Elementary Statistics
FORMULAE AND TABLES

For a sample $x_1, \ldots, x_n$, the *mean* is $\bar{x} = (1/n)(\sum x)$. The *sample variance* is

$$s^2 = \frac{1}{n}\sum(x-\bar{x})^2 = \left(\frac{1}{n}\sum x^2\right) - \bar{x}^2.$$

The *sample standard deviation* is s.
The *estimator for the population variance* is $\hat{s}^2 = \frac{n}{n-1}s^2$. The *median* is the middle value (or average of the middle two values) when the values are arranged in order of size.
The *least squares fit* to a set of data points $(x_1, y_1), \ldots, (x_n, y_n)$ is the line $y = mx + c$ where

$$m = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2} = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2}$$

and $\bar{y} = m\bar{x} + c$.
Hypothesis testing. If the null hypothesis is that the sample comes from a population with mean M and variance σ^2, then the standard error of the sample mean is $s_e = \sigma/\sqrt{n}$. Compare $|M - \bar{x}|/s_e$ with the table of significance levels.

Table 1 The area under a Gaussian (normal) curve corresponding to multiples of standard deviations (s.d.) from the mean. For example, the area lying within 0.5 standard deviations of the mean is 0.383.

No. of s.d.	Area	No of s.d.	Area	No. of s.d.	Area
0.00	0.000				
0.10	0.080	1.10	0.729	2.10	0.964
0.20	0.159	1.20	0.770	2.20	0.972
0.30	0.236	1.30	0.806	2.30	0.979
0.40	0.311	1.40	0.838	2.40	0.984
0.50	0.383	1.50	0.866	2.50	0.988
0.60	0.451	1.60	0.890	2.60	0.991
0.70	0.516	1.70	0.911	2.70	0.993
0.80	0.576	1.80	0.928	2.80	0.995
0.90	0.632	1.90	0.943	2.90	0.996
1.00	0.683	2.00	0.954	3.00	0.997

Table 2 Multiples of standard errors corresponding to various significance levels.

Significance (%)	Number of standard errors
10	1.64
5	1.96
1	2.58
0.1	3.29