Machine Learning Homework Week 1

Tô Đức Anh

September 2021

1 Problems

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?

2. Proof the following distributions are normalized then calculate the mean and standard deviation of these distribution:

   - Univariate normal distribution.
   - (Optional) Multivariate normal distribution.

2 Answers

2.1 Question 1

- Probability that someone testing positive for Hansen's disease under this new test actually has it:

$$P(actually\ has\ disease \mid tested) = \frac{P(tested \mid actually\ has\ disease) * P(actually\ has\ disease)}{P(tested)} \tag{1}$$

- Probability of people with Hansen's disease in the 5% are known to have Hansen's disease is:

$$5\% \times 98\% = 4.9\%$$

- Probability of people with Hansen's disease in the 95% are known to not have Hansen's disease is:

$$95\% \times 3\% = 2.85\%$$

- Probability of people testing positive for Hansen's disease is:

$$4.9\% + 2.85\% = 7.75\%$$

- Probability that someone testing positive for Hansen's disease under this new test actually has it is:

$$\frac{5\% \times 98\%}{7.75\%} \approx 63.2\%$$

2.2 Question 2

- Univariate normal distribution

2.2.1 Formula for univariate normal distribution

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

With respect to μ as Mean and σ as Standard Deviation

2.2.2 Prove that f(x) always $>$ 0 $\forall\mu$ and σ

We can see that $\frac{1}{\sigma\sqrt{2\pi}}$ will always $>$ 0 since the standard deviation σ always ≥ 0. Therefore $\frac{1}{\sigma\sqrt{2\pi}} \times e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$ will always > 0 $\forall\mu$ and σ

2.2.3 Prove that $\int_{-\infty}^{\infty} f(x)dx = 1$

2.2.4 Prove mean = μ:

We have:

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

So that expectation of f(x) is:

$$
\begin{aligned}
E[f(x)] &= \int_{-\infty}^{\infty} x \times \frac{1}{\sigma\sqrt{2\pi}} \times e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx \\
&= \frac{1}{\sigma\sqrt{2\pi}} \times \int_{-\infty}^{\infty} x \times e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx \\
&= \frac{\sqrt{2}\sigma}{\sigma\sqrt{2\pi}} \times \int_{-\infty}^{\infty} \left(\sqrt{2}\sigma t + \mu\right) \times e^{-t^2} dt \\
&= \frac{1}{\sqrt{\pi}} \times \left(\sqrt{2}\sigma \times \int_{-\infty}^{\infty} t \times e^{-t^2} dt + \mu \int_{-\infty}^{\infty} e^{-t^2} dt\right) \\
&= \frac{1}{\sqrt{\pi}} \times \left(\sqrt{2}\sigma \times \left[\frac{-1}{2} \times e^{-t^2}\right] + \mu \times \sqrt{\pi}\right) \\
&= \frac{\mu\sqrt{\pi}}{\sqrt{\pi}} \\
&= \mu
\end{aligned}
\tag{2}
$$

2.2.5 Prove Standard deviation = σ^2

$$ \tag{3} $$

- (Optional) Multivariate normal distribution.
  - Proof that the distribution are normalized:
  - Mean:
  - Standard Deviation: