UPOV

TC/32/6
ORIGINAL: English
DATE: September 11, 1995

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

GENEVA

TECHNICAL COMMITTEE

Thirty-second Session
Geneva, October 18 to 20, 1995

SEQUENTIAL ANALYSIS

Document prepared by the Office of the Chairman of the TWC with the help of experts from Denmark, Germany and the United Kingdom

SEQUENTIAL ANALYSIS

S GREGOIRE (France)
with the help of F LAIDIG (Germany), K KRISTENSEN (Denmark), M TALBOT (United Kingdom)

This document has been prepared by the TWC (Technical Working party on automation and Computer programs) for the Technical Committee.

The TWC was asked to present simply the principle of **sequential analysis** and to give an example.

This is done in three pages, this page plus the following two pages:
- PRINCIPLE OF THE SEQUENTIAL ANALYSIS METHOD (one page)
- ILLUSTRATION OF SEQUENTIAL ANALYSIS WITH AN EXAMPLE (one page)

In order to permit further investigation on the topic other parts have been added in the present document as an annex.
- GENERAL CONSIDERATIONS ON UPOV WORK (one page)
- COMPARISON OF DIFFERENT APPROACHES WITH EXAMPLES (7 pages)
- COMPLEMENTS CORRESPONDING TO THE DIFFERENT APPROACHES (9 pages)

GENERAL CONSIDERATIONS ON UPOV WORK (one page)

The general considerations on UPOV work is a page that considers the basic practice of the work carried out .
It is important to have in mind these elements when we think of methodology, but it is not necessary for the comprehension of the other parts of the document.
So this page may be read at any moment, or not read at all.

COMPARISON OF DIFFERENT APPROACHES WITH EXAMPLES (7 pages)

The comparison of different approaches with examples illustrate different ways to achieve the goal of a study.
Two examples are illustrating sequential analysis, four other examples illustrate other common practices such as "study during one year with a fixed sample size" for instance.
In this part the figures are given without reference on how they were computed.

COMPLEMENTS CORRESPONDING TO THE DIFFERENT APPROACHES (9 pages)

For those who wish to know how the figures for the different examples have been obtained, the complements corresponding to the different approaches give some more information.

PRINCIPLE OF THE SEQUENTIAL ANALYSIS METHOD

The **method originates from a practical observation.**
When we control objects; some of them are so good that we are able to accept them after a quick look, or so bad that with a quick look we are able to reject them. For other objects we need more information before to decide if they are good enough or not.
In other words, **we do not need to put the same effort on all the objects we control.**
As a result of this, we may save time or money if we decide for some objects at an early stage.

The general basis of sequential analysis is to define, -before the studies-, **sampling schemes which permit to take consistent[1] decisions at different stages of examination.**

The aim is to **compare**
"the **results** of examinations made on an object to control" **<-- to -->** "**decisional limits**"
in order to take a decision.

At each stage of examination the decision is
to accept the object if it is good enough with a chosen probability for acceptation
or **to reject** the object if it is too bad at a chosen probability for rejection
or **to continue** the study if we need more information before to fall in one of the two above categories.

At each stage **a pre-determined number of objects are examined** or measured. A decision rule indicates in which of the three above situations we are; according to the total number of objects examined so far.

Names given to different sampling schemes:
If the sampling method is defined with two stages it is called a "double" or "two-stages" sampling scheme.
If the sampling method is defined with more than two stages it is called "multiple" sampling scheme.
If at each new stage only one new object is observed the sampling scheme is called "progressive".
NB: In the case of a study in one stage with a pre-determined number of objects to examine, the sampling scheme is called a "fixed size" sampling scheme. This is not a sequential sampling scheme.

To be in the case of a sequential testing the following is required:
- there is more than one stage in the study
- the number of elements to observe is defined in advance for each stage
- the rule of decision at each stage is defined in advance
- the rule of decision at a given stage S is based on all the elements observed from stage one to this given stage S.

On the contrary a two stage study in which there are two independent tests with a decision for each stage; plus a decision rule which take into account the elementary decisions is not a sequential test.
(Example: we study for three years varieties for homogeneity and give a label "homogeneous" or "not enough homogeneous" after each year. Then we decide that the variety is homogeneous if it is homogeneous two years out of three. This is not sequential testing)

---
[1] "consistent decisions" mean that the risks to make a wrong decison is approximately the same at each stage of the study.

23

ILLUSTRATION OF SEQUENTIAL ANALYSIS EXAMPLE

goal: *We study varieties to check if the percentage of off-types is not too high.*

idea: *use an appropriate sequential test and the corresponding decision rule.*

Application of the idea: *We observe plants for different characters and count the number of off-types in our sample.*
We observe 90 plants in a first stage and take a decision for the samples with low or high number of off-types.
For the samples with an intermediate number of off-types, we examine 60 more plants and take our decision on the basis of the number of off-types in the 150 plants observed for the variety (90+60=150).

Decision rule in this example:

First stage: 90 plants have been examined, we have a number of off-types as a result of examination.
- *if there is 0 or 1 off-type we accept the variety*
- *if there are 2 or 3 off-types we continue the study with 60 more plants*
- *if there are 4 off-types or more we reject the variety (too much off-types)*

Second stage: (only if we have 2 or 3 off-types in the first 90 plants examined). 60 more plants are examined. We have a total number of off-types for the 150 plants observed.
- *We accept the variety if it has less than 4 off-types (out of 150 plants)*
- *We reject the variety if there is 4 off-types or more (out of 150 plants)*

Practical examples of application of the decision rule:

* Practical example for a sample of the variety "XXXXXX":
At first stage we observed 1 off-type out of 90 plants — *we accept the variety*

* Practical example for a sample of the variety "YYYYYY":
At first stage we observed 3 off-types — *we continue to stage 2*
At the second stage we observed 1 off-type — *we reject the variety*
(we have 4 off-types out of 150 plants)

* Practical example for a sample of the variety "ZZZZZZ":
At first stage we observed 3 off-types — *we continue to stage 2*
At the second stage we observed 0 off-type — *we accept the variety*
(we have 3 off-types out of 150 plants)

Should we use a classical one stage study or a two stages study ?(example):

If the preparation of our experiment is long and if the cost is not much affected by the number of plants to prepare, but our examination is quick and inexpensive; then we might prefer to use a one stage study. This could be the case for classical examinations in the field.
If the preparation of the sample is short and inexpensive, but the process for examination is laborious or expensive, or if the cost is much affected by the number of objects to examine; then we might prefer to use a two stage study, which will reduce the number of objects we examine. This could be the case for electrophoresis for instance.

NB: Sequential analysis is also applicable for other purposes than the control of the percentage of off-types. This is only an example.

