

IE 361 Module 20

Design and Analysis of Experiments: Part 1

Prof.s Stephen B. Vardeman and Max D. Morris

Reading: Section 6.1, *Statistical Quality Assurance Methods for Engineers*

After one brings a process to physical stability and quantifies what it is capable of doing, it's reasonable to consider fundamental changes to its configuration/how it is run. Intelligent/efficient data collection and analysis aimed at finding fundamental process improvements is the subject of the final set of modules of this course. The topic is the "design and analysis of experiments," with the goal of eventually addressing complex situations where there are many "process knobs" (factors), each with multiple settings (levels) and thus many many potential ways that things *could* be done, and the object is to find good combinations of levels of important factors.

Figure 1 illustrates the problem addressed in these modules. The noisy process output is affected by variables x_1, x_2, x_3 and potentially other variables (both recognized and unrecognized). The question is how to set up the "control panel" (the settings of the "knobs" or values of some variables x_1, x_2, x_3) to collect data and efficiently learn how to optimize the process to get desired values of y .

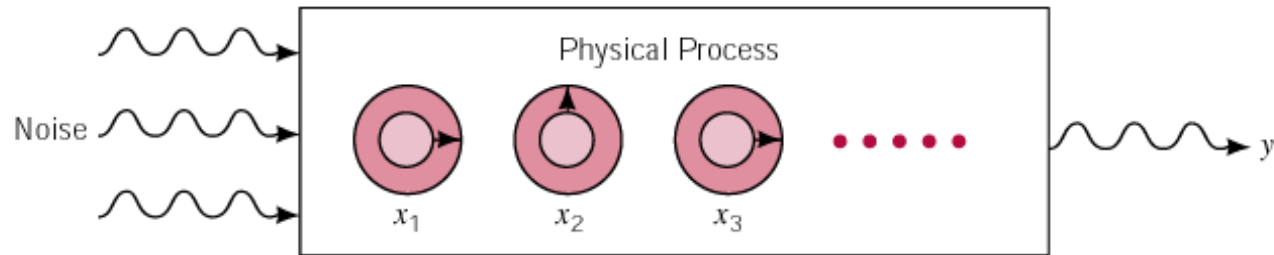


Figure 1: A Process With Many Inputs x or Factors Affecting a Response y

Samples from r Different Sets of Process Conditions and A One-Way Normal Model for Experimental Results

We begin with a most basic experimental scenario, where one has data consisting of observed responses, y , for some number, r , different processes conditions. We'll write

y_{ij} = the j th response in sample i (made under the i th set of process conditions)

where sample sizes are n_1, n_2, \dots, n_r .

Example 20-1 A classic data set from Devore's *Probability and Statistics for Engineering and the Sciences* concerns the current required to achieve a target

brightness on a type of television tube. All combinations of 2 types of glass and 3 types of phosphor created $r = 6$ types of tubes. Tests on 3 tubes of each type produced the measured current requirements y_{ij} (in μA) recorded in the following table with corresponding summary statistics.

Type 1 Tubes (Glass 1, Phosphor 1)	Type 2 Tubes (Glass 1, Phosphor 2)	Type 3 Tubes (Glass 1, Phosphor 3)
$y_{11} = 280$	$y_{21} = 300$	$y_{31} = 270$
$y_{12} = 290$	$y_{22} = 310$	$y_{32} = 285$
$y_{13} = 285$	$y_{23} = 295$	$y_{33} = 290$
$\bar{y}_1 = 285$ $s_1^2 = 25$	$\bar{y}_2 = 301.67$ $s_2^2 = 58.33$	$\bar{y}_3 = 281.67$ $s_3^2 = 108.33$
Type 4 Tubes (Glass 2, Phosphor 1)	Type 5 Tubes (Glass 2, Phosphor 2)	Type 6 Tubes (Glass 2, Phosphor 3)
$y_{41} = 230$	$y_{51} = 260$	$y_{61} = 220$
$y_{42} = 235$	$y_{52} = 240$	$y_{62} = 225$
$y_{43} = 240$	$y_{53} = 235$	$y_{63} = 230$
$\bar{y}_4 = 235$ $s_4^2 = 25$	$\bar{y}_5 = 245$ $s_5^2 = 175$	$\bar{y}_6 = 225$ $s_6^2 = 25$

It is often useful to model observations from r samples of respective sizes n_1, n_2, \dots, n_r as independent random samples from normal distributions with

possibly different means $\mu_1, \mu_2, \dots, \mu_r$ but a common standard deviations σ . Figure 2 illustrates these distributional assumptions.

