



Universidad Nacional del Litoral
Facultad de Ingeniería y Ciencias Hídricas

ESTADÍSTICA

Ingeniería Informática

TEORÍA

Mg.Ing. Susana Vanlesberg
Profesor Titular

UNIDAD 4

MODELOS PROBABILÍSTICOS

La aplicación de la teoría de probabilidades ha originado una serie de modelos. En diversas situaciones reales se realizan supuestos respecto de un problema, y esto lleva a descripciones análogas y a formas matemáticamente iguales a las de los modelos probabilísticos.

4.1 - MODELOS DE VARIABLE ALEATORIA DISCRETA

4.1.1- Modelo Bernoulli

Tal vez la situación más común que se presenta es aquella en la que los resultados de los experimentos pueden separarse en dos categorías mutuamente excluyentes: *éxito* o *fracaso*; por ejemplo prendido o apagado.

Puede entonces definirse la variable aleatoria x Bernoulli y asignársele valores (arbitrarios pero muy prácticos) a los eventos antes mencionados:

$x=0$ fracaso

$x=1$ éxito

La función masa de x es simplemente:

$$f(x) = \begin{cases} p & \text{si } x = 1 \\ 1 - p & \text{si } x = 0 \end{cases} \quad (1)$$

siendo p la probabilidad de éxito.

Esta función deberá cumplir con las condiciones de una función de probabilidad, una de ellas es que deberá ser igual a uno para todos los valores de la variable:

$$\sum_{x_i=0}^{x_i=1} f(x_i) = 1$$

$$p + (1 - p) = p + 1 - p = 1$$

Las características de este modelo son:

$$E(x) = \sum_{-x_i} x_i f(x_i) = 0 \cdot (1-p) + 1 \cdot p = p \quad (2)$$

$$\begin{aligned} V(x) &= \sum_{-x_i} (x_i - E(x))^2 f(x_i) = \\ &= (0-p)^2 (1-p) + (1-p)^2 \cdot p = \\ &= p^2 (1-p) + (1-p)^2 \cdot p = \\ &= p(1-p)(p+1-p) \end{aligned}$$

$$V(x) = p(1-p) \quad (3)$$

4.1.2 - Modelo Binomial

Si se realizan una serie de pruebas de tipo Bernoulli cuyos resultados sean mutuamente independientes y si la probabilidad de éxito permanece invariable en todas ellas se origina un nuevo modelo denominado BINOMIAL.

Se tratará de determinar la distribución del número total de éxitos en n pruebas de tipo Bernoulli cada una con probabilidad favorable igual a p. Para esto se considerarán 3 pruebas y se analizarán las probabilidades de ningún éxito, 1 éxito, 2 éxitos y 3 éxitos:

- Ningún éxito:

$$0; 0; 0; \quad (1-p)^3$$

- Un éxito:

$$1; 0; 0; \quad \text{ó} \quad 0; 1; 0; \quad \text{ó} \quad 0; 0; 1$$

cada secuencia es un evento, que entre sí son mutuamente excluyentes; cuya probabilidad es : $p(1-p)^2$; de esta manera la probabilidad de un éxito es :

$$3 p (1-p)^2$$

- Dos éxitos:

1 ; 1 ; 0 ; ó 1 ; 0 ; 1 ; ó 0 ; 1 ; 1 ;

nuevamente cada secuencia es un evento que es excluyente de los restantes:

$p^2 (1 - p)$ y el total: $3 p^2 (1 - p)$.

- Tres éxitos:

1 ; 1 ; 1 ; $p ; p ; p = p^3$

Esto permite generalizar y obtener la función masa de probabilidad de este modelo; ya que si se analizan los distintos casos se ve que se tiene en cuenta la forma de presentarse el resultado deseado, lo cual puede resumirse calculando los números combinatorios de n tomados de x (con $x = 0, 1, 2, 3$ en este caso):

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

$$\text{con el coeficiente binomial } \binom{n}{x} = \frac{n!}{x! (n-x)!}$$

(4)

que como se expresó tiene en cuenta el número de formas en que exactamente x éxitos se presenten en n pruebas . n y p son los parámetros de este modelo; n es claro que debe ser entero y p un número comprendido entre 0 y 1 .

El modelo se denomina BINOMIAL porque puede considerarse como el desarrollo del binomio $(p + q)^n$.

