Genomics Research Unit

BIOSTATISTICS

Lecture 6

One-sample Tests of Hypotheses about Means and Proportions

dr. Petr Nazarov

petr.nazarov@lih.lu

23-03-2018

OUTLINE

- **Hypothesis tests for means and proportions**
  - **hypotheses**
  - **developing hypotheses**
  - **types of errors**
  - **p-value**
  - **one-tail test**
  - **two-tail test and connection with interval estimation**
  - **hypothesis tests for proportion**
  - **errors of type II and power curve**

HYPOTHESES

Null and Alternative Hypotheses

Here we continue the discussion of statistical inference by showing how hypothesis testing can be used to determine whether a statement about the value of a population parameter should or should not be rejected.

In hypothesis testing we begin by making a tentative assumption about a population parameter, i.e. by formulation of a null hypothesis.

Null hypothesis
The hypothesis tentatively assumed true in the hypothesis testing procedure, H_0

Alternative hypothesis
The hypothesis concluded to be true if the null hypothesis is rejected, H_a

H_0: $\mu \leq$ const	H_0: $\mu \geq$ const	H_0: $\mu =$ const
H_a: $\mu >$ const	H_a: $\mu <$ const	H_a: $\mu \neq$ const

HYPOTHESES

Developing Null and Alternative Hypotheses: Example 1

Assume, that an average survival time for the glioblastoma patients (early state, age<50) is 18 months. You have developed a new treatment which should increase the survival time. Performing the clinical trial in order to determine the positive effect you obtained the average survival of 20 months. You would like to ensure that this effect is real, so you perform the hypothesis testing. As a general guideline, a research hypothesis should be stated as the alternative hypothesis. Hence, the appropriate null and alternative hypotheses for the study are

$$H_0: \mu \leq 18$$

$$H_a: \mu > 18$$

If the sample results indicate that H_0 cannot be rejected, researchers cannot conclude the new treatment is better. Perhaps more research and subsequent testing should be conducted. However, if the sample results indicate that H_0 can be rejected, researchers can make the inference that H_a: $\mu > 18$ is true. With this conclusion, the researchers gain the statistical support necessary to state that the new treatment increases survival time, and wide implementation of the treatment should be made.

HYPOTHESES

Developing Null and Alternative Hypotheses: Example 2

Consider the situation of a manufacturer of soft drinks who states that it fills two-liter containers of its products with an average of at least 67.6 fluid ounces. A sample of two-liter containers will be selected, and the contents will be measured to test the manufacturer's claim. In this type of hypothesis testing situation, we generally assume that the manufacturer's claim is true unless the sample evidence is contradictory. Using this approach for the soft-drink example, we would state the null and alternative hypotheses as follows.

$$H_0: \mu \geq 67.6$$

$$H_a: \mu < 67.6$$

If the sample results indicate H_0 cannot be rejected, the manufacturer's claim will not be challenged. However, if the sample results indicate H_0 can be rejected, the inference will be made that H_a: $\mu < 67.6$ is true. With this conclusion, statistical evidence indicates that the manufacturer's claim is incorrect and that the soft-drink containers are being filled with a mean less than the claimed 67.6 ounces. Appropriate action against the manufacturer may be considered.

HYPOTHESES

Developing Null and Alternative Hypotheses: Example 3

For example, on the basis of a sample of parts from a shipment just received, a quality control inspector must decide whether to accept the shipment or to return the shipment to the supplier because it does not meet specifications. Assume that specifications for a particular part require a mean length of two inches per part. If the mean length is greater or less than the two-inch standard, the parts will cause quality problems in the assembly operation. In this case, the null and alternative hypotheses would be formulated as follows.

$$H_0: \mu = 2$$

$$H_a: \mu \neq 2$$

If the sample results indicate H_0 cannot be rejected, the quality control inspector will have no reason to doubt that the shipment meets specifications, and the shipment will be accepted. However, if the sample results indicate H_0 should be rejected, the conclusion will be that the parts do not meet specifications. In this case, the quality control inspector will have sufficient evidence to return the shipment to the supplier.

HYPOTHESES

Type I Error

Type I error
The error of rejecting H_0 when it is true.

Type II error
The error of accepting H_0 when it is false.

Level of significance
The probability of making a Type I error when the null hypothesis is true as an equality, α

poor sensitivity

False Negative, β error

Population Condition

Conclusion		H_0 True	H_a True
	Accept H_0	Correct Conclusion	Type II Error
	Reject H_0	Type I Error	Correct Conclusion

False Positive, α error

poor specificity

HYPOTHESIS TESTING FOR MEAN

One-tailed Test

One-tailed test

A hypothesis test in which rejection of the null hypothesis occurs for values of the test statistic in one tail of its sampling distribution

H_0: $\mu \le \mu_0$ $\qquad$ H_0: $\mu \ge \mu_0$

H_a: $\mu > \mu_0$ $\qquad$ H_a: $\mu < \mu_0$

A Trade Commission (TC) periodically conducts statistical studies designed to test the claims that manufacturers make about their products. For example, the label on a large can of Hilltop Coffee states that the can contains 3 pounds of coffee. The TC knows that Hilltop's production process cannot place exactly 3 pounds of coffee in each can, even if the mean filling weight for the population of all cans filled is 3 pounds per can. However, as long as the population mean filling weight is at least 3 pounds per can, the rights of consumers will be protected. Thus, the TC interprets the label information on a large can of coffee as a claim by Hilltop that the population mean filling weight is at least 3 pounds per can. We will show how the TC can check Hilltop's claim by conducting a lower tail hypothesis test.

