CAPÍTULO 4
DISTRIBUCIONES DE PROBABILIDAD

4.1 INTRODUCCIÓN

En el capítulo 1 se introdujo el concepto de variable aleatoria cuantitativa asociada a una población, como una característica expresable numéricamente cuyo valor fluctúa de un individuo a otro de la misma.

Nota: A lo largo del resto de este texto se supondrá, salvo que explícitamente se indique lo contrario, que siempre que se menciona una variable aleatoria, ésta es de naturaleza cuantitativa. (De hecho, en los desarrollos clásicos del Cálculo de Probabilidades, el término de variables aleatorias se reserva sólo para las características aleatorias de naturaleza cuantitativa)

La pauta de variabilidad de una determinada variable aleatoria puede expresarse de forma sintética recurriendo a un tipo de modelo matemático: las **distribuciones de probabilidad**, cuyos conceptos básicos se exponen en este capítulo.

Cuando una variable aleatoria es **discreta**, es decir cuando su conjunto de valores posibles es discreto, su **distribución de probabilidad** viene caracterizada por la **función de probabilidad o de cuantía**, que da la probabilidad de cada uno de los valores posibles de la variable. Cuando el conjunto de valores posibles de la variable es un infinito continuo, la caracterización de dicha pauta de variabilidad se realiza mediante la **función de densidad**, que no es más que una idealización del concepto de histograma de frecuencias visto en el capítulo 2.

Los conceptos expuestos para variables aleatorias unidimensionales pueden generalizarse al caso **bidimensional**. Surgen en este contexto nuevos conceptos, como los de distribuciones **marginales** y distribuciones **condicionales.** Estrechamente relacionada con el concepto de independencia de sucesos, visto en el capítulo anterior, aparece también la idea de **independencia** de variables aleatorias.

Como el resultado de un proceso de abstracción, basado en la idea sencilla de lo que es una media aritmética, se introduce en el Cálculo de Probabilidades el concepto de **esperanza matemática** de una función de una variable aleatoria. A partir del mismo se definen los **momentos** de una distribución de probabilidad y, en particular, la **media** y la **varianza** de dicha distribución.

De acuerdo con el enfoque general del presente texto, el tratamiento de los conceptos expuestos se realiza a un nivel elemental e intuitivo, que consideramos suficiente para que un ingeniero pueda entenderlos y aplicarlos a los problemas reales con los que debe enfrentarse en su ejercicio profesional. Pese a dicho nivel elemental, hay que advertir que algunos de estos conceptos implican un cierto grado de complejidad y que algunos de los problemas tratados al final del capítulo no son triviales.

4.2 FUNCIÓN DE DISTRIBUCIÓN

En el capítulo 2 se introdujo el concepto de variable aleatoria asociada a una población, como una característica expresable numéricamente cuyo valor fluctúa según el individuo considerado de la población. La probabilidad de que dicha variable aleatoria tome valores comprendidos en un determinado intervalo puede interpretarse intuitivamente como la proporción de individuos de la población en los que el valor que toma dicha característica pertenece al intervalo considerado.

A toda variable aleatoria le corresponde, por tanto, una determinada forma de distribuirse dichas probabilidades en el conjunto de valores posibles. Se utiliza, en consecuencia, el término de **distribución de probabilidad** (o simplemente **distribución**) como sinónimo del término variable aleatoria

En el presente capítulo consideraremos en primer lugar el caso de variables aleatorias unidimensionales, en el que el conjunto de valores posibles se halla contenido en la recta real $\Re_1$.

Con el fin de caracterizar dicha distribución de probabilidad se define la **Función de Distribución** F(x) de una variable aleatoria **X** como sigue:

$$F(x) = P(\mathbf{X} \leq x)$$

Nota: debe tenerse claro que la "x" que aparece en F(x) corresponde a la que se halla a la derecha de la desigualdad en el segundo miembro. Para indicar que F es precisamente la función de distribución de la variable aleatoria **X**, puede utilizarse en los casos que puedan prestarse a confusión la terminología más precisa $F_X(x)$

De la definición dada se deduce que toda función de distribución F cumple las siguientes propiedades:

$F(-\infty) = P(\mathbf{X} \leq -\infty) = 0$

$F(\infty) = P(\mathbf{X} \leq \infty) = 1$

$0 \leq F(x) \leq 1$ por ser F(x) una probabilidad

F(x) es no decreciente.

Esta última propiedad se demuestra con facilidad si se tiene en cuenta que, siendo b > a, el suceso $\{\mathbf{X} \leq b\}$ puede descomponerse en la suma de los dos sucesos excluyentes {**X** # a} y $\{a < \mathbf{X} \leq b\}$

$$\{\mathbf{X} \leq b\} = \{\mathbf{X} \leq a\} + \{a < \mathbf{X} \leq b\}$$

de donde

$$P(\mathbf{X} \leq b) = P(\mathbf{X} \leq a) + P(a < \mathbf{X} \leq b)$$

o sea

$$F(b) = F(a) + P(a < \mathbf{X} \leq b)$$

y, como toda probabilidad es ≥ 0, se deduce que, si b > a, F(b) es $\geq$ F(a).

