Índice

Tema 4

Modelos de distribuciones

4.1 Introducción

En este tema estudiaremos algunas distribuciones asociadas a variables aleatorias discretas y continuas, que puedan ajustarse a una gran diversidad de problemas y campos científicos en los que se pueden aplicar. Podemos establecer la siguiente clasificación:

- Modelos de distribuciones discretas
- Modelos de distribuciones continuas

4.2 Modelos de distribuciones discretas

Para describir las variables aleatorias discretas asociadas a determinados experimentos basta con proporcionar la distribución de probabilidad; por ello vamos a estudiar las distribuciones de probabilidad más importantes que describan el comportamiento de muchas de las variables aleatorias discretas que se encuentran en la práctica y otras características asociadas.

4.2.1 Distribución Uniforme Discreta

Se considera la más simple de todas las distribuciones de probabilidad discretas, en la que la variable aleatoria asociada toma cada uno de los valores con una probabilidad idéntica.

Definición 4.1 *Diremos que una variable aleatoria discreta X tiene una distribución Uniforme de parámetro N, $N \geq 1$, si*

$$rg(X) = \{1, 2, ..., N\}$$

y la función de probabilidad puede escribirse

$$P[X = x] = \frac{1}{N} \qquad x = 1, 2, ..., N$$

En tal caso se denotará como $X \sim UD(N)$

Proposición 4.1 *Sea* $X \sim UD(N)$ *entonces*

$$E[X] = \frac{N+1}{2},$$

$$Var[X] = \frac{N^2-1}{12}.$$

Ejemplo 4.1 *Sea el experimento aleatorio de lanzar un dado. Definimos la variable* X *como la puntuación obtenida en dicho lanzamiento. Se tiene que*

$$rg(X) = \{1, 2, 3, 4, 5, 6\},$$

$$P[X = x] = \frac{1}{6} \qquad x \in rg(X).$$

Entonces $X \sim UD(6)$

Esta distribución también se puede definir para cualesquiera valores x_1, $x_2, ..., x_N$. Lo que la caracteriza es que todos los posibles valores que puede tomar la variable aleatoria tienen la misma probabilidad de ocurrir.

Ejemplo 4.2 *Se selecciona una bombilla al azar de una caja que contiene una bombilla de 40 vatios, una de 60, una de 75 y una de 100. Cada elemento del espacio muestral tiene una probabilidad de 1/4. Por tanto tenemos una distribución uniforme discreta donde:*

$$rg(X) = \{x_1, x_2, x_3, x_4\}$$

x_1= *"bombilla de 40 vatios";* x_2= *"bombilla de 60 vatios";* x_3= *"bombilla de 70 vatios";*
x_4= *"bombilla de 100 vatios"*

$$P[X = x_i] = \frac{1}{4}, \quad x_i \in rg(X).$$

Entonces, $X \sim UD(x_1, x_2, x_3, x_4)$.

4.2.2 Distribución de Bernoulli

Sea un experimento con dos posibles resultados, generalmente llamados éxito y fracaso para los que $P[\text{éxito}] = p$ y $P[\text{fracaso}] = 1 - p = q$. A este tipo de experimento se le denomina experimento o prueba de Bernoulli. Se le puede asociar una variable aleatoria de la siguiente forma:

$X : \Omega \to \{0, 1\}$ Fracaso $\to$ 0 Éxito $\to$ 1.

Por tanto, tenemos:

$P[X = 1] = P[\text{éxito}] = p$,

$P[X = 0] = P[\text{fracaso}] = q$.

Definición 4.2 *Una variable aleatoria discreta X se dice que sigue una distribución de Bernoulli de parámetro p, $X \sim Be(p)$, con $p \in [0,1]$; si se verifica:*

$$rg(X) = \{0, 1\}$$

y la función de probabilidad puede escribirse como:

$$P[X = x] = p^x(1-p)^{1-x} \quad x \in \{0, 1\}.$$

Proposición 4.2 *Si $X \sim Be(p)$ entonces*

$$\begin{aligned} E[X] &= p, \\ Var[X] &= pq. \end{aligned}$$

Como ejemplos de distribución de Bernoulli podemos citar un sistema eléctrico que funcione o no, consideramos éxito que el sistema funcione y fracaso que no funcione. De la misma forma podemos tener un conmutador en ON/OFF, un servidor con conexión o sin ella, un fusible defectuoso o no defectuoso, etc.

