# Linear Discriminant Analysis (2)

# Classification Problem

- Given a dataset that has $x$ and $y$ (class or label)

| x | y |
|---|---|
| 1.49 | 1 |
| 1.23 | 1 |
| 0.95 | 1 |
| 0.89 | 1 |
| 2.58 | 2 |
| 2.65 | 2 |
| 2.27 | 2 |
| 1.94 | 2 |
| 1.39 | 2 |
| 1.44 | 2 |

- Given a new value $x$, what class the it belongs to? (what is the predicted $y$ value)
- If $x = 1.4$, what is its associated $y$ value? (What class it

# Approach

- We will estimate two probabilities $p_1 = P(y = 1|x = 1.4)$ and $p_2 = P(y = 2|x = 1.4)$.
- If $p_1 > p_2$, we will classify the new point to class 1 and vice versa.

# Approach

We have, using the Bayes' Rule,

$$p_1 = P(y = 1|x = 1.4) = \frac{P(y = 1) * L(x = 1.4|y = 1)}{L(x = 1.4)}$$

where L(A) denotes the likelihood of the event A. Similarly,

$$p_2 = P(y = 2|x = 1.4) = \frac{P(y = 2) * L(x = 1.4|y = 1)}{L(x = 1.4)}$$

Since the denominator is the same we just need to compare the numerator.

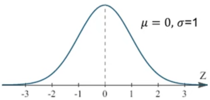
# LDA Assumptions

- $x$ is normally distributed in each class
- Assume that $x$ has the same variance in both classes

## Likelihood:

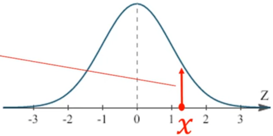
The *probability density function* for a normal distribution $N(\mu, \sigma^2)$ is:

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{\sigma^2}}$$




For a given distribution the *likelihood* of the distribution parameters being $\mu, \sigma^2$ given the observation $x$ is:

$$L(\mu, \sigma^2|x) = f(x|\mu, \sigma^2)$$



# Calculation

# Calculation

https://planetcalc.com/4986/

https://www.standarddeviationcalculator.io/normal-distribution-calculator

## Decision Boundary

- The decision boundary is where $p_1 = p_2$
- Thus,

$$
\begin{aligned}
\frac{P(y=1) * L(x=x_0|y=1)}{L(x=x_0)} &= \frac{P(y=2) * L(x=x_0|y=
}{L(x=x_0)} \\
\implies P(y=1) * L(x=x_0|y=1) &= P(y=2) * L(x=x_0|y= \\
\implies \pi_1 * L(x=x_0|y=1) &= \pi_2 * L(x=x_0|y= \\
\implies \pi_1 * f(x_0|\mu_1, \sigma^2) &= \pi_2 * f(x_0|\mu_2, \\
\implies \pi_1 * \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-(x_0-\mu_1)^2/(2\sigma^2)} &= \pi_2 * \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-(x_0-\mu_2)^2/( \\
\implies \pi_1 * \frac{1}{\sigma_1} e^{-(x_0-\mu_1)^2/(2\sigma_1^2)} &= \pi_2 * \frac{1}{\sigma_2} e^{-(x_0-\mu_2)^2/(
\end{aligned}
$$

$$
\implies 2x_0 = \mu_1 + \mu_2 + \frac{2\sigma^2(\ln \pi_2 - \ln \pi_1)}{\mu_1 - \mu_2}
$$