**Teoría de Detección de Señales**

Adriana F. Chávez De la Peña

[adrifelcha@gmail.com](mailto:adrifelcha@gmail.com)

-

Universidad Nacional Autónoma de México

Facultad de Psicología

Laboratorio 25

Proyecto PAPIME

1. **Introducción**

Con frecuencia nos enfrentamos a situaciones en que debemos decidir si ‘algo’ está o no ocurriendo para poder actuar en consecuencia, (e.g. ‘¿mi mamá está enojada?, ¿mi perro está enfermo?, ¿esta comida está pasada?). No parecería ser un gran problema si asumiéramos que somos infalibles en la detección de dichos casos, o bien, que aquello que nos interesa detectar es un evento tan particular que es completamente inconfundible con nada más en el mundo. Sin embargo, este casi nunca parece ser el caso, el mundo siempre está cargado de ruido e incertidumbre: tanto la información con base en la cual buscamos tomar una decisión, como la precisión con que nuestro sistema es capaz de evaluarla, son imperfectos.

La Teoría de Detección de Señales (TDS) constituye uno de los modelos estadísticos más sólidos y ampliamente estudiados en Psicología para dar cuenta del problema al que se enfrenta cualquier sistema ante este tipo de situaciones. Para ello, identifica como **‘señal’** a aquel evento que interesa detectar y define como **‘ruido’** al resto de los estímulos con que coexiste y que pueden llegar a confundirse con la misma.

Ejemplo:

Acabas de llegar a una fiesta y estás buscando a tu mejor amigo. Sabes cómo se ve y comienzas a buscar con la mirada a alguien que se le parezca, sin embargo, es muy probable que más de una persona comparta uno, o más, de sus rasgos, por lo que seguramente tendrás que mirar más de una vez antes de decidir que has encontrado a tu amigo y acercarte a saludarlo.

Señal: Tu mejor amigo.

Ruido: El resto de los invitados a la fiesta.

La TDS ha sido utilizada para describir un amplio número de fenómenos. Cuando hablamos de detección de señales, podemos referirnos a la señal tanto como un estímulo sensorial concreto (e.g. una luz, tono, u objeto particular), como una categoría más abstracta (e.g. una enfermedad, una emoción o un estado). (Ver la sección de lecturas recomendadas para más ejemplos).

De manera muy general, se puede hablar de dos grandes supuestos en torno a los cuales se des arrolla la TDS:

1. Hay variabilidad, siempre (Ver Fig. 1).
   1. Hay variabilidad en la señal

La idea central de variabilidad radica en la noción de que ningún estímulo se presenta ni se percibe exactamente igual cada vez que nos encontramos con él. Es decir, cada vez que nos encontramos con la señal en el mundo, ésta puede hacerlo dentro de un rango de posibilidades con cierta probabilidad. Esta idea se muestra gráficamente en la Figura 1, con la distribución normal azul identificada bajo la etiqueta de ‘Señal’. La idea es que la señal va adoptar una cierta forma de entre los puntos que abarca la distribución de probabilidad; siendo unas más probables que otras, conforme se aproximan a la media.

La variabilidad en la señal puede interpretarse en términos de dos fuentes: la percepción del sistema que ejecuta la tarea de detección, o la propia presentación estímulo en sí mismo. En el primer caso, se asume que cada vez que vemos un mismo estímulo que se mantiene constante en términos de sus propiedades físicas, (e.g. una luz o un tono), este puede ser percibido de manera distinta en cada presentación (i.e. unas veces parecerá un poco más intenso y otras, un poco menos). En el segundo caso, se asume que la señal puede tomar más de una forma, (e.g. si la señal es el enojo de un amigo, existen ciertos rasgos que son más o menos comúnmente asociados a su enfado; pero no siempre se va a ver exactamente igual).

* 1. Hay variabilidad en el entorno.

