TP3: Elements of Detection Theory

by Martino Ferrari

1 Elements of Detection Theory

Exercise 1

Given $X \sim \mathcal{N}(0,1)$ and $Y \sim \mathcal{N}(0,2)$:

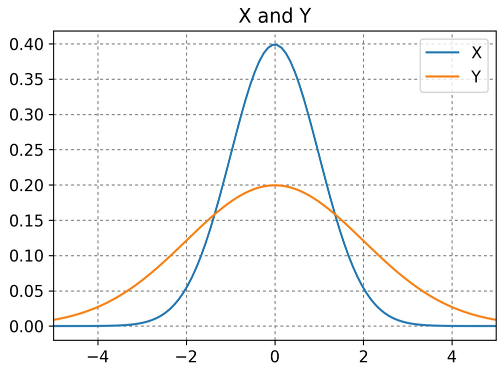

Figure 1. X and Y *pdf*

compute $P[-1 < Y \leqslant 1]$, $P[Y > 3.5]$, $P[-1 < X \leqslant 1]$ and $P[X > 3.5]$:

$$P[-1 < Y \leqslant 1] = \Phi\left(\tfrac{1}{2}\right) - \Phi\left(-\tfrac{1}{2}\right) = 0.393 \qquad P[Y > 3.5] = 1 - \Phi\left(\tfrac{3.5}{2}\right) = 0.0401$$
$$P[-1 < X \leqslant 1] = \Phi(1) - \Phi(-1) = 0.683 \qquad P[X > 3.5] = 1 - \Phi(3.5) = 0.000233$$

Table 1. Numerical results of exercise 2

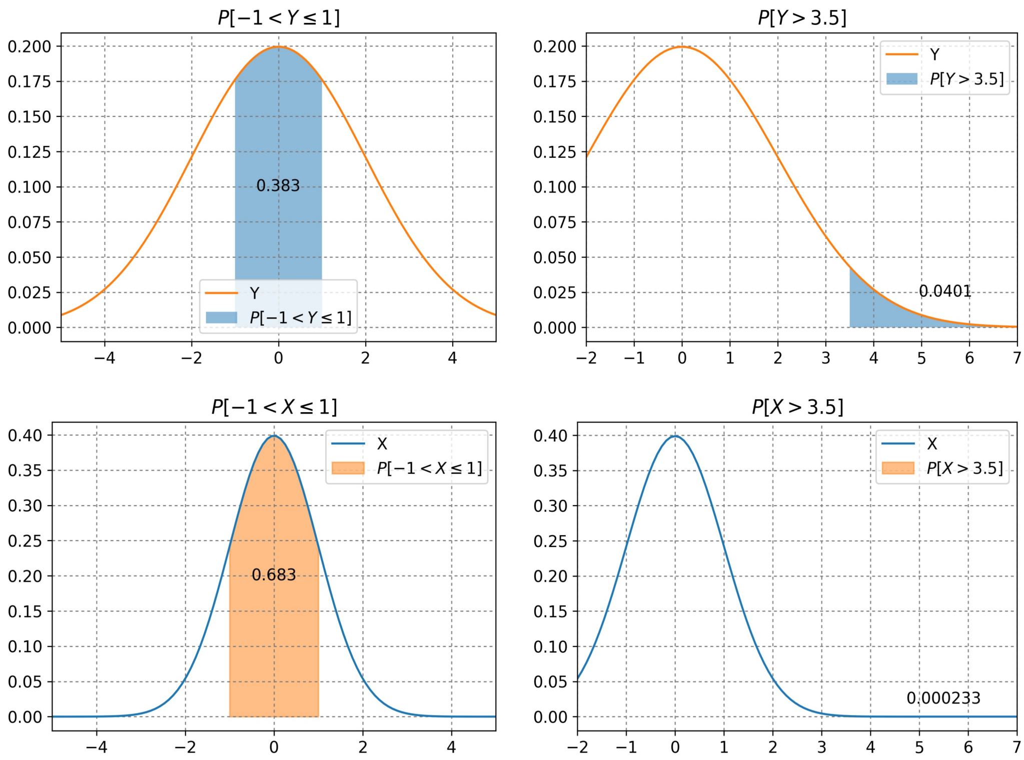

Figure 2. graphical results of exercise 1

Exercise 2

Given $X \sim \mathcal{N}(30, 10)$ that denote the peak temperature of Geneva in June, what is $P[X > 40]$, $P[X \leqslant 15]$ and $P[20 < X \leqslant 40]$:

