IEEE SW Test Workshop
Semiconductor Wafer Test Workshop

June 12 to 15, 2011
San Diego, CA

Statistical Analysis Fundamentals for Wafer Test

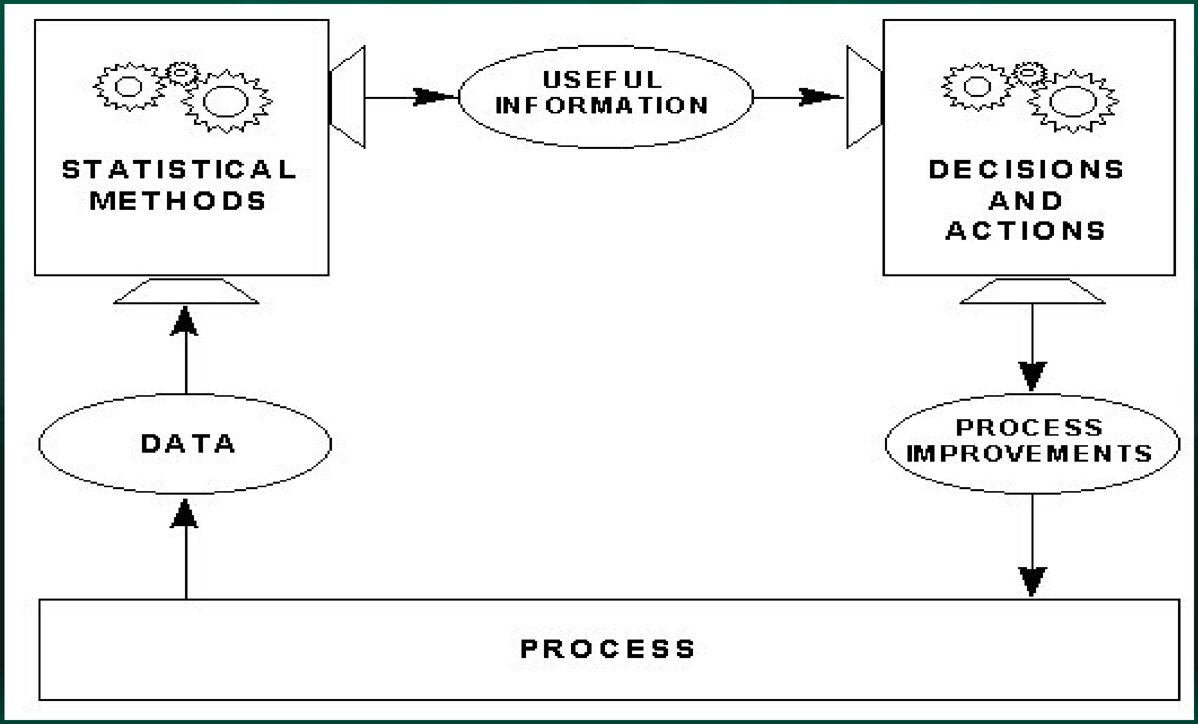
Lance Milner

Intel Fab 12, Chandler, AZ

- **Why Statistics?**
- **“In God we trust. All others bring data.”**
  - – W Edwards Deming (1900-1993)

Data Driven Decision Making

- **We use statistics to learn about a process or piece of equipment and to make decisions.**
  - Is this technology better than that one ?
  - Is the second metrology tool producing equivalent results to the first one?
  - Is the manufacturing process stable and capable ?



Statistical Tools, Techniques and Concepts

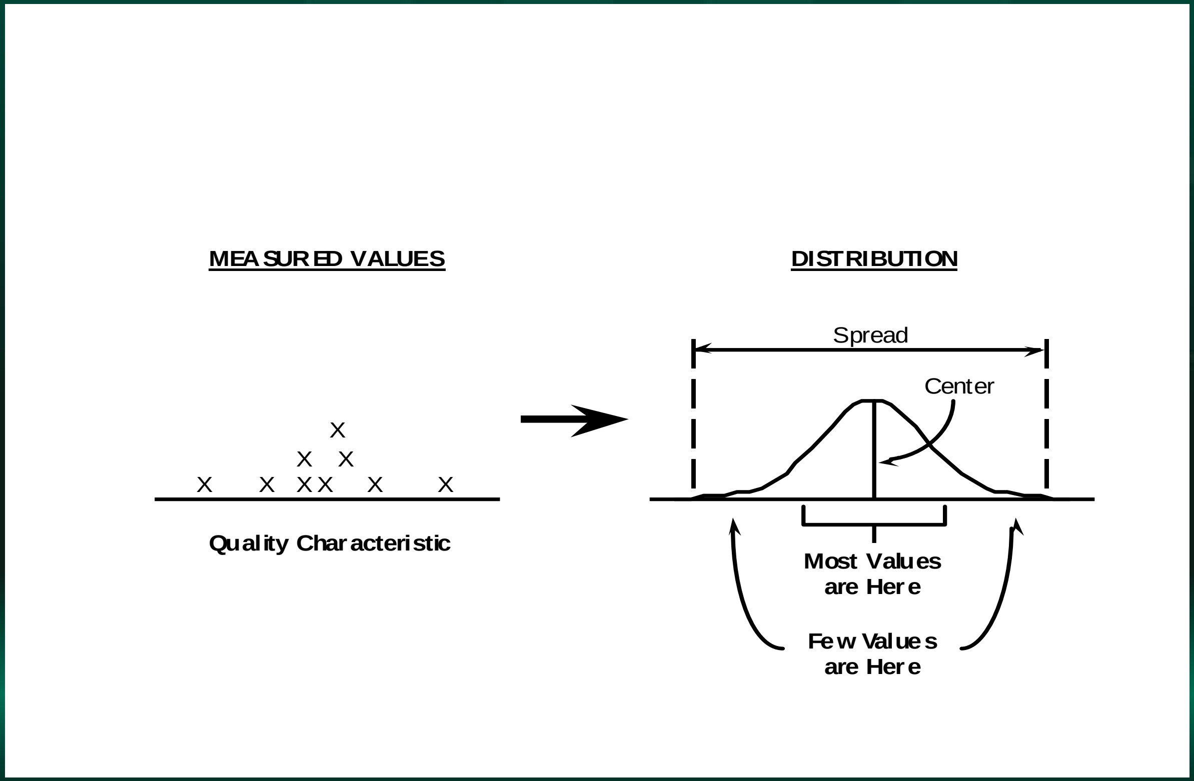
- **Graphical techniques to study data**
    - Distributions, Trend charts, box plots, probability plots, control charts etc

– Analytical techniques to process data
  - Summary and descriptive metrics (Mean, Standard Deviation etc)
  - Decision making and hypothesis testing

– Concepts and Models
  - Distributions and probability
  - Confidence intervals
  - Risk assessments

Visualizing Data : Distributions

- Distributions are fundamental to a lot of statistical analysis.
- They are also very useful to quickly understand the structure of a dataset.

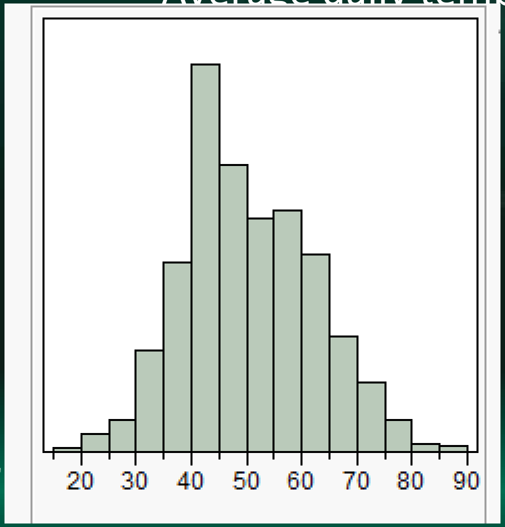


Distributions, Continued

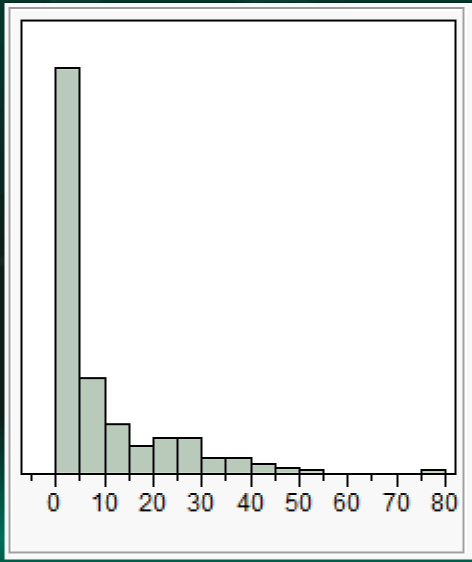
- **A distribution is described by its <u>shape</u>, <u>center</u>, and <u>spread</u>.**
  - Shape is determined by:
    - Symmetry
    - Modality
    - Outliers
  - Common measures of center:
    - Mean
    - Median
  - Common measures of spread:
    - Range, Interquartile Range (IQR)
    - Standard Deviation, Variance

Distribution Shape : Symmetry

- **Symmetric**
  - Portions above and below the center are roughly mirror images.
  - Examples:
    - Heights of females.
    - Measurement errors.
    - Average daily temperatures

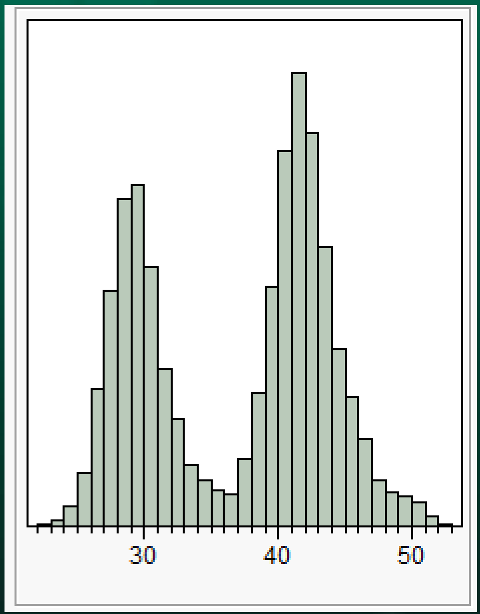


- Skewed
- Skewed left - long tail to the left.
- Skewed right - long tail to the right.
- Examples:
  - Total daily rainfall (right)
  - Annual salaries (right)



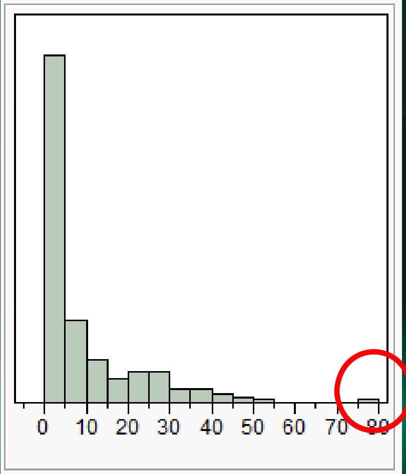
Distribution Shape - Modality

- **Modality**
  - Mode: Most frequent value.
  - Modality: Number of peaks in a dataset.
    - Bi-Modal – Two peaks
    - Tri-Modal – Three peaks
    - Multi-Modal – Many peaks
  - Usually the result of the combination of multiple groups of data
  - Examples :
    - Heights of American and Japanese males (bimodal)
    - Length measurements from 3 machines that have not been calibrated and matched



Distribution Shape - Outliers

- **Outliers are observations that fall outside the rest of the distribution.**
- **Outliers are usually the result of special causes.**
  - Do not simply delete and ignore outliers !
  - Their cause needs to be identified and understood.
  - Analysis should be done with and without outliers to understand their effect.

