TEMA 1

VARIABLE ALEATORIA DISCRETA

1.1. INTRODUCCIÓN

En la asignatura Estadística I hemos desarrollado los conceptos básicos de probabilidad sobre los resultados o sucesos de un experimento aleatorio. Pero los experimentos aleatorios son tales que los resultados a que dan lugar pueden ser de naturaleza cualitativa o cuantitativa. Así, por ejemplo, serían resultados cualitativos los derivados de los siguientes experimentos aleatorios: la calidad de las piezas fabricadas en una planta: buenas o defectuosas; o la preferencia de una persona sobre tres tipos de coches… Ejemplos de experimentos aleatorios cuyos resultados son cuantitativos son: el número de clientes que llegan a un comercio durante una hora; o el número de accidentes de automóvil en una ciudad en un mes dado.

1.2. CONCEPTO DE VARIABLE ALEATORIA

Pero trabajar con los resultados cualitativos de un experimento aleatorio introduce ciertas complicaciones, siendo de gran utilidad cuantificarlos, o lo que es lo mismo, asignar un valor numérico a cada suceso del espacio muestral correspondiente al experimento aleatorio considerado. Esta relación entre los sucesos del espacio muestral y el valor numérico que se les asigna la establecemos mediante la ***Variable Aleatoria***. Las variables aleatorias nos brindan así la oportunidad de trasladar los problemas relacionados con el espacio muestral a la recta real, con todas las ventajas que esto conlleva. Para ello necesitamos recordar el concepto de Espacio Probabilístico (visto en Estadística I):

Consideramos un experimento aleatorio y definimos **Espacio Probabilístico** como la terna $\mathbf{(\Omega, \wp(\Omega), P)}$ donde:

- Ω es el **Espacio muestral** correspondiente al experimento aleatorio
- $\wp(\Omega)$ es el **Conjunto de todos los posibles subconjuntos** del espacio muestral
- P es la **Función de probabilidad** asociada al experimento

Formalmente, sea Ω el espacio muestral de un experimento, una aplicación:

$$X: \Omega \to R$$
$$\omega \to X(\omega)$$

que a cada suceso elemental le asigna un número real, se dice que es una variable aleatoria, si para todo número real x, el conjunto $A = \{\omega \in \Omega: \ X(\omega) \leq x\}$ es un suceso, es decir, pertenece a $\wp(\Omega)$.

Ejemplo 1.1

1) X = Puntuación obtenida al lanzar un dado con 6 caras numeradas de 1 a 6
2) X = Número de clientes esperando a ser atendidos en un banco a las 11 de la mañana
3) X = Ventas totales realizadas por una empresa en un año
4) X = Hora a la que llega un avión a un aeropuerto

1.3 VARIABLES ALEATORIAS DISCRETAS Y CONTINUAS

Una variable aleatoria se dice ***Discreta*** si toma un conjunto de valores finito o infinito numerable, es decir, los valores se pueden contar con los números naturales.

Así pues, sería una *variable aleatoria de tipo discreto* el número de piezas defectuosas que aparecen en un proceso de fabricación, el número de llamadas telefónicas que son recibidas en una centralita durante un período de tiempo, el número de depósitos de una entidad bancaria, el número de trabajadores afectados por cierto convenio laboral, o el número de errores detectados en las cuentas de una compañía.

Una variable aleatoria es ***Continua*** cuando puede asumir una infinidad (un numero infinito no numerable) de valores, es decir, puede tomar cualquier valor en uno o más intervalos de la recta real.

Ejemplos de *variables aleatorias continuas* son la Renta Nacional de cierto país, el nivel de inflación acumulada en un mes, la cantidad de petróleo importado por Estados Unidos en un año, la variación en el precio de las acciones de IBM en un mes, el tiempo transcurrido desde la instalación de un nuevo componente hasta que falla o el tiempo que dedica un alumno a hacer un examen cuya duración máxima es de dos horas.

El conjunto de valores que puede tomar una variable aleatoria junto con sus respectivas probabilidades constituyen la denominada ***Distribución de Probabilidad*** de dicha variable aleatoria. Ésta da la máxima información posible de una variable aleatoria porque describe todo el fenómeno aleatorio y muestra como se distribuye o reparte la probabilidad en el espacio.