El resultado

$$P(a < \mathbf{X} \leq b) = F(b) - F(a)$$

indica que conociendo la función de distribución F(x) de una variable aleatoria, es posible calcular la probabilidad de que dicha variable se halle en cualquier intervalo que nos interese.

En consecuencia, dicha función de distribución caracteriza por completo la pauta de variabilidad de la variable aleatoria correspondiente.

*Autoevaluación: **el** lenguaje de programación BASIC tiene una función RND que genera un número "al azar" entre 0 y 1. ¿Qué crees que se entiende en este caso como un número "al azar" entre 0 y 1? ¿Cuál será la función de distribución de la variable aleatoria cuyos valores genera la función RND? (Ver respuesta en el Anejo al final del Tema)*

Una analogía mecánica, útil para comprender intuitivamente algunos conceptos, hace corresponder a cada variable aleatoria unidimensional una determinada forma de "distribuir" una masa total de 1 kg. en $\Re_1$. En esta analogía, el valor F(x) de la función de distribución en un punto x no es más que la masa total situada a la izquierda de dicho punto (incluyendo la del punto)

4.3 DISTRIBUCIONES DISCRETAS

Cuando el conjunto de valores posibles que puede tomar una variable aleatoria es **discreto**, es decir finito o infinito numerable, la forma más sencilla de caracterizar la distribución de probabilidad correspondiente es a partir de la **función de probabilidad**, también denominada a veces **función de cuantía** o **función de masa**, función que da la probabilidad de cada uno de los valores posibles.

En la analogía mecánica anteriormente expuesta, para una variable discreta 1kg. de masa se reparte en un conjunto discreto de puntos, y la función de probabilidad da la masa existente en cada punto (lo que justifica su denominación como función de masa)

Ejemplo:

Sea la población constituida por todos los posibles lanzamientos de dos monedas simétricas. A cada individuo de la población, es decir a cada lanzamiento de las dos monedas, se le asocia la variable aleatoria "número de caras obtenidas".

La variable aleatoria **X** considerada tiene, por tanto, sólo tres valores posibles: **X**=0, **X**=1 y **X**=2. Se trata, en consecuencia, de una variable aleatoria discreta.

La función de probabilidad P(**X**=x) de esta variable aleatoria será (por ser independientes los resultados de las dos tiradas):

$P(\mathbf{X}=0)$ = P(cruz en la 1ª)x P(cruz en la 2ª) = $\frac{1}{2} \times \frac{1}{2} = 0.25$

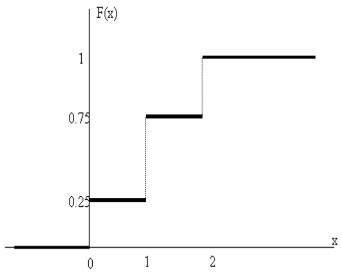
$P(\mathbf{X}=1)$ = P(cara en la 1ª)x P(cruz en la 2ª) + P(cruz en la 1ª)x P(cara en la 2ª) = $\frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = 0.50$

$P(\mathbf{X}=2)$ = P(cara en la 1ª)x P(cara en la 2ª) = $\frac{1}{2} \times \frac{1}{2} = 0.25$

P(**X**=x) = 0 para todo x diferente de 0, 1 ó 2

Vamos a calcular y dibujar la función de distribución correspondiente.

De la definición de la variable se deduce inmediatamente que:



si x < 0 $\Rightarrow$ F(x) = P(**X** $\leq$ x) = 0

si 0 $\leq$ x < 1 $\Rightarrow$ F(x) = P(**X** $\leq$ x) = 0.25

si 1 $\leq$ x < 2 $\Rightarrow$ F(x) = P(**X** $\leq$ x) = 0.75

si 2 $\leq$ x $\Rightarrow$ F(x) = P(**X** $\leq$ x) = 1

Por lo tanto, F(x) tiene una forma "en escalera", tal como se aprecia en la figura

Es importante notar que F(x) es continua por la derecha en todo punto, pero es discontinua por la izquierda en x=0, x=1 y x=2, o sea en aquellos puntos en los que la probabilidad de la variable es diferente de cero.

El anterior es un resultado general que deriva de la definición de función de distribución. Toda función de distribución es discontinua por la izquierda en aquellos puntos x en los que P(**X** = x) es no nula.

Dado que un conjunto es discreto si es finito o infinito numerable, una variable discreta puede tener un conjunto no finito de valores posibles, siempre que ese conjunto sea numerable

Autoevaluación: Considérese la variable aleatoria "número de veces que hay que lanzar un dado hasta obtener por primera vez un 5".¿Su conjunto de valores posibles es finito? ¿Es dicha variable aleatoria una variable discreta?