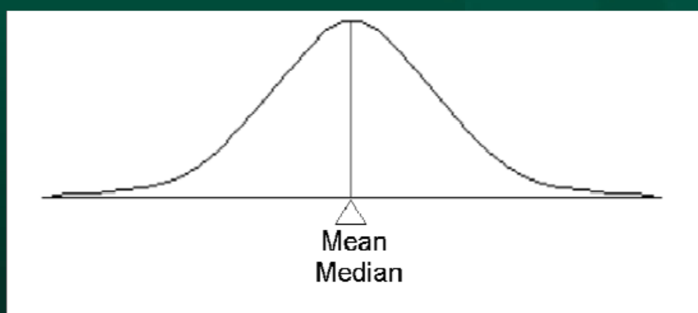


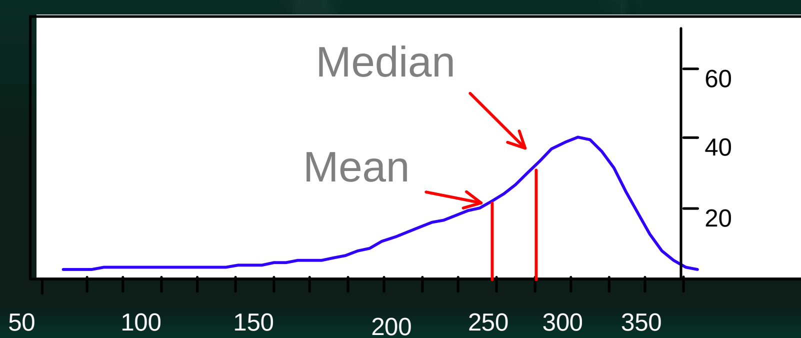
Outliers: Potential Mistakes

- Do not simply delete and ignore outliers !
- Their cause needs to be identified and understood. They almost always have *some* meaning.
- Analysis should be done with and without outliers to understand their effect.

Distributions: Measures of Center

Center	mean	sample average: avg, xbar $\overline{x} = \sum_{i=1}^{n} x_i \Big/ n$
	median	50th percentile: P50, med, m
	mode	peak, most frequently occurring value





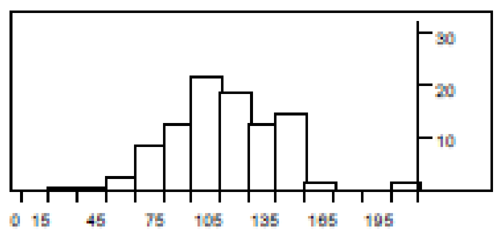
- **The mean and median are equivalent for symmetric distributions.**

- The mean is influenced by extreme values.
- Skewed left: mean < median.
- Skewed right: mean > median

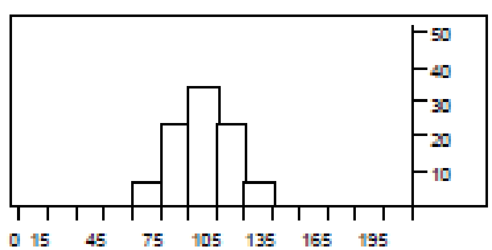
Distributions: Measures of Spread

range **R**	The range is the simplest measure of spread, calculated by subtracting the minimum value from the maximum value. The range was commonly used in the era before calculators and computers.
std dev **s**	computed by: $s = \sqrt{\sum (x_i - \bar{x})^2 / (N-1)}$
Variance **s^2**	The *variance* is the square of the std dev: $s^2 = \sum (x_i - \bar{x})^2 / (N-1)$

Examples



mean = 100
std dev = 30
range = 176



mean = 100
std dev = 16
range = 73

Distributions Wrapup

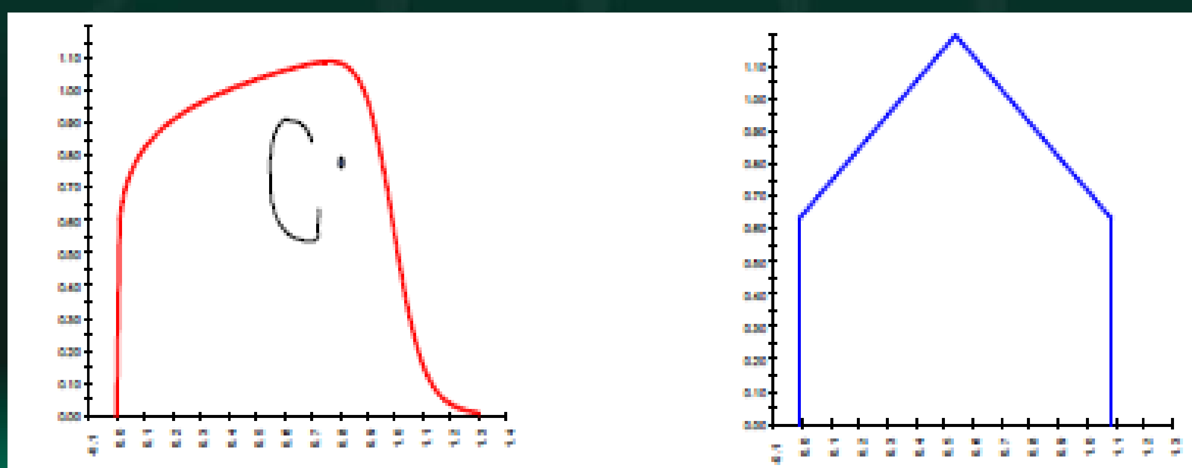
- Group of numbers put together is known as a 'distribution'.
- While many things are happening within a distribution, and they may be very large, the key information about a distribution can be summarized by three concepts
  - Shape
  - Center
  - Spread

Graphical Tools

Graphical Representation of Data

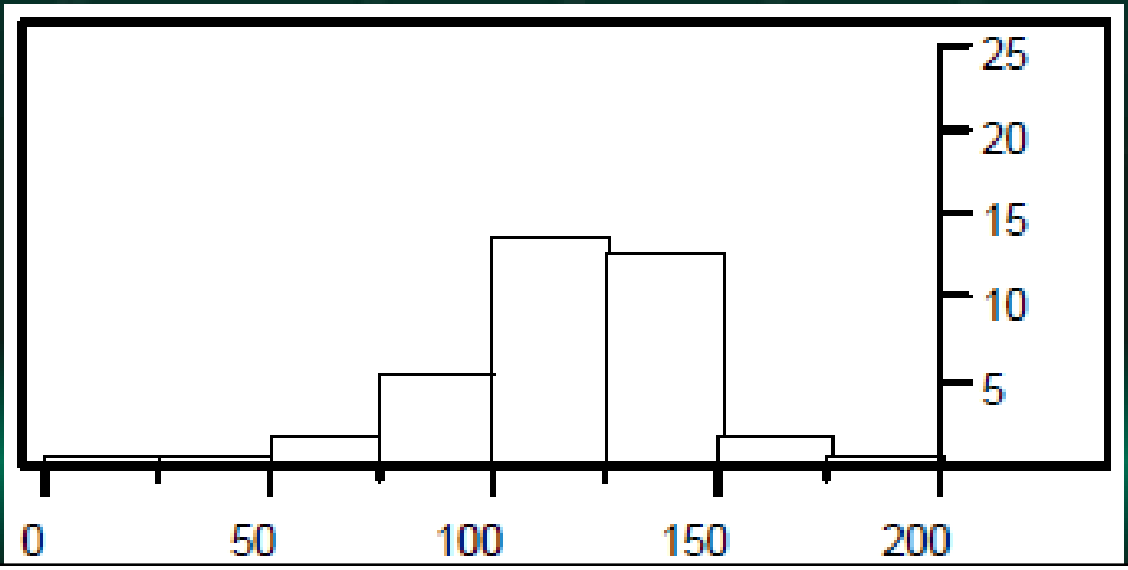
- **"You can tell a lot just by looking at it"**
  - Yogi Berra

avg	= 0.5387	avg	= 0.5387
stdev	= 0.2907	stdev	= 0.2906
skewness	= 0.0000	skewness	= 0.0000
kurtosis	= 2.0000	kurtosis	= 2.0000



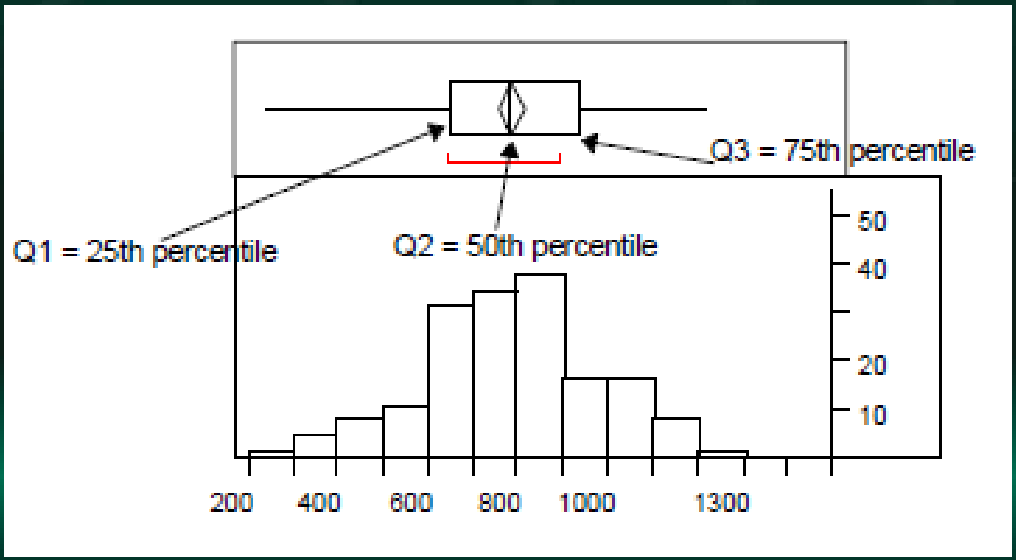
Graphical Tools: Histogram

- Most common and simple tool for data representation.
- Effective for identifying center, spread, and shape of distribution.
- Less effective for comparing multiple distributions.



Graphical Tools: Boxplots

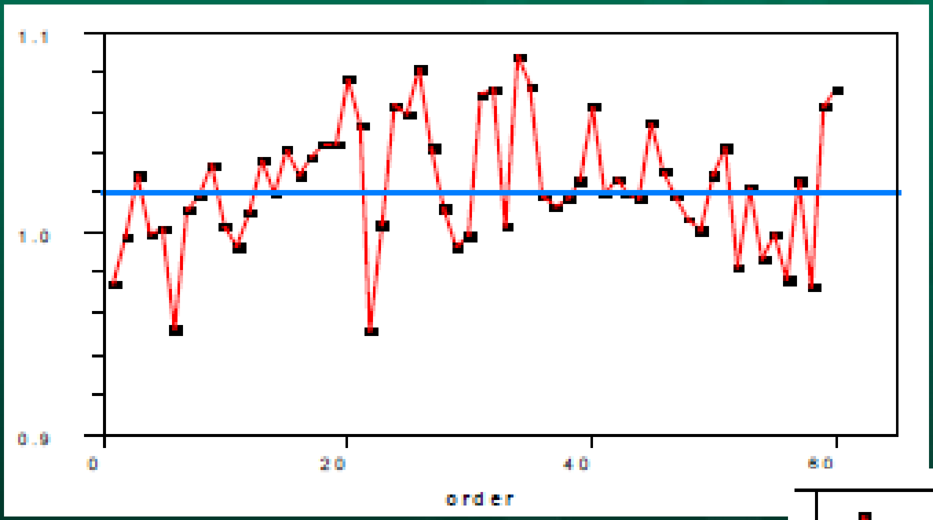
- Also known as a 5 number summary. Will show the Minimum, Maximum, Q1, Q3, and Median.
- Can loose some information of a distribution, but can be used to effectively demonstrate distribution differences.



Graphical Tools: Trend Plots

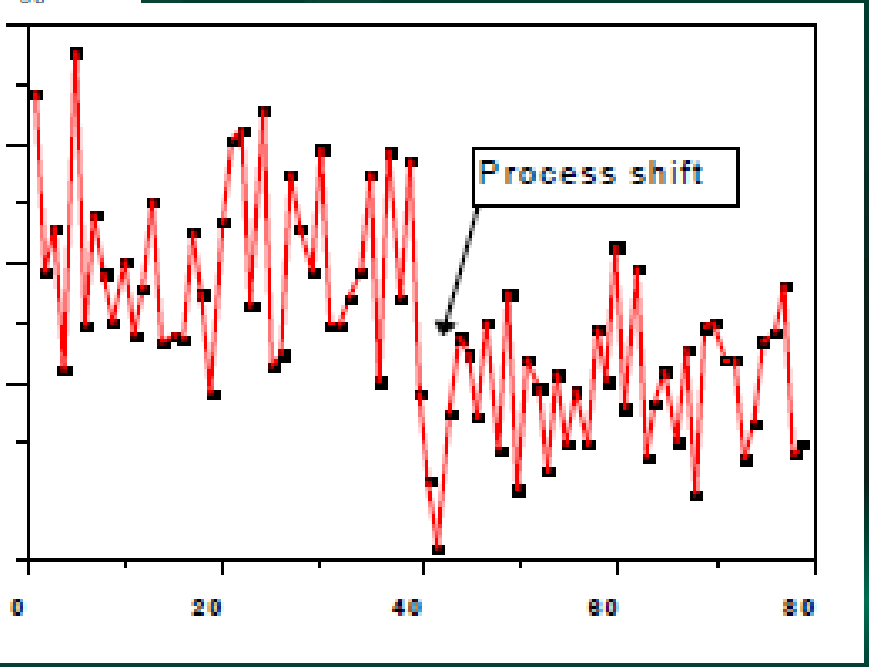
- Used to determine stability of a distribution over time. Can identify shifts in the data and detect subtle trends.
- Distribution of interest on Y axis, time on X axis.
- A stable process has mean and standard deviation consistent over a period of time.
- An extension of the Trend Plot is the Control Charts, as the basis for Statistical Process Control.

Graphical Tools: Trend Plots



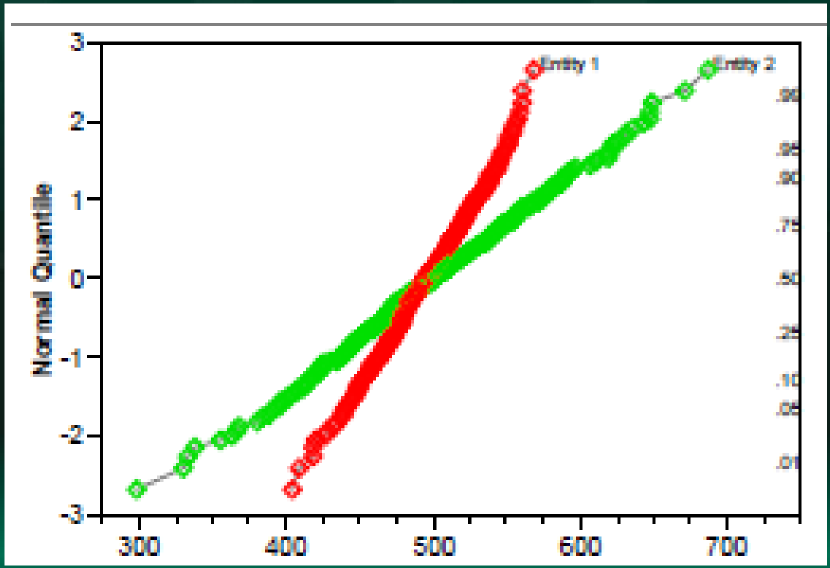
Is this process stable?

Precise timing of process shift identified.

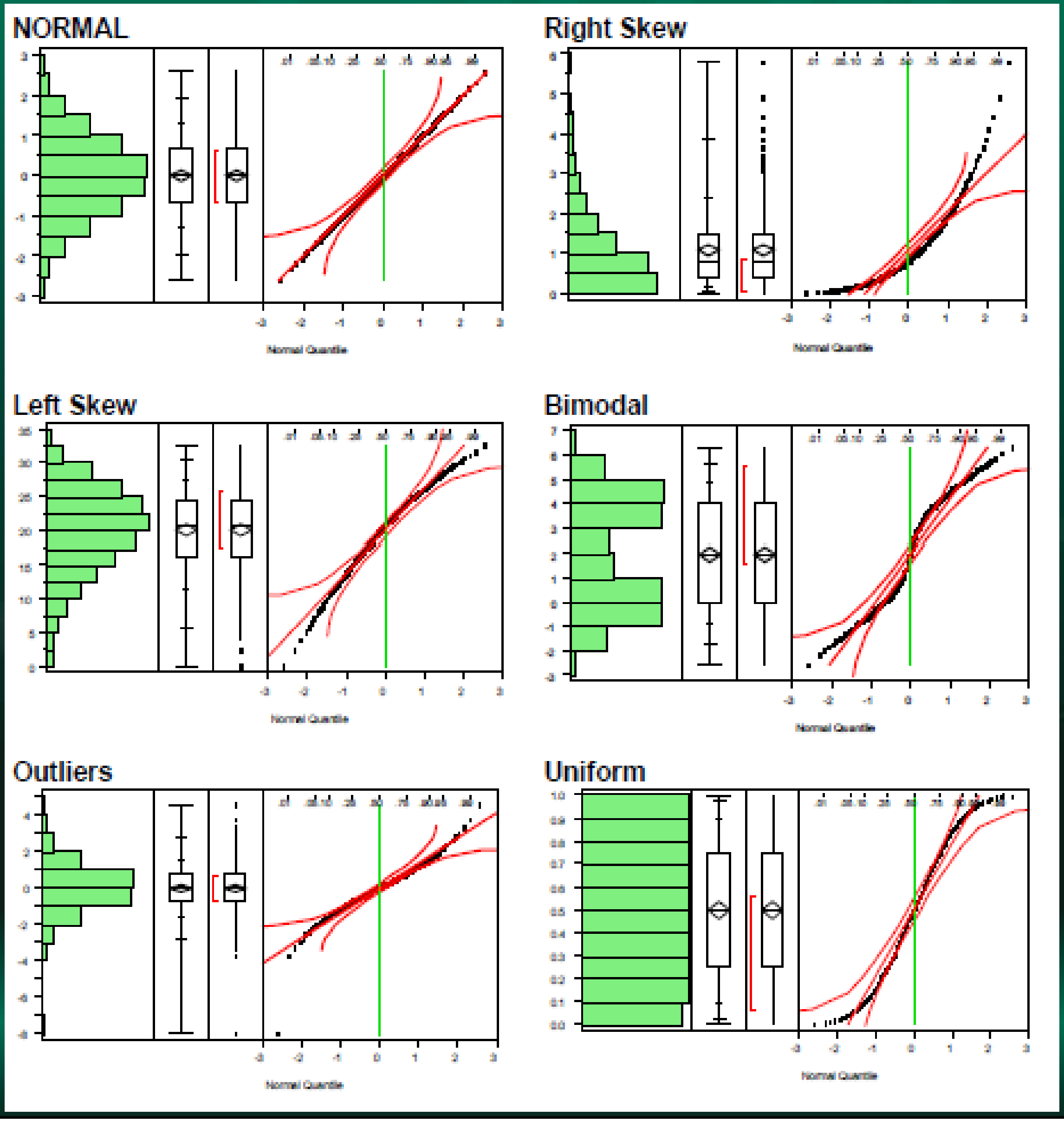


Graphical Tools: Normal Probability Plots

- Can be used to test ‘normality’ of distribution.
- Plots distribution on X axis, normalized value on Y axis.
- Effective for comparing multiple distributions at once.

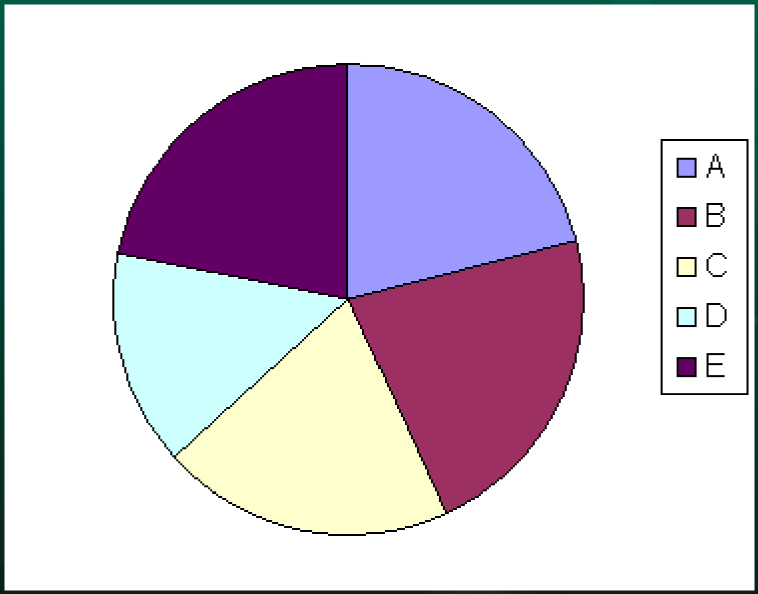


Graphical Tool Comparisons



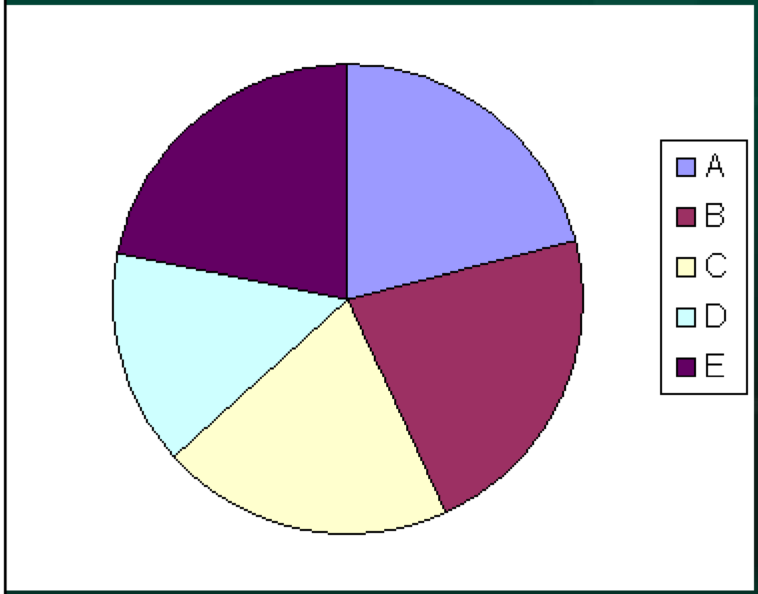
Graphical Mistakes

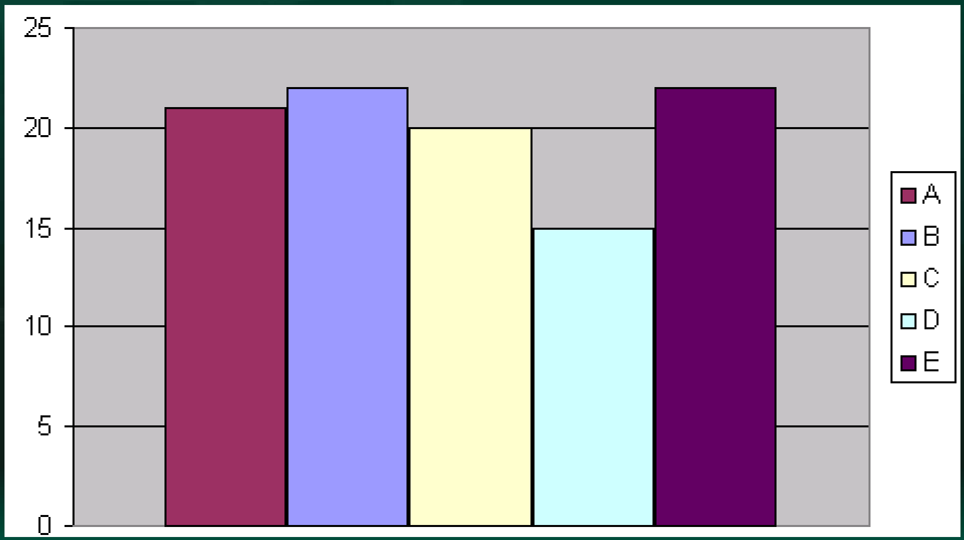
- **Which piece of the pie is bigger?**



Graphical Mistakes

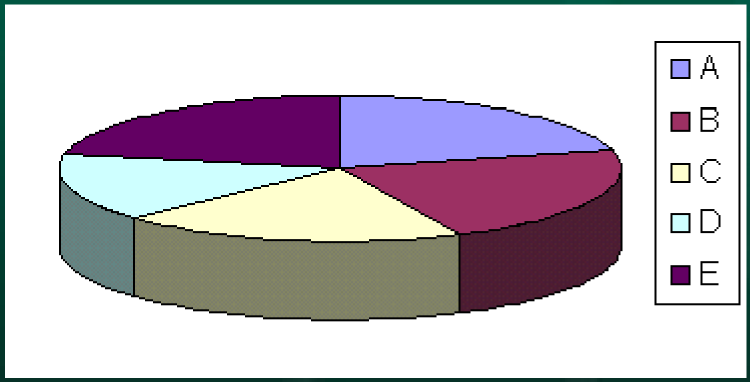
- The bar graph is better for categorical data. Arc length of circle very difficult to detect with human eye.

