

# **Métodos de Computación Científica**

**Proyecto de Cursado - Los Lirios de Fisher**



## **Integrantes:**

Fenoglio, Rubén Darío – LU 78734

García Suarez, José - LU 74964

Garat, Fabiana – LU 89108

López, Juan Matías – LU 83871

## Ejercicio 1

a)

Contiene 150 casos en total. Son 3 clases de 50 instancias cada una, donde cada clase representa un tipo de flor iris (Iris-setosa, Iris-versicolor, Iris-virginica).

b)

Hay 4 variables numéricas incluidas en los datos, en orden son:

- Longitud del sépalo en centímetros
- Ancho del sépalo en centímetros
- Longitud del pétalo en centímetros
- Ancho del pétalo en centímetros

Todas estas variables son variables continuas.

c)

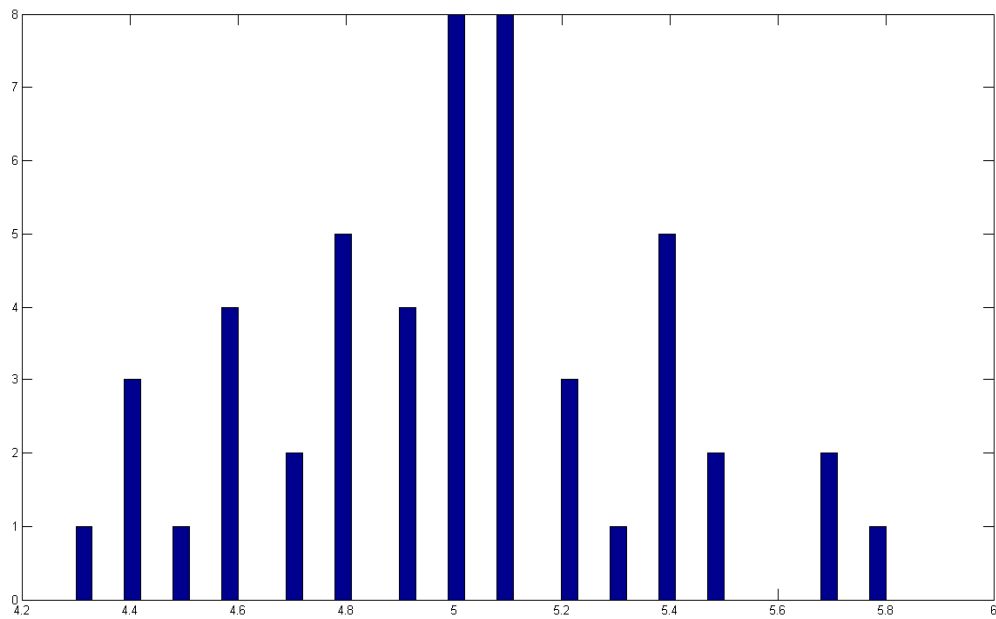
Hay solo una variable categórica. Las posibles categorías son los nombres de cada una de las 3 flores:

- Iris-Setosa
- Iris-Versicolor
- Iris-Virginica

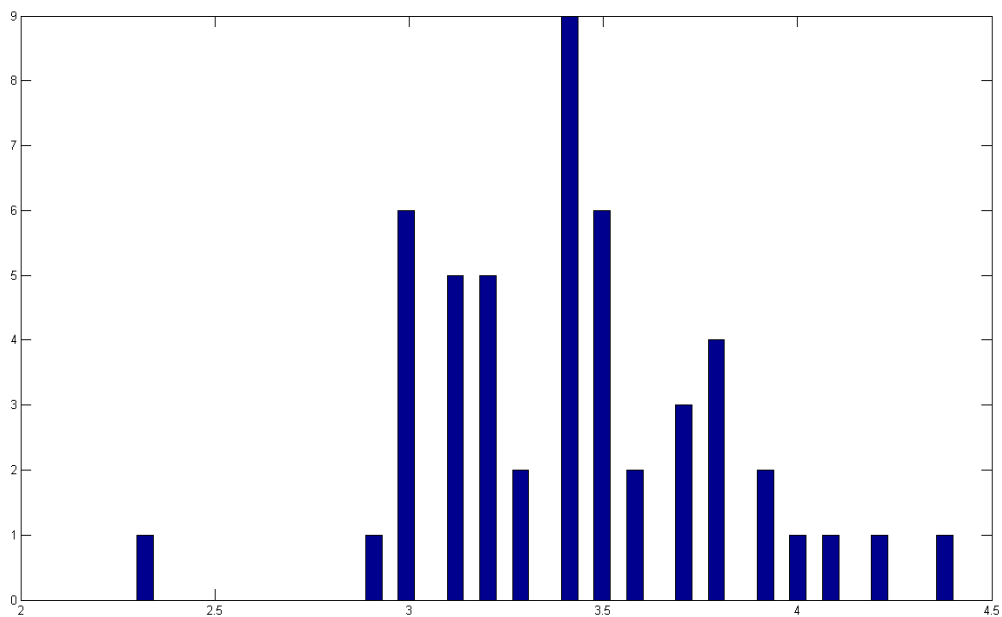
## Ejercicio 2

a) Usando la función predefinida de Matlab “hist(x)” y obteniendo previamente los vectores x correspondientes a la longitud y ancho del sépalo, y longitud y ancho del pétalo, se pueden obtener fácilmente los histogramas de frecuencia para estos cuatro datos de cada una de las diferentes flores. Usando “mean(x)” y “median(x)” se obtienen la media y la mediana respectivamente.

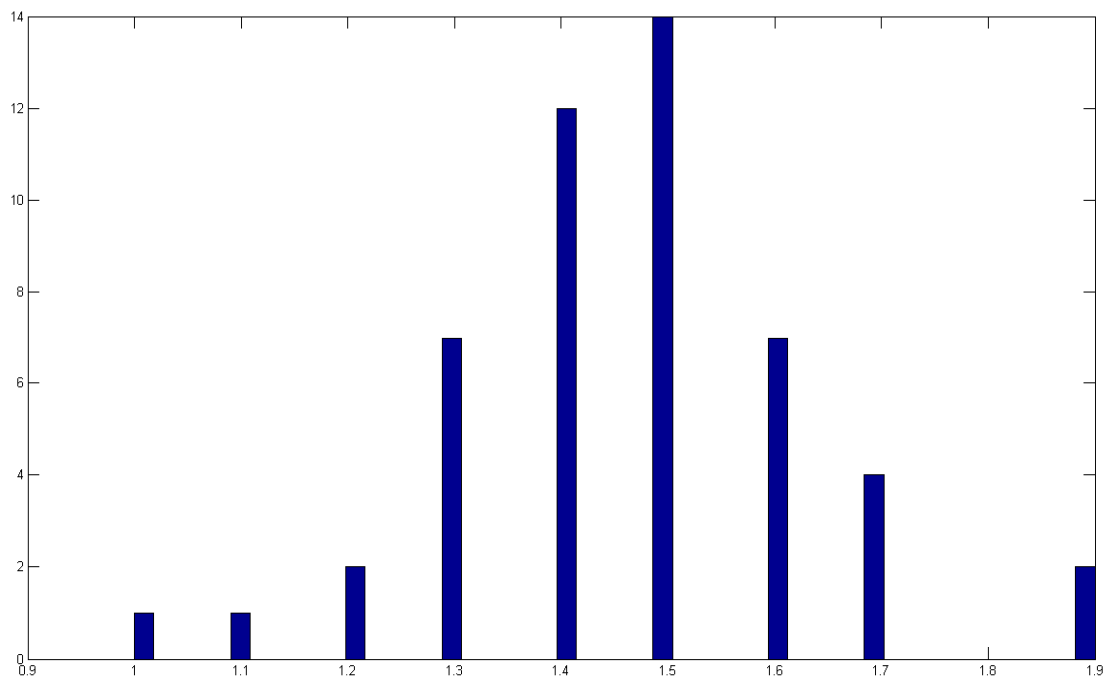
### Iris-Setosa:



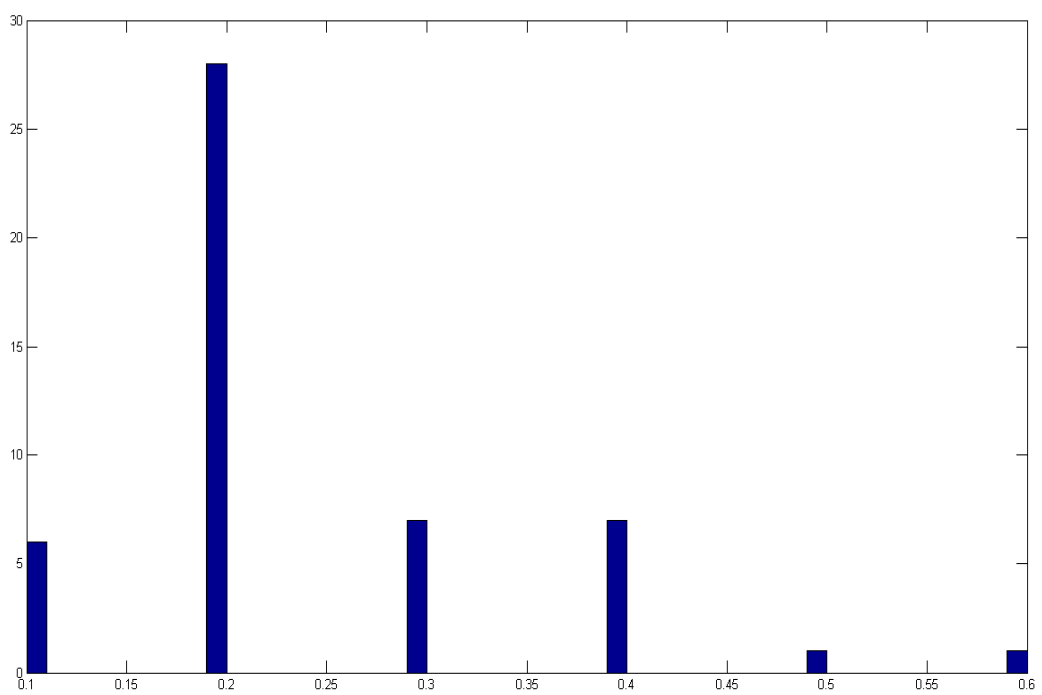
*Longitud del sépalo. Media: 5.0060, Mediana 5.0000*



