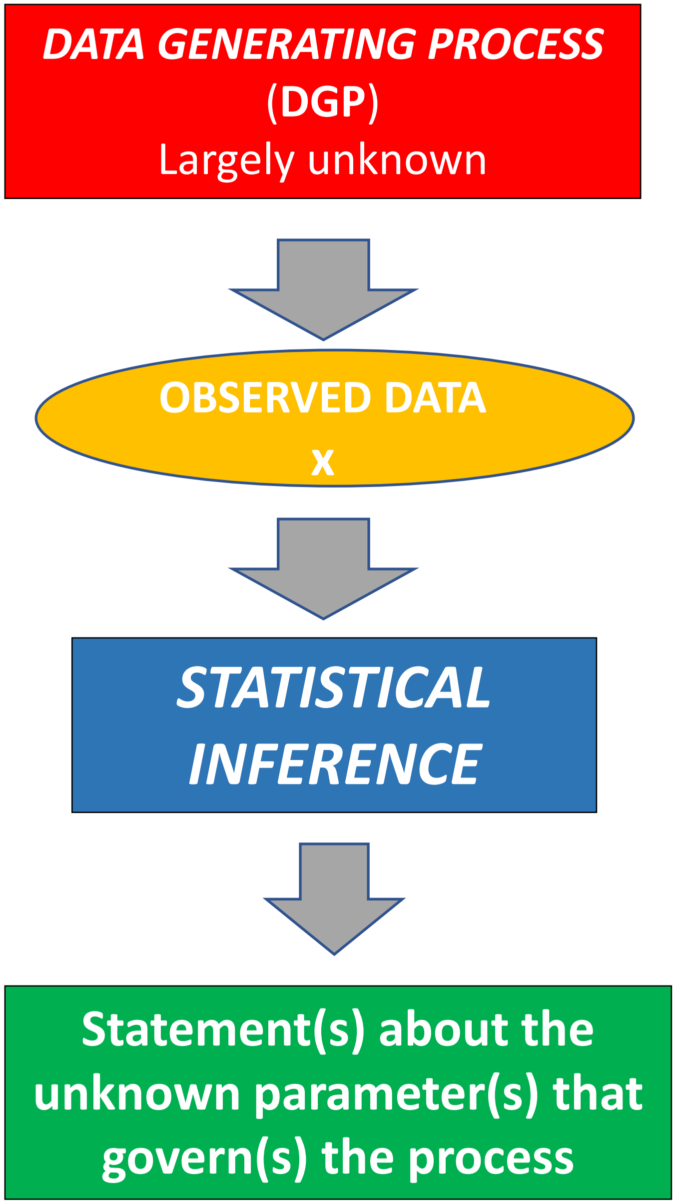
Statistical inference

The use of a set of data to draw information (make statements) about the process that generated those data is referred to as statistical inference.

The data generating process (DGP) can be modeled by a certain probability density function (PDF) $f(\mathbf{x}|\boldsymbol{\theta})$ with $\boldsymbol{\theta} = \theta_1, \theta_2, \ldots, \theta_k$ representing a vector of parameters and $\mathbf{x} = x_1, x_2, \ldots, x_n$ representing a vector of data.

If, for example, X is distributed according to a normal PDF, $N(\mu,\sigma^2)$, i.e., a two-parameter PDF, the observed data $(x_1,x_2,\ldots,x_n)$ are a random sample from a normal distribution with parameters μ *and* σ^2.

Note that if **one** (or more) **parameter**(s) is(are) left as **unknown**, not a unique distribution, but a family of distributions is specified.



Two major approaches to statistical inference are the following:

- Frequentist
- Bayesian

Note that, in the statistical context, a distinction is made between terms «likelihood» and «probability», depending on the aspect focused on, whether outcomes or parameters:

«probability» is used before data are available, to describe the plausibility of a future outcome, given a value for the parameter;

«likelihood» is used after data are available, to describe the plausibility of a parameter value.

The three major types of statistical inference are:

- ✓ Point Estimation (what single value of the parameter is most appropriate?)
- ✓ Interval Estimation (what interval of parameter values is most consistent with data?)
- ✓ Hypothesis Testing (which of two values of the parameter is most consistent with data?)

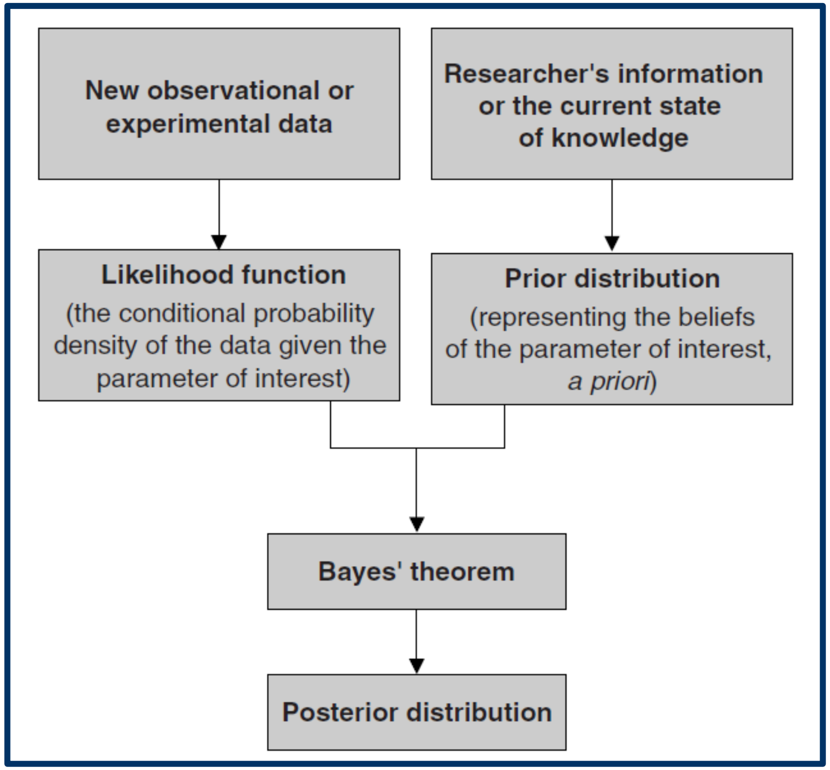
Frequentist approach

This approach judges inferences based on their performance in repeated sampling, i.e., based on the sampling distribution of the statistic used to make the inference. Several *ad hoc* methods are used to select the statistics used for inference.

Bayesian approach

This approach assumes that the inference problem is subjective and proceeds by:

- ✓ eliciting a prior distribution of the parameter;
- ✓ combining prior distribution with the density of data to obtain the joint distribution of the parameter and the data;
- ✓ using Bayes' Theorem to obtain the posterior distribution of the parameter, given the data.



Frequentist approach - point estimation

Let us assume that we have data $(x_1,x_2,\ldots,x_n)$, which represent a random sample from a normal distribution with parameters μ and σ^2 and, for simplicity, let us consider σ^2 as known.

The probability density function in this case is:

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\sigma^2}\right]$$

Available data have to be used to determine an estimate of μ.

The frequentist approach uses the sample mean of data, $\bar{x}$, as a point estimate of μ, based on the fact that the sampling distribution related to f(x) is itself a normal distribution centered on μ and having a sampling variance σ^2/n.