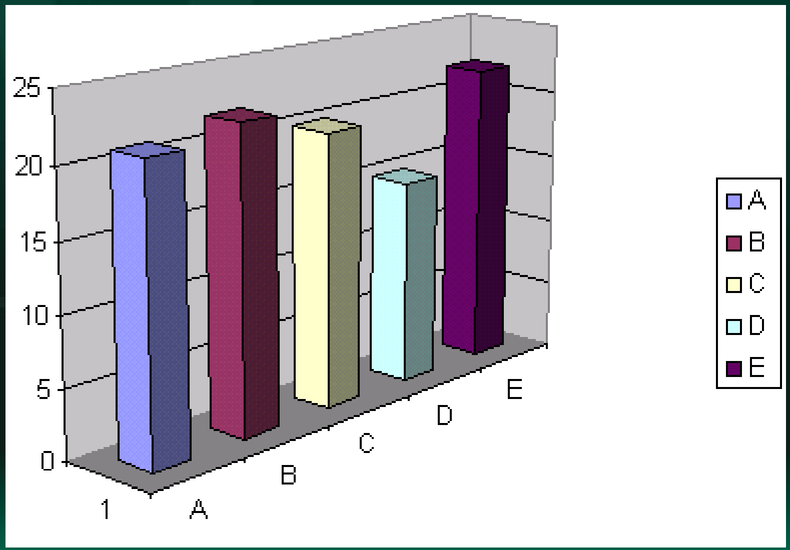




Graphical Mistakes

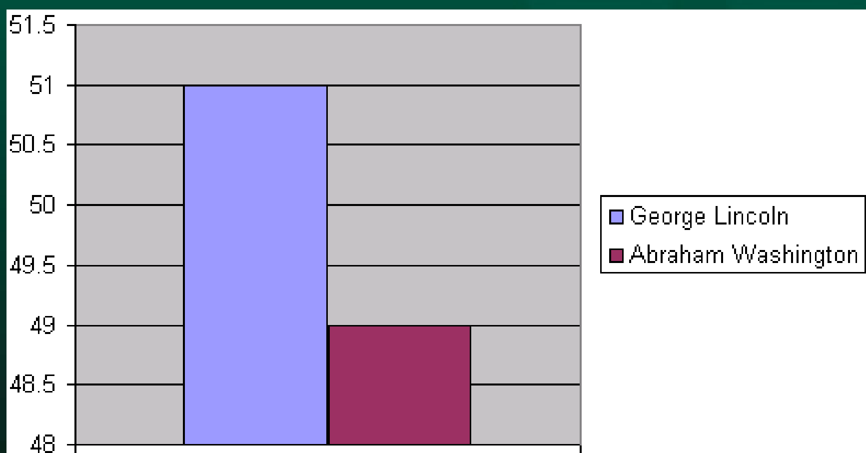
- **3D graphs can be difficult to read**

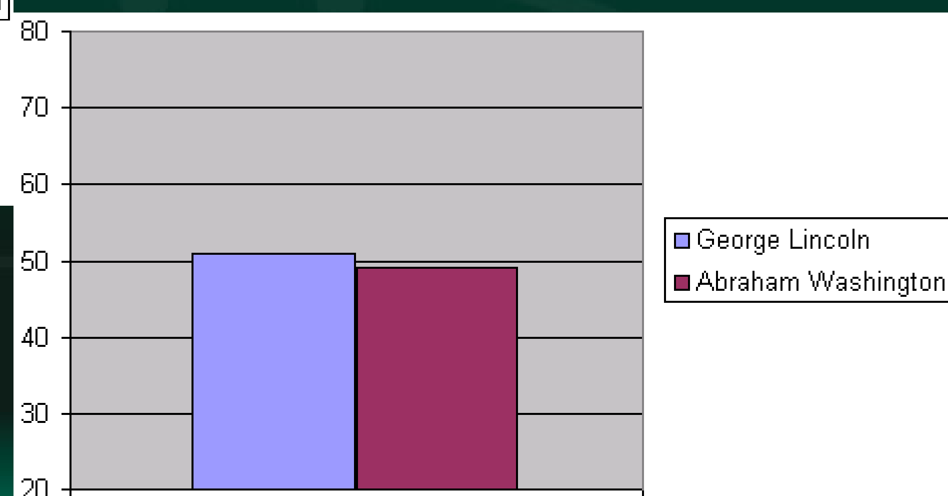




Graphical Mistakes

- Always beware of the scale of the graph. Which shows a larger gap?





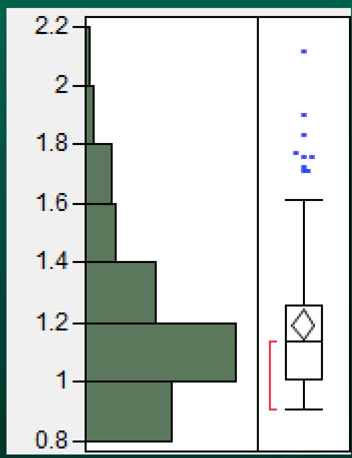
CPK Example Data

- Data in this example is a comparison of 3 different technologies (using made up data) comparing the CPK of probe marks.
- CPK is a measure of how close a distribution is to violating a spec limit. In this case, how close the probe marks are to the edge of the pad.
- CPK value of less than 1.0 is considered to indicate low level of quality. CPK value above 1.3 is an indication of higher quality.

