TPK4450: Semester work 2020

Information:

- The semester work is mandatory. It counts 30% to the final grade in the course.
- The semester work is individual. You must submit report individually.
- It consists in 3 parts: Fault diagnosis, failure prognosis, maintenance modelling
- You must submit a report for each part. You should include figures that the exercise asks for, with your comments, explanations and/or analysis.

FAULT DIAGNOSIS – PART I

Problem a)

We are monitoring an industrial process. This process is producing goods that are sensible to temperature variations. The temperature needs to be controlled in order to guarantee the quality of the products. A sensor is providing N = 10 temperature values for one production shift (one production shift = half-day = 4hours). The production process is considered to be viable if the mean temperature on this production shift interval does not deviate significantly from a reference value of $\mu_0 = 0$ °C.

Let us assume that at any point in time, the temperature value follows a normal distribution $X \sim \mathcal{N}(\mu_t, \sigma)$. We will use the sample mean S_N as an indicator of the good functioning of the production process. We need to set up alarm levels or fault detection thresholds based on this metric.

The reference values for the monitored random phenomena are: $\mu_0 = 0,\ \sigma = 2$, i.e. $X \sim \mathcal{N}(0,2)$.

1. If we use the sample mean as a decision metric and samples of size N = 10, and if we want to determine an alarm level for temperatures that may be too elevated, what alarm level should we consider if we want a probability of false alarms of no more than $\alpha = 0.05$?

   Hint: It can be shown that the sample mean has a distribution: $S_N \sim \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{N}}\right)$
2. Plot the PDF of the decision metric and the (upper) alarm level corresponding to probability $\alpha = 0.05$ of false alarms.
3. Generate nsamp = 100 samples from the sample mean distribution (with underlying random phenomena $X \sim \mathcal{N}(0,2)$) , classify each sample mean and count (empirically) the number of classifications as healthy samples and the number of false alarms. (Using the alarm level you have obtained in question 1).

Problem b)

Now suppose that we can propose a model for the failure mode of the temperature control system. Let us assume that if this control system is faulty, the temperature will rise to a new mean level: $\mu_1 = 2$°C. This means that temperature samples we could obtain from a faulty production process would be described by $X \sim \mathcal{N}(2,2)$.

4. Plot both the PDF corresponding to sample mean for a system respectively in faulty mode and healthy mode. If we keep the same alarm level as the one calculated in question 1, compute (empirically) the number of false alarms as well as non-detected faults. Use nsamp = 100 samples from each mode for this empirical estimation (100 from healthy, 100 from faulty)
5. Propose a decision structure to classify a sample mean based on likelihood ratio $\Lambda(s)$, keeping the probability of false alarms $\alpha = 0.05$ and considering that the hypotheses to be tested are:

$$H_0 : S \sim \mathcal{N}\left(\mu_0, \frac{\sigma}{\sqrt{N}}\right)$$

$$H_1 : S \sim \mathcal{N}\left(\mu_1, \frac{\sigma}{\sqrt{N}}\right)$$

6. Study the evolution of the probability of false alarms and non-detections as a function of the decision parameter λ_{NP}. Plot both curves.
7. Plot the so-called Receiver Operating Characteristic (ROC) curve by drafting the proportion of false positives (x-axis) against the proportion of true positives (y-axis). (detection vs false alarms)

Problem c)

We now assume that we monitor the evolution of the production process on 20 shifts (20 half-days) and that the temperature control system is slowly degrading. At time t = 20 shifts, the mean has reached 2°C, following to linear evolution.

8. Draw samples of size N = 10 for each date $t \in [0,20]$ and apply the test using the (upper) alarm level calculated in question 1. Make a plot of the classification of each production shift (e.g. healthy = 0, faulty = 1)
   Hint: you should calculate the empirical mean of the sample to classify it. Remember the alarm level is based on the sample mean.