*Ancho del sépalo. Media: 3.4180, Mediana 3.4000*

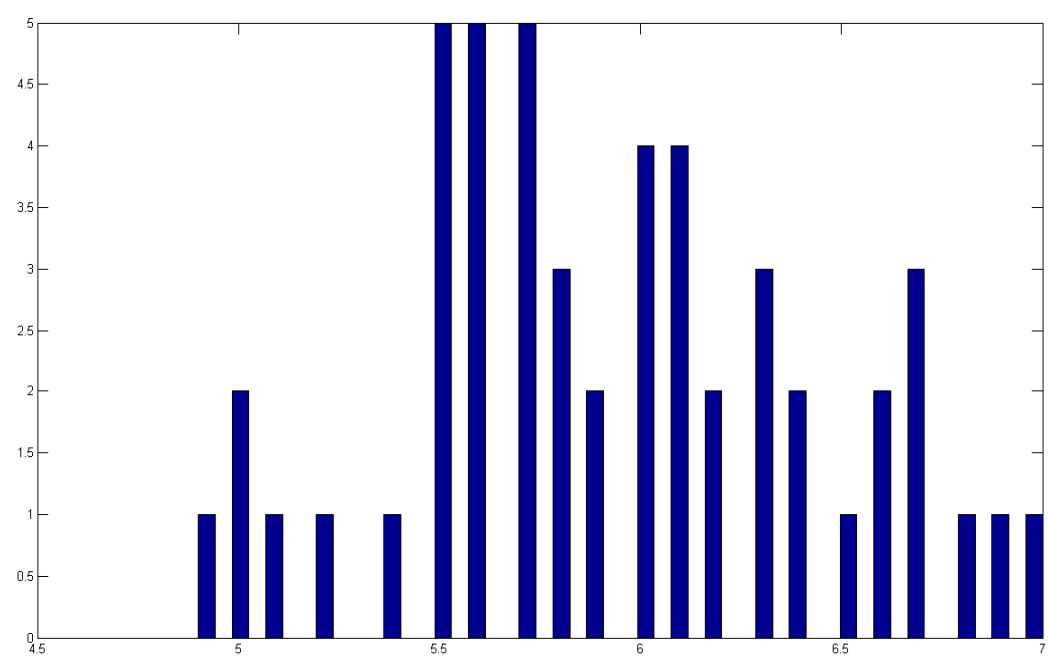


*Longitud del pétalo. Media 1.4640, Mediana 1.5000*

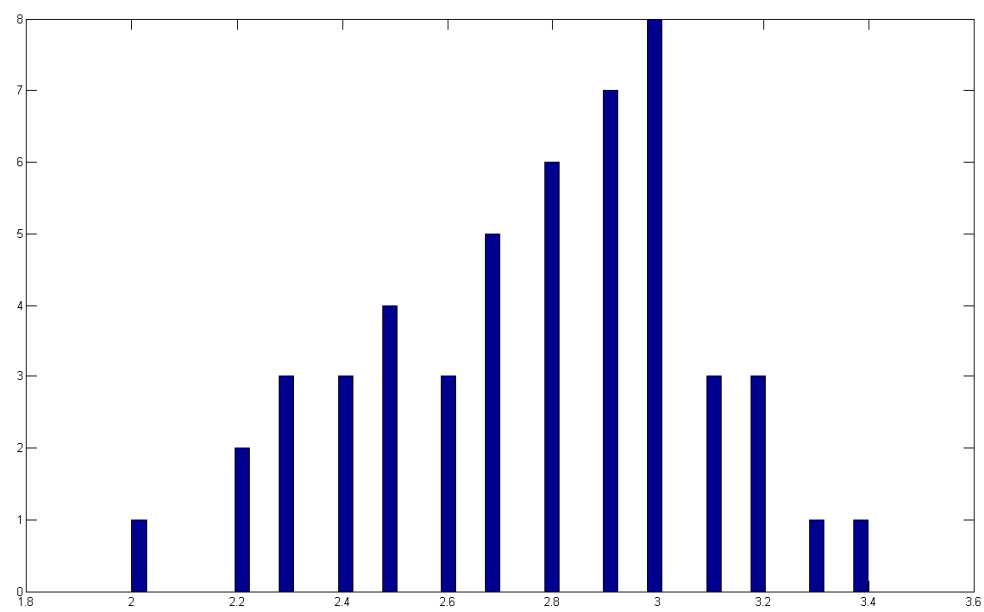


*Ancho del pétalo. Media 0.2440, Mediana 0.2000*

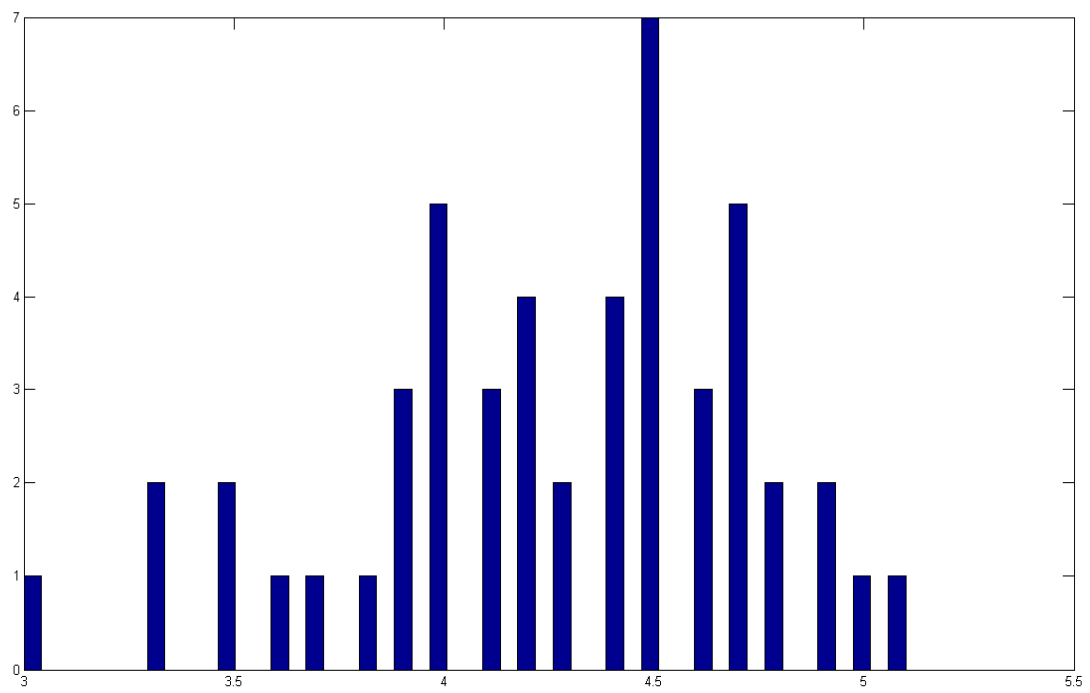
**Iris-Versicolor:**



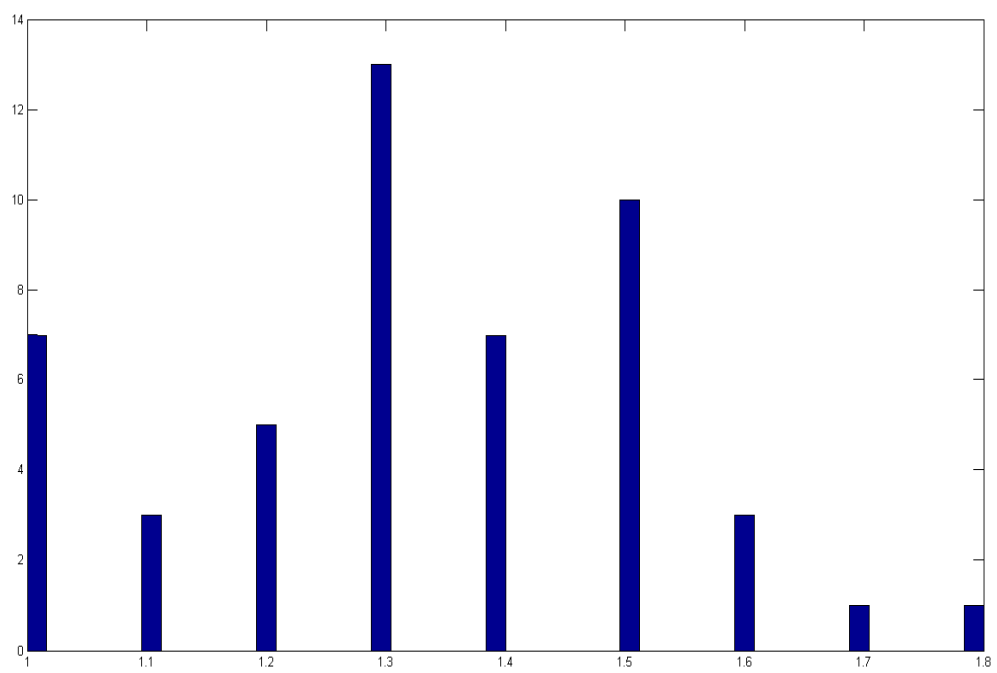
*Longitud del sépalo. Media: 5.9360, Mediana 5.9000*



