# Numerical Methods and Machine Learning for Image Processing

## Week 4, Class 1: Classification and Regression

October 18, 2021

Damon M. Chandler and Yi Zhang

# Today: Classification and Regression, Part 1

1. Basic definitions
2. Bayesian classification
    1. Problem formulation
    2. Bayes classification in Excel
    3. Bayes classification in Python

# Today: Classification and Regression, Part 1

1. Basic definitions
2. Bayesian classification
   1. Problem formulation
   2. Bayes classification in Excel
   3. Bayes classification in Python

# Basic definitions

- **Classification**
  - *Goal*: Predict a **class label** for each object/region/pixel
  - Prediction of a categorical variable
    - Input = features of an object
    - Output = class label for that object
- **Regression**
  - *Goal*: Predict a **numerical score** for each object/region/pixel
  - Prediction of a numerical (continuous) variable
    - Input = features of an object
    - Output = numerical score for that object

# Basic definitions

- Machine Learning
  - **Unsupervised** learning
    - Clustering (e.g., *k*-means)
    - Density estimation (e.g., GMM)
  - **Supervised** learning
    - **Learn** from **training data** (need ground-truth values/labels)
    - **Predict** from **testing data**
  - **Reinforcement** learning
    - No supervision, but uses a **reward function**
    - Used mostly in robotics

# Today: Classification and Regression, Part 1

1. Basic definitions
2. Bayesian classification
    1. Problem formulation
    2. Bayes classification in Excel
    3. Bayes classification in Python

# Bayes classification

Basic idea:

- Classify based on **most probable class** given the **feature measurements**
- $p(C_i|\mathbf{x})$ is the probability of class $C_i$ given that measured values are $\mathbf{x}$

$$p(C_i|\mathbf{x}) = \frac{p(C_i)p(\mathbf{x}|C_i)}{p(\mathbf{x})} = \frac{p(C_i)p(x_1|C_i)p(x_2|C_i)p(x_3|C_i)p(x_4|C_i)}{p(\mathbf{x})}$$

Steps:

- *Training*: Compute statistical models from training data (many $\mathbf{x}$'s)
- *Testing*:
    1. Input specimen $\mathbf{x}$
    2. Compute $p(C_1|\mathbf{x})$, $p(C_2|\mathbf{x})$, …, $p(C_N|\mathbf{x})$
    3. Classify as $C_1$ or $C_2$ or … or $C_N$ depending on which $p(C_i|\mathbf{x})$ is largest

# Bayes Classification

Definitions:

- $C_i$ is a particular class (e.g., $C_1 = \text{"Mac"}$, $C_2 = \text{"Linux"}$, $C_3 = \text{"Windows"}$)
- $\mathbf{x} = (x_1, x_2, \dots, x_N)$ is a vector of $N$ feature values (measurements)
- $p(\mathbf{x}) = p(x_1, x_2, \dots, x_N)$ is the (joint) probability of encountering the particular set of feature values $\mathbf{x}$
- $p(C_i)$ is the probability of encountering the particular class $C_i$
- $p(\mathbf{x}|C_i) = p(x_1, x_2, x_3, x_4|C_i)$ is the (conditional) probability of encountering the particular set of feature values $\mathbf{x}$ given that the class is $C_i$
- $p(C_i|\mathbf{x}) = p(C_i|x_1, x_2, x_3, x_4)$ is the (conditional) probability of encountering the particular class $C_i$ given that the feature values are $\mathbf{x}$

$$p(C_i|\mathbf{x}) = \frac{p(C_i)p(\mathbf{x}|C_i)}{p(\mathbf{x})}$$

# How to classify an iris into one of these three types?

Iris setosa

Iris versicolor

Iris virginica

https://archive.ics.uci.edu/ml/datasets/Iris



# Iris Data Set

*Download*: Data Folder, Data Set Description

**Abstract**: Famous database; from Fisher, 1936

| | | | | | |
|---|---|---|---|---|---|
| **Data Set Characteristics:** | Multivariate | **Number of Instances:** | 150 | **Area:** | Life |
| **Attribute Characteristics:** | Real | **Number of Attributes:** | 4 | **Date Donated** | 1988-07-01 |
| **Associated Tasks:** | Classification | **Missing Values?** | No | **Number of Web Hits:** | 4263892 |

### Source:

Creator:

R.A. Fisher

Donor:

Michael Marshall (MARSHALL%PLU '@' io.arc.nasa.gov)

### Data Set Information:

This is perhaps the best known database to be found in the pattern recognition literature. Fisher's paper is a classic in the field and is referenced frequently to this day. (See Duda & Hart, for example.) The data set contains 3 classes of 50 instances each, where each class refers to a type of iris plant. One class is linearly separable from the other 2; the latter are NOT linearly separable from each other.