GENERAL CONSIDERATIONS ON UPOV WORK

UPOV work is based upon examination of **plants or parts of plants**.
This is done

- to take **decisions on varieties** based on examination for **distinction, uniformity and stability,**
- to establish a **description form,**
- in order to give a **certificate of protection to the varieties.**

In most characteristics used, the values observed are influenced by the agro-climatic conditions. **UPOV aim is that decisions are likely to be the same for a given variety**

- if examination is done in **different countries**
- if the **years of examination differ.**

In all cases examinations are made on **samples** intended to be **representative of the variety** at the time of examination.
We never examine all the plants of a variety to take our decision.
(an examination of 100% of the objects produced is possible in the industry, for instance a control on each of the cars made in a factory. This is not possible for us).
The fact that the material received by the UPOV examiner is representative or is not representative of the variety is very important.

Using samples, we never have "the true value for the variety", but an estimation for the character we observe.
This estimation is generally a value which is the mean of the values observed.
This value is for instance "white flowers" for the colour of flowers or "12.45 cm" for a plant height.
In both examples the value seems very precise but in fact there is an uncertainty linked to the way we assessed this value.

For instance let us suppose we assess "White flowers" when more than 90% of the flowers are white.
If we observed 56 white flowers out of 60 examined we have 93.3% of white flowers.
If we make a new examination with another sample of the same variety we will not always see 56 white flowers out of 60; we may find 54, or 55, or 59 for instance.
This is more obvious in the case of a height. We will probably not find 12.45 cm if we observe the same character on another sample of the same variety.

The **quality of our estimation** is affected by

- the **representativeness** of our sample,
- the **variability** of the character observed,
- the **number of elements** we observe.

Very generally the three above elements affect the quality of the estimation as follow:
- If the sample is not representative of the variety, we may have a good precision of our estimation but this estimation is good for the specific sample, not for the variety.
- The less the character is variable for a variety the more precise is our estimate.
- The more elements we observe for a variety the more precise is our estimate.

The representativeness is generally assumed.

The amount of variability is usually a "biological fact" but we may have some control on it. For instance we may try to find locations and cultural practices which avoid to create heterogeneity within a variety. We may also take into account the variability to take our decisions.

It is the case for instance in the UPOV COYD method in which we are more prudent to assess distinction with characters which are variable among years than with characters which are consistent form one year to the other.

The number of elements to observe is described in UPOV recommendations.
The number of elements we observe, combined with the number of studies we do (the number of years most often) is very important for the precision of our estimation. It is also the easier element to control or change.

COMPARISON OF DIFFERENT APPROACHES WITH EXAMPLES
(sequential analysis among other methods)

1-Introduction:

in this document our goal is the following:

We want to check varieties for off-types.
We decide that
- the good varieties must not have more than 1% of off-types.
- we would like to avoid varieties having more than 5% of off-types.

We will in the following pages examine different ideas to organise a technical study in order to achieve this goal.

We will imagine different number of plants to observe (sampling scheme) and different ways to decide if we accept or reject a variety (decision rules).

We will imagine four different cases (-A- to -D) in which we try to define something that could fulfil our goal "from a common sense point of view".

For instance if our ***goal*** *is to check if there is 1 percent or less of white flowers in a pink variety, a "common sense" idea could be to observe one hundred plants and to count the white flowers. That is our* ***sampling scheme and observation****.*
Then we can decide to accept the variety if it has no white flower, and to reject the variety if it has more than 1 white flowers. That is our ***decision rule****.*

Then we will look at two cases (-E- and -F-) in which we define more precisely our goal and ask for a corresponding sampling technique to a computer program.

We can imagine many cases. The cases proposed here are only examples. They are not better or worse than other ideas.
For cases -A- to -D- the sample sizes and decision rules are arbitrary chosen by us. They are only the results of our thoughts, trying to find something which seems sensible.
For cases -E- and -F- a computer program will help us to choose which sample size and which decision rule to apply if we want to achieve our goal.

 23?

2-Presentation of the 6 example cases:

For the first two cases (-A-,-B-) the idea is to take a value mentioned in our goal and to imagine a simple decision rule as in the above example for white flowers.
In our goal we have two values
-1% of off-types for good varieties we wish to reject rarely (accept often)
-5% of off-types for bad varieties we wish to accept rarely (reject often)

A simple rule is to accept if in the sample we have up to 1% of off-types. That is the case -A- in which we will accept if there is 0,1 or 2 off-types out of 200 plants (2/200=1%).

An other simple rule is to reject if we have more than 5% of off-types. That is the case -B- in which we will reject if there is more than 10 off-types (10/200=5%).

For the case -C- the above ideas are kept but we try to make a decision after 50 plants. We accept those having very few off-types and reject those having a great number of off-types. Then we observe 150 more plants for the cases in which we were did not take decision after 50 plants.

For the case -D- the idea is to do the study two times (two years for instance) and to decide on the basis of the results obtained from each study.
We will observe 100 plants each time (100+100=200 at total).
If there is 0 or 1 off-type out of 100 plants we will tell that the variety is homogeneous for this study (1/100=1% as in case -A-)
After the two studies only, we will take a decision. If the result of the two studies is "homogeneous" one time or the two times we will accept the variety, otherwise we will reject it (= "heterogeneous" two times).

In case -E- and -F- we will be help by the computer to define a sample scheme and a decision rule.
In case -E- we will choose a one stage study. In case -F- we will choose a two-stage study.
case -E- is a fixed sample size sampling scheme
case -F- is a sequential test with a double sampling scheme.

For each case we will look at what will happen if we make 100 times our study on the same variety (using 100 representative samples). We will see the number of times we will accept or reject the variety if the true value of the percentage of off-types is 1% , and if the true value of the percentage of off-types is 5%.

NB: From all these cases only case -C-and case -F- are sequential tests.
Cases -A-, -B- and -E- are "fixed sample size" sampling schemes.
Case -D- is a two stage study but a decision is taken at each stage independently, then a decision rule is applied.

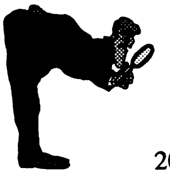
NB:
When the left part of the line is a black vertical line, the example is an example which refer to sequential analysis.

3-Results for the six example cases:

CASE -A-	IDEA

We take into account our objective to accept varieties having 1% of off-types or less.
We decide to accept the variety if the percentage of off-types is less or equal to 1 percent of the plants examined.

APPLICATION OF THE IDEA	-A-



200
2 OK 3 reject

We decide to examine 200 plants and

- to accept varieties if there is 0,1 or 2 off-types (2/200=1%).
- to reject the variety if there are 3 off-types or more.

WHAT WILL HAPPEN THEN ?	-A

if the true percentage of off-types is 1%
and we examine 100 times (100 samples) the same variety
we will accept it 68 times
and reject it 32 times

if the true percentage of off-types is 5%
and we examine 100 times (100 samples) the same variety
we will accept it 0 times
and reject it 100 times

CASE -B-	IDEA

We take into account our objective to accept varieties having 5% of off-types or less.
We decide to accept the variety if the percentage of off-types is less or equal to 5 percent of the plants examined.

