Biostatistics

Th2

Inferential Statistics

Carina Silva
(*carina.silva@estesl.ipl.pt*)
Mestrado em Farmácia

ESCOLA SUPERIOR DE TECNOLOGIA DA SAÚDE DE LISBOA
INSTITUTO POLITÉCNICO DE LISBOA

Contents

Contents

Contents

FAQ's

- What is a statistical hypothesis?

FAQ's

- What is a statistical hypothesis?
- What is the null hypothesis?

FAQ's

- What is a statistical hypothesis?
- What is the null hypothesis?
- What does a p-value mean?

FAQ's

- What is a statistical hypothesis?
- What is the null hypothesis?
- What does a p-value mean?
- What does it mean to say that a finding is *statistically significant*?

FAQ's

- What is a statistical hypothesis?
- What is the null hypothesis?
- What does a p-value mean?
- What does it mean to say that a finding is *statistically significant*?
- What does it mean "It was assumed a 5% significance level"?

FAQ's

- What is a statistical hypothesis?
- What is the null hypothesis?
- What does a p-value mean?
- What does it mean to say that a finding is *statistically significant*?
- What does it mean "It was assumed a 5% significance level"?
- What is the difference between a Hypothesis Test and a Confidence Interval?

Why do we need Inferential Statistics?

Suppose, you want to know the average salary of Pharmacy Technicians in Portugal. Which of the following methods can be used to calculate it?

1. Meet every Pharmacy Technician in Portugal. Note down their salaries and then calculate the total average?
2. Or hand pick a number of professionals in several cities. Note down their salaries and use it to calculate the Portuguese average.

Inferential Statistics

In simple language, Inferential Statistics is used to draw inferences beyond the immediate data available.
With the help of inferential statistics, we can answer the following questions:

- Making inferences about the population from the sample.
- Concluding whether a sample is significantly different from the population.

Pre-Requisites

To begin with Inferential Statistics, one must have a good grasp over the following concepts:

- Probability.
- Basic knowledge of Probability Distributions.
- Descriptive Statistics.

Population

- Aim of scientific study: make general statements on a process at the level of the population.
- E.g. assess if the cholesterol level is on average different between males and females who are elder than 25.

Population

- Population in statistics is a theoretical concept
  - It is in continuous evolution/change.
  - Often interest to generalize conclusions to future subject - so population cannot be entirely observed at the present.
  - Can typically be considered to be infinite.
- Population has to be clearly defined!

Population

- Inclusion criteria are characteristics a subject/experimental unit must have to belong to the population, e.g.
  - age above 25
  - normal BMI
- Exclusion criteria characteristics that the subject/experimental unit is not allowed to have to belong to the population, e.g.
  - pregnancy in study on new type of drug
  - diabetes, history of hard drugs, low health status when the aim is to delineate a range of normal values of blood pressure in a population of healthy individuals.

Random Variables

- Variables (e.g. direct cholesterol) vary in the population from subject to subject!
- Variables are thus random because their value changes in the population.
- Crucial question: How precise are the conclusions on the population based on a group of subjects in a sample!
- We will thus observe differences from sample to sample.
- Variability of the data plays a crucial role!

Random Variable

A **Random Variable** is a function, which assigns unique numerical values to all possible outcomes of a random experiment under fixed conditions.

Convention

- Use capital letters for a study characteristic (e.g. direct cholesterol) to indicate that it changes in the population without thinking about the observed value of a particular subject.
- Variable X is a random variable and is the result of random sampling of the characteristic from the population. Random variable X is thus the unknown variable that represent a measurement that we plan to collect on a random subject, but that we have not collected yet.

Convention

- Typically a sequence of random variables $X_1, , X_n$ will be collected in the study (with n subjects or experimental units).
- The concept of random variables is necessary to reason on how the results and conclusions change from sample to sample.
- Random variables can be qualitative, quantitative, discrete, continuous,

Probability

- **Definition of probability**:
  Is a set function $P : \mathcal{A} \rightarrow [0,1]$ which satisfies the axioms
  1. $P[A] \geq 0$ for every $A \in \mathcal{A}$.
  2. $P[\Omega] = 1$.
  3. $A_1, A_2, ...$ sequence of mutually exclusive events in $\mathcal{A}$, then

$$P\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} P[A_i] = P[A_1] + P[A_2] + ...$$

- **Definition of Probability space**
  Is the triplet $(\Omega, \mathcal{A}, P[.])$

Properties of $P(.)$

1. $P[\phi] = 0$
2. If $A_1, ..., A_n$ mutually exclusive events in $\mathcal{A}$, then

$$P\left[\bigcup_{i=1}^{n} A_i\right] = \sum_{i=1}^{n} P[A_i].$$

3. If $A \in \mathcal{A}$ then $P[A^c] = 1 - P[A]$.
4. If A and $B \in \mathcal{A}$, then $P[A \cup B] = P[A] + P[B] - P[A \cap B]$.
5. If A and $B \in \mathcal{A}$ and $A \subset B$, then $P[A] \leq P[B]$.
6. If $A_1, ..., A_n \in \mathcal{A}$, then

$$P\left[\bigcup_{i=1}^{n} A_i\right] \leq \sum_{i=1}^{n} P[A_i].$$

Example

A doctor knows that $P(\text{bacterial infection}) = 0.7$ and $P(\text{viral infection}) = 0.4$. What is $P(B \cap V)$ if $P(B \cup V) = 1$?

A: $P(B \cup V) = P(B) + P(V) - P(B \cap V) \iff P(B \cap V) = 0.7 + 0.4 - 1 = 0.1$

Conditional Probability

Let $(\Omega, \mathcal{A}, P[.])$ be a probability space

Definition: Conditional Probability

Let A and B two events in $\mathcal{A}$, $P[B] > 0$. Conditional probability of A given B

$$P[A|B] = \frac{P[A \cap B]}{P[B]}.$$

Conditional Probability

Example. In trying to determine the sex of a child, a pregnancy test called "starch gel electrophoresis"is used. This test may reveal the presence of a protein zone called the pregnancy zone. This zone is present in 43% of all pregnant women. Furthermore, it is known that 51% of all children born are male. 17% of all children born are male and the pregnancy zone is present.
Let $A1$ denote the event that the pregnancy zone is present, and $A2$ that the child is male. We know that, for a randomly selected pregnant woman:
$P(A1) = .43$, $P(A2) = .51$, $P(A1 \cap A2) = .17$

Q: What is the probability that the child is male given that the pregnancy zone is present?

Conditional Probability

A: The sample space no includes all pregnant women, but rather consists only of the 43% women with the pregnancy zone present. Of these, $.17/.43 = .395$ have male children. Thus

$$P(\text{male}|\text{zone present}) = P(A_2|A_1) = 0.395$$

Independence

Definition: Independent events

A and B two events in $\mathcal{A}$ are independent if one of the following is satisfied

(i) $P[A \cap B] = P[A]P[B]$

(ii) $P[A|B] = P[A]$ if $P[B] > 0$

(iii) $P[B|A] = P[B]$ if $P[A] > 0$

Remark: Independent events can only be mutually exclusive if the probability of at least one is zero.

Independence

Example. The probability that a couple heterozygous for eye color will parent a brown-eyed child is 3/4 for each child. Genetic studies indicate that the eye color of one child is independent of that of the other. Thus, if the couple has two children, then the probability that both will be brown-eyed is

$P(\text{first brown and second brown}) =$
$= \mathrm{P}(\text{first brown})\mathrm{P}(\text{second brown}) =$

Let $B_1, ..., B_n$ mutually disjoint (or mutually exclusive) in $\mathcal{A}$, $\Omega = \bigcup_{i=1}^{n} B_i$ and $P[B_i] > 0, \forall i$

Theorem - Theorem of total probabilities

For every $A \in \mathcal{A}$

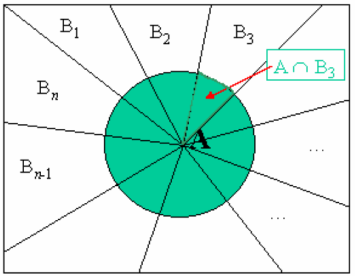
$$P[A] = \sum_{i=1}^{n} P[A|B_i]P[B_i].$$

Theorem - Bayes' Theorem

For every $A \in \mathcal{A}$ such that $P[A] > 0$

$$P[B_k|A] = \frac{P[A|B_k]P[B_k]}{\sum_{i=1}^{n} P[A|B_i]P[B_i]}.$$

Both theorems remain true if $n = \infty$.



Baye's Theorem

Example. The blood type distribution in the United States is type A, 41%; type B, 9%; type AB, 4%; and type O, 46%. It is estimated that during World War II, 4% of inductees with type O blood were typed as having type A; 88% of those with type A were correctly typed; 4% with type B blood were typed as A; and 10% with type AB were typed as A. A soldier was wounded and brought to surgery. He was typed as having type A blood. What is the probability that this is his true blood type?

Example (Cont.)

Let

$$\begin{aligned} A &: \text{ he has type A blood} \\ B &: \text{ he has type B blood} \\ AB &: \text{ he has type AB blood} \\ O &: \text{ he has type O blood} \\ TA &: \text{ he is typed as type A} \end{aligned}$$

We want to find $P(A|TA)$. We are given:

$$\begin{gathered} P(A) = .41; P(TA|A) = .88 \\ \mathrm{P(B)} = .09;\ \mathrm{P(TA|B)} = .04 \\ \mathrm{P(AB)} = .04;\ \mathrm{P(TA|AB)} = .10 \\ \mathrm{P(O)} = .46;\ \mathrm{P(TA|O)} = .04 \end{gathered}$$

Example (Cont.)

To get $P(A|TA)$, we can apply the definition of conditional probability:

$$P(A|TA) = \frac{P(A \cap TA)}{P(TA)}$$

By the multiplication rule:
$P(A \cap TA) = P(TA|A)P(A) = (.88)(.41) = .36$

Event TA can be partitioned into four mutually exclusive events, that is

$TA = (A \cap TA) \cup (B \cap TA) \cup (AB \cap TA) \cup (O \cap TA)$
By probability's properties:

$P(TA) = P(A \cap TA) + P(B \cap TA) + P(AB \cap TA) + P(O \cap TA)$

Example (Cont.)

Applying the multiplication rule to each of the terms on the right hand side:

$$P(TA) = P(TA|A)P(A) + P(TA|B)P(B)$$
$$+P(TA|AB)P(AB) + P(TA|O)P(O) =$$
$$= (.88)(.41) + (.04)(.09) + (.10)(.04) + (.04)(.46) = .39$$

Thus

$$P(A|TA) = \frac{P(A \cap TA)}{P(TA)} = \frac{.36}{.39} = .92$$

Exercise 1

Suppose that in a group of 502 people they were classified by blood group and sex as follows

	sex		
blood goup	Male	Female	Total
0	113	113	226
A	103	103	206
B	25	25	50
AB	10	10	20
Total	251	251	502

A person is picked at random from this group.