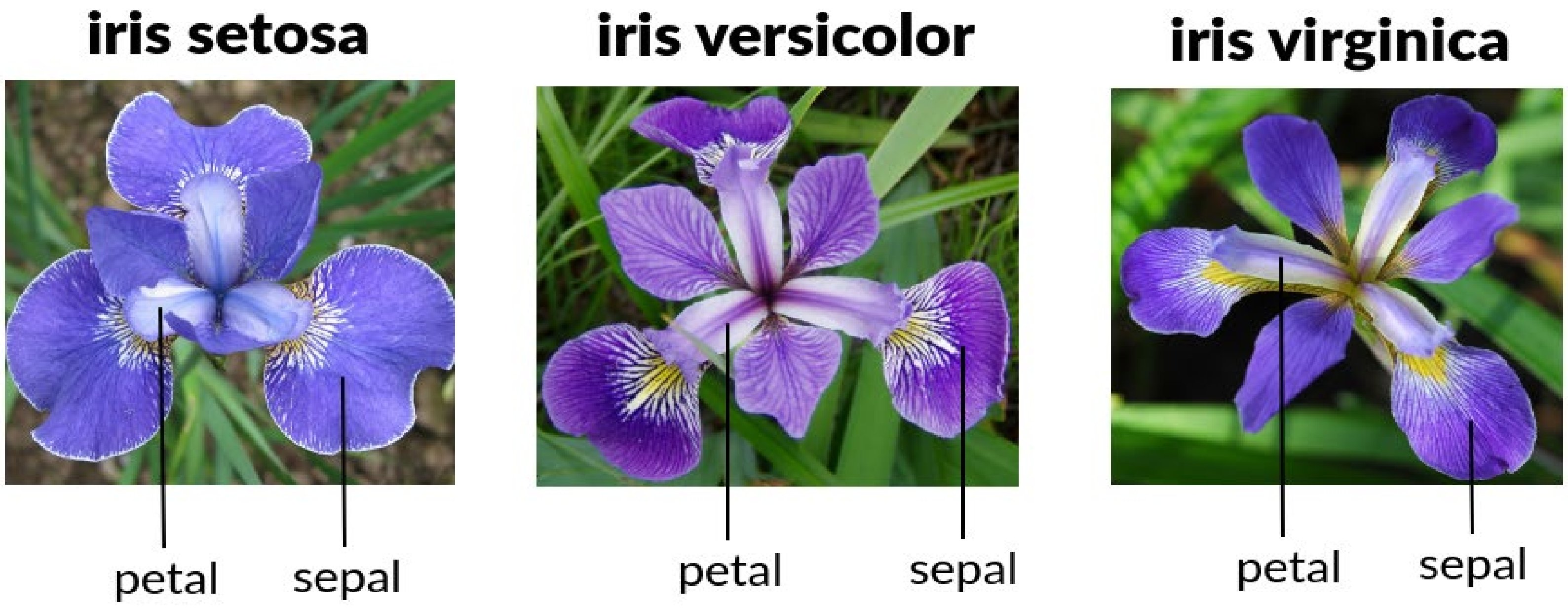
Predicted attribute: class of iris plant.

This is an exceedingly simple domain.

This data differs from the data presented in Fishers article (identified by Steve Chadwick, spchadwick '@' espeedaz.net ). The 35th sample should be: 4.9,3.1,1.5,0.2,"Iris-setosa" where the error is in the fourth feature. The 38th sample: 4.9,3.6,1.4,0.1,"Iris-setosa" where the errors are in the second and third features.

### Attribute Information:

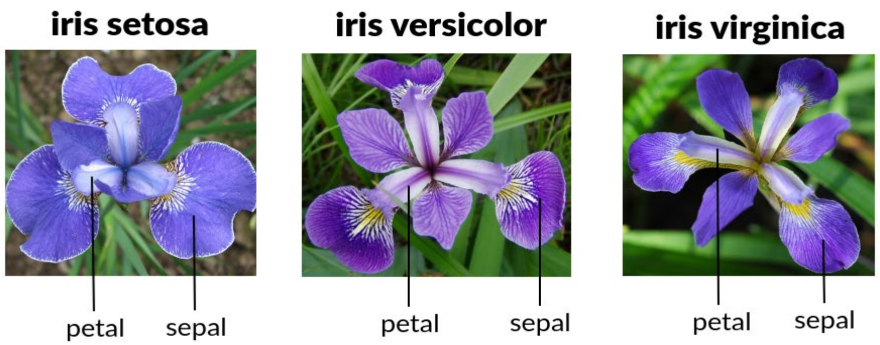
# Example: Bayes classification of flowers




Definitions:

- $C_i$ is a particular class ($C_1$ = "Setosa", $C_2$ = "Versicolor", $C_3$ = "Virginica")
- $\mathbf{x} = (x_1, x_2, x_3, x_4) = (\text{sepal length}, \text{sepal width}, \text{petal length}, \text{petal width})$ is a vector of 4 feature values (measurements)

# Example: Bayes classification of flowers




Definitions:

- $C_i$ is a particular class ($C_1 =$ "Setosa", $C_2 =$ "Versicolor", $C_3 =$ "Virginica")
- $\mathbf{x} = (x_1, x_2, x_3, x_4) = (\text{sepal length}, \text{sepal width}, \text{petal length}, \text{petal width})$ is a vector of 4 feature values (measurements)
- $p(\mathbf{x}) = p(x_1, x_2, x_3, x_4)$ is the (joint) probability of encountering the particular set of feature values $\mathbf{x}$
- $p(C_i)$ is the probability of encountering the particular class $C_i$
  - $p(C_1) = p(\text{"Setosa"}) = 1/3$
  - $p(C_2) = p(\text{"Versicolor"}) = 1/3$
  - $p(C_3) = p(\text{"Virginica"}) = 1/3$

# Example: Bayes classification of flowers

Definitions:

- $p(\mathbf{x}|C_i) = p(x_1, x_2, x_3, x_4|C_i)$ is the (conditional) probability of encountering the particular set of feature values $\mathbf{x}$ given that the class is $C_i$ (given that $\mathbf{x}$ was measured for a $C_i$-type specimen)
  - Assuming independence:
    $p(x_1, x_2, x_3, x_4|C_i) = p(x_1|C_i)p(x_2|C_i)p(x_3|C_i)p(x_4|C_i)$
    $= p(\text{sepal length}|C_i)p(\text{sepal width}|C_i)p(\text{petal length}|C_i)p(\text{petal width}|C_i)$
  - $p(x_n|C_i) = \mathcal{N}\left(\mu_{n,i}, \sigma_{n,i}^2\right) = \frac{1}{\sqrt{2\pi\sigma_{n,i}^2}} e^{-\frac{\left(x_n - \mu_{n,i}\right)^2}{2\sigma_{n,i}^2}}$
  - We can "learn" $\mu_{n,i}$ and $\sigma_{n,i}$ from training data
  - $\mu_{n,i}$ = average value of feature $x_n$ measured for class $C_i$
  - $\sigma_{n,i}$ = standard deviation of feature $x_n$ measured for class $C_i$

# Example: Bayes classification of flowers

Definitions:

- $p(C_i|\mathbf{x})$ is the (conditional) probability of encountering the particular class $C_i$ given that the feature values are $\mathbf{x}$

$$p(C_i|\mathbf{x}) = \frac{p(C_i)p(\mathbf{x}|C_i)}{p(\mathbf{x})} = \frac{p(C_i)p(x_1|C_i)p(x_2|C_i)p(x_3|C_i)p(x_4|C_i)}{p(\mathbf{x})}$$

$$= \frac{p(C_i)\mathcal{N}\left(\mu_{1,i},\sigma_{1,i}^2\right)\mathcal{N}\left(\mu_{2,i},\sigma_{2,i}^2\right)\mathcal{N}\left(\mu_{3,i},\sigma_{3,i}^2\right)\mathcal{N}\left(\mu_{4,i},\sigma_{4,i}^2\right)}{p(\mathbf{x})}$$

$$= \frac{\frac{1}{3}\left(\frac{1}{\sqrt{2\pi\sigma_{1,i}^2}}e^{-\frac{\left(x_1-\mu_{1,i}\right)^2}{2\sigma_{1,i}^2}}\right)\left(\frac{1}{\sqrt{2\pi\sigma_{2,i}^2}}e^{-\frac{\left(x_2-\mu_{2,i}\right)^2}{2\sigma_{2,i}^2}}\right)\left(\frac{1}{\sqrt{2\pi\sigma_{3,i}^2}}e^{-\frac{\left(x_3-\mu_{3,i}\right)^2}{2\sigma_{3,i}^2}}\right)\left(\frac{1}{\sqrt{2\pi\sigma_{4,i}^2}}e^{-\frac{\left(x_4-\mu_{4,i}\right)^2}{2\sigma_{4,i}^2}}\right)}{p(\mathbf{x})}$$

- Classify based on which is the largest: $p(C_1|\mathbf{x})$ or $p(C_2|\mathbf{x})$ or $p(C_3|\mathbf{x})$

# Example: Bayes classification of flowers

Decision:

- $p(C_i|\mathbf{x})$ is the (conditional) probability of encountering the particular class $C_i$ given that the feature values are $\mathbf{x}$

$$p(C_i|\mathbf{x}) = \frac{p(C_i)p(\mathbf{x}|C_i)}{p(\mathbf{x})} = \frac{p(C_i)p(x_1|C_i)p(x_2|C_i)p(x_3|C_i)p(x_4|C_i)}{p(\mathbf{x})}$$

- Classify based on which is the largest: $p(C_1|\mathbf{x})$ or $p(C_2|\mathbf{x})$ or $p(C_3|\mathbf{x})$

That is, we compare: $\dfrac{p(C_1)p(\mathbf{x}|C_1)}{p(\mathbf{x})}$ vs. $\dfrac{p(C_2)p(\mathbf{x}|C_2)}{p(\mathbf{x})}$ vs. $\dfrac{p(C_3)p(\mathbf{x}|C_3)}{p(\mathbf{x})}$

same as comparing: $p(C_1)p(\mathbf{x}|C_1)$ vs. $p(C_2)p(\mathbf{x}|C_2)$ vs. $p(C_3)p(\mathbf{x}|C_3)$

# Today: Classification and Regression, Part 1

1. Basic definitions
2. Bayesian classification
    1. Problem formulation
    2. Bayes classification in Excel
    3. Bayes classification in Python

# Learning the means and standard deviations from the *iris* training data

**Training Data for Class = Iris-setosa (1)**

| | x1 | x2 | x3 | x4 | C |
|---|---|---|---|---|---|
| | **sepal length** | **sepal width** | **petal length** | **petal width** | **class** |
| | 4.6 | 3.4 | 1.4 | 0.3 | Iris-setosa |
| | 4.8 | 3.1 | 1.6 | 0.2 | Iris-setosa |
| | 4.4 | 2.9 | 1.4 | 0.2 | Iris-setosa |
| | 5.7 | 4.4 | 1.5 | 0.4 | Iris-setosa |
| | 4.9 | 3 | 1.4 | 0.2 | Iris-setosa |
| | 5.1 | 3.8 | 1.9 | 0.4 | Iris-setosa |
| | 4.8 | 3.4 | 1.6 | 0.2 | Iris-setosa |
| | 4.6 | 3.1 | 1.5 | 0.2 | Iris-setosa |
| | 5.8 | 4 | 1.2 | 0.2 | Iris-setosa |
| | 5 | 3.5 | 1.6 | 0.6 | Iris-setosa |
| | 5 | 3.5 | 1.3 | 0.3 | Iris-setosa |
| | 4.6 | 3.6 | 1 | 0.2 | Iris-setosa |
| | 4.7 | 3.2 | 1.3 | 0.2 | Iris-setosa |
| | 5.7 | 3.8 | 1.7 | 0.3 | Iris-setosa |
| | 5.4 | 3.4 | 1.5 | 0.4 | Iris-setosa |
| | 4.7 | 3.2 | 1.6 | 0.2 | Iris-setosa |
| | 5.2 | 4.1 | 1.5 | 0.1 | Iris-setosa |
| | 4.8 | 3 | 1.4 | 0.1 | Iris-setosa |
| | 5.5 | 3.5 | 1.3 | 0.2 | Iris-setosa |
| | 5 | 3.3 | 1.4 | 0.2 | Iris-setosa |
| | 4.8 | 3.4 | 1.9 | 0.2 | Iris-setosa |
| | 4.5 | 2.3 | 1.3 | 0.3 | Iris-setosa |
| | 5.3 | 3.7 | 1.5 | 0.2 | Iris-setosa |
| | 4.9 | 3.1 | 1.5 | 0.1 | Iris-setosa |
| | 5.4 | 3.7 | 1.5 | 0.2 | Iris-setosa |
| | 4.3 | 3 | 1.1 | 0.1 | Iris-setosa |
| | 5.1 | 3.5 | 1.4 | 0.3 | Iris-setosa |
| ***mean*** | 4.985 | 3.404 | 1.456 | 0.241 | |
| ***std. dev.*** | 0.405 | 0.427 | 0.203 | 0.112 | |
| ***count*** | 27 | | | | |