4.4 DISTRIBUCIONES CONTINUAS. FUNCIÓN DE DENSIDAD

A nivel intuitivo podemos decir que una variable aleatoria es continua si su conjunto de valores posibles es un infinito continuo (en la práctica, si sus valores podrían apreciarse con un gran número de decimales si se dispusiera de un aparato de medida suficientemente preciso). Sin embargo, esta definición, que puede ser útil en la práctica en la inmensa mayoría de los casos, no es completamente correcta. A diferencia de lo que sucede con las variables discretas, cuyo tratamiento puede llevarse a cabo con herramientas matemáticas muy sencillas, el estudio de las variables aleatorias continuas exige el recurso a herramientas de análisis más sofisticadas, que en este texto se expondrán sólo de una forma elemental y poca rigurosa.

De una forma más precisa, se dice que una variable aleatoria es continua si su función de distribución F(x) es continua en todo punto (o sea en toda la recta real $\Re_1$) y derivable en todo $\Re_1$ excepto a lo sumo en un conjunto finito de puntos.

Del hecho de que toda función de distribución sea discontinua en todo punto cuya probabilidad sea diferente de cero, y del hecho de que la F(x) de cualquier variable continua ha de ser continua en todo punto, se deduce que **la probabilidad de que una variable continua tome un valor exacto cualquiera es siempre cero**.

Autoevaluación: la afirmación anterior puede parecer absurda la primera vez que se estudia. ¿Considera el lector que un modelo matemático que implica que la probabilidad de cualquier valor exacto es nula, va a ser de dudosa utilidad para el estudio de las variables aleatorias reales con que va a encontrarse en la práctica? ¿Cuál cree que es en la población de todos los jóvenes españoles la probabilidad de que la variable Estatura sea exactamente igual a 170 cms (es decir 170,00000000....... cms)?

Aunque son raras, en la práctica pueden presentarse variables aleatorias que no sean ni discretas (su conjunto de valores posibles no es discreto), ni continuas (existe algún valor de probabilidad no nula).

Autoevaluación: Poner un ejemplo de una variable aleatoria de interés práctico que no sea ni discreta ni continua. (Ver respuesta en el Anejo al final del Tema).

Siendo f(x) la derivada de F(x) se verifica que

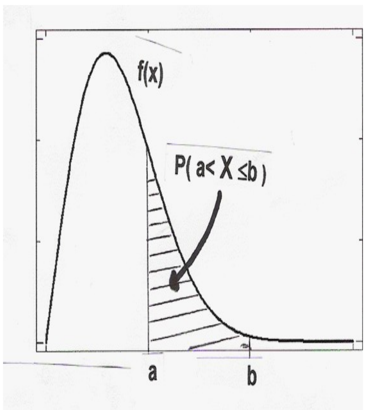
$$f(x) = \frac{d(F(x))}{dx} = \lim_{\Delta x \to o} \frac{F(x+\Delta x) - F(x)}{\Delta x} = \lim_{\Delta x \to o} \frac{P(X \le x + \Delta x) - P(X \le x)}{\Delta x} = \lim_{\Delta x \to o} \frac{P(x < X \le x + \Delta x)}{\Delta x}$$

El cociente P(x < **X** ≤ x+♠x) dividido por ♠x, da la "probabilidad por unidad de longitud", o densidad de probabilidad, en dicho intervalo. Su límite, cuando ♠x tiende a cero, puede por tanto interpretarse como la densidad de probabilidad existente en el punto x para la variable aleatoria considerada **X**.

A f(x) se le denomina **función de densidad** de la variable aleatoria, y se utiliza para caracterizar la pauta de variabilidad en el caso de variables aleatorias continuas. (Obsérvese que la función de densidad no existe en el caso de variables discretas pues, como se vio, la función de distribución no es continua, ni por tanto derivable, en aquellos puntos cuya probabilidad era no nula).

En la analogía mecánica anteriormente mencionada, a una variable continua le corresponderá una distribución de 1 kg. de masa en $\Re_1$, con una densidad que varía de un punto a otro de acuerdo con una función, que es precisamente la función de densidad que acabamos de definir. Obsérvese que la masa de un punto es siempre cero (al serlo la “longitud” del punto) y que la masa total en cualquier intervalo se obtendría, de acuerdo con resultados bien conocidos de la Mecánica, integrando la densidad en dicho intervalo.

Conocida f(x) es posible obtener la probabilidad de que **X** pertenezca a un intervalo cualquiera [a b] sin más que integrar dicha función de densidad en el intervalo.



En efecto, por ser F la primitiva de f:

$$\int_a^b f(x)dx = [F(x)]_a^b = F(b) - F(a) = P(a < X \le b)$$

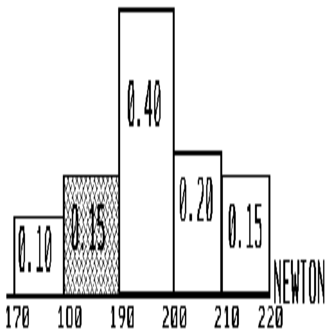
(<u>Nota</u>: en la expresión anterior el signo “≤” podría sustituirse por “<”, dado que para variables continuas la probabilidad de un valor exacto es nulo)

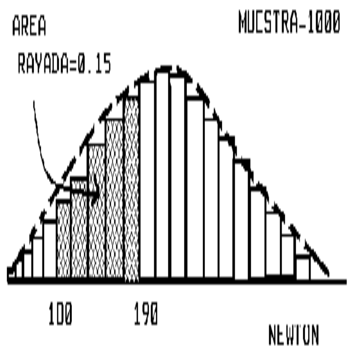
Por tanto, el área comprendida bajo la función de densidad de una variable aleatoria entre dos valores “a” y “b”, coincide con la probabilidad de que la variable aleatoria tome valores en dicho intervalo.

