Tema 7. Variables Aleatorias Continuas

Presentación y Objetivos.

En este tema se propone el estudio de las variables aleatorias continuas más importantes, desde la más simple incrementando el grado de sofisticación. Tras conocer en detalle cada una de ellas, se estudiará en qué condiciones se admiten aproximaciones a la distribución Normal, cuya importancia se verá reflejada y quedará justificada en los contenidos de los temas relacionados con la Inferencia Estadística.

Los **Objetivos** de esta Unidad Didáctica son:

1. Conocer a nivel conceptual y operativo las distribuciones continuas más importantes, motivadas a través de ejemplos.
2. Entender cuándo pueden obtenerse aproximaciones a la distribución Normal.

Esquema Inicial.

1. Distribución Uniforme.
2. Distribución Normal o Gaussiana.
3. Distribución Gamma.
4. Distribución Beta.

Desarrollo del Tema

1. Distribución Uniforme

Una variable aleatoria X se distribuye según una **distribución Uniforme o rectangular** en el intervalo $[a, b]$ con $-\infty < a < b < \infty$ si su función de densidad está dada por

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{en el resto} \end{cases}$$

Se escribe $X \sim U[a, b]$ o $X \in U[a, b]$ si X tiene una distribución Uniforme en $[a, b]$.

Es función de densidad ya que:

a) $f \geq 0 \;\; \forall x \in \mathbb{R}$ al ser $a < b$.
b) $\int_{-\infty}^{\infty} f(x)dx = \int_{a}^{b} \frac{1}{b-a} dx = \frac{1}{b-a}(b-a) = 1$

Su función de distribución es

$$F(x) = \begin{cases} 0 & \text{si } x < a \\ \dfrac{x-a}{b-a} & \text{si } a \le x \le b \\ 1 & \text{si } x > b \end{cases}$$

1.1 Medidas características

- Media:

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_a^b \frac{x}{b-a}dx = \frac{1}{b-a}\frac{x^2}{2}\Big]_a^b = \frac{1}{b-a}\frac{b^2-a^2}{2} = \frac{a+b}{2}$$

- Varianza: para calcularla, se obtendrá primero $E(X^2)$.

$$E(X^2) = \int_a^b \frac{x^2}{b-a}dx = \frac{1}{b-a}\frac{x^3}{3}\Big]_a^b = \frac{1}{b-a}\frac{b^3-a^3}{3} = \frac{b^2+ba+a^2}{3}$$

con lo que

$$\sigma^2 = V(X) = E(X^2) - E(X)^2 = \frac{b^2+ba+a^2}{3} - \frac{b^2+2ba+a^2}{4} = \frac{(b-a)^2}{12}$$

Además, una **propiedad** importante de la distribución Uniforme es la siguiente:

$$X \in U[a,b] \Rightarrow mX + n \in U[ma+n, mb+n]$$

2. Distribución Normal o Gaussiana

Esta distribución destaca por sus múltiples aplicaciones y por los teoremas centrales del límite que se estudiarán más adelante. Las aplicaciones más destacadas son las siguientes: para representar variables físicas (temperatura, altura, peso, …), errores de instrumentación, calificaciones en pruebas de aptitud, inferencia estadística (ya que muchos estadísticos tienden a la Normal conforme crece el tamaño de la muestra), tiempo de vida de un componente eléctrico.

Una variable aleatoria X tiene una **distribución Normal** con parámetros $\mu \in \mathbb{R}$ y $\sigma > 0$ ($X \sim N(\mu, \sigma)$) si su función de densidad es la siguiente:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}, \qquad x \in \mathbb{R}$$

La figura 1 muestra la representación gráfica de la función de densidad de la distribución Normal para diferentes valores de la desviación típica σ y con igual media $\mu = 0$.

