BASIC TOOLS

Measurement Systems Analysis — Revisited

GREGORY F. GRUSKA (REGION 10 COUNCILOR)
ORCHARD LAKE, MI

Discussions of the analysis of measurement systems and the effect of their variability on decisions is coming out of the metrology and statistics "backrooms" and becoming part of industrial and international standards. While everyone would agree the knowledge of the variation of one's measurement system is useful, how the analysis of that variation should be performed is becoming the subject of controversy.

This paper provides a collection of methods to study the stability and variability of a measurement system. Because the mechanics and computations associated with these methods are beyond the scope of this paper, interested readers are encouraged to consult the technical literature, including the references following this article, for detailed step by step "how to" guidance.

Too often individuals apply analysis techniques by rote; without understanding how their measurement system should perform. Before jumping into the statistical analysis of any process, we must first understand the purpose (aim) of the process and the sources of variation which could affect it.

Identification of Purpose
Activities requiring measurement systems fall into two general categories:
Category One: Analysis of the process distribution

This requires variable measurement. The historical "rule of thumb" guideline is that the discrimination be < 10% of tolerance. A common industrial requirement is based on the measurement system's variation as measured by the repeatability (equipment effect) and reproducibility (appraiser effect) also known as Gauge R&R, or GRR. This variation is evaluated by:

GRR < 10% of the expected variation is acceptable
The process distribution can be divided into a sufficient number of cells for purposes of analysis.

GRR > 30% of the expected variation is unacceptable
The process distribution will not be able to be divided into at least five cells for purposes of analysis.

If the GRR falls between 10% and 30%, subject matter expertise and process knowledge are necessary to determine if the measurement system will be adequate to analyze the process distribution.

Category Two: Control of the process variability
This requires the ability to identify that a change has occurred. This allows measurement to be:

variable
semi-variable (multiple categories), or
attribute (two categories)

Note that process control requires the limits of attribute and semi-variable measurement systems to be based on the process distribution, not on the specification. The specific amount of measurement variation to be allowed should be based on the total system variability and on the effect of a wrong decision.

Sources of Variation
The process to develop a study to analyze the variation of a measurement system should be the same as that used for any other system - the PDSA cycle (Plan - Do -Study - Act). Even though there are published procedures on measurement systems studies, the process team must always identify the potential sources of variation and determine if the standard study will be sufficient to encompass or quantify them.

Another concern in the design and selection of a measurement study deals with the fact that some measurement systems are "destructive". Measurement systems studies can be classified into studies in which:

factor levels can be replicated
a) Each cell within the experimental design can be assigned randomly and replicated.
b) Each cell can be replicated but not assigned randomly; i.e., there are physical or economic constraints which call for the use of split-plot or hierarchal designs.

factor levels cannot be replicated (i.e. the property changes)
(The part or property being measured is altered or destroyed by the measurement.)
a) the part (specimen) is destroyed (e.g., destructive testing, metallurgical (lab) testing, hardness testing, etc.)
the results can be re-viewed (i.e., can be remeasured; e.g., nugget size)
the results are ephemeral (e.g., spectral analyzer results)
b) the part changes gradually (e.g., powertrain testing, hot forgings, etc.)

The following is an annotated listing of measurement systems studies for stability and variability that encompass the commonality of most basic approaches.

Since the objective of measurement systems analysis is to quantify the common cause variation of the system, initial studies should verify the stability of the system. Consequently, stability studies (S) will be described first, followed by variability studies (V). The tags given to the various studies are arbitrary and do not indicate relative importance or suggested order of use.

CONTINUED ON PAGE 13

I Measurement Systems - Stability Studies (S)

S1 Single part (specimen), single measurement per cycle

a) A reference standard can be used if it is representative of the process.
b) The life of the characteristic (property) is known and extends beyond the expected duration of the study.
c) The measurement system is known (documented) to have a linear response over the expected range of the characteristic (property).

Analyze using:

individuals and moving range [x, mR] charts:

determine measurement system stability
1. compare the plotted points to the control limits
2. look for trends (x chart only)

compare $\sigma_e = \frac{\overline{R}}{d_2}$ with the repeatability estimate σ_r from

a variability study (see next section)
determine the bias if reference value is known: bias = reference value - $\overline{x}$.