μ_0 = 3 lbm

Suppose sample of n=36 coffee cans is selected. From the previous studies it's known that σ = 0.18 lbm

HYPOTHESIS TESTING FOR MEAN

One-tailed Test: Example

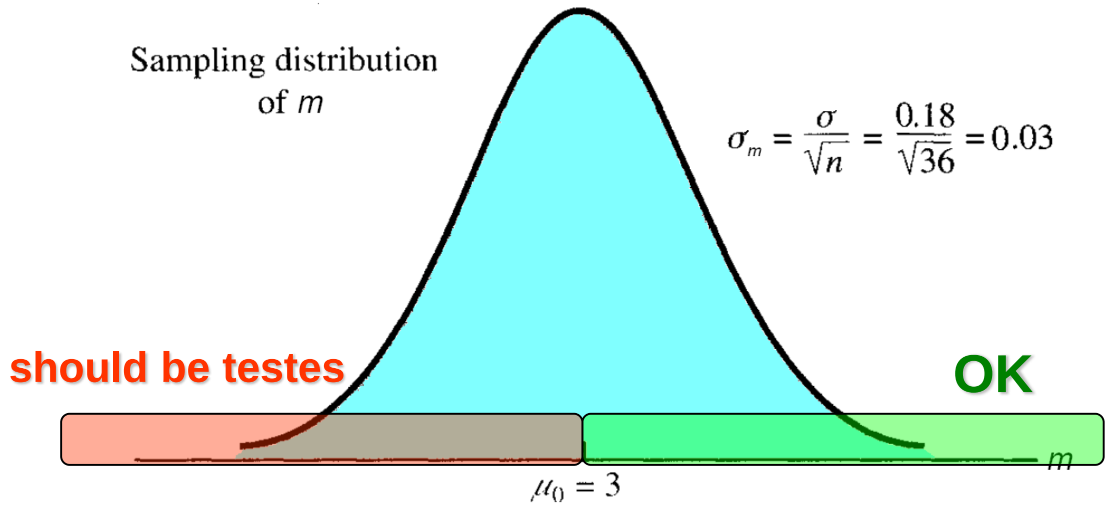
$\mu_0 = 3$ **lbm**

Suppose sample of $n = 36$ coffee cans is selected and $m = 2.92$ is observed. From the previous studies it's known that $\sigma = 0.18$ lbm

H_0: $\mu \geq 3$ **no action**

H_a: $\mu < 3$ **legal action**

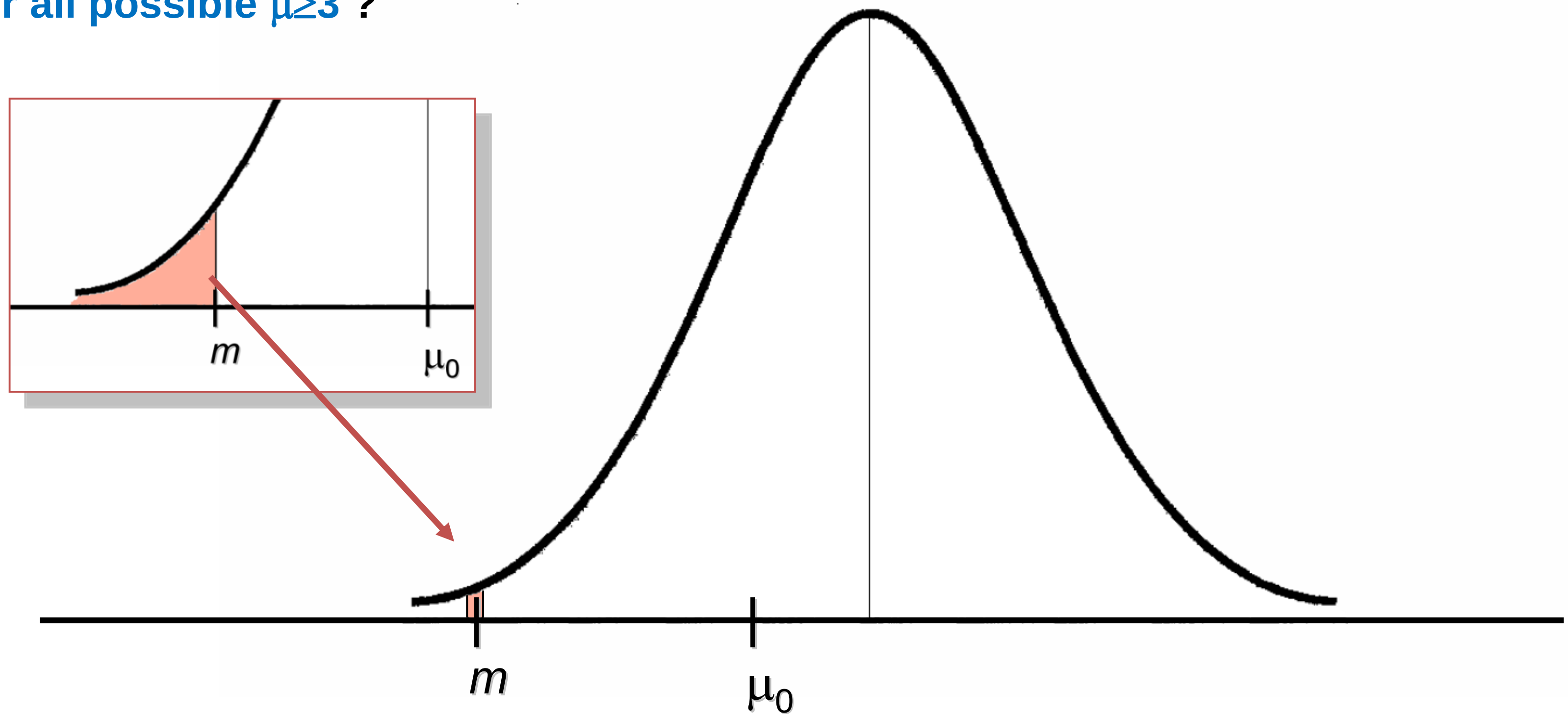
Let's say: in the extreme case, when μ=3, we would like to be 99% sure that we make no mistake, when starting legal actions against Hilltop Coffee. It means that selected significance level is $\alpha = 0.01$