The numerical results are:

$P[X>40] = 1 - \Phi\left(\frac{40-30}{10}\right) = 0.159$	$P[X \leqslant 15] = \Phi\left(\frac{15-30}{10}\right) = 0.0668$
$P[20 < X \leqslant 40] = \Phi\left(\frac{40-30}{10}\right) - \Phi\left(\frac{20-30}{10}\right) = 0.683$	

Table 2. Numerical results of exercise 2

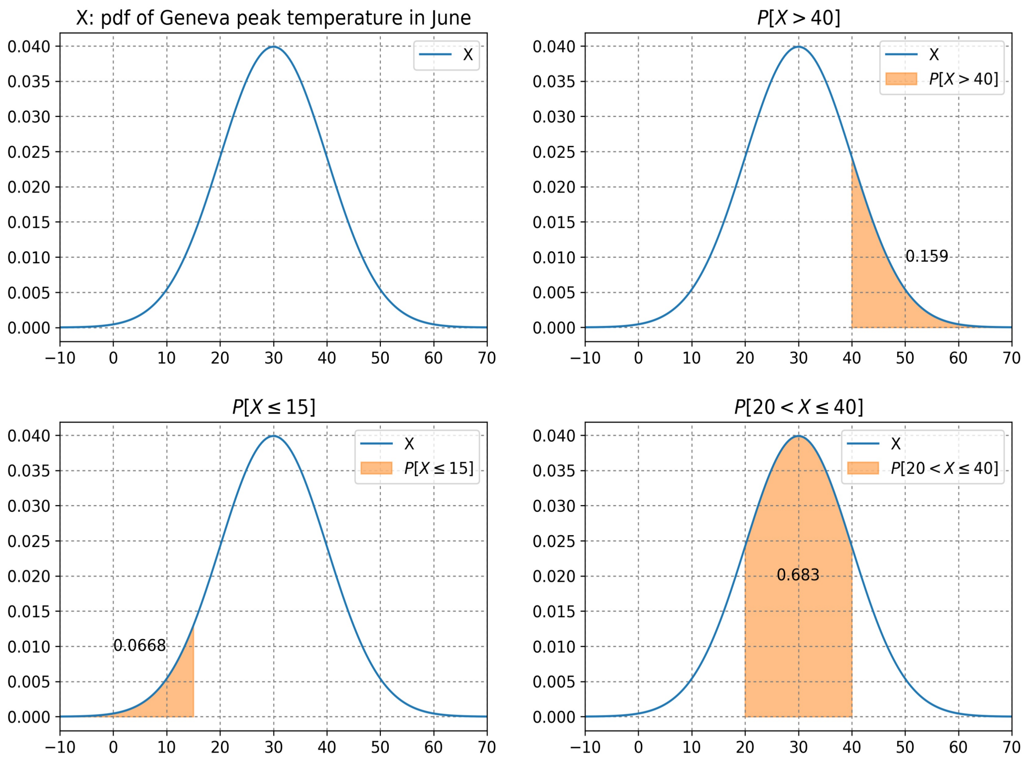

Figure 3. graphical results of exercise 2

Exercise 3

Let X be a Gaussian random variable, for which $E[X]=0$ and $P[|X|\leq 10]=0.1$. What is σ_X?