La función masa de probabilidad deberá ser igual a 1 para todos los valores que puede tomar la variable:

$$\sum_{x_i} f(x_i) = 1$$

$$\sum_{x_i} \binom{n}{x_i} p^{x_i} q^{n-x_i} =$$

$$= \sum_{x_i=0}^n \frac{n!}{x_i! (n-x_i)!} p^{x_i} q^{n-x_i} =$$

$$= \frac{n!}{0! n!} p^0 q^n + \frac{n!}{1! (n-1)!} p^1 q^{n-1} + \dots + \frac{n!}{n! 0!} p^n q^0 =$$

$$= q^n + \frac{n(n-1)!}{(n-1)!} p q^{n-1} + \dots + p^n = (p+q)^n = 1^n = 1$$

La función de distribución puede obtenerse haciendo la suma de todos los valores de la función masa menores o iguales que x:

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i) \quad (5)$$

Características del modelo

$$E(X) = \sum_{x_i=0}^n x_i f(x_i) =$$

$$= 0 \cdot \frac{n!}{0! (n-0)!} p^0 q^n + 1 \cdot \frac{n!}{1! (n-1)!} p^1 q^{n-1} + \dots + n \cdot \frac{n!}{n! (n-n)!} p^n q^0 =$$

$$= 0 + \frac{n(n-1)!}{(n-1)!} p q^{n-1} + \dots + n p^n q^0 = np(q^{n-1} + \dots + (n-1)p^{n-2}q + p^{n-1}) =$$

$$= np(p+q)^{n-1}$$

$$E(x) = np \quad (6)$$

La varianza se obtiene por considerar a la variable aleatoria X como la suma de n variables independientes idénticamente distribuidas como Bernoulli, con esperanza p y varianza p q:

$$Var(X) = \sum_{i=1}^n Var(x_i) = n p(1-p) = n p q \quad (7)$$

La distribución Binomial ha sido tabulada, tanto su función de cuantía o masa y la de distribución.

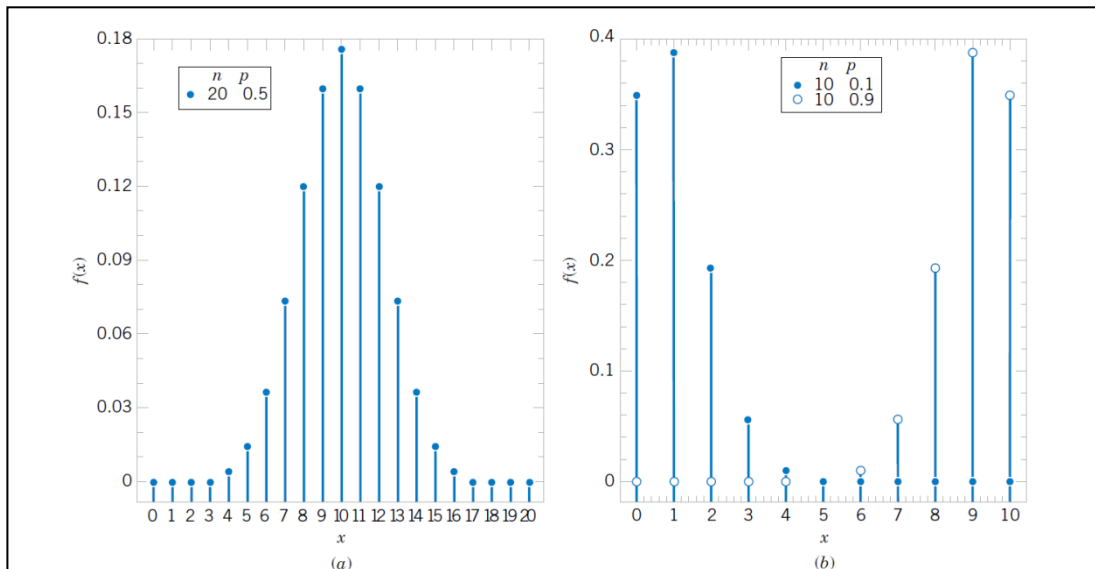


Figura N° 1 - Distribución Binomial para distintos valores de n y p

Ejemplo

Un estudio reciente reveló que el 60% de los conductores de automóviles se coloca el cinturón de seguridad al manejar. Si se selecciona una muestra de 10 automovilistas en una calle de nuestra ciudad, cuál es la probabilidad de que exactamente 7 se hayan colocado el cinturón.