S2 n ≥3 parts (specimens), single measurement per cycle per part (specimen)

a) Reference standards can be used if they are representative of the process.
b) The life of the characteristic (property) is known and extends beyond the expected duration of the study;
c) Measured parts (specimens) cover the expected range of the process variation of the characteristic.

Analyze using:

a [z, R] chart where $z_i = x_i - \mu_i$, and μ_i is the (reference) standard value, or determined by the average of a large number of successive readings of the part (specimen).

determine measurement system stability
1. compare the plotted points to the control limits
2. look for trends (x chart only)

compare $\sigma e = \frac{\overline{R}}{d_2}$ with the repeatability estimate σ_r
from a variability study
determine the bias if reference values are known: bias = reference value - $\overline{x}$.

determine the linearity if n ≥ 3 parts were used
the parts (specimens) must cover the expected range of the property
each part (specimen) should be analyzed separately for bias and repeatability.
quantify the linearity
1. plot the reference vs. measured values on a scatter diagram and compare to the 45 degree line.
2. determine consistency (homogeneity of variation) among the specimens; e.g. use max min F test, Levine's test, etc.
3. determine if all reference values have the same bias; i.e., the biases come from the same distribution.

(S1 & S2 can be used for measurement systems that are nondestructive and will be used with parts (specimens) with
(1) static properties, or
(2) dynamic (changing) properties which have been stabilized.

S3 Large sample from a stable process

The measurement system must be evaluating a homogeneous - independent, identically distributed (i.i.d.) sample (collected and isolated). The measurement of individual parts (specimens) are not replicated.

a) The life of the characteristic (property) is known and extends beyond the expected duration of the study.
b) Parts (specimens) cover the expected range of the process variation of the characteristic (property).
c) The measurement system's linearity is known (documented) over the expected range of the characteristic (property). (If the response is non-linear, the readings must be adjusted accordingly.)

Analyze by:

determining the total variability via a capability study with n ≥ 0 parts. (This preliminary study should also be used to verify the consistency of the sample: i.e. all parts (specimens) come from a unimodal distribution.)

$$\sigma^2_{total} = \sigma^2_{process} + \sigma^2_{ms}$$

measuring one or more individuals from the isolated sample per time period, use traditional [$\overline{x}$, R] or [x, mR] charts with the control limits determined by the capability study
1. compare the plotted points to the control limits
2. look for trends

Since the parts (specimens) do not change (an isolated sample), any indication of instability would be attributed to a change in the measurement system.

S4 Split specimens (general), single specimen per cycle (see also Wheeler)

a) The shelf life of the characteristic (property) is known and extends beyond the expected duration of the study.
b) Parts (specimens) cover the expected range of the process variation of the characteristic (property).
c) Specimens are split into m portions. With m = 2 portions, this is often called a test-retest study.

CONTINUED ON PAGE 14

Analyze using:
a range chart to track the consistency of the measurements (confounded with the "within lot" consistency)

compare* $\sigma_e = \frac{\bar{R}}{d_2}$ with the repeatability estimate σ_r from a variability study

This is an upper bound study: $\sigma_e^2 = \sigma_r^2 + \sigma_{btwn}^2$

[$\bar{x}$, mR] chart to track the consistency of the production process

S4a; S4 with pairs of consecutive (homogeneous) parts from different lots

This is an upper bound study: $\sigma_e^2 = \sigma_r^2 + \sigma_{btwn}^2 + \sigma_{lots}^2$

S5 Test Stands
In this situation, multiple measurement instruments (test stands) evaluate the same characteristics of a continual stream of product. The product stream is randomly assigned to individual stands.

S5a Attribute responses
Analyze using:

p charts

to determine consistency (of decisions) among the stands: a single chart including the results from all the stands
determine stability within individual stands: separate chart for each stand

Analyze the total system stability with a $[\bar{p}, \text{mR}]$ chart where $\bar{p}$ is the average p over all the test stands in a given day.

S5b Variable data responses

Analyze using ANOVA and graphical techniques *(see also James)*

Note: if more than one appraiser is involved in the data collection, then σ_e is affected also by the reproducibility of the measurement system. Quantify reproducibility by scatter and whisker plots indexed by appraiser.

calculate $\bar{x}$, s for each test stand (by characteristic), by time period

determine consistency among the stands: a single $[\bar{x}, s]$ chart including the results from all the stands
determine stability within individual stands: separate $[\bar{x}, s]$ chart for each stand

quantify the consistency (homogeneity of variation) among the stands; e.g., use max min F test, Levine's test, etc.

determine if all stands are on the same target by comparing stand averages; e.g., using a one-way ANOVA analysis. If any difference exists, isolate "different" stands by using, for example, Tukey's T-test.