1. What is the probability that she is female?
2. Given that she is female, what is the probability that her blood group is A or AB?

Exercise 2

Suppose that it is known that a fraction .001 of the people in a town have tuberculosis (TB). A TB test is given with the properties: If the person has TB the test will indicate it with probability .999. If does not have TB then there is a probability .002 that the test erroneously indicates that he does. For one person selected at random the test shows that he has TB. What is the probability that he really does?

Normal distribution

- Normal (Gaussian) $N(\mu, \sigma^2)$. $(\mu \in \Re, \sigma^2 > 0)$

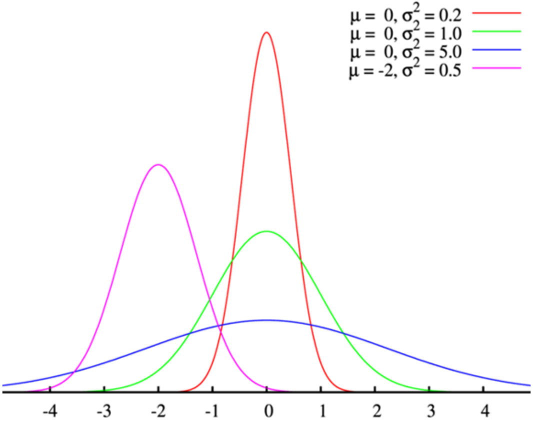
$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad -\infty < x < +\infty.$$

$$E[X] = \mu, \quad var[X] = \sigma^2.$$

μ is a location parameter and σ^2 is a scale parameter.

Note constants:
$\pi = 3.14159$
$e = 2.71828$

Normal distribution

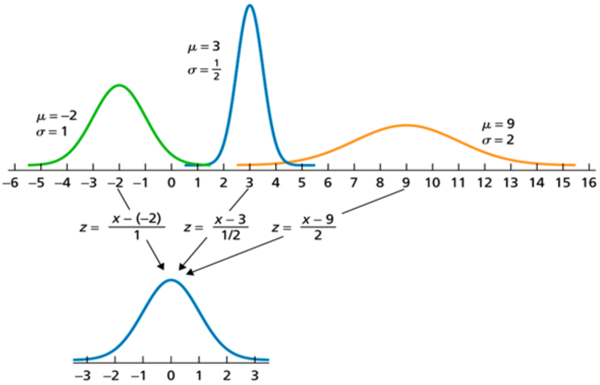


Standard Normal Distribution

All normal distributions can be converted into the standard normal curve by subtracting the mean and dividing by the standard deviation:

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

then Z has a Normal(0, 1) distribution. Such Z is called a standard normal random variable. The scale of Z has no units and it is called the standardized scale. It shows how many standard deviations from the mean the X variable is.


μ = 3
σ = 1/2
μ = −2
σ = 1
μ = 9
σ = 2
−6 −5 −4 −3 −2 −1 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
z = (x − (−2))/1
z = (x − 3)/(1/2)
z = (x − 9)/2
−3 −2 −1 0 1 2 3

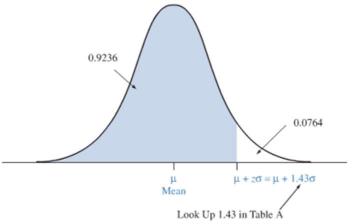


TABLE 6.3: Part of Table A for Normal Cumulative (Left-Tail) Probabilities

The top of the table gives the second digit for z. The table entry is the probability falling below $\mu + z\sigma$, for instance, 0.9236 below $\mu + 1.43\sigma$ for $z = 1.43$.

Second Decimal Place of z

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
...										
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9139	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9278	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441

- Table gives the areas under a standard normal density.
- It gives the area under the standard normal density curve to the left of a specified value $z \Longrightarrow P(Z < z)$.
- For example, to find the area under the standard normal density to the left of 1.43 we find the entry in the row labeled 1.4 under the column 0.03. It is 0.9236.

Example: The brain weights of adult Swedish males are *approximately* normally distributed with mean $\mu = 1{,}400$g and standard deviation $\sigma = 100$g.

Let X denote the brain weight of a randomly chosen person from this population. Calculate $P(X \leq 1,500)$.

Quantile

  - The $p \times 100^{th}$ quantile of a distribution is the value of x that gives $Pr(X \leq x) = p$.
- For example, the 75th quantile of the standard normal distribution is the value of z that has 75% of the area below it or Pr(Z$\leq z$) $= 0.75$.

  Example: What is the 10th percentile for the distribution of brain weights?

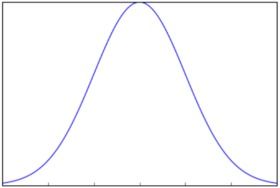
Exercise

The distribution of the expression values of the ALL patients on the Zyxin gene are distributed according to N(1.6; 0.42).
a) Compute the probability that the expression values are smaller than 1.2?
b) What is the probability that the expression values are between 1.2and 2.0?
c) What is the 15th quantile of that distribution. What does it mean?

