SmartCow
an integrated infrastructure for increased research capability and innovation in the European cattle sector

SmartCow project

Ethics in experiments on animals

Adjustment of animal numbers in experimentation

Question : Why is this important?

Patrick Gasqui

Tuesday September 22, 2020

INRAE – VetAgro Sup, Unité Mixte de Recherches d'Epidémiologie des maladies animales et zoonotique (EPIA), Centre de recherche de Clermont Auvergne-Rhône-Alpes, Département de Santé Animale (SA)

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité
INRAE
VetAgro Sup

SmartCow
an integrated infrastructure for increased research capability and innovation in the European cattle sector

Adjustment of animal numbers in experimentation

"Classical" statistical theory for using a test :

the initial framework of this presentation …

1 - as part of a one-sided hypothesis test:

H_0 : « $\mu_1 \leq \mu_0$ » *versus* H_1 : « $\mu_1 > \mu_0$ »

Adjustment of animal numbers in experimentation

"Classical" statistical theory for using a test :

the initial framework of this presentation …

1 - as part of a one-sided hypothesis test:

H_0 : « $\mu_1 \leq \mu_0$ » *versus* H_1 : « $\mu_1 > \mu_0$ »

for the user, in this presentation the hypothesis « H_1 » is the hypothesis of interest

Adjustment of animal numbers in experimentation

"Classical" statistical theory for using a test :

the initial framework of this presentation …

1 - as part of a one-sided hypothesis test:

H_0 : « $\mu_1 \leq \mu_0$ » ***versus*** H_1 : « $\mu_1 > \mu_0$ »

for the user, in this presentation the hypothesis « H_1 » is the hypothesis of interest

the user therefore wants to highlight a difference in means: « delta = μ_1 - μ_0 »

Adjustment of animal numbers in experimentation

the initial framework of this presentation …

2 - for a random variable of interest « Y » which follows a Gaussian distribution with standard deviation « σ_Y »

using a statistic of test: « T »
whose law (distribution) is known under H_0

and whose value is « T_{obs} » after the measurement results of a sample size « N »

example statistic « T »: the empirical mean

$$\overline{Y_N} = \frac{1}{N} . \sum_{i=1}^{i=N} Y_i$$

Adjustment of animal numbers in experimentation

the initial framework of this presentation …

2 - for a random variable of interest « Y » which follows a Gaussian distribution with standard deviation « σ_Y »

the standard deviation « σ_Y » is the precision of the studied variable « Y »

using a statistic of test: « T » whose law (distribution) is known under H_0

and whose value is « T_{obs} » after the measurement results of a sample size « N »

example statistic « T »: the empirical mean

$$\overline{Y_N} = \frac{1}{N} . \sum_{i=1}^{i=N} Y_i$$

Adjustment of animal numbers in experimentation

SmartCow
an integrated infrastructure for increased research capability and innovation in the European cattle sector

the initial framework of this presentation …

2 - for a random variable of interest « Y » which follows a Gaussian distribution with standard deviation « σ_Y »

using a statistic of test: « T » whose law (distribution) is known under H_0

and whose value is « T_{obs} » after the measurement results of a sample size « N »

example statistic « T »: the empirical mean

$$\overline{Y_N} = \frac{1}{N} . \sum_{i=1}^{i=N} Y_i$$

the standard deviation « σ_Y » is the precision of the studied variable « Y »

this precision « σ_Y » is important given that we want to be able to highlight a difference between two means « delta = µ1 - µ0 » with the variable « Y »

SmartCow
an integrated infrastructure for increased research capability and innovation in the European cattle sector

Adjustment of animal numbers in experimentation

the initial framework of this presentation …

2 - for a random variable of interest « Y » which follows a Gaussian distribution with standard deviation « σ_Y »

using a statistic of test: « T » whose law (distribution) is known under H_0

and whose val is « T_{obs} » after the measurement resu of a sample size « N »

example statistic « T »: the empirical mean

$$\overline{Y_N} = \frac{1}{N} . \sum_{i=1}^{i=N} Y_i$$

an example is the student test statistic when « Y » follows a Gaussian distribution

Adjustment of animal numbers in experimentation

the initial framework of this presentation …

3 - after choosing:

a risk of error of the first kind: « $\alpha = 0.05$ »

⟺

a confidence coefficient for the Confidence Interval (CI): « $\Upsilon = 0.95$ »

⟺

a particular value of the test statistic: « T_0 » under H_0

SmartCow
an integrated infrastructure for increased research capability and innovation in the European cattle sector

Adjustment of animal numbers in experimentation

the initial framework of this presentation …

3 - after choosing:

a risk of error of the first kind: « $\alpha = 0.05$ »

⇔

a confidence coefficient for the Confidence Interval (CI): « $\Upsilon = 0.95$ »

⇔

a particular value of the test statistic: « T_0 » under H_0

↓

« $\alpha = 0.05$ » is the risk of error conventionally taken in practice

Adjustment of animal numbers in experimentation

4 - we are led to take a "*decision*" from the result « T_{obs} » obtained on the sample :

either: *« we cannot reject H_0 »* ***which means: « $T_{obs} \leq T_0$ » or « Pvalue ≥ α »***
either: *« we accept H_1 »* ***which means: « $T_{obs} > T_0$ » or « Pvalue < α »***

without knowing the « reality » …

	« reality » H_0	« reality » H_1
« decision » H_0	« ϒ = 0.95 »	?
« decision » H_1	« α = 0.05 »	?

α = error of the first kind "controlled *a priori*"

⟺

it is the user who chooses its value *a priori*

$\alpha = 1 - \gamma$

α is the probability of concluding that there is a difference when there is none *[“notion of false positive”]*

Adjustment of animal numbers in experimentation

4 - we are led to take a "*decision*" from the result « T_{obs} » obtained on the sample :

either: *« we cannot reject H_0 »* ***which means: « $T_{obs} \leq T_0$ » or « Pvalue ≥ α »***
either: *« we accept H_1 »* ***which means: « $T_{obs} > T_0$ » or « Pvalue < α »***

without knowing the « reality » …

	« reality » H_0	« reality » H_1
« decision » H_0	« ϒ = 0.95 »	?
« decision » H_1	« α = 0.05 »	?