Por otro lado, es importante tomar en cuenta que las señales que interesa detectar coexisten en el mundo con otros estímulos; algunos de los cuales pueden llegar a producir una evidencia similar a la de nuestra señal y ser, por tanto, confundidos con la misma. Esta idea se representa en la Figura 1 con la distribución normal negra identificada bajo el nombre de ruido, que se traslapa con cierta probabilidad con la distribución de señal.

La Figura 1 corresponde a la representación gráfica asumida por la TDS para toda tarea de detección, e incorpora la noción de variabilidad previamente expuesta.

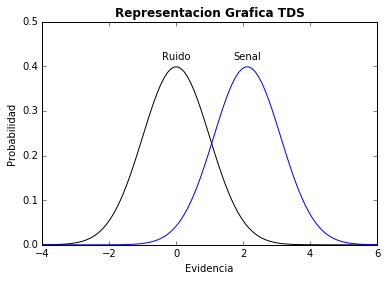


Fig 1. Representación gráfica de una tarea de detección, de acuerdo a la TDS.

El soporte de las distribuciones, identificado en la Figura 1 bajo el nombre de ‘Evidencia’ rara vez se define con precisión, teniendo una concepción más bien abstracta; La idea general es que cuando queremos detectar una señal particular, comenzamos a recolectar un tipo de evidencia específico a la tarea ante la que nos encontramos. Lo más importante, es que la señal siempre va a estar asociada en mayor medida con dicha evidencia, distribuyéndose siempre en valores situados por encima (a la derecha, en la Figura 1) del ruido.

Ejemplo:

Si queremos detectar si nuestro paciente tiene o no depresión a partir de la aplicación de un instrumento especializado, en el eje de la Evidencia tendríamos los posibles puntajes obtenibles en la prueba. En este escenario tendríamos que las personas con depresión obtienen cierto rango de puntajes con cierta probabilidad (sobre los cuales se despliega la distribución normal de señal), mientras que las personas sin depresión suelen caer en un rango de puntajes menor, pero que sin embargo en ocasiones, aunque con una probabilidad relativamente baja, puede coincidir con los puntajes obtenidos por personas con depresión.

Este primer supuesto de variabilidad, como algo inherente a todo estímulo y sistema, nos lleva a hablar de la **discriminabilidad de la señal**, o bien, de la **sensibilidad del sistema** ante la señal, que el modelo de detección de señales va a representar con un mismo parámetro: **d’**, que corresponde a la distancia entre las medias de las distribuciones de ruido y señal, y cuyo cómputo abordaremos más afondo más adelante con ayuda de nuestro graficador en Python.

1. Las consecuencias importan:

La TDS define toda tarea de detección como una tarea de decisión, donde el fin último por el cual el organismo se interesa en determinar si la señal está o no presente, es el de guiar su curso de acción. Es decir, el comportamiento de cualquier organismo va a depender de las señales que este detecta en su entorno.

Ejemplo:

1. A partir de la lectura de una tomografía, tengo que determinar si mi paciente tiene cáncer para comenzar un tratamiento.
2. Estoy buscando a mi novio en el aeropuerto, que acaba de regresar de un viaje de 3 meses, para recibirlo de manera efusiva.

Una consecuencia directa de la variabilidad involucrada en el entorno de decisión, es que el desempeño de todo sistema de detección es propenso a cometer errores y emitir un juicio de presencia o ausencia de la señal, que puede no coincidir con el estado del mundo. Dependiendo la correspondencia entre el estado del mundo y el juicio emitido por el sistema de detección, la TDS maneja las clasificaciones de respuesta mostradas en la Tabla 1; donde las celdas 2 y 3, corresponden a los errores posibles.

|  |  |  |  |
| --- | --- | --- | --- |
|  | | Estado real del mundo | |
| La señal está | La señal NO está |
| Juicio emitido por el sistema | “La señal sí está” | Hit  (1) | Falsa Alarma  (2) |
| “La señal no esta” | Omisión  (3) | Rechazo correcto  (4) |

Tabla 1.Clasificación de las posibles respuestas dadas ante una tarea de detección.

