TEMA 2

VARIABLE ALEATORIA CONTINUA

2.1. INTRODUCCIÓN

Como ya se comentó en el tema anterior, 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 estudio de la distribución de las variables continuas es más sutil que el de las discretas. Una de las causas es que para una variable continua la probabilidad de que la variable tome un determinado valor es cero, porque sería imposible en una suma infinita tomar el valor 1 que es la probabilidad total del espacio. Por tanto, aquí la noción de función de probabilidad no da resultados útiles. En su lugar utilizaremos la ***Función de Densidad*** de probabilidad que proporciona un medio para determinar la probabilidad de cualquier intervalo. A continuación, extenderemos el concepto de esperanza matemática al caso continuo, cuya definición e interpretación coincide con lo visto en el caso discreto y solo modifica su forma de calcularlo. Así tendremos las principales características de una variable aleatoria continua.

Por último, analizaremos los modelos probabilísticos continuos notables, cuyo interés reside en la capacidad de describir comportamientos genéricos de distintas magnitudes aleatorias que resultan semejantes según ciertas pautas. Nos encontramos así con grandes «familias probabilísticas» designadas con nombres propios que incluyen como casos particulares numerosos fenómenos, incorporando sus rasgos diferenciales mediante parámetros.

2.2. FUNCIÓN DE DENSIDAD

Para definir el concepto de función de densidad es conveniente visualizar la representación de una función de densidad como el límite del polígono de frecuencias relativas cuando el número de repeticiones del experimento aumenta. De esta forma utilizando el concepto frecuentista de la probabilidad, las frecuencias se estabilizan en torno a un valor que llamamos probabilidad y que mide la incertidumbre. Por otro lado, reduciremos la amplitud de las clases para conseguir una medida puntual en cada valor donde está definida la variable aleatoria.

Consideramos los tiempos de espera hasta que llegar un autobús que pasa cada diez minutos. El rango de valores está comprendido entre 0 y 10 minutos y comenzaremos observando 100 tiempos y aumentaremos hasta 50.000. Por otro lado, iremos disminuyendo la amplitud de las clases desde 1 minuto hasta 0,125 minutos.