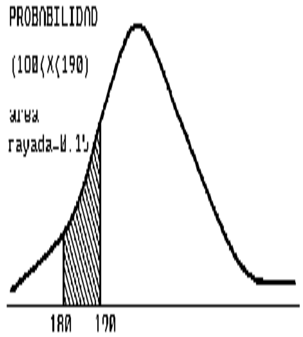
Autoevaluación: Constatar que la variable aleatoria generada por la función RND del BASIC (ver la Autoevaluación del apartado 4.2) es una variable aleatoria continua. Obtener la función de densidad de dicha variable

En las siguientes figuras se ilustra la relación entre el concepto de función de densidad y el de histograma de frecuencias visto en el capítulo 2.

Si concebimos un histograma de los valores existentes en la población en el que la barra que se traza sobre cada tramo tenga un **área** igual a la proporción de observaciones en dicho tramo, dicho histograma se irá aproximando a la función de densidad a medida que vaya aumentando el número de tramos.







Histograma para N=100 **Histograma para N=1000** **“Histograma” para N=∝**

4.5 VARIABLES BIDIMENSIONALES. INDEPENDENCIA

(<u>Nota</u>: este apartado se desarrollará sólo a un nivel muy elemental)

Si sobre cada individuo de una determinada población se observan los valores de dos características expresables numéricamente se tiene, según se expuso en el capítulo 2, una variable aleatoria bidimensional. Los valores posibles de dicha variable son, por tanto, puntos del plano $\Re_2$.

La función de distribución de una variable aleatoria bidimensional (**X,Y**) se define como F(x,y) = P({**X**≤x}{**Y**≤y})

Los dos tipos más importantes de variables bidimensionales son las discretas (ambas componentes discretas) y las continuas (ambas componentes continuas). En las primeras el conjunto de valores posibles de la variable es un conjunto discreto de puntos en $\Re_2$, cuyas probabilidades caracterizan por completo la distribución de probabilidad correspondiente.

Las variables aleatorias bidimensionales continuas vienen caracterizadas por su función de densidad $f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$ que recoge la "densidad de probabilidad" asociada a cada punto (x,y) del plano. La probabilidad de que la variable pertenezca a cualquier recinto S del plano puede obtenerse integrando, mediante una integral doble, f(x,y) en dicho recinto:

$$P((X,Y) \in S) = \iint_S f(x,y)dxdy$$

Toda variable aleatoria bidimensional **X**,**Y** tiene asociada dos distribuciones unidimensionales: la **distribución marginal** de **X** y la **distribución marginal** de **Y**, que no son más que las distribuciones (unidimensionales) que tienen cada una de las dos variables consideradas cuando se prescinde de los posibles valores de la otra variable.

También es posible definir para cada valor posible de una de las variables la **distribución condicional** de la otra. Así la distribución condicional (**Y/X**=x_0) puede definirse, de forma poco rigurosa pero intuitiva, como la distribución que tiene la variable **Y** si nos limitamos a considerar como población a aquellos individuos en los que la variable **X** vale x_0. De forma simétrica se definirían las distribuciones condicionales del tipo (**X/Y**=y_0)

Si todas las distribuciones condicionales (**Y/X**=x_0) son idénticas, es decir si no dependen del valor concreto x_0 que tome la variable condicionante, se dice que las dos variables aleatorias son **independientes**.

Autoevaluación: Razonar la lógica intuitiva de la definición que acaba de darse de independencia de dos variables aleatorias, basándose en los dos ejemplos siguientes:
1 – {Y: Coeficiente Intelectual X: Estatura} en la población de estudiantes universitarios españoles (Caso de probable independencia)
2 – {Y: Consumo de energía en calefacción una factoría X: Temperatura diaria} en la población de los días laborables de invierno (Caso de probable no independencia).

El concepto de independencia de variables aleatorias está estrechamente relacionado con el de independencia de sucesos visto en el capítulo anterior. Así, dos variables **X** e **Y** son independientes si cualquier pareja de sucesos, referido uno de ellos a la variable **X** y el otro a la variable **Y**, son independientes.

Digamos por último que los conceptos básicos expuestos en este apartado sobre variables bidimensionales, se generalizan con relativa facilidad para permitir la consideración de **variables aleatorias n-dimensionales**. En particular n variables aleatorias $\mathbf{X}_1$, $\mathbf{X}_2$,...,$\mathbf{X}_n$[1] son independientes si para todo punto $x_1,x_2,...,x_n$ los sucesos ($\mathbf{X}_1 \le x_1$),($\mathbf{X}_2 \le x_2$),...,($\mathbf{X}_n \le x_n$) son independientes.

