

EU 統計局

ワーキング・グループ “ 統計における品質評価 ”  
第 6 回会議，ルクセンブルグ，2003 年 10 月 2-3 日

ITEM 4.2C :  
方法論関係資料

## 統計品質に関する用語集（対訳）

（ 翻訳：独立行政法人 統計センター 研究センター，平成 17 年 1 月 ）



## 品質用語集

本用語集は、“How to make a Quality Report”(ドラフト 2003 年 10 月)と題する Eurostat のハンドブックで使用されている主要専門用語をカバーしている。本用語集は、意図的に各用語を短い文章で定義し、使用された原典を掲げている。より詳しく理解したい読者は、主要文献を参照すること。2002 年 5 月の Eurostat Glossary (Eurostat/A4/Quality/02/General/Glossary) は、本用語集に差替えられる。

訳注・用語の中には定訳がない、あるいは、対応する適切な日本語が存在しないか標準的なものがないと思われるものが多く、本資料で用いた訳語の中には適訳とは言い難いものが含まれている恐れがある。このため、対訳形式を採り、適宜、原文を参照できるようにした。また、原文には、誤りや、やや不適切と思われる箇所や、統一性を欠く用語がみられることや、直訳した場合誤解を招く恐れがある文章があるため、適宜、訳注を付した。

Name	Alternative name	Definition	Source	名称	別名	定義
Accessibility		Accessibility refers to the physical conditions in which users can obtain data: where to go, how to order, delivery time, clear pricing policy, convenient marketing conditions (copyright, etc.), availability of micro or macro data, various formats (paper, files, CD-ROM, Internet...), etc	Eurostat, Oct. 2003, "DEFINITION OF QUALITY IN STATISTICS", (Eurostat/A4/Quality/03/General/Definition)	アクセス可能性		アクセス可能性とは、ユーザーがデータを入手するための物理的条件に関連している。すなわち、どの場所に行き、どのように発注するか、引渡しまでの所要時間、明確な料金設定、マーケティングのための条件（著作権など）、マイクロデータ又はマクロデータの入手可能性、フォーマット（紙、ファイル、CD-ROM、インターネット、・・・）等である。
Accuracy		Accuracy in the general statistical sense denotes the closeness of computations or estimates to the exact or true values.	F.H.C. Marriott, 1990, A dictionary of statistical terms, Longman Scientific & Technical, Fifth edition, p.223	正確性		一般的に統計的意味での正確性は、計算値や推定値が、正確値又は真値との近さのことである。

Bias	Generally, an effect which deprives a statistical result of representativeness by systematically distorting it, as distinct from a random error which may distort on any one occasion but balances out on the average.	F.H.C. Marriott, 1990, A dictionary of statistical terms, Longman Scientific & Technical, Fifth edition, p.223	偏り	一般的に、系統的に歪めることによって統計的結果の代表性を奪う影響のことであり、それぞれの機会では歪める作用を有するが平均的にみると相殺されるランダムな誤差と区別される。
Clarity	Clarity refers to the data's information environment whether data are accompanied with appropriate metadata, illustrations such as graphs and maps, whether information on their quality also available (including limitation in use...) and the extent to which additional assistance is provided by the NSI.	Eurostat, Oct. 2003, "DEFINITION OF QUALITY IN STATISTICS", (Eurostat/A4/Quality/03/General/Definition)	明確性	明確性は、データの情報環境に関連している。データが適切なメタデータ、グラフやマップなどの図解を伴っているかどうか、データの品質に関する情報(利用上の制約などを含む)が入手可能かどうか、国家統計機関から追加的支援がどの程度提供されるか等である。
Coding	Coding is a technical procedure for converting verbal information into numbers or other symbols, which can be more easily counted and tabulated	Federal Committee on Statistical Methodology, December 1978, - Glossary of Nonsampling Error Terms: An Illustration of a Semantic Problem in Statistic, Statistical Policy Working Paper 4.	(分類) 格付, 符号化	格付は、テキスト情報を、数値やその他のシンボルに変換し、集計及び製表をより容易に行えるようにするための技術的手順のことである。
Coding error	The attribution of an incorrect code to a survey response.	Glossary, Previous version	格付誤り	調査回答に誤った分類区分に格付してしまうこと
Coefficient of variation	A measure of spread for a set of data defined as the ratio of the standard deviation to the mean.	B.S.EVERITT, 2002, The Cambridge dictionary of Statistics, Cambridge University Press. Second edition, p. 410	変動係数	標準偏差の平均値に対する比として定義され、データの分布の広がり度合いの尺度となる。
Coherence	Coherence of statistics is their adequacy to be reliably combined in different ways and for various uses. It is, however, generally easier to show cases of incoherence than to prove coherence. When	Eurostat, Oct. 2003, "DEFINITION OF QUALITY IN STATISTICS", (Eurostat/A4/Quality/03/General/Definition)	一貫性	統計間の一貫性は、様々な用途のために異なる方法で、それらを結合することの適切さを言う。ただし、一般的には、一貫性の存在を証明するよりも、