Solución

$A = \{\text{conductor que usa cinturón de seguridad}\}$

$$P(A) = 0.60$$

$X = \text{número de veces que se presenta el suceso } A.$

$$X = 0, 1, 2, 3, \dots$$

$$P(x = 7) = \binom{10}{7} (0.60)^7 (0.40)^{10-7} = 0.215$$

4.1.3 - Modelo Poisson

Considerando las condiciones que caracterizan al modelo Binomial, es necesario analizar que sucede con la distribución de la variable si el número de pruebas n se incrementa y asociado a esto la probabilidad p de éxito se hace cada vez más pequeña.

Es así que n se incrementa y p se hace pequeña, pero el número promedio de eventos en el intervalo total debe permanecer constante e igual a $p \cdot n$. Denominando a esta constante λ y considerando la función masa de probabilidad de x en el límite, esto es con $p \rightarrow 0$ y $n \rightarrow \infty$ y $\lambda = n \cdot p$ a partir de la función del modelo Binomial:

$$\begin{aligned} \lambda = n p &\rightarrow p = \frac{\lambda}{n} \\ f(x) &= \frac{n!}{x!(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} = \\ &= \frac{\lambda^x}{x!} \left(1 - \frac{\lambda}{n}\right)^n \frac{n!}{(n-x)! n^x \left(1 - \frac{\lambda}{n}\right)^x} = \\ &= \frac{\lambda^x}{x!} \left(1 - \frac{\lambda}{n}\right)^n \left[\frac{n(n-1)(n-2)\dots(n-(x+1))(n-x)!}{(n-x)! \left(n \left(1 - \frac{\lambda}{n}\right)\right)^x} \right] \end{aligned}$$

La expresión entre corchetes tiene x términos en el numerador y x términos en el denominador, y para un n suficientemente grande, este término es igual a :

$$n^x / n^x = 1$$

con lo cual queda por evaluar la expresión $(1 - \lambda/n)^n$ que cuando $n \rightarrow \infty$ es igual a la constante $e^{-\lambda}$, esto permite expresar a la función masa de probabilidad del modelo de Poisson de la siguiente forma:

$$f(x) = P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad (8)$$

siendo λ el parámetro del modelo.

Esta función deberá cumplir con la condición de ser igual a 1 para todos los valores que la variable puede tomar:

$$\begin{aligned} \sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} &= \\ &= e^{-\lambda} \left[\frac{\lambda^0}{0!} + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \dots + \frac{\lambda^n}{n!} \right] \\ \left[\frac{\lambda^0}{0!} + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \dots + \frac{\lambda^n}{n!} \right] &= e^{\lambda} \\ e^{-\lambda} \cdot e^{\lambda} &= e^{-\lambda + \lambda} = e^0 = 1 \end{aligned}$$

Características

$$E(X) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda} \lambda^x}{x!} =$$

$$\lambda \sum_{x=0}^{\infty} x \frac{e^{-\lambda} \lambda^{x-1}}{x(x-1)!}$$

$$\sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^{x-1}}{(x-1)!} = 1$$

$$E(X) = \lambda \quad (9)$$

$$\text{Var}(X) = E(X^2) - E^2(X)$$

$$E(X^2) = \sum_{x=0}^{\infty} x^2 \frac{e^{-\lambda} \lambda^x}{x!} = \lambda \left[\sum_{x=0}^{\infty} x \frac{\lambda^{x-1} e^{-\lambda}}{(x-1)!} \right] =$$

$$= \lambda \left[\sum_{x=1}^{\infty} \frac{x \lambda^{x-1} e^{-\lambda}}{(x-1)!} \right]$$

si $y = x - 1 \rightarrow y + 1 = x$, luego

$$\lambda \left[\sum_{y=0}^{\infty} (y+1) \frac{\lambda^y e^{-\lambda}}{y!} \right] = \lambda \left[\sum_{y=0}^{\infty} y \frac{e^{-\lambda} \lambda^y}{y!} + \sum_{y=0}^{\infty} \frac{e^{-\lambda} \lambda^y}{y!} \right]$$

$$\sum_{y=0}^{\infty} y \frac{e^{-\lambda} \lambda^y}{y!} = E(Y) = \lambda$$

$$\sum_{y=0}^{\infty} \frac{e^{-\lambda} \lambda^y}{y!} = 1$$

luego

$$E(X^2) = \lambda(\lambda + 1) = \lambda^2 + \lambda$$

$$Var(X) = \lambda^2 + \lambda - \lambda^2$$

$$Var(X) = \lambda \tag{10}$$

Generalmente este modelo se vincula a aquellos eventos que ocurren en una unidad de tiempo, luego el período de tiempo en el que se realiza el análisis constituye una secuencia de pruebas independientes cada una con distribución Binomial.