Frequentist approach - interval estimation (confidence interval)

In this approach an interval of parameter values consistent with data or supported by data has to be found.

At this aim Jerzy Neyman introduced confidence intervals in 1937, giving them the following definition:

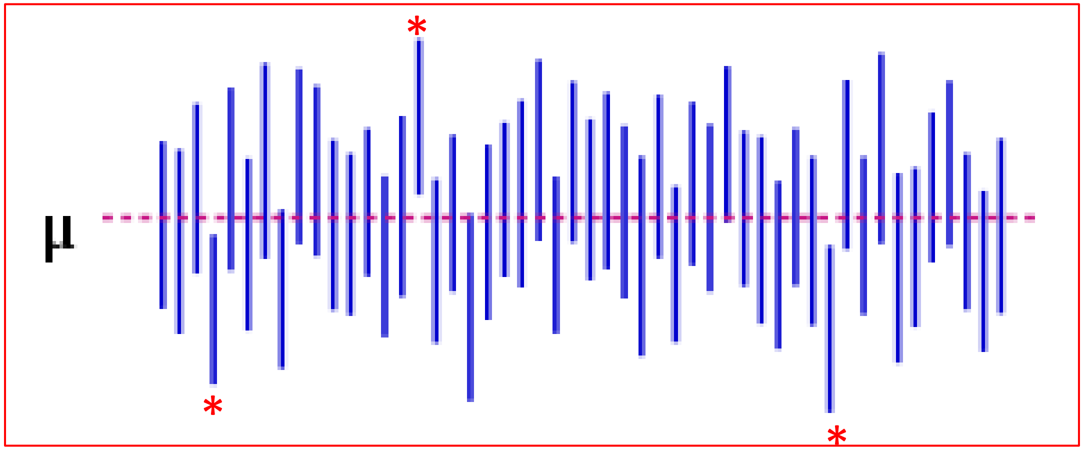
«A P% confidence interval for a parameter is an interval generated by a procedure that, on repeated sampling, has a P% of containing the true value of the parameter»

The confidence level indicates, then, the proportion of observed intervals, obtained from many separate data analyses of replicated experiments, that contain the true value of the parameter.

The P value is called the confidence level, usually indicated as $1-\alpha$, whereas α is called significance, or confidence, coefficient.

As a consequence of the original definition, a confidence interval among those calculated from several samples might not include the true value of the parameter.

In the following picture blue vertical line segments represent 50 calculations of confidence intervals for the population mean μ, represented as a red horizontal dashed line:



Note that some confidence intervals do not contain the population mean, as expected by definition.

If we randomly choose one realization, in 95% of cases we end up having chosen an interval that contains the parameter; however, we may be unlucky and pick the wrong one (although the parameter is not expected to be far from a “wrong” interval).

The procedure leading to the construction of the 50 intervals is referred to as confidence procedure.

Types of confidence interval for the mean when a normally-distributed variable is considered

First case: the variance, σ^2, is known

$$\overline{X} - \frac{z_{(1-\alpha/2)}\,\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + \frac{z_{(1-\alpha/2)}\,\sigma}{\sqrt{n}}$$

Second case: the variance, σ^2, is unknown; n > 30

$$\overline{X} - \frac{z_{(1-\alpha/2)}\,s}{\sqrt{n}} \leq \mu \leq \overline{X} + \frac{z_{(1-\alpha/2)}\,s}{\sqrt{n}}$$

Third case: the variance, σ^2, is unknown; n < 30

$$\overline{X} - \frac{t_{n-1\ (1-\alpha/2)}\,s}{\sqrt{n}} \leq \mu \leq \overline{X} + \frac{t_{n-1\ (1-\alpha/2)}\,s}{\sqrt{n}}$$

n-1	$t_{0.95}$	$t_{0.975}$	$t_{0.995}$
1	6.31	12.71	63.66
2	2.92	4.30	9.92
3	2.35	3.18	5.84
4	2.13	4.78	4.60
5	2.01	2.57	4.03
6	1.94	2.45	3.71
7	1.89	2.37	3.50
8	1.86	2.31	3.55
9	1.83	2.26	3.25
10	1.81	2.23	3.17
20	1.72	2.09	2.85
30	1.70	2.04	2.75
60	1.67	2.00	2.66
120	1.66	1.98	2.62
∞	1.645	1.96	2.58

Bayesian approach: Bayes' theorem

Thomas Bayes was an English statistician, philosopher and Presbyterian minister, who lived in the first half of the 18^{th} century.

The theorem named after him is stated mathematically using the following equation:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

where A and B are events and P are probabilities, and $P(B) \neq 0$.

P(A|B) is a conditional probability: the probability of event A to occur if B is true.

P(B|A) is another conditional probability: the probability of event B to occur if A is true

P(A) and P(B) are called marginal probabilities: the probabilities of observing A and B independently on each other.

A simple example of the Bayesian approach

Suppose that a test for using a particular drug is 99% sensitive and 99% specific.
This means that the test will produce 99% true positive results for drug users and 99% true negative results for non-drug users.
Suppose that 0.5% of people are users of the drug.
What is the probability that a randomly selected individual with a positive test is a drug user?

The calculation based on the Bayes' theorem is:

$$P(\text{User} \mid +) = \frac{P(+ \mid \text{User})P(\text{User})}{P(+)}$$

In this equation P(+|User), the probability that the test will be positive for a drug user, is known by hypothesis, i.e., 99% (0.99).
P(User), the probability of finding a drug user in the population, is also known by hypothesis, i.e., 0.5 % (0.005).

P(+) is the total probability that the test is positive.

P(+) has to account for true positive tests but also for false positive tests, i.e., cases in which the test is positive although the person is NOT a drug user. Then:

$$P(+) = \boxed{P(+ \mid \text{User})P(\text{User})} + \boxed{P(+ \mid \text{Non-user})P(\text{Non-user})}$$

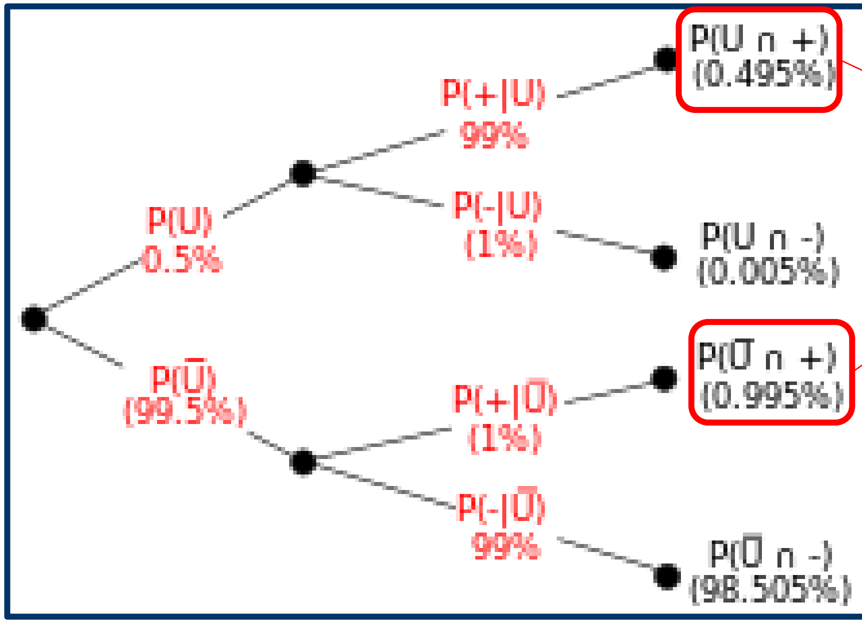
Under the initial hypothesys P(+| Non-user) is 1% (0.01) and P(Non-user) is 99.5% (0.995).

