**TEMA 3. Algunos modelos de probabilidad de tipo discreto**

En este capítulo se abordan «familias» muy específicas de probabilidad, que con cierta frecuencia se nos presentan en el mundo real. Van a ser distribuciones de tipo discreto y por lo ya estudiado hasta ahora, estaremos en condiciones de calcular sus características más representativas, como son: la media y la varianza.

Destacaremos por su especial interés, dos «tipos» de estas distribuciones: Las distribuciones tipo Binomial, y las distribuciones tipo Poisson.

Recordemos que al ser distribuciones de tipo discreto, no tiene sentido hablar en ellas de la función de densidad de probabilidad. Se expresarán a través de su función de cuantía, o bien de manera agregada a través de su función de distribución.

## 3.1 Al finalizar el tema el alumno debe conocer........

- ✓ Las condiciones bajo las cuales una variable aleatoria tiene una distribución de probabilidad de Bernoulli, binomial o Poisson.
- ✓ La expresión matemática y características de estas distribuciones de probabilidad .
- ✓ Distribución de Poisson como limite de la distribución binomialEl. Las medidas de posición, dispersión y forma de una variable aleatoria.

## 3.2 .Introducción.

Hasta ahora hemos estudiado, en términos generales, las variables aleatorias y sus distribuciones de probabilidad. Ahora vamos a describir familias específicas de distribuciones de probabilidad que ocurren con cierta frecuencia en el mundo real.

A la hora de estudiar estas distribuciones estudiaremos  distribuciones teóricas tanto de variable discreta, como  en el siguiente tema de variable continua. Para un adecuado estudio de las mismas es necesario conocer sus parámetros, esto es, las cantidades que son constantes para distribuciones particulares pero que pueden tomar diferentes valores para diferentes miembros de las familias de distribuciones de la misma clase, y una serie de funciones matemáticas tales como:

- $P[x]$ función de probabilidad.
- $f(x)$ función de densidad.
- $F(x)$ función de distribución.

Los parámetros más comunes son $\mu$ y $\sigma^2$, que como vimos eran la media y varianza de las distribuciones de probabilidad.

**Modelos de distribuciones de probabilidad de variables discretas**

1. Uniforme. Es la distribución donde todos resultados posibles tienen la misma probabilidad
2. Binomial. Es la que maneja la distribución de la probabilidad de obtener cierta cantidad de éxitos al repetir un experimento, con probabilidad de éxito constante y con resultados posibles del experimento independientes.
3. Geométrica. Es la distribución de la probabilidad de realizar cierto número de repeticiones de un experimento, antes de obtener el primer éxito.
4. Hipergeométrica. Es similar a la binomial, pero con un tamaño de muestra grande en relación al tamaño de la población.
5. Poisson. Es la distribución de la probabilidad de que ocurra un evento raro en un periodo de tiempo, un espacio o un lugar.

## 3.3 Distribución de Bernoulli.

Este modelo se utiliza principalmente en situaciones en las que sólo pueden ocurrir dos resultados posibles mutuamente excluyentes: uno de ellos de probabilidad $p$ y el otro de probabilidad $1-p$. Son ejemplos: El sexo de una persona: hombre o mujer, la respuesta correcta o incorrecta de un examen, etc.

Supongamos un experimento consistente en observar elementos de una población, con las siguientes características:

- Mediante esta observación los elementos sólo pueden clasificarse en dos categorías, que generalmente se llamaran "éxito" al suceso de probabilidad $p$ y "fracaso" al suceso de probabilidad $1-p$ (también llamado $q$).
- La proporción de sucesos "éxito" y "fracaso" en la población es constante y no se modifica cualquiera que sea la cantidad observada. Esto implica que los

elementos se reemplazan una vez observados en la población.

