Mathematics I

Exercise sheet 1 – Probability and the normal distribution

EM Lyon Business School – MSc Supply Chain & Operation Management – G. Metzler

Content. Course on the normal distribution (density, standardisation, probabilities, quantiles, safety stock), then exercises.
Prerequisites. Mean and standard deviation of a sample, reading a graph, basic Excel. **Tools.** Z table or Excel.

1 Course: the normal distribution

This first sheet focuses on the **normal (or Gaussian) distribution**, the most used probability model in statistics and in operations management. We recall what a continuous random variable is, then we study the normal law, learn how to compute probabilities and quantiles with a table or with Excel, and finish with a key supply chain application: the safety stock.

Why it matters in supply chain

The weight of a can leaving a filling line, the error of a demand forecast, the demand of a fast mover aggregated over a week, the average of many measurements… are well described by a normal law. Knowing it lets you answer questions such as: *what share of my production will be rejected? which machine setting keeps the rejects under 1 %? how much safety stock do I need for a 95 % service level?*

1.1 Continuous random variables: density and cumulative distribution function

A **random variable** X associates a number to the outcome of a random experiment (the weight of the next can, the lead time of the next delivery). It is **continuous** when it can take any value of an interval. Its behaviour is described by a **probability density function** (*pdf*) f: the probability that X falls in an interval is the **area under the curve** of f over this interval.

Definition 1.1: Density and cumulative distribution function

The **cumulative distribution function** (*cdf*) of X is the function

$$F(t) = \mathbb{P}[X \le t] = \int_{-\infty}^{t} f(x)\,dx.$$

The total area under f is equal to 1, and for a continuous variable $\mathbb{P}[X = t] = 0$: only intervals have a probability.

Three rules follow directly from the fact that the total area is 1 (Figure 2):

$$\mathbb{P}[X \ge t] = 1 - F(t), \qquad \mathbb{P}[t_1 \le X \le t_2] = F(t_2) - F(t_1), \qquad \mathbb{P}[X \le t] = \mathbb{P}[X < t].$$

Expectation and variance. As for a sample, a random variable has a mean, called its **expectation** $\mathbb{E}[X]$ (the long-run average of the values), and a **variance** $Var(X) = \mathbb{E}\big[(X - \mathbb{E}[X])^2\big]$ which measures the spread; the standard deviation is $\sigma = \sqrt{Var(X)}$. For all real numbers a:

(i) $\mathbb{E}[X + a] = \mathbb{E}[X] + a$, (ii) $\mathbb{E}[aX] = a\,\mathbb{E}[X]$,

(iii) $Var(X + a) = Var(X)$, (iv) $Var(aX) = a^2\,Var(X)$.

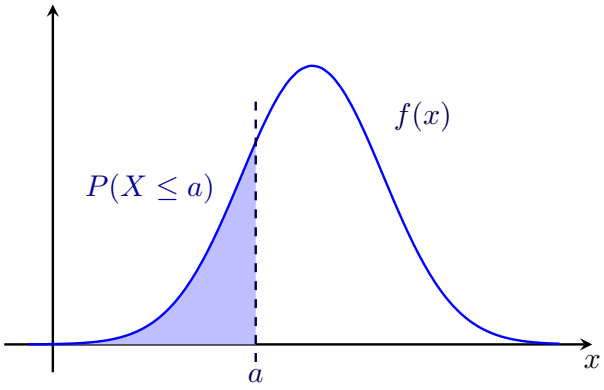

Figure 1: The probability $\mathbb{P}[X \leq a]$ is the area under the density, on the left of a.