Sampling Distribution and Central Limit Theorem

Suppose, you note down the salary of any 100 random Pharmacy Technicians in Portugal, calculate the mean and repeat the procedure for say like 200 times (arbitrarily).

When you plot a frequency graph of these 200 means, you are likely to get a curve similar to the one below.



This is called Sampling Distribution or the graph obtained by plotting sample means.

Sampling Distribution and Central Limit Theorem

A Sampling Distribution is a probability distribution of a statistic obtained through a large number of samples drawn from a specific population.

- The shape of the Sampling Distribution does not reveal anything about the shape of the population.
- Sampling Distribution helps to estimate the population statistic.

Central Limit Theorem

It states that when plotting a sampling distribution of means, the mean of sample means will be equal to the population mean. And the sampling distribution will approach a normal distribution with variance equal to $\frac{\sigma}{\sqrt{n}}$ where σ is the standard deviation of population and n is the sample size.

Central Limit Theorem

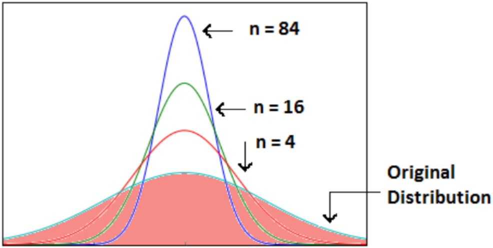
Points to note:

- Central Limit Theorem holds true irrespective of the type of distribution of the population.
- Now, we have a way to estimate the population mean by just making repeated observations of samples of a fixed size.
- Greater the sample size, lower the standard error and greater the accuracy in determining the population mean from the sample mean.

Central Limit Theorem

- This means No matter the shape of the population distribution, be it bi-modal, right skewed etc. The shape of the Sampling Distribution will remain the same (remember the normal curve- bell shaped). This gives us a mathematical advantage to estimate the population statistic no matter the shape of the population.
- The number of samples have to be sufficient to satisfactorily achieve a normal curve distribution. Also, care has to be taken to keep the sample size fixed since any change in sample size will change the shape of the sampling distribution and it will no longer be bell shaped.
- As we increase the sample size, the sampling distribution squeezes from both sides giving us a better estimate of the population statistic since it lies somewhere in the middle of the sampling distribution (generally).

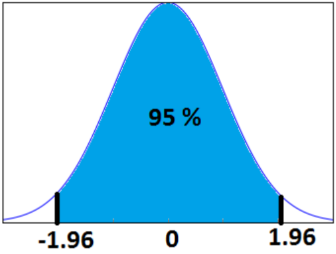
TLC



Now, since we have collected the samples and plotted their means, it is important to know where the population mean lies with respect to a particular sample mean and how confident can we be about it. This brings us to our next topic: **Confidence Interval**.

Confidence Interval

The confidence interval is a type of interval estimate from the sampling distribution which gives a range of values in which the population statistic (parameter) may lie.



Confidence Interval

Interesting points to note about Confidence Intervals:

- Confidence Intervals can be built with difference degrees of confidence suitable to a user's needs like 70 %, 90% etc.
- Greater the sample size, smaller the Confidence Interval, i.e more accurate determination of population mean from the sample means.
- There are different confidence intervals for different sample means. For example, a sample mean of 40 will have a difference confidence interval from a sample mean of 45.
- By 95% Confidence Interval, we do not mean that The probability of a population mean to lie in an interval is 95%. Instead, 95% C.I means that 95% of the Interval estimates will contain the population statistic.

Confidence Interval

In general, we will form a (1 - α)100% confidence interval for the parameter, where 1 - α represents the proportion of intervals that would contain the unknown parameter if this procedure were repeated on many different samples. The width of a (1 - α)100% confidence interval depends on:

- The confidence level (1 - α). As (1 - α) increases, so does the width of the interval. If we want to increase the confidence we have that the interval contains the parameter, we must increase the width of the interval.
- The sample size(s). The larger the sample size, the smaller the standard error of the estimator, and thus the smaller the interval.
- The standard deviations of the underlying distributions. If the standard deviations are large, then the standard error of the estimator will also be large.

Confidence Interval for the Mean Value

- $n \geq 30$, σ known, $X \sim \forall$

$$\bar{x} \pm z_{1-\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}}$$

Confidence Interval for the Mean Value

- $n \geq 30$, σ unknown, $X \sim \forall$

$$\bar{x} \pm z_{1-\frac{\alpha}{2}} * \frac{s}{\sqrt{n}}$$

Confidence Interval for the Mean Value

- $n < 30$, σ unknown, $X \sim N(\mu, \sigma)$

$$\bar{x} \pm t_{1-\frac{\alpha}{2};n-1} * \frac{s}{\sqrt{n}}$$

Exercise

The body temperature data set DataP4.txt (Link) contains the body temperature (Fahrenheit) and the gender of 130 volunteers, 65 men and 65 women. Is the mean body temperature of human adults really 98.6F (37C)? (Ignore differences due to gender).

- A **hypothesis test** is a procedure for determining if an assertion about a characteristic of a population is reasonable.

Example:
Someone says that the average weight of the mice used in experiments in ESTeSL's laboratories is 20g.
How would you decide whether this statement is true?

- find all the mice in the laboratories of ESTeSL and weight them all, or
- find out the average weight of mice at a small number of randomly chosen laboratories and compare the average weight to 20g.

Suppose your sample average was 19g. Is this 1g difference a significative result, or is the original assertion incorrect?

