergeométrica y de Poisson- que son contenido de la Estadística Inferencial que se enseñan en las carreras de Ingeniería en Sistemas de Información, Gestión Empresarial y Gestión de Empresas Turísticas e Industrias de la Recreación.

Adicionalmente, se emiten algunos criterios acerca de los procederes para la enseñanza de estos contenidos.

Un experimento es un proceso de intervención para obtener un resultado. En muchas ocasiones, el resultado depende del azar, por lo que predecirlo antes de aplicar el experimento tiene cierta incertidumbre que se representa por funciones de probabilidad.

Es necesario, entonces, evaluar todos los posibles resultados que un experimento pueda arrojar y las probabilidades de ocurrencia de cada uno de ellos. Así sería posible estimar cuál de todos es más probable que ocurra en ciertas condiciones.

Un experimento está formado por varios eventos o pruebas que son los que se evalúan como éxito o fracaso.

Un experimento se considera aleatorio cuando sus posibles resultados son conocidos previamente. Será el azar quien determine cuál resultado se obtendrá de cada prueba que se repita, siempre que todas se hagan en las mismas condiciones.

Partiendo de la conocida definición de muestra como el subconjunto seleccionado de una población que la representa dada una característica específica, los autores Guerra, Vargas y Walpole consideran un espacio muestral δ como el conjunto formado por todos los posibles resultados S_n con $n \in N^+$ de un experimento dado. Donde la función $f: \delta \to R$ Sdefine una variable aleatoria (Vargas Sabadías, 1995; Walpole, et al., 1999; Guerra Bustillo, et al., 2003).

De la mismo forma, para un espacio muestral dado δ de algún experimento, una **variable aleatoria** es cualquier regla que asocia un número con cada resultado en δ.

Una variable aleatoria discreta es aquella cuyos valores posibles bien constituyen un conjunto finito, o bien dichos valores pueden ser puestos en lista en una secuencia infinita en la cual existe un primer elemento, un segundo elemento, y así sucesivamente (según Devore, 2008, "contablemente" infinita).

Se define la variable aleatoria discreta con las letras mayúsculas del final del alfabeto latino. La notación $X(s)=x$ significa que x es el valor asociado cuando la variable aleatoria X tiene como resultado s.

DESARROLLO

Un experimento de Bernoulli es un experimento aleatorio donde pueden darse dos resultados: A (al que llamaremos éxito) o su contrario $\bar{A}$ (fracaso), de modo que la probabilidad de éxito es $P(A)=p$ y la probabilidad de fracaso es $P(\bar{A})=q=1-p$ (Vargas Sabadías, 1995).

Esta última definición es importante para caracterizar experimentos que sólo pueden tener dos resultados posibles, como el ejemplo típico de lanzar una moneda al aire y otros que pueden ser transformados a un experimento con dos resultados: éxito o fracaso.

Las probabilidades asignadas a varios resultados en δ determinan a su vez las probabilidades asociadas con los valores de cualquier variable aleatoria X particular. La distribución de probabilidad de X dice cómo está distribuida la probabilidad total de 1 entre los varios posibles valores de X. (Vargas Sabadías, 1995; Devore, 2008).

La función masa de probabilidad o distribución de probabilidad de una variable aleatoria discreta se define como como $p(x)=P(X=x)=P(\{s \in \delta : X(s)=x\}), \forall x \in \mathbb{R}$, lo que significa calcular la probabilidad de todas las $s \in \delta, X(s)=x$.

La función de distribución acumulativa de una variable aleatoria discreta X con función masa de probabilidad $p(x)$ se define como: $F(x)=P(X \leq x)=\sum_{y:y \leq x} p(y)$, lo que significa que $\forall x \in \mathbb{R}, F(x)$ es la probabilidad de que el valor observado de X como mucho sea x.

En este artículo se resumen las formas en que se ha abordado la enseñanza de las variables aleatorias como función, considerando para ello las distribuciones binomial, binomial negativa, hipergeométrica y de Poisson.

Como se está considerando la definición de experimento de Bernoulli, habrá que tener en cuenta que este proceso cumple con las siguientes condiciones iniciales:

1. La realización de pruebas sucesivas independientes.
2. Los resultados de cada una de las pruebas pueden ser calificados como éxito o fracaso.
3. La probabilidad de éxito permanece constante para todas las pruebas.

La variable aleatoria que representa al número de éxitos en n pruebas consecutivas independientes, recibe el nombre de variable aleatoria binomial, y la distribución de probabilidad de esta variable aleatoria discreta es llamada distribución binomial.