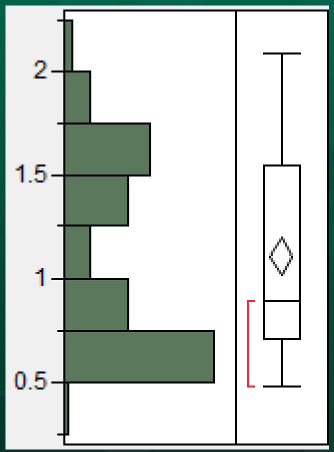
Summary data

Technology	Mean	SD
A	1.19	0.26
B	1.11	0.44
C	1.55	0.32

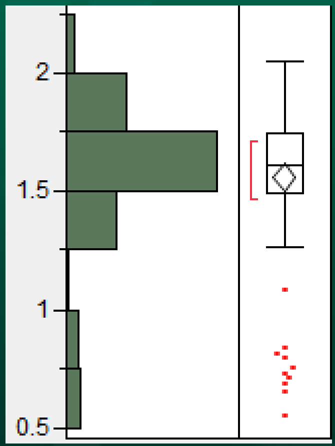
Histograms



A is skewed

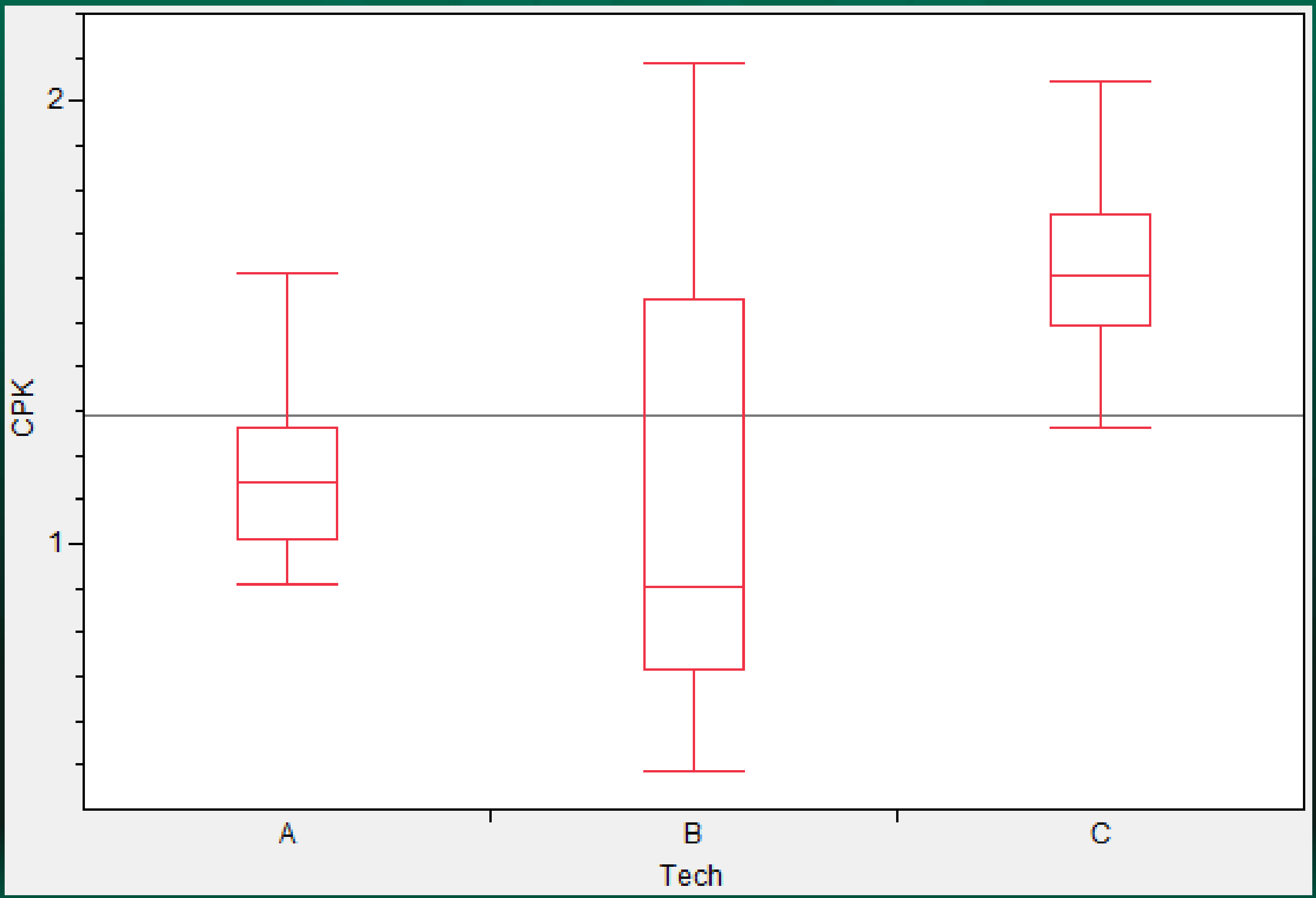


B is bi-modal

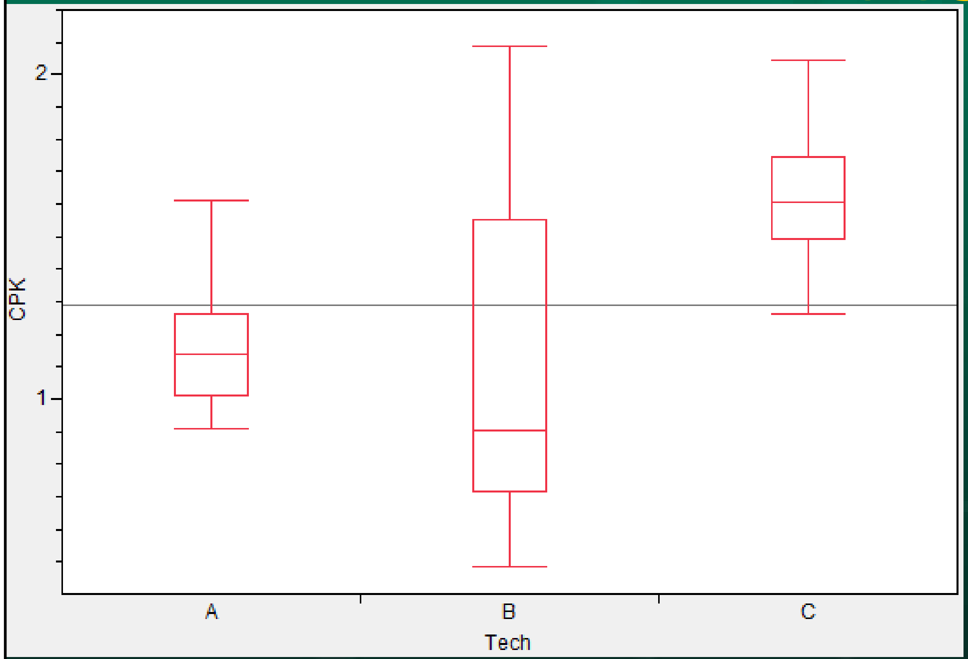


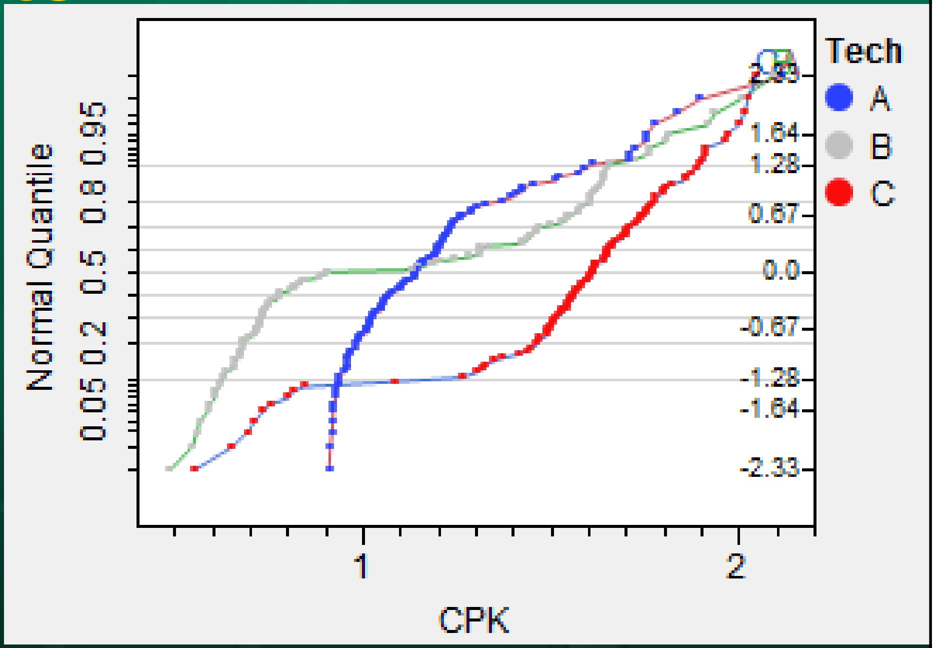
C has outliers

Box Plot Comparison



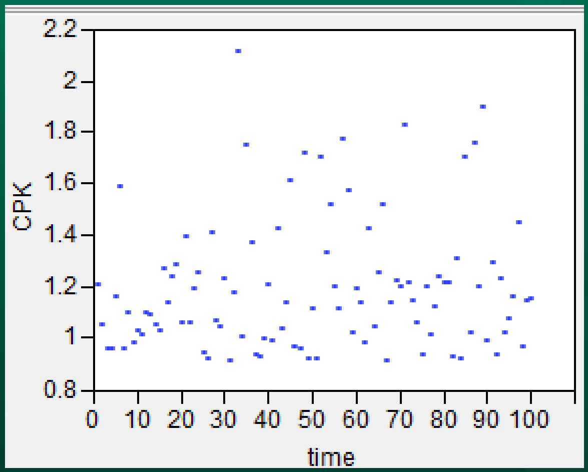
Box Plots and Normal Probability Plots

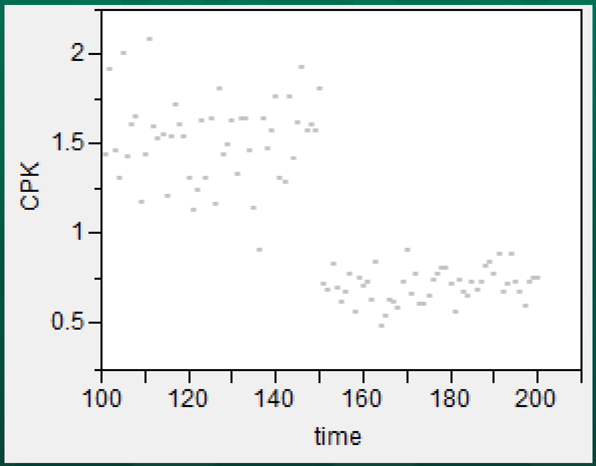


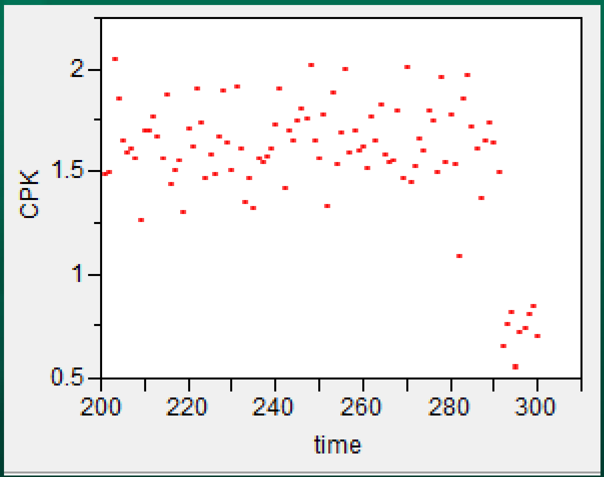


Probability plots can be effective at comparing multiple distributions. Probability plots can be 'busy' but effectively show the nuances of a distribution.

Time Plots







Relatively constant trend over time for A.

Clear shift seen with B. This information can be used to determine why the sift happened.

Timing of outliers can be seen with trend plots.

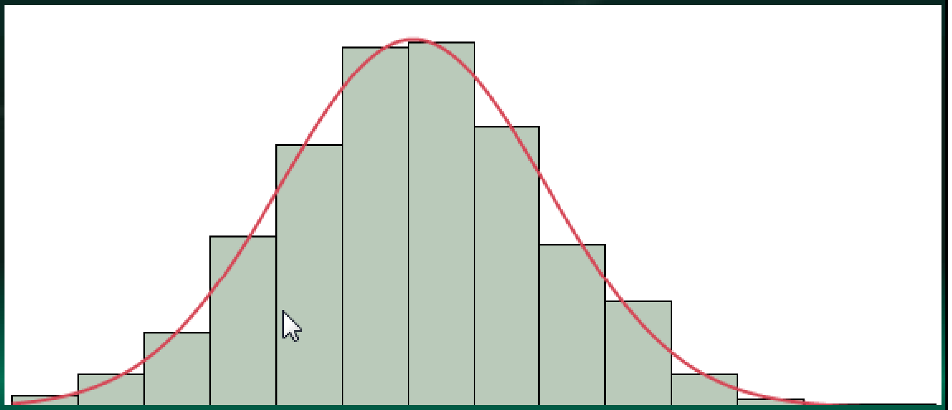
Graphics Wrapup

- Use histogram to get first idea of distribution shape, center, and spread.
- Prob plots and box plots can be used to compare multiple distributions at once.
- Once items of interest are identified within a distribution, look at trend plots to determine specific trends, shifts, or spikes in data.
- Always ask "What question are we asking the data?"

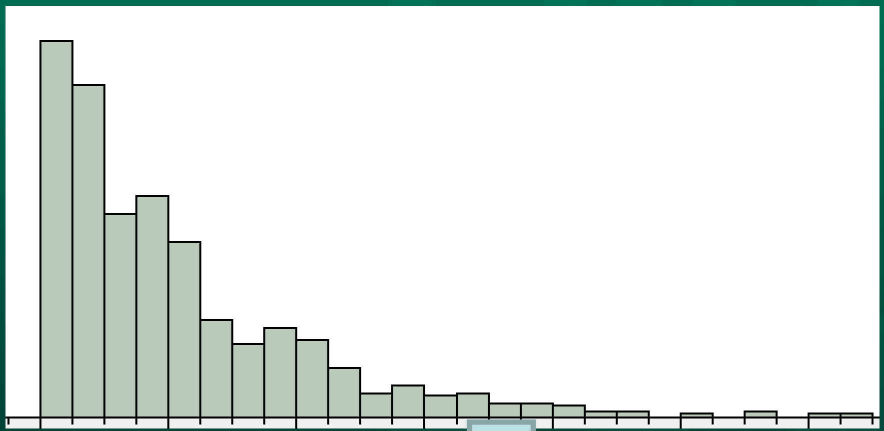
Normal Distribution

The Normal Distribution

- **Also known as the Bell Curve, or the Gaussian Distribution. Is critical to Statistical understanding for 3 reasons:**
  - Distribution occurs frequently in nature
  - Central Limit Theorem states that distribution of means of ANY distribution will be normal
  - Predictable percentiles.

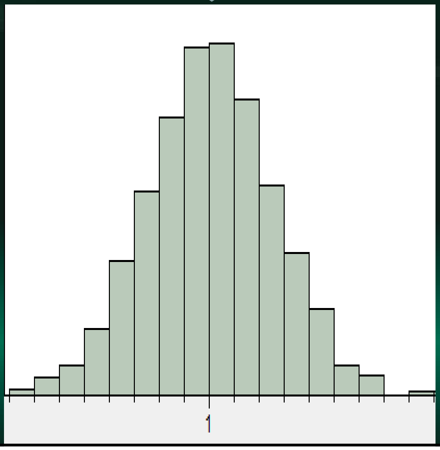


Normal Distribution: Central Limit Theorem



Distribution of unknown shape. Mean = 1.0, Standard deviation = 1.0

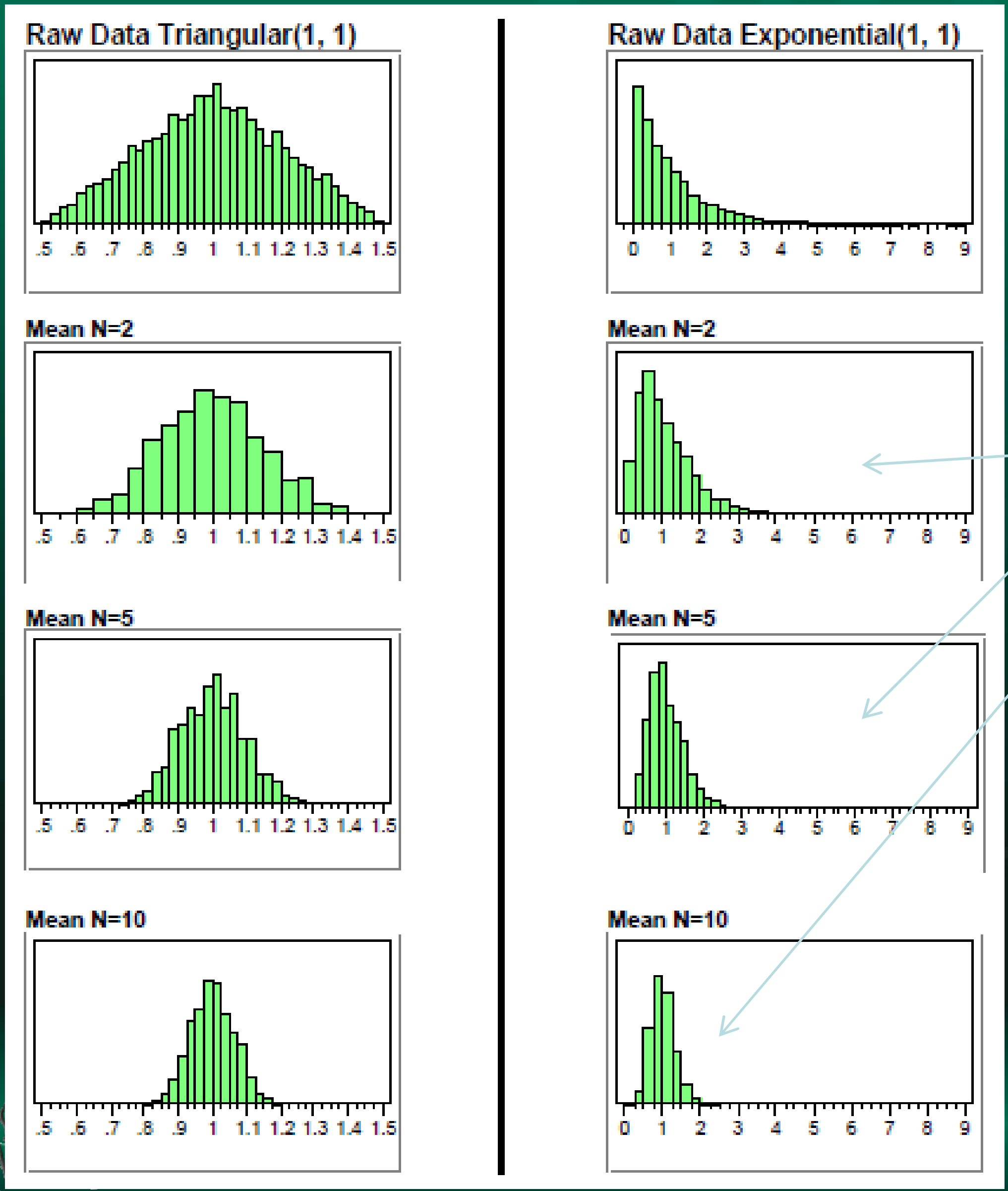
Randomly sample 16 data points at a time, calculate mean of each sample. Plot distribution of means.



