

Análisis de la curva de Roc en la decisión de diagnóstico clínico

Alejandra Ossa Yepes
Jóse Andres Carvajal Bautista

Resumen—The Roc curve is a statistical tool, which allows us to interpret data, in order to make the most convenient decision; In this article we will generate a view of this tool oriented to one of the most used branches, which is the clinical diagnostic analysis, in which a study of the behavior of a group of chosen data will be carried out in order to show the efficacy and good functioning of this statistical method.

I. INTRODUCCIÓN

La curva de Roc es una alternativa estadística que se implementa para al análisis de decisión con respecto a algún problema de este modo se espera relacionar y comprender el tema elegido con los soportes teóricos que se han obtenidos en un curso anterior, con lo que permita una mejor interpretación de los temas e información que se está mostrando. En el presente artículo, se realizó un análisis del funcionamiento de la Curva de Roc y su influencia en el diagnóstico clínico, en el cual se presenta una breve descripción de un caso médico y su relación con la metodología que ofrece para el análisis de decisiones. *La curva ROC es una herramienta estadística utilizada en el análisis de clasificar la capacidad discriminante de una prueba diagnóstica dicotómica. Es decir, una prueba, basada en una variable de decisión, cuyo objetivo es clasificar a los individuos de una población en dos grupos: uno que presente un evento de interés y otro que no.*

Como foco de interés del artículo, nos basaremos en el fundamento teórico que tiene esta técnica y su versatilidad en las diferentes aplicaciones, para aportar así una muestra de resultados de cómo es su funcionamiento y su desarrollo. Los datos corresponden a un experimento médico en el que

50 pacientes de los cuales 20 están enfermos, son sometidos a una prueba de detección mediante la cual se mide la concentración de una molécula viral.

Fueron ingenieros militares estadounidenses, durante la Segunda Guerra Mundial, los que empezaron a utilizar la curva Roc. Las usaron como herramienta para discernir en sus radares qué era una señal de ruido y qué era una señal de ofensiva militar de otro país (torpedos, misiles y similares).

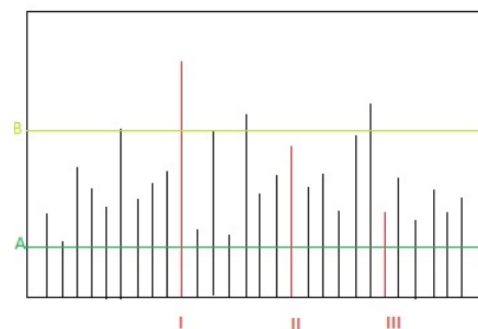


Figura 1. Ejemplo de receptor de señal en radar.[5]

La figura 1 representa la pantalla de un receptor de señal radar, las señales I,II y III corresponden a misiles, sin embargo el resto no y, por tanto, es ruido. Ante una nueva lectura deberíamos fijar un valor umbral a partir de cual considerar que esa señal corresponderá a un misil y por debajo no. Dicho valor umbral, según se escoja, nos puede dar muchos avisos falsos de misiles (falsos positivos) o, por el contrario, puede no avisar ante señales que correspondían a misiles (falsos negativos). [5]

II. DEFINICIÓN TEÓRICA DE LA CURVA DE ROC

Se tiene una medida determinada X (recordar que esta medida puede resultar de realizar una regresión logística sobre varias variables de distintas naturalezas) realizada sobre una población de positivos X_p , y otra de negativos X_n , con función de distribución G y F , respectivamente. Suponiendo que $E(X_n) \leq E(X_p)$, para clasificar a los individuos en uno u otro grupo, se debe fijar un criterio, punto de corte, a partir del cual un individuo será considerado positivo. Por tanto, fijado un punto de corte t , la sensibilidad de la prueba vendrá determinada por $1-G(t)$, siendo $F(t)$ su especificidad y quedando por tanto determinada la curva ROC por las coordenadas del vector $(1-F(t), 1-G(t))$, o, equivalentemente, por la función $1 - G(F^{-1}(1-t))$, $t \in [0, 1]$.