$P[|X|\leq 10]$ is equivalent to $P[-10<X\leq 10]$. To find σ_X we have to find:

$$\sigma_X\colon \left[\Phi\left(\frac{10}{\sigma_X}\right)-\Phi\left(\frac{-10}{\sigma_X}\right)\right]=0.1$$

This results in $\sigma_X=79.58$

Exercise 4

Next to the Q-function, the communications field uses the *complementary error function* (ERFC) for the tail probabilities of Gaussian random variables. It is defined as:

$$\operatorname{ercf}(n)=\frac{2}{\sqrt{\pi}}\int_n^\infty e^{-x^2}dx$$

Prove that:

$$Q(n)=\frac{1}{2}\operatorname{ercf}\left(\frac{n}{\sqrt{2}}\right)$$

$$Q(n)=\frac{1}{\sqrt{2\pi}}\int_n^\infty e^{-\frac{x^2}{2}}dx$$

$$\frac{1}{2}\operatorname{ercf}\left(\frac{n}{\sqrt{2}}\right)=\frac{1}{\sqrt{\pi}}\cdot\int_{\frac{n}{\sqrt{2}}}^\infty e^{-x^2}dx$$

$$x\rightarrow\frac{n}{\sqrt{2}}$$
$$u\rightarrow n$$

$$n = \sqrt{2}x \to x = \frac{n}{\sqrt{2}}$$

$$\frac{1}{\sqrt{\pi}} \cdot \frac{1}{\sqrt{2}} \cdot \int_n^{\infty} e^{-\left(\frac{u}{\sqrt{2}}\right)^2} du \qquad \frac{1}{\sqrt{2\pi}} \cdot \int_n^{\infty} e^{-\left(\frac{u}{\sqrt{2}}\right)^2} du$$

2 Bayesian and Neyman-Pearson Test

Exercise 1

Let there be two hypothesizes, H_0 and H_1:

$$\begin{cases} H_0 \colon X = Z \\ H_1 \colon X = \mu_1 + Z \end{cases}$$

where $Z \sim \mathcal{N}(0,1)$ and $\mu_1 = 1$.

The separation threshold τ for the MAP hypothesis is: $\tau = 0.5$ such that it minimize the $p_{\text{err}} = p_m + p_f$. In this specific case the two hypothesis have both same probability and distribution (only shifted of $\mu_1 - \mu_0$. Then $\tau = \frac{(\mu_1 + \mu_0)}{2} = 0.5$:

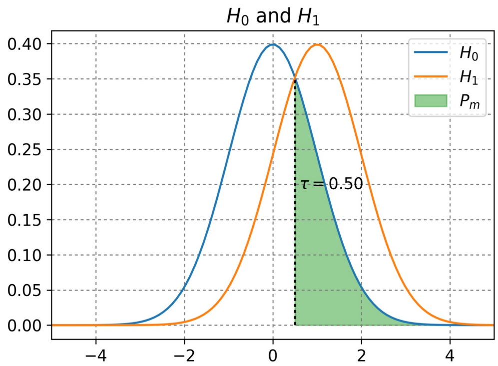

Figure 4. H_0 and H_1

To compute the p_m is enough to compute is simply equal to $p_m = Q(\tau)$, while the $p_f = \Phi(\tau - \mu_0)$ as $\sigma_X = 1$. However as explained before, in this particular case $p_m = p_f$, so:

$p_f = p_m = 0.309$

and

$p_d = (1 - p_m) \to 0.691$

Exercise 2

Let there be two hypothesizes, H_0 and H_1:

$$\begin{cases} H_0 \colon X = Z \\ H_1 \colon X = \mu_1 + Z \end{cases}$$

where $Z \sim \mathcal{N}(0,1)$ and $\mu_1 = \{0, 1, 2\}$.