4.6 ESPERANZA MATEMÁTICA

El concepto de media aritmética, o promedio, de un conjunto de valores observados, definido como la suma de todos ellos dividida por el número de valores, tiene una clara interpretación intuitiva. Una idealización de dicho concepto lleva a la definición de la **Esperanza Matemática**, o **Valor Medio**, de una función h(**X**) de una determinada variable aleatoria **X**.

[1] En realidad resultaría más preciso decir "las n componentes $\mathbf{X}_1,\mathbf{X}_2,...,\mathbf{X}_n$ de una variable aleatoria n-dimensional"

4.6.1 Variables discretas

En el caso de que **X** sea una variable aleatoria discreta, la esperanza matemática, o valor medio, de una función h(**X**) de dicha variable aleatoria se expresará como E(h(**X**)) y se define como:

$$E(h(X)) = \sum_i h(x_i)P(X = x_i)$$

sumatorio que se extiende para todos los valores x_i cuyas probabilidades no sean nulas.

*Autoevaluación: El origen de la denominación "Esperanza Matemática" proviene de las aplicaciones del Cálculo de Probabilidades a los juegos de azar que, como se mencionó en el apartado 1.2, fue un fenómeno que impulsó mucho el desarrollo de esta rama de las Matemáticas. Como aplicación de este concepto resolver el siguiente ejercicio elemental: Se propone a un jugador participar, pagando una cantidad, en un juego consistente en tirar un dado: si sale un número impar se gana tantos euros como el doble de puntos obtenidos, pero si sale un número par se pierde tantos euros como puntos han salido. ¿Qué serían en este ejemplo **X** y h(**X**)? ¿Cuánto es lo máximo que se puede aceptar pagar por partida para jugar a este juego?*

En el caso particular de que la función h(**X**) sea precisamente **X**, la expresión anterior da la **esperanza matemática, valor medio, o media** de una variable aleatoria **X** discreta. En este texto utilizaremos la letra **m** para referirnos a la media de una variable aleatoria en una población (recordemos que para referirnos a la media de una muestra utilizamos $\overline{x}$)

Media de **X** = E(**X**) = $\sum_i x_i P(X = x_i)$

Autoevaluación: El sumatorio que define el valor medio de una variable aleatoria discreta puede tener infinitos términos. En esos casos, el cálculo de su valor exigirá sumar una sucesión. Como ejercicio al respecto hallar el valor medio de la variable aleatoria definida en la Autoevaluación planteada al final del apartado 4.3

4.6.2 Variables continuas

Si la variable **X** es continua, en la expresión del valor medio hay que sustituir la probabilidad P(**X**=x_i) por el valor f(x) de la función de densidad en x, y el sumatorio por una integral definida

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

en la que los límites de integración se limitarán finalmente a la región en la que f(x) sea diferente de cero

*Autoevaluación: se selecciona al azar un número **X** entre 0 y 1. ¿Cuánto vale en promedio X^3?*

En el caso particular de que h(**X**) sea igual a **X**, se tiene la expresión para la media de una variable continua

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

*Autoevaluación: Una variable aleatoria **X** se dice que sigue una **distribución exponencial de parámetro** α si es siempre no negativa y verifica que $P(\mathbf{X}>x) = e^{-\alpha x}$. Hallar la función de distribución y la función de densidad de esta variable y comprobar que su valor medio es $1/\alpha$*

Una propiedad importante de la media, tanto para variables discretas como para continuas, es que es un operador lineal, o sea que la media de una combinación lineal de variables aleatorias es la combinación lineal de las medias de las mismas:

$$E(a_0+a_1\mathbf{X}_1+...+a_n\mathbf{X}_n) = a_0 + a_1E(\mathbf{X}_1) +...+ a_nE(\mathbf{X}_n)$$

4.7 MOMENTOS CENTRALES

Siendo m la media de una variable aleatoria **X**, al valor medio de la función $h(\mathbf{X}) = (\mathbf{X}-m)^\nu$ se le denomina "momento central del orden ν", simbolizándose por μ_ν.

Autoevaluación: demostrar que μ_1 es siempre igual a cero

Los momentos centrales μ_2, μ_3 y μ_4 son especialmente importantes (sobre todo μ_2) porque permiten definir la varianza y los coeficientes de asimetría y curtosis de una variable aleatoria

4.7.1 Varianza. Desviación Típica

Particularmente importante es el momento central de orden 2, al que se denomina **varianza**, a la que simbolizaremos siempre por σ^2

$$\text{Varianza} = \sigma^2 = E(\mathbf{X}-m)^2$$

A la raíz cuadrada positiva σ de la varianza se le denomina **desviación típica**.