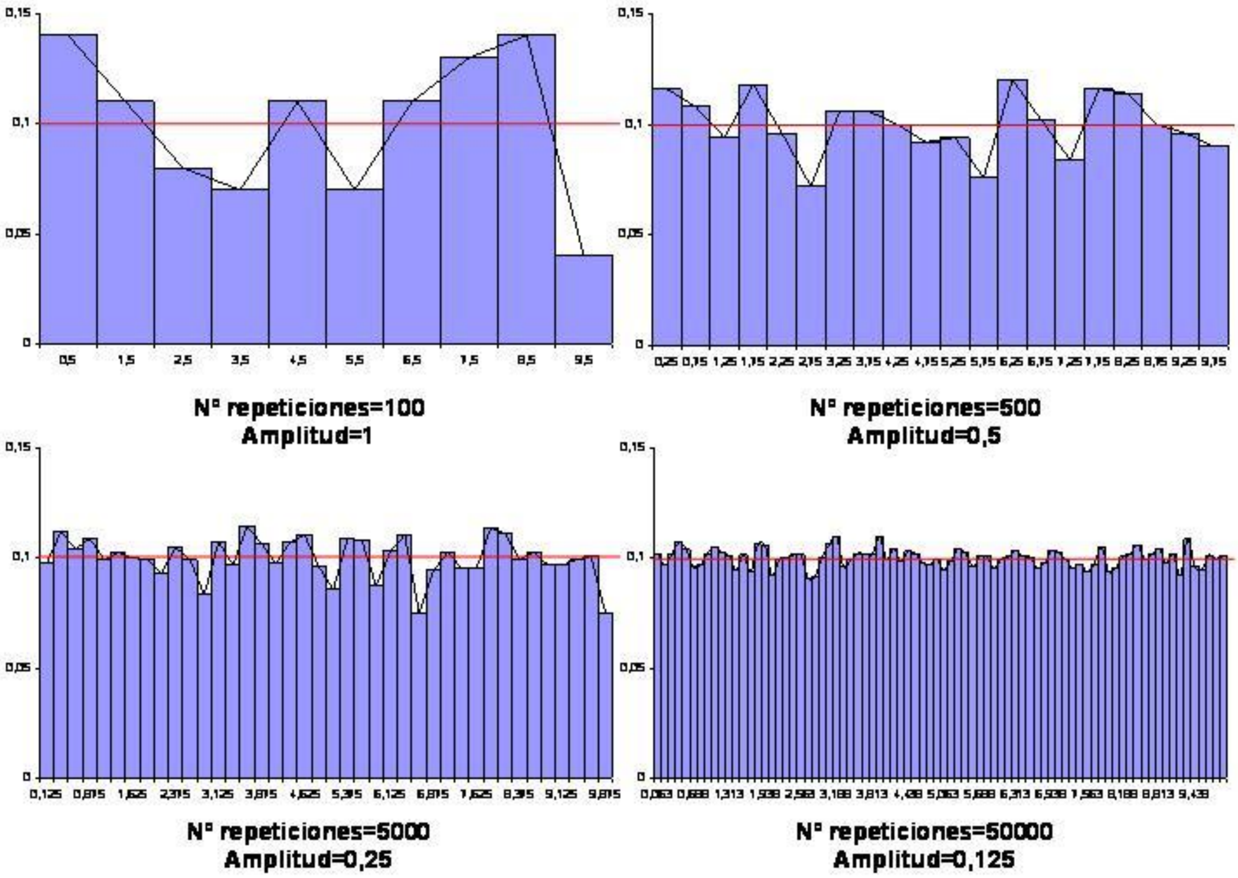


Figura 1.1: Aproximación de una función de densidad mediante histogramas

Se observa que el polígono de frecuencias se suaviza cada vez más y tiende a una función continua, cuando la amplitud de los intervalos tiende a cero, que llamaremos **función de densidad**.

El histograma representaba frecuencias mediante áreas, análogamente, la función de densidad expresa probabilidades por áreas. La frecuencia de un intervalo es el área del rectángulo del histograma que corresponde con dicho intervalo, así pues en el límite la probabilidad será el área encerrada por la densidad en el intervalo. Por lo tanto, si f(x) es la función de densidad de una variable aleatoria continua X entonces la forma de calcular la probabilidad, es decir, el área de un intervalo (a,b) será:

$$P\{a<X<B\}=\int_a^b f(x)dx$$

Además, la función de densidad conserva las propiedades básicas del polígono de frecuencias: es no negativa y el área total que contiene es uno. La función de densidad es la cantidad de probabilidad que existe por punto y se interpreta del siguiente modo: sus valores más altos corresponden a las zonas en las que es más probable que aparezcan resultados del fenómeno aleatorio; los valores con menor altura son aquellos donde es menos probable – o más raro – que tome valores la variable. Es erróneo entender la función de densidad como la probabilidad de que la variable tome un valor específico, pues ésta es siempre cero para cualquier variable continua (el área que queda encima de un punto es cero).

Más formalmente, podemos escribir que si f(x) es la función de densidad de una variable aleatoria continua x, se cumple que:

1. $f(x) \geq 0, \forall x \in R$
2. $\int_{-\infty}^{+\infty} f(x)dx = 1$
3. $P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b) = \int_a^b f(x)dx$

Ejemplo 2.1

Sea X = tiempo de espera (en minutos) de una persona en una estación de autobuses a la que llegan los buses con una frecuencia de 30 minutos

Se tiene que si $0 \leq a < b \leq 30$ entonces $P(a \leq X \leq b) = \dfrac{b-a}{30} = \int_a^b \dfrac{1}{30}dx$, por lo que X es una variable aleatoria continua con función de densidad $f(x) = \dfrac{1}{30}$ $0 \leq x \leq 30$

2.3. 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 en estudio tome un valor numérico igual o inferior a dicho punto.

