

## RESUMEN

La salud es parte fundamental de la vida de un país, estado o región, es un tema importante en la toma de decisiones de los presidentes, gobernadores y el área de salud.

Un problema que se da en muchos campos del conocimiento, es el de clasificar individuos con base en una o varias variables independientes, este tipo de problemas, hay dos tipos de errores: el que se comete cuando a un individuo no enfermo se le clasifica erróneamente y el segundo, el que se comete cuando a un individuo enfermo se le clasifica como no enfermo estos problemas se le denominan sensibilidad y especificidad

## MARCO TEÓRICO

- Una variable aleatoria Bernoulli que llamaremos  $D = \text{'estado'}$ :

$$D = \begin{cases} 0, & \text{si el individuo no presenta el evento} \\ 1, & \text{si el individuo presenta el evento.} \end{cases}$$

- Una variable aleatoria Bernoulli que llamaremos  $y = \text{'prueba'}$ :

$$Prueba = \begin{cases} y = 1, & \text{si la respuesta es positiva} \\ y = 0, & \text{si la respuesta es negativa.} \end{cases}$$

- Verdaderos positivos ( $V_+$ ): el número de casos con presencia de la condición y predicción correcta.
- Verdaderos negativos ( $V_-$ ): el número de casos con ausencia de la condición y predicción correcta.
- Falsos positivos ( $F_+$ ): el número de casos en los que está ausente la condición y con predicción incorrecta.
- Falsos negativos ( $F_-$ ): el número de casos en los que está presente la condición y con predicción incorrecta.

## REFERENCIAS

- [1] COVID19.SINAVE.GOB.MX. Disponible en: <https://covid19.sinave.gob.mx/>.
- [2] FRANCO N. M. y VIVO M. M., *Análisis de Curvas ROC Principios básicos y aplicaciones*. La Muralla, 2007.

## INTRODUCCIÓN

Las pruebas de diagnóstico médico están diseñadas para discriminar entre diferentes estados de salud, por ejemplo pacientes que están enfermos y los que no lo están (para aclarar conceptos, utilizaremos el término enfermo como una condición específica, independientemente de que la persona "no enferma" pueda tener otro tipo de problemas de salud). Antes de que las pruebas de diagnóstico se implementen en la práctica, es normal que se estudie su precisión o capacidad para discriminar. La precisión de la prueba de resultados binarios (es decir, positiva o negativa), se resume generalmente con la fracción de verdaderos positivos y la fracción de falsos positivos.

## RESULTADO 2

Resultado adicional:

La tasa de verosimilitud refleja el grado de evidencia de una respuesta del clasificador a favor de la presencia de la condición con respecto a la ausencia de la condición.

La tasa de verosimilitud positiva (LRP) dada por el cociente entre la sensibilidad y el complemento de la especificidad.

La tasa de verosimilitud negativa (LRN) es el cociente entre el complemento de la sensibilidad y la especificidad.

$$LRP = \frac{S}{1-E} = \frac{0.5129}{0.1471} = 3.48$$

$$LRN = \frac{1-S}{E} = \frac{0.4870}{0.8529} = 0.5709.$$

En general una tasa de verosimilitud (negativa o positiva) por encima a 10 o por debajo de 0.1 corresponde a clasificadores con una alta capacidad para detectar o descartar la presencia correctamente. Así LRP y LRN no reflejan ninguna predilección por la presencia de Covid-19 y por tanto, no descartan la enfermedad con confianza.

## FUTURA INVESTIGACIÓN

Como futuras investigaciones podemos hacer predicciones acerca de los pacientes con dicha enfermedad y explicar con detalle por que se cometen muchos errores al clasificar los casos.

## RESULTADO 1

Los datos están en la página [1] donde se obtiene de la semana 10 a la 30 del estado de Puebla una cifra alrededor de 20278 casos positivos, 19254 negativos y 3319 casos sospechosos de Covid-19 la pregunta es ¿cuál es la capacidad de una prueba para discriminar correctamente entre la presencia y ausencia de Covid-19?

Por lo cual optamos por calcular la sensibilidad es la probabilidad que, si tenemos un individuo enfermo, que la prueba lo clasifique como enfermo.

$$S = \frac{V_+}{V_+ + F_-} = \frac{20278}{39532} = 0.5129.$$

Por otro lado calculamos la especificidad es la

probabilidad que, si un individuo sano, que la prueba lo clasifique como sano.

$$E = \frac{V_-}{V_- + F_+} = \frac{19259}{22573} = 0.8529.$$

Por tanto, la prueba que se realiza detecta correctamente Covid-19 el 51.29% de las veces aunque el 14.71% de las veces provoca erróneamente una segunda prueba.

Dadas las graves consecuencias de la presencia de Covid-19 si no se detecta a tiempo, es deseable que la prueba diagnóstica empleada posea la mínima fracción de falsos negativos.

## CONCLUSIÓN

Como se menciono las tasas de verosimilitud toman un papel importante y es por ello que las tasas de verosimilitud LRP y LRN son utilizadas para comparar dos clasificadores en el siguiente sentido.

- Dados dos clasificadores X y Y, se dice que Y es mejor que X para confirmar la presencia de la condición, si las tasas de verosimilitud positivas verifican la siguiente relación  $LRP_Y > LRP_X$ .

- Dados dos clasificadores X y Y, se dice que Y es mejor que X para confirmar la ausencia de la condición, si las tasas de verosimilitud negativas verifican la siguiente relación  $LRN_Y > LRN_X$ .

Asimismo, se propone realizar la comparación de clasificadores a través de un gráfico en términos de sus tasas de verosimilitud.

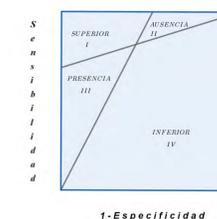


Figura 1: Regiones de comparación mediante tasas de verosimilitud.

Dicho de otra forma se puede expresar como en la Tabla 1.

TASA DE VEROSIMILITUD	COMPARACIÓN
$LRP_Y > LRP_X, LRN_Y < LRN_X$	Y, es S. (I)
$LRP_Y < LRP_X, LRN_Y < LRN_X$	Y, es S.C.A.C. (II)
$LRP_Y > LRP_X, LRN_Y > LRN_X$	Y, es S.C.P.C. (III)
$LRP_Y < LRP_X, LRN_Y > LRN_X$	Y, es I. (IV)

Tabla 1: Comparación de dos clasificadores.

De este modo, la comparación de los clasificadores resulta importante para ámbitos médicos.

## CONTACTO

Email [oscarmendezcu@gmail.com](mailto:oscarmendezcu@gmail.com)  
Phone 2227293339

Otro punto importante es desarrollar nuevas metodologías de estimación ya que son necesarias para el avance estadístico en la medicina.