1.4. DISTRIBUCIONES DISCRETAS

1.4.1. FUNCIÓN DE PROBABILIDAD

Para una variable aleatoria X discreta su distribución de probabilidad se denomina **función de probabilidad o función de cuantía** y la definimos como:

$p_i = P(X = x_i) \geq 0$, $i = 1,.., \infty$ siendo su soporte $D = \{x_1, x_2,\ldots,x_n,\ldots\}$, es decir, el conjunto de valores que puede tomar dicha variable. Esta función verifica que la suma de todas las probabilidades es igual a la unidad:

$$\sum_{i=1}^{\infty} p_i = 1$$

Igual que ocurría con las distribuciones de conjuntos de datos, resulta útil visualizar las funciones de probabilidad de las variables aleatorias discretas. El gráfico más habitual es el diagrama de barras donde se representan los valores de la variable en el eje horizontal y las probabilidades mediante alturas en el eje vertical.

Ejemplo 1.2.

Sea X = puntuación obtenida al lanzar un dado con 6 caras numeradas de 1 a 6

Si el dado está perfectamente construido entonces X es una variable aleatoria discreta que puede tomar los valores {1, 2, 3, 4, 5, 6} con probabilidades

$$p_i = P(X=i) = \frac{1}{6} \text{ para } i=1, \ldots, 6$$

Estos valores son la función de probabilidad de la variable aleatoria X.

1.4.2. FUNCIÓN DE DISTRIBUCIÓN

Asociada a cada distribución de probabilidad, ya sea de tipo discreto o continuo, tenemos una ***Función de Distribución***, en ocasiones también denominada, función de distribución acumulada. La función de distribución en un punto es la probabilidad de que la variable aleatoria tome un valor numérico igual o inferior a dicho punto.

La función de distribución de una variable aleatoria discreta es una función en forma de escalera, con saltos en los puntos de la recta real en los que hay asignada una probabilidad positiva. La magnitud del salto es igual a la cantidad de probabilidad en dichos puntos.

Es decir, a partir de la función de probabilidad construimos la **función de distribución** que viene dada por:

$$F(x) = P(X \le x)$$

En el caso de que la variable aleatoria X sea discreta: $F(x) = \sum_{x_j \le x} P(X = x_j) = \sum_{x_j \le x} p_j$

Se puede demostrar que la función de distribución verifica las siguientes propiedades:

1. $0 \le F(x) \le 1, \ \forall \ x \in R$
2. $\lim_{x \to -\infty} F(x) = 0$ y $\lim_{x \to \infty} F(x) = 1$
3. F(x) es monótona no decreciente
4. F(x) es continua por la derecha
5. $P(a < X \le b) = F(b) - F(a)$ (probabilidad de un intervalo)

Ejemplo 1.2 (continuación)

La función de distribución de la variable X viene dada por:

$$F(x) = \begin{cases} 0 & x < 1 \\ \dfrac{i}{6} & i \le x < i+1; i = 1,...,5 \\ 1 & x \ge 6 \end{cases}$$

Es, por lo tanto, una función constante a trozos con saltos en los números del 1 al 6 iguales a su probabilidad que es 1/6.

1.4.3. CARACTERÍSTICAS DE UNA DISTRIBUCIÓN: ESPERANZA MATEMÁTICA

La distribución de probabilidad de una variable aleatoria es un modelo para la distribución teórica de la variable. Por lo tanto, es de esperar, que cada distribución de probabilidad tenga asociadas medidas similares a las medidas descriptivas de las distribuciones de frecuencias. Y así, análogamente a como se hizo para una distribución de frecuencias, podemos construir medidas características o resumen del comportamiento de una variable aleatoria.

Estas medidas se basan en el concepto de ***Esperanza Matemática***. Sus orígenes se encuentran en los juegos de azar. Los apostadores deseaban saber cuál era su esperanza de ganar repetidamente un juego. En este sentido, el ***Valor Esperado*** representa la cantidad de dinero promedio que un jugador estaría dispuesto a ganar o perder después de un número muy grande de apuestas. Este significado también es válido para una variable aleatoria. Es decir, el valor promedio de una variable aleatoria después de un número grande de experimentos, es su valor esperado. La esperanza puede ser interpretada como «***centro de gravedad***» de una distribución de probabilidad, ya que, si asumiésemos dicho valor como único representante de la población, el error esperado sería nulo.

A la esperanza matemática de una variable aleatoria se le llama también media o valor esperado de la distribución de probabilidades de la variable. Se denota por:

$$\mathbf{E(x) = \mu}$$

Consideremos un experimento aleatorio de espacio muestral Ω, siendo X una variable aleatoria discreta con soporte D = $\{x_1, x_2, \ldots, x_n, \ldots\}$ y cuya función de probabilidad viene dada por $p_i = P(X = x_i) \geq 0$, $i = 1, .., \infty$. Se define la esperanza matemática o valor esperado de una v.a. discreta X al número:

$$E(X) = \sum_{i=1}^{\infty} x_i P(X = x_i) = \sum_{i=1}^{\infty} x_i p_i$$

Ejemplo 1.2 (continuación)

En este caso se tiene que E(X) = $1x\frac{1}{6} + 2x\frac{1}{6} + \ldots + 6x\frac{1}{6} = \frac{21}{6} = 3{,}5$

Podemos generalizar la definición de la esperanza de X a variables aleatorias que se obtienen de ella, es decir, a transformaciones g(X). Así, si X es una variable aleatoria discreta con soporte D = $\{x_1, x_2, \ldots, x_n, \ldots\}$, cuya función de probabilidad viene dada por $p_i = P(X = x_i) \geq 0$, $i = 1, .., \infty$ y g(X) es otra variable aleatoria, transformación de X, la esperanza matemática de la nueva variable aleatoria g(X) se define como:

$$E(g(X)) = \sum_{i=1}^{\infty} g(x_i)P(X = x_i) = \sum_{i=1}^{\infty} g(x_i)p_i$$

Ejemplo 1.2 (continuación)

En este caso se tiene que $E(X^2) = 1^2 x \frac{1}{6} + 2^2 x \frac{1}{6} + ... + 6^2 x \frac{1}{6} = \frac{91}{6} = 15{,}17$

Algunas propiedades interesantes de la esperanza matemática, que provienen de su definición como suma, son las siguientes:

1. La esperanza de una variable aleatoria constante es la propia constante E(a)=a.
2. $E(aX) = aE(X)$
3. $E(aX + b) = aE(X) + b$

Dadas dos variables aleatorias X e Y, se cumple, además:

4. $E(X + Y) = E(X) + E(Y)$
5. $E(X – Y) = E(X) - E(Y)$

Es importante complementar la información que proporciona la media sobre el valor esperado de la variable con una medida de la dispersión de los resultados del experimento alrededor de dicha media.

Así, dada la variable aleatoria X llamamos varianza de X y la denotaremos por Var(X) o con σ^2 a la siguiente expresión:

$$Var(X) = E((X-\mu)^2)=\sigma^2$$

Así, si la variable aleatoria X es discreta con valores $\{x_1, x_2,\ldots,x_n,\ldots\}$ y cuya función de probabilidad viene dada por $p_i = P(X = x_i) \geq 0,\ i = 1,.., \infty$ entonces la varianza viene dada por:

$$Var(X) = \sum_{i=1}^{\infty} (x_i - \mu)^2 P(X = x_i) = \sum_{i=1}^{\infty} (x_i - \mu)^2 p_i$$

Algunas propiedades interesantes de la varianza son las siguientes:

1. $Var(X) = E(X^2) – (E(X))^2$
2. La varianza de una variable aleatoria X es nula si y sólo si la variable aleatoria es constante.
3. $Var(X + k) = Var (X)$
4. $Var(kX) = k^2Var(X)$

5. $\text{Var}(aX + b) = a^2\text{Var}(X)$

Dadas dos variables aleatorias X e Y, que sean **independientes**, se cumple, además:

6. $\text{Var}(X + Y) = \text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y)$

También incluimos el concepto de desviación típica que es la raíz cuadrada positiva de la varianza. Se denota por σ y se define como:

$$\sigma = +\sqrt{Var(x)}$$

Ejemplo 1.2 (continuación)

La varianza de X vendrá dada por $\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{91}{6} - 3{,}5^2 = \frac{35}{12}$,

por lo tanto, la desviación típica será $D(X) = \sigma = \sqrt{\frac{35}{12}} = 1{,}71$

1.5. MODELOS NOTABLES

Una de las preocupaciones de los científicos dedicados al Cálculo de Probabilidades ha sido construir modelos teóricos de distribuciones de probabilidad que pudieran representar el comportamiento práctico de diferentes fenómenos aleatorios que aparecían en el mundo real.

A lo largo de este tema, y del siguiente, analizaremos los modelos de distribución de probabilidad que subyacen, más frecuentemente, en los fenómenos aleatorios que podemos tratar, comenzando por aquéllos donde la variable aleatoria descrita sea de tipo discreto para, en el tema siguiente mostrar los modelos continuos de probabilidad.

El esquema que seguiremos en cada modelo procurará contener los elementos básicos y necesarios, para el mejor conocimiento de cada distribución. Haremos una descripción del fenómeno que genera cada variable aleatoria, determinando su función de probabilidad y las características básicas que le correspondan (media, varianza, etc.), propondremos ejemplos de aplicación de cada modelo y cuando sea necesario, expondremos alguna propiedad importante.

