$$u = \frac{z - a_2}{\sigma_o} = \frac{\frac{a_1 + a_2}{2} - a_2}{\sigma_o} = \frac{a_1 - a_2}{2\sigma_o}$$

La ecuación 2.28 se puede escribir como

$$P_B = \int_{u=(a_1-a_2)/2\sigma_o}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{u^2}{2}\right] du = Q\left(\frac{a_1 - a_2}{2\sigma_o}\right) \qquad (2.29)$$

En donde $Q(x)$, es conocida como la **función de error complementario** o **función co-error**, y esta graficada en la figura 2.8.

Otra forma de función de co-error que es utilizada frecuentemente es

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} \exp\left[-u^2\right] du \qquad (2.30)$$

Las funciones de co-error, Q(x) y erfc(x) están relacionadas como sigue

$$\text{erfc}(x) = 2Q\left(x\sqrt{2}\right) \qquad (2.31)$$

$$Q(x) = \frac{1}{2}\text{erfc}\left(\frac{x}{\sqrt{2}}\right) \qquad (2.32)$$

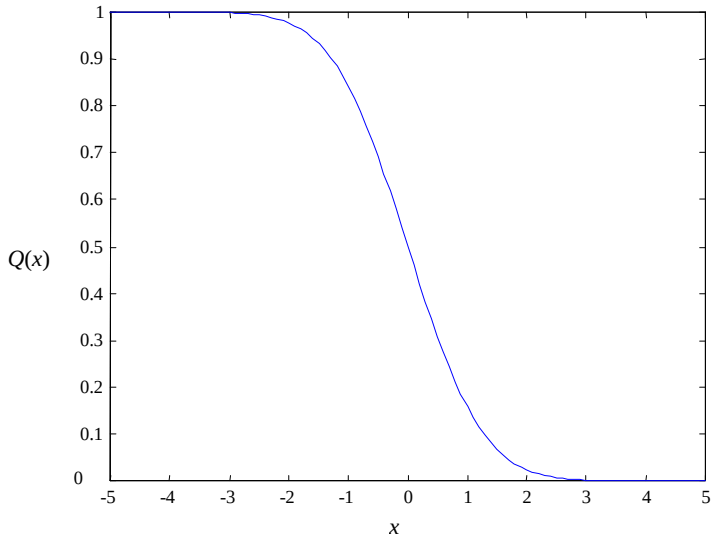

Figura 2.8. Función de error complementario $Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{u^2}{2}\right] du$

La figura 2.8 fue graficada en Matlab usando las siguientes instrucciones

```
» x = -5:0.1:5;
» q = 0.5*erfc(x/sqrt(2));
» plot(x,q)
```

Por lo tanto la regla de decisión para el caso analizado compara la muestra recibida, z, con un umbral de comparación óptimo γ_o.

2.2.4. Probabilidad de bit erróneo

Para el ejemplo binario de la sección 2.2.3, deseamos calcular la **probabilidad de bit erróneo**, P_B, utilizando la regla de decisión de la ecuación 2.27. La probabilidad de un error es calculada al asumir la probabilidad de las diferentes maneras en donde se pueda cometer un error. De la figura 2.7

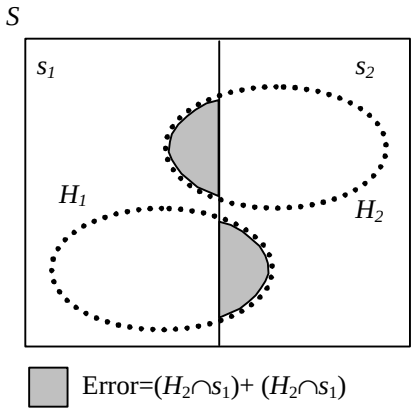

Figura 2.7. Diagrama de conjuntos para calcular P_B.

