Chapter 4
Statistics

I. Gaussian Distribution

II. Confidence Intervals

III. Q-Test

IV. Comparison of Means

V. Calibration Curves

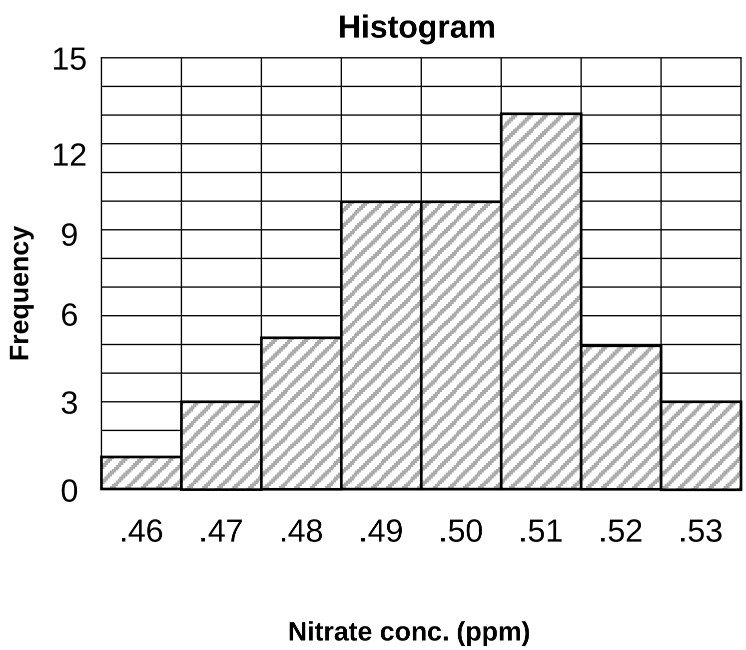
I. Gaussian Distribution

Ex: 50 Replicate Determinations of Nitrate (Handout)

.51	.51	.51	.50	.51
.51	.52	.53	.48	.49
.49	.48	.46	.49	.49
.51	.51	.51	.48	.50
.49	.52	.53	.50	.47
.50	.52	.49	.49	.50
.47	.50	.51	.49	.48
.52	.50	.50	.51	.51
.48	.49	.49	.51	.47
.51	.50	.50	.53	.52

Frequency Table

$[NO_3^-]$	Frequency
0.46	1
0.47	3
0.48	5
0.49	10
0.50	10
0.51	13
0.52	5
0.53	3



I. Gaussian Distribution (cont.)

Two Important Parameters:

1. μ = population mean
2. σ = population standard deviation

We cannot measure μ and σ!

We estimate with:

1. $\bar{x}$ = sample mean (average)
2. s = sample standard deviation

I. Gaussian Distribution (cont.)

Exercise 4-C, p 73: Suppose the mileage at which 10,000 sets of automobile brakes had been 80% worn through was recorded. The average was 62,700 miles, and the standard deviation was 10,400 miles

a) What fraction of brakes is expected to be 80% worn in less than 40,860 miles?

b) What fraction is expected to be 80% worn at a mileage between 57,500 and 71,020 miles?

I. Gaussian Distribution (cont.)

Exercise 4-C, p 73: Suppose the mileage at which 10,000 sets of automobile brakes had been 80% worn through was recorded. The average was 62,700 miles, and the standard deviation was 10,400 miles

a) What fraction of brakes is expected to be 80% worn in less than 40,860 miles?

b) What fraction is expected to be 80% worn at a mileage between 57,500 and 71,020 miles?

II. Confidence Intervals

$\bar{x}$ and s estimate the shape of the bell curve. How confident can we be that the sample mean is close to the population mean?

We can compute a “confidence interval.”