		originating from a single source, statistics are normally coherent in the sense that elementary results derived from the concerned survey can be reliably combined in numerous ways to produce more complex results. When originating from different sources, and in particular from statistical surveys of different nature and/or frequencies, statistics may not be completely coherent in the sense that they may be based on different approaches, classifications and methodological standards. Conveying neighbouring results, they may also convey not completely coherent messages, the possible effects of which, users should be clearly informed of.	/Definition)			矛盾例を示す方が容易である。単一のソースから作成された統計の場合には、当該調査から得られた基本結果を様々な方法によって組合せ、より複雑な結果を作り出すことができるという意味で、通常は一貫性があるといえる。複数のソースから作成された統計、特に、性質や周期の異なる統計調査をもとに作成された統計の場合には、それぞれが異なるアプローチや分類、基準に基づいているかもしれないという意味で完全な一貫性がないかもしれない。一致度が高いはずの統計の結果に一貫性がないこともありうるため、予想される影響について、ユーザーに明確に情報提供するべきである。
Comparability		Comparability is the extent to which differences between statistics from different geographical areas, non-geographical domains, or over time, can be attributed to differences between the true values of the statistics	Eurostat, Oct. 2003, "DEFINITION OF QUALITY IN STATISTICS", (Eurostat/A4/Quality/03/General/Definition)	比較可能性		比較可能性とは、異なる地域間、地域以外のドメイン間又は時点間における統計値の差違が、統計の真値の差違といえる範囲のことである。
Completeness		Completeness is the extent to which all statistics that are needed are available. It is usually described as a measure of the amount of available data from a statistical system compared to the amount that was expected to be obtained	Eurostat, Oct. 2003, "DEFINITION OF QUALITY IN STATISTICS", (Eurostat/A4/Quality/03/General/Definition)	完全性		完全性は、必要とされる統計のうち入手可能な範囲を言う。通常は、入手可能と期待されるデータの量に対する実際に統計システムから入手可能なデータの量の比を尺度として表される。
Confidence Interval		A a% confidence interval for an unknown population parameter $\theta$ , is an interval, calculated from sample values by a procedure such that if a large number of independent samples is taken, a% of the intervals obtained will contain $\theta$ .	Oxford dictionary of statistics"	信頼区間		未知の母集団パラメータ・ $\theta$ についての a% の信頼区間とは、多数の標本を互いに独立に抽出した場合に、その区間にパラメータ・ $\theta$ が含まれる確率が a% となるような手順で算出された区間

					のことである。
Consistent estimator		An estimator which converges in probability, as the sample size increases, to the parameter of which it is an estimator.	F.H.C. Marriott, 1990, A dictionary of statistical terms, Longman Scientific & Technical, Fifth edition, p.223	一致推定量	標本数が増加すると、パラメータの真値に確率収束する推定量
Constant scope		Constant scope refers to a frame population. It should not include births, deaths, mergers and demergers and the population should be identical for both periods from which indexes are estimated. The contrary of full scope.	Glossary, Previous version	コンスタント・スコープ	コンスタントスコープは、フレーム母集団に関連している。コンスタントスコープは、出生、死亡、合併、解散を含めず、指数を推定する両時点で同一となる。フルスコープの反対語。
Coverage error		Coverage error is the error associated with the failure to included some population units in the frame used for sample selection (undercoverage) and the error associated with the failure to identify units represented on the frame more than once (overcoverage). The source of coverage error is the sampling frame itself.	Statistical Policy working paper 31, "Measuring and reporting sources of error in survey", FCSM, Subcommittee on Measuring and Reporting the Quality of Survey Data	カバレッジ誤差	カバレッジ誤差は、標本抽出に使用するフレームに含まれなかった母集団ユニット(アンダーカバレッジ)が存在することに伴う誤差及びフレームに重複して含まれている母集団ユニット(オーバーカバレッジ)が存在することに伴う誤差のことである。カバレッジ誤差は、サンプリング・フレーム自身から発生する。
Cut-off threshold		A cut-off threshold is used, mainly for cost or burden reasons, to exclude from the target population (hence from the frame) units contributing very little to the requested statistics, small businesses for instance. the contribution from the population below the threshold can either be deemed negligible or be estimated by using a model	Glossary, Previous version (revised)	カットオフ閾値	カットオフ閾値は、主にコストと負担の理由から、目標母集団(すなわちフレーム)から、求められている統計への貢献度が極めて小さいユニット、例えば小企業を除外する場合に使用される。カットオフ閾値以下の母集団からの貢献は、無視できるとみなされるか、或いはモデルを使って推定される。
Editing, Edits		Data <i>editing</i> is the application of checks that identify missing, invalid or inconsistent entries or that point to data records that are potentially in error.	(Statistics Canada Quality Guidelines, 3rd edition, October 1998, page 35)	エディティング, エディツ	データエディティングとは、欠測、無効又は矛盾するデータを特定するため、又は誤りの可能性があるデータレ

				ト		コードを指摘するためにチェックを適用することである。
Electronic Data interchange (EDI)	Electronic document interchange	Electronic Data Interchange (EDI) is the computer-to-computer exchange of business data in a publicly published and globally standardized format.	Wikipedia, the free encyclopedia available at <a href="http://www.wikipedia.org/">http://www.wikipedia.org/</a>	電子データ交換 (EDI)		電子データ交換 (EDI) とは、国際的に標準化され、公表されているフォーマットによってビジネスデータをコンピューター間で交換することを言う。
Error		In general, a mistake or error in the colloquial sense. There may, for example, be a gross error or avoidable mistake; an error of reference, when data concerning one phenomenon are attributed to another; copying errors; an error of interpretation. In a more limited sense the word error is used in statistics to denote the difference between an occurring value and its true or expected value. There is here no imputation of mistakes on the part of a human agent; the deviation is a chance effect. In this sense we have, for example, errors of observations, errors in equations, errors of the first and second kinds in the testing hypothesis, and the error band surrounding an estimate; and also the Normal curve of errors itself.	F.H.C. Marriott, 1990, A dictionary of statistical terms, Longman Scientific & Technical, Fifth edition, p.223	誤差, 誤り		一般的には日常会話的意味の誤り又は間違いのことである。例えば、ひどい間違いや避けることのできた間違い; 調査対象の間違い; 転記ミス; 解釈の誤りなどがある。より限定的な意味では、統計で用いられる誤差という言葉は、計測値とその真値や期待値との差を意味する。この場合には、人間の側の間違いの含意はなく、偏差は偶然の効果である。この意味では、観測誤差、等式上の誤差、仮説検定における第1種と第2種の誤りなどや、推定値の周りの誤差の範囲、また誤差自身の正規曲線などがある。
Estimate		In the strict sense an estimate is the particular value yielded by an estimator in a given set of circumstances.	F.H.C. Marriott, 1990, A dictionary of statistical terms, Longman Scientific & Technical, Fifth edition, p.223	推定値		厳密な意味では、推定値とは、所定の条件の下で推定量により得られる特定の値のことである。
Estimator		An estimator is a rule or method of estimating a constant of a population. It is usually expressed as a function of sample values and hence is a variate whose distribution is of great importance in assessing the reliability of the estimate to which it leads.	F.H.C. Marriott, 1990, A dictionary of statistical terms, Longman Scientific & Technical, Fifth edition, p.223	推定量		推定量は、母集団の定数を推定する規則又は方法である。一般的には、標本値の関数としてあらわされる。従って、変量であり、その分布は、推定量から導出される推定値の信頼性を評価する上で重要である。