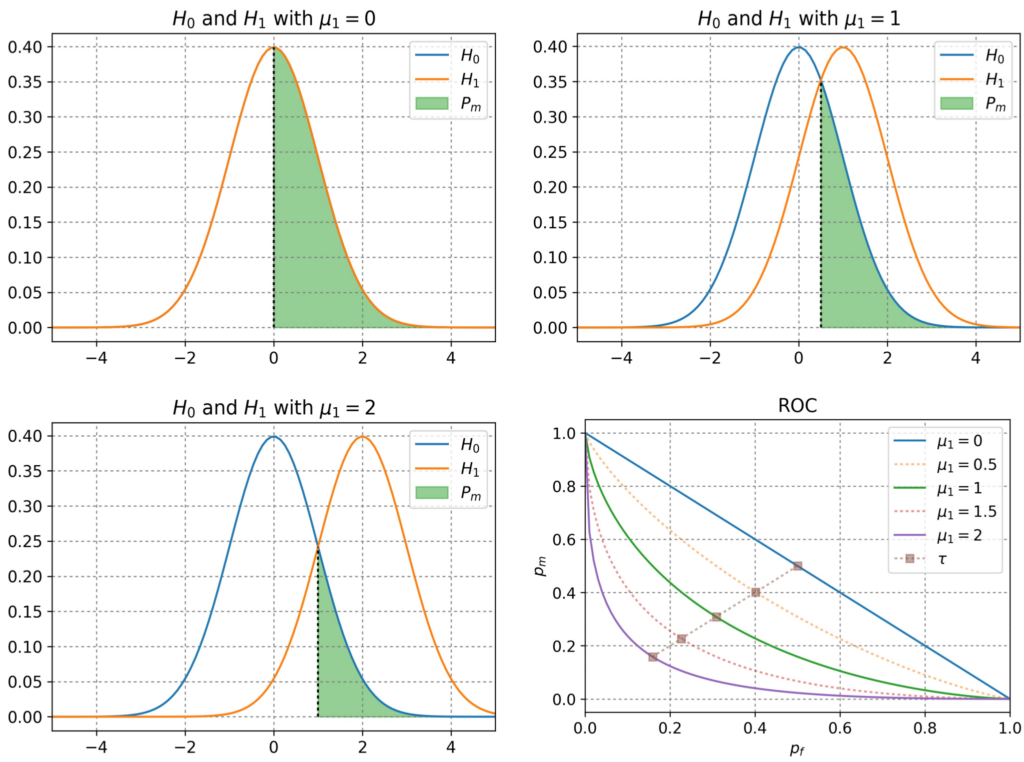

Figure 5. graphical results of exercise 3

As before to compute the distributions are equal and equiprobables and so $\tau = \frac{(\mu_1 + \mu_0)}{2}$ and $p_m = p_f$:

μ_1	τ	p_f	p_m	p_d
0	0	0.500	0.500	0.500
1	0.5	0.309	0.309	0.691
2	1	0.159	0.159	0.841

Table 3. Numerical results of exercise 3

Exercise 4

Given a binary communication system transmits a signal $X \sim \mathcal{B}(0.5)$. The receiver receives:

$$Y = VX + W \quad \wedge \quad W \perp V \perp X \quad \wedge \quad W, V \sim \frac{\mathcal{E}(1,1)}{\ln(1)}$$

In particular W and V has the following representation:

$$p_W(x), p_V(x) = \begin{cases} e^{-x} & x \geqslant 1 \\ 0 & x < 1 \end{cases}$$

However this is not a pdf because by definition the integral from $-\infty$ to $+\infty$ of the *pdf* has to be 1, and:

$$\int_{-\infty}^{+\infty} p_W(x)dx = \int_1^{\infty} e^{-x}dx = (-e^{-\infty} + e^{-1}) = \mathbf{e^{-1} \neq 1}$$

So this function has can be normalize as follow:

$$p_W(x), p_V(x) = \begin{cases} e^{(1-x)} & x \geqslant 1 \\ 0 & < 1 \end{cases}$$

This can be represented with two equiprobables hypothesis (as $X \sim B(p), p = 0.5$):

$$\begin{cases} H_0: W \\ H_1: W + V \end{cases}$$

Where the sum of $W + V$ is equivalent to the convolution of the two *pdf* but due the fact that the two signal have identical exponential low:

$$p_{[W+V]}(x) = \begin{cases} (x-2) \cdot e^{-(x-2)} & x \geqslant 2 \\ 0 & x < 2 \end{cases}$$

the *cdf* functions of H_0 and H_1 are:

$$\Phi_{H_0}(t) = \int_1^t e^{1-x} \mathrm{dx} = 1 - e^{1-t} \qquad t \geqslant 1$$