Table 4-2 Values of Student's *t*

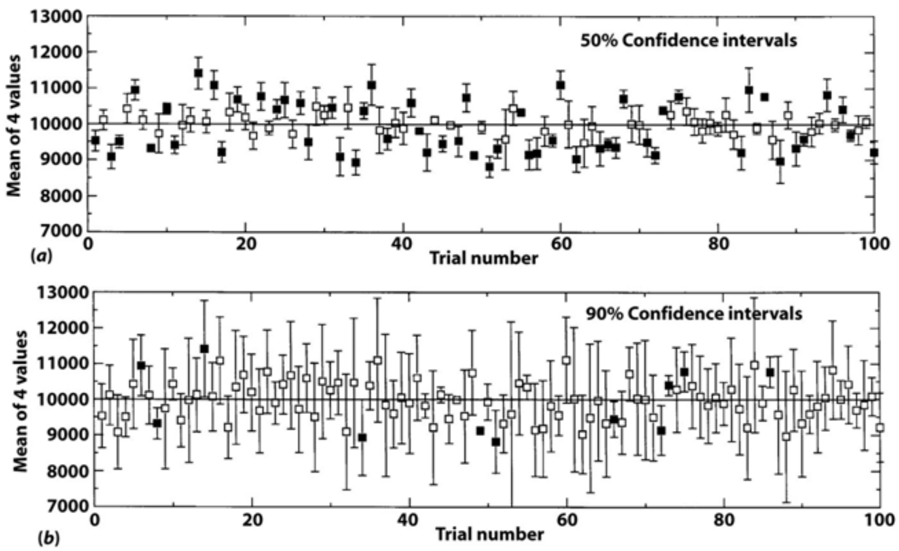
	Confidence level (%)						
Degrees of freedom	50	90	95	98	99	99.5	99.9
1	1.000	6.314	12.706	31.821	63.656	127.321	636.578
2	0.816	2.920	4.303	6.965	9.925	14.089	31.598
3	0.765	2.353	3.182	4.541	5.841	7.453	12.924
4	0.741	2.132	2.776	3.747	4.604	5.598	8.610
5	0.727	2.015	2.571	3.365	4.032	4.773	6.869
6	0.718	1.943	2.447	3.143	3.707	4.317	5.959
7	0.711	1.895	2.365	2.998	3.500	4.029	5.408
8	0.706	1.860	2.306	2.896	3.355	3.832	5.041
9	0.703	1.833	2.262	2.821	3.250	3.690	4.781
10	0.700	1.812	2.228	2.764	3.169	3.581	4.587
15	0.691	1.753	2.131	2.602	2.947	3.252	4.073
20	0.687	1.725	2.086	2.528	2.845	3.153	3.850
25	0.684	1.708	2.060	2.485	2.787	3.078	3.725
30	0.683	1.697	2.042	2.457	2.750	3.030	3.646
40	0.681	1.684	2.021	2.423	2.704	2.971	3.551
60	0.679	1.671	2.000	2.390	2.660	2.915	3.460
120	0.677	1.658	1.980	2.358	2.617	2.860	3.373
∞	0.674	1.645	1.960	2.326	2.576	2.807	3.291

NOTE: In calculating confidence intervals, σ may be substituted for s in Equation 4-6 if you have a great deal of experience with a particular method and have therefore determined its "true" population standard deviation. If σ is used instead of s, the value of t to use in Equation 4-6 comes from the bottom row of Table 4-2.

Table 4-2
Quantitative Chemical Analysis, Seventh Edition

What is the meaning of a confidence interval?

Example from book: Computer generates four random numbers (100 trials, μ = 10 000, σ = 1000)



III. Q-Test

<u>Q-Test for Outliers:</u>

1. Arrange the data in order of increasing value.
2. Calculate Q_{calc}
3. Look up Q_{table}
4. If $Q_{calc} > Q_{table}$, throw the point out.

Table 4-5 Values of Q for rejection of data

Q (90% confidence)[a]	Number of observations
0.76	4
0.64	5
0.56	6
0.51	7
0.47	8
0.44	9
0.41	10

IV. Comparison of Means

Handout w/ Excel

V. Calibration Curves

Some definitions:

1. analyte – component of a sample to be determined
2. standard – solution containing known amount of analyte
3. blank – solution with all reagents and solvents of sample, but no analyte
4. calibration curve – curve of “best fit” to response of standards and blank

Calculating the line of "best fit":

- "Linear regression" or "Method of least squares"
- Best line through points minimizes magnitude of (y_i – y_{calc}) for all *i*.
- Calculating the important parameters:
  - slope = *m* (Eqn. 4-16)
  - y-intercept = *b* (Eqn. 4-17)
  - standard deviation of *m* = s_m (Eqn. 4-21)
  - standard deviation of *b* = s_b (Eqn. 4-22)
  - standard deviation of y_{calc} = s_y (Eqn. 4-20)
  - standard deviation of x_{calc} = s_x (Eqn. 4-27)

USE EXCEL!!!