Process Capability Analysis Using Experiments

A designed experiment can aid in separating sources of variability in a quality characteristic.

Example: bottling soft drinks

Suppose the measured syrup content $B$ of a soft drink satisfies $B = \mu + M + H + A$, where:

- $\mu$ is the mean content;
- $M$ is a random effect associated with a particular machine;
- $H$ is a random effect associated with a particular filling head on a given machine;
- $A$ is a sample-to-sample random effect.
Then

$$\sigma_B^2 = \sigma_M^2 + \sigma_H^2 + \sigma_A^2.$$ 

If the overall variability $\sigma_B$ is unacceptably high, it could be reduced by focusing on any of the three components.

The components can be estimated separately from the Analysis of Variance in a factorial designed experiment.

Specific improvements could be made to reduce any large contributor.
Gauge and Measurement System Capability Studies

Basic concepts
In order to monitor, analyze, or control a quality characteristic, we must be able to measure it.

The generic measurement tool is a gauge (e.g. a tire pressure gauge, or a wire diameter gauge).

Studies of measurement systems often refer to gauge capability.
Gauge R & R

Repeatability: Do we get the same observed value if we measure the same unit several times under identical conditions?

Reproducibility: How much difference in observed values do we experience when units are measured under different conditions?
Example: Using control chart tools

Twenty parts are each measured twice ($m = 20, n = 2$). Is the gauge precise enough to distinguish part-to-part variability from within-part (measurement-to-measurement) variability?

In R:

gauge <- read.csv("Data/Table-08-06.csv")
library(qcc)
gaugeG <- with(gauge, qcc.groups(Measurement, Part))
summary(qcc(gaugeG, "R"))
summary(qcc(gaugeG, "xbar"))
The $R$ chart shows within-part variability; it shows stable variation and is in control, with standard deviation $\hat{\sigma}_{\text{Gauge}} = 0.887$.

The $\bar{x}$ chart has a different interpretation from the control context; the control limits are based on $\hat{\sigma}_{\text{Gauge}}$, and indicate how much variation in $\bar{x}$ can be attributed to measurement error.

The several points outside the control limits illustrate the **discriminating power** of the gauge: the ability to separate the part-to-part variability from measurement error.
Note

The control chart is designed to distinguish between a process being in statistical control and being out of control, not to assess gauge capability.

A more conventional way to compare within-sample variability and among-sample variability is using the Analysis of Variance:

```
summary(aov(Measurement ~ Part, data = gauge))
```

<table>
<thead>
<tr>
<th></th>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>factor(Part)</td>
<td>19</td>
<td>377.4</td>
<td>19.86</td>
<td>26.48</td>
<td>3.16e-10 ***</td>
</tr>
<tr>
<td>Residuals</td>
<td>20</td>
<td>15.0</td>
<td>0.75</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The $F$-value on the factor (Part) line tests the null hypothesis that there is no difference among parts. In the present context, we can interpret that as the null hypothesis that the differences among parts are too small to be detected using this gauge; that is, that the gauge capability is inadequate.

Since the null hypothesis is soundly rejected ($P = 3.16 \times 10^{-10}$), we infer that the gauge is indeed capable of separating the part-to-part variability from measurement error.
Precision-to-tolerance ratio

Gauge capability $\hat{\sigma}_{\text{Gauge}}$ can be compared with the tolerance implied by specification limits:

$$\frac{P}{T} = \frac{6\hat{\sigma}_{\text{Gauge}}}{\text{USL} - \text{LSL}}$$

where the familiar constant 6 is sometimes replaced by 5.15.

In the example, $\hat{\sigma}_{\text{Gauge}} = 0.866$, and the specification limits are USL = 60, LSL = 5, so $P/T = 0.0945$.

As a rule of thumb, $P/T \leq 0.1$ is often taken to mean adequate gauge capability, but this is only a rough guide.
Variance components

If we write a measurement as

\[ Y = X + \epsilon, \]

where \( X \) is the true value for a particular part, with standard deviation \( \sigma_P \), and \( \epsilon \) is the measurement error, with standard deviation \( \sigma_{\text{Gauge}} \), then the standard deviation of \( Y \) is \( \sigma_{\text{Total}} \), where

\[ \sigma_{\text{Total}}^2 = \sigma_P^2 + \sigma_{\text{Gauge}}^2. \]

We have an estimate of \( \sigma_{\text{Gauge}} \) from the control chart, and we can estimate \( \sigma_{\text{Total}} \) directly from the measurements, so we can estimate \( \sigma_P \) by subtraction.
In the example, $\hat{\sigma}_{\text{Total}} = 3.172$ and $\hat{\sigma}_{\text{Gauge}} = 0.887$, so

$$\hat{\sigma}_P = \sqrt{3.172^2 - 0.887^2} = 3.045.$$
These calculations are usually carried out using the Analysis of Variance:

```
summary(aov(Measurement ~ factor(Part), data = gauge))
```