Terminology

- The null hypothesis, H_0, is usually a hypothesis of agreement with conditions presumed to be true. A null hypothesis is either rejected or not rejected.
  - $H_0 : \mu = 20g$
- The alternative hypothesis, H_1, represents the statement that the researcher wants to prove. (It determines the rejection area).
  - $H_1 : \mu > 20g$
  - $H_1 : \mu < 20g$
  - $H_1 : \mu \neq 20g$

Terminology

- When H_0 is not rejected we say that ”the data do not give enough evidence to reject the null hypothesis”.

Terminology

- When H_0 is not rejected we say that ”the data do not give enough evidence to reject the null hypothesis”.
- When the H_0 is rejected we say that ”the data at hand are not compatible with the null hypothesis stated”, but are supportive of some other hypothesis which is the one which is established as the alternative hypothesis.

Terminology

- When H_0 is not rejected we say that "the data do not give enough evidence to reject the null hypothesis".
- When the H_0 is rejected we say that "the data at hand are not compatible with the null hypothesis stated", but are supportive of some other hypothesis which is the one which is established as the alternative hypothesis.
- The significance level alpha (α)
  - Type I error $= P[Rej H_0 | H_0 is True]$
  - Alpha $= 0.05$: the probability of incorrectly rejecting the null hypothesis when it is actually true is 5%.
  - If you need more protection from this error, then choose a lower value of alpha .

Error types

It is important to bear in mind the following:

The fact that a hypothesis is not rejected does not mean that it is true. We can only say that the hypothesis is supported by the available data.

When we formulate a hypothesis testing problem there are two types of errors we can commit.

- When we reject a true null hypothesis, we say that we commit a type I error. The probability of committing a type I error is represented by α.

Error types

It is important to bear in mind the following:

The fact that a hypothesis is not rejected does not mean that it is true. We can only say that the hypothesis is supported by the available data.

When we formulate a hypothesis testing problem there are two types of errors we can commit.

- When we reject a true null hypothesis, we say that we commit a type I error. The probability of committing a type I error is represented by α.
- When we fail to reject a false null hypothesis then we say that a type II error is committed. The probability of committing a type II error is represented by β.

Error types and Power of test

	Condition of null hypothesis	
Possible action	True	False
Fail to reject H_0	Correct $(1-\alpha)$	Type II error β
Reject H_0	Type I error α	Correct $(1-\beta)$

- Type I Error (alpha): calling genes as differentially expressed when they are NOT
- Type II Error (beta): NOT calling genes as differentially expressed when they ARE
- Power of a test: $1 - \beta$: The odds of confirming our theory correctly

Error type I vs Error type II

What we would expect is to have a lower alpha (α) and higher power (1-β) But:

- the lower the alpha , the lower the power; the higher the alpha, the higher the power. lower the alpha, the less likely it is that you will make a Type I Error (i.e., reject the null when it's true)
- the lower the alpha, the more "rigorous"the test.

Steps in hypothesis testing

- Determine the null and alternative hypothesis, using mathematical expressions if applicable.

Steps in hypothesis testing

- Determine the null and alternative hypothesis, using mathematical expressions if applicable.
- Select a significance level (alpha).

Steps in hypothesis testing

- Determine the null and alternative hypothesis, using mathematical expressions if applicable.
- Select a significance level (alpha).
- Take a random sample from the population of interest.

Steps in hypothesis testing

- Determine the null and alternative hypothesis, using mathematical expressions if applicable.
- Select a significance level (alpha).
- Take a random sample from the population of interest.
- Calculate a test statistic from the sample that provides information about the null hypothesis.

Steps in hypothesis testing

- Determine the null and alternative hypothesis, using mathematical expressions if applicable.
- Select a significance level (alpha).
- Take a random sample from the population of interest.
- Calculate a test statistic from the sample that provides information about the null hypothesis.
- Decision (by classical definition or with p-value).

Steps in hypothesis testing

- Determine the null and alternative hypothesis, using mathematical expressions if applicable.
- Select a significance level (alpha).
- Take a random sample from the population of interest.
- Calculate a test statistic from the sample that provides information about the null hypothesis.
- Decision (by classical definition or with p-value).
- Conclusion.

Decision

The decision as to whether H_0 is rejected or not rejected is made on the basis of data using the result of a *test statistic*, say $T(X_1, \ldots, X_n)$, or for short, T.
A good test statistic should be such that the probability of committing a type II error is as small as possible.

How to proceed once a test statistic T is chosen?

Rejection Regions

The set of possible values the test statistic T can take is divided into two regions

- The acceptance region $\mathfrak{A}$; observed values of the test statistic T, falling in this region lead to non-rejection of the null hypothesis.
- The rejection region R; observed values of the test statistic T falling in this region lead to the rejection of the null hypothesis.

Rejection Regions

- The alternative hypothesis tells the tale (1-tailed vs 2-tailed tests)

This is accomplished either with the knowledge of the exact sampling distribution of the test statistic, under the null hypothesis, or with the help of asymptotic theory.

