2

VARIABLES ALEATORIAS

2.1 INTRODUCCIÓN

La identificación de cada resultado, en algunos experimentos aleatorios, obedece a un reconocimiento de las propiedades que lo caracterizan. Por ejemplo, la condición de ser hembra en un recién nacido es un resultado cuya calificación depende de una serie de características cualitativas específicas. Igualmente la raza o la salud de un individuo dependen de la identificación de ciertas cualidades que ubican cada una de esas condiciones en una determinada categoría. En otros tipos de experimentos aleatorios no basta con calificar los resultados, sino que es necesario caracterizarlos cuantitativamente. En algunos casos esta cuantificación resulta de un proceso de conteo, así se habla del número de hijos, de dientes, de cromosomas, de espiráculos, de electrones, de emisiones radiactivas, etc. En otros casos, al determinar características como el peso, la talla, la temperatura o la concentración de alguna sustancia en ciertos objetos o elementos, se asigna a cada resultado un valor dentro de una escala de medición específica. Cada una de esas características cuantificables por conteo o por medición recibe el nombre genérico de variables aleatorias; son variables porque su valor cambia de un elemento a otro; y son aleatorias porque su comportamiento es impredecible. Las variables aleatorias son importantes porque ellas caracterizan los fenómenos o procesos naturales, por lo que resulta muy valioso comprender en la forma más completa posible sus propiedades y comportamiento. Una primera aproximación a éste conocimiento se logra estableciendo el conjunto de posibles valores que puede asumir la variable y su respectiva probabilidad de ocurrencia. El tratamiento de estos dos aspectos es el objetivo de éste capítulo.

2.2 DEFINICIÓN DE VARIABLE ALEATORIA

El espacio muestral de muchos experimentos aleatorios se pueden representar con símbolos. Por ejemplo al lanzar tres monedas, los ocho resultados posibles se pueden describir de la manera siguiente:

$$S = \{ccc, ccs, csc, scc, css, scs, ssc, sss\}$$

Este mismo espacio muestral se puede expresar en números. Para esto es necesario definir una regla o norma que al aplicarla le adjudique a cada resultado un valor. Por ejemplo se puede establecer la regla siguiente: *contar el número de sellos que aparecen en cada resultado del espacio muestral*. Así al evento *(ccc)* le corresponde el número cero porque no hay sellos en el mismo; a los tres eventos *(ccs, csc, scc)* les corresponde el número 1 porque en cada uno de ellos aparece sólo un sello. En la misma forma a los eventos *(css, scs, ssc)* les corresponde el 2 y al evento *(sss)* le corresponde el 3. En el esquema *a* de la Figura 2.1. se muestra la correspondencia entre el espacio muestral simbólico y el espacio muestral numérico. ¿Cuáles serán las reglas definidas para los esquemas *b* y *c* de la misma figura?

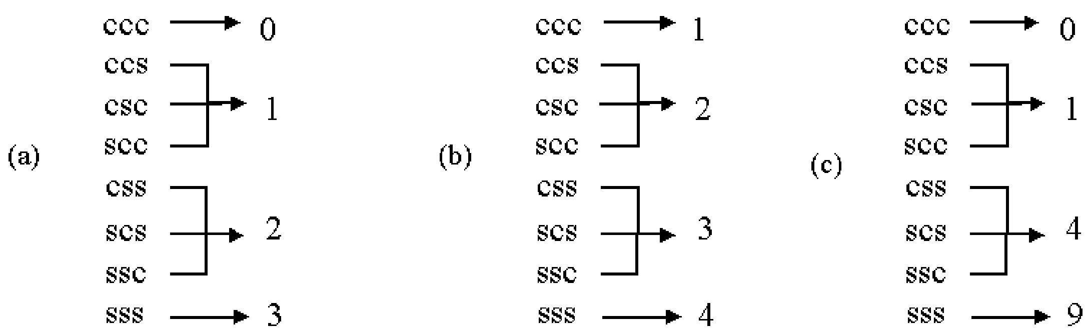


Figura 2.1: Correspondencia entre los resultados de un mismo espacio muestral simbólico con los valores de tres espacios muestrales numéricos diferentes.

Los tres espacios numéricos mostrados en la Figura 2.1 pueden ser expresados formalmente de la manera siguiente:

$S_1 = \{0, 1, 2, 3\}$ $\qquad$ $S_2 = \{1, 2, 3, 4\}$ $\qquad$ $S_3 = \{0, 1, 4, 9\}$

Igualmente las tres reglas definidas para la construcción de estos espacios numéricos, pueden establecerse con mayor rigor, asociándolas a una función matemática. Si la cantidad número de sellos se identifica con la letra x, las funciones que resultan son las siguientes:

$$1)\ f_{1(x)} = x \qquad 2)\ f_{2(x)} = x+1 \qquad 3)\ f_{3(x)} = x^2$$

A la primera regla "*número de sellos*" le corresponde la función $f_{1(x)} = x$

A la segunda regla "*número de sellos más uno*" le corresponde la función $f_{2(x)} = x+1\ +1$

A la tercera regla *"cuadrado del número de sellos"* le corresponde la función $f_{3(x)} = x^2$

Si cada una de estas reglas se define en forma genérica como una variable aleatoria, y a su vez sabemos que cada regla es una función matemática, la definición formal de variable aleatoria se puede enunciar de la manera siguiente:

"*Sea E un experimento aleatorio y S su espacio muestral, toda función que asigne a cada uno de los elementos $s \in S$, un número real X(s), se llama variable aleatoria.*" (Figura 2.2)

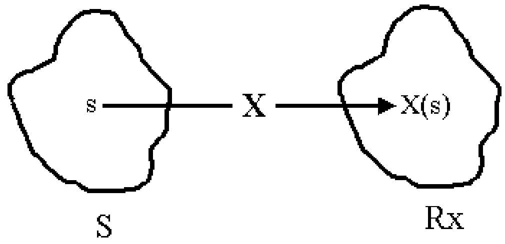


Figura 2.2: Variable aleatoria

Las variables aleatorias se identifican con letras mayúsculas. Aunque desde el punto de vista formal matemático definir una función como una variable no parece lo más adecuado, es la definición que se ha utilizado tradicionalmente. En nuestro ejemplo las tres variables aleatorias se pueden identificar de la manera siguiente:

X = número de sellos *Y = número de sellos + 1* *Z = cuadrado del número de sellos*

El resultado de definir una variable aleatoria es que se genera un nuevo espacio muestral numérico que se denomina recorrido o rango espacial y se identifica con la letra *R* (Figura 2.2). Los tres rangos espaciales originados son entonces:

$$R_x = \{0,1,2,3\} \qquad R_y = \{1,2,3,4\} \qquad R_z = \{0,1,4,9\}$$

Es importante puntualizar algunas cosas con relación al concepto de variable aleatoria:

1) Para un mismo experimento es posible definir diferentes variables aleatorias. En nuestro ejemplo se pudieron especificar otras variables aleatorias como el número de lanzamientos o la distancia entre las monedas.

2) En muchos casos el resultado de un experimento es directamente un número. Por ejemplo al determinar el sexo de un individuo se obtiene una categoría que luego puede transformarse en un número (0 = hembra y 1 = macho), pero si se mide la talla de un individuo se obtiene directamente un valor (Ej. 1,70 m)

3) En términos prácticos en el estudio de una variable aleatoria es más importante conocer los valores que ella asume que saber cuales son los elementos que conforman su espacio muestral. En otras palabras las diferentes formas de los resultados y la función como tal tienen poco interés en la praxis.

4) Los valores que asumen las variables aleatorias se identifican con letras minúsculas, por ejemplo x, y, z. Si se define como variable aleatoria *X = tamaño de una persona*, y se quiere indicar la probabilidad que esa persona supere determinada altura, este evento se puede expresar así $P(X > x)$, donde x asume el valor que se especifique.

2.3 VARIABLES ALEATORIAS DISCRETAS Y CONTINUAS

De acuerdo a las características del rango espacial, las variables aleatorias se clasifican en discretas y continuas.