Expected Value	Expectation value, expectation	The hypothetical averages from the conceived replicates of the survey all conducted under the same essential conditions.	Federal Committee on Statistical Methodology, December 1978, - Glossary of Nonsampling Error Terms: An Illustration of a Semantic Problem in Statistic, Statistical Policy Working Paper 4.	期待値	期待値 , 期待	仮に同一の基本条件の下で調査を繰り返した場合に得られる仮説的な平均値
Frame		The frame consists of previously available descriptions of the objects or material related to the physical field in the form of maps, lists, directories, etc., from which sampling units may be constructed and a set of sampling units selected; and also information on communications, transport, etc., which may be of value in improving the design for the choice of sampling units, and in the formation of strata, etc.	Federal Committee on Statistical Methodology, December 1978, - Glossary of Nonsampling Error Terms: An Illustration of a Semantic Problem in Statistic, Statistical Policy Working Paper 4.	フレーム , 枠		フレームは、地図、リスト、名簿などの形式で対象又はデータを記述したもので、事前に利用可能なものから成る。これらから、抽出対象ユニットが構成され、抽出対象ユニットの1組が選定される。連絡先、交通手段など、標本設計の改善や抽出層の設定に役立つ情報も含まれる。
Full scope		Full scope refers to a frame population, it should cover births, deaths, mergers and demergers, during the period from which indexes are estimated. The contrary of constant scope.	Glossary, Previous version	フルスコープ		フルスコープは、フレーム母集団に関連している。フルスコープは、指数推定の対象時点間の出生、死亡、合併、解散を含める。コンスタント・スコープの反対語。
Hedonic Method		The hedonic method is a regression technique used to estimate the prices of qualities or models that are not available on the market in particular periods, but whose prices in those periods are needed in order to be able to construct price relatives.	UN, Statistics Division, Glossary, available at <a href="http://unstats.un.org/unsd/sna1993/glossary.asp">http://unstats.un.org/unsd/sna1993/glossary.asp</a>	ヘドニック法		ヘドニック法は、特定時点において市場で入手できない品質又はモデルがあるものの、価格比を算出するにはその時点の価格が必要な場合に、回帰分析法により価格を推定する方法である。
Imputation		<i>Imputation</i> is the process used to resolve problems of missing, invalid or inconsistent responses identified during editing. This is done by changing some of the responses or missing values on the record being edited to ensure that a plausible,	(Statistics Canada Quality Guidelines, 3rd edition, October 1998, page 38)	補定		補定とは、エディティングにより特定された回答の欠落、無効又は矛盾を解決するための処理ことである。エディットしているレコード上の一部の回答又は欠測値を、妥当性が保証される



		internally coherent record is created.				ように変更し,レコード内が一貫性のあるものを作成する。
Inlier		An inlier is a data value that lies in the interior of a statistical distribution and is in error. Because inliers are difficult to distinguish from good data values they are sometimes difficult to find and correct. A simple example of an inlier might be a value in a record reported in the wrong units, say degrees Fahrenheit instead of degrees Celsius.	UNECE, CES, 1997, Work Session on Statistical Data Editing, Working Paper 22, PROBLEMS WITH INLIERS, US Bureau of Census, William E. Winkler.	インライア		インライアとは,統計分布の内側に位置するが,誤っているデータ値のことである。インライアを,正しいデータ値と区別することは困難であるため,これらを見出し訂正することが難しい場合がある。インライアの単純な例としては,摂氏の代わりに華氏を用いるなど,誤った単位で報告された数値を挙げることができる。
Interviewer error		Interviewer error are associated with effects on respondents' answers stemming from the different ways that interviewers administer the same survey. Examples of these errors include the failure to read the question correctly (leading to response errors by the respondent), delivery of the question with an intonation that influences the respondent's choice of answer, and failure to record the respondent's answer correctly.	Paul P. Biemer, Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz, Seymour Sudman,1991,"Measurement errors in survey", John Wiley & Sons, p. 760	調査員(面接者)誤差		調査員誤差とは,同一の調査であっても調査員のやり方の違いが回答者の回答に及ぼす効果に伴う誤差のことである。これらの誤差の例としては,正しく質問を読まないこと(回答者の回答誤りにつながる),回答者の回答の選択に影響を与えるようなイントネーションで質問を行なうこと,回答者の回答を正しく記録しないことなどが挙げられる。
Item non-response		Item non-response occurs when a respondent provides some, but not all, of the requested information, or if the reported information is not usable.	Statistical Policy working paper 31, "Measuring and reporting sources of error in survey", FCSM, Subcommittee on Measuring and Reporting the Quality of Survey Data	項目非回答		項目非回答は,回答者が,要求された情報の全部でなく一部しか提供しない場合,或いは報告された情報が使用不能な場合に発生する。
Item response rate		The item response rate is the ratio of the number of eligible units responding to an item to the number of responding units eligible to have responded to the item.	Madow, W., Nisselson, H., and Olkin, I., 1983, Incomplete Data in Sample Surveys, New York: Academic Press	項目回答率		項目回答率は,その項目に関する適格回答ユニット数に対する,その項目に回答した適格回答ユニット数の比を言う。