**Training Data for Class = Iris-versicolor (2)**

| | x1 | x2 | x3 | x4 | C |
|---|---|---|---|---|---|
| | **sepal length** | **sepal width** | **petal length** | **petal width** | **class** |
| | 6.3 | 2.3 | 4.4 | 1.3 | Iris-versicolor |
| | 6.2 | 2.9 | 4.3 | 1.3 | Iris-versicolor |
| | 5.6 | 2.9 | 3.6 | 1.3 | Iris-versicolor |
| | 6.3 | 2.5 | 4.9 | 1.5 | Iris-versicolor |
| | 6.1 | 2.8 | 4 | 1.3 | Iris-versicolor |
| | 5.2 | 2.7 | 3.9 | 1.4 | Iris-versicolor |
| | 5.7 | 2.8 | 4.1 | 1.3 | Iris-versicolor |
| | 7 | 3.2 | 4.7 | 1.4 | Iris-versicolor |
| | 5.7 | 2.9 | 4.2 | 1.3 | Iris-versicolor |
| | 5.7 | 3 | 4.2 | 1.2 | Iris-versicolor |
| | 6.4 | 3.2 | 4.5 | 1.5 | Iris-versicolor |
| | 6.2 | 2.2 | 4.5 | 1.5 | Iris-versicolor |
| | 5.8 | 2.6 | 4 | 1.2 | Iris-versicolor |
| | 6.7 | 3.1 | 4.4 | 1.4 | Iris-versicolor |
| | 5.8 | 2.7 | 3.9 | 1.2 | Iris-versicolor |
| | 5.5 | 2.5 | 4 | 1.3 | Iris-versicolor |
| | 6.5 | 2.8 | 4.6 | 1.5 | Iris-versicolor |
| | 5 | 2.3 | 3.3 | 1 | Iris-versicolor |
| | 6 | 2.9 | 4.5 | 1.5 | Iris-versicolor |
| | 6.1 | 2.8 | 4.7 | 1.2 | Iris-versicolor |
| | 6.9 | 3.1 | 4.9 | 1.5 | Iris-versicolor |
| | 5.5 | 2.4 | 3.8 | 1.1 | Iris-versicolor |
| | 6.3 | 3.3 | 4.7 | 1.6 | Iris-versicolor |
| | 6.6 | 3 | 4.4 | 1.4 | Iris-versicolor |
| | 5.8 | 2.7 | 4.1 | 1 | Iris-versicolor |
| | 6 | 3.4 | 4.5 | 1.6 | Iris-versicolor |
| | 5.1 | 2.5 | 3 | 1.1 | Iris-versicolor |
| ***mean*** | 6.000 | 2.796 | 4.226 | 1.330 | |
| ***std. dev.*** | 0.517 | 0.316 | 0.458 | 0.168 | |
| ***count*** | 27 | | | | |

**Training Data for Class = Iris-virginica (3)**

| | x1 | x2 | x3 | x4 | C |
|---|---|---|---|---|---|
| | **sepal length** | **sepal width** | **petal length** | **petal width** | **class** |
| | 5.8 | 2.7 | 5.1 | 1.9 | Iris-virginica |
| | 6.8 | 3.2 | 5.9 | 2.3 | Iris-virginica |
| | 5.9 | 3 | 5.1 | 1.8 | Iris-virginica |
| | 6.4 | 3.2 | 5.3 | 2.3 | Iris-virginica |
| | 6.9 | 3.1 | 5.1 | 2.3 | Iris-virginica |
| | 6.3 | 2.8 | 5.1 | 1.5 | Iris-virginica |
| | 5.6 | 2.8 | 4.9 | 2 | Iris-virginica |
| | 4.9 | 2.5 | 4.5 | 1.7 | Iris-virginica |
| | 7.9 | 3.8 | 6.4 | 2 | Iris-virginica |
| | 6.5 | 3 | 5.2 | 2 | Iris-virginica |
| | 7.7 | 2.8 | 6.7 | 2 | Iris-virginica |
| | 6.3 | 2.9 | 5.6 | 1.8 | Iris-virginica |
| | 6.7 | 2.5 | 5.8 | 1.8 | Iris-virginica |
| | 6.1 | 3 | 4.9 | 1.8 | Iris-virginica |
| | 7.7 | 2.6 | 6.9 | 2.3 | Iris-virginica |
| | 6.4 | 2.8 | 5.6 | 2.2 | Iris-virginica |
| | 6.3 | 2.7 | 4.9 | 1.8 | Iris-virginica |
| | 7.7 | 3 | 6.1 | 2.3 | Iris-virginica |
| | 7.3 | 2.9 | 6.3 | 1.8 | Iris-virginica |
| | 6.3 | 3.4 | 5.6 | 2.4 | Iris-virginica |
| | 6.9 | 3.2 | 5.7 | 2.3 | Iris-virginica |
| ***mean*** | 6.590 | 2.948 | 5.557 | 2.014 | |
| ***std. dev.*** | 0.767 | 0.308 | 0.645 | 0.257 | |
| ***count*** | 21 | | | | |

