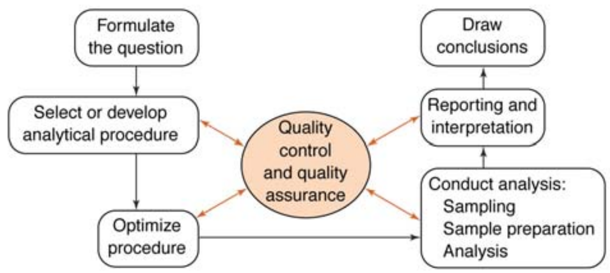
# Statistics and Chemical Measurements: Quantifying Uncertainty

*The bottom line:* ***Do we trust our results?***
*Should we (or anyone else)? Why?*




What is Quality Assurance?

What is Quality Control?



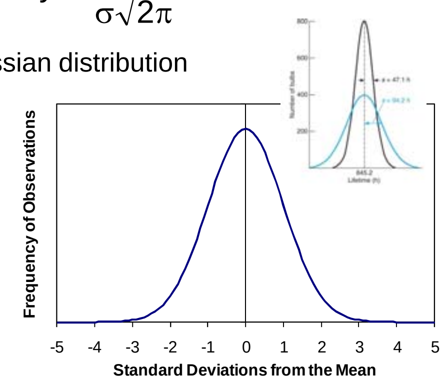
# Normal or Gaussian Distribution – The "Bell Curve"

**IF only *random errors* are present, data will follow a Gaussian Distribution**

This distribution is described by: $$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$$

Two important parameters for a Gaussian distribution

$\mu$: population *mean* or average
The mean defines




$\sigma$: population *standard deviation*
The standard deviation defines

## Normal or Gaussian Distribution – The "Bell Curve"

- Experimental determination of $\mu$ and $\sigma$ is unrealistic, because they are based on an infinite data set.
- SO, a more realistic goal is to calculate an arithmetic mean:

$$\bar{x} = \frac{\sum_i x_i}{n}$$ , where *n* is the number of samples.

- It is also more realistic to calculate a sample standard deviation:

$$s = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n-1}}$$

Why "n-1"?

Remember $e^2$?

Know how to calculate s on your calculator!

3

## Relating Standard Deviation, Gaussian Distribution and Probability

For ANY Gaussian curve ("normal distribution", random errors):

68.3% of measurements are within ± 1 std. dev. (  or )

95.5% of measurements are within ± 2 std. dev.

99.7% of measurements are within ± 3 std. dev.

We can predict the odds of finding a value within a specific range. It all boils down to area under the curve!

1. Pick a range on the x-axis of the curve
2. Integrate the area under this range (Table 4-1)
3. This area is the probability of observing a value somewhere in this range.

4

## Relating Standard Deviation, Gaussian Distribution and Probability

For example:

50% of the values should be > the mean, and 49.8650% should be between the mean and +3s.

Since 34.13% of the observations fall between the mean and +1s, and 47.73% fall between the mean and +2s, what fraction falls between +1s and +2s?

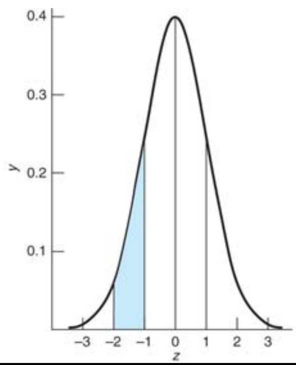
**TABLE 4-1** Ordinate and area for the normal (Gaussian) error curve, $y = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$

| \|z\| [a] | y | Area [b] | \|z\| | y | Area | \|z\| | y | Area |
|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.398 9 | 0.000 0 | 1.4 | 0.149 7 | 0.419 2 | 2.8 | 0.007 9 | 0.497 4 |
| 0.1 | 0.397 0 | 0.039 8 | 1.5 | 0.129 5 | 0.433 2 | 2.9 | 0.006 0 | 0.498 1 |
| 0.2 | 0.391 0 | 0.079 3 | 1.6 | 0.110 9 | 0.445 2 | 3.0 | 0.004 4 | 0.498 650 |
| 0.3 | 0.381 4 | 0.117 9 | 1.7 | 0.094 1 | 0.455 4 | 3.1 | 0.003 3 | 0.499 032 |
| 0.4 | 0.368 3 | 0.155 4 | 1.8 | 0.079 0 | 0.464 1 | 3.2 | 0.002 4 | 0.499 313 |
| 0.5 | 0.352 1 | 0.191 5 | 1.9 | 0.065 6 | 0.471 3 | 3.3 | 0.001 7 | 0.499 517 |
| 0.6 | 0.333 2 | 0.225 8 | 2.0 | 0.054 0 | 0.477 3 | 3.4 | 0.001 2 | 0.499 663 |
| 0.7 | 0.312 3 | 0.258 0 | 2.1 | 0.044 0 | 0.482 1 | 3.5 | 0.000 9 | 0.499 767 |
| 0.8 | 0.289 7 | 0.288 1 | 2.2 | 0.035 5 | 0.486 1 | 3.6 | 0.000 6 | 0.499 841 |
| 0.9 | 0.266 1 | 0.315 9 | 2.3 | 0.028 3 | 0.489 3 | 3.7 | 0.000 4 | 0.499 904 |
| 1.0 | 0.242 0 | 0.341 3 | 2.4 | 0.022 4 | 0.491 8 | 3.8 | 0.000 3 | 0.499 928 |
| 1.1 | 0.217 9 | 0.364 3 | 2.5 | 0.017 5 | 0.493 8 | 3.9 | 0.000 2 | 0.499 952 |
| 1.2 | 0.194 2 | 0.384 9 | 2.6 | 0.013 6 | 0.495 3 | 4.0 | 0.000 1 | 0.499 968 |
| 1.3 | 0.171 4 | 0.403 2 | 2.7 | 0.010 4 | 0.496 5 | ∞ | 0 | 0.5 |

