

Guía 2: Teoría de detección de señales (parte 1)

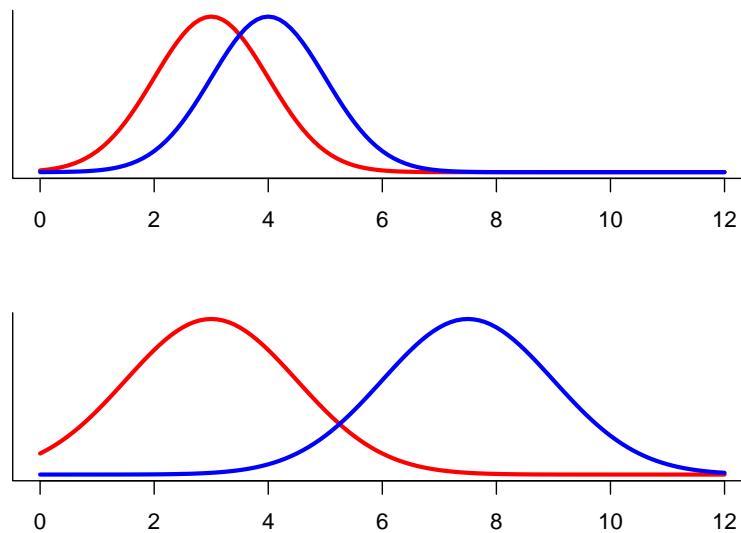
Toma de decisiones 2020

Guillermo Solovey

Ejercicio 1

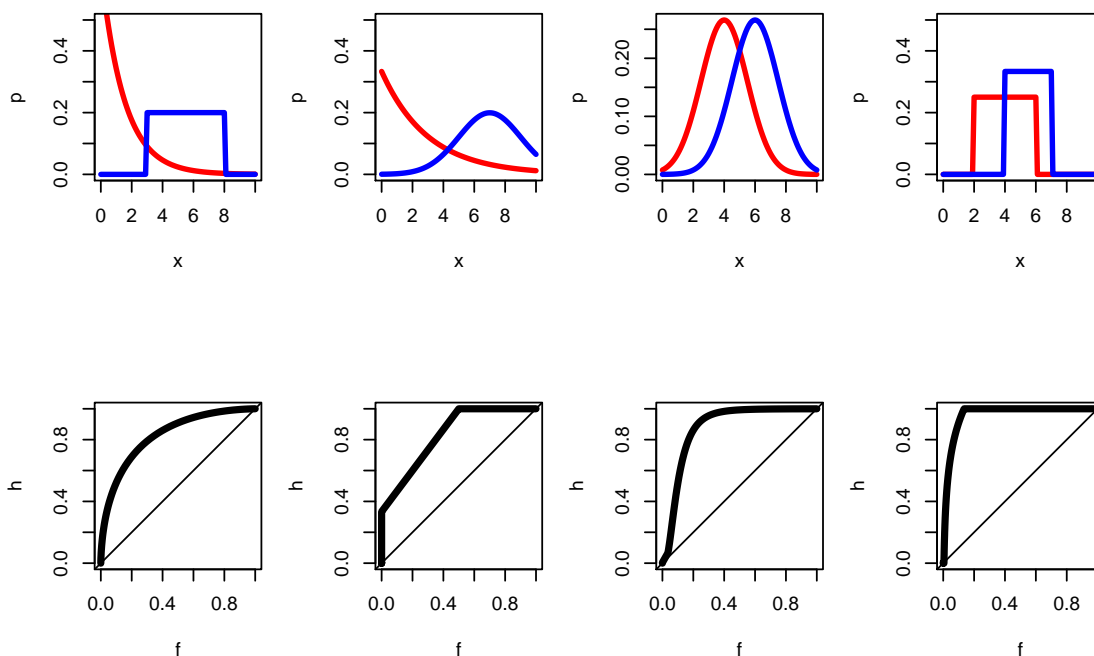
Para cada uno de los siguientes casos, elegir el par de distribuciones que mejor describe la tasa de hits (h) y falsas alarmas (f) y ubicar un criterio de decisión que podría corresponder aproximadamente esos valores.