The rejection regions relatively to a significance level α, are usually of one of the types,

$$\mathcal{R}_\alpha = \{t : t > t_\alpha\}, \qquad \mathcal{R}_\alpha = \{t : t < t_\alpha\},$$

$$\mathcal{R}_\alpha = \{t : t < t_1 \quad \text{or} \quad t > t_2\},$$

Rejection Regions

Usually, for reasons of symmetry, we choose equal probabilities on the tails and hence, we choose t_1 and t_2 such that

$$P(T < t_1 | H_0) = P(T > t_2 | H_0) = \frac{\alpha}{2}.$$

In a hypothesis testing problem as it was stated, the significance level α, which is a measure of the uncertainty associated with our inference, is fixed beforehand. The usual values chosen for α are 0.10, 0.05, 0.01. The size of the rejection region depends on this value. If for a certain test statistic with rejection region α, call it $\mathcal{R}_\alpha$ we have the relation:

$$\mathcal{R}_{\alpha_1} \supset \mathcal{R}_{\alpha_2} \Leftrightarrow \alpha_1 > \alpha_2.$$

P-value

- A p-value is a measure of how much evidence we have against the null hypothesis.

P-value

- A p-value is a measure of how much evidence we have against the null hypothesis.
- The p-value measures consistency by calculating the probability of observing the results from your sample of data or a sample with results more extreme, assuming the null hypothesis is true.

P-value

- A large p-value should not automatically be construed as evidence in support of the null hypothesis. Perhaps the failure to reject the null hypothesis was caused by an inadequate sample size.

P-value

- A large p-value should not automatically be construed as evidence in support of the null hypothesis. Perhaps the failure to reject the null hypothesis was caused by an inadequate sample size.
- When you see a large p-value in a research study, you should also look for one of two things:

P-value

- A large p-value should not automatically be construed as evidence in support of the null hypothesis. Perhaps the failure to reject the null hypothesis was caused by an inadequate sample size.
- When you see a large p-value in a research study, you should also look for one of two things:
  1. a power calculation that confirms that the sample size in that study was adequate for detecting a biological relevant difference; and/or

P-value

- A large p-value should not automatically be construed as evidence in support of the null hypothesis. Perhaps the failure to reject the null hypothesis was caused by an inadequate sample size.
- When you see a large p-value in a research study, you should also look for one of two things:
  1. a power calculation that confirms that the sample size in that study was adequate for detecting a biological relevant difference; and/or
  2. a confidence interval that lies entirely within the range of biological indifference.

- You should also be cautious about a small p-value, but for different reasons. In some situations, the sample size is so large that even differences that are trivial from a biological perspective can still achieve statistical significance.

Notes:
1.You should not interpret the p-value as the probability that the null hypothesis is true. Such an interpretation is problematic because a hypothesis is not a random event that can have a probability

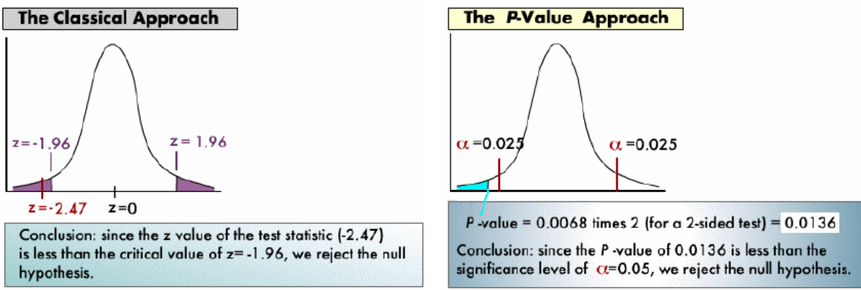
2.Bayesian statistics provides an alternative framework that allows you to assign probabilities to hypotheses and to modify these probabilities on the basis of the data that you collect.

P-value

If for some specific data we observe t_{obs} as the value for the test statistic, then the p-value is the probability of observing a value for the test statistic as "extreme" as t_{obs}. For each of the type of the rejection regions considered above we have:

$$
\begin{aligned}
p\text{-value} &= P(T > t_{obs} | H_0 \text{true}), \\
p\text{-value} &= P(T < t_{obs} | H_0 \text{true}), \\
p\text{-value} &= 2\ P(T > \text{abs}(t_{obs}) | H_0 \text{true})
\end{aligned}
$$

where $\text{abs}(t_{obs})$ means the absolute value of t_{obs}.


The Classical Approach
z=-1.96
z= 1.96
z=-2.47
z=0
Conclusion: since the z value of the test statistic (-2.47) is less than the critical value of z=-1.96, we reject the null hypothesis.
The P-Value Approach
α =0.025
α =0.025
P -value = 0.0068 times 2 (for a 2-sided test) = 0.0136
Conclusion: since the P -value of 0.0136 is less than the significance level of α=0.05, we reject the null hypothesis.

Confidence intervals and p-values

- The two statistical concepts are complementary.

Confidence intervals and p-values

- The two statistical concepts are complementary.
- Taken in isolation, p-values provide a measure of the statistical plausibility of a result.

Confidence intervals and p-values

- The two statistical concepts are complementary.
- Taken in isolation, p-values provide a measure of the statistical plausibility of a result.
- With a defined level of significance, p-values allow a decision about the rejection or maintenance of a previously formulated null hypothesis in confirmatory studies.

Confidence intervals and p-values

- The two statistical concepts are complementary.
- Taken in isolation, p-values provide a measure of the statistical plausibility of a result.
- With a defined level of significance, p-values allow a decision about the rejection or maintenance of a previously formulated null hypothesis in confirmatory studies.
- Confidence intervals provide an adequately plausible range for the true value related to the measurement of the point estimate.