a. $z = (x - \mu)/\sigma$.
b. The area refers to the area between $z = 0$ and $z =$ the value in the table. Thus the area from $z = 0$ to $z = 1.4$ is 0.419 2. The area from $z = -0.7$ to $z = 0$ is the same as from $z = 0$ to $z = 0.7$. The area from $z = -0.5$ to $z = +0.3$ is (0.191 5 + 0.117 9) = 0.309 4. The total area between $z = -\infty$ and $z = +\infty$ is unity.

*Harris, Quantitative Chemical Analysis, 8e*

© 2011 W. H. Freeman




5

## So just how good are your data? How do you know (statistically)?

When we determine an average (with some associated error), how sure are we that the "true value" is close to this average?

- What factors influence this confidence?

The most common statistical tool for determining that the "true" value is close to our calculated mean is the **confidence interval.**

$$\mu = \bar{x} \pm \frac{ts}{\sqrt{n}}$$

The *confidence interval* presents a range about the mean within which there is a fixed probability of finding $\mu$.

6

## Confidence Intervals

$$\mu = \bar{x} \pm \frac{ts}{\sqrt{n}}$$

- Values for t are tabulated based on several confidence levels and various numbers of degrees of freedom.

| | Confidence Level (%) | | | | | | |
|---|---|---|---|---|---|---|---|
| **Degrees of Freedom** | **50** | **90** | ***95*** | **98** | **99** | **99.5** | **99.9** |
| **3** | 0.765 | 2.353 | *3.182* | 4.541 | 5.841 | 7.453 | 12.924 |
| **120** | 0.677 | 1.658 | *1.980* | 2.358 | 2.617 | 2.860 | 3.373 |

- **NOTE:** even though the number of measurements (n) is used in the CI calculation, t is determined based on the *degrees of freedom* (n-1).
  - How can we work to minimize the range calculated at a given confidence interval?
  - How would you cut the CI in half experimentally?

7

## Are two sets of data really different? How do we tell?

- Generally base our determination of the 95% confidence interval.
- If there is greater than 95% probability that the data are the same, we say they do not differ. Less than 95% probability indicates statistically different results.
- Involve calculating a "t" ($t_{calculated}$) and comparing the result to tabulated values for t ($t_{table}$ or $t_{critical}$).
- "*Null Hypothesis*":

**Three different considerations:**

1. Comparing a measured result with a "Known" or "True" value.
2. Comparing two different methods.
3. Comparing differences of multiple samples and two or more methods.

8

## Comparing a measured result with a "Known" or "True" value.

Key question:

***Does the true value fall within our confidence limits?***

- Useful for comparing a result to a standard (i.e. SRM)
- Rearrange confidence limit calculation

$$\mu = \bar{x} \pm \frac{ts}{\sqrt{n}} \longrightarrow t_{calculated} = \frac{|\text{known value} - \bar{x}|}{s}\sqrt{n}$$

- If $t_{calculated} > t_{table}$ at 95% confidence, the results are *statistically different*.

9

## Comparing Two Different Methods

If the results of method A ($\bar{x}_1$, $s_1$) are different from the results of method B ($\bar{x}_2$, $s_2$), is this difference significant?

- Must consider both the *means* and *standard deviations*
- Still compare $t_{calculated}$ and $t_{table}$, but use new calculation

$$t_{calculated} = \frac{\bar{x}_1 - \bar{x}_2}{s_{pooled}}\sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

$$s_{pooled} = \sqrt{\frac{\sum_{set\,A}(x_i - \bar{x}_1)^2 + \sum_{set\,B}(x_j - \bar{x}_2)^2}{n_1 + n_2 - 2}} = \sqrt{\frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2}}$$

  - $n_1 + n_2 - 2$ = number of degrees of freedom
  - If $t_{calculated} > t_{table}$ at 95% confidence, the results *are statistically different*.