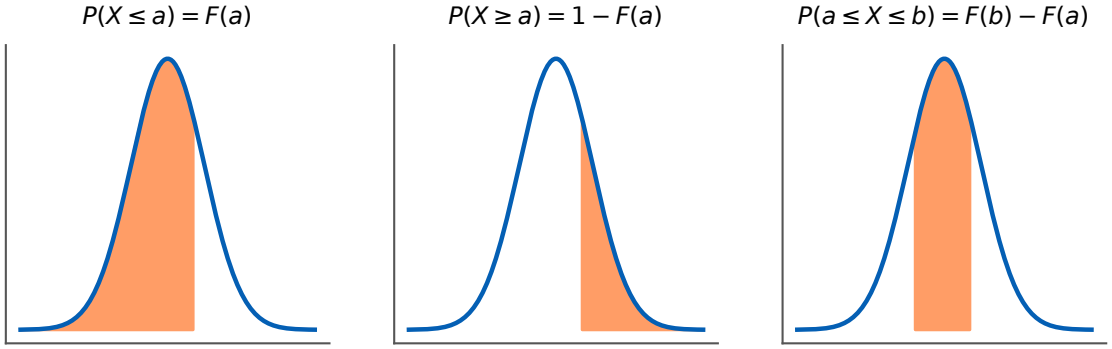

Figure 2: Every probability on a continuous variable is an area, computed from the cdf F.

Shifting all the values does not change their spread (iii); converting kilograms into grams multiplies the variance by 1000^2 (iv).

1.2 The normal distribution

Definition 1.2: Normal distribution

A random variable X follows a normal distribution with mean μ and standard deviation $\sigma > 0$, denoted $X \sim \mathcal{N}(\mu, \sigma^2)$, if its density is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \qquad x \in \mathbb{R}.$$

Then $\mathbb{E}[X] = \mu$ and $Var(X) = \sigma^2$.

Notation. We write $\mathcal{N}(\mu, \sigma^2)$ with the *variance*: $\mathcal{N}(85,\, 0.8^2)$ means a mean of 85 and a standard deviation of 0.8. Always check which convention a document or a software uses.

The density is the famous **bell curve** (Figure 3). Its main properties are:

- it is **symmetric** around μ, so **mean = median = mode** $= \mu$, and for all t: $\mathbb{P}[X \leq \mu - t] = \mathbb{P}[X \geq \mu + t]$;
- μ sets the **position** of the bell, σ its **width**: a larger σ gives a flatter, wider curve (Figure 4);
- the cdf has no explicit formula: probabilities are read in a **table** or computed by a software.

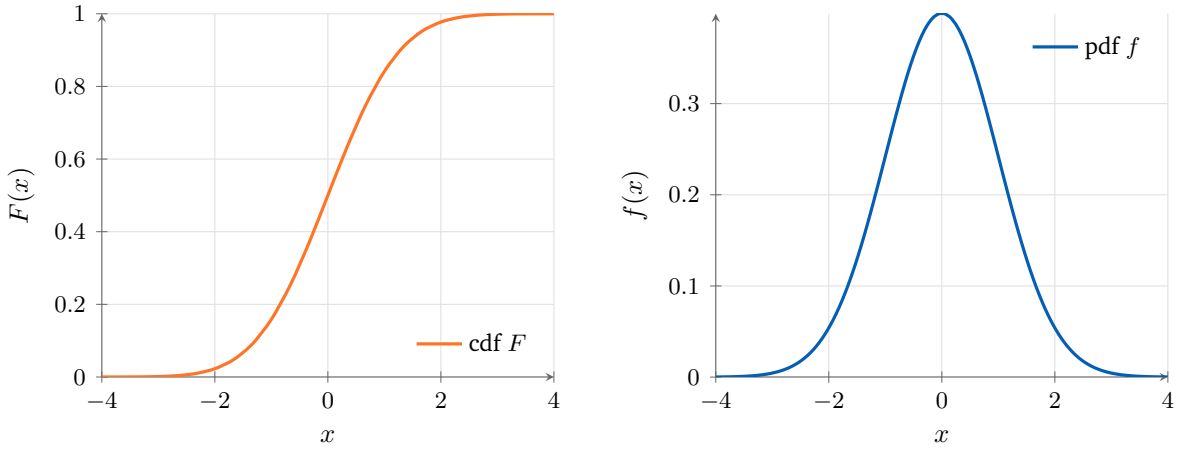

Figure 3: The cdf F (left) and the density f (right) of $X \sim \mathcal{N}(0,1)$. $F(t)$ is the area under f on the left of t; for instance $F(1.38) = \mathbb{P}[X \leq 1.38] = 0.9162$.

