

INTERNATIONAL STANDARD NORME INTERNATIONALE



3534

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION • МЕЖДУНАРОДНАЯ ОРГАНИЗАЦИЯ ПО СТАНДАРТИЗАЦИИ • ORGANISATION INTERNATIONALE DE NORMALISATION

Statistics — Vocabulary and symbols

First edition — 1977-07-01

Statistique — Vocabulaire et symboles

Première édition — 1977-07-01

UDC/CDU 311 : 519 : 001.4

Ref. No./Réf. n° : ISO 3534-1977 (E/F)

Descriptors : statistics, probability theory, sampling, quality control, vocabulary, symbols/**Descripteurs** : statistique, calcul des probabilités, échantillonnage, contrôle de qualité, vocabulaire, symbole.

Price based on 47 pages/Prix basé sur 47 pages

FOREWORD

ISO (the International Organization for Standardization) is a worldwide federation of national standards institutes (ISO member bodies). The work of developing International Standards is carried out through ISO technical committees. Every member body interested in a subject for which a technical committee has been set up has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work.

Draft International Standards adopted by the technical committees are circulated to the member bodies for approval before their acceptance as International Standards by the ISO Council.

International Standard ISO 3534 was developed by Technical Committee ISO/TC 69, *Applications of statistical methods*. It results from the merging into one single document of draft International Standards ISO/DIS 3534 and 3725, which were circulated to the member bodies in February 1975 and April 1975 respectively.

ISO/DIS 3534 has been approved by the member bodies of the following countries :

Australia	Germany	Romania
Austria	Hungary	South Africa, Rep. of
Belgium	India	Spain
Brazil	Israel	Sweden
Bulgaria	Mexico	Switzerland
Canada	Netherlands	Turkey
Chile	New Zealand	United Kingdom
Czechoslovakia	Poland	U.S.S.R.
France	Portugal	Yugoslavia

ISO/DIS 3725 has been approved by the member bodies of the following countries :

Australia	Hungary	South Africa, Rep. of
Belgium	Israel	Sweden
Brazil	Mexico	Switzerland
Canada	Netherlands	Turkey
Chile	New Zealand	United Kingdom
Czechoslovakia	Poland	U.S.S.R.
France	Portugal	Yugoslavia
Germany	Romania	

The member bodies of the following countries expressed disapproval of both documents on technical grounds :

Japan
U.S.A.

This International Standard cancels and replaces ISO/R 645-1967 and ISO/R 1786-1970.

AVANT-PROPOS

L'ISO (Organisation internationale de normalisation) est une fédération mondiale d'organismes nationaux de normalisation (comités membres de l'ISO). L'élaboration des Normes internationales est confiée aux comités techniques de l'ISO. Chaque comité membre intéressé par une étude a le droit de faire partie du comité technique correspondant. Les organisations internationales, gouvernementales et non gouvernementales, en liaison avec l'ISO, participent également aux travaux.

Les projets de Normes internationales adoptés par les comités techniques sont soumis aux comités membres pour approbation, avant leur acceptation comme Normes internationales par le Conseil de l'ISO.

La Norme internationale ISO 3534 a été élaborée par le comité technique ISO/TC 69, *Application des méthodes statistiques*. Elle résulte de la fusion en un seul document des projets de Normes internationales ISO/DIS 3534 et 3725, qui ont été soumis aux comités membres en février 1975 et en avril 1975 respectivement.

L'ISO/DIS 3534 a été approuvé par les comités membres des pays suivants :

Afrique du Sud, Rép. d'	Espagne	Portugal
Allemagne	France	Roumanie
Australie	Hongrie	Royaume-Uni
Autriche	Inde	Suède
Belgique	Israël	Suisse
Brésil	Mexique	Tchécoslovaquie
Bulgarie	Nouvelle-Zélande	Turquie
Canada	Pays-Bas	U.R.S.S.
Chili	Pologne	Yougoslavie

L'ISO/DIS 3725 a été approuvé par les comités membres des pays suivants :

Afrique du Sud, Rép. d'	Hongrie	Suède
Allemagne	Israël	Suisse
Australie	Mexique	Tchécoslovaquie
Belgique	Nouvelle-Zélande	Turquie
Brésil	Pays-Bas	Royaume-Uni
Canada	Pologne	U.R.S.S.
Chili	Portugal	Yougoslavie
France	Roumanie	

Les comités membres des pays suivants ont désapprouvé les deux documents pour des raisons techniques :

Japon
U.S.A.

La présente Norme internationale annule et remplace l'ISO/R 645-1967 et l'ISO/R 1786-1970.

CONTENTS

	Page
Scope and field of application	1
1 Terms used in theory of probability	2
2 General statistical terms	12
3 General terms relating to methods of sampling	29
4 Terms relating to sampling inspection	32
5 Symbols.	39
English alphabetical index.	42

STANDARDSISO.COM : Click to view the full PDF of ISO 3534:1977

SOMMAIRE

	Page
Objet et domaine d'application	1
1 Termes utilisés en calcul des probabilités	2
2 Termes statistiques généraux	12
3 Termes généraux relatifs aux méthodes d'échantillonnage	29
4 Termes relatifs au contrôle par échantillonnage	32
5 Symboles	39
Répertoire alphabétique français	45

STANDARDSISO.COM : Click to view the full PDF of ISO 3534:1977

STANDARDSISO.COM : Click to view the full PDF of ISO 3534:1977

Statistics — Vocabulary and symbols

Statistique — Vocabulaire et symboles

SCOPE AND FIELD OF APPLICATION

This International Standard gives definitions, in English and French, of statistical terms which may be used in the drafting of other International Standards. In addition, it defines symbols for a limited number of these terms.

The terms are classified under the following main headings :

- 1 Terms used in theory of probability
- 2 General statistical terms
- 3 General terms relating to methods of sampling
- 4 Terms relating to sampling inspection

OBJET ET DOMAINE D'APPLICATION

La présente Norme internationale donne les définitions, en anglais et en français, de termes statistiques susceptibles d'être utilisés dans la rédaction d'autres Normes internationales. En outre, elle définit un ensemble de symboles pour un nombre limité de ces termes.

Les termes sont classés sous les principales rubriques suivantes :

- 1 Termes utilisés en calcul des probabilités
- 2 Termes statistiques généraux
- 3 Termes généraux relatifs aux méthodes d'échantillonnage
- 4 Termes relatifs au contrôle par échantillonnage.

1 TERMS USED IN THEORY OF PROBABILITY

A great number of terms are given both in this clause and in the following clause ("General statistical terms"). Nevertheless, it seems useful to define them separately to draw attention of the reader to the fact that :

- a) the terms in their probabilistic sense apply to principles, independent of any practical application;
- b) the terms in their statistical sense apply to the observations to which they relate : these definitions are of a specifically operational character.

NOTE ON THE NOTION OF PROBABILITY

The concept of probability may be introduced in either of two forms, depending on whether it is intended to designate a degree of belief or whether it is considered as the limit value of a frequency. In both cases, its introduction requires that some precautions be taken which cannot be developed within the context of an International Standard and for which users should refer to specialized literature.

For practical reasons, however, it may be considered that, whenever the conditions of a test can be reproduced, the probability $Pr(E)$ of an event E occurring is the value around which the occurrence frequency of the latter oscillates and towards which it tends when the number of tests is indefinitely increased.

1.1 random variable; variate : A variable which may take any of the values of a specified set of values and with which is associated a probability distribution (see 1.2).

A random variable which may take only isolated values is said to be "discrete". A random variable which may take all the values of a finite or infinite interval is said to be "continuous".

1.2 probability distribution of a random variable : A function which determines the probability that a random variable takes any given value or belongs to a given set of values.

The probability on the whole interval of variation of the variate equals 1.

1.3 distribution function : A function giving, for every value x , the probability that the random variable X be less than or equal to x :

$$F(x) = Pr [X \leq x]$$

1.4 probability density function for a continuous random variable : The derivative (when it exists) of the distribution function.

$$f(x) = \frac{dF(x)}{dx}$$

NOTE — $f(x) dx$ is the "probability element".

$$f(x) dx = Pr [x < X < x + dx]$$

1 TERMES UTILISÉS EN CALCUL DES PROBABILITÉS

De nombreux termes figurent à la fois dans ce chapitre et dans le chapitre suivant («Termes statistiques généraux»). Il a cependant paru utile de les définir séparément pour attirer l'attention du lecteur sur le fait :

- a) que les termes probabilistes s'appliquent à des concepts, indépendamment de toute réalisation;
- b) que les termes statistiques s'appliquent à des observations et aux calculs qui y sont relatifs : ces termes ont un caractère opérationnel que précise la définition.

NOTE SUR LA NOTION DE PROBABILITÉ

La notion de probabilité peut être introduite sous deux formes, soit qu'on veuille s'en servir pour désigner un degré de croyance, soit qu'on la considère comme la valeur limite d'une fréquence. Dans les deux cas, son introduction nécessite certaines précautions qui ne peuvent être développées dans le cadre d'une Norme Internationale et pour lesquelles il convient de se référer aux ouvrages spécialisés.

Pour les besoins de la pratique, on peut toutefois considérer que, dans la mesure où l'on est en état de reproduire les conditions d'un essai, la probabilité $Pr(E)$ d'un événement E susceptible de se produire est la valeur autour de laquelle sa fréquence d'apparition oscille et vers laquelle elle tend, lorsque le nombre d'essais augmente indéfiniment.

1.1 variable aléatoire : Variable pouvant prendre n'importe quelle valeur d'un ensemble déterminé de valeurs, et à laquelle est associée une loi de probabilité (voir 1.2).

Une variable aléatoire qui ne peut prendre que des valeurs isolées est dite «discrète». Une variable aléatoire qui peut prendre toutes les valeurs d'un intervalle fini ou infini est dite «continue».

1.2 loi de probabilité d'une variable aléatoire : Fonction déterminant la probabilité qu'une variable aléatoire prenne une valeur donnée quelconque ou appartienne à un ensemble donné de valeurs.

La probabilité sur tout le domaine de variation de la variable est égale à l'unité.

1.3 fonction de répartition : Fonction donnant, pour toute valeur x , la probabilité que la variable aléatoire X soit inférieure ou égale à x :

$$F(x) = Pr [X \leq x]$$

1.4 densité de probabilité pour une variable aléatoire continue : Dérivée (lorsqu'elle existe) de la fonction de répartition.

$$f(x) = \frac{dF(x)}{dx}$$

NOTE — $f(x) dx$ s'appelle la «probabilité élémentaire».

$$f(x) dx = Pr [x < X < x + dx]$$

1.5 probability for a discrete random variable : x_i being one of the values which can be taken by the discrete random variable X , the probability p_i is :

$$p_i = \Pr [X = x_i]$$

1.6 bivariate distribution : A distribution which determines the probability that a pair of variates takes any given values or belongs to a given set of values.

1.7 multivariate distribution : A distribution which determines the probability that several variates considered simultaneously take any given values or belong to a given set of values.

1.8 marginal distribution : In the case of a probability distribution with k variates, a distribution of a subset of p variates, the other $(k - p)$ variates taking any values in their interval of variation.

Example : In a probability distribution with three variates, X, Y, Z , there are :

- three bivariate marginal distributions : distribution of the pairs $(X, Y), (X, Z), (Y, Z)$;
- three univariate marginal distributions : distributions of X, Y and Z .

1.9 conditional distribution : In the case of a probability distribution with k variates, a distribution of a subset of p variates, when the other $(k - p)$ variates have fixed values.

Example : In a probability distribution with two variates X, Y , we have :

- conditional distribution of X . A specific distribution is expressed as "distribution of X for $Y = y$ ";
- conditional distribution of Y . A specific distribution is expressed as "distribution of Y for $X = x$ ".

1.10 correlation : The inter-dependent relationship between two or several variates when such a relationship includes a random part.

NOTE — In this case, it is said that a "stochastic link" exists between the variates.

1.11 fractile (or quantile) of a probability distribution : If p is a number between 0 and 1, the p -fractile is the value of the random variable for which the distribution function equals p or "jumps" from a value less than or equal to p to a value greater than p .

It is possible that the distribution function equals p throughout the interval between consecutive possible values of the variate. In this case, any value in this interval may be considered as the p -fractile.

1.5 probabilité pour une variable aléatoire discrète : x_i étant l'une des valeurs pouvant être prises par la variable aléatoire discrète X , la probabilité p_i est telle que

$$p_i = \Pr [X = x_i]$$

1.6 loi de probabilité à deux variables : Loi déterminant la probabilité qu'un couple de variables aléatoires prenne des valeurs données quelconques ou appartienne à un ensemble donné de valeurs.

1.7 loi de probabilité à plusieurs variables : Loi déterminant la probabilité que plusieurs variables aléatoires considérées simultanément prennent des valeurs données quelconques ou appartiennent à un ensemble donné de valeurs.

1.8 loi marginale : Dans une loi de probabilité à k variables, loi d'un sous-ensemble de p variables, les $(k - p)$ autres variables pouvant prendre des valeurs quelconques dans leur domaine de variation.

Exemple : Dans une loi de probabilité à trois variables, X, Y, Z , on distingue

- trois lois marginales à deux variables : loi des couples $(X, Y), (X, Z), (Y, Z)$;
- trois lois marginales à une variable : lois de X , de Y et de Z .

1.9 loi conditionnelle : Dans une loi de probabilité à k variables, loi d'un sous-ensemble de p variables, lorsque les $(k - p)$ autres variables ont des valeurs données.

Exemple : Dans une loi de probabilité à deux variables X, Y , on distingue

- les lois conditionnelles de X . Une loi particulière s'énonce «loi de X pour $Y = y$ »;
- les lois conditionnelles de Y . Une loi particulière s'énonce «loi de Y pour $X = x$ ».

1.10 corrélation : Liaison mutuelle entre deux ou plusieurs variables lorsque cette liaison comporte une partie aléatoire.

NOTE — Dans ce cas, on dit qu'il existe une «liaison stochastique» entre les variables.

1.11 fractile (ou quantile) d'une loi de probabilité : p étant un nombre compris entre 0 et 1, le fractile d'ordre p est la valeur de la variable aléatoire pour laquelle la fonction de répartition prend la valeur p ou «saute» d'une valeur inférieure ou égale à p à une valeur supérieure à p .

Il se peut que la fonction de répartition soit égale à p dans tout l'intervalle entre deux valeurs possibles consécutives de la variable. Dans ce cas, toute valeur de cet intervalle peut être considérée comme fractile d'ordre p .

1.12 median: The 0,50 fractile of the probability distribution.

1.13 mode: The value(s) of a random variable such that the probability (discrete variate) or the density (continuous variate) has a maximum for this value (or these values).

If there is one mode, the probability distribution of the variate is said to be "unimodal"; in the other case, it is said to be "multimodal" (bimodal if there are two modes).

1.14 expectation (mean) of a random variable :

a) For a discrete random variable X taking the values x_i with the probabilities p_i , the expectation is defined by

$$E(X) = \sum p_i x_i$$

the sum being extended for all the values x_i which can be taken by X .

b) For a continuous random variable X having the density $f(x)$, the expectation is defined by

$$E(X) = \int x f(x) dx$$

the integral being extended for all values of the interval of variation of X .

No distinction is made between the expectation of a random variable and that of a probability distribution.

NOTES

1 marginal expectation: The expectation of the marginal distribution of a random variable.

2 conditional expectation: The expectation over a conditional distribution of a random variable.

1.15 centred random variable: A random variable the expectation of which equals zero.

NOTE — If the random variable X has an expectation equal to $E(X)$, the corresponding centred variate is $X - E(X)$.

1.16 variance (of a random variable, or of a probability distribution) : The expectation of the square of the centred variate. (See 1.22.)

1.17 standard deviation (of a random variable, or of a probability distribution) : The positive square root of the variance.

1.18 coefficient of variation (of a random variable, or of a probability distribution) : The ratio of the standard deviation to the absolute value of the expectation.

1.12 médiane : Fractile d'ordre $p = 0,50$ de la loi de probabilité.

1.13 mode : Valeur(s) d'une variable aléatoire telle que la probabilité (variable discrète) ou la densité (variable continue) ait un maximum pour cette valeur (ou ces valeurs).

S'il y a un seul mode, la loi de probabilité de la variable est dite «unimodale»; dans le cas contraire, elle est dite «multimodale» (bimodale s'il y a deux modes).

1.14 espérance mathématique (moyenne) d'une variable aléatoire :

a) Pour une variable aléatoire discrète X prenant les valeurs x_i avec les probabilités p_i , l'espérance mathématique est définie par

$$E(X) = \sum p_i x_i$$

la sommation étant étendue à toutes les valeurs x_i susceptibles d'être prises par X .

b) Pour une variable aléatoire continue X de densité $f(x)$, l'espérance mathématique est définie par

$$E(X) = \int x f(x) dx$$

l'intégrale étant étendue à tout le domaine de variation de X .

On dit indifféremment espérance mathématique d'une variable aléatoire ou d'une loi de probabilité.

NOTES

1 espérance mathématique marginale : Espérance mathématique de la loi marginale d'une variable aléatoire.

2 espérance mathématique conditionnelle : Espérance mathématique de la loi conditionnelle d'une variable aléatoire.

1.15 variable aléatoire centrée : Variable aléatoire dont l'espérance mathématique est égale à zéro.

NOTE — Si la variable aléatoire X a pour espérance mathématique $E(X)$, la variable centrée correspondante est $X - E(X)$.

1.16 variance (d'une variable aléatoire ou d'une loi de probabilité) : Espérance mathématique du carré de la variable centrée. (Voir 1.22.)

1.17 écart-type (d'une variable aléatoire ou d'une loi de probabilité) : Racine carrée positive de la variance.

1.18 coefficient de variation (d'une variable aléatoire ou d'une loi de probabilité) : Rapport de l'écart-type à la valeur absolue de l'espérance mathématique.

1.19 standardized variate: A random variable the expectation of which equals zero and the standard deviation of which equals 1.

NOTES

1 If the random variable X has an expectation equal to $E(X)$ and a standard deviation equal to σ , the corresponding standardized variate is the variate

$$\frac{X - E(X)}{\sigma}$$

The distribution of the standardized variate is called "standardized distribution".

2 The concept of standardized variate can be generalized to that of a "reduced variate" which is defined using another location and another scale parameter.

1.20 moment of order q about the origin: * In a univariate distribution, the expectation of the q th power of the variate:

$$E[X^q]$$

NOTE — The moment of order 1 is the expectation of the variate X .

1.21 moment of order q about an origin a : * In a univariate distribution, the expectation of the q th power of the variate $(X - a)$:

$$E[(X - a)^q]$$

1.22 centred moment of order q : * In a univariate distribution, the expectation of the q th power of the central variate $[X - E(X)]$:

$$E\{[X - E(X)]^q\}$$

NOTE — The centred moment of order 2 is the variance of variate X .

1.23 joint moment of orders q and s about the origin: * In a bivariate distribution, the expectation of the product of the q th power of variate X by the s th power of variate Y :

$$E[X^q Y^s]$$

NOTE — The moment of orders 1 and 0 is the expectation of the marginal distribution of X .

The moment of orders 0 and 1 is the expectation of the marginal distribution of Y .

1.24 joint moment of orders q and s about an origin (a, b) : * In a bivariate distribution, the expectation of the product of the q th power of the variate $(X - a)$ by the s th power of the variate $(Y - b)$:

$$E[(X - a)^q (Y - b)^s]$$

* If, in the definition of the moments, the quantities X , $X - a$, Y , $Y - b$, etc., are replaced by their absolute values, i.e. $|X|$, $|X - a|$, $|Y|$, $|Y - b|$, etc., other moments called absolute moments are defined.

1.19 variable aléatoire réduite: Variable aléatoire dont l'espérance mathématique est égale à zéro et dont l'écart-type est égal à l'unité.

NOTES

1 Si la variable aléatoire X a pour espérance mathématique $E(X)$ et pour écart-type σ , la variable réduite correspondante est la variable

$$\frac{X - E(X)}{\sigma}$$

La loi de la variable réduite est appelée «loi réduite».

2 Le terme variable aléatoire réduite peut désigner également une variable transformée de la variable en utilisant une autre valeur centrale et un autre paramètre d'échelle.

1.20 moment d'ordre q par rapport à l'origine: * Dans une loi de probabilité à une variable, espérance mathématique de la $q^{\text{ème}}$ puissance de la variable:

$$E[X^q]$$

NOTE — Le moment d'ordre 1 est l'espérance mathématique de la variable aléatoire X .

1.21 moment d'ordre q par rapport à une origine a : * Dans une loi de probabilité à une variable, espérance mathématique de la $q^{\text{ème}}$ puissance de la variable $(X - a)$:

$$E[(X - a)^q]$$

1.22 moment centré d'ordre q : * Dans une loi de probabilité à une variable, espérance mathématique de la $q^{\text{ème}}$ puissance de la variable centrée $[X - E(X)]$:

$$E\{[X - E(X)]^q\}$$

NOTE — Le moment centré d'ordre 2 est la variance de la variable X .

1.23 moment d'ordres q et s par rapport à l'origine: * Dans une loi de probabilité à deux variables, espérance mathématique du produit de la $q^{\text{ème}}$ puissance de la variable X par la $s^{\text{ème}}$ puissance de la variable Y :

$$E[X^q Y^s]$$

NOTE — Le moment d'ordres 1 et 0 est l'espérance mathématique de la loi marginale de X .

Le moment d'ordres 0 et 1 est l'espérance mathématique de la loi marginale de Y .

1.24 moment d'ordres q et s par rapport à une origine (a, b) : * Dans une loi de probabilité à deux variables, espérance mathématique du produit de la $q^{\text{ème}}$ puissance de la variable $(X - a)$ par la $s^{\text{ème}}$ puissance de la variable $(Y - b)$:

$$E[(X - a)^q (Y - b)^s]$$

* Si, dans la définition des moments, on remplace les quantités X , $X - a$, Y , $Y - b$, etc. par leurs valeurs absolues, c'est-à-dire $|X|$, $|X - a|$, $|Y|$, $|Y - b|$, etc., on définit d'autres moments appelés moments absolus.

1.25 joint centred moment of orders q and s :* In a bivariate distribution, the expectation of the product of the q th power of the centred variate $[X - E(X)]$ by the s th power of the centred variate $[Y - E(Y)]$:

$$E \{ [X - E(X)]^q [Y - E(Y)]^s \}$$

NOTE — The joint centred moment of orders 2 and 0 is the variance of the marginal distribution of X .

The joint centred moment of orders 0 and 2 is the variance of the marginal distribution of Y .

1.26 covariance : The joint centred moment of orders 1 and 1.

$$E \{ [X - E(X)] [Y - E(Y)] \}$$

1.27 coefficient of correlation : The ratio of the covariance of two variates to the product of their standard deviations.

1.28 regression curve : In the case of two variates, the curve defined by the expectation of Y for $X = x$, as x varies, is called the "regression curve of Y on X ".

When the regression curve of Y on X is a straight line, the regression is called "linear".

In this case, the coefficient of linear regression of Y on X is the coefficient of x (slope) in the equation of the regression line.

1.29 regression surface : In the case of three variates, the surface defined by the expectation of Z for $X = x$ and $Y = y$, as x and y vary, is called "the regression surface of Z on X and Y ".

When the regression surface is a plane surface, the regression is called "linear".

In this case, the coefficient of linear regression of Z on X is the coefficient of x in the equation of the regression plane surface.

NOTE — The above definitions may be extended to more than three variates.

1.30 uniform distribution; rectangular distribution : The probability distribution of a continuous variate for which the probability density function is constant within a finite interval and zero outside this interval.

1.25 moment centré d'ordres q et s :* Dans une loi de probabilité à deux variables, espérance mathématique du produit de la $q^{\text{ème}}$ puissance de la variable centrée $[X - E(X)]$ par la $s^{\text{ème}}$ puissance de la variable centrée $[Y - E(Y)]$:

$$E \{ [X - E(X)]^q [Y - E(Y)]^s \}$$

NOTE — Le moment centré d'ordres 2 et 0 est la variance de la loi marginale de X .

Le moment centré d'ordres 0 et 2 est la variance de la loi marginale de Y .

1.26 covariance : Moment centré d'ordres 1 et 1.

$$E \{ [X - E(X)] [Y - E(Y)] \}$$

1.27 coefficient de corrélation : Quotient de la covariance de deux variables par le produit de leurs écarts-types.

1.28 courbe de régression : Dans le cas de deux variables, la courbe définie par l'espérance mathématique de Y pour $X = x$, lorsque x varie, s'appelle «courbe de régression de Y en X ».

Lorsque la courbe de régression de Y en X est une droite, la régression est dite «linéaire».

Dans ce cas, le coefficient de régression linéaire de Y en X est le coefficient de x (pente) dans l'équation de la droite de régression.

1.29 surface de régression : Dans le cas de trois variables, la surface définie par l'espérance mathématique de Z pour $X = x$ et $Y = y$, lorsque x et y varient, s'appelle «surface de régression de Z en X et Y ».

Lorsque la surface de régression est un plan, la régression est dite «linéaire».

Dans ce cas, le coefficient de régression linéaire partielle de Z par rapport à X est le coefficient de x dans l'équation du plan de régression.

NOTE — Les définitions ci-dessus s'étendent à plus de trois variables.

1.30 loi uniforme; loi rectangulaire : Loi de probabilité d'une variable aléatoire continue, telle que la densité de probabilité est constante dans un intervalle fini et nulle hors de cet intervalle.

* See foot-note on page 5.

* Voir note du bas de la page 5.

1.31 normal distribution; Laplace-Gauss distribution : The probability distribution of a continuous random variable X such that, if x is any real number, the probability density is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right]$$

$$-\infty < x < +\infty$$

NOTE — m is the expectation and σ is the standard deviation of the normal distribution.

1.32 standardized normal distribution : The probability distribution of the standardized normal variate.

For a normal variate X having m and σ as parameters, the standardized random variable is

$$U = \frac{X-m}{\sigma}$$

and the probability density of the U distributed variate is

$$f(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right)$$

$$-\infty < u < +\infty$$

1.33 chi-squared distribution : The distribution of the sum of the squares of independent standardized normal variates.

The number of these variates is the number ν of degrees of freedom of the χ^2 distributed variate, a parameter of the distribution.

The probability density function of the χ^2 distributed variate is

$$f(\chi^2, \nu) = \frac{(\chi^2)^{(\nu/2)-1}}{2^{\nu/2}\Gamma(\frac{\nu}{2})} \exp\left(-\frac{\chi^2}{2}\right)$$

$$\chi^2 \geq 0$$

NOTE — The probability distribution of the variate $\chi^2/2$ is a gamma distribution of parameter $m = \nu/2$. (See 1.38.)

1.34 t-distribution; Student distribution : The distribution of a quotient of independent random variables, the numerator of which is a standardized normal variate, and the denominator of which is the positive square root of the quotient of a χ^2 distributed variate and its number of degrees of freedom.

The number of degrees of freedom of χ^2 is the number ν of degrees of freedom of the t -distributed variate.

The probability density function of the t -distributed variate is

$$f(t, \nu) = \frac{1}{\sqrt{\pi}\nu} \frac{\Gamma[(\nu+1)/2]}{\Gamma(\nu/2)} \frac{1}{(1+t^2/\nu)^{(\nu+1)/2}}$$

$$-\infty < t < +\infty$$

1.31 loi normale; loi de Laplace-Gauss : Loi de probabilité d'une variable aléatoire continue X , telle que, si x est un nombre réel quelconque, la densité de probabilité est

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right]$$

$$-\infty < x < +\infty$$

NOTE — m est l'espérance mathématique et σ est l'écart-type de la loi normale.

1.32 loi normale réduite : Loi de probabilité de la variable normale réduite.

Pour une variable normale X de paramètres m et σ , la variable aléatoire réduite est

$$U = \frac{X-m}{\sigma}$$

et la densité de probabilité de la variable U est

$$f(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right)$$

$$-\infty < u < +\infty$$

1.33 loi de χ^2 : Loi de la somme des carrés de variables normales réduites indépendantes.

Le nombre de ces variables est le nombre ν de degrés de liberté de la variable χ^2 , paramètre de cette loi.

La densité de probabilité de la variable χ^2 est

$$f(\chi^2, \nu) = \frac{(\chi^2)^{(\nu/2)-1}}{2^{\nu/2}\Gamma(\frac{\nu}{2})} \exp\left(-\frac{\chi^2}{2}\right)$$

$$\chi^2 \geq 0$$

NOTE — La loi de probabilité de la variable $\chi^2/2$ est une loi gamma de paramètre $m = \nu/2$. (Voir 1.38.)

1.34 loi de t ; loi de Student : Loi du quotient de variables aléatoires indépendantes, le numérateur étant une variable normale réduite et le dénominateur étant la racine carrée positive du quotient d'une variable χ^2 par son nombre de degrés de liberté.

Ce nombre de degrés de liberté de χ^2 est le nombre ν de degrés de liberté de la variable t .

La densité de probabilité de la variable t est

$$f(t, \nu) = \frac{1}{\sqrt{\pi}\nu} \frac{\Gamma[(\nu+1)/2]}{\Gamma(\nu/2)} \frac{1}{(1+t^2/\nu)^{(\nu+1)/2}}$$

$$-\infty < t < +\infty$$

1.35 F-distribution : The distribution of the quotient of two independent χ^2 distributed variates, each one divided by its number of degrees of freedom. The numbers of degrees of freedom of the χ^2 distributed variates of the numerator ν_1 and of the denominator ν_2 are, in this order, the numbers of degrees of freedom of the F -distributed variate.

NOTE — The probability density function of the F -distributed variate is

$$f(F, \nu_1, \nu_2) = \frac{\Gamma[(\nu_1 + \nu_2)/2]}{\Gamma(\nu_1/2) \Gamma(\nu_2/2)} (\nu_1)^{\nu_1/2} (\nu_2)^{\nu_2/2} \frac{F^{(\nu_1/2)-1}}{(\nu_1 F + \nu_2)^{(\nu_1 + \nu_2)/2}}$$

$$F \geq 0$$

1.36 log-normal distribution : The probability distribution of a continuous random variable X which can take any value from a to infinity, with the probability density

$$f(x) = \frac{1}{(x-a) \sigma \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{\log_e (x-a) - m}{\sigma} \right)^2 \right]$$

$$x > a$$

m and σ being respectively the mean and the standard deviation of $\log_e (X-a)$.

NOTES

1 The probability distribution of the variate $\log_e (X-a)$ is a normal distribution; m and σ are respectively the expectation and the standard deviation of this variate.

2 The $\log_{10}$ is often used instead of the $\log_e$. In this case :

$$f(x) = \frac{0,434 \ 3}{(x-a) \sigma \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{\log_{10} (x-a) - m}{\sigma} \right)^2 \right]$$

m and σ being respectively the mean and the standard deviation of $\log_{10} (X-a)$.

1.37 exponential distribution : The probability distribution of a continuous random variable X which can take any value from 0 to $+\infty$, the distribution function of which is

$$F(x) = 1 - e^{-\lambda x}$$

$$\lambda > 0$$

This probability distribution may be generalized by substituting $x - a$ for x (with $x \geq a$).

1.38 gamma distribution : The probability distribution of a continuous random variable X which can take any value from 0 to $+\infty$ with the probability density

$$f(x) = \frac{e^{-x} x^{m-1}}{\Gamma(m)}$$

with $m > 0, x \geq 0$

and $\Gamma(m) = \int_0^\infty e^{-x} x^{m-1} dx$

m , a positive constant, determines the form of the distribution. When m is an integer, we have

$$\Gamma(m) = (m-1)!$$

1.35 loi de F : Loi du quotient de deux variables χ^2 indépendantes, chacune étant divisée par son nombre de degrés de liberté. Les nombres de degrés de liberté des variables χ^2 du numérateur, ν_1 , et du dénominateur, ν_2 , sont, dans cet ordre, les nombres de degrés de liberté de la variable F .

NOTE — La densité de probabilité de la variable F est

$$f(F, \nu_1, \nu_2) = \frac{\Gamma[(\nu_1 + \nu_2)/2]}{\Gamma(\nu_1/2) \Gamma(\nu_2/2)} (\nu_1)^{\nu_1/2} (\nu_2)^{\nu_2/2} \frac{F^{(\nu_1/2)-1}}{(\nu_1 F + \nu_2)^{(\nu_1 + \nu_2)/2}}$$

$$F \geq 0$$

1.36 loi log-normale : Loi de probabilité d'une variable aléatoire continue X , pouvant prendre toute valeur depuis a jusqu'à l'infini, avec la densité de probabilité

$$f(x) = \frac{1}{(x-a) \sigma \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{\log_e (x-a) - m}{\sigma} \right)^2 \right]$$

$$x > a$$

m et σ étant respectivement la moyenne et l'écart-type de $\log_e (X-a)$.

NOTES

1 La loi de probabilité de la variable $\log_e (X-a)$ est une loi normale; m et σ sont respectivement l'espérance mathématique et l'écart-type de cette variable.

2 Le $\log_{10}$ est fréquemment employé à la place du $\log_e$. Dans ce cas,

$$f(x) = \frac{0,434 \ 3}{(x-a) \sigma \sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{\log_{10} (x-a) - m}{\sigma} \right)^2 \right]$$

m et σ étant respectivement la moyenne et l'écart-type de $\log_{10} (X-a)$.

1.37 loi exponentielle : Loi de probabilité d'une variable aléatoire continue X , pouvant prendre toute valeur de 0 à $+\infty$, dont la fonction de répartition est

$$F(x) = 1 - e^{-\lambda x}$$

$$\lambda > 0$$

Cette loi de probabilité peut être généralisée en remplaçant x par $x - a$ (avec $x \geq a$).

1.38 loi gamma : Loi de probabilité d'une variable aléatoire continue X , pouvant prendre toute valeur de 0 à $+\infty$, avec la densité de probabilité

$$f(x) = \frac{e^{-x} x^{m-1}}{\Gamma(m)}$$

avec $m > 0, x \geq 0$

et $\Gamma(m) = \int_0^\infty e^{-x} x^{m-1} dx$

m , constante positive, détermine la forme de la distribution. Lorsque m est un nombre entier, on a

$$\Gamma(m) = (m-1)!$$

For $m = 1$, the gamma distribution becomes an exponential distribution.

This probability distribution may be generalized by substituting $(x - a)/b$ for x (with $x \geq a$ and $b > 0$).

1.39 Gumbel distribution (type I extreme value distribution) : The probability distribution of a continuous random variable X , the distribution function of which is

$$F(x) = \exp(-e^{-y})$$

with $y = (x - a)/b, b > 0, -\infty < x < +\infty$

1.40 Fréchet distribution (type II extreme value distribution) : The probability distribution of a continuous random variable X , the distribution function of which is

$$F(x) = \exp(-y^k)$$

with $y = (x - a)/b, b > 0, x \geq a$

k , a positive constant, determines the shape of the distribution.

1.41 Weibull distribution (type III extreme value distribution) : The probability distribution of a continuous random variable X , the distribution function of which is

$$F(x) = 1 - \exp(-y^k)$$

with $y = (x - a)/b, b > 0, x \geq a$

k , a positive constant, determines the shape of the distribution.

1.42 binomial distribution : The probability distribution of a discrete random variable such that, if $x \in \{0, 1, 2, \dots, n\}$, then

$$Pr[X = x] = \frac{n!}{x! (n-x)!} p^x (1-p)^{n-x}$$

$$0 < p < 1$$

1.43 negative binomial distribution : The probability distribution of a discrete random variable X such that, if $x \in \{0, 1, 2, \dots\}$, then

$$Pr[X = x] = \frac{c(c+1) \dots (c+x-1)}{x!} p^c (1-p)^{x-c}$$

with $c > 0$ and $0 < p < 1$

NOTE — The name "negative binomial distribution" is derived from the fact that the successive probabilities for $x = 0, 1, 2, \dots$, are obtained by developing the binomial of negative exponent $-c$:

$$p^c [1 - (1-p)]^{-c}$$

following the positive and integral powers of $1 - p$.

Pour $m = 1$, la loi gamma devient une loi exponentielle.

Cette loi de probabilité peut être généralisée en remplaçant x par $(x - a)/b$ (avec $x \geq a$ et $b > 0$).

1.39 loi de Gumbel (loi des valeurs extrêmes du type I) : Loi de probabilité d'une variable aléatoire continue X dont la fonction de répartition est

$$F(x) = \exp(-e^{-y})$$

avec $y = (x - a)/b, b > 0, -\infty < x < +\infty$

1.40 loi de Fréchet (loi des valeurs extrêmes du type II) : Loi de probabilité d'une variable aléatoire continue X dont la fonction de répartition est

$$F(x) = \exp(-y^k)$$

avec $y = (x - a)/b, b > 0, x \geq a$

k , constante positive, détermine la forme de la distribution.

1.41 loi de Weibull (loi des valeurs extrêmes du type III) : Loi de probabilité d'une variable aléatoire continue X dont la fonction de répartition est

$$F(x) = 1 - \exp(-y^k)$$

avec $y = (x - a)/b, b > 0, x \geq a$

k , constante positive, détermine la forme de la distribution.