$\alpha = 1 - \gamma$

α = error of the first kind "controlled *a priori*"

« α » and « γ » with the concept of confidence interval (CI) :

ϒ = « probability that the CI:] -∞, T_0] contains the observed value: T_{obs} under H_0 »

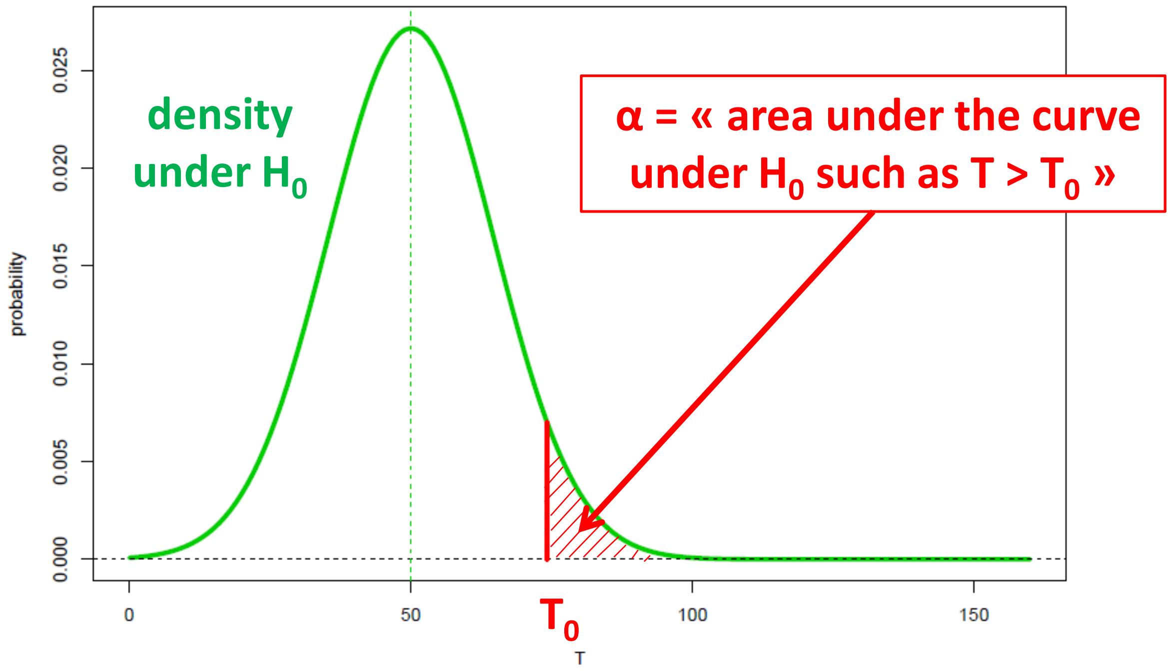
α = « probability that the CI:] -∞, T_0] no contains the observed value: T_{obs} under H_0 »

it is the notion of equivalence between doing a test or calculating a confidence interval

Adjustment of animal numbers in experimentation

We can visualize this with the density distribution of "T" under H_0 with « N = 6 » and « σ_Y = 36 »

	« reality » H_0	« reality » H_1
« decision » H_0	« ϒ = 0.95 »	?
« decision » H_1	« α = 0.05 »	?



we are led to take a "*decision*" from the result « T_{obs} » obtained on the sample :

if « $T_{obs} \leq T_0$ »
or « Pvalue ≥ α »
« we cannot reject H_0 »
H_0 : « $\mu_1 \leq \mu_0$ »

if « $T_{obs} > T_0$ »
or « Pvalue < α »
« we accept H_1 »
H_1 : « $\mu_1 > \mu_0$ »

Adjustment of animal numbers in experimentation

therefore, it is classically the user who chooses a priori:

- ✓ **the hypotheses to be tested: H_0 : « $\mu_1 \leq \mu_0$ » versus H_1 : « $\mu_1 > \mu_0$ »**
- ✓ **the test and therefore the test statistic “T” under H_0**
- ✓ **a sample size « N »**
- ✓ **a risk of error of the first kind: « $\alpha = 0.05$ », therefore the confidence coefficient « $\gamma = 1 - \alpha$ » and therefore a particular value of the test statistic: « T_0 » under H_0.**

from the observation « T_{obs} » obtained on the sample the user has his result

Ok, but where is the problem ?

Adjustment of animal numbers in experimentation

therefore, it is classically the user who chooses a priori:

- ✓ **the hypotheses to be tested: H_0 : « $\mu_1 \leq \mu_0$ » versus H_1 : « $\mu_1 > \mu_0$ »**
- ✓ **the test and therefore the test statistic "T" under H_0**
- ✓ **a sample size « N »**
- ✓ **a risk of error of the first kind: « $\alpha = 0.05$ », therefore the confidence coefficient « $\gamma = 1 - \alpha$ » and therefore a particular value of the test statistic: « T_0 » under H_0.**

from the observation « T_{obs} » obtained on the sample the user has his result

Ok, but where is the problem ?

The problem is that you have to define the "sensitivity" of a test, that is, its ability to identify a difference when it exists.

Note: We have the same problem when we take a measurement with a device. We try to know its sensitivity before using it.

Adjustment of animal numbers in experimentation

5 – on the other hand, we have "no" a priori information on "reality" under H_1 , while we also have a possible error (called « β »), *either that of « deciding H_0 » while the « reality is H_1 »*.

?

	« reality » H_0	« reality » H_1
« decision » H_0	« ϒ = 0.95 »	« β » ?
« decision » H_1	« α = 0.05 »	« 1 – β » ?

β = second kind error "not known *a priori*"
because in general we do not know
the distribution of T under H_1
and because we do not know the value of « μ_1 »

1 - β = power of the test
= « ability of the test to detect
a difference when it exists »

β is the probability of concluding that there is no difference when there is one *[“notion of false negative”]*

The « power of the test » is ultimately the "sensitivity" of the statistical test to be able to "detect" a difference.