En distribuciones continuas la función de distribución en un punto viene medida por el área bajo la función de densidad a la izquierda del punto, y se obtiene integrando la función de densidad entre el extremo inferior del soporte de la distribución de probabilidad, y el punto.

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt$$

o equivalentemente, se tiene que:

$$f(x) = \frac{dF(x)}{dx} = \lim_{h \to 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \to 0} \frac{P(x \leq X \leq x+h)}{h}$$

Se puede demostrar que la función de distribución verifica las mismas propiedades que en el caso discreto:

1. $0 \leq F(x) \leq 1, \ \forall \ x \in R$
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. $P(a < X \leq b) = F(b) - F(a)$ (probabilidad de un intervalo) , aunque teniendo en cuenta que en el caso continuo la probabilidad de un punto concreto es cero entonces se escribe como:

$$P(a < X \le b) = P(a \le X < b) = P(a \le X \le b) = P(a < X < b) = F(b) - F(a) =$$

$$\int_{-\infty}^{b} f(t)dt - \int_{-\infty}^{a} f(t)dt$$

Ejemplo 2.1 (continuación)

En este caso la función de distribución de X viene dada por:

$$F(x) = \begin{cases} 0 & 0 < x \\ \int_0^x \frac{1}{30} du = \frac{x}{30} & 0 \le x < 30 \\ 1 & x \ge 30 \end{cases}$$

Es, por lo tanto, una función continua no decreciente y que su derivada verifica que

$$F'(x) = \frac{1}{30} = f(x) \qquad 0 \le x \le 30$$

2.4. CARACTERÍSTICAS DE UNA DISTRIBUCIÓN: ESPERANZA MATEMÁTICA

Cada distribución continua de probabilidad tiene asociadas medidas similares a las medidas descriptivas de las distribuciones de frecuencias que resume su comportamiento y nos permite reconocer sus características.

Estas medidas se basan en el concepto de ***Esperanza Matemática***, cuya definición e interpretación coincide con la dada para las variables aleatorias discretas. 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:

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

Si X es una variable aleatoria continua con función de densidad f(x), la esperanza matemática se define:

$$E(X) = \int_{-\infty}^{+\infty} xf(x)dx$$

Ejemplo 2.1 (continuación)

En este caso se tiene que $E(X) = \int_0^{30} x \frac{1}{30} dx = \left[\frac{x^2}{60}\right]_0^{30} = \frac{900}{60} = 15$

Podemos generalizar la definición de la esperanza de X a variables aleatorias que se obtienen de ella, es decir, a transformaciones g(X). En el caso de que X sea una variable aleatoria continua con función de densidad f(x), entonces:

$$E(g(X)) = \int_{-\infty}^{+\infty} g(x)f(x)dx$$

Ejemplo 2.1 (continuación)

En este caso se tiene que $E(X^2) = \int_0^{30} x^2 \frac{1}{30} dx = \left[\frac{x^3}{90}\right]_0^{30} = \frac{27000}{90} = 300$

Algunas propiedades interesantes de la esperanza matemática, que provienen de su definición como suma o integral, 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
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:

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

Si la variable es continua se obtiene:

$$Var(X) = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)$$

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. $Var(aX + b) = a^2 Var(X)$
6. $Var(X \pm Y) = Var(X)+Var(Y)$ si X e Y son variables aleatórias independientes

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 2.1 (continuación)

La varianza de X vendrá dada por $Var(X) = E(X^2) - (E(X))^2 = 300 - 15^2 = 75$ y, por lo tanto, la desviación típica será $D(X) = \sqrt{75} = 8{,}66$

2.5. DISTRIBUCIONES CONTINUAS NOTABLES

Cuando estudiamos variables aleatorias continuas existirán distribuciones, igual que ocurría en el caso discreto, que se repiten o son más habituales que otras y cuya diferencia reside en unos parámetros que modifican la escala o la forma de la función de densidad. Por lo tanto, en este apartado presentaremos las distribuciones notables que describen comportamientos genéricos de distintas magnitudes aleatorias que resultan semejantes según ciertas pautas. Nos encontramos así con grandes «familias probabilísticas» designadas con nombres propios que incluyen como casos particulares numerosos fenómenos, incorporando sus rasgos diferenciales mediante parámetros.