2.3.1 Variables aleatorias discretas: Una variable aleatoria se denomina discreta si el rango espacial esta constituido por un número finito o infinito contable de valores:

$$R_x = \{x_1, x_2, x_3, ..., x_r, ...x_n, ...\}$$

Estas variables se generan fundamentalmente de experimentos cuyos resultados son evaluados mediante la enumeración o contaje de elementos o cosas: número de hijos, número de estomas, número de partículas, número de átomos, etc.

Ejemplo 2.1.
Se registra el número de varones nacidos en los primeros cuatro partos ocurridos el primer día del año.

Si se considera que en cada parto la probabilidad de nacer varón o nacer hembra es $1/2$, el espacio muestral *S* estará formado por 16 resultados igualmente posibles. Al definir como variable aleatoria el *número de varones* se origina un espacio *Rx* formado por cinco resultados que son valores numerables. Estos valores resultaron del contaje de una característica del experimento (Figura 2.3)

$$S = \begin{pmatrix} \text{HHHH} \\ \text{VHHH, HVHH, HHVH, HHHV} \\ \text{VVHH, VHVH, VHHV, HVHV, HHVV, HVVH} \\ \text{VVVH, VVHV, VHVV, HVVV} \\ \text{VVVV} \end{pmatrix} \xrightarrow{X = \text{Numero de varones}} Rx = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$

Figura 2.3: Rango espacial (Rx) de la variable aleatoria número de varones

2.3.2 Variables aleatorias continuas. Una variable aleatoria es continua si el rango espacial está constituido por un número infinito de valores en un intervalo dado:

$$R_x = \{X_{(s)} = x \,/\, x_1 \le X \le x_2\}$$

Este tipo de variables son generadas por la medición de magnitudes como la longitud, el peso, el volumen, la densidad, la concentración, la temperatura, etc.

Ejemplo 2.2.
Se capturó una trucha en un río y se le determinó la talla.

En éste experimento el espacio numérico Rx se origina inmediatamente como resultado de determinar la longitud del cuerpo del pez, que es una característica propia de cada individuo. El rango espacial Rx está formado por infinitos resultados dentro de un determinado intervalo (Figura 2.4)

$$S = \{\text{Talla de la truchas}\}$$

$$\Downarrow$$

$$Rx = \{x_i = \text{talla} \;/\; 10 \text{ cm} \leq x_i \leq 15 \text{ cm}\}$$

Figura 2.4: Espacio muestral (S) y rango espacial (Rx) de la variable aleatoria longitud del cuerpo de una especie de trucha

2.4 FUNCIÓN DE PROBABILIDAD DE UNA VARIABLE ALEATORIA

Al introducir éste capítulo se dijo que para tener un buen conocimiento de una variable aleatoria no basta con saber cuales son los valores que puede asumir sino también es necesario describir su comportamiento en términos de probabilidades. Para ello se requiere generar una nueva función, conocida como función de probabilidad, de densidad o de cuantía, con la cual es posible asignar un valor de probabilidad a cada resultado de un rango espacial (Figura 2.3).

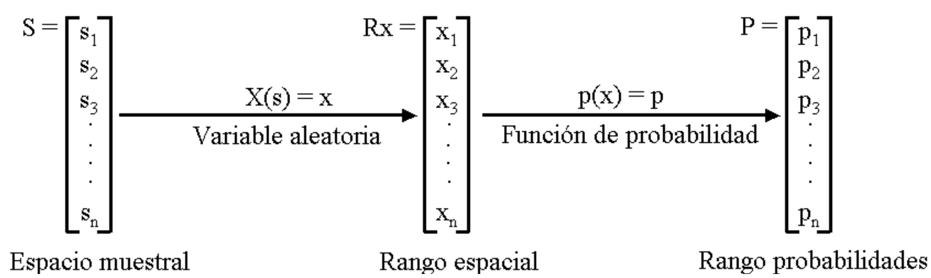


Figura 2.5: Relación entre la variable aleatoria y la función de probabilidad. *X(s)* = variable aleatoria; x_i = valor de la variable aleatoria; *p(x)* = función de probabilidad; p_i = valor de probabilidad de x_i.

2.4.1. Función de probabilidad de una variable aleatoria discreta.

En el caso de variables discretas la función de probabilidad se denota como $p_{(x)}$ y se interpreta como la probabilidad de que la variable aleatoria tome el valor x_i.

$$p_{(x)} = P(X = x_i)$$

Obviamente como la función de probabilidad genera un valor de probabilidad, estos números deben satisfacer las condiciones siguientes:

$$1)\ 0 \leq p_{(x)} \leq 1 \qquad 2)\ \sum_{R_x} p_{(x)} = 1 \qquad 3)\ P(x_1 \leq X \leq x_2) = \sum_{x_1}^{x_2} p_{(x)}$$

Las dos primeras condiciones son equivalentes con los axiomas probabilísticos de positividad y certidumbre. La tercera propiedad simplemente establece que si se conoce la función de probabilidad de una variable aleatoria discreta, entonces se puede calcular la probabilidad correspondiente a cualquier intervalo abierto o cerrado entre dos puntos x_1 y x_2.

En la Figura 2.6 se ilustra, con el ejemplo de lanzar una moneda hasta obtener cara por primera vez, como una variable discreta es descrita en términos de sus valores y las probabilidades con las cuales ellos ocurren.

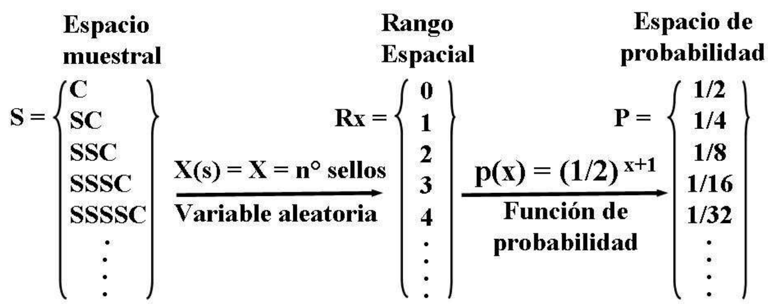


Figura 2.6: Descripción numérica y probabilística de los resultados que se pueden obtener al lanzar una moneda hasta que aparece cara por primera vez.

El conjunto de pares ordenados: $[x_i , p_{(xi)}]$ para una variable discreta se denomina distribución de probabilidad. En el caso de la Figura 2.6, la distribución de probabilidad está representada por el conjunto de pares de valores siguientes:

x_i	$p_{(xi)}$
0	0,5000
1	0,2500
2	0,1250
3	0,0625
4	0,0313
.	.
.	.

En la Figura 2.7 se representa gráficamente ésta distribución. Una vez conocidos los valores de la variable y sus respectivas probabilidades se tiene un conocimiento bastante completo del comportamiento de la variable. En este momento es fácil responder a interrogantes relativas a la variable aleatoria. Por ejemplo en el caso de la Figura 2.7 ¿Cuál es la probabilidad de obtener menos de tres sellos?. La respuesta se tiene sumando las probabilidades que hay en el espacio de probabilidad.

$$P(X < 3) = P(X \le 2) = p_{(0)} + p_{(1)} + p_{(2)} = 0,50 + 0,25 + 0,125 = 0,875$$

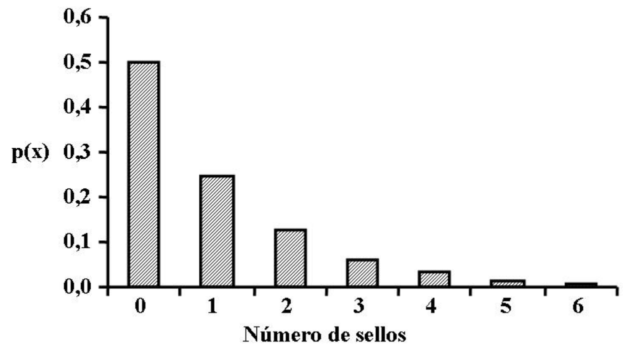


Figura 2.7: Distribución probabilística de la variable aleatoria número de sellos

Veamos otra aplicación que se puede obtener a partir de la distribución de probabilidad de una variable aleatoria.

Ejemplo 2.3.
En un estudio de campo se determinó para cierta región, el número de crías por madriguera para una determinada especie de roedor y la probabilidad con la cual esto ocurre (Tabla 2.1). Suponiendo que ésta situación se mantiene estable, y algún tiempo después se revisan N madrigueras de la especie en cuestión y en la misma región, es posible estimar en forma aproximada el número de individuos por madriguera que se espera encontrar. Simplemente se obtiene el producto p(x)N. Si el número de madrigueras revisadas es N = 300, las frecuencias esperadas para cada categoría son las que se muestran en la Tabla 2.1.

