

Trabajo publicado en www.ilustrados.com
La mayor Comunidad de difusión del conocimiento

TRABAJO DE INVESTIGACIÓN

***MOMENTOS NO CENTRADOS DE LA DISTRIBUCIÓN
HIPEGEOMETRICA***

Presentado por:

***Lic. Sandra Salazar Palomino
Br. Wilbert Colque Candia***

APURÍMAC – PERU

II.- INTRODUCCIÓN

En el presente trabajo nos dedicaremos al estudio de los Momentos no Centrados de la Distribución hipergeométrica.

De esta manera daremos un concepto matemático importante que tiene diversas aplicaciones en los modelos probabilísticos.

Con el propósito de presentar un desarrollo riguroso de este tema, se requerirán conocimientos, matemáticos de un nivel considerablemente mayor del que estemos suponiendo aquí pero que el transcurso del desarrollo del proyecto iremos tocando cada uno de estos términos.

En este trabajo de investigación desarrollaremos la relación que tienen cada uno de los momentos no centrados de la Distribución hipergeométrica que también se los conoce con el nombre de momentos potenciales.

Entre las Esperanzas matemáticas que son de especial importancia en la estadística están los momentos de una variable aleatoria que en este caso para nuestro estudio será la variable aleatoria de la Distribución hipergeométrica y veremos como mediante la función generadora de momentos existe relación con cada una de estos momentos. De esta manera el lector podrá preguntarse si este método es del todo útil y que servirá o no en la concepción de conocimientos en nuestra formación profesional.

III.- PLANEAMIENTO DEL PROBLEMA

Los momentos no centrados de la Distribución Hipergeométrica se han escogido en virtud a la gran importancia para el desarrollo de las características y propiedades de la Distribución hipergeométrica.

En tal sentido indagaremos si los momentos no centrados de esta distribución nos ayudaran en la verificación de estas propiedades y características, mediante la función generadora de momentos.

IV.- HIPÓTESIS

Utilizando la relación existente entre los $k+1$ primeros momentos centrados realizadas por Pearson probaremos que también existe relación entre los $k+1$ primeros momentos no centrados de la Distribución hipergeométrica.

V.- FORMULACIÓN DE OBJETOS

5.1. OBJETIVOS GENERALES

- Determinar la relación que existe entre los $k+1$ primeros momentos no centrados de la Distribución hipergeométrica.

5.2. OBJETIVOS ESPECÍFICOS

- Definir la función de probabilidad de la Distribución hipergeométrica.
- Definir la función Generatriz de Probabilidad de la distribución en estudio.
- Precisar los momentos no centrados (momentos respecto al origen)

VI.- JUSTIFICACIÓN

El presente trabajo de investigación, pretende contribuir con el avance de las ciencias estadísticas en virtud a la gran importancia para el desarrollo de las características y propiedad de la Distribución Hipergeométrica, además puede ser una herramienta muy poderosa para estudiar diversos aspectos de las diferentes distribuciones de probabilidades.

VII.- METODOLOGÍA

El método que utilizamos en este proyecto es el método Mixto Analítico y Deductivo.

VIII. LIMITACIONES DE LA INVESTIGACIÓN

En el proceso de investigación se han experimentado algunos obstáculos que dificultaron el normal avance del presente trabajo, tales como:

- La falta del material bibliográfico acerca del trabajo de investigación.
- Falta de experiencia en este tipo de investigaciones, que no nos permite actuar con pericia y eficacia de los experimentos por ser este el primer proyecto de investigación que lo realizo, sobre temas relacionado con la estadística.

IX MARCO TEÓRICO CONCEPTUAL

IX.1.ANTECEDENTES DE LA INVESTIGACIÓN

Se encontró los siguientes antecedentes: Huygens propone 5 problemas de gran importancia en el desarrollo de la teoría de la probabilidad. De Moivre, en 1711, resuelve el cuarto utilizando lo que hoy conocemos como función de probabilidad Hipergeométrica, siendo esta la primera referencia histórica que sobre la citada distribución aparece en la literatura probabilística.

