

# Proyecto de Análisis Estadístico de Datos Biomédicos

334753 Alberto Abaroa Villanueva  
338386 Clifford Kennedy Negrete

05 de Junio 2017

## 1 Resumen

Se obtuvieron datos de 500 pacientes ambulatorios del Hospital General de Ensenada con respecto a 4 variables, su nivel de glucosa (glu), edad (eda), sexo (sex) y su índice de masa corporal (imc). Dichos datos son utilizados para clasificar a los pacientes en tres grupos (niveles diabéticos): Sin diabetes (SD), pre-diabetes (PD) y con diabetes (CD); los grupos son evaluados entre sí con respecto a sus medias. Así como también se evaluó la influencia del imc en la ocurrencia de la diabetes, a través de un análisis ANOVA.

## 2 Introducción

La diabetes se está convirtiendo rápidamente en la epidemia del siglo XXI y en un reto de salud global. Estimaciones de la Organización de la Salud (OMS) indican que a nivel mundial, de 1995 a la fecha casi se ha triplicado el número de personas que viven con diabetes, con cifra actual con más de 415 millones de personas con diabetes [Cho et al., 2015]. De acuerdo con la Federación Internacional de Diabetes, China, India, Estados Unidos, Brasil, Rusia y México, son (en ese orden) los países con mayor número de diabéticos [Hernández-Ávila et al., 2013].

La diabetes es una enfermedad crónica de causas múltiples. En su etapa inicial no produce síntomas y cuando se detecta tardíamente y no se trata

adecuadamente ocasiona complicaciones de salud graves como infarto del corazón, ceguera, falla renal, amputación de las extremidades inferiores y muerte prematura [Hernández-Ávila et al., 2013]. Se ha estimado que la esperanza de vida de individuos con diabetes se reduce entre 5 y 10 años [Paneni et al., 2013]. En México la edad promedio de las personas que murieron por diabetes en 2010 fue de 66.7 años, lo que sugiere una reducción de 10 años.

Los estilos de vida poco saludables son altamente prevalentes entre niños, adolescentes y adultos mexicanos, propiciando un aumento importante de la obesidad y sobre peso, principal factor de riesgo modificable de la diabetes. Así, la prevalencia de la diabetes en esta población ha incrementado sustancialmente en las últimas décadas: en 1993 la prevalencia de los diabéticos con diagnóstico conocido en población mayor de 20 años fue de 4.0%, mientras que en 2000 y 2007 se describió una prevalencia del 5.8 y 7 %, respectivamente [Hernández-Ávila et al., 2013].

## **2.1 Tipos de diabetes**

Es bien conocido que la diabetes es una afección crónica que se desencadena cuando el organismo pierde su capacidad de producir suficiente insulina o de utilizarla con eficacia; debido a esto las personas que padecen de diabetes no absorben glucosa adecuadamente, de modo que ésta queda circulando en la sangre (hiperglucemia) y dañando los tejidos con el paso del tiempo. Existen tres tipos principales de diabetes:

- Diabetes tipo I
- Diabetes tipo II
- Diabetes mellitus gestacional (DMG)

### **2.1.1 Diabetes tipo I**

La diabetes tipo I es causada por una reacción autoinmune, en la que el sistema de defensas del organismo ataca las células productoras de insulina del páncreas. Como resultado, el organismo deja de producir la insulina que necesita. La razón por la que esto sucede continúa en incertidumbre. La

enfermedad puede afectar a personas de cualquier edad, pero suele aparecer en niños o jóvenes adultos. Las personas con esta forma de diabetes necesitan inyecciones de insulina a diario con el fin de controlar sus niveles de glucosa en sangre. Sin insulina, una persona con diabetes tipo I morirá [López et al., 2006].

Las personas con diabetes tipo I pueden llevar una vida típica y saludable mediante una combinación de terapia diaria de insulina, estrecha monitorización y ejercicio físico habitual

La diabetes tipo I suele desarrollarse repentinamente y podrían presentarse síntomas como:

- Sed anormal y sequedad de boca
- Micción frecuente
- Cansancio extremo
- Apetito constante
- Pérdida de peso repentina
- Lentitud en la curación de las heridas
- Infecciones recurrentes
- Visión borrosa

### **2.1.2 Diabetes tipo II**

La diabetes tipo II es el tipo más común de diabetes. Suele aparecer en adultos, pero cada vez, hay más casos de niños y adolescentes. En el diabetes tipo II, el organismo puede producir insulina pero, o no es suficiente, o el organismo no responde a sus efectos, provocando una acumulación de glucosa en la sangre.

Las personas con diabetes tipo II podrían pasar mucho tiempo sin saber de su enfermedad debido a que los síntomas podrían tardar años en aparecer o en reconocerse, tiempo durante el cual el organismo se va deteriorando

debido al exceso de glucosa en sangre [López et al., 2006].