2.5.1. DISTRIBUCIÓN UNIFORME

Iniciamos estos modelos de distribuciones de probabilidad con la *distribución Uniforme Continua*, la cual es quizás la más sencilla entre las distribuciones continuas. Una variable aleatoria sigue una distribución uniforme si su masa de probabilidad está repartida uniformemente a lo largo de su soporte. En el caso de una distribución de tipo continuo, esto significa que su función de densidad es constante.

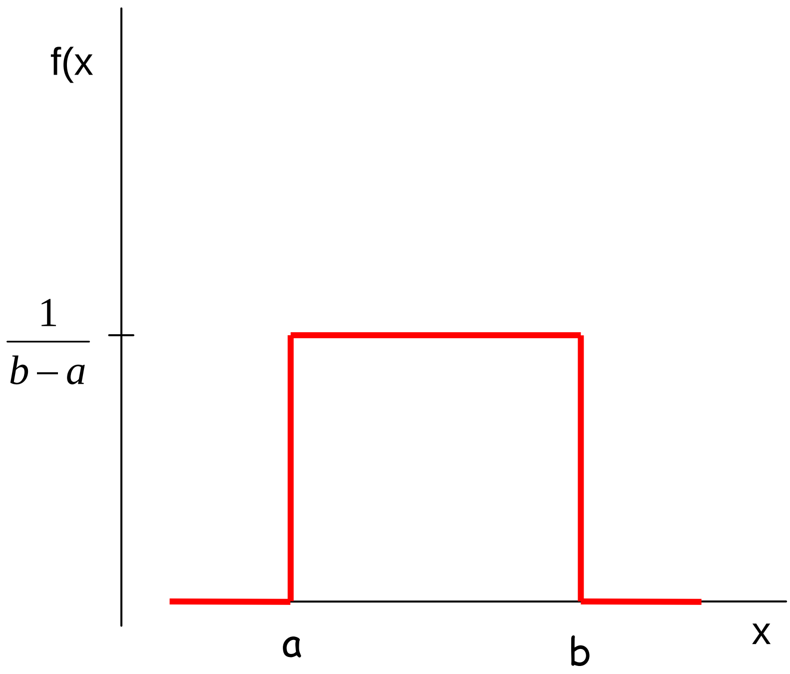
Esta distribución será útil para modelar situaciones en las que no existe ninguna evidencia a favor de determinados resultados, por lo cual resulta aplicable el principio de indiferencia (el tiempo en minutos que debe esperar al autobús una persona cuando va a trabajar en un día de tráfico normal sigue una distribución uniforme en un intervalo de 0 a 15 minutos, por ejemplo).

Muchas disposiciones de datos numéricos en el mundo de los negocios concuerdan con una distribución uniforme, pues proporciona una representación adecuada para redondear las diferencias que surgen al medir cantidades físicas entre los valores observados y los reales. Por ejemplo, si el peso de un determinado elemento se redondea al kilogramo más cercano, es común que el error de redondeo se encuentre distribuido uniformemente en el intervalo (-0,5; 0,5).

En general, si a < b$\in$**R** diremos que una variable aleatoria X continua se *distribuye uniformemente en (a,b)* y lo pondremos como X ~ U(a,b) si su soporte es (a,b) y su función de densidad viene dada por:

$$f(x) = \frac{1}{b-a} \text{ si } x \in (a,b) \text{ y 0 en el resto}$$

Su gráfico es el siguiente:



Su media y su varianza vienen dadas

$$E(X) = \frac{a+b}{2} \text{ y } V(X) = \frac{(b-a)^2}{12}$$

Un resultado notable respecto a esta distribución, que constituye la raíz de todo método de simulación estocástica es el siguiente:

Si X es una variable aleatoria con función de distribución F(x) y hacemos la transformación U = F(X) entonces la nueva variable aleatoria U sigue una distribución uniforme continua en el intervalo (0,1) y la representamos por U ~ U(0,1).