Los momentos han sido estudiados, principalmente, por Karl Pearson y V. Romanovsky, en sucesivos artículos publicados entre 1899 y 1925.

En este último año Pearson encuentra la relación entre sus $k+1$ primeros momentos centrales.

A nivel local no se encontró una investigación parecida referente al tema que tratamos.

Juan Antonio García Ramos publica un artículo por Internet referente al tema es estudio.

IX.2.- CONTENIDO

Capítulo I: DEFINICIÓN DE TÉRMINOS BÁSICOS

- I.1. Población
- I.2. Muestra
- I.3. Elemento
- I.4. Experimento aleatorio
- I.5. Espacio Muestral
- I.6. Evento
- I.7. Probabilidad
- I.8. Ensayos sin reposición
- I.9. Ensayos con reposición
- I.10. Variable aleatoria (Discreta y continua)

Capítulo II: CONCEPTOS BÁSICOS

- II.1. La Distribución Hipergeométrica
- II.2. Valor Esperado
- II.3. Momento no centralizado
- II.4. Función Generatriz de Momentos
- II.5. Función Característica
- II.6. Serie de Taylor
- II.7. Propiedades de la función Generadora de Momentos.
- II.8. Serie Hipergeométrica

Capítulo III: RESUMEN DE LA INVESTIGACIÓN

CONCLUSIONES

X.- CRONOGRAMA

ACTIVIDADES	ENERO 2004	FEBRERO 2004	MARZO 2004
Desarrollo del capítulo I	X		
Desarrollo del capítulo II		X	
Desarrollo del caso práctico			X

XI.- BIBLIOGRAFÍA

- 1.- Probabilidad y Aplicaciones Estadísticas. por: Paul L. Meyer.
- 2.- Estadísticas I: Probabilidad y Estadísticas. por: Javier Martín Ruego y Luis Ruiz – Maya Pérez.
- 3.- Probabilidad y estadísticas Para ingeniería y ciencias por: Gabriel Velasco Sotomayor Marian Wisniewski.

CAPITULO I: DEFINICIÓN DE TÉRMINOS BÁSICOS

I.1.Población.- Conjunto de todos los individuos u objetos que poseen una característica común ejemplo: la población peruana, la población de estudiantes de la UNSAAC, los estudiantes de la carrera de Matemáticas.

I.2.Muestra.- Está constituida por una parte de los individuos que componen la población y que permite sacar resultados referenciales de la población a partir de su análisis.

Ejm. Los estudiantes del curso de Análisis de regresión, la carrera profesional de Matemáticas constituye una muestra.

I.3.Elemento.- Son las unidades para los cuales se desea obtener información. En un estudio puede haber más de un elemento.

I.4.Experimento Aleatorio.- No es algo que se haga necesariamente en la practica o en un laboratorio; al contrario casi siempre es imaginario.

Las características principales de un experimento aleatorio son las siguientes:

- a.- Se puede repetir las veces que uno quiere o está sujeto a repetición.
- b.- Su resultado final esta fuera de control, esto es, depende del azar.
- c.- Si embargo, se puede anticipar la totalidad de resultados posibles.
- d.- Se observa cierto patrón de regularidad a medida que aumentan las repeticiones.

I.5.Espacio Muestral.- es la totalidad de resultados posibles en un experimento aleatorio, es decir, el conjunto de todas las posibles ocurrencias.

Generalmente se denota por la letra Ω .

$$\Omega = \{w_1, w_2, w_3, \dots\}$$

I.6.Evento.- Es cualquiera de los conjuntos que comprenden el espacio muestral.

I.7.Probabilidad.- La probabilidad de un suceso (o evento) es un concepto abstracto que se usa para describir el grado de certidumbre o verosimilitud de ese evento, de acuerdo con las condiciones bajo las que se realiza algún experimento aleatorio.

