

MODELOS DE PROBABILIDAD I

MODELOS DE PROBABILIDAD (INTRODUCCIÓN)

PROCESOS EXPERIMENTALES

DISTRIBUCIÓN CAUSAL

DISTRIBUCIÓN DICOTÓMICA

DISTRIBUCIÓN BINOMIAL

Una distribución de probabilidad queda definida y caracterizada por:

- 1.- la especificación de la variable aleatoria y su campo de variación.
- 2.- la especificación de su asignación de probabilidades, mediante la función de distribución.(Alternativamente mediante la f.cuantía o densidad , la F.C. o la F.G.M.(si existe).(Estas son las FUNCIONES DE DEFINICIÓN)

Si un conjunto dado de distribuciones tiene sus funciones de distribución con la misma ESTRUCTURA FUNCIONAL, diremos que pertenece a la misma FAMILIA DE DISTRIBUCIONES, al mismo MODELO DE PROBABILIDAD o a la misma DISTRIBUCIÓN-TIPO.

p.ej : Todas las distribuciones que están definidas sobre una v.a. continua de modo que para $x \geq 0$ la función de densidad es : $f(x) = \alpha e^{-\alpha x}$ siendo α un real positivo (alternativamente: $F(x) = 1 - e^{-\alpha x}$; $\phi(t) = (1 - t/\alpha)^{-1}$; son equivalentes las tres caracterizaciones), pertenecen a la misma familia, modelo o tipo (el exponencial).

La estructura matemática de las funciones de definición que caracterizan un modelo de probabilidad suelen depender de uno o más parámetros. Estos parámetros son los PARÁMETROS DE LA DISTRIBUCIÓN(TIPO), y tienen un importancia fundamental, en Estadística matemática y sobre todo en INFERENCIA ESTADÍSTICA.

Muchos modelos de probabilidad pueden establecerse teóricamente sin necesidad de recurrir a un sistema de aleatorización racional .Sin embargo , en muchos casos resulta conveniente definir los modelos de probabilidad recurriendo a un claro sistema de aleatorización sobre determinado tipo de fenómeno aleatorio .Procediendo de esta manera podremos disponer de un sistema para identificar el modelo a aplicar en un gran número de situaciones prácticas semejantes.

El procedimiento es sencillo : primero haremos una clasificación de los fenómenos aleatorios de más fácil determinación (procesos experimentales), después determinaremos algunas aleatorizaciones que nos generan variables aleatorias cargadas de gran significado práctico y , por último , obtendremos la estructura funcional de las funciones de definición de su distribución , partiendo , para ello, de la probabilidad inducida para la variable por el fenómeno aleatorio.

Nos apoyamos , por tanto ,en el concepto de proceso experimental para definir muchos de los modelos de probabilidad que vamos a estudiar.

Un proceso experimental es el conjunto de características que rigen la realización de un determinado fenómeno aleatorio. Un proceso quedará definido por una serie de características o hipótesis que puedan aplicarse a cierta categoría de experimentos o experiencias en las que participa el azar. Cada proceso dará cuenta de un conjunto de fenómenos similares que se producen con las mismas características o bajo las mismas hipótesis.

A partir de las características del fenómeno que analicemos (partiendo del proceso experimental del que se trate) podremos , identificando la variable aleatoria que nos interesa , estudiar y determinar la estructura matemática de su distribución .Podremos agrupar los modelos de probabilidad a aplicar.

CLASIFICACIÓN DE LOS PROCESOS EXPERIMENTALES

De una manera un exhaustiva podemos clasificar los procesos experimentales en tres grandes grupos:

- **Procesos experimentales puros**
- **Procesos experimentales de observación**
- **Procesos experimentales de selección o extracción aleatoria.**

Cabría considerar un cuarto grupo que supondrían aquellos procesos experimentales de "salto al límite" de cualquiera de los anteriores, cuando las características propias del fenómeno considerado tomen valores tan elevados que puedan considerarse tendentes a infinito , serían procesos que podemos denominar Gaussianos en honor al investigador de la distribución normal , ya que es a esta distribución o modelo al que suelen converger en su "salto al límite" los modelos o distribuciones de otros procesos .