APPLICATION OF THE IDEA	-B-

200
10 OK 11 reject

We decide to examine 200 plants and

- to accept varieties if there is 0 to 10 off-types (10/200=5%).
- to reject the variety if there are 11 off-types or more.

WHAT WILL HAPPEN THEN ?	-B-

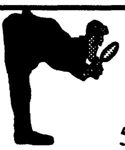
if the true percentage of off-types is 1%
and we examine 100 times (100 samples) the same variety
we will accept it 100 times
and reject it 0 times

if the true percentage of off-types is 5%
and we examine 100 times (100 samples) the same variety
we will accept it 58 times
and reject it 42 times

Case -C- **IDEA**

**We decide to make a two-stage study,
in which we take a decision for some cases
and continue in other cases.**

APPLICATION OF THE IDEA -C-



50
0 OK/1-5 continue / 6 reject

150
10/200 OK 11/200 reject

We decide
- to study first 50 plants and to accept if there are no off-types, to reject if there are 6 off-types or more

- to look at 150 other plants if there is 1 to 5 off-types, then to reject the variety if more than 10 plants are off-type out of 200 plants

WHAT WILL HAPPEN THEN? -C-

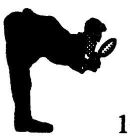
if the true percentage of off-types is 1%
and we examine 100 times (100 samples) the same variety
we will accept it 100 times.
we will reject it 0 times

if the true percentage of off-types is 5%
and we examine 100 times (100 samples) the same variety
we will accept it 59 times
and reject it 41 times

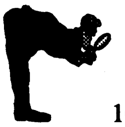
Case -D- **IDEA**

**we decide to observe 100 plants,
then to always make another study on 100 other plants.**

APPLICATION OF THE IDEA -D-



100
1 homo 2 hetero



100
1 homo 2 hetero

We declare the variety homogeneous in one study if the number of off-types is 0 or 1 out of 100 plants and heterogeneous if there are 2 off-types or more.

We reject the variety if it is heterogeneous the two times.
We accept the variety if it is homogeneous in the two studies, or heterogeneous in only one study.

WHAT WILL HAPPEN THEN? -D-

if the true percentage of off-types is 1%
and we examine 100 times (100 samples * 2 studies) the same variety
we will accept it 93 times
and reject it 7 times.

if the true percentage of off-types is 5%
and we examine 100 times (100 samples * 2 studies) the same variety
we will accept it 7 times
and reject it 93 times

CASE -E- **IDEA**

- the good varieties must not have more than 1% of off-types.
- we would like to avoid varieties having more than 5% of off-types.
- we want to reject varieties which in fact have 1% of off-types only 5 times out of 100 samples of the same variety.
- we want to accept varieties which in fact have 5% of off-types only 10 times out of 100 samples of the same variety.

APPLICATION OF THE IDEA -E-

130
3 OK 4 reject

We want to make our study in one stage

Computations tells us that the more appropriate way to achieve our goal is to observe 130 plants and to reject a variety which have 4 off-types or more.

WHAT WILL HAPPEN THEN? -E-

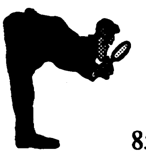
if the true percentage of off-types is 1%
and we examine 100 times (100 samples) the same variety
we will accept it 96 times
and reject it 4 times

if the true percentage of off-types is 5%
and we examine 100 times (100 samples) the same variety
we will accept it 10 times
and reject it 90 times

case -F- **IDEA**

- the good varieties must not have more than 1% of off-types.
- we would like to avoid varieties having more than 5% of off-types.
- we want to reject varieties which in fact have 1% of off-types only 5 times out of 100 samples of the same variety.
- we want to accept varieties which in fact have 5% of off-types only 10 times out of 100 samples of the same variety.

APPLICATION OF THE IDEA -F-



85
1 OK/ 2-3 continue/ 4 reject

60
3/145 OK 4/145 reject

We want to make our study in two stages

Computations tells us that an appropriate way to achieve our goal is to observe 85 plants in the first stage and 60 more plants in the second stage.
We will
-accept the variety if it has 0 or 1 off-type at the first stage,
-reject the variety if is has 4 off-types or more at the first stage,
-continue for the second stage in other cases
if a second stage is done we will reject the variety if it has (at total on the two stages) 4 off-types or more.

WHAT WILL HAPPEN THEN? -F-

if the true percentage of off-types is 1%
and we examine 100 times (100 samples) the same variety
we will accept it 95 times
and reject it 5 times

if the true percentage of off-types is 5%
and we examine 100 times (100 samples) the same variety
we will accept it 10 times
and reject it 90 times

4-Summary (If the same variety is examined 100 times (100 different samples of the same variety))

explanation of columns GA to BR:

column GA=good accepted =times the variety will be accepted if there is in fact 1% off-types
- if we often accept good varieties (having 1% off-types) we are happy ☺
- if we rarely accept good varieties (having 1% off-types) we are not happy☹

column GR=good rejected =times the variety will be rejected if there is in fact 1% off-types
- if we often reject good varieties (having 1% off-types) we are not happy ☹
- if we rarely reject good varieties (having 1% off-types) we are happy ☺

column BA=bad accepted=times the variety will be accepted if there is in fact 5% off-types
- if we often accept bad varieties (having 5% off-types) we are not happy ☹
- if we rarely accept bad varieties (having 5% off-types) we are happy ☺

column BR=bad rejected =times the variety will be rejected if there is in fact 5% off-types
- if we often reject bad varieties (having 5% off-types) we are happy ☺
- if we rarely reject bad varieties (having 5% off-types) we are not happy ☹

Table 1

case	Idea	Application	decision	GA	GR	BA	BR
-A-	accept if the % observed is less than 1%	observe 200 plants	-accept if 0 or 1 off-type -reject if 2 off-types or more	68/100 ☹	32/100 ☹	0/100 ☺	100/100 ☺
-B-	accept if the % observed is less than 5%	observe 200 plants	-accept if 0 to 10 off-types -reject if 11 off-types or more	100/100 ☺	0/100 ☺	58/100 ☹	42/100 ☹
-C-	Observe on a limited number of plants, and on more plants if necessary	first: 50 plants then 150 more plants if necessary and decision on 200 plants (50+150)	first 50 plants -accept if no off-types -reject if 6 off-types or more -if 1 to 5 off-types continue if necessary 150 more plants -accept if 0 to 10 off-types out of 200 plants -reject if > 10 off-types out of 200 plants	100/100 ☺	0/100 ☺	59/100 ☹	41/100 ☹
-D-	Observe two times	first time 100 plants second time 100 plants	declare homogeneous for one study if 0 or 1 off-type -accept if 1 or 2 times homogeneous -reject if 2 times heterogeneous	93/100 ☺	7/100 ☺	7/100 ☺	93/100 ☺
-E-	Observe one time	observe 130 plants	-accept if 0 to 3 off-types -reject if 4 off-types or more	96/100 ☺	4/100 ☺	10/100 ☺	90/100 ☺
-F-	Observe on a limited number of plants, and on more plants if necessary	first 85 plants then 60 more plants if necessary	first 85 plants -accept if 0 or 1 off-types -reject if 4 off-types or more -if 2 or 3 off-types continue if necessary 60 more plants -accept up to 4 off-types out of 145 plants -reject if 4 off-types or more (145 plants)	95/100 ☺	5/100 ☺	10/100 ☺	90/100 ☺