Como siempre, el problema surge cuando se desconocen las distribuciones de la variable en las poblaciones de positivos y negativos y deben estimarse a partir de muestras aleatorias. Una de las posibilidades es suponer que las poblaciones siguen algún modelo paramétrico, el gaussiano usualmente, o bien, aplicar algún método no paramétrico, siendo los más frecuentes sustituir las funciones de distribución desconocidas por sus Funciones de Distribución Empíricas (FDE) o por las Funciones de Distribución Empírica Suavizadas (FDES).

- El uso de la distribución Binomial para el comportamiento de la gráfica genera más comodidad a la hora de clasificar las variables operativas y permiten el análisis de datos de esta distribución, además el modelo Binomial es muy fácil de manejar algebraicamente.

II-A. MODELO BINOMIAL

El modelo binormal es introducido inicialmente por Dorfman & Alf (1969) y tratado más recientemente por Hsieh & Turnbull (1996). En él se supone que existe una función monótona H de modo que transforme a las dos poblaciones de origen en normales, esto es, tanto $H(X_n)$ como $H(X_p)$ siguen una a distribución normal con medias μ_1 y μ_2 y varianzas σ_1^2 y σ_2^2 , respectivamente. Así las cosas,

si se denota por Φ la función de distribución normal estándar, la curva ROC tiene la expresión. [1]

$$\begin{aligned} ROC(t) &= 1 - G(F^{-1}(1-t)) \\ &= 1 - G(H^{-1}F^{-1}(1-t)) \\ &= 1 - \Phi\left(\frac{\mu_1 + \sigma_1\Phi^{-1}(1-t) - \mu_2}{\sigma_2}\right) \end{aligned}$$

III. LA CURVA DE ROC

Sea la variable aleatoria $D = [\text{estado}]$ que sigue una distribución Bernoulli de parámetro p que llamaremos prevalencia del evento sobre la población. Dicha variable toma los valores: [5]

$$D \sim Be(p) D = \begin{cases} 0 & \text{cuando el individuo no presenta el evento} \\ 1 & \text{cuando el individuo presenta el evento} \end{cases}$$

Por tanto,

$$\begin{aligned} \text{prevalencia} &= p \\ &= Pr(\text{"estado"} = 1) \\ &= Pr(D = 1) \\ &= Pr(\text{presentar el evento}) \end{aligned}$$

y puesto que ambos estados forman un espacio de sucesos:

$$\begin{aligned} 1 - p &= Pr(\text{"estado"} = 0) \\ &= Pr(D = 0) \\ &= Pr(\text{no presentar el evento}) \end{aligned}$$

IV. CARACTERÍSTICAS OPERATIVAS

Yerushalmy (1947) introduce los términos de sensibilidad y especificidad como indicadores estadísticos que evalúan el grado de eficacia inherente a una prueba diagnóstica (resumen de la precisión diagnóstica).

- Se establece la siguiente relación de nomenclaturas para la comprensión del texto

Resultado de la prueba en estudio	Estado respecto a la enfermedad según el estándar de referencia	
	Enfermo	Sano
Positivo	Verdadero positivo (VP)	Falso positivo (FP)
Negativo	Falso negativo (FN)	Verdadero negativo (VN)

Figura 2. La generación de los conceptos de VP, VN, FP y FN. [4]

IV-A. SENSIBILIDAD

La *sensibilidad* es un parámetro que se mide en el grupo de sujetos que verdaderamente están enfermos. Es el cociente entre verdaderos positivos y el total de las personas enfermas. Por tanto, es la probabilidad de obtener un resultado positivo cuando el individuo tiene la enfermedad, o la proporción de verdaderos positivos. Según Pepe (2003) el termino proporción puede causar confusión en epidemiología resultando más apropiado sustituirlo por fracción. Se denota entonces la sensibilidad como $T P F$ (*True Positive Fraction*). La sensibilidad es especialmente importante cuando una enfermedad no debe pasar desapercibida y cuando el pronóstico mejora mucho con el tratamiento precoz.

IV-B. ESPECIFICIDAD

La *especificidad* es un parámetro que se mide en el grupo de sujetos no enfermos. Es el cociente entre verdaderos negativos y el total de no enfermos. Por tanto, es la probabilidad de obtener un resultado negativo cuando el individuo no tiene la enfermedad. Se denota la especificidad como $T N F$ (*True Negative Fraction*). La especificidad es especialmente importante cuando la enfermedad cambia la vida del enfermo o tiene cierto estigma (no se quiere preocupar inútilmente o estigmatizar erróneamente a alguien con un falso positivo) y también cuando las consecuencias de un tratamiento suponen un riesgo para el enfermo (amputación, etc.) [2]

Sensibilidad	$= VP / (VP + FN) = FVP$ (fracción de verdaderos positivos)
Especificidad	$= VN / (VN + FP) = FVN$ (fracción de verdaderos negativos) $= 1 - FFP$ (fracción de falsos positivos)

Figura 3. Resultado de una prueba y su estado respecto a la enfermedad. [3]

Resultado de la prueba en estudio	Estado respecto a la enfermedad según el estándar de referencia	
	Presente	Ausente
Positivo	a (enfermos con prueba +)	b (no enfermos con prueba +)
Negativo	c (enfermos con prueba -)	d (no enfermos con prueba -)

Figura 4. La generación posterior de las celdas con las que se generan los cálculos de sensibilidad, especificidad y valores predictivos. [4]

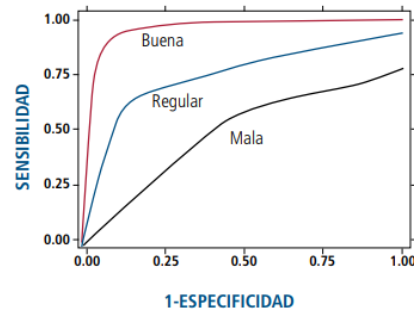


Figura 5. Esquema explicativo de distintas posibilidades de curva de Roc. [4]

IV-C. VALOR PREDICTIVO POSITIVO

El valor predictivo positivo (VPP) de una prueba diagnóstica (PD) corresponde a la proporción de individuos con una prueba positiva para una enfermedad o evento de interés determinado, que están realmente enfermos de ella. El VPP de una prueba se puede explicar con el siguiente escenario: si el resultado de una PD es positivo ¿qué probabilidad tiene el paciente de presentar la enfermedad en estudio?.

Dicho de otra manera, el VPP de una PD se define por la proporción de resultados positivos de la prueba diagnóstica en quienes la enfermedad está presente (confirmada por el estándar de referencia); o, por la probabilidad de que un paciente sea un VP teniendo el resultado de la PD positiva; o, por la probabilidad de tener la enfermedad dado que el resultado de la PD fue positiva.

De esto se desprende que para calcular el VPP de una PD se ha de dividir el número enfermos con prueba positiva por la sumatoria de los enfermos con prueba positiva y los sujetos “no enfermos” con prueba positiva; es decir $a / (a + b)$; ó $VP / VP + FP$.

IV-D. VALOR PREDICTIVO NEGATIVO

Por su parte, el valor predictivo negativo (VPN) de una prueba diagnóstica corresponde a la proporción de individuos con una prueba negativa para una determinada enfermedad o evento de interés en estudio, que no tienen la enfermedad o evento de interés. Así, este concepto se podría explicar con el siguiente escenario: si el resultado de una prueba es negativo, ¿cuál es la probabilidad que tiene el

paciente de no presentar la enfermedad en estudio? Dicho de otra manera, el VPN de una PD se define por la proporción de personas con resultados negativos de la PD en quienes la enfermedad está ausente; por la probabilidad de que el paciente sea un VN teniendo una PD negativa; y por la probabilidad de que la enfermedad esté ausente dado un resultado negativo de la PD.

De lo anteriormente expuesto, se puede colegir que para calcular el VPN de una PD se ha de dividir el número enfermos con prueba negativa por la sumatoria de los enfermos con prueba negativa y los sujetos “no enfermos” con prueba negativa; es decir $d / (c + d)$; ó $VN / FN + VN$. [4]

Sensibilidad = $\frac{a}{a+c}$	Especificidad = $\frac{d}{b+d}$
VPP = $\frac{a}{a+b}$	VPN = $\frac{d}{c+d}$

Figura 6. Formula para la realización de los calculos de sensibilidad, especificidad, valores predictivos y razones de probabilidad. [4]

La Curva ROC permite describir que tan separadas están las distribuciones de la sensibilidad y la especificidad de un test, por ejemplo; si las funciones de densidad se cruzan la curva ROC resultante será similar a la mostrada en Figura 7, si las distribuciones se solapan casi por completo, entonces la curva ROC está representada por la línea diagonal como se muestra en la Figura 8 indicando que el test no es informativo. Si las funciones de densidad están simétricamente distantes como se muestra en la Figura 9 entonces la curva ROC será las líneas que forman la esquina superior izquierda, indicando un test perfecto. [2]

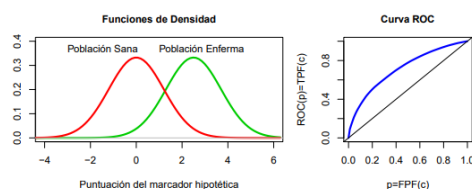


Figura 7. Ejemplo de funciones de densidad de la población sana y enferma que generan una curva ROC. [2]

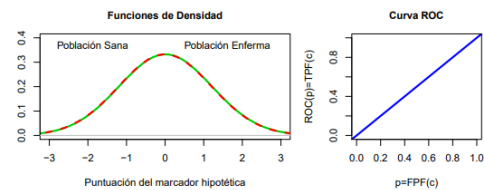


Figura 8. Ejemplo de funciones de densidad de la población sana y enferma que generan una curva ROC para un test no informativo.[2]

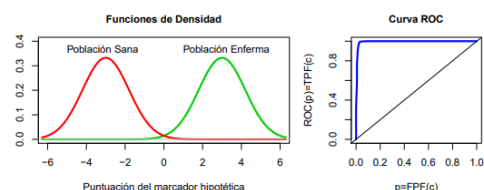


Figura 9. Ejemplo de funciones de densidad de la población sana y enferma que generan una curva ROC para un test casi perfecto.[2]

V. ÁREA BAJO LA CURVA ROC

El Área Bajo la Curva ROC (*Area Under the Curve, AUC*) estima la capacidad de distinguir o de “discriminar” entre enfermos y no enfermos que tiene una prueba diagnóstica. Se define como:

$$AUC = \int_0^1 ROC(p) dp$$

Si el área bajo la curva valiese 1 (100 %, la curva ROC se dirige hacia la esquina superior izquierda como en la Figura 6) la prueba sería perfecta, ya que clasificaría al 100 % de los enfermos como enfermos y al 100 % de los exentos de enfermedad como sanos. En cambio, si el área bajo la curva valiese 0,5 (50 %, área debajo de la línea diagonal representada en la Figura 5) existiría la misma probabilidad de clasificar a un enfermo como sano que como enfermo. Un área de 0,5 bajo la curva equivale a no discriminar, se interpreta como una prueba “no informativa”. Esta curva ROC sería, de hecho, aquella que tendría valores de sensibilidad iguales a 1-especificidad en todos los posibles puntos de corte. En estas circunstancias se dice que la prueba es “inútil” ya que no reduce nada el grado de incertidumbre previo acerca de si el paciente tiene o no la enfermedad. La diagonal del gráfico que expresa lo que sería una prueba

inútil se suele representar junto al área bajo la curva ROC. Al representar la línea diagonal se ve en cuánto supera la prueba que se valora a lo que sería una prueba no informativa. Un área menor de 0,5 requiere invertir los criterios de diagnósticos: es decir, considerar los negativos como positivos y viceversa. Al invertirlos se comprueba que la prueba discrimina más del 50 %. La construcción de curvas ROC permite además la comparación de varias pruebas diagnósticas. El área bajo la curva por ser una mitad global implica pérdida de información, por lo que no debe considerarse aisladamente sin examinar la curva. A veces, curvas ROCs de trazados muy distintos pueden tener áreas semejantes (Borgueño et al. (1995)).

Entonces solo es posible afirmar que un si una curva ROC es uniformemente mejor que otra tendrá mayor área bajo ella y el test correspondiente será la prueba de elección, debido a su superior capacidad de discriminación. El área bajo la curva ROC tiene una interpretación interesante. Según Bamber (1975) y Hanley and McNeil (1982), se define como la probabilidad de clasificar correctamente un par de individuos sano y enfermo, seleccionados al azar de la población, mediante los resultados obtenidos al aplicar la prueba diagnóstica. [2]

VI. CRITERIOS A CONSIDERAR EN LA VALORACIÓN DE UNA PD

Características de la población: La sensibilidad o especificidad de una prueba dependen de las características de la población estudiada.

Si se altera o cambia la población en estudio, cambiarán también estos índices. Los datos informados de sensibilidad y especificidad, que son evaluados en poblaciones con una tasa significativa de enfermedad, pueden no ser aplicables en otras poblaciones diferentes en las que se utilice la prueba. Para que este criterio se cumpla, el artículo debe contener información sobre los siguientes aspectos: género y edad de los sujetos en evaluación, resumen de los síntomas clínicos iniciales o estadio de la enfermedad, y criterios de elección para los sujetos que son enrolados en el estudio.

Subgrupos adecuados: La sensibilidad y la

especificidad pueden representar valores promedios para una población determinada. A menos que el problema para el cual se utiliza la prueba haya sido definido con mucha precisión, aquellas pueden variar en diferentes subgrupos poblacionales. Para que la prueba pueda ser utilizada con éxito deberían tenerse en cuenta distintos niveles de precisión según los distintos subgrupos existentes en la población estudiada. Este criterio se cumple cuando se informa sobre la precisión de la prueba en relación con cualquier subgrupo demográfico o clínico (por ejemplo en sujetos sintomáticos y sujetos asintomáticos).

Sesgo de selección: Puede producirse cuando los sujetos con los resultados positivos o negativos de una prueba son derivados de forma preferente para verificar el diagnóstico mediante otra prueba considerada el estándar de referencia. Para que este criterio se cumpla, todos los sujetos deberían de haber sido asignados para recibir tanto la prueba diagnóstica en estudio como el estándar de referencia a través de un procedimiento directo o mediante el seguimiento clínico.

Sesgo de medición: Podría introducirse si la PD o el estándar de referencia se realizan sin tomar precauciones para garantizar la objetividad de su interpretación (similar al enmascaramiento utilizado en los ensayos clínicos para tratamiento). Se puede obviar si la PD en evaluación y el estándar de referencia son interpretadas de forma separada y enmascarada por personas independientes que desconocen los resultados de una y otra.

Precisión de los resultados: La precisión de la sensibilidad y la especificidad depende del número de pacientes evaluados. Igual que otras medidas, el resultado estimado debe tener los intervalos de confianza o el error estándar reportados independientemente de la magnitud encontrada.

Presentación de resultados indeterminados: No todas las PD dan lugar a un sí o un no como respuesta, a veces dan lugar a resultados equívocos o indeterminados. La frecuencia de resultados indeterminados limitará la aplicabilidad de la prueba o la hará más cara si da lugar a

otros procedimientos diagnósticos posteriores. La frecuencia de resultados indefinidos y el modo en el que se usan en el cálculo de la precisión de la prueba constituyen una información de importancia crítica para conocer la eficacia de la misma. Para que este criterio se cumpla el trabajo debe reflejar de forma apropiada todos los resultados positivos, negativos o indeterminados generados durante el estudio, así como si los resultados indeterminados se incluyeron o excluyeron al calcular los indicadores de precisión de la prueba.

Reproducibilidad de la prueba: Las pruebas no siempre dan el mismo resultado, por motivos relacionados con la variabilidad de éstas o de la interpretación del observador. Los motivos y el impacto de este asunto deben ser tenidos en cuenta. Para que se cumpla este criterio en pruebas que requieren interpretación del observador, al menos alguna de las pruebas debería ser evaluada con alguna medida que resuma la variabilidad interobservador. Para pruebas sin interpretación del observador, el criterio se cumple cuando se refleja una media que resuma la variabilidad del instrumento. [4]

VII. APORTES

Se analizó los datos de un experimento médico con 50 pacientes, de los cuales 20 están enfermos, y estos son sometidos a una prueba de detección, en la cual se mide la concentración de una molécula viral. En la siguiente tabla se muestra la cantidad de pacientes enfermos los cuales se denotan con + y los que no con -, al frente de cada paciente se encuentra el resultado de la prueba. [6]

enfermo	concentración	enfermo	concentración
-	0.55	-	0.53
-	0.54	-	0.56
-	0.59	-	0.59
-	0.65	-	0.66
-	0.69	-	0.69
-	0.72	-	0.72
-	0.73	-	0.73
-	0.78	-	0.81
-	0.82	-	0.82
-	0.90	-	0.94
-	0.95	-	0.96
-	1.14	-	1.16
-	1.18	-	1.20
-	1.23	-	1.25
-	1.25	-	1.30

enfermo	concentración	enfermo	concentración
+	0.90	+	0.98
+	1.04	+	1.05
+	1.12	+	1.12
+	1.30	+	1.31
+	1.35	+	1.36
+	1.37	+	1.38
+	1.40	+	1.40
+	1.42	+	1.43
+	1.46	+	1.46
+	1.48	+	1.50

Cuadro I
DATOS DE LOS PACIENTES

En la siguiente tabla se muestra los intervalos de confianza con respecto a la sensibilidad y a la especificidad para cada valor estimado respectivamente, la primera columna representa el recorrido de valores de sensibilidad o especificidad, va desde cero hasta uno, la segunda y la cuarta corresponde a los extremos de los intervalos y la tercera columna es una media de las columnas anteriores, en estos intervalos se puede ver que para valores altos de sensibilidad se obtiene valores bajos de especificidad y viceversa.

```

===== 100%
95% CI (2000 stratified bootstrap replicates):
se sp.low sp.median sp.high
0.0 1.0000 1.0000 1.0000
0.1 1.0000 1.0000 1.0000
0.2 1.0000 1.0000 1.0000
0.3 1.0000 1.0000 1.0000
0.4 1.0000 1.0000 1.0000
0.5 0.9667 1.0000 1.0000
0.6 0.6333 1.0000 1.0000
0.7 0.6000 0.8333 1.0000
0.8 0.5667 0.7667 1.0000
0.9 0.5333 0.7333 0.9667
1.0 0.4333 0.6333 0.8667
> ci.se(ROC2)
===== 100%
95% CI (2000 stratified bootstrap replicates):
sp se.low se.median se.high
0.0 1.0000 1.00 1.00
0.1 1.0000 1.00 1.00
0.2 1.0000 1.00 1.00
0.3 1.0000 1.00 1.00
0.4 1.0000 1.00 1.00
0.5 0.9331 1.00 1.00
0.6 0.6500 1.00 1.00
0.7 0.5500 0.90 1.00
0.8 0.5000 0.75 1.00
0.9 0.5000 0.70 0.95
1.0 0.4500 0.65 0.90

```

Figura 10. Intervalos de Confianza

- **Ci.sp:** Intervalo de confianza de especificidad
- **Ci.se:** Intervalo de confianza de sensibilidad

Mediante un análisis que se realizó en R, el área bajo una curva es 0.9133 este resultado es bueno ya que es muy próximo a 1, además también en la grafica se puede observar que la curva de ROC esta mucho mas arriba que la recta $y=x$, lo cual representa que la prueba que se hizo fue buena.

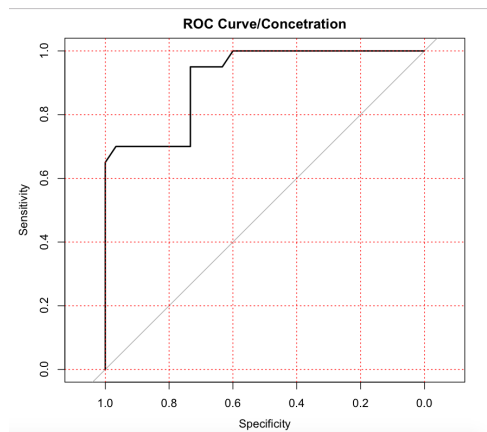


Figura 11. Curva de Roc

- **Área bajo la curva:** 0.9133
- **Varianza:** 0.001523659
- **Intervalo de confianza:** [0.8368,0.9898]
- **Máximos Locales de la sensibilidad**(1.0,0.95),(0.7,0.65)
- **Máximos Locales de la especificidad**(0.60,0.733),(0.9667,1.0)

En la siguiente figura se muestra la curva de ROC suavizada, la cual tiene un área de 0.8837, claramente el área es menor que la curva original ya que cuando se suaviza se eliminan detalles innecesarios y bucles, esto hace que el área que comprende la curva sea menor.

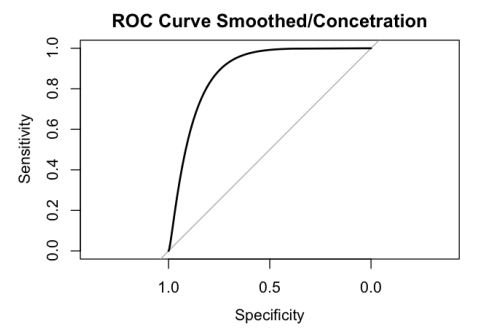


Figura 12. Curva de Roc suavizada

- **Área bajo la curva:** 0.8837

Resultados: Con los anteriores datos se puede concluir que el diagnostico es bueno, se tiene un buen rendimiento de la curva, el intervalo de sensibilidad es bueno ya que esta muy alejado de 0.5 y muy próximo a 1. en los datos anteriores se puede ver cual es

el máximo de la sensibilidad y de la especificidad, la varianza es muy pequeña luego los datos obtenidos no son muy dispersos.

VIII. CONCLUSIÓN

A la hora de realizar una acción ante un problema es fundamental asegurar su veracidad, tener una idea clara de que tan factible es esta solución y conveniente para el caso que se presenta, en donde se evidencia el claro desempeño que tiene este modelo estadístico para la comparación de factores en este caso con un propósito clínico; pero es fácil imaginar las numerosas formas de aplicaciones de esta herramienta para asegurar un buen análisis de decisión entorno a lo que se quiere.

La curva Roc tiene una ventaja muy grande en su uso, debido a que posee características muy básicas y de fácil manejo, con lo cual permite que su manipulación sea muy rápida y se pueda evidenciar su resultado con mayor eficacia, además hay muchas herramientas estadísticas, que poseen esta característica y permiten su análisis fácilmente..

REFERENCIAS

- [1] Title = Comparación de pruebas diagnósticas desde la curva ROC, Author = Pablo Martínez-Cambor, Year= 2010, Note= <http://www.kurims.kyoto-u.ac.jp/EMIS/journals/RCE/V30/v30-2body/v30n2a01Martinez.pdf>
- [2] Title = Curvas ROC para Datos de Supervivencia. Aplicación a Datos Biomédicos, Author = Anaderli Torres Ortiz, Year= 2010, Note= http://eio.usc.es/pub/mte/descargas/ProyectosFinMaster/Proyecto_407.pdf
- [3] Title = Curvas ROC, Author = Anaderli Torres Ortiz, Year= 2010, Note= https://www.fisterra.com/mbe/investiga/curvas_roc/curvas_roc2.pdf
- [4] Title = Como interpretar un artículo sobre pruebas diagnósticas, author = DR. CARLOS MANTEROLA D., year= 2009, Note= http://paginas.facmed.unam.mx/deptos/sp/wp-content/uploads/2015/11/U10_compl_manterolac_epiclin.pdf
- [5] Title = Curvas ROC (Receiver-Operating-Characteristic), Author = Ana Rocío del Valle Benavides, Year= S.F, Note= <https://idus.us.es/xmlui/bitstream/handle/11441/63201/Valle%20Benavides%20Ana%20Roc%C3%ADo%20del%20TFG.pdf?sequence=1>
- [6] Title= Curvas ROC: tutorial en Excel Author= Xlstat Year= 2017 Note= <https://help.xlstat.com/customer/es/portal/articles/2062277-curvas-roc-tutorial-en-excel>