Tabla 2.1. Distribución de probabilidades y frecuencia esperadas del número de crías por madriguera para cierta especie de roedor.

Número de crías	Probabilidad $p_{(x)}$	Frecuencia esperada $[\, p_{(x)}N \,]$
1	0,250	75
2	0,400	120
3	0,200	60
4	0,080	24
5	0,05	15
6	0,02	6

* Sí el valor de f no es entero, como se trata de una variable discreta, en la práctica las frecuencias deben aproximarse al valor entero más próximo. En estos casos es posible que la suma total de las frecuencias difiera levemente de N.

2.4.1.1 Parámetros de la distribución de una variable aleatoria discreta

La mayoría de las veces resulta impráctico manejar toda la distribución de probabilidades para determinar el comportamiento de una variable, por lo que es conveniente conocer algunos parámetros que caracterizan la variable aleatoria. Esta idea se aprecia claramente cuando se tiene una función determinística como es la ecuación de una recta de regresión, $f_{(x)} = \alpha + \beta x$ caracterizada por los parámetros α y β. Conocidos estos dos parámetros es fácil definir la recta. La función $f_{(x)} = e^{-rx}$ tiene un único parámetro, la tasa intrínseca de crecimiento r de una población biológica. Dos de los parámetros más importante para caracterizar las variables aleatorias son el valor promedio y la varianza, los que proporcionan una rápida visión de la naturaleza de la variable.

2.4.1.1.1 Valor promedio

Si en el ejemplo 2.3 se quisiera conocer el número promedio de crías por madriguera, el número de crías se multiplica por la frecuencia esperada y su total se divide entre el total de madrigueras (Tabla 2.2)

Tabla 2.2. Distribución de frecuencias y cómputos necesarios para calcular el valor promedio.

Número de crías (x)	Frecuencia (f)	fx	fx / $\mathbf{N}$
1	75	600	2,00
2	120	1200	4,00
3	60	720	2,40
4	24	336	1,12
5	15	240	0,80
6	6	108	0,36
Total	300	3204	10,68

$$\overline{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i} = \frac{3204}{300} = 10{,}68$$

El promedio del número de crías en cada madriguera es de 10,68 individuos. Un resultado similar se encuentra si se divide cada producto fx entre N y luego se suman todos los resultados. En tal caso la suma proporcionará directamente el valor medio de la talla. (Ver Tabla 2.2). Si en la fórmula de cálculo de $\overline{x}$ se sustituye $\sum_{i=1}^{n} f_i$ por N y se aplica el concepto de frecuencia relativa a la probabilidad que establece que $fr_{(x)} = p_{(x)}$, se obtiene una nueva fórmula de cálculo para la media a partir de los valores de probabilidad:

$$\overline{x} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i} = \frac{\sum_{i=1}^{n} f_i x_i}{N} = \sum_{i=1}^{n} \frac{f_i x_i}{N} = \sum_{i=1}^{n} \frac{f_i}{N}(x) = \sum_{i=1}^{n} f_{r_i}(x) = \sum_{i=1}^{n} p(x_i)x$$

La conclusión es que el valor promedio de la distribución de una variable discreta es igual a la suma del producto de cada valor de la variable por su probabilidad de ocurrencia. Si este concepto lo extrapolamos de la muestra a la población el valor promedio de la distribución de valores de una variable discreta es:

$$\mu = \sum_{i=1}^{n} p(x_i)x$$

Este valor promedio de la distribución de una variable aleatoria también se conoce como Esperanza matemática o Valor esperado y se denota como E(x), así se tiene que:

$$E(x) = \mu = \sum_{i=1}^{n} p(x_i)x$$

En la Tabla 2.3 se muestra como se organizan los valores de una variable aleatoria y sus probabilidades para obtener el valor esperado y la varianza.

Tabla 2.3. Organización de valores y operaciones necesarias para obtener el valor esperado y la varianza de una variable aleatoria

(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
x_i	$p_{(x_i)}$	$x_i\ p_{(x_i)}$	$x_i - \mu$	$(x_i - \mu)^2$	$p(x_i)(x_i - \mu)^2$
x_1	$p_{(x_1)}$	$x_1\ p_{(x_1)}$	$x_1 - \mu$	$(x_1 - \mu)^2$	$p(x_1)(x_1 - \mu)^2$
x_2	$p_{(x_2)}$	$x_2\ p_{(x_2)}$	$x_2 - \mu$	$(x_2 - \mu)^2$	$p(x_2)(x_2 - \mu)^2$
x_3	$p_{(x_3)}$	$x_3\ p_{(x_3)}$	$x_3 - \mu$	$(x_3 - \mu)^2$	$p(x_3)(x_3 - \mu)^2$
.	.	.	.	.	.
.	.	.	.	.	.
x_n	$p_{(x_n)}$	$x_n\ p_{(x_n)}$	$x_n - \mu$	$(x_n - \mu)^2$	$p(x_n)(x_n - \mu)^2$
		$\sum p(x_i)x_i$			$\sum p(x_i)(x_i - \mu)^2$

2.4.1.1.2 Varianza

Si por ejemplo, de una población humana se extrae un niño y se le determina el número de caries, cabrían las preguntas siguientes: ¿El número de caries será igual al valor promedio de la población? ¿El valor estará cercano o alejado al valor promedio?. Si sólo conocemos el valor promedio no podremos responder ninguna de estas preguntas. Sólo sabemos que el niño tendrá un número de caries mayor o menor al promedio y que sus probabilidades de ocurrencia dependen de la forma de la distribución de la variable. De modo que no basta conocer el valor medio de una variable aleatoria para poder describir desde un punto de vista práctico alguna de sus características más interesantes. Las preguntas hechas anteriormente hacen pensar que se requiere otro tipo de medida que cuantifique la dispersión de valores alrededor del valor medio. Lo más simple sería determinar la desviación de cada valor respecto al valor medio, es decir sería necesario obtener para cada x_i la desviación $x_i - \mu$ (Columna 4, Tabla 2.3). Como se quiere tener un valor de desviación para toda la distribución, la suma de las mismas podría representar una medida general de desviación. Sin embargo como el término $\sum(x_i - \mu) = 0$, la mejor manera de evadir este problema es elevando al cuadrado cada desviación de forma que $(x_i - \mu)^2$ (Columna 5, Tabla 2.3), de modo que el valor promedio de todas las diferencias $(x_i - \mu)^2$ se podría usar como esa medida única de dispersión. Puesto que $(x_i - \mu)^2$ es también una variable aleatoria, su valor promedio será:

$$\sigma^2 = \sum_{i=1}^{n} p(x_i)(x_i - \mu)^2 \quad \text{(Columna 6, Tabla 2.3)}$$

Esta medida de dispersión se denomina varianza y se identifica por el símbolo σ^2 que es el cuadrado de la letra griega σ. Una fórmula más simple para el cálculo de σ^2 se obtiene al desarrollar el producto notable $(x_i - \mu)^2$ presente en la fórmula anterior. As se tiene que:

$$\sigma^2 = \left[\sum_{i=1}^{n} x_i^2 \, p(x_i) \right] - \mu^2$$

Ejemplo 2.4.
En un estudio odontológico de una población se determinó que el número de caries por niños tiene la distribución de probabilidades siguiente:

N° caries	0	1	2	3	4	5	6	7
p(x)	0,19	0,29	0,21	0,15	0,09	0,04	0,02	0,01

Se quiere conocer: a) la probabilidad de que un niño tenga más de 2 y menos de 6 caries; b) el número promedio de caries por niño, y c) la varianza de la distribución.

a) Para responder la primera pregunta se requiere encontrar la $P(2 < X < 6)$ para lo cual se procede de la manera siguiente:

$$P(2 < X < 6) = P(3 \leq X \leq 5) = p_{(3)} + p_{(4)} + p_{(5)} = 0{,}15 + 0{,}09 + 0{,}04 = 0{,}28$$

Para obtener los valores del promedio y de la varianza del número de caries, los datos se organizan como se muestra en la Tabla 2.3.