It is important that the power of the test is as great as possible

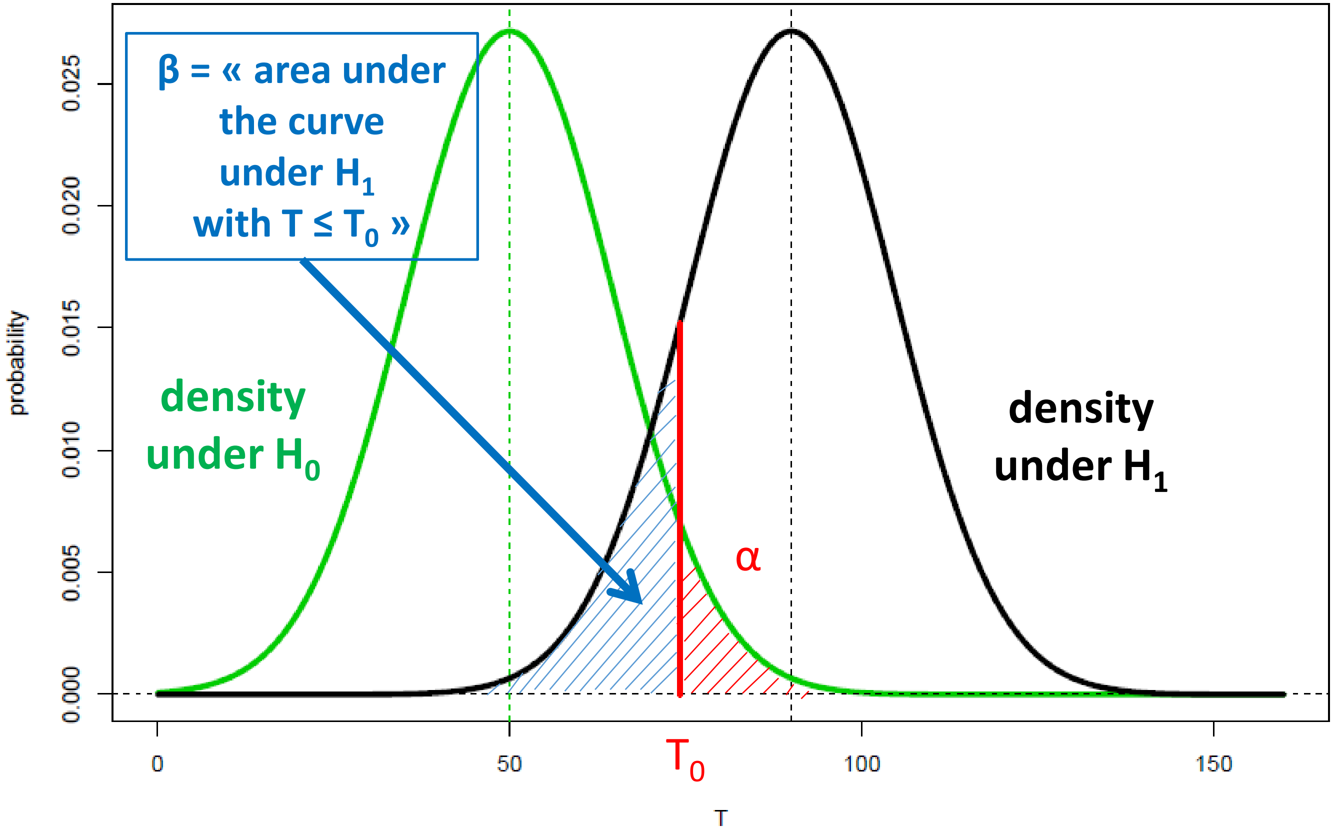
Adjustment of animal numbers in experimentation

example of density distribution of statistical « T » under H_0 and under H_1

with: delta = $\mu_1 - \mu_0$

with N = 6 and σ_Y = 36

if delta = 40 → *then « β » can be estimated*



If "delta" is fixed, then we can estimate "β" and also plot the distribution under H_1.

	« reality » H_0	« reality » H_1
« decision » H_0	« ϒ = 0.95 »	*« β ≈ 0.14 »*
« decision » H_1	« α = 0.05 »	*« 1 – β ≈ 0.86 »*

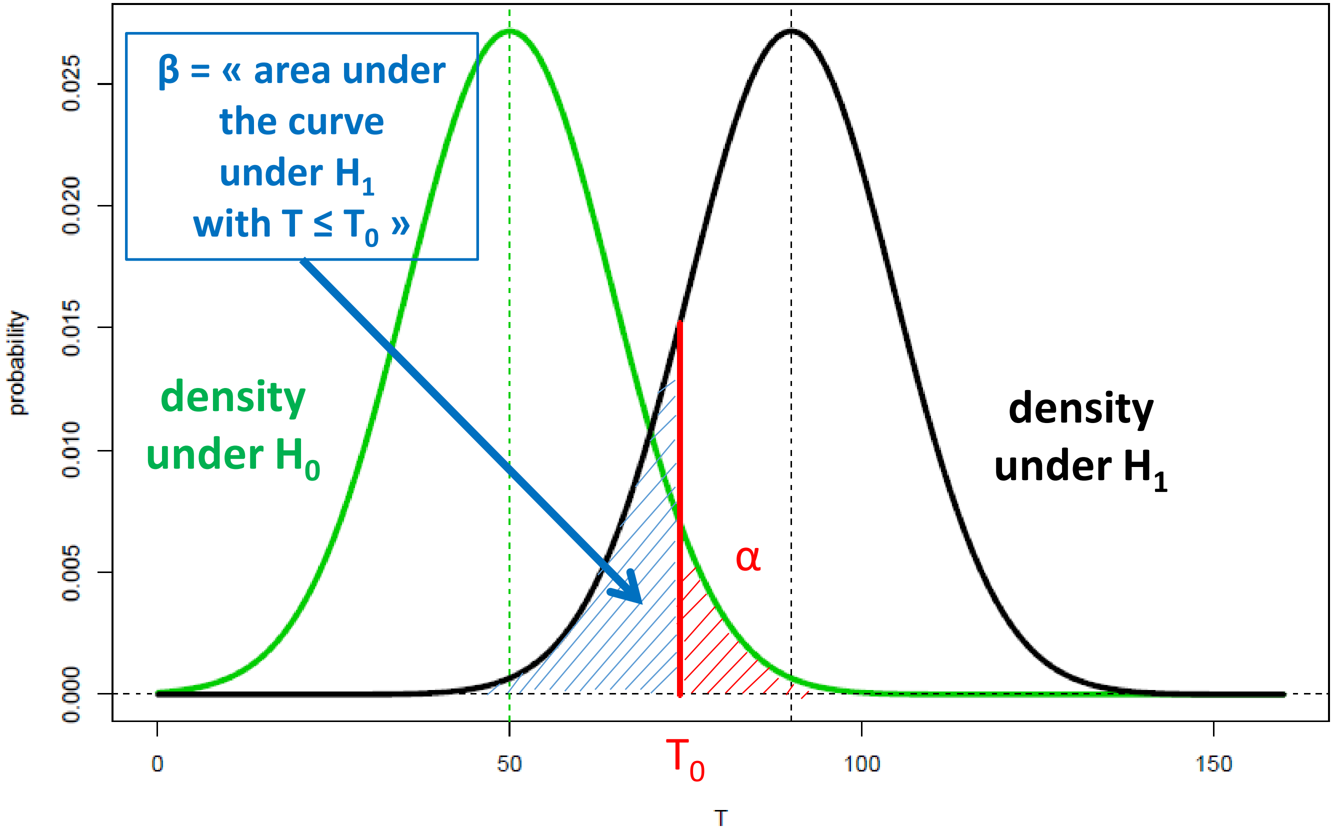
Adjustment of animal numbers in experimentation