$$P_B=P(H_2\cap s_1)+P(H_1\cap s_2)$$

$$P_B=P(H_2|s_1)P(s_1)+P(H_1|s_2)P(s_2)$$

Esto es, dado que la clase de señal $s_1(t)$ fue transmitida, un error resulta si la hipótesis H_2 es elegida; o, dado que la clase $s_2(t)$ fue transmitida, un error resulta si se elige la hipótesis H_1. Para el caso especial de pdf's simétricas y señales equiprobables $P(s_1)=P(s_2)=0.5$, podemos escribir

$$P_B=P(H_2|s_1)=P(H_1|s_2)$$

Entonces P_B es numéricamente igual al área debajo de la cola de cualquier pdf, $p(z|s_1)$ o $p(z|s_2)$, evaluando en el lado incorrecto del umbral. Es decir

$$P_B = P(\gamma_o < z \le \infty) = \int_{z=\gamma_o=(a_1+a_2)/2}^{\infty} p(z|s2)dz$$

$$P_B = \int_{z=\gamma_o=(a_1+a_2)/2}^{\infty} \frac{1}{\sigma_o\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{z-a_2}{\sigma_o}\right)^2\right]dz \qquad (2.28)$$

haciendo el cambio de variable

$$u=\frac{z-a_2}{\sigma_o}; \qquad z=u\sigma_o+a_2; \qquad dz=\sigma_o du$$

y cambiando la variable del límite inferior de la integral

$$L(z)=\frac{\frac{1}{\sigma_o\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{z-a_1}{\sigma_o}\right)^2\right]}{\frac{1}{\sigma_o\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{z-a_2}{\sigma_o}\right)^2\right]}$$

$$L(z)=\frac{\exp\left[-\frac{z^2}{2\sigma_o{}^2}\right]\exp\left[-\frac{a_1{}^2}{2\sigma_o{}^2}\right]\exp\left[\frac{2za_1}{2\sigma_o{}^2}\right]}{\exp\left[-\frac{z^2}{2\sigma_o{}^2}\right]\exp\left[-\frac{a_2{}^2}{2\sigma_o{}^2}\right]\exp\left[\frac{2za_2}{2\sigma_o{}^2}\right]}$$

$$L(z)=\exp\left[\frac{z(a_1-a_2)}{\sigma_o{}^2}-\frac{a_1{}^2-a_2{}^2}{2\sigma_o{}^2}\right]\begin{matrix}H_1\\ >\\ <\\ H_2\end{matrix}\frac{P(s_2)}{P(s_1)} \qquad (2.26)$$

En donde a_1 es la componente de la señal de salida en el receptor cuando $s_1(t)$ es enviada y a_2 es la componente de la señal de salida cuando $s_2(t)$ es enviada. La inecuación en 2.24 se mantiene al aplicar transformaciones monótonamente ascendentes. Por lo tanto, tomando el logaritmo natural en ambos lados de la inecuación 2.26

$$l(z)=\frac{z(a_1-a_2)}{\sigma_o{}^2}-\frac{a_1{}^2-a_2{}^2}{2\sigma_o{}^2}\begin{matrix}H_1\\ >\\ <\\ H_2\end{matrix}\ln\left(\frac{P(s_2)}{P(s_1)}\right)$$

Cuando las clases de señal son igualmente probables

$$\ln\left(\frac{P(s_2)}{P(s_1)}\right)=0$$

así es que

$$z\begin{matrix}H_1\\ >\\ <\\ H_2\end{matrix}\frac{a_1{}^2-a_2{}^2}{2(a_1-a_2)}$$

simplificando y definiendo el **umbral óptimo** γ_o

$$z\begin{matrix}H_1\\ >\\ <\\ H_2\end{matrix}\frac{a_1+a_2}{2}=\gamma_o \qquad (2.27)$$

Entonces ahora tenemos una regla de decisión en términos de las pdf's. Si ordenamos la ecuación 2.23, tenemos la ecuación conocida como **prueba de relación de probabilidades**

$$\frac{p(z|s_1)}{p(z|s_2)} \underset{H_2}{\overset{H_1}{\gtrless}} \frac{P(s_2)}{P(s_1)} \qquad (2.24)$$

en donde

$$\text{Relación de probabilidades} \frac{p(z|s_1)}{p(z|s_2)}.$$

La ecuación 2.24 corresponde a hacer una decisión basada en la comparación de una medición de la señal recibida contra un umbral. Ya que la prueba esta basada en la elección de una clase de señal con máxima probabilidad a posteriori, el criterio de decisión es conocido como criterio **máximo a posteriori** (MAP). También es conocido como **criterio de mínimo error**, ya que en el promedio, este criterio proporciona el mínimo de errores en decisiones incorrectas.