*Ancho del sépalo. Media: 2.7700, Mediana 2.8000*

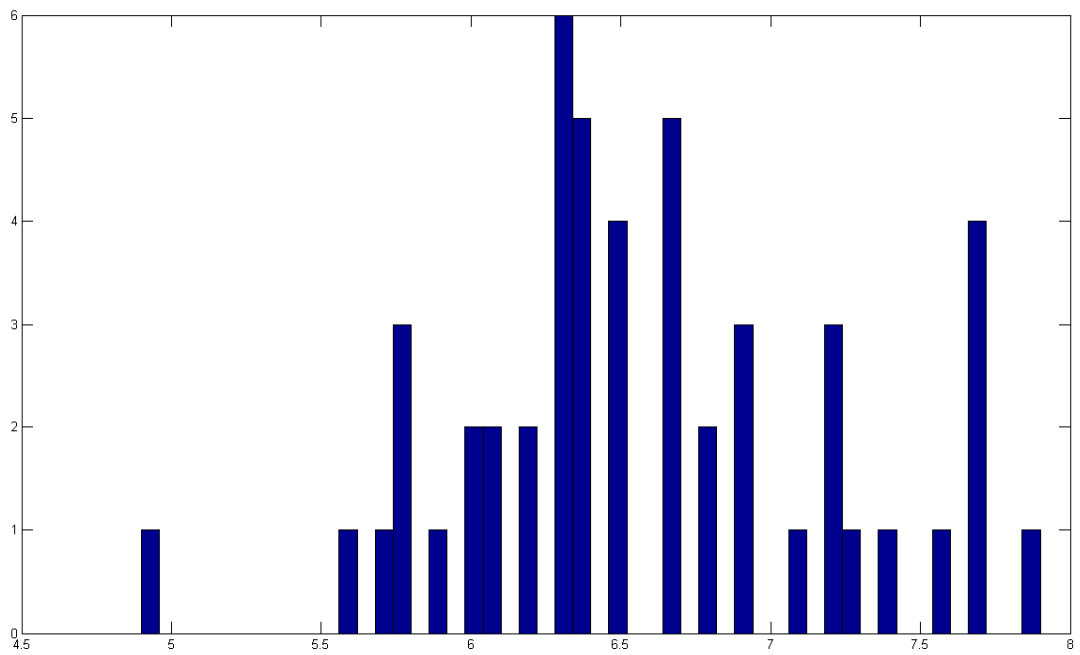


*Longitud del pétalo. Media 4.2600, Mediana 4.3500*

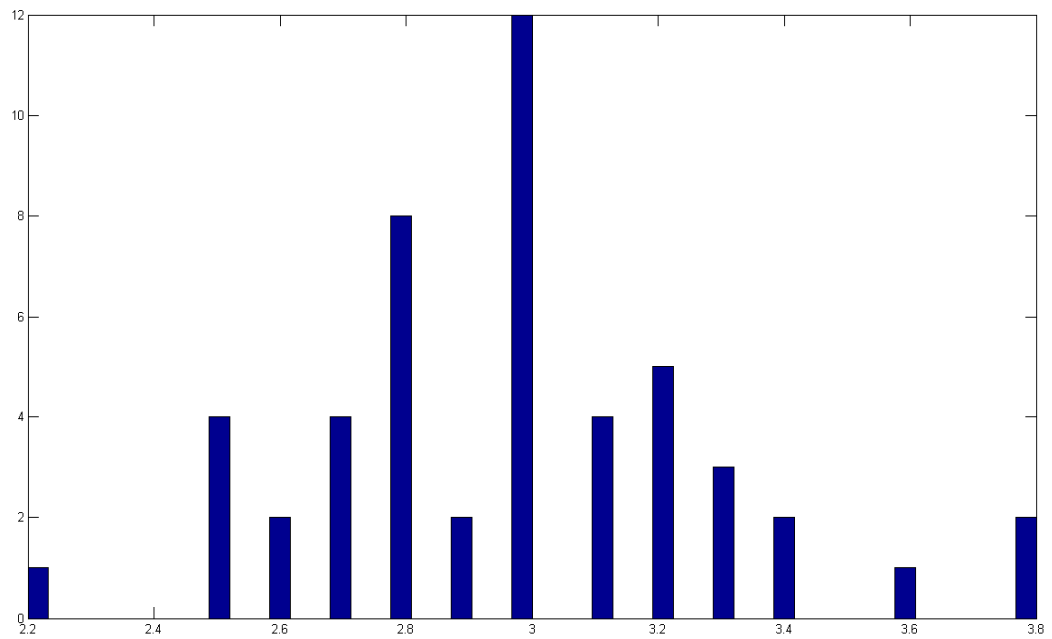


*Ancho del pétalo. Media 1.3260, Mediana 1.3000*

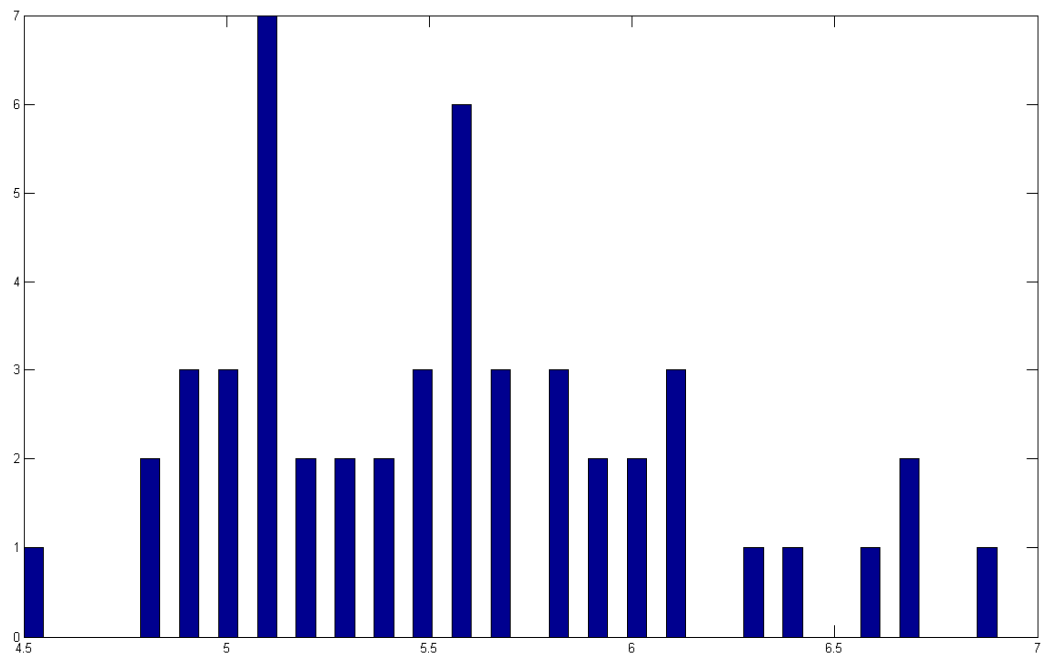
**Iris-Virginica:**



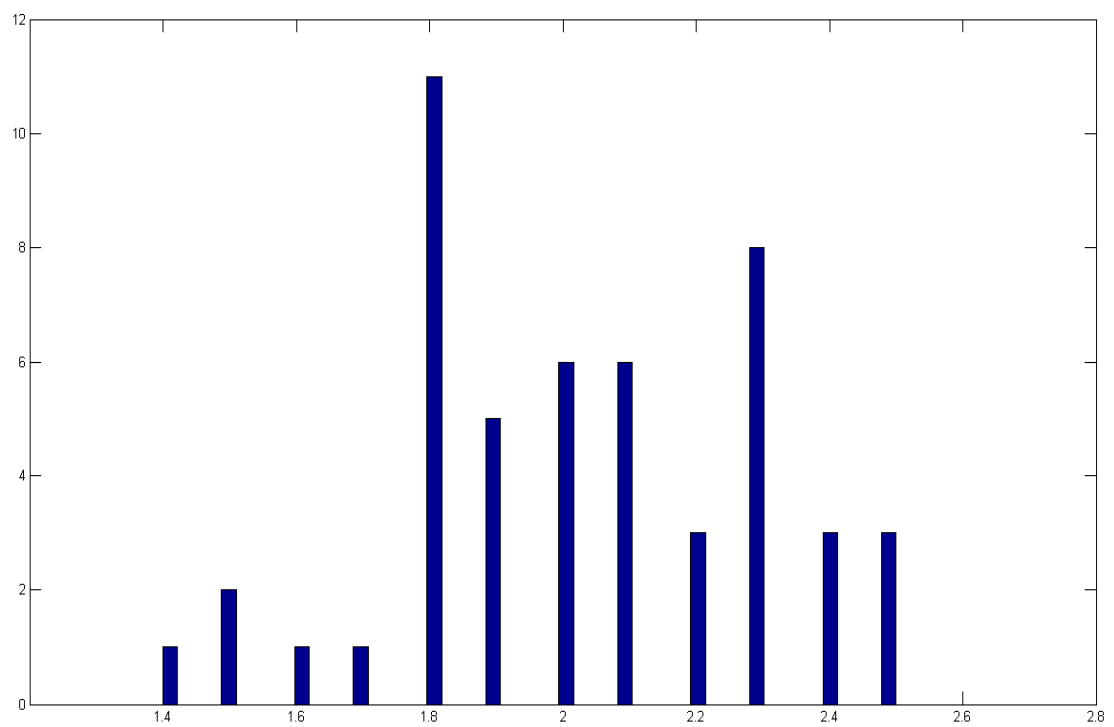