Ejemplo 4.3 *Consideramos el lanzamiento de una moneda. Definimos Éxito="Obtener cara al lanzar la moneda" y Fracaso="Obtener cruz al lanzar la moneda". En tal caso la variable X tiene una probabilidad de éxito $\frac{1}{2}$ y por tanto $X \sim Be(1/2)$ Calculamos la media y varianza*

$$\begin{aligned} E[X] &= \frac{1}{2}, \\ Var[X] &= \frac{1}{4}. \end{aligned}$$

4.2.3 Distribución Binomial

Consideramos n experimentos de Bernoulli independientes e idénticamente distribuidos y sea la variable aleatoria X definida como el número de éxitos en n pruebas de Bernoulli. Diremos que X sigue una distribución binomial de parámetros n y p,

$$X \sim B(n, p),$$

donde n es el número de pruebas, p probabilidad de éxito en una prueba.

Definición 4.3 *Se dice que una variable aleatoria discreta X sigue una distribución binomial de parámetros $n \in N$ y $p \in [0,1]$ si se verifica:*

$$rg(X) = \{0, 1, ..., n\}$$

y la función de probabilidad puede escribirse como:

$$P[X = k] = \binom{n}{k} p^k q^{(n-k)}, \quad k = 0, 1, \ldots, n.$$

Proposición 4.3 *Si* $X \sim B(n,p)$ *podemos expresar* $X = \sum_{i=1}^{n} I_i$ *con*

$$I_i = \begin{cases} 1 & \textit{si se tiene éxito en la prueba } i \\ 0 & \textit{si se tiene fracaso en la prueba } i\,, \end{cases}$$

es decir, como una suma de n *variables aleatorias independientes e idénticamente distribuidas (i.i.d) según* $I_i \sim Be(p)$.

De la Proposición 4.3 se deducen las siguientes propiedades.

Proposición 4.4 *Si* $X \sim B(n,p)$ *entonces*

$$\begin{aligned} E[X] &= np, \\ Var[X] &= npq. \end{aligned}$$

Proposición 4.5 *Dadas* $X \sim B(n,p)$, $Y \sim B(m,p)$ *independientes entonces* $X+Y \sim B(n+m,p)$.

Como ejemplos de variables que siguen una distribución binomial podemos citar: número de veces que aparece el resultado "cara" al lanzar una moneda diez veces; número de éxitos en la recepción de un mensaje enviado a 100 destinatarios; número de ordenadores en una subred que han sido infectados por un determinado virus, etc.

Ejemplo 4.4 *La probabilidad de que cierta clase de componente sobreviva a una prueba de choque dada es 3/4. Calcular la probabilidad de que sobrevivan exactamente dos de los cuatro componentes que se prueban.*

Supongamos que las pruebas realizadas son independientes y con probabilidad de éxito p=3/4 para cada una de las 4 pruebas, por tanto la variable $X \sim B(4, 3/4)$; *tenemos que calcular*

$$P\left[X=2\right] = \binom{4}{2}\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^2 = \frac{4!}{2!2!} \times \frac{3^2}{4^4} = \frac{27}{128} = 0.2109,$$

Podemos calcular también la media y la varianza

$$E[X] = 3,$$
$$Var[X] = \frac{3}{4}.$$

Nota 4.1 *La distribución binomial encuentra aplicaciones en muchos campos científicos. Dentro de la ingeniería es normal estudiar la proporción de elementos defectuosos dentro de un proceso industrial; las mediciones de control de calidad y los esquemas de muestreo se basan en la distribución binomial. Ésta se aplica en cualquier situación industrial donde el resultado es un proceso dicotómico y los resultados del proceso son independientes y la probabilidad de éxito es constante de una prueba a otra.*

4.2.4 Distribución Multinomial

En muchas aplicaciones hay más de dos resultados posibles. A menudo la dicotomía "defectuoso" o "no defectuoso" en situaciones de ingeniería es una simplificación de la realidad, donde suele haber más de dos categorías que caracterizan artículos o partes de una línea de producción. Para estas situaciones el experimento binomial se convierte en experimento multinomial cuando cada prueba tiene más de dos resultados posibles. Si un experimento puede tener como consecuencia k posibles resultados $E_1,E_{2,...,}E_k$ con probabilidades $p_1,p_{2,...,}p_k$, entonces la distribución multinomial dará la probabilidad de que E_1 ocurra x_1 veces, E_2 ocurra x_2 veces,..., E_k ocurra x_k veces en n pruebas independientes, donde