HYPOTHESIS TESTING FOR MEAN

Let's Try to Understand…

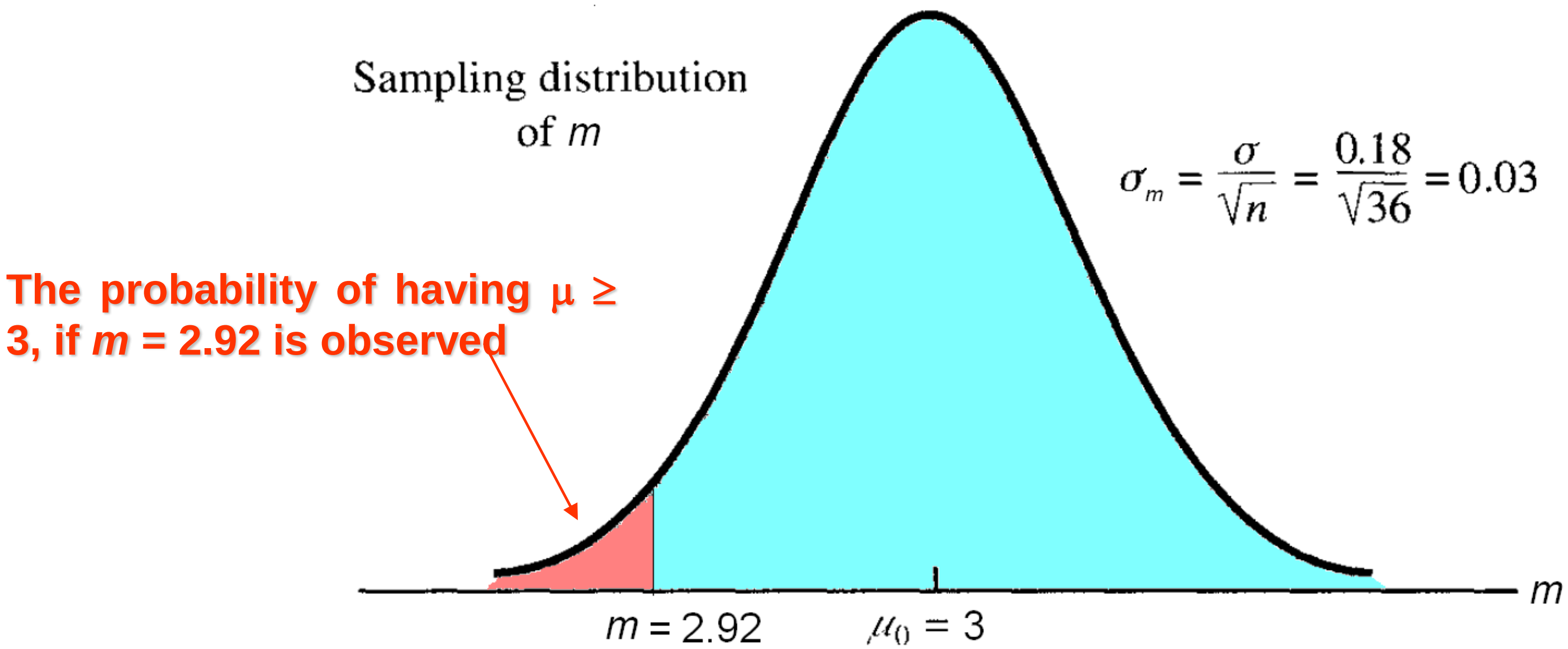
Let's find the probability of observation *m* for all possible $\mu \geq 3$. We start from an extreme case ($\mu=3$) and then probe all possible $\mu>3$. See the behavior of the small probability area around measured *m*. What you will get if you summarize its area for all possible $\mu \geq 3$?



$P(m)$ for all possible $\mu \geq \mu_0$ is equal to $P(x<m)$ for an extreme case of $\mu=\mu_0$

HYPOTHESIS TESTING FOR MEAN

Let's Try to Understand…



In other words, red area characterizes the probability of the null hypothesis.