Tabla 2.3. Cómputos necesarios para calcular los parámetros de una distribución de probabilidades de una variable discreta.

Distribución de probabilidades		Cálculo de μ usando $\mu=\sum_{i=1}^{n} p(x_i)x_i$	Calculo de σ² usando $\sigma^2=\sum_{i=1}^{n} p(x_i)(x_i-\mu)^2$			Cálculo de σ² usando $\sigma^2=\left[\sum_{i=1}^{n} p(x_i)x_i^2\right]-\mu^2$
x_i	$p(x_i)$	$p(x_i)x_i$	$(x_i-\mu)$	$(x_i-\mu)^2$	$p(x_i)(x_i-\mu)^2$	$p(x_i)x_i^2$
0	0,19	0,00	-1,91	3,65	0,69	0,00
1	0,29	0,29	-0,91	0,83	0,24	0,29
2	0,21	0,42	0,09	0,01	0,00	0,84
3	0,15	0,45	1,09	1,19	0,18	1,35
4	0,09	0,36	2,09	4,37	0,39	1,44
5	0,04	0,20	3,09	9,55	0,38	1,00
6	0,02	0,12	4,09	16,73	0,33	0,72
7	0,01	0,07	5,09	25,91	0,26	0,49
Total		1,91			2.48	6,13
		$\mu = 1{,}91$			$\sigma^2=2{,}48$	$\sigma^2=\left[\sum_{i=1}^{n} p(x_i)x_i^2\right]-\mu^2=2{,}48$

2.4.2. Función de probabilidad de una variable aleatoria continua.
En el caso de variables aleatorias continuas la función de probabilidad se identifica como $f_{(x)}$. Para las variables continuas no tiene sentido encontrar la probabilidad exacta de un valor puntual puesto que su rango espacial esta formado infinitos valores, de modo que la expresión $P(X \le x_i)$ carece de sentido porque siempre será igual a cero. Esto puede ilustrarse si se considera el experimento de medir la temperatura del agua en la superficie de un lago (Figura 2.8).

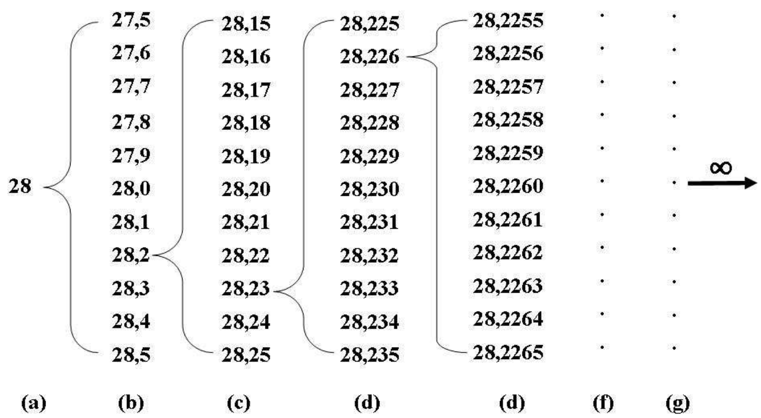


Figura 2.8: Sucesivos incrementos de apreciación en la lectura de la temperatura, la cual teóricamente puede extenderse hasta infinito.

Un termómetro con una apreciación en grados puede determinar que la temperatura del agua es 28°C (Figura 2.8a). Sin embargo debido a la apreciación tan gruesa, cualquier valor entre 27,5°C y 28,5°C el instrumento lo aprecia como 28°C. Si se cambia el termómetro por otro con una apreciación de 0,1°C, el valor de temperatura tendrá una décima más de apreciación, digamos que fue de 28,2°C (Figura 2.8b). Pero la incertidumbre se mantiene porque cualquier valor entre 28,15° y 28,25 es medido como 28,2°C. Esta falta de seguridad sobre cual es el verdadero valor de temperatura del agua siempre estará presente, en primer lugar porque teóricamente la apreciación del termómetro puede incrementarse indefinidamente (Figuras 2.8c,d,e,f, g) y en segundo término porque el rango espacial de la temperatura, igual que el de todas las variables continuas, esta formado por infinitos valores.
Al no poderse definirse para una variable aleatoria continua una función $p_{(x)}$ que asigne una probabilidad a cada valor x_i de su rango espacial, es necesario establecer una nueva función $f_{(x)}$ que fije la probabilidad de todos los valores x_i. Esta función debe satisfacer las condiciones siguientes:

1) $0 \le f_{(x)} \le 1$ 2) $\int_{x \le x_i} f_{(x)} dx = 1$ 3) $P(a \le X \le b) = \int_a^b f_{(x)} dx$

Para el caso de las variables continuas no es posible establecer la distribución de probabilidades como un conjunto de pares de valores porque como ya vimos, no es posible asignar un valor de probabilidad a un resultado puntual. Sin embargo, se puede considerar que la propia función *f(x)* representa la distribución de probabilidad. Cualquier función *f(x)* puede representarse gráficamente como una curva haciendo el área bajo la curva equivalente a su probabilidad de ocurrencia. Bajo esta perspectiva la curva de una determinada función debe cumplir con las dos propiedades siguientes: i) toda el área bajo la curva es igual a la unidad y ii) la probabilidad de ocurrencia de la variable entre dos valores dados es igual al área bajo la curva entre esos dos valores (Figura 2.9)

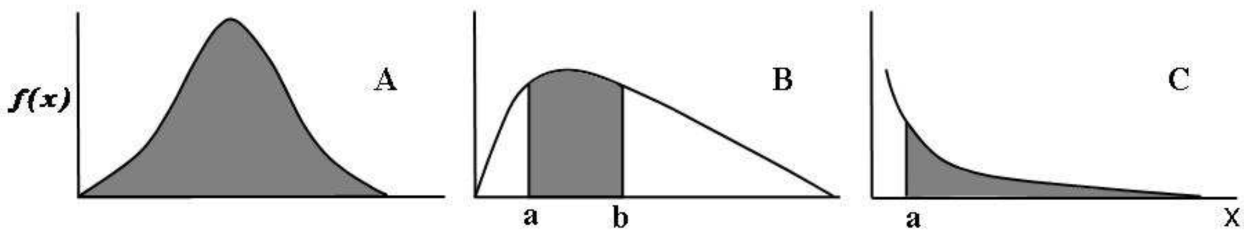


Figura 2.9. Distribuciones de probabilidad de tres variables aleatorias continuas. A) El área sombreada bajo la curva de f(x) es igual a la unidad. B) El área bajo la curva entre los valores a y b representa la probabilidad de que la variable se encuentre entre esos dos valores. C) El área bajo la curva representa la probabilidad de que la variable sea igual o mayor al valor a.

Una consecuencia de la naturaleza de las variables continuas, es que las probabilidades siguientes todas son iguales

$$P(a < X < b); P(a < X \leq b); P(a \leq X < b); P(a \leq X \leq b)$$

Ejemplo 2.5.
Encuentre la probabilidad de que una variable aleatoria sea mayor a 2 y menor a 4 si se sabe que su función de probabilidad es

$$f(x) = \begin{cases} xe^{-x};\ para\ x > 0 \\ \\ 0\ para\ cualquier\ otro\ valor\ de\ x \end{cases}$$

Para poder tener idea del comportamiento de f(x), se puede graficar la curva de la distribución de valores de X. Para ello, se deben encontrar algunos valores de f(x) para diferentes valores de X, que permitan hacer un bosquejo de la curva.

x	0	1	2	3	4	5	6	7	8
f(x)	0	0,37	0,27	0,15	0,07	0,03	0,01	0,01	0,00

En la Figura 2.10 se muestra la forma de la curva.

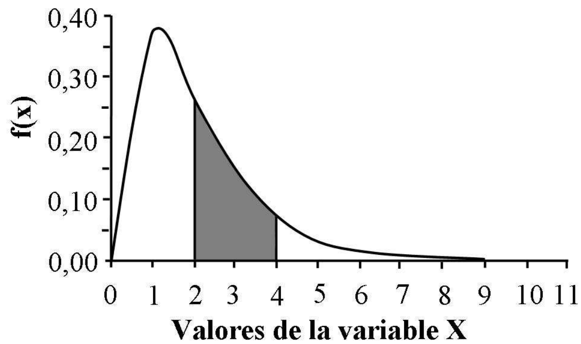


Figura 2.10: Distribución de probabilidades de la función $f_{(x)} = xe^{-x}$. El área sombreada representa la probabilidad que tiene la función de asumir valores en el intervalo (2, 4).