Distribution is normal. Mean = 1.0, Standard error =

$$\frac{s}{\sqrt{n}} = \frac{1}{\sqrt{16}} = 0.25$$

Normal Distribution: CLT

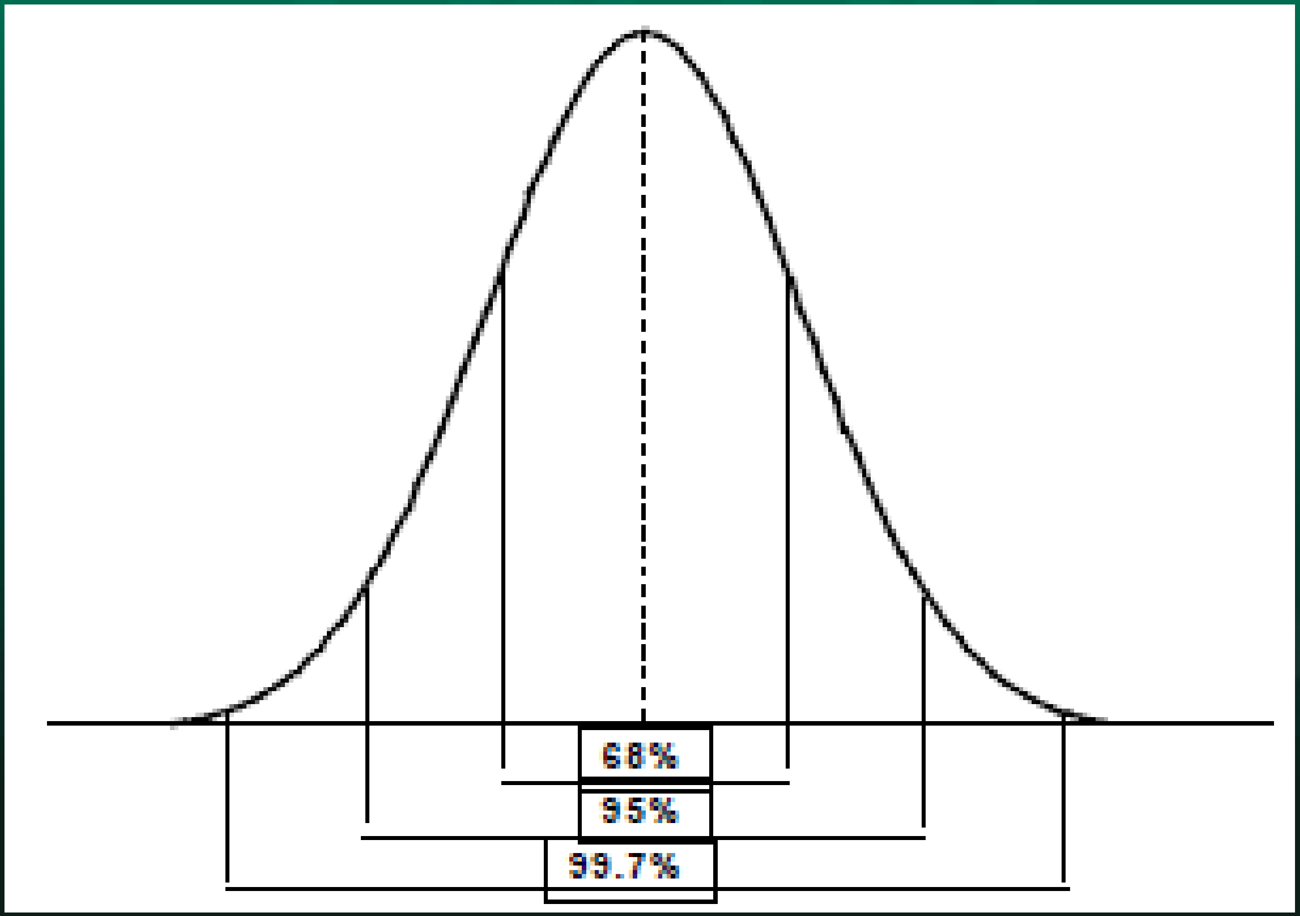


Averages of random sample of size N.

Normal Probabilities

- It is known that the region within +/- 1 standard deviation of the mean of a normal distribution will contain 68% of the data.
- The region within +/- 2 standard deviations of the mean will contain 95% of the data.
- The region within +/- 3 standard deviations of the mean will contain 99.7% of the data.
- Other percentiles also determined by Z table.

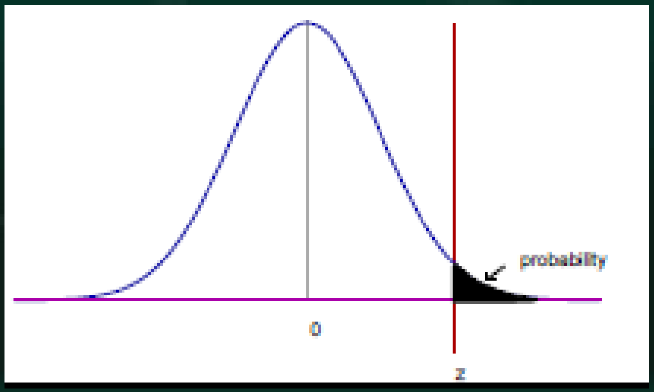
Normal Probabilities



Normal Probability: Z-score

- Percent of distribution from a single data point can be determined by normalizing the data point to the mean and standard deviation of the distribution.
- Uses the Z-score formula.

$$Z = \frac{x - \mu}{\sigma}$$



Normal Probability: Z-score example

- Measured values are normally distributed with a mean of 5400 and a standard deviation of 50. What is the probability of the next value chosen being 5487, assuming the process is problem free?

$$Z = \frac{x-\mu}{\sigma} = \frac{5487-5400}{50} = 1.74$$

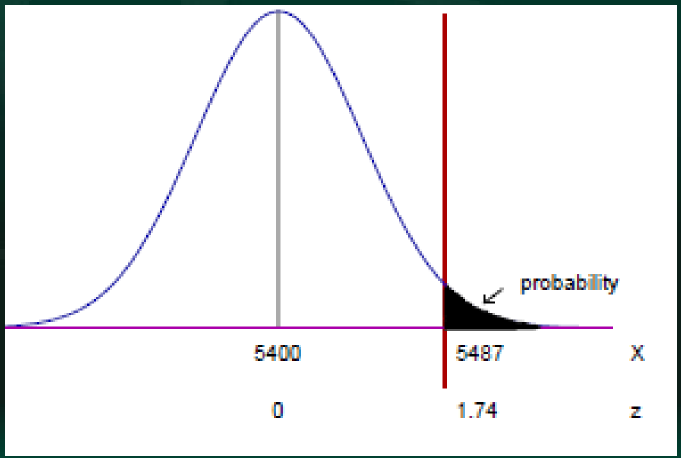
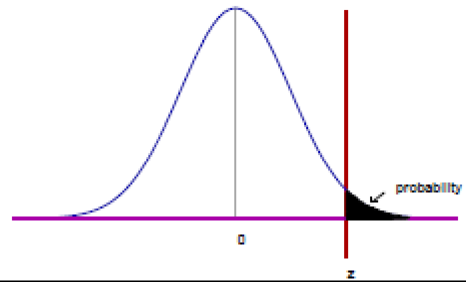


Table A: Standard Normal
Tail Area Probability

Generated using RS/1 v4.3.1 function $PROBNORM, and rounding to 4 decimal digits.



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
3.5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
3.6	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
3.7	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
3.8	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
3.9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

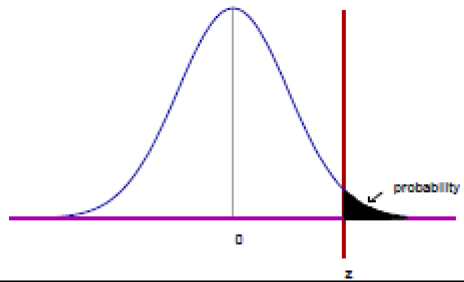
$$P(X > 5487) =$$

$$P(Z > 1.74) =$$

$$0.0409$$

Table A: Standard Normal Tail Area Probability

Generated using RS/1 v4.3.1 function $PROBNORM, and rounding to 4 decimal digits.



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	[illegible]	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
3.5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
3.6	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
3.7	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
3.8	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
3.9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Normal Distribution Mistakes

- Do not always assume that distributions are normal. Dangerous to apply normal probabilities and Z-scores to non-normal data.
- CLT makes the distributions of larger sample means normal, but not the original distribution.

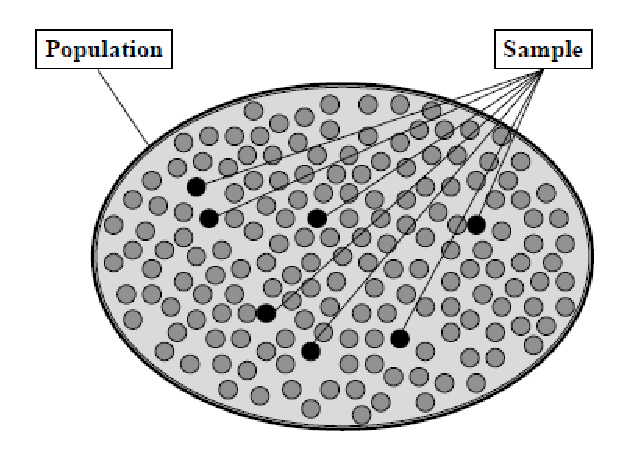
Normal Distribution Review

- Normal Distribution key to Statistical understanding
- Common in nature
- Central Limit Theorem states that distribution of sample means will be normal.
- Normal probability means that with only the overall mean, standard deviation, and a single point, one can determine the percentile of that single point.

Statistical Sampling

Population vs. Sample

- **Population: Consists of ALL measurements for a specific group.**
- **Sample: A representative subset of a population.**
- **Sampling is a more cost effective and efficient way of measuring an entire population, and *nearly* as accurate.**



Population vs. Sample

Population	Sample
The entire population of the United States of America (census)	A marketing study of 1000 Americans for a certain test product
Every single Canadian Goose	A sample of 200 Canadian Geese being tracked in a migration study
Planarity and orientation of every single probe within a probe card array.	Measurements of the planarity and orientation of a smaller subset of the probes used to estimate the overall probe card health

Good sampling

- Sampling can be a highly effective way of measuring an infinitely large population.
- Sample must be indicative of overall population.
- Sample should be sufficiently random.
- Sample size should be adequate.
- Avoid convenience sampling.

Good Sampling: Randomization

- If sample is chosen from a population in a way that each member of the population has an equal way of being chosen, it is random.
- Non-random samples contain bias.

Example of Bad Randomization

	A	B	C	D	E
1	A1 8	B1 19	C1 13	D1 7	E1 8
2	A2 11	B2 25	C2 32	D2 22	E2 6
3	A3 13	B3 27	C3 45	D3 21	E3 9
4	A4 8	B4 32	C4 24	D4 21	E4 4
5	A5 13	B5 10	C5 18	D5 16	E5 12

Sample Size

- **"33 1/3 of the mice used in the experiment were cured by the test drug; 33 1/3 of the test population were unaffected by the drug and remained in a moribund condition; the third mouse got away."**
  - Erwin Neter

Sample Size

- **Logically, the larger the sample size, the closer the results will be to the population.**
- **Solid estimates can be obtained with sample sizes of around 20 to 30.**
- **Sampling should be based on measurement level with highest level of variability.**
  - Ex: Variability likely larger from one probe card to another rather than within a single probe card. Better to sample more cards and fewer probes within card.