La TDS asume que el organismo fija un **criterio de elección** a lo largo del eje de la Evidencia, que va a determinar a partir de cuánta evidencia va a juzgar la señal como presente. Dicho criterio se va a representar como una línea transversal que atraviesa ambas distribuciones en una determinada altura, y se le va a identificar con el parámetro **k**. La TDS asume que los organismos van a fijar esta regla de elección, ponderando la información a la que tienen acceso con la información que poseen sobre la estructura de la tarea (i.e. cómo suele presentarse la señal, qué tan probable es que se presente, etc.)

* 1. Sesgo

Sin embargo, no todos los errores tienen el mismo costo. Imaginemos el caso de una presa en potencia que busca determinar si el sonido que acaba de escuchar en la maleza corresponde o no con el de un depredador; no hay tiempo que perder, y el costo que dicho organismo tendría que pagar por cometer una falsa alarma (gasto innecesario de energía) o una omisión (morir devorado) es sustancialmente diferente. En este escenario particular, es muy probable que la presa sea mucho más propensa a correr por su vida, juzgando la presencia del depredador a partir de valores menores de evidencia.

Esta discrepancia en el peso que se le da a las consecuencias posibles de emitir una u otra respuesta y obtener uno de los cuatro posibles resultados, suele representarse en términos de una matriz de pagos, que nos ayude a definir cuáles son las consecuencias que el organismo buscará evitar o promover, según sea el caso, en mayor medida.

Ejemplo:

¿Eso que escuché en la maleza es un depredador?

|  |  |  |
| --- | --- | --- |
|  | Es un depredador | No es nada |
| ‘Sí’ | Me salvo | Gasto de energía innecesario |
| ‘No’ | Muerte | Vivo |

Ya sea por los distintos pesos que tengan las posibles consecuencias para el organismo, o porque se tiene una preferencia o predisposición inherente a decretar la presencia o ausencia de la señal, la TDS asume que el desempeño de los organismos que se enfrentan a tareas de detección de señales va a depender tanto de la calidad de la información a la que se tiene acceso (dentro de lo que se incluye la importancia de la variabilidad, que determina tanto la discriminabilidad de la señal como la sensibilidad del sistema ante la misma), como de un sesgo de elección.

La localización del criterio en nuestro eje de evidencia recolectada va a estar altamente influida por el sesgo que tenga nuestro sistema. Podemos hablar entonces de dos tipos distintos de sesgo: conservador y liberal. El primero, favorece la emisión de respuestas negativas al desplazar el criterio a la derecha y requerir al sistema la recolección de mayores niveles de evidencia antes de dar por detectada la señal. El segundo, promueve la detección de la señal, situando el criterio de elección hacia la izquierda, emitiendo un juicio de detección con valores menores de evidencia. Nótese que un sistema carente de sesgo, sería aquel que situara su criterio de elección justo en el punto en que las dos distribuciones se juntan, donde la probabilidad de cometer cualquiera de los tipos de acierto y errores, son iguales entre sí.

Para cuantificar el **sesgo del sistema**, la TDS nos proporciona dos medidas: la primera de ellas corresponde a la distancia entre el punto de sesgo-neutro (i.e. el punto donde se interceptan ambas distribuciones) y la localización del criterio (**c**) y la segunda, a la razón entre el punto en que el criterio toca la distribución de la señal y la distribución de ruido (**β**). Dichos parámetros no sólo permiten saber cuán grande es el sesgo del sistema, sino que facilitan su clasificación en las categorías previamente expuestas, siendo el caso que si β<1 o C<0, sabemos se trata de un sesgo liberal y si β>1, C>0, hablamos de un sesgo conservador.