*Longitud del sépalo. Media: 6.5880, Mediana 6.5000*



*Ancho del sépalo. Media: 2.9740, Mediana 3.0000*



*Longitud del pétalo. Media 5.5520, Mediana 5.5500*

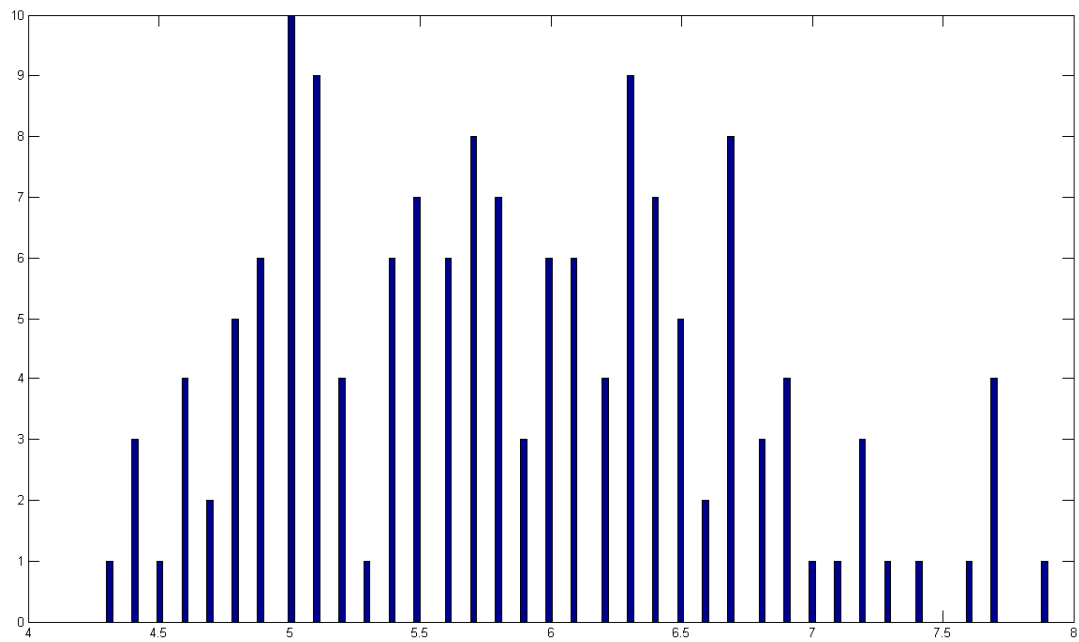


*Ancho del pétalo. Media 2.0260, Mediana 2.0000*

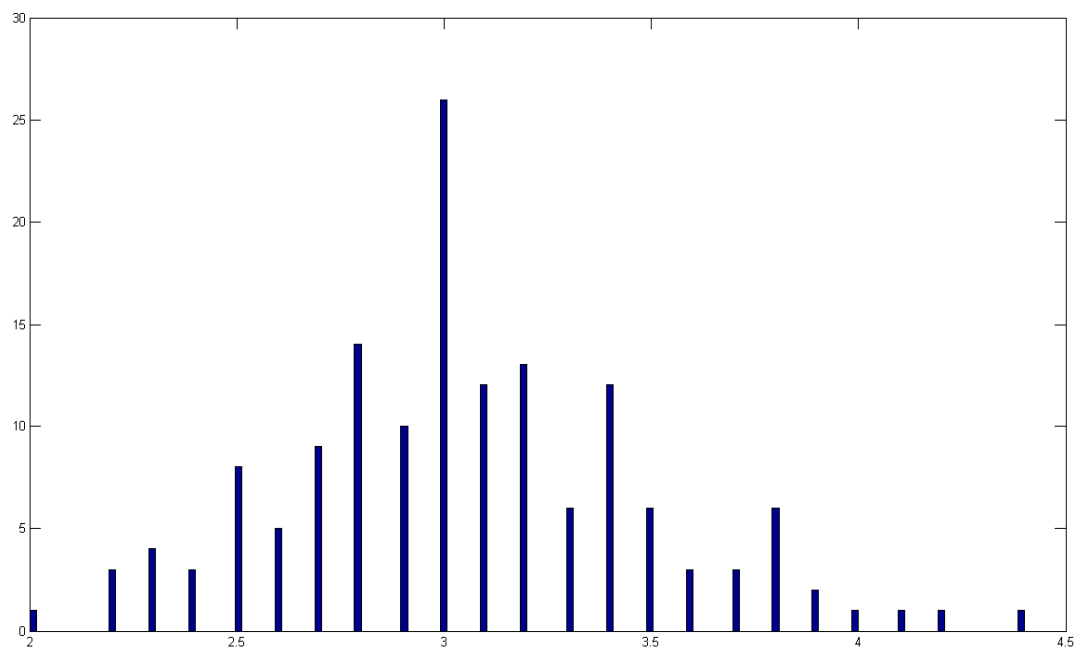


Además se realizaron los mismos graficos para el total de los datos incluyendo las 3 flores iris.