Marco-editing		A macro-edit detects individual errors by checks on aggregated data, or checks applied to the whole body of records. The checks are typically based on the models, either graphical or numerical formula based, that determine the impact of specific fields in individual records on the aggregate estimates.	UNECE, Glossary of Terms on Statistical Data Editing, CES Methodological material, Geneva, 2000, also available at <a href="http://amrads.jrc.cec.eu.int/kb/se/glossary/g-source.htm">http://amrads.jrc.cec.eu.int/kb/se/glossary/g-source.htm</a>	マクロ・エディティング		マクロ・エディットは、集計値のチェック或いは全レコードを対象としたチェックにより個々のレコードの誤りを検出する。これらのチェックは、典型的にはグラフィカルあるいは数式ベースのモデルに基づいて行なわれ、集計値に対する個々のレコードの影響を決定する。
Mean square error		The expected value of the square of the difference between an estimator and the true value of a parameter. If the estimator is unbiased then the mean square error is simply the variance of the estimator. For a biased estimator the mean squared error is equal to the sum of the variance and the square of the bias.	B.S.EVERITT, 2002, The Cambridge dictionary of Statistics, Cambridge University Press. Second edition, p. 410	平均平方誤差		母数の推定量と真値の差の2乗の期待値。推定量が不偏である場合は、平均平方誤差は単純に推定量の分散である。偏りのある推定量の場合には、平均平方誤差は、分散と偏りの2乗の合計に等しい。
Measurement error		Measurement error refers to error in survey responses arising from the method of data collection, the respondent, or the questionnaire (or other instrument). It includes the error in a survey response as a result of respondent confusion, ignorance, carelessness, or dishonesty; the error attributable to the interviewer, perhaps as a consequence of poor or inadequate training, prior expectations regarding respondents' responses, or deliberate errors; and error attributable to the wording of the questions in the questionnaire, the order or context in which the questions are presented, and the method used to obtain the responses.	Paul P. Biemer, Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz, Seymour Sudman, 1991, "Measurement errors in survey", John Wiley & Sons, p. 760	測定誤差		測定誤差は、データ収集方法、回答者又は調査票(又はその他の調査インストルメント)から生じる調査回答上の誤りに関連する。測定誤差には、回答者の混乱、無知、不注意、不誠実の結果によるもの、不十分な調査員研修、回答者の回答に対する事前の期待や故意の誤りなど調査員に帰されるもの、調査票上の質問の言葉づかいや質問順序、回答を得るために用いられた方法などに帰されるものが含まれる。
Micro editing		Finding errors by inspection of individual observations. Editing done at the record, or questionnaire level.	UNECE, Glossary of Terms on Statistical Data Editing, CES Methodological material, Geneva, 2000, also available at	ミクロ・エディティング		個々の観測値の検査による誤りの発見。レコードもしくは調査票レベルで行なわれるエディティング。

			<a href="http://amrads.jrc.cec.eu.int/kbase/glossary/g-source.htm">http://amrads.jrc.cec.eu.int/kbase/glossary/g-source.htm</a>			
Misclassification		Misclassification refers to allocating statistical units in a wrong class, in relation to a given classification. For instance, a business is classified in Trade instead of Industry.	Glossary, Previous version	誤分類		誤分類は、所定の分類に関して、誤った分類区分に統計ユニットを格付することを言う。例えば、ある企業を、工業の代わりに(誤って)商業に格付してしまうことである。
Model assumption error		Model assumption are errors that occur due the use of methods, such as calibration, generalized regression estimator, calculation based on full scope or constant scope, benchmarking, seasonal adjustment and other models not included in the preceding accuracy components, in order to calculate statistics or indexes.	Glossary, Previous version	モデル前提誤差		モデル前提誤差とは、統計や指数を算出するために、キャリブレーション、一般化回帰推定量、フルスコープ又はコンスタントスコープに基づく計算、ベンチマーキング、季節調整やその他のモデルを用いることによって生じる誤差のことであり、他の種類の誤差に該当しないものである。(訳注・原文にはやや不適切な箇所がある。)
Non probability sample	Non-random Sample	Non-probability Sample is a sample in which the selection of units is based on factors other than random chance, e.g. convenience, prior experience or the judgment of the researcher. Examples of nonprobability samples are: convenience, judgmental, quota and snowball.	ESOMAR, Glossary of Marketing Research Terms, <a href="http://www.esomar.nl">http://www.esomar.nl</a>	非確率的標本	非無作為標本	非確率的標本は、ユニットの選定が、便宜的、過去の経験、研究者の判断など無作為抽出以外のファクターに基づいている標本である。非確率的標本の例としては、便宜的、判断、割当、雪だるま方式などがある。
Non response		Non response is a form of nonobservation present in most surveys. Nonresponse means failure to obtain a measurement on one or more study variables for one or more elements k selected for the survey. The term encompasses a wide variety of reasons for nonobservation: "impossible to contact", "not at home", "unable to answer", "incapacity", "hard core refusal", "inaccessible", "unreturned questionnaire", and others. In the first two cases contact with the selected element is never established.	Carl-Erik Sarndal, Bengt Swensson, Jan Wretman, 1992, "Model assisted survey sampling", Springer - Verlag New York, p.694	非回答		非回答は、大部分の調査で存在する非観測の一形態である。非回答は、調査のために選定されたひとつ又は複数の要素(訳注・ユニットのことか?)についてのひとつ又は複数の調査項目の測定値を得ることができなかったことを意味する。この用語には、非観測の様々な理由が含まれている。すなわち、“接触不能”、“不在”、“回答不能”、“能力欠如”、“拒否”“アクセ