The final calculation is, then

$$\begin{aligned} P(\text{User} \mid +) &= \frac{P(+ \mid \text{User})P(\text{User})}{P(+)} \\ &= \frac{P(+ \mid \text{User})P(\text{User})}{P(+ \mid \text{User})P(\text{User}) + P(+ \mid \text{Non-user})P(\text{Non-user})} \\ &= \frac{0.99 \times 0.005}{0.99 \times 0.005 + 0.01 \times 0.995} \approx 33.2\% \end{aligned}$$

Note that the key probability, in this case, is that of false positives, P(+|Non-user), since it is multiplied by a number close to unity (i.e., the probability of finding a non-user, which is high), thus, if it is not small, the denominator is increased, and the P(User|+) is decreased.

The calculation can be visualized by the scheme below, in which U are users and Ū are non-users:



$$\frac{0.99 \times 0.005}{0.99 \times 0.005 + 0.01 \times 0.995} \approx 33.2\%$$

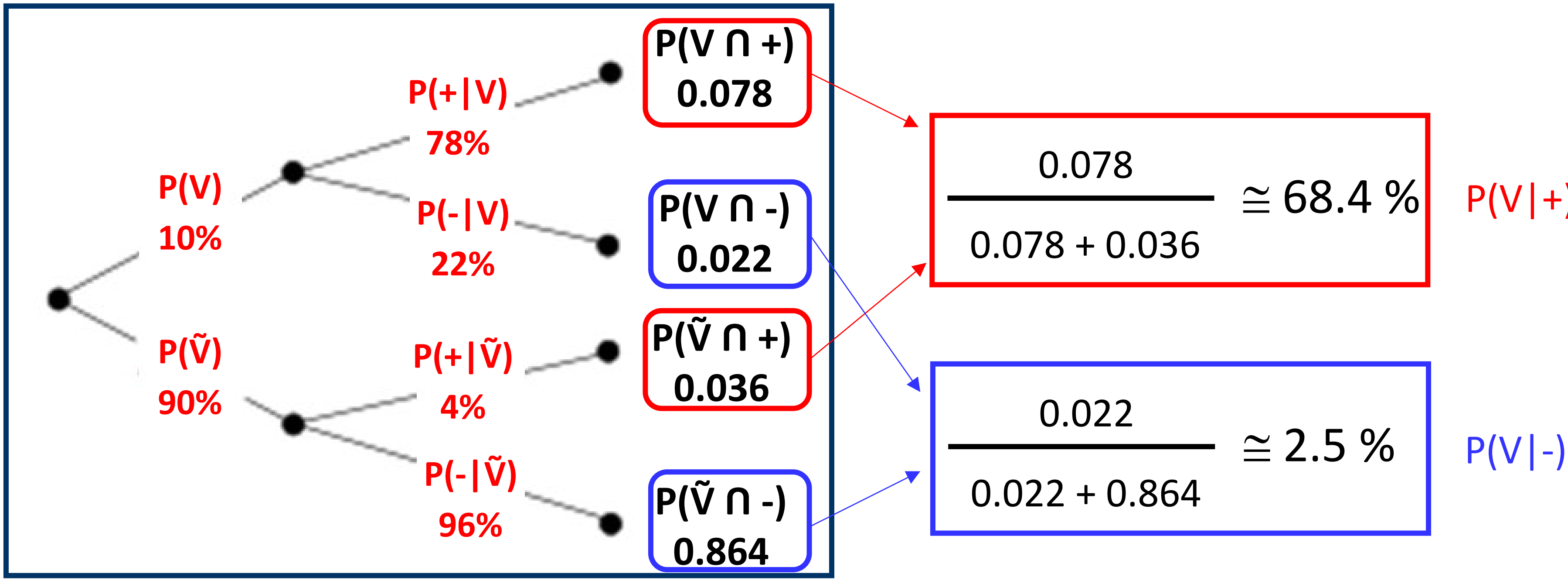
Note that the final probability, 33.2%, arises from the ratio between P(U ∩ +) and the sum between it and P(Ū ∩ +), where the symbol ∩ indicates that both U and + events have to occur.

It is clear that only by reducing further P(+|Ū), the probability of false positives, the P(Ū ∩ +) can be reduced and thus the P(U ∩ +) has a greater weight on the total of positive results.

Application of the Bayes' theorem to fast antigenic tests for SARS-CoV 2

Let us consider that the sensitivity, indicated as P(+|V), and the selectivity, indicated as P(-|Ṽ), of fast antigenic tests for SARS-CoV 2 are 78 and 96%, respectively, which are values averaged for tests provided by different manufacturers, and that the actual incidence of the virus in the population is 10%, indicated as P(V).

Calculations based on the Bayes' theorem are:



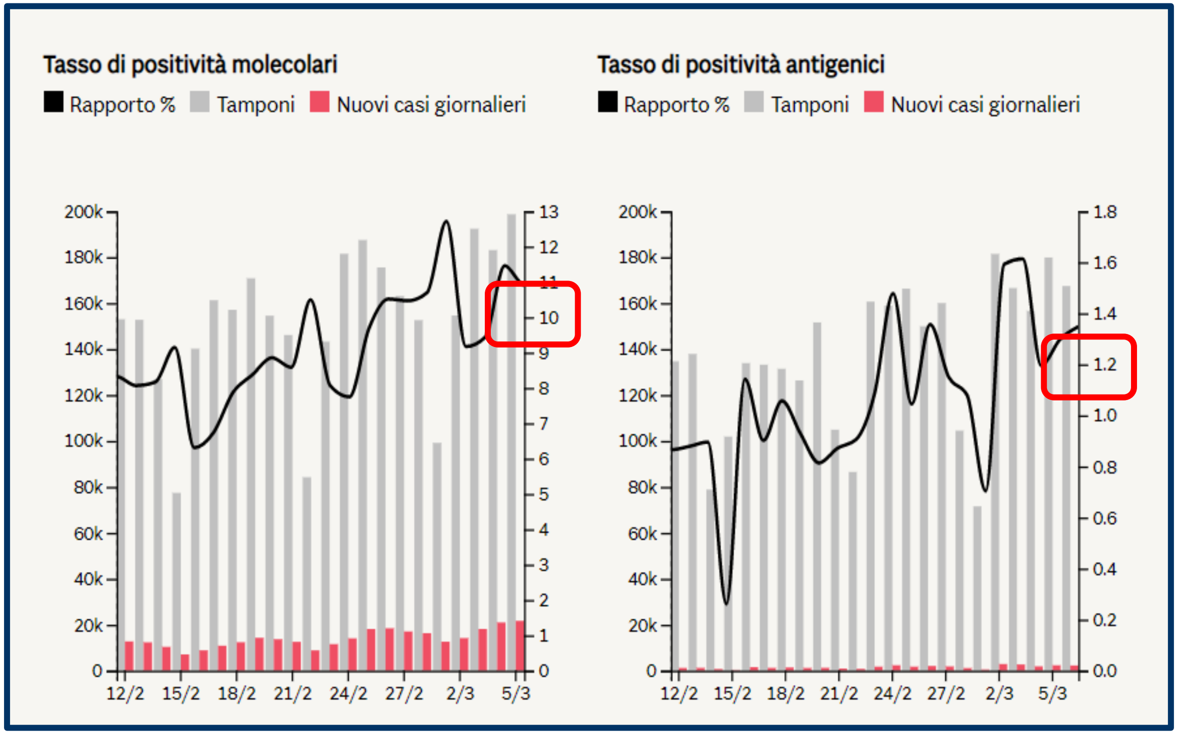
The result indicates that the probability that a randomly selected individual with a positive test is actually affected by SARS-CoV2 is ca. 68%, whereas the probability that a negative test is obtained despite the fact that the virus is present is 2.5%.