2.2.3. Ejemplo de detección de señales

El ejemplo de la figura 2.4 trata con pdf's triangulares. La figura 2.6 ilustra las pdf's para las señales binarias de salida perturbadas por ruido gausiano, $z(T)=a_1+n_o$ y $z(T)=a_2+n_o$ de un receptor típico. Las señales a_1 y a_2, son mutuamente exclusivas e igualmente probables. El ruido, n_o, se asume como una variable aleatoria gausiana independiente con media cero, varianza σ_o^2, y pdf $p(n_o)$ dada por

$$p(n_o) = \frac{1}{\sigma_o\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{n_o^2}{\sigma_o^2}\right)\right] \qquad (2.25)$$

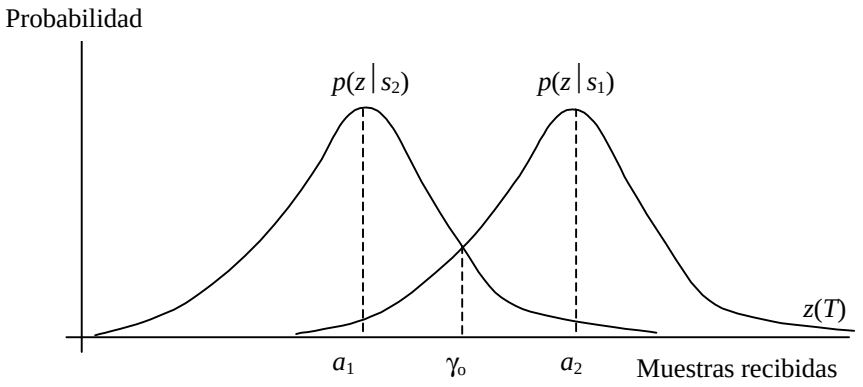

Figura 2.6. Probabilidades condicionales para un receptor binario típico.

Obteniendo la relación de probabilidades

$$L(z) = \frac{p(z|s_1)}{p(z|s_2)}$$

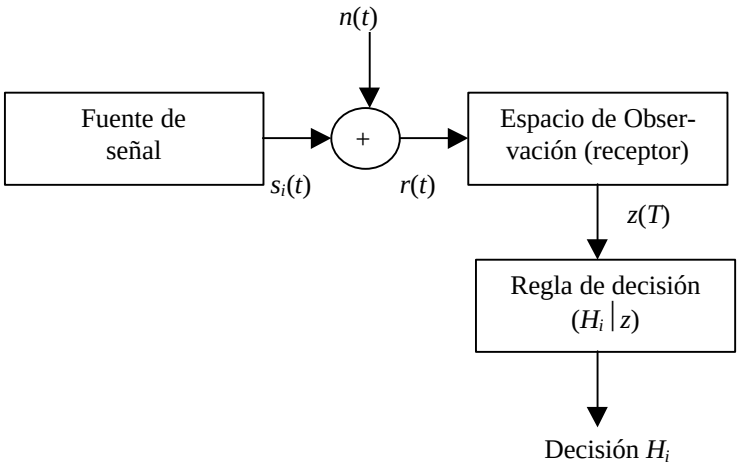

Figura 2.5. Elementos de la teoría de decisión.

En donde la relación del conjunto de hipótesis H_i con las clases de señal s_i se muestra a continuación

$$M \text{ hipótesis} \begin{cases} H_1 : s_1 \\ H_2 : s_2 \\ . \\ . \\ . \\ H_M : s_M \end{cases} M \text{ clases de señal}$$

La muestra $z(T)$ esta constituida de una componente de señal, $a_i(T)$, y una componente de ruido $n_o(T)$. El tiempo T es la duración del símbolo. En cada kT, en donde k es un entero, el receptor usa la regla de decisión para decidir que clase de señal se recibió. Por simplicidad de notación, normalizando T=1s, la ecuación 2.21 se puede escribir como $z=a_i+n_o$.

Un punto razonable para establecer la regla de decisión en el receptor para el caso de dos clases de señal se muestra a continuación

$$P(s_1|z) \underset{H_2}{\overset{H_1}{>}} P(s_2|z) \qquad (2.22)$$

La ecuación 2.22 muestra que debemos elegir la hipótesis H_1 si la probabilidad a posteriori $P(s_1 | z)$ es mayor que la probabilidad a posteriori $P(s_2 | z)$. De otra forma, debemos elegir H_2.