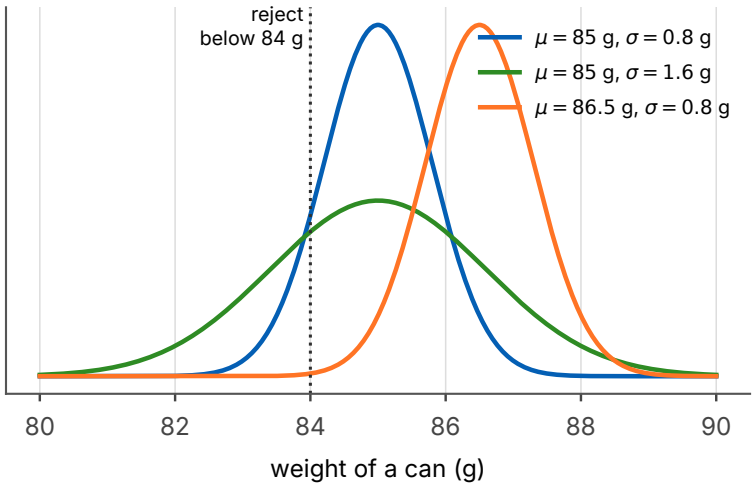

Figure 4: Weight of 85 g cans for three settings of a filling machine. Increasing μ moves the curve to the right (fewer cans under 84 g, but more product given away); a less precise machine (larger σ) spreads the weights and increases the rejects on both sides.

Some important figures. Whatever μ and σ, the same proportions of values lie within a given number of standard deviations around the mean (Figures 5 and 5):

- about 68 % of the values of $X \sim \mathcal{N}(\mu, \sigma^2)$ fall in $[\mu - \sigma, \mu + \sigma]$;
- 95 % fall in $[\mu - 1.96\sigma, \mu + 1.96\sigma]$ (about $\mu \pm 2\sigma$);
- 99 % fall in $[\mu - 2.58\sigma, \mu + 2.58\sigma]$ and 99.7 % in $[\mu - 3\sigma, \mu + 3\sigma]$.

In practice: the "Six Sigma" vocabulary

Quality engineers express the performance of a process by the number of standard deviations between the mean and the nearest tolerance limit. A process "at 3σ" produces about 0.27 % of out-of-tolerance parts (2 700 per million); the *Six Sigma* target (6σ, allowing a drift of the mean of 1.5σ) corresponds to 3.4 defects per million. Reducing σ is often cheaper than paying for rejects.

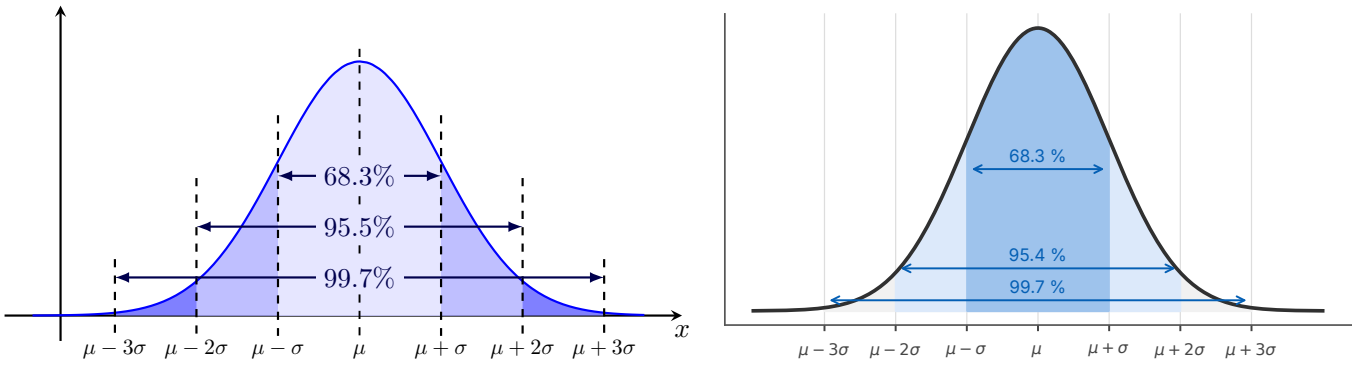

Figure 5: Proportion of the values of a normal variable within 1, 2 and 3 standard deviations of the mean.