					ス不能”“質問票未回収”その他である。最初の2つのケースは、選定されたエレメントとの接触が全くとれなかったことを意味する。
Non response bias		see non response error		非回答バイアス	非回答誤差を参照のこと
Non response error		Nonresponse errors, which occur when the survey fails to get a response to one, or possibly all, of the questions. Nonresponse causes both an increase in variance, due to the decrease in the effective sample size and/or due to the use of imputation, and may cause a bias if the nonrespondents and respondents differ with respect to the characteristic of interest.	Statistics Canada, October 1998, "statistics Canada Quality Guidelines", 3 <sup>rd</sup> edition, Catalogue no. 12-539-XIE, p. 101	非回答誤差	非回答誤差とは、調査において、質問のひとつ又は全部に対する回答を得られなかったことによって生じる誤差のことである。非回答は、有効な標本数の減少及び/若しくは補定の利用による分散拡大の原因となる。また、非回答者と回答者の特性が異なる場合には、偏りが発生する可能性がある。
Non sampling error		An error in sample estimates which cannot be attributed to sampling fluctuations. Such errors may arise from many different sources such as defects in the frame, faulty demarcation of sample units, defects in the selection of sample units, mistakes in the collection of data due to personal variations or misunderstandings or bias or negligence or dishonesty on the part of the investigator or of the interviewee, mistakes at the stage of the processing of the data, etc.	F.H.C. Marriott, 1990, A dictionary of statistical terms, Longman Scientific & Technical, Fifth edition, p.223	非標本誤差	サンプリングの変動に理由を求められない標本推定値の誤差。このような誤差は、フレームの欠陥、抽出対象ユニットの誤った区分、抽出対象ユニットの選定の欠陥、調査員や被調査者の個人差、誤解、偏り、怠慢、不誠実によるデータ収集上の誤り、データ処理段階の誤りなど多くの異なる原因から発生する可能性がある。
Out of scope units		Units that should not be included in the sampling frame because they do not belong to the target population in the reference period. If enumerated, they cause over-coverage.	Glossary, Previous version	調査対象外のユニット	調査対象時点の目標母集団に属していないために、抽出フレームに含めてはならないユニット。

Outlier		In a sample of $n$ observations it is possible for a limited number to be so far separated in value from the remainder that they give rise to the question whether they are not from a different population, or that the sampling technique is a fault. Such values are called outliers.	F.H.C. Marriott, 1990, A dictionary of statistical terms, Longman Scientific & Technical, Fifth edition, p.223	外れ値		観測サンプル $n$ 個において、他から大きく外れた少数の数値が存在することがありうる。この場合、この大きく外れた値は、異なる母集団に属するかもしれないといった疑問や、サンプリング法に誤りがあるかもしれないといった疑問が生じる。このような値は外れ値と呼ばれる。(訳注・原文に誤りがあると思われる。)
Over coverage		Over-Coverage arises from the presence in the frame of units not belonging to the target population and of units belonging to the target population that appear in the frame more than once.	Ad-Hoc	オーバー・カバレッジ		オーバー・カバレッジは、目標母集団に属していないユニットがフレーム内に含まれている場合や、重複して含まれている場合に発生する。
Population		<i>Population</i> is the total membership or population or "universe" of a defined class of people, objects or events. There are two types of population, viz, target population and survey population. A target population is the population outlined in the survey objects about which information is to be sought and a survey population is the population from which information can be obtained in the survey. The target population is also known as the scope of the survey and the survey population is also known as the coverage of the survey. For administrative records the corresponding populations are: the "target" population as defined by the relevant legislation and regulations, and the actual "client population".	(United Nations Glossary of Classification Terms. Prepared by the Expert Group on International Economic and Social Classifications. Available at <a href="http://www.un.org/Depts/unsd/class/glossary_short.htm">http://www.un.org/Depts/unsd/class/glossary_short.htm</a> )	母集団		母集団とは、所定の種類の人、対象物、事象の全個体群或いは全体を言う。母集団には2種類ある。すなわち、目標母集団と調査母集団である。目標母集団は、情報が求められている調査目的において輪郭が描かれる集団であり、調査母集団は、調査によって情報を得ることのできる集団である。目標母集団は、また、調査対象とも呼ばれており、調査母集団は、調査カバレッジとも呼ばれている。行政記録において対応する集団は、該当法および規則により規定されている“目標”母集団と実際の“顧客母集団”である。
Probability sample		Probability sampling is an approach to sample selection that satisfies certain conditions, which, for the case of selecting elements directly from the population, are described as follows: 1)we can define the set of sample, $\{S1, S2, \dots, Sm\}$ , that are possible	Carl-Erik Sarndal, Bengt Swensson, Jan Wretman, 1992, "Model assisted survey sampling", Springer - Verlag New York, p.694	確率的標本		確率的サンプリングは、所定の条件を満たす標本抽出アプローチを言う。母集団から直接エレメント(訳注・ユニットのことか?)を抽出する場合には、次のように説明される。すなわち、

		to obtain with the sampling procedure. 2)a known probability of selection $p(s)$ is associated with each possible sample $s$ . 3) the procedure gives every element in the population a nonzero probability of selection. 4)we select one sample by a random mechanism under which each possible $s$ receives exactly the probability $p(s)$ . A sample realized under these four requirements is called probability sample.				1) 抽出手順により得られる可能性のある標本セット $\{S1, S2, \dots, Sm\}$ を定義することが可能である。2) 可能性のある各標本セットの抽出確率 $p(s)$ が分かっている。3) 母集団のすべてのエレメントの抽出確率がゼロでない。4) 各標本セットが選定される確率が厳密に $p(s)$ であるようなランダムなメカニズムによりひとつの標本セットを選定する。この4つの条件にもとづき実現された標本を確率的標本と呼ぶ。
Processing error		Once data have been collected, they pass through a range of processes before the final estimates are produced: coding, keying, editing, weighting, tabulating, etc. Errors introduced at these stages are called processing errors.	Glossary, Previous version	処理誤差		データの回収後、符号格付、キー入力、エディティング、ウェイト調整、表作成など様々な処理を経て、最終推定値が出される。これらの段階で生じた誤差を処理誤差と呼ぶ。
Punctuality		Punctuality refers to the time lag existing between the actual delivery date of data and the target date when it should have been delivered, for instance, with reference to dates announced in some official release calendar, laid down by Regulations or previously agreed among partners.	Eurostat, Oct. 2003, "DEFINITION OF QUALITY IN STATISTICS", (Eurostat/A4/Quality/03/General/Definition)	時間厳守性		時間厳守性は、実際のデータ公表日とそのデータが公表されるべきであった公表予定日のタイムラグを意味する。例えば、公式な公表予定表で告知されている日付、規則により定められている日付、当事者間で事前に合意している日付などが公表予定日である。
Quality control survey		A replicated survey carried out on a small scale by very experienced staff in order to obtain some "zero-default" results with which the actual results of the survey can be compared.	Glossary, Previous version	品質管理調査		“ゼロ・デフォルト”(訳注: zero-defect(欠陥がない)の間違えがもしれない。)の結果を出すために経験豊富なスタッフが小規模に実施する模写調査で、その調査結果と実際の調査結果を比較することができる。

Quality index		A one-dimension synthetic information on quality, possibly calculated as a weighted mean of all available quality indicators.	Glossary, Previous version	品質指数		品質に関する1次元の合成情報で、入手可能なすべての品質指標の加重平均として計算されることが多い。
Reference period	Reference time	The period of time for which data are collected	F.H.C. Marriott, 1990, A dictionary of statistical terms, Longman Scientific & Technical, Fifth edition, p.223	対象期間	対象時点	データが収集される時点(訳注:調査票を回収する期間という意味ではなく、例えば、10月1日午前0:00時点の年齢のように、調査の参照時点)
Refusal rate		In the sampling of human populations, the proportion of individuals who, through successfully contacted, refuse to give the information sought. The proportion is usually and preferably calculated by dividing the number of refusals by the total number of the sample which it was originally desired to achieve.	F.H.C. Marriott, 1990, A dictionary of statistical terms, Longman Scientific & Technical, Fifth edition, p.223	拒否率		人間を母集団とする標本抽出において、接触が取れたものの、求めている情報の提供を拒否した者の割合。この割合は、通常、拒否した者の数を最初に回答が得られると期待された総標本数で割って算出する。
Register		A written and complete record containing regular entries of items and details on particular set of objects. Administrative registers come from administrative sources and become statistical registers after passing through statistical processing in order to make it fit for statistical purposes (production of register based statistics, frame creation, etc.)	COMMON TERMINOLOGY OF METIS, Version of: 29 September, 1999 Page: 7 Editor: Daniel W. Gillman & ad hoc addition	登録		特定の対象に関する項目及び詳細を組織立って記入する書面による完全な記録。行政登録は、行政ソースから得られ、統計の目的に適合させるために、統計的処理を経た後に統計的な登録となる。(登録ベースの統計の作成、フレーム作成、など)
Relative standard error		The relative standard error (RSE) is a measure of an estimate's reliability. The RSE of an estimate is obtained by dividing the standard error of the estimate (SE(r)) by the estimate itself (r). This quantity is expressed as a percent of the estimate and is calculated as follows: $RSE=100 \times (SE(r)/r)$ .	National center for health statistics, "Data definitions", <a href="http://www.cdc.gov/nchs/dataw/h/nchsdefs/list.htm">http://www.cdc.gov/nchs/dataw/h/nchsdefs/list.htm</a>	相対的標準誤差		相対的標準誤差(RSE)は、推定値の信頼性の尺度である。推定値のRSEは、推定値の標準誤差(SE(r))を推定値自身(r)で割って求める。この数値は、推定値に対する百分率で表され、次のように計算される。 $RSE = 100 \times (SE(r)/r)$

Relevance		Relevance is the degree to which statistics meet current and potential users' needs. It refers to whether all statistics that are needed are produced and the extent to which concepts used (definitions, classifications etc.) reflect user needs.	Statistics Canada, October 1998, "statistics Canada Quality Guidelines", 3 <sup>rd</sup> edition, Catalogue no. 12-539-XIE, p. 101	ニーズ適合性		ニーズ適合性は、統計が、現在及び将来のユーザーのニーズを満たしている程度を言う。これは、必要とされる統計がすべて作成されているかどうか、使用されている概念(定義、分類など)がユーザーのニーズを反映している程度に関連している。
Reweighting		Reweighting consists of raising the original weights for the respondent values when estimates are computed. Reweighting concerns mainly unit non-response. It may also be used to increase precision through the use of auxiliary information. Standard methods include post-stratification, calibration and response propensity modelling.	Glossary, Previous version	ウェイト調整		ウェイト調整は、推定値を計算する時に回答者の数値に割り当てたオリジナル・ウェイトを引き上げることを言う。ウェイト調整は、主としてユニット非回答に関連している。また、補助的情報の利用により精度を高めるために利用されることもある。標準的方法には、事後層化、キャリブレーション、回答傾向モデリングなどがある。
Sampling error		The part of the difference between a population value and an estimate thereof, derived from a random sample, which is due to the fact that only a sample of values is observed; as distinct from errors due to imperfect selection, bias in response or estimation, errors of observation and recording, etc.	F.H.C. Marriott, 1990, A dictionary of statistical terms, Longman Scientific & Technical, Fifth edition, p.223	標本誤差		母集団の値とその推定値の間の差のうち、無作為サンプルを用いることによって生じる部分。ひとつのサンプル値のみしか観測されないという事実に起因する。不完全な選定、回答或いは推定の偏りや、観測、記録時の誤差などと区別される。
Standard error		The positive square root of the variance of the sampling distribution of a statistic.	F.H.C. Marriott, 1990, A dictionary of statistical terms, Longman Scientific & Technical, Fifth edition, p.223	標準誤差		統計の標本分布の分散の正の平方根
Statistical characteristics		A numerical value (like turnover, average income) defined by a statistical measure that is used to summarise the values for a specific quantitative variable (like turnover, disposable income) for all statistical units in a specific group.	Glossary, Previous version	統計的特性		特定グループ内のすべての統計ユニットの特定の数量(取引高、可処分所得など)の値を要約するために使われる統計的尺度により定義される数値(取引高、平均所得など)