Ejemplo 2.2

Un hombre acude a una estación de autobuses en la que pasa un autobús cada media hora. Sea X = tiempo de espera (en minutos) ~ U(0,30). Calcular la probabilidad de que tenga que esperar más de 10 minutos

$$P(X>10) = \int_{10}^{30} \frac{1}{30}\,dx = \frac{20}{30} = \frac{2}{3}$$

2.5.2. DISTRIBUCIÓN EXPONENCIAL

Esta distribución aparece asociada a experimentos de Poisson que se desarrollan en el tiempo. Corresponde a la distribución de probabilidad de X = tiempo transcurrido entre dos ocurrencias de un suceso o también al tiempo transcurrido hasta la primera ocurrencia del suceso. Para deducir su distribución partimos de la variable Y= número de ocurrencias en una unidad de tiempo y su distribución es Poisson de media λ. A continuación, calculamos la probabilidad de que pase más de x minutos hasta la ocurrencia del suceso:

P(X > x) = P(hay que esperar más de x unidades de tiempo hasta la ocurrencia)

Si tenemos que esperar más de x unidades de tiempo significa que en x unidades de tiempo no ha ocurrido ningún suceso, por lo tanto, podremos calcular esa probabilidad con la Poisson. Sea Y_X=número de sucesos en x unidades de tiempo y su distribución es Poisson de media λx. Por lo tanto, utilizando la función de probabilidad de la Poisson tenemos:

$P(X > x) = P(Y_X = 0) = e^{-\lambda x}$

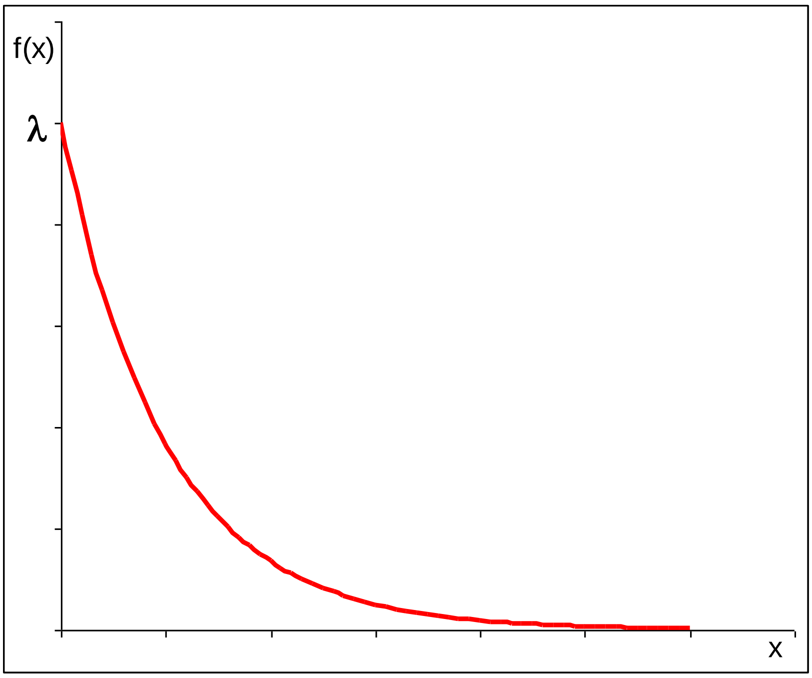
La función de distribución de X es el complementario hasta 1 de esta probabilidad, es decir,

$F(x) = P(X \leq x) = 1\text{-}P(X > x) = 1\text{-} e^{-\lambda x}$ siempre que x>0

Dada la función de distribución, podemos calcular la derivada y obtenemos la función de densidad de X:

$$f(x) = \lambda e^{-\lambda x} \text{ si } x>0$$

cuya gráfica corresponde con una exponencial y de ahí su nombre.



