



CIMAT

Centro de Investigación en Matemáticas, A.C.

---

# UN ALGORITMO DE CLASIFICACIÓN COMO HERRAMIENTA DE DIAGNÓSTICO DE DIABETES CON DATOS DE LA PRUEBA ORAL DE TOLERANCIA A LA GLUCOSA

**T E S I S**

Que para obtener el grado de

**Maestro en Ciencias**

con Especialidad en

**Computación y Matemáticas Industriales**

Presenta

Paola Vargas Bernal

Directores de Tesis:

Dr. Miguel Ángel Moreles Vázquez

Dr. Joaquín Peña Acevedo

---

Autorización de la versión final



# Agradecimientos

Quiero agradecer:

- A Dios por hacer lo imposible, posible.
- A mis asesores el Dr. Miguel Ángel Moreles Vázquez y al Dr. Joaquín Peña Acevedo por permitirme trabajar con ustedes, por lo que me aportaron y me enseñaron para realizar este trabajo de investigación, por su apoyo, dedicación y buen consejo.
- A mis sinodales Dr. Arturo Hernández Aguirre y Dra. Adriana Monroy Guzmán por evaluar mi tesis y tomarse el tiempo para escucharme.
- Al Centro de Investigación en Matemáticas, A.C. (CIMAT) por darme la oportunidad de realizar mis estudios de maestría, por hacerme parte de la institución y su apoyo incondicional. Además, agradezco a mis profesores, compañeros, personal administrativo por su ayuda y apoyo en mi proceso académico.
- Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por su financiamiento en la realización de mis estudios y de mi estancia en México.
- A mis padres Alfonso Vargas y Cecilia Bernal, y a mi hermana Fernanda Vargas por apoyarme, escucharme y aconsejarme durante esta experiencia, gracias por su amor y entrega incondicional.
- Y a mis amigos por acogerme y echarme la mano.



# Índice general

Lista de figuras	7
<b>1. Introducción</b>	<b>9</b>
<b>2. Modelación de la interacción glucosa/insulina durante la prueba oral de tolerancia a la glucosa</b>	<b>13</b>
2.1. Sobre la enfermedad . . . . .	13
2.2. Modelos de interacción glucosa/insulina . . . . .	15
2.2.1. El modelo básico de interacción glucosa/insulina . . . . .	16
2.3. Otros Modelos . . . . .	20
<b>3. Ajuste del modelo básico de interacción glucosa/insulina a datos reales de OGTT</b>	<b>23</b>
3.1. Análisis de datos . . . . .	23
3.2. Preliminares sobre estimación bayesiana . . . . .	25
3.2.1. Metodología de estimación . . . . .	25
3.2.2. Estimadores puntuales . . . . .	29
3.3. Aplicación de la metodología a los datos de OGTT . . . . .	30
3.3.1. Función de Verosimilitud . . . . .	30
3.3.2. Distribuciones a priori . . . . .	32
3.3.3. MCMC . . . . .	33
3.4. Resultados del ajuste . . . . .	34
<b>4. Un algoritmo de clasificación para diagnóstico de diabetes</b>	<b>41</b>
4.1. Máquina de soporte vectorial . . . . .	41
4.1.1. Clasificación usando un hiperplano separador . . . . .	41
4.1.2. Clasificador de máximo margen . . . . .	43
4.1.3. Clasificador de soporte vectorial . . . . .	46

4.2. Exploración de parámetros clasificadores . . . . .	48
4.3. Propuesta de parámetros clasificadores . . . . .	49
<b>5. Conclusiones y trabajo futuro</b>	<b>53</b>
<b>A. Datos</b>	<b>55</b>
<b>Bibliografía</b>	<b>61</b>

# Índice de figuras

3.1. Concentración de glucosa en un paciente de cada clase. . . . .	24
3.2. Dispersión de los datos en cada una de las clases. . . . .	26
3.3. Histogramas de las distribuciones marginales para cada parámetro de los datos que se presentan en la figura 3.4(a)-(b) . . . . .	35
3.4. Curvas obtenidas de pacientes sanos . . . . .	37
3.5. Curvas obtenidas de pacientes enfermos . . . . .	38
3.6. Histogramas del error relativo medio en el caso 1 . . . . .	39
3.7. Histogramas del error relativo medio en los tres casos . . . . .	39
4.1. Parámetros obtenidos en los 80 pacientes . . . . .	49
4.2. Parámetro contra parámetro de los 80 datos. . . . .	50
4.3. A contra $\alpha$ de los 80 datos. . . . .	51
4.4. Clasificación de personas sanas y personas enfermas o con alguna alteración . . . . .	52
5.1. Ajustes de pacientes por mejorar. . . . .	54





# Capítulo 1

## Introducción

Según la Organización Mundial de la salud (OMS) [10], la diabetes es una enfermedad crónica que ha aumentado significativamente en el número de personas que la padecen, de pasar de 108 millones de personas en 1980 a 422 millones en 2014; proliferándose con mayor rapidez en los países de ingresos bajos y medios. Aunque no se pueda determinar con exactitud las causas de este incremento, si se ve relacionado con el aumento del número de personas con obesidad y sobrepeso, la inactividad física y el consumo de tabaco.

Los distintos tipos de diabetes pueden causar complicaciones a la salud, tales como ceguera, amputación de los miembros inferiores, insuficiencia renal, accidente cerebrovascular, infarto de miocardio, e incrementar el riesgo de muerte prematura. Se estima que esta enfermedad causó 1.6 millones de muertes en el 2015 y los altos niveles de glucemia causaron 2.2 millones de muertes en el 2012. La OMS proyecta que en el 2030 la diabetes será la séptima causa de mortalidad en el mundo. No obstante, sus consecuencias se pueden tratar, evitar o retrasar manteniendo una dieta saludable, realizando ejercicio, evitando el consumo de tabaco y haciéndose exámenes periódicos para detectar esta enfermedad.

Adicionalmente, las personas que sufren esta enfermedad y sus familias tienen repercusiones económicas por los costos médicos y la pérdida de sus trabajos. Así mismo, para los sistemas de salud y las economías de los países por los costos de la atención hospitalaria y los tratamientos necesarios para esta enfermedad.

Dado a este contexto mundial sobre la diabetes y las consecuencias que tiene para la salud de la población, la investigación es necesaria en todas las disciplinas.

Desde el campo de la modelación numérica, se pueden proponer modelos que nos permitan profundizar en la comprensión de la enfermedad explorando sus propiedades con herramientas numéricas de solución.

Desafortunadamente, estos modelos involucran parámetros los cuales no son conocidos, o difíciles de obtener. Una manera de proceder, es formular la estimación de parámetros como un problema de optimización. Se postula una función objetivo y se proponen métodos numéricos de solución.

La modelación y estimación de parámetros en modelos de diabetes es un campo activo de estudio. Por ejemplo, en [2] se hace una revisión de la modelación en donde se describe la cinética del sistema glucosa/insulina usando pruebas de tolerancia a la glucosa intravenosa y oral. En particular el caso que nos concierne la interacción glucosa/insulina usando la prueba oral de tolerancia a la glucosa (OGTT). Se presenta también un análisis de los parámetros y su uso en el diagnóstico de la enfermedad. En [3] se presenta un modelo minimal con interacción de glucosa/insulina/glucagón usando la OGTT. En los dos modelos anteriores se utiliza la estimación bayesiana para encontrar los parámetros desconocidos.

En este trabajo nos interesa no solo la modelación y estimación de parámetros, sino el uso de estos últimos como herramienta de diagnóstico. Se suponen conocidos los datos de la prueba oral de tolerancia a la glucosa (OGTT) y dado un modelo de la cinética glucosa/insulina, se propone una metodología de estimación. De nuestra investigación proponemos dos parámetros del modelo como herramienta de diagnóstico. Para este último se propone una estrategia de separación basada en máquinas de soporte vectorial.

La estructura de la tesis es la siguiente: en el *capítulo uno* se encuentra la introducción. En el *capítulo dos* se describe brevemente en qué consiste la diabetes, los distintos tipos que existen, las diferentes alteraciones de la glucosa y se explica sobre la prueba oral de tolerancia a la glucosa (OGTT). También, se describen algunos modelos minimales que describen la cinética del sistema glucosa/insulina durante la OGTT. En el *capítulo tres* se propone un ajuste del modelo básico de interacción glucosa/insulina como un problema de estimación de los parámetros del modelo. Para la estimación se utiliza la teoría bayesiana, por lo que en este capítulo se da una explicación breve sobre los pasos a seguir para aplicar la estimación bayesiana y se hacen las propuestas para el modelo con el que se trabajó. Para

las pruebas se utilizan datos obtenidos en la prueba oral de tolerancia a la glucosa (OGTT) que se le realizó a 80 pacientes. En el *capítulo cuatro* se da una breve explicación sobre el método de clasificación basado en una máquina de soporte vectorial cuando los datos a separar son linealmente separables y linealmente no separables, luego se realiza una exploración con los parámetros obtenidos en el capítulo tres y se encuentran los mejores parámetros que separan al conjunto de datos en dos grupos: pacientes sanos y pacientes que presentan alguna alteración o diabetes tipo 2. Con este conjunto de datos se aplica el método de máquina de soporte vectorial (SVM) para encontrar las dos regiones deseadas. En el *capítulo cinco* se presentan las conclusiones finales y el trabajo a futuro que se puede realizar.



## Capítulo 2

# Modelación de la interacción glucosa/insulina durante la prueba oral de tolerancia a la glucosa

En este capítulo se describe brevemente en qué consiste la enfermedad de la diabetes, los distintos tipos que existen de ésta, las diferentes alteraciones de la glucosa y se menciona sobre la herramienta que se usa para su detección. También, se describirán algunos modelos que describen la cinética del sistema glucosa/insulina durante la OGTT.

### 2.1. Sobre la enfermedad

El alimento que consume una persona se transforma y se utiliza por el cuerpo para generar energía. Cuando se digiere el alimento ocurren dos cosas:

- El azúcar conocido como glucosa entra en el torrente sanguíneo. El organismo utiliza la glucosa como fuente de energía.
- El páncreas produce hormonas como la insulina, la cual transportan la glucosa del torrente sanguíneo hacia los músculos, la grasa y las células hepáticas donde es almacenada o utilizada para producir energía.

La diabetes mellitus es una enfermedad en la cual se presenta un desorden metabólico que se caracteriza por el aumento de la cantidad de glucosa en la sangre

debido a que el páncreas no produce la insulina necesaria, la insulina producida no se utiliza de forma eficaz, o ambas alteraciones. Algunos de los síntomas que se presentan cuando se tiene la enfermedad son sed, excreción abundante de orina, visión borrosa y pérdida de peso.

La diabetes se puede clasificar por tipo como:

- Diabetes de tipo 1 (llamada anteriormente diabetes mellitus insulino - dependiente), la cual se caracteriza por la carencia absoluta de insulina. Es más común en niños y adolescentes.
- Diabetes de tipo 2 (llamada anteriormente diabetes mellitus insulino - no dependiente), en la cual hay un defecto en la secreción de insulina. Tiende a aparecer en la edad adulta, aunque está aumentando en niños y adolescentes. Se relaciona con la obesidad, la inactividad física y la mala alimentación.
- Diabetes del embarazo, la cual se caracteriza por el aumento de la cantidad de glucosa en la sangre de forma variable durante el embarazo. Lo más normal es que desaparezca en las siguientes 6 semanas después del parto.

También está la alteración de glucosa de ayuno (IFG) y la intolerancia a la glucosa (IGT) que son estados intermedios de hiperglucemia (aumento anormal de la cantidad de glucosa en la sangre). En la alteración de glucosa de ayuno, la cantidad de glucosa en la sangre estando en ayunas es mayor a lo normal, pero no llega a los valores para considerarse diabetes. En la intolerancia a la glucosa, la cantidad de glucosa en la sangre a las 2 horas de haberse tomado una dosis de glucosa es mayor a lo normal, pero no llega a los valores para considerarse diabetes. Estas alteraciones son estados entre la normalidad y la diabetes, pero quienes lo presentan tienen mayor riesgo de sufrir diabetes de tipo 2.

Los niveles normales de glucosa de una persona en ayunas deben estar entre los 70 y 100 mg/dl, si la concentración de glucosa se aleja excesivamente de estos valores puede llevar a grandes consecuencias, por ejemplo, si los niveles de glucosa son bajos puede causar convulsiones, pérdida de conciencia e incluso la muerte, y si los niveles de glucosa son elevados puede causar daño, disfunción y falla de distintos órganos y sistemas como los ojos, los riñones, el corazón, los vasos sanguíneos y los nervios.

Una herramienta para el diagnóstico de la diabetes es la prueba oral de tolerancia a la glucosa (OGTT por sus siglas en inglés). Esta prueba se realiza después de tres días en los que se ha tenido dieta normal y actividad física usual. El paciente debe presentarse en ayunas de toda la noche, en el cual se permite el consumo de agua. Se toma una muestra de sangre en ayunas, después se ingiere una dosis de 75 gr de glucosa en 250 ml de agua y se empiezan a tomar muestras de sangre por las siguientes dos horas cada treinta minutos.

En [9] comentan que una vez se toma la muestra de sangre, esta es guardada en un tubo que contiene fluoruro de sodio (6 mg por ml de sangre) e inmediatamente es centrifugada para separar el plasma. El plasma es congelado hasta que se pueda estimar la concentración de glucosa. Con esta información, se interpretan los resultados como se muestra en el cuadro 2.1.