To be completely correct, the red area gives us a probability of making an error when rejecting the null hypothesis, or the <u>p-value.</u>

HYPOTHESIS TESTING FOR MEAN

One-tailed Test

here $\mu_0 = 3$

H_0: $\mu \geq 3$

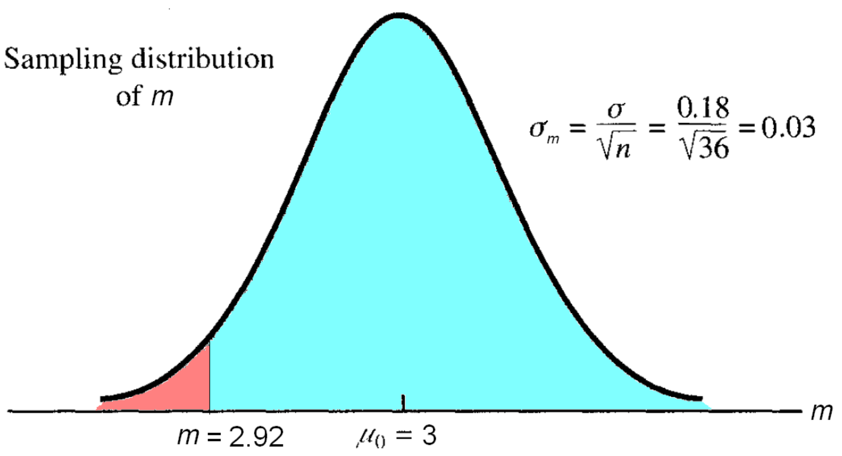
H_a: $\mu < 3$

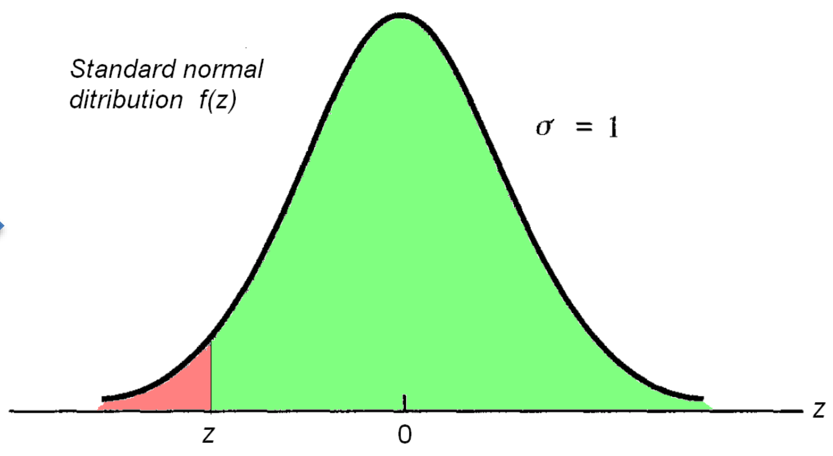
Step 1. Introduce the test statistics

Test statistic

A statistic whose value helps determine whether a null hypothesis can be rejected

$$z = \frac{m - \mu_0}{\sigma / \sqrt{n}}$$





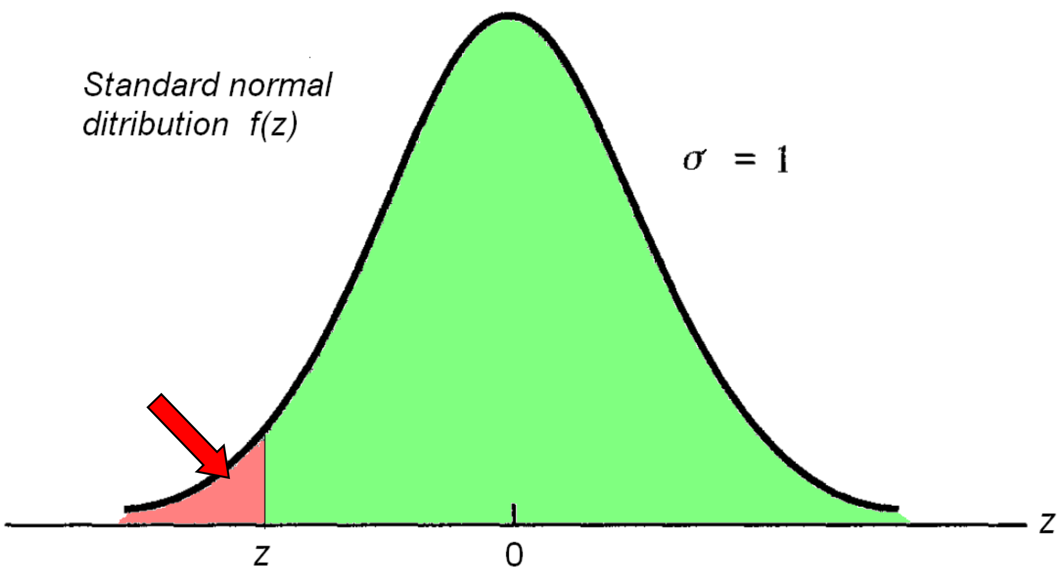
HYPOTHESIS TESTING FOR MEAN

One-tailed Test

Step 2. Calculate p-value and compare it with α

p-value

A probability, computed using the test statistic, that measures the support (or lack of support) provided by the sample for the null hypothesis. It is a probability of making error of type I



$$p > \alpha \Rightarrow H_0$$
$$p < \alpha \Rightarrow H_a$$

m = AVERAGE(…)

n = number of experiments

σ = given constant

μ_0 = given constant

z = (m- μ_0)/ σ *SQRT(n)