La distribución de X se denomina *distribución exponencial de parámetro* λ y lo denotaremos como $X\sim Exp(\lambda)$ con $\lambda>0$ representando número medio de éxitos por unidad de tiempo.

Este tipo de distribución es recomendable cuando el planteamiento resulta habitual en el ámbito económico-empresarial, en el que frecuentemente interesa conocer el período de tiempo necesario hasta que se presenta un determinado acontecimiento: aprobación de presupuestos, salida de una empresa a bolsa, contratación de un trabajador,…

Su media y su varianza vienen dados por:

$$E(X) = \frac{1}{\lambda} \text{ y } V(X) = \frac{1}{\lambda^2}.$$

Si el parámetro λ expresa la intensidad de aparición del fenómeno para el que estamos registrando los tiempos de espera, es natural que el tiempo medio entre dos apariciones del suceso sea el inverso a dicha intensidad.

La distribución exponencial resulta, también, muy útil para atacar problemas de filas de espera y colas. Cuando el tiempo de espera para pagar en la caja de un supermercado es aleatorio, esta incertidumbre puede representarse a menudo mediante una distribución exponencial. Así, se pueden modelar mediante la distribución exponencial las siguientes situaciones: duración de la prestación de un servicio, el tiempo entre llegadas sucesivas a una cola o punto de servicio, el tiempo de duración de algunos equipos, etc.

La distribución exponencial se trata de una distribución *sin memoria*, esto es, la probabilidad de que no se produzca un suceso durante un intervalo es independiente de que haya tenido lugar antes. Matemáticamente:

$$P(X > x+h \mid X > x) = P(X > h) \quad \forall x, h > 0$$

Supongamos, por ejemplo, la existencia de dos cabinas telefónicas, una de las cuales estaba ocupada bastante tiempo antes de llegar y la otra se había ocupado muy recientemente, entonces ambas cabinas tienen la misma probabilidad de desocuparse en cualquier momento, olvidando el tiempo T que llevaban ocupadas. En definitiva, la ley exponencial describe el tiempo de vida de cierto sistema en el que no hay envejecimiento o desgaste.

La comprobación de esta propiedad es muy sencilla, basta recordar como se calcula la probabilidad condicionada y la probabilidad de los sucesos bajo el modelo exponencial. La probabilidad condicionada se calcula como el cociente entre la probabilidad de la intersección de ambos resultados y la probabilidad del resultado que condiciona, es decir:

$$P(X > x+h \mid X > x) = \frac{P(X > x+h \cap X > x)}{P(X > x)} = \frac{P(X > x+h)}{P(X > x)}$$

Ahora, utilizando la probabilidad de la distribución exponencial, sustituimos y simplificamos:

$$\frac{P(X > x+h)}{P(X > x)} = \frac{e^{-\lambda(x+h)}}{e^{-\lambda x}} = e^{-\lambda h} = P(X > h)$$

Ejemplo 2.3

El tiempo que un dependiente de una tienda tarda en atender a un cliente se distribuye según una exponencial con media 10 minutos. Calcular la probabilidad de que el dependiente tarde más de media hora en atender a un cliente.

Sea X = tiempo de atención a un cliente (en minutos) ~ Exp(0.1)