1. **Parámetros de la TDS (Ejercicios con Graficador)**

Los archivos cuyo título comienza con ‘TeoremadeBayes\_Graficador\_’, contienen el código que estaremos utilizando como material de apoyo para entender los parámetros involucrados en la Teoría de Detección de Señales. Para tener acceso a los graficadores interactivos, recuerda que debes correr el código desde la consola; puedes acceder y modificar el código y los comentarios correspondientes desde cualquier IDE de Python o editor de texto.

Antes de ahondar a detalle en los parámetros, hay que declarar un par de supuestos formales que hace la Teoría para facilitar la representación gráfica del modelo y la estimación paramétrica:

1. En su forma clásica, la TDS asume que las distribuciones de ruido y señal son distribuciones normales.
2. La TDS asume equivarianza entre las distribuciones de ruido y señal. Es decir, asume que la dispersión de ambas distribuciones es la misma, fijando la desviación estándar a 1.
3. Para facilitar la estimación paramétrica, a la distribución de ruido (que por definición debe aparecer siempre a la izquierda de la señal) se le asigna una media de 0.

Una revisión un poco más profunda en la literatura, (sobre todo en literatura más formal y especializada) nos demuestra que, si se cuenta con información suficiente, los supuestos 1 y 2 pueden violarse. Por ejemplo, en estudios de memoria de reconocimiento, donde se les pide a los participantes que discriminen entre estímulos que les fueron presentados en una etapa previa y estímulos completamente nuevos, los resultados demuestran consistentemente que la distribución de señal (de estímulos previamente vistos), tiene una mayor varianza que la distribución de ruido; dicho eso, si se piensa utilizar la TDS como modelo de referencia para una tarea de memoria de reconocimiento, se puede hacer caso omiso del supuesto de equivarianza.

**2.a Tasas de ejecución**

Como se mencionó previamente, al realizar una tarea de detección existen dos posibles tipos de aciertos: al detectar la señal (Hits) y al rechazar el ruido (Rechazos), y dos posibles tipos de errores: los falsos positivos (Falsas alarmas) y los falsos negativos (Omisiones). La materia prima con base en la cual funciona el modelo propuesto por la TDS, son las tasas de aciertos y errores cometidos durante la tarea, de manera que por cada participante que pasa por una tarea de detección, tenemos cuatro tasas que describen su ejecución:

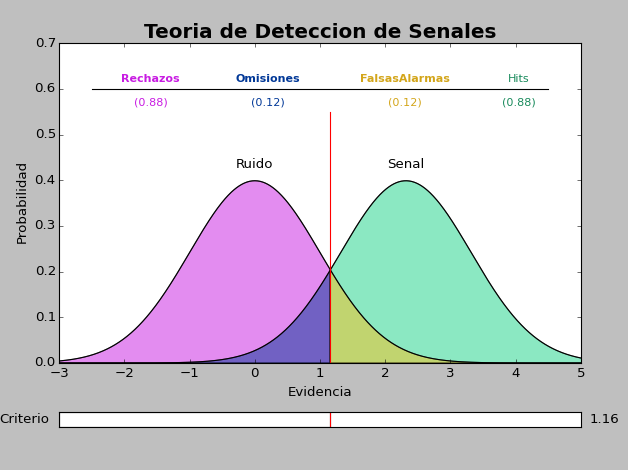
La Tabla 2 ilustra el cómputo de las cuatro tasas de ejecución, como una relación entre el resultado obtenido y el tipo de ensayo con base en el que se le definió como tal. Es decir, tenemos dos tasas definidas en relación al número total de ensayos con la señal (la tasa de hits y la tasa de omisiones) que nos dicen qué proporción de los ensayos con señal fueron detectados correctamente y cuáles se dejaron pasar; y tenemos dos tasas definidas en relación al total de ensayos con ruido (la tasa de falsas alarmas y la tasa de rechazos correctos) que nos describen la relación de los ensayos con ruido que fueron discriminados correctamente y aquellos que se confundieron con la señal.

|  |  |  |  |
| --- | --- | --- | --- |
|  | | ¿La señal estaba presente? | |
| Sí: SEÑAL | No: RUIDO |
| ¿Qué respondió  el sujeto? | Sí |  |  |
| No |  |  |

Tabla 2. Ilustración de las cuatro tasas de ejecución obtenidas por participante en tareas de detección.

