Seminar - Bridge Game

Interactive Exercises：Classification

张绍群

2023/11/27

Machine Learning —— Classification

给定一个示例，将其归为垃圾邮件或者非垃圾邮件

例 0.40 (贝叶斯 Spam 过滤器) 如何确定一个电子邮件是 Spam?

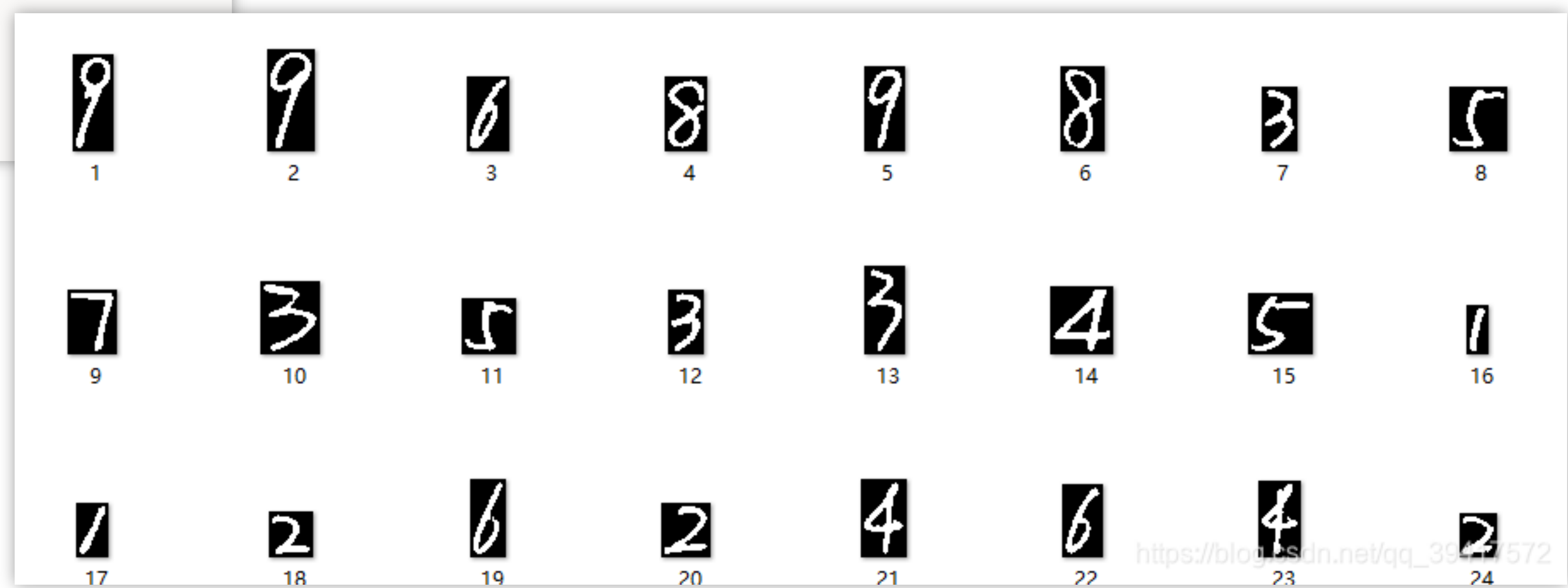
- 假设我们有一个垃圾邮件的集合 B 和一个不是垃圾的邮件集合 G. 利用贝叶斯公式来预测一个新的电子邮件是 Spam 的概率.
- 考察一个特定的单词 ω, 统计该单词在集合 B 和 G 中出现的次数分别为 $n_B(\omega)$ 和 $n_G(\omega)$.
- 设 S 是事件：邮件为 Spam, E 是事件：邮件内容含单词 ω. 需计算 $P(S \mid E)$.

思路: 根据贝叶斯公式, 我们需要分别估算

- Spam 邮件中含有单词 ω 的概率 $P(E \mid S)$
- 非 Spam 邮件中含有单词 ω 的概率 $P(E \mid \bar{S})$
- 比较这两者的大小

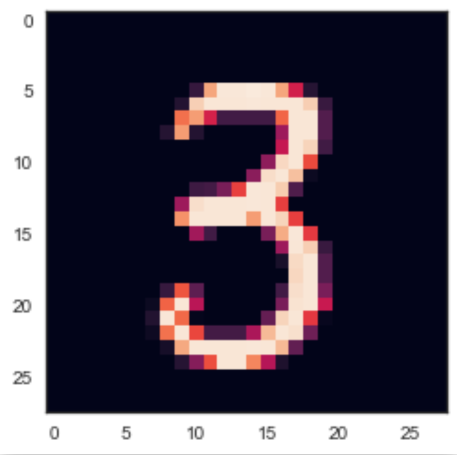
给定 something —— 称为输入样本 (Input)

给定一个手写字符的图片，将其分类为一个已知字符



输出一个类别 —— Label

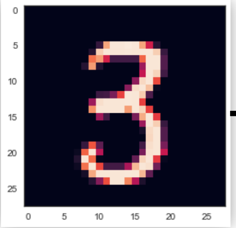
Machine Learning —— Classification



- **对于 Input 来讲，最重要的是 features**
  - 在垃圾邮件分类任务中，所要查找的单词，feature 是 $\boldsymbol{\omega}$
  - 在图像识别任务中，输入的是一张图片，features 是相对应的像素

- **对于 Output**
  - 在垃圾邮件分类任务中，“**是**” or “**否**” 垃圾邮件
  - 在图像识别任务中，输出是相对应的**数字**，*e.g.*，1，2，3，⋯

Input



ML model？

Output

3

Binary Classification with Two Features

- 抽象：$(X, Y) \rightarrow Z$

Alibaba Cloud | TIANCHI 天池

Index	Height(Inches)	Weight(Pounds)	Gender
1	65.78331	112.9925	1
2	71.51521	136.4873	0
3	69.39874	153.0269	0
4	68.2166	142.3354	1
5	67.78781	144.2971	1
6	68.69784	123.3024	1
7	69.80204	141.4947	1
8	70.01472	136.4623	1