# Using the Bayes classifier to predict the *iris* testing data

| | x1 | x2 | x3 | x4 | C | Predicted C | Class = Iris-setosa (1) | | | | | | Class = Iris-versicolor (2) | | | | | | Class = Iris-virginica (3) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | sepal length | sepal width | petal length | petal width | class | predicted class | P(x1\|C=1) | P(x2\|C=1) | P(x3\|C=1) | P(x4\|C=1) | P(C=1) | P(C=1\|x) | P(x1\|C=2) | P(x2\|C=2) | P(x3\|C=2) | P(x4\|C=2) | P(C=2) | P(C=2\|x) | P(x1\|C=3) | P(x2\|C=3) | P(x3\|C=3) | P(x4\|C=3) | P(C=3) | P(C=3\|x) | Pmax | Decision |
| 1 | 5.1 | 3.5 | 1.4 | 0.2 | Iris-setosa | Iris-setosa | 0.95 | 0.91 | 1.90 | 3.34 | 0.333 | **1.82** | 0.17 | 0.11 | 0.00 | 0.00 | 0.333 | **0.00** | 0.08 | 0.26 | 0.00 | 0.00 | 0.333 | **0.00** | 1.82 | Iris-setosa |
| 2 | 5.6 | 2.7 | 4.2 | 1.3 | Iris-versicolor | Iris-versicolor | 0.31 | 0.24 | 0.00 | 0.00 | 0.333 | **0.00** | 0.57 | 1.21 | 0.87 | 2.33 | 0.333 | **0.47** | 0.23 | 0.94 | 0.07 | 0.03 | 0.333 | **0.00** | 0.47 | Iris-versicolor |
| 3 | 6.7 | 3 | 5.2 | 2.3 | Iris-virginica | Iris-virginica | 0.00 | 0.60 | 0.00 | 0.00 | 0.333 | **0.00** | 0.31 | 1.03 | 0.09 | 0.00 | 0.333 | **0.00** | 0.52 | 1.28 | 0.53 | 0.84 | 0.333 | **0.10** | 0.10 | Iris-virginica |
| 4 | 6.1 | 3 | 4.6 | 1.4 | Iris-versicolor | Iris-versicolor | 0.02 | 0.60 | 0.00 | 0.00 | 0.333 | **0.00** | 0.76 | 1.03 | 0.62 | 2.17 | 0.333 | **0.35** | 0.42 | 1.28 | 0.21 | 0.09 | 0.333 | **0.00** | 0.35 | Iris-versicolor |
| 5 | 4.4 | 3.2 | 1.3 | 0.2 | Iris-setosa | Iris-setosa | 0.35 | 0.83 | 1.47 | 3.34 | 0.333 | **0.47** | 0.01 | 0.56 | 0.00 | 0.00 | 0.333 | **0.00** | 0.01 | 0.93 | 0.00 | 0.00 | 0.333 | **0.00** | 0.47 | Iris-setosa |
| 6 | 4.9 | 3.1 | 1.5 | 0.1 | Iris-setosa | Iris-setosa | 0.96 | 0.73 | 1.92 | 1.62 | 0.333 | **0.72** | 0.08 | 0.80 | 0.00 | 0.00 | 0.333 | **0.00** | 0.05 | 1.15 | 0.00 | 0.00 | 0.333 | **0.00** | 0.72 | Iris-setosa |
| 7 | 6.1 | 2.9 | 4.7 | 1.4 | Iris-versicolor | Iris-versicolor | 0.02 | 0.47 | 0.00 | 0.00 | 0.333 | **0.00** | 0.76 | 1.20 | 0.51 | 2.17 | 0.333 | **0.34** | 0.42 | 1.28 | 0.26 | 0.09 | 0.333 | **0.00** | 0.34 | Iris-versicolor |
| 8 | 7.2 | 3.2 | 6 | 1.8 | Iris-virginica | Iris-virginica | 0.00 | 0.83 | 0.00 | 0.00 | 0.333 | **0.00** | 0.05 | 0.56 | 0.00 | 0.05 | 0.333 | **0.00** | 0.38 | 0.93 | 0.49 | 1.10 | 0.333 | **0.06** | 0.06 | Iris-virginica |
| 9 | 5.4 | 3.9 | 1.7 | 0.4 | Iris-setosa | Iris-setosa | 0.58 | 0.48 | 0.95 | 1.29 | 0.333 | **0.11** | 0.39 | 0.00 | 0.00 | 0.00 | 0.333 | **0.00** | 0.16 | 0.01 | 0.00 | 0.00 | 0.333 | **0.00** | 0.11 | Iris-setosa |
| 10 | 6.7 | 3.1 | 4.7 | 1.5 | Iris-versicolor | Iris-versicolor | 0.00 | 0.73 | 0.00 | 0.00 | 0.333 | **0.00** | 0.31 | 0.80 | 0.51 | 1.42 | 0.333 | **0.06** | 0.52 | 1.15 | 0.26 | 0.21 | 0.333 | **0.01** | 0.06 | Iris-versicolor |
| 11 | 5.6 | 3 | 4.1 | 1.3 | Iris-versicolor | Iris-versicolor | 0.31 | 0.60 | 0.00 | 0.00 | 0.333 | **0.00** | 0.57 | 1.03 | 0.84 | 2.33 | 0.333 | **0.38** | 0.23 | 1.28 | 0.05 | 0.03 | 0.333 | **0.00** | 0.38 | Iris-versicolor |
| 12 | 5 | 3.6 | 1.4 | 0.2 | Iris-setosa | Iris-setosa | 0.98 | 0.84 | 1.90 | 3.34 | 0.333 | **1.75** | 0.12 | 0.05 | 0.00 | 0.00 | 0.333 | **0.00** | 0.06 | 0.14 | 0.00 | 0.00 | 0.333 | **0.00** | 1.75 | Iris-setosa |
| 13 | 6.7 | 3.3 | 5.7 | 2.5 | Iris-virginica | Iris-virginica | 0.00 | 0.91 | 0.00 | 0.00 | 0.333 | **0.00** | 0.31 | 0.35 | 0.00 | 0.00 | 0.333 | **0.00** | 0.52 | 0.67 | 0.60 | 0.26 | 0.333 | **0.02** | 0.02 | Iris-virginica |
| 14 | 4.8 | 3 | 1.4 | 0.3 | Iris-setosa | Iris-setosa | 0.89 | 0.60 | 1.90 | 3.10 | 0.333 | **1.04** | 0.05 | 1.03 | 0.00 | 0.00 | 0.333 | **0.00** | 0.03 | 1.28 | 0.00 | 0.00 | 0.333 | **0.00** | 1.04 | Iris-setosa |
| 15 | 5.4 | 3.9 | 1.3 | 0.4 | Iris-setosa | Iris-setosa | 0.58 | 0.48 | 1.47 | 1.29 | 0.333 | **0.18** | 0.39 | 0.00 | 0.00 | 0.00 | 0.333 | **0.00** | 0.16 | 0.01 | 0.00 | 0.00 | 0.333 | **0.00** | 0.18 | Iris-setosa |