Procesos experimentales puros son aquellos en los que se considera la realización de una prueba o experimento una o más veces. Cada prueba realizada podrá darnos un cierto número de resultados posibles , siempre susceptibles de convertirse en dos únicos complementarios : éxito o fracaso. Cada resultado tendrá un probabilidad de ocurrir. Dependiendo de las características de estas probabilidades (constantes o no a lo largo del proceso) , del número de pruebas (una , varias o un número indeterminado) , y sobre todo de la aleatorización que consideremos , que dependerá de las pretensiones de nuestra experimentación ,podremos derivar distintas distribuciones de probabilidad. Cuando cada prueba puede dar tan sólo uno de dos resultados posibles (éxito o fracaso) suele hablarse de que se trata de una experiencia de *Bernouilli*. En honor a este autor podemos llamar a todos estos procesos (tengan o no sus pruebas dos únicos resultados posibles) procesos de *Bernouilli*.

De los procesos de Bernouilli podemos hacer derivar distribuciones de variable discreta muy importantes como la dicotómica, binomial , la geométrica, binomial negativa , hipergeométrica , la hipergeométrica negativa, la polinomial o la hipergeométrica de varias variables. Cada una de ellas podrá deducirse dependiendo de las características de las pruebas (dos o más resultados) de la naturaleza de las probabilidades (constantes o no) y del número de pruebas; y por supuesto de la aleatorización considerada que dependerá de nuestros intereses prácticos.

Los procesos experimentales de observación engloban situaciones y fenómenos en los que se observa la naturaleza (o, por decir lo de una manera más amplia, la realidad) a la espera de que se produzca un hecho, durante un determinado periodo experimental o a lo largo de un determinado espacio de experimentación (durante un intervalo de tiempo o de espacio). El hecho sujeto a estudio puede o no producirse, escapándose su realización al control causal ; esto es, es aleatorio. Igualmente, puede producirse el hecho ninguna , una o más veces, durante el periodo experimental. Ejemplos de estos tipo de hechos serían el desencadenamiento de un accidente o un fallo, una llamada telefónica, un siniestro, la llegada de un cliente a una oficina etc.

Es fundamental distinguir entre los procesos experimentales puros (de Bernouilli) y los procesos de observación : En los primeros un experimentador realiza una o varias pruebas, en los segundos se limita a observa que un evento se produzca (o no).

Si el objeto de nuestro interés es el número de hechos que se producen en periodo experimental una adecuada aleatorización nos llevará a la consideración de una variable aleatoria discreta, fácilmente definible como "número de hechos ocurridos" .Bajo ciertas condiciones, podremos derivar una adecuada distribución para esta variable. La distribución más importante que puede derivarse para este tipo de casos es la distribución de Poisson.

Si, por el contrario, nos interesa determinar el tiempo (o el espacio) necesarios para que se produzca el hecho que consideramos ; el tiempo (o el espacio) para que se produzca el hecho será una variable aleatoria continua .Las distribuciones ,derivadas, serán las asociadas a los fenómenos de espera, o de fiabilidad, que estudiaremos posteriormente. (La más, importante de ellas, es la distribución exponencial).

Los procesos de selección aleatoria se caracterizan por la extracción aleatoria de uno o más individuos de entre el conjunto de los que constituyen la población estudiada. De los individuos seleccionados se podrán analizar características cuantitativas o cuantitizables. Estas características serán variables aleatorias que dependiendo de su propia , naturaleza y de las hipótesis del proceso tendrán una u otra distribución y seguirán uno u otro modelo. Los procesos de selección aleatoria son fundamentales en la inferencia estadística ya que a partir de ellos puede deducirse toda la teoría del muestreo. Estos procesos, serán considerados con detalle en el estudio del muestreo aleatorio. Digamos ahora, que todas las distribuciones de probabilidad pueden ser derivadas de procesos de este tipo .En cualquier caso, la selección por antonomasia genera las distribuciones uniforme de variable discreta y continua, y desde el punto de vista práctico, las hipótesis de este tipo de procesos nos conducen en muchos casos a la distribución Normal o algunas de sus distribuciones derivadas.

DISTRIBUCIÓN CAUSAL

La distribución causal no es, propiamente una distribución de probabilidad. La razón es que es aplicable en aquellos *casos* en que la probabilidad de un único y cierto valor de la variable es 1 y, para todo el resto de valores reales la probabilidad es cero. En consecuencia esta distribución sólo puede dar cuenta de fenómenos en los que la ocurrencia de un cierto suceso es segura y la ocurrencia de cualquier otro es imposible. En estos casos, obviamente, no existe azar, pero a través de esta distribución, nos resulta fácil ver la causalidad o el determinismo como un caso particular de azar o aleatoriedad.