Para realizar el análisis de datos, bajo el marco de la TDS, sólo necesitaremos un par de estas tasas: la tasa de hits y la tasa de falsas alarmas. Esto bajo el entendido de que las tasas de omisión y rechazos correctos no son más que su complemento, respectivamente, y que estas dos tasas contienen toda la información que necesitamos sobre el desempeño de los participantes.

La idea general de la importancia de estas tasas de ejecución, es que cada una representa el área de las distribuciones de ruido y señal que cae a la izquierda o derecha del criterio de decisión. El archivo ‘SDT\_Graficador\_Tasas.py’ presenta de manera gráfica esta idea (Ver Fig. 2)

Fig. 2. El Graficador de Tasas de Ejecución ilustra la idea de que, dependiendo la localización del criterio de decisión que esté usando el participante, cambia la proporción de aciertos y errores que se puedan cometer en función al área bajo la curva. En la parte superior del simulador se muestra la proporción de cada distribución que cae bajo cada clasificación hecha por el modelo. El slider colocado en la parte inferior del graficador permite al usuario alterar la posición del criterio sobre el eje de decisión y alterar así la probabilidad de obtener cada outcome posible. Con fines ilustrativos, en este graficador el valor de d’ se mantiene constante y lo único que se altera es la localización de la línea que atraviesa ambas distribuciones (y que simula al criterio de elección).

La Fig. 2 presenta una vista previa del Graficador. En ella, se puede observar cómo la distribución de señal y la distribución de ruido se dividen a ambos lados del criterio, en los aciertos y errores correspondientes. El supuesto descriptivo que hace la teoría, es que el organismo computa la evidencia que observa con la información que tiene sobre las consecuencias de cometer uno u otro posible error para colocar un criterio de elección que maximice sus ganancias, o bien, minimice sus pérdidas. Esto se ilustra en el Graficador con el slider ubicado en la parte inferior, con el que se puede alterar la posición del criterio sobre el eje de decisión y observar los cambios en la probabilidad de cometer ciertos aciertos o ciertos errores que se dan en consecuencia.

Para la estimación paramétrica se utiliza la misma lógica, pero se sigue el procedimiento inverso. Dado que no podemos observar ni cuantificar de manera directa el criterio usado por los participantes para responder a la tarea, qué tan juntas o separadas se encuentran las distribuciones de ruido y señal para cada participante o qué tipo de sesgo pudieran estar siguiendo, utilizamos las tasas de ejecución para hacer inferencias sobre la localización del criterio, la diferencia entre las medias de ambas distribuciones y el grado en que una respuesta se favorece sobre otra.

A partir de ahora comenzaremos a hablar sobre cómo se calculan cada uno de los parámetros del modelo, de acuerdo a la teoría clásica que sigue los supuestos estadísticos previamente descritos. Es importante aclarar que el Graficador de Tasas previamente expuesto no representa la teoría con entera precisión; el propósito de ese primer Graficador es simplemente ilustrar cómo describe la TDS el comportamiento de un sistema que se enfrenta ante una tarea de detección, donde existen dos distribuciones que se sobreponen. El Graficador permite manipular directamente la localización del criterio, con la simpleza que implicaría desplazar una línea vertical sobre el eje de decisión y ver qué consecuencias tiene sobre la probabilidad de obtener un tipo particular de acierto o error.

**2.b Criterio k**

Una vez que hemos resumido el desempeño de nuestro participante en la tarea de detección, el parámetro cuya estimación resulta más sencilla y directa es el Criterio (k). Entender cómo se computa el parámetro nos requiere únicamente de mantener presente el supuesto de que el Ruido se distribuye normalmente y se va a localizar siempre a la izquierda de la señal, por lo que le asignamos una media de cero para tener un punto de referencia para estimar el espacio en que se desarrollan el resto de los parámetros.