Podemos reemplazar las probabilidades a posteriori de la ecuación 2.22 con sus expresiones equivalentes del teorema de Bayes, ecuación 2.20

$$p(z|s_1)P(s_1) \underset{H_2}{\overset{H_1}{>}} p(z|s_2)P(s_2) \qquad (2.23)$$

cualquier valor en el eje z. Si las pdf's no se solapan, nosotros podemos clasificar la señal con certeza. Para el ejemplo de la figura 2.4, necesitamos una regla que nos ayude a clasificar las señales recibidas, ya que algunas señales caerán en la región en donde las dos pdf se solapan. Consideremos una señal recibida z_a. Asumamos que las dos clases de señal s_1 y s_2, son igualmente probables, y calculemos las dos probabilidades a posteriori alternativas. Además, sugiramos una regla que el receptor pueda utilizar para decidir a que clase de señal pertenece z_a.

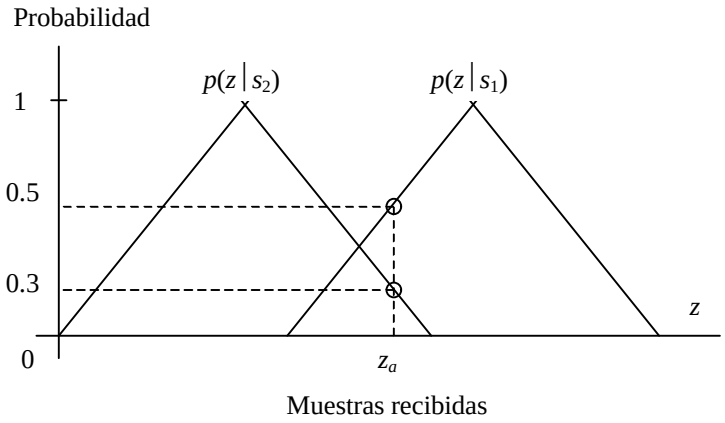

Figura 2.4. Representación gráfica del teorema de Bayes.

Solución

De la figura podemos observar que $p(z_a \mid s_1)=0.5$ y $p(z_a \mid s_2)=0.3$. Sustituyendo en 2.20 con $P(s_1)=P(s_2)=0.5$

$$P(s_1|z_a) = \frac{(0.5)(0.5)}{(0.5)(0.5)+(0.3)(0.5)} = \frac{5}{8}$$

$$P(s_1|z_a) = \frac{(0.5)(0.5)}{(0.5)(0.5)+(0.3)(0.5)} = \frac{5}{8}$$

Una regla es decidir que la señal recibida pertenece a la clase con máxima probabilidad a posteriori (clase s_1). Examine la figura 2.4 y notará que esta regla de máxima probabilidad va con el sentido común.

2.2.2. Elementos de la teoría de la decisión

Los elementos de la teoría de la decisión se muestran en la figura 2.5. La fuente de señal en el transmisor consiste del conjunto de formas de onda $\{s_i(t)\}$, i=1, ..., M. Una forma de onda $r(t)=s_i(t)+n(t)$ se recibe, en donde $n(t)$ es ruido blanco aditivo gausiano (AWGN) introducido por el canal de comunicaciones. En el receptor, la forma de onda es reducida a un número simple, $z(t=T)$, que puede aparecer en cualquier lugar del eje z. Debido a que el ruido es un proceso gausiano y se asume un receptor lineal, la salida, $z(t)$, es también un proceso gausiano, y el número $z(T)$ es una variable aleatoria continua.

$$z(T)=a_i(T)+n_o(T) \qquad (2.21)$$

aleatoria, se especifica el ancho de su pdf. La varianza y el valor cuadrático medio están relacionados por

$$\sigma_x^2 = E(X^2 - 2m_x X + m_x^2)$$

$$\sigma_x^2 = E(X^2) - 2m_x E(X) + m_x^2$$

$$\sigma_x^2 = E(X^2) - m_x^2$$

2.2. TEORÍA DE LA DECISIÓN

Los fundamentos matemáticos de las pruebas de hipótesis descansan en el teorema de Bayes, que se deriva a partir de la definición de la relación entre las probabilidades condicionales y conjuntas, como se demostró en la ecuación 2.7 y que repetimos aquí por comodidad cambiando la notación

$$P(s_i|z_j) = \frac{P(z_j|s_i)P(s_i)}{\sum_{i=1}^{M} P(z_j|s_i)P(s_i)} \qquad i=1, ..., M \qquad (2.19)$$

