Statistical Analysis

For all Sciences

Error Calculations

- Error Calculations are a form of statistics that are used to mathematically determine if some trend is accurately predicating a set of outcomes.
- Each Calculation has a limited but accepted range of use.
- Scientists (Natural and Social) and Mathematicians have come to a common agreement about when and where to use each analysis.
- IN GENERAL- Each calculationis giving us Accuracy and/or Precision

Error Calculations

Percent Error

- (Accuracy Calculation)
- Used to determine the percent difference between the theoretical and experimental values

$$\%Error = \frac{Theoretical - Exprimental}{Theoretical} x100\%$$

accepted value

your value

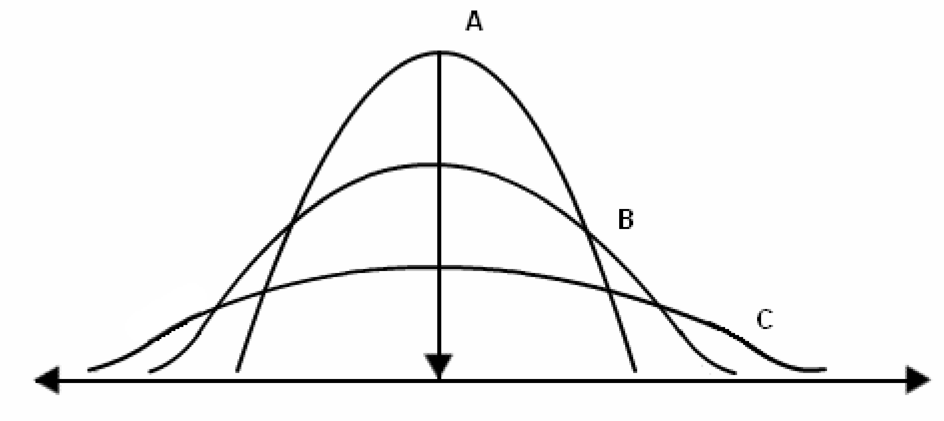
Error Calculations

- Mean Absolute Deviation - "Average Deviation from the mean" (a.d.)
- (Precision Calculation)
  - Used to determine how precise a given measurement is compared to other measurements
  - Calculate the absolute deviation from the average
  - Average the deviations

Error Calculations

- (Accuracy Calculation)
- Average Deviation of the mean (A.D.)
  - Used to calculate how accurate a set of measurements are compared to the theoretical value
  - (a.d.) / sqrt (the number of time the measurement was made)

Interval Range



- The interval range for Mean Absolute Deviation (a.d.) is 57.7%
  - This means 57.5% of all data randomly collected should fall in the range of the a.d.

- The interval range of the Average Deviation from the mean is 50%
  - This means that means that there is a 50% chance that the accepted (or true) value falls in the range.

Variance

- Statistical variance gives a measure of how the data distributes itself about the mean or expected value. Unlike range that only looks at the extremes, the variance looks at all the data points and then determines their distribution.
- Although variance could describe the a.d., it usually is represented by the following equation.

$$\sigma^2 = \frac{\sum(X-\mu)^2}{N}$$

  - σ^2 = variance
  - $\sum (X - \mu)^2$ = The sum of $(X - \mu)^2$ for all data points
  - X = individual data points
  - μ = mean of the data
  - N = number of data points

Standard Deviation

- The most common statistical predictor of precision

$$\sigma = \sqrt{\frac{\sum(X-\mu)^2}{n}}$$

  - X = individual data points
  - μ = mean of the data
  - N = number of data points

Note: Although it is the most common,
it may not be the best.

Standard Deviation

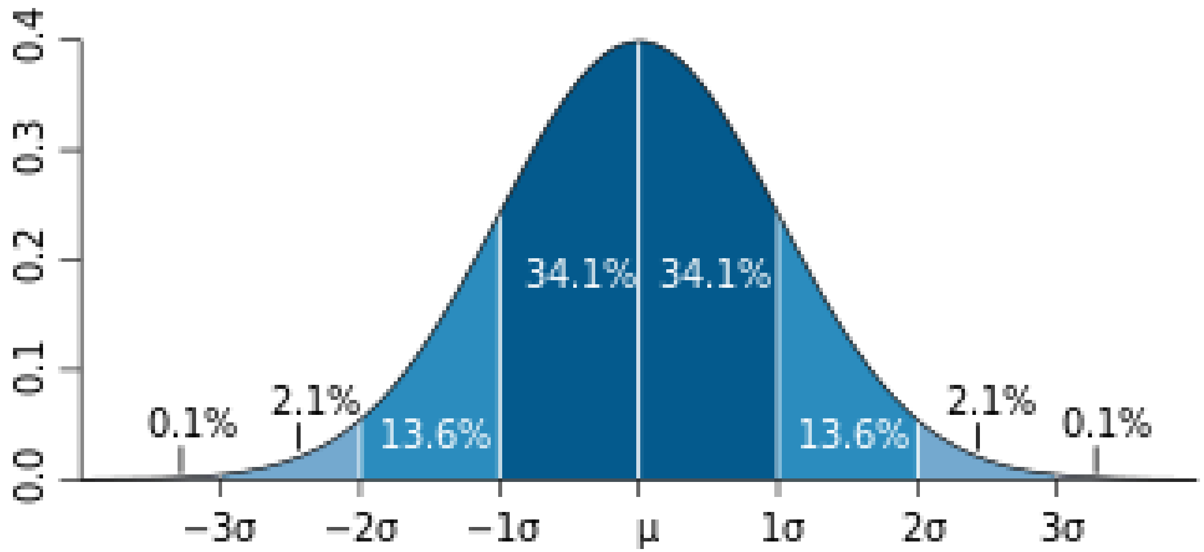
- The most common statistical predictor of precision