This assumes $\sigma$ is the "same" for both data sets. If not, the calculation changes. How do we know? **F-Test**

10

## F-Test for Comparing Standard Deviations

$$F_{calc} = \frac{(s_1)^2}{(s_2)^2} \qquad F \text{ always } \geq 1$$

Compare $F_{calc}$ with $F_{table,}$ if $F_{calc} > F_{table,}$ difference is significant!

**TABLE 4-4** Critical values of $F = s_1^2/s_2^2$ at 95% confidence level

| Degrees of freedom for $s_2$ | Degrees of freedom for $s_1$ | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 12 | 15 | 20 | 30 | ∞ |
| 2 | 19.0 | 19.2 | 19.2 | 19.3 | 19.3 | 19.4 | 19.4 | 19.4 | 19.4 | 19.4 | 19.4 | 19.4 | 19.5 | 19.5 |
| 3 | 9.55 | 9.28 | 9.12 | 9.01 | 8.94 | 8.89 | 8.84 | 8.81 | 8.79 | 8.74 | 8.70 | 8.66 | 8.62 | 8.53 |
| 4 | 6.94 | 6.59 | 6.39 | 6.26 | 6.16 | 6.09 | 6.04 | 6.00 | 5.96 | 5.91 | 5.86 | 5.80 | 5.75 | 5.63 |
| 5 | 5.79 | 5.41 | 5.19 | 5.05 | 4.95 | 4.88 | 4.82 | 4.77 | 4.74 | 4.68 | 4.62 | 4.56 | 4.50 | 4.36 |
| 6 | 5.14 | 4.76 | 4.53 | 4.39 | 4.28 | 4.21 | 4.15 | 4.10 | 4.06 | 4.00 | 3.94 | 3.87 | 3.81 | 3.67 |
| 7 | 4.74 | 4.35 | 4.12 | 3.97 | 3.87 | 3.79 | 3.73 | 3.68 | 3.64 | 3.58 | 3.51 | 3.44 | 3.38 | 3.23 |
| 8 | 4.46 | 4.07 | 3.84 | 3.69 | 3.58 | 3.50 | 3.44 | 3.39 | 3.35 | 3.28 | 3.22 | 3.15 | 3.08 | 2.93 |
| 9 | 4.26 | 3.86 | 3.63 | 3.48 | 3.37 | 3.29 | 3.23 | 3.18 | 3.14 | 3.07 | 3.01 | 2.94 | 2.86 | 2.71 |
| 10 | 4.10 | 3.71 | 3.48 | 3.33 | 3.22 | 3.14 | 3.07 | 3.02 | 2.98 | 2.91 | 2.84 | 2.77 | 2.70 | 2.54 |
| 11 | 3.98 | 3.59 | 3.36 | 3.20 | 3.10 | 3.01 | 2.95 | 2.90 | 2.85 | 2.79 | 2.72 | 2.65 | 2.57 | 2.40 |
| 12 | 3.88 | 3.49 | 3.26 | 3.11 | 3.00 | 2.91 | 2.85 | 2.80 | 2.75 | 2.69 | 2.62 | 2.54 | 2.47 | 2.30 |
| 13 | 3.81 | 3.41 | 3.18 | 3.02 | 2.92 | 2.83 | 2.77 | 2.71 | 2.67 | 2.60 | 2.53 | 2.46 | 2.38 | 2.21 |
| 14 | 3.74 | 3.34 | 3.11 | 2.96 | 2.85 | 2.76 | 2.70 | 2.65 | 2.60 | 2.53 | 2.46 | 2.39 | 2.31 | 2.13 |
| 15 | 3.68 | 3.29 | 3.06 | 2.90 | 2.79 | 2.71 | 2.64 | 2.59 | 2.54 | 2.48 | 2.40 | 2.33 | 2.25 | 2.07 |
| 16 | 3.63 | 3.24 | 3.01 | 2.85 | 2.74 | 2.66 | 2.59 | 2.54 | 2.49 | 2.42 | 2.35 | 2.28 | 2.19 | 2.01 |
| 17 | 3.59 | 3.20 | 2.96 | 2.81 | 2.70 | 2.61 | 2.55 | 2.49 | 2.45 | 2.38 | 2.31 | 2.23 | 2.15 | 1.96 |
| 18 | 3.56 | 3.16 | 2.93 | 2.77 | 2.66 | 2.58 | 2.51 | 2.46 | 2.41 | 2.34 | 2.27 | 2.19 | 2.11 | 1.92 |
| 19 | 3.52 | 3.13 | 2.90 | 2.74 | 2.63 | 2.54 | 2.48 | 2.42 | 2.38 | 2.31 | 2.23 | 2.16 | 2.07 | 1.88 |
| 20 | 3.49 | 3.10 | 2.87 | 2.71 | 2.60 | 2.51 | 2.45 | 2.39 | 2.35 | 2.28 | 2.20 | 2.12 | 2.04 | 1.84 |
| 30 | 3.32 | 2.92 | 2.69 | 2.53 | 2.42 | 2.33 | 2.27 | 2.21 | 2.16 | 2.09 | 2.01 | 1.93 | 1.84 | 1.62 |
| ∞ | 3.00 | 2.60 | 2.37 | 2.21 | 2.10 | 2.01 | 1.94 | 1.88 | 1.83 | 1.75 | 1.67 | 1.57 | 1.46 | 1.00 |

