U-Pb Geochronology Practical: Background

Basic Concepts:

accuracy: measure of the difference between an experimental measurement and the "true" value

precision: measure of the reproducibility of the experimental result

random error: unpredictable fluctuations in observations that yield results that differ from experiment to experiment

systematic error: errors that cause the measurement to differ from the "true" values with reproducible discrepancy

Sample versus Population

A **sample** is a group of data which is a subset of an infinite **population**.

population distribution: the "true" distribution of values of a given quantity

sample distribution: the measured distribution of those values; in the limit of infinite measurements, the sample distribution approaches the parent distribution

How do we describe the sample versus population distributions? Often with concepts like the **mean**, the **variance**, and the **standard deviation**.

Parametric statistics

The **mode** is the most probable value of **x**

The **median** ($x_{1/2}$) is the value for which half the observations are greater and half less

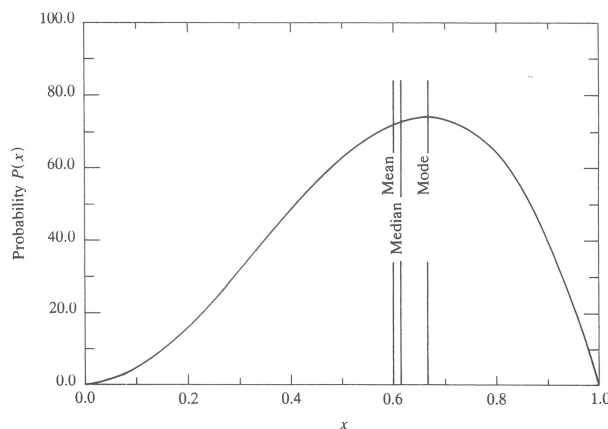The **mean** (**x**) is the average value of a quantity **x**

FIGURE 1.3
Asymmetric distribution illustrating the positions of the mean, median, and most probable value of the variable.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \text{(sample)} \qquad \mu = \lim_{N \to \infty} \frac{1}{N}\sum_{i} x_i \quad \text{(population)}$$

The **deviation** (d_i) of a measurement x_i from the mean μ of the parent distribution is the difference between x_i and μ

$$d_i = x_i - \bar{x} \text{ (sample)} \qquad d_i = x_i - \mu \text{ (population)}$$

The **average deviation**, α, is the (in the limit of infinite measurements) the average of the absolute values of the deviations from the mean μ … but the absolute value is computationally inconvenient…

$$\alpha = \frac{1}{n} \sum |x_i - \bar{x}| \text{ (sample)} \qquad a = \lim_{N \to \infty} \frac{1}{N} \sum_i |x_i - \mu| \text{ (population)}$$

The **variance**, σ^2, is (in the limit of infinite measurements) the average of the squares of the deviations from the mean μ

$$\sigma^2 = \frac{1}{n-1} \sum_i (x_i - \bar{x})^2 \text{ (sample)} \qquad s^2 = \lim_{N \to \infty} \frac{1}{N} \sum_i (x_i - \mu)^2 \text{ (population)}$$

The **standard deviation**, σ, is (in the limit of infinite measurements) the square root of the variance; the standard deviation estimates the likelihood of a measurement falling within a certain interval about the mean:

$$\sigma = \sqrt{\sigma^2} \text{ (sample)} \qquad s = \sqrt{s^2} \text{ (population)}$$

The **standard error** of the mean is a measure of how well the sample mean, **x**, estimates the population mean, μ … it **does not** represent the same thing as the standard deviation! It is calculated from the standard deviation divided by the square root of the number of measurements:

$$\sigma_{mean} = \frac{\sigma}{\sqrt{n}}$$

Note that as n approaches infinity, the standard error approaches zero, e.g. the sample mean must approach the population mean.

Distributions

Recall that a **histogram** is a graph of the frequency of binned measurements; the apparent shapes of distributions illustrated in histograms are highly sensitive to the choice of bin size…

A **probability density function** appears as the smooth curve over the columns of a histogram, in the limit of infinite measurements. The shape of the probability density function is independent of the choice of bin size, and thus always "looks" the same unlike associated histograms.

The most commonly observed (or assumed) distribution in science and engineering is the **Gaussian** or **normal distribution**, which is the symmetric bell-shaped curve arising when the number of different possible outcomes is infinite but the probability of each is finite:

$$P_G = \left(\frac{1}{\sigma_i \sqrt{2\pi}}\right) \exp\left[-\frac{1}{2}\left(\frac{x_i - \bar{x}}{\sigma_i}\right)^2\right]$$ (sample)

$$P_G = \lim_{N \to \infty}\left(\frac{1}{s_i \sqrt{2\pi}}\right) \exp\left[-\frac{1}{2}\left(\frac{x_i - \mu}{s_i}\right)^2\right]$$ (population)