$$\Phi_{H_1}(t) = \int_2^t (x-2) e^{2-x} \mathrm{dx} = 1 - (t-1) e^{2-t}$$

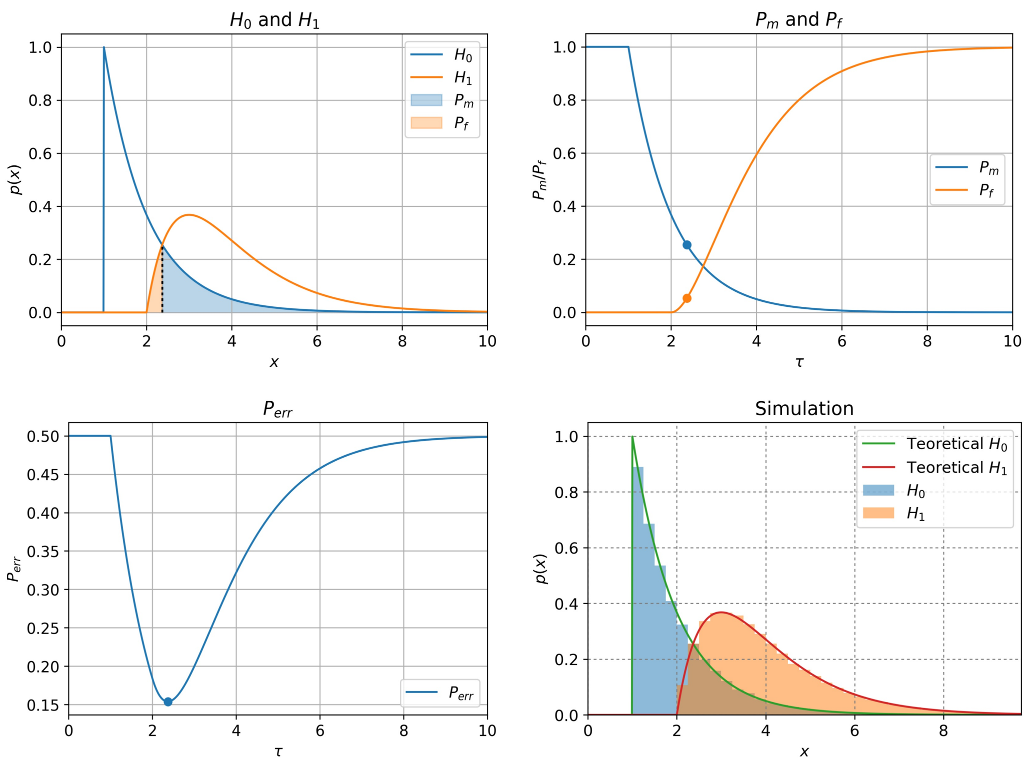

Figure 6. graphical results of exercise 2

The threshold τ is such that minimizes the $p_{\mathrm{err}}(x) = P(H_0) \cdot p_m(x) + P(H_1) \cdot p_f(x)$. However the two hypotheses are equiprobables ($P(H_0) = p, P(H_1) = 1 - p, p = 0.5$) and so we can simplify in:

$$\tau = \arg\min_x \left(p_m(x) + p_f(x)\right) = \arg\min_x \left((1 - \Phi_{H_0}(x)) + \Phi_{H_1}(x)\right)$$

in the range from 2 to infinite this can be expressed as

$$\tau = \arg\min_{x \in [2, \infty[} ((1 - (1 - e^{1-x})) + (1 - (x - 1)e^{2-x}))$$

$$\arg\min_{x \in [2, \infty[} (-e^{1-x} + 1 - (x - 1)e^{2-x})$$

It is now possible to find the minimum deriving:

$$p'_{\mathrm{err}} = e^{1-x} + (e^{2-x} - x\, e^{2-x}) + e^{2-x} = e^{1-x} + e^{2-x}(2 - x)$$

$$p'_{\mathrm{err}} = 0 \rightarrow \tau = x = e^{-1}(2e + 1)$$

resulting in $\tau = 2.37$ with a total $p_{\mathrm{err}} = 0.154$, $p_m = 0.249$ and $p_f = 0.054$.

To verify the hypothesis and the choice of τ, I chose to simulate 10^5 realization of Y and first display the histogram of the two hypothesis (last plot of figure 6) and then to compute the p_{err} of the classified events, $p_{\mathrm{err}} = \frac{\mathrm{error}}{N_{\mathrm{events}}} = 0.154$.

3 Code

The code is written in *Python 3*, using the mathematical libraries *numpy* and *scipy*. Attached there is both the *Jupyter notebook* (http://jupyter.org/) and the converted *python* source.