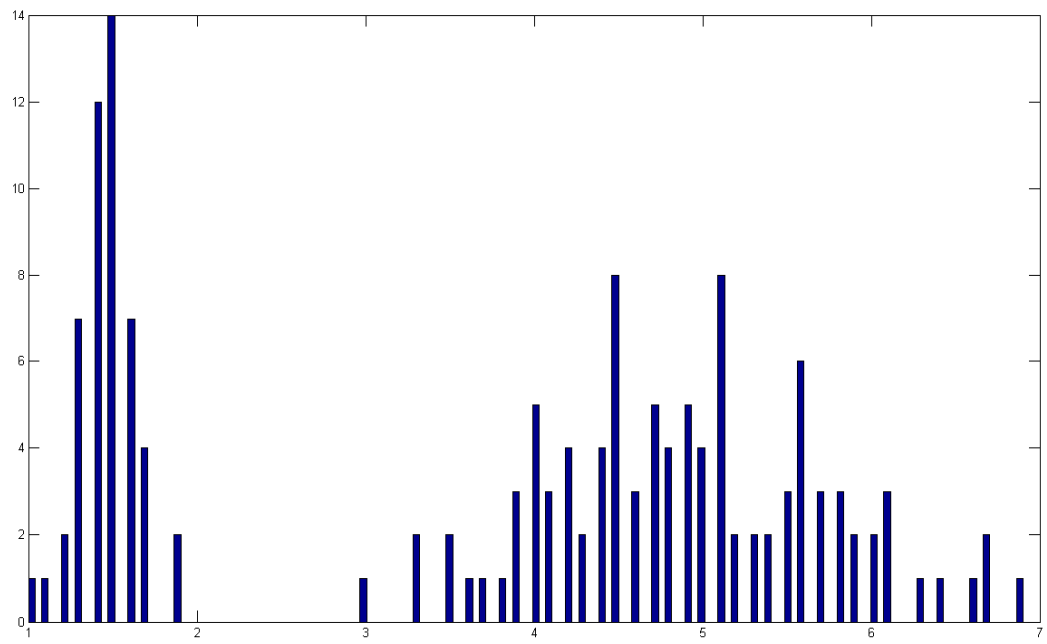
**Todas las flores:**



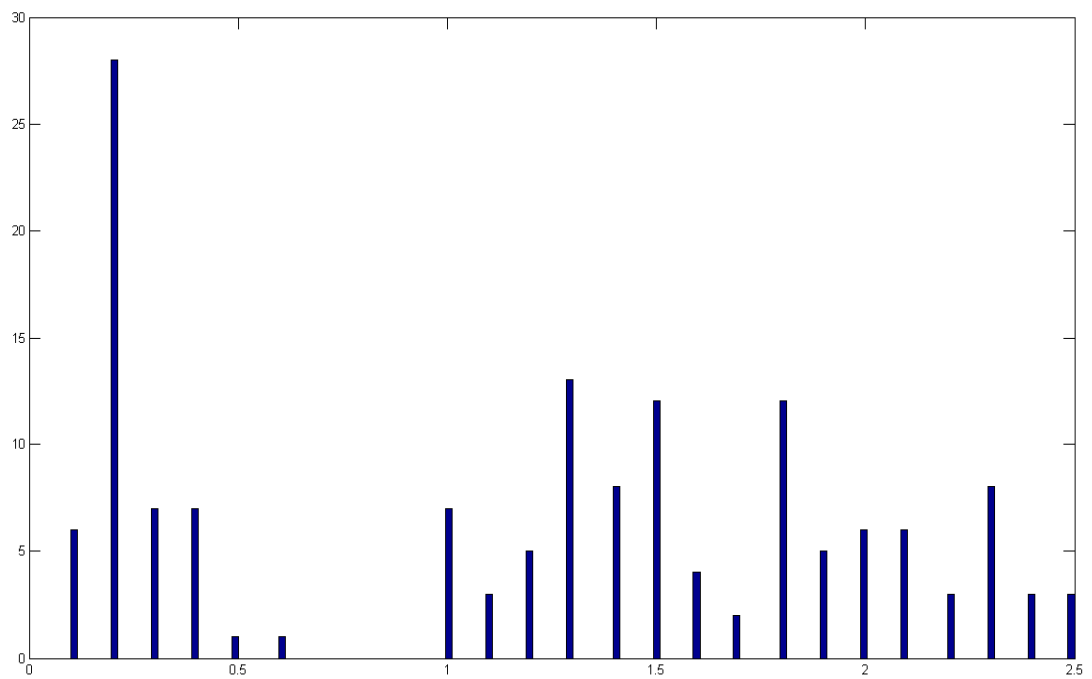
*Longitud del sépalo. Media: 5.8433, Mediana 5.8000*



*Ancho del sépalo. Media: 3.0540, Mediana 3.0000*



*Longitud del pétalo. Media 3.7587, Mediana 4.3500*



*Ancho del pétalo. Media 1.9787, Mediana 1.3000*

b) Para completar las cuatro tablas (una por cada dato analizado) se uso el script siguiente, el cual hace uso de las funciones predefinidas de matlab: “mean(x)”, “sum(x)”, y “std(x)”.

```
%punto 2 b

%Imprime por pantalla la media, suma de cuadrados, media de los cuadrados, y
%desviacion standard de cada columna de cada flor.

function[]= punto2b ()

    for Col=0:1:3

        fprintf('DATOS PARA LA COLUMNA %s\n',nombreColumna(Col));

        fprintf('""'\n\n');

        for flor=0:1:2

            x=vectorColumna(flor,Col);

            media=mean(x);

            fprintf('Media de la flor %s es %.4f\n',nombreFlor(flor),
                media);

            x2=x.^2;

            sumaCuadrados=sum(x);

            fprintf('Suma Cuadrados de la flor %s es
                %.4f\n',nombreFlor(flor),
                    sumaCuadrados);

            mediaCuadrados = mean(x2);

            fprintf('Media de los Cuadrados de la flor %s es
                %.4f\n',nombreFlor(flor), mediaCuadrados);

            desvStd = std(x);

            fprintf('Desviacion standar de la flor %s es
                %.4f\n',nombreFlor(flor), desvStd);

            fprintf('-----\n');

        end

        fprintf('\n');

        fprintf('\n');

    end

end

end
```

**Tablas:**

<i>Long-Sepalo</i>	<b>Media</b>	<b>Suma Cuadrados</b>	<b>Media Cuadrados</b>	<b>Desviación Std.</b>
<b>Iris-Virginica</b>	6.5880	329.4000	43.7980	0.6359
<b>Iris-Versicolor</b>	5.9360	296.8000	35.4972	0.5162
<b>Iris-Setosa</b>	5.0060	250.3000	25.1818	0.3525

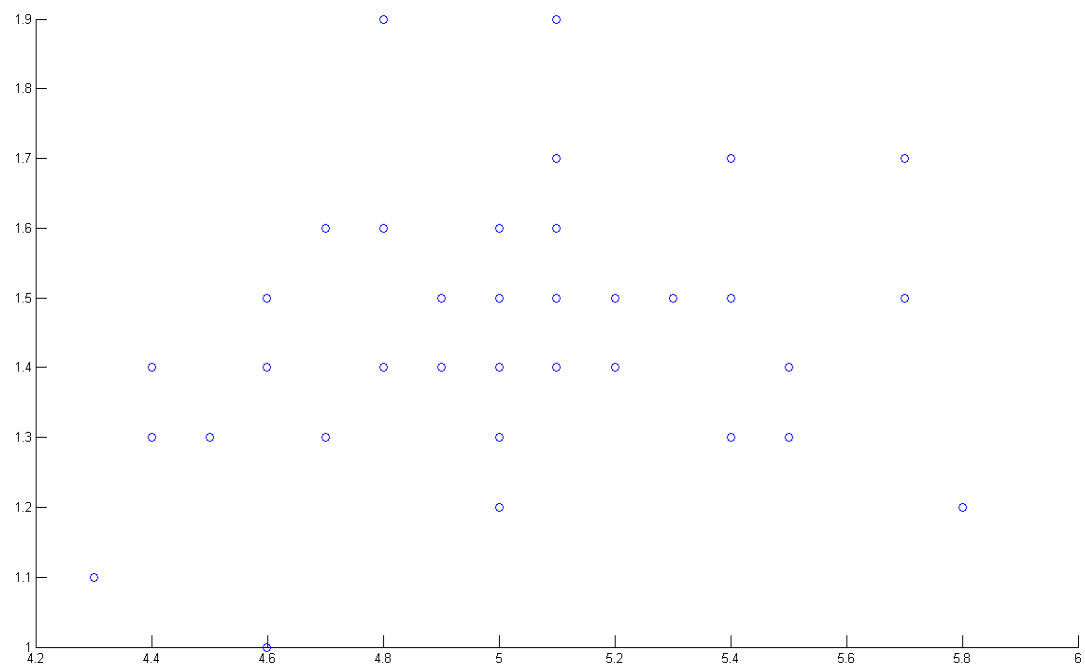
<i>Ancho-Sepalo</i>	<b>Media</b>	<b>Suma Cuadrados</b>	<b>Media Cuadrados</b>	<b>Desviación Std.</b>
<b>Iris-Virginica</b>	2.9740	148.7000	8.9466	0.3225
<b>Iris-Versicolor</b>	2.7700	138.5000	7.7694	0.3138
<b>Iris-Setosa</b>	3.4180	170.9000	11.8250	0.3810

<i>Long-Petalo</i>	<b>Media</b>	<b>Suma Cuadrados</b>	<b>Media Cuadrados</b>	<b>Desviación Std.</b>
<b>Iris-Virginica</b>	5.5520	277.6000	31.1232	0.5519
<b>Iris-Versicolor</b>	4.2600	213.0000	18.3640	0.4699
<b>Iris-Setosa</b>	1.4640	73.2000	2.1728	0.1735

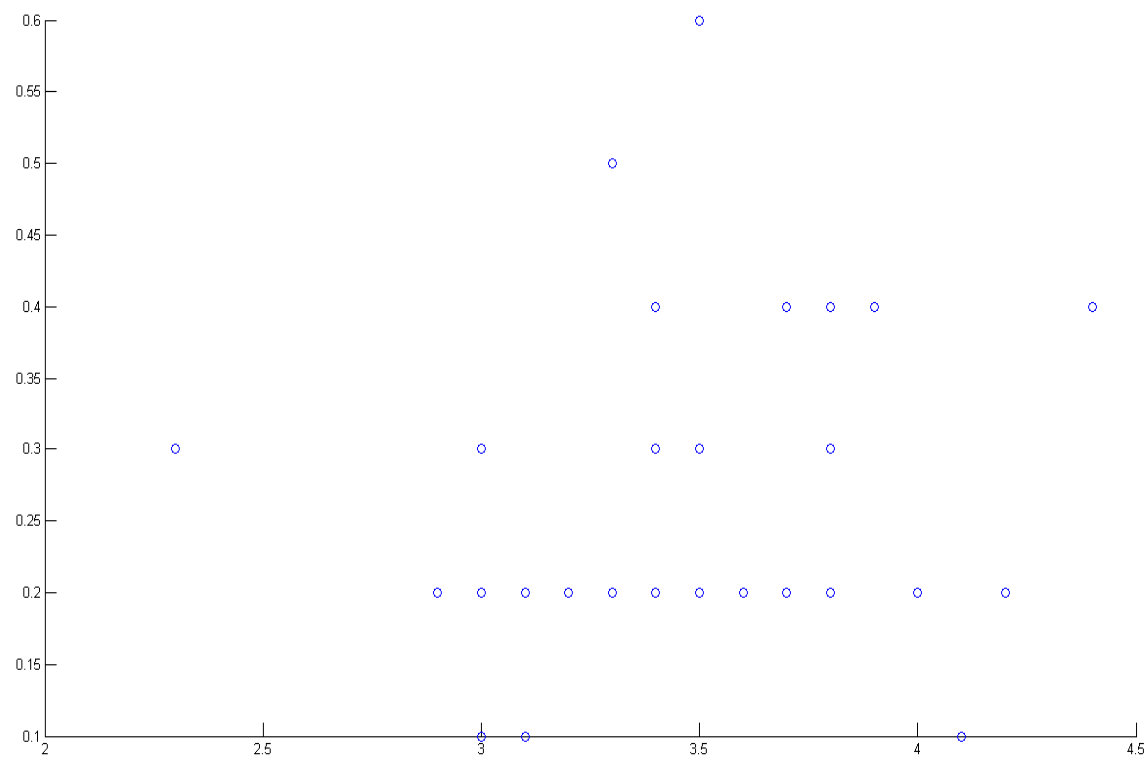
<i>Ancho-Petalo</i>	<b>Media</b>	<b>Suma Cuadrados</b>	<b>Media Cuadrados</b>	<b>Desviación Std.</b>
<b>Iris-Virginica</b>	2.0260	101.3000	4.1786	0.2747
<b>Iris-Versicolor</b>	1.3260	66.3000	1.7966	0.1978
<b>Iris-Setosa</b>	0.2440	12.2000	0.0708	0.1072

c) Para realizar los gráficos de dispersión, MatLab provee de la función “scatter(x,y)” donde x, e y son los vectores que contienen los datos que se quieren comparar.