1.3 Computing probabilities: the standardisation

Tables only exist for one normal law: the **standard normal distribution** $\mathcal{N}(0,1)$, whose cdf is denoted Φ. Every other normal law is brought back to it.

Theorem 1.1: Standardisation

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then the random variable

$$Z = \frac{X-\mu}{\sigma}$$

follows a normal distribution with mean 0 and standard deviation 1. Consequently, for every real number t,

$$\mathbb{P}[X \leq t] = \mathbb{P}\left[Z \leq \frac{t-\mu}{\sigma}\right] = \Phi\left(\frac{t-\mu}{\sigma}\right).$$

Proof. Using (i) and (iii), the variable $X' = X - \mu$ has expectation 0 and the same variance σ^2 as X. Using (ii) and (iv), $Z = X'/\sigma$ has expectation 0 and variance $\sigma^2/\sigma^2 = 1$. A linear transformation of a normal variable remains normal. □

We say that we have *centred* and *reduced* (or *standardised*) X. The value $z = \frac{t-\mu}{\sigma}$ is called the z**-score** of t: it counts how many standard deviations t is away from the mean. Figure 6 shows the two steps.

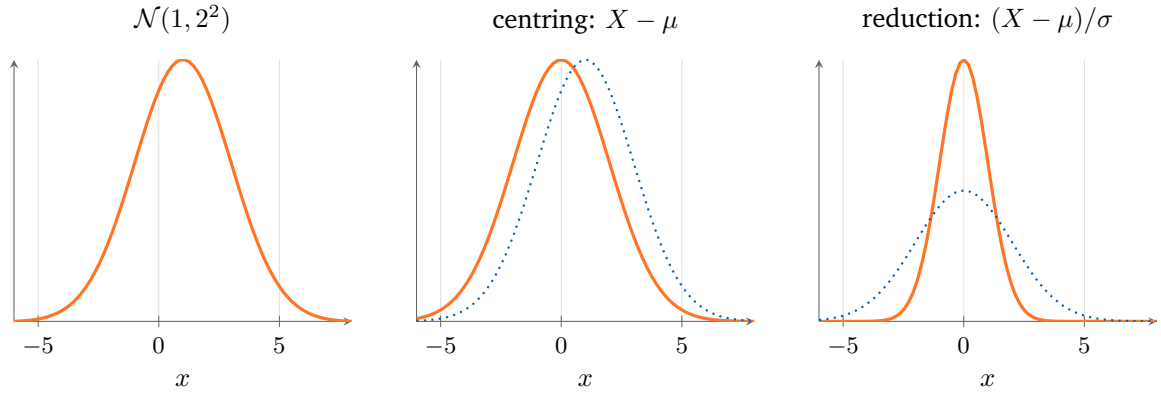

Figure 6: The standardisation steps. Left: the initial distribution. Middle: centring moves the curve to 0. Right: reducing (dividing by σ) gives a variance equal to 1. Dotted lines: the distribution at the previous step.

The method, step by step. To compute $\mathbb{P}[X \leq t]$, $\mathbb{P}[X \geq t]$ or $\mathbb{P}[t_1 \leq X \leq t_2]$:

1. write the event with inequalities "$\leq$" only, using $\mathbb{P}[X \geq t] = 1 - \mathbb{P}[X \leq t]$ and $\mathbb{P}[t_1 \leq X \leq t_2] = \mathbb{P}[X \leq t_2] - \mathbb{P}[X \leq t_1]$;
2. standardise each bound: $z = \frac{t-\mu}{\sigma}$;
3. read $\Phi(z)$ in the table; for a negative z, use the symmetry $\Phi(-z) = 1 - \Phi(z)$.