Como podrá comprobarse, en cada modelo existirán un conjunto de valores inherentes a cada variable que se denominan parámetros del modelo. Dichos parámetros identifican a cada modelo, de forma que, al asignarles valores numéricos concretos, será posible calcular probabilidades asociadas a cada suceso que nos interese.

1.5.1. DISTRIBUCIÓN UNIFORME DISCRETA

Existen ciertas magnitudes aleatorias en las que no aparece ninguna evidencia a favor de determinados resultados, por lo cual resulta aplicable el principio de indiferencia. Este sería el caso cuando lanzamos un dado, extraemos una bola de un bombo de lotería, seleccionamos al azar una carta de la baraja o elegimos al azar a una persona de un grupo. Estos ejemplos, conducen al *Modelo Uniforme Discreto*, cuya distribución se corresponde con un reparto equitativo de la probabilidad. Supongamos que un empresario convocado a una entrevista se dispone a acudir a la entrevista utilizando el metro. Consultando el plano, observa que puede optar entre tres líneas alternativas que le conducen hasta la sede del periódico. Por tanto, la variable: Línea de metro elegida, al azar, por el empresario sigue una distribución uniforme, donde se asigna a cada una de las tres líneas la misma probabilidad 1/3.

Una **Distribución Uniforme Discreta** describe el comportamiento de una variable aleatoria X que puede tomar los valores 1 a N con la misma probabilidad cada uno de ellos. X es, por lo tanto, una variable aleatoria discreta con soporte D = {1,…,N} y con función de probabilidad dada por:

$$P(X = i) = \frac{1}{N} \qquad i = 1,..,N$$

La distribución de X recibe el nombre de *distribución uniforme discreta* y la denotaremos como $X \sim U_D(N)$. Tiene un único parámetro que es N e indica el número de resultados posibles, que tendremos que particularizar en cada situación estudiada.
La media y la varianza de esta distribución vienen dadas por:

$$E(X) = \frac{1}{N}\sum_{i=1}^{N} i = \frac{N+1}{2}$$

$$Var(X) = \frac{1}{N}\sum_{i=1}^{N} i^2 - E(X)^2 = \frac{N^2-1}{12}$$

Ejemplo 1.3.

El número de coches que vende diariamente un concesionario oscila entre 21 y 40. Determina la probabilidad de que un día cualquiera venda más de 35 coches. Calcula, además, su media y su varianza.

El número de coches vendidos es 20+X, siendo X un valor aleatorio entre 1 y 20, cuya distribución es uniforme discreta.

En este caso, el parámetro es N = 20 y la variable aleatoria $X \sim U_D(20)$. Nos piden

$$P(X+20>35)=P(X>15) = P(X\geq 16) = \frac{5}{20} = 0{,}25$$

La media y la varianza vendrán dadas por

$E(X) = \frac{20+1}{2} = 10{,}5$ → El número medio de coches vendidos es 20+10,5=30,5 coches

$Var(X) = \frac{20^2-1}{12} = \frac{399}{12} = 33{,}25$ → La varianza del nº de coches vendidos es 33,25.

1.5.2 DISTRIBUCIONES ASOCIADAS AL EXPERIMENTO DE BERNOULLI

Un *experimento de Bernoulli* es un experimento aleatorio que tiene sólo dos resultados posibles, que denominaremos «*éxito*» y «*fracaso*», los cuales son mutuamente excluyentes y conjuntamente exhaustivos, y que es susceptible de ser replicado indefinidamente de forma independiente de forma que las probabilidades de obtener un éxito, p ($0<p<1$), y un fracaso, $q = 1-p$ se mantienen constantes en todas las replicaciones.

Es frecuente hallar en el mundo real fenómenos dicotómicos, por ejemplo, en una tirada de una moneda sacar cara o cruz, en un dado obtener par o impar, que una pieza fabricada sea defectuosa o no, que un ciudadano vote a un candidato o no, que una factura sea correcta o incorrecta, que un domicilio tenga teléfono o no, etc. Como puede comprobarse son muy variados los casos en que nuestras observaciones pueden proceder de un mundo dicotómico que permita el uso de las siguientes distribuciones.

1.5.2.1 DISTRIBUCIÓN DE BERNOULLI

La distribución de Bernoulli corresponde a la variable aleatoria X = número de éxitos obtenidos al realizar un experimento de Bernoulli una sola vez. En este caso estamos ante una variable aleatoria discreta cuyo soporte es D = {0,1} y cuya función de probabilidad puede expresarse como:

$$P(X = x) = p^x(1 - p)^{1-x} \quad \text{para } x = 0 \text{ y } x = 1$$

Diremos que *X tiene una distribución de Bernoulli de parámetro p* y lo pondremos como X ~ Be(p). La media y la varianza de esta distribución son:

$$E(X) = p \quad \text{y} \quad V(x) = p(1 - p)$$

Ejemplo 1.4

Se extrae una persona al azar y se define que X = 0 si es hombre y X = 1 si es mujer. X ~ Be(p) donde p es la proporción (en tanto por uno) de mujeres que hay en la población objeto de estudio.