Para calcular el criterio lo único que necesitamos es conocer la tasa de Falsas Alarmas, que tal y como mencionábamos en el segmento anterior, nos indica qué proporción de la distribución de ruido cae a la derecha del criterio. Dado que a la distribución de ruido, le fue asignada arbitrariamente una media de cero, podemos asignar un valor al punto en que el criterio corta la distribución de ruido y define las tasas de Rechazos y Falsas Alarmas obtenidas por el participante. Conociendo el área de la distribución de Ruido que cae bajo el criterio, (el complemento de la tasa de Falsas Alarmas, o bien, la Tasa de Rechazos correctos), y sabiendo que la distribución tiene una desviación estándar de 1, podemos convertir el valor de la tasa (que corresponde a la probabilidad de cometer un rechazo correcto, de acuerdo al área bajo la curva) en Puntajes Z y conocer la localización del criterio.

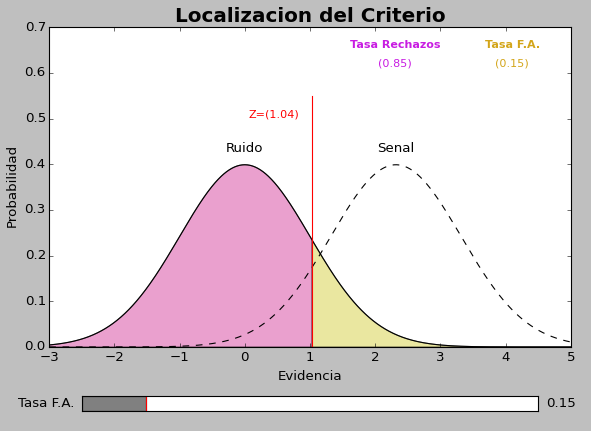
El archivo titulado “SDT\_Graficador\_Criterio.py” contiene el código para correr un graficador que ilustra el cálculo del criterio como una función de la tasa de Falsas Alarmas obtenida por el participante. El parámetro k, por lo general, va estar representado por un número natural (un número positivo), que indica en términos de Puntajes Z la posición del criterio sobre el eje de decisión, relativo a la distribución de ruido con media cero. El criterio sólo tiene valores positivos, porque normalmente se espera que la tasa de falsas alarmas nunca tenga un valor mayor a 0.5 (las consecuencias de una tasa de Falsas Alarmas tan alta, se expondrán con más claridad en el apartado correspondiente a la d’)

Fig 3. Graficador del cálculo de la localización del criterio sobre el eje de decisión. El graficador ilustra la manera en que la tasa de falsas alarmas (manipulable por el usuario) refleja la localizacón del criterio y permite su cuantificación a partir de la conversión de los valores p a Puntajes Z.

**2.c Discriminabilidad d’**

**2.d Sesgo: C**

**2.e Sesgo: β**

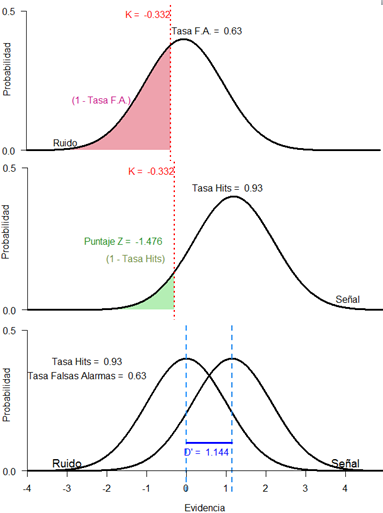
* Discriminabilidad (d’). Para encontrar la distancia entre las medias de la distribución de ruido y señal, necesitamos saber el punto en que el criterio toca cada distribución. Para ello, calculamos las probabilidades complementarias a las tasas de hits y falsas alarmas y las traducimos a puntajes Z (Ver Fig. 3). Dado que el puntaje Z funciona como una medida de dispersión de la media, basta con restar el puntaje Z de la intersección del criterio con la distribución de señal a el puntaje Z de intersección con la distribución de ruido para conocer la localización de la media de la señal. Por definición, d’ sólo puede tener valores positivos ya que la teoría asume que la distribución de señal siempre está a la derecha de la distribución de ruido porque contiene una mayor cantidad de la evidencia con base en la cual se hace el juicio de detección de la señal.