**Iris-Setosa:**

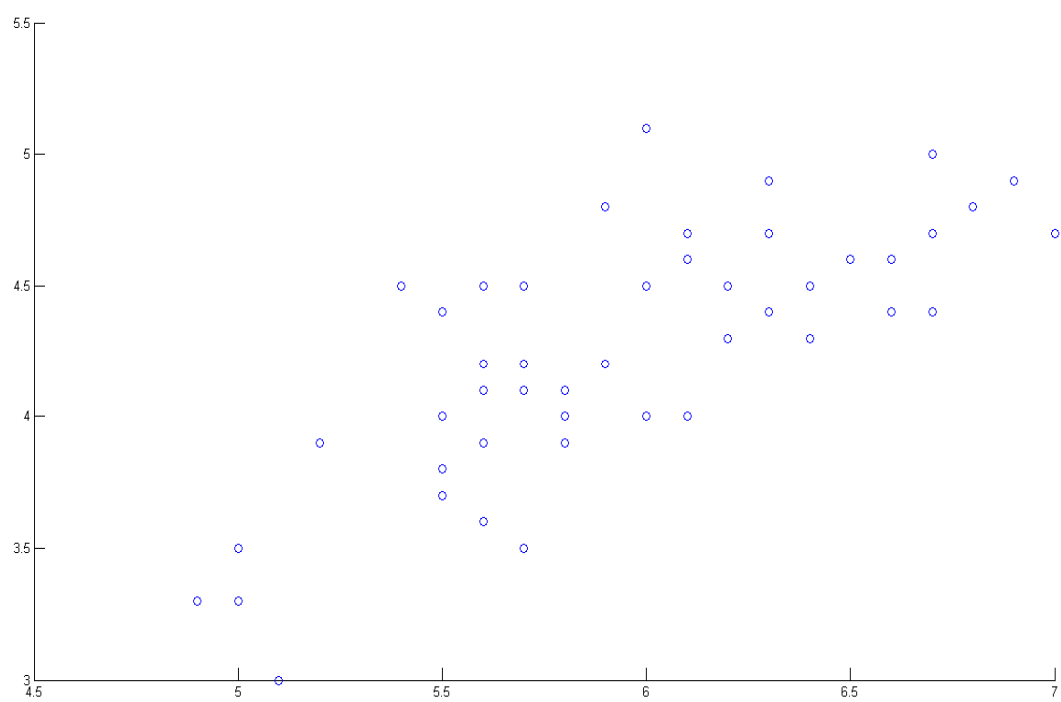


*Longitud del Sépalo (eje X) VS Longitud del Pétalo (eje Y)*

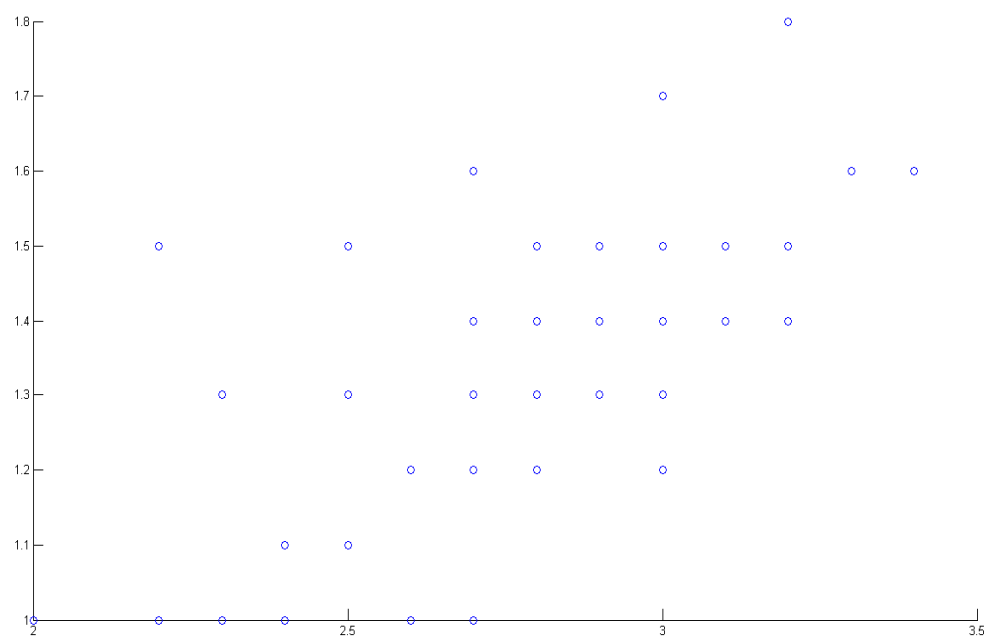


*Ancho del Sépalo (eje X) VS Ancho del Pétalo (eje Y)*

**Iris-Versicolor:**

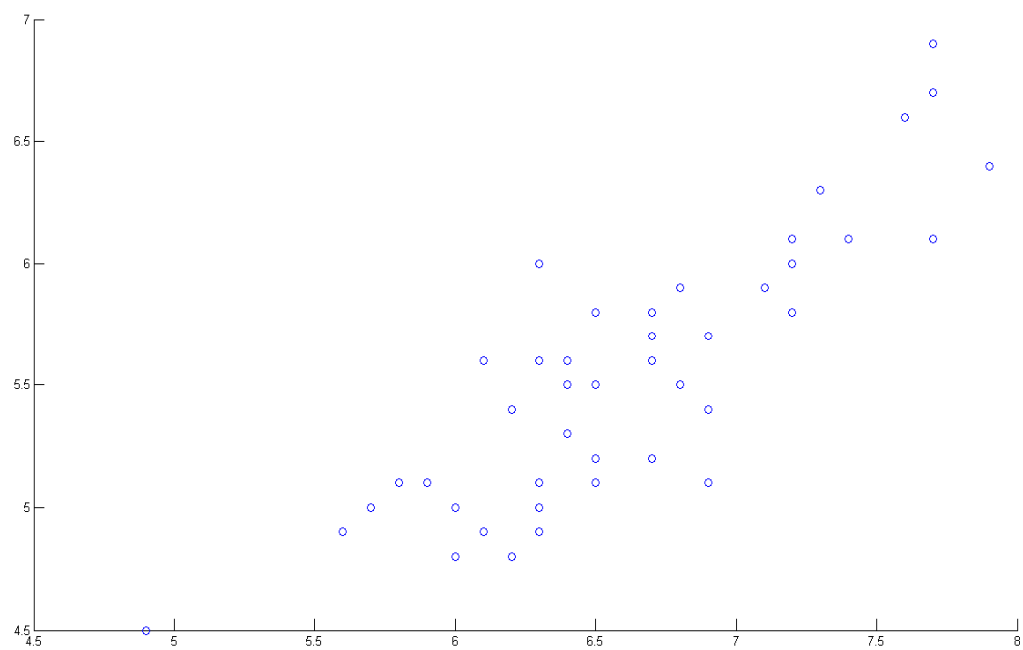


*Longitud del Sépalo (eje X) VS Longitud del Pétalo (eje Y)*

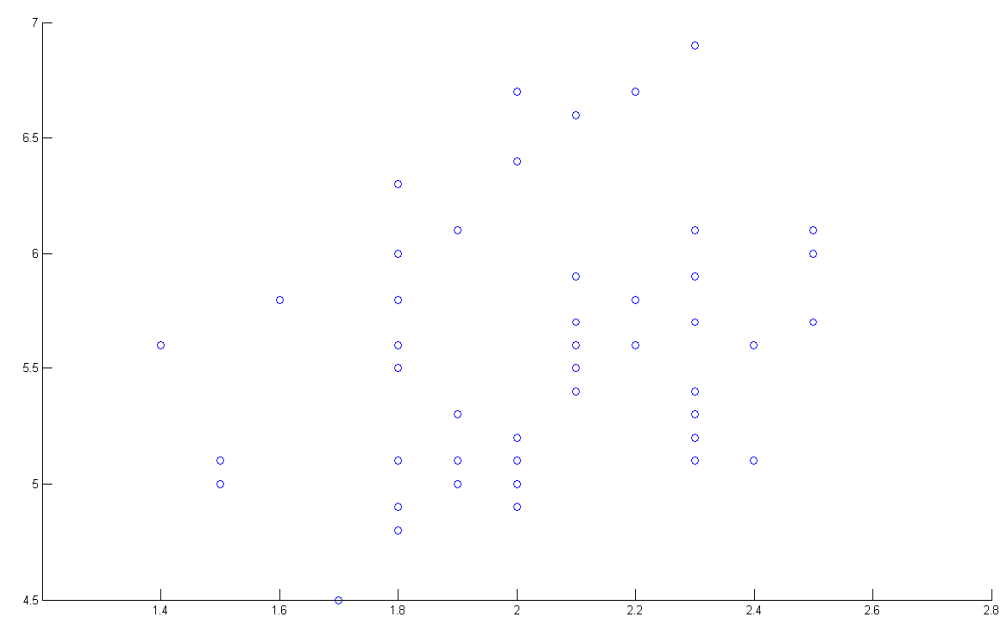


*Ancho del Sépalo (eje X) VS Ancho del Pétalo (eje Y)*

**Iris-Virginica:**



*Longitud del Sépalo (eje X) VS Longitud del Pétalo (eje Y)*



*Ancho del Sépalo (eje X) VS Ancho del Pétalo (eje Y)*

Como se puede ver de los gráficos, tanto la flor Iris-Versicolor, como la flor Iris-Virginica, tienden a incrementar su tamaño de forma uniforme, es decir que a medida que se incrementa el tamaño de sus sépalos, el tamaño de sus pétalos crece proporcionalmente.

Por otra parte, la flor Iris-Setosa pareciera que tiende a un valor constante de la longitud de su pétalo a medida que crece su sépalo. De la misma forma el ancho del pétalo tiende a un valor constante a medida que crece el ancho del sépalo, o por lo menos crece más lentamente el pétalo que el sépalo.

### Ejercicio 3

Para el efectuar las predicciones utilizamos una función Matlab denominada ***predictions*** cuyos parámetros de entrada son:

**longitudSepalo:** vector con las longitudes de todos los sépalos a estudiar.

**longitudPétalo:** vector con las longitudes de todos los pétalos a estudiar.

**p:** valor para el cual se quiere obtener una predicción.

**g:** grado del polinomio que ajustara los valores, para obtener la predicción.

Como salida obtendremos dos valores la **predicción** propiamente dicha y **polinomio** que será el polinomio de ajuste de grado g.