1.5.2.2 DISTRIBUCIÓN BINOMIAL

Si bien el modelo anterior permite describir gran cantidad de fenómenos, en general las investigaciones no se limitarán a una prueba única, siendo frecuentes en el ámbito económico los estudios donde se observa cierto resultado en una serie de pruebas repetidas.

Por ejemplo, si tenemos un total de 20 empresarios convocados para una entrevista, cada uno puede decidir si acudirá o no a ella. Como consecuencia de la repetición de la experiencia, definiremos ahora una variable aleatoria que designe el número de éxitos obtenidos a lo largo de la investigación (en nuestro ejemplo «*número de empresarios que acceden a la entrevista*»). El rasgo distintivo de esta variable con respecto al modelo Bernoulli es la presencia de un nuevo parámetro que indica el número de observaciones llevadas a cabo.

La distribución binomial surge, por lo tanto, al repetir un experimento de Bernoulli un número dado de veces n, de forma independiente. Sea X = número de éxitos obtenido. X es una variable aleatoria discreta cuyo soporte es D = {0, 1, … n} y cuya función de probabilidad viene dada por:

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} \qquad x = 0,1,...,n$$

La distribución de X recibe el nombre de *distribución binomial de parámetros n y p* y lo pondremos como X ~ Bi(n,p). Para calcular las probabilidades acumuladas utilizaremos las tablas estadísticas que nos proporcionan la probabilidad hasta un valor concreto: $P\left(X \leq k\right) = \sum_{i=0}^{k} P\left(X = i\right)$.

La media y la varianza de dicha distribución vienen dadas por.

$$E(X) = np \qquad y \qquad V(x) = np(1-p)$$

El hecho de contar éxitos o fracasos es arbitrario y dependerá del contexto del problema. Por lo tanto, es evidente que la v.a. Y = número de fracasos en las n repeticiones también sigue una distribución binomial cuyos parámetros son n y 1-p: Y~Bi(n, 1-p). Se cumple como es obvio que la suma de éxitos y fracasos es igual al número de repeticiones, es decir, X+Y = n.

Por último, si el experimento se repite una sola vez entonces estamos en una v.a. Bernoulli: Bi(1, p) = Be(p). También parece lógico entender que una distribución Binomial es la suma de los resultados de n v.a. Bernoulli independientes: Bi(n,p)=Be(p)+…Be(p), con n sumandos, cada uno representa un experimento dicotómico que se repite de forma independiente. Por lo tanto, es obvio entender que si sumamos dos v.a. Binomiales independientes es como si el experimento dicotómico se repitiese más veces, es decir: $Bi(n_1,p)+Bi(n_2,p)=B(n_1+n_2,p)$.

La aplicación de la distribución binomial al mundo de la empresa es amplia, por ejemplo se utiliza en situaciones de toma de decisiones, control de calidad, en esquemas de muestreo con reemplazamiento, ventas, mercadotecnia, investigación de opiniones, etc., siempre y cuando pueda suponerse que el proceso se ajusta al modelo teórico.

Ejemplo 1.4 (continuación)

Si ahora extraemos una muestra de 30 personas al azar y contamos cuantas mujeres hay en la muestra con la variable X, entonces X~Bi(30, p) si la población de personas es muy grande.

1.5.3. DISTRIBUCIÓN HIPERGEOMÉTRICA

Esta distribución aparece asociada a la realización de experimentos aleatorios dicotómicos en los que desaparece la condición de independencia. Más concretamente es la distribución asociada al *muestreo sin reemplazamiento en poblaciones finitas*.

Las condiciones en las que se define este modelo de probabilidad son las siguientes: consideramos una población (empresarios, alumnos presentados a un examen, población activa de cierto sector económico,…) de tamaño N sobre los que nos interesa estudiar determinada característica, que podríamos seguir denominando «éxito» (acceder a la entrevista, aprobar el examen, obtener un trabajo, o en general cualquier rasgo distintivo). Supongamos clasificados los integrantes de la población según la característica de interés, de manera que D elementos de la misma presentan el rasgo estudiado y los N-D restantes no lo presentan (D debe ser un número mayor que cero y menor que N). Si de la población total representada seleccionamos aleatoriamente y sin reposición una muestra de n elementos, el número X de ellos que presentan la característica analizada es una variable aleatoria discreta cuya distribución recibe el nombre de distribución Hipergeométrica de parámetros N, D y n y lo expresamos como X ~ H(N, n, D). Su soporte viene dado por D = $\{\max\{0, n-N+D\}, ..., \min\{n, D\}\}$ y su función de probabilidad es:

$$p(x) = \frac{\binom{D}{x}\binom{N-D}{n-x}}{\binom{N}{n}} \quad \text{para } x \in \{\max\{0, n-N+D\}, ..., \min\{n, D\}\}$$