Se representa y se define por

$$b(x; n, p) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, x = 0, 1, 2, \ldots, n.$$

La distribución de probabilidad acumulada binomial se representa y define como:

$$B(x; n, p) = P(X \leq x) = \sum_{y=0}^{x} b(y; n, p)$$

Una de las razones que argumentan los estudiantes para rechazar la estadística es lo complejo y difícil de comprender del contenido, y lo tedioso y reiterativo de los cálculos. Esta situación se constituye en un reto para los profesores, que tienen que encontrar la manera de conseguir que los estudiantes comprendan los contenidos, y sean capaces de aplicarlos a nuevas situaciones de su vida profesional.

Para lograrlo existen varias condiciones que se complementan entre sí:

1. El aseguramiento del nivel de partida de los contenidos.
2. La explicación estricta y consciente del modelo matemático a seguir de cada una de las distribuciones abordadas en clase.
3. El uso de ejemplos prácticos pertinentes a los programas de estudio.
4. El uso de paquetes de programas estadísticos para agilizar los cálculos en clase.

Los profesores deben asegurarse del conocimiento previo del auditorio de estudiantes, sobre cada uno de los precedentes necesarios para el aprendizaje de los contenidos referidos a distribuciones aleatorias discretas. Si fuera preciso, puede dedicarse una o más clases a ello. Es preferible invertir tiempo en una actividad docente recordatoria de cada uno de los conocimientos precedentes necesarios, que interrumpir constantemente una explicación, para aclarar las dudas emanadas del no dominio u olvido de dichos conocimientos por parte del auditorio.

De manera general, se reconocen como competencias previas al aprendizaje de las funciones aleatorias discretas, las siguientes:

- Teoría de conjuntos y técnicas de conteo.
- Funciones algebraicas y álgebra de funciones.
- Series y operaciones con series.
- Límites y continuidad.
- Derivadas y reglas de derivación.
- Métodos para calcular máximos y mínimos.
- Uso de graficadores y análisis de datos en hojas de cálculo.
- Entre otras.

Cuando el trabajo inter-, intra- y transdisciplinario de los profesores del programa sea fuerte y sistemático el diagnóstico del dominio grupal de los conocimientos precedentes recae en el propio colectivo de profesores, ya sea del semestre o de las disciplinas. Quedaría solo a discreción del profesor el diagnóstico de las individualidades.

Cuando no es bueno el trabajo metodológico de colectivos de semestre o de disciplina, en cualquiera de sus modalidades, es el profesor de la asignatura quien deberá asegurar el conocimiento previo de los contenidos precedentes para el aprendizaje de los nuevos. Es muy recomendable que se realice al menos una actividad docente recordatoria de los conocimientos precedentes al aprendizaje de funciones aleatorias discretas.

El pleno dominio por el profesor de los contenidos que se explican, es una condición obligatoria para el éxito en el proceso de enseñanza de las estadísticas en general, y de las funciones aleatorias discretas en particular. Este sería el momento para explicar un ejemplo práctico de aplicación del modelo específico que se enseña y cuyas características se abordarán más adelante.

Primeramente, se explicará cada uno de los componentes del modelo de cada función de distribución aleatoria discreta. Este nivel de detalle es quien garantiza posteriormente la pertinente interpretación de los resultados de su aplicación a diversos ejemplos prácticos y la futura experiencia empírica del estudiante en el ejercicio de su profesión.

A continuación, el profesor deberá abordar las condiciones iniciales para la aplicación de cada modelo de distribución aleatoria discreta, sustentando su explicación en que dichas condiciones emanan de la generalización de experiencias empíricas y no lo contrario.

La intencionada comparación entre las diferentes distribuciones contenido de estudio, le facilitará al estudiante definir el tipo de distribución a aplicar en cada ejemplo o problema específico. Ello garantizará que sea capaz de reconocer qué tipo de distribución sería la adecuada de aplicar en cada caso práctico futuro; y que su atención, durante el proceso de aprendizaje, se concentre más en la interpretación de los resultados y no en la aplicación mecánica de un modelo o de un software para realizar un cálculo. Para lograrlo, el profesor puede valerse de un cuadro comparativo de distribuciones aleatorias discretas que se elabora de conjunto con los estudiantes durante el desarrollo de los contenidos referidos a este tema. El inicio de cada conferencia sería la revisión del cuadro confeccionado hasta el momento y la conclusión sería la elaboración conjunta de los nuevos contenidos abordados con la nueva distribución estudiada.

El uso de ejemplos adecuados a cada auditorio o grupo de estudiantes exige de un fuerte trabajo de preparación por parte del docente, que deberá encontrar ejemplos reales con información creíble y adecuada al futuro ámbito de desempeño de los estudiantes de su auditorio. No es concebible que a un grupo de alumnos de programas como Gestión Empresarial o Sistemas de Información se les expliquen ejemplos de medicina o pedagogía (por lo demás inusitadamente frecuentes en los libros de texto) en lugar de casos provenientes del ámbito empresarial, que sería su más probable ambiente laboral futuro.