La función, además, grafica los valores recibidos junto con el polinomio y la predicción obtenida por la misma.

```
function [prediccion, polinomio] = predictions(longitudSepalo, longitudPétalo, p, g)

polinomio=polyfit(longitudSepalo, longitudPétalo, g);
prediccion= polyval(polinomio, p);

matriz = [longitudSepalo, longitudPétalo];
matrizOrdenada = sortrows(matriz);

holdon
plot(matrizOrdenada(1:50, 1), matrizOrdenada(1:50, 2), '.r');
plot(matrizOrdenada(1:50, 1), polyval(polinomio, matrizOrdenada(1:50,1)), 'g');
plot(p, prediccion, '<');
holdoff

end
```

Decidimos utilizar 3 polinomios para cada caso. Es decir, se realizó el cálculo para polinomios de grado 1 hasta 3.

- a. Estudio para la clase Setosa en función de las longitudes del pétalo y el sépalo, para predecir cuánto debe medir la longitud del pétalo para un sépalo de longitud 7.



```

>>longitudSepaloSetosa
=[5.1;4.9;4.7;4.6;5.0;5.4;4.6;5.0;4.4;4.9;5.4;4.8;4.8;4.3;5.8;5.7;5.4;5.1;5.7;5.1;5.4;5.1;4.6;5.1;4.8;5.0;5.0;5.2;5.2;
4.7;4.8;5.4;5.2;5.5;4.9;5.0;5.5;4.9;4.4;5.1;5.0;4.5;4.4;5.0;5.1;4.8;5.1;4.6;5.3;5.0];
>>longitudPetalSetosa
=[1.4;1.4;1.3;1.5;1.4;1.7;1.4;1.5;1.4;1.5;1.5;1.6;1.4;1.1;1.2;1.5;1.3;1.4;1.7;1.5;1.7;1.5;1.0;1.7;1.9;1.6;1.6;1.5;1.4;
1.6;1.6;1.5;1.5;1.4;1.5;1.2;1.3;1.5;1.3;1.5;1.3;1.3;1.3;1.6;1.9;1.4;1.6;1.4;1.5;1.4];

>>prediccion=predictions(longitudSepaloSetosa, longitudPetalSetosa, 7, 1)

prediccion =

    1.7230

>>prediccion=predictions(longitudSepaloSetosa, longitudPetalSetosa, 7, 2)

prediccion =

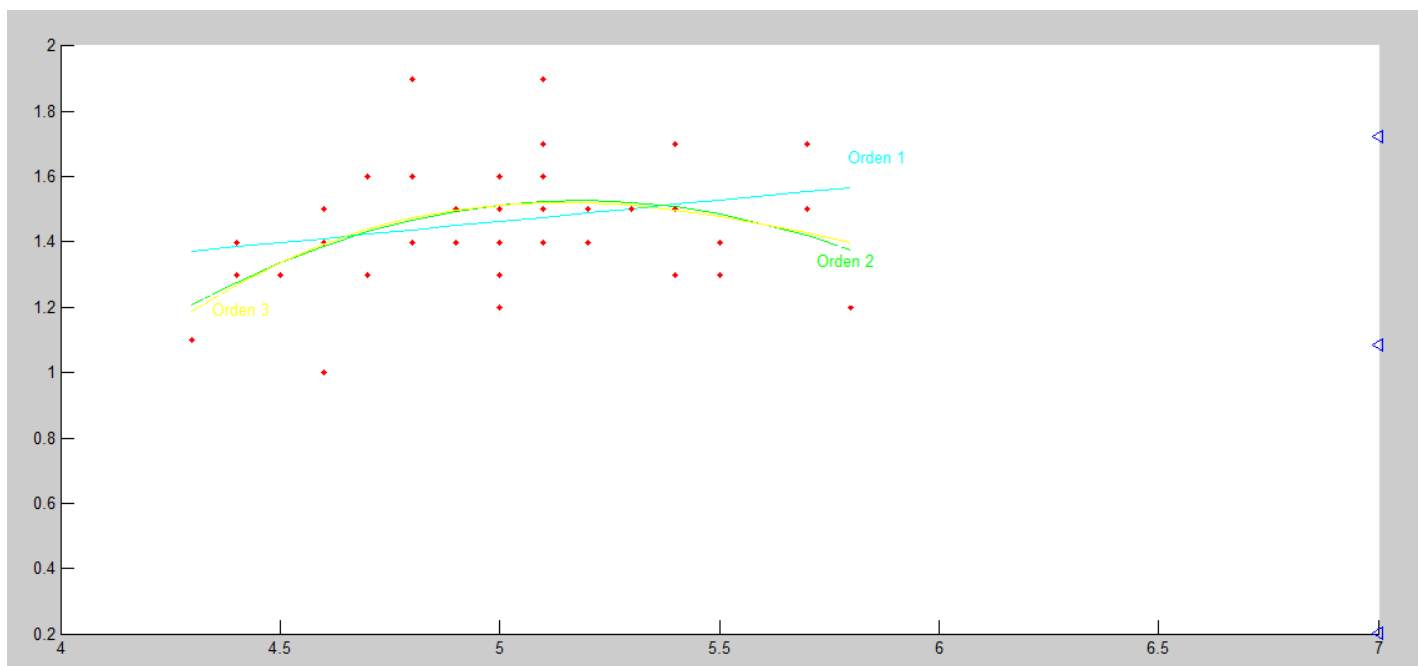
    0.2017

>>prediccion=predictions(longitudSepaloSetosa, longitudPetalSetosa, 7, 3)

prediccion =

    1.0834

```



Obs: al introducir grados mayores a 3 para el polinomio los resultados se van de los esperados