Su media y su varianza vienen dadas por:

$$E(X) = n\frac{D}{N} \quad \text{y } V(X) = n\frac{D}{N}\left(1-\frac{D}{N}\right)\left(\frac{N-n}{N-1}\right)$$

Conviene hacer notar que si el muestreo fuera con reemplazamiento (independencia en las repeticiones) la distribución de X sería una $Bi\left(n, \frac{D}{N}\right)$.

Observar, finalmente, que si N→ ∞ y n<<N, es decir, si la población es muy grande y el tamaño de la muestra es una fracción muy pequeña de la misma, (Habitualmente se

supone N > 50 y n < 10%N) se verifica que H(N, n, D) ≈ $\mathrm{Bi}\left(n, \frac{D}{N}\right)$ y, por lo tanto, el muestreo con y sin reemplazamiento serán prácticamente equivalentes.

Ejemplo 1.4 (continuación)

La población analizada consta de 25 personas de las cuales 20 son mujeres y se extrae una muestra sin reemplazamiento de tamaño 10. Calcula la probabilidad de que 7 sean mujeres.

X = número de mujeres en la muestra ~ H(N = 25, n = 10, D = 20).

$$P(X=7)=\frac{\binom{20}{7}\binom{5}{3}}{\binom{25}{10}}=\frac{\frac{20x19x18x17x16x15x14}{7!}x\frac{5x4}{2}}{\frac{25x24x...x21x20x...x16}{10!}}=\frac{15x14x10}{\frac{25x24x23x22x21}{10x9x8}}=0.2372$$

Si el muestreo es con reemplazamiento, calcula la probabilidad de que 7 sean mujeres.

Si es un muestreo con reemplazamiento, es decir, se puede repetir una misma persona entonces X ~ Bi(n=10,p=20/25) = Bi(10, 0.8)

$$P(X=7)=\binom{10}{7}0.8^7x0.2^3=0.2013$$

1.5.4. DISTRIBUCIÓN DE POISSON

A menudo nos interesa estudiar sucesos que, aunque no resultan frecuentes, pueden presentarse en el transcurso del tiempo o del espacio. Estas situaciones del tipo «número de casas incendiadas en un año», «erratas en la página de un periódico», «llamadas de teléfono equivocadas», «atracos en una sucursal bancaria»,… se adaptan bien al modelo probabilístico denominado de *Poisson* o «*ley de los sucesos raros*». Esta distribución viene caracterizada por un único parámetro que representa el número medio de sucesos por unidad de tiempo o espacio. Como consecuencia, el valor del parámetro cambia según cuál sea la «unidad» adoptada, esto es, en función de la amplitud del intervalo espacial o temporal en el que nos movemos.

La distribución de Poisson aparece asociada al llamado *experimento de Poisson*. Dicho experimento consiste en analizar la ocurrencia de un suceso, que denominaremos éxito, en el transcurso del tiempo o del espacio. Supondremos, en lo que sigue, que dicho soporte es el tiempo y que se cumplen las siguientes hipótesis:

a) Para intervalos de tiempo suficientemente pequeños la probabilidad de que ocurran dos o más éxitos es despreciable y la probabilidad de que ocurra un éxito es de un orden de magnitud menor que la de que no ocurra ninguno

b) Para intervalos de tiempo que no se solapan los resultados obtenidos en cada uno de ellos son independientes

c) El número medio de éxitos es proporcional al intervalo de tiempo considerado y crece a un ritmo constante $\lambda > 0$

En este contexto sea X = número de éxitos ocurridos en 1 unidad de tiempo. X es una variable discreta de soporte D = {0, 1, 2, …., ∞} y cuya función de probabilidad viene dada por:

$$p(x) = e^{-\lambda} \frac{\lambda^x}{x!} \text{ si } x \in D = \{0, 1, 2, \ldots, \infty\}$$

La distribución de X recibe el nombre de *distribución de Poisson de parámetro λ* y lo pondremos como $X \sim P(\lambda)$. Para calcular las probabilidades acumuladas utilizaremos las tablas estadísticas que nos proporcionan la probabilidad hasta un valor concreto: $P(X \leq k) = \sum_{i=0}^{k} P(X = i)$.