Adicionalmente, no todos los grupos de estudiantes tienen las mismas características. Hay grupos con mayor o menor preparación previa en matemáticas, además de otros aspectos que pudieran contribuir en su actuar como son los horarios de clases, el ambiente de las aulas, la conformación del grupo, etc. Quiere decir, que el ejemplo que en un grupo haya dado resultado, tal vez en el próximo -en lugar de acelerar- ralentice la comprensión del contenido. Queda en las manos del profesor la elección del ejemplo y el método que utilice en la enseñanza.

Al estudiante se lo prepara para la profesión. Todo acercamiento a condiciones similares a las de su futuro desempeño empírico contribuye a una mejor formación. Los sistemas automatizados de gestión empresarial son cada vez más populares, el trabajo administrativo empresarial transfiere cada vez más las operaciones de cálculo a las computadoras, y puntualiza en los análisis de resultados para la toma de decisiones.

Garantizar habilidades de interpretación en los estudiantes, a partir del análisis de resultados de cálculos de softwares, propiciaría motivarles a su uso y lograr que estén familiarizados con las tecnologías de la informática y las comunicaciones, de manera que puedan resolver problemas de la práctica en su área específica del conocimiento. Adicionalmente, les prepara para los ejercicios de investigación científica que usualmente constituyen las evaluaciones integradoras finales.

Las herramientas informáticas en la enseñanza de las matemáticas y las estadísticas se usan para lograr que los estudiantes complementen su formación en la relación modelo matemático y software específico. Con ello, el profesor logra –por una parte- concentrar más la atención del auditorio de estudiantes en el análisis e interpretación de resultados y -por otra- elevar la complejidad de los ejemplos a utilizar pues, generalmente, los softwares tienen facilidades para graficar y con ello se aumentan las facilidades para profundizar los análisis.

Una buena táctica didáctica sería la utilización de ejemplos sencillos usando el método demostrativo y el cálculo conjunto de ejemplos más complejos con ayuda del software.

Al enseñar la distribución binomial, el profesor deberá abordar las condiciones iniciales para su aplicación. Para ello es recomendable que se comience presentando un ejemplo proveniente de la producción, en este caso editorial, como sigue:

> En una editorial se ha comprobado que uno de cada cinco ejemplares de las encuadernaciones de carátula dura, no pasa una prueba de resistencia. Se decide seleccionar al azar una muestra de 15 ejemplares de la producción de un turno de trabajo.
>
> Las preguntas a responder serían:

a. ¿Qué probabilidad habrá de que exactamente fallen 8 de los 15 seleccionados?

b. ¿Qué probabilidad habría de que no más de 8 de los 15 no pasen la prueba de resistencia?

Para explicarlos se seguirá un conjunto de pasos:

Paso No. 1. Identificación de los componentes del problema

- Experimento: que será la selección de 15 ejemplares encuadernados (muestra) de toda la producción de un turno de trabajo, para someterlos a una prueba de resistencia,
- El espacio muestral está constituido por todos los posibles resultados del experimento. Habrá que ejemplificar algunos de estos resultados, para la mejor comprensión del auditorio de estudiantes. Puede considerarse que un evento en el experimento tendrá éxito cuando se detecte un fallo y cuando la prueba muestre que la encuadernación es resistente se considera como un evento fallido.

  Tomando como ejemplo que fallen 8 eventos de los 15 que se evalúan, habrá una cantidad de resultados posibles en el espacio muestral que deben ser considerados.

  Lo primero es que el estudiante sepa calcular la cantidad de posibles resultados del experimento que se está evaluando como

  $\|\delta\| = \binom{15}{8} = C\ (15;8)\ =\ 6\,435$ resultados diferentes.

  Adicionalmente, el estudiante deberá reconocer la diferencia entre la cantidad de éxitos y fracasos y el orden en que ocurran.

  El profesor puede auxiliarse de una representación visual de dos probables resultados para 8 fallos, de manera que los estudiantes lo comprendan, como se muestra a continuación.

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

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

Figura 1. Diferencia entre dos muestras del espacio muestral δ.

- La variable aleatoria X es la cantidad de libros que no pasan la prueba de resistencia de encuadernación. Es recomendable que el profesor consiga que sean los estudiantes quienes la identifiquen como una variable aleatoria discreta.

Con los componentes identificados se puede trabajar a la par con un paquete de programas estadístico –en este caso se utiliza el SPSS. El paso de identificación de los componentes aparece en el SPSS como se muestra a continuación.