If we take line -A- of table 1
we might think that we are happy that we almost never accept a variety which has 5% off-types.
But we might think that it is not correct to reject 32 times out of 100 a variety which has 1% of off-types.

If we take line -B- of table 1
We might think that we are happy that we "always" accept a variety which has 1% off-types.
But we might think that it is not correct to accept 58 times out of 100 a variety which has 5% of off-types.

If we take line -C- of table 1
We might think that we are happy to "always" accept a variety which has 1% off-types.
But we might think that it is not correct to accept 59 times out of 100 a variety which has 5% of off-types.

If we take line -D- of table 1
We might think that we are happy to often (93/100) accept a variety which has 1% off-types.
We might think that we are happy to often (93/100) reject a variety which has 5% off-types.

If we take line -E- of table 1
We might think that we are happy to often (95/100) accept a variety which has 1% off-types.
We might think that we are happy to often (90/100) reject a variety which has 5% off-types.

If we take line -F- of table 1
We might think that we are happy to often (95/100) accept a variety which has 1% off-types.
We might think that we are happy to often (90/100) reject a variety which has 5% off-types.

-5 Conclusions:

We see from these examples

- that we are not able "by common sense" to guess the adequation of our ideas to our goal .

- that having imagined a way to work we are able to compute the figures and think if we are happy or not with the values that indicate the adequation to our goal.

- that computer programs are able to help us to find an appropriate solution.

- that no method (1 year fixed sample size, two years fixed sample size, two years sequential test,...) is better or worse than an other. For each method we may find a solution which is appropriate to our goal, for each method we can choose a solution which in fact is not appropriate to our goal.

NB:
It is only by chance that the results are in accordance with our goal in case -D-. If we would like to take a decision after one year of test with our rule the risks would not be in accordance with our goal in case -D-.

In case -F- the risks will be in accordance with our goal after the first step and also after the second step. It is an advantage of sequential analysis versus the case -D-

COMPLEMENTS CORRESPONDING TO THE DIFFERENT EXAMPLES

The figures are outputs of Qalstat. Qalstat is a computer program which help to define sample schemes, in particular with sequential testing, but also for fixed sample size.

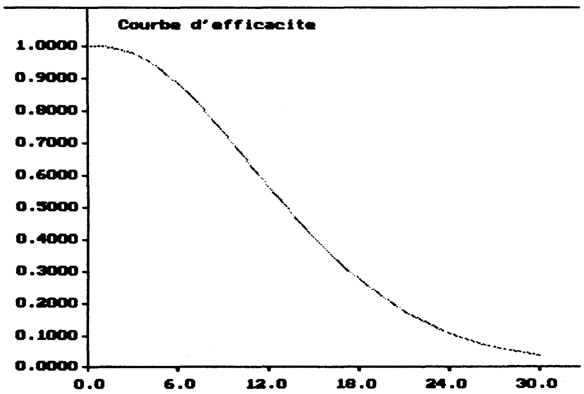
Two types of figures are given:
- curves
- tables

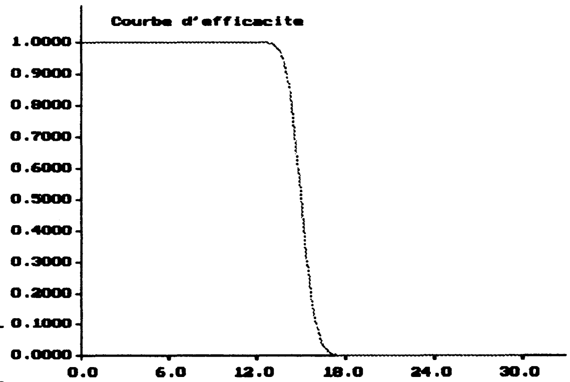
The curves show the probability to accept a sample for a given true percentage of off-types in a variety.
This probability vary from 0.0000 (never accept) to 1.0000 (always accept) on the Y axis.
The true percentage of off-types in the variety vary in the examples form 0.0 (0% of off-types) to 30.0 (30% of off-types) on the X axis.

A curve is specific of a given sampling scheme. All the curves have the same shape and the same scale, but the values are different from one example to the other.

General shape of the curves:

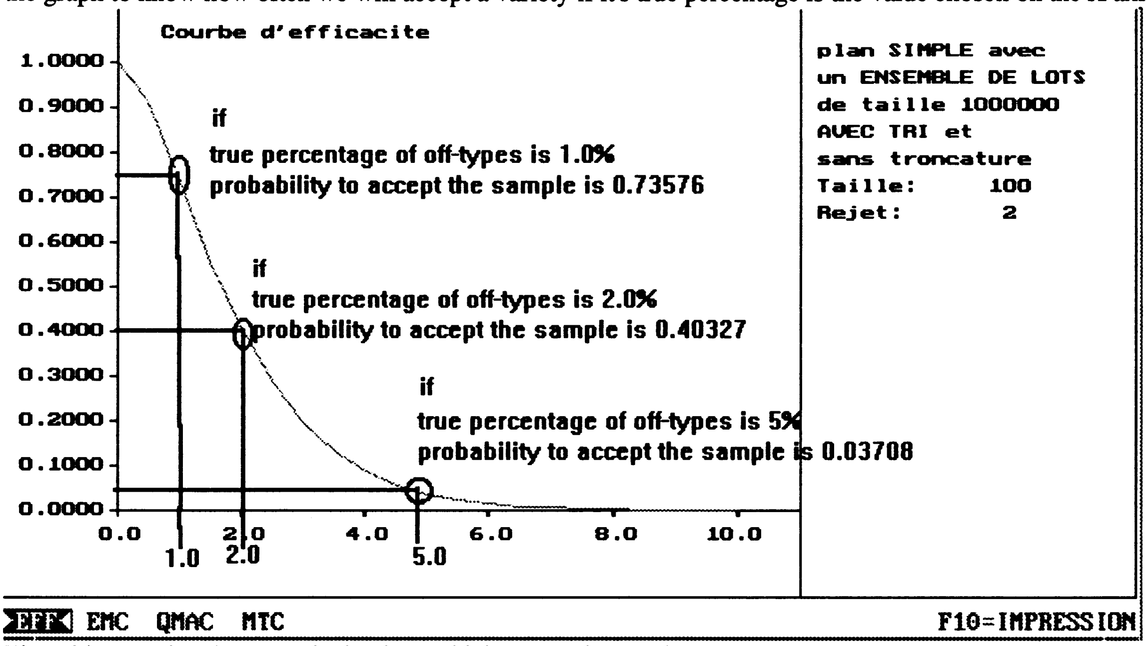
All sampling schemes start at the point X=0.0 Y=1,0000 (upper left point of the curve)
That mean that if the variety has no off-type at all, we will never find off-types and then we will always accept the variety.
On the oter end of the curve there is a common point X=100.0 Y=0.0000 (not seen on the graphs)
This point is not on the curves because the scale end at 30.0 (30% of off-types in the variety)
That point mean that if there is 100% of off-types we will find only off-types and then always reject the variety.
In between
- the less off-types will be present in the variety, the more often we will accept it (upper part of the curve)
- if the percentage of off-types increase, we are less likely to accept the variety (lower part of the curve)