$$x_1 + x_2 + ... + x_k = n.$$

La función de probabilidad de las variables aleatorias $X_1,X_{2,...,}X_k$, que representan el número de ocurrencias para $E_1,E_{2,...,}E_k$ en n pruebas independientes es

$$P[X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k] = \frac{n!}{x_1!x_2!\ldots x_k!}p_1^{x_1}p_2^{x_2}\ldots p_k^{x_k}$$

con

$$\sum_{i=1}^{k} x_i = n$$

y

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

A la variable $(X_1, \ldots, X_n)$ se le dice que sigue una distribución multinomial y se denota por $\mathcal{M}(p_1, p_2, \ldots, p_k; n)$

Ejemplo 4.5 *Los fallos de impresión de un libro se pueden clasificar en erratas tipográficas (E), mala impresión (M) y hojas en blanco (H). Un editor presenta en los fallos de sus publicaciones un 80% de erratas, un 15% de hojas mal impresas y sólo un 5% de hojas en blanco. Calcular la probabilidad de que de 10 fallos encontrados en un libro, 6 sean erratas y 3 carencias de impresión.*

Al considerar 10 fallos, 6 corresponden a erratas, 3 a carencias de impresión, el fallo restante corresponde a una hoja en blanco. Por lo tanto

$$P\left[X_1 = 6; X_2 = 3; X_3 = 1\right] = \frac{10!}{6!3!1!} \times 0.80^6 \times 0.15^3 \times 0.05 = 0.03716.$$

4.2.5 Distribución Hipergeométrica

La diferencia entre la distribución binomial y la distribución hipergeométrica está en la forma en que se realiza el muestreo. En el caso de la binomial se requiere independencia entre las pruebas, es decir, el muestreo se realiza con reemplazamiento. Por otro lado, la distribución hipergeométrica no requiere independencia y el muestreo se realiza sin reemplazamiento. Las aplicaciones de la distribución hipergeométrica se encuentran en muchas áreas, con gran uso en control de calidad.

Consideramos un conjunto de N elementos en el que tenemos N_1 elementos del tipo 1 y $N - N_1$ elementos de tipo 2. Extraemos n **sin reemplazamiento** y definimos la variable aleatoria:

X = "número de elementos del tipo 1 en los n extraídos". El modelo de distribución que sigue esta variable aleatoria es una hipergeométrica.

Definición 4.4 *Se dice que una variable aleatoria discreta X sigue una distribución hipergeométrica de parámetros N, N_1 y n con $n \leq N$, si se verifica:*

$$rg(X) = \{máx\,(0, n-(N-N_1)), ..., mín(n, N_1)\}$$

y la función de probabilidad de X es

$$P[X=k] = \frac{\binom{N_1}{k}\binom{N-N_1}{n-k}}{\binom{N}{n}}, \qquad k \in rg(X)$$

con $0 \leq k \leq N_1$, $0 \leq n-k \leq N-N_1$. Se denota $X \sim \mathcal{H}(N, N_1, n)$.

Proposición 4.6 *Si $X \sim \mathcal{H}(N, N_1, n)$ entonces*

$$\begin{aligned} E[X] &= n\frac{N_1}{N}, \\ Var[X] &= n\frac{N_1}{N}\left(1-\frac{N_1}{N}\right)\left(\frac{N-n}{N-1}\right) . \end{aligned}$$

Ejemplo 4.6 *Un fabricante de chips de silicio los empaqueta en lotes de 25. El comprador los inspecciona tomando 3 chips y acepta un lote si encuentra menos de 2 chips defectuosos.*

(a) Calcular la probabilidad de que el comprador acepte un lote con 6 chips defectuosos.

(b) ¿Cuál es el número esperado y la varianza de chips defectuosos en los 3 inspeccionados?