1.42 loi binomiale : Loi de probabilité d'une variable aléatoire discrète X , telle que, si $x \in \{0, 1, 2, \dots, n\}$, on a

$$Pr[X = x] = \frac{n!}{x! (n-x)!} p^x (1-p)^{n-x}$$

$$0 < p < 1$$

1.43 loi binomiale négative : Loi de probabilité d'une variable aléatoire discrète X , telle que, si $x \in \{0, 1, 2, \dots\}$, on a

$$Pr[X = x] = \frac{c(c+1) \dots (c+x-1)}{x!} p^c (1-p)^{x-c}$$

avec $c > 0$ et $0 < p < 1$

NOTE — L'appellation de «loi binomiale négative» vient de ce que les probabilités successives, pour $x = 0, 1, 2, \dots$, s'obtiennent en développant le binôme d'exposant négatif $-c$:

$$p^c [1 - (1-p)]^{-c}$$

suivant les puissances entières et positives de $1 - p$.

1.44 Poisson distribution : The probability distribution of a discrete random variable X such that, if $x \in \{0, 1, 2, \dots\}$, then

$$Pr [X = x] = e^{-m} \frac{m^x}{x!}$$

where m is a positive parameter.

NOTE — The expectation and the variance of the Poisson distribution are both equal to m .

1.45 hypergeometric distribution : If three positive or null integers N , n and d are given such that the numbers in the table below are positive or null integers,

N	d	$N - d$
n	x	$n - x$
$N - n$	$d - x$	$N - n - d + x$

then

$$Pr [X = x] = \frac{n! (N - n)! d! (N - d)!}{N! x! (n - x)! (d - x)! (N - n - d + x)!}$$

1.46 bivariate normal distribution; bivariate Laplace-Gauss distribution : The probability distribution of two continuous variates X and Y , such that the probability density function is

$$f(x, y) = \frac{1}{2\pi \sigma_X \sigma_Y \sqrt{1 - \rho^2}} \exp \left\{ -\frac{1}{2(1 - \rho^2)} \left[\left(\frac{x - m_X}{\sigma_X} \right)^2 - 2\rho \left(\frac{x - m_X}{\sigma_X} \right) \left(\frac{y - m_Y}{\sigma_Y} \right) + \left(\frac{y - m_Y}{\sigma_Y} \right)^2 \right] \right\}$$

with $-\infty < x < +\infty$

and $-\infty < y < +\infty$

NOTES

1 m_X and m_Y are the mathematical expectations and σ_X and σ_Y are the standard deviations of the marginal distributions of X and Y which are normal distributions; ρ is the coefficient of correlation of X and Y .

2 This concept can be extended to a distribution with more than two variates.

1.47 standardized bivariate normal distribution; standardized Laplace-Gauss distribution : The probability distribution of a pair of standardized normal variates. For a pair of normal variates (X, Y) of parameters (m_X, m_Y) and (σ_X, σ_Y) , the corresponding standardized variates are

$$U = \frac{X - m_X}{\sigma_X} \text{ and } V = \frac{Y - m_Y}{\sigma_Y}$$

1.44 loi de Poisson : Loi de probabilité d'une variable aléatoire discrète X , telle que, si $x \in \{0, 1, 2, \dots\}$, on a

$$Pr [X = x] = e^{-m} \frac{m^x}{x!}$$

où m est un paramètre positif.

NOTE — L'espérance mathématique et la variance de la loi de Poisson sont toutes deux égales à la valeur du paramètre m .

1.45 loi hypergéométrique : Étant donné trois entiers positifs ou nuls N , n et d tels que les nombres du tableau suivant soient des entiers positifs ou nuls,

N	d	$N - d$
n	x	$n - x$
$N - n$	$d - x$	$N - n - d + x$

on a

$$Pr [X = x] = \frac{n! (N - n)! d! (N - d)!}{N! x! (n - x)! (d - x)! (N - n - d + x)!}$$

1.46 loi normale à deux variables; loi de Laplace-Gauss à deux variables : Loi de probabilité à deux variables aléatoires continues X et Y , telle que la densité de probabilité soit

avec $-\infty < x < +\infty$

et $-\infty < y < +\infty$

NOTES

1 m_X et m_Y sont les espérances mathématiques et σ_X et σ_Y sont les écarts-types des lois marginales de X et de Y qui sont des lois normales; ρ est le coefficient de corrélation de X et Y .

2 Ce concept s'étend à une distribution à plus de deux variables.

1.47 loi normale réduite à deux variables; loi de Laplace-Gauss réduite à deux variables : Loi de probabilité d'un couple de variables normales réduites. Pour un couple de variables normales (X, Y) de paramètres (m_X, m_Y) et (σ_X, σ_Y) , les variables réduites correspondantes sont

$$U = \frac{X - m_X}{\sigma_X} \text{ et } V = \frac{Y - m_Y}{\sigma_Y}$$

and the probability density function is :

$$f(u, v) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left[-\frac{1}{2(1-\rho^2)} (u^2 - 2\rho uv + v^2) \right]$$

with $-\infty < u < +\infty$

and $-\infty < v < +\infty$

(ρ is the coefficient of correlation of X and Y , and also of U and V .)

NOTE — This concept can be extended to a distribution with more than two variates.

1.48 multinomial distribution : The probability distribution of k discrete variates $X_1, X_2, \dots, X_k$, such that, if $x_1, x_2, \dots, x_k$ are integers taken from 0, 1, 2, 3, ..., n , and such that their sum is $x_1 + x_2 + \dots + x_k = n$, then

$$Pr [X_1 = x_1, X_2 = x_2, \dots, X_k = x_k] = \frac{n!}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}$$

with $p_i \geq 0 (i = 1, 2, \dots, k)$

and

$$\sum_{i=1}^k p_i = 1$$

et la densité de probabilité est

$$f(u, v) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left[-\frac{1}{2(1-\rho^2)} (u^2 - 2\rho uv + v^2) \right]$$

avec $-\infty < u < +\infty$

et $-\infty < v < +\infty$

(ρ est le coefficient de corrélation de X et Y aussi bien que de U et V .)

NOTE — Ce concept s'étend à une distribution à plus de deux variables.

1.48 loi multinomiale : Loi de probabilité à k variables discrètes $X_1, X_2, \dots, X_k$, telle que, si $x_1, x_2, \dots, x_k$ sont des entiers pris parmi 0, 1, 2, 3, ..., n et tels que leur somme $x_1 + x_2 + \dots + x_k = n$, on a

avec $p_i \geq 0 (i = 1, 2, \dots, k)$

et

$$\sum_{i=1}^k p_i = 1$$

2 GENERAL STATISTICAL TERMS

In order to avoid any confusion with similar terms in clause 1 ("Terms used in theory of probability"), certain definitions in this clause which relate to the concept of sampling may be qualified as empirical. However, this qualifying word is omitted where there is no fear of ambiguity. Furthermore, some terms relating to sampling are used in this clause, which are defined in clause 3 ("General terms relating to methods of sampling").

2.1 item :

- a) An actual or conventional object on which a set of observations may be made, or
- b) A defined quantity of material, on which a set of observations may be made, or
- c) An observed value, either qualitative (attributes) or quantitative (measures).

NOTE — The English terms "individual" and "unit" are sometimes used as synonyms for "item".

2.2 population : The totality of items under consideration.

Every clearly defined part of a population is called a subpopulation.

In the case of a random variable, the probability distribution (1.2) is considered as defining the population of that variable.

2.3 characteristic : A property which helps to differentiate between the items of a given population.

The differentiation may be either quantitative (by variables) or qualitative (by attributes).

2.4 test : An operation made in order to measure or classify a characteristic.

2.5 observed value : The value of a characteristic determined as the result of an observation or test.

2.6 absolute difference : The absolute value of the difference between two values.

2.7 range : The difference between the greatest and the smallest observed values of a quantitative characteristic.

2.8 mid-range : The arithmetic mean of the greatest and the smallest observed values of a quantitative characteristic.

2.9 class : In the case of quantitative characteristics, each of the consecutive intervals into which the total interval of variation is divided.

2 TERMES STATISTIQUES GÉNÉRAUX

Pour éviter toute confusion avec les termes analogues du chapitre 1 («Termes utilisés en calcul des probabilités»), certains termes du présent chapitre relatifs à la notion d'échantillonnage peuvent être qualifiés d'empiriques. Toutefois, ce qualificatif est omis lorsqu'aucune ambiguïté n'est à craindre. D'autre part, il est fait usage dans le présent chapitre, de certains termes relatifs à l'échantillonnage, qui sont définis dans le chapitre 3 («Termes généraux relatifs aux modes d'échantillonnage»).

2.1 individu :

- a) Objet concret ou conventionnel sur lequel un ensemble d'observations peut être fait, ou
- b) Quantité définie de matière, sur laquelle un ensemble d'observations peut être fait, ou
- c) Valeur observée, qu'elle soit qualitative (attributs) ou quantitative (mesures).

NOTE — Les termes anglais «individual» et «unit» sont quelquefois utilisés comme synonymes de «individu».

2.2 population : Totalité des individus pris en considération.

Toute partie nettement définie d'une population est une sous-population.

Dans le cas d'une variable aléatoire, la loi de probabilité (1.2) est considérée comme définissant la population de cette variable.

2.3 caractère : Propriété servant à différencier les individus d'une population donnée.

La différenciation entre individus peut être soit quantitative (par mesures ou par variables), soit qualitative (par attributs).

2.4 essai : Opération faite en vue de mesurer ou de classer un caractère.

2.5 valeur observée : Valeur d'un caractère donnée sous la forme du résultat d'une observation ou d'un essai.

2.6 écart : Valeur absolue de la différence entre deux valeurs.

2.7 étendue : Écart entre la plus grande et la plus petite des valeurs observées d'un caractère quantitatif.

2.8 milieu de l'étendue : Moyenne arithmétique de la plus grande et de la plus petite des valeurs observées d'un caractère quantitatif.

2.9 classes : Dans le cas de caractères quantitatifs, chacun des intervalles contigus en lesquels on a partagé l'intervalle total de variation.

2.10 class limits : The values defining the upper and lower bounds of a class.

NOTE — It should be specified which of the two limits is considered as belonging to the class.

2.11 mid-point of class : The arithmetic mean of the upper and lower limits of a class.

2.12 class interval : The difference between the upper limit and the lower limit of a class.

2.13 absolute frequency; size : The number of items (of a population, lot, sample, class, etc.).

2.14 cumulative absolute frequency : In the case of quantitative characteristics, the number of items the values of which are lower than or equal to a given value, or lower than or equal to the upper limit of a given class.

2.15 relative frequency : The ratio of the number of times a particular value (or a value falling within a given class) is observed to the total number of observations.

2.16 cumulative relative frequency : In the case of quantitative characteristics, the relative frequency of items the values of which are lower than or equal to a given value, or lower than or equal to the upper limit of a given class.

2.17 frequency distribution : The relationship between the values of a characteristic and their absolute or relative frequencies.

The distribution is often presented as a table with special groupings (classes) if the values are measured on a continuous scale.

2.18 univariate distribution : In the case of one variate, the relation between the values, classes or forms of this characteristic on the one hand and their sizes or frequencies on the other hand.

2.19 histogram : A graphical representation of the frequency distribution of a continuous variable.

Classes are represented by contiguous intervals on one of the axes on a linear scale. The absolute (or relative) frequencies of the classes are represented by rectangles having these intervals as bases and areas which are proportional to the absolute (or relative) frequencies of the classes.

2.20 bar diagram : A graphical representation of the frequency distribution of a discrete variate. Straight-line segments (bars), proportional to the absolute (or relative) frequencies of the variates, are drawn perpendicular to an axis on which the values of the variate are represented on a linear scale.

2.21 cumulative absolute (relative) frequency polygon : The polygonal line obtained by joining the points the

2.10 limites de classe : Valeurs qui définissent les bornes supérieure et inférieure d'une classe.

NOTE — On doit préciser laquelle des deux limites est considérée comme appartenant à la classe.

2.11 centre de classe : Moyenne arithmétique des limites supérieure et inférieure d'une classe.

2.12 intervalle de classe : Différence entre la limite supérieure et la limite inférieure d'une classe.

2.13 effectif : Nombre d'individus (d'une population, d'un lot, d'un échantillon, d'une classe, etc.).

2.14 effectif cumulé : Dans le cas de caractères quantitatifs, nombre des individus dont la valeur est inférieure ou égale à une valeur donnée, ou inférieure ou égale à la limite supérieure d'une classe donnée.

2.15 fréquence : Rapport du nombre de fois où une valeur particulière (ou une valeur appartenant à une classe donnée) est observée, au nombre total d'observations.

2.16 fréquence cumulée : Dans le cas de caractères quantitatifs, fréquence des individus dont la valeur est inférieure ou égale à une valeur donnée, ou inférieure ou égale à la limite supérieure d'une classe donnée.

2.17 distribution de fréquence : Relation entre les valeurs d'un caractère et leurs effectifs ou leurs fréquences.

La distribution est souvent présentée sous forme d'un tableau avec des groupements spéciaux (classes) si les valeurs sont mesurées avec une échelle continue.

2.18 distribution à une variable : Dans le cas d'une seule variable, relation entre les valeurs, classes ou modalités de ce caractère d'une part, et leurs effectifs ou leurs fréquences d'autre part.

2.19 histogramme : Représentation graphique d'une distribution de fréquence d'une variable continue.

Les classes sont représentées par des intervalles contigus sur l'un des axes suivant une échelle linéaire. Les effectifs (ou fréquences) des classes sont représentés par des rectangles dont les bases sont les intervalles précédents et dont les aires sont proportionnelles aux effectifs (ou aux fréquences) des classes.

2.20 diagramme en bâtons : Représentation graphique d'une distribution de fréquence d'une variable discrète. Perpendiculairement à un axe sur lequel sont représentées les valeurs de la variable, suivant une échelle linéaire, sont élevés des segments rectilignes (bâtons) proportionnels aux effectifs (ou aux fréquences) de ces valeurs.

2.21 polygone des effectifs (ou des fréquences) cumulés : Ligne polygonale obtenue en joignant les points ayant pour

abscissa of which is the upper limit on each class and the ordinate of which is either the cumulative absolute frequency or the cumulative relative frequency.

2.22 bivariate distribution : In the case of two variates considered simultaneously, the relation between the values, classes or forms of characteristics on the one hand and their sizes or frequencies on the other hand.

2.23 multivariate distribution : In the case of several variates considered simultaneously, the relation between the values, classes or forms of characteristics on the one hand and their sizes or frequencies on the other hand.

2.24 two-way table : A table used to represent a distribution with two characteristics and in which the rows and the columns correspond to the values or classes of the first characteristic and second characteristic respectively.

At the intersection of a row and a column appears the number or the frequency corresponding to the combination of values, classes or forms thus marked.

2.25 contingency table : A two-way table in the case of two qualitative characteristics.

The concept may be generalized to the case of more than two characteristics.

2.26 marginal distribution : In a distribution with two variates, the distribution of one of the variates, the other being ignored. It may be obtained by adding for each value or class of the variate under consideration the corresponding absolute (or relative) frequencies of the other variate.

The notion of marginal distribution may be extended to more than two variates. In this case, a marginal distribution is a univariate or a multivariate distribution, the other variates being integrated out.

2.27 conditional distribution : In a distribution with two variates, the distribution of one of the variates when the other one is fixed.

There are, for each variate, as many conditional distributions as there are values or classes for the other variate.

A conditional distribution is expressed by indicating the variate to which it applies and the value or class fixed for the other variate : for example, conditional distribution of the variate Y given $X = x$.

Conditional distributions of absolute frequencies are read directly in the rows and columns of the two-way table of frequencies. The conditional distributions of relative frequencies are obtained by dividing the numbers in each row (or column) by the total number in the relevant row (or column).

abscisse la limite supérieure de chaque classe et pour ordonnée soit l'effectif cumulé, soit la fréquence cumulée.

2.22 distribution à deux variables : Dans le cas de deux variables considérées simultanément, relation entre les valeurs, classes ou modalités de ces caractères d'une part, et leurs effectifs ou leurs fréquences d'autre part.

2.23 distribution à plusieurs variables : Dans le cas de plusieurs variables considérées simultanément, relation entre les valeurs, classes ou modalités de ces caractères d'une part, et leurs effectifs ou leurs fréquences d'autre part.

2.24 tableau à double entrée : Tableau utilisé pour représenter une distribution à deux caractères et dont les lignes d'une part, les colonnes d'autre part, correspondent respectivement aux valeurs, classes ou modalités de l'un et de l'autre caractère.

À l'intersection d'une ligne et d'une colonne figure l'effectif ou la fréquence correspondant à la combinaison de valeurs, classes ou modalités ainsi repérée.

2.25 tableau de contingence : Tableau à double entrée dans le cas de deux caractères qualitatifs.

La notion s'étend au cas de plus de deux caractères.

2.26 distribution marginale : Dans une distribution à deux variables, distribution de l'une des variables lorsqu'on ne tient pas compte de l'autre. Elle peut s'obtenir en faisant pour chaque valeur ou classe de la variable considérée la somme des effectifs (ou des fréquences) correspondant à l'autre variable.

La notion de distribution marginale s'étend à un nombre de variables supérieur à deux. Une distribution marginale est alors une distribution à une ou plusieurs variables, les autres variables n'étant pas prises en considération.

2.27 distribution conventionnelle : Dans une distribution à deux variables, distribution de l'une des variables lorsque l'autre est fixée.

Il y a, pour chaque variable autant de distributions conditionnelles qu'il y a de valeurs ou classes pour l'autre variable.

Une distribution conditionnelle se désigne en indiquant la variable à laquelle elle s'applique et la valeur ou la classe fixée pour l'autre variable : par exemple, distribution conditionnelle de la variable Y pour $X = x$.

Les distributions conditionnelles des effectifs se lisent directement dans les lignes et les colonnes du tableau à double entrée des effectifs. Les distributions conditionnelles des fréquences s'obtiennent en divisant les effectifs de chaque ligne (ou colonne) par l'effectif total de la ligne (ou colonne).

The idea of a conditional distribution may be extended when the number of variates is greater than two. The conditional distributions are then the distributions of one or more variates when the other variate or variates are given.

2.28 arithmetic mean : The sum of values divided by their number.

2.29 (arithmetic) weighted average : Each of the values being assigned a non-negative coefficient called the weight, the sum of the products of each value by its weight, divided by the sum of the weights.

2.30 median : If n values are arranged in increasing order of algebraic magnitude and numbered 1 to n , the median of these n values is the $\left(\frac{n+1}{2}\right)$ th value, if n is odd.

If n is even the median lies between the $\left(\frac{n}{2}\right)$ th and $\left(\frac{n}{2} + 1\right)$ th values and is not defined uniquely. Unless otherwise specified, it may be taken to be the arithmetic mean of these two values.

2.31 mode : The value of the variate which occurs with the greatest absolute (or relative) frequency.

A distribution may possess several local modes, in which case it is called a multimodal distribution.

2.32 mean deviation : The arithmetic mean of the deviations from an origin when all deviations are given a positive sign.

