Homework 1

Atlas42

September 2021

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

From the question, we can infer these information:
The probability of selecting a person with Hansen is:

$$P(hasHansen) = 0.05$$

The probability of selecting a person without Hansen is:

$$P(doesnthaveHansen) = 0.05$$

The probability of selecting a person with Hansen from those that have Hansen with the test is:

$$P(test \mid hasHansen) = 0.98$$

The probability of selecting a person with Hansen from those that don't have Hansen with the test is:

$$P(test \mid doesnthaveHansen) = 0.3$$

The question at hand:
$P(hasHansen \mid test) = ?$
Applying Bayes Theorem:

$$P(A|B) = P(A)\frac{P(B|A)}{P(B)}$$

where A, B = events

$P(A \mid B)$ = probability of A given B is true

$P(B \mid A)$ = probability of B given A is true

$P(A), P(B)$ = the independent probabilities of A and B

Thus, we have:

$$P(hasHansen \mid test) = P(test \mid hasHansen)\frac{P(hasHansen)}{P(test)}$$

$$= 0.98\frac{0.05}{0.05\times0.98+0.95\times0.03}$$

$$\approx 0.6322$$

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

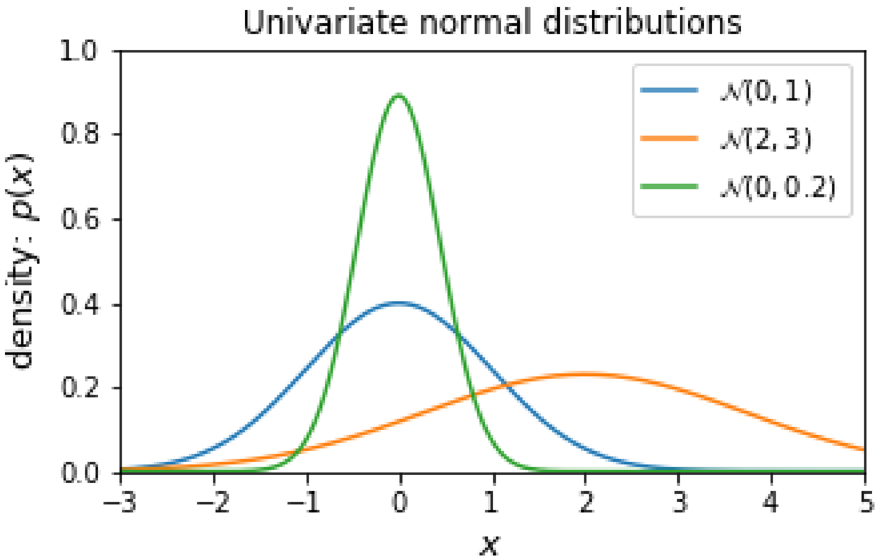
a. The normal distribution, also known as Gaussian distribution, is defined by two parameters, mean μ, which is expected value of the distribution and standard deviation which corresponds to the expected squared deviation from the mean. Mean, μ controls the Gaussian's center position and the standard deviation controls the shape of the distribution. The square of standard deviation is typically referred to as the variance σ^2. We denote this distribution as $N(\mu, \sigma^2)$.

Given the mean and variance, one can calculate probability distribution function of normal distribution with a normalised

Gaussian function for a value x, the density is:

$$P(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

We call this distribution univariate because it consists of one random variable.



Now, we calculate the mean and standard deviation of univariate normal distribution:
First, we have the mean:
$E\left(X\right) = \mu_X = \int_{-\infty}^{\infty} x \cdot f\left(x\right) dx.$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int\limits_{-\infty}^{\infty} \left(z + \mu\right) e^{-\frac{z^2}{2\sigma^2}} dz$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int\limits_{-\infty}^{\infty} z e^{-\frac{z^2}{2\sigma^2}} dz + \mu \left[\frac{1}{\sqrt{2\pi\sigma^2}} \int\limits_{-\infty}^{\infty} e^{-\frac{z^2}{2\sigma^2}} dz\right]$$

The first integral evaluate to 0 because the integrand is an odd function and the integration can be split in two simmetric halves (with respect to the y axis), which are both convergent (i.e. the limits $\lim_{a\to\infty}\int_0^a f(x)\,dx$ and $\lim_{a\to-\infty}\int_a^0 f(x)\,dx$ that define them exist). The second integral (within square brackets) evaluates to 1 (it is a Gaussian pdf), so you are left with:

$$E[X] = \mu$$

Next up, we calculate the variance of the univariate normal distribution:

X $= \int_{-\infty}^{\infty} x^2 f_X x x - X^2$

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

$= \frac{\sqrt{2}\sigma}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \sqrt{2}\sigma t + \mu^2 \exp -t^2 t - \mu^2$

$= \frac{1}{\sqrt{\pi}} 2\sigma^2 \int_{-\infty}^{\infty} t^2 \exp -t^2 t + 2\sqrt{2}\sigma\mu \int_{-\infty}^{\infty} t \exp -t^2 t + \mu^2 \int_{-\infty}^{\infty} \exp -t^2 t - \mu^2$

$= \frac{1}{\sqrt{\pi}} 2\sigma^2 \int_{-\infty}^{\infty} t^2 \exp -t^2 t + 2\sqrt{2}\sigma\mu - \frac{1}{2} \exp -t^2 - \infty\infty + \mu^2\sqrt{\pi} - \mu^2$

$= \frac{1}{\sqrt{\pi}} 2\sigma^2 \int_{-\infty}^{\infty} t^2 \exp -t^2 t + 2\sqrt{2}\sigma\mu \cdot 0 + \mu^2 - \mu^2$

$= \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^2 \exp -t^2 t$

$= \frac{2\sigma^2}{\sqrt{\pi}} - \frac{t}{2} \exp -t^2 - \infty\infty + \frac{1}{2} \int_{-\infty}^{\infty} \exp -t^2 t$

$= \frac{2\sigma^2}{\sqrt{\pi}} \cdot \frac{1}{2} \int_{-\infty}^{\infty} \exp -t^2 t$

$= \frac{2\sigma^2\sqrt{\pi}}{2\sqrt{\pi}}$

$= \sigma^2$