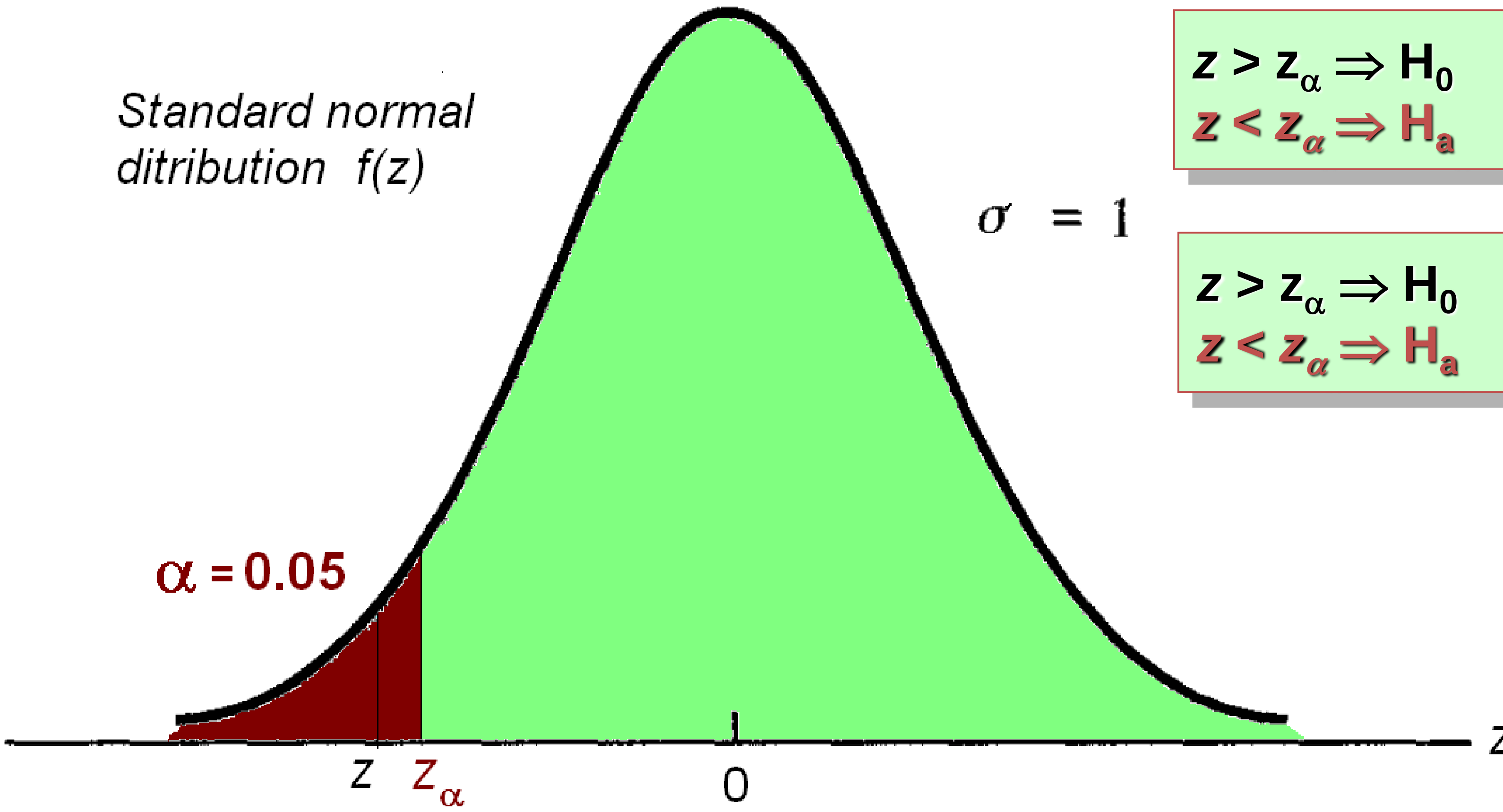
p-value = NORM.S.DIST(-ABS(z), true)

HYPOTHESIS TESTING FOR MEAN

One-tailed Test

Critical value

A value that is compared with the test statistic to determine whether H_0 should be rejected



