

IE 361 students Brown, Freyer, and Parayno worked on the placement of “center” buttons on some circular disks. They quantified the goodness of button placement by measuring

$$y = \text{disk radial run-out}$$

(in inches). The first step in the students’ work was a measurement study.

a) One student took a single disk, placed it in the measuring device and measured y . She then removed the disk, placed it a second time in the device, and measured y . She repeated this (with the same disk) a total of 10 times. The values she obtained had $R = .0045$ in, $s = .0012$ in, and $\bar{y} = .0350$ in.

i) The values R and s measure what kind of variation? (**Circle exactly one** of the following):

repeatability variation

reproducibility variation

both repeatability and reproducibility

ii) Here, R is several times s . This should be no surprise. **Explain why** in 20 words or less.

iii) Use the value of s above and **find 90% confidence limits** for the “population” standard deviation, σ , corresponding to s . (Plug correct numbers into a correct formula, but don’t bother to simplify.)

b) Each of 15 company technicians individually placed a second disk (all used the same one) in the measuring device and measured y . The sample standard deviation of the values they obtained was .0050 in.

i) In terms of the variances $\sigma_\alpha^2, \sigma_\beta^2, \sigma_{\alpha\beta}^2$, and σ^2 that are parameters of the two-way random effects model often used to describe a standard gauge R&R study, the standard deviation above is an empirical approximation of what? (**Give an expression** involving some or all of these that .0050 in estimates.)

ii) If engineering specifications on run-out are 0 to .0600 in and one takes the .0050 in value as indicating “gauge R&R variability,” what is an estimated gauge capability ratio (or “precision to tolerance ratio”) here? Is it a “good” one? (**Give a number and say** whether or not it is favorable.)

At a later date, run-out values were measured once (by a fixed technician) on each of 20 other disks produced over a very short period. These values had a sample standard deviation of .0040 in .

c) Based on the .0040 in figure and the results of the measurement study, **give an estimate** of the actual disk-to-disk short term standard deviation of run-out for this process (not including measurement error).

d) Your answer to c) is itself subject to uncertainty, that can be quantified in terms of a standard error (an estimated standard deviation). **Give this standard error** for your answer to c). (Plug in, but you need not simplify.)

Suppose now that it is desirable to set up a process monitoring scheme for disk run-out. Samples of $n = 5$ disks will be taken hourly from production and run-out measured. Under a scenario in which there is a single technician that does all the measuring, a “stable process” model (stable production process and stable measuring process model) with “ σ ” of about .0040 in is perhaps appropriate. Suppose further that a standard mean run-out of $\mu = .0350$ in is realistic.

e) Find appropriate standards-given control chart limits for \bar{x} and s under this scenario.

$$UCL_{\bar{x}} = \underline{\hspace{2cm}}$$

$$UCL_s = \underline{\hspace{2cm}}$$

$$LCL_{\bar{x}} = \underline{\hspace{2cm}}$$

$$LCL_s = \underline{\hspace{2cm}}$$

Here are \bar{x} and s values for run-outs from 10 consecutive periods.

Period	1	2	3	4	5	6	7	8	9	10	
\bar{x}	.0257	.0412	.0428	.0324	.0416	.0348	.0347	.0340	.0420	.0323	$\sum \bar{x} = .3615$
s	.0052	.0048	.0010	.0046	.0077	.0044	.0050	.0028	.0030	.0016	$\sum s = .0401$

f) What do the limits of part **e)** applied to the values above indicate about the stability of the (production plus measuring) process?

g) Do retrospective \bar{x} chart control limits produce a much different outcome from that in part **f)**? (**Compute** the limits and **say** whether there is a fundamental change in conclusion.)

$$UCL_{\bar{x}} = \underline{\hspace{2cm}}$$

$$LCL_{\bar{x}} = \underline{\hspace{2cm}}$$

h) As a matter of fact, a different technician made the measurements in each of the 10 periods represented in the values in the table. **How** might this help explain the results in parts **f)** and **g)**?

i) The ideal is to eliminate technician differences in measuring technique. But if it is impossible to do so, one can take them into account in setting up process monitoring schemes. If “ σ ” represents stable “production process plus single-technician measuring process” variability, and $\sigma_{\text{reproducibility}}$ represents technician-to-technician variability, a sensible standard deviation for \bar{y} becomes $\sqrt{\sigma_{\text{reproducibility}}^2 + \sigma^2 / n}$. In light of this, what limits for an \bar{x} chart would use in place of limits from part **e)** if $\sigma_{\text{reproducibility}} = .0049$ in ?