| 16 | 5.1 | 3.3 | 1.7 | 0.5 | Iris-setosa | Iris-setosa | 0.95 | 0.91 | 0.95 | 0.24 | 0.333 | **0.07** | 0.17 | 0.35 | 0.00 | 0.00 | 0.333 | **0.00** | 0.08 | 0.67 | 0.00 | 0.00 | 0.333 | **0.00** | 0.07 | Iris-setosa |
| 17 | 6 | 3 | 4.8 | 1.8 | Iris-virginica | Iris-virginica | 0.04 | 0.60 | 0.00 | 0.00 | 0.333 | **0.00** | 0.77 | 1.03 | 0.40 | 0.05 | 0.333 | **0.01** | 0.39 | 1.28 | 0.31 | 1.10 | 0.333 | **0.06** | 0.06 | Iris-virginica |
| 18 | 5.9 | 3 | 4.2 | 1.5 | Iris-versicolor | Iris-versicolor | 0.08 | 0.60 | 0.00 | 0.00 | 0.333 | **0.00** | 0.76 | 1.03 | 0.87 | 1.42 | 0.333 | **0.32** | 0.35 | 1.28 | 0.07 | 0.21 | 0.333 | **0.00** | 0.32 | Iris-versicolor |
| 19 | 4.6 | 3.2 | 1.4 | 0.2 | Iris-setosa | Iris-setosa | 0.63 | 0.83 | 1.90 | 3.34 | 0.333 | **1.10** | 0.02 | 0.56 | 0.00 | 0.00 | 0.333 | **0.00** | 0.02 | 0.93 | 0.00 | 0.00 | 0.333 | **0.00** | 1.10 | Iris-setosa |
| 20 | 6.7 | 3 | 5 | 1.7 | Iris-versicolor | Iris-virginica | 0.00 | 0.60 | 0.00 | 0.00 | 0.333 | **0.00** | 0.31 | 1.03 | 0.21 | 0.21 | 0.333 | **0.00** | 0.52 | 1.28 | 0.43 | 0.74 | 0.333 | **0.07** | 0.07 | Iris-virginica |
| 21 | 6.8 | 3 | 5.5 | 2.1 | Iris-virginica | Iris-virginica | 0.00 | 0.60 | 0.00 | 0.00 | 0.333 | **0.00** | 0.23 | 1.03 | 0.02 | 0.00 | 0.333 | **0.00** | 0.50 | 1.28 | 0.62 | 1.47 | 0.333 | **0.19** | 0.19 | Iris-virginica |
| 22 | 6.9 | 3.1 | 5.4 | 2.1 | Iris-virginica | Iris-virginica | 0.00 | 0.73 | 0.00 | 0.00 | 0.333 | **0.00** | 0.17 | 0.80 | 0.03 | 0.00 | 0.333 | **0.00** | 0.48 | 1.15 | 0.60 | 1.47 | 0.333 | **0.16** | 0.16 | Iris-virginica |
| 23 | 6.4 | 2.7 | 5.3 | 1.9 | Iris-virginica | Iris-virginica | 0.00 | 0.24 | 0.00 | 0.00 | 0.333 | **0.00** | 0.57 | 1.21 | 0.06 | 0.01 | 0.333 | **0.00** | 0.50 | 0.94 | 0.57 | 1.40 | 0.333 | **0.13** | 0.13 | Iris-virginica |
| 24 | 5.2 | 3.4 | 1.4 | 0.2 | Iris-setosa | Iris-setosa | 0.86 | 0.94 | 1.90 | 3.34 | 0.333 | **1.69** | 0.23 | 0.20 | 0.00 | 0.00 | 0.333 | **0.00** | 0.10 | 0.44 | 0.00 | 0.00 | 0.333 | **0.00** | 1.69 | Iris-setosa |
| | | | | | | | | | | | | | … | | | | | | | | | | | | | |
| 51 | 5.4 | 3 | 4.5 | 1.5 | Iris-versicolor | Iris-versicolor | 0.58 | 0.60 | 0.00 | 0.00 | 0.333 | **0.00** | 0.39 | 1.03 | 0.73 | 1.42 | 0.333 | **0.14** | 0.16 | 1.28 | 0.16 | 0.21 | 0.333 | **0.00** | 0.14 | Iris-versicolor |
| 52 | 7.2 | 3.6 | 6.1 | 2.5 | Iris-virginica | Iris-virginica | 0.00 | 0.84 | 0.00 | 0.00 | 0.333 | **0.00** | 0.05 | 0.05 | 0.00 | 0.00 | 0.333 | **0.00** | 0.38 | 0.14 | 0.43 | 0.26 | 0.333 | **0.00** | 0.00 | Iris-virginica |
| 53 | 5.7 | 2.6 | 3.5 | 1 | Iris-versicolor | Iris-versicolor | 0.21 | 0.16 | 0.00 | 0.00 | 0.333 | **0.00** | 0.65 | 1.04 | 0.25 | 0.35 | 0.333 | **0.02** | 0.27 | 0.68 | 0.00 | 0.00 | 0.333 | **0.00** | 0.02 | Iris-versicolor |
| 54 | 6.8 | 2.8 | 4.8 | 1.4 | Iris-versicolor | Iris-versicolor | 0.00 | 0.34 | 0.00 | 0.00 | 0.333 | **0.00** | 0.23 | 1.26 | 0.40 | 2.17 | 0.333 | **0.08** | 0.50 | 1.16 | 0.31 | 0.09 | 0.333 | **0.01** | 0.08 | Iris-versicolor |
| 55 | 6.7 | 3.3 | 5.7 | 2.1 | Iris-virginica | Iris-virginica | 0.00 | 0.91 | 0.00 | 0.00 | 0.333 | **0.00** | 0.31 | 0.35 | 0.00 | 0.00 | 0.333 | **0.00** | 0.52 | 0.67 | 0.60 | 1.47 | 0.333 | **0.10** | 0.10 | Iris-virginica |
| 56 | 5.7 | 2.5 | 5 | 2 | Iris-virginica | Iris-virginica | 0.21 | 0.10 | 0.00 | 0.00 | 0.333 | **0.00** | 0.65 | 0.81 | 0.21 | 0.00 | 0.333 | **0.00** | 0.27 | 0.45 | 0.43 | 1.55 | 0.333 | **0.03** | 0.03 | Iris-virginica |
| 57 | 6 | 2.7 | 5.1 | 1.6 | Iris-versicolor | Iris-versicolor | 0.04 | 0.24 | 0.00 | 0.00 | 0.333 | **0.00** | 0.77 | 1.21 | 0.14 | 0.65 | 0.333 | **0.03** | 0.39 | 0.94 | 0.48 | 0.42 | 0.333 | **0.02** | 0.03 | Iris-versicolor |
| 58 | 5.5 | 2.4 | 3.7 | 1 | Iris-versicolor | Iris-versicolor | 0.44 | 0.06 | 0.00 | 0.00 | 0.333 | **0.00** | 0.48 | 0.57 | 0.45 | 0.35 | 0.333 | **0.01** | 0.19 | 0.27 | 0.01 | 0.00 | 0.333 | **0.00** | 0.01 | Iris-versicolor |