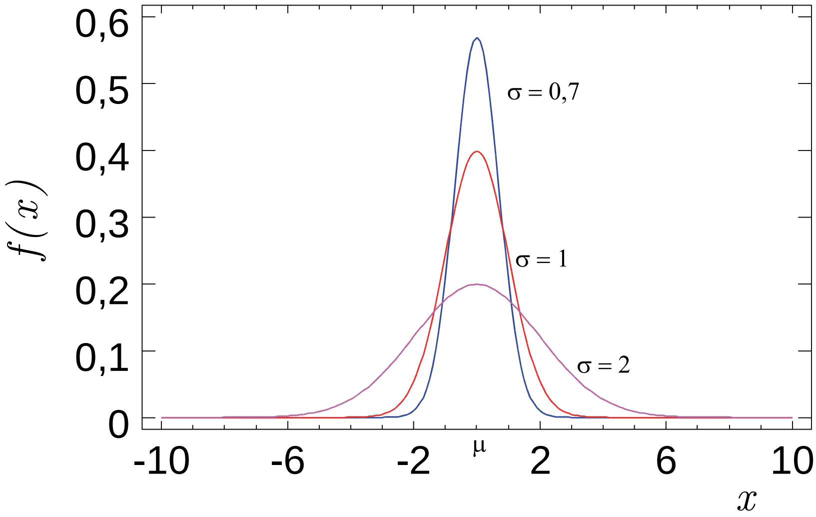


Figura 1: Distribución Normal para diferentes valores de σ

Esta función tiene las siguientes características:

- f tiene como asíntota a $y = 0$ ya que $\lim_{x\to\pm\infty} f(x) = 0$.
- f es simétrica respecto a μ: $f(\mu - a) = f(\mu + a) \quad \forall a \in \mathbb{R}$.
- f tiene un máximo en $x = \mu$, $y = \frac{1}{\sigma\sqrt{2\pi}}$.
- f tiene dos puntos de inflexión: $x = \mu \pm \sigma$.

Observación: La distribución Normal $N(\mu, \sigma)$ es reproductiva respecto de μ y σ^2, es decir, dadas X, Y variables aleatorias independientes con $X\sim N(\mu_1, \sigma_1)$, $Y\sim N(\mu_2, \sigma_2)$ entonces

$$X + Y\sim N\left(\mu_1 + \mu_2, \sqrt{{\sigma_1}^2 + \sigma_2^2}\right)$$

La función de distribución es

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} dx$$

Esta función existe pero no se puede poner en función de las conocidas. Se transforma en una distribución $N(0,1)$ haciendo el cambio $\frac{x-\mu}{\sigma} = y$:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-\frac{y^2}{2}} dy$$

que es la función de distribución de una $N(0,1)$ en $\frac{x-\mu}{\sigma}$, que está tabulada. Si $Z\sim N(0,1)$ (habitualmente se utiliza Z para denotar una distribución $N(0,1)$), se tiene

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

es decir

$$F_{N(\mu,\sigma)}(x) = F_{N(0,1)}\left(\frac{x-\mu}{\sigma}\right)$$

2.1 Medidas características

- Media: $E(X) = \mu$
- Varianza: $V(X) = \sigma^2$

2.2 Relación entre $N(0,1)$ y $N(\mu,\sigma)$

- Si $Z \in N(0,1)$ entonces $X = \sigma Z + \mu$ con $\sigma > 0$, es $N(\mu,\sigma)$.
- Inversamente, si $X \in N(\mu,\sigma)$ entonces $Z = \frac{X-\mu}{\sigma} \in N(0,1)$. A esta transformación se le llama **tipificación**.

Así,

$$X \in N(\mu,\sigma) \Leftrightarrow \frac{X-\mu}{\sigma} \in N(0,1)$$

Ejemplo 1: Dada una variable aleatoria $X \sim N(\mu,\sigma)$ se calcula la probabilidad de un intervalo (a,b) de la siguiente forma:

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

donde $Z \sim N(0,1)$.

Ejemplo 2: Sea $X \sim N(\mu,\sigma)$. Se pueden conocer exactamente las siguientes probabilidades:

$$P(\mu - \sigma \leq X \leq \mu + \sigma) = P(-1 \leq Z \leq 1) = 0{,}6826$$

$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = P(-2 \leq Z \leq 2) = 0{,}9544$$

$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = P(-3 \leq Z \leq 3) = 0{,}9974$$

Estas probabilidades indican que existe gran concentración de valores alrededor de la media.