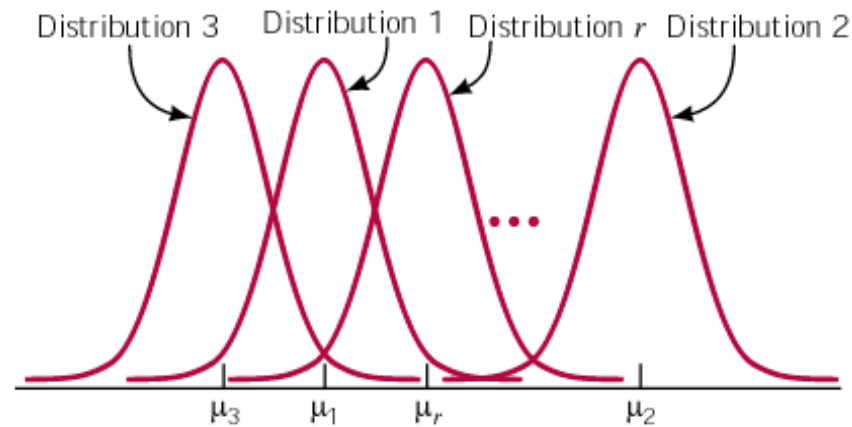


Figure 2: Distributions of Responses Under r Different Sets of Process Conditions

This basic "one-way normal model" is sometimes expressed in symbolic form as

$$y_{ij} = \mu_i + \epsilon_{ij}$$

where the ϵ_{ij} are independent normal random variables with mean 0 and standard deviation σ .

Section 6.1 of *SQAME* discusses ways (based on examination of residuals much as in the regression analysis of Stat 231) for investigating the reasonableness of the one way normal model in a particular application. In this module and the ones that follow, we will take for granted that such work has been taken care of, and consider what then can be done in the way of statistical inference and planning.

Estimation of σ and Linear Combinations of Means

Where the one-way normal model is appropriate, it makes sense to pool together the r standard deviations s_1, s_2, \dots, s_r to make a single pooled estimate of the common group standard deviation, σ . The way we will do this is to use

$$\begin{aligned} s_{\text{pooled}}^2 &= \frac{(n_1 - 1) s_1^2 + (n_2 - 1) s_2^2 + \dots + (n_r - 1) s_r^2}{(n_1 - 1) + (n_2 - 1) + \dots + (n_r - 1)} \\ &= \frac{(n_1 - 1) s_1^2 + (n_2 - 1) s_2^2 + \dots + (n_r - 1) s_r^2}{n - r} \end{aligned}$$

as an estimate of σ^2 , where $n = n_1 + n_2 + \dots + n_r$ is the total number of observations in the study. This estimate of σ^2 is a weighted average of the r

sample variances. Corresponding to it is the estimate

$$s_{\text{pooled}} = \sqrt{s_{\text{pooled}}^2}$$

of σ , the standard deviation of responses for any fixed one of the conditions $1, 2, \dots, r$. This pooled sample standard deviation can be used to make confidence limits for σ . These are

$$s_{\text{pooled}} \sqrt{\frac{n-r}{\chi_{\text{upper}}^2}} \quad \text{and} \quad s_{\text{pooled}} \sqrt{\frac{n-r}{\chi_{\text{lower}}^2}}$$

where the appropriate degrees of freedom are $\nu = n - r$.