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida	Rol
1	x	Numérico	8	0	Cantidad de éxitos entre n eventos consecutivos.	Ninguna	Ninguna	8	Derecha	Desconocido	Entrada
2	n	Numérico	8	0	Cantidad de eventos del experimento a realizarse.	Ninguna	Ninguna	8	Derecha	Desconocido	Entrada
3	p	Numérico	8	2	La probabilidad de éxito de un evento.	Ninguna	Ninguna	8	Derecha	Desconocido	Entrada
4	b	Numérico	8	3	b(x;n,p): función de probabilidad binomial	Ninguna	Ninguna	8	Derecha	Desconocido	Entrada

Figura 2. Identificación de los componentes del problema. Vista de variables.

Paso No. 2. Identificación del tipo de distribución de probabilidad para la variable discreta que se ajuste

De acuerdo a lo explicado anteriormente, es este el momento de justificar el uso de la distribución binomial como un proceso de Bernoulli y convertir las condiciones iniciales que ya fueron explicadas teóricamente a este caso particular, a saber:

1. Se realizan 15 eventos sucesivos (n=15),
2. Constituyen eventos independientes, que significa que la probabilidad de ocurrencia de uno no influye la del siguiente,
3. Cada prueba solo tiene dos posibles resultados:

   A= no pasan la prueba de resistencia de encuadernación (éxito) o

   $\bar{A}$ = pasan las pruebas de encuadernación (fracaso).
4. Todas las pruebas tienen la misma probabilidad de que su resultado sea exitoso.

Como ya está identificado y debidamente formulado en el SPSS toda la estructuración del caso de estudio que se ha tomado como ejemplo, es posible facilitarle los datos específicos al paquete de programas SPSS para dar respuesta al inciso a) ¿Qué probabilidad habrá de que exactamente fallen 8 de las 15 seleccionadas?, según se muestra en la siguiente figura.

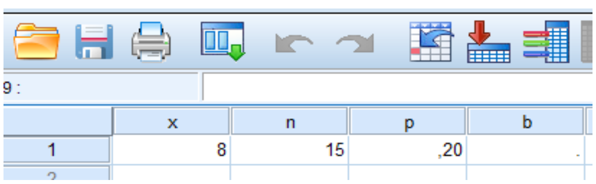

Figura 3. Identificación del tipo de distribución de probabilidad para la variable discreta que se ajuste. Vista de datos.

La variable aleatoria discreta tiene una distribución binomial y se representa como X~Bin(n,p), donde n=15 es la cantidad de experimentos que se van a realizar, p=0,2 es la probabilidad de que cada experimento obtenga como resultado un éxito.

El planteamiento en SPSS se muestra en la figura que sigue.

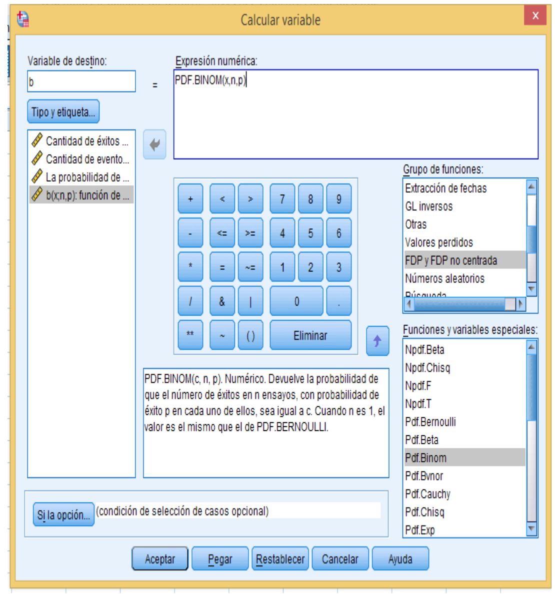

Figura 4. Identificación del tipo de distribución de probabilidad para la variable discreta que se ajuste. Expresión numérica.

Paso No. 3. Planteamiento del modelo matemático correspondiente

La primera pregunta que se plantea para el caso de estudio es: ¿Qué probabilidad habrá de que exactamente fallen 8 de las 15 seleccionadas?

Esto quiere decir que se necesita hallar la probabilidad de que de 15 eventos realizados 8 no pasen la prueba de resistencia, por tanto se quiere hallar la siguiente probabilidad:

$$P(X=8)= P(X\leq 8)-P(X\leq 7)=B(8;15{,}0.2)-B(7;15{,}0.2)= b(8;15{,}0.2)$$

Esta expresión significa que la probabilidad de que x=8 se puede calcular tanto por la función de distribución de probabilidad acumulada, como por la función de distribución de probabilidad, siendo ambas identidades.

