Menú principal

Principales Distribuciones de una variable aleatoria discreta.

Esquema

Como se sabe, la ley de probabilidad de una variable aleatoria discreta X está bien definida si se conoce su distribución de probabilidad $P(X=x_i)$; i = 1, 2, …, n. O bien, si se conoce su función de distribución F(x), teniendo que:

$$\sum_i P(X=x_i)=1; \qquad F(x)=P(X\le x)=\sum_{x_i\le x} P(X=x_i)$$

Distribución Uniforme.

Una variable aleatoria discreta X que toma valores enteros 1, 2, …, n con probabilidades: P(X=k)=1/n; k=1, 2, …, n recibe el nombre de variable uniforme discreta y su distribución de probabilidad distribución uniforme discreta.

a) Depende de un solo parámetro n.

b) Su media y varianza son:

$$\mu = \frac{n+1}{2}; \qquad \sigma^2 = \frac{n^2-1}{12}$$

[Volver al principio]

Distribución Binomial.

Supongamos un experimento aleatorio S y consideremos asociado con él un suceso A de probabilidad p y sea A* el suceso contrario, cuya probabilidad será q=1-p. Para distinguirlos con mayor facilidad, al suceso A lo llamaremos éxito, y al suceso A* fracaso.

Repitamos el experimento en las mismas condiciones n veces y supongamos que:

a) El resultado de cada ensayo es independiente de los resultados anteriores.

b) La probabilidad de A permanece constante y no varía de una prueba a otra..

Todo experimento que tenga esta características diremos que sigue el modelo de la **distribución binomial.**

A la variable X que expresa el número de éxitos obtenidos en las n pruebas del experimento aleatorio la llamaremos variable aleatoria binomial. Representaremos por **B(n, p)** a la variable de la distribución, siendo **n** y **p** los parámetros.

Función de probabilidad de la distribución Binomial.

La variable aleatoria que expresa el número de éxitos en n pruebas, es una variable aleatoria discreta cuyos valores varían de 0, 1, 2, ..., n. Y las probabilidades con que toman dichos valores se calculan de la siguiente forma: $p_r = \binom{n}{r} p^r q^{n-r}$

Media y varianza de la distribución binomial.

a) Media:

$$\mu = n.p$$

b) Varianza:

$$\sigma^2 = n.p.q$$

En ella hay que observar:

c) Es simétrica si p=q. Si p<q asimetría a la derecha; si p>q, asimetría a la izquierda. Cuando n $\rightarrow \infty$ la distribución tiende a ser simétrica (distribución normal).

d) Los valores de P(X=k) están tabulados para valores de p comprendidos entre 0 y 0.5. Si p>0.5, hay que tener en cuenta que B(n, k, p) = B(n, n-k, q).

Ajuste.

El problema del ajuste es el de buscar una distribución binomial que "*mejor*" se adapta a unos datos, es decir, se ha de calcular el valor del parámetro p para el cual las probabilidades binomiales obtenidas sean las mas cercanas posibles a los datos. El procedimiento que se sigue es igualar la media de la muestra a la media de la población, con lo que tenemos:

$$\bar{x} = np, \quad luego \ \ tomaremos \quad p = \frac{\bar{x}}{n}$$

Con este valor de p se calculan las correspondientes probabilidades y se comprueba si el ajuste es bueno. Para medir la bondad del ajuste de una manera objetiva existen métodos matemáticos que se tratarán más adelante.

[Volver al principio]

Distribución polinomial.

Es una generalización de la distribución binomial para el caso de que en cada prueba se consideren k sucesos excluyentes $A_1, A_2, \ldots, A_k$ con probabilidades respectivas $p_1, p_2, \ldots, p_k$ siendo la suma de todas igual a la unidad.

Supongamos que se realizan sucesivamente n pruebas independientes de este tipo y consideremos las variables X_i = "Número de veces que ocurre el suceso A_i en las n pruebas".