The more plants we will observe, the more abrupt will be the decrease of the curve. In the two above examples we examine 20 plants (left curve) or 2100 plants (right curve).
Visually we see that
with 20 plants
we will "always" accept (Y near 1.0000) varieties having in fact 0% of off-types
we will "always" reject (Y near 0.0000) varieties ahaving in fact 30% of off-types or more
with 2100 plants
we will "always" accept (Y near 1.0000) varieties having in fact 0% to 14% of off-types
we will "always" reject (Y near 0.0000) varieties ahaving in fact more than 16% of off-types
That also mean that we are able to work more precisely if we observe 2100 plants than if we observe 20 plants.

The Y axis is the probability to accept a sample
The X axis is the true percentage of off-types in the variety. We never know this percentage. But we are able with the graph to know how often we will accept a variety if it's true percentage is the value chosen on the X axis.



The table contains the numerical values which are on the graph
the first column (% non conf.) is the equivalent of the X axis (true percentage of off-types)
the second column (Efficacité) is the equivalent of the Y axis (probability to accept the sample)

% non confo.	Efficacité	Eff.Moy.Contrô.	Qual.Moy.A.C	Moy.Tot.Contrô.
0.000	1.00000	100.0	0.000	100.0
1.000	0.73576	100.0	0.736	264311.6
2.000	0.40327	100.0	0.807	596768.6
3.000	0.19462	100.0	0.584	805397.3
4.000	0.08716	100.0	0.349	912845.4
5.000	0.03708	100.0	0.185	962922.5
6.000	0.01517	100.0	0.091	984830.4
7.000	0.00601	100.0	0.042	993987.7
8.000	0.00232	100.0	0.019	997680.9
9.000	0.00087	100.0	0.008	999126.8
10.000	0.00032	100.0	0.003	999678.4

ESC Sortie PGU Page précédente PGD Page suivante

If we take the value 1,000 (1% of off-types is the true percentage of off-types in the variety), then 0.73576 is the probability to accept the sample with this sampling scheme
The corresponding sentence is:

"if the true percentage of off-types is 1% and we examine 100 times (100 samples) the same variety we will accept it 74 times and reject it 26 times"

74 times is 0.73576 rounded to 74/100
26 times is the probability to reject which is (100-74)=26

In the examples 1% and 5% of off-types are taken as illustrations because they are the limits mentionned in the goal. In fact we can take any value we wish as the true percentage of off-types and find the corresponding probability. For instance for 2 % of off-types we would accept the sample 40 times out of 100 (0.40327)

CASE -A-

IDEA

We take into account our objective to accept varieties having 1% of off-types or less. We decide to accept the variety if the percentage of off-types is less or equal to 1 percent of the plants examined.

APPLICATION OF THE IDEA

case 1

We decide to examine 200 plants and

to accept varieties if there is 0,1 or 2 off-types.

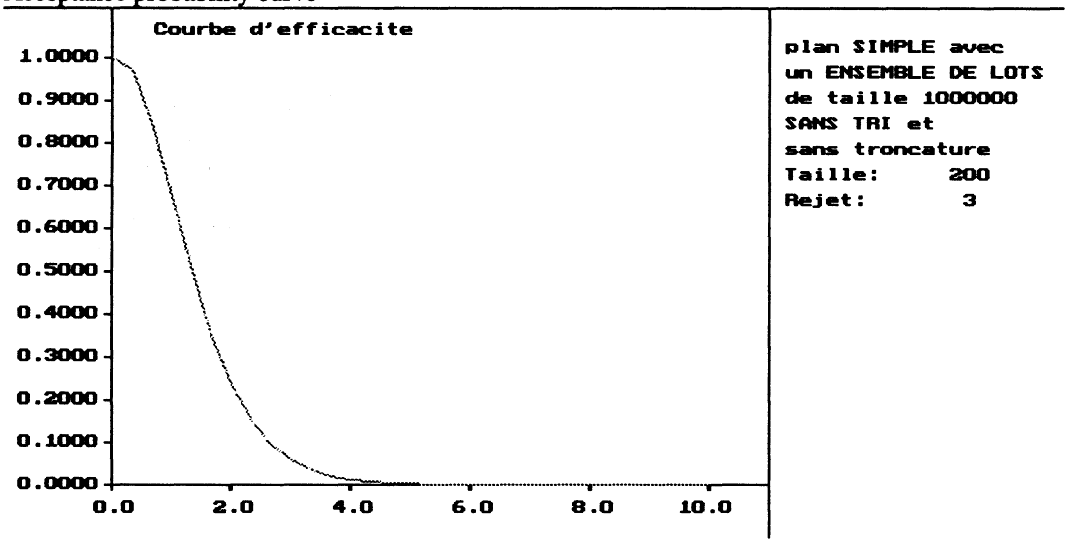
to reject the variety if there are 3 off-types or more.

WHAT WILL HAPPEN THEN ? -A-

if the true percentage of off-types is 1% and we examine 100 times (100 samples) the same variety we will accept it 68 times and reject it 32 times

if the true percentage of off-types is 5% and we examine 100 times (100 samples) the same variety we will accept it 0 times and reject it 100 times

Acceptance probability curve



corresponding values:

% off-types % non confo.	prob.to accept the variety Efficacité
0,000	1,00000
1,000	0,67668
2,000	0,23515
3,000	0,05929
4,000	0,01249
5,000	0,00234
6,000	0,00040
7,000	0,00006
8,000	0,00001
9,000	0,00000
10,000	0,00000

CASE -B-

IDEA

We take into account our objective to accept varieties having 5% of off-types or less. We decide to accept the variety if the percentage of off-types is less or equal to 5 percent of the plants examined.