Su media y su varianza vienen dadas por:

$$E(X) = \lambda \text{ y } V(X) = \lambda.$$

Algunas propiedades interesantes y que son útiles conocer son las siguientes:

1. Si cambiamos el intervalo de estudio entonces el parámetro se transforma proporcionalmente, es decir, si X = número de ocurrencias en 1 unidad de tiempo y es $X \sim P(\lambda)$ entonces Y = número de ocurrencias en x unidades de tiempo tiene una distribución Poisson de parámetro λx: $Y \sim P(\lambda x)$.
2. La suma de dos v.a. de tipo Poisson independientes sigue otra distribución Poisson cuyo parámetro es la suma de los parámetros de cada sumando: $P(\lambda_1)+P(\lambda_2)=P(\lambda_1+\lambda_2)$.

La distribución de Poisson ha resultado ser muy útil en problemas de *líneas de espera o colas*. Los clientes llegan a una máquina fotocopiadora con una tasa media de dos cada cinco minutos. En la práctica, se pueden representar los procesos de llegada de esta clase mediante una distribución de Poisson. Asumiendo que éste es el caso, la variable que mide el número de llegadas de clientes en un período de cinco mínutos tiene una distribución de Poisson con media 2. Como vemos la distribución de Poisson aparece de manera natural para representar el número de ocurrencias de un suceso en un período de tiempo. El intervalo de tiempo dado puede ser de cualquier magnitud; un minuto, una hora, un día, una semana, etc.; y el experimento de Poisson puede producir observaciones para la variable aleatoria que representen, por ejemplo: el número de llegadas por hora de enfermos al servicio de urgencias de un hospital, el número de huelgas anuales en una fábrica, el número de clientes que llegan a un banco durante cinco minutos, etc.

Ejemplo 1.5

El número de aviones que llegan a un aeropuerto en 1 hora sigue una distribución de Poisson de media 20. Calcula la probabilidad de que en media hora lleguen más de 10 aviones

Si X = número de aviones que llegan en media hora $\sim P(0.5x20) = P(10)$

Nos piden $P(X > 10) = 1 - P(X \leq 10) = 0{,}4169$

1.6. APROXIMACIÓN BINOMIAL-POISSON

Si en una distribución binomial el número de experimentos realizados, n, es muy grande y la probabilidad de éxito, p, es pequeña, las probabilidades calculadas por dicha distribución se aproximarán a las de una distribución de Poisson. De ahí el nombre que ésta recibe de *ley de los sucesos raros*.

Más concretamente se verifica que si $X\sim Bi(n,p)$

$$\forall a,b\in N \quad P(a\leq X\leq b) \approx P(a\leq Y\leq b) \text{ donde } Y\sim P(\lambda)$$

siempre que $n\to\infty$ y $np\to\lambda$ constante.

Esta aproximación suele funcionar bien si $p\leq 0.1$.

Ejemplo 1.6.

La probabilidad de que una determinada póliza de seguros sea reclamada a lo largo de un año es de una entre 10000. Si una empresa de seguros tiene 1000 pólizas de este tipo, calcula la probabilidad de que haya más de 1 reclamación al año.

Sea X = número de pólizas reclamadas $\sim Bi(1000,0.0001) \approx P(0.1)$

$P(X > 1) = 1 - P(X \leq 1) = 0,0047$

APENDICE: OTRAS CARACTERÍSTICAS DE UNA DISTRIBUCIÓN DE PROBABILIDAD

Además de la media, la varianza y la desviación típica existen otras características de una distribución de probabilidad que, al igual que ocurría con las distribuciones estadísticas de frecuencias, informan sobre aspectos complementarios de la misma como, por ejemplo, su forma y/o proporcionan medidas alternativas de posición y dispersión.

En este apéndice revisamos, brevemente, algunas de dichas características

A.1. Momentos

Sea X una variable aleatoria discreta con función de probabilidad p y soporte D

Se define el momento de orden $k \in \mathbf{N}$ respecto al origen como

$$\alpha_k = E(X^k) = \sum_{x \in D} x^k p(x)$$

Observar, en particular, que $\alpha_1 = E[X]$ es la media de la distribución. Un resultado importante relacionado con ellos es que caracterizan, de forma única, a la distribución, es decir, que si dos variables aleatorias tienen todos sus momentos respecto al origen, iguales entonces sus distribuciones son iguales.