		Plasma venoso (mg dl <sup>-1</sup> )
Diabetes Mellitus	Glucosa en ayunas	$\geq 126$
	Glucosa a las 2h	$\geq 200$
Alteración de glucosa de ayuno (IFG)	Glucosa en ayunas	$\geq 110$ y $< 126$
	Glucosa a las 2h (si se mide)	$< 140$
Intolerancia a la glucosa (IGT)	Glucosa en ayunas (si se mide)	$< 126$
	Glucosa a las 2h	$\geq 140$

Cuadro 2.1: Valores de diagnóstico de diabetes mellitus y otras categorías de hiperglucemia.

Con el fin de mejorar el diagnóstico se han propuesto varios modelos en la literatura para describir la cinética del sistema glucosa/insulina durante la OGTT. Una estrategia natural es usar los parámetros de los modelos para determinar las diferentes patologías diabéticas. En la siguiente sección analizamos algunos modelos comunes en la literatura.

## 2.2. Modelos de interacción glucosa/insulina

La interacción entre la glucosa y la insulina se modela mediante sistemas en ecuaciones diferenciales y pueden ser muy complejos y de gran dimensión. Describimos algunos modelos minimales que se han utilizado en la práctica con éxito razonable. Utilizaremos la notación que se usa en las referencias consultadas.

### 2.2.1. El modelo básico de interacción glucosa/insulina

El modelo de diabetes de interés con el que se trabajó, que se tomo de [1], se basa en dos principios: todas las personas tienen un nivel óptimo de glucosa y que los niveles de glucosa forman parte de un proceso autorregulador en que también interfieren distintas hormonas como la insulina y el glucagón. La insulina, secreta por las células  $\beta$  del páncreas, transporta la glucosa del torrente sanguíneo hacia los músculos, el tejido adiposo y las células hepáticas.

El modelo que explica el proceso regulador de los niveles de glucosa durante una prueba oral de tolerancia a la glucosa, está descrito por las siguientes ecuaciones diferenciales

$$\begin{aligned}\dot{G} &= F_1(G, H) + J(t) \\ \dot{H} &= F_2(G, H)\end{aligned}\tag{2.2.1}$$

donde  $G$  representa la concentración de glucosa en la sangre en el tiempo  $t$ ,  $H$  representa la concentración de insulina en el tiempo  $t$ . La función  $J(t)$  representa la rapidez a la cual incrementa la concentración de glucosa por una entrada externa y la dependencia con respecto a  $G$  y  $H$  de las funciones  $F_1$  y  $F_2$  se debe a que los cambios en  $G$  y  $H$  están determinandos por los valores de  $G$  y de  $H$ .

Se supone que los valores  $G_0$  y  $H_0$  son los valores estables que alcanza  $G$  y  $H$  cuando el paciente llega en ayunas a realizarse la prueba, por lo tanto, se tiene que  $F_1(G_0, H_0) = 0$  y  $F_2(G_0, H_0) = 0$ .

Como los valores que son de interés son las desviaciones que tiene  $G$  y  $H$  con respecto a los valores estables  $G_0$  y  $H_0$  en un cierto tiempo  $t$ , entonces se realiza la siguiente sustitución

$$\begin{aligned}g &= G - G_0 \\ h &= H - H_0\end{aligned}\tag{2.2.2}$$

Luego, se obtiene que el sistema (2.2.1) queda de la siguiente manera

$$\begin{aligned}\dot{g} &= F_1(G_0 + g, H_0 + h) + J \\ \dot{h} &= F_2(G_0 + g, H_0 + h)\end{aligned}\tag{2.2.3}$$



Realizando el desarrollo de Taylor a las funciones  $F_1$  y  $F_2$  se obtiene que

$$F_1(G_0 + g, H_0 + h) = F_1(G_0, H_0) + \frac{\partial F_1(G_0, H_0)}{\partial G} g + \frac{\partial F_1(G_0, H_0)}{\partial H} h + e_1$$

$$F_2(G_0 + g, H_0 + h) = F_2(G_0, H_0) + \frac{\partial F_2(G_0, H_0)}{\partial G} g + \frac{\partial F_2(G_0, H_0)}{\partial H} h + e_2$$

donde  $e_1$  y  $e_2$  son muy pequeñas con respecto a  $g$  y  $h$ .

Se supone que los valores de  $G$  y  $H$  no difieren mucho de los valores estables  $G_0$  y  $H_0$ , entonces podemos omitir  $e_1$  y  $e_2$  del desarrollo anterior, obteniendo

$$\begin{aligned} \dot{g} &= \frac{\partial F_1(G_0, H_0)}{\partial G} g + \frac{\partial F_1(G_0, H_0)}{\partial H} h + J \\ \dot{h} &= \frac{\partial F_2(G_0, H_0)}{\partial G} g + \frac{\partial F_2(G_0, H_0)}{\partial H} h \end{aligned} \quad (2.2.4)$$

Se puede determinar los signos de los términos que acompañan a  $g$  y  $h$  en (2.2.4).

- Si  $h = 0$  y  $g > 0$ , entonces  $\dot{g} < 0$ , ya que la concentración de glucosa en la sangre disminuye debido a que ésta pasa a los músculos, el tejido adiposo y las células hepáticas. Por lo tanto,  $\frac{\partial F_1(G_0, H_0)}{\partial G} < 0$ .
- Si  $h > 0$ , entonces el páncreas secreta insulina para disminuir la concentración de glucosa en la sangre. Por lo tanto,  $\frac{\partial F_1(G_0, H_0)}{\partial H} < 0$ .
- Si  $g > 0$ , entonces el cuerpo produce hormonas para regular la concentración de glucosa, lo que incrementa  $H$ . Por lo tanto,  $\frac{\partial F_2(G_0, H_0)}{\partial G} > 0$ .
- Por el metabolismo hormonal del cuerpo la concentración de hormonas en la sangre debe disminuir. Por lo tanto,  $\frac{\partial F_2(G_0, H_0)}{\partial H} < 0$ .

Por lo tanto, el modelo de diabetes queda de la siguiente manera

$$\dot{g} = -m_1g - m_2h + J \quad (2.2.5)$$

$$\dot{h} = -m_3h + m_4g \quad (2.2.6)$$

donde  $m_1, m_2, m_3$  y  $m_4$  son constantes positivas.

### Reducción del modelo

Las ecuaciones (2.2.5) y (2.2.6) son dos ecuaciones de primer orden que están en términos de  $g$  y  $h$ . Como la información que se requiere es la concentración de glucosa en la sangre, entonces se busca eliminar la variable  $h$  y obtener una ecuación que solo dependa de  $g$ . Para obtener esto, se empieza derivando la ecuación (2.2.5) con respecto al tiempo y se obtiene

$$\ddot{g} = -m_1\dot{g} - m_2\dot{h} + \dot{J} \quad (2.2.7)$$

sustituyendo (2.2.6) en (2.2.7) se tiene

$$\ddot{g} = -m_1\dot{g} + m_2m_3h - m_2m_4g + \dot{J}. \quad (2.2.8)$$

De la ecuación (2.2.5) se puede escribir

$$m_2h = -\dot{g} - m_1g + J.$$

Por lo tanto,  $g$  satisface la ecuación diferencial de segundo orden

$$\ddot{g} + (m_1 + m_3)\dot{g} + (m_1m_3 + m_2m_4)g = m_3J + \dot{J}. \quad (2.2.9)$$

El lado derecho de la ecuación (2.2.9) será cero salvo en un pequeño instante de tiempo, el momento en el que la persona ingiere la dosis de glucosa requerida para realizar la prueba. Si suponemos que  $t = 0$  es el tiempo en el que la glucosa se ingiere completamente, para tiempos  $t > 0$ ,  $g$  satisface la siguiente ecuación diferencial lineal homogénea de segundo orden

$$\ddot{g} + 2\alpha\dot{g} + \omega_0^2g = 0 \quad (2.2.10)$$

donde  $\alpha = \frac{m_1 + m_3}{2}$  y  $\omega_0^2 = m_1m_3 + m_2m_4$ .

Las soluciones de  $g(t)$  de la ecuación (2.2.10) son de tres formas distintas, cuando  $\alpha^2 - \omega_0^2$  es positivo, negativo o cero. Veamos la solución de  $g(t)$  en los tres casos:

1.  $\alpha^2 - \omega_0^2$  es negativo.

$$\begin{aligned} g(t) &= e^{-\alpha t}(a \cos \omega t + b \sin \omega t) \\ &= \sqrt{a^2 + b^2} e^{-\alpha t} \left( \frac{a}{\sqrt{a^2 + b^2}} \cos \omega t + \frac{b}{\sqrt{a^2 + b^2}} \sin \omega t \right) \\ &= A e^{-\alpha t} (\cos \delta \cos \omega t + \operatorname{sen} \delta \sin \omega t) \\ &= A e^{-\alpha t} \cos(\omega t - \delta) \end{aligned}$$

Por lo tanto,

$$G(t) = G_0 + A e^{-\alpha t} \cos(\omega t - \delta), \quad \text{con } \omega^2 = \omega_0^2 - \alpha^2. \quad (2.2.11)$$

2.  $\alpha^2 - \omega_0^2$  es positivo.

$$\begin{aligned} g(t) &= a e^{-(\alpha-\omega)t} + b e^{-(\alpha+\omega)t} \\ &= \sqrt{a^2 + b^2} e^{-\alpha t} \left( \frac{a}{\sqrt{a^2 + b^2}} e^{\omega t} + \frac{b}{\sqrt{a^2 + b^2}} e^{-\omega t} \right) \\ &= A e^{-\alpha t} (e^{\omega t} \cos \delta + e^{-\omega t} \operatorname{sen} \delta) \end{aligned}$$

Por lo tanto,

$$G(t) = G_0 + A e^{-\alpha t} (e^{\omega t} \cos \delta + e^{-\omega t} \operatorname{sen} \delta), \quad \text{con } \omega^2 = \alpha^2 - \omega_0^2. \quad (2.2.12)$$

3.  $\alpha^2 - \omega_0^2 = 0$  es igual a cero.

$$\begin{aligned} g(t) &= e^{-\alpha t}(a + bt) \\ &= \sqrt{a^2 + b^2} e^{-\alpha t} \left( \frac{a}{\sqrt{a^2 + b^2}} + \frac{b}{\sqrt{a^2 + b^2}} t \right) \\ &= A e^{-\alpha t} (\cos \delta + t \operatorname{sen} \delta) \end{aligned}$$

Por lo tanto,

$$G(t) = G_0 + A e^{-\alpha t} (\cos \delta + t \operatorname{sen} \delta). \quad (2.2.13)$$

**Comentario.** Mostraremos en capítulos subsecuentes que a pesar de su simplicidad, el modelo anterior ha sido satisfactorio para una clasificación adecuada de patologías diabéticas.

## 2.3. Otros Modelos

Con el fin de ilustrar los diferentes factores en la modelación de la interacción glucosa/insulina, describimos otros modelos a continuación.

### Un modelo minimal glucosa/insulina

Un modelo minimal de glucosa/insulina descrito en [2], se basa en la cinética de la concentración plásmatica de glucosa,  $G$  (El valor basal será  $G_b$ ), y la acción de la insulina,  $X$ . El modelo de ecuaciones será

$$\begin{aligned}\frac{dG}{dt} &= -(p_1 + X(t))G(t) + p_1G_b + \frac{Ra(t)}{V_G}, \quad G(0) = G_b, \\ \frac{dX}{dt} &= -p_2X(t) + p_3(I(t) - I_b), \quad X(0) = 0,\end{aligned}$$

donde  $I$  es la concentración plásmatica de insulina (El valor basal será  $I_b$ ). Los datos de la concentración de insulina se suponen libres de errores y se utilizan para construir una función de entrada del sistema.  $Ra$  es la razón en la que aparece la glucosa oral en el plasma, y  $p_1, p_2, p_3$  y  $V_g$  son parámetros que tienen el mismo significado que en el modelo minimal con IVGTT.

### Un modelo minimal con interacción de glucagón

El modelo minimal de interacción glucosa (G)/insulina (I)/glucagón (L) descrito en [3], se basa en 5 ecuaciones diferenciales. Este modelo tiene 4 parámetros desconocidos los cuales son estimados usando un enfoque bayesiano. El modelo de ecuaciones será

$$\begin{aligned}\frac{dG}{dt} &= L - I + \frac{D}{\theta_2}, \\ \frac{dI}{dt} &= \theta_0(G - G_b)^+ - \frac{I}{a}, \\ \frac{dI}{dt} &= \theta_1(G_b - G)^+ - \frac{L}{b}, \\ \frac{dD}{dt} &= -\frac{D}{\theta_2} + \frac{2V}{c}, \\ \frac{dV}{dt} &= -\frac{2V}{c},\end{aligned}$$

donde  $D$  es la glucosa en el sistema digestivo,  $V$  es la glucosa de la solución que se toma al ser transferida al sistema digestivo,  $a$  y  $b$  es la vida media de depuración de la insulina y el glucagón,  $c$  es el tiempo que el sujeto demora en tomarse la solución de glucosa,  $\theta_0$  es la sensibilidad de la insulina en los tejidos,  $\theta_1$  es la sensibilidad del glucagón en el hígado y  $\theta_2$  es la vida media de la glucosa en el sistema digestivo. Los parámetros  $\theta_0, \theta_1$  y  $\theta_2$  son parámetros a estimar.