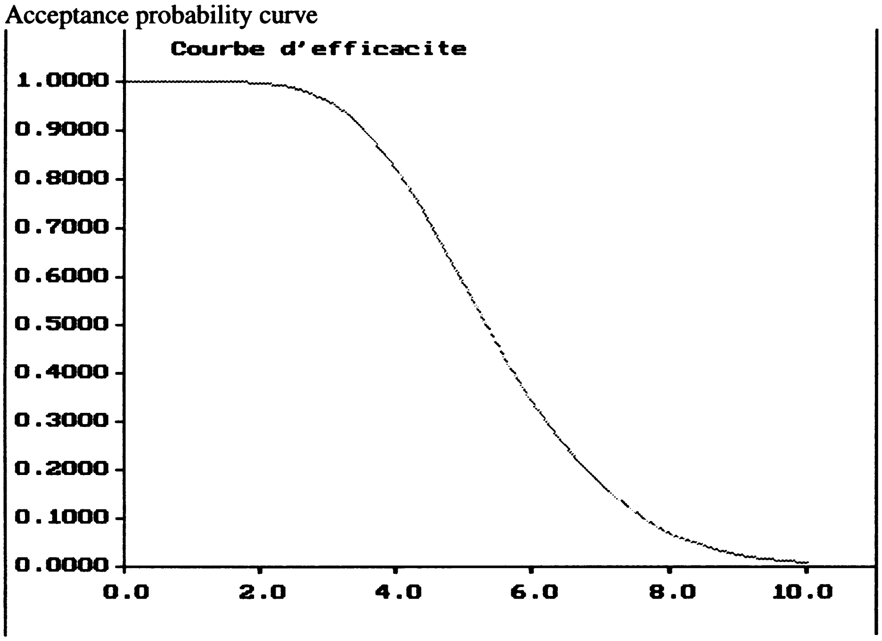
APPLICATION OF THE IDEA

We decide to examine 200 plants and
to accept varieties if there is 0 to 10 off-types.
to reject the variety if there are 11 off-types or more.

WHAT WILL HAPPEN THEN ? -B-

if the true percentage of off-types is 1% and we examine 100 times (100 samples) the same variety we will accept it 100 times and reject it 0 times
if the true percentage of off-types is 5% and we examine 100 times (100 samples) the same variety we will accept it 58 times and reject it 42 times

Acceptance probability curve



plan SIMPLE avec
un ENSEMBLE DE LOTS
de taille 1000000
SANS TRI et
sans troncature
Taille: 200
Rejet: 11

corresponding values:
% off-types prob.to accept the variety

% non confo.	Efficacité
0,000	1,00000
1,000	0,99999
2,000	0,99747
3,000	0,95987
4,000	0,81998
5,000	0,58307
6,000	0,34071
7,000	0,16613
8,000	0,06913
9,000	0,02507
10,000	0,00807

CASE -C-

IDEA

**We decide to make a two-stage study,
in which we take a decision for some cases
and continue in other cases.**

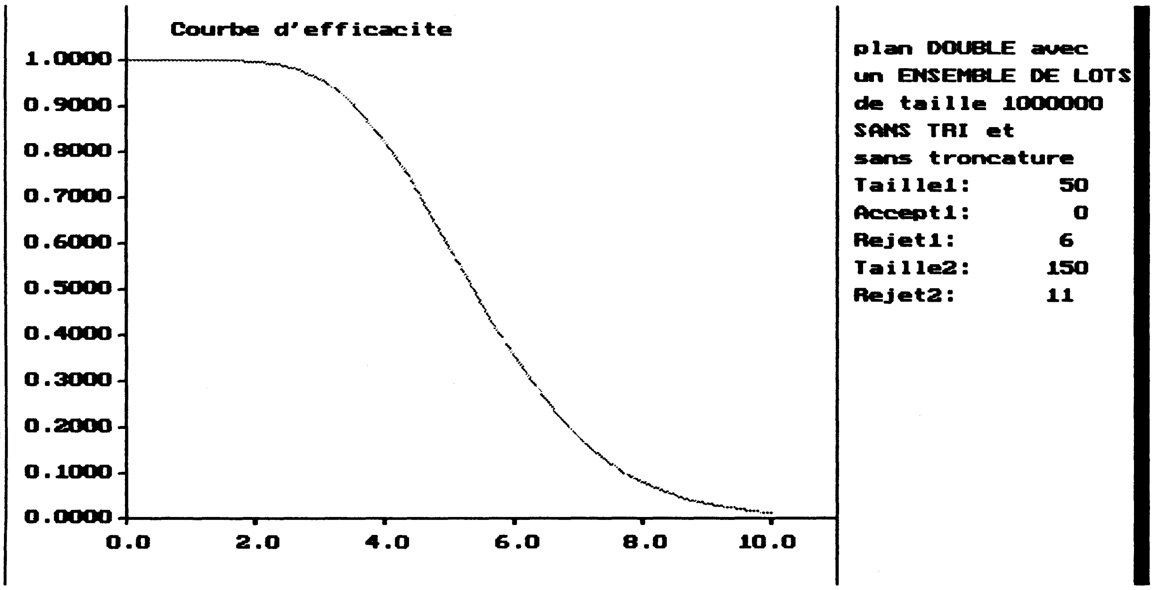
APPLICATION OF THE IDEA

We decide
- to study first 50 plants and to accept if there are no off-types, to reject if there are 6 off-types or more
- to look at 150 other plants if there is 1 to 5 off-types, then to reject the variety if more than 10 plants are off-type out of 200 plants

WHAT WILL HAPPEN THEN ? -C-

if the true percentage of off-types is 1% and we examine 100 times (100 samples) the same variety we will accept it 100 times and reject it 0 times.
if the true percentage of off-types is 5% and we examine 100 times (100 samples) the same variety we will accept it 59 times and reject it 41 times.

Acceptance probability curve



corresponding values:
% off-types prob.to accept the variety

% non confo.	Efficacité
0,000	1,00000
1,000	0,99998
2,000	0,99718
3,000	0,95932
4,000	0,82161
5,000	0,58946
6,000	0,35103
7,000	0,17730
8,000	0,07861
9,000	0,03196
10,000	0,01258

CASE -D-

IDEA

we decide to observe 100 plants, then to always make another study on 100 other plants.