*Critical values of F for a one-tailed test of the hypothesis that $s_1 > s_2$. There is a 5% probability of observing F above the tabulated value.*

*You can compute F for a chosen level of confidence with the Excel function FINV(probability,deg_freedom1,deg_freedom2). The statement "=FINV(0.05,7,6)" reproduces the value F = 4.21 in this table. The statement "=FINV(0.1,7,6)" gives F = 3.01 for 90% confidence.*

Harris, *Quantitative Chemical Analysis*, 8e

© 2011 W. H. Freeman

11

## Comparing Differences of Multiple Samples and Two or More Methods.

Only individual samples have been run, no replicates.

The basis for our decision becomes the average difference between the two methods.

$$t_{calculated} = \frac{\bar{d}}{s_d}\sqrt{n} \qquad s_d = \sqrt{\frac{\sum_i (d_i - \bar{d})^2}{n-1}}$$

If $t_{calculated} > t_{table}$ at 95% confidence, the results *are statistically different*.

12

## Tests for Data Validity: Testing for "outliers"

Useful when one piece of data appears to be outside a reasonable range.

- Tests for statistical probability that the outlier is a member of the same population of the consistent data
- These are statistical tests, but are still subjective and should be used carefully to avoid eliminating useful data!!!

**I. Q-Test**

$$Q_{calculated} = \frac{|gap|}{|range|}$$

*gap* is the difference b/w outlier and nearest value

*range* is total spread of the data.

Compare $Q_{calculated}$ with $Q_{table}$ (typically use 90% confidence)

- If $Q_{calculated}$ is *greater* than $Q_{table}$, there is a statistical probability that the outlier is an invalid data point and may be discarded.
- If $Q_{calculated}$ is *less* than $Q_{table}$, the data point should be retained.

| Number of Observations | $Q_{critical}$ At 90% confidence |
|---|---|
| 3 | 0.94 |
| 4 | 0.76 |
| 5 | 0.64 |
| 6 | 0.56 |
| 7 | 0.51 |
| 8 | 0.47 |
| 9 | 0.44 |
| 10 | 0.41 |

13

## Tests for Data Validity: Testing for "Outliers"

**II. Grubbs Test**

$$G_{calculated} = \frac{|\text{suspect value} - \bar{x}|}{s}$$

Compare $G_{calculated}$ with $G_{table}$

- If $G_{calculated}$ is *greater* than $G_{table}$, there is a statistical probability that the outlier is an invalid data point and should be discarded.
- If $G_{calculated}$ is *less* than $G_{table}$, the data point should be retained.

**TABLE 4-5** Critical values of $G$ for rejection of outlier

| Number of observations | $G$ (95% confidence) |
|---|---|
| 4 | 1.463 |
| 5 | 1.672 |
| 6 | 1.822 |
| 7 | 1.938 |
| 8 | 2.032 |
| 9 | 2.110 |
| 10 | 2.176 |
| 11 | 2.234 |
| 12 | 2.285 |
| 15 | 2.409 |
| 20 | 2.557 |

**Care must be taken to avoid dismissing useful data!**

**Common Sense should be the guide!**

14

## Spreadsheet Tips and Hints

**Excel is great, but no amount of calculation can salvage bad data!**

- When entering calculations, use parentheses at will!
  
  SQRT(23+A5/2) is different than SQRT((25+A5)/2)!!
- Be sure order of operations will be followed correctly
  1. Exponents
  2. Multiplication and Division (left to right)
  3. Addition and Subtraction
- Document your spreadsheet by including cell formulas for critical calculations
- Use absolute references when helpful
  - The dollar sign "locks" a row or column
  - i.e. $B$5 will refer to cell B5 in any calculation, but B$5 will allow the column to vary while the row stays locked at 5
- Learn common built-in functions
  - Things like SUM, STDEV, AVERAGE
  - Check out the Insert→Function menu in Excel
  - "Help" or right-clicking can come in handy, too!

15