Partial credit will be assigned based upon the correctness, completeness, and clarity of your answers. ***Correct answers without proper justification will not receive any credit***.

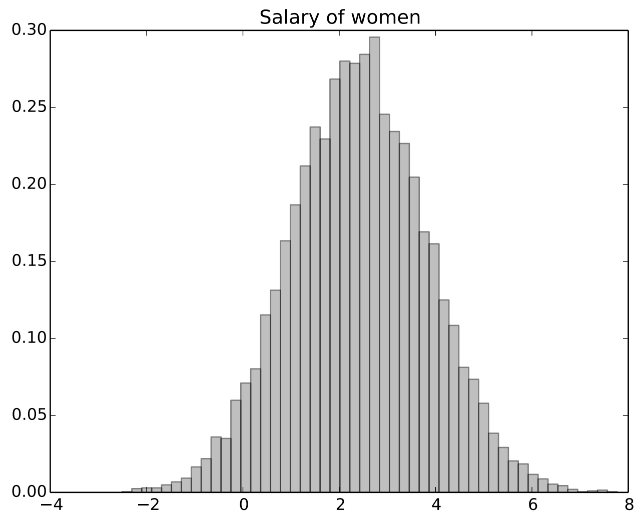
Exercise 1. Let $n \in \mathbb{N}^\star$ and:

- $(X_1, \ldots, X_n)$ be a random sample of the $\mathcal{N}(\mu_1, 2)$ distribution,
- $(Y_1, \ldots, Y_n)$ be a random sample of the $\mathcal{N}(\mu_2, 2)$ distribution,

where μ_1 and μ_2 are two unknown parameters. Assume that the samples are independent. The aim of the exercise is to find a confidence interval for $\mu_2 - \mu_1$.
Let:

$$Z = \sqrt{n}\ \frac{\left(\overline{X}_n - \mu_1\right) - \left(\overline{Y}_n - \mu_2\right)}{2}.$$

1. Show that:
$$\sqrt{n}\ \frac{\overline{X}_n - \mu_1}{\sqrt{2}} \sim \mathcal{N}(0,1).$$
2. Deduce that $Z \sim \mathcal{N}(0,1)$.
3. Find $u > 0$ such that:
$$\mathbb{P}\left(-u \leqslant Z \leqslant u\right) = 99\%.$$
4. Deduce that:
$$\mathbb{P}\left(\mu_2 - \mu_1 \in \left[\overline{Y}_n - \overline{X}_n - \frac{2u}{\sqrt{n}}; \overline{Y}_n - \overline{X}_n + \frac{2u}{\sqrt{n}}\right]\right) = 99\%.$$
5. Application. The salaries of 500 men and 500 women in an entreprise are shown in the following histograms:



The observed average and variance for the salary of men are, respectively:

$$\bar{x} = 2.5\,\text{K\$}, \quad s_n = 1.413.$$

The observed average and variance for the salary of women are, respectively:

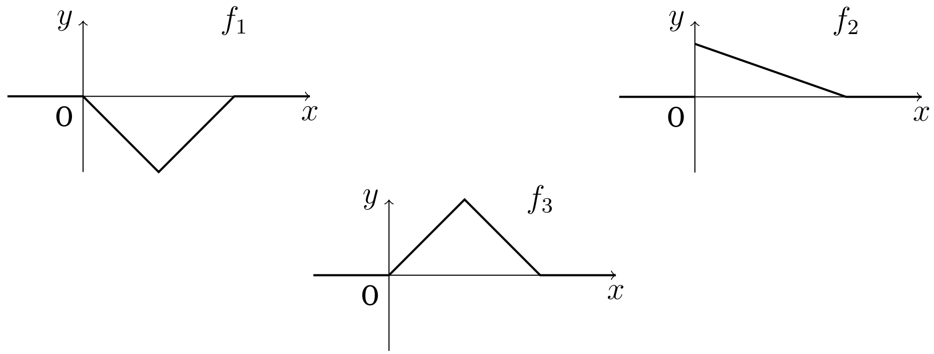
$$\bar{y} = 2.4\,\text{K\$}, \quad s_n = 1.414.$$

(a) Explain why one can apply what has been done in the previous questions to these observed samples.

(b) Give a 99%-confidence interval for the difference of salary between men and women.

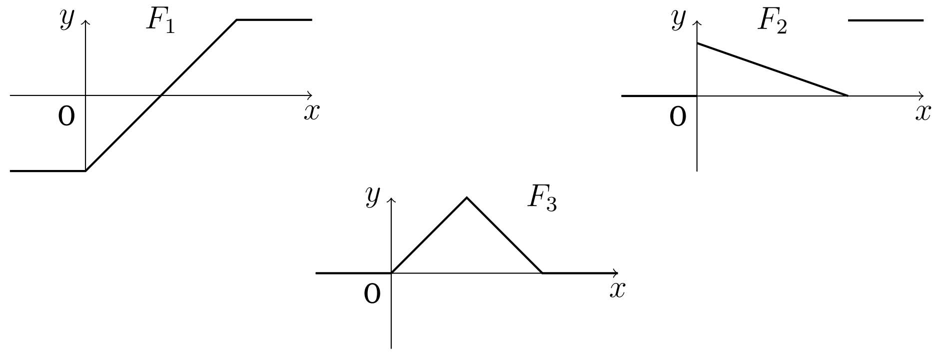
(c) Is there strong evidence that women in this company earn less than men (you don't have to test $\mu_1 \leqslant \mu_2$ against $\mu_1 > \mu_2$: give an answer based on the result of the previous question)?

Exercise 2. At least one of the following functions is not a probability density function. Which one(s) and why (the graphs are not scaled[1])?



Exercise 3. At least one of the following functions is not a cumulative density function. Which one(s) and why (the graphs are not scaled)?

[1]les figures ne sont pas à l'échelle



Exercise 4. In a scratchcard[2], one has to scratch a table with 3 rows and 3 columns. It reveals 3 stars and the ticket is a winning one when the stars are aligned (same row/column or diagonally).

If the stars are (uniformly) randomly printed on each card, what is the probability of winning?

Below are an example of a losing and a winning ticket:

	A	B	C
1	★		
2	★		
3		★	

	A	B	C
1			
2	★	★	★
3			

Exercise 5. Let $U \sim \mathcal{U}([0,1])$ and let $X = 3U + 1$.

1. Compute F_U, the cumulative distribution function of U.
2. Denote by F_X the cumulative distribution function of X.
   (a) Compute $F_X(t)$ for all $t < 1$.
   (b) Compute $F_X(t)$ for all $t \in [1, 4]$.
   (c) Compute $F_X(t)$ for all $t > 4$.
3. Plot F_X.
4. What is the probability distribution of X?

[2]jeu à gratter