En una aplicación de comunicaciones, s_i es la i-esima clase de señal de un conjunto de M clases, y z_j es la j-esima muestra de la señal recibida. La ecuación 2.19 puede ser vista como la descripción de un experimento involucrando una muestra recibida y algún conocimiento estadístico de las clases de señal a las cuales la muestra recibida pertenece. La probabilidad de ocurrencia de la i-esima clase de señal, $P(s_i)$, antes del experimento, se llama **probabilidad a priori**. Como un resultado de examinar una muestra particular recibida, z_j, podemos encontrar una medida estadística de la posibilidad de que z_j pertenezca a la clase s_i a partir de las pdf's $P(z_j | s_i)$. Después del experimento, podemos calcular la **probabilidad a posteriori**, $P(s_i | z_j)$, que puede ser vista como un refinamiento de nuestro conocimiento previo. Entonces entramos a un experimento con algún conocimiento previo concerniente a la probabilidad del estado por naturaleza, y después de examinar la muestra de la señal, estaremos provistos de una probabilidad a posteriori. El parámetro $P(z_j)$ es la probabilidad de la muestra recibida, z_j, sobre el espacio entero de clases de señal. Este término, $P(z_j)$, puede ser visto como un factor de escala, ya que su valor es el mismo para cada clase de señal.

2.2.1. Forma mezclada del teorema de Bayes

Para la mayoría de las aplicaciones de ingeniería de comunicaciones, los valores posibles de las muestras recibidas son continuas en rango, debido a la presencia de ruido aditivo gausiano en el canal. Por lo tanto, la forma más útil del teorema de Bayes contiene valores continuos de la pdf en lugar de los discretos de la ecuación 2.19. Escribiendo nuevamente 2.19 y enfatizando este cambio

$$P(s_i|z) = \frac{p(z|s_i)P(s_i)}{\sum_{i=1}^{M} p(z|s_i)P(s_i)} \qquad i=1, ..., M \qquad (2.20)$$

En donde $p(z | s_i)$ es la pdf condicional de la muestra continua recibida, z, condicionada a que se transmitió la clase de señal s_i.

Ejemplo.

Consideremos las dos clases de señal s_1 y s_2, caracterizadas por sus pdf's condicionales de forma triangular $p(z | s_1)$ y $p(z | s_2)$, mostradas en la figura 2.4. Una señal se recibe; esta puede tener

$$p(x) = \frac{dF(x)}{dx} \qquad (-\infty \le x \le \infty) \qquad (2.12)$$

o en forma equivalente

$$F(x) = \int_{-\infty}^{x} p(u)du \qquad (-\infty \le x \le \infty) \qquad (2.13)$$

Frecuentemente estaremos enfrentados al problema de determinar la probabilidad de que una variable aleatoria X caiga dentro de un intervalo (x_1, x_2), en donde $x_2 > x_1$. Para determinar la probabilidad de este evento, empecemos con el evento $\{X \le x_2\}$. El evento puede expresarse siempre como la unión de dos eventos mutuamente exclusivos $\{X \le x_1\}$ y $\{x_1 < X \le x_2\}$. Por lo tanto, la probabilidad del evento puede expresarse siempre como la suma de las probabilidades de los eventos mutuamente exclusivos

$$P(X \le x_2) = P(X \le x_1) + P(x_1 < X \le x_2)$$

$$F(x_2) = F(x_1) + P(x_1 < X \le x_2)$$

$$P(x_1 < X \le x_2) = F(x_2) - F(x_1) = \int_{x1}^{x2} p(x)dx \qquad (2.14)$$

En otras palabras, la probabilidad del evento $\{x_1 < X \le x_2\}$ es simplemente el área debajo de la pdf en el rango $x_1 < X \le x_2$.

