

# Accuracy and precision

In a set of measurements, **accuracy** is closeness of the measurements to a specific value, while **precision** is the closeness of the measurements to each other.

*Accuracy* has two definitions:

1. More commonly, it is a description of *systematic errors*, a measure of statistical bias; low accuracy causes a difference between a result and a "true" value. ISO calls this *trueness*.
2. Alternatively, ISO defines<sup>[1]</sup> accuracy as describing a combination of both types of observational error above (random and systematic), so high accuracy requires both high precision and high trueness.

*Precision* is a description of *random errors*, a measure of statistical variability.

In simpler terms, given a set of data points from repeated measurements of the same quantity, the set can be said to be *accurate* if their average is close to the *true value* of the quantity being measured, while the set can be said to be *precise* if the values are close to each other. In the first, more common definition of "accuracy" above, the two concepts are independent of each other, so a particular set of data can be said to be either accurate, or precise, or both, or neither.

## Contents

### Common technical definition

Quantification

### ISO definition (ISO 5725)

### In binary classification

### In psychometrics and psychophysics

### In logic simulation

### In information systems

### See also

### References

### External links

## Common technical definition

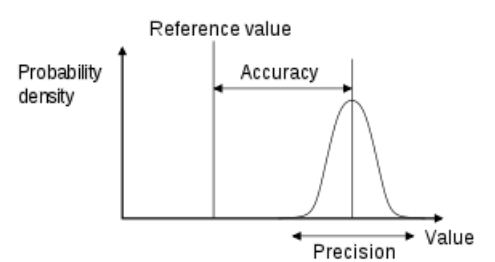
In the fields of science and engineering, the accuracy of a measurement system is the degree of closeness of measurements of a quantity to that quantity's true value.<sup>[2]</sup> The precision of a measurement system, related to reproducibility and repeatability, is the degree to which repeated measurements under unchanged conditions show the same results.<sup>[2][3]</sup> Although the two words precision and accuracy can be synonymous in colloquial use, they are deliberately contrasted in the context of the scientific method.

The field of statistics, where the interpretation of measurements plays a central role, prefers to use the terms *bias* and *variability* instead of accuracy and precision: bias is the amount of inaccuracy and variability is the amount of imprecision.

A measurement system can be accurate but not precise, precise but not accurate, neither, or both. For example, if an experiment contains a systematic error, then increasing the sample size generally increases precision but does not improve accuracy. The result would be a consistent yet inaccurate string of results from the flawed experiment. Eliminating the systematic error improves accuracy but does not change precision.

A measurement system is considered *valid* if it is both *accurate* and *precise*. Related terms include *bias* (non-random or directed effects caused by a factor or factors unrelated to the independent variable) and *error* (random variability).

The terminology is also applied to indirect measurements—that is, values obtained by a computational procedure from observed data.



*Accuracy* is the proximity of measurement results to the true value; *precision* is the degree to which repeated (or reproducible) measurements under unchanged conditions show the same results.

In addition to accuracy and precision, measurements may also have a measurement resolution, which is the smallest change in the underlying physical quantity that produces a response in the measurement.

In numerical analysis, accuracy is also the nearness of a calculation to the true value; while precision is the resolution of the representation, typically defined by the number of decimal or binary digits.

In military terms, accuracy refers primarily to the accuracy of fire (*justesse de tir*), the precision of fire expressed by the closeness of a grouping of shots at and around the centre of the target.<sup>[4]</sup>

## Quantification

In industrial instrumentation, accuracy is the measurement tolerance, or transmission of the instrument and defines the limits of the errors made when the instrument is used in normal operating conditions.<sup>[5]</sup>

Ideally a measurement device is both accurate and precise, with measurements all close to and tightly clustered around the true value. The accuracy and precision of a measurement process is usually established by repeatedly measuring some traceable reference standard. Such standards are defined in the International System of Units (abbreviated SI from French: *Système international d'unités*) and maintained by national standards organizations such as the National Institute of Standards and Technology in the United States.

This also applies when measurements are repeated and averaged. In that case, the term standard error is properly applied: the precision of the average is equal to the known standard deviation of the process divided by the square root of the number of measurements averaged. Further, the central limit theorem shows that the probability distribution of the averaged measurements will be closer to a normal distribution than that of individual measurements.

With regard to accuracy we can distinguish:

- the difference between the mean of the measurements and the reference value, the bias. Establishing and correcting for bias is necessary for calibration.
- the combined effect of that and precision.