| 59 | 7.1 | 3 | 5.9 | 2.1 | Iris-virginica | Iris-virginica | 0.00 | 0.60 | 0.00 | 0.00 | 0.333 | **0.00** | 0.08 | 1.03 | 0.00 | 0.00 | 0.333 | **0.00** | 0.42 | 1.28 | 0.54 | 1.47 | 0.333 | **0.14** | 0.14 | Iris-virginica |
| 60 | 5.6 | 3 | 4.5 | 1.5 | Iris-versicolor | Iris-versicolor | 0.31 | 0.60 | 0.00 | 0.00 | 0.333 | **0.00** | 0.57 | 1.03 | 0.73 | 1.42 | 0.333 | **0.20** | 0.23 | 1.28 | 0.16 | 0.21 | 0.333 | **0.00** | 0.20 | Iris-versicolor |
| 61 | 7.6 | 3 | 6.6 | 2.1 | Iris-virginica | Iris-virginica | 0.00 | 0.60 | 0.00 | 0.00 | 0.333 | **0.00** | 0.01 | 1.03 | 0.00 | 0.00 | 0.333 | **0.00** | 0.22 | 1.28 | 0.17 | 1.47 | 0.333 | **0.02** | 0.02 | Iris-virginica |
| 62 | 5.1 | 3.8 | 1.5 | 0.3 | Iris-setosa | Iris-setosa | 0.95 | 0.61 | 1.92 | 3.10 | 0.333 | **1.14** | 0.17 | 0.01 | 0.00 | 0.00 | 0.333 | **0.00** | 0.08 | 0.03 | 0.00 | 0.00 | 0.333 | **0.00** | 1.14 | Iris-setosa |
| 63 | 6.5 | 3 | 5.8 | 2.2 | Iris-virginica | Iris-virginica | 0.00 | 0.60 | 0.00 | 0.00 | 0.333 | **0.00** | 0.48 | 1.03 | 0.00 | 0.00 | 0.333 | **0.00** | 0.52 | 1.28 | 0.58 | 1.19 | 0.333 | **0.15** | 0.15 | Iris-virginica |
| 64 | 7.7 | 3.8 | 6.7 | 2.2 | Iris-virginica | Iris-virginica | 0.00 | 0.61 | 0.00 | 0.00 | 0.333 | **0.00** | 0.00 | 0.01 | 0.00 | 0.00 | 0.333 | **0.00** | 0.18 | 0.03 | 0.13 | 1.19 | 0.333 | **0.00** | 0.00 | Iris-virginica |
| 65 | 5.8 | 2.7 | 5.1 | 1.9 | Iris-virginica | Iris-virginica | 0.13 | 0.24 | 0.00 | 0.00 | 0.333 | **0.00** | 0.72 | 1.21 | 0.14 | 0.01 | 0.333 | **0.00** | 0.31 | 0.94 | 0.48 | 1.40 | 0.333 | **0.06** | 0.06 | Iris-virginica |
| 66 | 6 | 2.2 | 5 | 1.5 | Iris-virginica | Iris-versicolor | 0.04 | 0.02 | 0.00 | 0.00 | 0.333 | **0.00** | 0.77 | 0.21 | 0.21 | 1.42 | 0.333 | **0.02** | 0.39 | 0.07 | 0.43 | 0.21 | 0.333 | **0.00** | 0.02 | Iris-versicolor |
| 67 | 5.1 | 3.7 | 1.5 | 0.4 | Iris-setosa | Iris-setosa | 0.95 | 0.73 | 1.92 | 1.29 | 0.333 | **0.58** | 0.17 | 0.02 | 0.00 | 0.00 | 0.333 | **0.00** | 0.08 | 0.07 | 0.00 | 0.00 | 0.333 | **0.00** | 0.58 | Iris-setosa |
| 68 | 6.2 | 3.4 | 5.4 | 2.3 | Iris-virginica | Iris-virginica | 0.01 | 0.94 | 0.00 | 0.00 | 0.333 | **0.00** | 0.72 | 0.20 | 0.03 | 0.00 | 0.333 | **0.00** | 0.46 | 0.44 | 0.60 | 0.84 | 0.333 | **0.03** | 0.03 | Iris-virginica |
| 69 | 6.6 | 2.9 | 4.6 | 1.3 | Iris-versicolor | Iris-versicolor | 0.00 | 0.47 | 0.00 | 0.00 | 0.333 | **0.00** | 0.39 | 1.20 | 0.62 | 2.33 | 0.333 | **0.23** | 0.52 | 1.28 | 0.21 | 0.03 | 0.333 | **0.00** | 0.23 | Iris-versicolor |
| 70 | 5.5 | 2.3 | 4 | 1.3 | Iris-versicolor | Iris-versicolor | 0.44 | 0.03 | 0.00 | 0.00 | 0.333 | **0.00** | 0.48 | 0.37 | 0.77 | 2.33 | 0.333 | **0.11** | 0.19 | 0.14 | 0.03 | 0.03 | 0.333 | **0.00** | 0.11 | Iris-versicolor |
| 71 | 5.5 | 4.2 | 1.4 | 0.2 | Iris-setosa | Iris-setosa | 0.44 | 0.16 | 1.90 | 3.34 | 0.333 | **0.15** | 0.48 | 0.00 | 0.00 | 0.00 | 0.333 | **0.00** | 0.19 | 0.00 | 0.00 | 0.00 | 0.333 | **0.00** | 0.15 | Iris-setosa |
| 72 | 6.3 | 3.3 | 6 | 2.5 | Iris-virginica | Iris-virginica | 0.01 | 0.91 | 0.00 | 0.00 | 0.333 | **0.00** | 0.65 | 0.35 | 0.00 | 0.00 | 0.333 | **0.00** | 0.48 | 0.67 | 0.49 | 0.26 | 0.333 | **0.01** | 0.01 | Iris-virginica |
| 73 | 5.6 | 2.5 | 3.9 | 1.1 | Iris-versicolor | Iris-versicolor | 0.31 | 0.10 | 0.00 | 0.00 | 0.333 | **0.00** | 0.57 | 0.81 | 0.68 | 0.93 | 0.333 | **0.10** | 0.23 | 0.45 | 0.02 | 0.00 | 0.333 | **0.00** | 0.10 | Iris-versicolor |
| 74 | 5.8 | 2.8 | 5.1 | 2.4 | Iris-virginica | Iris-virginica | 0.13 | 0.34 | 0.00 | 0.00 | 0.333 | **0.00** | 0.72 | 1.26 | 0.14 | 0.00 | 0.333 | **0.00** | 0.31 | 1.16 | 0.48 | 0.50 | 0.333 | **0.03** | 0.03 | Iris-virginica |
| 75 | 4.9 | 2.4 | 3.3 | 1 | Iris-versicolor | Iris-versicolor | 0.96 | 0.06 | 0.00 | 0.00 | 0.333 | **0.00** | 0.08 | 0.57 | 0.11 | 0.35 | 0.333 | **0.00** | 0.05 | 0.27 | 0.00 | 0.00 | 0.333 | **0.00** | 0.00 | Iris-versicolor |

# Today: Classification and Regression, Part 1

1. Basic definitions
2. Bayesian classification
    1. Problem formulation
    2. Bayes classification in Excel
    3. Bayes classification in Python

# Bayes classification in Python

```python
from sklearn.naive_bayes import GaussianNB
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
import numpy as np
import pandas as pd

#%% load the data
data = pd.read_excel("iris.mini.xlsx",
  sheet_name="iris (all)", usecols="F:J", header=1)
# use only 2 features (sepal length, sepal width)
# in order to make the data 2D to allow visualization
data = data.drop(["petal length", "petal width"], axis=1)
data.head()


#%% separate into features and labels; training and testing data
features = data.drop(["class"], axis=1)
X = np.array(features)
y = np.array(data["class"])

X_trn, X_tst, y_trn, y_tst = train_test_split(X, y, test_size=0.5)


#%% fit the model and test the prediction
gnb = GaussianNB()
gnb.fit(X_trn, y_trn)

print("means:", gnb.theta_, "\n")
print("stds:", np.sqrt(gnb.sigma_), "\n")

y_trn_prd = gnb.predict(X_trn)
y_tst_prd = gnb.predict(X_tst)

print("trn acc =", accuracy_score(y_true=y_trn, y_pred=y_trn_prd))
print("tst acc =", accuracy_score(y_true=y_tst, y_pred=y_tst_prd))
```

| | sepal length | sepal width | class |
|---|---|---|---|
| 0 | 5.8 | 2.7 | Iris-virginica |
| 1 | 4.6 | 3.4 | Iris-setosa |
| 2 | 6.3 | 2.3 | Iris-versicolor |
| 3 | 6.8 | 3.2 | Iris-virginica |
| 4 | 6.2 | 2.9 | Iris-versicolor |

```
means: [[4.95833333 3.32916667]
 [6.00384615 2.8       ]
 [6.508      2.996     ]]

stds: [[0.35227436 0.38131261]
 [0.50419835 0.30382181]
 [0.64244533 0.33163836]]

trn acc = 0.76
tst acc = 0.8133333333333334
```

# Bayes classification in Python

```
#%% plot data and the decision areas

c_dict = {
  "Iris-setosa": 0,
  "Iris-versicolor": 0.5,
  "Iris-virginica": 1
}

clrs_y = np.zeros(y.shape)
for idx in range(len(y)):
  clrs_y[idx] = c_dict[y[idx]]

ax = plt.subplot(111)
plt.xlabel(features.columns[0])
plt.ylabel(features.columns[1])

x_min = X[:,0].min() - 0.5
x_max = X[:,0].max() + 0.5
y_min = X[:,1].min() - 0.5
y_max = X[:,1].max() + 0.5
xx, yy = np.meshgrid(
    np.arange(x_min, x_max, 0.02),
    np.arange(y_min, y_max, 0.02)
    )

Z = gnb.predict_proba(np.c_[xx.ravel(), yy.ravel()])
Z = Z.argmax(axis=1)
Z = Z.reshape(xx.shape)

ax.contourf(xx, yy, Z, cmap="plasma", alpha=0.5)
ax.scatter(X[:,0], X[:,1], c=clrs_y, cmap="plasma",
  alpha=0.5, edgecolors="k", s=40)
```