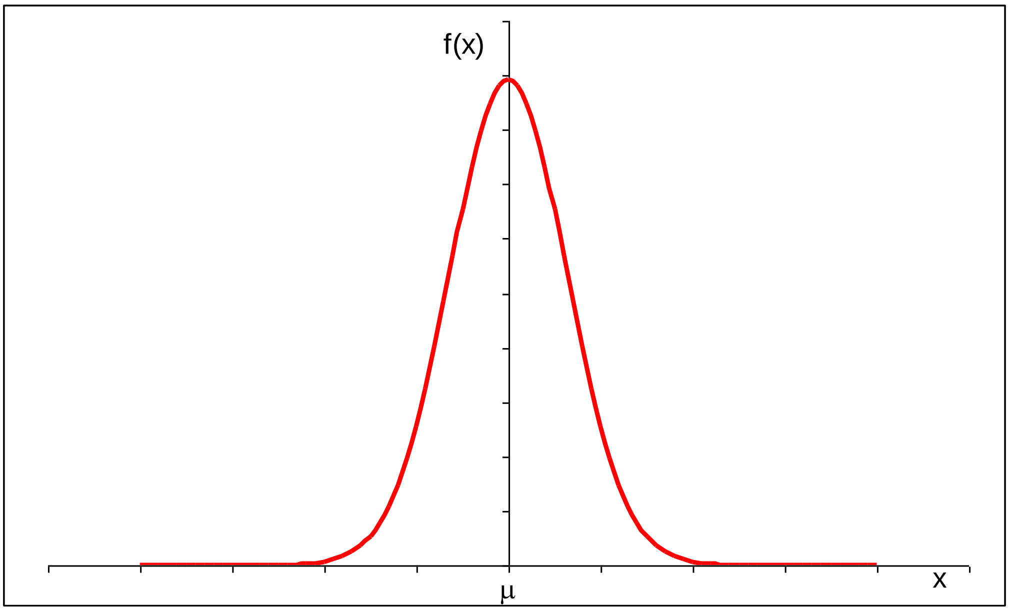
P(X>30) = $\int_{30}^{\infty} 0.1e^{-0.1x}dx = \left[-e^{-0.1x}\right]_{30}^{\infty}$ = e^{-3} = 0,0498

2.5.3. DISTRIBUCIÓN NORMAL

Diremos que una variable aleatoria X se distribuye según una *distribución normal de media* $\mu\in\mathbf{R}$ *y desviación típica* $\sigma > 0$ y lo pondremos como X ~ N(μ,σ) si su soporte es **R** y su función de densidad es:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \quad \text{si } x\in(-\infty,\infty)$$

Su gráfico es el siguiente:



Su media y su varianza vienen dadas por:

$$E(X) = \mu \text{ y } V(X) = \sigma^2$$

por lo tanto, está centrada en μ y σ es su desviación típica, indicando su dispersión.

Esta distribución tiene las siguientes propiedades

1) La curva de densidad es simétrica en torno a x = μ, teniendo un único máximo en dicho punto. Por lo tanto μ es, además, la mediana y la moda de esta distribución
2) La familia de distribuciones normales es cerrada bajo transformaciones lineales verificándose que, si Y = aX + b con a,b∈**R** se tiene que Y ~ N(aμ+b,|a|σ)

3) En particular, se sigue que, si tipificamos X, es decir, si $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$, distribución que recibe el nombre de *distribución normal estándar*. Por lo tanto, $X = \mu+\sigma Z$ y, por lo tanto, toda distribución normal se puede obtener, mediante un cambio de origen y escala, a partir de la distribución N(0,1).

4) La suma de v.a. Normales independientes es otra v.a. Normal cuya media es la suma de medias y cuya varianza es la suma de varianzas (¡tener cuidado porque utilizamos la desviación típica!):

$$\left.\begin{array}{r} X \sim N(\mu_X, \sigma_X) \\ Y \sim N(\mu_Y, \sigma_Y) \end{array}\right\} \text{v.a.independientes} \Rightarrow \mathrm{X} + \mathrm{Y} \sim \mathrm{N}\left(\mu_X + \mu_Y, \sqrt{\sigma_X^2 + \sigma_Y^2}\right)$$

La distribución normal proporciona una adecuada representación de las distribuciones de una gran cantidad de variables físicas. Algunos ejemplos son: los errores de medida de una balanza de precisión, la temperatura y el nivel de precipitaciones en Meteorología, características físicas y psicológicas de un individuo, etc. Además, esta distribución es un modelo muy útil para múltiples aplicaciones, como decisiones de inversión, gráficos para la mejora de la calidad de procesos, productos y servicios.