$z > z_\alpha \Rightarrow H_0$
$z < z_\alpha \Rightarrow H_a$

$z > z_\alpha \Rightarrow H_0$
$z < z_\alpha \Rightarrow H_a$

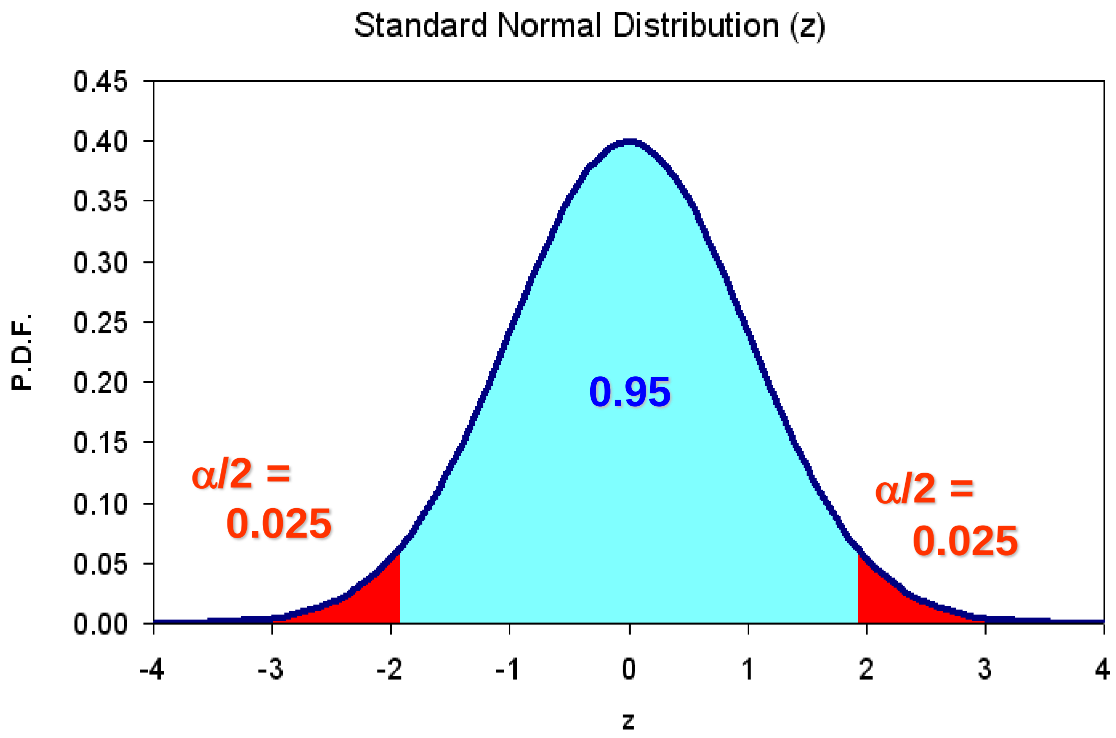
HYPOTHESIS TESTING FOR THE MEAN

Two-tailed test

A hypothesis test in which rejection of the null hypothesis occurs for values of the test statistic in either tail of its sampling distribution.

$$H_0: \mu = \mu_0$$

$$H_a: \mu \neq \mu_0$$



HYPOTHESIS TESTING FOR THE MEAN

σ is Unknown

if σ in unknown:

$\sigma \rightarrow s$

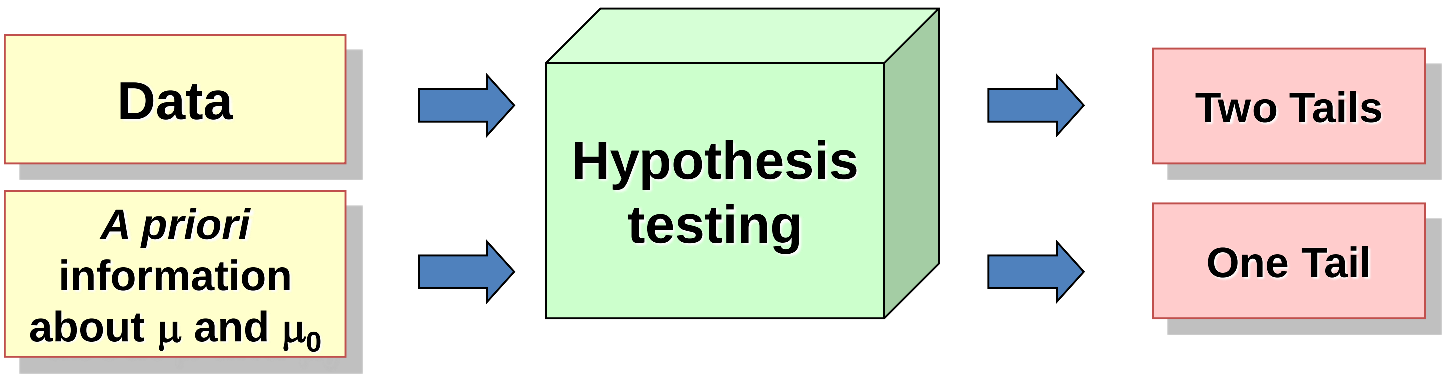
$z \rightarrow t$

	Lower Tail Test	Upper Tail Test	Two-Tailed Test
Hypotheses	$H_0 : \mu \geq \mu_0$ $H_a : \mu < \mu_0$	$H_0 : \mu \leq \mu_0$ $H_a : \mu > \mu_0$	$H_0 : \mu = \mu_0$ $H_a : \mu \neq \mu_0$
Test Statistic	$t = \frac{m - \mu_0}{s/\sqrt{n}}$	$t = \frac{m - \mu_0}{s/\sqrt{n}}$	$t = \frac{m - \mu_0}{s/\sqrt{n}}$
Rejection Rule: p-Value Approach	Reject H_0 if p-value $\leq \alpha$	Reject H_0 if p-value $\leq \alpha$	Reject H_0 if p-value $\leq \alpha$
Rejection Rule: Critical Value Approach	Reject H_0 if $t \leq -t_\alpha$	Reject H_0 if $t \geq t_\alpha$	Reject H_0 if $t \leq -t_{\alpha/2}$ or if $t \geq t_{\alpha/2}$

HYPOTHESIS TESTING FOR THE MEAN

One Tail Test vs. Two Tail Test

There is a raging controversy (for about the last hundred years) on whether or not it is ever appropriate to use a one-tailed test. The rationale is that if you already know the direction of the difference, why bother doing any statistical tests. While it is **generally safest to use a two-tailed tests**, there are situations where a one-tailed test seems more appropriate. The bottom line is that **it is the choice of the researcher** whether to use one-tailed or two-tailed research questions.



$$2\times\text{p-value}_{(1\ \text{tail})} = \text{p-value}_{(2\ \text{tails})}$$

HYPOTHESIS TESTING FOR THE MEAN

Example

Number of living cells in 5 wells under some conditions are given in the table, with average value of 4705. In a reference literature source authors clamed a mean quantity of 5000 living cells under the same conditions. Is our result significantly different?