example of density distribution of statistical « T » under H_0 and under H_1

with: delta = μ_1 - μ_0

with N = 6 and σ_Y = 36

if delta = 40 → *then « β » can be estimated*

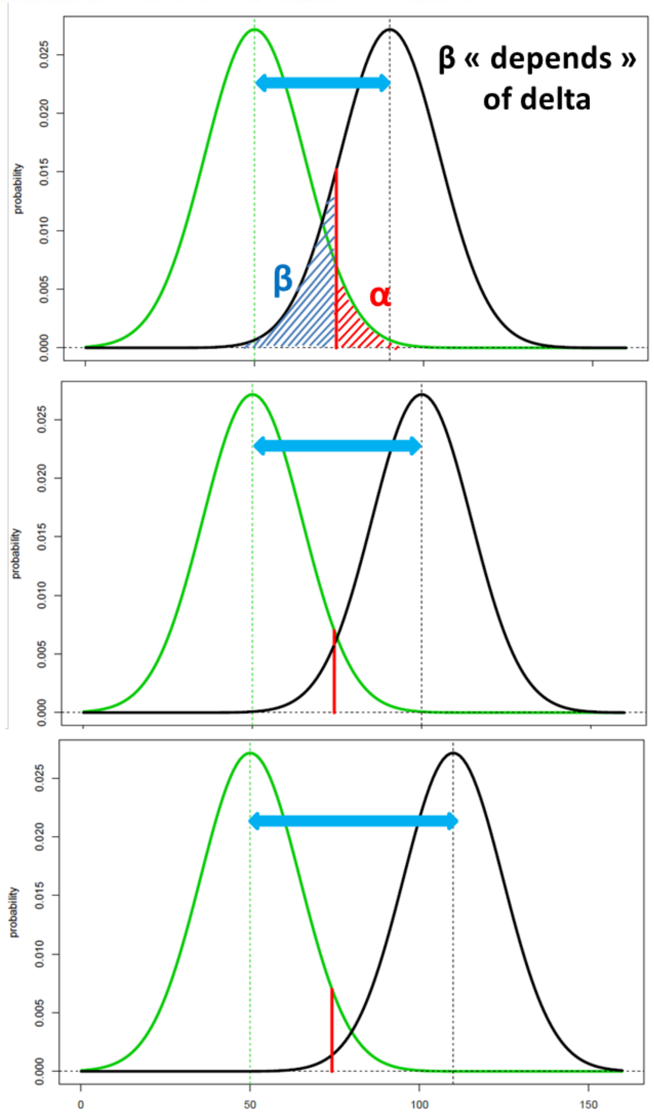


if the parameters: « α », « delta », « N » and « σ_Y » are fixed a priori, then we can estimate « β », « 1-β » and also plot the distribution under H1.

SmartCow
an integrated infrastructure for increased research capability and innovation in the European cattle sector

Adjustment of animal numbers in experimentation

with delta = $\mu_1 - \mu_0$ and $\sigma_Y = 36$



if delta = 40 — *with N = 6*

	« reality » H_0	« reality » H_1
H_0	« ϒ = 0.95 »	*« β ≈ 0.14 »*
H_1	« α = 0.05 »	*« 1 – β ≈ 0.86 »*

if delta = 50 — *with N = 6*

	« reality » H_0	« reality » H_1
H_0	« ϒ = 0.95 »	*« β ≈ 0.04 »*
H_1	« α = 0.05 »	*« 1 – β ≈ 0.96 »*

if delta = 60 — *with N = 6*

	« reality » H_0	« reality » H_1
H_0	« ϒ = 0.95 »	*« β ≈ 0.01 »*
H_1	« α = 0.05 »	*« 1 – β ≈ 0.99 »*

For a fixed sample size “N”,
as "delta" increases,

the second kind error "β" decreases
and
the power of the test”1-β” increases

If we want a power of the test at least 0.99,
with a sample size “N = 6”,
we can only identify that
a difference in means “delta” greater than 60.

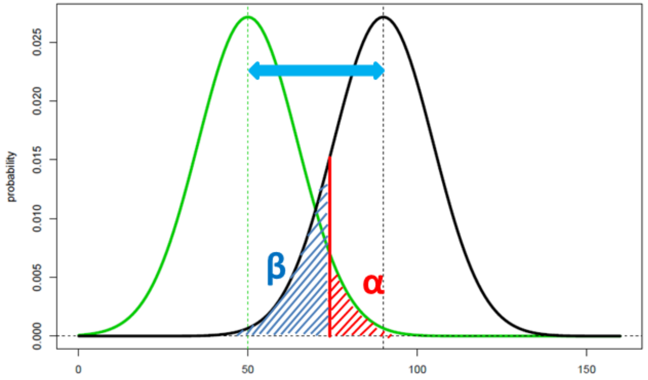
But we want to be able to identify a difference in averages of the order of 40. How do we do this?