(a) Se tiene un conjunto de $N = 25$ chips, en los que hay $N_1 = 6$ defectuosos, y $N - N_1 = 19$ no defectuosos. Extraemos $n = 3$ sin reemplazamiento. Consideramos la variable aleatoria X = "número de chips defectuosos en los 3 seleccionados",

$$X \sim \mathcal{H}(N=25, N_1=6, n=3) .$$

La probabilidad de que el comprador acepte el lote es:

$$P[X<2] = P[X=0] + P[X=1] = \frac{\binom{6}{0}\binom{19}{3}}{\binom{25}{3}} + \frac{\binom{6}{1}\binom{19}{2}}{\binom{25}{3}} = 0.8674,$$

$$\begin{aligned} E[X] &= 3 \times \frac{6}{25} = 0.72 , \\ Var[X] &= 3 \times \left(\frac{6}{25}\right) \times \left(\frac{19}{25}\right) \times \left(\frac{22}{24}\right) = 0.5016 . \end{aligned}$$

Ejemplo 4.7 *Se dispone de una urna con 3 bolas blancas y 7 negras. Extraemos 4 bolas al azar y queremos contar el número de bolas blancas en las 4 extracciones. Enunciar los modelos que se obtendrían al realizar el experimento con y sin reemplazamiento.*

Si existe reemplazamiento, la probabilidad de extraer una bola blanca no varia en el tiempo, y por tanto

$$X \sim B\left(4, \frac{3}{10}\right).$$

Si no se realiza reemplazamiento, las probabilidades variarán en el tiempo ya que cambia la composición de la urna y, por lo tanto,

$$X \sim \mathcal{H}(10, 3, 4) \ .$$

4.2.6 Distribución Geométrica

Realizamos una serie de pruebas de Bernoulli independientes e idénticamente distribuidas con la misma probabilidad de éxito, p, en cada prueba. Definimos:

$X =$ "número de fracasos antes de obtener el primer éxito". Entonces se tiene un tipo de distribución denominada geométrica.

Definición 4.5 *Se dice que una variable aleatoria discreta X sigue una distribución geométrica de parámetro $p \in (0, 1]$ si se verifica:*

$$rg(X) = \{0, 1, 2, \ldots\}$$

y la función de probabilidad es:

$$P[X = k] = q^k p \ , \quad k = 0, 1, ..., \quad k \in rg(X).$$

Se denotará como $X \sim Ge(p)$, con $0 < p \leq 1$.

Comprobamos que es función de probabilidad:

$$\sum_{k=0}^{\infty} P[X = k] = \sum_{k=0}^{\infty} q^k p = p \sum_{k=0}^{\infty} q^k = p \frac{1}{1-q} = p \frac{1}{p} = 1 \ ,$$

pues $0 \leq q < 1$.

Proposición 4.7 *Si $X \sim Ge(p)$, entonces*

$$E[X] = \frac{q}{p} \ ,$$

$$Var[X] = \frac{q}{p^2} \ .$$

Proposición 4.8 *Si $X \sim Ge(p)$, entonces*

$$P\left[X \geq k + j | X \geq k\right] = P[X \geq j] = q^j \ , \qquad k, j = 0, 1, 2, ... \ .$$

Se dice que la distribución geométrica carece de memoria.

Ejemplo 4.8 *En un proceso de fabricación se puede suponer por estudios previos que la probabilidad de que un artículo sea defectuoso es $p = 0.01$. Inspeccionamos artículos secuencialmente hasta encontrar uno defectuoso. Calcular la probabilidad de que el quinto artículo seleccionado sea el primero defectuoso.*

Definimos $X =$ "número de artículos seleccionados antes de encontrar el primero defectuoso", $X \sim Ge(0.01)$.

$$\begin{aligned} P[\textit{el quinto artículo seleccionado sea el primero defectuoso}] &= P[X=4] \\ &= 0.99^4 \times 0.01 = 0.0096\ . \end{aligned}$$

Calculamos también la media y la varianza

$$\begin{aligned} E[X] &= \frac{q}{p} = \frac{0.99}{0.01} = 99\ , \\ Var[X] &= \frac{q}{p^2} = \frac{0.99}{(0.01)^2} = 9900\ . \end{aligned}$$

Nota 4.2 *Algunos autores definen la distribución geométrica como $Y =$ "número de la prueba en que obtengo el primer éxito". Nótese que $Y = X + 1$, siendo X la variable con la que hemos trabajado antes, por lo que la función de probabilidad y características se deducen de forma inmediata.*

$$\begin{aligned} P[Y=k] &= P[X=k-1]\ , \\ E[Y] &= E[X]+1\ , \\ Var[Y] &= Var[X+1] = Var[X]\ . \end{aligned}$$

Nota 4.3 *Cómo área de aplicación de la distribución geométrica, podemos citar situaciones donde los ingenieros intentan determinar la "eficacia" de un sistema de conmutación telefónica durante periodos ocupados. En este caso, el éxito representa la conexión al sistema de conmutación. La variable sería el número de intentos fracasados antes de una conexión al sistema (éxito).*

4.2.7 Distribución de Poisson

La distribución de Poisson es una distribución discreta de gran utilidad donde la variable aleatoria suele estudiar el número de eventos independientes que ocurren a velocidad constante en un intervalo de tiempo o espacio. El intervalo puede ser de cualquier longitud, como un minuto, un día, una semana. Por ejemplo, el número de llamada telefónicas por hora que llegan a una oficina, el número de mensajes que llegan a un servidor de correos durante una hora, etc.