Para demostrar a los estudiantes la veracidad del anterior planteamiento se explicará que:

$$\begin{aligned}&B(8;15{,}0.2)-B(7;15{,}0.2)=b(8;15{,}0.2)=\mathit{Miembro\ Izquierdo\ de\ la\ ecuación}\\&=B(8;15{,}0.2)-B(7;15{,}0.2)=\textstyle\sum_{y=0}^{8}b(y;15{,}0.2)-\sum_{y=0}^{7}b(y;15{,}0.2)=[b(0;15{,}0.2)+b(1;15{,}0.2)+\\&b(2;15{,}0.2)+b(3;15{,}0.2)+b(4;15{,}0.2)+b(5;15{,}0.2)+b(6;15{,}0.2)+b(7;15{,}0.2)+b(8;15{,}0.2)]-\\&[b(0;15{,}0.2)+b(1;15{,}0.2)+b(2;15{,}0.2)+b(3;15{,}0.2)+b(4;15{,}0.2)+b(5;15{,}0.2)+b(6;15{,}0.2)+\\&b(7;15{,}0.2)]=b(8;15{,}0.2)=\mathit{Miembro\ Derecho\ de\ la\ ecuación}\end{aligned}$$

Por lo que queda demostrado que el miembro izquierdo es igual que el miembro derecho.

Es sabido además que una distribución de probabilidad puede calcularse por una suma de distribución de probabilidades acumuladas.

Se calcularía entonces la expresión obtenida como:

$$b(8;15{,}0.2)=\binom{15}{8}0.2^{8}(1-0.2)^{15-8}=\frac{15!}{(15-8)!\,8!}0.2^{8}(0.8)^{7}$$

Ya planteado el modelo matemático para resolver la primera pregunta del ejemplo objeto de análisis el inciso a), se puede observar que los cálculos son un poco tediosos

Para poder calcular el resultado de planteamiento anterior están las funciones en el SPSS como se muestra a continuación:

$$B(8;15{,}0.2)-B(7;15{,}0.2)\sim CDF.BINOM(8{,}15{,}0.2)-CDF.BINOM(7{,}15{,}0.2)\approx 0{,}003$$
$$b(8;15{,}0.2)\sim PDF.BINOM(8{,}15{,}0.2)\approx 0{,}003$$

El mismo resultado en SPSS se muestra a continuación.

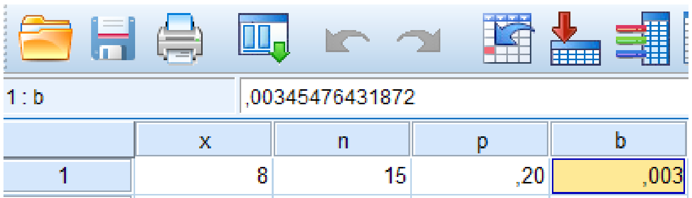

Figura 5. Planteamiento del modelo matemático correspondiente. Resultado final inciso a).

Esto se interpretaría como: de 15 muestras seleccionadas para la revisión, hay una probabilidad baja de 0,003, que exactamente 8 de ellas no pasen la prueba de resistencia de encuadernación.

De la misma manera se procedería para el inciso b) del ejercicio usado como ejemplo demostrativo.

$$(X\le 8)=B(8;15{,}0.2)=\sum_{x=0}^{8}b(x;15{,}0.2)=\sum_{x=0}^{8}\binom{15}{x}(0.2)^{x}(1-0.2)^{15-x}$$
$$=\binom{15}{0}(0.2)^{0}(1-0.2)^{15-0}+\binom{15}{1}(0.2)^{1}(1-0.2)^{15-1}+\binom{15}{2}(0.2)^{2}(1-0.2)^{15-2}+\cdots$$
$$+\binom{15}{8}(0.2)^{8}(1-0.2)^{15-8}$$

Y el resultado según el SPSS se muestra en la siguiente figura.

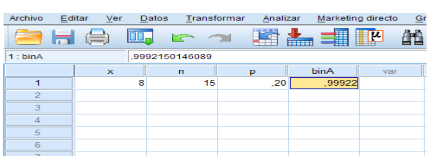

Figura 6. Planteamiento del modelo matemático correspondiente. Resultado final inciso b).

El resultado quiere decir que se halló la probabilidad de $0 \vee 1 \vee \ldots \vee 8$ de los ensayos hayan fracasado.

La función de distribución hipergeométrica se diferencia de la binomial en la forma de realizar el muestreo. En la función de distribución binomial el muestreo se realiza con reemplazo, que indica que hay independencia entre las muestras. En la función de distribución hipergeométrica, a cambio, se requiere independencia en los experimentos que se realizan, por lo que el muestreo será sin reemplazo.