Para encontrar la probabilidad solicitada es necesario hallar el área bajo la curva localizada entre los valores 2 y 4. Par ello se procede a integrar por partes la función, f(x)= xe^{-x}, entre esos dos valores.

$$P(2 \le X \le 4) = \int_2^4 xe^{-x} = \left[-xe^{-x} - e^{-x}\right]_2^4 = \left[-e^{-x}(x+1)\right]_2^4 = -5e^{-4} + 3e^{-3} = 0,3144$$

2.4.2.1 Parámetros de la distribución de una variable continua.

El significado de la media y la varianza de la distribución de una variable aleatoria continua sigue siendo el mismo que tienen para las variables aleatorias discretas, sólo que en lugar de sumar un número definido de valores enteros, es necesario sumar infinitos valores, de modo que sus fórmulas de cálculo son las siguientes:

$$\mu = \int_{-\infty}^{+\infty} x\, f(x)dx \qquad\qquad \sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx = E(X^2) - \mu^2$$

Ejemplo 2.6.

Encontrar la media y la varianza de la distribución de la variable aleatoria del Ejemplo 2.5 cuya función es $f_{(x)} = xe^{-x}$.

$$\mu = E(X) = \int_0^{+\infty} x\, f_{(x)}dx = \int_0^{+\infty} x(xe^{-x})dx = \int_0^{+\infty} x^2e^{-x}dx = 2$$

$$\sigma^2 = E(X^2) - \mu^2 = \int_0^{+\infty} x\,(x^2e^{-x})dx - \mu^2 = \int_0^{+\infty} x^3e^{-x}dx - \mu^2 = 6 - 2^2 = 6 - 4 = 2$$

2.5. FUNCIÓN DE DISTRIBUCIÓN ACUMULADA

Posiblemente la función de distribución de probabilidades acumuladas sea una de las funciones con más aplicación en la práctica estadística porque la mayoría de las tablas usadas en ésta disciplina se generan a partir de funciones acumuladas.

2.5.1 Función acumulada para variables discretas.
Al rango espacial de cualquier experimento se puede asociar otra función que cuantifica la probabilidad de que la variable aleatoria (**X**) asuma un valor igual o menor a x_i. Esta función se simboliza como F(x) y se denomina función de distribución acumulativa (Figura 2.11).

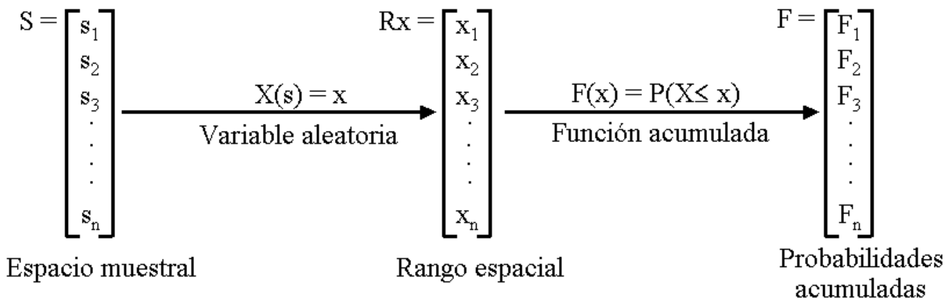


Figura 2.11: Función acumulada

Para el caso de variables aleatorias discretas la función acumulativa queda definida de la manera siguiente:

$$F(x) = P(X \leq x_i) = \sum_{x \leq x_i} p(x)$$

Ejemplo 2.7.
De un lote de 100 ratones integrado por 50 hembras y 50 machos, un biólogo extrae aleatoriamente un animal del lote y registra su sexo. Cuando el ratón es macho lo regresa al lote y continúa con la extracción de animales hasta conseguir la primera hembra. Construya el espacio muestral (S), el rango espacial (Rx), el rango de probabilidad (P) y el espacio de probabilidades acumuladas (F) de la variable aleatoria X = número de ratones machos. Luego, usando la función acumulada calcule la probabilidad de que antes de aparecer la primera hembra se hayan seleccionado a) menos de tres machos y b) al menos dos machos.

En la Figura 2.12 se muestran los espacios solicitados. Para calcular la $P(X = x_i)$ se recurre a la regla de probabilidad de eventos independientes, $P(A \cap B) = P(A)P(B)$, de modo que las probabilidades de cada resultado son las siguientes:

P(h)= P(X = 0)= 1/21 = 0,500
P(mh)= P(X = 1)= 1/2 x 1/2 = 1/22 = 0,250
P(mmh)= P(X = 2)= 1/2 x 1/2 x 1/2 = 1/23 = 0,125
P(mmmh)= P(X = 3)= 1/2 x 1/2 x 1/2 x 1/2 = 1/24 = 0,0625
P(mmmmh)= P(X = 4)= 1/2 x 1/2 x 1/2 x 1/2 x 1/2 = 1/25 = 0,03125
. . . .
. . . .

Si se pone un poco más de atención a los cálculos anteriores se puede ver que la fórmula general para calcular la probabilidad de cada resultado es $p(x) = 1/2^{(x+1)}$. Esta fórmula es la manera formal de expresar la función de probabilidad para éste u otros experimentos similares.

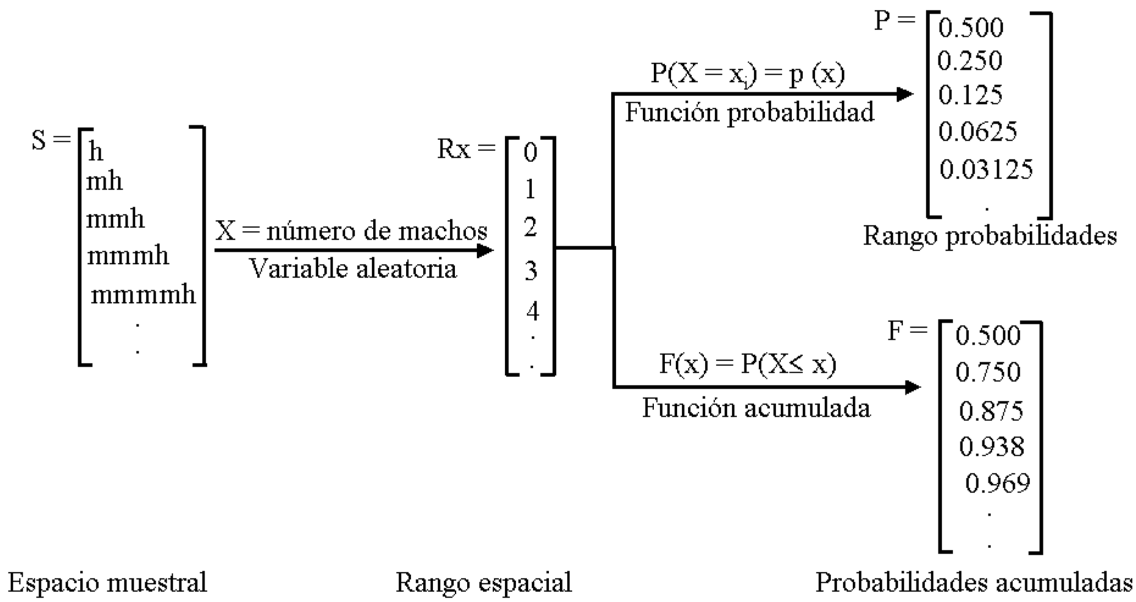


Figura 2.12: Espacio Muestral (S), Rango Espacial (Rx), Rango de Probabilidad (P) y el Espacio de Probabilidades Acumuladas (F) de la variable aleatoria número de ratones machos.