Definición 4.6 *Una variable aleatoria X discreta, sigue una distribución de Poisson de parámetro $\lambda > 0$, $X \sim \mathcal{P}(\lambda)$, si:*

$$rg(X) = \{0, 1, 2, \ldots\}$$

y la función de probabilidad viene dada por:

$$P[X=k] = e^{-\lambda}\frac{\lambda^k}{k!}\ , \quad k = 0, 1, 2, \ldots\ ,$$

Comprobamos que es función de probabilidad:

$$\sum_{k=0}^{\infty} P[X=k] = \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^{-\lambda}e^{\lambda} = 1\ .$$

Proposición 4.9 *Si $X \sim \mathcal{P}(\lambda)$ entonces*

$$E[X] = \lambda,$$
$$Var[X] = \lambda.$$

Proposición 4.10 *Dadas $X \sim \mathcal{P}(\lambda_1)$ e $Y \sim \mathcal{P}(\lambda_2)$ independientes entonces $X + Y \sim \mathcal{P}(\lambda_1 + \lambda_2)$.*

Corolario 4.1 *Dadas $X_1 \sim \mathcal{P}(\lambda)$,..., $X_n \sim \mathcal{P}(\lambda)$ independientes e idénticamente distribuidas entonces $S = X_1 + .. + X_n \sim \mathcal{P}(n\lambda)$.*

Ejemplo 4.9 *El número medio de camiones que llega cada día a la dársena de descarga de un puerto es 10, ($\lambda = 10$). Con las instalaciones de que dispone el puerto no se pueden descargar más de 15 camiones en un día. Calcular la probabilidad de que en un día determinado haya camiones que no puedan ser descargados.*

Sea X = "número de camiones que llegan en un día". La variable aleatoria $X \sim \mathcal{P}(\lambda = 10)$.

$$P[\textit{en un día determinado haya camiones que no puedan ser descargados}] =$$

$$P[X > 15] \; = \; 1 - P[X \leq 15] \; = \; 1 - 0.9513 \; = \; 0.0487 \, .$$

Nota 4.4 *La distribución de Poisson se utiliza a menudo para aproximar las probabilidades de una $B(n, p)$, cuando n es grande y p pequeño, aproximándose por una $\mathcal{P}(\lambda = np)$. En la práctica, la aproximación es buena cuando $p < 0.1$ y $np \leq 5$*

En la siguiente tabla representamos en la segunda columna la $P[X = x]$ para una distribución Binomial con parámetros $n = 40$ y $p = 0.05$, la tercera columna corresponde a la probabilidad para la distribución de Poisson con parámetro $\lambda = 40 \times 0.05 = 2$.

x	$B(n = 40,\ p = 0.05)$	$P(\lambda = 2)$
0	0.1285	0.1353
1	0.2706	0.2707
2	0.2777	0.2707
3	0.1851	0.1804
4	0.0901	0.0902
5	0.0342	0.0361
6	0.0105	0.0120

Tabla 4.1: Aproximación de la binomial por la Poisson

Ejemplo 4.10 *En un proceso de fabricación de productos de vidrio se producen rara vez burbujas que dejan al producto defectuoso para su venta. Se ha observado que uno de cada 1000 artículos que se producen tienen burbujas, es decir, la probabilidad de que un producto sea defectuoso podemos suponer que es $p = 0.001$). ¿Cuál es la probabilidad de que en un lote de 8000 artículos haya menos de 4 defectuosos?*

Sea X =" número de artículos defectuosos en 8000".