Proceso experimental del que se puede hacer derivar.

Esta distribución puede hacerse derivar de un proceso puro o de Bernoulli si consideramos las siguientes

- **Se realiza un número definido de pruebas, N.**
- **Cada una de las pruebas puede dar dos resultados mutuamente excluyentes : A y no A**
- **La probabilidad en cada prueba de obtener un resultado A es siempre la misma (constante) y es igual a 1 , lógicamente la probabilidad de obtener en cualquier prueba un resultado no A será cero.**
- **(Derivación de la distribución) si en estas circunstancias aleatorizamos de forma que la variable aleatoria X se defina como " el número de resultados A obtenidos en N pruebas" la variable aleatoria X tendrá una distribución causal (de parámetro N)**

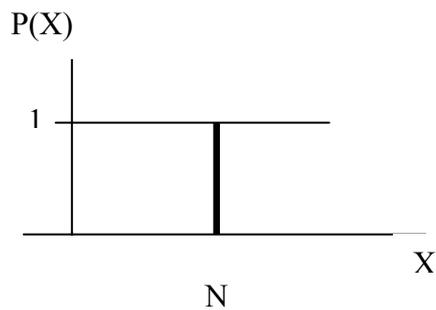
Características analíticas ,de la distribución

Es fácil comprobar que la función de cuantía de esta distribución será :

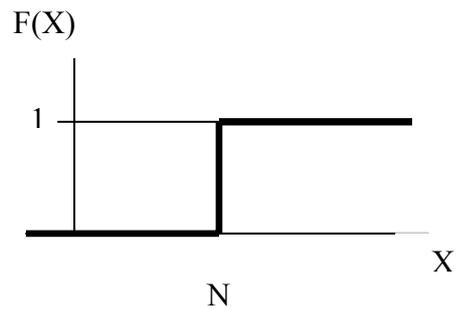
$$\begin{array}{l} x \\ x = N \\ x = N \end{array} \quad \begin{array}{l} P(x) \\ 1 \\ 0 \end{array}$$

$$\begin{array}{l} x \\ x < N \\ x \geq N \end{array} \quad \begin{array}{l} F(x) \\ 0 \\ 1 \end{array}$$

Siendo la función de distribución



función de cuantía



función de distribución

La media del modelo causal será

$$\mu = E[x] = \sum_{\forall x} x_i P(x_i) = N$$

Siendo la varianza

$$\sigma^2 = \alpha_2 - \mu^2 = N^2 - N^2 = 0 \text{ dado que}$$

$$\alpha_2 = E[x^2] = \sum_{\forall x} x_i^2 P(x_i) = N^2$$

DISTRIBUCIÓN DICOTÓMICA.(Bernouilli).

El campo de variación de la variable es : $\{0,1\}$. y la función de cuantía es :

$$P(X=0) = q = 1-p$$

$$P(X=1) = p .$$

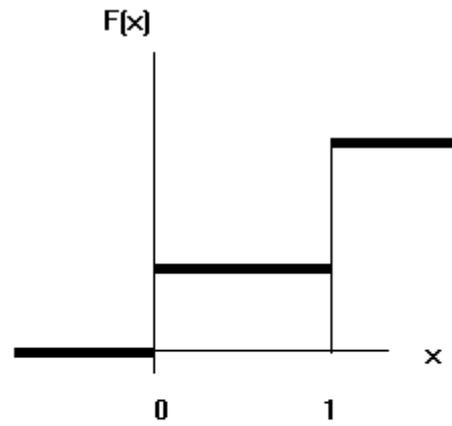
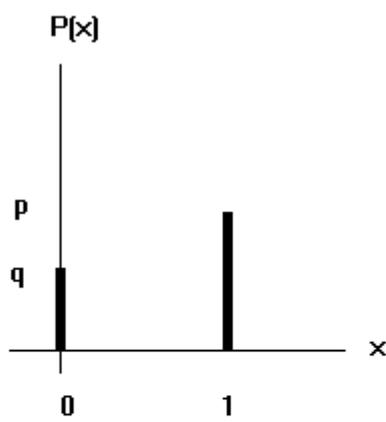
Si una variable aleatoria X sigue o tiene una distribución dicotómica de parámetro p

se expresará como $x \Rightarrow D(p)$

Modeliza situaciones en las que :

- Se realiza una prueba
- Que sólo puede dar dos resultados posibles: A y A
- La probabilidad del resultado A es $P(A) = p$ y la del resultado A es $P(A) = q = 1-p$.

- En estas circunstancias la variable aleatoria X significa "nº de resultados A que se obtienen.



x	$P[x]$
0	q
1	p
	<hr/>
	1

x	$F[x]$
$x < 0$	0
$0 \leq x < 1$	q
$x \geq 1$	$1 = p+q$

La media de la distribución será: $\mu = E[x] = \sum_{\forall x} x_i P(x_i) = 0 \cdot q + 1 \cdot p = p$

La varianza de la distribución: $\sigma^2 = \alpha_2 - \mu^2$

con : $\alpha_2 = \sum x^2 \cdot P(x) = 0 \cdot q + 1 \cdot p = p$

$$\sigma^2 = \alpha_2 - \mu^2 = p - p^2 = p(1-p) = p \cdot q$$

Y la F.G.M.:

$$\phi(t) = E(e^{tx}) = \sum e^{tx} P(x) = e^0 q + e^t p = pe^t + q$$

Es fácil comprobar que todos los momentos ordinarios de orden mayor o igual a 1 son iguales a p.

DISTRIBUCIÓN BINOMIAL.

Nos encontramos con un modelo derivado de un proceso experimental puro, en el que se plantean las siguientes circunstancias.

- Se realiza un número n de pruebas (separadas o separables).
- Cada prueba puede dar dos únicos resultados A y \bar{A}
- La probabilidad de obtener un resultado A es p y la de obtener un resultado \bar{A} es q , con $q= 1-p$, en todas las pruebas. Esto implica que las pruebas se realizan exactamente en las mismas condiciones y son , por tanto ,independientes en sus resultados. Si se trata de extracciones, (muestreo), las extracciones deberán ser con devolución (reemplazamiento) , o bien población grande (M.A.S). A este respecto hagamos una consideración: si el proceso consiste en extraer individuos de una población y observar si poseen cierta característica: el parámetro n será el número de extracciones (tamaño muestral) y el parámetro p la proporción de individuos de la población que poseen la característica en cuestión. Se ha comentado que para que la probabilidad, de que en cada extracción obtengamos un individuo poseedor de la característica sea constante en todas la pruebas es necesario que las proporciones poblacionales no cambien tras cada extracción es decir se reemplace cada individuo extraído .Sin embargo si la población es muy grande, aunque no reemplacemos los individuos extraídos las variaciones en las proporciones de la población restante serán muy pequeñas y, aunque de hecho las probabilidades de, obtener un éxito varíen tras cada prueba, esta variación será muy pequeña y podremos considerar que son constantes .

Ilustremos con un ejemplo:

Supongamos que una ciudad hay 1000000 de habitantes de los cuales 450000 son varones y 550000 son mujeres . Si extraemos un individuo al azar la probabilidad. de que sea mujer será.

$$p = \frac{50000}{100000} = 0,55$$

Si repetimos esta prueba varias veces y no reponemos "en el saco" al sujeto extraído la probabilidad de obtener una mujer en cada siguiente extracción variará, al variar la composición por sexos de la población restante. Sin embargo, al ser la población tan grande, la variación de esta probabilidad con cada sucesiva prueba será prácticamente despreciable y podremos considerar, en la práctica que las probabilidades son constantes: en efecto:

Si, en la primera prueba obtenemos una mujer y no la reintegramos a la población la de probabilidad de obtener una mujer en la segunda prueba será:

$$p_1 = \frac{54999}{99999} = 0,5499999 \approx 0,55$$

Si por el contrario en la primera prueba se obtiene un varón la probabilidad de obtener una mujer el siguiente será: $p_1 = \frac{54999}{100000} = 0,550000001 \approx 0,55$

Por lo tanto bien podríamos considerar que la probabilidad de extraer una mujer en sucesivas elecciones aleatorias es constante. En consecuencia, si consideráramos el muestreo de 10 individuos de esa ciudad, aunque no reemplazáramos las extracciones, la variable aleatoria x = número de mujeres obtenidas en las diez extracciones, seguiría una distribución binomial de parámetros $n = 10$ Y $p= 0.55$.