Promedios estadísticos

El **valor medio**, m_x, o **valor esperado** de una variable aleatoria X esta definido por

$$m_x = E(X) = \int_{-\infty}^{\infty} xp(x)dx \qquad (2.15)$$

En donde $E(.)$ es el operador valor esperado. El enésimo momento de una distribución de probabilidad de la variable aleatoria X esta definido como

$$E(X^n) = \int_{-\infty}^{\infty} x^n p(x)dx \qquad (2.16)$$

Para el propósito de análisis en los sistemas de comunicaciones, los momentos más importantes de X son los dos primeros momentos. Entonces, n=1 en la ecuación 2.16 proporciona m_x como en 2.15, y n=2 proporciona el **valor cuadrático medio** de X como en la siguiente ecuación

$$E(X^2) = \int_{-\infty}^{\infty} x^2 p(x)dx \qquad (2.17)$$

En forma adicional se pueden definir los momentos centrales, que son los momentos de la diferencia entre X y m_x. El segundo momento central, conocido como la **varianza** de X, σ_x^2, esta definido por

$$\mathrm{var}(X) = \sigma_x^2 = E((X - m_x)^2) = \int_{-\infty}^{\infty} (x - mx)^2 p(x)dx \qquad (2.18)$$

La raíz cuadrada de la varianza es conocida como la **desviación estándar** de X, σ_x. La varianza es una medida de la aleatoriedad de la variable aleatoria X. Al especificar la varianza de una variable

Dada una variable aleatoria $X=X(s)$, consideremos el evento $\{X \leq x\}$ en donde x es cualquier número en el intervalo $(-\infty, \infty)$. Podemos escribir la probabilidad de este evento como $P(X \leq x)$ y su notación simple como $F(x)$

$$F(x)= P(X \leq x) \qquad (-\infty \leq x \leq \infty) \qquad (2.11)$$

La función $F(x)$ es la **función de distribución de probabilidad** de la variable aleatoria X. También es conocida como la **función acumulativa de probabilidad**, (cdf). Dado que $F(x)$ es una probabilidad, su rango esta limitado a $0 \leq F(x) \leq 1$. De echo, $F(-\infty)=0$ y $F(\infty)=1$. Por ejemplo, la variable aleatoria $X(s)=s$ generada por el lanzamiento del dado es mostrada en la figura 2.2.

s	$X=X(s)=s$	$P(X)$	$F(x)=P(X \leq x)$
1	1	1/6	1/6
2	2	1/6	2/6
3	3	1/6	3/6
4	4	1/6	4/6
5	5	1/6	5/6
6	6	1/6	6/6

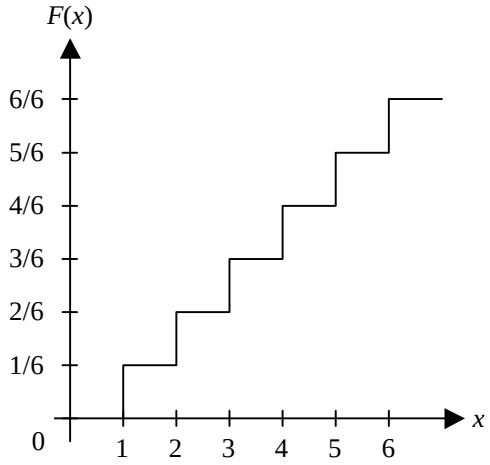

Figura 2.2. Ejemplo de cdf para el experimento del dado.

La cdf de una variable aleatoria continua y típicamente aparece como en la figura 2.3. Esta es una función de x lisa, no decreciente.

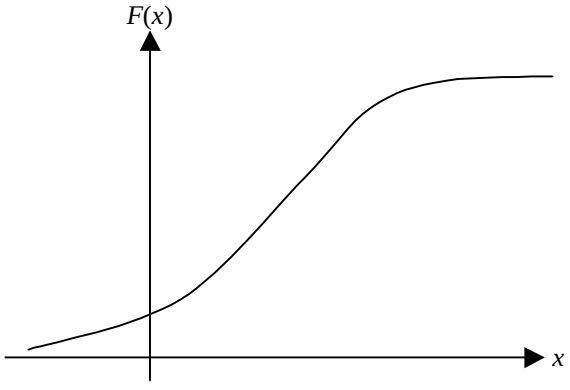

Figura 2.3. Ejemplo de cdf de una variable aleatoria continua.