Así, basta estudiar la $N(0,1)$, cuya función de densidad es

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \qquad \forall z \in \mathbb{R}$$

La función de distribución es

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{z^2}{2}} dz$$

que verifica $F(z) = 1 - F(-z)$ por ser simétrica respecto del 0.

Como $e^{-\frac{1}{2}z^2}$ no tiene primitiva, la función de distribución está tabulada para distintos valores de z, en las denominadas tablas de áreas acumuladas. Se pueden encontrar distintos tipos de tablas:

1. Tablas de colas:
   Dan las áreas de las colas de la $N(0,1)$.
   $$P(|Z| \geq z_1) = P(-\infty < Z \leq -z_1) + P(z_1 \leq Z < \infty) = \frac{2}{\sqrt{2\pi}} \int_{z_1}^{\infty} e^{-\frac{z^2}{2}} dz$$
2. Tablas de áreas acumuladas:
   a) Áreas de la izquierda: $F(z_1) = P(Z \leq z_1)$
   b) Áreas de la derecha: $P(Z \geq z_1)$
3. Tablas de otro tipo: $P(0 < Z \leq z_1)$

Conocida una tabla, se pueden construir las otras. En las tablas suele aparecer en la columna décimas y en las filas centésimas, que forman la abscisa positiva de la $N(0,1)$.

Ejemplo 3: Conocida $F(z) = P(Z \leq z)$ se puede calcular:

$$P(Z \leq -z) = P(Z \geq z) = 1 - P(Z \leq z) = 1 - F(z)$$

$$P(Z \geq z) = 1 - F(z)$$

$$P(Z \geq -z) = P(Z \leq z) = F(z)$$

$$P(-z \leq Z \leq z) = F(z) - F(-z) = F(z) - \big(1 - F(z)\big) = 2F(z) - 1$$

Ejemplo 4: Un tubo electrónico tiene una distribución de vida Normal 280 horas de media y desviación típica σ. ¿Cuál debe ser el valor máximo que debe alcanzar σ si se quiere que el tubo tenga una probabilidad 0,8 de vivir entre 240 y 320 horas?

Se define la variable aleatoria

$$X = tiempo\ de\ vida\ de\ un\ tubo\ electrónico\ (en\ horas)$$

cuya distribución es $N(280, \sigma)$ y hay que obtener el valor de σ tal que

$$P(240 < X < 320) = 0{,}8$$

Tipificando se obtiene

$$P(240 < X < 320) = P\left(\frac{240 - 280}{\sigma} < \frac{X - 280}{\sigma} < \frac{320 - 280}{\sigma}\right) =$$

$$= P\left(\frac{40}{\sigma} < Z < -\frac{40}{\sigma}\right) = 0{,}8$$

Como esa probabilidad es igual a

$$2P\left(Z \leq \frac{40}{\sigma}\right) - 1 = 0{,}8$$

se busca σ tal que $P\left(Z \leq \frac{40}{\sigma}\right) = F_Z\left(\frac{40}{\sigma}\right) = 0{,}9$. En la tabla de la $N(0{,}1)$ se debe buscar el valor $z_{0,9}$ que verifica $F_Z\left(z_{0,9}\right) = 0{,}9$. Se obtiene $z_{0,9} = 1{,}2816$ (se ha utilizado una tabla con la función de distribución y se ha interpolado linealmente entre los valores $z_{0,8997} = 1{,}28$ y $z_{0,9015} = 1{,}29$), y

$$\frac{40}{\sigma} = 1{,}2816 \Rightarrow \sigma = 31{,}21098$$

2.3 Relación entre Binomial, Poisson y Normal

La distribución Normal puede utilizarse para aproximar probabilidades de variables Binomiales y de Poisson.

$$X \in B(n,p) \Longrightarrow \frac{X - np}{\sqrt{npq}} \xrightarrow[n\to\infty]{} Z \in N(0{,}1)$$