I.8.Ensayos sin Reposición.- Suponga que se tiene una población de N objetos que son de sólo t tipos diferentes denotados por tipo 1, tipo 2, tipo 3, ..., tipo t . Suponga además, que se escoge de ahí una muestra de tamaño "n" sin reposición. Si en la población hay K_1 objetos del tipo 1, ..., K_t objetos hasta t entonces la probabilidad de que en la muestra haya exactamente K_1 objetos del tipo 1; K_2 objeto del tipo 2, ..., hasta X_t objetos del tipo t , viene dada por la expresión siguiente.

$$\frac{\binom{k_1}{x_1} \binom{k_2}{x_2} \dots \binom{k_t}{x_t}}{\binom{N}{n}}, \text{, donde } \sum_{i=1}^t x_i = n \quad \text{y} \quad \sum_{i=1}^t k_i = N$$

I.9.Ensayos con reposición (ENSAYOS DE BERNOULLI).- si un experimentado Aleatorio consiste de "n" ensayos repetidos independientes, de los cuales sólo hay dos tipos que son: Éxito (con probabilidad individual igual a p) y fracaso (con probabilidad individual a $q = 1 - p$), entonces la probabilidad de que en "n" ensayos se obtengan exactamente K éxitos y $n-k$ fracasos está dada por.

$$\binom{n}{k} p^k = q^{n-k}$$

I.10.VARIABLE ALEATORIA.- Si Ω es un espacio muestral con una medida de probabilidad y X es una función con valor real definida con

respecto a los elementos de Ω , entonces X se denomina variable aleatoria.

Variable aleatoria Discreta (V.A.D.)- es cuando su campo de variación se compone de un número finito o numerable de puntos.

Variable aleatoria Continua (V.A.C.)- es continua si su función de distribución, $F(x)$, es continua, la primera derivada existe y es continua.

Función de densidad de probabilidad de una V.A.D.- si " X " es una V.A.D, la función dada por:

$f(x)=P(X=x)$ para cada x contiene en el intervalo de X se denomina función de Densidad de probabilidad, o distribución de probabilidad de X .

CAPITULO II: CONCEPTOS BASICOS

II.1. LA DISTRIBUCIÓN HIPERGEOMETRICA

Aparece la distribución hipergeométrica al considerar una población de N elementos, dividida en dos clases exhaustivas formadas por A y B puntos muestrales, respectivamente.

Consideremos los elementos de la primera clase como "éxitos" y los de la segunda como "fracasos", se define la variable aleatoria " X " como el número de éxito o puntos de la primera clase obtenidos al realizar " n " pruebas o extracciones sin emplazamiento. La función de probabilidad de la variable " X " es:

$$h_x = p(X = x) = \frac{\binom{A}{x} \binom{B}{n-x}}{\binom{A+B}{n}}, \quad \text{para } \max\{0, n-B\} \leq x \leq \min\{n, A\}$$

Si llamamos N al tamaño de la población, es decir, $N=A+B$, y p y q a las probabilidades de éxito y fracaso iniciales, respectivamente $p=A/N$ y $q=B/N$ Podemos escribir:

$$h_x = p(X = x) = \frac{\binom{Np}{x} \binom{Nq}{n-x}}{\binom{N}{n}} = \frac{\binom{n}{x} \binom{N-n}{Np-x}}{\binom{N}{Np}} \quad \text{para } \max\{0, n-B\} \leq x \leq \min\{n, A\}$$

II.2. VALOR ESPERADO

Sea "X" una variable aleatoria discreta con valores posibles $x_1, x_2, \dots, x_n, \dots$. Y sea $P(X=x_i)$, $i=1, 2, \dots, n, \dots$. Entonces el valor esperado de "X" (o esperanza matemática de x), que se denota con $E(x)$, se define como:

$$E(x) = \sum_{i=1}^{\infty} X_i p(x = x_i)$$

II.3. MOMENTO NO CENTRADO

Sea k un número entero positivo. Se define el momento no centrado (o momento respecto al origen) de orden k de la variable, como el valor esperado de X^k . Lo denotaremos:

$$E(x^k) = u'_k(x) = u'_k \quad . \quad \text{Se verifica que } u'_0 = 1 \text{ y } u'_1 = E[x] = u$$

II.4. FUNCIÓN GENERATRIZ DE MOMENTOS

Supóngase que "X" es una variable aleatoria, es decir, "X" es una función del espacio muestral a los números reales. Al calcular diversas características de la variable aleatoria "X" como $E(x)$, $V(x)$, trabajamos directamente con la distribución de probabilidades de "X" (La distribución de probabilidad esta dada por una función: la fdp en el caso continuo, o las probabilidades puntuales $P(x_i) = p(x = x_i)$ en el caso discreto. La ultima también se puede considerar como una función que toma valores distintos de cero solo si $x=x_i$, $i=1, 2, \dots$)

Posiblemente podemos presentar otra función y hacer los cálculos necesarios mediante ella.