A common convention in science and engineering is to express accuracy and/or precision implicitly by means of significant figures. Where not explicitly stated, the margin of error is understood to be one-half the value of the last significant place. For instance, a recording of 843.6 m, or 843.0 m, or 800.0 m would imply a margin of 0.05 m (the last significant place is the tenths place), while a recording of 843 m would imply a margin of error of 0.5 m (the last significant digits are the units).

A reading of 8,000 m, with trailing zeros and no decimal point, is ambiguous; the trailing zeros may or may not be intended as significant figures. To avoid this ambiguity, the number could be represented in scientific notation:  $8.0 \times 10^3$  m indicates that the first zero is significant (hence a margin of 50 m) while  $8.000 \times 10^3$  m indicates that all three zeros are significant, giving a margin of 0.5 m. Similarly, one can use a multiple of the basic measurement unit: 8.0 km is equivalent to  $8.0 \times 10^3$  m. It indicates a margin of 0.05 km (50 m). However, reliance on this convention can lead to false precision errors when accepting data from sources that do not obey it. For example, a source reporting a number like 153,753 with precision +/- 5,000 looks like it has precision +/- 0.5. Under the convention it would have been rounded to 154,000.

Alternatively, in a scientific context, if it is desired to indicate the margin of error with more precision, one can use a notation such as  $7.54398(23) \times 10^{-10}$  m, meaning a range of between  $7.54375$  and  $7.54421 \times 10^{-10}$  m.

Precision includes:

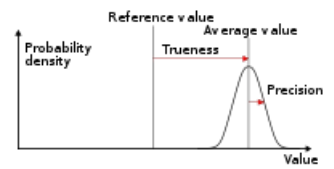
- *repeatability* — the variation arising when all efforts are made to keep conditions constant by using the same instrument and operator, and repeating during a short time period; and
- *reproducibility* — the variation arising using the same measurement process among different instruments and operators, and over longer time periods.

In engineering, precision is often taken as three times Standard Deviation of measurements taken, representing the range that 99.73% of measurements can occur within.<sup>[6]</sup> For example, an ergonomist measuring the human body can be confident that 99.73% of their extracted measurements fall within  $\pm 0.7$ cm - if using the GRYPHON processing system - or  $\pm 13$ cm - if using unprocessed data.<sup>[7]</sup>

## ISO definition (ISO 5725)

A shift in the meaning of these terms appeared with the publication of the ISO 5725 series of standards in 1994, which is also reflected in the 2008 issue of the "BIPM International Vocabulary of Metrology" (VIM), items 2.13 and 2.14.<sup>[2]</sup>

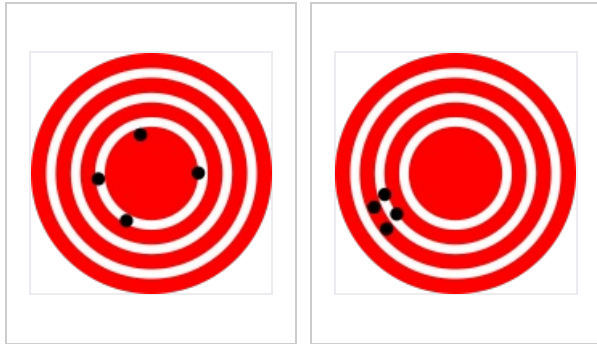
According to ISO 5725-1,<sup>[1]</sup> the general term "accuracy" is used to describe the closeness of a measurement to the true value. When the term is applied to sets of measurements of the same measurand, it involves a component of random error and a component of systematic error. In this case trueness is the closeness of the mean of a set of measurement results to the actual (true) value and precision is the closeness of agreement among a set of results.



ISO 5725-1 and VIM also avoid the use of the term "bias", previously specified in BS 5497-1,<sup>[8]</sup> because it has different connotations outside the fields of science and engineering, as in medicine and law.