Expected mean squares are

\[ E(MS_{\text{Residuals}}) = \sigma_{\text{Gauge}}^2 \]
\[ E(MS_{\text{Part}}) = \sigma_{\text{Gauge}}^2 + n\sigma_P^2 \]

where \( n = 2 \) is the sample size.
We estimate $\sigma_P$ by

$$\hat{\sigma}_P = \sqrt{\frac{1}{n}(MS_{Part} - MS_{Residuals})}$$

$$= 3.091,$$

essentially the same value as obtained from $\bar{R}$. 
Note

This way of estimating variance components works only for balanced data like these.

It fails if, for instance, the sample sizes are not all equal.

Tools for working with “mixed models” are needed for unbalanced data.

In R:

```r
library(lme4)
summary(lmer(Measurement ~ (1 | Part), data = gauge))
```

The output gives $\hat{\sigma}_{\text{Part}} = 3.091$, and $\hat{\sigma}_{\text{Residual}} = 0.866$, with no need for solving equations.
Gauge R & R using ANOVA

Example: \( p = 10 \) parts measured by \( o = 3 \) “operators” (inspectors), each making \( n = 3 \) measurements.

Difference among measurements for a given inspector reflect \textit{repeatability}.

Difference among inspectors reflect \textit{reproducibility}.

Random effects statistical model:

\[
Y_{ijk} = \mu + P_i + O_j + (PO)_{ij} + \epsilon_{ijk},
\]

\( i = 1, \ldots, p, \quad j = 1, \ldots, o, \quad k = 1, \ldots, n. \)
**In R:**

```r
thermal <- read.csv("Data/Table-08-07.csv")
summary(aov(Impedance ~ factor(Part) * factor(Inspector),
    data = thermal))
```

Expected mean squares are

\[
E(\text{MS}_{\text{Residuals}}) = \sigma^2,
\]

\[
E(\text{MS}_{\text{Part}}) = \sigma^2 + n\sigma^2_P + on\sigma^2_P
\]

\[
E(\text{MS}_{\text{Inspector}}) = \sigma^2 + n\sigma^2_P + pn\sigma^2_O
\]

\[
E(\text{MS}_{\text{Part:Inspector}}) = \sigma^2 + n\sigma^2_P
\]
Solve for estimates of the variance components:

\[ \hat{\sigma}^2 = MS_{\text{Residuals}} \]
\[ = 0.511 \]

\[ \hat{\sigma}_P^2 = \frac{1}{on}(MS_{\text{Part}} - MS_{\text{Part:Inspector}}) \]
\[ = 48.293 \]

\[ \hat{\sigma}_O^2 = \frac{1}{pn}(MS_{\text{Inspector}} - MS_{\text{Part:Inspector}}) \]
\[ = 0.565 \]

\[ \hat{\sigma}_{PO}^2 = \frac{1}{n}(MS_{\text{Part:Inspector}} - MS_{\text{Residuals}}) \]
\[ = 0.728. \]
The gauge **repeatability** $\sigma^2_{\text{Repeatability}}$ is measured by $\sigma^2$:

$$\hat{\sigma}_{\text{Repeatability}} = \hat{\sigma} = 0.715.$$ 

The gauge **reproducibility** $\sigma^2_{\text{Reproducibility}}$ is measured by $\sigma^2_O + \sigma^2_{PO}$:

$$\hat{\sigma}_{\text{Reproducibility}} = 1.137.$$ 

The gauge **capability** $\sigma^2_{\text{Gauge}}$ is the sum of $\sigma^2_{\text{Repeatability}}$ and $\sigma^2_{\text{Reproducibility}}$:

$$\hat{\sigma}_{\text{Gauge}} = 1.343.$$ 

The specification limits are $\text{LSL} = 18$ and $\text{USL} = 58$, so the $P/T$ ratio is estimated to be

$$P/T = \frac{6\hat{\sigma}_{\text{Gauge}}}{\text{USL} - \text{LSL}} = \frac{6 \times 1.343}{58 - 18} = 0.201.$$
As before, tools for working with “mixed models” are more convenient and more general, in that unlike ANOVA they work with unbalanced data.

In R:

```r
library(lme4)
summary(lmer(Impedance ~ (1 | Part) + (1 | Inspector) + 
               (1 | Inspector : Part), data = thermal))
```

The output gives $\hat{\sigma}_{\text{Inspector:Part}}$, $\hat{\sigma}_\text{Part}$, $\hat{\sigma}_\text{Inspector}$, and $\hat{\sigma}_\text{Residual}$, with no need for solving equations.