How to read the Z table. The table gives $\Phi(z)$ for $z \geq 0$: the row gives the units and the first decimal, the column the second decimal. For $z = 1.38$: row 1.3, column 0.08, $\Phi(1.38) = 0.9162$. Values to know: $\Phi(1) = 0.8413$, $\Phi(1.645) = 0.95$, $\Phi(1.96) = 0.975$, $\Phi(2.326) = 0.99$, $\Phi(2.576) = 0.995$.

Example 1.1. *Let $X \sim \mathcal{N}(5, 2^2)$. Let us compute the probability that X takes values lower or equal than 1, i.e. $\mathbb{P}[X \leq 1]$.*

$$
\begin{aligned}
&\quad \downarrow \textit{ we subtract the mean of } X \textit{ and divide by its standard deviation on both sides of the inequality} \\
\mathbb{P}[X \leq 1] &= \mathbb{P}\left[\frac{X-5}{2} \leq \frac{1-5}{2}\right], \\
&\quad \downarrow \textit{ we simplify the right hand-side} \\
&= \mathbb{P}\left[\frac{X-5}{2} \leq -2\right], \\
&\quad \downarrow \textit{ we set } Z = \frac{X-5}{2} \textit{ and } Z \sim \mathcal{N}(0,1) \\
&= \mathbb{P}\left[Z \leq -2\right], \\
&\quad \downarrow \textit{ we use the symmetry} \\
&= \mathbb{P}\left[Z \geq 2\right], \\
&\quad \downarrow \textit{ we use the fact that } \mathbb{P}[Z \geq t] = 1 - \mathbb{P}[Z \leq t] \\
&= 1 - \mathbb{P}\left[Z \leq 2\right], \\
&\quad \downarrow \textit{ we read the value in the } Z \textit{ table or use Excel} \\
&= 1 - 0.977 = 0.023.
\end{aligned}
$$

Example 1.2 (Supply chain example: transit time of a carrier)**.** *The transit time of a road carrier between Lyon and Hamburg is $T \sim \mathcal{N}(48,\, 6^2)$ hours.*

- *Probability that a shipment arrives in less than 36 h:* $\mathbb{P}[T \leq 36] = \Phi\left(\frac{36-48}{6}\right) = \Phi(-2) = 1 - 0.977 = 0.023$.
- *Probability that it arrives late, after 60 h:* $\mathbb{P}[T > 60] = 1 - \Phi(2) = 0.023$ *(by symmetry, the same value).*
- *Probability that it arrives between 42 h and 54 h:* $\Phi(1) - \Phi(-1) = 2 \times 0.8413 - 1 = 0.683$*: about two shipments out of three.*

With Excel: probabilities

Version	Function
ENGLISH	`NORM.DIST(`t`; `μ`; `σ`; TRUE)` and `NORM.S.DIST(`z`; TRUE)` for $\mathcal{N}(0,1)$
FRENCH	`LOI.NORMALE.N(`t`; `μ`; `σ`; VRAI)` and `LOI.NORMALE.STANDARD.N(`z`; VRAI)`

where μ and σ are the mean and the **standard deviation** (not the variance!). The last parameter is called *cumulative*: **TRUE** computes $\mathbb{P}[X \leq t] = F(t)$, **FALSE** computes the value of the

density $f(t)$ (rarely useful). Example: `=NORM.DIST(36;48;6;TRUE)` returns 0.0228.

1.4 Computing quantiles: working backwards

It is often necessary to do the reverse operation: we know a **probability** and we look for the corresponding **value**. This is what we do when we build confidence intervals and tests (sheets 2 and 3), when we set a machine to keep the rejects under a target, or when we size a safety stock.

Definition 1.3: Quantile

The quantile of order $\alpha \in]0,1[$ of X is the value t_α such that $\mathbb{P}[X \leq t_\alpha] = \alpha$, *i.e.* $t_\alpha = F^{-1}(\alpha)$. For $X \sim \mathcal{N}(\mu, \sigma^2)$:

$$t_\alpha = \mu + z_\alpha\, \sigma,$$

where z_α is the quantile of order α of $\mathcal{N}(0,1)$. By symmetry, $z_{1-\alpha} = -z_\alpha$.