Ejemplo 5: Una variable aleatoria $X \in B(100; 0{,}06)$ se aproximaría a

$$Y \in N(np = 100 \cdot 0{,}06 = 6, \sqrt{npq} = \sqrt{100 \cdot 0{,}06 \cdot 0{,}94} = 2{,}37)$$

Como la distribución Normal es continua, ¿cómo se pueden calcular probabilidades discretas? Se utiliza la corrección de continuidad que tiene en cuenta que el número n equivale al intervalo continuo $(n - 0{,}5;\ n + 0{,}5)$. Así, $X \in B(n,p)$ se aproxima a $Y \in N\left(np, \sqrt{npq}\right)$ significa en la práctica:

$$P(a \leq X \leq b) \cong P(a - 0{,}5 \leq Y \leq b + 0{,}5)$$

Además,

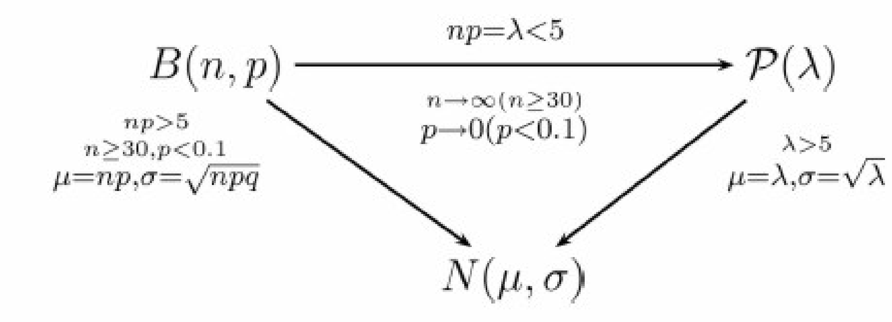
$$P(X \leq b) \cong P(Y \leq b + 0{,}5)$$

$$P(X \geq a) \cong P(Y \geq a - 0{,}5)$$

$$P(X = a) \cong P(a - 0{,}5 \leq Y \leq a + 0{,}5)$$

Cuando en la distribución de Poisson $\mathcal{P}(\lambda)$, el parámetro λ es superior a 5, $X \sim \mathcal{P}(\lambda)$, se puede aproximar con la variable aleatoria $Y \in N\left(\mu = \lambda, \sigma = \sqrt{\lambda}\right)$.

Resumiendo,



Ejemplo 6: Un vivero prepara pedidos de la planta Actinida (kiwi). Se necesita planta macho y planta hembra para la fructificación. El vivero estima que el 6% de las plantas son machos. Si se realiza un pedido de 100 plantas,

a) ¿Cuál es la probabilidad de que no se consiga la fructificación?
b) ¿Cuál es el número esperado de plantas macho?
c) Si para asegurar la fructificación de todas las plantas hembra se necesita que al menos el 10% de las plantas sean machos, ¿cuál es la probabilidad de que fructifiquen todas las plantas hembras?

Se define la variable aleatoria

$$X = número\ de\ plantas\ macho\ que\ hay\ entre\ las\ 100\ plantas \sim B(100; 0{,}06)$$

a) $P(\text{no conseguir fructificación}) = P(\text{ningún macho o todos machos}) = P(X=0) + P(X=100) = 0{,}94^{100} + 0{,}06^{100} = 0{,}0020548$
b) $E(X) = np = 100 \cdot 0{,}06 = 6$ plantas
c) El 10% de 100 es 10, con lo que se pide $P(X \geq 10)$, probabilidad que se puede calcular utilizando la aproximación por la distribución Normal, ya que se cumplen las condiciones necesarias: $n = 100 > 30$, $p = 0{,}06 < 0{,}1$ y $np = 100 \cdot 0{,}06 = 6 > 5$. Así, si $Y \sim N(6, \sqrt{npq} = \sqrt{5{,}64} = 2{,}3748)$ se tiene

$$P(X \geq 10) \cong P(Y \geq 9{,}5) = P\left(Z \geq \frac{9{,}5-6}{\sqrt{5{,}64}}\right) = P(Z \geq 1{,}4737) = 0{,}0708$$

2.4 Distribución Normal Truncada

Como ya se ha mencionado, una de las aplicaciones más habituales de la distribución Normal es la representación de medidas físicas. Muchas de estas medidas sólo están definidas para los valores positivos o en un determinado intervalo. Para estos casos se utiliza la Normal truncada. A continuación se muestra, en general, cómo se define la distribución de una variable aleatoria truncada.

