# Machine-Learning-with-Python

Nguyen Minh Anh

September 2021

## 1 Probability

**Exercise 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?
Solution 1 :
P(disease) = 0.05
$P(positive \mid disease) = 0.98$
$P(positive \mid notdisease) = 0.03$
$P(disease \mid positive) =??$

$$P(disease \mid positive) = \frac{P(positive \mid disease) \times P(disease)}{P(positive)}$$

$P(positive) = P(positive \mid disease) \times P(disease) + P(positive \mid notdisease) \times P(notdisease)$
$= 0.98 \times 0.05 + 0.95 \times 0.03$
$= 0.0775$

$$=> P(disease \mid positive) = \frac{0.98 \times 0.05}{0.0775} = 0.632$$

**Exercise 2** 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.
Solution 2:
1.1 Normalized
1.2 Calculate mean, variance
$X \sim N(\mu, \sigma^2)$
Proof $E(X) = \mu$

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

$E(Z) = 0$

Inverting this formula we have $X = \sigma Z + \mu$

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

Proof $Var(X) = \sigma^2$
Letting $u = \frac{x-\mu}{\sigma}$
Formula $Var(X) = \int_{\infty}^{\infty} \sum_x (x-\mu)^2 f(x) dx$

$$Var(X) = E((X-\mu)^2)) = \int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u^2 \mathrm{e}^{\frac{-u^2}{2}} dz$$

$u = t => du = dt$
$dv = t\mathrm{e}^{\frac{-t^2}{2}} => v = \mathrm{e}^{\frac{-t^2}{2}}$

$$Var(X) = \frac{\sigma^2}{\sqrt{2\pi}} \left[ (-t\mathrm{e}^{\frac{-t^2}{2}}) \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \mathbf{e}^{\frac{-t^2}{2}} dt \right] = \sigma^2$$