Sin embargo, si la población es pequeña, las variaciones de la probabilidad de éxito con cada prueba serán importantes sino se devuelve a la población original cada sujeto extraído. En este caso, no podremos considerar que p y q son constantes a lo largo de todo el proceso y el número de éxitos obtenidos en n pruebas será una variable aleatoria que no seguirá una distribución binomial sino una nueva distribución que estudiaremos, más tarde llamada hipergeométrica.

- En estas circunstancias se aleatoriza de forma que variable aleatoria signifique: $X = n^{\circ}$ de resultados A que se obtienen en las n pruebas

Se plantean dos valores con variación por lo que tendremos dos parámetros p y n , por lo que la distribución binomial se explicitará: $X \Rightarrow B(n, p)$

El campo de variación de la variable será $\{0, 1, 2, 3, \dots, n\}$, por lo que no es necesario comentar que es de carácter discreto. Así tendremos que si queremos calcular la probabilidad de que $X=1$ en n pruebas, tendríamos 1 resultado A y $n-1$ resultados no A.

Prueba 1	Prueba 2	Prueba 3		Prueba n	Prueba
No A	No A	No A	No A	No A	A	Resultado
q	q	q	q	q	p	Probabilidad

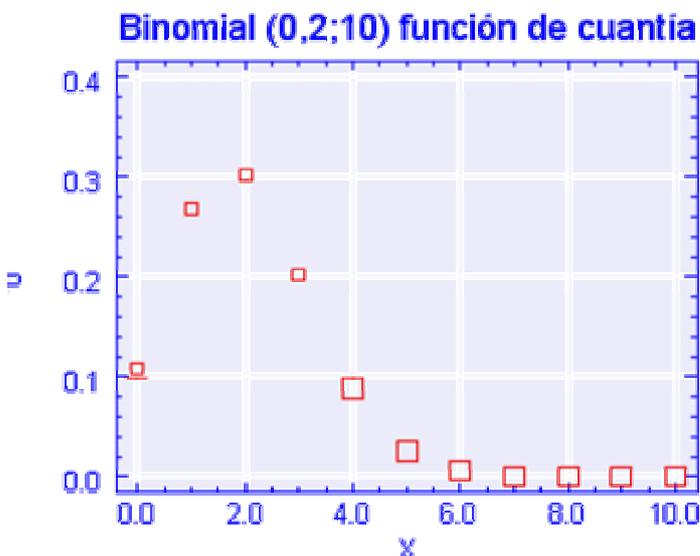
Por lo que la probabilidad de conseguir un resultado A sería en principio, y dado que las pruebas son independientes $P(X=1) = \underbrace{q \cdot q \cdot \dots \cdot q}_{n-1 \text{ veces}} \cdot p = q^{n-1} \cdot p$. Lo que sería correcto si

el resultado A, lo fuera en la última prueba. Dado que nos es indiferente en que prueba sea tendríamos que multiplicar esta probabilidad por $\binom{n}{1}$. Por lo que tendríamos que:

$P(X=1) = \binom{n}{1} \underbrace{q \cdot q \cdot \dots \cdot q}_{n-1 \text{ veces}} \cdot p = \binom{n}{1} q^{n-1} \cdot p$ generalizando para cualquier valor de X no que daría que la función de cuantía de la binomial.

tiene como expresión

$$P(x) = \binom{n}{x} p^x q^{n-x}$$



Gráficamente la función de cuantía quedaría de la siguiente manera, siempre dependiendo de los valores que tomen los parámetros.

En este caso se trata de una binomial $B(0,2;10)$. Siendo esta la representación gráfica de su función de cuantía

La función de distribución será para cada valor k de la variable $F(x = k) = \sum_{i=1}^K P(x_i)$

Los indicadores-momentos (media y varianza) pueden obtenerse a partir de la función de cuantía (operador esperanza) o a a partir de F.G.M.:

Siendo la F.G.M.: $\varphi(t) = E[e^{tx_i}] = \sum_{\forall i} e^{tx_i} \cdot P(x_i)$

$$= \sum_{x=1}^n \left(\frac{n!}{x!(n-x)!} p^x q^{n-x} e^{tx} \right) = \sum_{x=1}^n \left(\frac{n!}{x!(n-x)!} (pe^t)^x q^{n-x} \right)$$

desarrollando el Binomio de Newton obtendríamos, $\varphi(t) = (pe^t + q)^n$

También podríamos haber obtenido la F.G.M., partiendo del hecho de que la distribución o modelo binomial es la reiteración n veces del modelo dicotómico, Así si:

$Y \Rightarrow D(p)$ es decir realizamos una prueba con probabilidad de éxito p, siendo la aleatorización el número de resultados A en esa prueba. Tendremos que un serie de pruebas independientes (n) con probabilidad de A en cada prueba, será la agregación de n dicotómicas. Si X es el número de resultados A en esas n pruebas y por tanto una Binomial n,p. Así

$$X \Rightarrow B(n, p) \equiv \underbrace{Y + Y \dots Y}_{n \text{ veces}} \equiv \underbrace{D(p) + D(p) \dots + D(p)}_{n \text{ veces}}$$

De esta manera si conocemos que la F.G.M de la dicotómica (Y) es $\varphi(t) = pe^t + q$

$$\text{dado que } X \Rightarrow B(n, p) \equiv \underbrace{Y + Y \dots + Y}_{n \text{ veces}}$$

la F.G.M. de X (binomial) será el producto, n veces, de la F.G.M de Y (dicotómica) dado que las Y son independientes. Así

$$\varphi_x(t) = \underbrace{\varphi_y(t) \cdot \varphi_y(t) \cdots \varphi_y(t)}_{n \text{ veces}} = \underbrace{(pe^t + q) \cdot (pe^t + q) \cdots (pe^t + q)}_{n \text{ veces}} = (pe^t + q)^n$$

Mediante la aplicación del teorema de los momentos podemos calcular las ratios del modelo.

La media sería $\mu = \alpha_1 = \varphi'(t)$
 $t \rightarrow 0$

$$\varphi'(t) = n \left(pe^t + q \right)^{n-1} \cdot pe^t \quad \text{si hacemos } t=0 \quad \varphi'(t=0) = n \left(p+q \right)^{n-1} \cdot p$$

por lo que la media sería $\mu = n \cdot (1)^{n-1} \cdot p = np \quad \mu = np$

Podríamos haber calculado la media partiendo de la binomial como reiteración de dicotómicas ; así

Si $Y \Rightarrow B(n, P)$ e $Y \Rightarrow D(p)$ conociendo que la media de $Y=p$

$$E[x] = E \left[\underbrace{Y + Y + \dots + Y}_{n \text{ veces}} \right] = \underbrace{E[y] + E[y] + \dots + E[y]}_{n \text{ veces}} = \underbrace{p + p + \dots + p}_{n \text{ veces}} = np$$

que lógicamente nos resulta con el mismo resultado

La varianza de la Binomial vendría dada por $\sigma^2 = \alpha_2 - \mu^2$

Siendo $\alpha_2 = \varphi''(t)$
 $t \rightarrow 0$

$$\varphi''(t) = n \cdot (n-1) \left(pe^t + q \right)^{n-2} \cdot pe^t \cdot pe^t + pe^t \cdot n \cdot \left(pe^t + q \right)^{n-1}$$

si hacemos $t=0 \quad \varphi''(t=0) = n \cdot (n-1) (1)^{n-2} \cdot p \cdot p + p \cdot n \cdot (1)^{n-1} = n \cdot (n-1) \cdot p^2 + pn$

o lo que es lo mismo

$$\varphi''(t=0) = n \cdot (n-1) \cdot p^2 + pn = n^2 p^2 - np^2 + np = \alpha_2$$

Así $\sigma^2 = \alpha_2 - \mu^2 = n^2 p^2 - np^2 + np - n^2 p^2 = -np^2 + np = np(1-p) = npq$

Luego la varianza sería $\sigma^2 = npq$

Calculando la varianza de la binomial como reiteración de dicotómicas sería

Si $Y \Rightarrow B(n, P)$ e $Y \Rightarrow D(p)$ conociendo que la varianza de $Y=pq$

$$D^2[x] = D^2 \left[\underbrace{Y + Y + \dots + Y}_{n \text{ veces}} \right] = \underbrace{D^2[y] + D^2[y] + \dots + D^2[y]}_{n \text{ veces}} = \underbrace{pq + pq + \dots + pq}_{n \text{ veces}} = npq$$

Partiendo de la F.G.M obtendríamos los diversos momentos ordinarios necesarios para el cálculo de otros coeficientes como el de asimetría y curtosis. Cuyas expresiones son las siguientes.