Las suposiciones que conducen a una función de distribución hipergeométrica son:

1. La población o conjunto que se va a muestrear es finita, conocida y se compone de individuos,
2. Cada individuo puede ser caracterizado como éxito (E) o fracaso (F) y hay M éxitos en la población. Lo que indica que la estructura de la población es conocida.
3. Se selecciona una muestra de n individuos sin reemplazo, de tal modo que cada subconjunto de tamaño tiene igual probabilidad de ser seleccionado.
4. El tamaño de muestra n debe ser mayor que el 5% de la población N.
5. La variable aleatoria de interés X es el número de éxitos en la muestra.

Al considerar las definiciones sobre distribución de probabilidad y distribución de probabilidad acumulada hipergeométrica abordados por algunos autores consultados (Walpole, et al., 1999;Lind, Marchal & Mason, 2004; Devore, 2008), puede concluirse entonces, que:

6. Si es el número de éxitos (E) en una muestra completamente aleatoria de tamaño extraída de la población compuesta de M éxitos y $(N\text{-}M)$ fallas, entonces la distribución de probabilidad de llamada distribución hipergeométrica, es:

$$P(X=x)=h(x;n,M,N)=\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}},x\in\mathbb{Z}: Max(0,n-N+M)\le x\le \mathrm{Min}\,(n,M)$$

- La distribución de probabilidad acumulada hipergeométrica se representa y define como:

$$H(x;n,M,N)=P(X\le x)=\sum_{y=0}^{x}h(y;n,M,N)$$

Dando cumplimiento a los aspectos sugeridos anteriormente, el ejemplo práctico que se sugiere para la explicación de la función de distribución hipergeométrica sería una adecuación del utilizado en la explicación de la función de distribución binomial. Con ello se lograría que la atención del estudiante se concentre en el nuevo reto analítico que se plantee en una situación familiar. Sería:

> En la misma editorial se sabe que en un turno de trabajo se producen 25 ejemplares de un libro de texto de Administración. De ellos, 20 ejemplares pasan la prueba de resistencia y los restantes 5 (20%) suelen no pasarla. Se decide seleccionar al azar una muestra de 5 ejemplares de la producción de un turno de trabajo.
>
> Las preguntas a responder serían:

a. ¿Qué probabilidad habrá de que exactamente fallen 3 de los 5 seleccionados?

b. ¿Qué probabilidad habría de que no más de 3 no pasen la prueba de resistencia?

Como en el caso discutido anteriormente para la distribución binomial, se seguirá un conjunto de pasos:

Paso No. 1. Identificación de los componentes del problema

El profesor deberá lograr que sea el auditorio de estudiantes quien encuentre las diferencias entre el caso tratado con la distribución binomial y la nueva situación. Para ello recapitula la manera en que se identificaron los componentes del problema anterior para el nuevo caso de estudio, como sigue:

- Experimento: seleccionar 5 libros de la producción de un turno de trabajo (población conocida N=25) para aplicarles una prueba de resistencia,
- Resultado: cálculo de la probabilidad de existencia de una cantidad especificada de individuos de la muestra seleccionada (n), que cumplan alguna característica de la población. Equivale a la cantidad de éxitos en la muestra,
- Existe en la población M individuos con las características que se evalúan (cantidad de éxitos E=5),
- La variable aleatoria X es el número de éxitos en la muestra M.

Los componentes identificados se facilitan al paquete de programas estadístico SPSS, como se muestra a continuación.

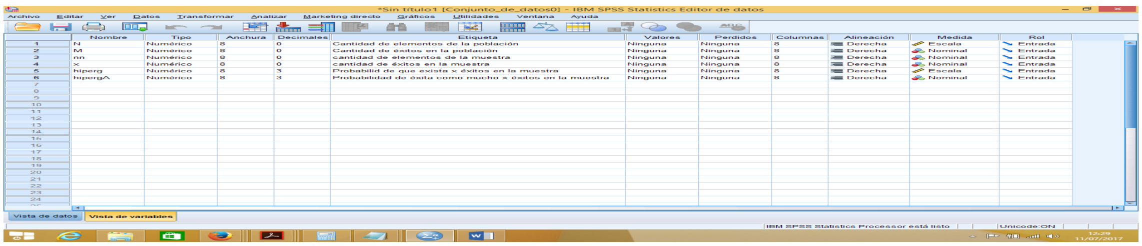

Figura 7. Identificación de los componentes del problema. Vista de variables.

Paso No. 2. Identificación del tipo de distribución de probabilidad para la variable discreta X que se ajuste

Este sería el momento de presentar la distribución hipergeométrica –variante de la binomial- como mejor aproximación a la situación planteada. Para ello se revisarían -de conjunto con el auditorio de estudiantes- las suposiciones que conducen a una función de distribución hipergeométrica adecuadas al ejemplo:

1. La población que se va a muestrear es N=25,
2. Cada individuo puede ser caracterizado como éxito (E) los que no pasan la prueba de resistencia, o fracaso (F) los que pasan la prueba de resistencia. Hay $M=5$ éxitos en la población,
3. Se selecciona una muestra de $n=5$ individuos,
4. El tamaño de muestra $n=5>0{,}05*25$,
5. La variable aleatoria X es la cantidad de libros que no pasan la prueba de resistencia en la muestra elegida.