Well	Cells
1	5128
2	4806
3	5037
4	4231
5	4322

Two Tails

H_0: $\mu = 5000$

H_a: $\mu \neq 5000$

Let's use α=0.05

$$t = \frac{m - \mu_0}{s/\sqrt{n}}$$

m = AVERAGE(A2:A6)

s = STDEV(A2:A6)

μ_0 = 5000

t = (m- μ_0)/s*SQRT(5)

p-value = T.DIST(-ABS(t);5-1;true)

n	5
mean	4704.8
stdev	409.49
mu	5000
t	-1.612
p-value 2 t	0.1823
p-value 1 t	0.0911

Two tail using Excel 2010

p-value = 2*T.DIST(**-**ABS(z), d.f., true**)**

p-value = T.DIST.2T(ABS(z), d.f.**)**

Two tail using Excel 2003

p-value = TDIST(ABS(z), d.f., tails**)**

HYPOTHESIS TESTING FOR PROPORTION

Hypotheses for Proportions

Proportions

π – population proportion

p – sample proportion

π_0 – testing value

For the proportions:
1) use z-statistics
2) use proper equation for σ_p

$np \geq 5$, $n(1-p) \geq 5$

	Lower Tail Test	Upper Tail Test	Two-Tailed Test
Hypotheses	$H_0: \pi \geq \pi_0$ $H_a: \pi < \pi_0$	$H_0: \pi \leq \pi_0$ $H_a: \pi > \pi_0$	$H_0: \pi = \pi_0$ $H_a: \pi \neq \pi_0$
Test Statistic	$z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$	$z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$	$z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$
Rejection Rule: p-Value Approach	Reject H_0 if p-value $\leq \alpha$	Reject H_0 if p-value $\leq \alpha$	Reject H_0 if p-value $\leq \alpha$
Rejection Rule: Critical Value Approach	Reject H_0 if $z \leq -z_\alpha$	Reject H_0 if $z \geq z_\alpha$	Reject H_0 if $z \leq -z_{\alpha/2}$ or if $z \geq z_{\alpha/2}$

Two tail: p-value = 2*NORM.S.DIST(-ABS(z),TRUE)

HYPOTHESIS TESTING FOR MEAN

An Easy Way to Two Tail Hypothesis

1) **Formulate the hypothesis ($m = \mu_0$)**
2) **Select α**
3) **Calculate the interval estimation for 1-α confidence**
4) **Check whether your μ_0 is inside the interval**

Well	Cells
1	5128
2	4806
3	5037
4	4231
5	4322

E = CONFIDENCE.T (α, s, n)

m=	4705
E=	359
m-E=	4346
m+E=	5064

ONE SAMPLE HYPOTHESES

Formulate Hypotheses

Proportion

$\pi <>= \pi_0$ (const)

Ensure that

$np \geq 5$

$n(1-p) \geq 5$

Mean

$\mu <>= \mu_0$ (const)

Standard deviation σ is known

Standard deviation σ is unknown (use s)

Calculate statistics ***z***

Calculate statistics ***t***

$$z = \frac{p - \pi_0}{\sqrt{\dfrac{\pi_0(1-\pi_0)}{n}}}$$

$$z = \frac{m - \mu_0}{\sigma / \sqrt{n}}$$

$$t = \frac{m - \mu_0}{s / \sqrt{n}}$$

Calculate p-value

◆= τ * NORM.S.DIST(-abs(z))

◆= τ * T.DIST(-abs(t),n-1)

τ = 1 for “</>” hypotheses (one tail)
τ = 2 for “=” hypothesis (two tails)

HYPOTHESES

Type II Error

Type I error
The error of rejecting H_0 when it is true.

Type II error
The error of accepting H_0 when it is false.

Level of significance
The probability of making a Type I error when the null hypothesis is true as an equality

poor sensitivity

False Negative, β error

Population Condition

Conclusion		H_0 True	H_a True
	Accept H_0	Correct Conclusion	Type II Error
	Reject H_0	Type I Error	Correct Conclusion