DEFINICIÓN.- Sea "x" una variable aleatoria discreta con distribución de probabilidad $p(x_i) = p(x = x_i)$, $i=1, 2, \dots$. La función M_x , llamada FUNCIÓN GENERADORA DE MOMENTOS DE X, se define como:

$$M_x(t) = \sum_{j=1}^{\infty} e^{tx_j} p(x_j)$$

Si X es una variable aleatoria continua con fdp f, definimos la función Generadora de momentos como:

$$M_x(t) = \int_{-\infty}^{+\infty} e^{tx} f(x) dx$$

OBSERVACIONES

- tanto en el caso discreto como en el continuo, $M_x(t)$ es simplemente el valor esperado de e^{tx} por tanto, podemos combinar las expresiones anteriores y escribir.
- $M_x(t) = E(e^{tx})$
- La función Generatriz de Momentos se escribe como una serie infinita o integral (impropia), dependiendo si la variable aleatoria es discreta o continua. Tal serie puede no existir siempre (es decir, convergir a un valor infinito). Para todos los valores de t . sin embargo, no nos interesara esta posible dificultad. Cada vez que hagamos uso de la función generadora de momentos siempre supondremos que existe. (Para $t=0$, la función Generadora de momentos siempre existe y es igual a 1)

II.5. FUNCIÓN CARACTERÍSTICA

Hay otra función muy relacionada con la función generatriz de momentos que a menudo se usa en su lugar, se llama Función Característica, se denota con C_x y se define como:

$$C_x(t) = E[e^{itx}] \quad \text{donde : } i = \sqrt{-1} \quad \text{la unidad imaginaria.}$$

Por razones teóricas, hay una ventaja considerable al usar $C_x(t)$ en vez de $M_x(t)$ (por esta razón, $C_x(t)$ siempre existe para todos los valores de t). Sin embargo a fin de evitar cálculos con números complejos restringimos nuestra exposición a la función Generadora de Momentos.

II.6. SERIE DE TAYLOR

Alrededor de $x=a$ se puede representar $f(x)$ de manera exacta en la vecindad $x=a$ la serie de Taylor se define como una serie de potencias de la forma. Siguiente:

$$f(x) = f(a) + hf'(a) + \frac{h^2}{2!} f''(a) + \dots + \frac{h^n}{n!} f^{(n)}(a)$$

De donde si evaluamos la n -ésima derivada de la función f calculada en $x=0$. Se obtiene la llamada serie de Maclaurin, es decir:

Sea la función e^x entonces si evaluamos la derivada n -ésima en $x=0$ se obtiene:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + \dots$$

II.7. PROPIEDADES DE LA FUNCIÓN GENERADORA DE MOMENTOS

Daremos ahora la razón para llamar M_x función generadora de momentos. Recordemos el desarrollo en serie de Maclaurin de la función de e^x :

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + \dots$$

(Se sabe que esta serie converge para todos los valores de X.)

$$\text{Así, } e^{tx} = 1 + tx + \frac{(tx)^2}{2!} + \dots + \frac{(tx)^n}{n!} + \dots$$

Ahora:

$$M_x(t) = E(e^{tx}) = E\left(1 + tx + \frac{(tx)^2}{2!} + \dots + \frac{(tx)^n}{n!} + \dots\right)$$

Donde el valor esperado de una suma finita es igual a la suma de los valores esperados. Sin embargo, considerando una suma infinita como la anterior no podemos aplicar de inmediato tal resultado. No obstante, resulta que en condiciones justamente generales esta operación es todavía válida.