Coeficiente de Asimetría $\gamma_1 = \frac{q-p}{\sqrt{npq}}$ en el que se puede comprobar que si $p=q=0,5$ su valor es cero, lo que nos indica simetría para este valor del parámetro p , siendo el que fuere n

Coeficiente de Curtosis $\gamma_2 = \frac{1-6pq}{npq}$ siendo mínimo para $p=q=0,5$

En cuanto a la Moda, dado que es el valor de la variable con mayor probabilidad (maximiza su función de probabilidad). Tendremos que si X_0 es el valor modal

$P(X_0) \geq P(X_i)$ para cualquier valor de i , en particular se verificará que

$P(X_0) \geq P(X_0 + 1)$ y $P(X_0) \geq P(X_0 - 1)$ es decir que el valor de la función de cuantía para el valor modal será mayor que dicho valor para los valores anterior y posterior a la moda.

Sustituyendo por las correspondientes funciones de cuantía en ambas inecuaciones, llegaremos a la expresión $pn - q \leq X_0 \leq pn + p$

Siendo X_0 el valor modal

Generalmente será un único valor (la parte entera de la media), y podrán ser dos valores modales cuando $pn + p$ (ó $pn - q$) sea un número entero. Como ejemplo: si lanzamos un dado 10 veces y queremos saber cuál es el número más probable de ases que conseguiremos tendremos que calcular la moda en un modelo $B(1/6; 10)$. Aplicando la desigualdad anterior tendríamos: $1/6 \cdot 10 - 5/6 \leq Mo \leq 1/6 \cdot 10 + 1/6$. Es decir $5/6 \leq Mo \leq 11/6$. Sólo existe un número natural que verifique esa cota, el 1. Para que hubiera más de una moda, ambos valores de la desigualdad debieran ser números naturales como los valores de la variable X con distribución binomial

Teorema de adición en la binomial

Se dice que una distribución verifica el Teorema de adición, para alguno de sus parámetros, o que es reproductiva, si dadas 2 o más variables aleatorias independientes que siguen todas ellas una distribución de ese tipo con parámetros distintos, la variable suma de todas ellas sigue, también una distribución de ese tipo con parámetros la suma de los parámetros de las variables originales.

Este teorema se prueba siempre a partir de la F.G.M. Se procede obteniendo la F.G.M. de la distribución de la variable suma, que por ser las variables originales independientes será el producto de las F.G.M. de las distribuciones de las variables originales. Una vez obtenida la F.G.M. comprobaremos si efectivamente se trata de F.G.M. de una distribución de ese tipo con parámetros la suma de los parámetros de las distribuciones de las variables originales. En virtud del carácter recursivo de la operación suma, para demostrar el teorema basta con demostrar que se cumple para la

suma de dos variables aleatorias.

Pues bien, la distribución binomial verifica el teorema de adición para el parámetro n , aunque no lo verifica para el parámetro p (además es necesario para que se verifique el teorema que el parámetro p de las distribuciones de las variables originales sea el mismo):

"La variable suma de dos o más variables binomiales independientes de parámetros $(n_1, p) : (n_2, p) : \dots$ se distribuye como una distribución binomial de parámetros $(n_1+n_2+\dots, p)$ "

En efecto :

sean X e Y dos variables aleatorias independientes tales que:

$$X \Rightarrow B(n_1, p)$$

$$Y \Rightarrow B(n_2, p)$$

pretendemos probar que $Z = X + Y \Rightarrow B(n_1 + n_2, p)$

Conocemos que la F.G.M del modelo binomial es para X $\varphi_x(t) = (pe^t + q)^{n_1}$

$$\text{siendo para } Y \quad \varphi_y(t) = (pe^t + q)^{n_2}$$

Dado que X e Y son independientes la F.G.M. de la distribución de la variables suma será el producto de las dos F.G.M.

$$\varphi_z(t) = \varphi_x(t) \cdot \varphi_y(t) = (pe^t + q)^{n_1} (pe^t + q)^{n_2} = (pe^t + q)^{n_1+n_2}$$

que es la F.G.M. de una $B(n_1+n_2, p)$.

Obsérvese que este teorema se utilizó para el cálculo de la F.G.M. de la Binomial como reiteración (suma) de Dicotómicas, con la salvedad de tomar a la Dicotómica como una Binomial de $n = 1$, es decir adición de n binomiales $B(1, p)$