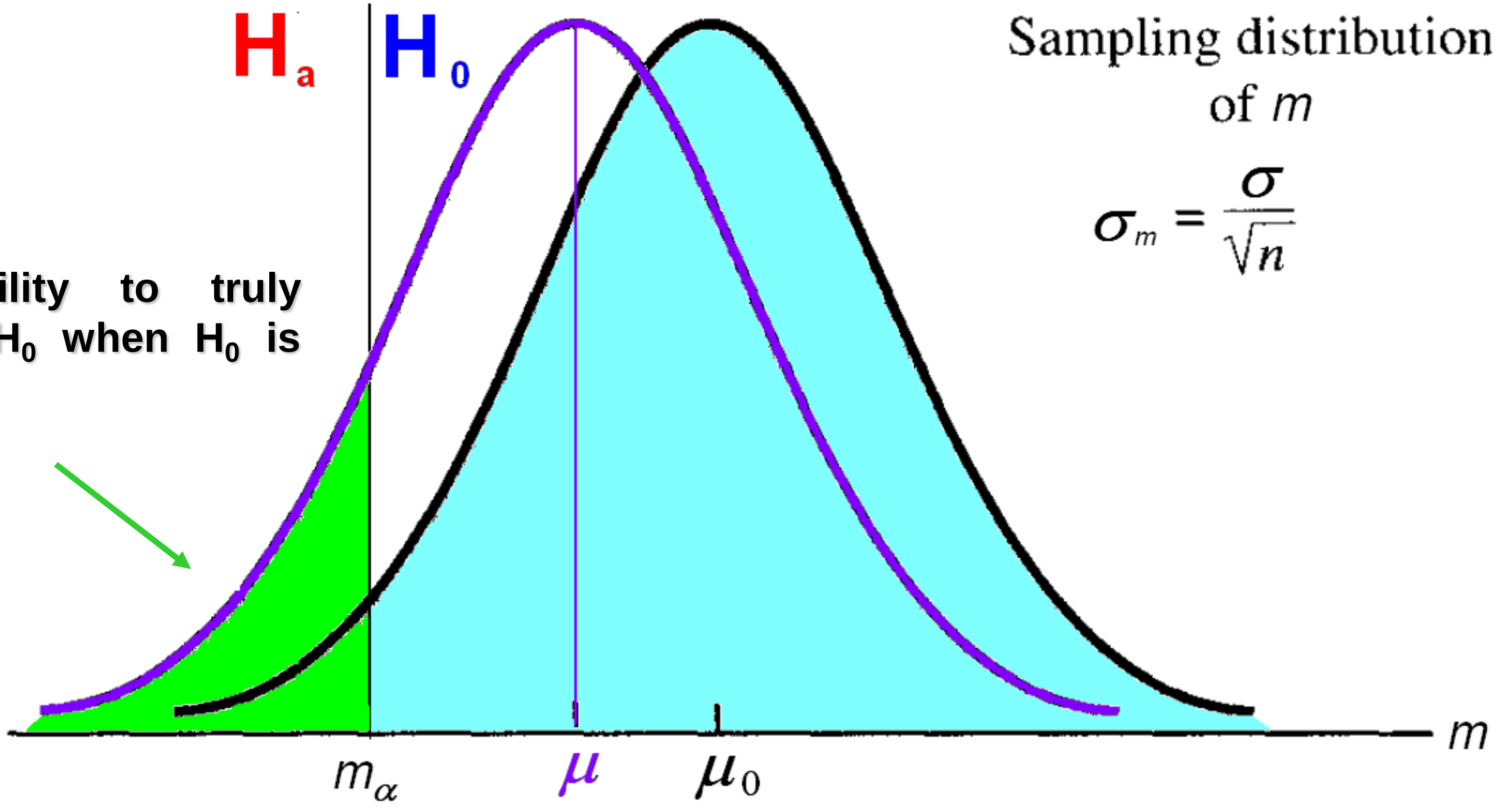
False Positive, α error

poor specificity

HYPOTHESES

Power Curve

Probability to truly reject H_0 when H_0 is false.



HYPOTHESES

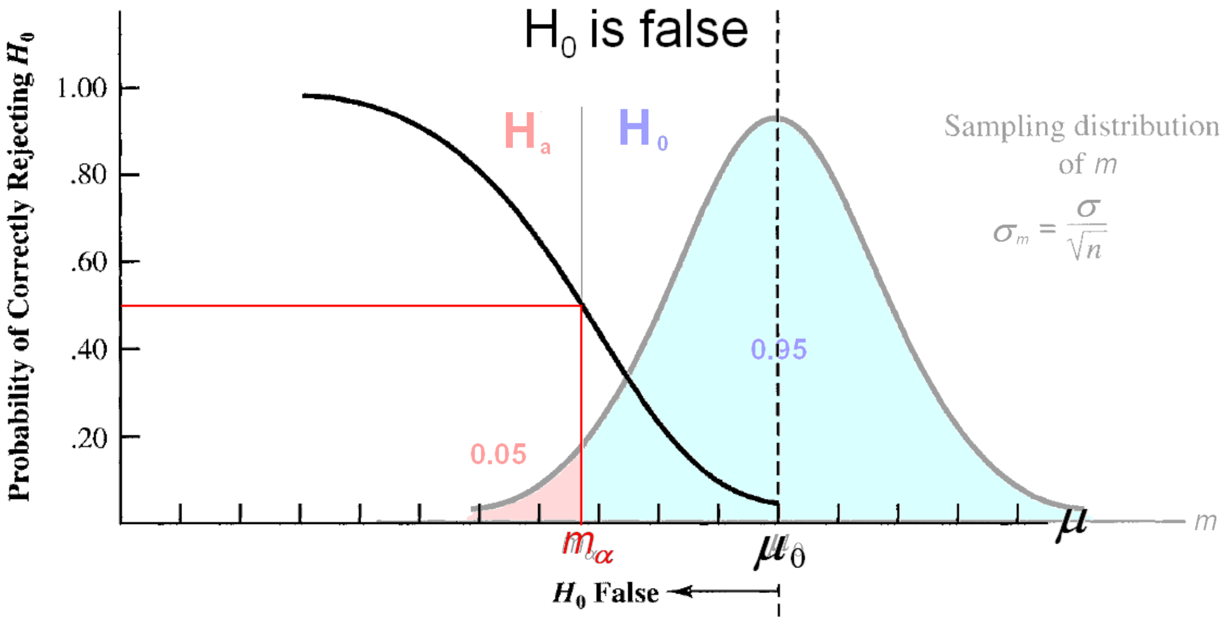
Power Curve

Power

The probability of correctly rejecting H_0 when it is false

Power curve

A graph of the probability of rejecting H_0 for all possible values of the population parameter not satisfying the null hypothesis. The power curve provides the probability of correctly rejecting the null hypothesis



UNIVERSITÉ DU LUXEMBOURG

QUESTIONS ?

Thank you for your attention

To be continued….