

## Chapter 4: Statistical tests

### 1 Session 1

*Objective: understand the basics of statistical testing; compute and interpret p-values, for a normal variable, and for a proportion.*

**Steps of the session:**

1. Individual work: Practice with the instructions given on the web site.
2. Share and work: discuss what you understand.
3. Individual work: Practice with the exercises of Session 1 (at least exercises 1 and 3).
4. Share and work: discuss your scripts together.
5. Individual work at home: Finish exercises of this session.
6. Individual work at home: Read the chapter *One Sample tests*, sections *Test on a Gaussian sample*; *Test of a proportion*; *Goodness-of-fit tests*.

**Exercise 1.1.** A packaging machine is supposed to produce packs of 1 kg. The actual weight of a pack is modeled by a random variable following a normal distribution, with a standard deviation of 20 g. However, it is possible to tune the mean weight of the packs.

1. The production manager decides not to distribute packs with weight too far away from the prescribed value of 1 kg. What hypotheses  $H_0$  and  $H_1$  should he test? Establish the decision rule for that test at thresholds 5% and 1%.
2. The company manager thinks that the packs going out on sale are too heavy, causing money loss. What hypotheses  $H_0$  and  $H_1$  should the production manager use to answer the criticism? Establish the decision rule for that test at thresholds 5% and 1%.
3. A pack has been weighed at 1018 grams. What is the p-value for the test of the previous question? What is the p-value for the test of the first question?
4. A consumers' association sues the company for selling packs that are too light. What hypotheses  $H_0$  and  $H_1$  should the production manager use to answer? Establish the decision rule for that test at thresholds 5% and 1%.
5. A pack has been weighed at 982 grams. What is the p-value for the test of the previous question? What is the p-value for the test of the first question?

**Exercise 1.2.** A paracetamol concentration of more than 150 mg per kilogram body weight is considered as dangerous. The measurements of paracetamol in blood tests are modelled by a random variable with normal distribution  $\mathcal{N}(\mu, \sigma)$ . The standard-deviation, linked to the testing method, is supposed to be known and equal to 5 mg.

1. Give the hypotheses and the decision rule for the test deciding, at threshold 5%, whether a patient is at risk, from the results of one blood test (you are a cautious doctor).
2. A patient arrives at the hospital with signs of paracetamol poisoning. A blood test is made and a concentration of 140 mg is found. Give the p-value for the test of the previous question. Should that patient be considered at risk?

**Exercise 1.3.** Throughout the exercise, the variable of interest is supposed to follow a normal distribution. For each of the situations below:

- identify the null hypothesis and the alternative;
- compute a p-value using the function `pnorm`;
- answer the question by making a decision at threshold 5%;
- does your decision change at threshold 1%?
- Download and source the function `normal.pvalue`. It takes a value `x`, a `mean`, a standard deviation `sd`, and an alternative in `"two.sided"`, `"less"`, `"greater"`; Compute again the p-value using that function.

1. The red blood cell count (RBC) of a healthy individual is 5 billions per liter, with a standard deviation of 0.04.
  - (a) Is a value of 5.07 unusual?
  - (b) Is it too large?
  - (c) A patient has a RBC at 4.93. Is he anemic?
2. The total nutritional intake should be 3000 Kcal per day, with standard deviation 250.
  - (a) Is an intake of 3300 Kcal per day reasonable?
  - (b) Is an intake of 3500 Kcal per day too much?
  - (c) Is an intake of 2500 Kcal per day sufficient?
3. The focus of a normal human eye, measured in diopters, is 0 with a standard deviation of 1.4. Nearsighted people have a negative focus, farsighted people have a positive focus.

- (a) Should a focus of  $-2$  be corrected?
  - (b) Is a person with focus measured at  $-3$  nearsighted?
  - (c) Is a person with focus measured at  $+2$  farsighted?
4. Automatic speed camera measurements have a standard deviation of 4% of the measured speed.
- (a) Should a driver measured at 54 Km/h in a city (speed limit 50) be fined?
  - (b) Should a driver measured at 134 Km/h on a freeway (speed limit 130) be fined?
  - (c) Should a driver measured at 100 Km/h on a highway (speed limit 90) be fined?

**Exercise 1.4.** Recall that if  $p$  is the probability of an event, the distribution of the relative frequency over a sample of size  $n$ , can be approximated by the normal distribution with mean  $p$ , standard deviation  $\sqrt{p(1-p)/n}$ . For each of the situations below:

- identify the null hypothesis and the alternative;
  - compute a p-value using the function `pnorm`;
  - answer the question by making a decision at threshold 0.05;
  - does your decision change at threshold 1%?
1. At the European roulette, there are 18 black numbers, 18 red numbers and the zero, which is green. Only bets on black or red are allowed, so that the probability of winning is  $18/37$ .
- (a) Is loosing 20 times out of 30 games unusual?
  - (b) Is winning 30 times out of 40 games unusual?
  - (c) Over a whole night, a gambler has won 55 of his 100 games. Did he cheat?
  - (d) Over a whole season, a very wealthy gambler has lost 552 of his 935 games. Is he hopelessly unlucky?
2. For a certain disease, there exists a treatment that cures 70% of the cases. A laboratory proposes a new treatment claiming that it is better than the previous one. As the expert in charge of deciding whether the new treatment should be authorized, what is your conclusion?
- (a) Out of 100 patients having received the new treatment, 74 of them have been cured.

- (b) Out of 200 patients having received the new treatment, 148 of them have been cured.
  - (c) Out of 1000 patients having received the new treatment, 740 of them have been cured.
3. The percentage of 35-year-old women having wrinkles is 25%. An anti-wrinkle treatment is proposed. Do you consider the treatment as efficient?
- (a) Out of 100 women having followed the treatment, 20 of them still had wrinkles.
  - (b) Out of 200 women having followed the treatment, 40 of them still had wrinkles.
  - (c) Out of 1000 women having followed the treatment, 200 of them still had wrinkles.
4. In order to study the influence of X-rays on the spermatogenesis of *Bombyx mori*, males have been exposed to radiation on the second day and on the fourth day of the larval stage. These males have been mated with non exposed females, and the number of fertile eggs laid by the females have been counted: out of a total of 5646 eggs laid, 4998 were fertile. In a control group of non exposed males and females, 5834 fertile eggs out of 6221 were obtained.
- (a) Is the proportion of fertile eggs from exposed males smaller than 0.9?
  - (b) Is the proportion of fertile eggs from non exposed males larger than 0.9?