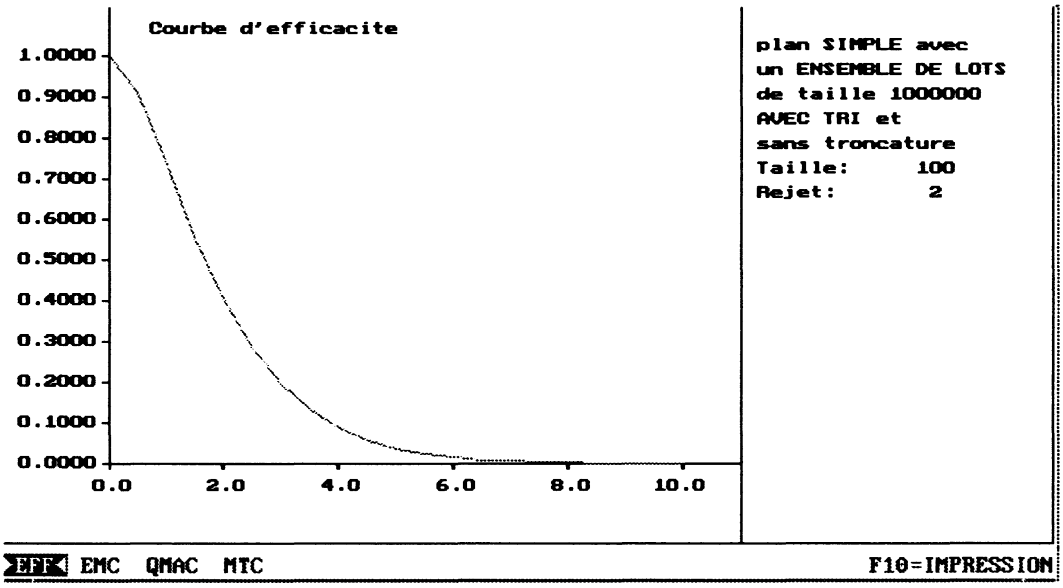
APPLICATION OF THE IDEA

We declare the variety homogeneous in one study if the number of off-types is 0 or 1 out of 100 plants and heterogeneous if there are 2 off-types or more.
Then the variety is rejected if it is heterogeneous the two times
We accept the variety if it is homogeneous in the two studies, or heterogeneous in only one study

WHAT WILL HAPPEN THEN -D-

if the true percentage of off-types is 1% and we examine 100 times (100 samples * 2 studies) the same variety we will accept it 93 times and reject it 7 times.
if the true percentage of off-types is 5% and we examine 100 times (100 samples * 2 studies) the same variety we will accept it 7 times and reject it 93 times.

Acceptance probability curve



corresponding values:
% off-types prob.to accept the variety

% non confo.	Efficacité
0,000	1,00000
1,000	0,73576
2,000	0,40327
3,000	0,19462
4,000	0,08716
5,000	0,03708
6,000	0,01517
7,000	0,00601
8,000	0,00232
9,000	0,00087
10,000	0,00032

NB: the above table give the figure for ONE independant single test.
We can see that for one trial, the probabilities are not in accordance with our goal. We would reject 26 times out of 100 a variety having 1% of off-types (100%-0,73576).
To compute the probabilities after two years we apply the formulas described in TWC/13/8. This correspond to a probability to accept the variety of 93% if there is 1% of off-types (93%=1-((1-0.73)*(1-0.73))
and of 7% if there is 5% of off-types (7%=0.037*(2-0.037))

CASE -E-

IDEA

- the good varieties must not have more than 1% of off-types.
- we would like to avoid varieties having more than 5% of off-types.
- we want to reject varieties which in fact have 1% of off-types only 5 times out of 100 samples of the same variety.
- we want to accept varieties which in fact have 5% of off-types only 10 times out of 100 samples of the same variety.

APPLICATION OF THE IDEA

We want to make our study in one stage

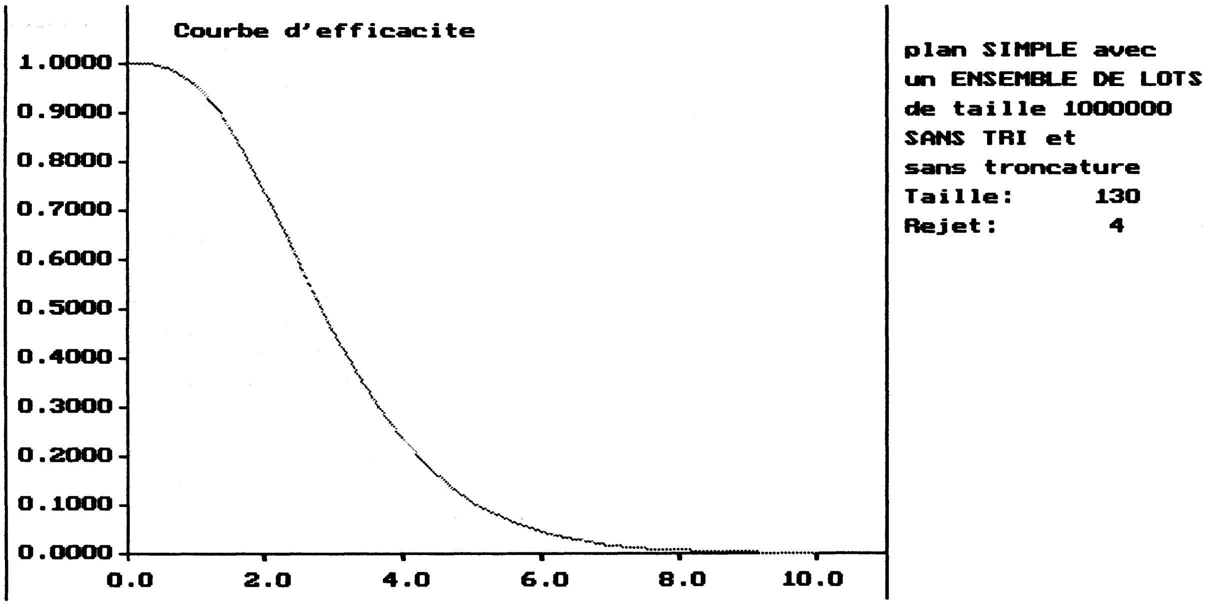
Computations tells us that the more appropriate way to achieve our goal is to observe 130 plants and to reject a variety which have 4 off-types or more

WHAT WILL HAPPEN THEN -E-

if the true percentage of off-types is 1% and we examine 100 times (100 samples) the same varietywe will accept it 96 times and reject it 4 times.

if the true percentage of off-types is 5% and we examine 100 times (100 samples) the same variety we will accept it 10 times and reject it 90 times

Acceptance probability curve



corresponding values:

% off-types % non confo.	prob.to accept the variety Efficacité
0,000	1,00000
1,000	0,95776
2,000	0,73689
3,000	0,45050
4,000	0,23227
5,000	0,10576
6,000	0,04383
7,000	0,01686
8,000	0,00610
9,000	0,00209
10,000	0,00069

CASE -F-

IDEA

- the good varieties must not have more than 1% of off-types.
- we would like to avoid varieties having more than 5% of off-types.
- we want to reject varieties which in fact have 1% of off-types only 5 times out of 100 samples of the same variety.
- we want to accept varieties which in fact have 5% of off-types only 10 times out of 100 samples of the same variety.