Dos propiedades importantes de la varianza son las siguientes:

1 - $\sigma^2(a + b\mathbf{X}) = b^2\sigma^2(\mathbf{X})$

2 - Si **X** e **Y** son independientes $\sigma^2(\mathbf{X}+\mathbf{Y}) = \sigma^2(\mathbf{X}) + \sigma^2(\mathbf{Y})$

Nota Técnica: realmente para que se verifique la segunda propiedad basta con que **X** e **Y** estén incorrelacionadas. Ello es así porque la expresión general de la varianza de una combinación lineal de dos variables aleatorias es

$$\sigma^2(a\mathbf{X}+b\mathbf{Y}) = a^2\sigma^2(\mathbf{X}) + b^2\sigma^2(\mathbf{Y}) + 2ab\,cov(\mathbf{X},\mathbf{Y})$$

donde $cov(\mathbf{X},\mathbf{Y})$ es la covarianza entre **X** e **Y**, que se define com $E(\mathbf{X}-m_x)(\mathbf{Y}-m_y)$

4.7.2 Coeficientes de asimetría y de Curtosis

De acuerdo con los conceptos que ya fueron introducidos a nivel descriptivo en el capítulo 2, a partir de los momentos centrales de orden 3 y 4 se definen, como sigue, los **coeficientes de asimetría** y de **curtosis** de la distribución de una variable aleatoria **X**

Coeficiente de Asimetría: $CA = \frac{\mu_3}{\sigma^3} = \frac{E(X-m)^3}{\sigma^3}$

Coeficiente de Curtosis: $CC = \frac{\mu_4}{\sigma^4} - 3 = \frac{E(X-m)^4}{\sigma^4} - 3$

4.A AUTOEVALUACIONES RESUELTAS Y EJERCICIOS

4.A.1 Respuesta a algunas Autoevaluaciones

*Autoevaluación: **el** lenguaje de programación BASIC tiene una función RND que genera un número "al azar" entre 0 y 1. ¿Qué crees que se entiende en este caso como un número "al azar" entre 0 y 1? ¿Cuál será la función de distribución de la variable aleatoria cuyos valores genera la función RND?*

La expresión "generar al azar un número entre 0 y 1" es poco precisa (por ejemplo, el inverso del número de puntos que se obtiene al tirar un dado, sería un número entre 0 y 1 resultante del azar). Lo que hace la función RND es tomar un valor en [0 1], de forma que la probabilidad de que caiga en un intervalo cualquiera [a b] interior a [0 1] sea proporcional a la longitud (b-a) del intervalo, sin depender de la posición de éste. (Por ejemplo, es igual de probable obtener un número entre 0.2 y 0.5 que entre 0.6 y 0.9, siendo en ambos casos la mitad que la de obtener un valor entre 0.1 y 0.7).

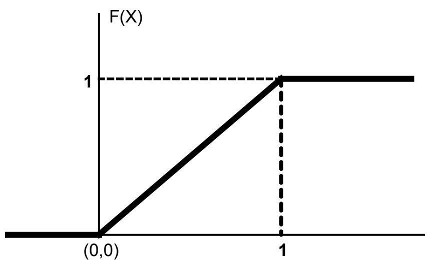
De acuerdo con esta idea, la función de distribución F(x) de esta variable aleatoria será:

Si $x < 0 \Rightarrow F(x) = P(\mathbf{X} \le x) = 0$ porque **X** no puede ser negativo

Si $x > 1 \Rightarrow F(x) = P(\mathbf{X} \le x) = 1$ porque **X** siempre será ≤ 1

Si $0 \le x \le 1 \Rightarrow F(x) = P(\mathbf{X} \le x) = x$ (por ser dicha probabilidad proporcional a la longitud del intervalo [0 x] y ser la constante de proporcionalidad = 1 porque $P(0 \le x \le 1) = 1$)

La función de distribución resultante, que se refleja en la figura adjunta, tiene forma de una línea quebrada



Puede constatarse que dicha función es continua en todo punto, y derivable excepto en los puntos x=0 y x=1, puntos cuya probabilidad es cero (como la de cualquier otro punto, por ser la probabilidad proporcional a la longitud de los intervalos y tener un punto "longitud" nula)

Autoevaluación: Poner un ejemplo de una variable aleatoria de interés práctico que no sea ni discreta ni continua.

Sea **X** la variable aleatoria: {cantidad de lluvia, en litros/m^2, caída diariamente en una ciudad}. Esta cantidad es una variable de naturaleza continua que podría medirse con tantos decimales con permitiera el instrumento de medida. La probabilidad de cualquier valor exacto (diferente de cero) puede considerarse nula (¿cuál sería la frecuencia relativa de los días en los que llueve 10 l/m^2 exactamente, o sea 10.00000..... l/m^2?) Sin embargo existe un valor, X=0, para el cual la probabilidad no es nula, porque la probabilidad de que un día no llueva nada en absoluto no es cero (por ejemplo, en esa ciudad el 70% de los días no llueve)

La variable anterior podría considerarse una "mixtura" de dos variables: Una discreta $\mathbf{X_1}$ con un único valor posible ($\mathbf{X_1}$=0) y probabilidad 0.7, y otra continua $\mathbf{X_2}$ con probabilidad total 0.3 y una distribución continua (por ejemplo, una exponencial de parámetro α=0.1). La función de distribución de **X** podría obtenerse en cada punto x como un promedio ponderado de las correspondientes a $\mathbf{X_1}$ y $\mathbf{X_2}$: $F_X(x) = 0.7 \times F_{X_1}(x) + 0.3 \times F_{X_2}(x)$