- Las observaciones son independientes, la probabilidad de "éxito" es siempre la misma no se modifica.

| | | |
|---|---|---|
| **Experimento da lugar a dos posibles resultados:** | **Éxito** | **p** |
| | **Fracaso** | **(1 - p ) = q** |

Diremos que una variable aleatoria $X$ sigue una distribución de Bernoulli de parámetro $p$ si su distribución de probabilidad está dada por:

$$P(x) = P(X = x) = p^x (1-p)^{1-x} \quad para \;\; x = 0, 1, 2, \cdots$$

La variable aleatoria $x$ quedará definida como:

$$x = \begin{cases} 0 & si \quad "fracaso" \\ 1 & si \quad "éxito" \end{cases}$$

Abreviadamente esta distribución la indicaremos por: $x \to B(p)$ o bien $x \to B(1, p)$

Características de ésta distribución:

1. Función de distribución:

$$F(x) = P(X \le x) = \begin{cases} 0 & si \quad x < 0 \\ 1-p & si \quad 0 \le x < 1 \\ 1 & si \quad x \ge 1 \end{cases}$$

2. Media :

   Se calcula como vimos anteriormente en el valor esperado de una variable aleatoria discreta; $E[x] = p$

3. Varianza:

   Se calcula como vimos anteriormente en el caso de una variable aleatoria discreta; $\mu_2 = E[x^2] - (E[x])^2 = \alpha_2 - (\alpha_1)^2 = Var[x] = p(1-p) = pq$

   Los parámetros de esta distribución son:

| Distribución Bernoulli | Parámetros |
|---|---|
| Media | $p$ |
| Varianza | $pq$ |
| Desviación típica | $\sqrt{pq}$ |

### 3.4 Distribución binomial.

La repetición de un experimento juega un papel muy importante en probabilidad, y en estadística. Una generalización de la distribución de Bernoulli se obtiene cuando el experimento o prueba de Bernoulli se repite varias veces. Por tanto, estamos ante un experimento binomial cuando repetimos "n" veces de forma independiente un ensayo de Bernoulli. Conceptualmente la distribución $B(n, p)$ describe situaciones que se pueden presentar si un mismo suceso dicotómico se observa o se repite n-veces y si los posibles resultados en cada ocasión son independientes de los que puedan lograrse en las demás.

Definimos la variable aleatoria binomial $x$, como el número de éxitos que tienen lugar cuando se realizan n repeticiones independientes de un experimento o prueba de Bernoulli.

Para obtener su función de probabilidad, consideramos que al realizar n repeticiones independientes del experimento hemos obtenido $x$ resultados de éxito (con probabilidad $p$ ) y $n-x$ resultados de fracaso (con probabilidad q) :

$$\underbrace{p\ p \ldots p}_{x} \quad \underbrace{q\, q \cdots q}_{n-x} = p^x\ q^{n-x} = p^x (1-p)^{n-x}$$

La probabilidad de $x$ elementos de éxito en cualquier orden, requiere sumar las probabilidades de todos los sucesos mutuamente excluyentes que verifican esta condición. Estos sucesos se obtienen permutando las letras anteriores de todas las posibles formas:

$$\frac{n!}{x!\,(n-x)!} = \binom{n}{x}$$

Diremos que una variable aleatoria $X$ sigue una distribución de binomial de parámetros $n$ y $p$ si su distribución de probabilidad está dada por:

$$P(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \quad para \quad x = 0, 1, 2 \cdots n$$

Abreviadamente esta distribución la indicaremos por: $x \to B(n, p)$ o bien $x \to B(x, n, p)$.

Características de ésta distribución:

1. Función de distribución:

$$F(x) = P(X \le x) = \begin{cases} 0 \quad si \quad x < 0 \\ \sum_{i=0}^{x} \binom{n}{i} \; p^i \; (1-p)^{n-i} \quad si \quad 0 \le x < n \\ 1 \quad si \quad x \ge n \end{cases}$$

1. Media :

   Para obtener la media y la varianza de la distribución binomial tenemos que tener en cuenta que la variable aleatoria binomial está definida como el número de éxitos que tienen lugar cuando se realizan $n$ repeticiones independientes de un experimento o prueba de Bernoulli, es decir, como suma de $n$ variables independientes: $x = x_x + x_2 + \cdots + x_n$

   Si utilizamos las propiedades de la esperanza:

$$E[x] = E[x_1 + x_2 + \cdots + x_n] = E[x_1] + E[x_2] + \cdots + E[x_n] = p + \cdots + p = np$$

2. Varianza:

$$Var[x] = Var[x_1 + x_2 + \cdots + x_n] = Var[x_1] + Var[x_2] + \cdots + Var[x_n] =$$

$$= pq + \cdots pq = npq$$

3. Propiedad reproductiva:

Si $x_1$ y $x_2$ son dos variables aleatorias independientes distribuidas $x_1 \to B(n_1, p)$ y $x_2 \to B(n_2, p)$, entonces la variable aleatoria: $x = x_1 + x_2$ se distribuye según una $B(n_1 + n_2, p)$.