Sample Size

Detectable Difference	Minimal Sample Size
0.35 standard deviations	109
0.5 standard deviations	54
0.75 standard deviations	26
1.0 standard deviations	16
1.5 standard deviations	8
2.0 standard deviations	6
Gross Reality Check (GRC)	5 or less

Note: numbers are taken from a larger sample size table for a 1 sample hypothesis test assuming alpha and beta values of 0.05.

Parameters vs. Statistics

- Parameter is a number describing a characteristic of a population.
- Statistic is a number describing a characteristic of a sample.
- In short, study of Statistics is study of using sample data as estimate for unknown population parameter.

Calculation	Parameter	Statistic
Mean	μ	$\bar{x}$
Standard Deviation	σ	s

Sampling Wrapup

- Populations contain every possible set of the data in question, but is often very difficult or impossible to completely measure.
- Sampling is an effective way to measure part of a distribution which will represent the population.
- Some of the mistakes made in sampling include not randomizing, using an insufficient sample size, or pulling non-representative samples.

IEEE SW Test Workshop
Semiconductor Wafer Test Workshop

Measurement Capability

Metrology

- Wafer Test is a Metrology – the product of testing wafers is DATA.
- We don't actually change the Silicon, we just measure and report on it.

Measurement Statistics concepts

- **Bias: How close is measurement to 'real value'.**
- **Precision: Measure of the variability of measurements. Defined as how closely multiple measurements of same thing resemble each other.**

Accuracy: Bias

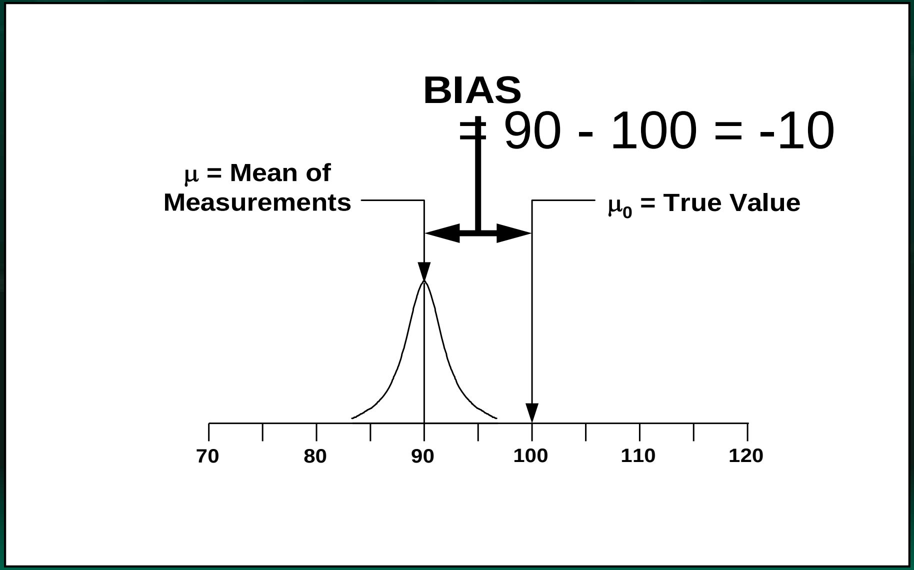
- ***Bias*** **is:**
  - The numerical value used to quantify accuracy.
  - The difference between the mean value of all the measurements (μ) and the true value (μ_0).
  - Note: Bias is sometimes called "offset."

$$\boxed{\text{Bias} = \mu - \mu_0}$$

Bias is not number of significant digits of a measurement. If something is measured to 15 decimal places it has a high level of resolution, not necessarily high accuracy.

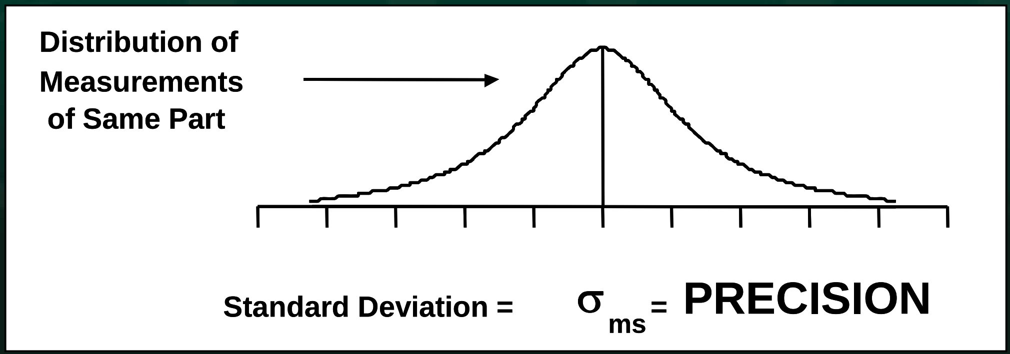
Accuracy: Bias

- **Bias quantifies the amount by which a metrology is consistently off target from the true value.**
- **True value must be based on known standard.**
  - Note: Bias can be positive or negative.



Precision

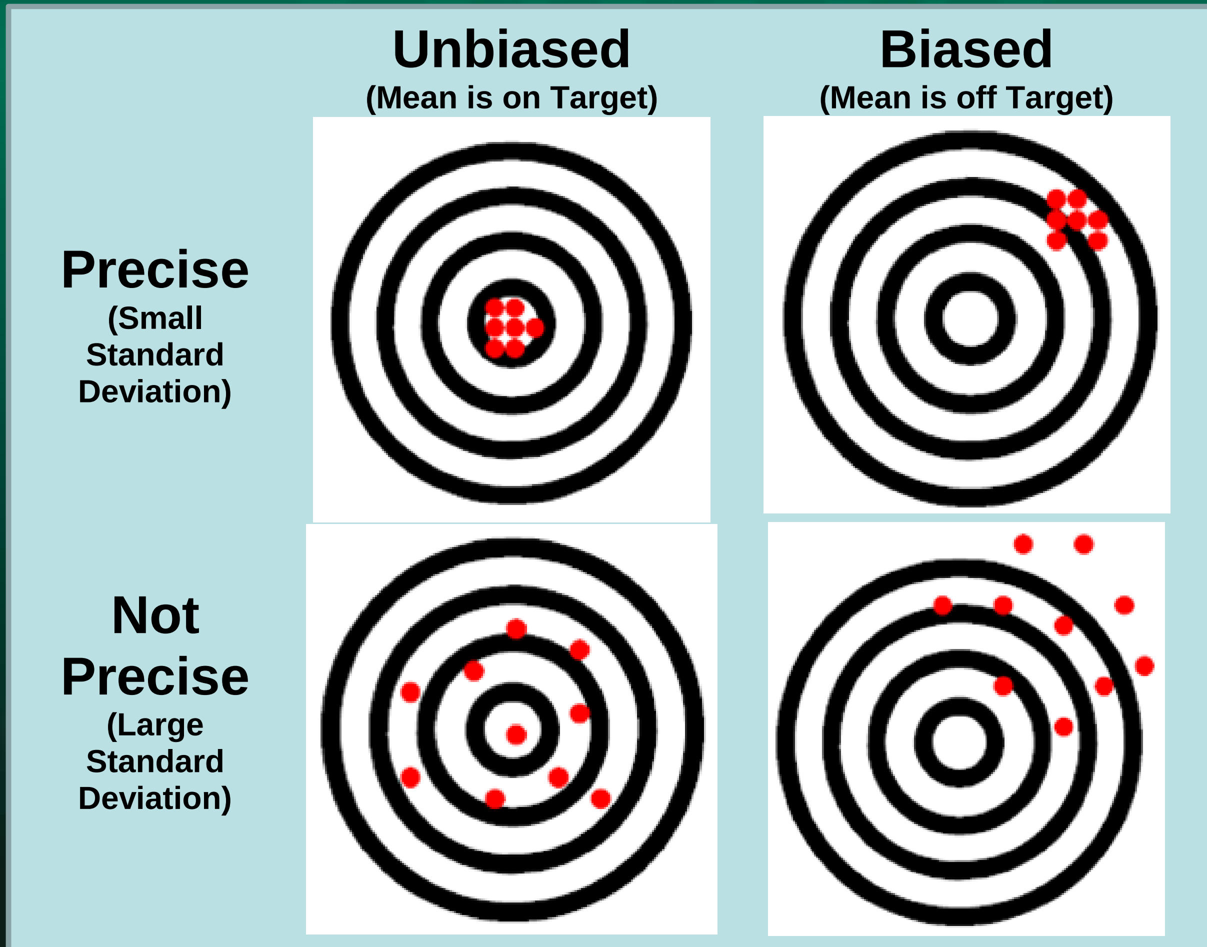
- **Precision measures the natural variation of repeated measurements.**
  - The total variation in the measurement system is quantified by σ_{ms} (standard deviation of the measurement system).
  - The ***smaller*** the standard deviation, the ***better*** the precision
- **Precision is also known as spread or noise.**



Accuracy vs. Precision

- **The difference between precision and accuracy:**
  - **Accuracy** describes how close the measurements are to the ***truth.***
  - **Precision** describes how close the measurements are to ***each other.***

Precision and Bias



Intel metrology almost always assesses precision and bias relative to a specified “reference tool,” rather than a “NIST reference standard”

Two Components of Precision

- **Precision can be separated into two components, called *repeatability* and *reproducibility* (abbreviated "R&R").**
- **The variance of the measurement system is the sum of the repeatability variance and the reproducibility variance.**

Repeatability

Repeatability is:

- **The variation that results when repeated measurements are made of a parameter under *identical* conditions:**
  - Same operator
  - Same set-up procedure
  - Same environmental conditions
  - During a very short period of time
- **The *"inherent"* variability of the measurement equipment and method.**
- **Measured by σ_{rpt}, the standard deviation of the distribution of repeated measurements.**

Dynamic vs. Static Repeatability

- **Static Repeatability: Measures the "inherent" variability in the measurement tool itself.**
  - Variation from repeated measurements in which the part is ***not removed*** from the tool between measurements.

- **Dynamic Repeatability: Measures the "inherent" variability of the tool *and* the measurement method.**
  - Variation from repeated measurements in which the part is ***removed*** and re-fixtured between measurements.

Examples: Static vs. Dynamic

- **Static Repeatability:**

Unit is Placed in Socket → Thermal Resistance is Measured 30 Times → Unit is Removed from Socket

- **Dynamic Repeatability:**
  - Repeat 30 times

Unit is Placed in Socket → Thermal Resistance is Measured Once → Unit is Removed from Socket

Dynamic or Static?

- **Dynamic repeatability will usually discover problems with the measurement tool *sooner rather than later.***
  - Can be used to check robustness of other elements of the measurement process including procedures and fixtures.

- **Static repeatability is always smaller than dynamic repeatability.**
  - If dynamic repeatability is too large, a small additional experiment can be done to investigate if the large variation is due to the static repeatability or to other elements.

Dynamic repeatability is always preferred.

Reproducibility

Reproducibility is:

- **The variation that results when *different* conditions are used to make the same measurement:**
  - Different operators
  - Different set-ups
  - Different positions
  - Different measurement media or fixtures
  - Different environmental conditions
  - Different times
- **Measured by σ_{rpd}, which is approximately the standard deviation of all measurements taken under different conditions.**

Two Components of Precision

- **The variance of the measurement system is the sum of the repeatability variance and the reproducibility variance.**

$$\sigma^2_{\text{Measurement System (ms)}} = \sigma^2_{\text{Repeatability (rpt)}} + \sigma^2_{\text{Reproducibility (rpd)}}$$

- **Recall that:**

$$\sigma^2_{(1+2)} = \sigma^2_1 + \sigma^2_2$$

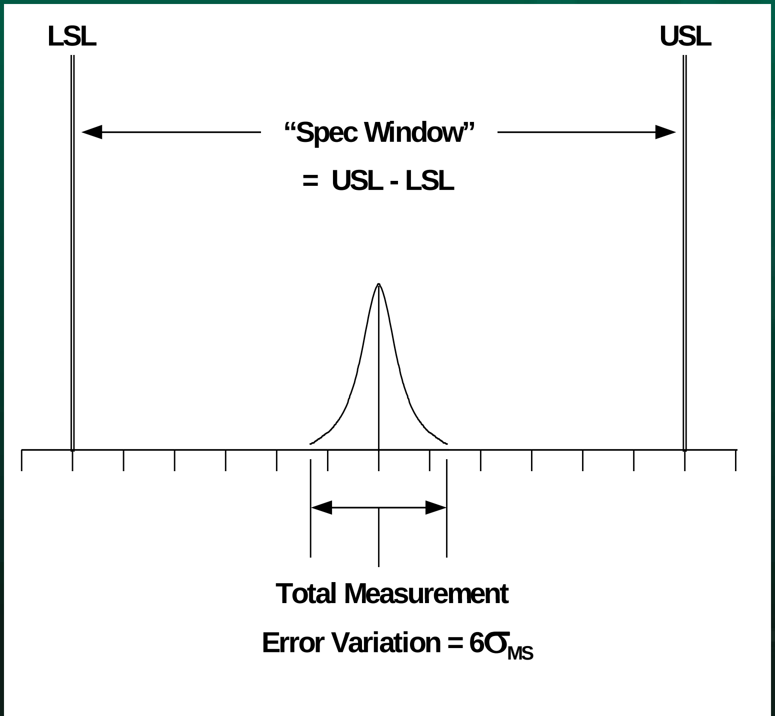
Capability

- The capability of a *manufacturing process* is its ability to meet specifications.