Si se tomara para el análisis un intervalo de tiempo igual al doble o al triple del inicial se verá que el parámetro es también igual al doble, al triple, etc., marcando esto la dependencia del tiempo de este modelo y por ello vinculado a los procesos estocásticos. Se entiende por procesos estocásticos a aquellos en los que interesa la secuencia, en el tiempo, de ocurrencia de eventos.

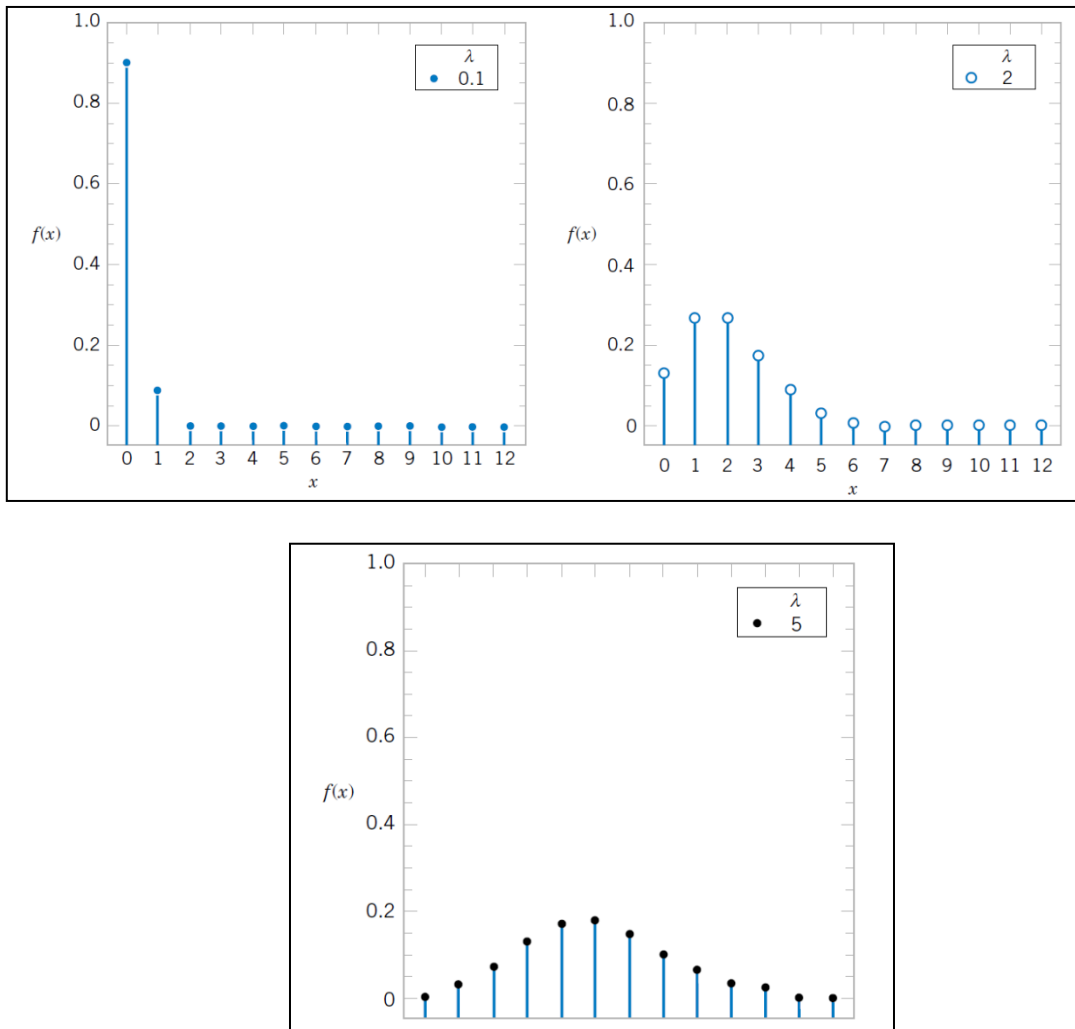


Figura N° 2 - Distribución de Poisson para distintos valores del parámetro

Ejemplo

Un estudio de filas de clientes en las cajas registradoras en Supermercados en una zona dada reveló que durante cierto período (entre las 5 y las 9 p.m.) en los fines de semana, el número medio de clientes en espera fue igual a 4. Determinar la probabilidad de que al visitar un supermercado en ese lapso de tiempo se encuentre que no hay clientes en la fila de una de las cajas.

Solución

$$P(X = 0) = \frac{e^{-\lambda} \lambda^0}{0!} = \frac{e^{-4} 4^0}{0!} = 0,018$$

4.1.4 - Modelo hipergeométrico

Este modelo surge cuando se realiza un muestreo sin reposición de una población finita con sus elementos clasificados en dos categorías.

Si N es el total de elementos de los cuales hay k de una categoría y N-k de otra, al realizar una extracción de n elementos, sin reposición, cada extracción que se realice posteriormente es dependiente del resultado de la extracción anterior con lo cual va cambiando la probabilidad de éxito.

Para derivar la función correspondiente a la variable aleatoria x: número de éxitos o elementos pertenecientes a la categoría que se estudia, en una extracción, se deberán considerar todas las maneras posibles o combinaciones, de extraer x elementos de la categoría deseada, de los n extraídos y los restantes que pertenezcan a la otra categoría. El total de casos se obtiene de las combinaciones del total N extraídos de a n. Luego la función masa de probabilidad para este modelo es la siguiente:

$$P(X = x) = f(x) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}} \quad (11)$$

Características

Para obtener las características es posible decir que la variable aleatoria X es la suma de n variables x_i como en el caso Binomial, pero con la diferencia que aquí las x_i son dependientes. pero como para sumar las esperanzas no se necesita que las variables aleatorias sean independientes es posible obtener la esperanza de la siguiente forma:

$$E(X) = E(x_1) + E(x_2) + \dots + E(x_n)$$

donde cada $E(x_i)$ es el producto de la probabilidad de x en la iésima prueba: k/N , si no se sabe que ha ocurrido en pruebas anteriores o posteriores por la cantidad de pruebas:

$$E(X) = n p = n \left(\frac{k}{N} \right) \quad (12)$$

La varianza en cambio no es aditiva para variables dependientes. Pero se obtiene la siguiente expresión:

$$Var(X) = \frac{n k (N - k)}{N^2} \left(\frac{N - n}{N - 1} \right) \quad (13)$$

$$Var(X) = n * p * q * \left(\frac{N - n}{N - 1} \right)$$

siendo $\left(\frac{N - n}{N - 1} \right)$ el factor de corrección por muestreo sin reposición y población finita.

Cuando $\left(\frac{n}{N} \right) \leq 0.05$ la distribución Hipergeométrica se aproxima a la Binomial.

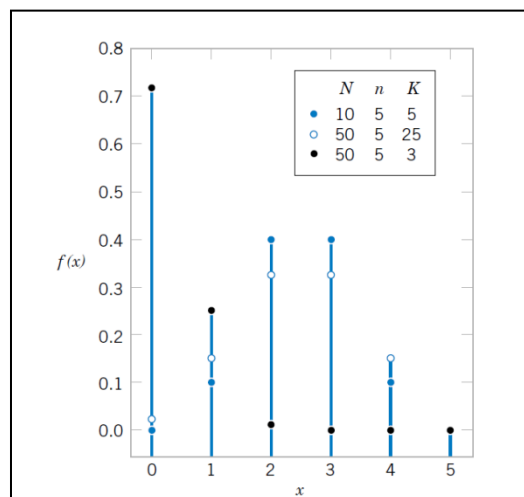


Figura N° 3 - Distribución Hipergeométrica para distintos valores de N, n y K

Ejemplo

Se sabe que de una caja que contiene 20 diskettes, 13 son nuevos y 7 son usados. Si se seleccionan al azar 3 de ellos; determinar la probabilidad de que 2 sean nuevos.

Solución.

$$P(x = 2) = \frac{\binom{13}{2} \binom{7}{1}}{\binom{20}{3}}$$