Como otro ejemplo de variable aleatoria mixta ver el segundo ejercicio resuelto en 4.A.2

4.A.2 Ejercicios resueltos

Un comerciante compra cierto tipo de equipos usados a 100 € y, tras una revisión que le cuesta 25 €, los revende a 220 €. Sin embargo, si el equipo vendido resulta defectuoso debe devolver la cantidad cobrada y pagar además una indemnización de 100 €. Por sus estadísticas el comerciante sabe que el 80% de los equipos que se le presentan para comprar son correctos y el 20% restante son defectuosos. Para comprobar si el equipo funciona correctamente, el comerciante tiene la posibilidad (tras haberlo comprado y revisado pero antes de venderlo) de realizar una prueba, que con una probabilidad P diagnostica correctamente su estado (tanto cuando el equipo es correcto como cuando es defectuoso), lo que le posibilita no venderlo si la prueba diagnostica el equipo como defectuoso
Determinar en función de P el coste máximo C que el comerciante puede pagar por la prueba, para que sea rentable realizarla

Supongamos en primer lugar que no se realiza la prueba. El beneficio medio por equipo adquirido será:

E(Beneficio) = 0.80x(220-100-25) + 0.20x(-125-100) = 31 €

Si se hace la prueba, hay 4 situaciones posibles cuyas probabilidades y resultados económicos serán:

- Equipo correcto y diagnóstico correcto: Probabilidad = 0.8xP Beneficio = 95 – C
- Equipo correcto y diagnóstico erróneo: Probabilidad = 0.8x(1-P) Beneficio = -125 – C
- Equipo defectuoso y diagnóstico correcto: Probabilidad = 0.2xP Beneficio = -125 - C
- Equipo defectuoso y diagnóstico erróneo: Probabilidad = 0.2x(1-P) Beneficio = -225 - C

El beneficio medio usando la prueba será, por tanto,

E(Benef.) = 0.8xPx(95-C) + 0.8x(1-P)x(-125-C) + 0.2xPx(-125-C) + 0.2x(1-P)x(-225-C) =196xP -145 –C

Para que la hacer la prueba sea rentable deberá ser 196xP – 145 – C > 31 $\Rightarrow$ <u>C < 196xP -176</u>

(Como era de esperar, cuanto más segura es la prueba más se puede pagar por realizarla. Por una prueba absolutamente segura (P=1) se podría pagar hasta 20 €, mientras que una prueba con P < 176/196 = 0.898 nunca sería rentable realizarla)

*Un comerciante compra diariamente 800 kgs de cierto producto perecedero a 1 €/kg para revenderlo a 1.5 €/kg. La demanda diaria **X** se distribuye exponencialmente con media 1000 kgs, y las cantidades que se quedan si vender cada día se pierden. ¿Qué cantidad C debe comprar diariamente?*

Asumiremos, en primer lugar, que debe buscarse C de forma que maximice el valor medio del beneficio (<u>Nota</u>: este supuesto sería discutible, puesto que, por ejemplo, no tiene en cuenta el coste e medio plazo de no satisfacer ciertos días la demanda de los clientes, que podrían decidir dejar de comprar en ese comercio)

La variable beneficio diario es una variable "mixta", pues según que la demanda **X** resulte mayor o menor que C se tendrá:

Si **X** > C, el comerciante venderá los C kgs y su beneficio será B = Cx0.50. La probabilidad de que pase esto será $P(\mathbf{X}>C) = e^{-0.001C}$ (pues el parámetro α de la distribución exponencial de la demanda es 1/media = 1/1000 = 0.001)

Si **X** $\leq$ C, se venderá la cantidad **X**, con un beneficio 0.50x**X**, quedándose sin vender los C-**X** kgs restantes, que ocasionarán un coste de 1x(C-**X**). La densidad de probabilidad asociada a cada valor posible x será la de la función de densidad de la exponencial que es $0.001e^{-0.001x}$

El beneficio medio será, por tanto (aplicando para la parte discreta la expresión para la media de una variable discreta y para la continua la correspondiente a variables continuas):

$$E(\text{Beneficio}) = 0.5Ce^{-0.001C} + \int_0^C \left(0.5x - 1(C-x)\right)0.001e^{-0.001x}dx$$

Para hallar el valor de C que maximiza E(Beneficio) se iguala a cero la derivada respecto a C de la expresión anterior (recordar las reglas para derivar una integral respecto a un parámetro que aparece en el integrando y en el límite superior de integración)

Llamando, para aligerar la escritura, U a $e^{-0.001C}$

$$\frac{dE(\text{Beneficio})}{dC} = 0.5U - 0.5xCx0.001xU + 0.5xCx0.001xU - \int_0^C 0.001e^{-0.001x}dx = 0.5U - 1 + U$$

e igualando a cero: U = 1/1.5 o sea $e^{-0.001C} = 0.6667 \Rightarrow C = -\frac{\log(0.6667)}{0.001} = 405.4$ kgs

Un dispositivo está formado por dos componentes independientes conectadas ***en serie****. La duración de cada componente se distribuye de forma exponencial con media igual a 100 horas. Hallar la vida media del dispositivo.*

Sea X_1 la duración de la 1ª componente, X_2 la de la 2ª e Y la duración del dispositivo.