A la variable k-dimensional $(X_1, X_2, \ldots, X_k)$ se le denomina variable polinomial o multinomial.

Para hallar su función de probabilidad $P(X_1=n_1, X_2=n_2, \ldots, X_k=n_k)$ con

$$\sum_i n_i = n$$

consideremos uno de sus sucesos favorables:

$$\overbrace{A_1 A_1 \ldots A_1}^{n_1} \overbrace{A_2 A_2 \ldots A_2}^{n_2} \ldots \overbrace{A_k A_k \ldots A_k}^{n_k}$$

$$\begin{cases} probabilidad = p_1^{n_1} \cdot p_2^{n_2} \ldots p_k^{n_k} \\ núm.\ de\ sucesos = \dfrac{n!}{n_1!\,n_2!\ldots n_k!} \end{cases} \quad Luego \quad P(X_1 = n_1, X_2 = n_2, \ldots, X_k = n_k) =$$

$$\frac{n!}{n_1!\,n_2!\ldots n_k!} p_1^{n_1} \cdot p_2^{n_2} \ldots p_k^{n_k}$$

En ella hay que observar:

a) Depende de k parámetros n, p_1, p_2, …, p_{k-1}.

b) Si el experimento consiste en extracciones de una urna, estas han de ser con reemplazamiento para mantener las probabilidades de los sucesos A_i constantes a lo largo de todas las pruebas.

[Volver al principio]

Distribución de Poisson.

Una variable aleatoria X sigue una distribución de Poisson si puede tomat valores enteros 0, 1, 2, …, n, …, con probabilidades:

$$P(X=k)=\frac{\lambda^k}{k!}.e^{-\lambda}\begin{cases}\lambda>0\\ k=0,1,2,\ldots\\ e=2{,}71828\ldots\end{cases}$$

En esta distribución hay que observar:

a) Depende de un solo parámetro.

b) Su media, varianza son

$$\mu=\lambda,\quad \sigma^2=\lambda$$

c) Es una buena aproximación de la binomial cuando n grande y p pequeño:

$$Si\begin{cases}n\to\infty\\ p\to 0\\ np\to\lambda(finito)\end{cases}\quad \text{entonces}\quad \binom{n}{k}p^k q^{n-k}\to\frac{\lambda^k}{k!}.e^{-\lambda}$$

En general si n>50 y p<0.1 ó n.p<5 la distribución de Poisson es una buena aproximación de la Binomial.

d) La distribución de Poisson presenta una ligera asimetría hacia la izquierda. Cuando n tiende hacia infinito tiende a ser simétrica (distribución normal).

[Volver al principio]

Distribución Hipergeométrica.

Consideremos una población de N elementos de dos clases A y A* excluyentes, de los cuales n_A son de la clase A y n_{A^*} son de la Clase A*, con $n_A + n_{A^*} = N$.

Al tomar un elemento de esta población la probabilidad de que proceda de una u otra clase es:

$$p(A) = \frac{n_A}{N} = p \Rightarrow n_A = N.p$$

$$p(A^*) = \frac{n_{A^*}}{N} = q \Rightarrow n_{A^*} = N.q,\ siendo\ q = 1 - p$$

Sea ahora el experimento de tomar n elementos consecutivos de una población **sin reemplazamiento.** A la variable aleatoria X= " número de elementos de la clase A en la muestra de tamaño n" se le denomina *variable hipergeométrica*.

Para hallar su función de probabilidad P(X=k) consideremos uno de los sucesos favorables:

$$AA\overset{k}{\ldots}AA^*A^*\overset{n-k}{\ldots}A^* = \begin{cases} \Pr ob. = \dfrac{n_A}{N}\dfrac{n_A - 1}{N-1}\cdots\dfrac{n_A-(k-1)}{N-(k-1)}\dfrac{n_{A^*}}{N-k}\dfrac{n_{A^*}-1}{N-k-1}\cdots\dfrac{n_{A^*}-n+k+1}{N-n+1} \\ número\ de\ sucesos \quad = \dfrac{n!}{k!\,(n-k)!} \end{cases}$$

Luego

$$P(X=k) = \frac{\binom{n_A}{k}\binom{n_{A^*}}{n-k}}{\binom{N}{n}} = \frac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}}$$

En ella hay que observar

a) Depende de los parámetros N, n, p

b) Su media y varianza son:

$$\mu = np, \quad \sigma^2 = npq.\frac{N-n}{N-1}$$

c) Se diferencia de la binomial en que, en aquella las probabilidades son constantes a lo largo de las pruebas(extracciones con reemplazamiento) mientras que en la hipergeométrica, varían de una prueba a otra (extracciones sin reemplazamiento).

d) Si N es grande respecto de n, las probabilidades varían muy poco de una prueba a la siguiente, por lo que en estos casos se puede decir que la variable hipergeométrica sigue aproximadamente una distribución binomial, es decir:

$$P(X=k) = \frac{\binom{n_A}{k}\binom{n_{A^*}}{n-k}}{\binom{N}{n}} = \frac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}} \to \binom{n}{k}p^k q^{n-k}$$

Se conviene en sustituirla cuando $n/N<0.1$

Por otra parte y de manera análoga a la distribución polinomial, si tenemos una población con N elementos repartidos en k clases excluyentes $A_1, A_2, \ldots, A_k$ con N_1 elementos de la clase A_1, N_2 elementos de la clase A_2, …, N_k elementos de la clase A_k siendo $N_1+N_2+\ldots+N_k = N$, al tomar consecutivamente n elementos sin reemplazamiento y llamando X_i = "Número de elementos que hay de la clase A_i en la muestra de tamaño n" la variable k-dimensional $(X_1, X_2, \ldots, X_k)$ tiene por función de probabilidad:

$$P(X_1 = n_1; X_2 = n_2; \ldots; X_k = n_k) = \frac{\binom{N_1}{n_1}\binom{N_2}{n_2}\cdots\binom{N_k}{n_k}}{\binom{N}{n}} \quad \text{siendo } N_1 + N_2 + \ldots + N_k = N$$

$$y \quad n_1 + n_2 + \ldots + n_k = n$$

[Volver al principio]

Distribución geométrica o de Pascal

Sea el experimento consistente en la realización sucesiva de pruebas de Bernoulli. A la variable X= "número de la prueba en la que aparece por primera vez el suceso A" se le denomina *variable geométrica*. Ara hallar la probabilidad de P(X=k), hay que notar que esta probabilidad es la del suceso $A^*A^*\overset{k-1}{\ldots}A^*A$ luego $P(X=k)=q^{k-1}.p$

a) Esta distribución depende de un solo parámetro p.

b) Su media y varianza son

$$\mu = \frac{1}{p}, \quad \sigma^2 = \frac{q}{p^2}$$

c) Si el experimento consiste en extracciones de una urna estas han de ser con reeemplazamiento.

[Volver al principio]

Distribución Binomial negativa.

Sea X = "número de pruebas en que aparece el suceso A* hasta la enésima aparición del suceso A".

Para hallar la probabilidad del suceso P(X=k), consideremos uno de los sucesos favorables:

$$AA\overset{n-1}{\ldots}AA*A*\overset{k}{\ldots}A*A\begin{cases} \Pr ob. = p^n q^k \\ núm.\,de\ sucesos = \dfrac{(n+k-1)!}{k!\,(n-1)!} = \dbinom{n+k-1}{k} \end{cases}$$

luego:

$$P(X = k) = \binom{n+k-1}{k} p^n q^k$$

En ella hay que observar:

a) Depende de dos parámetros n y p.

b) Su media y varianza son:

$$\mu = \frac{nq}{p} \qquad \sigma^2 = \frac{nq}{p^2}$$

c) Si el experimento consiste en extracciones de una urna, estas han de ser con reemplazamiento.

[Volver al principio]