Sea X variable aleatoria y sea T un subconjunto de $\mathbb{R}$, **la función de distribución de *X* truncada en *T*** es $P(X \leq x | X \in T)$ que será:

a) Si X es discreta:

$$P(X = x | X \in T) = \frac{P(X = x, X \in T)}{P(X \in T)} = \frac{P(X = x)}{\sum_{t \in T \cap D_x} p(t)} \quad \text{si } x \in T$$

donde D_X es el soporte de X, es decir, $D_X = \{x \in \mathbb{R} \,/\, p(x) > 0\}$.

En el resto es 0.

b) Si X es continua:

$$P(X \leq x | X \in T) = \frac{P(X \leq x, X \in T)}{P(X \in T)} = \frac{\int_{(-\infty,x] \cap T} f(y)dy}{\int_T f(y)dy}$$

Es decir, la función de densidad de la distribución truncada es

$$\frac{f(x)}{\int_T f(y)dy} \quad \text{si } x \in T$$

Particularizando para $X \sim N(\mu, \sigma)$ se tiene la distribución Normal truncada.

Ejemplo 7: Obtener la función de densidad de la Normal truncada en el 0, es decir, la distribución de $X | X \geq 0$.

$$f(x) = \begin{cases} k \dfrac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} & \text{si } x \geq 0 \\ 0 & \text{en el resto} \end{cases}$$

donde

$$k = \frac{1}{\int_0^\infty \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(t-\mu)^2}{\sigma^2}} dt}$$

Si $\mu > 3\sigma$ entonces $k \approx 1$ y se omite la constante k.

Ejemplo 8: En una población el cociente intelectual sigue una distribución Normal de media 100 y de varianza 256. Se considera que un estudiante de Informática terminará la carrera si su cociente intelectual es superior a 110. Por otra parte se clasifica a una persona como muy inteligente si su cociente es superior a 132. Calcular la proporción de *muy inteligentes* entre los ingenieros en Informática.

Se define la variable aleatoria

$$X = cociente\ intelectual\ de\ la\ población \sim N(100, \sqrt{256} = 16)$$

y sea A el suceso que representa “terminar la carrera de Informática”. Se pide $P(X \geq 132|A)$.

$$P(X \geq 132|A) = P(X \geq 132|X \geq 110) = \frac{P(X \geq 132, X \geq 110)}{P(X \geq 110)} = \frac{P(X \geq 132)}{P(X \geq 110)} =$$

$$= \frac{P\left(\frac{X-100}{16} \geq \frac{132-100}{16}\right)}{P\left(\frac{X-100}{16} \geq \frac{110-100}{16}\right)} = \frac{P(Z \geq 2)}{P(Z \geq 0{,}625)} = \frac{1-0{,}9772}{1-0{,}7324} = 0{,}0852$$

3. Distribución Gamma

Se dice que una variable aleatoria X sigue una **distribución Gamma** de parámetros $\lambda, p > 0$ ($\Leftrightarrow X \in \gamma(\lambda, p)$) si su función de densidad es de la forma:

$$f(x) = \frac{\lambda^p}{\Gamma(p)} x^{p-1} e^{-\lambda x} \qquad \text{si} \quad x > 0$$

donde $\Gamma(p)$ es la función Gamma definida por

$$\Gamma(p) = \int_0^{\infty} x^{p-1}\, e^{-x} dx$$

Se puede demostrar que esta función es continua, que converge para $p > 0$ y que converge uniformemente respecto a p. Además, tiene las siguientes propiedades:

a) $\Gamma(1) = 1$
b) $\Gamma(p) = (p-1)\Gamma(p-1)$
c) $\Gamma(p) = (p-1)!$ si $p \in \mathbb{N}$
d) $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$
e) $\int_0^{\infty} x^{p-1}\, e^{-\lambda x} dx = \frac{\Gamma(p)}{\lambda^p}$ siendo $\lambda = b + ic$ con $b > 0$

De esta última propiedad se deduce fácilmente que $f(x)$ es función de densidad.

La representación de f depende de los valores de λ y p, como se puede observar en la figura 2. Si $p \leq 1$ tiene perfil de una J transpuesta y si $p > 1$ tiene un pico en $x = \frac{1}{\lambda}(p-1)$. p es un factor de forma y λ de escala.

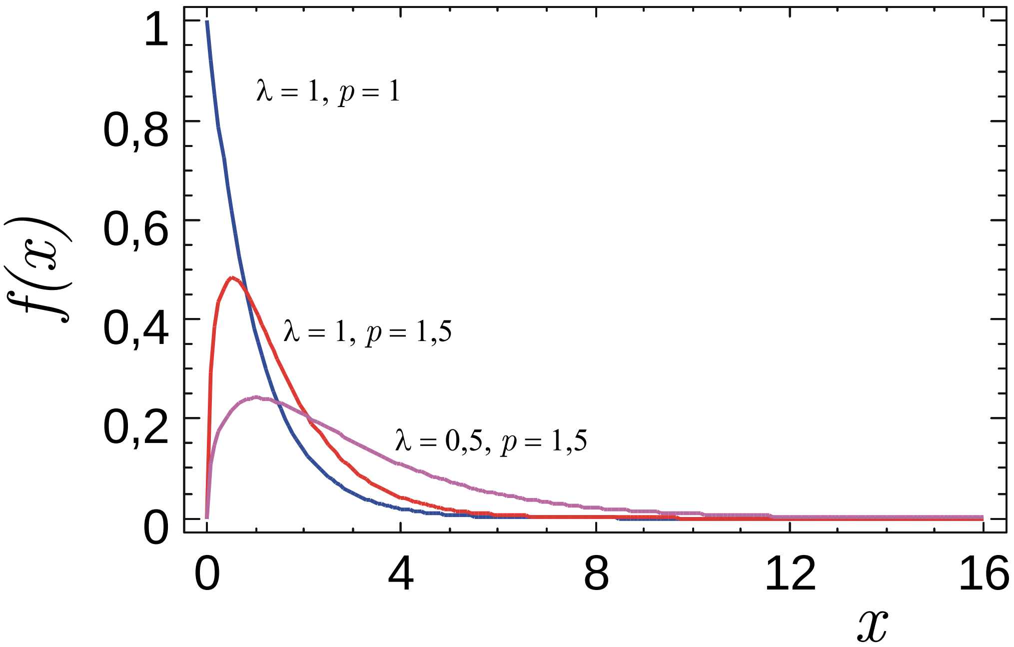


Figura 2: Distribución Gamma para diferentes valores de λ y p.

Observación: La distribución Gamma $\gamma(\lambda, p)$ es reproductiva respecto de p, es decir, dadas dos variables aleatorias X, Y independientes con $X \sim \gamma(\lambda, p_1)$, $Y \sim \gamma(\lambda, p_2)$ entonces

$$X + Y \sim \gamma(\lambda, p_1 + p_2)$$

Su función de distribución es

$$F(x) = \frac{\lambda^p}{\Gamma(p)} \int_0^x t^{p-1} e^{-\lambda t} dt \ , \quad 0 < x < \infty$$

3.1 Medidas características

- Media:

$$E(X) = \frac{p}{\lambda}$$

- Varianza:

$$V(X) = \frac{p}{\lambda^2}$$

3.2 Distribución de Erlang

Si p es entero, la distribución Gamma se conoce como **distribución de Erlang** y la función de distribución se puede expresar como

$$F(x) = 1 - \left(1 + \lambda x + \frac{(\lambda x)^2}{2!} + \cdots + \frac{(\lambda x)^{p-1}}{(p-1)!}\right) e^{-\lambda x}$$

integrando por partes.

Esta distribución se utilizó por primera vez en problemas de tráfico en líneas telefónicas.