Los parámetros básicos de la distribución son:

| Distribución Binomial | Parámetros |
|---|---|
| Media | $np$ |
| Varianza | $npq$ |
| Desviación típica | $\sqrt{n\ p\ q}$ |

### 3.5 Distribución de Poisson.

Consideremos un experimento en el que observamos la aparición de sucesos puntuales sobre un soporte continuo; el número de clientes que llegan a una sucursal bancaria durante cinco minutos, el número de averías que sufre una máquina durante un año, el número de órdenes de devolución de piezas que recibe una empresa en una semana, etc. Suponemos que el proceso se caracteriza por:

- Es estable, produce a largo plazo un número medio de sucesos constante $\lambda$ por unidad de tiempo o espacio o área….
- Los sucesos aparecen aleatoriamente de forma independiente, es decir, el proceso no tiene memoria: conocer el número de sucesos en un intervalo no ayuda a predecir el número de sucesos en el siguiente.

Definimos una variable aleatoria de Poisson como el número de sucesos en un intervalo de longitud fija. En este caso, la distribución de probabilidad de la variable aleatoria de Poisson depende solamente del número medio $\lambda$ de resultados que ocurren en un intervalo.

Diremos que la variable aleatoria $X$ sigue una distribución de Poisson de parámetro $\lambda(\lambda > 0)$si su distribución de probabilidad es:

$$P(x) = P(X = x) = \frac{\lambda^{x}}{x!} e^{-\lambda} \quad si \quad x = 0, 1, \cdots$$

Abreviadamente esta distribución la indicaremos por: $x \to P(\lambda)$

Características de ésta distribución:

1. Función de distribución:

$$F(x) = P(X \le x) = \begin{cases} 0 \quad si \quad x < 0 \\ \sum_{i=0}^{x} \frac{\lambda^{i}}{i!} e^{-\lambda} \quad x = 0, 1, 2.. \end{cases}$$

2. Media :

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

3. Varianza:

$E[x^2]=\sum_{x=0}^{\infty} x^2 P(X=x)$, si realizamos este calculo llegamos a $E[x^2]=\lambda^2+\lambda$

Y como sabemos que: $Var[x]=E[x^2]-(E[x])^2=\lambda^2+\lambda-\lambda^2=\lambda$

Vemos por tanto que en la distribución de Poisson la media y la varianza valen lo mismo.

4. Propiedad reproductiva:

Si $x_1$ y $x_2$ son dos variables aleatorias independientes distribuidas $x_1 \to P(\lambda_1)$ y $x_2 \to P(\lambda_2)$, entonces la variable aleatoria: $x=x_1+x_2$ se distribuye según una $P(\lambda_1+\lambda_2)$.

Los parámetros básicos de la distribución son:

| **Distribución Poisson** | **Parámetros** |
|---|---|
| Media | $\lambda$ |
| Varianza | $\lambda$ |
| Desviación típica | $\sqrt{\lambda}$ |

**Distribución de Poisson como límite de la Binomial.**

La distribución de Poisson se obtiene como limite de la distribución binomial cuando $n \to \infty$, de manera que podemos considerar un soporte continuo, y $p \to 0$, para que el número medio de sucesos, $np$, permanezca constante. En este caso $np=\lambda$, es decir el número de éxitos es una variable aleatoria que tiene una distribución de Poisson con el parámetro $\lambda$.

En la práctica la distribución binomial se puede aproximar a una de Poisson cuando

$n < 30$ $y$ $p \leq 0{,}1$ . Por tanto, se utilizará la distribución de Poisson en situaciones reales donde la probabilidad de éxito del suceso es muy pequeña y donde se producen un gran número de repeticiones.