Probability Machine Learning

thiencao

September 2021

Exercise1: To evaluate a new test to detect Hansen's disease, a group of 5 percent of those known to have Hansen's disease were tested. The test found Hansen's disease in 98 percent of people with the disease and 3 percent of those who don't. What is the probability that someone who tests positive for Hansen's disease on this new test actually has the disease?

SOLVED:
A: known to have Hansen disease
B: known to have no Hansen disease
H: Positive Hansen disease result
$P(A) = 0.05$
$P(B) = 0.95$
$P(H \mid A) = 0.98$
$P(H \mid B) = 0.03$

So the probability that someone who tests positive for Hansen' disease on this new test actually has the disease is:

$$(A|H) = \frac{P(H|A) \times P(A)}{P(H)}$$

$$P(A|H) = \frac{P(H|A) \times P(A)}{P(H|A) \times P(A) + P(H|B) * P(B)}$$

$$(A|H) = \frac{0.98 \times 0.05}{0.98 \times 0.05 + 0.03 \times 0.95}$$

$$P(A|H) = \frac{0.049}{0.0775} \approx 0.632$$

Exercise2: Proof the following distributions are normalized then calculate the mean and standard deviation of these distribution:

1. Univariate normal distribution.

2. (Optional) Multivariate normal distribution.

<u>SOLVED:</u>

1.Univariate normal distribution:

+) Univariate normal distribution have density function:

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

Proof the Univariate is normalize

$$\iff \int_{-\infty}^{+\infty} f(x)dx = 1$$

$$\iff \int_{-\infty}^{+\infty} \frac{1}{2\pi\sigma^2} \times e^{\frac{(x-\mu)}{-2\sigma^2}} dx = 1$$

$$\iff \frac{1}{\sqrt{2\pi\sigma^2}} \times \int_{-\infty}^{+\infty} e^{\frac{(x-\mu)}{-2\sigma^2}} dx = 1(*)$$

Making the substitution

$$y = \frac{(x-\mu)}{\sigma}$$

, we have:

$$\frac{1}{\sqrt{2\pi}} \times \int_{-\infty}^{+\infty} e^{\frac{-y^2}{2}} dy$$

let

$$I = \int_{-\infty}^{+\infty} e^{\frac{-y^2}{2}} dy$$

. Then:

$$I^2 = \int_{-\infty}^{+\infty} e^{\frac{-y^2}{2}} dy \times \int_{-\infty}^{+\infty} e^{\frac{-x^2}{2}} dx$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{\frac{-(y^2+x^2)}{2}} dydx$$

Let

$$x = rcos\phi, y = rsin\phi$$

$$dydx = rd\theta dr$$

We Have:

$$I^2 = \int_0^{+\infty} \int_0^{2\pi} e\frac{-r^2}{2} rd\phi dr$$

$$= 2\pi \int_0^{+\infty} re\frac{-r^2}{2} dr$$

$$= -2\pi e\frac{-r^2}{2} \mid_0^{+\infty}$$

$$= 2\pi$$

$$\rightarrow I = \sqrt{2\pi}$$

Substituting in (*) we have :

$$\iff \frac{1}{\sqrt{2\pi\sigma^2}} \times \int_{-\infty}^{+\infty} e^{\frac{(x-\mu)}{-2\sigma^2}} dx = 1(Proofed)$$

Calculatuing Mean: First: We let

$$Z = \frac{(X-\mu)}{\sigma}$$

We have:

$$E[Z] = \int_{-\infty}^{+\infty} xf_Z(x)dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} xe \frac{-x^2}{2} dx$$

$$= \frac{-1}{\sqrt{2\pi}} e \frac{-x^2}{2} \mid_{-\infty}^{+\infty}$$

$$= 0$$

Because :

$$X = \mu + \sigma Z$$

So:

$$E(X) = E(\mu) + E(\sigma Z)$$

$$\iff E(X) = \mu + E(Z)E(\sigma)$$

But: E(Z) = 0

$$\rightarrow E(X) = \mu$$

Calculating Variance:

$$Var(Z) = E(Z^2) - ((E(Z))^2 = E(Z^2)$$

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

We let:

$$u = x, dv = xe^{\frac{-x^2}{2}}$$

$$Var(Z) = \frac{1}{\sqrt{2\pi}} (-xe^{\frac{-x^2}{2}} |_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} e^{\frac{-x^2}{2}} dx)$$

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

$$= 1$$

But:

$$X = \mu + \sigma Z$$

$$\iff Var(X) = Var(\mu) + \sigma^2 Var(Z)$$

$$\rightarrow Var(X) = 0 + \sigma^2 1$$

$$\rightarrow Var(X) = \sigma^2$$

2.Multivariate normal distribution:

+) Multivariate normal distribution have density function:

$$\phi(X) = \prod_{j=1}^{p} \frac{1}{2\pi\sigma_j^2} e^{\frac{-(x_j - \mu_j)^2}{2\sigma_j^2}}$$

We have to proof that:

$$\int$$