Note that the percentage of positive tests (with respect to all tests performed) obtained under the described conditions would be:

$$\frac{0.078 + 0.036}{0.078 + 0.036 + 0.022 + 0.864} \cong 11.4\ \%$$

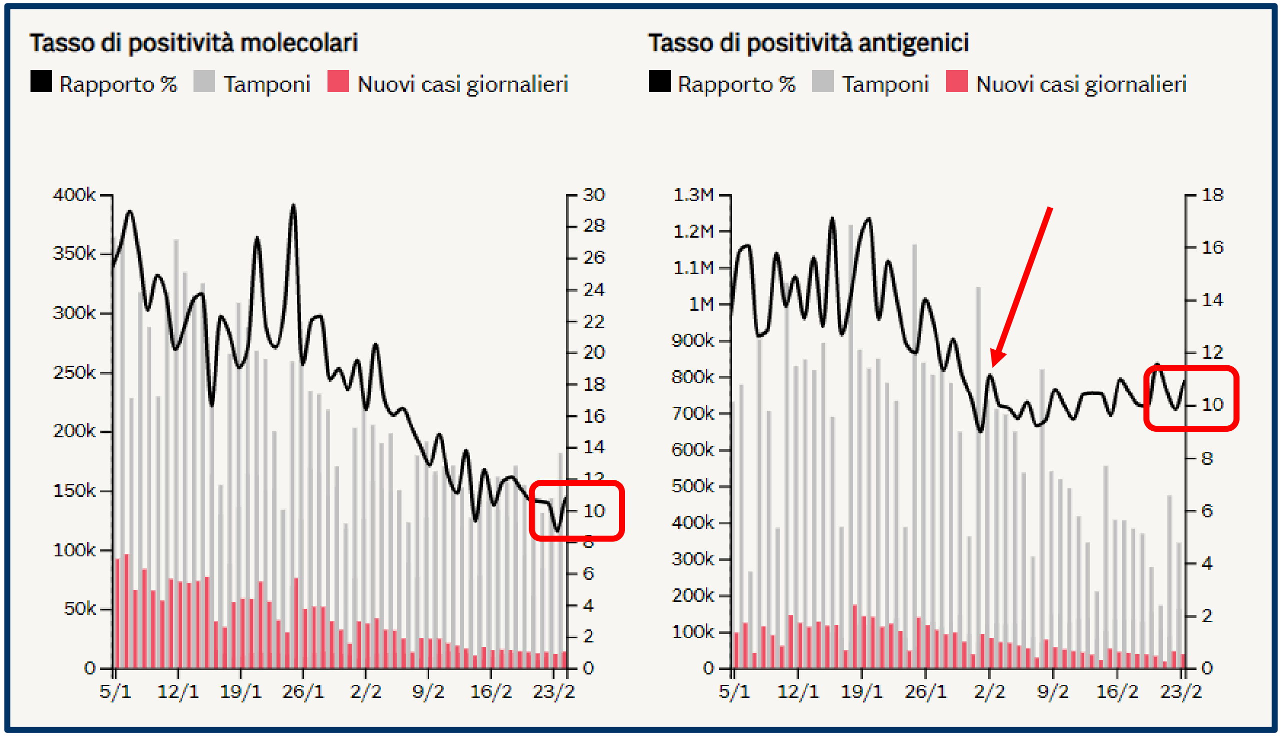
Interestingly, the percentage of positive results with respect to antigenic tests obtained in Italy between February 12th and March 6th 2021 was just slightly higher than 1%.

This result could be interpreted with the application of antigenic tests to a population with a much lower viral incidence.



Source: https://lab24.ilsole24ore.com/coronavirus/

The scenario has completely changed one year after, as shown by the percentage of positive tests obtained between January 5th and February 23rd, 2022:

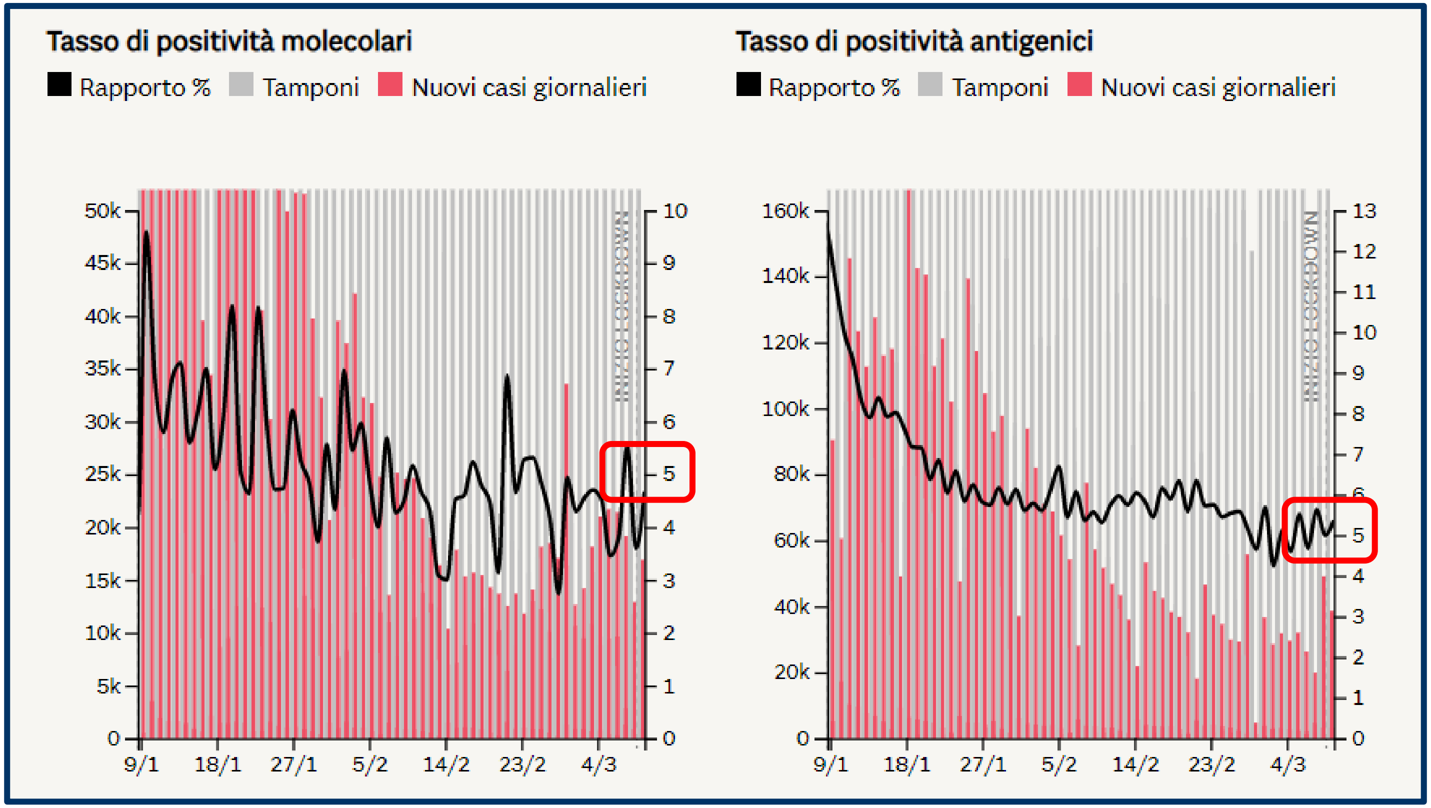


Source: https://lab24.ilsole24ore.com/coronavirus/

A possible interpretation of this result is that antigenic tests were used on a population in which the virus circulation was more similar to that occurring in the population to which molecular tests are applied.

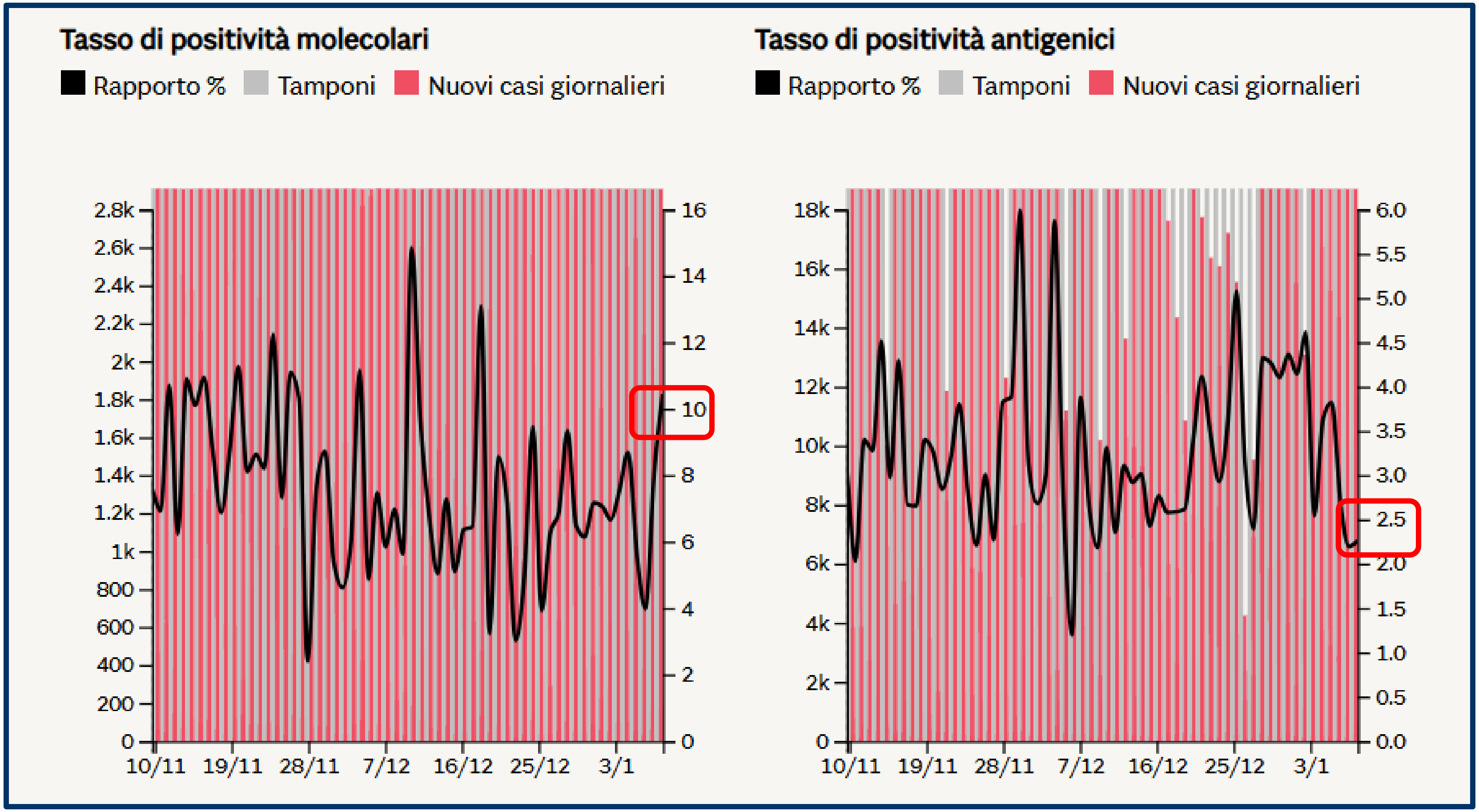
This change seems to have occurred since the beginning of February 2022.

A good consistency between the percentage of positive tests obtained using the two types of testing can be inferred also for the time interval January-March 2023 (note that, due to an internal mistake in the auto-scaling of the vertical axis, the actual number of daily cases is much lower than the one apparently deducible from red bars):



Source: https://lab24.ilsole24ore.com/coronavirus/

The last available data, dating back to the interval November 2024-January 2025, indicate a new significant difference in the apparent positivity rate between molecular and antigenic tests.



Source: https://lab24.ilsole24ore.com/coronavirus/

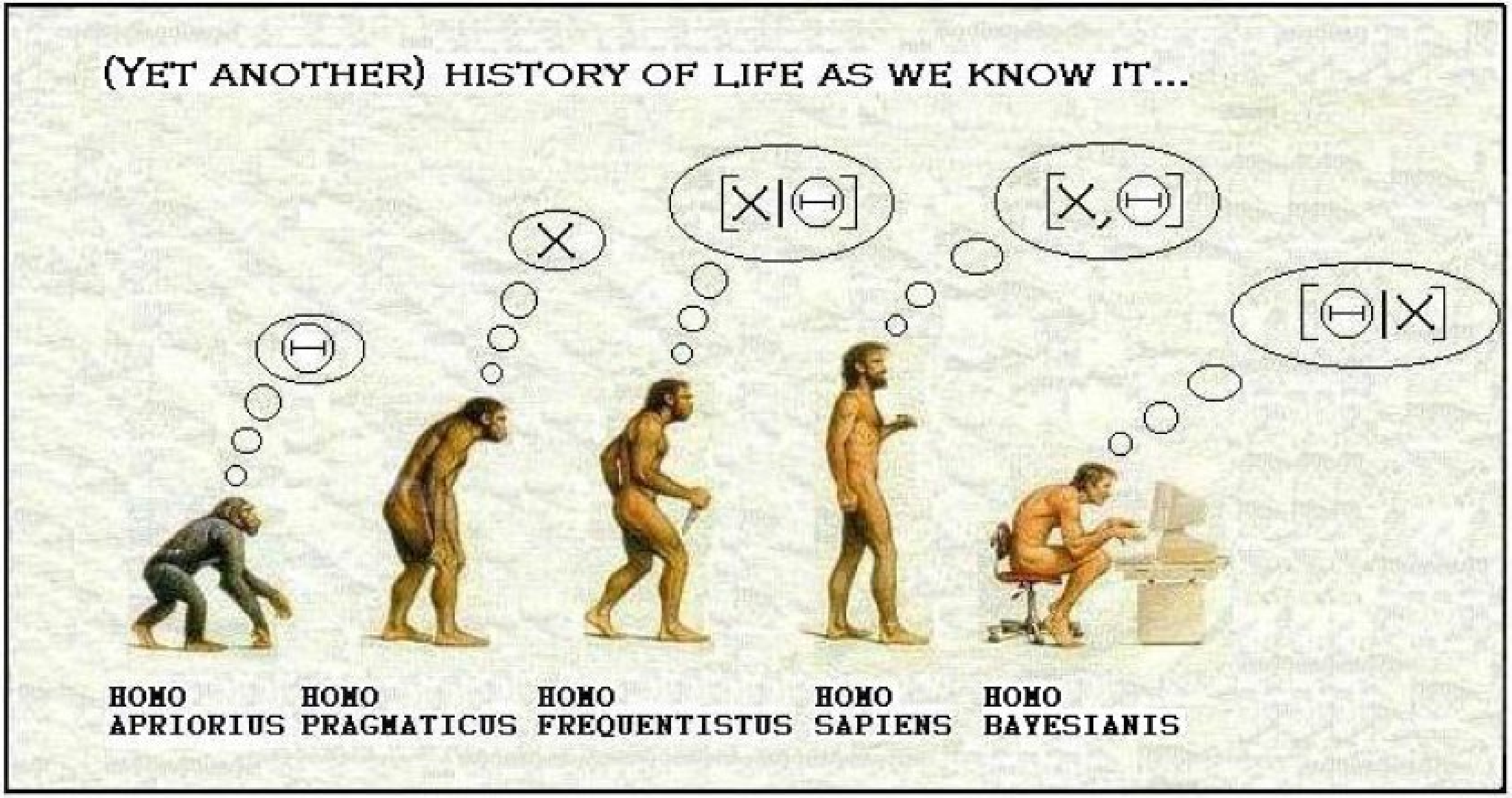
Bayesian approach – credible intervals

The Bayesian approach provides an alternative to confidence intervals, represented by credible intervals.

According to the definition, a 95% credible interval is the interval of values in which we are 95% certain that an unobserved parameter falls, based on sample data of size n.

The credible interval approach starts from sample data to arrive to the population parameter, whereas, according to statisticians not favorable to the confidence interval approach, i.e., to the frequentist approach, the latter appears to do the reverse.

An ironic representation of this controversy is given by the following cartoon, in which symbols [X|Θ] and [Θ|X] are intended to indicate, respectively, that X, the sample data, are inferred if Θ, the parameter value, is known, and viceversa.



German Molina & Enrique ter Horst, 2001

Calculation of a credible interval

Let us assume that we have data $(x_1,x_2,\ldots,x_n)$, that can be referred to as a vector **x**, obtained from a random variable X, and suppose that this variable is distributed according to a parameter θ.

In Bayesian analysis also this parameter is treated as a random variable, taking values in a space Θ and supposed to be distributed with a probability density function $h(\theta)$, which is known as prior distribution.

The joint probability density function, i.e., the probability that a data vector **x** will be obtained if a specific value of the parameter θ is true, $P(\mathbf{x}, \theta)$,

can be calculated from the product: $h(\theta)\, f(\mathbf{x}|\theta)$

where $f(\mathbf{x}|\theta)$ is called conditional probability density function of **x**.

Additionally, the (unconditional) probability density function of **x** is:

$$f(x) = \sum_{\theta \in \Theta} h(\theta) f(x \mid \theta),$$ if the parameter has a discrete distribution

$$f(\boldsymbol{x}) = \int_{\Theta} h(\theta) f(\boldsymbol{x} \mid \theta)\, d\theta,$$ if the parameter has a continuous distribution

Finally, starting from the Bayes's theorem, the posterior probability density function of θ can be determined by the following equation:

$$h(\theta|\mathbf{x}) = \frac{h(\theta)f(\mathbf{x}|\theta)}{f(\mathbf{x})}$$

In one of the most common Bayesian approaches, a credible interval (or Bayesian confidence set) at a 1-α level of confidence, indicated as C(**x**), is a subset of the parameter space that depends on the data vector **x** so that:

$P[\theta \in C(\mathbf{x})] = 1-\alpha$

In this definition θ is random and the interval can be obtained only if $h(\theta|\mathbf{x})$ is known.

An example of the calculation is given in the following for the case of Normal Distribution.

Credible intervals: the case of Normal Distribution

Suppose that **x** = (x_1,x_2,…,x_n) is a random sample of size n obtained from a normal distribution with unknown mean μ and known variance σ^2:

$$g(x \mid \mu) = \frac{1}{\sqrt{2\,\pi\sigma}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

Suppose, also, that μ has a normal distribution h(μ) with mean *a* and variance *b* (representing the prior distribution):

$$h(\mu) = \frac{1}{\sqrt{2\pi b}} \exp\left[-\frac{1}{2}\left(\frac{\mu-a}{b}\right)^2\right]$$

Under these conditions the following equation can be written:

$$f(\boldsymbol{x} \mid \mu) = g(x_1 \mid \mu) g(x_2 \mid \mu) \cdots g(x_n \mid \mu) = \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i - \mu}{\sigma}\right)^2\right]$$

$$= \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left[-\frac{1}{2\sigma^2}(w^2 - 2\mu y + n\mu^2)\right]$$

with:

$$w^2 = \sum_{i=1}^{n} x_i^2 \qquad y = \sum_{i=1}^{n} x_i$$

On the other hand:

$$h(\mu) = \frac{1}{\sqrt{2\pi}b}\exp\left[-\frac{1}{2}\left(\frac{\mu - a}{b}\right)^2\right] = \frac{1}{\sqrt{2\pi}b}\exp\left[-\frac{1}{2b^2}(\mu^2 - 2a\mu + a^2)\right]$$

Therefore:

$$h(\mu)f(\boldsymbol{x} \mid \mu) = \frac{1}{(2\pi)^{n/2}\sigma^n}\ \frac{1}{\sqrt{2\pi}b}\ \exp\left[-\frac{1}{2b^2}(\mu^2 - 2a\mu + a^2) - \frac{1}{2\sigma^2}(w^2 - 2\mu y + n\mu^2)\right]$$