Esta distribución es, indudablemente, la más importante y la de mayor uso de todas las distribuciones continuas de probabilidad, motivado esto, entre otras, por las siguientes razones:

- Es la piedra angular en la aplicación de la inferencia estadística en el análisis de datos, puesto que las distribuciones de muchos estadísticos muestrales tienden hacia la distribución normal conforme crece el tamaño de la muestra.

- El Cálculo de Probabilidades para muchas leyes puede aproximarse bajo ciertas condiciones por la distribución normal.

- En gran cantidad de fenómenos empíricos, la distribución normal se aproxima a las distribuciones de frecuencias observadas.

Ejemplo 2.4

La rentabilidad diaria (en %) de una cotización del Banco A se distribuye según una normal de media 0 y desviación típica 10%

Calcular la probabilidad de que un día se obtenga una rentabilidad superior al 20%

Sea X = rentabilidad diaria ~ N(0,10). Se tiene que $Z = \frac{X-0}{10} \sim N(0,1)$

Nos piden P(X>20) = P(Z>2) = 1 – P(Z≤ 2) = 1 – 0,9772 = 0,0228

2.6. APROXIMACIÓN DE DISTRIBUCIONES DISCRETAS A UNA NORMAL

Una de las aplicaciones más importantes de la distribución normal, es que puede aproximar, bajo condiciones límite, a la distribución de gran cantidad de distribuciones discretas y continuas. Dos de las aproximaciones más utilizadas son la aproximación binomial-normal y Poisson-normal que describimos, brevemente, a continuación.

2.6.1 Aproximación Binomial-Normal

Sea X ~ Bi(n,p) y sea Y ~ $N\left(np, \sqrt{npq}\right)$ con q = 1-p.

Sean a, b $\in$ {0, 1, …, n}. Se verifica que:

$$P(a \leq X \leq b) \approx P(a - 0.5 \leq Y \leq b + 0.5)$$

si $n \to \infty$. En la práctica esta aproximación funciona bien si $n \geq 30$, $np > 5$ y $nq > 5$

Ejemplo 2.5

Un Banco recibe un 80% de solicitudes de trabajo por parte de personas solteras.

Si extraemos 200 solicitudes al azar, calcular la probabilidad de que más de 180 correspondan a personas solteras

Sea X = número de solicitudes de personas solteras ~Bi(200,0,8) ≈ $N\left(200x0{,}8, \sqrt{200x0{,}8x0{,}2}\right) = N(160{,}5{,}66)$

P(X>180) = 1 - P(X≤180) ≈ 1 - $P\left(Z \leq \frac{180{,}5-160}{5{,}66}\right) = 1$ - P(Z≤3,62) = 1 – 0,9999 = 0,0001

2.6.2 Aproximación Poisson-Normal

Sea $X \sim P(\lambda)$ y sea $Y \sim N(\lambda, \sqrt{\lambda})$

Sean $a,b \in \{0, 1, \ldots, \infty\}$. Se verifica que:

$$P(a \leq X \leq b) \approx P(a-0.5 \leq Y \leq b+0.5)$$

si $\lambda \to \infty$. Aproximaremos a la normal siempre que no lo tengamos en tablas y $\lambda > 30$.

Ejemplo 2.6

El número de aviones que llega a un aeropuerto en 1 hora se distribuye según una Poisson con media 20 aviones. Calcular la probabilidad de que en un día lleguen al aeropuerto más de 500 aviones

Sea X = número de aviones que llegan al aeropuerto en 1 día $\sim P(20x24) = P(480) \approx N(480, \sqrt{480}) = N(480{,}21{,}91)$

$P(X>500) = 1 - P(X \leq 500) \approx P\left(Z \leq \frac{500{,}5 - 480}{21{,}91}\right) = 1 - P(Z \leq 0{,}94) = 1 - 0{,}8264 =$

$0{,}1736$