Como ya está identificada y debidamente formulada en el SPSS toda la estructuración del caso de estudio que se ha tomado como ejemplo, es posible facilitarle los datos específicos al paquete de programas SPSS para dar respuesta a los dos incisos que se plantean según se muestra en la siguiente figura.

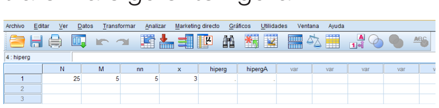

Figura 8. Identificación del tipo de distribución de probabilidad para la variable discreta que se ajuste. Vista de datos.

La variable aleatoria discreta X tiene una distribución hipergeométrica y se representa como $X\sim Hyp\ (n,M,N)$, donde $n=5$ es la muestra seleccionada, $M=5$ son los éxitos en la población y la cantidad de elementos de la población de estudio.

El planteamiento en SPSS se muestra en la figura que sigue.

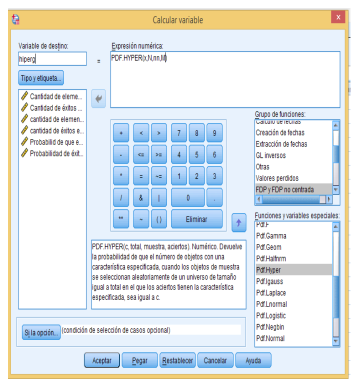

Figura 9. Identificación del tipo de distribución de probabilidad para la variable discreta X que se ajuste. Expresión numérica.

Paso No. 3. Planteamiento del modelo matemático correspondiente

La primera pregunta que se plantea para el caso de estudio es: ¿Qué probabilidad habrá de que exactamente fallen 3 de los 5 seleccionados?

Esto quiere decir que se necesita hallar la probabilidad de que 5 libros seleccionados, exactamente 3 no pasen la prueba de resistencia; por tanto, se quiere hallar la siguiente probabilidad:

$$P(X=3)=h(3;5{,}5{,}25)=\frac{\binom{5}{3}\binom{25-5}{5-3}}{\binom{25}{5}}=\frac{\frac{5!}{(5-3)!*3!}*\frac{20!}{(20-2)!*2!}}{\frac{25!}{(25-5)!*5!}}=\frac{\frac{5!*20!}{2!*3!*18!*2!}}{\frac{25!}{20!*5!}}$$

La segunda pregunta que se plantea para el caso de estudio es: ¿Qué probabilidad habría de que no más de 3 no pasen la prueba de resistencia?

Esto quiere decir que se necesita hallar la probabilidad de que 5 libros seleccionados, como máximo 3 no pasen la prueba de resistencia; por tanto, se quiere hallar la siguiente probabilidad:

$$P(X \leq 3) = P(X=0) + P(X=1) + P(X=2) + P(X=3) = H(3;5,5,25) = \sum_{x=0}^{3} h(x;5,5,25)$$

$$= \frac{\binom{5}{0}\binom{25-5}{5-0}}{\binom{25}{5}} + \frac{\binom{5}{1}\binom{25-5}{5-1}}{\binom{25}{5}} + \frac{\binom{5}{2}\binom{25-5}{5-2}}{\binom{25}{5}} + \frac{\binom{5}{3}\binom{25-5}{5-3}}{\binom{25}{5}}$$

$$= \frac{\frac{5!}{(5-0)!*0!} * \frac{20!}{(20-0)!*0!}}{\frac{25!}{(25-5)!*5!}} + \frac{\frac{5!}{(5-1)!*1!} * \frac{20!}{(20-1)!*1!}}{\frac{25!}{(25-5)!*5!}} + \frac{\frac{5!}{(5-2)!*2!} * \frac{20!}{(20-2)!*2!}}{\frac{25!}{(25-5)!*5!}}$$

$$+ \frac{\frac{5!}{(5-3)!*3!} * \frac{20!}{(20-3)!*3!}}{\frac{25!}{(25-5)!*5!}}$$

$$= \frac{\frac{5!}{5!*0!} * \frac{20!}{20!*0!} + \frac{5!}{4!*1!} * \frac{20!}{19!*1!} + \frac{5!}{3!*2!} * \frac{20!}{18!*2!} + \frac{5!}{2!*3!} * \frac{20!}{17!*3!}}{\frac{25!}{(25-5)!*5!}}$$

Una vez planteado el modelo matemático para resolver las preguntas del ejemplo se procede a realizar los cálculos con ayuda del SPSS, como se muestra a continuación.