If the pre-exponential term is condensed into a term C together with parts of the exponential function not including μ, the following equation is obtained:

$$h(\mu)f(\boldsymbol{x} \mid \mu) = \frac{1}{(2\pi)^{n/2}\sigma^n}\ \frac{1}{\sqrt{2\pi}b}\ \exp\left[-\frac{a^2}{2b^2}\right]\exp\left[-\frac{1}{2\sigma^2}w^2\right]\exp\left[-\frac{1}{2b^2}(\mu^2 - 2a\mu) - \frac{1}{2\sigma^2}(-2\mu y + n\mu^2)\right]$$

$$h(\mu)f(\boldsymbol{x} \mid \mu) = C\exp\left\{-\frac{1}{2}\left[\left(\frac{1}{b^2} + \frac{n}{\sigma^2}\right)\mu^2 - 2\left(\frac{a}{b^2} + \frac{y}{\sigma^2}\right)\mu\right]\right\}$$

The expression can be rearranged as follows:

$$h(\mu)f(\boldsymbol{x} \mid \mu) = C\exp\left\{-\frac{1}{2}\left[\left(\frac{1}{b^2}+\frac{n}{\sigma^2}\right)\mu^2 - 2\left(\frac{a}{b^2}+\frac{y}{\sigma^2}\right)\mu\right]\right\} =$$

$$= C\exp\left\{-\frac{1}{2}\left(\frac{1}{b^2}+\frac{n}{\sigma^2}\right)\left[\mu^2 - 2\left(\frac{a/b^2+y/\sigma^2}{1/b^2+n/\sigma^2}\right)\mu\right]\right\} =$$

$$= C\exp\left\{-\frac{1}{2}\left(\frac{1}{b^2}+\frac{n}{\sigma^2}\right)\left[-\left(\frac{a/b^2+y/\sigma^2}{1/b^2+n/\sigma^2}\right)^2 + \mu^2 - 2\left(\frac{a/b^2+y/\sigma^2}{1/b^2+n/\sigma^2}\right)\mu + \left(\frac{a/b^2+y/\sigma^2}{1/b^2+n/\sigma^2}\right)^2\right]\right\} =$$

$$= C\exp\left\{-\frac{1}{2}\left(\frac{1}{b^2}+\frac{n}{\sigma^2}\right)\left[-\left(\frac{a/b^2+y/\sigma^2}{1/b^2+n/\sigma^2}\right)^2 + \left(\mu - \frac{a/b^2+y/\sigma^2}{1/b^2+n/\sigma^2}\right)^2\right]\right\} =$$

$$= C\exp\left\{+\frac{1}{2}\left(\frac{1}{b^2}+\frac{n}{\sigma^2}\right)\left(\frac{a/b^2+y/\sigma^2}{1/b^2+n/\sigma^2}\right)^2 - \frac{1}{2}\left(\frac{1}{b^2}+\frac{n}{\sigma^2}\right)\left(\mu - \frac{a/b^2+y/\sigma^2}{1/b^2+n/\sigma^2}\right)^2\right\} =$$

$$= K\exp\left[-\frac{1}{2}\left(\frac{1}{b^2}+\frac{n}{\sigma^2}\right)\left(\mu - \frac{a/b^2+y/\sigma^2}{1/b^2+n/\sigma^2}\right)^2\right]$$

According to the Bayes' theorem, the h(μ)f(**x**|μ) product should be divided by f(x) to obtain h(μ|x):

$$h(\mu|\mathbf{x}) = \frac{h(\mu)f(\mathbf{x}|\mu)}{f(\mathbf{x})}$$

Considering that f(**x**) can be considered as a normalizing factor (i.e., a numerical value), the resulting expression, corresponding to the posterior distribution h(μ|**x**), is proportional to a normal distribution :

$$K\exp\left[-\frac{1}{2}\left(\frac{1}{b^2}+\frac{n}{\sigma^2}\right)\left(\mu-\frac{a/b^2+y/\sigma^2}{1/b^2+n/\sigma^2}\right)^2\right]$$

with the following mean and variance

mean

$$\frac{a/b^2+y/\sigma^2}{1/b^2+n/\sigma^2} = \frac{yb^2+a\sigma^2}{nb^2+\sigma^2}$$

variance

$$\frac{1}{1/b^2+n/\sigma^2} = \frac{\sigma^2 b^2}{\sigma^2+nb^2}$$

This distribution is said to be conjugate to the normal distribution with unknown mean and known variance.

Note that the variance of the posterior distribution is deterministic, since it depends on data only through the sample size n.

In the special case in which b = σ, the posterior distribution is a normal with mean (y+a)/(n+1) and variance $\sigma^2/(n+1)$.

A numerical comparison between frequentist and Bayesian confidence intervals

Initial conditions

The length of a certain machined part is supposed to be 10 centimeters but due to imperfections in the manufacturing process the actual length is normally distributed with mean μ and variance σ^2.

The variance is due to inherent factors in the process, which remain fairly stable over time. From historical data, it is known that σ = 0.3 cm. On the other hand, μ may be set by adjusting various parameters in the process and hence may change to an unknown value fairly frequently. Thus, suppose that we consider for μ a prior normal distribution with mean 10 and standard deviation 0.3; moreover, a sample of 100 parts has mean 10.2 cm.

Calculation of the frequentist 95% confidence interval

In this case:

$$\bar{X} \quad \frac{z_{(1-\alpha/2)}\,\sigma}{\sqrt{n}} \qquad \bar{X} \quad \frac{z_{(1-\alpha/2)}\,\sigma}{\sqrt{n}}$$

Since $z_{0.975} = 1.96$, the interval is:

$$10.2 - (1.96 \times 0.3)/10 \leq \mu \leq 10.2 + (1.96 \times 0.3)/10 \Rightarrow 10.1412 \leq \mu \leq 10.2588$$

Bayesian 95% confidence interval

In the specific example, the standard deviation of the prior distribution (b) and σ are the same, so the posterior distribution, that has to be used to calculate the confidence interval, has the following parameters:

Mean) $(y+a)/(n+1)$

y corresponds to the sum of values obtained experimentally: $10.2 \times 100 = 1020$
a is the mean of the prior distribution, equal to 10
n is the number of parts subjected to length measurements: 100.

The mean of posterior distribution is, then, equal to 1030/101 = 10.198

Standard deviation) $[\sigma^2/(n+1)]^{1/2}$

$\sigma = 0.3$, thus the standard deviation of the posterior distribution is equal to:

$[0.3^2/101]^{1/2} = 0.0299$

Since $z_{0.975} = 1.96$, the Bayesian 95% confidence interval is:

$$10.198 - (1.96 \times 0.0299) \le \mu \le 10.198 + (1.96 \times 0.0299) \Rightarrow 10.1394 \le \mu \le 10.2566$$

The comparison between the two types of intervals, before rounding off decimal figures:

Frequentist) $10.1412 \le \mu \le 10.2588$

Bayesian) $10.1394 \le \mu \le 10.2566$

shows that there can be cases in which the two intervals may become almost identical once the rounding of figures is performed, although this cannot be considered as a general rule.

Conceptual comparison between frequentist and Bayesian approaches

The difference in the approach to statistical inference followed by frequentists and Bayesians can be conceptualized as follows.

A frequentist believes that probabilities are only defined as the quantities obtained in the limit after the number of independent trials tends to infinity. For example, if an unbiased coin is tossed over numerous trials, the probability 1/2 represents the value to which the ratio between heads (or tails) and the total number of trials will converge as the number of trials tends to infinity.