Statistical measure		A summary (mean, mode, total, index, etc.) of the individual quantitative variable values for the statistical units in a specific group (study domains).	Glossary, Previous version	統計的尺度		特定グループ(研究ドメイン)内の統計ユニット個々の数量値の要約(平均値, 最頻値, 総計, 指数など)
Statistical unit		An object of statistical survey and the bearer of statistical characteristics. The statistical unit is the basic unit of statistical observation within a statistical survey	COMMON TERMINOLOGY OF METIS, Version of: 29 September, 1999 Page: 7 Editor: Daniel W. Gillman	統計ユニット		統計調査の対象物及び統計的特性を有するもの。統計ユニットは, 統計調査における統計的観測の基本単位である。(訳注・定義欄で用いられている“エレメント(element)”は統計ユニットと同一とみられる。)
Study domains		Statistics are presented for different sub-groups of the population, so called study domains. These study domains are usually defined according to some classification (e.g. territorial units, economic activity etc.)	Glossary, Previous version	研究ドメイン		統計は, 研究ドメインと呼ばれる母集団を構成するサブグループ別に表章される。これらの研究ドメインは, 通常は, ある分類(例えば, 地域, 経済活動など)に従って定義されている。
Survey design	Sampling plan, Sample design	Survey design defines the fixed properties of the data collection over all possible implementations within a fixed measurement environment. The usage is not uniform as regards the precise meaning of this and similar terms like "sample plan," "survey design," "sample plan" or "sampling design." These cover one or more parts constituting the entire planning of a (sample) survey inclusive of processing, etc. The term "sampling plan" may be restricted to mean all steps taken in selecting the sample; the term "sample design" may cover in addition the method of estimation; and "survey design" may cover also other aspects of the survey, e.g. choice and training of interviewers, tabulation plans, etc. "Sample design" is sometimes used in a clearly defined sense, with reference to a given frame, as the set of rules or specifications for the drawing of a sample in an unequivocal manner.	1)Model assisted survey sampling", Carl-Erik Sarndal, Bengt Swensson, Jan Wretman., & 2)Federal Committee on Statistical Methodology, December 1978, - Glossary of Nonsampling Error Terms: An Illustration of a Semantic Problem in Statistic, Statistical Policy Working Paper 4.	調査設計	標本計画, 標本設計	調査設計とは, 一定の計測環境の下で可能な全てのデータ収集実行選択肢から一つを定めることである。調査設計や類似語の“標本計画”“標本設計”の厳密な定義に関しては, 使い方が統一されていない。これらの言葉は, データ処理等を含めた(標本)調査全体の計画を構成するひとつ又は複数の部分をカバーしている。“標本計画”という言葉は, 標本選定に関する段階に限定されるかもしれない。“標本設計”は, 推定方法もカバーしているかもしれない, “調査設計”は, さらに調査のその他の側面, 例えば, 調査員の選定や研修, 製表計画などもカバーしているかもしれない。“標本設計”は, 時に明確な定義の下で使われることがある。それは, 所与のフレームに

						において、明確な方法で標本を抽出する規則あるいは仕様を指す場合である。
Target population		The <i>target population</i> is the population we wish to study, that is, the set of elements about which estimates are required.	Carl-Erik Sarndal, Bengt Swensson, Jan Wretman, 1992, "Model assisted survey sampling", Springer - Verlag New York, p.694	目標母集団		目標母集団は、研究したい母集団、すなわち、それに関する推定値が必要とされているエレメント(訳注:統計ユニット?)の集合を言う。
Timeliness		Timeliness of information reflects the length of time between its availability and the event or phenomenon it describes	Eurostat, Oct. 2003, "DEFINITION OF QUALITY IN STATISTICS", (Eurostat/A4/Quality/03/General/Definition)	適時性		情報の適時性は、その情報が記述している事象又は現象と、その情報が利用可能となるまでの期間の長さを反映している。
True value		The actual population value that would be obtained with perfect measuring instruments and without committing any error of any type, both in collecting the primary data and in carrying out mathematical operations.	Glossary, Previous version	真値		母集団の実際の値。基本データの収集段階においても、数学的処理の実行段階においても、完全な測定器具を用い、いかなる種類の誤差も発生させなかった場合に得られるであろう値。
Under coverage		Undercoverage results from the omission from the frame of units belonging to the target population.	Ad-hoc	アンダー・カバレッジ		アンダー・カバレッジは、目標母集団に属するユニット・フレームからの遺漏により生じる
Unit non response		Unit non response is a complete failure to obtain data from a sample unit.	Statistical Policy working paper 31, "Measuring and reporting sources of error in survey", FCSM, Subcommittee on Measuring and Reporting the Quality of Survey Data	ユニット非回答		ユニット非回答は、標本ユニットからデータが全く得られなかったことを言う。
Unit response rate		The ratio, expressed in percentage of the number of interviews to the number of eligible units in the sample. The weighted response rate calculates the ratio using the inverse probability of inclusion in the sample as a weight for each unit. In some occasions a value that reflects the importance of the unit is also	Ad-hoc	ユニット回答率		標本の中の適格ユニット数に対する調査ユニット数の比率。加重回答率は、各ユニットについて標本に選定される確率の逆数で加重して比率を計算したものである。場合によっては、ユニットの重要性を反映した値をウ