According to ISO 5725-1, Accuracy consists of trueness (proximity of measurement results to the true value) and precision (repeatability or reproducibility of the measurement)

### Accuracy of a target grouping according to BIPM and ISO 5725



Low accuracy due to poor precision

Low accuracy due to poor trueness

## In binary classification

*Accuracy* is also used as a statistical measure of how well a binary classification test correctly identifies or excludes a condition. That is, the accuracy is the proportion of correct predictions (both true positives and true negatives) among the total number of cases examined.<sup>[9]</sup> As such, it compares estimates of pre- and post-test probability. To make the context clear by the semantics, it is often referred to as the "Rand accuracy" or "Rand index",<sup>[10][11][12]</sup> It is a parameter of the test. The formula for quantifying binary accuracy is:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

where TP = True positive; FP = False positive; TN = True negative; FN = False negative

Note that, in this context, the concepts of trueness and precision as defined by ISO 5725-1 are not applicable. One reason is that there is not a single "true value" of a quantity, but rather two possible true values for every case, while accuracy is an average across all cases and therefore takes into account both values. However, the term *precision* is used in this context to mean a different metric originating from the field of information retrieval (see below).

## In psychometrics and psychophysics

In psychometrics and psychophysics, the term *accuracy* is interchangeably used with validity and *constant error*. *Precision* is a synonym for reliability and *variable error*. The validity of a measurement instrument or psychological test is established through experiment or correlation with behavior. Reliability is established with a variety of statistical techniques, classically through an internal consistency test like Cronbach's alpha to ensure sets of related questions have related responses, and then comparison of those related question between reference and target population.

## In logic simulation

In logic simulation, a common mistake in evaluation of accurate models is to compare a logic simulation model to a transistor circuit simulation model. This is a comparison of differences in precision, not accuracy. Precision is measured with respect to detail and accuracy is measured with respect to reality.<sup>[13][14]</sup>

## In information systems

Information retrieval systems, such as databases and web search engines, are evaluated by many different metrics, some of which are derived from the confusion matrix, which divides results into true positives (documents correctly retrieved), true negatives (documents correctly not retrieved), false positives (documents incorrectly retrieved), and false negatives (documents incorrectly not retrieved). Commonly used metrics include the notions of precision and recall. In this context, precision is defined as the fraction of retrieved documents which are relevant to the query (true positives divided by true+false positives), using a set of ground truth relevant results selected by humans. Recall is defined as the fraction of relevant documents retrieved compared to the total number of relevant documents (true positives divided by true positives+false negatives). Less commonly, the metric of accuracy is used, is defined as the total number of correct classifications (true positives plus true negatives) divided by the total number of documents.

None of these metrics take into account the ranking of results. Ranking is very important for web search engines because readers seldom go past the first page of results, and there are too many documents on the web to manually classify all of them as to whether they should be included or excluded from a given search. Adding a cutoff at a particular number of results takes ranking into account to some degree. The measure precision at k, for example, is a measure of precision looking only at the top ten ( $k=10$ ) search results. More sophisticated metrics, such as discounted cumulative gain, take into account each individual ranking, and are more commonly used where this is important.

## See also

- Bias-variance tradeoff in statistics and machine learning
- Accepted and experimental value
- Data quality
- Engineering tolerance
- Exactness (disambiguation)
- Experimental uncertainty analysis
- F-score
- Hypothesis tests for accuracy
- Information quality
- Measurement uncertainty
- Precision (statistics)
- Probability
- Random and systematic errors
- Sensitivity and specificity
- Significant figures
- Statistical significance

## References

- BS ISO 5725-1: "Accuracy (trueness and precision) of measurement methods and results - Part 1: General principles and definitions.", p.1 (1994)
- JCGM 200:2008 International vocabulary of metrology* ([http://www.bipm.org/utis/common/documents/jcgm/JCGM\\_200\\_2008.pdf](http://www.bipm.org/utis/common/documents/jcgm/JCGM_200_2008.pdf)) — *Basic and general concepts and associated terms (VIM)*
- Taylor, John Robert (1999). *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements* (<https://books.google.com/books?id=giFQcZub80oC&pg=PA128>). University Science Books. pp. 128–129. ISBN 0-935702-75-X.
- North Atlantic Treaty Organization, Nato Standardization Agency AAP-6 - Glossary of terms and definitions, p 43.
- Creus, Antonio. *Instrumentación Industrial*
- author., Black, J. Temple (21 July 2020). *DeGarmo's materials and processes in manufacturing* (<http://worldcat.org/oclc/1246529321>). ISBN 978-1-119-72329-5. OCLC 1246529321 (<https://www.worldcat.org/oclc/1246529321>).
- Parker, Christopher J.; Gill, Simeon; Harwood, Adrian; Hayes, Steven G.; Ahmed, Maryam (2021-05-19). "A Method for Increasing 3D Body Scanning's Precision: Gryphon and Consecutive Scanning" (<https://www.tandfonline.com/doi/full/10.1080/00140139.2021.1931473>). *Ergonomics*: 1–47. doi:10.1080/00140139.2021.1931473 (<https://doi.org/10.1080/00140139.2021.1931473>). ISSN 0014-0139 (<https://www.worldcat.org/issn/0014-0139>). PMID 34006206 (<https://pubmed.ncbi.nlm.nih.gov/34006206>).
- BS 5497-1: "Precision of test methods. Guide for the determination of repeatability and reproducibility for a standard test method." (1979)
- Metz, CE (October 1978). "Basic principles of ROC analysis" ([http://www.umich.edu/~ners580/ners-bioe\\_481/lecture\\_s/pdfs/1978-10-semNucMed\\_Metz-basicROC.pdf](http://www.umich.edu/~ners580/ners-bioe_481/lecture_s/pdfs/1978-10-semNucMed_Metz-basicROC.pdf)) (PDF). *Semin Nucl Med*. **8** (4): 283–98. doi:10.1016/s0001-2998(78)80014-2 (<https://doi.org/10.1016%2Fs0001-2998%2878%2980014-2>). PMID 112681 (<https://pubmed.ncbi.nlm.nih.gov/112681>).
- "Archived copy" ([https://web.archive.org/web/20150311073014/http://alta.asn.au/events/altss\\_w2003\\_proc/altss/courses/powers/ALTSS2003-Val+Eval-L3.pdf](https://web.archive.org/web/20150311073014/http://alta.asn.au/events/altss_w2003_proc/altss/courses/powers/ALTSS2003-Val+Eval-L3.pdf)) (PDF). Archived from the original ([http://www.alta.asn.au/events/altss\\_w2003\\_proc/altss/courses/powers/ALTSS2003-Val+Eval-L3.pdf](http://www.alta.asn.au/events/altss_w2003_proc/altss/courses/powers/ALTSS2003-Val+Eval-L3.pdf)) (PDF) on 2015-03-11. Retrieved 2015-08-09.
- Powers, David M. W. (2015). "What the F-measure doesn't measure". *arXiv:1503.06410* (<https://arxiv.org/abs/1503.06410>) [cs.LR (<https://arxiv.org/archive/cs.LR>)].
- David M W Powers "The Problem with Kappa" (<http://www.anthology.aclweb.org/F/E12/F12-1035.pdf>) (PDF)

2. David H. W. Powers. The Problem with Rappa (<http://www.anthology.aclweb.org/L/E12/E12-1999.pdf>) (1999). *Anthology.aclweb.org*. Retrieved 11 December 2017.
3. Acken, John M. (1997). "none". *Encyclopedia of Computer Science and Technology*. **36**: 281–306.
4. Glasser, Mark; Mathews, Rob; Acken, John M. (June 1990). "1990 Workshop on Logic-Level Modelling for ASICS". *SIGDA Newsletter*. **20** (1).





## External links

---

- BIPM - Guides in metrology (<http://www.bipm.org/en/publications/guides/>), *Guide to the Expression of Uncertainty in Measurement (GUM)* and *International Vocabulary of Metrology (VIM)*
  - "Beyond NIST Traceability: What really creates accuracy" (<http://img.en25.com/Web/Vaisala/NIST-article.pdf>), *Controlled Environments* magazine
  - Precision and Accuracy with Three Psychophysical Methods (<http://www.yorku.ca/psycho>)
  - Appendix D.1: Terminology (<http://physics.nist.gov/Pubs/guidelines/appd.1.html>), *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*
  - Accuracy and Precision (<http://digipac.ca/chemical/sigfigs/contents.htm>)
  - Accuracy vs Precision ([https://www.youtube.com/watch?v=\\_LL0uiOgh1E](https://www.youtube.com/watch?v=_LL0uiOgh1E)) — a brief video by Matt Parker
  - What's the difference between accuracy and precision? (<https://www.youtube.com/watch?v=hRAFPdDppzs>) by Matt Anticole at TED-Ed
  - Precision and Accuracy exam study guide (<https://www.examprof.com/2020/07/precision-and-accuracy.html>)
- 

Retrieved from "[https://en.wikipedia.org/w/index.php?title=Accuracy\\_and\\_precision&oldid=1030951007](https://en.wikipedia.org/w/index.php?title=Accuracy_and_precision&oldid=1030951007)"

---

**This page was last edited on 28 June 2021, at 23:23 (UTC).**

Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.