Adjustment of animal numbers in experimentation

SmartCow
an integrated infrastructure for increased research capability and innovation in the European cattle sector

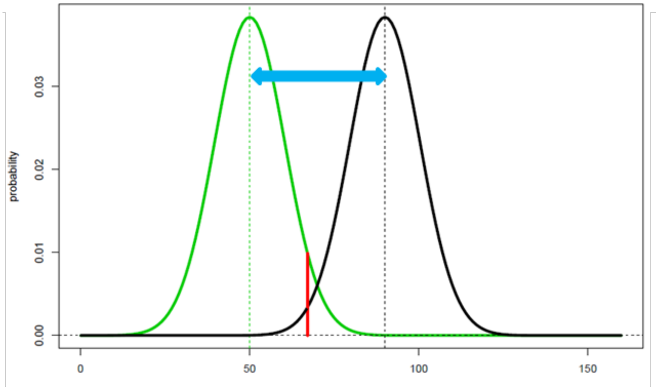
so if we want to keep a good ability to detect a « *given difference* », and so if we want to have a « *minimum sensitivity a priori* », we have two main complementary solutions:

with a « minimum delta » of 40



with N = 6

	« reality » H_0	« reality » H_1
H_0	« ϒ = 0.95 »	***« β ≈ 0.141 »***
H_1	**« α = 0.05 »**	***« 1 – β ≈ 0.859 »***



with N = 12

	« reality » H_0	« reality » H_1
H_0	« ϒ = 0.95 »	***« β ≈ 0.014 »***
H_1	**« α = 0.05 »**	***« 1 – β ≈ 0.986 »***

1st solution :
we « increase »
the size « N » of the sample

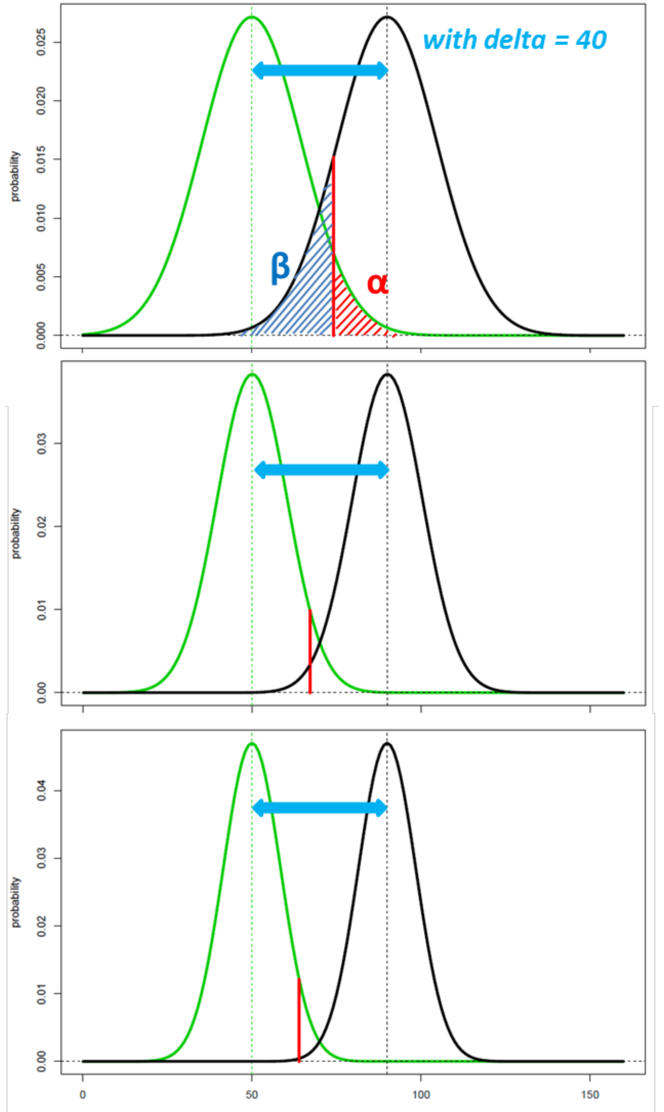
$$Var(\bar{Y}) = \frac{1}{N}.Var(Y) = \frac{1}{N}.\sigma_Y{}^2$$

For a fixed "minimum delta"
as sample size "N" increases,
the second kind error "β" decreases
and
the power of the test"1-β" increases

If we want a power of the test at least 0.986,
and
identify a difference in means "delta = 40",
We just need to take a sample size "N = 12".

SmartCow
an integrated infrastructure for increased research capability and innovation in the European cattle sector

Adjustment of animal numbers in experimentation



with N = 6

	« reality » H_0	« reality » H_1
H_0	« ϒ = 0.95 »	« β ≈ 0.141 »
H_1	« α = 0.05 »	« 1 – β ≈ 0.859 »

with N = 12

	« reality » H_0	« reality » H_1
H_0	« ϒ = 0.95 »	« β ≈ 0.014»
H_1	« α = 0.05 »	« 1 – β ≈ 0.986 »

with N = 18

	« reality » H_0	« reality » H_1
H_0	« ϒ = 0.95 »	« β ≈ 0.001 »
H_1	« α = 0.05 »	« 1 – β ≈ 0.999»

**For a fixed "minimum delta"
as sample size "N" increases,
the second kind error "β" decreases
and
the power of the test"1-β" increases**

$$Var(\bar{Y}) = \frac{1}{N}.Var(Y) = \frac{1}{N}.\sigma_Y{}^2$$

**In fact, the decrease in "β"
and the increase in the power of the test "1-β"
are very rapid [in (1/N)],
as the sample size "N" increases.**

SmartCow
an integrated infrastructure for increased research capability and innovation in the European cattle sector

Adjustment of animal numbers in experimentation

formula for N_{min}

u_p is the quantile of probability "**p**" for the Gaussian distribution *N*(0,1).

for $\alpha = 0.05$ we have **$u_{1-\alpha} = 1.644854$**

and for $\beta = 0.01$ or $(1-\beta) = 0.99$ we have **$u_\beta = -2.326348$**

The minimum « N » size allowing the desired « precision », ie « $(1-\beta) \geq 0.99$ » is such that :

$$N \geq \left[\frac{\sigma_Y \, . \left(u_{1-\alpha} - u_\beta\right)}{delta}\right]^2$$

with « delta = $\mu_1 - \mu_0$ » for the minimum detectable deviation sought
and « σ_Y » for the standard deviation of the Gaussian law of the studied and measured variable Y.

For the previous example we had « delta = 40 » and « σ_Y = 36 »
which gives « $N \geq 12.77406$ » or « $N_{min} = 13$ » for a power « $(1-\beta) \geq 0.99$ ».

SmartCow
an integrated infrastructure for increased research capability and innovation in the European cattle sector

Adjustment of animal numbers in experimentation

formula for N_{min}

u_p is the quantile of probability "**p**" for the Gaussian distribution *N*(0,1).

for α = 0.05 we have **$u_{1-\alpha}$ = 1.644854**

and for β = 0.01 or (1-β) = 0.99 we have **u_β = -2.326348**

The minimum « N » size allowing the desired « precision », ie « (1-β) ≥ 0.99 » is such that :

$$N \geq \left[\frac{\sigma_Y \,.\, (u_{1-\alpha} - u_\beta)}{delta}\right]^2$$

the parameters :
« α », « β », « N », « delta » and « σ_Y »
are well linked.

with « delta = μ_1 - μ_0 » for the minimum detectable deviation sought
and « σ_Y » for the standard deviation of the Gaussian law of the studied and measured variable Y.

Adjustment of animal numbers in experimentation

formula for N_{min}

u_p is the quantile of probability "**p**" for the Gaussian distribution *N*(0,1).

for $\alpha = 0.05$ we have **$u_{1-\alpha} = 1.644854$**
and for $\beta = 0.01$ or $(1-\beta) = 0.99$ we have **$u_\beta = -2.326348$**

The minimum « N » size allowing the desired « precision », ie « $(1-\beta) \geq 0.99$ » is such that :

$$N \geq \left[\frac{\sigma_Y \cdot (u_{1-\alpha} - u_\beta)}{delta}\right]^2$$

with « delta = $\mu_1 - \mu_0$ » for the minimum detectable deviation sought
and « σ_Y » for the standard deviation of the Gaussian law of the studied and measured variable Y.

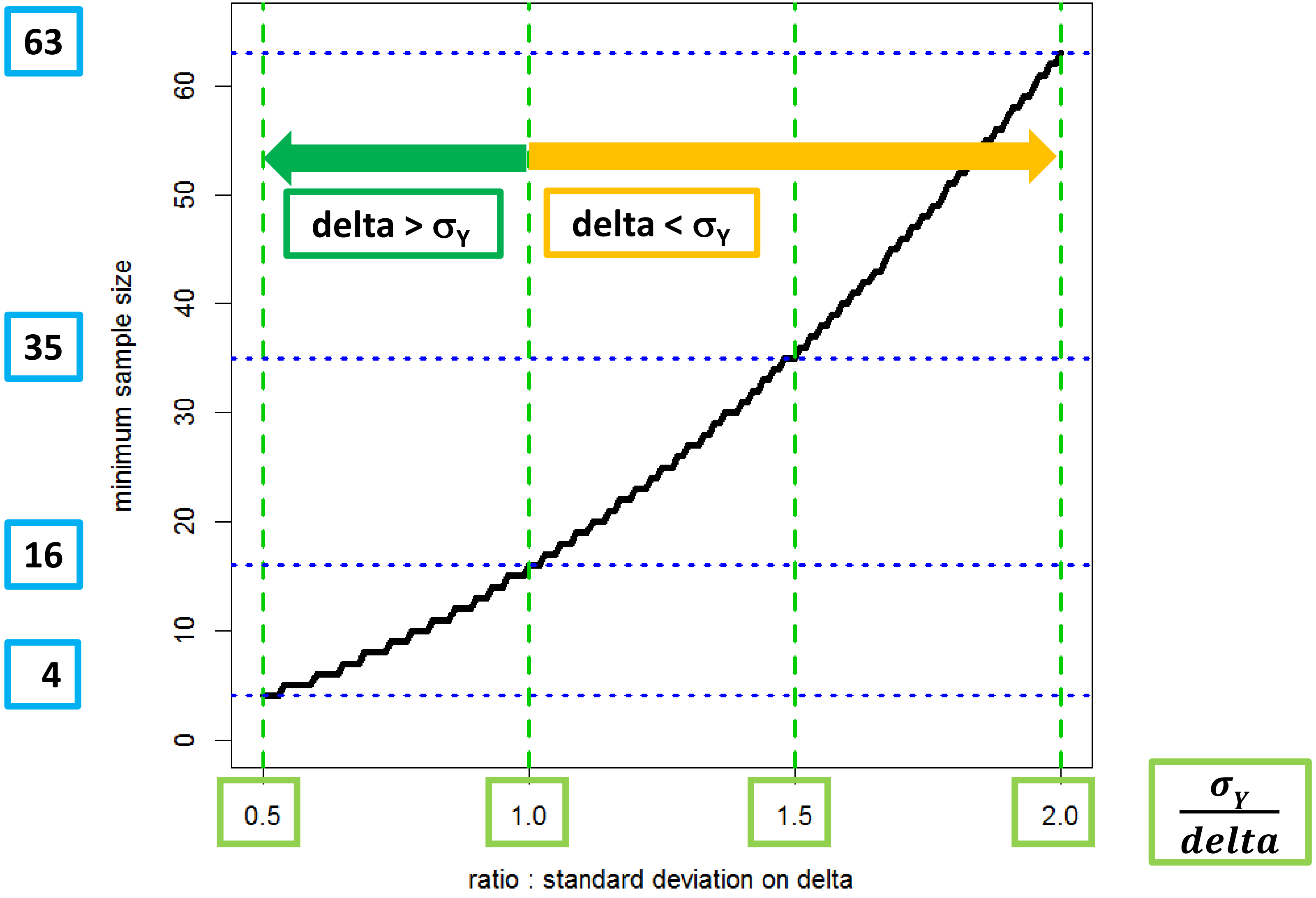
Note : as long as « delta > σ_Y » the minimum number will remain « reasonable ».
On the other hand in the opposite case, the minimum workforce may quickly « explode »
because we will be in the case where we are looking
for a « sensitivity » lower than the « measurement accuracy ».