Las variables  $a$  y  $b$  se toman como 0,6 horas y  $c$  como  $\frac{5}{60}$  horas. En el artículo se menciona que para la estimación bayesiana tomaron las a priori de los parámetros  $\theta_0, \theta_1$  como una distribución Gamma con parámetro de forma 2 y parámetro de escala  $\frac{1}{4}$  y para  $\theta_2$  tomaron una distribución Gamma con parámetro de forma 10 y parámetro de escala 20, truncada en los valores extremos  $0,16 < \theta_2 < 2$ .



# Capítulo 3

## Ajuste del modelo básico de interacción glucosa/insulina a datos reales de OGTT

En este capítulo se propone el ajuste del modelo básico de interacción glucosa/insulina (2.2.10) como un problema de estimación de los parámetros del modelo. El problema se formula como uno de estimación bayesiana de los parámetros  $A$ ,  $\alpha$ ,  $\delta$  y/o  $\omega_0$ . Se utilizan datos obtenidos en la prueba oral de tolerancia a la glucosa (OGTT) que se le realizó a 80 pacientes.

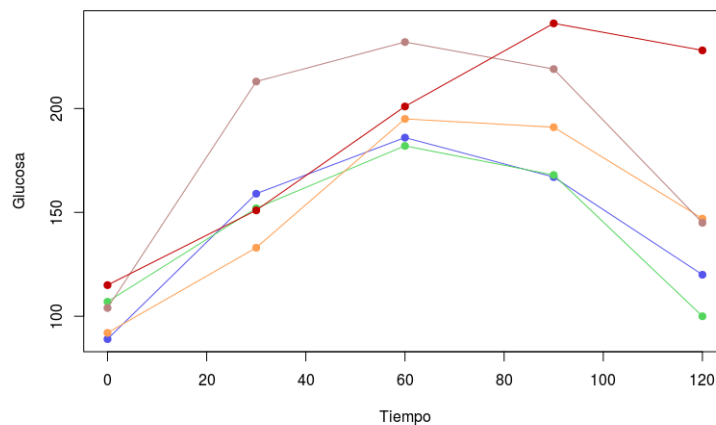
### 3.1. Análisis de datos

Para la estimación de parámetros que se realizó, se usaron los datos proporcionados por la Dra. Adriana Monroy Guzmán, al aplicar la prueba oral de tolerancia a la glucosa (OGTT) en 80 pacientes mujeres en el Hospital General de la Ciudad de México. Los resultados que se obtuvieron a partir de la prueba fueron:

- 51 pacientes sanos.
- 4 con alteración de glucosa de ayuno (IFG), es decir, la concentración de glucosa en ayuno es mayor o igual a 100.
- 15 con intolerancia a la glucosa (IGT), es decir, la concentración de la glucosa a las dos horas de iniciada la prueba es mayor o igual a 140.

- 7 con las dos alteraciones anteriores (IFG - IGT)
- 3 con diabetes de tipo 2 (T2DM), es decir, la concentración de glucosa en ayuno es mayor o igual a 126 y la concentración de la glucosa a las dos horas de iniciada la prueba es mayor o igual a 200.

En la figura 3.1 tomamos los resultados obtenidos de la prueba oral de tolerancia a la glucosa (OGTT) por un paciente de cada una de las clases mencionadas anteriormente. La curva azul es de un paciente sano, la curva verde es de un paciente con IFG, la curva naranja es de un paciente con IGT, la curva café es de un paciente con IFG e IGT, y la curva roja es de un paciente con diabetes tipo 2. Como se puede observar no hay un patrón distintivo, lo cual nos muestra que solo con la información que se tiene de la concentración de glucosa en la sangre en distintos tiempos, no es suficiente para una clasificación.



**Fuente:** Elaboración propia.

Figura 3.1: Concentración de glucosa en un paciente de cada clase.

En la figura 3.2 se muestra la dispersión de los datos en cada una de las clases mencionadas y en los distintos tiempos que se tomó la prueba. Se puede observar que la concentración de glucosa de las personas con diabetes de tipo 2 en todos los tiempos es mayor que en el resto de las clases. También, se ve que para los pacientes sanos los valores más altos de la concentración de glucosa se encuentran en el tiempo 30, para los pacientes con IFG en 60, para los pacientes con IGT en



60, para los pacientes con IFG-IGT en 90 y para los pacientes con diabetes tipo 2 en 90. Si trazáramos una curva por las medianas, se tendría que la curva entre IFG e IGT tienen un compartamiento similar, pero la de IGT se demora más en caer. Las curvas entre IFG-IGT y diabetes tipo 2 también tienen un compartamiento similar, pero la caída de la curva de diabetes tipo 2 es más lenta.

En el apéndice A se encuentran las tablas con los valores de la concentración de glucosa en los tiempos 0,30,60,90 y 120 de los 80 pacientes.

## 3.2. Preliminares sobre estimación bayesiana

Un problema bien planteado es aquel que tiene existencia, unicidad y continuidad de la solución con respecto a los datos. Hay una relación de causalidad datos a solución. Un problema mal planteado es aquel que no satisface alguna de las propiedades. Un problema inverso es en general un problema mal planteado. Se encuentra típicamente en situaciones en donde uno puede hacer observaciones indirectas de un cantidad de interés.

En los problemas inversos clásicos, se mide una cantidad  $y \in \mathbb{R}^m$  para obtener información acerca de otra cantidad  $x \in \mathbb{R}^n$ . Para relacionar ambas cantidades, se plantea un modelo en donde hay dependencia entre éstas. El modelo puede no ser correcto y contener otros parámetros que no son muy bien conocidos. Además, las cantidades medidas  $y$  siempre contienen ruido. El modelo de un problema inverso lo podemos escribir como

$$y = f(x, e) \tag{3.2.1}$$

donde  $f : \mathbb{R}^n \times \mathbb{R}^k \rightarrow \mathbb{R}^m$  es la función del modelo y  $e \in \mathbb{R}^k$  es un vector que contiene los parámetros que no son muy bien conocidos y las mediciones de ruido.

### 3.2.1. Metodología de estimación

Los métodos de inversión estadísticos buscan reformular el problema inverso en una búsqueda estadística de información. Se tienen cantidades que pueden ser observadas y otras que no se pueden observar. En los problemas inversos, las cantidades que no pueden ser observadas son de gran interés y éstas dependen de las

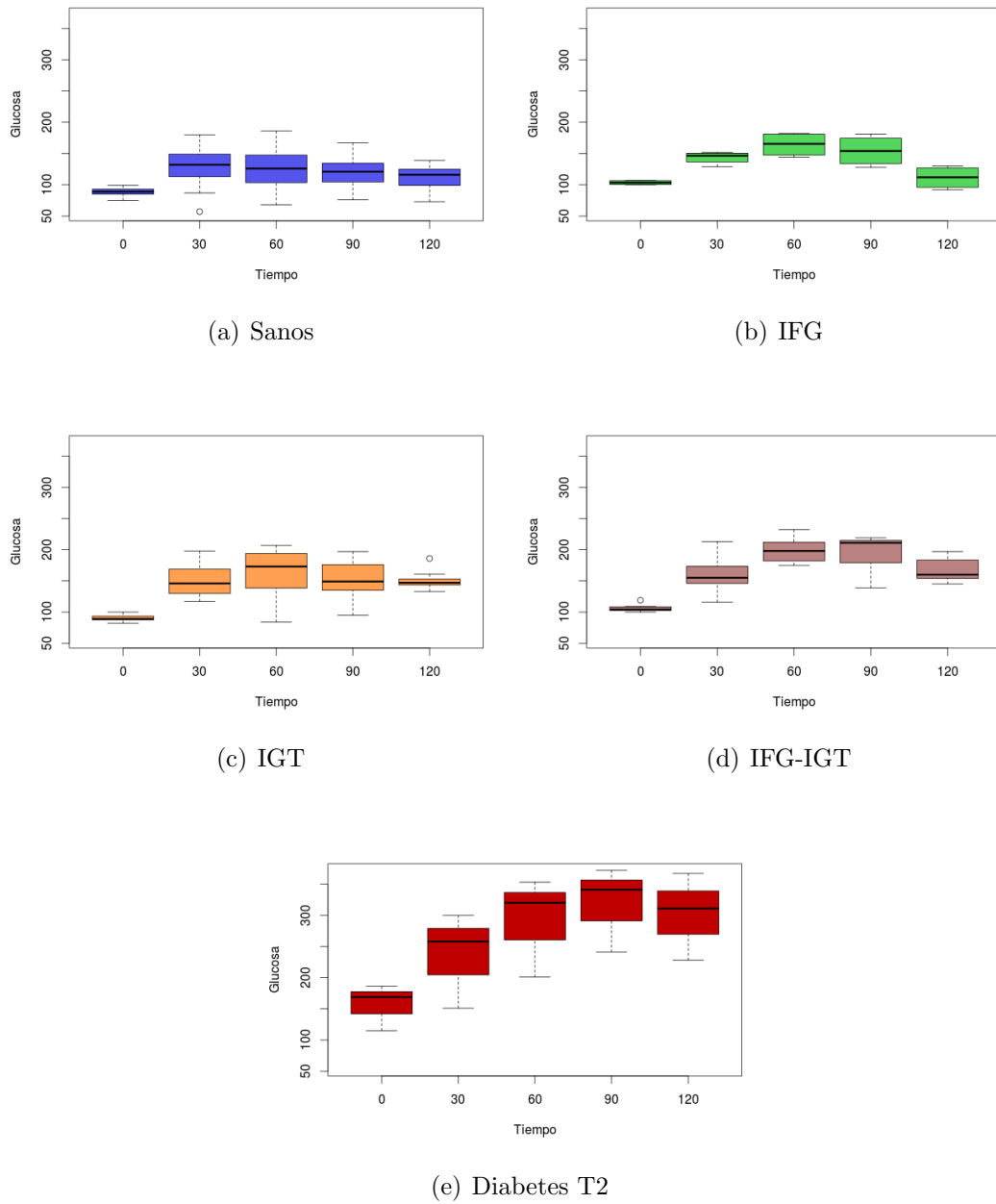


Figura 3.2: Dispersión de los datos en cada una de las clases.

cantidades que pueden ser observadas a través de modelos. El objetivo de la inversión estadística es extraer información y evaluar la incertidumbre acerca de las

variables basado en el conocimiento durante el proceso de medición, y de la información y el modelo de las variables desconocidas que se tiene antes de la medición.

El enfoque de la inversión estadística se basa en los siguientes principios:

1. Todas las variables que estén en el modelo se consideran como variables aleatorias.
2. La aleatoriedad describirá el grado de información sobre las realizaciones.
3. El grado de información que concierne a estos valores se codifica en distribuciones de probabilidad.
4. La solución del problema inverso será la distribución de probabilidad posterior.

Como en los problemas inversos estadísticos todos los parámetros son variables aleatorias, se denotarán con letra mayúscula y las realizaciones se denotarán con letra minúscula, luego el modelo (3.2.1) quedará

$$Y = f(X, E). \quad (3.2.2)$$

Está es una relación entre las tres variable aleatorias  $Y$ ,  $X$  y  $E$  y sus distribuciones aleatorias dependen unas de otras. Ahora, se introducirá la nomenclatura con la que se trabajará.

La variable aleatoria observable  $Y$  la llamaremos **la medición** y su realización  $Y = y_{obs}$  en el proceso de medición será **el dato**. La variable aleatoria no observable  $X$  la llamaremos **la variable desconocida**. Las variables que no son ni las observable ni las no observables se llaman **parámetros** o **ruido**, dependiendo del ajuste. Antes de realizar las mediciones de  $Y$ , se tiene una información acerca de la variable  $X$ . Esta información se codificará en la función de densidad  $x \mapsto \pi_{pr}(x)$  que llamaremos **densidad a priori**.

Después de analizar el ajuste de medición y la información disponible alrededor de las variables, se encuentra la función de densidad conjunta de  $X$  y  $Y$ , la cual denotamos por  $\pi(x, y)$ . Entonces, la densidad marginal de la variable desconocida  $X$  debe ser

$$\pi_{pr}(x) = \int_{\mathbb{R}^m} \pi(x, y) dy.$$

Por otro lado, si conocemos el valor de la variable desconocida,  $X = x$ , la función de densidad condicional de  $Y$  da esta información, la cual será

$$\pi(x|y) = \frac{\pi(x, y)}{\pi_{pr}(x)}, \quad \text{si } \pi_{pr}(x) \neq 0.$$

La probabilidad condicional de  $Y$  se llamará **función de verosimilitud**, porque esto expresa la verosimilitud de los resultados de las distintas mediciones con  $X = x$  dado.

Supongamos que los datos medidos  $Y = y_{obs}$  son dados. La distribución de probabilidad condicional

$$\pi(x|y_{obs}) = \frac{\pi(x, y_{obs})}{\pi(y_{obs})}, \quad \text{si } \pi(y_{obs}) = \int_{\mathbb{R}^n} \pi(x, y_{obs}) dx \neq 0$$

se llamará **distribución posterior** de  $X$ . Esta distribución expresa que conoceremos a cerca de  $X$  después de la observación realizada  $Y = y_{obs}$ .

El enfoque bayesiano de los problemas inversos se resume en: Dado los datos  $Y = y_{obs}$ , encontrar la distribución de probabilidad condicional  $\pi(x|y_{obs})$  de la variable  $X$ . El siguiente teorema es el *teorema de Bayes de problemas inversos*.

**Teorema 3.2.1.** *Consideremos que la variable aleatoria  $X \in \mathbb{R}^n$  tiene una función de densidad a priori  $\pi_{pr}(x)$  y los datos consisten de valores observados  $y_{obs}$  de una variable aleatoria observable  $Y \in \mathbb{R}^k$  tal que  $\pi(y_{obs}) > 0$ . Entonces, la función de distribución posterior de  $X$ , dado los datos  $y_{obs}$  es*

$$\pi_{post}(x) = \pi(x|y_{obs}) = \frac{\pi_{pr}(x)\pi(y_{obs}|x)}{\pi(y_{obs})} \tag{3.2.3}$$

En resumen resolver un problema inverso en el sentido bayesiano se puede dividir en tres tareas:

1. Con la información de de la variable desconocida  $X$ , proponer una función de densidad a priori,  $\pi_{pr}(x)$ , que tenga toda la información a priori.
2. Encontrar la función de verosimilitud  $\pi(y|x)$  que describe la relación entre las observaciones y las variables desconocidas.
3. Desarrollar métodos que exploren la función de densidad posterior.

Para el paso 3 utilizaremos el metodo Markov Chain Monte Carlo (MCMC) con la impletación del módulo de python `emcee`. [ref tal].

### 3.2.2. Estimadores puntuales

La solución del problema inverso mediante el enfoque bayesiano será una distribución posterior. Si la variable desconocida es una variable aleatoria que tiene pocas entradas, la visualización de la función de densidad como una función no negativa de estas variables es posible. En los problemas inversos que se presentan normalmente la dimensionalidad del problema es grande, y la distribución posterior estará en un espacio de dimensión alta donde la visualización no es posible. Aunque, si se conoce la distribución posterior, se puede calcular estimadores puntuales o estimadores de propagación o de intervalo.

Cuando se trabaja con un estimador puntual se busca responder a la pregunta: “Dado un conjunto de observaciones y una información a priori, ¿Cuál es el valor más probable de la variable desconocida  $X$ ? ”, mientras que cuando se utilizan estimadores de intervalo se responde a la pregunta: “Dado un conjunto de observaciones y la información a priori, ¿En qué intervalo están los valores de la variable desconocida con 95 % de probabilidad?”.

Uno de los estimadores estadísticos más comunes es el estimador **máximo posterior (MAP)**. Dada la función de densidad posterior  $\pi(x|y)$ , de la variable desconocida  $X \in \mathbb{R}^n$ , el estimador MAP satisface,

$$x_{MAP} = \arg \max_{x \in \mathbb{R}^n} \pi(x|y), \quad (3.2.4)$$

siempre y cuando el maximizador exista. Aunque este maximizador exista, puede que no sea único. La posible no existencia y no unicidad muestran que el estimador individual basado en un enfoque de problema inverso puede ser insatisfactorio. Para encontrar el estimador MAP se debe hallar la solución de un problema de optimización. Lo más común es utilizar métodos iterativos basados en gradientes para la búsqueda del maximizador. Se ha visto, que resolver este problema computacionalmente puede ser similar a resolver los métodos clásicos de regularización.

Otro estimador puntual que es muy popular es la **media condicional (MC)** de la variable desconocida  $X$  condicionada sobre los datos  $y$ , y se define como

$$x_{MC} = E(x|y) = \int_{\mathbb{R}^n} x \pi(x|y) dx, \quad (3.2.5)$$

siempre y cuando la integral converja. Para encontrar el estimador MC se debe

resolver un problema de integración. Una ventaja del estimador MC es que las propiedades de suavidad no son tan importantes como con el problema de optimización del estimador MAP, más aún cuando se utilizan métodos de gradiente para solucionar éste. El principal problema que tienen el estimador MC es que la integración casi siempre se hace sobre un espacio de alta dimensión, y en este caso los métodos comunes de cuadratura no se pueden aplicar.

### 3.3. Aplicación de la metodología a los datos de OGTT

#### 3.3.1. Función de Verosimilitud

La construcción de la función de verosimilitud es una de las partes más sencillas de los problemas inversos estadísticos. La función de verosimilitud contiene la información del modelo directo que se usa en las técnicas de inversión, además de la información del ruido, de otras mediciones que se usen y la incertidumbre del modelado.

Uno de los modelos más usados es el de adicionar el ruido y considerarlo independiente de la variable desconocida  $X$ . Entonces, el modelo estocástico será

$$Y = f(X) + E, \tag{3.3.6}$$

con  $X \in \mathbb{R}^n$ ,  $Y, E \in \mathbb{R}^m$ , y  $X$  y  $E$  independientes entre si. Consideremos que la distribución de probabilidad del ruido  $E$  es conocida, entonces se tiene que

$$\mu_E(B) = P\{E \in B\} = \int_B \pi_{ruido}(e)de.$$

Si se fija  $X = x$ , por la independencia entre  $X$  y  $E$ , la función de densidad de  $E$  no tendrá alteraciones cuando se condiciona sobre  $X = x$ . Por lo tanto, cuando  $Y$  se condicione sobre  $X$  se distribuirá igual que  $E$ , la función de densidad será trasladada por  $f(x)$ , es decir, la función de verosimilitud será

$$\pi(y|x) = \pi_{ruido}(y - f(x)). \tag{3.3.7}$$

Entonces, si la función de densidad a priori de la variable desconocida  $X$  es

$\pi_{pr}$ , de la fórmula de Bayes (3.2.3), se tiene que

$$\pi(x|y) \propto \pi_{pr}(x)\pi_{ruido}(y - f(x)).$$

Ahora, consideraremos que la variable desconocida  $X$  y el ruido  $E$  no son independientes. Para este caso, es necesario conocer la función de densidad condicional del ruido, es decir,

$$\mu_E(B|x) = \int_B \pi_{ruido}(B|x)de.$$

Se puede escribir la función de verosimilitud como,

$$\pi(y|x) = \int_{\mathbb{R}^m} \pi(y|x, e)\pi_{ruido}(e|x)de. \quad (3.3.8)$$

Cuando se fija  $X = x$  y  $E = e$ , se especifica  $Y$ , es decir,  $Y = y = f(x) + e$ , entonces

$$\pi(y|x, e) = \delta(y - f(x) - e),$$

donde  $\delta(\cdot)$  es la delta de Dirac.

Sustituyendo la ecuación anterior en (3.3.8) se obtiene que,

$$\begin{aligned} \pi(y|x) &= \int_{\mathbb{R}^m} \delta(y - f(x) - e)\pi_{ruido}(e|x)de \\ &= \pi_{ruido}(y - f(x)|x) \end{aligned}$$

y por lo tanto,

$$\pi(x|y) \propto \pi_{pr}(x)\pi_{ruido}(y - f(x)|x).$$

Definiremos la función de verosimilitud,  $\pi(y|x)$ , que se usará para la estimación de parámetros usando métodos bayesianos.

Tomamos el siguiente modelo

$$Y = G(X, t) + E \quad (3.3.9)$$

donde  $G(X, t)$  es la concentración de glucosa en la sangre de una persona en el tiempo  $t$ , y  $E \in \mathbb{R}^5$  es el ruido que se distribuye  $\mathcal{N}(0, \sigma^2 I)$ .

De la prueba oral de tolerancia a la glucosa, tenemos las mediciones  $y_i$  para  $i \in \{0, 1, 2, 3, 4\}$ , que corresponden a las concentraciones de glucosa en los tiempos  $t \in \{0, 30, 60, 90, 120\}$ . La función  $G(X, t)$  depende del vector de parámetros  $X$  y del tiempo  $t$ . Los tiempos que se tienen en cuenta son los mencionados anteriormente.

Supongamos que  $X$  y  $E$  son independientes entonces se tiene que  $y_i \sim N(G(X, t_i), \sigma^2)$ . Por lo tanto,

$$\pi(y|x) = \exp\left(-\frac{1}{2\sigma^2}\|Y - G(X, t)\|^2\right) \quad (3.3.10)$$

donde

$$Y = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} \quad y \quad G(X, t) = \begin{bmatrix} G(X, t_0) \\ G(X, t_1) \\ G(X, t_2) \\ G(X, t_3) \\ G(X, t_4) \end{bmatrix}.$$

### 3.3.2. Distribuciones a priori

En los problemas inversos con el enfoque estadístico, uno de los pasos más cruciales es la construcción de la función de densidad a priori y frecuentemente esta parte es la más variante de la solución. El mayor problema para encontrar una función de distribución a priori cae en la naturaleza de la información a priori. En la mayoría de los casos el conocimiento a priori que se tiene de la variable desconocida es de naturaleza cualitativa y el problema consiste en transformar ese tipo de información en forma cuantitativa, para poder codificar la función de densidad a priori.

El objetivo es designar una función de densidad a priori,  $\pi_{pr}(x)$ , a las variables desconocidas, con la siguiente propiedad: Si  $E$  es una colección de vectores *esperados*  $x$ , que representan las posibles realizaciones de la variable desconocida  $X$  y  $U$  es una colección de vectores *inesperados*, se tiene

$$\pi_{pr}(x) \gg \pi_{pr}(x') \quad \text{con } x \in E \text{ y } x' \in U.$$

Por lo tanto, la distribución de probabilidad a priori se concentrará en los valores de  $x$  que se esperan ver y se le asigna una probabilidad más alta a estos valores



que a los valores que no se esperan ver.

Las distribuciones a priori que se tomaron para los cuatro parámetros,  $A$ ,  $\alpha$ ,  $\omega_0$  y  $\delta$ , que se quieren estimar fueron todas distribuciones uniformes. Con esta distribución se busca tener información acerca de los parámetros y de esta forma no sesgar los resultados. Los intervalos que se tomaron en cada uno de los parámetros dependieron de la interpretación que se le puede dar a cada parámetro.

- El parámetro  $A$  es la amplitud. El intervalo que se tomo fue entre  $0,5 \text{ MaxVal}$  y  $2,5 \text{ MaxVal} + 150$ , donde  $\text{MaxVal} = \max(G(t_i) - G(t_0))$  para  $t_i \in \{30, 60, 90, 120\}$ .
- El parámetro  $\alpha$  se relaciona con la rapidez de caída. Este tiene que ser mayor a cero. Se observó que el valor de  $\alpha$  era muy pequeño, por lo tanto el intervalo que se tomo fue entre 0 y 0,1.
- El parámetro  $\omega_0$  es la frecuencia. Este tiene que ver con la cantidad de ciclos que va a dar la curva, por esto debe estar entre 0 y un valor que represente un ciclo. Se observó que aproximadamente 0,15 era ese valor.
- El parámetro  $\delta$  es el ángulo de fase. Este puede tomar valores entre  $-2\pi$  y  $2\pi$ .

### 3.3.3. MCMC

MCMC es un método en el que se genera una caminata aleatoria en el espacio de parámetros, y en el tiempo, va a generar un conjunto representativo de muestras de la distribución. Cada punto en la cadena de Markov dependerá únicamente de la posición anterior.

Para la estimación de los parámetros usando estimación bayesiana se usó el módulo de python `emcee`, el cual implementa un algoritmo de muestreo MCMC. Este módulo genera muestras de los parámetros a estimar,  $X \in \mathbb{R}^n$ , y devuelve  $M$  muestras de la función de densidad posterior de  $X$ . El algoritmo que utiliza para realizar esto es una modificación del método de Metropolis - Hastings. [Artículo `emcee`]

Para aplicar el método de Cadenas de Markov Monte Carlo se tomó 1 caminata aleatoria que realizó 50000 pasos. Para determinar la distribución posterior se descartaron los primeros 20000 pasos de la caminata. Para que el muestreador pueda generar una de las realizaciones de la distribución posterior se usa la solución analítica de la ecuación diferencial o se tiene que resolver usando técnicas numéricas en el intervalo 0 a 120.

### 3.4. Resultados del ajuste

Se aplicó el método MCMC a las 80 muestras, tomando a la función  $G(X, t)$  de (3.3.9) como en la ecuación (2.2.11). Con las muestras obtenidas de la función de densidad posterior de los parámetros estimados  $A, \alpha, \delta$  y/o  $\omega_0$  se calculó el estimador MAP y el estimador MC.

Tanto para los pacientes sanos como para los pacientes que presentan una alteración de la glucosa o diabetes de tipo 2, estudiamos ajustes tanto con el estimador MAP como con el estimador MC. En la figura 3.3 se presentan los histogramas de las distribuciones marginales de cada uno de los parámetros que se estimadores para el paciente cuyos datos se presentan en la figura 3.4(a)-(b). En cada histograma se marco con un linea roja el valor que se obtuvo del estimador MAP y con una linea azul el valor que se encontro del estimador MC.

En la figura 3.4 se muestran algunas curvas de personas sanas y en la figura 3.5 se muestran algunas curvas de personas con alguna alteración de la glucosas o con diabetes tipo 2 utilizando los parámetros obtenidos con el estimador MAP y con el estimador MC, cuando se tomo a  $G(X, t)$  de (3.3.9) como en la ecuación (2.2.11).

Para cada uno de los 80 pacientes se calculó el error relativo medio entre el valor de concentración de glucosa en los tiempos  $t \in \{0, 30, 60, 90, 120\}$  y el valor de la curva obtenida con los estimadores MAP y MC en los tiempos anteriores. En la figura 3.6 se muestra el histograma del error relativo medio con el estimador MAP y con el estimador MC.

En el histograma del error relativo se observa que hay menos dispersión con el estimador MAP que con el estimador MC, ya que la mayoría de los datos con el estimador MAP su error cae entre 0 y 0,05 mientras que con el estimador MC la mayoría de los datos tienen un error entre 0 y 0,1. Esto se ve reflejado en las

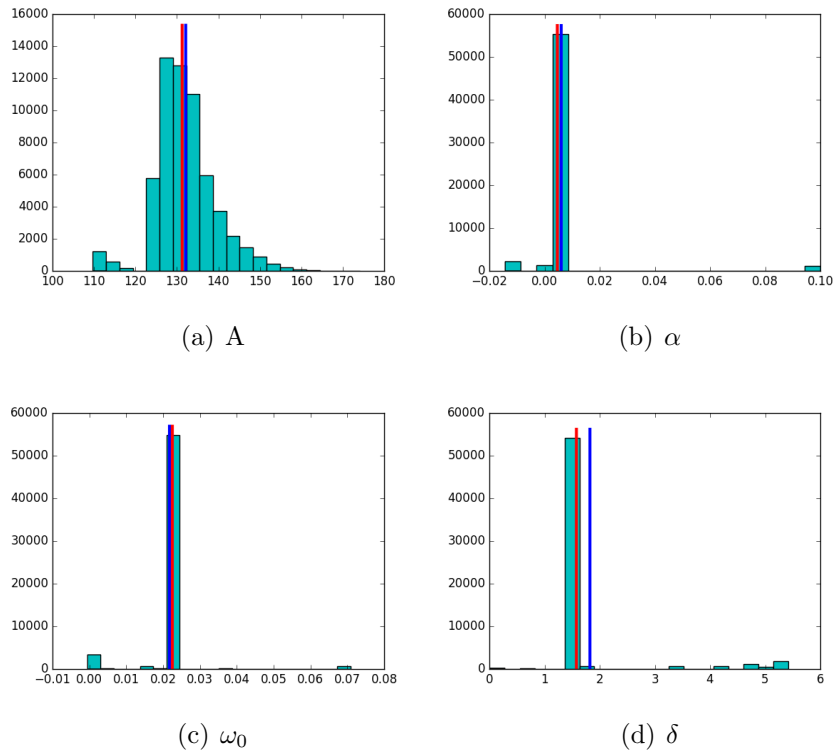


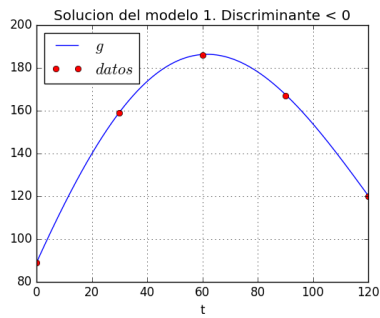
Figura 3.3: Histogramas de las distribuciones marginales para cada parámetro de los datos que se presentan en la figura 3.4(a)-(b)

figuras 3.4 y 3.5, ya que en los casos en que el estimador MAP ajuste muy bien, el estimador MC no lo hace tan bien, y en los otros casos el ajuste con ambos estimadores es muy similar. Por esto se tomaron los valores que se obtuvieron con el estimador MAP para buscar algún criterio de clasificación.

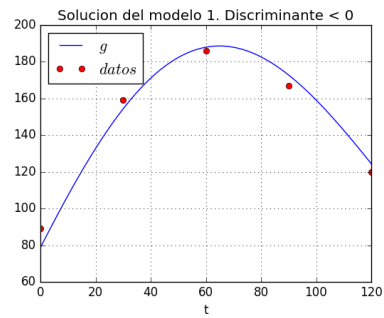
Después se aplicó el método MCMC a las 80 muestras, tomando a la función  $G(X, t)$  de (3.3.9) como en las ecuaciones (2.2.12), (2.2.13). Para estos dos casos se observaron que las curvas que se obtuvieron con el estimador MAP no ajustaban tan bien como lo hacía el primer caso, por esto se generaron los histogramas con el error relativo medio, para ver la dispersión del error en los tres casos. En la figura 3.7 se muestran dichos histogramas.

En los histogramas se observa que cuando se toma a  $G(X, t)$  como en la ecuación (2.2.11) se presenta menor dispersión que en los otros dos casos. Cuando el

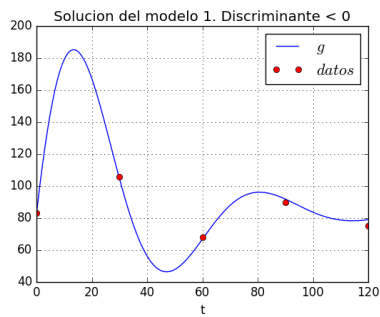
discriminante es menor a cero, 51 datos tienen un error entre 0 y 0,05 y 23 datos entre 0,05 y 0,1, cuando el discriminante es mayor a cero, 39 datos tienen un error entre 0 y 0,05 y 30 datos entre 0,05 y 0,1 y cuando el discriminante es cero 31, datos tienen un error entre 0 y 0,05 y 38 datos entre 0,05 y 0,1. Por lo tanto, para la clasificación se considerará solo los parámetros encontrados en el primer caso.



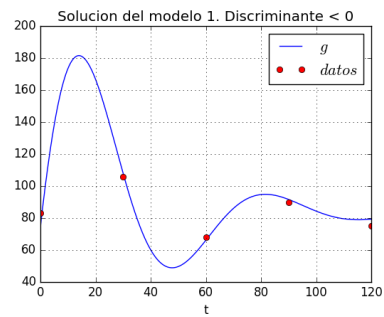
(a) MAP



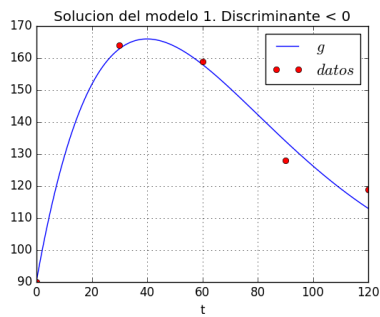
(b) MC



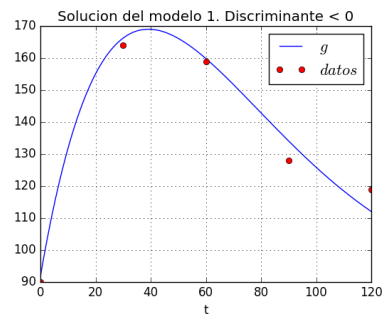
(c) MAP



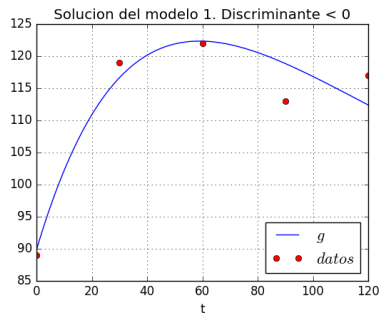
(d) MC



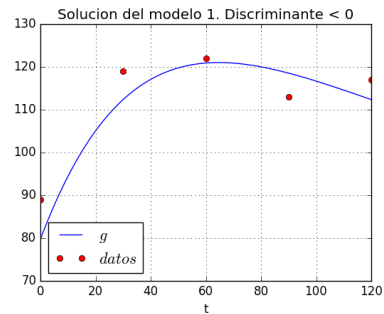
(e) MAP



(f) MC

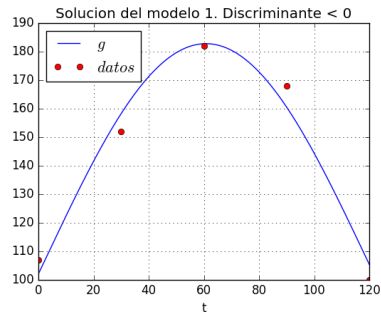


(g) MAP

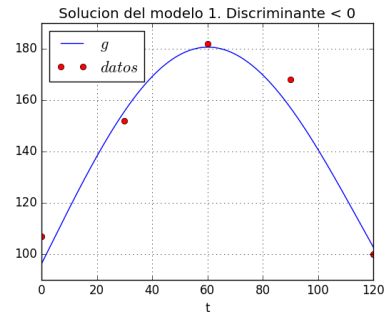


(h) MC

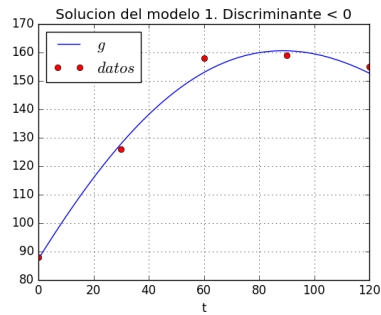
Figura 3.4: Curvas obtenidas de pacientes sanos



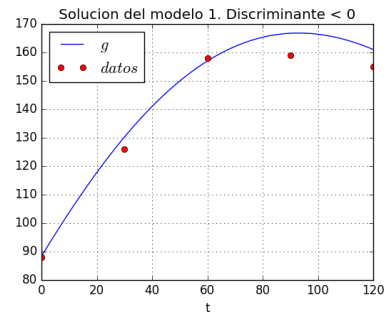
(a) MAP



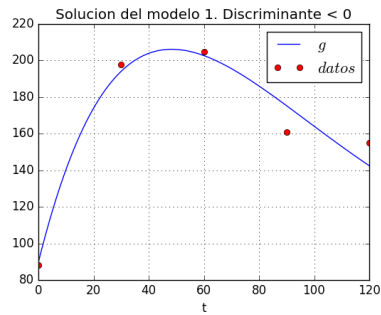
(b) MC



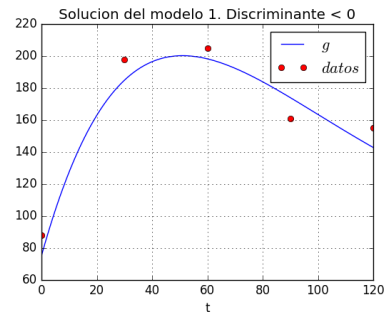
(c) MAP



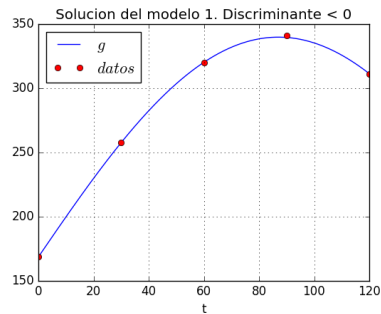
(d) MC



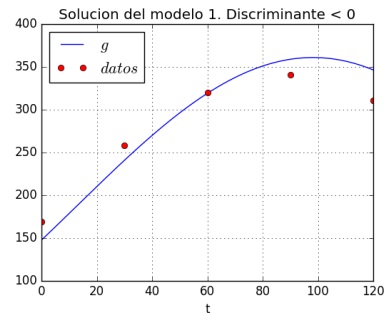
(e) MAP



(f) MC

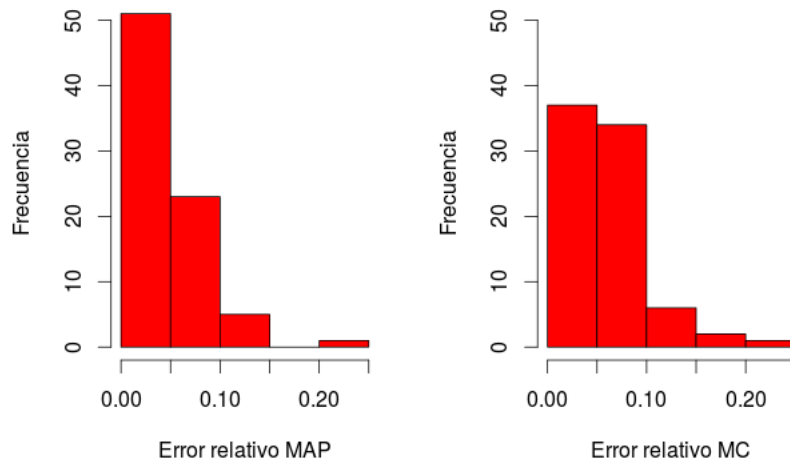


(g) MAP



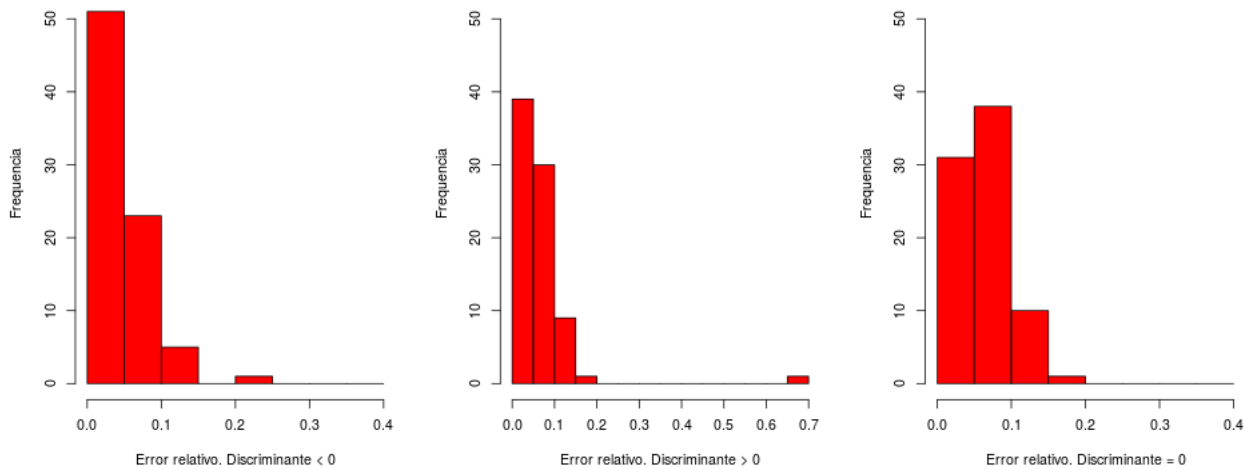
(h) MC

Figura 3.5: Curvas obtenidas de pacientes enfermos



**Fuente:** Elaboración propia.

Figura 3.6: Histogramas del error relativo medio en el caso 1



**Fuente:** Elaboración propia.

Figura 3.7: Histogramas del error relativo medio en los tres casos





# Capítulo 4

## Un algoritmo de clasificación para diagnóstico de diabetes

En este capítulo se consideran los parámetros que se encontraron en el capítulo anterior para explorar sus bondades de separación entre el grupo de pacientes sanos y pacientes que presentan alguna alteración o diabetes tipo 2. Con el conjunto de parámetros que separan a estos grupos se aplicará el método de máquina de soporte vectorial (SVM) para encontrar las dos regiones deseadas.

### 4.1. Máquina de soporte vectorial

La máquina de soporte vectorial es un método de clasificación en donde se busca, a través de algún tipo de kernel, separar un conjunto de datos en dos grupos. Cuando el conjunto de datos se puede separar por medio de un hiperplano se le conoce como clasificador de máximo margen. Como no todos los conjuntos de datos son separables, el método anterior se puede extender a este tipo de datos y encontrar el mejor hiperplano que los divida en dos grupos, este método se conoce como clasificador de soporte vectorial.

#### 4.1.1. Clasificación usando un hiperplano separador

En un espacio de dimensión  $p$ , se define un **hiperplano** como un subespacio afín plano de dimensión  $p - 1$ . Por ejemplo, si tenemos un espacio en dos dimensiones, entonces un hiperplano es un subespacio de dimensión uno, es decir, una recta.

En dos dimensiones, la ecuación del hiperplano viene dada por

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 = 0, \quad (4.1.1)$$

para los parámetros  $\beta_0, \beta_1$  y  $\beta_2$ . Si un punto  $X = (X_1, X_2)^T$  satisface la ecuación (4.1.1), entonces este punto pertenece al hiperplano. Observemos que (4.1.1) es la ecuación de una recta.

Generalizando lo anterior, tendremos que en dimensión  $p$ , la ecuación del hiperplano viene dada por

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p = 0 \quad (4.1.2)$$

lo que significa que si un vector  $p$ -dimensional  $X = (X_1, X_2, \dots, X_p)^T$  satisface la ecuación (4.1.2), entonces el vector  $X$  pertenece al hiperplano.

Cuando un vector  $X$  no cumple la ecuación (4.1.2), entonces puede satisfacer que

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p < 0, \quad (4.1.3)$$

lo que indica que el vector  $X$  cae a un lado del hiperplano, o satisfacer

$$\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p > 0, \quad (4.1.4)$$

lo que indica que el vector  $X$  cae al otro lado del hiperplano. Por lo tanto, el hiperplano divide al espacio de dimensión  $p$  en dos partes, y para saber en que lado del hiperplano cae un vector  $X$  solo falta con evaluar el signo del lado izquierdo de la ecuación (4.1.2).

Ahora, vamos a hablar de cómo clasificar un conjunto de datos en un espacio de dimensión  $p$  usando un hiperplano separador.

Se tiene una matriz  $\mathbf{X}$  de tamaño  $n \times p$  que estará conformada por  $n$  observación en un espacio de dimensión  $p$

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}, \quad (4.1.5)$$

donde cada una de estas observaciones pertenece a una clase. Etiquetaremos con -1 a una clase y con 1 a la otra, entonces  $y_1, y_2, \dots, y_n \in \{-1, 1\}$ . También, se tendrá un dato prueba que será un vector de dimensión  $p$  con observaciones,  $x^* = (x_1^*, x_2^*, \dots, x_p^*)^T$ . El objetivo es crear un clasificador usando los datos de entrenamiento de manera que clasifique de manera correcta el dato prueba. Para esto, se utilizará la idea de hiperplano separador.

Supongamos que se puede construir un hiperplano de forma que separe de forma correcta los datos de entrenamiento según la etiqueta de su clase. Un **hiperplano separador** debe cumplir que

$$\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} < 0 \quad \text{si} \quad y_i = -1, \quad (4.1.6)$$

y

$$\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} > 0 \quad \text{si} \quad y_i = 1. \quad (4.1.7)$$

Por lo tanto, un hiperplano separador tendrá la propiedad de que

$$y_i (\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}) > 0 \quad (4.1.8)$$

para todo  $i \in \{1, 2, \dots, n\}$ .

Si existe un hiperplano separador entonces el clasificador que tendremos consiste en tomar una observación prueba y colocarlo en una clase dependiendo en que lado del hiperplano caiga. La observación prueba  $x^*$  se clasificará según el signo de la función

$$f(x^*) = \beta_0 + \beta_1 x_1^* + \beta_2 x_2^* + \dots + \beta_p x_p^*. \quad (4.1.9)$$

Si  $f(x^*)$  es negativo, entonces la observación prueba estará en la clase -1 y si  $f(x^*)$  es positivo, entonces la observación prueba estará en la clase 1.

### 4.1.2. Clasificador de máximo margen

Si existe un hiperplano que separa perfectamente el conjunto de datos que se tiene, dicho hiperplano no es único. Esto pasa ya que si suponemos que tenemos un hiperplano separador podemos desplazar este hacia arriba o hacia abajo, o rotarlo de forma que no toque a ninguna de las observaciones de nuestro conjunto de

datos. Para construir el clasificador debemos elegir un hiperplano separador, por esto nos preguntamos, ¿Cuál de los infinitos hiperplanos que existen usamos?.

El hiperplano que usaremos es el hiperplano de máximo margen, también conocido como hiperplano de óptima separación. Este hiperplano separador es aquel que está más lejos del conjunto de observaciones. Si tomamos la distancia perpendicular de cada dato al hiperplano, la menor de estas distancias será la distancia mínima de las observaciones al hiperplano y se le conocerá como **margen**. Luego, el hiperplano de máximo margen es un hiperplano separador donde el margen es el más grande, es decir, el hiperplano cuya distancia mínima es la más lejana al conjunto de datos.

Una vez encontrado el hiperplano de máximo margen, podemos clasificar una observación prueba observando en que lado del hiperplano cae. A esto se le conocerá como **clasificador de máximo margen**. Se espera que si el hiperplano encontrado clasifica bien el conjunto de datos muestra, lo hará también de forma correcta con los datos prueba.

Sea  $\beta_0, \beta_1, \dots, \beta_p$  los coeficientes del hiperplano de máximo margen, entonces el clasificador de máximo margen seleccionará la clase a la que pertenece una observación prueba  $x^*$  según el signo de la función

$$f(x^*) = \beta_0 + \beta_1 x_1^* + \beta_2 x_2^* + \dots + \beta_p x_p^*.$$

Las observaciones que están equidistantes al hiperplano de óptima separación y su distancia al hiperplano es el margen se conocen como **vectores de soporte**. Estos vectores de dimensión  $p$  soportan al hiperplano, es decir, si hay movimiento de estos vectores el hiperplano también se moverá. Luego, el hiperplano de máximo margen solo depende de estos vectores, ya que si los otros datos se mueven siempre y cuando no crucen el margen, el hiperplano no cambia.

Ahora, se planteará el problema de optimización que se debe resolver para encontrar los parámetros que definen al hiperplano de máximo margen.

Se tienen  $n$  observaciones  $x_1, x_2, \dots, x_n \in \mathbb{R}^p$ , cada observación tiene una etiqueta según a la clase que pertenezca,  $y_1, y_2, \dots, y_n \in \{-1, 1\}$ . Luego, el hiperplano de máximo margen,  $\mathcal{H}(x)$ , será la solución del siguiente problema de optimización con restricciones

$$\begin{array}{ll} \text{máx} & M \\ \beta_0, \beta_1, \dots, \beta_p & \end{array} \quad (4.1.10)$$

$$\text{sujeto a } \sum_{j=1}^p \beta_j^2 = 1 \quad (4.1.11)$$

$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}) \geq M \quad \forall i = 1, 2, \dots, n. \quad (4.1.12)$$

Para garantizar que una observación esta en el lado correcto del hiperplano, basta con que cumpla que  $y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}) > 0$ . Las restricciones (4.1.12) se colocan para que todas las observaciones esten en el lado correcto del hiperplano, con un amortiguamiento de  $M$ , y  $M$  positivo.

La ecuación (4.1.11) no es una restricción del hiperplano, ya que si  $\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} = 0$  define un hiperplano, entonces  $k(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}) = 0$  también lo hará, pero se coloca como restricción para garantizar que la distancia perpendicular de la  $i$ -ésima observación al hiperplano esté dada por

$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip})$$

ya que

$$\begin{aligned} d(x_i, \mathcal{H}(x)) &= \frac{|\mathcal{H}(x_i)|}{\|(\beta_1, \beta_2, \dots, \beta_p)\|} \\ &= \frac{y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip})}{\sqrt{\sum_{j=1}^p \beta_j^2}} \\ &= y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}). \end{aligned}$$

Por lo tanto, las restricciones (4.1.11) y (4.1.12) garantizan que todo el conjunto de datos este en el lado correcto del plano y al menos a una distancia  $M$  del hiperplano. Notese, que  $M$  es el margen de nuestro hiperplano y el problema de optimización a resolver es encontrar los valores de  $\beta_0, \beta_1, \dots, \beta_p$  que maximicen el margen  $M$ .

### 4.1.3. Clasificador de soporte vectorial

El clasificador de máximo margen es muy intuitivo cuando el conjunto de datos que se tiene es perfectamente separable. Sin embargo, muchas veces no existe el hiperplano de máximo margen debido a que el conjunto de datos no es separable, y el problema de optimización (4.1.10) - (4.1.12) no tiene solución para  $M > 0$ . El clasificador de máximo margen se puede generalizar para el caso en el que el conjunto de datos sea no separable y se conoce como clasificador de soporte vectorial. Con este clasificador se busca un hiperplano que separe la mayoría de las observaciones, usando un margen suave.

Para el caso en el que el conjunto de datos no es separable, se debe considerar un clasificador que se base en usar un hiperplano que no separe perfectamente a los dos clases que se tienen, pero que se tenga

- Mayor robustez en la observación individual, y
- Mejor clasificación en la mayoría de las observaciones.

Se puede considerar clasificar de forma errónea alguna de las observaciones de entrenamiento, si se logra tener una mejor clasificación con el resto de observaciones.

El clasificador de soporte vectorial, también conocido como clasificador de margen suave, realiza lo anteriormente mencionado. Éste no va a buscar maximizar el margen de modo que todas las observaciones estén en el lado correcto del hiperplano, más aún, del lado correcto del margen, si no que va a permitir que algunas observaciones estén en el lado incorrecto del margen o hasta que estén en el lado incorrecto del hiperplano. El margen va a ser suave debido a que se puede violar por algunas observaciones.

Se buscará un hiperplano que separe a la mayoría de las observaciones de entrenamiento de forma correcta en las dos clases que se tienen, pero se puede permitir que mal clasifique a unos cuantas observaciones. Luego, el hiperplano que escogemos será la solución del siguiente problema de optimización con restricciones

$$\max_{\beta_0, \beta_1, \dots, \beta_p, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n} M \quad (4.1.13)$$

$$\text{sujeto a } \sum_{j=1}^p \beta_j^2 = 1 \quad (4.1.14)$$

$$y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}) \geq M(1 - \varepsilon_i) \quad (4.1.15)$$

$$\varepsilon_i \geq 0, \quad \sum_{i=1}^n \varepsilon_i \leq C, \quad (4.1.16)$$

donde  $C$  es un parámetro de sintonización no negativo. Como en (4.1.10),  $M$  es la anchura del margen y se busca que este valor sea tan grande como sea posible. En (4.1.15),  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ , son variables de holgura que permite que cada observación individualmente este en el lado incorrecto del margen o del hiperplano.

Cuando se resuelve el problema de optimización y se encuentran los parámetros  $\beta_0, \beta_1, \dots, \beta_p$ , para clasificar una observación prueba,  $x_*$ , lo haremos como se hacia con el clasificador de máximo margen, determinando el signo de la función

$$f(x^*) = \beta_0 + \beta_1 x_1^* + \beta_2 x_2^* + \dots + \beta_p x_p^*,$$

y de este modo ver a que lado del hiperplano se encuentra.

Veamos unas observaciones sobre el problema de optimización. Se puede ver que la variable de holgura  $\varepsilon_i$  dice en que parte está ubicada la observación  $i$ -ésima con respecto al hiperplano y al margen. Si  $\varepsilon_i = 0$ , la observación está en el lado correcto del margen. Si  $\varepsilon_i > 0$ , la observación está en el lado incorrecto del margen. Si  $\varepsilon_i > 0$ , la observación está en el lado incorrecto del hiperplano.

Ahora, se hablará del papel que juega la variable de sintonización  $C$ . La restricción (4.1.16) muestra que el parámetro  $C$  acota la suma de las variables de holgura  $\varepsilon_i$ 's. Este va a determinar el número y la severidad con la que se permitirá a las observaciones violar el margen o el hiperplano. Se puede pensar a  $C$  como un coste que entre las  $n$  observaciones se pueden utilizar para poder violar el margen. Si  $C = 0$  no se permite violar el margen, por lo que  $\varepsilon_1 = \varepsilon_2 = \dots = \varepsilon_n = 0$  y el problema de optimización (4.1.13) - (4.1.16) se convierte en el problema (4.1.10) - (4.1.12), es decir, que se resolverá clasificación de máximo margen (El hiperplano existe solo si las clases son separables). Cuando  $C > 0$ , no más de  $C$  observaciones

van a estar en el lado incorrecto del hiperplano, ya que  $\varepsilon_i > 1$  para dichas observaciones y por (4.1.16) la suma de los  $\varepsilon_i$ 's no puede superar  $C$ . Observemos que si  $C$  se aumenta se tiene mayor tolerancia para violar el margen, por lo que el margen será más amplio. De forma similar, si  $C$  se disminuye se tiene menor tolerancia para violar el margen, por lo que el margen será más estrecho.

El problema de optimización (4.1.13) - (4.1.16) tiene una propiedad interesante, la selección del hiperplano solo va a depender de las observaciones que caen sobre el margen o las que lo violan, y por lo tanto el clasificador solo se verá afectado por estas observaciones. Las observaciones que caen en el lado correcto del margen no afectan al clasificador de soporte vectorial, si alguna de estas observaciones es movida sin cruzar el margen, el hiperplano obtenido será el mismo que el anterior. Los datos que caen sobre el margen o en el lado incorrecto son conocidos como **vectores de soporte** y afectan directamente al hiperplano.

Cuando el valor del parámetro de sintonización  $C$  es grande, se obtiene un margen más amplio y muchas más observaciones pueden violar el margen, por lo que el número de vectores de soporte será grande. Esto implica que muchas observaciones determinan el hiperplano separador. A diferencia, de cuando  $C$  es pequeño, el margen será más estrecho, y el número de vectores de soporte será menor.

Como para el clasificador de soporte vectorial, la selección del hiperplano separador solo depende de un subconjunto de los datos, es decir, de los vectores de soporte, entonces este clasificador será un método bastante robusto para las observaciones que están lejos del hiperplano.

## 4.2. Exploración de parámetros clasificadores

En la figura se muestra el valor que se obtuvo de cada uno de los parámetros en los 80 pacientes. En estas gráficas se observa que los parámetros no toman un valor distintivo para cada una de las categorías y por tanto un sólo parámetro no basta para hacer la clasificación. Los puntos azules representan a los pacientes sanos, los puntos verdes representan a los pacientes con IFG, los puntos naranja a los pacientes con IGT, los puntos cafés a los pacientes con IFG e IGT, y los puntos rojos a los pacientes con diabetes tipo 2.



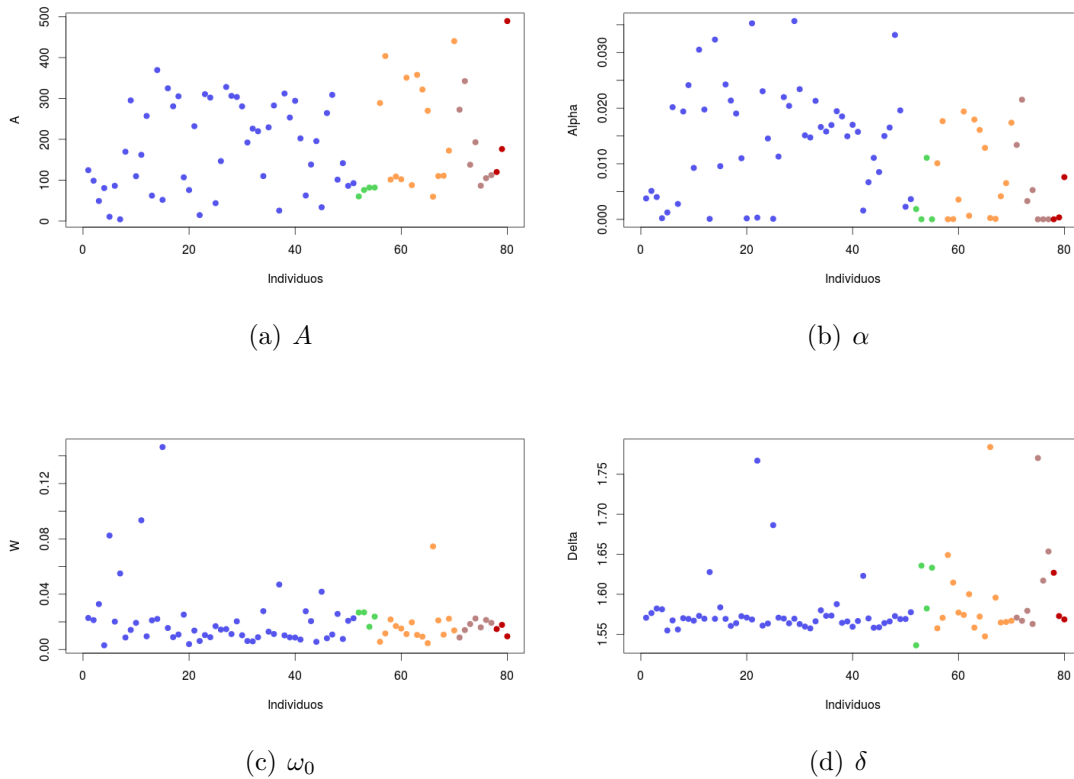


Figura 4.1: Parámetros obtenidos en los 80 pacientes

Luego, se realizaron gráficas dos a dos de los cuatros parámetros,  $A$ ,  $\alpha$ ,  $\omega$  y  $\delta$  que se encontraron en el capítulo anterior. En la figura 4.2 se observan los parámetros que enfrentados no dan un buena gráfica de separación ya que los pacientes sanos y los que tienen alguna alteración o están enfermos se mezclan entre sí.

### 4.3. Propuesta de parámetros clasificadores

En la figura 4.3 se muestra  $A$  contra  $\alpha$  de los 80 datos que se tienen. Se observó que los parámetros  $A$  y  $\alpha$  eran buenos parámetros de clasificación, donde  $A$  representa la amplitud y  $\alpha$  la rapidez de caída de las curvas encontradas.

Se aplicó el clasificador de soporte vectorial, se etiquetó con 1 a los pacientes sanos y con -1 a los pacientes que presentan alguna alteración o tienen diabetes

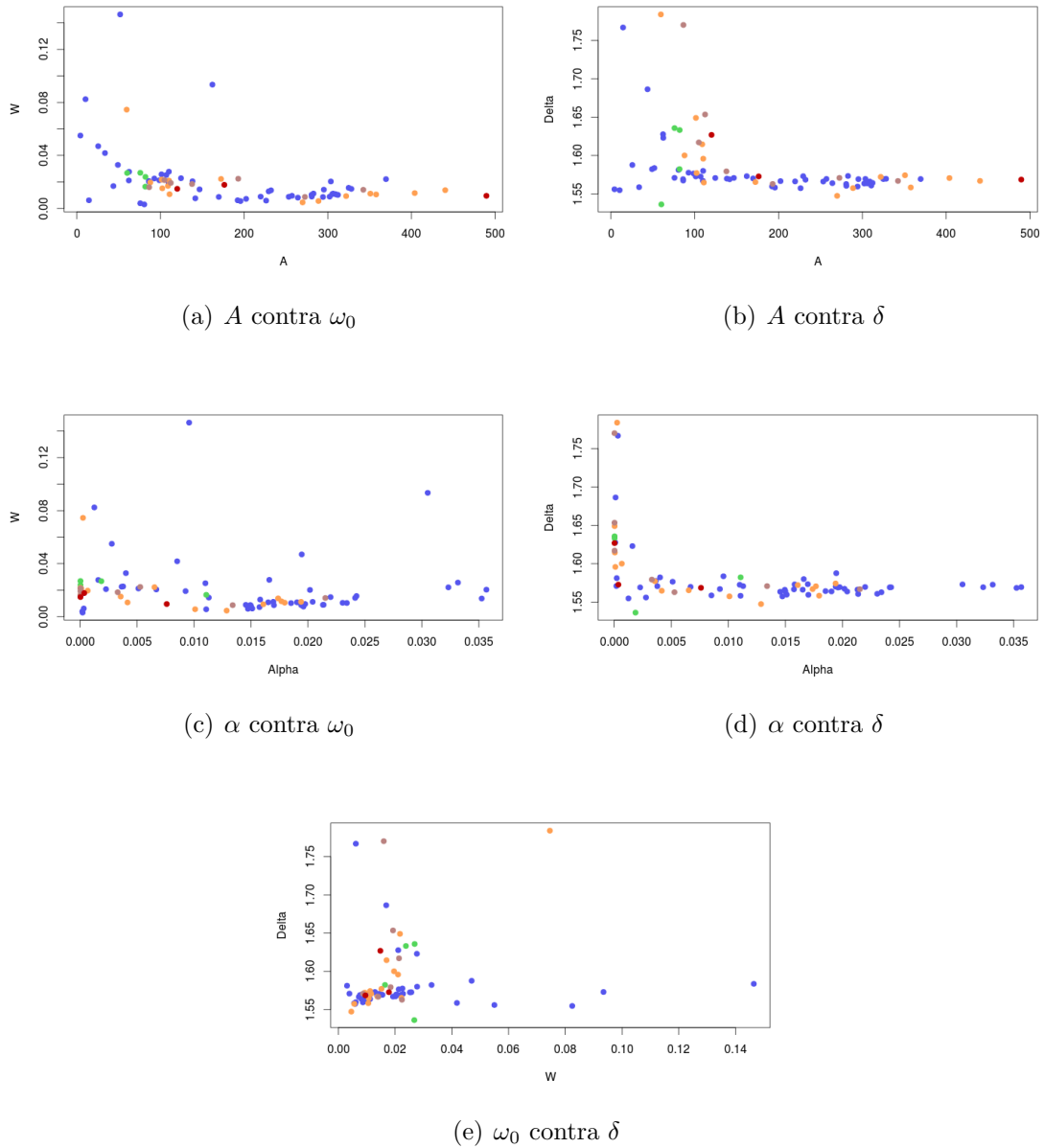
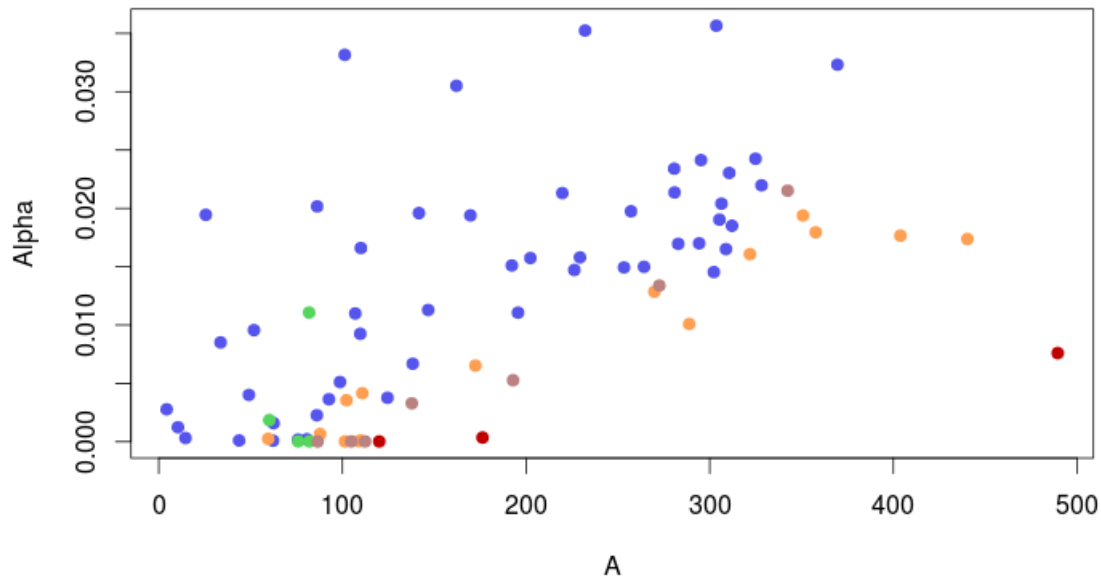


Figura 4.2: Parámetro contra parámetro de los 80 datos.

tipo 2. La figura 4.4 muestra las dos regiones de separación donde la región azul es la zona de personas sanas y la región rosada es la zona de personas enfermas o con alguna alteración.