APPLICATION OF THE IDEA

We want to make our study in two stages.

Computations tells us that an appropriate way to achieve our goal is to observe 85 plants in the first stage and 60 more plants in the second stage.

We will

-accept the variety if it has 0 or 1 off-type at the first stage,

-reject the variety if is has 4 off-types or more at the first stage,

-continue for the second stage in other cases

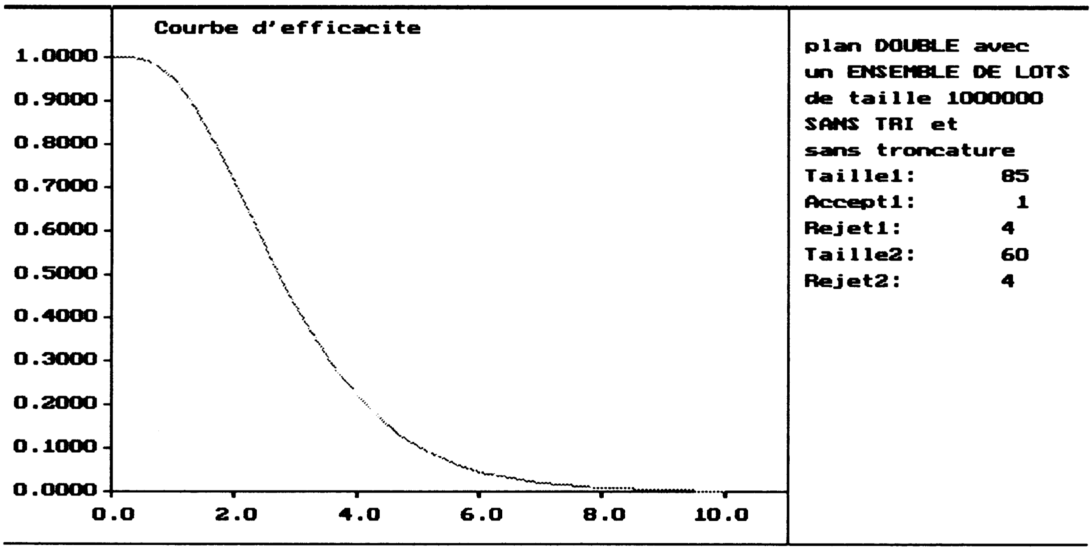
if a second stage is done we will reject the variety if it has (at total on the two stages) 4 off-types or more.

WHAT WILL HAPPEN THEN -F-

if the true percentage of off-types is 1% and we examine 100 times (100 samples) the same variety we will accept it 95 times and reject it 5 times.

if the true percentage of off-types is 5% and we examine 100 times (100 samples) the same variety we will accept it 10 times and reject it 90 times.

Acceptance probability curve



corresponding values:

% off-types prob.to accept the variety

% non confo.	Efficacité
0,000	1,00000
1,000	0,95093
2,000	0,71263
3,000	0,42554
4,000	0,21884
5,000	0,10264
6,000	0,04558
7,000	0,01963
8,000	0,00832
9,000	0,00349
10,000	0,00146

DO I REDUCE THE NUMBER OF PLANTS TO EXAMINE IF I USE A TWO STAGE SCHEME INSTEAD OF A ONE STAGE SAMPLE SCHEME?

It is an advantage of sequential testing versus one stage sample size to reduce the number of plants we examine.
To see how important this reduction is, we will look at an example.
As cases -E- and -F- fullfil exactly the same goal, we are able to compare them.

Let us look at the following tables for case -E- (one stage) and case -F- (two-stages).
The third column "Eff. Moy. Contrô." contains the average number of plants we will examine.

% non confo.	Efficacité	Eff.Moy.Contrô.
0,000	1,00000	130,0
1,000	0,95776	130,0
2,000	0,73689	130,0
3,000	0,45050	130,0
4,000	0,23227	130,0
5,000	0,10576	130,0
6,000	0,04383	130,0
7,000	0,01686	130,0
8,000	0,00610	130,0
9,000	0,00209	130,0
10,000	0,00069	130,0

case -E- one stage

% non confo.	Efficacité	Eff.Moy.Contrô.
0,000	1,00000	85,0
1,000	0,95093	96,9
2,000	0,71263	110,1
3,000	0,42554	113,6
4,000	0,21884	109,9
5,000	0,10264	103,6
6,000	0,04558	97,5
7,000	0,01963	92,8
8,000	0,00832	89,6
9,000	0,00349	87,6
10,000	0,00146	86,4

case -F- two stages

In case -E- (one stage) we always look at 130 plants.
If we read column 1 (% non confo.=true % of off-types) and 3 (Eff. Moy. Contrô.= average number of plantts examined) of the table on the left we see the foolowing:
- If the true percentage of off-types is 0% and we examine 100 samples of the same variety we will look at 130 plants by sample. For all the following lines we have the same value (130 plants).

In case -F- (two stages) the average number of plants we will examine will vary according to the actual percentage of off-types in our variety. If a variety has few or many off-types we will often stop after the first stage, and then observe only 85 plants.
In case -F- if we read column 1 (% non confo.=true % of off-types) and 3 (Eff. Moy. Contrô.= average number of plants examined) of the table on the right we see the foolowing:
- If the true percentage of off-types is 0% and we examine 100 samples of the same variety we will look at 85 plants by sample.
- If the true percentage of off-types is 1% and we examine 100 samples of the same variety we will look at "in average" 97 plants by sample.
- If the true percentage of off-types is 2% and we examine 100 samples of the same variety we will look at "in average" 110 plants by sample.
- etc..
- If the true percentage of off-types is 10% and we examine 100 samples of the same variety we will look at "in average" 86 plants by sample.

We see in this example that the average number of plants by sample we will examine is between 85 (if there is 0% of off-types or very many off-types, in these two cases we will always stop at first stage=85 plants) and 114 plants (if there is in fact 3% of off-types in the variety).
That clearly indicate that we can reduce the number of plants observed when we choose a two-stage sample scheme instead of a one stage sample scheme

We must however keep in mind that sometimes there might be some practical difficulties to make two-stages studies.
For instance if the preparation of the plants before the study is time consuming,
- it might be easier to prepare always 130 plants;
- than to prepare 85 plants for all varieties, and 60 more only for the varieties which had 2 or 3 off-types.

NB: what does mean "in average"? (explanation through an example)
If we examine 100 samples of a given variety and
- we conclude for 82 samples at first stage
- and at second stage for the other 18 samples.
we have examined 82 times 85 plants and 18 times 145 plants. The average number of plants is 95.8
(82*85)+(18*145)=9580 plants for 100 samples.