Adjustment of animal numbers in experimentation

$$N \geq \left[\frac{\sigma_Y \, . \left(u_{1-\alpha} - u_{\beta}\right)}{delta}\right]^2$$

for $\alpha = 0.05$
and for $\beta = 0.01$ or $(1-\beta) = 0.99$

N_{min}



« N_{min} » depends on the ratio « σ_Y / delta »

SmartCow
an integrated infrastructure for increased research capability and innovation in the European cattle sector

Adjustment of animal numbers in experimentation

$$N \geq \left[\frac{\sigma_Y \,.\, (u_{1-\alpha} - u_\beta)}{delta}\right]^2$$

for α = 0.05
and for β = 0.01 or (1-β) = 0.99

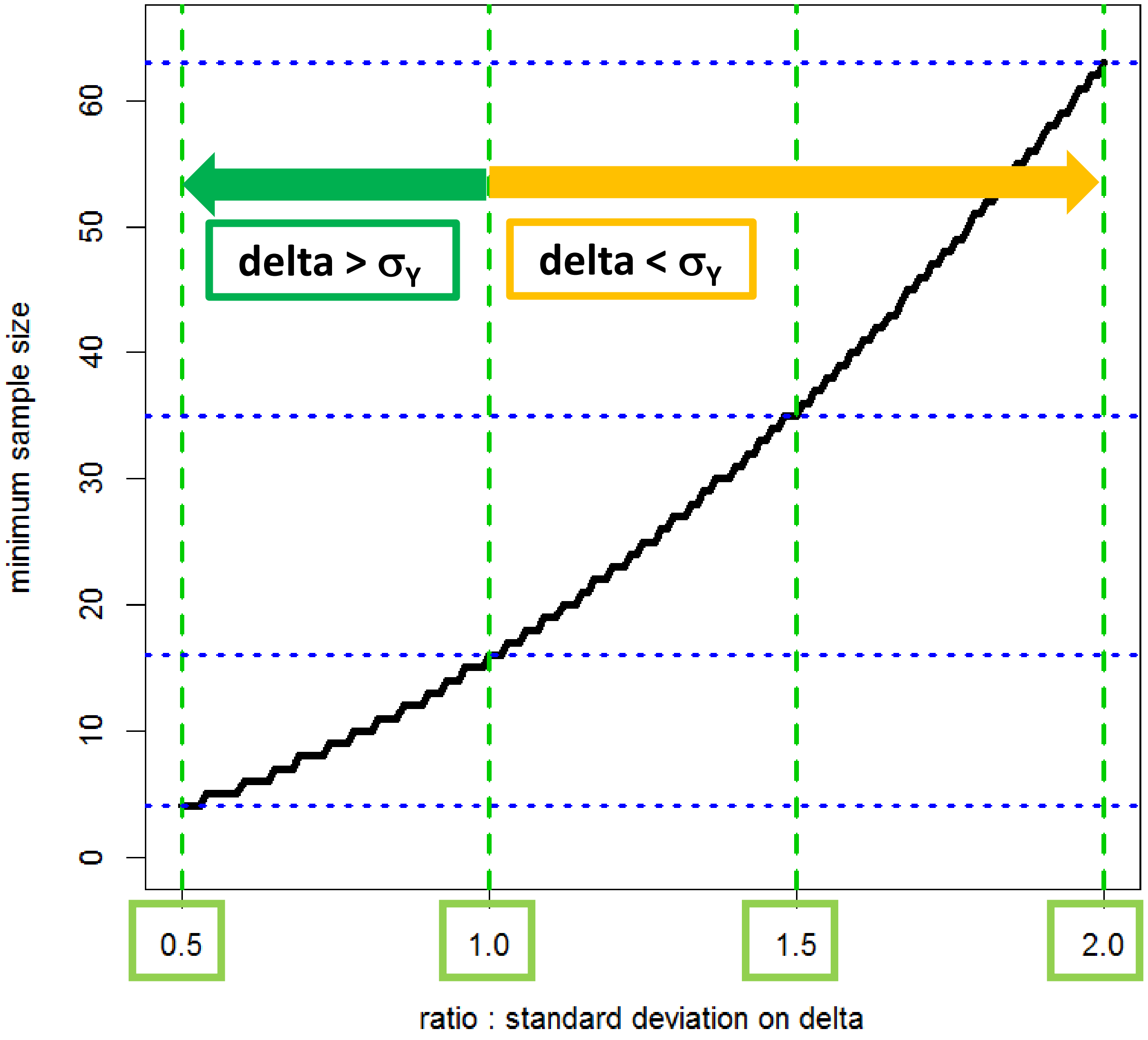
for a given sample size N a priori we can "correctly" identify a difference of at least

$$delta \geq \frac{\sigma_Y . (u_{1-\alpha} - u_\beta)}{\sqrt{N}}$$

for example, with N = 6 we have delta ≥ 1.6 σ_Y or (σ_Y / delta) ≤ 0.62

N_{min}

63
35
16
4



$$\frac{\sigma_Y}{delta}$$

SmartCow
an integrated infrastructure for increased research capability and innovation in the European cattle sector

Adjustment of animal numbers in experimentation

so with the 1st solution:

when we "increase" the size "N" of the sample
the variance of the estimator "decreases"
the risk of error β "decreases"
and the power of the test "increases".