Podemos suponer

$$X \sim B(n = 8000, p = 0.001)\ .$$

Puesto que n *es muy grande y* p *está muy próximo a cero podemos aproximar* $B(n,p) \approx P(\lambda = np)$, *es decir,*

$$X \sim P(\lambda = n \times p = 8)$$

$P[\textit{haya menos de 4 defectuosos}] \ = \ P[X < 4] = \sum_{k=0}^{3} P[X = k] = 0.0424.$

4.3 Modelos de distribuciones continuas

En esta parte del tema se estudiarán algunas de las distribuciones más comunes, que se pueden aplicar a una gran diversidad de situaciones en los que se estudian problemas de naturaleza continua. Para describir este tipo de variables aleatorias basta proporcionar la función de densidad.

4.3.1 Distribución Uniforme

Definición 4.7 *Se dice que una variable aleatoria continua X sigue una distribución uniforme en el intervalo (a,b), si su función de densidad está definida como*

$$f(x) = \begin{cases} \frac{1}{b-a} & si \quad a < x < b \\ 0 & c.c. \end{cases}$$

Se denota por $X \sim U(a,b)$

Proposición 4.11 *Si* $X \sim U(a,b)$ *entonces*

$$E\,[X] = \frac{a+b}{2},$$

$$Var[X] = \frac{(b-a)^2}{12}.$$

.

Ejemplo 4.11 *El volumen de ventas de un almacén se distribuye uniformemente entre 38 y 120 miles de euros. Calcular*

(a) *La probabilidad de que las ventas sean superiores a 100000 euros.*

(b) *La esperanza matemática y varianza de las ventas*

Sea X= "Volumen de ventas del almacén en miles de euros" $\sim U(38,120)$.

(a) La función de densidad de esta variable

$$f(x) = \frac{1}{120-38} = \frac{1}{82} \qquad si \quad 38 < x < 120$$

Por tanto, como la variable está definida en miles de euros

$$P[X > 100] = \int_{100}^{+\infty} f(x)dx = \int_{100}^{120} \frac{1}{82}dx = \frac{1}{82}[x]_{100}^{120} = \frac{20}{82} = 0.2439$$

(b)

$$E[X] = \frac{a+b}{2} = \frac{38+120}{2} = 79$$

.

$$Var[X] = \frac{(b-a)^2}{12} = \frac{(120-38)^2}{12} = 560.333$$

4.3.2 Distribución Exponencial

Definición 4.8 *Se dice que la variable aleatoria continua X sigue una distribución exponencial de parámetro λ, con $\lambda > 0$ si su función de densidad es*

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & si \quad x > 0 \\ 0 & c.c. \end{cases}$$

Se denota por $X \sim Exp(\lambda)$

Proposición 4.12 *Si $X \sim Exp(\lambda)$ entonces*

$$E[X] = \frac{1}{\lambda},$$
$$Var[X] = \frac{1}{\lambda^2}.$$

Como ejemplos de una distribución exponencial podemos citar el tiempo que tarda en fallar un sistema electrónico, tiempo transcurrido entre la llegada de dos mensajes de correos consecutivos, etc.

Proposición 4.13 *Si $X \sim Exp(\lambda)$ entonces*

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & si \quad x > 0 \\ 0 & c.c. \end{cases}$$

$$P[X > x] = e^{-\lambda x}, \quad x > 0.$$

Proposición 4.14 *Si $X \sim Exp(\lambda)$ entonces*

$$P[X \geq x + t | X \geq t] = P[X \geq x] = e^{-\lambda x}, \qquad x, t > 0.$$

Se dice que la distribución exponencial no tiene memoria. Por ello, esta distribución se suele utilizar cuando medimos el tiempo de vida restante de poblaciones que no envejecen con el tiempo.

Ejemplo 4.12 *Se ha comprobado que la duración de vida de cierto tipo de circuitos sigue una distribución exponencial con media 8 años. Calcular:*

(a) La probabilidad de que un circuito tenga una vida entre 3 y 12 años.

(b) La probabilidad de que un circuito que ha vivido más de 10 años, viva 15 años más.

(a) Sea X="Tiempo de vida de los circuitos", como la media de la variable aleatoria es 8,

$X \sim Exp(\lambda = \frac{1}{8})$*, por tanto la función de densidad*

$$f(t) = \frac{1}{8} e^{-\frac{t}{8}}, \quad t > 0$$

donde t va medido en años. Entonces:

$$P[3 < X < 12] = \int_3^{12} \frac{1}{8} e^{-\frac{t}{8}} dt = \left[-e^{-\frac{t}{8}} \right]_3^{12} = 0.69 - 0.22 = 0.46$$

(b) Utilizando la Proposición 1.14:

$$P[X > 25 | X > 10] = \frac{P[X > 25]}{P[X > 10]} = e^{-\lambda 15}$$

Por tanto

$$P[X > 25 | X > 10] = P[X > 15]$$

Luego podemos afirmar que la variable aleatoria X no tiene memoria.