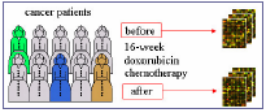
Confidence intervals and p-values

- The two statistical concepts are complementary.
- Taken in isolation, p-values provide a measure of the statistical plausibility of a result.
- With a defined level of significance, p-values allow a decision about the rejection or maintenance of a previously formulated null hypothesis in confirmatory studies.
- Confidence intervals provide an adequately plausible range for the true value related to the measurement of the point estimate.
- Confidence intervals gives an estimate of the precision with which a statistic estimates a population value.

Confidence intervals and p-values

- The two statistical concepts are complementary.
- Taken in isolation, p-values provide a measure of the statistical plausibility of a result.
- With a defined level of significance, p-values allow a decision about the rejection or maintenance of a previously formulated null hypothesis in confirmatory studies.
- Confidence intervals provide an adequately plausible range for the true value related to the measurement of the point estimate.
- Confidence intervals gives an estimate of the precision with which a statistic estimates a population value.
- If the alternative hypothesis is unilateral it is not possible compare the results with confidence intervals.

Types of hypothesis tests



Dependent samples

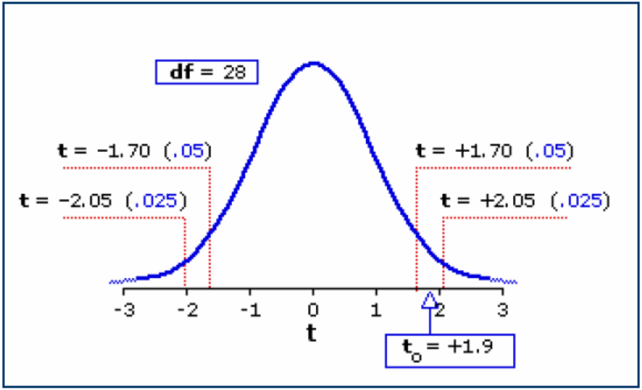
Independent samples

Comparison	Two Groups		More than two Grou
Hypothesis Testing	**Paired data**	**Unpaired data**	**Complex data**
Parametric (variance equal)	One sample t-test	Two-sample t-test	One-Way Analysis of Varian (ANOVA)
Parametric (variance not equal)	Welch t-test		Welch ANOVA
Non-Parametric	Wilcoxon Signed-Rank Test	Wilcoxon Rank-Sum Test (Mann-Whitney U Test)	Kruskal-Wallis Test

One-sample t-test

- The One-Sample t-test compares the mean score of a sample to a known value. Usually, the known value is a population mean.
- Assumptions:
  - n?
  - the variable is normally distributed.
  - $H_0 : \mu = \mu_0$
  - Statistical test: $Z_0 = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ or $T_0 = \frac{\bar{X}-\mu_0}{S/\sqrt{n-1}} \sim t_{n-1}$
  - Reject H_0 if $|z_0| > z_{1-\alpha/2}$ or Reject H_0 if $|t_0| > t_{1-\frac{\alpha}{2};n-1}$ (2-tailed)
  - p-value=$2(|Z| > z_0)$, $X \sim N(0,1)$ (2-tailed)
    p-value=$2(|T| > t_0)$, $T \sim t_{n-1}$ (2-tailed)

- n=29, $X \sim$ Normal distributed, $H_1 : \mu \neq 0$
- t_0=1,9
- using $\alpha = 0.05$ or 0.1



PERCENTAGE POINTS OF THE T DISTRIBUTION

Tail Probabilities									
One Tail		0.10	0.05	0.025	0.01	0.005	0.001	0.0005	
Two Tails		0.20	0.10	0.05	0.02	0.01	0.002	0.001	
D	1	3.078	6.314	12.71	31.82	63.66	318.3	637	1
E	2	1.886	2.920	4.303	6.965	9.925	22.330	31.6	2
G	3	1.638	2.353	3.182	4.541	5.841	10.210	12.92	3
R	4	1.533	2.132	2.776	3.747	4.604	7.173	8.610	4
E	5	1.476	2.015	2.571	3.365	4.032	5.893	6.869	5
E	6	1.440	1.943	2.447	3.143	3.707	5.208	5.959	6
S	7	1.415	1.895	2.365	2.998	3.499	4.785	5.408	7
	8	1.397	1.860	2.306	2.896	3.355	4.501	5.041	8
O	9	1.383	1.833	2.262	2.821	3.250	4.297	4.781	9
F	10	1.372	1.812	2.228	2.764	3.169	4.144	4.587	10
	11	1.363	1.796	2.201	2.718	3.106	4.025	4.437	11
F	12	1.356	1.782	2.179	2.681	3.055	3.930	4.318	12
R	13	1.350	1.771	2.160	2.650	3.012	3.852	4.221	13
E	14	1.345	1.761	2.145	2.624	2.977	3.787	4.140	14
E	15	1.341	1.753	2.131	2.602	2.947	3.733	4.073	15
D	16	1.337	1.746	2.120	2.583	2.921	3.686	4.015	16
O	17	1.333	1.740	2.110	2.567	2.898	3.646	3.965	17
M	18	1.330	1.734	2.101	2.552	2.878	3.610	3.922	18
	19	1.328	1.729	2.093	2.539	2.861	3.579	3.883	19
	20	1.325	1.725	2.086	2.528	2.845	3.552	3.850	20
	21	1.323	1.721	2.080	2.518	2.831	3.527	3.819	21
	22	1.321	1.717	2.074	2.508	2.819	3.505	3.792	22
	23	1.319	1.714	2.069	2.500	2.807	3.485	3.768	23
	24	1.318	1.711	2.064	2.492	2.797	3.467	3.745	24
	25	1.316	1.708	2.060	2.485	2.787	3.450	3.725	25
	26	1.315	1.706	2.056	2.479	2.779	3.435	3.707	26
	27	1.314	1.703	2.052	2.473	2.771	3.421	3.690	27
	28	1.313	1.701	2.048	2.467	2.763	3.408	3.674	28
	29	1.311	1.699	2.045	2.462	2.756	3.396	3.659	29
	30	1.310	1.697	2.042	2.457	2.750	3.385	3.646	30
	32	1.309	1.694	2.037	2.449	2.738	3.365	3.622	32
	34	1.307	1.691	2.032	2.441	2.728	3.348	3.601	34
	36	1.306	1.688	2.028	2.434	2.719	3.333	3.582	36
	38	1.304	1.686	2.024	2.429	2.712	3.319	3.566	38
	40	1.303	1.684	2.021	2.423	2.704	3.307	3.551	40
	42	1.302	1.682	2.018	2.418	2.698	3.296	3.538	42
	44	1.301	1.680	2.015	2.414	2.692	3.286	3.526	44
	46	1.300	1.679	2.013	2.410	2.687	3.277	3.515	46
	48	1.299	1.677	2.011	2.407	2.682	3.269	3.505	48
	50	1.299	1.676	2.009	2.403	2.678	3.261	3.496	50