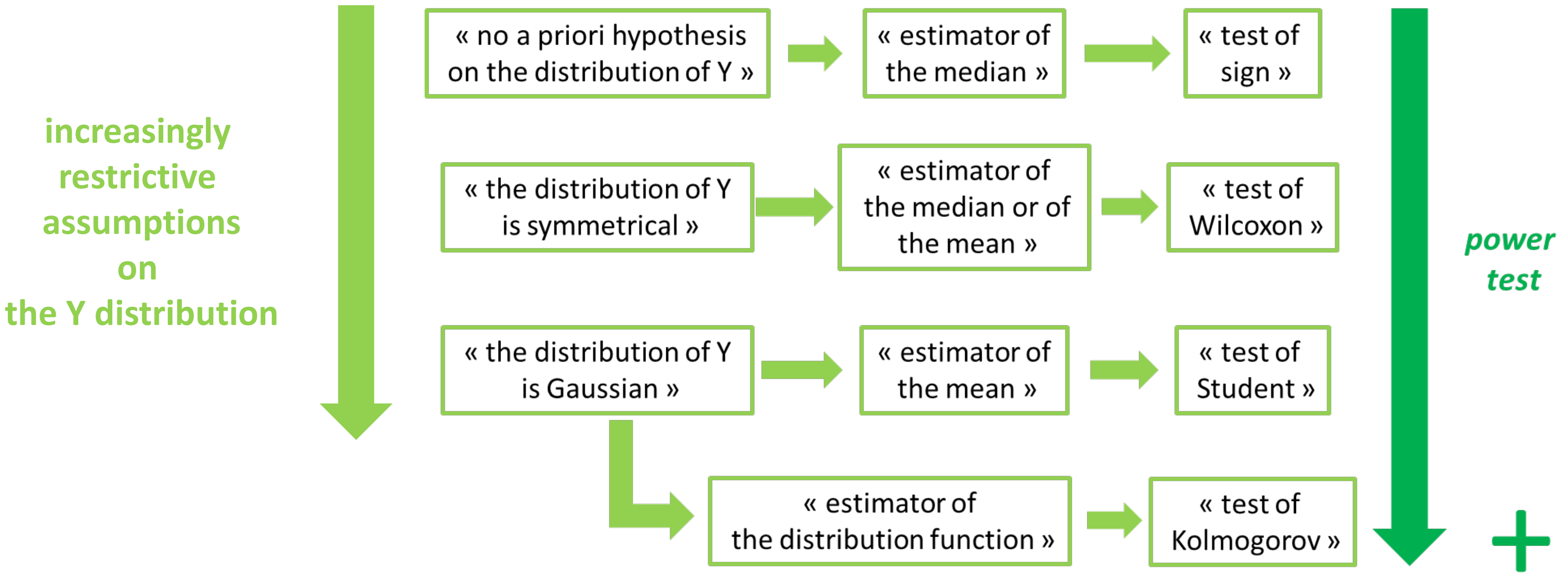
$$Var(\bar{Y}) = \frac{1}{N}.Var(Y) = \frac{1}{N}.\sigma_Y{}^2$$

therefore by choosing "N" according to the "minimum difference" (or "delta")
that we want to be able to highlight *a priori*,
we "select" *a priori* the "power of the test" that we want,
and therefore the "expected sensitivity" of the test is thus determined.

Adjustment of animal numbers in experimentation

2nd complementary solution: choose the most suitable test to the assumptions that can be made a priori on the distribution of the measured variable "Y".

for the same value of "delta" a priori, and the same sample size "N" a priori

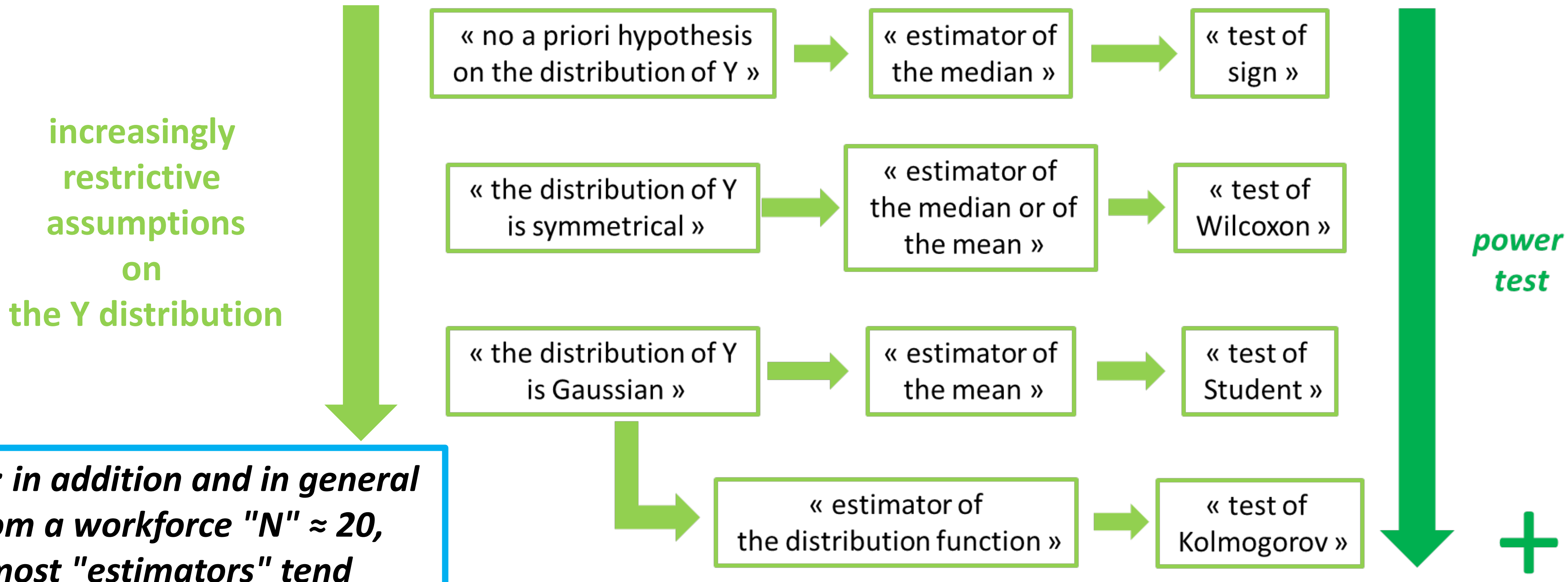


Adjustment of animal numbers in experimentation

SmartCow
an integrated infrastructure for increased research capability and innovation in the European cattle sector

2nd complementary solution: choose the most suitable test to the assumptions that can be made a priori on the distribution of the measured variable "Y“.

for the same value of “delta” a priori, and the same sample size “N” a priori



Note: in addition and in general from a workforce "N" ≈ 20, most "estimators" tend to a Gaussian distribution.

Adjustment of animal numbers in experimentation

in conclusion

Question: Why is this important?

Answer: choose the size of the sample N a priori
allows to "calibrate" a priori
the "desired sensitivity" of the statistical test to be used
taking into account a priori knowledge such as a minimum standard deviation σ_Y.

the smaller σ_Y
the smaller the sample size N required

“a statistician is not a magician”

Thank you for your attention.