La derivada de $F(x)$, denotada como $p(x)$, es conocida como la **función de densidad de probabilidad** (pdf) de la variable aleatoria X. Entonces tenemos que

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{P(AB)}{P(A)} \qquad (2.5)$$

Escribiendo 2.4 y 2.5 de una forma diferente

$$P(AB)=P(A \mid B)P(B)=P(B \mid A)P(A) \qquad (2.6)$$

Una relación extremadamente útil para probabilidades condicionales es el **teorema de Bayes** que muestra que la probabilidad para los eventos mutuamente exclusivos A_i, i=1, 2, …, n, es

$$P(A_i|B) = \frac{P(A_iB)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{n} P(B|A_j)P(A_j)} \qquad (2.7)$$

Demostración: Para el numerador de 2.7

$$P(A_iB)=P(B \mid A_i)P(A_i)$$

Para el denominador, de 2.2 y 2.6

$$P(B) = \sum_{j=1}^{n} P(A_jB) = \sum_{j=1}^{n} P(B|A_j)P(A_j)$$

Otro concepto importante en la teoría de la probabilidad es la **independencia estadística**. Para explicar este concepto, consideremos los eventos A y B y su probabilidad condicional $P(A \mid B)$, que es la probabilidad de la ocurrencia de A dado que B ha ocurrido. Suponga que la ocurrencia de A no depende de la ocurrencia de B. Esto es

$$P(A \mid B)=P(A) \qquad (2.8)$$

Sustituyendo 2.8 en 2.6

$$P(AB)=P(A)P(B) \qquad (2.9)$$

Esto significa que la probabilidad conjunta de los eventos A y B se puede factorizar en el producto de sus probabilidades marginales $P(A)$, $P(B)$ respectivamente. De esta forma, la expresión general de 2.9 es

$$P\left(\bigcap_i A_i\right) = \prod_i P(A_i) \qquad (2.10)$$

2.1.3. Variables aleatorias, distribución de probabilidad y densidad de probabilidad

Dado un experimento cuyo espacio de eventos es S y elementos $s \in S$, se define una función $X(s)$ cuyo dominio es S y cuyo rango es un conjunto de números en el eje de los números reales. La función $X(s)$ es la **variable aleatoria**. Por ejemplo, en el experimento del lanzamiento del dado S={1, 2, 3, 4, 5, 6}. Una variable aleatoria en este espacio de eventos puede ser $X(s)=s$, entonces los resultados del experimento son mapeados en los enteros 1, ..., 6.

$$P(B) = \frac{3}{6}$$

Entonces de 2.1, la probabilidad de la unión de estos dos eventos mutuamente exclusivos es

$$P(A \cup B) = P(A) + P(B) = \frac{2}{6} + \frac{3}{6} = \frac{5}{6}$$

2.1.2. Eventos conjuntos y probabilidad condicional

En lugar de tratar con un experimento sencillo, asumamos dos eventos y consideremos sus resultados. En general, si un experimento tiene los posibles resultados A_i, i=1, 2, …, n, y un segundo experimento tiene los posibles resultados B_j, j=1, 2, …, m, entonces el experimento combinado tiene los posibles resultados conjuntos $(A_iB_j)=(A_i \cap B_j)$, i=1, 2, …, n, j=1, 2, …, m. Asociado a cada resultado del conjunto (A_iB_j) existe la **probabilidad conjunta** $P(A_iB_j)$ que satisface la condición

$$0 \leq P(A_iB_j) \leq 1$$

Asumiendo que los resultados A_i, i=1, 2, …, n, son mutuamente exclusivos, es fácil mostrar de la figura 2.1 que

$$P(B_j) = \sum_{i=1}^{n} P(A_iB_j) \qquad (2.2)$$

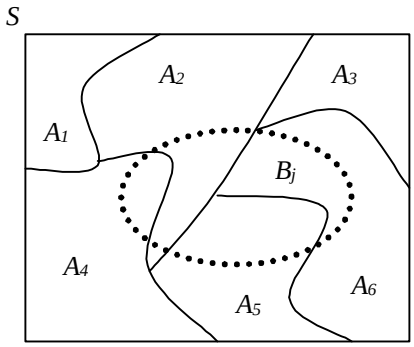

Figura 2.1. Eventos conjuntos.