Por el enunciado del problema, al ser X_1 y X_2 exponenciales de media 100, se tiene $P(X_1>x) = P(X_2>x) = e^{-0.01x}$

La probabilidad de que Y sea mayor que un valor dado "y" será, por estar las componentes en serie y ser sus vidas independientes,

$P(Y>y) = P(X_1>y)P(X_2>y) = e^{-0.01y}e^{-0.01y} = e^{-0.02y}$

lo que demuestra que Y sigue también una distribución exponencial de parámetro α = 0.02. El valor medio de Y será, por tanto

E(Y) = 1/0.02 = 50 horas

(Observar que, como era de esperar, la vida media del dispositivo es menor que la de las componentes al estar éstas conectadas en serie)

Un dispositivo está formado por dos componentes A y B conectadas ***en paralelo****. Las vidas de ambas componentes siguen distribuciones exponenciales independientes de medias 100 horas para A y 50 horas para B. Sea* ***Y*** *la variable aleatoria: vida del dispositivo*
*Hallara la vida media E(****Y****) del dispositivo*

Sea X_1 la duración de la 1ª componente, X_2 la de la 2ª e Y la duración del dispositivo.

Por el enunciado del problema, al ser X_1 y X_2 exponenciales de media 100 y 50, se tiene $P(X_1>x) = e^{-0.01x}$ y $P(X_2>x) = e^{-0.02x}$

La probabilidad de que Y sea mayor que un valor dado "y" será, por estar las componentes en paralelo y ser sus vidas independientes,

$P(Y>y) = P(\{X_1>y\} + \{X_2>y\}) = P(X_1>y) + P(X_2>y) - P(X_1>y)P(X_2>y) = e^{-0.01y} + e^{-0.02y} - e^{-0.03y}$

que en este caso no corresponde a la expresión de una distribución exponencial.

Para hallar E(Y) utilizaremos directamente la expresión $E(Y) = \int_{-\infty}^{+\infty} y f(y)dy$ lo que exige obtener previamente la función de densidad de Y

$f(y) = \dfrac{dF(y)}{dy}$ $\qquad F(y) = P(Y\leq y) = 1 - P(Y>y) = 1 - (e^{-0.01y} + e^{-0.02y} - e^{-0.03y})$ $\qquad$ (para y>0)

$f(y) = 0.01e^{-0.01y} + 0.02e^{-0.02y} - 0.03e^{-0.03y} \Rightarrow E(Y) = \int_0^{\infty} y\,(0.01e^{-0.01y} + 0.02e^{-0.02y} - 0.03e^{-0.03y})dy$

$= \dfrac{1}{0.01} + \dfrac{1}{0.02} - \dfrac{1}{0.03} = 116.67$ horas

(Observar que, como era de esperar, la vida media del dispositivo es mayor que la de las componentes, al estar éstas conectadas en paralelo)

4.A.3 Ejercicios adicionales

La vida ***X*** *de las pilas eléctricas de una determinada marca se distribuye exponencialmente con media 50 horas.*
a) Si una pila lleva ya funcionando sin fallo 50 horas, ¿cuál es la probabilidad de que falle a lo largo de las 10 horas siguientes?
b) Y si una pila lleva ya funcionando sin fallo 100 horas, ¿cuál es la probabilidad de que falle a lo largo de las 10 horas siguientes?
Nota: los resultados anteriores justifican que a la distribución exponencial se la denomine distribución "sin memoria"

Un dispositivo está formado por ***n*** *componentes independientes conectadas en paralelo. La vida de cada una de dichas componentes se distribuye exponencialmente con media 100 horas. Calcular el valor mínimo que debe tener* ***n*** *si se desea que la fiabilidad del dispositivo a las 200 horas sea superior al 95%*

Sean $X_1, \ldots, X_n$ n números seleccionados independientemente al azar entre 0 y 1
a) Sea ***Y*** *= máximo($X_1, \ldots, X_n$). ¿Cuánto valdrá en promedio* ***Y****?*
b) Sea ***Z*** *= mínimo($X_1, \ldots, X_n$). ¿Cuánto valdrá en promedio* ***Z****?*

Sea ***U*** *una variable aleatoria distribuida uniformemente en el intervalo [0 1] (Esta variable ya se ha manejado a lo largo del capítulo como la correspondiente a un número "al azar" entre 0 y 1). Sea* ***X*** *una variable aleatoria continua cualquiera de función de distribución $F_X(x)$, y sea F_X^{-1}*

*la función inversa de F_X. Demostrar que la variable aleatoria **Y**=F_X^{-1} (**U**) se distribuye como la variable **X**. (Nota;: la propiedad anterior es la base para generar en el ordenador variables aleatorias que sigan una distribución cualquiera, a partir de la función RND comentada en el capítulo)*