II Measurement Systems - Variability Studies (V)

Note: All descriptive studies are *enumerative* in nature in that they describe the measurement system (including the effects of the environment) during the study. Since measurement systems are to be used in making future decisions about products, processes, or services, an *analytic* conclusion about the measurement system is necessary. The transition from enumerative to analytic results require subject matter knowledge and expertise to:
1. assure that all expected measurement sources of variation are considered in the design and execution of the study, and
2. analyze the results (data) in light of the expected uses, environment, control, maintenance, etc.

V1 Standard GRR techniques (see also ASQC)

These gauge repeatability and reproducibility studies include graphical analysis as well as numerical analysis. *(see also James)*

V1a Range method (R&R)

V1b Range method (R&R and within part)

V1c ANOVA method

V1d modified ANOVA/Range method *(see also Wheeler)*

V2 Multiple readings with p ≥ 2 instruments

Analyze using:

Grubb's (or Thompson's) estimates *(see also Grubbs, Thompson)*

process variability
instrument variability = repeatability
confidence region calculations are available

V3 Split specimens (m = 2)

a) The sample must cover the expected process variability.
b) Split each specimen into m = 2 portions (i.e., test-retest portions).

Analyze using:

regression techniques
estimate repeatability with the error term: $\sigma_r = \sigma_e$
linearity (by comparing estimated line with 45° line

V3a; V3 with pairs of consecutive parts

CONTINUED ON PAGE 15

This is an upper bound study: $\sigma_r^2 \leq \sigma_e^2 + \sigma_{btwn}^2$

V4 Split specimens (general)

a) (same as V3-a)
b) split specimens into m portions where m = 0 mod 2 or 3; m ≥ 2 (e.g., m = 3,4,6,9, ...)

Analyze using:

standard GRR techniques including graphics
ANOVA-Randomized Block Design (two-way ANOVA)

V4a, V4 with pairs of consecutive (homogeneous) parts from different lots

$$\sigma_r^2 \leq \sigma_e^2 + \sigma_{parts}^2 + \sigma_{lots}^2$$

This is an upper bound study

Note: V3 and V4 can be also used to analyze measurement systems with dynamic characteristics

Assuming the part (specimen) characteristic (property) is dynamic

V5; V1 with stabilized parts

E.g., engines which are 'broken-in' versus 'green' engines

V6 Time series analysis

a) repeated readings over specified time intervals

Analyze by: determining the degradation model for each sample part

$\sigma_r = \sigma_e$
consistency of degradation (if n ≥ 2)

V7 Linear analysis

a) repeated readings over specified time intervals
b) degradation in the measurement system is known (documented) to have a linear response over the specified time intervals.

Analyze by: linear regression

$\sigma_r = \sigma_e$
consistency of degradation (if n ≥ 2)

V7a; V7 with a homogeneous sample

a,b (as in V7)

Analyze by: linear regression

$$\sigma_r^2 \leq \sigma_e^2 + \sigma_{btwn}^2$$

This is an upper bound study

V8 Time versus characteristic (property) degradation

V6 & V7 can be modified to determine if the degradation is time (i.e., shelf life) or activity dependent.

References

ASQC/*AIAG Measurement Systems Analysis Manual*, AIAG, Detroit, MI, 1990.

Grubbs, F. E., On Estimating Precision of Measuring Instruments and Product Variability, *JASA, 43* , 1948.

Gruska, G. F., and Heaphy, M. S., *Measurement Systems Analysis*, tbp

James, P. D., Graphical Displays of Gage R&R Data, *AQC Transactions*, ASQC, 1991.

Heaphy, M. S., et. al., Measurement System Parameters, *SME - IQ81-154*, 1981

Lipson, C. and Sheth, N. J., *Engineering Experiment*, McGraw-Hill, NY, 1973.

Thompson, W. A., Precision of Simultaneous Measurement Procedures, *JASA 58*, 1963.

Wheeler, D. J., Evaluating the Measurement Process When the Testing is Destructive, TAPPI Conference, 1991.