En forma similar, si los resultados B_j, j=1, 2, …, m, son mutuamente exclusivos tenemos

$$P(A_i) = \sum_{j=1}^{m} P(A_iB_j) \qquad (2.3)$$

La **probabilidad condicional** de un evento A asumiendo B, denotada por $P(A|B)$, es por definición

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(AB)}{P(B)} \qquad (2.4)$$

En donde $A \cap B = AB$ y $P(B) \neq 0$. En forma similar, la probabilidad condicional de un evento B asumiendo A es

$$\overline{A} = \{1, 3, 5, 6\}$$

Se dice que dos eventos son **mutuamente exclusivos** si ellos no tienen eventos elementales en común. Por ejemplo, *A* y *B* son eventos mutuamente exclusivos y *A* y $\overline{A}$ también lo son.

La **unión** (suma) de dos eventos es un evento que consiste de todos los eventos elementales en los dos eventos. Por ejemplo, la unión de *B* y *C*, denotada como $B \cup C$, es el evento *D*

$$D = B \cup C = \{1, 2, 3, 6\}$$

Similarmente

$$A \cup \overline{A} = S$$

Por otra parte, la **intersección** de dos eventos es un evento que consiste de los eventos elementales que son comunes a ambos eventos. Entonces, si $E = B \cap C$ representa la intersección de los eventos *B* y *C*, entonces

$$E = B \cap C = \{1, 3\}$$

Cuando dos eventos son **mutuamente exclusivos**, la intersección es el conjunto vacío. Por ejemplo, *A* y *B* son eventos mutuamente exclusivos, y *A* $\overline{A}$ y también lo son ya que

$$A \cap B = \varnothing$$

$$A \cap \overline{A} = \varnothing$$

Las definiciones de unión e intersección se pueden extender a más de dos eventos de forma similar.

2.1.1. Los axiomas

Asociado a cada evento *A* contenido en el evento seguro *S*, esta un número $P(A)$ llamado la **probabilidad** del evento *A*. Este número se elige de manera que satisface las siguientes tres condiciones:

$$P(A) \geq 0$$

$$P(S) = 1$$

Sí $A \cap B = \varnothing$ entonces $P(A \cup B) = P(A) + P(B)$

En donde la forma general del tercer postulado se puede extender a más de dos eventos de la siguiente manera: Suponga que A_i, $i = 1, 2, \ldots$, son eventos mutuamente exclusivos de tal forma que

$$A_i \cap A_j = \varnothing \qquad i \neq j = 1, 2, \ldots$$

Entonces la probabilidad de la unión de estos eventos mutuamente exclusivos es

$$P\left(\bigcup_i A_i\right) = \sum_i P(A_i) \qquad (2.1)$$

En el ejemplo anterior, en donde $A = \{2, 4\}$, $B = \{1, 3, 6\}$ y $A \cap B = \varnothing$ tenemos que

$$P(A) = \frac{2}{6}$$

Capítulo 2
Fundamentos de la teoría de decisión

Los fundamentos de la teoría de la decisión están construidos sobre las disciplinas matemáticas de la probabilidad y las variables aleatorias.

Los elementos básicos del problema de la decisión son: 1) Un conjunto de hipótesis que caractericen los posibles estados verdaderos por naturaleza, 2) Una prueba en dónde los datos se obtienen a partir de aquellos que deseamos inferir la verdad, 3) Una regla de decisión que opera sobre los datos para decidir, de forma óptima, cual hipótesis describe mejor el estado verdadero por naturaleza, y 4) Un criterio de optimización. El criterio de optimización que eligiremos para la regla de decisión es minimizar la probabilidad de hacer una decisión errónea.

2.1. FUNDAMENTOS DE PROBABILIDAD

En teoría de probabilidad se utiliza la siguiente terminología:

- El espacio *S* es conocido como el **evento seguro** o espacio de muestreo y sus elementos son resultados experimentales.
- El **conjunto vacío** $\varnothing$ es el evento imposible.
- Un evento, por ejemplo *A*, es un **subconjunto** de *S* y puede consistir de eventos elementales del evento seguro.

Por ejemplo en el experimento del lanzamiento de un dado y observar el resultado

$$S=\{1, 2, 3, 4, 5, 6\}$$

En donde los enteros 1, ..., 6 representan el número de puntos en las seis caras del dado. Por otro lado, se pueden definir los eventos *A*, *B* y *C* como

$$A=\{2, 4\}$$

$$B=\{1, 3, 6\}$$

$$C=\{1, 2, 3\}$$

Adicionalmente, el complemento del evento *A* denotado como $\overline{A}$, consiste de todos los eventos elementales en el evento seguro *S* que no están en *A*