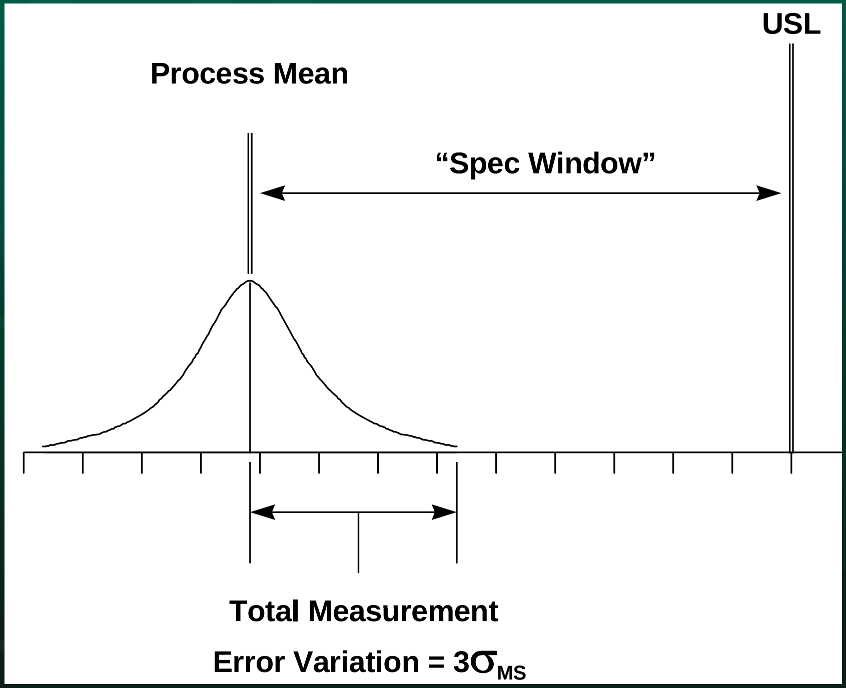
- In contrast, the capability of a *measurement system* is the amount of the spec window that is lost to measurement variation.

Measurement System Capability

Two-Sided Specs



Single-Sided Spec



P/T Ratio

- The capability of the measurement system is quantified by the *Precision-to-Tolerance (P/T) Ratio.*

- The P/T ratio expresses the percentage of the spec window that is lost to measurement error.

P/T Ratio Formula

- For *two-sided specs* (both USL and LSL):

$$P/T = \frac{6\sigma_{MS}}{USL - LSL} \times 100\%$$

- For *one-sided specs* (either USL or LSL):

$$P/T = \frac{3\sigma_{MS}}{TOL} \times 100\%$$

  – Where:
    - TOL = (Process Mean - LSL) for LSL only
    - TOL = (USL - Process Mean) for USL only

P/T Ratio Criteria

- **Smaller P/T values are desirable.**
  - P/T ≤ 30%: Measurement system is ***capable.***
  - P/T > 30%: Measurement system is ***not capable*** (not precise enough).

- **P/T ratios can also be computed for repeatability, by replacing σ_{ms} with σ_{rpt} (denoted as P/T_{rpt}).**

IEEE SW Test Workshop
Semiconductor Wafer Test Workshop

Confidence Interval

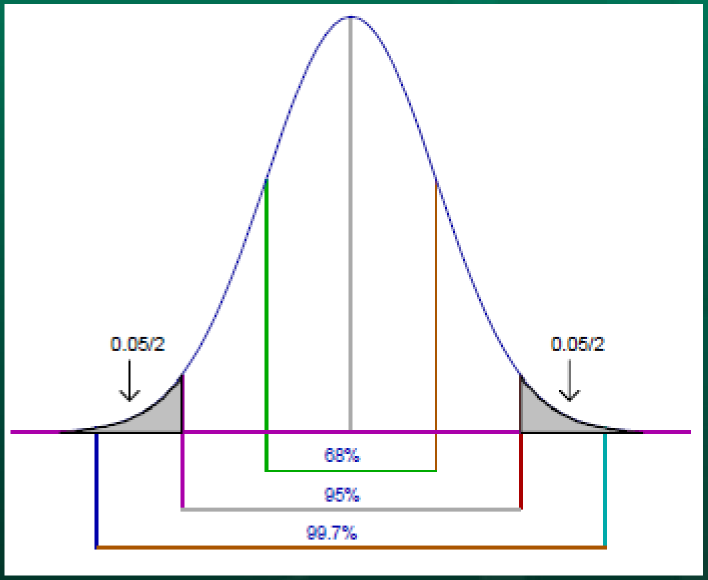
Confidence Interval

- On April 16, 2011 Pat’s Run (the Pat Tillman Memorial Run) was held in Tempe, AZ with over 22,000 participants.
- A random sample of 75 results show that the average finishing time was 48:30 (with a standard deviation of 460 seconds).
- Can we believe that the actual mean of all 22,000 finishers was EXACTLY 48:30?
- Can we believe that the actual value was close? How close?

Confidence Interval

$$\bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

95% Confidence Interval



$$\bar{x} \pm 2\frac{\sigma}{\sqrt{n}}$$

Race Time Confidence Interval Calculations

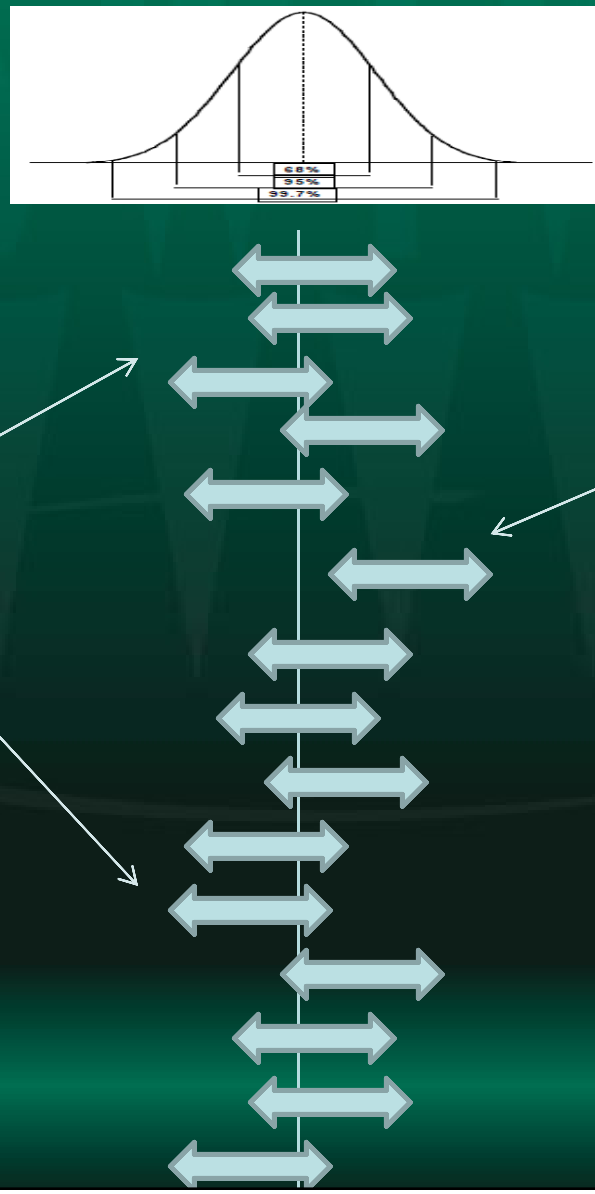
$$\bar{x} \pm 2\frac{\sigma}{\sqrt{n}} =$$

$$2910 \pm 2 \cdot \frac{460}{\sqrt{75}} =$$

$$\{2804, 3016\}$$

- X-bar=48:30 = 2910 seconds, s=460 seconds, n=75.
- Calculated 95% Confidence Interval between **46:44 and 50:16.**
- Recall that with sampling, we using partial data to estimate a true mean value which we cannot know at this time.
- 95% Confidence Interval means that **"We can be 95% certain that the true mean value lies within the interval".**

Confidence Interval



Most of the time, a sample will create a confidence interval which contains the true value. (95% of the time for a 95% CI)

1 in 20 samples will by random chance create a CI which will not contain the true mean.

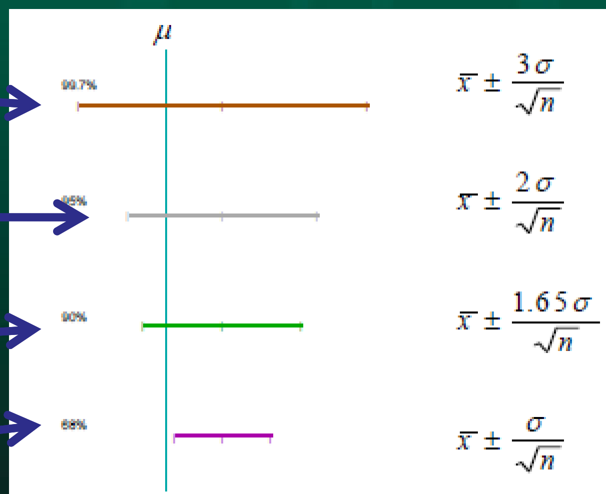
Confidence Intervals

99.7%

95%

90%

68%



Adjusting the confidence level leads to changing size interval. Tradeoff of more confidence against less useful information.

Confidence Interval

Lower process variation leads to tighter confidence

$$\bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Increasing confidence level leads to lower probability of being wrong.

Higher sample size leads to tighter confidence

Confidence Interval: T Distribution

- Use of the Z score (standard normal table) requires sufficiently large sample size and a known process standard deviation (sigma).
- Since these assumptions are almost always not met, best to use the T-distribution.
- T-distribution utilizes information on sample size, and will slightly 'penalize' the confidence interval size for smaller samples.

Confidence Interval based on T

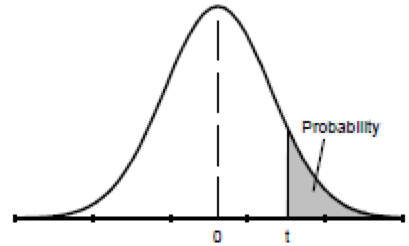
Sample standard deviation

$$\bar{x} \pm t_{\alpha/2,df} \cdot \frac{s}{\sqrt{n}}$$

$$df = n - 1$$

Table B: t Distribution
One Tail Area Probability

From Box, Hunter, and Hunter, *Statistics for Experimenters*, p 631.



df	Tail Area Probability									
ν	0.4	0.25	0.1	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	14.089	22.326	31.598
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	7.453	10.213	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	0.258	0.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	0.257	0.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	0.257	0.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	0.257	0.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	0.257	0.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	0.256	0.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	0.256	0.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	0.256	0.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	0.256	0.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	0.256	0.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	0.256	0.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	0.256	0.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	0.256	0.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	0.255	0.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	0.254	0.679	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
120	0.254	0.677	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	0.253	0.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

Confidence Interval Wrapup

- Confidence intervals can be used to determine how closely a statistic estimate is to the population parameter.
- With a 95% CI, we are 95% confident that the true mean lies somewhere within the interval.
- Best to use the T distribution since actual sigma is often not known.

Acknowledgements

- **Bulk of material adapted from the following Intel Statistical training material:**
  - Stat Basics Rev 3.2
  - Survival Statistics Rev 4
  - Measurement Capability Analysis Rev 5.1.2
- **Thanks for some data and guidance from Tom S Leahy, Intel STTD, Hillsboro, OR.**