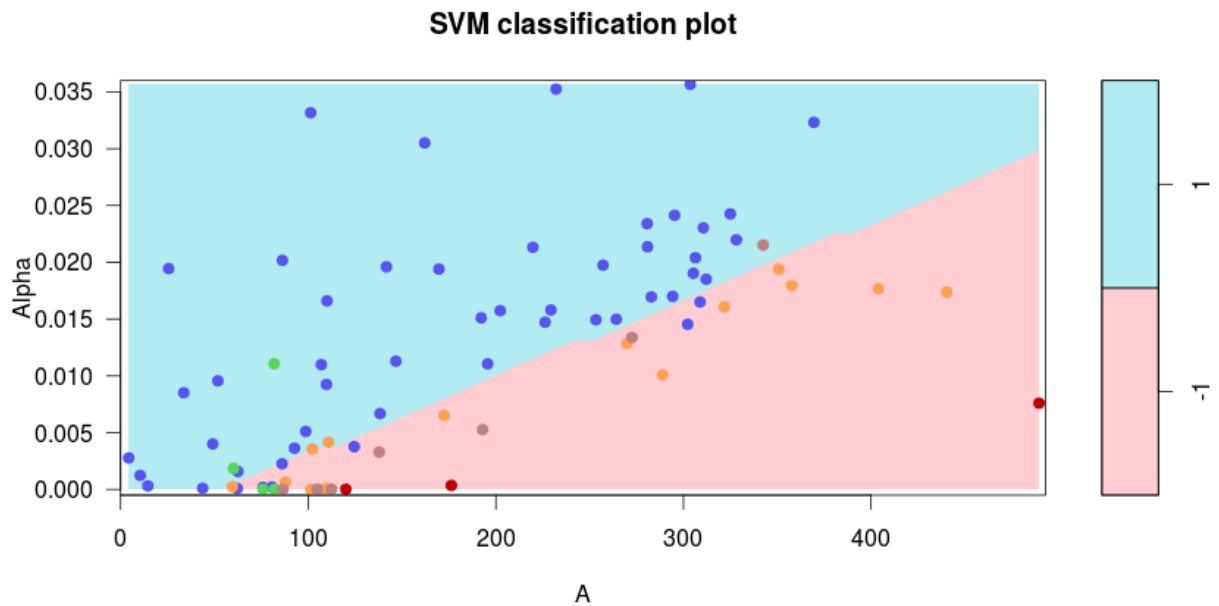


**Fuente:** Elaboración propia.

Figura 4.3: A contra  $\alpha$  de los 80 datos.

De la figura 4.4 se observa que se obtuvo una buena clasificación de los datos. Se tiene que:

- 45 de 51 sanos bien clasificados.
- 23 de 29 enfermos bien clasificados.
- 6 de 51 sanos mal clasificados.
- 2 de 4 con alteración de glucosa de ayuno (IFG) mal clasificados.
- 3 de 15 con intolerancia a la glucosa IGT mal clasificados.
- 1 de 7 con alteración de glucosa de ayuno y con intolerancia a la glucosa (IFG - IGT) mal clasificado.



**Fuente:** Elaboración propia.

Figura 4.4: Clasificación de personas sanas y personas enfermas o con alguna alteración

En la figura 4.4 se identifica que los pacientes que quedaron mal clasificados están cerca de la línea de separación, esto nos sugiere cuales son los casos pre-diabéticos. De esta manera, se puede recomendar con anterioridad a estos pacientes tomar medidas preventivas, como por ejemplo mejorar sus hábitos alimenticios y realizar ejercicio, para evitar que padezcan a largo plazo de diabetes. Por lo tanto, el algoritmo de clasificación nos podría dar una herramienta para la detección de pacientes pre-diabéticos, siendo de gran utilidad en la medicina, ya que de esta manera se puede prevenir la enfermedad en lugar de tratarla. Por otra parte, los resultados arrojados en la figura 4.4 nos muestra que se puede dar continuidad a la investigación con mejores muestreos y un mayor número de datos.

# Capítulo 5

## Conclusiones y trabajo futuro

En este trabajo se buscaba hacer el ajuste del modelo básico de interacción glucosa/insulina a datos reales de OGTT, como un problema de estimación de parámetros. A partir de la información obtenida se pretendía hacer una búsqueda para encontrar los mejores parámetros que permitiera la clasificación para el diagnóstico de la diabetes.

En la estimación de parámetros se encontraron buenos ajustes en un alto porcentaje de pacientes. En la figura 5.1 se muestra unos ejemplos de pacientes sanos y con alteraciones de glucosa a los que no se les pudo encontrar parámetros que resultaran en un buen ajuste. Esto puede ocurrir por que el modelo básico de interacción glucosa/insulina con el que se trabajó, no puede explicar todas las mediciones de los datos obtenidos de la prueba oral de tolerancia a la glucosa, y estos fueron los mejores resultados que dió este modelo.

A partir de los parámetros que se obtuvieron en la estimación bayesiana se pudo encontrar dos buenos parámetros clasificadores como lo fueron  $A$ , la amplitud de la curva y  $\alpha$  la rapidez de caída de la curva. Aunque la gráfica de separación que se obtuvo no generaba dos grupos linealmente separables, si se pudo aplicar el clasificador de soporte vectorial, con el cual se puede encontrar un hiperplano separador que permite que algunos datos esten en la zona incorrecta. Lo bueno de la clasificación que se encontró fue que los datos que quedaron en el lado incorrecto estuvieron cerca de la línea separadora. Las personas sanas que quedaron mal clasificadas son personas a las que se le puede recomendar tener una mejor alimentación y que realicen ejercicio para prevenir que padezcan la diabetes.

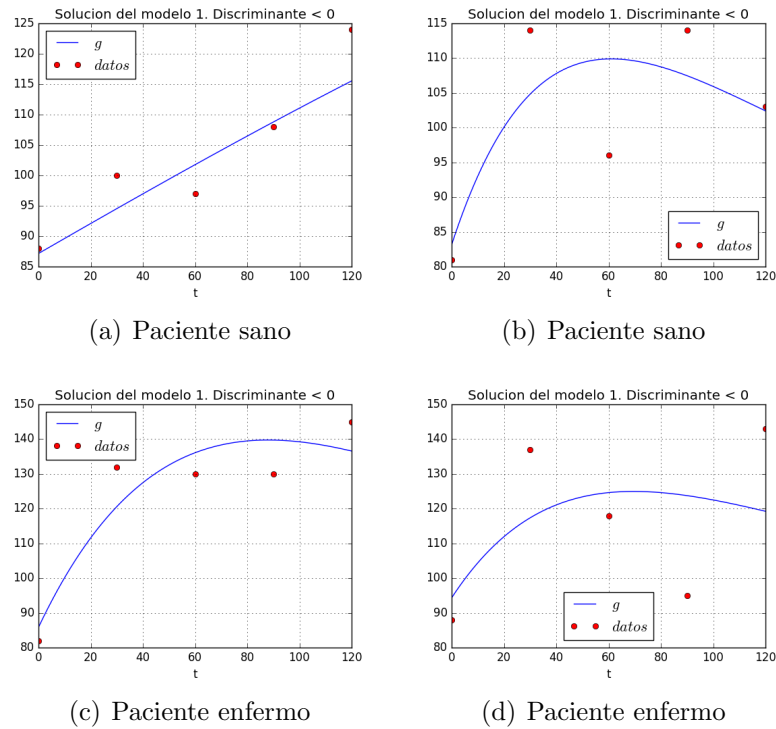


Figura 5.1: Ajustes de pacientes por mejorar.

Entre los trabajos futuros que se podrían realizar serian probar otros modelos más complejos de interacción glucosa/insulina, para ver si estos modelos logran explicar mejor los datos con los que se trabajo. Una vez seleccionado un modelo de glucosa/insulina, se pasaría a estimar los parámetros del modelo con la estimación bayesiana, donde la concentración de glucosa se obtendría resolviendo un sistema de ecuaciones. Con esta información se pasaría a la exploración de parámetros clasificadores, tratando de buscar la amplitud y la rapidez de caída de la curva, a partir de la información obtenida con la estimación bayesiana. Si esto sucede, se identificaría si estos mismos parámetros vuelven a ser buenos parámetros de separación o de lo contrario se deben buscar unos nuevos.

Por otra parte, sería bueno obtener más datos de pacientes que se hayan realizado la prueba oral de tolerancia a la glucosa para tener mejores resultados en el momento de realizar la clasificación de los datos.



# Apéndice A

## Datos

Paciente	G(0)	G(30)	G(60)	G(90)	G(120)
1	89	159	186	167	120
2	75	123	148	131	106
3	95	130	133	100	74
4	88	100	97	108	124
5	79	87	76	95	78
6	81	116	104	94	78
7	99	102	98	94	100
8	85	110	114	101	103
9	95	154	155	107	128
10	81	129	135	132	106
11	83	106	68	90	75
12	86	126	135	104	118
13	87	115	147	149	123
14	91	180	139	109	103
15	93	57	113	107	79
16	97	169	161	122	125
17	86	141	105	121	114
18	90	154	143	128	131
19	85	137	142	114	89
20	89	108	86	127	122
21	93	130	101	115	97
22	98	96	93	105	108
23	94	160	111	142	117
24	86	148	138	151	133
25	92	103	121	146	128
26	92	136	148	144	129
27	90	164	159	128	119



Paciente	G(0)	G(30)	G(60)	G(90)	G(120)
28	87	151	136	121	125
29	91	154	120	103	105
30	95	150	115	132	116
31	81	114	96	114	103
32	78	117	99	107	110
33	96	132	119	122	113
34	90	135	140	94	94
35	93	144	162	137	131
36	87	138	164	124	130
37	82	96	88	76	85
38	85	151	120	157	107
39	88	136	132	140	124
40	87	147	126	130	129
41	89	119	122	113	117
42	92	125	163	117	89
43	85	152	170	160	123
44	92	127	119	127	130
45	86	112	96	78	73
46	94	141	143	131	138
47	93	159	158	148	139
48	82	107	101	77	90
49	83	102	103	93	98
50	89	137	159	157	128
51	99	148	175	155	125

Cuadro A.1: Pacientes sanos.

Paciente	G(0)	G(30)	G(60)	G(90)	G(120)
52	100	149	144	140	92
53	107	152	182	168	100
54	106	129	151	128	130
55	101	144	180	181	124

Cuadro A.2: Pacientes con alteración de glucosa en ayuno (IFG).

Paciente	G(0)	G(30)	G(60)	G(90)	G(120)
56	82	132	130	130	145
57	89	167	193	136	151
58	92	133	195	191	147
59	89	128	185	197	186
60	88	126	158	159	155
61	99	155	191	120	148
62	82	120	170	156	143
63	87	171	145	136	144
64	95	146	173	134	146
65	88	137	118	95	143
66	90	175	84	136	141
67	96	152	206	197	161
68	82	117	132	149	147
69	100	192	207	192	133
70	88	198	205	161	155

Cuadro A.3: Pacientes con intolerancia a la glucosa (IGT).

Paciente	G(0)	G(30)	G(60)	G(90)	G(120)
71	109	155	175	160	160
72	108	186	181	139	150
73	100	161	208	198	177
74	104	213	232	219	145
75	119	116	183	217	197
76	102	155	198	211	158
77	103	137	216	213	190

Cuadro A.4: Pacientes con alteración de glucosa en ayuno e intolerancia a la glucosa (IFG - IGT).

---

Paciente	G(0)	G(30)	G(60)	G(90)	G(120)
78	115	151	201	241	228
79	169	258	320	341	311
80	186	300	353	372	367

Cuadro A.5: Pacientes con diabetes de tipo 2 (T2DM).



# Bibliografía

- [1] Ackerman, E., Gatewook, L., Rosevear, J., & Molnar, Gl. (1969). Blood glucose regulation and diabetes. *Concepts and Models of Biomathematics*, pp.131-156.
- [2] Palumbo, P., Ditlevsen, S., Bertuzzi, A., & De Gaetano, A. (2013). Mathematical modeling of the glucose-insulin system: A review. *Mathematical biosciences 244(2)*, pp. 69-81.
- [3] Christen, J. A., Capistrán, M., Monroy, A., Alavez, S., Vargas, S. Q., Flores-Arguedas, H. A. & Kuschinski, N. (2016). A Diabetes minimal model for Oral Glucose Tolerance Tests. *arXiv preprint arXiv:1601.04753*.
- [4] Davies, M. J.(1976). A Differential Model of Diabetes Mellitus. *Mathematical Modelling Applied Mathematics Department, University College of Wales*, pp. 116-126.
- [5] Kaipio, J., & Somersalo, E. (2006). *Statistical and computational inverse problems* (Vol. 160). Springer Science & Business Media.
- [6] Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. (2013). emcee:the MCMC hammer. *Publications of the Astronomical Society of the Pacific, 125(925)*, 306.
- [7] Gilks, W. R., Richardson, S., & Spiegelhalter, D. (Eds.). (1995). *Markov chain Monte Carlo in practice*. CRC press.
- [8] James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112). New York: springer.
- [9] World Health Organization. (1999). Definition, diagnosis and classification of diabetes mellitus and its complications: report of a WHO consultation. Part 1, Diagnosis and classification of diabetes mellitus.

- [10] Organización mundial de la salud. (2016). Informe mundial sobre la diabetes. Recuperado en: [http://apps.who.int/iris/bitstream/10665/204877/1/WHO\\_NMH\\_NVI\\_16.3\\_spa.pdf](http://apps.who.int/iris/bitstream/10665/204877/1/WHO_NMH_NVI_16.3_spa.pdf)
- [11] World Health Organization (WHO). Diabetes. Recuperado en: <http://www.who.int/mediacentre/factsheets/fs312/en/>