Los valores de F(x) se obtuvieron de la manera siguiente:

$F_{(0)} = P(X \leq 0) = p_{(0)} = 0{,}50$

$F_{(1)} = P(X \leq 1) = p_{(0)} + p_{(1)} = 0{,}50 + 0{,}25 = 0{,}75$

$F_{(2)} = P(X \leq 2) = p_{(0)} + p_{(1)} + p_{(2)} = 0{,}50 + 0{,}25 + 0{,}125 = 0{,}875$

$F_{(3)} = P(X \leq 3) = p_{(0)} + p_{(1)} + p_{(2)} + p_{(3)} = 0{,}50 + 0{,}25 + 0{,}125 + 0{,}0625 = 0{,}938$

$F_{(4)} = P(X \leq 4) = p_{(0)} + p_{(1)} + p_{(2)} + p_{(3)} + p_{(4)} = 0{,}50 + 0{,}25 + 0{,}125 + 0{,}0625 + 0{,}03125 = 0{,}969$

Las probabilidades solicitadas se calculan de la manera siguiente:

a) Probabilidad de extraer menos de tres machos:

$$P(X < 3) = P(X \leq 2) = F_{(2)} = 0{,}875$$

b) Probabilidad de extraer al menos dos machos:

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - F_{(1)} = 1 - 0{,}75 = 0{,}25$$

Ejemplo 2.8.
Sea la variable aleatoria **X** = la suma de la cara de dos dados, determine la distribución de probabilidades acumuladas y calcule las probabilidades siguientes:

1) $P(X \leq 6)$
2) $P(3 \leq X \leq 8)$
3) $P(X > 3)$
4) $P(2 < X < 8)$
5) $P(2 \leq X \leq 8 \text{ y } 5 \leq X \leq 10)$
6) $P(X > 8 \text{ ó } X<4)$
7) $P[(5 < X < 10) \text{ ó } X>7]$
8) $P[(4 \leq X \leq 7) / X \leq 6]$

a) El espacio muestral está formado por 36 posibles resultados equiprobables

$$S = \begin{Bmatrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{Bmatrix}$$

b) El rango espacial de la variable aleatoria es el siguiente:

$$Rx = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

c) Las distribuciones de probabilidad y acumulativa serán:

x_i	$p(x_i)$	$F(x_i)$
2	0,02778	0,02778
3	0,05556	0,08333
4	0,08333	0,16667
5	0,11111	0,27778
6	0,13889	0,41667
7	0,16667	0,58333
8	0,13889	0,72222
9	0,11111	0,83333
10	0,08333	0,91667
11	0,05556	0,97222
12	0,02778	1,00000

Las probabilidades solicitadas serán:

$$1)\ P(X \leq 6) = F_{(6)} = 0{,}41667$$

$$2)\ P(3 \leq X \leq 8) = P(X \leq 8) - P(X \leq 2) = F_{(8)} - F_{(2)} = 0{,}72222 - 0{,}02778 = 0{,}69$$

$$3)\ P(X > 3) = 1 - P(X \leq 3) = 1 - F_{(3)} = 1 - 0{,}08333 = 0{,}92$$

$$4)\ P(2 < X < 8) = P(3 \leq X \leq 7) = P(X \leq 7) - P(X \leq 2) = F_{(7)} - F_{(2)} = 0{,}58333 - 0{,}02778 = 0{,}56$$

5) $P(3 \le X \le 8 \text{ y } 5 \le X \le 10) = P(5 \le X \le 8 = P(X \le 8) - P(X \le 4) = F_{(8)} - F_{(4)} = 0{,}7222 - 0{,}16667 = 0{,}56$

6) $P(X > 8 \text{ ó } X<4) = P(X \ge 9) + P(X \le 3) = 1 - P(X \le 8) + P(X \le 3) = 1 - F_{(8)} + F_{(3)} = 1 - 0{,}7222 + 0{,}08333 = 0{,}36$

7) $P[(5 < X < 10) \text{ ó } X>7] = P(6 \le X \le 9) + P(X \ge 8) - P(8 \le X \le 9) =$

$= P(X \le 9) - P(X \le 5) + 1 - P(X \le 7) - P(X \le 9) + P(X \le 7) = 1 - P(X \le 5) = 1 - F_{(5)} = 1 - 0{,}28 = 0{,}72$

8) $P\left[(4 \le X \le 7) / X \le 6\right] = \dfrac{P(4 \le X \le 6)}{P(X \le 6)} = \dfrac{P(X \le 6) - P(X \le 3)}{P(X \le 6)} = \dfrac{F_{(6)} - F_{(3)}}{F_{(6)}} = \dfrac{0{,}41667 - 0{,}08333}{0{,}41667} = 0{,}80$

2.5.2 Función acumulada para variables continuas.

Cuando se trata de variables continuas la función acumulativa se define como:

$$\Phi(x_i) = P(X \le x_i) = \int_{x \le x_i} f(x)dx$$

En el caso de variables discretas la $P(X \le x_i)$ se obtiene sumando los valores de probabilidad de todos los resultados iguales o menores a x_i. Para las variables continuas ésta probabilidad se obtiene calculando el área que se encuentra por debajo de $f_{(x)}$ y a la izquierda del valor x_i (él inclusive) (Figura 2.13)

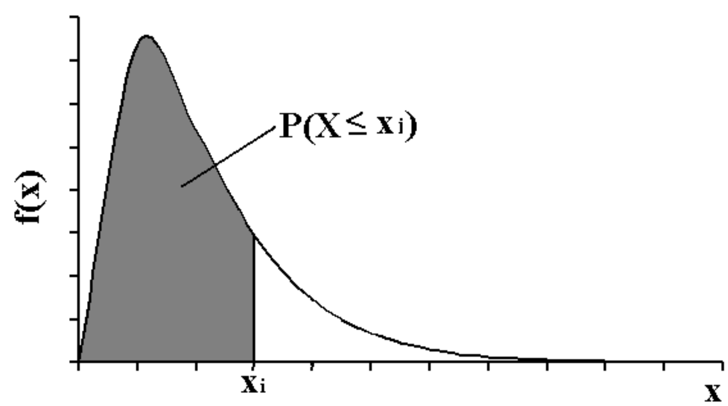


Figura 2.13: El área sombreada representa la probabilidad que la variable aleatoria continua asuma un valor en el intervalo (0, x)

Para determinar dicha área se debe integrar la función $f_{(x)}$ para todo el intervalo cuyo límite superior es x_i. Dado que estos cómputos pueden llegar a ser bastantes complejos dependiendo de la naturaleza de $f_{(x)}$, se han desarrollado para las funciones de probabilidad más usadas, tablas con las probabilidades acumuladas. Estas facilitan el cálculo de probabilidades. En la Tabla 2.4 se muestran las probabilidades acumuladas para la función del Ejemplo 2.5.

Tabla 2.4. Probabilidades acumuladas, $\Phi(x) = \int_0^{\infty} xe^{-x}$

x	$\phi_{(x)}$	x	$\phi_{(x)}$	x	$\phi_{(x)}$	x	$\phi_{(x)}$	x	$\phi_{(x)}$
0,0	0,0000	2,0	0,5940	4,0	0,9084	6,0	0,9826	8,0	0,9970
0,1	0,0047	2,1	0,6204	4,1	0,9155	6,1	0,9841	8,1	0,9972
0,2	0,0175	2,2	0,6454	4,2	0,9220	6,2	0,9854	8,2	0,9975
0,3	0,0369	2,3	0,6691	4,3	0,9281	6,3	0,9866	8,3	0,9977
0,4	0,0616	2,4	0,6916	4,4	0,9337	6,4	0,9877	8,4	0,9979
0,5	0,0902	2,5	0,7127	4,5	0,9389	6,5	0,9887	8,5	0,9981
0,6	0,1219	2,6	0,7326	4,6	0,9437	6,6	0,9897	8,6	0,9982
0,7	0,1558	2,7	0,7513	4,7	0,9482	6,7	0,9905	8,7	0,9984
0,8	0,1912	2,8	0,7689	4,8	0,9523	6,8	0,9913	8,8	0,9985
0,9	0,2275	2,9	0,7854	4,9	0,9561	6,9	0,9920	8,9	0,9986
1,0	0,2642	3,0	0,8009	5,0	0,9596	7,0	0,9927	9,0	0,9988
1,1	0,3010	3,1	0,8153	5,1	0,9628	7,1	0,9933	9,1	0,9989
1,2	0,3374	3,2	0,8288	5,2	0,9658	7,2	0,9939	9,2	0,9990
1,3	0,3732	3,3	0,8414	5,3	0,9686	7,3	0,9944	9,3	0,9991
1,4	0,4082	3,4	0,8532	5,4	0,9711	7,4	0,9949	9,4	0,9991
1,5	0,4422	3,5	0,8641	5,5	0,9734	7,5	0,9953	9,5	0,9992
1,6	0,4751	3,6	0,8743	5,6	0,9756	7,6	0,9957	9,6	0,9993
1,7	0,5068	3,7	0,8838	5,7	0,9776	7,7	0,9961	9,7	0,9993
1,8	0,5372	3,8	0,8926	5,8	0,9794	7,8	0,9964	9,8	0,9994
1,9	0,5663	3,9	0,9008	5,9	0,9811	7,9	0,9967	9,9	0,9995

Ejemplo 2.9.
Usando la Tabla 2.4, determinar la probabilidad que la variable aleatoria con función de probabilidad $f_{(x)} = xe^{-x}$ sea mayor a 2 y menor a 4 (Figura 2.10 del Ejemplo 2.5).