- b. Estudio para la clase Versicolor en función de las longitudes del pétalo y el sépalo, para predecir cuánto debe medir la longitud del pétalo para un sépalo de longitud 10.

```

>>longitudSepaloVersicolor
=[7.0;6.4;6.9;5.5;6.5;5.7;6.3;4.9;6.6;5.2;5.0;5.9;6.0;6.1;5.6;6.7;5.6;5.8;6.2;5.6;5.9;6.1;6.3;6.1;6.4;6.6;6.8;
6.7;6.0;5.7;5.5;5.5;5.8;6.0;5.4;6.0;6.7;6.3;5.6;5.5;5.5;6.1;5.8;5.0;5.6;5.7;5.7;6.2;5.1;5.7];
>>longitudPetalovVersicolor
=[4.7;4.5;4.9;4.0;4.6;4.5;4.7;3.3;4.6;3.9;3.5;4.2;4.0;4.7;3.6;4.4;4.5;4.1;4.5;3.9;4.8;4.0;4.9;4.7;4.3;4.4;4.8;
5.0;4.5;3.5;3.8;3.7;3.9;5.1;4.5;4.5;4.7;4.4;4.1;4.0;4.4;4.6;4.0;3.3;4.2;4.2;4.2;4.3;5.3;4.1];
>>prediccion=predictions(longitudSepaloVersicolor, longitudPetalovVersicolor, 10, 1)

prediccion =

    6.4973

>>prediccion=predictions(longitudSepaloVersicolor, longitudPetalovVersicolor, 10, 2)

prediccion =

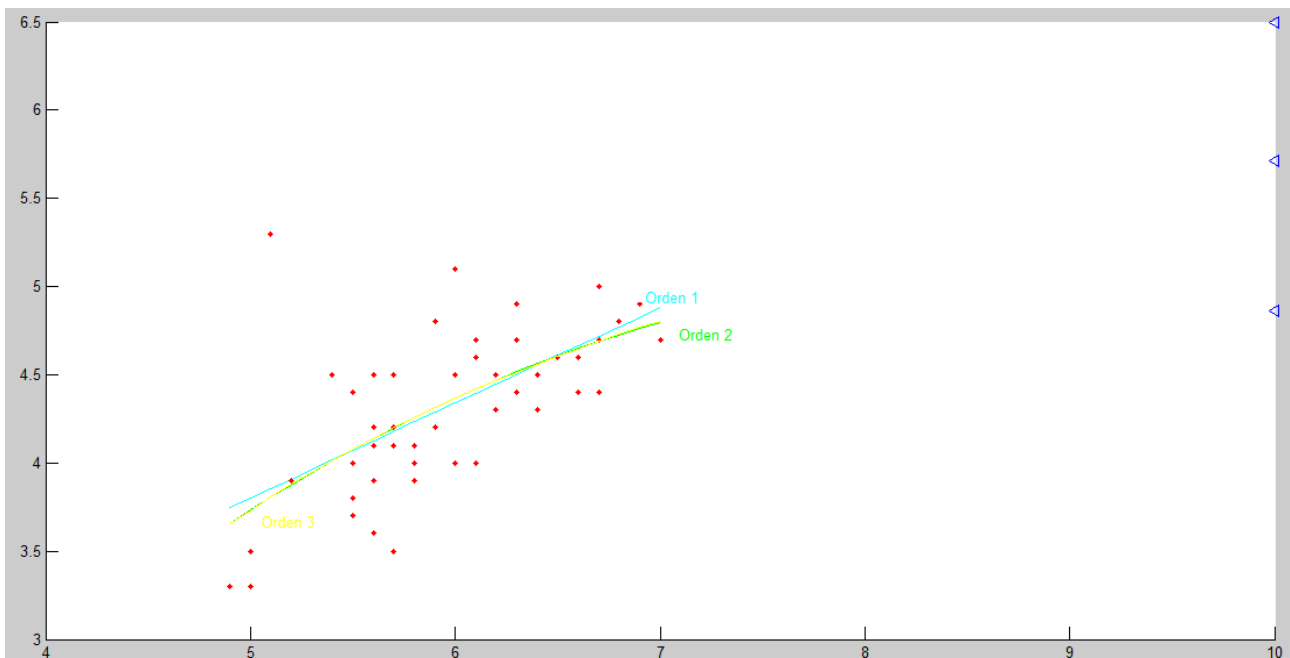
    4.8590

>>prediccion=predictions(longitudSepaloVersicolor, longitudPetalovVersicolor, 10, 3)

prediccion =

    5.7113

```



Obs: al introducir grados mayores a 3 para el polinomio los resultados se van de los esperados

- c. Estudio para la clase Versicolor en función de las longitudes del pétalo y el sépalo, para predecir cuánto debe medir la longitud del pétalo para un sépalo de longitud 8.

```

>>longitudSepaloVirginica
=[6.3;5.8;7.1;6.3;6.5;7.6;4.9;7.3;6.7;7.2;6.5;6.4;6.8;5.7;5.8;6.4;6.5;7.7;7.7;6.0;6.9;5.6;7.7;6.3;6.7;7.2;6.2;6.1;6.
4;7.2;7.4;7.9;6.4;6.3;6.1;7.7;6.3;6.4;6.0;6.9;6.7;6.9;5.8;6.8;6.7;6.7;6.3;6.5;6.2;5.9];
>>longitudPetalovirginica
=[6.0;5.1;5.9;5.6;5.8;6.6;4.5;6.3;5.8;6.1;5.1;5.3;5.5;5.0;5.1;5.3;5.5;6.7;6.9;5.0;5.7;4.9;6.7;4.9;5.7;6.0;4.8;4.9;5.
6;5.8;6.1;6.4;5.6;5.1;5.6;6.1;5.6;5.5;4.8;5.4;5.6;5.1;5.1;5.9;5.7;5.2;5.0;5.2;5.4;5.1];
>>prediccion=predictions(longitudSepaloVirginica, longitudPetalovirginica, 8, 1)

prediccion =

    6.6111

>>prediccion=predictions(longitudSepaloVirginica, longitudPetalovirginica, 8, 2)

prediccion =

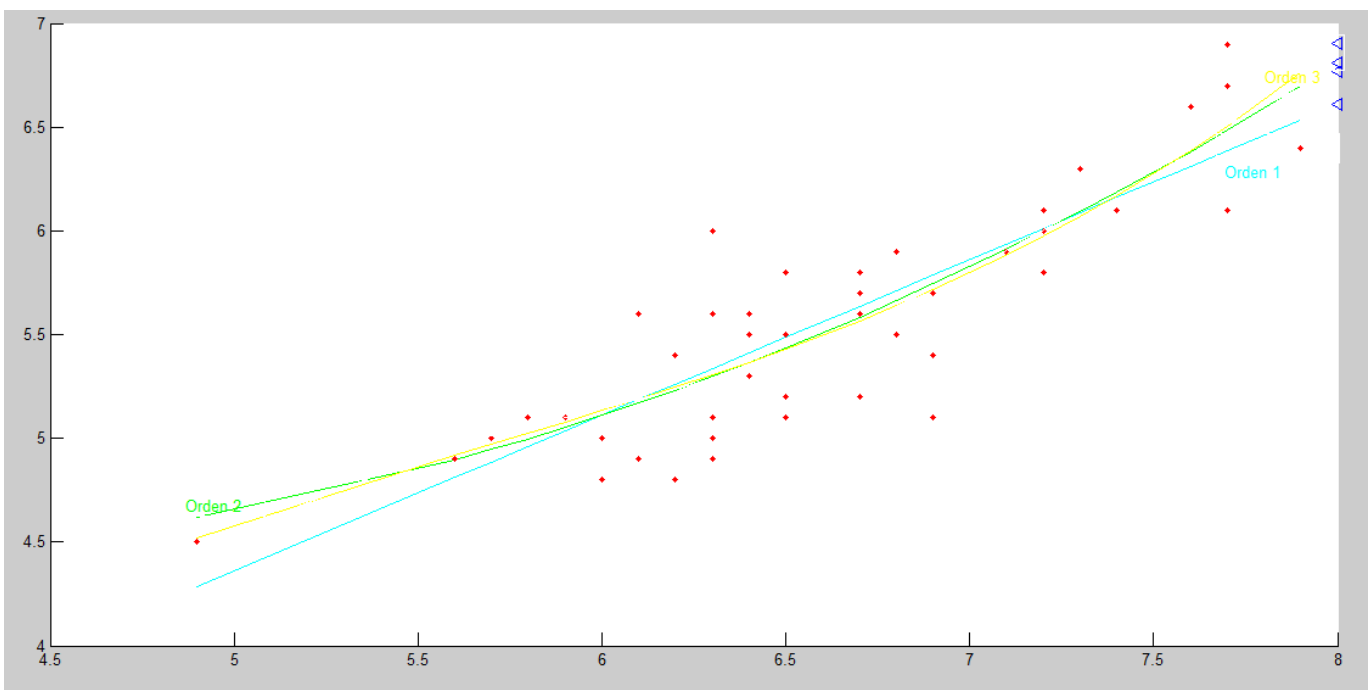
    6.8064

>>prediccion=predictions(longitudSepaloVirginica, longitudPetalovirginica, 8, 3)

prediccion =

    6.9044

```



Obs: al introducir grados mayores a 3 para el polinomio los resultados obtenidos son “parecidos”, es decir varían en algunos decimales al obtenido con un polinomio de grado 3

## **Conclusiones Extraídas de los Resultados**

**a)** La longitud del pétalo de la Iris-Setosa tiende a ciertos valores constantes a medida que crece la longitud del sépalo, pero, a medida que dicho valor aumenta, se observa una desviación pronunciada de las curvas de predicción de los valores constantes. El efecto de esta desviación se observa al intentar predecir el valor en números superiores (como, por ejemplo, de 7.0 u 8.0), donde el resultado del cálculo devuelve valores reducidos o negativos, lo que representa un absurdo.

En consecuencia, concluimos que las predicciones no son confiables para la búsqueda de los valores deseados.

**b)** A medida que se incrementa el tamaño del sépalo de la Iris-Versicolor, el tamaño de sus pétalos se incrementa proporcionalmente. Las curvas de predicción obtenidas representan este comportamiento adecuadamente según los resultados obtenidos en el ejercicio número dos, por lo cual podemos concluir que las curvas son correctas y que la predicción resulta confiable.

**c)** De la misma manera que sucede con la Iris-Versicolor, las curvas obtenidas para la predicción de los atributos de la Iris-Virginica se corresponden con los resultados obtenidos en el ejercicio dos, de acuerdo a su comportamiento de crecimiento proporcional, con lo cual podemos aseverar que las predicciones son confiables.