Assumptions

- The distribution of the data being tested is normal.
  - For paired t-test, it is the distribution of the subtracted data that must be normal.

Assumptions

- The distribution of the data being tested is normal.
  - For paired t-test, it is the distribution of the subtracted data that must be normal.
  - For unpaired t-test, the distribution of both data sets must be normal.

Assumptions

- The distribution of the data being tested is normal.
  - For paired t-test, it is the distribution of the subtracted data that must be normal.
  - For unpaired t-test, the distribution of both data sets must be normal.
- Plots: Histogram, Density Plot, QQplot,

Assumptions

- The distribution of the data being tested is normal.
  - For paired t-test, it is the distribution of the subtracted data that must be normal.
  - For unpaired t-test, the distribution of both data sets must be normal.
- Plots: Histogram, Density Plot, QQplot,
- Test for Normality: Kolmogorov-Smirnov test, Shapiro-Wilk test.

Assumptions

- The distribution of the data being tested is normal.
  - For paired t-test, it is the distribution of the subtracted data that must be normal.
  - For unpaired t-test, the distribution of both data sets must be normal.
- Plots: Histogram, Density Plot, QQplot,
- Test for Normality: Kolmogorov-Smirnov test, Shapiro-Wilk test.
- Homogeneous: the variances of the two population are equal

Assumptions

- The distribution of the data being tested is normal.
  - For paired t-test, it is the distribution of the subtracted data that must be normal.
  - For unpaired t-test, the distribution of both data sets must be normal.
- Plots: Histogram, Density Plot, QQplot,
- Test for Normality: Kolmogorov-Smirnov test, Shapiro-Wilk test.
- Homogeneous: the variances of the two population are equal
  - If the two populations are symmetric, and if the variances are equal, then the t test may be used.

Assumptions

- The distribution of the data being tested is normal.
  - For paired t-test, it is the distribution of the subtracted data that must be normal.
  - For unpaired t-test, the distribution of both data sets must be normal.
- Plots: Histogram, Density Plot, QQplot,
- Test for Normality: Kolmogorov-Smirnov test, Shapiro-Wilk test.
- Homogeneous: the variances of the two population are equal
  - If the two populations are symmetric, and if the variances are equal, then the t test may be used.
  - If the two populations are symmetric, and the variances are not equal, then use the two-sample unequal variance t-test or Welch's t test.

Assumptions

- The distribution of the data being tested is normal.
  - For paired t-test, it is the distribution of the subtracted data that must be normal.
  - For unpaired t-test, the distribution of both data sets must be normal.
- Plots: Histogram, Density Plot, QQplot,
- Test for Normality: Kolmogorov-Smirnov test, Shapiro-Wilk test.
- Homogeneous: the variances of the two population are equal
  - If the two populations are symmetric, and if the variances are equal, then the t test may be used.
  - If the two populations are symmetric, and the variances are not equal, then use the two-sample unequal variance t-test or Welch's t test.
- Test for equality of the two variances: Variance ratio F-test.

Non-parametric tests

- Constitute a flexible alternative to the t-tests where there is no distributional assumption.
- In cases where a parametric test would be appropriate, non-parametric tests have less power.

Parametric Tests

- Only for quantitative data.
- Assume that the data follows a certain distribution (normal distribution).
- Assuming equal variances and Unequal variances.
- More powerful.

Non-parametric tests:

- Does not assume normal distribution
- No variance assumption
- Decrease effects of outliers (Robust)
- Not recommended if there is less than 5 replicates per group
- Less powerful

References

- Ewens, W.J. and Grant, G.R. (2001) *Statistical Methods in Bioinformatics: An Introduction.* Springer, New York.
- Krijnen, W. (2009) Applied Statistics for Bioinformatics Using R. [http://cran.r-project.org/doc/contrib/Krijnen-IntroBioInfStatistics.pdf]