NOTE — Generally, the chosen origin is the arithmetic mean although the mean deviation is minimal when the median is taken as an origin.

2.33 variance : A measure of dispersion based on the mean squared deviation from the arithmetic mean.

Depending upon the cases considered, it may be advantageous to divide the sum of the squared deviations from the arithmetic mean by the number of deviations or by that number minus 1.

Thus, for a series of n observations $x_1, x_2, \dots, x_n$ with mean $\bar{x} = \frac{1}{n} \sum x_i$, one of the expressions

$$\frac{1}{n} \sum (x_i - \bar{x})^2 \text{ or } \frac{1}{n-1} \sum (x_i - \bar{x})^2$$

would be used.

It is important to make clear which definition is adopted.

NOTE — The second expression is generally denoted by s^2 . (See 2.38 and 2.58.)

2.34 standard deviation : The positive square root of the variance.

La notion de distribution conditionnelle s'étend à un nombre de variables supérieur à deux. Les distributions conditionnelles sont alors des distributions à une ou plusieurs variables, l'autre ou les autres variables étant fixées.

2.28 moyenne arithmétique : Quotient d'une somme de valeurs par leur nombre.

2.29 moyenne (arithmétique) pondérée : Les valeurs étant affectées chacune d'un coefficient positif ou nul appelé poids, somme des produits de chaque valeur par son poids, divisée par la somme des poids.

2.30 médiane : n valeurs étant rangées par ordre de grandeur algébrique non décroissant et numérotées de 1 à n , la médiane de ces n valeurs est la $\left(\frac{n+1}{2}\right)$ ème valeurs si n est impair.

Si n est pair, la médiane est comprise entre la $\left(\frac{n}{2}\right)$ ème et la $\left(\frac{n}{2} + 1\right)$ ème valeurs et n'est pas définie de manière unique. Sauf indication contraire, on peut prendre pour médiane la moyenne arithmétique de ces deux valeurs.

2.31 mode : Valeur de la variable à laquelle correspond le plus grand effectif (ou la plus grande fréquence).

Une distribution peut présenter plusieurs modes locaux; dans ce cas, elle est dite multimodale.

2.32 écart moyen : Moyenne arithmétique des écarts par rapport à une origine; les écarts sont pris en valeur absolue.

NOTE — Généralement, l'origine choisie est la moyenne arithmétique bien que l'écart moyen soit minimal quand on prend la médiane pour origine.

2.33 variance : Mesure de la dispersion basée sur le carré moyen des écarts à la moyenne arithmétique.

Selon les cas envisagés, il peut être avantageux de diviser la somme des carrés des écarts à la moyenne arithmétique par le nombre de carrés ou par ce nombre diminué d'une unité.

Ainsi, pour une série de n observations $x_1, x_2, \dots, x_n$, de moyenne $\bar{x} = \frac{1}{n} \sum x_i$, on utilise l'une des expressions

$$\frac{1}{n} \sum (x_i - \bar{x})^2 \text{ ou } \frac{1}{n-1} \sum (x_i - \bar{x})^2$$

Il importe donc toujours de préciser la définition adoptée.

NOTE — La deuxième expression citée ci-dessus est généralement notée s^2 . (Voir 2.38 et 2.58.)

2.34 écart-type : Racine carrée positive de la variance.

2.35 coefficient of variation : The ratio of the standard deviation to the absolute value of the arithmetic mean.

NOTE — The ratio may be expressed as a percentage.

2.36 moment of order q about the origin : In a distribution of a single variate, the arithmetic mean of the q th power of the observed values :

$$\frac{1}{n} \sum_i n_i x_i^q$$

where

n_i is the number of observations with value x_i ;

$n = \sum_i n_i$ is the total number of observations.

NOTE — The moment of order 1 is the arithmetic mean of the observed values.

2.37 moment of order q about an origin a : In a distribution of a single variate, the arithmetic mean of the q th power of the differences between the observed values and the value a :

$$\frac{1}{n} \sum_i n_i (x_i - a)^q$$

where

n_i is the number of observations with value x_i ;

$n = \sum_i n_i$ is the total number of observations.

2.38 centred moment of order q : In a distribution of a single variate, the arithmetic mean of the q th power of the differences between the observed values and their arithmetic mean $\bar{x}$:

$$\frac{1}{n} \sum_i n_i (x_i - \bar{x})^q$$

where

n_i is the number of observations with value x_i ;

$n = \sum_i n_i$ is the total number of observations.

NOTE — The central moment of order 1 is equal to zero. The central moment of order 2 is the variance of the observed values.

2.39 joint moment of orders q and s about the origin : In a distribution of two variates, the arithmetic mean of the products of the q th power of x_i by the s th power of y_j for all paired values (x_i, y_j) constituting the distribution of variates X, Y :

$$\frac{1}{n} \sum_{ij} n_{ij} x_i^q y_j^s$$

2.35 coefficient de variation : Rapport de l'écart-type à la valeur absolue de la moyenne arithmétique.

NOTE — Ce rapport peut être exprimé en pourcentage.

2.36 moment d'ordre q par rapport à l'origine : Dans une distribution à une variable, moyenne arithmétique de la $q^{\text{ème}}$ puissance des valeurs observées :

$$\frac{1}{n} \sum_i n_i x_i^q$$

où

n_i est l'effectif des observations de valeur x_i ;

$n = \sum_i n_i$ est l'effectif total observé.

NOTE — Le moment d'ordre 1 est la moyenne arithmétique des valeurs observées.

2.37 moment d'ordre q par rapport à une origine a : Dans une distribution à une variable, moyenne arithmétique de la $q^{\text{ème}}$ puissance des différences entre les valeurs observées et la valeur a :

$$\frac{1}{n} \sum_i n_i (x_i - a)^q$$

où

n_i est l'effectif des observations de valeur x_i ;

$n = \sum_i n_i$ est l'effectif total observé.

2.38 moment centré d'ordre q : Dans une distribution à une variable, moyenne arithmétique de la $q^{\text{ème}}$ puissance des différences entre les valeurs observées et leur moyenne arithmétique $\bar{x}$:

$$\frac{1}{n} \sum_i n_i (x_i - \bar{x})^q$$

où

n_i est l'effectif des observations de valeur x_i ;

$n = \sum_i n_i$ est l'effectif total observé.

NOTE — Le moment centré d'ordre 1 est égal à zéro. Le moment centré d'ordre 2 est la variance des valeurs observées.

2.39 moment d'ordres q et s par rapport à l'origine : Dans une distribution à deux variables, moyenne arithmétique des produits de la $q^{\text{ème}}$ puissance de x_i par la $s^{\text{ème}}$ puissance de y_j , pour tous les couples de valeurs (x_i, y_j) constituant la distribution des variables X, Y :

$$\frac{1}{n} \sum_{ij} n_{ij} x_i^q y_j^s$$

where

n_{ij} is the observed frequency for the pair (x_i, y_j) ;

$n = \sum_{ij} n_{ij}$ is the total number of observed pairs.

NOTES

1 The joint moment of orders q and s is one of the moments of order $(q + s)$.

2 The moment of orders 1 and 0 is the mean of the marginal distribution of X and the moment of orders 0 and 1 is the mean of the marginal distribution of Y .

2.40 joint moment of orders q and s about an origin (a, b) : In a distribution of two variates, the arithmetic mean of the products of the q th power of the difference between x_i and value a , by the s th power of the difference between y_j and value b , for all paired values (x_i, y_j) constituting the distribution of variates X, Y :

$$\frac{1}{n} \sum_{ij} n_{ij} (x_i - a)^q (y_j - b)^s$$

where

n_{ij} is the observed frequency for the pair (x_i, y_j) ;

$n = \sum_{ij} n_{ij}$ is the total number of observed pairs.

2.41 joint centred moment of orders q and s : In a distribution of two variates, the arithmetic mean of the products of the q th power of the difference between x_i and its mean value $\bar{x}$, by the s th power of the difference between y_j and its mean value $\bar{y}$, for all paired values (x_i, y_j) constituting the distribution of variates X, Y :

$$\frac{1}{n} \sum_{ij} n_{ij} (x_i - \bar{x})^q (y_j - \bar{y})^s$$

where

n_{ij} is the observed frequency for the pair (x_i, y_j) ;

$n = \sum_{ij} n_{ij}$ is the total number of observed pairs.

NOTE — The centred moment of orders 2 and 0 is the variance of the marginal distribution of X . The centred moment of orders 0 and 2 is the variance of the marginal distribution of Y .

2.42 covariance: The centred moment of orders 1 and 1:

$$\frac{1}{n} \sum_{ij} n_{ij} (x_i - \bar{x}) (y_j - \bar{y})$$

where

n_{ij} is the observed frequency for the pair (x_i, y_j) ;

$n = \sum_{ij} n_{ij}$ is the total number of observed pairs.

où

n_{ij} est l'effectif observé pour le couple (x_i, y_j) ;

$n = \sum_{ij} n_{ij}$ est l'effectif total observé.

NOTES

1 Le moment d'ordres q et s est l'un des moments d'ordre $(q + s)$.

2 Le moment d'ordres 1 et 0 est la moyenne de la distribution marginale de X et le moment d'ordres 0 et 1 est la moyenne de la distribution marginale de Y .

2.40 moment d'ordres q et s par rapport à une origine (a, b) : Dans une distribution à deux variables, moyenne arithmétique des produits de la $q^{\text{ème}}$ puissance de la différence entre x_i et la valeur a , par la $s^{\text{ème}}$ puissance de la différence entre y_j et la valeur b , pour tous les couples de valeurs (x_i, y_j) constituant la distribution des variables X, Y :

$$\frac{1}{n} \sum_{ij} n_{ij} (x_i - a)^q (y_j - b)^s$$

où

n_{ij} est l'effectif observé pour le couple (x_i, y_j) ;

$n = \sum_{ij} n_{ij}$ est l'effectif observé.

2.41 moment centré d'ordres q et s : Dans une distribution à deux variables, moyenne arithmétique des produits de la $q^{\text{ème}}$ puissance de la différence entre x_i et sa valeur moyenne $\bar{x}$, par la $s^{\text{ème}}$ puissance de la différence entre y_j et sa valeur moyenne $\bar{y}$, pour tous les couples de valeurs (x_i, y_j) constituant la distribution des variables X, Y :

$$\frac{1}{n} \sum_{ij} n_{ij} (x_i - \bar{x})^q (y_j - \bar{y})^s$$

où

n_{ij} est l'effectif observé pour le couple (x_i, y_j) ;

$n = \sum_{ij} n_{ij}$ est l'effectif total observé.

NOTE — Le moment centré d'ordres 2 et 0 est la variance de la distribution marginale de X . Le moment centré d'ordres 0 et 2 est la variance de la distribution marginale de Y .

2.42 covariance: Moment centré d'ordres 1 et 1:

$$\frac{1}{n} \sum_{ij} n_{ij} (x_i - \bar{x}) (y_j - \bar{y})$$

où

n_{ij} est l'effectif observé pour le couple (x_i, y_j) ;

$n = \sum_{ij} n_{ij}$ est l'effectif total observé.

2.43 coefficient of correlation : The ratio of the covariance of two variables to the product of their standard deviations.

NOTE — This coefficient is frequently used as a numerical expression for the interdependence between two variates or between series of paired observations.

Its value will always lie between -1 and $+1$. When the coefficient of correlation equals one of these limits, this means that there exists an exact linear relationship between the two variates or between the paired observations.

If two variates are statistically independent, the coefficient of correlation is zero, but the converse is not true; a value zero does not prove independence.

2.44 regression curve : In the case of two variates X and Y , the (empirical) regression curve of Y on X provides a correspondence between each value x of the variate X and a central value $y(x)$ of the variate Y .

The regression is called a "straight line" regression, when the regression curve may be considered as a straight line.

In this case, the regression coefficient of Y on X is the coefficient of x in the equation $y = y(x)$ of the regression line of Y on X fitted to the empirical regression curve.

2.45 regression surface : In the case of three variates X , Y and Z , the (empirical) regression surface of Z on X and Y provides a correspondence between each pair (x, y) of values of the variate X and the variate Y and a central value $z(x, y)$ of the characteristic Z .

The regression is called "rectilinear" when the regression surface may be considered as a plane surface.

In this case, the partial regression coefficient of Z on X is the coefficient of x in the equation $z = z(x, y)$ of the regression plane surface Z on X and Y fitted to the empirical regression surface.

NOTE — The above concepts can be extended to situations in which there are more than three variates.

2.46 population parameter : A quantity used to describe the distribution of a characteristic in the population.

2.47 statistic : A function of the observed values derived from a sample.

2.48 order statistic : When the values in a sample are arrayed in order of algebraic magnitude, each of these ordered values is known as an order statistic. An example is the smallest value in a sample. More generally, any statistic based on order statistics in this narrower sense is also called an order statistic, for example the range.

2.49 likelihood : The probability of obtaining some set of observed values, when it is supposed that these are obtained from a specified probability distribution depending on parameters.

2.43 coefficient de corrélation : Quotient de la covariance de deux variables par le produit de leurs écarts-types.

NOTE — Ce coefficient est fréquemment utilisé comme une expression numérique de l'interdépendance entre deux variables aléatoires ou entre des séries d'observations appariées.

Sa valeur sera toujours comprise entre -1 et $+1$. Quand le coefficient de corrélation est égal à l'une de ces limites, cela signifie qu'il existe une relation linéaire exacte entre les deux variables aléatoires ou entre les observations appariées.

Si deux variables aléatoires sont statistiquement indépendantes, le coefficient de corrélation est nul, mais la réciproque n'est pas vraie; un coefficient de corrélation nul ne prouve pas l'indépendance.

2.44 courbe de régression : Dans le cas de deux variables X et Y , la courbe de régression (empirique) de Y en X fait correspondre à toute valeur x de la variable X une valeur centrale $y(x)$ de la variable Y .

La régression est dite «linéaire» lorsque la courbe de régression peut être considérée comme une droite.

Dans ce cas, le coefficient de régression de Y en X est le coefficient de x dans l'équation $y = y(x)$ de la droite de régression de Y en X ajustée à la courbe empirique de régression.

2.45 surface de régression : Dans le cas de trois variables X , Y et Z , la surface de régression (empirique) de Z en X et Y fait correspondre à tout couple (x, y) de valeurs de la variable X et de la variable Y une valeur centrale $z(x, y)$ du caractère Z .

La régression est dite «linéaire» lorsque la surface de régression peut être considérée comme un plan.

Dans ce cas, le coefficient de régression linéaire partielle de Z par rapport à X est le coefficient de x dans l'équation $z = z(x, y)$ du plan de régression de Z en X et Y ajusté à la surface empirique de régression.

NOTE — Les concepts ci-dessus s'étendent à plus de trois variables.

2.46 paramètre de population : Grandeur contribuant à définir la distribution d'un caractère dans la population.

2.47 statistique : Fonction des valeurs observées provenant d'un échantillon.

2.48 statistique d'ordre : Quand les valeurs d'un échantillon sont classées par ordre de grandeur algébrique, chacune de ces valeurs ordonnées est appelée statistique d'ordre; par exemple la plus petite valeur d'un échantillon. Plus généralement, une statistique basée sur des statistiques d'ordre dans ce sens étroit est aussi appelée une statistique d'ordre; par exemple l'étendue.

2.49 vraisemblance : Probabilité d'obtenir un certain ensemble de valeurs observées dans l'hypothèse où celles-ci suivent une loi de probabilité spécifiée dépendant de paramètres.

2.50 estimation; estimate (of parameters) :

a) **estimation** : The operation made for the purpose of assigning, from the values observed in a sample, numerical values to the parameters of a distribution chosen as the statistical model of the population from which this sample is taken.

b) **estimate** : The result of this operation.

This result may be expressed either as a single numerical value (point estimation) or as a confidence interval. (See 2.59 and 2.60.)

2.51 estimator : A statistic intended to estimate a population parameter.

2.52 total estimation error : In the estimation of a parameter, the difference between the calculated value of the estimator and the true value of this parameter.

NOTE — Total estimation error may be due to

- sampling error;
- measurement error;
- rounding-off of values or subdividing into classes;
- a bias of the estimator;
- other errors.

2.53 sampling error : Part of the total estimation error of a parameter due to the random nature of the sample. (See 2.56.)

2.54 bias of estimator : The deviation of the expectation of an estimator of a parameter from the true value of this parameter.

This expression may also be used in a wider sense to designate the non-coincidence of the expectation of an estimator with the true value of the parameter.

2.55 unbiased estimator : An estimator of a parameter such that its expectation equals the true value of this parameter.

2.56 standard error : The standard deviation of an estimator; the standard error provides an estimation of the random part of the total estimation error involved in estimating a population parameter from a sample.

2.57 estimation of the mean of a population (or of the expectation of a probability distribution) : The operation made for the purpose of estimating the mean of a population or the expectation of the probability distribution considered as representing the population.

NOTE — The arithmetic mean of a simple random sample taken from a population is an unbiased estimator of the mean of this population.

2.50 estimation (de paramètres) :

a) **Opération** ayant pour but, à partir des valeurs observées dans un échantillon, d'attribuer des valeurs numériques aux paramètres d'une loi prise comme modèle statistique de la population dont est issu cet échantillon.

b) **Résultat** de cette opération.

Ce résultat peut s'exprimer soit par une simple valeur numérique (estimation ponctuelle), soit par un intervalle de confiance. (Voir 2.59 et 2.60.)

2.51 estimateur : Statistique utilisée pour estimer un paramètre de la population.

2.52 erreur totale d'estimation : Dans l'estimation d'un paramètre, différence entre la valeur calculée de l'estimateur et la valeur vraie de ce paramètre.

NOTE — L'erreur totale d'estimation peut être due

- à l'erreur d'échantillonnage;
- à l'erreur de mesure;
- à l'arrondissement des valeurs ou au rangement en classes;
- à un biais de l'estimateur;
- à d'autres erreurs.

2.53 erreur d'échantillonnage : Partie de l'erreur totale d'estimation d'un paramètre due à la nature aléatoire de l'échantillon. (Voir 2.56.)

2.54 biais d'un estimateur : Différence entre l'espérance mathématique d'un estimateur d'un paramètre et la valeur vraie de ce paramètre.

Cette expression peut aussi être utilisée dans un sens plus général pour désigner la non-coïncidence de l'espérance mathématique de l'estimateur et de la valeur vraie du paramètre.

2.55 estimateur sans biais : Estimateur d'un paramètre tel que son espérance mathématique soit égale à la valeur vraie de ce paramètre.

2.56 erreur-type : Écart-type d'un estimateur; l'erreur-type permet une estimation de la partie aléatoire de l'erreur totale d'estimation faite en estimant un paramètre de la population à partir d'un échantillon.

2.57 estimation de la moyenne d'une population (ou de l'espérance mathématique d'une loi de probabilité) : Opération ayant pour but d'estimer la moyenne d'une population, ou l'espérance mathématique de la loi de probabilité considérée comme représentant cette population.

NOTE — La moyenne arithmétique d'un échantillon simple ou hasard, issu d'une population est un estimateur sans biais de la moyenne de cette population.

2.58 estimation of the variance of a population (or of a probability distribution): The operation made for the purpose of estimating the variance of a population or of a probability distribution considered as representing this population.

NOTES

1 An unbiased estimator of the variance of a probability distribution obtained from a sample of n independent observations is obtained by dividing by $n-1$ the sum of the squares of the deviations of the observed values from the arithmetic mean of this sample.

2 The positive square root of the unbiased estimator of the variance is often used as an estimator of the standard deviation. However, it ought to be noted that this estimator is not unbiased.

2.59 two-sided confidence interval: When T_1 and T_2 are two functions of the observed values such that, θ being a population parameter to be estimated, the probability $Pr[T_1 \leq \theta \leq T_2]$ is equal to $1 - \alpha$ (where $1 - \alpha$ is a fixed number, positive and less than 1), the interval between T_1 and T_2 is a two-sided $(1 - \alpha)$ confidence interval for θ .

The limits T_1 and T_2 of the confidence interval are random variables and as such will generally assume different values in every sample.

In a long series of samples, the relative frequency of cases where the interval includes θ would be approximately equal to $1 - \alpha$.

2.60 one-sided confidence interval: When T is a function of the observed values, such that, θ being a population parameter to be estimated, the probability $Pr[T \leq \theta]$ or the probability $Pr[T \geq \theta]$ is equal to $1 - \alpha$ (where $1 - \alpha$ is a fixed number, positive and less than 1), the interval from the smallest possible value of θ up to T , or the interval between T and the greatest possible value of θ , is a one-sided $(1 - \alpha)$ confidence interval for θ .

The limit T of the confidence interval is a random variable and as such will assume different values in every sample.

In a long series of samples, the relative frequency of cases where the interval includes θ would be approximately equal to $1 - \alpha$.

2.61 confidence coefficient; confidence level: The value $1 - \alpha$ of the probability associated with a confidence interval or a statistical tolerance interval. (See 2.63.)

2.62 confidence limit: Each of the lower and upper limits, T_1 and T_2 , of the two-sided confidence interval. For a one-sided interval, the single limit T of this interval.

2.58 estimation de la variance d'une population (ou d'une loi de probabilité): Opération ayant pour but d'estimer la variance d'une population, ou de la loi de probabilité considérée comme représentant cette population.

NOTES

1 Un estimateur sans biais de la variance d'une loi de probabilité à partir d'un échantillon de n observations indépendantes s'obtient en divisant par $n-1$ la somme des carrés des écarts des valeurs observées par rapport à la moyenne arithmétique de cet échantillon.

2 La racine carrée positive de l'estimateur sans biais de la variance est souvent utilisée comme estimateur de l'écart-type. Il faut noter pourtant que cet estimateur n'est pas sans biais.

2.59 intervalle de confiance bilatéral: Quand T_1 et T_2 sont deux fonctions des valeurs observées telles que, θ étant un paramètre de population à estimer, la probabilité $Pr[T_1 \leq \theta \leq T_2]$ soit égale à $1 - \alpha$ ($1 - \alpha$ étant un nombre fixé, positif et inférieur à 1), l'intervalle entre T_1 et T_2 est un intervalle de confiance bilatéral à $(1 - \alpha)$ pour θ .

Les limites T_1 et T_2 de l'intervalle de confiance sont des variables aléatoires qui, comme telles, auront en général des valeurs différentes pour chaque échantillon.

Dans une grande série d'échantillons, la fréquence des cas où l'intervalle comprendra θ sera approximativement égale à $1 - \alpha$.

2.60 intervalle de confiance unilatéral: Quand T est une fonction des valeurs observées telle que, θ étant un paramètre de la population à estimer, la probabilité $Pr[T \leq \theta]$ ou la probabilité $Pr[T \geq \theta]$ soit égale à $1 - \alpha$ ($1 - \alpha$ étant un nombre fixé, positif et inférieur à 1), l'intervalle s'étendant depuis la plus petite valeur possible pour θ jusqu'à T , ou l'intervalle s'étendant depuis T jusqu'à la plus grande valeur possible pour θ , est un intervalle de confiance unilatéral à $(1 - \alpha)$ pour θ .

La limite T de l'intervalle de confiance est une variable aléatoire qui, comme telle, aura des valeurs différentes pour chaque échantillon.

Dans une grande série d'échantillons, la fréquence des cas où l'intervalle comprendra θ sera approximativement égale à $1 - \alpha$.

2.61 niveau de confiance: Valeur $1 - \alpha$ de la probabilité associée à un intervalle de confiance ou à un intervalle statistique de dispersion. (Voir 2.63.)

2.62 limite de confiance: Chacune des limites inférieure et supérieure, T_1 et T_2 , d'un intervalle de confiance bilatéral. Pour un intervalle unilatéral, limite unique T de cet intervalle.

2.63 statistical tolerance interval : An interval for which it can be stated with a given level of confidence that it contains at least a specified proportion of the population.

When both limits are defined by statistics, the interval is two-sided. When one of the two limits is not finite or consists of the absolute boundary of the variable, the interval is one-sided.

2.64 statistical tolerance limits : For a two-sided statistical tolerance interval, the lower and upper limits of this interval. For a one-sided statistical tolerance interval, the limit of this interval.

2.65 goodness of fit of a distribution : A measure of the agreement between an observed distribution and a theoretical distribution, specified *a priori* or fitted to the observations.

2.66 statistical test : A procedure that is intended to decide whether a hypothesis about the distribution of one or more populations should be rejected or not rejected (accepted).

The decision taken is a result of the value of an appropriate statistic or statistics, calculated from values observed in samples taken from the populations under consideration. As the value of the statistic is subject to random variations, there is some risk of error when the decision is taken.

It is important to note that, generally speaking, a test assumes *a priori* that certain assumptions are fulfilled (for example, assumption of independence of the observations, assumption of normality, etc.). These assumptions serve as a basis of the test.

2.67 null hypothesis — alternative hypothesis

null hypothesis : The hypothesis to be rejected or not rejected (accepted) at the outcome of the test and designated H_0 .

alternative hypothesis : The hypothesis, usually composite (see 2.68), which is opposed to the null hypothesis and designated H_1 .

Examples

1 Test of the hypothesis that the mean m of a variate X in a population is not less than a given number, m_0 :

$$H_0 (m \geq m_0) \text{ and } H_1 (m < m_0)$$

2 Test of the hypothesis that the proportions of defective pieces in two lots, p_1 and p_2 , have the same value (unspecified) :

$$H_0 (p_1 = p_2) \text{ and } H_1 (p_1 \neq p_2)$$

3 Test of the hypothesis that a variate X is distributed normally (with unspecified parameters). Alternative hypothesis : the distribution is not normal.

2.63 intervalle statistique de dispersion : Intervalle pour lequel on peut affirmer, avec un niveau de confiance donné, qu'il contient au moins une proportion également donnée de la population.

Lorsque les deux limites sont définies par des statistiques, l'intervalle est bilatéral. Lorsque l'une des deux limites est infinie ou est constituée par une borne absolue de la variable, l'intervalle est unilatéral.

2.64 limites statistiques de dispersion : Pour un intervalle statistique bilatéral de dispersion, limites inférieure et supérieure de cet intervalle. Pour un intervalle statistique unilatéral de dispersion, limite de cet intervalle.

2.65 adéquation d'une distribution; validité de l'ajustement : Accord entre une distribution observée et une distribution théorique spécifiée *a priori* ou ajustée aux observations.

2.66 test statistique : Procédure ayant pour objet de décider du rejet ou du non-rejet (acceptation) d'une hypothèse relative à la distribution d'une ou plusieurs populations.

La décision adoptée résulte de la valeur d'une ou plusieurs statistiques appropriées, calculée à partir des valeurs observées sur des échantillons prélevés dans les populations considérées. La valeur de la statistique étant sujette à des variations aléatoires, la prise de décision comporte certains risques d'erreur.

Il est important de noter que, d'une façon tout à fait générale, un test suppose *a priori* que certaines hypothèses soient vérifiées (par exemple hypothèse d'indépendance des observations, hypothèse de normalité, etc.). Ces hypothèses sous-jacentes servent de base au test.

2.67 hypothèse nulle — hypothèse alternative

hypothèse nulle : Hypothèse que l'on doit rejeter ou ne pas rejeter (accepter) à l'issue du test, et notée H_0 .

hypothèse alternative : Hypothèse, généralement composite (voir 2.68), que l'on oppose à l'hypothèse nulle, et notée H_1 .

Exemples

1 Test de l'hypothèse que la moyenne m d'une variable X dans une population est au moins égale à un nombre donné, m_0 :

$$H_0 (m \geq m_0) \text{ et } H_1 (m < m_0)$$

2 Test de l'hypothèse que les proportions de pièces défectueuses dans deux lots, p_1 et p_2 , ont la même valeur (non spécifiée) :

$$H_0 (p_1 = p_2) \text{ et } H_1 (p_1 \neq p_2)$$

3 Test de l'hypothèse qu'une variable X est distribuée normalement (avec des paramètres non spécifiés). Hypothèse alternative : la distribution n'est pas normale.

2.68 simple hypothesis — composite hypothesis

simple hypothesis : A hypothesis which specifies the population distribution completely.

composite hypothesis : A hypothesis which does not specify the population distribution completely.

In practice, it is more often that not a hypothesis that comprises a generally infinite collection of simple hypotheses.

Examples

1 Assuming a normal distribution, the hypothesis $m = m_0$ is simple if the standard deviation is known, but composite if it is unknown.

2 All the hypotheses in the examples given in 2.67 are composite.

2.69 parametric test — distribution free test — non-parametric test

parametric test : A test in which the distribution function of the test statistic depends on the form of the distribution function of the observations.

Example : Student test of the null hypothesis that the mean of a population with unknown variance is equal to a given value, where the distribution of the population must be supposed normal.

distribution free test : A test in which the distribution function of the test statistic does not depend on the form of the distribution function of the observations.

NOTE — If the hypothesis does not explicitly make an assertion about a parameter, the test is sometimes non-parametric.

Example : Test of the null hypothesis that the observations of a given series are independently distributed with the same unspecified distribution function.

2.70 critical region — significance level

critical region : The set of possible values of the statistic used such that, if the value of the statistic which results from the observed values belongs to the set, the null hypothesis will be rejected, whereas it will not be rejected (accepted) if the opposite is the case.

significance level (of a test) : The given value, designated α , which limits the probability of the null hypothesis being rejected, if the null hypothesis is true.

Critical regions are determined in such a way that if the null hypothesis is true, the probability of this null hypothesis being rejected should be not more than this given value α , which is in general low (for example 5 % or 1 %).

Example : A classic way of testing the null hypothesis about the mean of a normal distribution with known standard deviation σ , $H_0 (m \geq m_0)$, against the alternative $H_1 (m < m_0)$, uses the statistic $\bar{x}$, the mean of the sample.

2.68 hypothèse simple — hypothèse composite

hypothèse simple : Hypothèse qui spécifie complètement la distribution de la population.

hypothèse composite : Hypothèse qui ne spécifie pas complètement la distribution de la population.

En pratique, il s'agit le plus souvent d'une hypothèse qui comprend un ensemble généralement infini d'hypothèses simples.

Exemples

1 Pour une distribution supposée normale, l'hypothèse $m = m_0$ est simple si l'écart-type est connu, mais elle est composite s'il est inconnu.

2 Toutes les hypothèses des exemples donnés en 2.67 sont composites.

2.69 test paramétrique — test non paramétrique

test paramétrique : Test dans lequel la fonction de répartition de la statistique dépend de la forme de la fonction de répartition des observations.

Exemple : Test de Student de l'hypothèse nulle que la moyenne d'une population de variance inconnue est égale à une valeur donnée, la répartition de la population étant obligatoirement normale.

test non paramétrique : Test dans lequel la fonction de répartition de la statistique ne dépend pas de la forme de la fonction de répartition des observations.

NOTE — Ce terme recouvre, en anglais, les notions correspondant à «distribution free test» et «non-parametric test».

Exemple : Test de l'hypothèse nulle que les observations d'une série donnée sont réparties d'une manière indépendante selon la même fonction de répartition non spécifiée.

2.70 région critique — niveau de signification

région critique : Ensemble des valeurs possibles de la statistique utilisée tel que, si la valeur de cette statistique résultant des valeurs observées appartient à cet ensemble, on rejettera l'hypothèse nulle, tandis qu'on ne la rejettera pas (acceptera) dans le cas contraire.

niveau de signification (d'un test) : Valeur donnée, notée α , limitant la probabilité de rejet d'une hypothèse nulle, quand celle-ci est vraie.

Les régions critiques sont déterminées de façon que, si l'hypothèse nulle est vraie, la probabilité de rejet de cette hypothèse nulle soit au plus égale à cette valeur donnée, α , généralement faible (par exemple 5 % ou 1 %).

Exemple : Une façon classique de tester l'hypothèse nulle relative à la moyenne d'une distribution normale d'écart-type connu σ , $H_0 (m \geq m_0)$, contre l'alternative $H_1 (m < m_0)$ consiste à utiliser la statistique $\bar{x}$, moyenne de l'échantillon.

The critical region is the set of values less than $A = m_0 + u_\alpha \sigma / \sqrt{n}$

where

n is the sample number;

u_α is the fractile of order α of the unit normal variable.

If the calculated value of $\bar{x}$ is less than A , H_0 is rejected. If the opposite is the case, H_0 is not rejected (accepted).

2.71 one-sided test — two-sided test — critical value(s)

one-sided test : A test in which the statistic used is one-dimensional and the critical region is the set of values lower than a given number (or the set of values greater than a given number).

two-sided test : A test in which the statistic used is one-dimensional and in which the critical region is the set of values lower than a first given number and the set of values greater than a second given number.

critical values : The given number(s) which limit the critical region.

NOTE — The choice between a one-sided test and a two-sided test is determined by the alternative hypothesis. In the example given in 2.70, the test is one-sided and the critical value is A .

2.72 error of the first kind — type I risk (size of the test)

error of the first kind : The error committed in rejecting the null hypothesis (because the statistic takes a value which belongs to the critical region), when the null hypothesis is true.

type I risk (size of the test) : The probability of committing the error of the first kind, which varies according to the real situation (within the framework of the null hypothesis). Its maximum value is the significance level of the test. (See 2.70.)

Example : In the case envisaged in the example given in 2.70, the error of the first kind would consist in rejecting H_0 ($m < m_0$), because $\bar{x}$ would be less than A , whereas, in effect, m would be equal to or greater than m_0 . The probability of such an error is α if $m = m_0$ and decreases as m increases.

2.73 error of the second kind — type II risk — power of the test

error of the second kind : The error committed in failing to reject (accept) the null hypothesis (because the value of the statistic does not belong to the critical region), when the null hypothesis is not true (the alternative hypothesis therefore being true).

La région critique est l'ensemble des valeurs inférieures à $A = m_0 + u_\alpha \sigma / \sqrt{n}$

où

n est l'effectif de l'échantillon;

u_α est le fractile d'ordre α de la variable normale réduite.

Si la valeur calculée de $\bar{x}$ est inférieure à A , on rejettera H_0 . Dans le cas contraire, on ne rejettera pas (acceptera) H_0 .

2.71 test unilatéral — test bilatéral — valeur(s) critique(s)

test unilatéral : Test dans lequel la statistique utilisée est unidimensionnelle et la région critique comprend uniquement l'ensemble des valeurs inférieures à un nombre donné (ou l'ensemble des valeurs supérieures à un nombre donné).

test bilatéral : Test dans lequel la statistique utilisée est unidimensionnelle et la région critique comprend à la fois l'ensemble des valeurs inférieures à un premier nombre donné et l'ensemble des valeurs supérieures à un second nombre donné.

valeur(s) critique(s) : Nombre(s) donné(s) limitant la région critique.

NOTE — Le choix entre un test unilatéral et un test bilatéral est déterminé par l'hypothèse alternative. Dans l'exemple donné en 2.70, le test est unilatéral et la valeur critique est A .

2.72 erreur de première espèce — risque de première espèce

erreur de première espèce : Erreur commise lorsqu'on décide de rejeter l'hypothèse nulle (parce que la statistique prend une valeur appartenant à la région critique), alors que cette hypothèse nulle est vraie.

risque de première espèce : Probabilité de commettre l'erreur de première espèce qui varie suivant la situation réelle (dans le cadre de l'hypothèse nulle). Sa valeur maximale est le niveau de signification du test. (Voir 2.70.)

Exemple : Dans le cas envisagé dans l'exemple donné en 2.70, l'erreur de première espèce consisterait à rejeter H_0 ($m \geq m_0$), parce que $\bar{x}$ serait inférieure à A , alors que, effectivement, m serait égal ou supérieur à m_0 . La probabilité d'une telle erreur est égale à α si $m = m_0$ et décroît lorsque m croît.

2.73 erreur de seconde espèce — risque de seconde espèce — puissance du test

erreur de seconde espèce : Erreur commise lorsqu'on décide de ne pas rejeter (accepter) l'hypothèse nulle (parce que la statistique prend une valeur n'appartenant pas à la région critique), alors que cette hypothèse nulle n'est pas vraie (l'hypothèse alternative l'étant donc).

type II risk : The probability, designated β , of committing the error of the second kind. Its value depends on the real situation and can only be calculated if the alternative hypothesis is adequately specified.

power of test : The probability of not committing the error of the second kind. It is equal to $1 - \beta$. Thus, it is the probability of rejecting the null hypothesis when this hypothesis is false.

Examples

1 In the example given in 2.70, the error of second kind would consist in accepting H_0 ($m \geq m_0$), because $\bar{x}$ would be greater than A , whereas, in effect, m would be less than m_0 . The probability β of such an error depends on the effective value of m : the lower m is than m_0 , and the weaker it is, the closer to 1, on the other hand, is the power.

2 In the test of hypothesis H_0 (the population is distributed normally) against the alternative H_1 (the population is not distributed normally) (see example 3 in 2.67), it is usually not possible to express β as a function of the alternative hypothesis as this alternative is too imprecise.

2.74 power curve of a test : The curve which, when the hypothesis concerns the value of a one-dimensional parameter, expresses graphically the variation of the power in terms of the value of the parameter.

Most often, the only part of the curve of interest is that which corresponds to values of the parameter realizing the alternative hypothesis. This represents the probability, in terms of the value of the parameter, of accepting the alternative hypothesis when it is true.

Examples

1 Figure 1 illustrates the power curve for the test of the hypothesis H_0 ($m \geq m_0$) (see example in 2.70) against H_1 ($m < m_0$) as a function of m , the population mean.

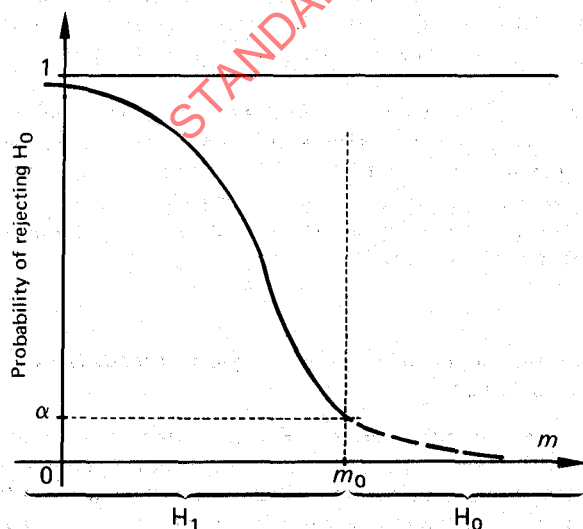


FIGURE 1 — Power curve

risque de seconde espèce : Probabilité, notée β , de commettre l'erreur de seconde espèce. Sa valeur dépend de la situation réelle et ne peut être calculée que si l'hypothèse alternative est suffisamment spécifiée.

puissance du test : Probabilité de ne pas commettre l'erreur de seconde espèce. Elle est égale à $1 - \beta$. C'est donc la probabilité de rejeter l'hypothèse nulle quand celle-ci est fausse.

Exemples

1 À nouveau, dans le cas envisagé dans l'exemple donné en 2.70, l'erreur de seconde espèce consisterait à accepter H_0 ($m \geq m_0$), parce que $\bar{x}$ serait supérieure à A , alors que, effectivement, m serait inférieur à m_0 . La probabilité β d'une telle erreur dépend de la valeur effective de m : plus m est inférieur à m_0 , plus elle est faible, plus, au contraire, la puissance est voisine de 1.

2 Dans le test de l'hypothèse nulle H_0 (la population est distribuée normalement) contre l'alternative H_1 (la population n'est pas distribuée normalement) (voir exemple 3 en 2.67), il n'est généralement pas possible d'exprimer β en fonction de l'hypothèse alternative, celle-ci étant trop imprecise.

2.74 courbe de puissance d'un test : Courbe qui, lorsque l'hypothèse porte sur la valeur d'un paramètre unidimensionnel, représente graphiquement la variation de la puissance d'un test en fonction de la valeur de ce paramètre.

Le plus souvent, on ne s'intéresse qu'à la partie de la courbe qui correspond à des valeurs du paramètre réalisant l'hypothèse alternative. Elle représente la probabilité, en fonction de la valeur du paramètre, d'accepter l'hypothèse alternative, quand celle-ci est vraie.

Exemples

1 La figure 1 représente la courbe de puissance du test de l'hypothèse H_0 ($m \geq m_0$) contre H_1 ($m < m_0$) (voir exemple en 2.70) en fonction de m , moyenne de la population.

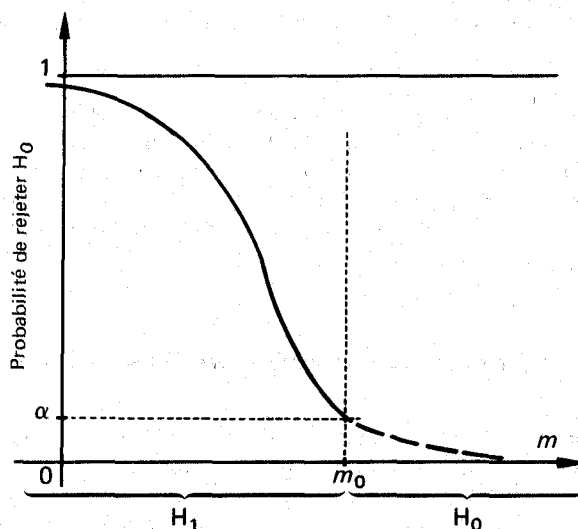


FIGURE 1 — Courbe de puissance

2 Figure 2 illustrates the power curve for the test of the hypothesis H_0 ($p < p_0$) against H_1 ($p \geq p_0$) as a function of p , the proportion of defectives in a lot subjected to control.

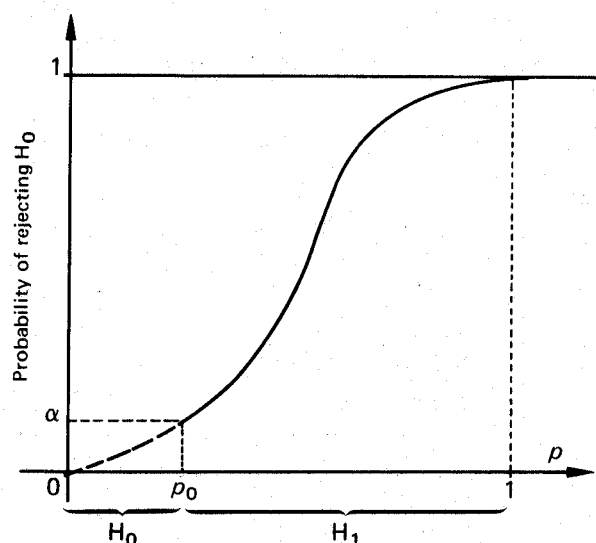


FIGURE 2 — Power curve

2 La figure 2 représente la courbe de puissance du test de l'hypothèse H_0 ($p < p_0$) contre H_1 ($p \geq p_0$) en fonction de p , proportion des pièces défectueuses dans un lot soumis au contrôle.

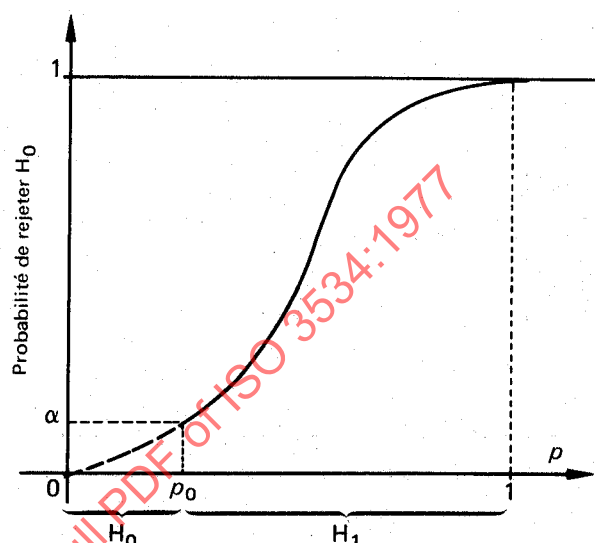


FIGURE 2 — Courbe de puissance

2.75 operating characteristic curve of a test (OC curve of a test): The curve which, when the hypothesis concerns the value of a one-dimensional parameter, expresses graphically the variation of the probability of not rejecting (accepting) the null hypothesis in terms of the value of the parameter.

Examples

1 Figure 3 illustrates the OC curve for the test of the hypothesis H_0 ($m \geq m_0$) against H_1 ($m < m_0$) (see 2.70) as a function of m , the population mean.

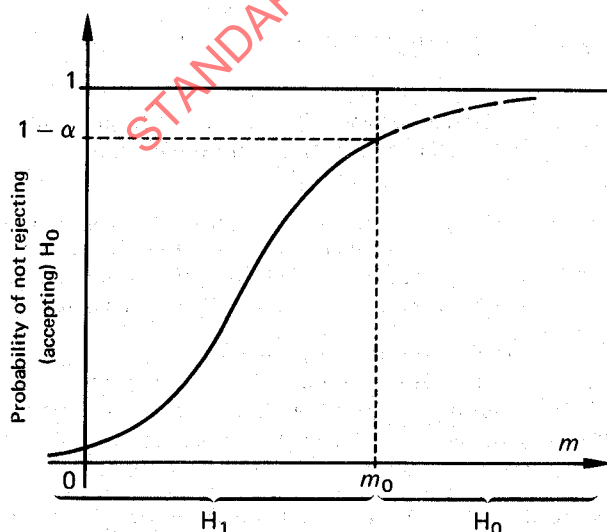


FIGURE 3 — OC curve

2.75 courbe d'efficacité d'un test : Courbe qui, lorsque l'hypothèse porte sur la valeur d'un paramètre unidimensionnel, représente graphiquement la variation de la probabilité de ne pas rejeter (accepter) l'hypothèse nulle en fonction de la valeur de ce paramètre.

Exemples

1 La figure 3 représente la courbe d'efficacité du test de l'hypothèse H_0 ($m \geq m_0$) contre H_1 ($m < m_0$) (voir exemple en 2.70) en fonction de m , moyenne de la population.

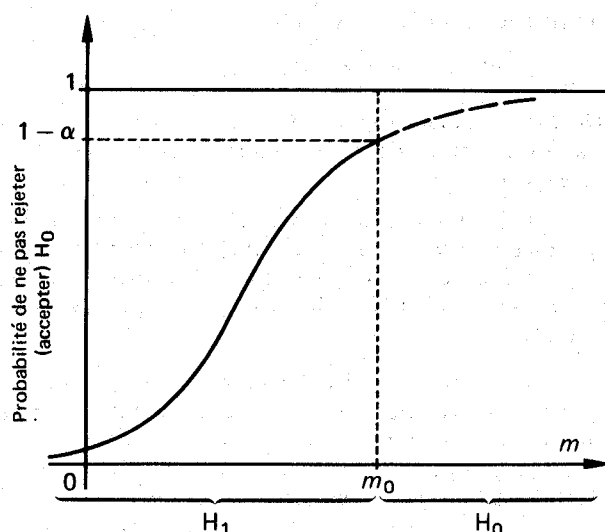


FIGURE 3 — Courbe d'efficacité

2 Figure 4 illustrates the OC curve for the test of the hypothesis H_0 ($p < p_0$) against H_1 ($p \geq p_0$) as a function of p , the proportion of defectives in a lot subjected to control.

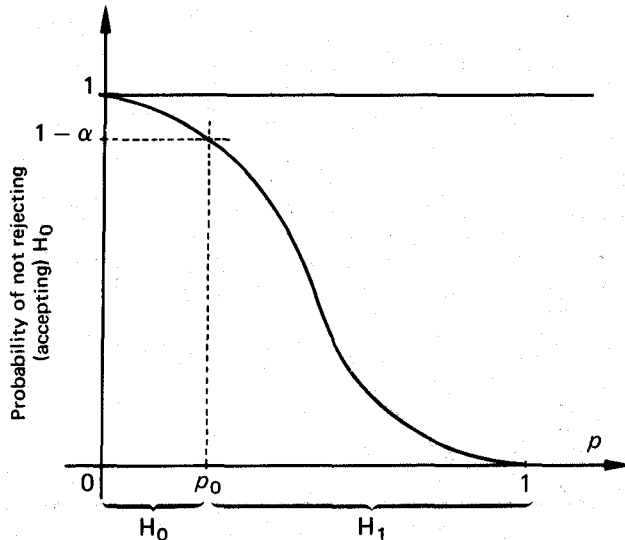


FIGURE 4 – OC curve

2.76 significant result (at the chosen level α) : The result of a test which leads to the rejection of the null hypothesis. If the hypothesis is not rejected, the result is not significant.

Example : For the example given in 2.70, if $\bar{x}$ is less than $A = m_0 + u_\alpha \sigma / \sqrt{n}$, $\bar{x}$ is said to be significantly lower than m_0 at the level α .

2.77 chi-squared test : A test in which the statistic used has a χ^2 distribution, occurring, for instance, in the following problems :

- test of the equality between the variance of a normal population and a specified value, based on the variance estimated from a sample taken from this population;
- comparison between the observed sizes and the theoretical sizes corresponding to a distribution specified in advance or defined from observed values;
- tests of independence and of homogeneity.

2.78 t-test; Student test : A test in which the statistic used has a t -(Student) distribution, occurring, for instance, in the following problems :

- test of the equality between the mean of a normal population and a specified value, based on the mean observed on a sample taken from this population, the variance of the population being estimated from the sample;
- test of the equality between the means of two normal populations having the same variance based on the means observed on two independent samples taken from these populations, the common variance being estimated from these samples;
- test applying to the value of a linear regression or correlation coefficient.

2 La figure 4 représente la courbe d'efficacité du test de l'hypothèse H_0 ($p < p_0$) contre H_1 ($p \geq p_0$) en fonction de p , proportion des pièces défectueuses dans un lot soumis au contrôle.

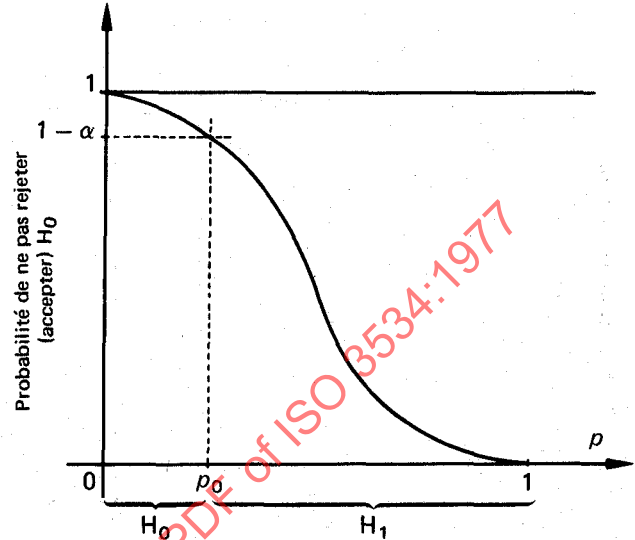


FIGURE 4 – Courbe d'efficacité

2.76 résultat significatif (au niveau de signification α choisi) : Résultat d'un test, qui a conduit au rejet de l'hypothèse nulle. Dans le cas contraire, le résultat est non significatif.

Exemple : Pour l'exemple donné en 2.70, si $\bar{x}$ est inférieure à $A = m_0 + u_\alpha \sigma / \sqrt{n}$, on dira que $\bar{x}$ est significativement inférieure à m_0 au niveau α .

2.77 test χ^2 : Test où la statistique utilisée suit une loi de χ^2 et intervenant notamment dans les problèmes suivants :

- test de l'égalité de la variance d'une population normale à une valeur spécifiée, à partir de la variance estimée sur un échantillon issu de cette population;
- comparaison des effectifs observés avec les effectifs théoriques correspondant à une loi spécifiée *a priori* ou définie à partir des valeurs observées;
- tests d'indépendance et d'homogénéité.

2.78 test t ; test de Student : Test où la statistique utilisée suit une loi de t (de Student) et intervenant notamment dans les problèmes suivants :

- test de l'égalité de la moyenne d'une population normale à une valeur spécifiée, à partir de la moyenne observée sur un échantillon issu de cette population, la variance de la population étant estimée sur l'échantillon;
- test de l'égalité des moyennes de deux populations normales de même variance, à partir des moyennes observées sur deux échantillons indépendants issus de ces populations, la variance commune étant estimée sur ces échantillons;
- tests relatifs à la valeur d'un coefficient de régression ou de corrélation linéaire.

2.79 *F*-test : A test in which the statistic used has an *F*-distribution, occurring, for instance, in the following problems :

- test of the equality of the variances of two normal populations based on the variances estimated from two independent samples taken from these populations;
- tests occurring in the analysis of variance.

2.80 a) repetition : A term denoting the execution of a statistical inquiry at different points in space or time, usually as part of a co-ordinated programme, as distinct from replication.

b) replication : During the course of an experiment or survey, replication is the determination of a value more than once so as to obtain a better estimation of the variation. Replication should be distinguished from repetition by the fact that replication of an experiment denotes repeated determinations carried out, as far as possible, at one place and one period of time. The successive determinations, including the first, are called replicates.

NOTES — The English terms "repetition" and "replication" are equivalent to the French term "répétition". The above definitions are equivalent to the corresponding French definition.

2.81 randomization

to draw at random : If, from a population of N items, a single item is drawn in such a way that all items have an equal chance of being drawn, that item is said to be drawn at random.

randomization : If, from a population consisting of the natural numbers 1 to n , these are drawn at random one by one successively, without replacement until the population is exhausted, the numbers are said to be drawn "in random order".

If these n numbers have been associated in advance with n distinct objects or n distinct operations which are then rearranged in the order in which the numbers are drawn, the order of the objects or operations is said to be randomized.

This process is called randomization.

2.82 true value : The value which characterizes a quantity perfectly defined in the conditions which exist at the moment when that quantity is observed (or the subject of a determination).

It is an ideal value which could be arrived at only if all causes of measurement error were eliminated and the population was infinite.

2.83 accuracy of the mean : The closeness of agreement between the true value and the mean result which would be obtained by applying the experimental procedure a very large number of times.

The smaller the systematic part of the experimental errors which affect the results, the more accurate is the procedure.

2.79 test *F* : Test où la statistique utilisée suit une loi de *F* et intervenant notamment dans les problèmes suivants :

- test de l'égalité des variances de deux populations normales, à partir des variances estimées sur deux échantillons indépendants issus de ces populations;
- tests intervenant dans l'analyse de la variance.

2.80 répétition : Chaque réalisation, y compris la première, dans le cas d'une expérience ou d'une partie d'expérience réalisée plus d'une fois dans les mêmes conditions.

NOTE — Le terme français «répétition» correspond aux deux termes anglais «repetition» et «replication». La définition ci-dessus recouvre les deux définitions correspondantes anglaises.

2.81 randomisation

tirer au hasard : Si, dans une population de N individus, on tire un seul individu de telle façon que tous les individus aient une chance égale d'être tirés, on dit que cet individu a été tiré au hasard.

randomisation : Si, dans une population composée des nombres naturels 1 à n , ceux-ci sont tirés successivement un par un au hasard sans remise jusqu'à épuisement, on dit que les nombres ont été tirés «dans un ordre au hasard».

Si ces nombres ont été associés à l'avance à n objets distincts ou n opérations distinctes, qui sont alors mis dans l'ordre dans lequel les nombres sont tirés, on dit que l'ordre des objets ou des opérations a été «randomisé».

Ce procédé s'appelle la randomisation.

2.82 valeur vraie : Valeur qui caractérise une grandeur parfaitement définie dans les conditions qui existent au moment où cette grandeur est examinée (ou fait l'objet d'une détermination).

C'est une valeur idéale qu'on ne pourrait atteindre que si l'on pouvait éliminer toutes les causes d'erreur de mesure et si la population était infinie.

2.83 justesse : Étroitesse de l'accord entre la valeur vraie et le résultat moyen qui serait obtenu en appliquant le procédé expérimental un grand nombre de fois.

Le procédé est d'autant plus juste que la partie systématique des erreurs expérimentales qui affectent les résultats est moindre.

2.84 precision : The closeness of agreement between the results obtained by applying the experimental procedure several times under prescribed conditions.

The smaller the random part of the experimental errors which affect the results, the more precise is the procedure.