Figura 10. Planteamiento del modelo matemático correspondiente. Resultado final.

Esto se interpretaría como:

- De 5 libros seleccionados para la revisión, hay una probabilidad de que exactamente 3 de ellas no pasen la prueba de resistencia de encuadernación.
- De 5 libros seleccionados para la revisión, hay una probabilidad de 0,998 de que cómo máximo 3 de ellos no pasen la prueba de resistencia de encuadernación.

El profesor debe insistir en que los estudiantes noten la diferencia entre los resultados obtenidos con anterioridad, lo que creará un precedente analítico para la elaboración de la tabla resumen.

De la misma manera se procedería para el resto de los incisos planteados en el ejercicio usado como ejemplo demostrativo.

En este trabajo se han abordado los casos de las distribuciones binomial e hipergeométrica, pero en realidad se enseñan otras como la binomial negativa, la Poisson. Incluso hay carreras en que la enseñanza de funciones como la de Weibull, la hiperexponencial y la hiperexponencial negativa son necesarias para el desempeño futuro de la profesión. Como se insistía anteriormente, es muy importante que los estudiantes identifiquen las características de cada una de estas funciones bajo la guía del profesor.

Para ello se sugiere que las conclusiones de cada clase -donde se enseña cada vez una nueva distribución- se hagan a través de la elaboración conjunta de un cuadro comparativo, como se muestra en la siguiente tabla.

Tabla 1. Tabla de comparación de las distribuciones.

Distribuciones	Condiciones iniciales para identificación de la distribución	Variable aleatoria discreta	Función de distribución probabilidad	Significado	Función de distribución de probabilidad acumulada	Significado

La introducción de cada clase, donde se aborde una nueva distribución, comenzará con la revisión de las funciones anteriormente estudiadas con ayuda de la propia tabla.

CONCLUSIONES

Mantener la motivación de los estudiantes depende de las destrezas del profesor y de su disposición a enseñar. El aseguramiento del nivel de partida de los contenidos, la explicación estricta y consciente del modelo matemático a seguir de cada una de las distribuciones abordadas en clase, el uso de ejemplos prácticos pertinentes a los programas de estudio, y el uso de paquetes de programas estadísticos para agilizar los cálculos en clase, permiten lograr la motivación de los estudiantes y elevar su aprovechamiento docente en el aprendizaje de la estadística. Por ello, los autores consideran pertinente puntualizar en los siguientes aspectos:

- La claridad con que se expliquen las definiciones previas a los contenidos objeto de estudio en cada clase, garantiza que el estudiante pueda entender los nuevos conceptos de experimento, variable aleatoria como función, espacio muestral y otros.
- La explicación correcta y detallada de experimento y proceso de Bernoulli garantiza la base del entendimiento de las funciones de distribución de probabilidad más utilizadas por los estudiantes en las asignaturas de Estadística de las carreras de Gestión Empresarial, Sistemas de Información y Gestión de Empresas Turísticas e Industrias de la Recreación.
- Un claro fundamento teórico y la formulación en pasos mediante un ejercicio demostrativo, posibilita que el estudiante conciba un algoritmo lógico que le permita afrontar exitosamente situaciones de su desempeño profesional.
- El uso de paquetes estadísticos -como puede ser el SPSS- constituye una eficaz herramienta de apoyo a los cálculos, y permite que en las clases se aproveche más el tiempo en el análisis e interpretación de los

resultados. Asimismo, posibilita aumentar la cantidad y complejidad de los ejercicios prácticos.

- La elaboración conjunta de una tabla resumen después de cada clase, donde se aborden las diferentes funciones de distribución de probabilidad, así como su recordatorio antes de cada clase, permite que el estudiante sintetice los conocimientos a través de la comparación de las diferentes distribuciones abordadas.

REFERENCIAS BIBLIOGRÁFICAS

Devore, J. L. (2008). ***Probabilidad y estadística para ingeniería y ciencias.*** México, D.F.: Cengage Learning Editores, S. A. de C.V.

Guerra Bustillo, C., et al. (2003). ***Estadística.*** La Habana: Félix Varela.

International Business Machines Corporation. (2017). *SPSS Statictic Base V.22*. Recuperado de http://www-03.ibm.com/software/products/es/spss-stats-base/

Lind, D. A., Marchal, W., & Mason, R. D. (2004). ***Estadística para administración y economía.*** México, D.F.: Alfaomega Grupo Editor, S.A. de C. V.

Vargas Sabadías, A. (1995). ***Estadística Descriptiva e Inferencial.*** Cuenca: Compobell, S. L.

Walpole, R. E., et al. (1999). ***Probabilidad y estadística para ingenieros***. México: Prentice-Hall Hispanoamericana, S. A.