For instance $\alpha = 0.5$ gives the median μ. Usual quantiles of $\mathcal{N}(0,1)$: $z_{0.90} = 1.282$, $z_{0.95} = 1.645$, $z_{0.975} = 1.960$, $z_{0.99} = 2.326$, $z_{0.995} = 2.576$ (Figure 7).

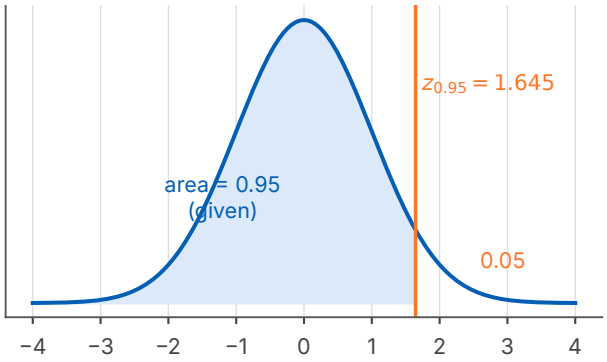

Figure 7: The quantile reads the table "backwards": the area is given, the value z is unknown.

Example 1.3 (Setting a filling machine)**.** *Bottles are filled with* $X \sim \mathcal{N}(\mu, 4^2)$ *ml and rejected below 495 ml. Which setting* μ *keeps the rejects at* $1\,\%$*?*

$$\mathbb{P}[X < 495] = 0.01 \iff \frac{495-\mu}{4} = z_{0.01} = -2.326 \iff \mu = 495 + 2.326 \times 4 = 504.3 \text{ ml}.$$

With $\mu = 500$ *ml, the rejects would be* $\Phi(-1.25) = 10.6\,\%$*. Each extra millilitre is product given away: the quantile gives the* minimum *setting compatible with the target.*

With Excel: quantiles

Version	Function
ENGLISH	`NORM.INV(`p`;` μ`;` σ`)` and `NORM.S.INV(`p`)` for $\mathcal{N}(0,1)$
FRENCH	`LOI.NORMALE.INVERSE.N(`p`;` μ`;` σ`)` and `LOI.NORMALE.STANDARD.INVERSE.N(`p`)`

where $p \in]0,1[$ is the order of the quantile. Example: `=NORM.INV(0.95;48;6)` returns 57.9: 95 % of the shipments of the previous carrier arrive within 57.9 h, a reasonable service promise.

1.5 Application: safety stock and service level

During the replenishment lead time L (days), the demand is uncertain. If the daily demands are independent with mean μ and standard deviation σ, the demand over the lead time has mean $\mu_L = L\mu$ and standard deviation $\sigma_L = \sigma\sqrt{L}$ (variances add up, standard deviations do not), and it is close to a normal law.

Definition 1.4: Reorder point and safety stock

To reach a **cycle service level** CSL (probability of no stock-out during a replenishment cycle), we reorder when the stock falls to

$$ROP = \mu_L + SS, \qquad \text{with the safety stock } SS = z_{CSL}\,\sigma_L.$$

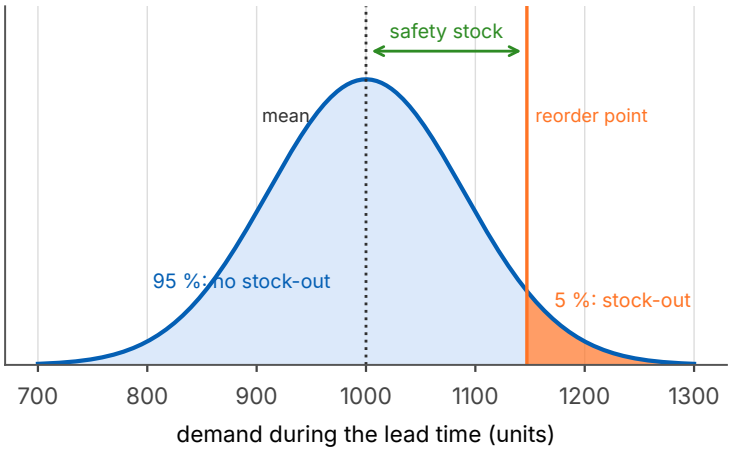

Figure 8: Demand during the lead time. The reorder point is the quantile of order CSL; the safety stock is the distance between this quantile and the mean.