2.85 repeatability :

a) *Qualitatively*

The closeness of agreement between successive results obtained with the same method on identical test material, under the same conditions (same operator, same apparatus, same laboratory and short intervals of time).

NOTE — The representative parameters of the dispersion of the population which may be associated to the results, are qualified by the term "repeatability".

Exemple : Standard deviation of repeatability, variance of repeatability, ...

b) *Quantitatively*

The value below which the absolute difference between two single test results obtained in the above conditions may be expected to lie with a specified probability.

In the absence of other indication, the probability is 95 %.

2.86 reproducibility :

a) *Qualitatively*

The closeness of agreement between individual results obtained with the same method on identical test material but under different conditions (different operators, different apparatus, different laboratories and/or different times).

NOTE — The representative parameters of the dispersion of the population which may be associated to the results are qualified by the term "reproducibility".

Exemple : Standard deviation of reproducibility, variance of reproducibility, ...

b) *Quantitatively*

The value below which the absolute difference between two single test results on identical material obtained by operators in different laboratories, using the standardized test method, may be expected to lie with specified probability.

In the absence of other indication, the probability is 95 %.

NOTE — If the terms 2.83 to 2.86 are used with other meanings than these given above, the terms should be specially and precisely defined, as experience has shown that ambiguities can arise.

2.84 fidélité : Étroitesse de l'accord entre les résultats obtenus en appliquant le procédé expérimental à plusieurs reprises dans des conditions déterminées.

Le procédé est d'autant plus fidèle que la partie aléatoire des erreurs expérimentales qui affectent les résultats est moindre.

2.85 répétabilité :

a) *Qualitativement*

Étroitesse de l'accord entre les résultats successifs obtenus avec la même méthode sur une matière identique soumise à l'essai dans les mêmes conditions (même opérateur, même appareil, même laboratoire et court intervalle de temps).

NOTE — Les paramètres représentatifs de la dispersion de la population qui peut être associée aux résultats sont qualifiés du terme «répétabilité».

Exemple : Écart-type de répétabilité, variance de répétabilité, ...

b) *Quantitativement*

Valeur au-dessous de laquelle est située, avec une probabilité spécifiée, la valeur absolue de la différence entre deux résultats individuels obtenus dans les conditions ci-dessus.

En l'absence d'indication, la probabilité est de 95 %.

2.86 reproductibilité :

a) *Qualitativement*

Étroitesse de l'accord entre les résultats individuels obtenus avec la même méthode sur une matière identique soumise à l'essai, mais dans des conditions différentes (opérateurs différents, appareils différents, laboratoires différents et/ou époques différentes).

NOTE — Les paramètres représentatifs de la dispersion de la population qui peut être associée aux résultats sont qualifiés du terme «reproductibilité».

Exemple : Écart-type de reproductibilité, variance de reproductibilité, ...

b) *Quantitativement*

Valeur au-dessous de laquelle est située, avec une probabilité spécifiée, la valeur absolue de la différence entre deux résultats individuels obtenus dans les conditions ci-dessus.

En l'absence d'indication, la probabilité est de 95 %.

NOTE — Si les termes 2.83 à 2.86 sont employés dans un sens autre que celui donné ci-dessus, ils doivent être spécialement et parfaitement définis, car l'expérience a montré que des ambiguïtés pouvaient survenir.

3 GENERAL TERMS RELATING TO METHODS OF SAMPLING

3.1 sampling : The procedure used to draw or constitute a sample.

NOTE — The sampling of bulk materials such as ores, etc., and continuous materials such as sheet metal, etc., requires the use of specific terms which are defined from 3.17 to 3.22.

3.2 sample : One or more items taken from a population and intended to provide information on the population and possibly to serve as a basis for a decision on the population or on the process which had produced it.

3.3 sampling unit : For the purpose of sampling, an item or, in a multistage sampling, a group of items, taken from the population.

3.4 sampling with replacement : A method of sampling in which each item taken and observed is returned to the population before the next item is taken.

The same item may appear several times in the sample.

3.5 sampling without replacement : A method of sampling in which items are taken from the population once only or successively without being returned to the population.

3.6 simple random sampling : The taking of n items from a population of N items in such a way that all possible combinations of n items have the same probability of being chosen.

3.7 subsample : A sample obtained by sampling in which the items to be taken are randomly distributed in parts of equal size which are prepared and tested separately.

The mean of the results of these tests is the sample mean. The total number of items to be taken should be, in this case, a multiple of the number of the subsamples.

3.8 duplicated sample : A subsample containing only two subsamples.

3.9 stratified sampling : Of a population which can be divided into different subpopulations (called strata), sampling carried out in such a way that specified proportions of the sample are drawn from the different strata.

3 TERMES GÉNÉRAUX RELATIFS AUX MÉTHODES D'ÉCHANTILLONNAGE

3.1 échantillonnage : Procédure utilisée pour tirer ou constituer un échantillon.

NOTE — L'échantillonnage des matières en vrac telles que des minerais, etc., et des matières continues telles que des tôles, etc., nécessite l'emploi de termes spécifiques qui sont définis de 3.17 à 3.22.

3.2 échantillon : Un ou plusieurs individus prélevés dans une population et destinés à fournir une information sur la population et, éventuellement, à servir de base à une décision concernant la population ou le procédé qui l'a produite.

3.3 unité d'échantillonnage : En vue d'un échantillonnage, individu ou, dans un échantillonnage à plusieurs degrés, groupe d'individus, prélevés dans la population.

3.4 échantillonnage avec remise (non exhaustif) : Mode d'échantillonnage selon lequel chaque individu prélevé et observé, est remis dans la population avant prélèvement de l'individu suivant.

Le même individu peut figurer plusieurs fois dans l'échantillon.

3.5 échantillonnage sans remise (exhaustif) : Mode d'échantillonnage selon lequel les individus sont prélevés en une seule fois ou successivement dans la population, mais sans remise dans celle-ci.

3.6 échantillonnage simple au hasard : Mode de prélèvement de n individus dans une population de N individus tel que toutes les combinaisons possibles de n individus aient la même probabilité d'être prélevées.

3.7 échantillon subdivisé : Échantillon obtenu au cours d'un échantillonnage dans lequel les individus à prélever sont répartis au hasard en n sous-échantillons d'égal effectif qui font respectivement l'objet d'une préparation et d'un essai séparés.

La moyenne des résultats de ces essais est la moyenne de l'échantillon. Le nombre total d'individus à prélever doit être, dans ce cas, un multiple du nombre de sous-échantillons.

3.8 échantillon dédoublé : Échantillon subdivisé comprenant seulement deux sous-échantillons.

3.9 échantillonnage stratifié : Les strates étant les différentes sous-populations en lesquelles la population à échantillonner peut être divisée, l'échantillonnage stratifié est effectué de sorte que des proportions spécifiées de l'échantillon soient tirées des différentes strates de la population totale.

3.10 periodic systematic sampling : If the N items in a population have been arranged on a systematic basis (for example in order of production) and numbered 1 to N , a periodic systematic sampling of n items consists in the taking of the n items numbered

$$h, h + k, h + 2k, \dots, h + (n-1)k$$

where h and k are whole numbers satisfying the relation

$$h + (n-1)k \leq N < h + nk$$

h being generally taken at random from the first k whole numbers.

NOTE — The word "periodic" may be omitted when there is no risk of confusion.

3.11 cluster sampling : A method of sampling in which the population is divided into aggregates (or clusters) of items bound together in a certain manner. A sample of these clusters is taken at random and all the items which constitute them are included in the sample.

NOTE

multistage cluster sampling : Cluster sampling with more than two stages, each sampling being made on aggregates (or clusters in which the clusters already obtained by the preceding sampling have been divided.

3.12 multistage sampling : A type of sampling in which the sample is selected by stages, the sampling units at each stage being subsampled from the larger units chosen at the previous stage.

3.13 primary sample : A sample taken from a population during the first stage of multistage sampling. (See 3.12.)

3.14 secondary sample : A sample taken from the primary sample during the second stage of multistage sampling. (See 3.12.)

3.15 final sample : A sample obtained at the final stage of multistage sampling. (See 3.12.)

3.16 sampling fraction : The ratio of the size of a sample to the size of the population, or subpopulation, from which it has been taken.

3.17 increment : A quantity of material taken at one time from a larger body of material.

3.18 test piece : A part obtained from an increment or from a specimen for the purpose of testing.

3.19 bulk sample; gross sample : An aggregation of the increments.

3.10 échantillonnage systématique périodique : Si les N individus d'une population ont été disposés d'après une certaine base systématique (par exemple suivant l'ordre de la production) et s'ils ont été numérotés de 1 à N , l'échantillonnage systématique périodique de n individus consiste à prélever les n individus numérotés

$$h, h + k, h + 2k, \dots, h + (n-1)k$$

où h et k sont des nombres entiers satisfaisant à la relation

$$h + (n-1)k \leq N < h + nk$$

h étant généralement pris au hasard parmi les k premiers nombres entiers.

NOTE — Le mot «périodique» peut être omis quand aucune confusion n'est à craindre.

3.11 échantillonnage en grappes à un degré : Mode d'échantillonnage selon lequel la population est divisée en sous-ensembles (ou grappes) d'individus liés ensemble d'une certaine manière. Un échantillon de ces grappes est prélevé au hasard et tous les individus qui les constituent sont compris dans l'échantillon.

NOTE

échantillonnages en grappes à plusieurs degrés : Échantillonnage en grappes à plus de deux degrés, chaque tirage portant sur des sous-ensembles (ou grappes) en lesquels ont été partagées les grappes déjà obtenues au tirage précédent.

3.12 échantillonnage à plusieurs degrés : Mode d'échantillonnage selon lequel l'échantillon est prélevé par degrés, les unités élémentaires d'échantillonnage à chaque degré étant sous-échantillonnées à partir des unités plus importantes obtenues au degré précédent.

3.13 échantillon primaire : Échantillon pris dans une population au premier degré d'un échantillonnage à plusieurs degrés. (Voir 3.12.)

3.14 échantillon secondaire : Échantillon pris dans l'échantillon primaire, au second degré d'un échantillonnage à plusieurs degrés. (Voir 3.12.)

3.15 échantillon final : Échantillon obtenu au dernier degré d'un échantillonnage à plusieurs degrés. (Voir 3.12.)

3.16 taux d'échantillonnage; fraction de sondage : Quotient de l'effectif d'un échantillon par l'effectif de la population, ou sous-population, dans laquelle il a été prélevé.

3.17 prélèvement élémentaire : Quantité de matière prélevée en une seule fois dans une quantité de matière plus importante.

3.18 éprouvette : Pièce obtenue à partir d'un prélèvement élémentaire ou d'un spécimen et destinée à être essayée.

3.19 échantillon global : Ensemble des prélèvements élémentaires.

3.20 sample preparation : For bulk materials, the set of material operations such as reduction of sizes, mixing, dividing, etc., necessary to transform a bulk sample into a laboratory or test sample.

NOTE — The sample preparation should not, as far as possible, modify its representativity of the sampled lot.

3.21 laboratory sample : A sample intended for laboratory inspection or testing.

3.22 test sample : A sample as prepared for testing.

3.20 préparation d'un échantillon : Pour une matière en vrac, ensemble des opérations matérielles telles que réduction de dimensions, mélange, division, etc., nécessaires pour faire passer l'échantillon global à l'état d'échantillon pour laboratoire ou pour essai.

NOTE — La préparation d'un échantillon ne doit pas, autant qu'il est possible, modifier sa représentativité du lot échantillonné.

3.21 échantillon pour laboratoire : Échantillon destiné à être utilisé pour un contrôle ou pour des essais en laboratoire.

3.22 échantillon pour essai : Échantillon dans l'état de préparation où il est soumis à l'essai.

STANDARDSISO.COM : Click to view the full PDF of ISO 3534-1977

4 TERMS RELATING TO SAMPLING INSPECTION

4.1 inspection : The process of measuring, examining, testing, gauging or otherwise comparing the unit with the applicable requirements.

4.2 quality control : The set of operations (programming, co-ordinating, carrying out) intended to maintain or to improve quality, and to set up the production at the most economical level which allows for customer satisfaction.

4.3 statistical quality control : Quality control using statistical methods (such as control charts and sampling plans).

4.4 sampling inspection : The inspection of a limited number of units, or of a limited quantity of material, taken at random from the lot or process according to a prescribed sampling plan.

4.5 batch; lot : A definite quantity of some commodity manufactured or produced under conditions which are presumed uniform.

4.6 inspection lot : A definite quantity of items submitted for inspection normally manufactured by one supplier under conditions that are presumed uniform.

4.7 consignment : A quantity of some commodity delivered at one time. The consignment may consist of one or more lots or parts of lots.

4.8 order : A quantity of a commodity ordered at one time from one producer. An order may consist of one or more consignments.

4.9 specification : A detailed statement of a set of requirements to be satisfied by a product, a material or a process, indicating, whenever appropriate, the procedure by means of which it may be determined whether the requirements given are satisfied.

NOTES

1 A specification may be a standard, a part of a standard, or independent of a standard.

2 As far as practicable, it is desirable that the requirements be expressed numerically in terms of appropriate units together with their limits.

4.10 tolerance zone : The zone of values in which a measurable characteristic is in conformity with its specification.

4 TERMES RELATIFS AU CONTRÔLE PAR ÉCHANTILLONNAGE

4.1 inspection; contrôle : Mesure, examen, essai, étalonnage ou toute autre comparaison de l'individu avec les spécifications exigées.

4.2 contrôle de la qualité : Gestion de la qualité, c'est-à-dire ensemble des opérations (prévision, coordination, réalisation) destinées à maintenir ou à améliorer la qualité, et à établir la production au niveau le plus économique qui tienne compte de la satisfaction de l'utilisateur.

4.3 contrôle statistique de la qualité : Contrôle de la qualité utilisant les méthodes statistiques (telles que les cartes de contrôle et les plans d'échantillonnage).

4.4 contrôle par échantillonnage : Contrôle d'un nombre limité d'individus, ou d'une quantité limitée de matière, prélevés au hasard dans un lot ou au cours d'une fabrication suivant un plan d'échantillonnage déterminé.

4.5 lot : Quantité définie d'une marchandise déterminée, fabriquée ou produite dans des conditions présumées uniformes.

4.6 lot pour inspection (de contrôle) : Nombre déterminé d'individus soumis à l'inspection (au contrôle) fabriqués normalement par un fournisseur dans des conditions présumées uniformes.

4.7 livraison : Quantité de marchandise transférée en une seule fois. La livraison peut être composée d'un ou plusieurs lots ou parties de lots.

4.8 commande : Quantité de marchandise commandée en même temps à un fournisseur. Une commande peut comporter une ou plusieurs livraisons.

4.9 spécification : Indication précise d'un ensemble de conditions à remplir par un produit, un matériau ou un procédé, incluant, si nécessaire, les méthodes qui permettent de déterminer si ces conditions sont remplies.

NOTES

1 Une spécification peut être une norme ou une partie de norme, ou encore être distincte d'une norme.

2 Dans la mesure où cela est réalisable, il est souhaitable que les conditions requises soient exprimées numériquement, avec leurs unités convenables, et en y incluant les limites.

4.10 zone de tolérance : Zone de valeurs à l'intérieur de laquelle un caractère mesurable est conforme à sa spécification.

4.11 tolerance limit : The limiting value (lower or upper) specified for a quantitative characteristic.

When there is one specified limit, it is called "single tolerance limit". When there are two limits, upper and lower, they are respectively called "upper tolerance limit" and "lower tolerance limit".

4.12 tolerance : The difference between the upper and lower tolerance limits.

4.13 lot inspection; acceptance inspection; receiving inspection

lot inspection : Any inspection of a lot carried out to determine its quality and take corresponding action.

receiving inspection : Lot inspection by a consumer on a lot delivered or offered for delivery.

acceptance inspection : Inspection to determine whether an item delivered or offered for delivery is acceptable.

NOTE — These terms are not mutually exclusive : receiving inspection is often acceptance inspection as well.

4.14 final inspection : Lot inspection carried out at the end of a production line.

4.15 original inspection : The term "original" applies to lots submitted to inspection for the first time, in order to distinguish them from lots which are to be re-submitted.

NOTE — In practice, when a lot has been rejected, the producer may be authorized to re-submit this lot after it has been modified in order to improve its quality (screening inspection, repairing, etc.).

4.16 100 % inspection : Inspection of all the items or of the whole of the material in a lot.

4.17 inspection by attributes : A method which consists in taking note, for every item of a population or of a sample taken from this population, of the presence or absence of a certain qualitative characteristic (attribute) and in counting how many items have or do not have this characteristic.

4.18 inspection by defect-counting : A method which consists in counting, over the items of a population or over a sample taken from this population, the number of defects (in the sense of 4.20) per item or per 100 items.

4.19 inspection by variables : A method which consists in measuring a quantitative characteristic for each item of a population or of a sample taken from this population.

4.20 defect : Any non-conformance of an item with specified requirements.

4.11 limite de tolérance : Valeur limite spécifiée (inférieure ou supérieure) d'un caractère quantitatif.

Lorsqu'il n'y a qu'une seule limite spécifiée, on l'appelle «limite simple de tolérance». Lorsqu'il y a deux limites, supérieure et inférieure, on les appelle respectivement «limite supérieure de tolérance» et «limite inférieure de tolérance».

4.12 tolérance : Écart entre les limites supérieure et inférieure de tolérance.

4.13 contrôle de réception :

a) Tout contrôle d'un lot effectué pour déterminer sa qualité et prendre la décision correspondante.

b) Contrôle effectué par un client sur un lot livré ou présenté à la livraison.

c) Contrôle pour déterminer si un article livré ou présenté à la livraison est acceptable.

NOTE — Le terme français «contrôle de réception» correspond aux trois termes anglais «lot inspection», «receiving inspection» et «acceptance inspection». Ces termes anglais ne sont d'ailleurs pas nettement distinctifs : on utilise indifféremment «receiving inspection» ou «acceptance inspection».

4.14 contrôle final : Contrôle effectué à la fin d'un processus de production.

4.15 contrôle en première présentation : Le terme «en première présentation» s'applique aux lots soumis pour la première fois au contrôle, afin de les distinguer des lots qui pourraient être présentés à nouveau.

NOTE — En pratique, lorsqu'un lot a été rejeté, le fournisseur peut être autorisé à le présenter à nouveau après modifications destinées à améliorer sa qualité (tri, réparation, etc.).

4.16 contrôle à 100 % : Contrôle de toutes les pièces ou de la totalité de la matière d'un lot.

4.17 contrôle par attributs : Contrôle dans lequel on constate, pour chacun des individus d'une population ou d'un échantillon prélevé dans cette population, la présence ou l'absence d'un certain caractère qualitatif (attribut) et on compte combien d'entre eux possèdent ou non ce caractère.