A Bayesian interprets probabilities as the degree of belief in a hypothesis. Under this philosophy, it is perfectly valid to begin with a prior distribution, gather a few observations, and then make decisions based on the resulting posterior distribution from applying Bayes' theorem.

Bayesian credible intervals treat their bounds as fixed and the estimated parameter as a random variable, whereas frequentist confidence intervals treat their bounds as random variables and the parameter as a fixed value.

Apart from special cases, Bayesian credible intervals do not coincide with frequentist confidence intervals for two reasons:

- ✓ credible intervals incorporate problem-specific contextual information from the prior distribution, whereas confidence intervals are based only on the data;
- ✓ credible intervals and confidence intervals treat nuisance parameters in radically different ways.

Note that a nuisance parameter is any quantity whose value is not relevant to the goal of an analysis but is nevertheless required to determine some quantity of interest. For example, σ is a nuisance parameter when μ is the quantity or interest.

An example of application of Bayesian statistics to clinical trials

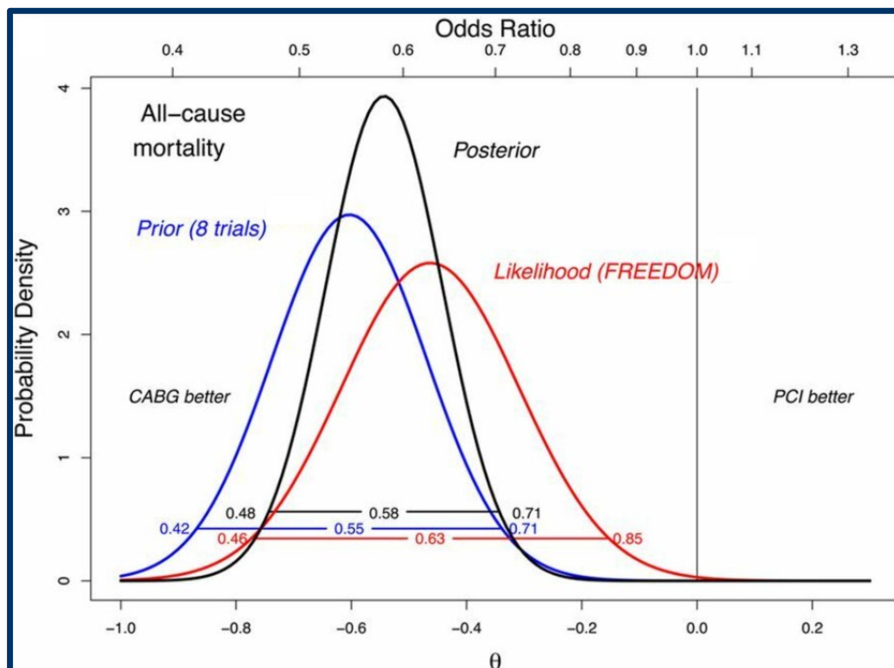
Bayesian statistics have been increasingly adopted in clinical trials in last years, showing some advantages over frequentist approaches.

An example of their application has been reported to compare coronary artery bypass graft (CABG) with percutaneous coronary intervention (PCI) as treatments for diabetic patients with multivessel coronary artery disease.

Specifically, the all-cause mortality was compared between CABG and PCI, expressed as odds ratio (OR) or as its logarithm (θ = log OR).

In this case a gaussian prior PDF (i.e., $h(\theta)$) was hypothesised starting from 8 previous clinical trials, showing a lower mortality for CABG compared to PCI.

The Likelihood gaussian PDF (i.e., $f(\mathbf{x}|\theta)$) was obtained from the results of the FREEDOM (Future Revascularization Evaluation in Patients with Diabetes Mellitus: Optimal Management of Multivessel Disease) trial.

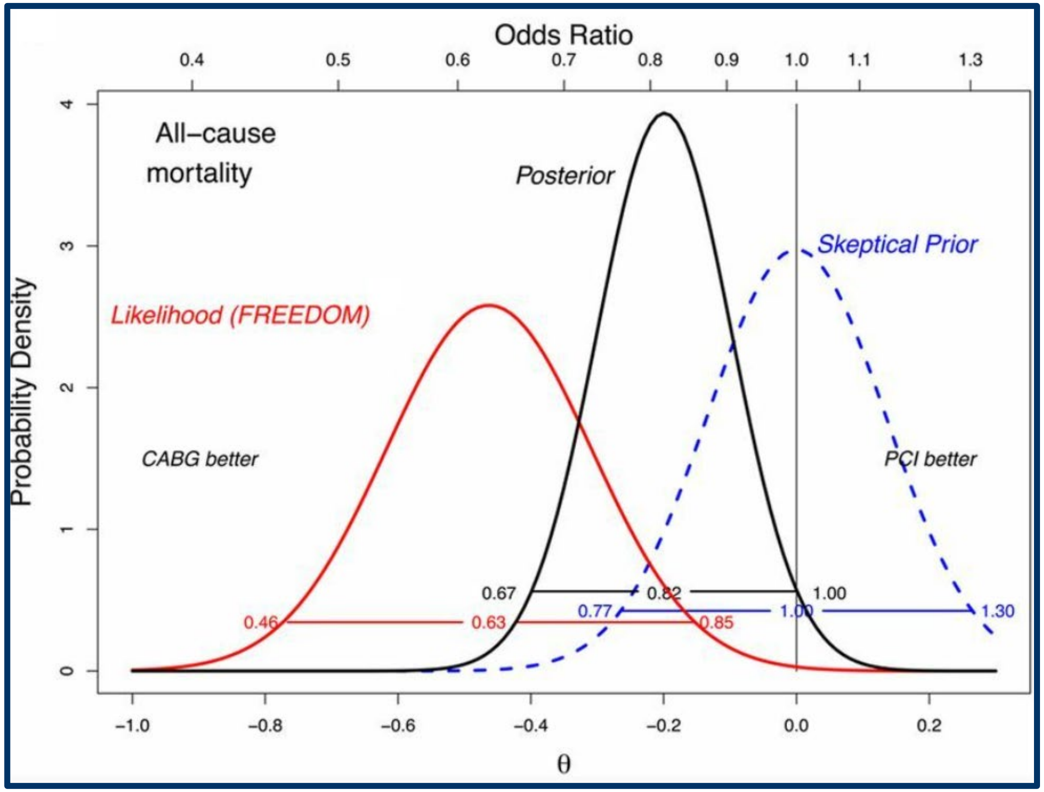


Starting from Bayes' theorem and using a computational approach, the posterior gaussian PDF (i.e., $h(\theta|\mathbf{x})$) was obtained.

As shown in the previous figure, the posterior inference contains a maximum (mode) at 0.58 with a 95% Bayesian credible interval (BCI) that extends from 0.48 to 0.71, thus confirming the advantage of CABG over PCI.

The Authors of the same paper also made a different calculation, based on a «skeptical» prior PDF, i.e., a PDF centered on a OR = 1 (indicating no mortality risk difference between CABG and PCI).

As shown in the figure, even in this case a posterior PDF centered on a OR lower than 1 and leading to a credible interval not exceeding OR = 1 was obtained.



Figures adapted from: J.A. Bittl and Y. He, *Circulation: Cardiovascular Quality and Outcomes*, 2017, 10:e003563.