APPENDIX A - TESTS OF SIGNIFICANCE / HYPOTHESIS TESTING

Description

Tests of significance are used to determine if relative differences exist between sample statistics (or population parameters) and between different samples (or populations). A statistical Hypothesis is a statement about the probability distribution of a random variable. The hypotheses are always statements about the population or distribution under study (including parameters), NOT statements about the sample.

Purpose

A statistical test provides a mechanism for making quantitative decisions about a process or product. The intent is to determine whether there is enough evidence to "reject" a conjecture or hypothesis about the process or product. The conjecture is called the null hypothesis.

Types of tests:

F Test: Test for homogeneity of variances.
Chi-Square Test: Test for equality of standard deviations or variances from a Normal distribution.
t or Z Test: Test for equality of means.
Goodness of Fit: Test to assess distribution (probability model) use.

Benefits

Using a data-driven and logical approach assists in the following:

- Not making an incorrect decision about a comparison that is being analyzed and the cost penalties associated with it.
- Making a correct decision about a comparison that is being analyzed.

Implementation

Required: Formal coursework in Probability and Statistics (with estimation and confidence interval theory, Tests of Significance).
Assumption: That the random variables to be compared are independent.

The value of Test of Significance output is only as good as the effort and amount of data put into the process (i.e., a good assessment of variability is required, which generally means a reasonably adequate sample size).

Process Flow

1. Determine the comparison to be made :
   - Population means/Sample means.
   - Population standard deviations/Sample standard deviations.
   - Population variances/Sample variances.
   - Population proportions/Sample proportions, etc.
2. Determine the comparison “type” to be made:
   - Sample to sample.
   - Population to population.
3. Determine the “scope” of the comparison to be made :
   - One sided (tailed) and direction.
   - Two sided (tailed).
4. Define Null and Alternative Hypothesis. Decide on the level of alpha (Type I error) and beta (Type II error)that can be tolerated. Therefore, assess the “power” of thetest.
5. Determine the sample size necessary to meet the alpha and beta requirements (if required).
   - Obtain data (equal / unequal sample sizes) from random samples.
   - Estimate statistics / obtain parameters for comparison.
6. Determine “acceptance” and “rejection” criteria. The “cutoff “ points/-regions from thesamples and associated tables.
7. Compare and determine acceptance or rejection of the hypothesis.

Example Format

Let H_0 be the NULL hypothesis an H_1 or H_A the ALTERNATIVE hypothesis , then

TYPE I ERROR : $\alpha = P[\text{Reject } H_0 \mid H_0 \text{ is True}]$

TYPE II ERROR : $\beta = P[\text{Accept } H_0 \mid H_A \text{ Is True}]$

SO	$\alpha =$	P(Type I error)	"Significance"
	$\beta =$	P(Type II error)	
	$1.0 - \alpha =$	P(Avoiding a Type I error)	"Confidence"
	$1.0 - \beta =$	P(Avoiding a Type II error)	"Power"

DECISION

TRUTH	Accept H_0	Reject H_0
H_0 is True	$(1.0 - \alpha)$	α
H_A is True	β	$(1.0 - \beta)$

There are many "types" of test :

One sided Lower / Upper

$H_0 : A \prec B$ or $H_0 : A \succ B$

$H_A : A \geq B$ or $H_A : A \leq B$

Two sided Lower and Upper

$H_0 : A \equiv B$

$H_A : A \leq B$ or $H_A : A \geq B$

where A , B are parameters of interest.

Example:

Manufacturing has been having an issue with scrap generated as a result of a given machined feature that is crucial to the performance of finished product in the field. As a result, they implemented a control that supposedly stabilizes the surface of the product prior to machining. They took measurements before and after the control and wanted to determine if there was a significant difference and to what degree improvement took place. The specification for the machined surface is 0.00 to 0.04. The following data set was generated for the analysis:

Table A.1. Data Set for Scrap Example

1 Without control	2 With control	3 Reduction
0.02802	0.01407	0.01395
0.03024	0.01925	0.01099
0.05109	0.02138	0.02971
0.03964	0.02214	0.01750
0.03747	0.00690	0.03057
0.04211	0.01660	0.02551
0.05570	0.02353	0.03217
0.01909	0.02063	-0.00154
0.02639	0.01714	0.00925
0.03470	0.01154	0.02316
0.03528	0.01726	0.01802
0.02997	0.01566	0.01431
0.03008	0.01284	0.01724
0.02174	0.01260	0.00914
0.02814	0.01222	0.01592
0.03608	0.01835	0.01773
0.06406	0.01582	0.04824
0.02474	0.01679	0.00795
0.02495	0.01705	0.00790
0.03373	0.01415	0.01958
0.02669	0.01070	0.01599
0.04598	0.02101	0.02497
0.02877	0.01277	0.01600
0.01904	0.01463	0.00441
0.01949	0.01220	0.00729
0.03640	0.02137	0.01503
0.03727	0.01212	0.02515
0.03197	0.01297	0.01900
0.03228	0.03096	0.00132

H0 : Significant difference exists between the sample with control vs. sample without control at the alpha = .05 level.Ha : Not so.

The following results were obtained:

Table A.2. Results of Testing

Variable	T-test for Dependent Samples Marked differences are significant at p < .05000 Mean	Std.Dv.	N	Diff.	Std.Dv. Diff.	t	df	p
Without control	0.033487	0.010642						
With control	0.016367	0.004872	29	0.017119	0.010339	8.917053	28	0.000000

Conclusion: Significant differences exist between the two samples. To get an assessment of these marked differences, reference the following box plot (Figure 1):

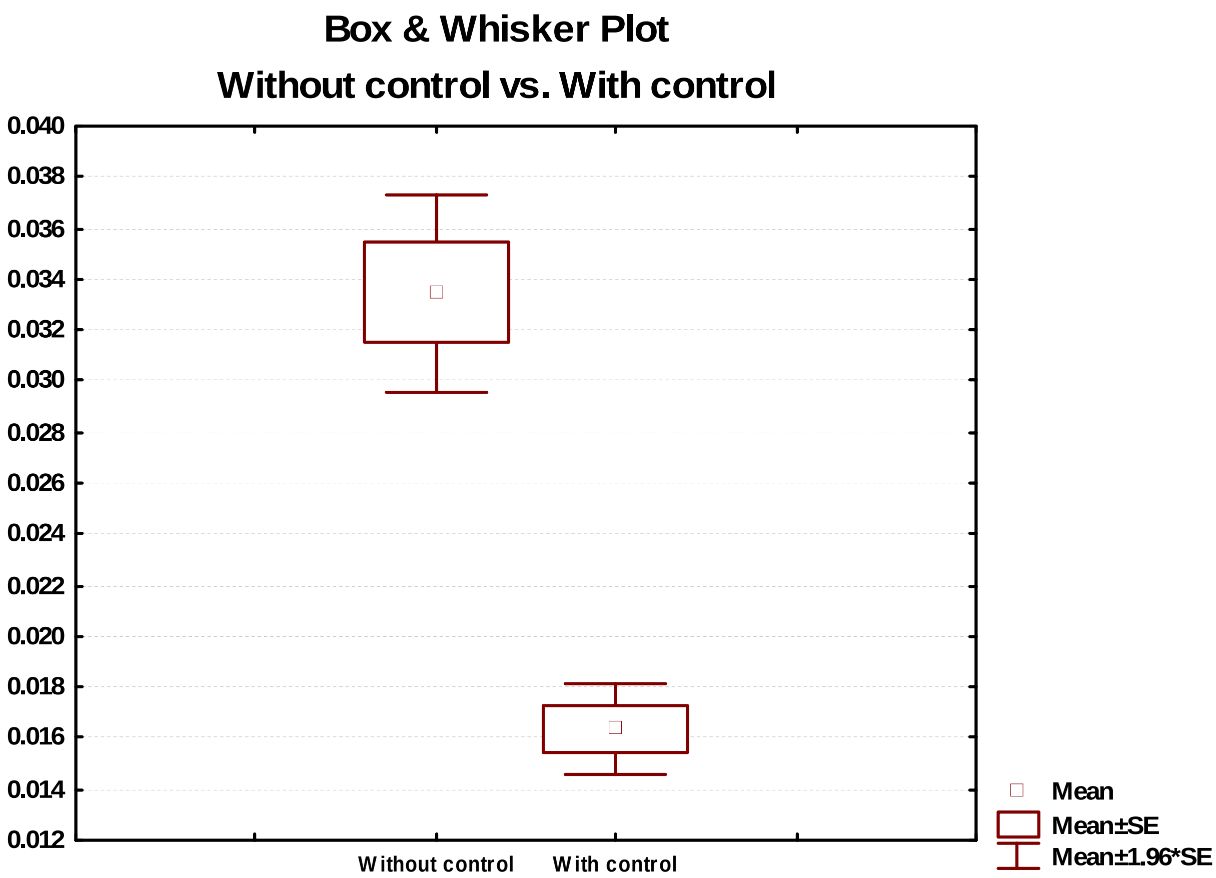

Figure A.1. Box and Whisker Plot

The following chart captures the differences (reduction in machined surface dimension) with the control:

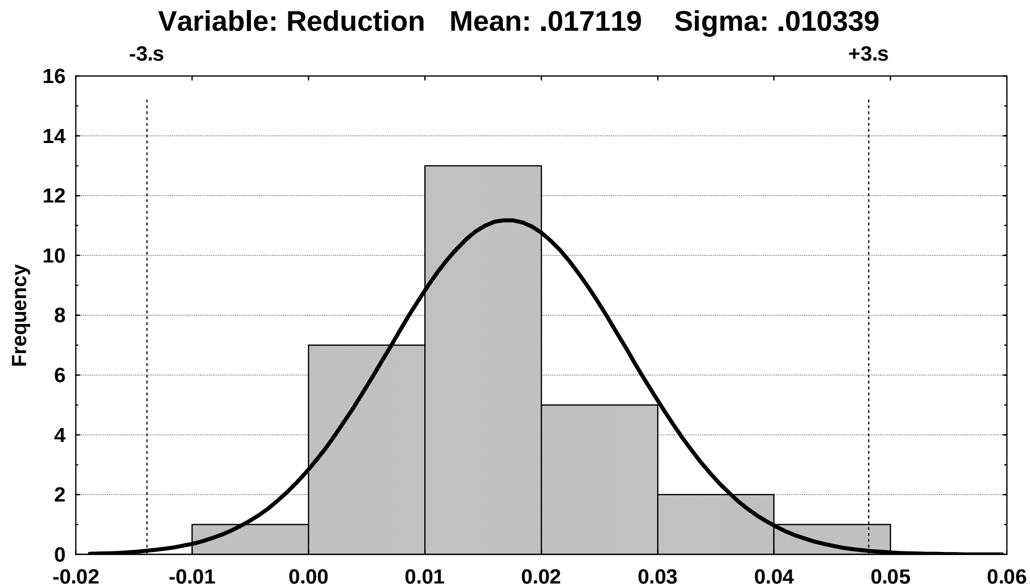

Figure A.2. Histogram of Machined Differences

What is the capability now? Analyzing the data with the control we find that we are still non-capable to the specification (see Figure 3 below). The next step in the improvement process should entail forming a team and brainstorming potential influencing parameters for a designed experiment (DOE).

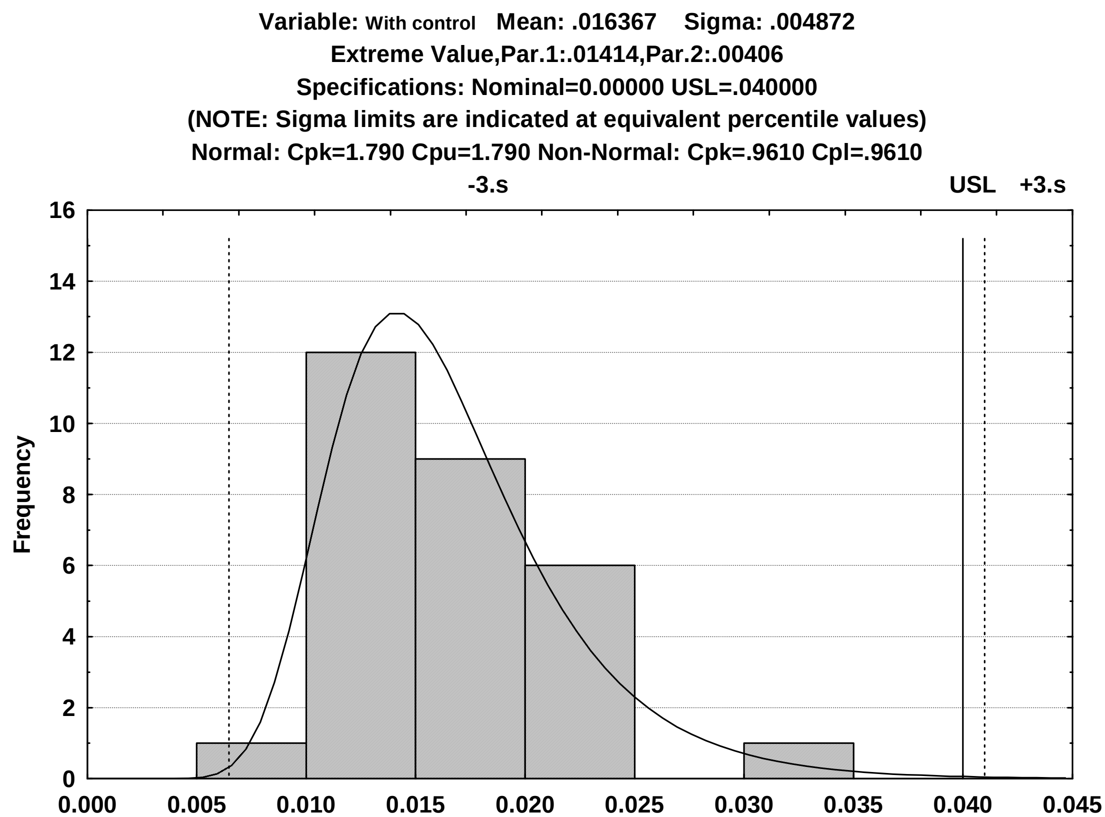

Figure A.3. Distribution of measurements with control mechanism

General Comments

Tests of significance play an important role in analysis involving comparisons. They are an essential step in the continuous improvement process.

References

Bain and Engelhardt. *Introduction to Probability and Mathematical Statistics,* PWS-Kent, 1989.

Lipson and Sheth. *Statistical Design and Analysis of Engineering Experiments,* McGraw Hill, 1983 .