4.3.3 Distribución Normal

Es la más importante de las distribuciones continuas ya que permite describir un número muy grande de fenómenos aleatorios, como por ejemplo aquellos en los que intervienen un número elevado de factores no controlables, que actúan de manera independiente y con efectos pequeños.

Definición 4.9 *Se dice que una variable aleatoria X sigue una distribución normal de parámetros* $\mu \in R$ *y* $\sigma^2 > 0$*, si su función de densidad es*

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2} \right], \qquad x \in \mathbb{R},$$

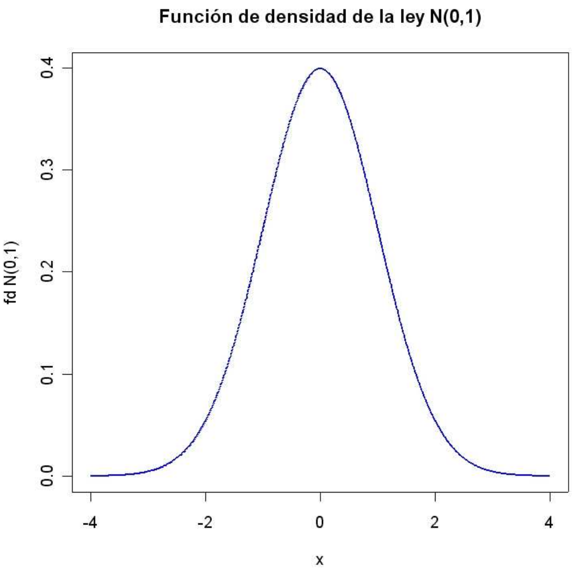
se denota por $X \sim N(\mu, \sigma^2)$

Proposición 4.15 *Si* $X \sim N(\mu, \sigma^2)$ *entonces*

$$E[X] = \mu,$$
$$Var[X] = \sigma^2.$$

La gráfica de la función de densidad, se denomina curva normal (o campana de Gauss), es una curva con forma de campana, la cual describe aproximadamente muchos fenómenos que se utilizan en la vida cotidiana. Por ejemplo, las mediciones físicas en áreas como los experimentos metereológicos, estudios de lluvia, errores en las mediciones científicas, etc.

A continuación se muestra la gráfica de la función de densidad para una distribución normal con media 0 y desviación típica igual a 1.



Gráficamente se observa que presenta un máximo en μ, tiene puntos de inflexión en $\mu - \sigma$ y $\mu + \sigma$ y es simétrica respecto a μ. Se aproxima al eje horizontal de manera asintótica conforme nos alejamos de la media en cualquier dirección. El área total entre la curva y el eje horizontal es igual a 1.

Proposición 4.16

(a)Si $X \sim N(\mu, \sigma^2)$ entonces $aX + b \sim N(a\mu + b, a^2\sigma^2)$ con $a, b \in \mathrm{R}$, $a \neq 0$.

(b) En particular, si $X \sim N(\mu, \sigma^2)$ entonces $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

Por tanto, la función de distribución de cualquier $N(\mu, \sigma^2)$ se puede calcular utilizando la distribución $N(0,1)$ de la siguiente forma

$$P\left[X \leq x\right] = P\left[\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right] = P\left[Z \leq \frac{x-\mu}{\sigma}\right]$$

donde $Z = \frac{X-\mu}{\sigma}$ es la normal tipificada. Generalmente, se denota $\Phi(z) = P\left[Z \leq z\right]$.

Esta propiedad permite emplear las tablas de la $N(0,1)$ para obtener la función de distribución de cualquier distribución normal.