4.18 contrôle par nombre de défauts : Contrôle dans lequel on compte, sur les individus d'une population ou d'un échantillon prélevé dans cette population, le nombre de défauts (dans le sens de 4.20) par individu ou par 100 individus.

4.19 contrôle par mesures : Contrôle dans lequel on mesure un caractère quantitatif lié à chacun des individus d'une population ou d'un échantillon prélevé dans cette population.

4.20 défaut : Non-conformité d'un individu aux prescriptions imposées pour un caractère.

4.21 defective; defective item : An item containing one or more defects.

4.22 critical defect : A defect that, according to judgment and experience, is likely to result in hazardous or unsafe conditions for individuals using, maintaining, or depending upon the considered product, or that is likely to prevent performance of the function of a major end item.

4.23 major defect : A defect, other than critical, that is likely to result in a failure or to reduce materially the usability of the considered product for its intended purpose.

4.24 minor defect : A defect that is not likely to reduce materially the usability of the considered product for its intended purpose, or that is a departure from established specifications having little bearing on the effective use or operation of this product.

4.25 critical defective : An item which contains one or more critical defects; it may also contain major or minor defects.

4.26 major defective : An item which contains one or more major defects; it may also contain minor defects but no critical defects.

4.27 minor defective : An item which contains one or more minor defects but no critical or major defects.

4.28 mean number of defects per unit : The number of defects divided by the number of units on which it has been observed. Multiplied by 100, it gives the "mean number of defects per 100 units".

4.29 fraction defective; per cent defective : The number of defective items divided by the total number of items. Multiplied by 100, the fraction defective gives the "per cent defective".

4.30 sampling plan : A plan according to which one or more samples are taken in order to obtain information and possibly to reach a decision.

4.31 single sampling : A type of sampling which consists in taking only one sample per lot.

4.32 double sampling : A type of sampling which consists in taking possibly a second sample according to the information given by the first.

4.33 multiple sampling : A type of sampling which consists in taking possibly up to k successive samples, the decision on taking the i th sample ($i \leq k$) being dependent on the information given by the $(i - 1)$ previous samples.

4.21 défectueux; individu défectueux : Individu présentant un ou plusieurs défauts.

4.22 défaut critique : Défaut qui, d'après le jugement et l'expérience, est susceptible de conduire à un manque de sécurité ou à des risques d'accident pour les utilisateurs, le personnel d'entretien, ou ceux qui dépendent du produit considéré, ou bien qui pourrait empêcher l'accomplissement de la fonction d'un produit final plus important.

4.23 défaut majeur : Défaut qui, sans être critique, risque de provoquer une défaillance, ou bien de réduire de façon importante la possibilité d'utilisation du produit considéré pour le but qui lui est assigné.

4.24 défaut mineur : Défaut qui ne réduit vraisemblablement pas beaucoup la possibilité d'utilisation du produit considéré pour le but qui lui est assigné, ou qui traduit, par rapport aux spécifications fixées, une divergence n'entraînant que peu de conséquences appréciables sur l'utilisation ou le fonctionnement efficace de ce produit.

4.25 défectueux critique : Individu qui présente un ou plusieurs défauts critiques; il peut également présenter des défauts majeurs ou mineurs.

4.26 défectueux majeur : Individu qui présente un ou plusieurs défauts majeurs; il peut également présenter des défauts mineurs, mais aucun défaut critique.

4.27 défectueux mineur : Individu qui présente un ou plusieurs défauts mineurs, mais aucun défaut critique ou majeur.

4.28 nombre moyen de défauts par unité : Nombre de défauts divisé par le nombre d'unités sur lesquelles ils ont été observés. Multiplié par 100, il donne le «nombre moyen de défauts par 100 unités».

4.29 proportion de défectueux; pourcentage de défectueux : Nombre d'individus défectueux divisé par le nombre total d'individus. Multipliée par 100, la proportion de défectueux donne le «pourcentage de défectueux».

4.30 plan d'échantillonnage : Plan selon lequel on prélève un ou plusieurs échantillons en vue d'une information à obtenir et, éventuellement, d'une décision à prendre.

4.31 échantillonnage simple : Mode d'échantillonnage qui consiste à ne prélever qu'un seul échantillon par lot.

4.32 échantillonnage double : Mode d'échantillonnage qui consiste à prélever éventuellement un second échantillon selon l'information donnée par le premier.

4.33 échantillonnage multiple : Mode d'échantillonnage qui consiste à prélever éventuellement jusqu'à k échantillons successifs, la décision d'en prélever un j ème ($j \leq k$) étant fonction des informations données par les $(j - 1)$ échantillons précédents.

4.34 sequential sampling : A type of sampling which consists in taking successive items, or sometimes successive groups of items, but without fixing their number in advance, the decision to accept or reject the lot being taken as soon as the results permit it, according to rules laid down in advance.

4.35 curtailed inspection : In the case of single, double or multiple sampling, an inspection stopped as soon as the decision (acceptance or rejection) may be taken.

In the case of sequential sampling, inspection is stopped after inspecting a predetermined number even though the decision to accept or reject has not been reached. This decision is then taken in accordance with another predetermined rule.

4.36 continuous sampling plan : A sampling plan intended to cater to the particular circumstances of continuous production.

4.37 acceptance sampling plan : A sampling plan intended for determining the acceptance or the rejection of a lot.

4.38 acceptance : Agreement to take a lot as offered.

4.39 rejection : Refusal to take a lot as offered.

NOTE — When it is applied to the consignment of a producer, the term "rejection" means in a more liberal sense "non-acceptance of the lot under the conditions of the contract" (for example the lot may be placed in another class, subjected to a lowering of price, etc.).

4.40 acceptance number : In sampling inspection by attributes, the upper value of the number of defects or defectives found in the sample which involves the acceptance of the lot.

NOTE — In sampling by variables, other acceptance numbers may be defined in terms of the observed values and the permitted tolerances.

4.41 rejection number : In sampling inspection by attributes, the lower value of the number of defects or defectives found in the sample which involves the acceptance of the lot.

NOTE — In sampling by variables, the rejection number may be defined in terms of the observed values and the permitted tolerances.

4.42 probability of acceptance : The probability that a lot of a given quality will be accepted by a given sampling plan.

4.43 probability of rejection : The probability that a lot of a given quality will be rejected by a given sampling plan.

4.34 échantillonnage progressif : Mode d'échantillonnage qui consiste à prélever des individus successifs, ou quelquefois des groupes successifs d'individus, mais sans fixer leur nombre à l'avance, la décision d'acceptation ou de rejet du lot étant prise dès que les résultats obtenus le permettent, selon des règles fixées à l'avance.

4.35 contrôle tronqué : Dans le cas d'échantillonnage simple, double ou multiple, contrôle arrêté dès que la décision (acceptation ou rejet) peut être prise.

Dans le cas d'échantillonnage progressif, contrôle arrêté suivant une règle fixée à l'avance, bien que la décision d'acceptation ou de rejet n'ait pu être prise. Cette décision est alors prise selon une règle également fixée à l'avance.

4.36 plan d'échantillonnage continu : Plan d'échantillonnage adapté aux circonstances particulières d'une production continue.

4.37 plan de réception par échantillonnage : Plan d'échantillonnage destiné à déterminer l'acceptation ou le rejet d'un lot.

4.38 acceptation : Consentement à prendre un lot tel qu'il se présente.

4.39 rejet : Refus de prendre un lot tel qu'il se présente.

NOTE — Lorsqu'il s'applique à la livraison d'un fournisseur, le terme «rejet» s'entend au sens plus large de «non-acceptation du lot aux conditions du contrat» (par exemple, le lot peut être déclassé, subir une diminution de prix, etc.).

4.40 critère d'acceptation : Dans un contrôle par échantillonnage par attributs, plus grande valeur du nombre de défauts ou de défectueux trouvés dans l'échantillon contrôlé entraînant l'acceptation du lot.

NOTE — Dans un contrôle par échantillonnage par mesures, des critères d'acceptation peuvent être définis en fonction des valeurs observées et des tolérances admises.

4.41 critère de rejet : Dans un contrôle par échantillonnage par attributs, plus petite valeur du nombre de défauts ou de défectueux trouvés dans l'échantillon contrôlé entraînant le rejet du lot.

NOTE — Dans un contrôle par échantillonnage par mesures, des critères de rejet peuvent être définis en fonction des valeurs observées et des tolérances admises.

4.42 probabilité d'acceptation : Probabilité qu'un lot de qualité donnée soit accepté par application d'un plan d'échantillonnage déterminé.

4.43 probabilité de rejet : Probabilité qu'un lot de qualité donnée soit rejeté par application d'un plan d'échantillonnage déterminé.

4.44 operating characteristic curve for a sampling plan (OC curve) : A curve showing, for a given sampling plan, the probability of acceptance of a lot as a function of its actual quality.

4.45 producer's risk point : A point on the operating characteristic curve corresponding to some predetermined and usually low probability of rejection. This probability of rejection is called the "producer's risk".

In practical applications the producer's risk may be assigned a specified value (5 %) and the corresponding lot quality is then used as a parameter for choosing a sampling plan.

In other cases the lot quality may be specified, for example as an acceptable quality level or a process average, and the producer's risk will then be a variable quantity but of direct interest to the producer.

4.46 consumer's risk point : A point on the operating characteristic curve corresponding to a predetermined and usually low probability of acceptance. This probability is then called the "consumer's risk" and the corresponding lot quality is called the limiting quality.

In practical applications the consumer's risk is usually specified at 10 % or 5 %.

4.47 indifference point (point of control) (of the OC curve) : The point of the operating characteristic curve corresponding to probabilities of acceptance and rejection equal to 0,50.

4.48 process average : The quality of a production estimated from successive samples taken from lots in the original inspection.

4.49 limiting quality : In a sampling plan, a quality level which corresponds to a specified and relatively low probability of acceptance. It is the limiting lot quality characteristic that the consumer is willing to accept with a low probability that a lot of this quality would occur.

When expressed as per cent defective, it may be referred to as lot tolerance per cent defective.

4.50 lot tolerance per cent defective : A quality level which in a sampling plan corresponds to a specified and relatively low probability of acceptance (usually 10 %). (See 4.49.)

4.51 acceptable quality level : A quality level which in a sampling plan corresponds to a specified but relatively high probability of acceptance.

It is the maximum per cent defective (or the maximum number of defects per hundred units) that, for purposes of sampling inspection, can be considered satisfactory as a process average.

4.44 courbe d'efficacité d'un plan d'échantillonnage : Courbe montrant, pour un plan d'échantillonnage donné, la probabilité d'accepter un lot en fonction de sa qualité réelle.

4.45 point du risque du fournisseur : Point, sur la courbe d'efficacité, correspondant à une probabilité de rejet fixée à l'avance et généralement faible. Cette probabilité de rejet est appelée «risque du fournisseur».

En pratique, on peut donner au risque du fournisseur une valeur de 5 % et la qualité du lot correspondant est alors utilisée comme paramètre pour le choix d'un plan d'échantillonnage.

Dans d'autres cas, la qualité du lot peut être spécifiée, par exemple, comme niveau de qualité acceptable ou comme la moyenne d'un processus et le risque du fournisseur sera alors une variable, mais d'un intérêt direct pour le fournisseur.

4.46 point du risque du client : Point, sur la courbe d'efficacité, correspondant à une probabilité d'acceptation fixée à l'avance et généralement faible. Cette probabilité est alors appelée «risque du client» et la qualité du lot correspondant est appelée qualité limite.

En pratique, le risque du client est fixé à 10 % ou à 5 %.

4.47 point d'indifférence (de la courbe d'efficacité) : Point de la courbe d'efficacité où les probabilités d'acceptation et de rejets sont égales à 0,50.

4.48 qualité moyenne d'une fabrication : Qualité d'une fabrication estimée à partir d'échantillons successifs prélevés sur des lots contrôlés en première présentation.

4.49 qualité limite : Dans un plan d'échantillonnage, niveau de qualité qui correspond à une probabilité d'acceptation spécifiée et relativement faible. C'est la caractéristique limite de qualité d'un lot que le client ne veut accepter qu'avec une faible probabilité.

Lorsque cette caractéristique est exprimée en pourcentage de défectueux, on l'appelle niveau de qualité toléré.

4.50 niveau de qualité toléré : Niveau de qualité qui, dans un plan d'échantillonnage, correspond à une probabilité d'acceptation spécifiée et relativement faible (généralement 10 %). (Voir 4.49.)

4.51 niveau de qualité acceptable : Niveau de qualité qui, dans un plan d'échantillonnage, correspond à une probabilité d'acceptation spécifiée, mais relativement forte.

C'est le pourcentage maximal de défectueux (ou le nombre maximal de défauts par cent unités) qui, pour le contrôle par échantillonnage, peut être considéré comme satisfaisant en tant que caractéristique moyenne de la qualité de la production.

4.52 average outgoing quality : The average quality, expressed in per cent defectives or in number of defects per 100 items, of the product obtained after inspection, including not only lots accepted by the plan but also lots rejected by the plan which have been given 100 % inspection and in which all defectives have been replaced by non-defectives.

The average outgoing quality depends on the chosen sampling plan. It is a function of the quality of the lots submitted to inspection, and may be shown by a curve.

4.53 average outgoing quality limit : The maximum value of the average outgoing quality such as it was defined in 4.52.

4.54 inspection level : A characteristic of a sampling plan, chosen in advance and connecting the size of the sample(s) to the lot size.

4.55 normal inspection : The inspection which is used when there is no reason to think that the quality level of the production differs from the acceptable level provided for.

4.56 reduced inspection : The inspection, less severe than the normal inspection, to which the latter is switched when the inspection results of a number of lots allow the belief that the quality level of the production is high.

4.57 tightened inspection : The inspection, more severe than the normal inspection, to which the latter is switched when the inspection results of a number of lots allow the belief that the quality level of the production is poor.

4.58 average sample number : The average number of sample units inspected per lot in reaching a decision to accept or reject.

It is dependent on the actual quality level of the submitted lot.

4.59 assignable cause : A cause of systematic variation revealed by statistical methods.

4.60 process under control : A process the mean and variability of which remain stable.

4.61 control chart : A chart on which limits are drawn and on which are plotted values of any statistic computed from successive samples of a production.

The control chart is used to investigate if a process may be considered to be under control.

The statistics which are used (mean, range, per cent defective, etc.) define the different kinds of control charts.

4.52 qualité moyenne après contrôle : Qualité moyenne, exprimée en pourcentage de pièces défectueuses ou en nombre de défauts par 100 individus, du produit obtenu après contrôle et comprenant, non seulement les lots acceptés par le plan, mais aussi les lots refusés par le plan après contrôle à 100 % et remplacement des pièces défectueuses par des non défectueuses.

La qualité moyenne après contrôle dépend du plan d'échantillonnage choisi. Elle est fonction de la qualité des lots soumis au contrôle et peut être représentée par une courbe.

4.53 limite de la qualité moyenne après contrôle : Valeur maximale de la qualité moyenne après contrôle, telle qu'elle a été définie en 4.52.

4.54 niveau de contrôle : Caractéristique d'un plan d'échantillonnage, choisie *a priori* et reliant l'effectif du ou des échantillons à celui du lot.

4.55 contrôle normal : Contrôle utilisé lorsqu'il n'y a pas de raison de penser que le niveau de qualité de la fabrication diffère du niveau acceptable prévu.

4.56 contrôle réduit : Contrôle, moins sévère que le contrôle normal, auquel on passe lorsque les résultats du contrôle d'un certain nombre de lots permettent de penser que le niveau de qualité de la fabrication est élevé.

4.57 contrôle renforcé : Contrôle, plus sévère que le contrôle normal, auquel on passe lorsque les résultats du contrôle d'un certain nombre de lots permettent de penser que le niveau de qualité de la fabrication est bas.

4.58 effectif moyen contrôlé : Valeur moyenne du nombre d'unités contrôlées par lot pour prendre la décision d'accepter ou de rejeter.

Cette valeur dépend du niveau de qualité réel des lots présentés au contrôle.

4.59 cause assignable : Cause de variation systématique, mise en évidence par des méthodes statistiques.

4.60 processus sous contrôle : Processus dont la moyenne et la variabilité restent stables.

4.61 carte de contrôle : Graphique sur lequel sont tracées des limites et où l'on reporte les valeurs d'une statistique calculées sur des échantillons successifs d'une fabrication.

La carte de contrôle est utilisée pour s'assurer qu'un processus peut être considéré comme étant sous contrôle.

Les statistiques utilisées (moyenne, étendue, proportion de défectueux, etc.) définissent les différents types de cartes de contrôle.

4.62 cumulative sum chart; cusum chart : A control chart on which are plotted calculated values of a statistic for a series of successive samples. The ordinate of each point plotted represents the ordinate of the immediately preceding point increased by the value of the statistic computed from the last sample.

Cusum charts are interpreted by comparison with limits which may vary in position for each point.

4.63 upper and/or lower control limits : In a control chart, the limit below which (upper limit) or above which (lower limit), or the limits between which the statistic under consideration lies with a very high probability when the process is under control.

When the value of the statistic computed from a sample is above the upper limit or below the lower limit, a corrective action should generally be made on the process of production.

4.64 upper and/or lower warning limits : In a control chart, the limit below which (upper limit) or above which (lower limit), or the limits between which the statistic under consideration lies with a high probability when the process is under control.

When the value of the statistic computed from a sample is outside the warning limits but inside the control limits, an increased supervision of the process of production is generally necessary.

4.62 carte à somme cumulée : Carte de contrôle sur laquelle sont portées les valeurs calculées d'une statistique pour une série d'échantillons successifs. L'ordonnée de chaque point est l'ordonnée du point précédent augmentée de la valeur de la statistique calculée depuis le dernier échantillon.

Les cartes à somme cumulée sont interprétées par comparaison avec des limites qui peuvent varier pour chaque point.

4.63 limites supérieure et/ou inférieure de contrôle : Dans une carte de contrôle, limite en deçà de laquelle (limite supérieure) ou au-delà de laquelle (limite inférieure), ou limites entre lesquelles la statistique considérée a une très forte probabilité de se situer lorsque le processus est sous contrôle.

Lorsque la valeur de la statistique, calculée sur un échantillon, s'inscrit au-dessus de la limite supérieure ou au-dessous de la limite inférieure, une action correctrice doit généralement être exercée sur le processus de fabrication.

4.64 limites supérieure et/ou inférieure de surveillance : Dans une carte de contrôle, limite en deçà de laquelle (limite supérieure) ou au-delà de laquelle (limite inférieure), ou limites entre lesquelles la statistique considérée a une forte probabilité de se situer lorsque le processus est sous contrôle.

Lorsque la valeur de la statistique, calculée sur un échantillon, s'inscrit hors des limites de surveillance tout en restant entre les limites de contrôle, une surveillance accrue du processus de fabrication est généralement nécessaire.