Así mismo se define el momento central de orden μ_k como

$$\mu_k = E[(X-\alpha_1)^k]$$

Se tiene, en particular, que:

$$\mu_1 = 0, \ \mu_2 = Var(X)$$

Los momentos centrales y no centrales están muy relacionados entre si y se puede probar que:

$$\mu_k = \sum_{u=0}^{k} (-1)^u \binom{k}{u} \alpha_u \alpha_{k-u}$$

A.1.1. Medidas de forma

Dos momentos centrales muy utilizados son μ_3 y μ_4 a partir de los cuales se cuantifica la forma de la distribución de probabilidad mediante los coeficientes de asimetría (CA) y curtosis (CK) que vienen dados por:

$$CA = \frac{\mu_3}{\sigma^3}$$

de forma que, al igual que ocurría con las distribuciones estadísticas de frecuencias:

- Si CA > 0 la distribución es asimétrica a derechas con respecto a su media
- Si CA = 0 la distribución es simétrica con respecto a su media
- Si CA < 0 la distribución es asimétrica a izquierdas con respecto a la media

El coeficiente de curtosis se calcula:

$$CK = \frac{\mu_4}{\sigma^4}$$

y dependiendo de su valor podemos comparar su distribución con la normal que es la de referencia:

- Si CK > 3 la distribución es más apuntada que la distribución normal y se dice que es leptocúrtica
- Si CK = 3 la distribución es igual de apuntada que la distribución normal y se dice que es mesocúrtica
- Si CK < 3 la distribución es menos apuntada que la distribución normal y se dice que es platicúrtica

Tanto la asimetría como la curtosis informan de la capacidad que tiene la distribución de probabilidad de generar observaciones atípicas. Cuanto más asimétrica es una distribución de probabilidad, mayor es la probabilidad de que genere observaciones muy alejadas de la media, bien sea por la cola derecha (CA>0) o por la izquierda (CA<0). Si es simétrica todo dependerá de su grado de apuntamiento. Así, cuanto mayor es el grado de apuntamiento de una distribución (es decir, mayor es el grado de leptocurtosis), más apuntada es la distribución en torno a su media y mayor probabilidad hay de que se generen observaciones muy alejadas de la media por, al menos, una de las dos colas. Por el contrario, cuanto menor es el grado de apuntamiento (es decir, mayor es el grado de platicurtosis) más plana es la distribución de probabilidad de forma que tiende a dar la misma probabilidad a todos los valores del soporte siendo, por tanto, menor su probabilidad de producir valores atípicos, siendo la distribución uniforme la que menor grado de apuntamiento tiene.

A.2. Cuantiles

Otras características notables de una distribución de probabilidad son los cuantiles. Sea $0<p<1$. Definiremos el cuantil de orden p de la distribución, x_p, como aquél valor del soporte que verifica que

$$P(X \leq x_p) \geq p \text{ y } P(X \geq x_p) \geq 1-p$$

En particular se tiene que:

$x_{0.5}$ = Me es la mediana de la distribución

$x_{0.25} = C_1$, $x_{0.5} = C_2$ y $x_{0.75} = C_3$ son los cuartiles de la distribución

$x_{0.1i} = D_i$; i = 1, …., 9 son los deciles de la distribución

$x_{0.01i} = P_i$; i = 1, …, 99 son los percentiles de la distribución

Los cuantiles se suele utilizar para la elaboración de intervalos predictivos del valor de X para una confianza $0<\alpha<1$. Un intervalo muy utilizado es el que tiene la forma $\left(x_{\frac{\alpha}{2}}, x_{1-\frac{\alpha}{2}} \right]$ para $0<\alpha<0.5$. Así, por ejemplo, el intervalo intercuartílico $[C_1,C_3]$ contiene el valor de la variable X con un nivel de confianza del 50%, mientras que el intervalo interdecílico $[D_1, D_9]$ lo contiene con un nivel de confianza del 80% y el percentílico $[P_1,P_{99}]$ con un nivel del 98%.

A.3. Otras características

Una característica notable de una distribución de probabilidad es la moda, Mo, que se define como el valor más probable de la distribución de probabilidad. Puede además, existir varias modas que se corresponderían con máximos relativos de p, en cuyo caso la distribución se denomina bimodal, trimodal, etc. La existencia de varias modas suele venir asociada a la existencia de varios mecanismos aleatorios subyacentes a la variable X y, en estos casos, el interés del estudio radica en descubrir cuáles son.

Finalmente también se utilizan las desviaciones absolutas respecto a la media, mediana o moda que vienen dadas por $D_{\alpha_1} = E[|X-\alpha_1|]$, D_{Me} = E[|X-Me|] y D_{Mo} = E[|X-Mo|] y que se utilizan como medidas de dispersión robustas respecto a cada una de dichas medidas.