La probabilidad buscada es:

$$P(2 \leq X \leq 4) = P(X \leq 4) - P(X \leq 2) = \phi_{(4)} - \phi_{(2)} = 0{,}9084 - 0{,}5940 = 0{,}3144$$

Los valores de probabilidades acumuladas $\phi_{(4)}$ y $\phi_{(2)}$ se buscaron en la Tabla 2.4.4 y se sustituyeron en la expresión anterior.

Ejemplo 2.10.
Supongamos que la variable X, contenido de plomo en la sangre de personas tiene la función de probabilidades,

$$f(x) = \begin{cases} \left(\dfrac{1}{\sigma\sqrt{2\pi}}\right) e^{-1/2\left[\frac{x-\mu}{\sigma}\right]^2} \\ 0 \text{ para } x < 0 \end{cases}$$

Usando las probabilidades acumuladas (Tabla 2.5), calcule la probabilidad de que un individuo seleccionado aleatoriamente a) tenga una concentración superior a 0,40 ppm y b) tenga una concentración menor a 0,30 ppm si se sabe que forma parte de un grupo de personas cuya concentración de plomo en la sangre se encuentra entre 0,25 y 0,45 ppm.

Tabla 2.5. Probabilidades acumuladas $\Phi(x) = \int_0^{¥} \left(1/\sigma\sqrt{2\pi}\right) e^{-1/2[x-\mu/\sigma]^2}$

x	$\phi_{(x)}$	x	$\phi_{(x)}$	x	$\phi_{(x)}$	x	$\phi_{(x)}$
0,00	0,011521	0,20	0,324718	0,40	0,913659	0,60	0,999268
0,01	0,014561	0,21	0,358065	0,41	0,927102	0,61	0,999467
0,02	0,018268	0,22	0,392531	0,42	0,938882	0,62	0,999615
0,03	0,022750	0,23	0,427863	0,43	0,949118	0,63	0,999724
0,04	0,028125	0,24	0,463782	0,44	0,957941	0,64	0,999804
0,05	0,034518	0,25	0,500000	0,45	0,965482	0,65	0,999862
0,06	0,042059	0,26	0,536218	0,46	0,971875	0,66	0,999903
0,07	0,050882	0,27	0,572137	0,47	0,977250	0,67	0,999933
0,08	0,061118	0,28	0,607469	0,48	0,981732	0,68	0,999954
0,09	0,072898	0,29	0,641935	0,49	0,985439	0,69	0,999968
0,10	0,086341	0,30	0,675282	0,50	0,988479	0,70	0,999979
0,11	0,101557	0,31	0,707280	0,51	0,990952	0,71	0,999986
0,12	0,118639	0,32	0,737730	0,52	0,992947	0,72	0,999990
0,13	0,137656	0,33	0,766471	0,53	0,994543	0,73	0,999994
0,14	0,158655	0,34	0,793373	0,54	0,995810	0,74	0,999996
0,15	0,181651	0,35	0,818349	0,55	0,996807	0,75	0,999997
0,16	0,206627	0,36	0,841345	0,56	0,997585	0,76	0,999998
0,17	0,233529	0,37	0,862344	0,57	0,998188	0,77	0,999999
0,18	0,262270	0,38	0,881361	0,58	0,998650	0,78	0,999999
0,19	0,292720	0,39	0,898443	0,59	0,999002	0,79	1,000000

a) La probabilidad $P(X \geq 0{,}40)$ se obtiene calculando el área debajo de $f_{(x)}$ por encima de 0,40 como se observa en la Figura 2.14.

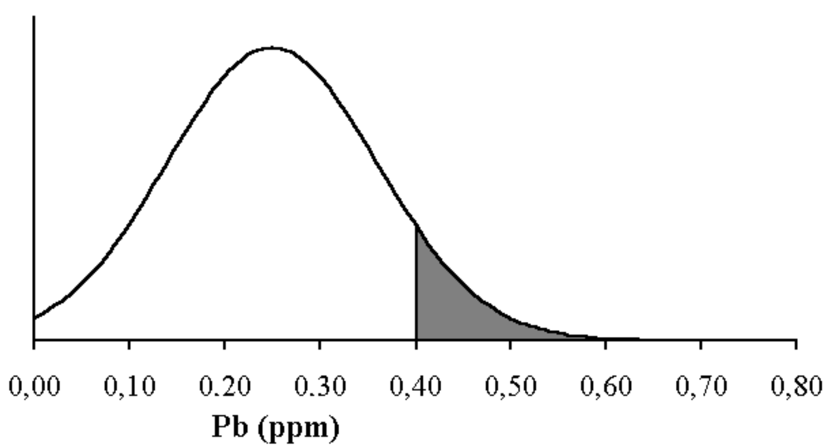


Figura 2.14: El área sombreada representa la probabilidad de que una persona tenga una concentración de plomo igual o superior 0,40 ppm.

Usando la Tabla 2.5 se puede calcular fácilmente esa área, que es equivalente a la probabilidad solicitada.

$$P(X \geq 40) = 1 - P(X \leq 0{,}40) = 1 - \phi_{(0{,}40)} = 1 - 0{,}913659 = 0{,}08634$$

En términos prácticos se puede decir que aproximadamente el 8,6% de los individuos de esa población tienen más de 0,40 ppm de plomo en la sangre.

d) La segunda probabilidad solicitada es condicionada (Figura 2.15).

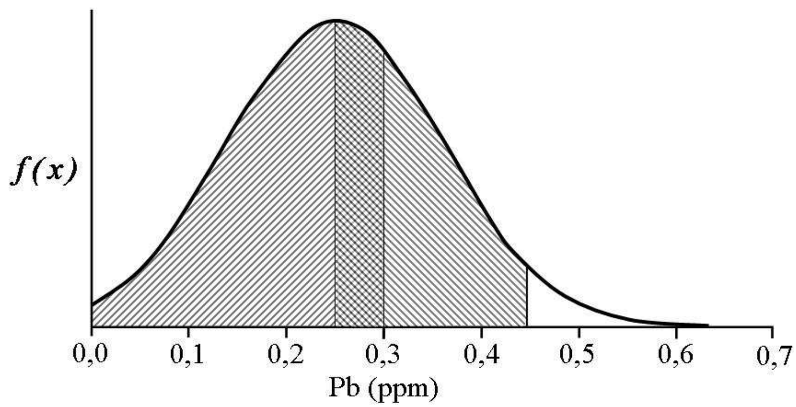


Figura 2.15: El área con el rayado descendente representa la probabilidad que tiene la variable aleatoria de asumir un valor entre 0,25 y 0,45. El área con el rayado ascendente representa la probabilidad que el valor de la variable aleatoria sea menor a 0,30.