- Existe una asociación entre los modelos de Poisson y Erlang:
  Si el número de sucesos aleatorios independientes que ocurren en un lapso específico, es una variable de Poisson con frecuencia constante de ocurrencia igual a λ, entonces, dado $p \in \mathbb{N}$, el tiempo de espera hasta que ocurre el p-ésimo suceso de Poisson, tiene una distribución de Erlang de parámetros λ y p.
  Se tiene que la distribución de Erlang $\gamma(\lambda, p)$ es el modelo para el tiempo de espera hasta que ocurre el p-ésimo evento de Poisson, y la distribución de Poisson $\mathcal{P}(\lambda x)$ es el modelo para el número de sucesos independientes que ocurren en un tiempo x, encontrándose éste distribuido de acuerdo con el modelo de Erlang.

Algunos ejemplos de modelos que siguen una distribución Gamma, son los siguientes:

- Tiempo aleatorio de fallo de un sistema que falla sólo si de manera exacta los componentes fallan y el fallo de cada componente ocurre a una frecuencia constante λ por unidad de tiempo.
- Problemas de líneas de espera para representar el intervalo total para completar una reparación si ésta se realiza en subestaciones de manera independiente y con una frecuencia constante λ.
- Si se considera una pieza metálica que se encuentra sometida a cierta fuerza, de manera que se romperá después de aplicar un número específico de ciclos de fuerza, donde los ciclos ocurren de manera independiente y a una frecuencia promedio; el tiempo que debe transcurrir antes de que el material se rompa se distribuye según una gamma.

3.3 Distribución Exponencial

Es también un caso particular de la distribución Gamma y es muy importante por sus múltiples usos y aplicaciones.

Una variable aleatoria X sigue una **distribución Exponencial** de parámetros $\lambda > 0$ ($\Leftrightarrow X \in Exp(\lambda)$) si su función de densidad es

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{en el resto} \end{cases}$$

Como se puede observar, la distribución exponencial es una distribución Gamma con $p = 1$, es decir, $Exp(\lambda) \equiv \gamma(\lambda, p = 1)$.

La función de distribución es $F(x) = 1 - e^{-\lambda x}$ si $x > 0$ (0 en el resto).

Sus medidas características son:

- Media:

$$E(X) = \frac{1}{\lambda}$$

- Varianza:

$$V(X) = \frac{1}{\lambda^2}$$

La distribución exponencial resulta al considerar en un proceso de Poisson la variable continua $T = tiempo\ entre\ la\ ocurrencia\ de\ dos\ sucesos\ consecutivos$:

$$P(T > t) = P\big(0 \text{ sucesos en } (0, t)\big) = e^{-\lambda t} \Longrightarrow F(t) = 1 - e^{-\lambda t} \Longrightarrow f(t) = F'(t) = \lambda e^{-\lambda t}$$

- Una propiedad fundamental de la distribución exponencial es que no tiene memoria: La probabilidad de ocurrencia de sucesos presentes o futuros no depende de lo que haya ocurrido en el pasado.
  Así, que una unidad falle en un lapso específico no depende del tiempo que la unidad haya estado en operación, sino sólo depende de la duración del lapso. Esto es, dada $X \sim Exp(\lambda)$,

$$P(X \geq x + h | X \geq x) = P(X \geq h)$$

Ejemplo 9: El número de personas necesarias para la visita turística a una cueva es de 25. Si el número de personas que solicita dicho servicio sigue una distribución de Poisson de media 90 personas por hora y el último grupo partió a las siete, ¿a qué hora se espera que salga el siguiente? Si alguien llega a las 7:25 y con él hay 24 personas, ¿cuál es la probabilidad de que la visita tarde más de dos minutos en empezar?