- (a) $h = 0.84$; $f = 0.5$
- (b) $h = 0.56$; $f = 0.2$
- (c) $h = 0.93$; $f = 0.07$



Ejercicio 2

En la primera fila se muestran las distribuciones de la respuesta interna para cada una de las dos alternativas del estímulo en una tarea de decisión binaria. En la segunda fila se ven las curvas ROC correspondientes. ¿A qué distribuciones corresponden las curvas ROC?



Ejercicio 3

Crear una función llamada `fun_basicsdt` que tenga como parámetros la fracción de hits (h) y de falsas alarmas (f) y devuelva el factor de sensibilidad (o discriminabilidad) d' y el de sesgo en la respuesta, c :

$$d' = z(h) - z(f)$$

$$c = -0.5 (z(h) + z(f))$$

La función `z` está en el paquete 'lestat'. con la siguiente línea cargamos el paquete
`library(lestat)`

Si obtienen un error, es porque el paquete no está instalado. Para eso,
hacen (por única vez)
`install.packages("lestat")`

Ahora sí, pueden crear la función.

```
fun_basicsdt <- function(h,f){
  # La función 'invcdf(normal(), p)' calcula z(p)
  ---
  return(____)
}
```

Ejercicio 4

Escuchar voces en sonidos ruidosas es un fenómeno que los parapsicólogos llaman voz electrónica (o, en inglés, electronic voice phenomena, EVP). El Dr. M diseña un experimento que busca determinar la capacidad de detectar estas voces. Para el experimento, convoca a dos personas una de las cuales es un espiritista que se

considera experto en EVP. Cada participante escucha 100 fragmentos de un audio de 5 segundos. La mitad de estos fragmentos eran ruido aleatorio y la otra mitad tenían una voz muy tenue por sobre el ruido. Luego de escuchar cada uno de estos fragmentos, el participante reportó si escuchó o no una voz. El espiritista detectó una voz en el 92% de los fragmentos que de hecho contenían una voz frente a un 58% del participante escéptico. En cambio, en los fragmentos que no contenían ninguna voz, el espiritista reportó presencia de voz el 48% de las veces y el escéptico reportó haber escuchado una voz sólo en el 9% de los audios. Según los datos recolectados por el Dr. M: ¿Es cierto que el espiritista tiene una capacidad mayor para detectar voces en un contexto de ruido?

Referencia: Este problema viene de “Anderson, N. D. (2015). Teaching signal detection theory with pseudoscience. *Frontiers in psychology*, 6, 762”.

Ejercicio 5

Para investigar un extraño caso de un grupo de 50 turistas perdidos durante una travesía, el departamento de policía convoca a un detective vidente. Durante la consulta, el vidente dijo que 31 de los turistas se encontraban a orillas de un río cercano. Horas más tarde, los policías encontraron a menos de 500 metros de un río a 17 de las personas que el vidente predijo que estarían cerca del río. Sin embargo, 11 de los turistas que el vidente dijo que estarían lejos del río, fueron encontrados también a menos de 500 metros de un río. ¿Tenía el detective una sensibilidad especial para discriminar la ubicación de los turistas perdidos?

ayuda: el vidente puede considerarse un participante en un experimento en el que se presentan N casos que tiene que clasificar entre ‘cerca’ o ‘lejos’.

Referencia: Este problema viene de “Anderson, N. D. (2015). Teaching signal detection theory with pseudoscience. *Frontiers in psychology*, 6, 762”.

Ejercicio 6

Se evalúan dos tests que serían capaces de predecir la aparición de un síntoma psicológico. El primer test se prueba en 200 personas y un año más tarde se encuentra que el test había reconocido correctamente a 25 de las 30 personas que desarrollaron el síntoma. Un estudio similar usando el segundo test fue realizado en un grupo diferente de 150 personas, identificando exitosamente a 15 de las 28 que terminaron desarrollando el síntoma. Explicar por qué estos datos son insuficientes para concluir cuál test es mejor para identificar a las personas que pueden desarrollar el síntoma.

Ejercicio 7

Simular un experimento de detección ¹

En este experimento los participantes deben identificar la presencia de una señal visual ruidosa. En cada realización (*trial*), se le presenta al participante ‘ruido’ o ‘señal+ruido’ con igual probabilidad. En los *trials* correspondientes a ‘señal+ruido’, el participante recibe un input sensorial que se obtiene sorteando un número al azar de una distribución Normal. En los *trials* de ‘ruido’ el input sensorial también es Normal pero con un valor medio menor. El participante fija un criterio de decisión de modo que si la señal sensorial es menor que ese criterio responde “no” y si es mayor responde “si”.

A- Fijar los parámetros del modelo de acuerdo a los siguientes datos. El participante tiene $d' = 1.5$. ¿Qué desviación estándar elegiría para las distribuciones? Establecer un criterio no sesgado justo en la mitad de los valores medios de las distribuciones.

¹Este ejercicio fue propuesto por Megan Peters y Brian Odegaard en su tutorial ‘Computational models of perception, metacognition, and consciousness: a hand-on tutorial’ dictado en el marco de la conferencia ASSC2018.

```

dp          <- ___
criterio    <- ___

sigmaS <- ___      # desviación estándar de la distribución correspondiente a "señal+ruido"
sigmaN <- ___      # desviación estándar de la distribución correspondiente a "ruido"
muS      <- ___
muN      <- ___

```

B- Usando la función `rnorm` generar 500 muestras de la distribución “señal+ruido” y 500 de la distribución “ruido”. Guardar la información de qué tipo de trial es cada muestra.

```

ntrialsS <- ___
ntrialsN <- ___

# La información de qué estímulo corresponde a cada trial la vamos a guardar en la
# variable Label:
# señal+ruido ----> 1
# ruido          ----> 0
# Crear un vector de Labels para cada tipo de estímulo. El largo de cada vector debe ser
# ntrialsS y ntrialsN
signalLabels <- rep(___)
noiseLabels  <- rep(___)

# Tomar las muestras de cada distribución.
signalSamples <- rnorm(___)
noiseSamples  <- rnorm(___)

# Poner todas las muestras juntas, primero las de 'señal+ruido', luego las de 'ruido'
allSamples <- append(___, ___)

# Hacer lo mismo con los vectores de Labels
allLabels <- ____

```

C- Para cada muestra, simular la decisión del participante: ¿había un target presente o no? Si la muestra es mayor que el criterio fijado, la respuesta es “sí”. Caso contrario, es “no”.

```

decision <- ___

```

D- Determinar en cada caso si la respuesta es correcta. Luego calculá el % de respuestas correctas de este participante.

```

isCorrect      <- ___
percentCorrect <- ___

```

E- Calcular la tasa de hits y de falsas alarmas.

```

# definir un vector isHit y otro isFA que evalúa si cada trial es un hit o una
# falsa alarma, respectivamente.
isHit <- ___
isFA  <- ___

# suma los vectores anteriores para tener el número de hits y falsas alarmas
nHits <- ____;
nFAs  <- ____;

# calcula la tasa de hits y de falsas alarmas
h <- ____;

```

```
f <- ___;
```

F- Calcular d' y el sesgo de respuesta que exhibe el participante, c .

Importante: a veces el sesgo de respuesta se llama criterio, pero no hay que confundirlo con el criterio de decisión que definimos antes. En este cálculo estamos determinando cuánto se desvía el criterio del observador respecto del punto medio de las dos distribuciones. Por lo tanto, si el observador es perfectamente no sesgado, “ c ” debería ser 0. Este “ c ” no coincide con el criterio porque el valor del criterio que definimos antes depende de la elección que hayamos hecho para definir las distribuciones y hacer la simulación.

```
dprime_estimado <- ___  
c_estimado      <- ___
```

G- Comparar los valores de d' y c calculados en F con los del modelo que generó los datos. ¿Coincide el d' ? ¿Es realmente no sesgado el observador?

H- ¿Qué pasa si se reduce el número de *trials*? ¿Qué tan estable es la predicción de la simulación (d' y c) de una simulación a la siguiente?