Aunque las razones para desarrollar diabetes tipo II aún no son conocidas, son varios los factores de riesgo relevantes, tales como:

- Obesidad
- Origen étnico
- Edad avanzada
- Mala alimentación
- Falta de actividad física
- Antecedentes familiares de diabetes

En contraste con las personas con diabetes tipo I, la mayoría de quienes tienen diabetes tipo II no suelen necesitar dosis diarias de insulina para sobrevivir. Sin embargo, para controlar la afección se podría recetar insulina unida a una medicación oral, una dieta sana y el aumento de la actividad física [López et al., 2006].

### 3 Planteamiento del problema

Se obtuvo una muestra de 500 pacientes diabéticos donde estos son caracterizados en distintos grupos debido a sus deficiencias o diagnósticos previos, por lo que se tiene dividido lo que se espera obtener, como los siguientes análisis que se buscan practicar como la diferencia en las muestras con diabetes y sin diabetes, que tanto influye el IMC en el diabetes, ANOVA de los grupos y edades, si hay sobre peso que tanta cantidad se tienen con diabetes (regresión lineal); todo esto a través de la herramienta computacional **R**.

### 4 Objetivos

1. Obtener datos de estadística descriptiva de cada tipo de dato (edad, IMC, glucosa)

2. Determinar la presencia o ausencia de una relación lineal del IMC y edad con el diabetes
3. Análisis ANOVA para determinar las diferencias entre las muestras

## 5 Descripción de los datos

Mediante un estudio de tipo descriptivo, a través de un método de muestreo no probabilístico se adquirieron 500 historias clínicas (206 hombres y 294 mujeres) de pacientes ambulatorios que asistieron al laboratorio clínico del Hospital General de Ensenada, con edades comprendidas entre 18 y 96 años, de las cuales se tomó la información de: nivel de glucosa, edad, género e IMC.

Los pacientes se clasificaron por grupos de edad, 85 pacientes menores o iguales a 30 años, 96 pacientes entre 31 y 35 años, 76 pacientes entre 36 y 45 años, 103 pacientes entre 46 y 55 años y por último 140 pacientes con 56 años o más.

## 6 Análisis estadístico

Los resultados obtenidos a través del análisis descriptivo de las edades mostró que la edad promedio de los pacientes es de 57 años, la desviación estándar ( $\sigma$ ) en la edad, es de 23.65 años.

Para el análisis descriptivo de la glucosa se obtuvo que la cantidad promedio fue de 181.65 mg/dl, la desviación estándar ( $\sigma$ ) es de 66.64 mg/dl.

Por último tenemos el análisis descriptivo de los datos de IMC, en donde se tiene que el promedio en los pacientes es de  $26.43 \text{ kg/m}^2$  con una desviación estándar ( $\sigma$ ) de  $5.14 \text{ kg/m}^2$ .

Valor	Edad (años)	Glucosa (mg/dl)	IMC ( $kg/m^2$ )
Min	18	70	17
Max	96	299	36
Promedio	57.7	181.65	26.43
$\sigma$	23.65	66.64	5.14
$\sigma^2$	559.54	4,441.11	26.49

Los datos se clasificaron según su nivel de glucosa en sin diabetes (SD,  $\leq 99$  mg/dl), pre-diabetes (PD, 100 - 125 mg/dl) y con diabetes (CD,  $\geq 126$  mg/dl), con lo que se hizo la siguiente tabla:

Edades	SD	PD	CD	<b>Total</b>
$\leq 30$	08	15	62	85
31 - 45	14	14	68	96
46 - 60	13	09	54	76
61 - 75	17	07	79	103
$\geq 76$	15	13	112	140
<b>Total</b>	67	58	375	<b>500</b>

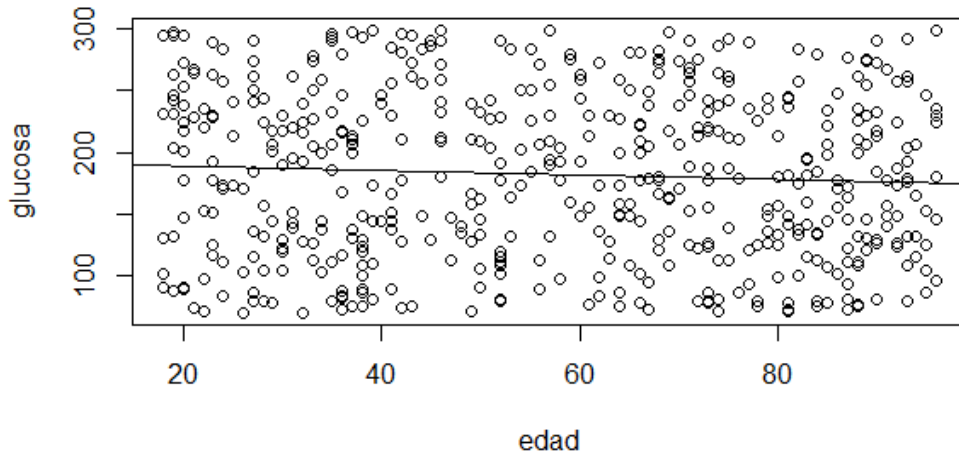
La tabla de la estadística de prueba se muestra a continuación:

Fuente de variación	Suma de cuadrados	G. L.	Cuadrado medio	Razón de variación
Tratamientos	2,605.79	02	1,302.89	0.3302
Bloques	269.46	04	67.36	
Residuales	31,566.85	08	3,945.86	
Total	34,442.10	14		

Utilizando un  $\alpha = 0.05$ , obtenemos un valor crítico de 3.2873 en la distribución F, por lo que la hipótesis nula, la cual dice que los efectos de los tratamientos son similares, se acepta, ya que la razón de variación es menor a nuestro valor crítico.