$$\sigma = \sqrt{\frac{\Sigma(X-\mu)^2}{n}}$$

  - X = individual data points
  - μ = mean of the data
  - N = number of data points

Note: Although it is the most common,
it may not be the best.

Interval Range



- 1 Standard Deviation = 68.27%
- 2 Standard Deviations = 95.45%
- 3 Standard Deviations = 99.73%
- 4 Standard Deviations = 99.99%

Standard Deviation vs. Mean Absolute Deviation

- Standard deviation exaggerates the impact of larger deviations. This emphasizes point that are outliers.
- Mean Absolute reduces the weight of outliers. And therefore the value does not reflect the impact of larger scatter or dispersion properly.
- For Data Sets with wide ranges of data, Standard Deviation is Better.
- For Sets of data where errors may be greater than 3% Mean Absolute Deviation is better.
  - Note in labs we usually assume error as 0% so Standard Deviation is used more often.

Standard Deviation of the Mean

- Similar to Average Deviation of the Mean in that it tells the Accuracy of our data.

$$\boxed{\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}}$$

- It's interval range is 50%

$$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2}$$

Sample vs Population

- When we are using all of the data (The population) we have we use:

$$\sigma^2 = \frac{\Sigma(X-\mu)^2}{N}$$

- Sometimes we need to take a sample data from a larger population. To err on the side of caution we increase our error by subtracting 1 from N.

$$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2}$$

- The mean is now **x** (for sample mean) instead of **μ** (the population mean),
- And the answer is **s** (for Sample Standard Deviation) instead of **σ**.

Addition information on Standard Deviation

- Instructional Video
  - http://www.youtube.com/watch?v=HvDqbzu0i0E

- Example Problems
  - http://www.mathsisfun.com/data/standard-deviation-formulas.html

- Example Physics Problem
  - http://www.batesville.k12.in.us/physics/apphynet/Measurement/standard_deviation.htm

Student T-Test

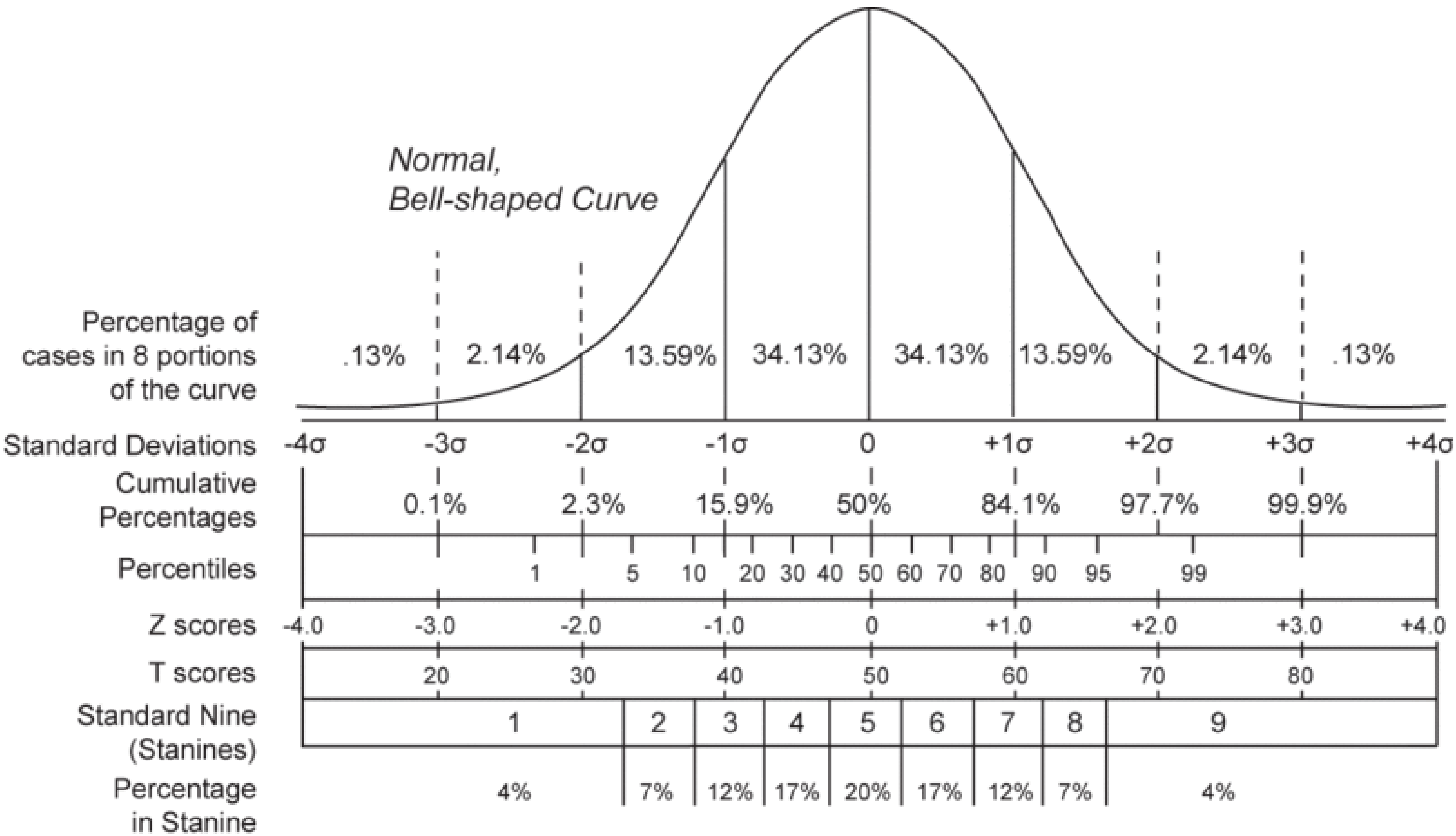
- The T-Test test can be used to test the accuracy of your data to the theoretical value.