This equation yields the unit normal distribution; the integral under the curve is of unit area (=1). It can be subsequently scaled to *n* measurements.
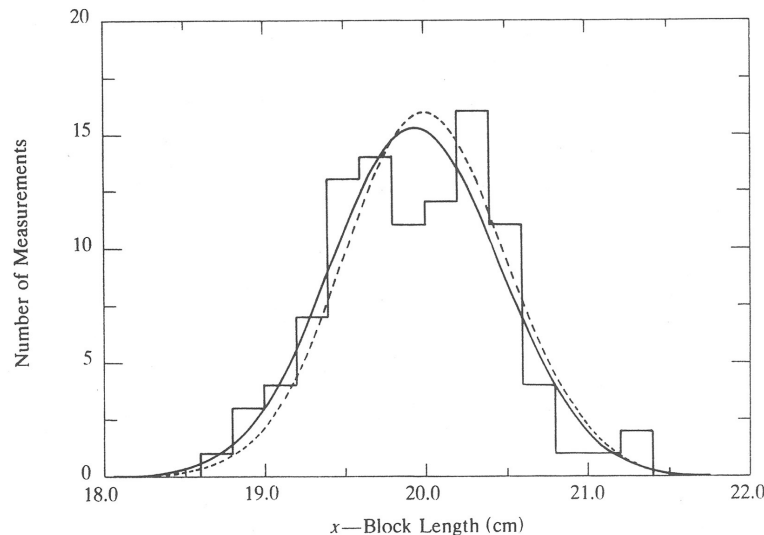

FIGURE 1.2
Histogram of measurements of the length of a block. The solid Gaussian curve was calculated from the mean ($\bar{x} = 19.9$ cm) and standard deviation ($s = 0.52$ cm) estimated from these measurements. The dashed curve represents the parent distribution with mean $\mu = 20.0$ cm and standard deviation $\sigma = 0.50$ cm.
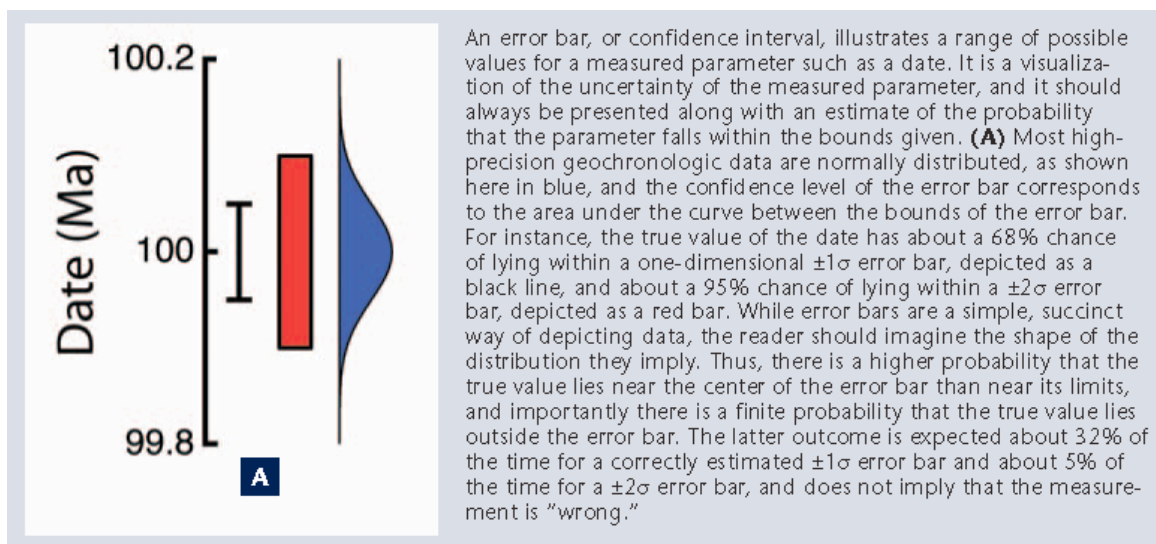
Confidence Intervals

The probability that a measurement will differ from **x** by some amount Δ**x** is called the confidence interval; it is simply the area under the unit probability density function, bound by ±Δ**x**:

$$A_G = \int_{\bar{x}-\Delta x}^{\bar{x}+\Delta x} \left(\frac{1}{\sigma_i \sqrt{2\pi}} \right) \exp\left[-\frac{1}{2}\left(\frac{x_i - \bar{x}}{\sigma_i} \right)^2 \right] dx$$

For a Gaussian probability function,

      If Δ**x** = σ, then A_G = 0.68269 ~ 68% … the "1σ confidence interval"

      If Δ**x** = 2σ, then A_G = 0.95449 ~ 95% … the "2σ confidence interval"

An error bar, or confidence interval, illustrates a range of possible values for a measured parameter such as a date. It is a visualization of the uncertainty of the measured parameter, and it should always be presented along with an estimate of the probability that the parameter falls within the bounds given. **(A)** Most high-precision geochronologic data are normally distributed, as shown here in blue, and the confidence level of the error bar corresponds to the area under the curve between the bounds of the error bar. For instance, the true value of the date has about a 68% chance of lying within a one-dimensional ±1σ error bar, depicted as a black line, and about a 95% chance of lying within a ±2σ error bar, depicted as a red bar. While error bars are a simple, succinct way of depicting data, the reader should imagine the shape of the distribution they imply. Thus, there is a higher probability that the true value lies near the center of the error bar than near its limits, and importantly there is a finite probability that the true value lies outside the error bar. The latter outcome is expected about 32% of the time for a correctly estimated ±1σ error bar and about 5% of the time for a ±2σ error bar, and does not imply that the measurement is "wrong."