		used as a weighting factor (like size of workforce for establishments).				エイト付けに用いることもある。(事業所の従業員規模など)
User satisfaction survey		A statistical survey aiming to assess the satisfaction of users of statistics.	Glossary, Previous version	ユーザー満足度調査		統計に対するユーザーの満足度を評価するための統計調査
Variance		The variance is the mean square deviation of the variable around the average value. It reflects the dispersion of the empirical values around its mean.	Glossary, Previous version	分散		分散は、変数の平均値周辺の平均平方偏差である。分散は、平均周辺の経験値の散乱度合いを反映している。
Variance estimation		The task of estimating the value of the variance of an estimate. Methods employed are commonly classified in: - analytic methods: use and compute the proper formulae of the variance - approximate methods: methods which use approximations for complex and multi-stage sample designs. These methods can be classified in three categories: - simplifying assumptions: setting assumptions which allow use of straightforward formulae (as if an analytic method could be applied). - Taylor's linearisation techniques: estimation of non-linear statistics is simplified in using a Taylor's development of the concerned statistics. - Replication or re-sampling methods: a large number of subsamples are derived from the initial sample and variance is estimated from the variability of the results in this set of subsamples. Several methods are based on re-sampling, such as the jackknife, the bootstrap and the balanced repeated replication methods (BRR).	Glossary, Previous version	分散推定		推定値の分散を推定する作業。使用される方法は、一般的に次のように分類される。すなわち、解析的方法：適切な分散算出公式の使用と計算，近似法：複雑で多段階の標本設計のために近似を利用する方法。近似法は、3つのカテゴリーに分類することができる。すなわち、前提の単純化：単純な公式が利用できるように(解析的方法が適用可能となる)前提を設定。テイラーの線形化技法：非線形統計の推定を、テイラー展開を利用して単純化。反復又はリサンプリング方法：最初の標本から多数の副標本を発生させ、この副標本から得た結果の変動から分散を推定。ジャックナイフ、ブートストラップ、均衡反復複製法(BRR)などいくつかの方法がリサンプリングに基づく方法である。
Working group "Assessment of Quality in Statistics"		Eurostat modified in spring 1998 the terms of reference of the former Working Group on the quality evaluation of structural business statistics, thereby creating the « Working Group on Assessment of Quality in Statistics ». The mandate of the WG is to:	Glossary, Previous version	“統計における品質評価”ワーキンググループ		Eurostat は、1998年春に、構造企業統計の品質評価に関するワーキンググループを改編し、<統計における品質評価に関するワーキンググループ>を設置した。このワーキンググループ

		1. harmonise the definition of quality in statistics, 2. standardise Quality Reports, 3. address methodological problems for measuring the quality of statistics, and 4. co-ordinate all activities related to quality within Eurostat and the NSIs.		ループ		の任務は次のとおりである。1.統計における品質の定義の調和, 2.品質報告書の標準化, 3.統計の品質の測定に関する方法論的問題を扱う, 4.Eurostat 及び NSI における品質関連のすべての活動の調整
--	--	--	--	-----	--	---

## 使用した原典

1. F.H.C. Marriott, 1990, A dictionary of statistical terms, Longman Scientific & Technical, Fifth edition, pp.223
2. B.S.EVERITT, 2002, The Cambridge dictionary of Statistics, Cambridge University Press. Second edition, pp. 410
3. Graham Upton, Ian Cook, 2002, A dictionary of statistics, Oxford University Press, First edition, pp.420
4. Johanna Laiho, Leena Hietaniemi, 2002, Quality Guidelines for Official Statistics, Statistics Finland, pp.134
5. FCSM, July 2001, Measuring and Reporting Sources of Error in Surveys, Subcommittee on Measuring and Reporting the Quality of Survey Data, Statistical working paper 31
6. Carl-Erik Sarndal, Bengt Swensson, Jan Wretman, 1992, "Model assisted survey sampling", Springer - Verlag New York, pp.694
7. Paul P. Biemer, Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz, Seymour Sudman, 1991, "Measurement errors in survey", John Wiley & Sons, pp. 760
8. National center for health statistics, "Data definitions", available at <http://www.cdc.gov/nchs/datawh/nchsdefs/list.htm>
9. Federal Committee on Statistical Methodology, December 1978, - Glossary of Nonsampling Error Terms: An Illustration of a Semantic Problem in Statistic, Statistical Policy Working Paper 4.
10. Statistics Canada, October 1998, "Statistics Canada Quality Guidelines", 3rd edition, Catalogue no. 12-539-XIE, pp . 101
11. Daniel W. Gillman, 1999, COMMON TERMINOLOGY OF METIS, STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE, Version of: 29 September, 1999, Paper No. 10, pp. 45
12. Eurostat, Oct. 2003, "DEFINITION OF QUALITY IN STATISTICS", (Eurostat/A4/Quality/03/General/Definition)sixth meeting Luxembourg – 2-3 October 2003.
13. Handbook of Household Surveys, Revised Edition, Studies in Methods, Series F, No. 31, United Nations, New York, 1984, Para. 8.3
14. EPA/Office of Research and Development: Guidance for Quality Assurance Project Plans (QA/G-5), EPA/ 600/R-98/018
15. Wikipedia, the free encyclopaedia available at <http://www.wikipedia.org>
16. UN, Statistics Division, Glossary, available at <http://unstats.un.org/unsd/sna1993/glossary.asp>

17. William E. Winkler, 1997, "PROBLEMS WITH INLIERS" Work Session on Statistical Data Editing, US Bureau of Census, Working Paper 22,
18. Madow, W., Nisselson, H., and Olkin, I., 1983, Incomplete Data in Sample Surveys, New York: Academic Press
19. UNECE, Glossary of Terms on Statistical Data Editing, CES Methodological material, Geneva, 2000, also available at <http://amrads.jrc.cec.eu.int/k-base/glossary/g-source.htm>
20. ESOMAR, Glossary of Marketing Research Terms, available at <http://www.esomar.nl>
21. United Nations Glossary of Classification Terms. Prepared by the Expert Group on International Economic and Social Classifications. Available at: [http://www.un.org/Depts/unsd/class/glossary\\_short.htm](http://www.un.org/Depts/unsd/class/glossary_short.htm)