Example 1.4 (Safety stock of a fast mover)**.** *Daily demand* $\mathcal{N}(200, 40^2)$ *units, lead time* $L = 5$ *days:* $\mu_L = 1000$ *and* $\sigma_L = 40\sqrt{5} = 89.4$ *units. For a* $95\,\%$ *service level,* $SS = 1.645 \times 89.4 = 147$ *units and* $ROP = 1\,147$ *units. For* $99\,\%$*,* $SS = 2.326 \times 89.4 = 208$ *units: four points of service cost* $41\,\%$ *more safety stock.*

1.6 Is my variable normal?

All the computations above assume a normal law. It is a good model for measurements around a target, for errors and for totals or averages of many values, but **not** for individual lead times or delays (skewed to the right), for the demand of slow movers (many zeros) or for counts of rare events (Poisson law). Before using it, check:

- the **histogram**: one symmetric bump? the **boxplot**: median in the middle of the box?
- mean $\approx$ median, skewness close to 0, about $95\,\%$ of the values in $\bar{x} \pm 2s$;
- the **QQ-plot** (Figure 9): the sorted data against the quantiles of a normal law should lie on a straight line.

Important to remember

- $X \sim \mathcal{N}(\mu, \sigma^2)$: symmetric bell curve, mean = median = μ, spread σ; probabilities are areas.

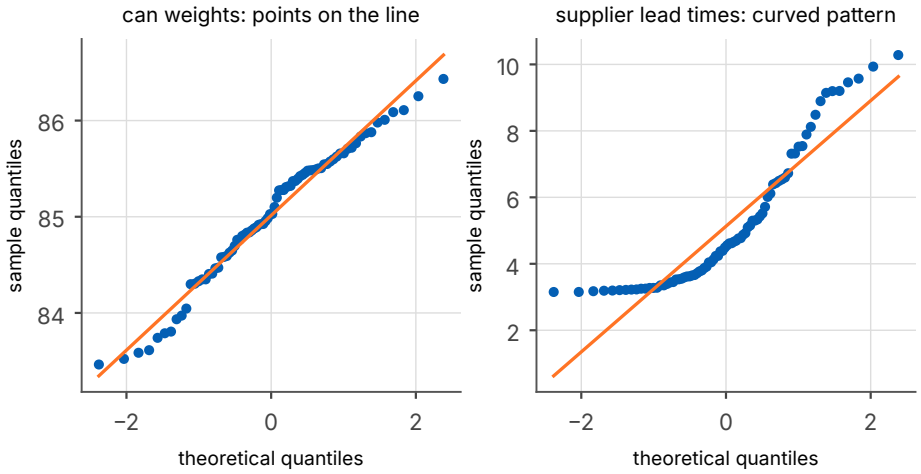

Figure 9: QQ-plots. Left: weights of cans, compatible with a normal law. Right: lead times of a supplier, clearly skewed.

- $\mathbb{P}[X \leq t] = \Phi\left(\frac{t-\mu}{\sigma}\right)$; $\mathbb{P}[X \geq t] = 1 - \mathbb{P}[X \leq t]$; $\Phi(-z) = 1 - \Phi(z)$.
- Quantile: $t_\alpha = \mu + z_\alpha \sigma$, with $z_{0.95} = 1.645$, $z_{0.975} = 1.96$, $z_{0.99} = 2.326$.
- $95\,\%$ of the values in $\mu \pm 1.96\sigma$, $99\,\%$ in $\mu \pm 2.58\sigma$.
- Safety stock: $SS = z_{CSL}\,\sigma\sqrt{L}$.
- Excel: `NORM.DIST`, `NORM.INV` (French `LOI.NORMALE.N`, `LOI.NORMALE.INVERSE.N`). Check normality before using them.