X 表示 Height
Y 表示 Weight
features

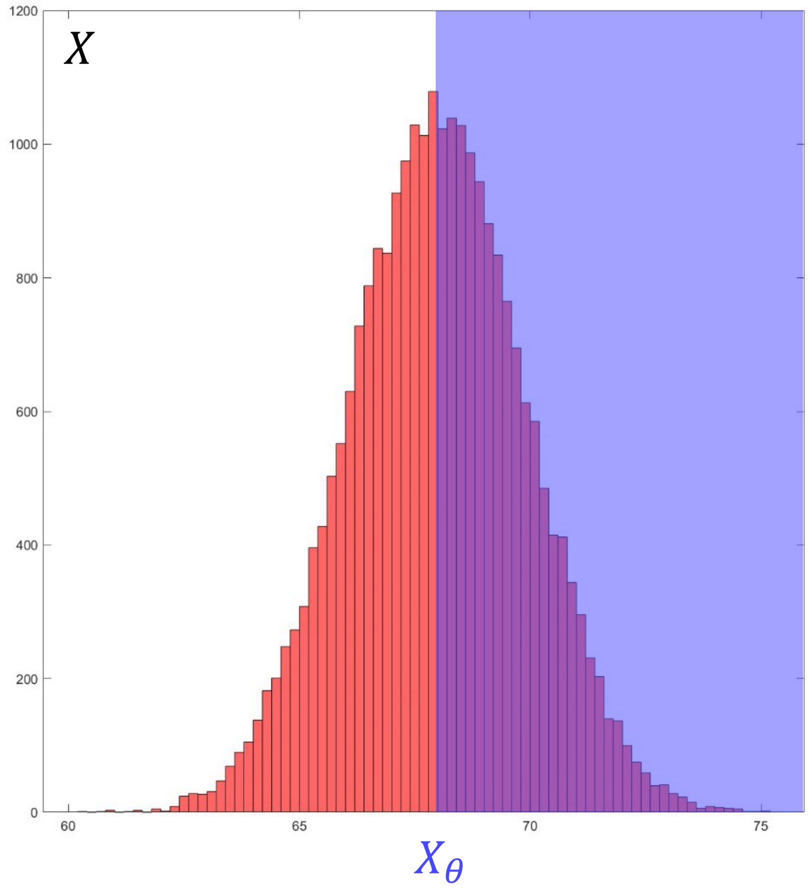
Z 表示 Gender (0 表示女生，1 表示男生)

- 根据《概率论与数理统计》所学的知识，有哪些方法可以实现这个分类任务？

比如：给定 $(X = 70.1111, Y = 141.3333)$，问：对应的 Z 应该是多少？
- $Z = 1$ 的概率？
- $Z = 0$ 的概率？

Classification with Univariate Feature

- $X \rightarrow Z \cap Y \rightarrow Z$



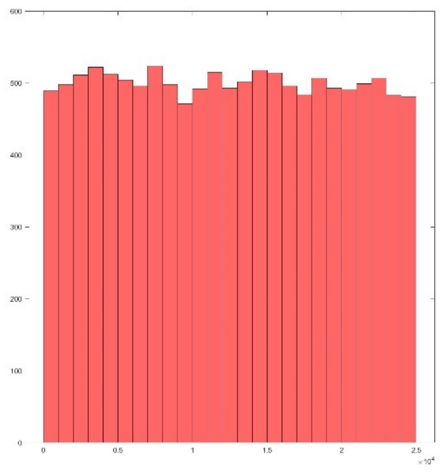
IF Red refers to 1, while Blue corresponds to 0

Then we have

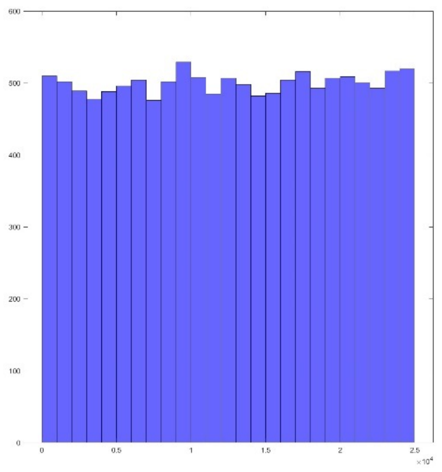
$$X \geq X_\theta \rightarrow Z = 1$$
$$X < X_\theta \rightarrow Z = 0$$

Obviously, this case is ideal.

Girl



Boy

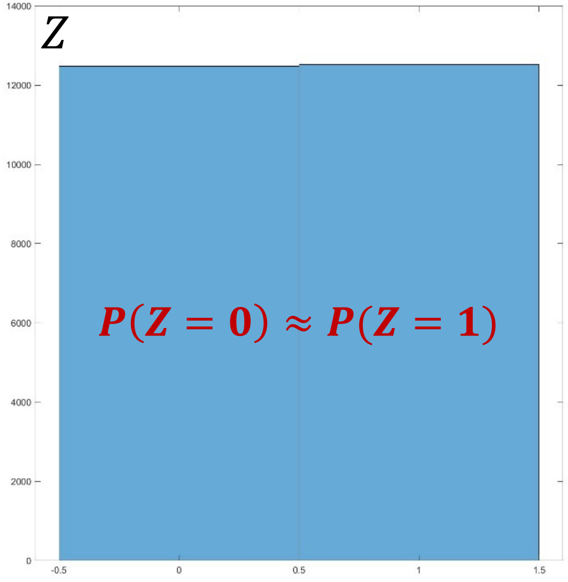


Classification with Univariate Feature

- Conditional Probability:
  - First, compute $\mathrm{P}(\mathrm{Z} \mid \mathrm{X} = 70.1111)$
  - *Second, make a decision by compare* $\mathrm{P}(\mathrm{Z} = 1 \mid \mathrm{X} = 70.1111)$ with $\mathrm{P}(\mathrm{Z} = 0 \mid \mathrm{X} = 70.1111)$

$$\mathrm{P}(\mathrm{Z} = 1 \mid \mathrm{X} = 70.1111) = \frac{P(Z = 1, X = 70.1111)}{P(X = 70.1111)} = \frac{P(Z = 1)P(X = 70.1111 \mid Z = 1)}{P(X = 70.1111)}$$

$$\mathrm{P}(\mathrm{Z} = 0 \mid \mathrm{X} = 70.1111) = \frac{P(Z = 0, X = 70.1111)}{P(X = 70.1111)} = \frac{P(Z = 0)P(X = 70.1111 \mid Z = 0)}{P(X = 70.1111)}$$



Compare $P(X = 70.1111 \mid Z = 0)$ with $P(X = 70.1111 \mid Z = 1)$

$P(X = 70.1111 \mid Z = 0) = 0$ $\quad$ $P(X = 70.1111 \mid Z = 1) = 0$

What ?!!!

Classification with Univariate Feature

- Conditional Probability:
  - First, compute $\mathrm{P}(\mathrm{Z} \mid \mathrm{X} = 70.1111)$
  - Second, make a decision by compare $\mathrm{P}(\mathrm{Z} = 1 \mid \mathrm{X} = 70.1111)$ with $\mathrm{P}(\mathrm{Z} = 0 \mid \mathrm{X} = 70.1111)$
- Improvement of Discretization :
  - **Discrete $\boldsymbol{X}$, Discrete $\boldsymbol{Z}$**
  - Partition the range of variable X as $x_1, x_2, \dots, x_n$. For example, horizon $\delta = 0.5$, $x_1 = [60, 60.5)$
  - Find $70.1111 \in x_k$
  - Compute $\mathrm{P}(\mathrm{Z} = 1 \mid \mathrm{X} = x_\mathrm{k})$ replacing $\mathrm{P}(\mathrm{Z} = 1 \mid \mathrm{X} = 70.1111)$
  - Compare $P(X = x_k \mid Z = 0)$ with $P(X = x_k \mid Z = 1)$

$P(X = x_k \mid Z = 0) = 0.0439$ **<** $P(X = x_k \mid Z = 1) = 0.0586$ Identify $Z = 1$?

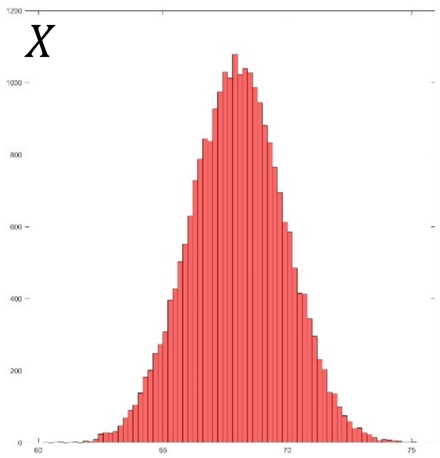
- Depend heavily on the setting of the discretization horizon

horizon 0.2: 0.0213 < 0.0254
horizon 0.12: 0.0144 > 0.0141

Classification with Univariate Feature

- Conditional Probability:
  - First, compute $\mathrm{P}(\mathrm{Z} \mid \mathrm{X} = 70.1111)$
  - Second, make a decision by compare $\mathrm{P}(\mathrm{Z} = 1 \mid \mathrm{X} = 70.1111)$ with $\mathrm{P}(\mathrm{Z} = 0 \mid \mathrm{X} = 70.1111)$
- Improvement of Continuum:
  - Model (X, Z)
  - $\mathrm{P}(\mathrm{Z} \leq z \mid \mathrm{X} = x) = \int_{-\infty}^{z} \frac{f(x,z)}{f_X(x)} dz$?
  - We have challenges:
    - Z is discrete
    - It is hard to obtain $f(x, z)$.

Assume Gaussian



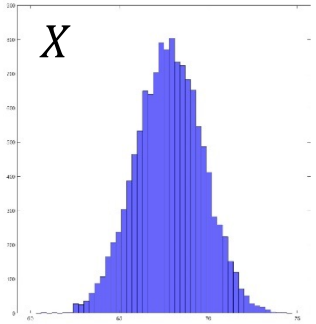
Classification with Univariate Feature

- Conditional Probability:
  - First, compute P(Z | X = 70.1111)
  - Second, make a decision by compare P(Z = 1 | X = 70.1111) with P(Z = 0 | X = 70.1111)
- Improvement of Continuum:
  - Model (X, Z)
  - **Continuous $\boldsymbol{X}$, Discrete $\boldsymbol{Z}$**

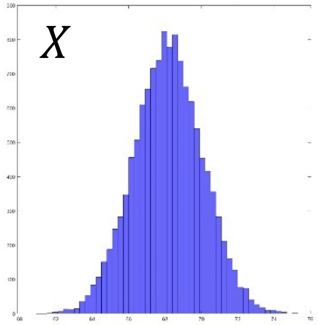
$$\mathrm{P}(\mathrm{Z}=1\mid \mathrm{X}=x)=\frac{P(Z=1,X=x)}{P(X=x)}=\frac{P(Z=1)P(X=x\mid Z=1)}{P(X=x)}$$

$$=\frac{P(Z=1)\int_{x}^{x+\delta} f_{X|Z}(x|z=1)dx}{P(X=x)}$$

$\boldsymbol{Z=0}$



$\boldsymbol{Z=1}$



Assume Gaussian

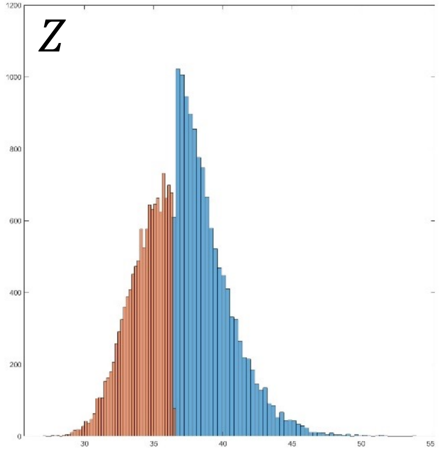
  - Steps:
    - ➢ X is Gaussian, $\boldsymbol{\mathcal{N}}(67.9931, 3.6164)$
    - ➢ $X|Z=0$ is Gaussian, $\boldsymbol{\mathcal{N}}(67.8077, 3.5384)$
    - ➢ $X|Z=1$ is Gaussian, $\boldsymbol{\mathcal{N}}(68.1786, 3.6258)$
    - ➢ Next ?

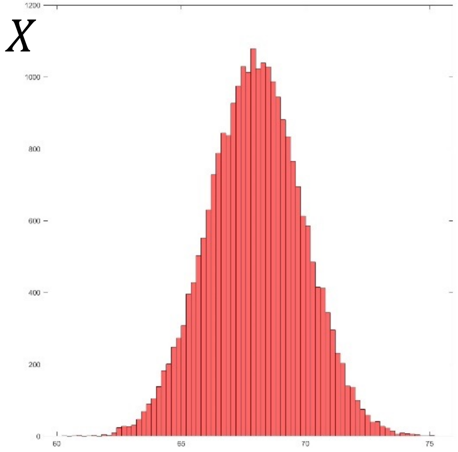
$$\frac{x-67.8007}{\sqrt{3.5384}} \quad \text{VS} \quad \frac{x-68.1786}{\sqrt{3.6258}}$$

$$1.2282-0 \quad > \quad 1.0148-0 \quad \Rightarrow \quad Z=0$$

Classification with Univariate Feature

- Conditional Probability:
  - First, compute P(Z | X = 70.1111)
  - Second, make a decision by compare P(Z = 1 | X = 70.1111) with P(Z = 0 | X = 70.1111)
- Improvement of Continuum:
  - Model (X, Z)
  - **Continuous $\boldsymbol{X}$, Continuous $\boldsymbol{Z}$**
  - $\mathrm{E}(\mathrm{Z} \mid \mathrm{X} = x) = \int_{-\infty}^{+\infty} \frac{f(x,z)}{f_X(x)} dz = \int_{-\infty}^{+\infty} z f_{Z|X}(z|X = x)\, dz$
  - Steps:
    - X is Gaussian, $\mathcal{N}(67.9931, 3.6164)$
    - Z also is continuous
    - compute $f_{Z|X}(z|x) = ?$





Assume Gaussian

Even though we assume a continous $\boldsymbol{Z}$,
we cannot obtain the correlation coefficients
between $\boldsymbol{X}$ and $\boldsymbol{Z}$, RIGHT?

条件期望的性质

- 线性性. 对任意常数 a, b 有 $\mathbb{E}(aX_1+bX_2|Y)=a\mathbb{E}(X_1|Y)+b\mathbb{E}(X_2|Y)$;

- 函数型. 对离散型随机向量 (X,Y) 和函数 $g(X)$, 有

$$\mathbb{E}(g(X)|Y)=\sum_i g(x_i)P(X=x_i|Y=y)$$

  对连续型随机向量 (X,Y) 和函数 $g(X)$, 有

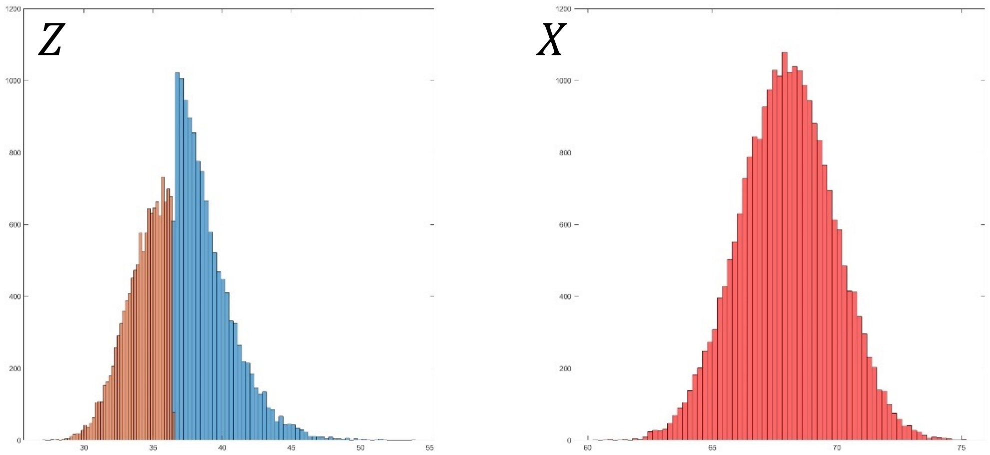
$$\mathbb{E}(g(X)|Y)=\int_{-\infty}^{+\infty} g(x)f(x|Y=y)dx$$

- 若随机向量 $(X,Y)\sim\mathcal{N}(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho)$, 则在 $Y=y$ 的条件下随机变量 X 服从正态分布 $\mathcal{N}(\mu_x-\rho\sigma_x(y-\mu_y)/\sigma_y,(1-\rho^2)\sigma_x^2)$, 由此可得

$$\mathbb{E}(X|y)=\mu_x-\frac{\rho\sigma_x(y-\mu_y)}{\sigma_y}$$

Classification with Univariate Feature

- Conditional Probability:
  - First, compute $\mathrm{P}(\mathrm{Z} \mid \mathrm{X} = 70.1111)$
  - Second, make a decision by compare $\mathrm{P}(\mathrm{Z} = 1 \mid \mathrm{X} = 70.1111)$ with $\mathrm{P}(\mathrm{Z} = 0 \mid \mathrm{X} = 70.1111)$
- Improvement of Continuum:
  - Model (X, Z)
  - **Continuous $\boldsymbol{X}$, Continuous $\boldsymbol{Z}$**
  - $\mathrm{E}(\mathrm{Z} \mid \mathrm{X} = x) = \int_{-\infty}^{+\infty} z \frac{f(x,z)}{f_X(x)} dz$
  - Steps:
    - X is Gaussian, $\mathcal{N}(67.9931, 3.6164)$
    - Z is Gaussian, $\mathcal{N}(0.2027, 8.9741)$
    - we can get the correlation coefficient $\rho = 0.1103$
    - compute $(Z|X = x) \in \mathcal{N}(0.2027 - 0.1103 \times \frac{2.9957}{1.9017} \times (x - 67.9931), (1 - 0.1103^2) \times 8.9741)$
    - compute $(Z|X = 70.1111) \in \mathcal{N}(0.1653, 8.8645)$
    - Next ?



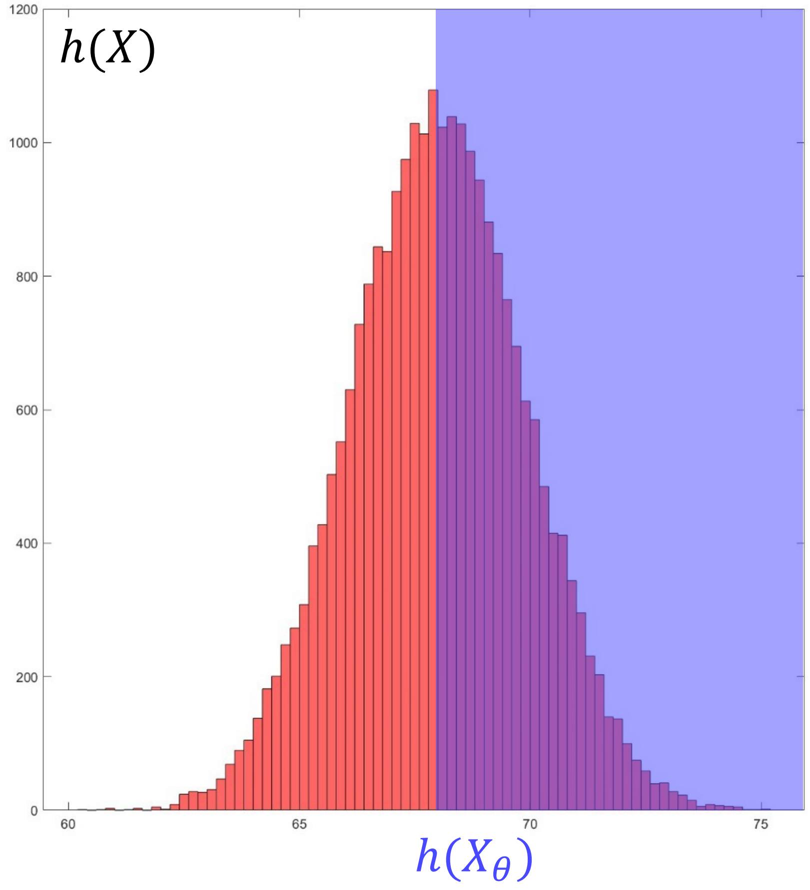
Assume Gaussian

Classification with Univariate Feature

- Possible: Is there a function h such that $h(X)$ is obviously separable?

$$X \rightarrow h(X) \rightarrow Z$$

$\boxed{\text{Q 1}}$

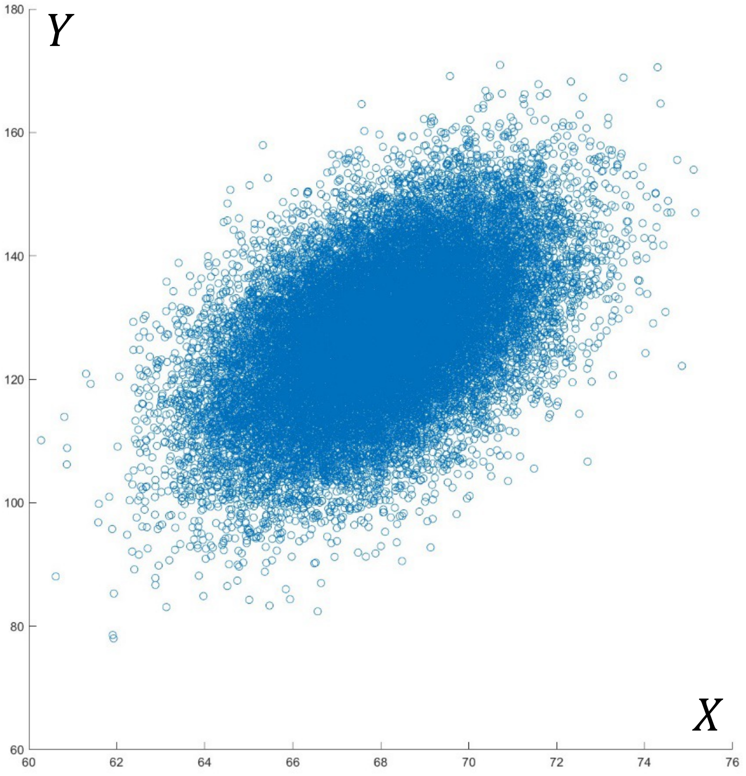


Classification with Multivariate Features

- $(X, Y) \to Z$?
  - Is X independent to Y?
  - IF X is **independent** to Y, THEN we should make a classification, separately
  - Select the maximum possibility

Classification with Multivariate Features

- $(X, Y) \to Z$?
  - Is X independent to Y?
  - It is observed that X is correlated to Y.

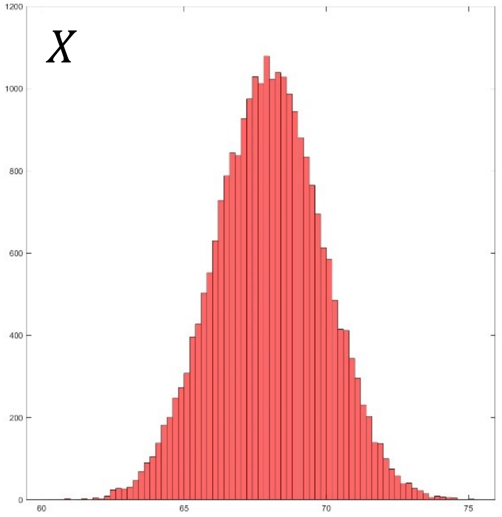


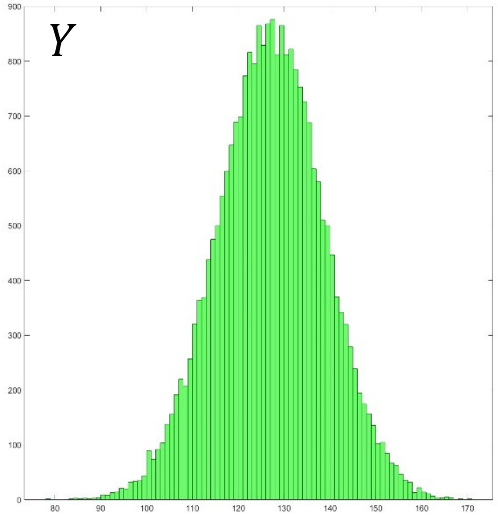
There are correlation, but we still cannot identify independency?

$$\mathbf{R}_{cor} = \begin{bmatrix} 1 & 0.529 \\ 0.529 & 1 \end{bmatrix}$$

$$\boldsymbol{\lambda}_1 = \frac{1}{2}, u_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\boldsymbol{\lambda}_2 = \frac{3}{2}, u_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$





Assume Gaussian

Classification with Multivariate Features

- $(X, Y) \rightarrow Z$?
  - IF X is **dependent** to Y, THEN we should make a classification jointly
  - **Discrete $(\boldsymbol{X}, \boldsymbol{Y})$, Discrete $\boldsymbol{Z}$**
    - Partition the range of variable X as $x_1, x_2, \dots, x_n$. For example, horizon $\delta_1 = 0.5$, $x_1 = [60, 60.5)$
      Partition the range of variable Y as $y_1, y_2, \dots, y_n$. For example, horizon $\delta_2 = 2$, $y_1 = [78, 80)$
    - Find $(70.1111, 141.3333) \in (x_i, y_j)$
    - Compute $\mathrm{P}(\mathrm{Z} = 0 \mid \mathrm{X} = x_i, \mathrm{Y} = y_j)$ and $\mathrm{P}(\mathrm{Z} = 1 \mid \mathrm{X} = x_i, \mathrm{Y} = y_j)$
    - Compare $P\left(\mathrm{X} = x_i, \mathrm{Y} = y_j \middle| Z = 0\right)$ with $P\left(\mathrm{X} = x_i, \mathrm{Y} = y_j \middle| Z = 1\right)$

$\Downarrow$ $\Downarrow$

0.0066 **>** 0 Identify $Z = 0$?

- Depend heavily on the setting of the discretization horizons

horizons (0.5,1.5): 0.050 > 0
horizons (0.5,10): 0.0196 > 0
horizons (0.2,15): 0.0161 > 0.0088

Classification with Multivariate Features

- $(X, Y) \rightarrow Z$?
  - IF X is **dependent** to Y, THEN we should make a classification jointly
  - Provided Gaussian assumptions, we can obtain the $f(x, y)$
    - X is Gaussian, $\mathcal{N}(67.9931, 3.6164)$
    - Y is Gaussian, $\mathcal{N}(127.0794, 135.9765)$
    - we can get the correlation coefficient $\rho = 0.529$
    - compute $(X, Y) \in \mathcal{N}(\,(67.9931{,}127.0794)^T, \mathbf{M}_{cov})$ $\qquad \mathbf{M}_{cov} = \begin{bmatrix} 3.6164 & 0.529 \\ 0.529 & 135.9765 \end{bmatrix}$
    - Next ?

- $(X, Y) \rightarrow Z$?
  - Compare the possibility
    $\mathrm{P}(X = 70.1111, Y = 141.3333 \mid Z = 0)$ and $\mathrm{P}(X = 70.1111, Y = 141.3333 \mid Z = 1)$
  - It is observed that both $(X, Y)|Z = 0$ and $(X, Y)|Z = 1$ obey Gaussian

Classification with Multivariate Features

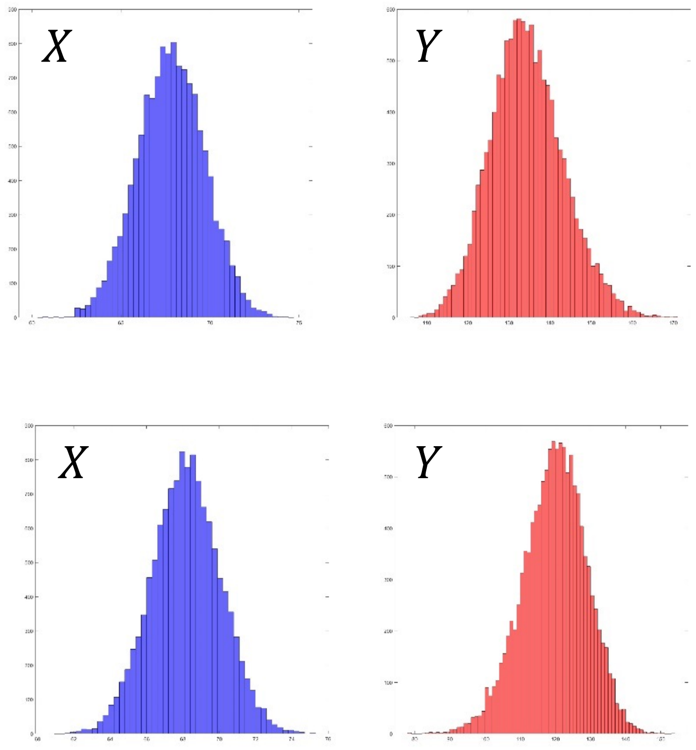
$Z = 0$

$\mathcal{N}\left(\left(\frac{x-67.8007}{\sqrt{3.5384}}, \frac{y-137.4883}{\sqrt{79.1715}}\right), ?\right)$ VS $\mathcal{N}\left(\left(\frac{x-68.1786}{\sqrt{3.6258}}, \frac{y-119.6706}{\sqrt{83.0022}}\right), ?\right)$

- $(X, Y) \rightarrow Z$?
  - Continuous $(\boldsymbol{X}, \boldsymbol{Y})$, Discrete $\boldsymbol{Z}$
  - Compare the possibility
    $\mathrm{P}(X = 70.1111, Y = 141.3333 \mid Z = 0)$ and $\mathrm{P}(X = 70.1111, Y = 141.3333 \mid Z = 1)$
  - It is observed that both $(X, Y)|Z = 0$ and $(X, Y)|Z = 1$ obey Gaussian

$\boldsymbol{Z = 0}$

$\boldsymbol{Z = 1}$



- Provided Gaussian assumptions, we can obtain the $f(x, y|Z = 0)$
  - $X|Z = 0$ is Gaussian, $\mathcal{N}(67.8077, 3.5384)$
  - $Y|Z = 0$ is Gaussian, $\mathcal{N}(137.4883, 79.1715)$
  - we can get the correlation coefficient $\rho = 0.729$
  - compute $(X|Z = 0, Y|Z = 0) \in \mathcal{N}\left((67.8077, 137.4883)^T, \begin{pmatrix} 3.5384 & 12.2023 \\ 12.2023 & 79.1715 \end{pmatrix}\right)$
- Provided Gaussian assumptions, we can obtain the $f(x, y|Z = 1)$
  - $X|Z = 1$ is Gaussian, $\mathcal{N}(68.1786, 3.6258)$
  - $Y|Z = 1$ is Gaussian, $\mathcal{N}(119.6706, 83.0022)$
  - we can get the correlation coefficient $\rho = 0.7407$
  - compute $(X|Z = 1, Y|Z = 1) \in \mathcal{N}\left((68.1786, 119.6706)^T, \begin{pmatrix} 3.6258 & 12.8488 \\ 12.8488 & 83.0022 \end{pmatrix}\right)$

$$\int_{70.1111}^{70.1111+\delta_1} \int_{141.3333}^{141.3333+\delta_2} f(x, y|Z = 0)\, dxdy \text{ VS } \int_{70.1111}^{70.1111+\delta_1} \int_{141.3333}^{141.3333+\delta_2} f(x, y|Z = 1)\, dxdy$$

Classification with Multivariate Features

$Z = 1$

- $(X, Y) \rightarrow Z$?

$$\mathcal{N}\left(\left(\frac{x - 67.8007}{\sqrt{3.5384}}, \frac{y - 137.4883}{\sqrt{79.1715}}\right), ?\right) \quad \text{VS} \quad \mathcal{N}\left(\left(\frac{x - 68.1786}{\sqrt{3.6258}}, \frac{y - 119.6706}{\sqrt{83.0022}}\right), ?\right)$$

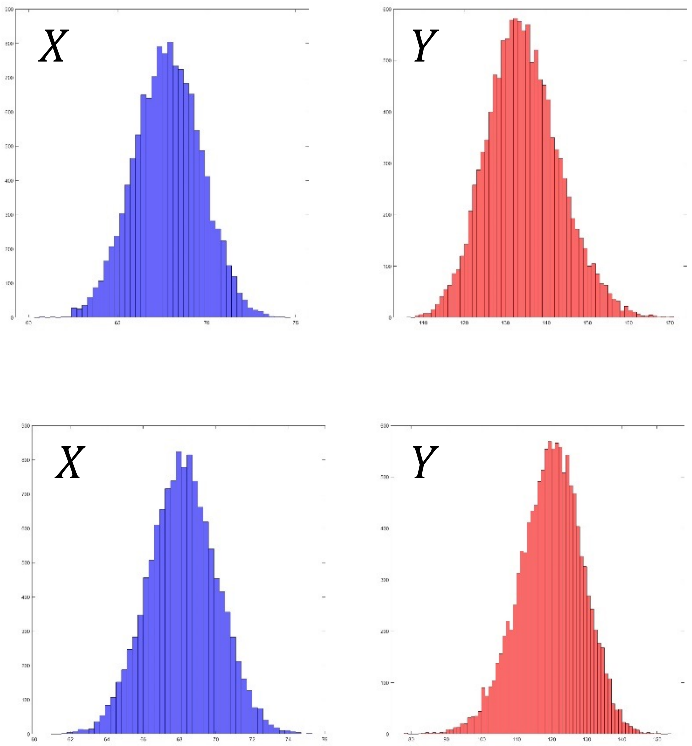
  - Compare the possibility

    $P(X = 70.1111, Y = 141.3333 \mid Z = 0)$ and $P(X = 70.1111, Y = 141.3333 \mid Z = 1)$

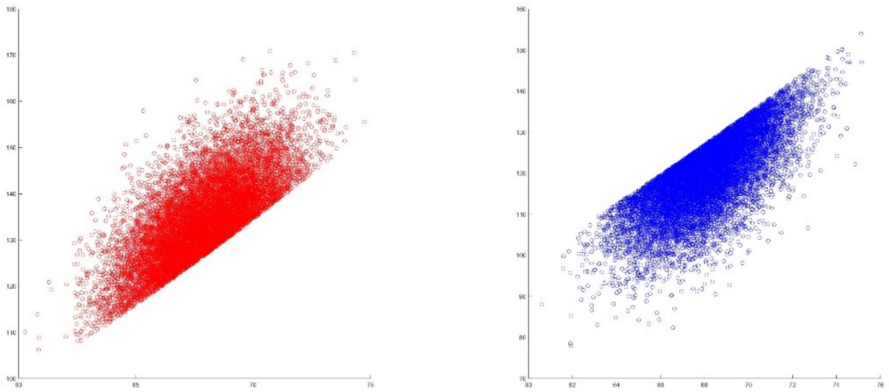
  - It is observed that both $(X, Y)|Z = 0$ and $(X, Y)|Z = 1$ obey Gaussian

$\boldsymbol{Z = 0}$

$\boldsymbol{Z = 1}$



$\boldsymbol{Z = 0}$ $\boldsymbol{Z = 1}$



Why? (思考题)

Classification with Multivariate Features

- $(X, Y) \rightarrow Z$?
  - IF X is **dependent** to Y, THEN we should make a classification jointly
  - Provided Gaussian assumptions, we can obtain the $f(x, y)$
    - X is Gaussian, $\mathcal{N}(67.9931, 3.6164)$
    - Y is Gaussian, $\mathcal{N}(127.0794, 135.9765)$
    - we can get the correlation coefficient $\rho = 0.529$
    - compute $(X, Y) \in \mathcal{N}(\,(67.9931, 127.0794)^T, \mathbf{M}_{cov})$ $\quad \mathbf{M}_{cov} = \begin{bmatrix} 3.6164 & 0.529 \\ 0.529 & 135.9765 \end{bmatrix}$
    - Next ?

- $(X, Y) \rightarrow Z$?
  - **Continuous $(\boldsymbol{X}, \boldsymbol{Y})$, Continuous $\boldsymbol{Z}$**
  - 三维高斯建模

Classification with Multivariate Features

- $(X, Y) \rightarrow Z$?
  - IF X is **dependent** to Y, THEN we should make a classification jointly
  - Provided Gaussian assumptions, we can obtain the $f(x, y)$
    - X is Gaussian, $\mathcal{N}(67.9931, 3.6164)$
    - Y is Gaussian, $\mathcal{N}(127.0794, 135.9765)$
    - we can get the correlation coefficient $\rho = 0.529$
    - compute $(X, Y) \in \mathcal{N}(\,(67.9931, 127.0794)^T, \mathbf{M}_{cov})$ $\qquad \mathbf{M}_{cov} = \begin{bmatrix} 3.6164 & 0.529 \\ 0.529 & 135.9765 \end{bmatrix}$
    - Next ?

- $(X, Y) \rightarrow Z$?
  - Image: If $\mathrm{g(X, Y)} \rightarrow Z$, compute $P(Z)$

Q 2

Summary

Bayesian formula

- Univariate Feature
  - Discrete X, Discrete Z — $Z=1$
  - Continuous X, Discrete Z — $Z=0$
  - Continuous X, Continuous Z — $Z=0$
- Multivariate Features
  - Discrete (X,Y), Discrete Z — $Z=0$
  - Continuous (X,Y), Discrete Z — $Z=0$
  - Continuous (X,Y), Continuous Z — $Z=1$
- Q3:
  - 1. Which result is reliable ?
  - 2. Which way is efficient?

$$\mathrm{P}(\mathrm{Z}=1 \mid \mathrm{X}=x_k)=\boxed{\frac{P(Z=1,X=x_k)}{P(X=x_k)}}=\boxed{\frac{P(Z=1)P(X=x_k \mid Z=1)}{P(X=x_k)}}$$

$$\mathrm{P}(\mathrm{Z}=1 \mid \mathrm{X}=x)=\boxed{\frac{P(Z=1,X=x)}{P(X=x)}}$$

$$=\boxed{\frac{P(Z=1)P(X=x \mid Z=1)}{P(X=x)}}=\frac{P(Z=1)\int_{-\infty}^{x} f_{X|Z}(x|z=1)dx}{P(X=x)}$$

$$\mathrm{E}(\mathrm{Z} \mid \mathrm{X}=x)=\int_{-\infty}^{+\infty} z\boxed{\frac{f(x,z)}{f_X(x)}}dz=\int_{-\infty}^{+\infty} z\boxed{f_{Z|X}(z|X=x)}dz$$

$$\mathrm{P}(\mathrm{Z}=1 \mid \mathrm{X}=x_k,Y=y_k)=\boxed{\frac{P(Z=1,X=x_k,Y=y_k)}{P(X=x_k,Y=y_k)}}=\boxed{\frac{P(Z=1)P(X=x_k\,,Y=y_k| \,Z=1)}{P(X=x_k,Y=y_k)}}$$

$$\mathrm{P}(\mathrm{Z}=1 \mid \mathrm{X}=x,Y=\mathrm{y})=\boxed{\frac{P(Z=1,X=x,Y=y)}{P(X=x,Y=y)}}$$

$$=\boxed{\frac{P(Z=1)P(X=x,Y=y \mid Z=1)}{P(X=x,Y=y)}}=\frac{P(Z=1)\int_{-\infty}^{y}\int_{-\infty}^{x} f_{X,Y|Z}(x,y|z=1)\,dydx}{P(X=x,Y=y)}$$

$$\mathrm{E}(\mathrm{Z} \mid \mathrm{X}=x,\mathrm{Y}=\mathrm{y})=\int_{-\infty}^{+\infty} z\boxed{\frac{f(x,y,z)}{f_{X,Y}(x,y)}}dz=\int_{-\infty}^{+\infty} z\boxed{f_{Z|X,Y}(z|X=x,Y=y)}dz$$

Summary

- **Q1:** Is there a function h such that $h(X)$ is obviously separable? $\boxed{X \rightarrow h(X)} \rightarrow Z$
- **Q2:** Is there an apposite function g such that $\boxed{(\mathrm{X},\mathrm{Y}) \rightarrow g(\mathrm{X},\mathrm{Y})} \rightarrow Z$, compute $P(Z)$

  - Statistics: given an explicit function before classification
  - Classical ML: define a kernel function before classification
  - Modern ML: represent the concerned function by model and data
    — representation learning
  - Other ways?: directly construct the concerned function by ergodic algorithms e.g., the genetic algorithm
    — evolutionary learning

Summary - genetic algorithm

- Q1: Is there a function h such that $h(X)$ is obviously separable? $\boxed{X \rightarrow h(X)} \rightarrow Z$
- Q2: Is there an apposite function g such that $\boxed{(\mathrm{X}, \mathrm{Y}) \rightarrow g(\mathrm{X}, \mathrm{Y})} \rightarrow Z$, compute $P(Z)$

  - Other ways?: directly construct the concerned function by ergodic algorithms e.g., the genetic algorithm

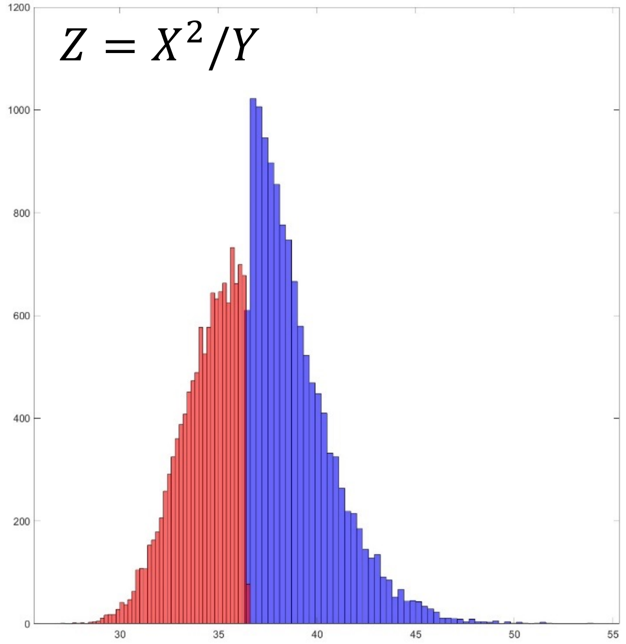
    — evolutionary learning

    — $Z = X^2/Y$ is apposite

$$\frac{70.1111^2}{141.3333} = 34.7800 < E(\frac{X^2}{Y}) \Rightarrow Z = 0$$

— match the case of "Discrete (X, Y), Discrete Z"

Why?



Any Questions?