Example 20-1 continued The $r = 6$ sample standard deviations for the different tube types in the glass-phosphor study are pooled to make

$$\begin{aligned} s_{\text{pooled}} &= \sqrt{\frac{2(25) + 2(58.33) + 2(108.33) + 2(25) + 2(175) + 2(25)}{18 - 6}} \\ &= 8.3 \mu\text{A} \end{aligned}$$

This intends to measure the variation in current required to produce the standard brightness in tubes of any single type. Appropriate degrees of freedom for this estimate are $\nu = 18 - 6 = 12$ and 95% confidence limits are

$$8.3\sqrt{\frac{12}{23.337}} \quad \text{and} \quad 8.3\sqrt{\frac{12}{4.404}}$$

that is

$$6.0 \mu\text{A} \quad \text{and} \quad 13.8 \mu\text{A}$$

σ (and its estimate, s_{pooled}) is a measure of basic background noise in an experiment, a baseline against which any apparent differences in average response for different conditions are to be measured. One specific way in which these comparisons can be made, is to make confidence limits for interesting linear combinations of means $\mu_1, \mu_2, \dots, \mu_r$. That is, we'll let

$$L = c_1\mu_1 + c_2\mu_2 + \dots + c_r\mu_r$$

stand for an arbitrary linear combination of group mean responses. The corresponding linear combination of sample means is

$$\hat{L} = c_1\bar{y}_1 + c_2\bar{y}_2 + \cdots + c_r\bar{y}_r$$

and is an obvious estimate of L . Useful special cases of this formalism are:

$$\begin{array}{lll} c_i = 1 \text{ and all others } 0 & L = \mu_i & \hat{L} = \bar{y}_i \\ c_i = 1, c_{i'} = -1 \text{ and all others } 0 & L = \mu_i - \mu_{i'} & \hat{L} = \bar{y}_i - \bar{y}_{i'} \end{array}$$

The first of these is the mean for condition i and the second is the difference between the condition i and condition i' response means.

Confidence limits for L can be based on \hat{L} and s_{pooled} as

$$\hat{L} \pm t s_{\text{pooled}} \sqrt{\frac{c_1^2}{n_1} + \frac{c_2^2}{n_2} + \cdots + \frac{c_r^2}{n_r}}$$

The degrees of freedom for t are those associated with s_{pooled} , namely $\nu = n - r$. The quantity $s_{\text{pooled}} \sqrt{\frac{c_1^2}{n_1} + \frac{c_2^2}{n_2} + \dots + \frac{c_r^2}{n_r}}$ is an estimate of the standard deviation of \hat{L} and t times this is a kind of "margin of error" for estimating L .

Example 20-1 continued In the glass-phosphor study, 95% confidence limits for the mean current requirement for tube type i are

$$\bar{y}_i \pm t s_{\text{pooled}} \sqrt{\frac{(1)^2}{n_i}}$$

that is

$$\bar{y}_i \pm 2.179 (8.3) \sqrt{\frac{1}{3}} \quad \text{or} \quad \bar{y}_i \pm 10.44 \mu\text{A}$$

That is, each of the 6 sample means is in some sense "good to within $10.44 \mu\text{A}$ " as representing the corresponding tube mean current requirement.

As a second use of the formula for confidence limits for L , consider the estimation of the difference in current requirement means for tube type i tube type i' . Limits are

$$\bar{y}_i - \bar{y}_{i'} \pm t s_{\text{pooled}} \sqrt{\frac{(1)^2}{n_i} + \frac{(-1)^2}{n_{i'}}$$

So 95% confidence limits for $\mu_i - \mu_{i'}$ in the glass phosphor problem are

$$\bar{y}_i - \bar{y}_{i'} \pm 2.179 (8.3) \sqrt{\frac{1}{3} + \frac{1}{3}} \text{ or } \bar{y}_i - \bar{y}_{i'} \pm 14.77$$

So, for example, limits for comparing tube types 1 and 2 are

$$(285 - 301.67) \pm 14.77 \text{ i.e. } -16.67 \pm 14.77$$

and there is clear evidence of a difference between current requirement means for tube types 1 and 2 since the margin of error of estimation (14.77) is smaller than the absolute value of the observed difference in sample means (16.67).

On the other hand, 95% confidence limits for comparing tube types 1 and 3 are

$$(285 - 281.67) \pm 14.77 \text{ i.e. } 3.33 \pm 14.77$$

and since the "margin of error in estimation" (14.77) is larger than the observed absolute difference in sample means (3.33)