2 Exercises

Exercise 2.1. *Ages in a city are considered to be distributed according to a normal distribution $\mathcal{N}(\mu, \sigma^2)$, with the values of μ and σ depending on the city. An individual is considered young if he is under 25 and old if he's over 50. Finally, an individual whose age is between 30 and 40 will be considered to be in "golden age".*

We assume that the age of the population in Strasbourg follows a normal distribution with parameters $\mu = 33$ and $\sigma = 20$, whereas in Lyon, the population follows a normal distribution with parameters $\mu = 35$ and $\sigma = 22$.

a) What are the limitations of this model?

b) In which city can we find the most young people?

c) Conversely, in which city can we find the most old people?

d) Finally, in which city can we find the most people in their prime?

Exercise 2.2 (Probabilities and Costs)**.**
We want to study the distribution of supply times (expressed in hours) between a supplier and a set of stores for a particular brand for the coming year.
Studies carried out the previous year showed that these times, are distributed according to a normal law of mean $\mu = 1$ (and standard deviation $\sigma = 0.1$.

What is the probability that the supply time is:

1. *lower than one hour*
2. *between 0.95 and 1 hour*
3. *between 1 and 1.05 hour*
4. *lower than 0.95 or greater than 1.05.*

It is assumed that the supply cost is time dependent and represented by a random variable:

- *the cost is equal to $1,000\$$ when the the supply time is lower than 0.95*
- *the cost is equal to $1,200\$$ when the the supply time is between than 0.95 and 1*
- *the cost is equal to $1,400\$$ when the the supply time is between than 1 and 1.05*
- *the cost is equal to $1,500\$$ when the the supply time is greater than 1.05*

Determine the mean average cost using the previous questions' results.

Exercise 2.3 (Study of an Assembly line)**.**
We consider the production of $85g$ cans. The process includes: filling in the cans, sealing, weighing and disposal if the weight is less than $85g$ (production waste), sorting, labeling. Even if the automatons are of very good quality, it is impossible to obtain cans weighing exactly $85g$. There is a standard deviation in packing the cans related to the technical processes, which follows a normal distribution of the standard deviation equal to $0.8g$.

1. *We aim to determinate the mean value of packing the cans that has to be used so that the waste related to the weight of package is less than 2%.*

We are now interested in the production of $100g$ *packages. Once the mixture has been prepared, it is packaged by the automatons. The process includes: packing, sealing, weighing and disposal if the weight is less than* $98.5g$ *or bigger than* $101.5g$*, sorting, labeling. The automatons are the same as those used for the* $85g$ *packages. The standard deviation in packing, related to the technical processes, follows, also for the* $100g$ *packages, a normal distribution of the standard error equal to* $0.8g$*.*

2. *Knowing that the automatons are set to a mean* $\mu = 100g$*, we are interested in the percentage of packages that are then discarded as waste, statistically speaking.*
3. *Would it be appropriate to set the machines to a mean* $\mu = 100.5$*?*

Exercise 2.4 (Probabilities)**.** *You were told that the amount of time lapsed between consecutive commands on a website followed a normal distribution with a mean of* 15 *seconds. You were also told that the probability that the time lapsed between two consecutive commands to fall between* 16 *to* 17 *seconds was* 13%*. The probability that the time lapsed between two consecutive commands would fall below* 13 *seconds was* 7%*.*

What is the probability that the time lapsed between two consecutive commands will be between 13 *and* 16 *seconds?*

Exercise 2.5. *It is assumed that the average distance covered by a delivery driver during a delivery day follows a Gaussian distribution. During a first period of measure, we find that* 10% *of all journeys are longer than* $250km$ *and exactly* 25% *of the journey are less longer than* $150km$*.*

Compute the average distance covered by the delivery driver, and the standard deviation of this distance.