- T-Tests are used for data sets of 30 pieces of data or less. There is another test called the Z-Test for data sets larger than 30.

- The equation for the t-test is
- $t = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = \frac{\bar{x}-\mu}{\sigma_{\overline{x}}}$

T-Test Table

α	0.25	0.20	0.15	0.1	0.05	0.025	0.01	0.005	0.0025	0.0010	0.0
One Sided	75%	80%	85%	90%	95%	97.50%	99%	99.50%	99.75%	99.90%	99.
Two Sided	50%	60%	70%	80%	90%	95%	98%	99%	99.50%	99.80%	99.
1	1	1.376	1.963	3.078	6.314	12.71	31.82	63.66	127.3	318.3	63
2	0.816	1.061	1.386	1.886	2.92	4.303	6.965	9.925	14.09	22.33	3
3	0.765	0.978	1.25	1.638	2.353	3.182	4.541	5.841	7.453	10.21	12
4	0.741	0.941	1.19	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.
5	0.727	0.92	1.156	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.
6	0.718	0.906	1.134	1.44	1.943	2.447	3.143	3.707	4.317	5.208	5.
7	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.
8	0.706	0.889	1.108	1.397	1.86	2.306	2.896	3.355	3.833	4.501	5.
9	0.703	0.883	1.1	1.383	1.833	2.262	2.821	3.25	3.69	4.297	4.
10	0.7	0.879	1.093	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.
11	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.
12	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.428	3.93	4.
13	0.694	0.87	1.079	1.35	1.771	2.16	2.65	3.012	3.372	3.852	4.
14	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.
15	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.
16	0.69	0.865	1.071	1.337	1.746	2.12	2.583	2.921	3.252	3.686	4.
17	0.689	0.863	1.069	1.333	1.74	2.11	2.567	2.898	3.222	3.646	3.
18	0.688	0.862	1.067	1.33	1.734	2.101	2.552	2.878	3.197	3.61	3.
19	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.
20	0.687	0.86	1.064	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.
21	0.686	0.859	1.063	1.323	1.721	2.08	2.518	2.831	3.135	3.527	3.
22	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.
23	0.685	0.858	1.06	1.319	1.714	2.069	2.5	2.807	3.104	3.485	3.
24	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.
25	0.684	0.856	1.058	1.316	1.708	2.06	2.485	2.787	3.078	3.45	3.
26	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.
27	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.
28	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.
29	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.
30	0.683	0.854	1.055	1.31	1.697	2.042	2.457	2.75	3.03	3.385	3.


Normal,
Bell-shaped Curve
Percentage of cases in 8 portions of the curve
.13%
2.14%
13.59%
34.13%
34.13%
13.59%
2.14%
.13%
Standard Deviations
-4σ -3σ -2σ -1σ 0 +1σ +2σ +3σ +4σ
Cumulative Percentages
0.1% 2.3% 15.9% 50% 84.1% 97.7% 99.9%
Percentiles
1 5 10 20 30 40 50 60 70 80 90 95 99
Z scores
-4.0 -3.0 -2.0 -1.0 0 +1.0 +2.0 +3.0 +4.0
T scores
20 30 40 50 60 70 80
Standard Nine (Stanines)
1 2 3 4 5 6 7 8 9
Percentage in Stanine
4% 7% 12% 17% 20% 17% 12% 7% 4%