Se definen las variables aleatorias

$$X = número\ de\ personas\ que\ llegan\ por\ minuto \sim \mathcal{P}\left(\lambda = \frac{90}{60} = \frac{3}{2} = 1{,}5\right)$$

$$Y = tiempo\ que\ tardan\ en\ juntarse\ 25\ personas \sim \gamma(1{,}5; 25)$$

Se observa que Y sigue una distribución de Erlang de parámetros $\lambda = 1{,}5$ y $p = 25$ ya que es el tiempo de espera hasta que ocurre el suceso 25 de Poisson. Su esperanza es

$$E(Y) = \frac{p}{\lambda} = \frac{25}{1{,}5} = 16{,}6 \text{ minutos}$$

Con lo que se espera que el siguiente grupo salga entre las 7:16 y 7:17 horas.

Se considera ahora la variable aleatoria

$$T = tiempo\ (en\ minutos) que\ tarda\ en\ empezar\ la\ visita$$

que sigue una distribución Exponencial de parámetro $\lambda = 1{,}5$, ya que representa el tiempo que tarda en llegar una nueva persona, es decir, el tiempo hasta la ocurrencia del siguiente suceso de Poisson. Así, la probabilidad de que la visita tarde más de dos minutos en empezar es

$$P(T > 2) = 1 - F_T(2) = e^{-1{,}5 \cdot 2} = e^{-3} = 0{,}049787$$

4. Distribución Beta

La distribución Beta juega un gran papel en la estadística bayesiana que se comentará en unidades posteriores. Se utiliza para representar variables ficticias cuyos valores se encuentran restringidos a un intervalo de longitud finita. Otros ejemplos de su uso son:

- Distribución de artículos defectuosos sobre un intervalo de tiempo específico.
- Evaluación de programas y técnicas de revisión.
- Distribución de la proporción de valores que deben caer entre dos observaciones extremas.

Una variable aleatoria X sigue una **distribución Beta** de parámetros $p, q > 0$ ($X \in Beta(p,q)$) si su función de densidad es

$$f(x) = \frac{1}{\beta(p,q)} x^{p-1}(1-x)^{q-1} \text{ si } 0 < x < 1$$

y 0 en el resto, donde $\beta(p,q)$ es la función beta definida por

$$\beta(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1} dx \quad \text{con } p,q > 0$$

y que tiene las siguientes propiedades:

a) $\beta(1,1) = 1$
b) $\beta\left(\frac{1}{2}, \frac{1}{2}\right) = \pi$

c) $\beta(p,q) = \beta(q,p)$

d) $\beta(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$

La representación gráfica de la función de densidad de la distribución Beta para diferentes valores de los parámetros se muestra en la figura 5. p y q son parámetros de perfil. La distribución es simétrica sólo si $p = q$.

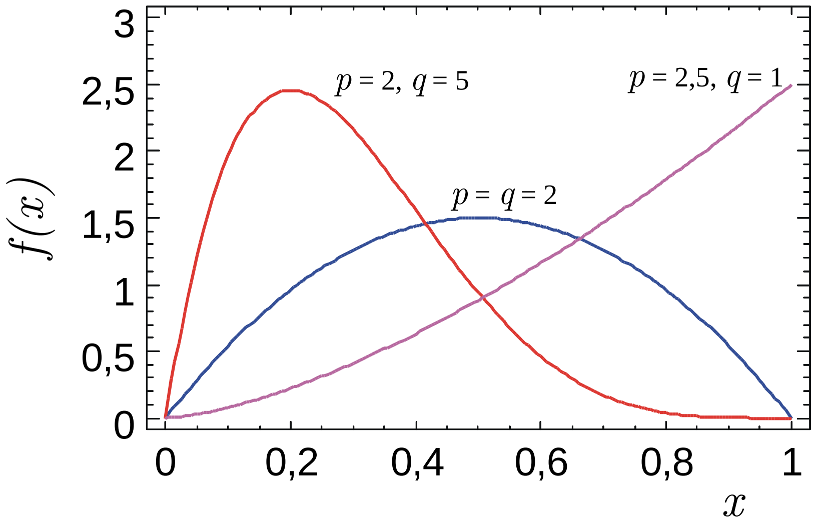


Figura 3: Distribución Beta para diferentes valores de p y q

4.1 Medidas características

- Media:

$$E(X) = \frac{p}{p+q}$$

- Varianza:

$$V(X) = \frac{pq}{(p+q)^2(p+q+1)}$$