Proposición 4.17 *Sea $Z \sim N(0,1)$ con función de distribución Φ. Si $X \sim N(\mu, \sigma^2)$ entonces*

(a)

$$P\left[X \leq x\right] = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

(b)

$$P\left[a < X < b\right] = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

Teorema 4.1 *Sea $Z \sim N(0,1)$ con función de distribución Φ. Entonces $\Phi(-z) = 1 - \Phi(z), \qquad z \in \mathbb{R}$*

Teorema 4.2 *Sean $X \sim N(\mu_1, \sigma_1^2)$ e $Y \sim N(\mu_2, \sigma_2^2)$ independientes. Entonces*

(a)
$$X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

(b)
$$X - Y \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$$

Corolario 4.2 *Dadas $X_{1,\ldots,}X_n$ variables aleatorias independientes e idénticamente distribuidas según $N(\mu, \sigma^2)$ entonces*

$$\textstyle\sum_{i=1}^{n} X_i \sim N(n\mu, n\sigma^2).$$

Ejemplo 4.13 *Dada una distribución normal con $\mu = 50$ y $\sigma = 10$, calcular la probabilidad de que X tome una valor entre 45 y 62.*

$$\begin{aligned} P[45 < X < 62] &= P\left[\frac{45-50}{10} < Z < \frac{62-50}{10}\right] = P\left[-0.5 < Z < 1.2\right] \\ &= \Phi(1.2) - \Phi(-0.5) \\ &= 0.8849 - 0.3085 = 0.5764 \end{aligned}$$

Ejemplo 4.14 *Sea una distribución normal con $\mu = 300$ y $\sigma = 50$, obtener la probabilidad de que X tome un valor mayor que 362.*

$$P\left[X > 362\right] = P\left[Z > \frac{362-300}{50}\right] = P\left[Z > 1.24\right] = 1 - \Phi(1.24) = 1 - 0.8925 = 0.1075$$

La distribución normal encuentra una gran aplicación como distribución límite, es decir, bajo ciertas condiciones la distribución normal proporciona una buena aproximación continua a otras distribuciones. Puede utilizarse para aproximar probabilidades de variables binomiales cuando n es grande, y p no esté muy cercano a cero o uno, en general esta aproximación se utiliza para $np > 5$. Entonces aproximamos por una $N(\mu = np, \sigma^2 = npq)$. La distribución normal también puede aproximar a la distribución de Poisson, $\mathcal{P}(\lambda)$, cuando λ es elevado, considerando $N(\mu = \lambda, \sigma^2 = \lambda)$.

4.3.4 Otras Distribuciones

Existen otras distribuciones que aunque no vamos a desarrollar en el tema pueden tener cierto interés por su utilidad dentro de algunas áreas de la ingeniería como el control de calidad y fiabilidad de procesos. La distribución gamma es una generalización de la distribución exponencial y representa el tiempo transcurrido hasta que cierto suceso aparece k veces. Por tanto la distribución gamma se puede expresar como suma de exponenciales. La distribución gamma no tiene en cuenta que el envejecimiento de un sistema, incrementa su probabilidad de "fallo". Para incluir esta condición en el modelo probabilístico resulta útil emplear la distribución de Weibull. Esta distribución se emplea a menudo en ingeniería ya que, seleccionado los valores adecuados para los parámetros asociados a esta distribución, se puede caracterizar la fiabilidad de un gran número de sistemas.

4.4 Apéndice: Corrección de continuidad

Cuando aproximamos una distribución discreta (binomial o Poisson) por una distribución normal es aconsejable realizar una corrección por continuidad. Sabemos que en una variable aleatoria discreta la $P[X = x] > 0, x \in rg(X)$, en cambio, cuando la variable es continua la $P[X = x] = 0$. Por ello al aproximar una variable aleatoria discreta por una continua debemos evitar los problemas en los puntos, considerando intervalos de la siguiente forma:

$$P[X = a] \approx P[a - 1/2 < X \leq a + 1/2]$$

A continuación se detalla un ejemplo donde incluimos la corrección de continuidad:

Ejemplo 4.15 *Unos grandes almacenes estiman que el* 60% *de sus clientes paga con dinero en efectivo, y el resto recurre a otros medios. Calcular la probabilidad de que de 30 clientes exactamente 20 paguen en efectivo.*

Sea X="Número de clientes que pagan en efectivo entre 30 clientes", por tanto la variable aleatoria X *sigue una distribución binomial que la vamos a aproximar a una distribución normal, es decir,* $X \sim B(30, 0.6) \simeq N(18, 7.2)$.

Tenemos que calcular la $P[X = 20]$*, como hemos pasado de una distribución discreta a una continua aplicamos la corrección de continuidad, luego*

$$\begin{aligned} P[X = 20] &= P[19.5 < X \leq 20.5] = P[X \leq 20.5] - P[X \leq 19.5] = \\ &= \Phi\left(\frac{20.5 - 18}{2.68}\right) - \Phi\left(\frac{19.5 - 18}{2.68}\right) = 0.115. \end{aligned}$$