Interesa obtener dos áreas, la que se encuentra entre 0,25 y 0,30, que representa la intersección de los dos eventos y el área entre 0,25 y 0,45 que es el nuevo espacio muestral reducido. En términos de probabilidades se tiene:

$$P\left(X \leq 30 \big/ 0{,}25 \leq X \leq 0{,}45\right) = \frac{P\left[\left(X \leq 30\right) \cap \left(0{,}25 \leq X \leq 0{,}45\right)\right]}{P\left(0{,}25 \leq X \leq 0{,}45\right)} = \frac{P\left(0{,}25 \leq X \leq 0{,}30\right)}{P\left(0{,}25 \leq X \leq 0{,}45\right)} =$$

$$= \frac{P(X \leq 0{,}30) - P(X \leq 0{,}25)}{P(X \leq 0{,}45) - P(X \leq 0{,}25)} = \frac{\Phi(0{,}30) - \Phi(0{,}25)}{\Phi(0{,}45) - \Phi(0{,}25)} = \frac{0{,}685272 - 0{,}500}{0{,}965482 - 0{,}500} = \frac{0{,}1853}{0{,}4655} = 0{,}3980$$

2.6 EJERCICIOS

1. En los siguientes enunciados indique con una C, dentro del paréntesis, si se trata de una variable aleatoria continua y con un D si se trata de una variable aleatoria discreta:
   a. La temperatura del agua en un río ( ).
   b. El tiempo de vida de un hombre. ( )
   c. La toneladas de maíz producidos en cierta región ( ).
   d. Las calorías consumidas diariamente por una gaviota ( ).
   e. El total de presas consumidas en una semana por un depredador ( ).

f. La cantidad de niños vacunados contra la hepatitis B en un hospital ( ).
g. Las pulsaciones de un ratón ( ).
h. La tensión arterial de un conejo ( ).
i. La profundidad de una laguna ( ).
j. La concentración de amilasa en la saliva ( ).
k.

2. Para cada uno de los experimentos que se enumeran a continuación identifique la variable aleatoria que los caracteriza, señale el tipo de variable y construya el respectivo rango espacial.
   a. Cinco niños son vacunados contra la viruela y al cabo de 24 horas se determina individualmente si les dio fiebre.
   b. Se cruzaron dos cobayos: un macho de pelo blanco y una hembra de pelo negro, obteniéndose una primera camada de 12 crías. Se contó el número de crías discriminándolas por el color del pelo.
   c. El ritmo cardíaco en personas adultas normales puede variar entre 60 y 100 pulsaciones por minuto. Con el propósito de conocer el efecto de un fármaco sobre el ritmo cardíaco se cuenta el número de pulsaciones en un minuto en 25 adultos normales antes de aplicársele la droga.
   d. En un lugar ubicado a nivel del mar, se colocaron 10 cubos de hielo en un recipiente e inmediatamente se comenzó a calentar, acción que continuó en forma constante hasta que el agua obtenida de los cubos de hielo se evaporó totalmente. En cualquier momento de éste proceso se midió una sola vez la temperatura.
   e. En el mismo experimento anterior se determinó el tiempo que transcurrió desde el inicio del experimento hasta que el agua se evaporó totalmente.
   f. Una manera de evaluar el contenido estomacal de los peces es mediante el método de los puntos. Este consiste en ponderar determinadas proporciones de llenura estomacal con un número. En la tabla siguiente se muestra un ejemplo:

Proporción de llenado	Ponderación
Vacío	0
¼ de estomago lleno	1
½ estomago lleno	2
¾ de estomago lleno	3
Completo	4

   Suponga que en un determinado estudio se colectaron 25 peces y a cada uno se le determinó la llenura del estomago con el método de los puntos usando los valores de ponderación de la tabla anterior. Los registros individuales se sumaron y se obtuvo el total de puntos para todos los peces.
   g. Se sabe que la distancia máxima que los ejemplares de una especie de roedor se alejan de su madriguera es equivalente a un radio de 30 m. En una determinada fecha se capturó un ejemplar de esta especie y se midió la distancia entre la madriguera y el punto de captura.
   h. En un estudio sobre los animales dispersores de semillas, se colocaron 10 g de un tipo de semilla en un recipiente. A dicho envase solo tienen acceso las hormigas. El

recipiente fue expuesto a estos insectos por 24 horas. Transcurrido este lapso se obtuvo el peso total de las semillas remanentes.

3. Se capturaron tres zorros en una selva nublada donde habitan dos especies de éste mamífero. En la medida que fueron capturados se le determinó la especie y luego se liberaron. Si se sabe que de cada tres zorros capturados dos perteneces a la especie Z1 y uno pertenece a la Z2, determine la distribución de probabilidades para el número de zorros de cada especie y represente dicha distribución gráficamente.

4. Los aminoácidos Valina, Glicina y Alanina se unieron mediante una reacción enzimática. Si la molécula resultante (tripéptido) se le determina la secuencia de unión de los aminoácidos, determine el espacio muestral de este experimento y encuentre los valores de la variable aleatoria X = número de la posición ocupada por cada aminoácido en el tripéptido formado.

5. Sea X la variable aleatoria que determina el número de recién nacidos con el factor Rh+ en la sangre. Si se sabe que la aparición del factor Rh+ ocurre con una probabilidad de 8/10 y la del factor Rh- con una probabilidad de 2/10 y se determina el tipo de factor sanguíneo en los primeros cuatro recién nacidos de un día cualquiera, haga lo siguiente:
   a. Establezca la distribución de probabilidades de la variable aleatoria.
   b. Calcule el valor de μ.
   c. Calcule el valor de σ

6. La función de probabilidad para una variable aleatoria discreta *X* viene dada por:

$$p_{(x)} = \begin{cases} \dfrac{x}{6} & cuando\ x = 1, 2, 3 \\ 0\ para\ cualquier\ otro\ valor\ de\ x \end{cases}$$

   a. Encuentre la distribución de probabilidades.
   b. Calcule $E(x)$ y σ^2
   c. Calcule y grafique los valores de la función acumulada *F(x)*.

7. Se sabe que la variable aleatoria X tiene la función de probabilidad siguiente:

$$f_{(x)} = \begin{cases} 2x\ cuando\ 0 \le x \le 1 \\ 0\ para\ cualquier\ otro\ valor\ de\ x \end{cases}$$

   Sabiendo que el área de un triángulo rectángulo se calcula como la mitad del producto de la base por la altura ($A = \frac{1}{2}bh$), calcule las probabilidades siguientes:
   a. $P(1/2 \le x \le 3/4)$
   b. $P(-1/2 \le x \le 1/2)$

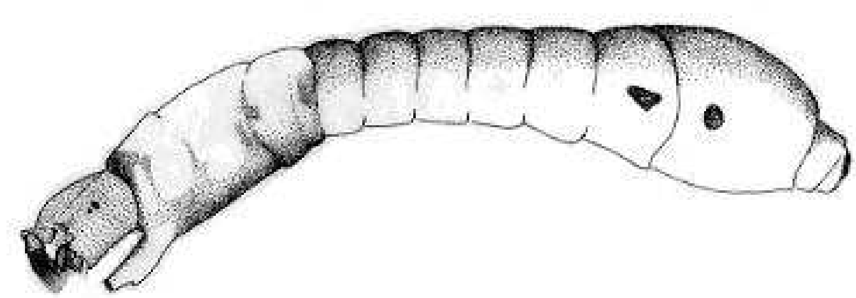
8. Determine el valor de la constante *k* para que la función $f_{(x)} = k(x^2 + 4)$ pueda se usada para describir la distribución de probabilidades del número de presas consumidas por un depredador en un día cualquiera, si se sabe que el máximo de presas consumidas es igual o menor a 3.

9. La función de densidad de una variable aleatoria continua *X* viene dada por:

$$f_{(x)} = \begin{cases} kx^2(x-1) \text{ cuando } 0 \le x \le 1 \\ 0 \text{ para cualquier otro valor de } x \end{cases}$$

   a. Calcular el valor *k* para que $f_{(x)}$ sea función de densidad.
   b. Determinar la distribución de probabilidades.
   c. Calcular la media y la varianza de X.

10. La longitud del pico de la especie de ave *Piquitus chiquitus* varía entre 1,5 y 2,4 cm. La función que describe la distribución de probabilidades de la longitud del pico para esta especie es la siguiente: $f_{(x)} = k(x-1)(3-x)$, se pide lo siguiente:
    a. Calcular el valor de k para que $f_{(x)}$ sea una función de probabilidad o densidad.
    b. Si se sabe que los adultos de la especie tienen picos con longitudes mayores a 1,8 cm, calcule la probabilidad de que un individuo capturado con un muestreo aleatorio sea adulto.
    c. Calcule el valor promedio y desviación del tamaño de los picos de esta especie de ave.



Larva de la Familia Simulidae (Insecta: Diptera)