Schoene et al. (2013)

Note that in high-precision geochronology the tradition is to tabulate and illustrate analytical data with their 2σ (95%) confidence interval uncertainties.

By contrast, *in situ* geochronological methods (ion probe, LA-ICPMS) traditionally illustrate their data with 1σ (68%) confidence interval uncertainties, e.g.

Constructing Probability Functions (and Histograms)

Two approaches:

1) Assume that the *total* sample distribution can be modeled as a specific probability density function. Then, the area under a normalized probability density function is made equal to the area of the equivalent histogram. For example:

    a) Calculate the unit Gaussian probability function from the appropriate eqn.

    b) Normalize the probability function (Pg) by:

$$y(x_i) = P_G(x_i) * n\Delta x$$

2) Assume that *each measurement* has a normally distributed uncertainty with a known variance. Then you can assign a unit Gaussian curve to each measurement, and then sum all these up along the x-axis. Again the area under the probability density function is equal to the area under the histogram. The benefit of this approach is that it can handle non-uniform variances for each observation.

Weighted Mean and Variance

Following from the idea that the individual observations within a sample may have non-uniform variance, it is common to weight each observation by the inverse of its variance when calculating the mean of the sample distribution:

$$\bar{x} = \frac{\sum (x_i / \sigma_i^2)}{\sum (1/\sigma_i^2)} \qquad\qquad \sigma_{\bar{x}}^2 = \frac{1}{\sum (1/\sigma_i^2)}$$

The standard error of the weighted mean decreases with the number of measurements.

Statistical evaluation of the best-fit solution

How do we evaluate how well a best-fit solution, in this case the weighted mean of a normal distribution, describes our data distribution?

We can define our goodness-of-fit parameter, χ^2, as the sum of the squares of the deviations of our observations from our model mean value, weighted by the variances of those observations:

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{\sigma_i} \right)^2$$

In geochronology, we have a tradition of normalizing the χ^2 statistic by the degrees of freedom of the system to derive the **Mean Squared Weighted Deviation** or **MSWD** (e.g. Wendt and Carl, 1991). This statistic quantifies the extent to which data scatter from the best-fit solution beyond stated uncertainties.

$$MSWD = \frac{\chi^2}{f} = \frac{\chi^2}{n-1}$$

f is the degrees of freedom of the model, e.g. the number of experimental parameters (measurements) – number of model parameters (unknowns)

The expectation (mean) value of the MSWD is 1, in other words:

1) **MSWD = 1**… scatter of data about the model is accurately described by analytical uncertainty

2) **MSWD >> 1**… scatter of data about the model is beyond analytical uncertainties; the model is a poor fit to the data or the experimental error is underestimated

3) **MSWD << 1**… scatter of data about the model is much less than the analytical uncertainty; experimental error is probably overestimated

Wendt and Carl (1991) established the frequency distribution of the MSWD about the mean value of 1, as a simplified function of f:

$$\sigma_{MSWD} = \sqrt{\frac{2}{f}} = \sqrt{\frac{2}{n-2}}$$

References and further reading

Bevington, P.R., and Robinson, D.K., 2003, Data Reduction and Error Analysis for the Physical Sciences, 3rd Ed.: McGraw-Hill, New York, 320 p.

Ludwig, K., 2000, Decay constant errors in U–Pb concordia-intercept ages: Chemical Geology, v. 166, no. 3-4, p. 315–318.

Ludwig, K., 1998, On the treatment of concordant uranium-lead ages: Geochimica et Cosmochimica Acta, v. 62, no. 4, p. 665–676.

Ludwig, K.R., 2003, Isoplot 3.00: a geochronological toolkit for Microsoft Excel: Berkeley Geochronology Center Special Publication, v. 4, p. 1–71.

McLean, N.M., Bowring, J.F., and Bowring, S.A., 2011, An algorithm for U-Pb isotope dilution data reduction and uncertainty propagation: Geochemistry Geophysics Geosystems, v. 12, no. null, p. Q0AA18, doi: 10.1029/2010GC003478.

Schmitz, M.D., and Schoene, B., 2007, Derivation of isotope ratios, errors, and error correlations for U-Pb geochronology using 205Pb- 235U-(233U)-spiked isotope dilution thermal ionization mass spectrometric data: Geochemistry Geophysics Geosystems, v. 8, no. 8, p. 1–20, doi: 10.1029/2006GC001492.

Schoene, B., Condon, D.J., Morgan, L., and McLean, N., 2013, Precision and Accuracy in Geochronology: Elements, v. 9, no. 1, p. 19–24, doi: 10.2113/gselements.9.1.19.

Wendt, I., and Carl, C., 1991, The statistical distribution of the mean squared weighted deviation: Chemical geology. Isotope geoscience section, v. 86, no. 4, p. 275–285.