Fig. 3. Ilustración del cálculo de d’, para un caso donde la tasa de Hits es de 0.93 y la tasa de Falsas alarmas, 0.63, a partir de la resta del puntaje Z correspondiente a la intersección del criterio con la distribución de señal (la distribución de en medio) del puntaje Z correspondiente a su intersección con la distribución de ruido (i.e. k; distribución superior). d’ se señala como la distancia entre las distribuciones del panel inferior.

* Sesgo (β). La medida más comúnmente usada para evaluar el sesgo corresponde a la razón entre el punto en que el criterio toca a las distribuciones de señal y ruido. En otras palabras, es la razón de verosimilitud a la altura del criterio. Para calcular β, hay que encontrar el valor ordinal (O) correspondiente al punto en que el criterio cruza cada una de las distribuciones y calcular la razón entre el valor correspondiente a la distribución de señal y la distribución de ruido. El valor de Beta nos indica directamente el tipo de sesgo que muestra el sistema. Si Beta es mayor a uno, quiere decir que el criterio está situado en un punto donde la distribución de señal tiene valores más altos que la distribución de ruido, mostrando un sesgo conservador; por el contrario, si Beta es menor a uno, asumimos que el criterio cae en un punto donde la distribución de ruido tiene valores mayores a la de señal, mostrando un sesgo liberal.

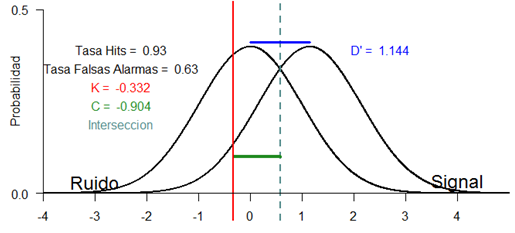
* Sesgo (C). Una segunda medida de sesgo propuesta por el modelo, es la distancia entre el criterio y el punto en que ambas distribuciones se intersectan. Partiendo del supuesto de que las distribuciones comparten una misma desviación estándar, la localización la intersección se obtiene dividiendo d’ entre dos. Finalmente, restamos ésta cifra a la localización del criterio. Si C tiene un valor positivo quiere decir que el criterio está localizado a la derecha del punto de intersección y por tanto, hablamos de un sistema con sesgo conservador; en tanto que si C es negativa, sabemos que el criterio se sitúa a la izquierda y el sistema está respondiendo de acuerdo a un sesgo liberal.

Fig. 4. Ejemplificación, a partir de los valores de d’ (en azul) y k (en rojo) previamente estimados, de la medida de sesgo C (en verde).

1. Ejercicio – (Simulació n de un experimento)

Lecturas recomendadas:

Material amigable para principiantes

Killeen, P (2014). Signal Detection Theory.

Lee, M., Wagenmakers, E. (2014) Bayesian Cognitive Modeling: A practical course. Chapter 11: Signal detection theory.

Swets, J. (1973) The relative operating characteristic in Psychology. Science. V. 182, 990-1000.Killeen

Swets, J., Tanner, W. & Birdsall, T. (1961) Decision processes in perception. Psychological Review. 68, 301-340.

Materiales un poco más formales y especializados:

Gescheider, G. (1997) Psychophysics: The fundamentals. Lawrence Erlbaum Associates, publishers.

Wickens, T. (2002) Elementary Signal Detection Theory. Oxford University Press.

Ejemplos de aplicación de la TDS en ámbitos diversos: