

# MACHINE LEARNING

## Homework Week 1

September 14, 2021

Nguyen Son Tung

---

1. To evaluate a new test for detecting Hansen's disease, a group of people 5% of which are known to have Hansen's disease are tested. The test finds Hansen's disease among 98% of those with the disease and 3% of those who don't. What is the probability that someone testing positive for Hansen's disease under this new test actually has it?

### Answer

- Denoted those people who got Hansen's disease as H (0: Not infected; 1: Infected)

- Denoted the outcome of the test as T (0: Negative; 1: Positive)

We already know that:

$$\begin{aligned}P(H = 1) &= 0.05 \Rightarrow P(H = 0) = 1 - 0.05 = 0.95 \\P(T = 1|H = 1) &= 0.98 \\P(T = 1|H = 0) &= 0.03\end{aligned}$$

---

$$P(H = 1|T = 1) = \frac{P(T = 1|H = 1) \cdot P(H = 1)}{P(T = 1)}$$

$$P(H = 1|T = 1) = \frac{P(T = 1|H = 1) \cdot P(H = 1)}{P(T = 1|H = 1) \cdot P(H = 1) + P(T = 1|H = 0) \cdot P(H = 0)}$$

$$P(H = 1|T = 1) = \frac{0.98 \cdot 0.05}{0.98 \cdot 0.05 + 0.03 \cdot 0.95}$$

$$P(H = 1|T = 1) = \frac{98}{155} \approx 63.22\%$$

2. Proof the following distributions are normalized then calculate the mean and standard deviation of these distribution:
- (a) Univariate normal distribution
  - (b) Multivariate normal distribution (Optional)

**Answer**

- (a) The probability density function of the Univariate Gaussian Distribution is given by:

$$\varphi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}x^2} \quad -\infty < x < \infty$$

To prove that the above expression is normalized, we have to show that

$$\varphi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}x^2} = 1$$

or

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}x^2} = \sqrt{2\pi\sigma^2}$$

*Proof.* Let

$$I = \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}x^2} dx$$

Squaring the above expression,

$$I^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}x^2 - \frac{1}{2\sigma^2}y^2} dx dy \quad (1)$$

To integrate this expression we make the transformation from Cartesian coordinates (x, y) to polar coordinates (r,  $\theta$ ), which is defined by

$$x = r \cos \theta$$

$$y = r \sin \theta$$

We have:

$$\cos^2 \theta + \sin^2 \theta = 1 \Rightarrow x^2 + y^2 = r^2$$

The Jacobian of the change of variables is given by,

$$\begin{aligned}
\frac{\partial(x, y)}{\partial(r, \theta)} &= \begin{vmatrix} \frac{\partial(x)}{\partial(r)} & \frac{\partial(x)}{\partial(\theta)} \\ \frac{\partial(y)}{\partial(r)} & \frac{\partial(y)}{\partial(\theta)} \end{vmatrix} \\
&= \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} \\
&= r \cos^2 \theta + r \sin^2 \theta \\
&= r
\end{aligned}$$

Equation (1) can be rewritten as

$$\begin{aligned}
I^2 &= \int_0^{2\pi} \int_0^\infty e^{-\frac{r^2}{2\sigma^2}} r \, dr \, d\theta \\
&= 2\pi \int_0^\infty e^{-\frac{r^2}{2\sigma^2}} r \, dr \\
&= 2\pi \int_0^\infty \frac{1}{2} e^{-\frac{u}{2\sigma^2}} \, du \\
&= \pi \left[ e^{-\frac{u}{2\sigma^2}} (-2\sigma^2) \right]_0^\infty \\
&= 2\pi\sigma^2 \\
\Rightarrow I &= \sqrt{2\pi\sigma^2}
\end{aligned}$$

To prove that  $\mathcal{N}(x|\mu, \theta^2)$  is normalized, we make the transformation  $y = x - \mu$ , so that

$$\begin{aligned}
\int_{-\infty}^\infty \mathcal{N}(x|\mu, \theta^2) dx &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty e^{-\frac{y^2}{2\sigma^2}} dy \\
&= \frac{I}{\sqrt{2\pi\sigma^2}} \\
&= \frac{\sqrt{2\pi\sigma^2}}{\sqrt{2\pi\sigma^2}} \\
&= 1
\end{aligned}$$