Supondremos que las condiciones pedidas se satisfacen y por consiguiente procederemos.

Recordemos que t es una constante respecto a la esperanza y podemos escribir:

$$M_x(t) = 1 + tE(x) + t^2 \frac{E(x^2)}{2!} + \dots + \frac{t^n E(x^n)}{n!} + \dots$$

Si $M_x(t)$ existe en un intervalo $[t] < T$, con $T > 0$, se verifica que $u'_k = E(X^k)$ es el cociente de $t^k/k!$ en el desarrollo de Taylor de $M_x(t)$.

$$M_x(t) = E(e^{tx}) = \sum_{k=0}^{\infty} u'_k \frac{t^k}{k!}$$

II.8. SERIE HIPEREGOMÉTRICA

Se conoce como serie hipergeométrica o serie hipergeométrica de Gauss a la siguiente:

$${}_2F_1(a; b; c; x) = \sum_{i=0}^{\infty} \frac{a_{(i)} b_{(i)} c^i}{c_{(i)} i!} \quad \text{donde } c \neq 0, -1, -2, \dots$$

y los $m_{(i)}$ son los símbolos de pochhammer
 $m_{(i)} = m(m+1) \dots (m+i-1)$, $m_{(0)} = 1$

Si a ó b son enteras negativos la serie tiene un número finito de términos no nulos.

Es absolutamente convergente para $|x| < 1$ y divergente para $|x| > 1$.

Si $|x| = 1$ es:

- a) Absolutamente convergente sí $c-a-b > 0$,
- b) Condicionalmente convergente sí $-1 < c - a - b \leq 0$ y $x = -1$,
- c) Divergente sí $c-a-b \leq -1$.

CAPITULO III: RESUMEN DE LA INVESTIGACIÓN

La función generatriz de probabilidad de la distribución hipergeométrica es :

$$G_n(z) = \sum_{x=\max\{0, n-Nq\}}^{\min\{n, Np\}} P(X = x)z^x$$

$$= h_m z^m {}_2F_1(-n+m, -Np+m; Nq-n+1+2m; z),$$

siendo $m = \max\{0, n - Nq\}$.

Consideremos la expresión de $Mx(t)$ (que en lo sucesivo denotaremos simplemente como $M(t)$ para la distribución hipergeométrica.

$$M(t) = E(e^{tX}) = \sum_{x=\max\{0, n-Nq\}}^{\min\{n, Np\}} P(X = x)e^{tx}$$

$$= h_m e^{tm} {}_2F_1(-n+m, -Np+m; Nq-n+1+2m; e^t),$$

siendo $m = \max\{0, n - Nq\}$.

La serie ${}_2F_1(a, b; c; x)$ es una solución de la "ecuación de Gauss" :

$$x(1-x) \frac{d^2 {}_2F_1(x)}{dx^2} + [c - (a+b+1)x] \frac{d {}_2F_1(x)}{dx} - ab {}_2F_1(x) = 0$$

Si hacemos

$$x = e^t, \quad \frac{d_2 F_1(x)}{dx} = \frac{dF}{dt} e^{-t} \frac{d^2_2 F_1(x)}{dx^2} = \frac{d_2 F}{dt^2} e^{-2t} - \frac{dF}{dt} e^{-2t}$$

como $M(t) = h_m e^{tm}$ se tiene que $F = (1/h_m) e^{-tm} M$ y además

$$F' = \frac{e^{-tm}}{h_m} (M' - nM)$$

$$F'' = \frac{e^{-tm}}{h_m} (M'' - 2mM' + m^2 M)$$

Luego, sustituyendo en la ecuación

$$e^t (1 - e^t) \left[\frac{e^{-tm}}{h_m} \left(M'' - 2mM' + m^2 M e^{-2t} - \frac{e^{-tm}}{hm} (M' - nM) e^{-2t} \right) \right]$$

$$+ \left[Nq - n + 1 + 2m - (-n + m - Np + m + 1) e^t \right] \frac{e^{-tm}}{hm} (M' - nM) e^{-t}$$

$$- (-n + m) (-Np + m) \frac{e^{-tm}}{hm} M = 0$$

y simplificando convenientemente, se obtiene:

$$(1 - e^t) (M'' - emM' + m^2 M - M' + mM)$$

$$+ [Nq - n + 1 + 2m - (-n + 2m - Np + 1) e^t] (M' - nM) - (n - m) (Np - m) e^t M = 0$$

o incluso

$$M'' (1 - e^t) + M' [Nq - n + e^t (n + Np)] - M (m^2 + Nqm - nm + nNpe^t) = 0$$

y recordando que

$$\begin{aligned}
M(t) &= 1 + \mu_1' \frac{t}{1!} + \mu_2' \frac{t^2}{2!} + \dots \\
M'(t) &= \mu_1' + \mu_2' t + \mu_3' \frac{t^2}{2!} + \dots \\
M''(t) &= \mu_2' + \mu_3' t + \mu_4' \frac{t^2}{2!} + \dots \\
et &= 1 + \frac{t}{1!} + \frac{t^2}{2!} + \dots
\end{aligned}$$

Resulta

$$\begin{aligned}
&\left(\mu_2' + \mu_3' \frac{t}{1!} + \mu_4' \frac{t^2}{2!} + \dots \right) \left(-\frac{t}{1!} - \frac{t^2}{2!} - \frac{3}{3!} - \dots \right) \\
&+ \left(\mu_1' + \mu_2' t + \frac{t}{1!} \mu_3' + \frac{t^2}{2!} + \dots \right) \left[Nq - n \left(1 + \frac{t}{1!} + \frac{t^2}{2!} + \dots \right) (n + Np) \right] \\
&- nNp \left(1 + \mu_1' \frac{t}{1!} + \mu_2' \frac{t^2}{2!} + \dots \right) \left(1 + -\frac{t}{1!} - \frac{t^2}{2!} + \dots \right) = 0
\end{aligned}$$

Igualando ahora a cero el coeficiente de t^k resulta

$$\sum_{i=1}^k \binom{k}{i} \left[\mu_{k+2-i}' - (n + Np) \mu_{k+1-i}' + nNp \mu_{k+i}' \right] = N \mu_{k+i}' - nNp \mu_k' \quad (1)$$

y utilizando el operador desplazamiento E, $E^p(\mu_s') = \mu_{s+p}'$, tenemos

$$\left[(1 + E)^k - E^k \right] \left[\mu_2' - (n + Np) \mu_1' + nNp \mu_0' \right] = N \mu_{k+1}' - nNp \mu_k' \quad (2)$$

resulta válido para $k \leq 0$.

La expresión (1) puede escribirse también en la forma

$$\sum_{i=1}^k \binom{k}{i-1} \left[\mu_{i+1}' - (n + Np) \mu_i' + nNp \mu_{i-1}' \right] = N \mu_{k+1}' - nNp \mu_k'$$

y de aquí, si convenimos que $\binom{k}{-2} = \binom{k}{-1} = 0$ para todo k, podemos obtener el valor del $(k+1)$ -ésimo momento en función de los k + 1 precedentes:

$$\mu_{k+1}^{\cdot} = \sum_{i=0}^k \mu_i^{\cdot} \left[\binom{k}{i-2} - (n+Np) \binom{k}{i-1} + nNp \binom{k}{i} \right] \quad (3)$$

valido para $k \geq 0$.

Para finalizar, si aplicamos los resultados (1), (2) o (3) para los primeros valores de k se obtiene:

$$\mu_0^{\cdot} = 1$$

$$\mu_1^{\cdot} = np$$

$$\mu_2^{\cdot} = \frac{np(nNp - n - Np + N)}{N - 1}$$

$$\mu_3^{\cdot} = \frac{np(2n^2 + N^2 + 3nNp + 3nN^2p + n^2N^2p^2 + 2N^2p^2 - 3n^2Np - 3nN^2p^2 - 3nN - 3N_2p)}{(N - 1)(N - 2)}$$

CONCLUSIÓN.

De la exposición realizada podemos observar que si existe relación entre los k+1 primeros momentos no centrados de la distribución, por lo tanto podemos realizar o describir las características de ésta distribución.