Una vez que se han calculado las sumas de cuadrados, las medias cuadráticas, los grados de libertad y la F, se procede a elaborar la tabla de información ANOVA.

Figure 1: Regresión lineal de los valores de glucosa y edad

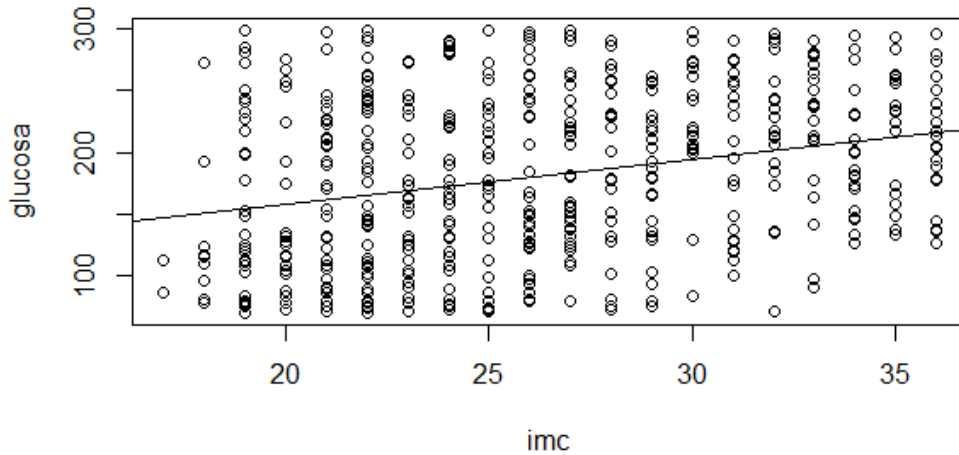


Fuente de variación	Suma de cuadrados	G. L.	Cuadrado medio	Razón de variación
Entre muestras	13,022.96	02	6,514.45	36.59
Dentro de muestras	2,136.40	12	178.03	
Total	15,165.33	14		

Utilizando un  $\alpha = 0.05$ , obtenemos un valor crítico de 3.2873 en la distribución F, por lo que la hipótesis nula, la cual dice que nuestras medias son iguales, se rechaza, ya que la razón de variación es mayor a nuestro valor crítico.

Por último, se realizó en **R** un análisis de regresión lineal simple para determinar si existe una asociación lineal entre la glucosa y el IMC, así como también entre la glucosa y la edad, tal y como se puede apreciar en la figura 1 y 2.

Figure 2: Regresión lineal de los valores de glucosa e IMC



## 7 Conclusiones

Desde el punto de vista estadístico, se determinó que tanto la edad, clasificada en categorías, como el IMC resultaron ser significativos con respecto a los niveles de glucosa, aunque el análisis de regresión lineal simple no evidenció relación alguna (lineal) entre los niveles de glucosa y la edad del paciente. A mayor edad del paciente no podemos asegurar que su nivel de glucosa se incrementará en una cierta cantidad promedio o por el contrario, garantizar que esta disminuya. En consecuencia, la edad y el nivel de glucosa están relacionados estadísticamente a reserva de una forma lineal.

Cabe mencionar que los valores utilizados en el trabajo son considerados una muestra representativa de la ciudad de Ensenada, sin embargo, no es posible generalizar los resultados debido a las diferencias que existen en las distintas regiones con respecto al clima, alimentación, costumbres, etc.



## References

- [Cho et al., 2015] Cho, N., Whiting, D., Guariguata, L., Montoya, P., and Forouhi, N. (2015). *Atlas de la Diabetes de la FID*. Federation international Diabetes, 7 edition.
- [Hernández-Ávila et al., 2013] Hernández-Ávila, M., Gutiérrez, J. P., and Reynoso-Noverón, N. (2013). Diabetes mellitus en México. El estado de la epidemia. *Salud pública de México*, 55(1):129–136.
- [López et al., 2006] López, J., Luna, M., Navarro, M., Pallardo, L., and Sancho, M. A. (2006). Guía asistencial de diabetes mellitus y embarazo. *Revista Española*, 22(1):73–87.
- [Paneni et al., 2013] Paneni, F., Beckman, J. A., Creager, M. A., and Cosentino, F. (2013). Diabetes and vascular disease: Pathophysiology, clinical consequences, and medical therapy: Part i.