1. A case study in *Creating Quality* by Kolarik concerns the monitoring of a die-casting process for a 35-mm camera body. Over 20 shifts of production, a total of 2735 camera bodies were made and inspected. A total of 215 nonconformities were identified on those 2735 bodies, and 97 of the bodies were judged to be nonconforming. A retrospective analysis of the data from the 20 shifts found no evidence of process instability over that period, and so these historical rates are taken as standard values.

a) In a subsequent shift, 150 camera bodies are produced. A total of 12 nonconformities are identified on these, and 9 of the bodies are judged to be nonconforming. Do either of these counts provide clear evidence of process change? (Compare the two counts to appropriate control limits and interpret your result.)

12 nonconformities



9 nonconforming

Interpretation:

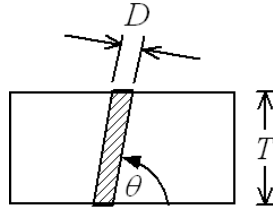
b) Suppose that at some point in the future, the standard nonconformity rate is reduced, production is increased to 200 bodies per shift, and a process monitoring system is adopted that signals “out of control”/“intervention warranted” if there are 3 or more nonconformities observed on a shift. Evaluate the ARL of this monitoring system if in fact the nonconformity rate is .01 per camera body. Then interpret the value in 25 words or less in the context of the application. (In case either is relevant, the binomial pmf is

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x} \text{ and the Poisson pmf is } f(x) = \frac{e^{-\lambda} \lambda^x}{x!} .)$$

ARL = _____

Interpretation:

2. A hole with a circular cross section of diameter D is to be drilled in a rectangular metal block of thickness T . While the hole axis is supposed to be perpendicular to the top face of the block, it may in fact make an angle of θ (different from $\pi/2$) with the faces of the block. This is pictured below.



The amount of metal removed in drilling is then

$$R = \left(\frac{\pi}{4}\right) \frac{D^2 T}{\sin \theta}$$

In drilling, all of D, T , and θ are subject to variation. Suppose that means and standard deviations are

$\mu_D = 3$ and $\sigma_D = .3$, $\mu_T = 50$ and $\sigma_T = .5$, while $\mu_\theta = \frac{\pi}{2}$ and $\sigma_\theta = .01\pi$ (the linear measures are in mm and the angle is in radians). Find an approximate standard deviation of R .

3. The same air supply runs both heads of a pneumatic wrench used to simultaneously tighten two nominally identical bolts in the assembly of a hydraulic pump. Process standards for the torque required to loosen either of those bolts are $\mu = 40$ ft lbs and $\sigma = 2$ ft lb, and the standard for correlation between the two torques is $\rho = .8$. A check of a single newly assembled pump produces a pair of measured torques of 37 ft lbs and 43 ft lbs. Does this clearly indicate a change from process standard values? (Show appropriate calculations.)

4. Lengths of steel shafts are measured by a laser gauge that produces a coded voltage proportional to shaft length (above some reference length). In the units produced by the gauge, shafts of a particular type have length specifications 2300 ± 20 . Below are measured lengths of $n = 10$ such shafts.

2298, 2301, 2298, 2289, 2291, 2290, 2290, 2287, 2280, 2289

Assume that these shafts were produced by a physically stable process.

a) How confident are you that an 11th shaft produced by this process would measure at least 2280? (Give a number.)

Henceforth assume that it is sensible to model shaft length as normally distributed. For the lengths listed above, $\bar{x} = 2291.3$ and $s = 6.2$.

b) Give 90% confidence limits for a process capability index that measures process *potential*.

c) Give two-sided limits that you are 95% sure contain at least 99% of all shaft lengths.

d) Give a 95% lower confidence bound for a process capability index that measures current process performance.

Below are 20 multiple choice questions. Each is worth 5 points. Choose the *single* best answer for each, and fill in the corresponding circle on the answer sheet *using pencil*.

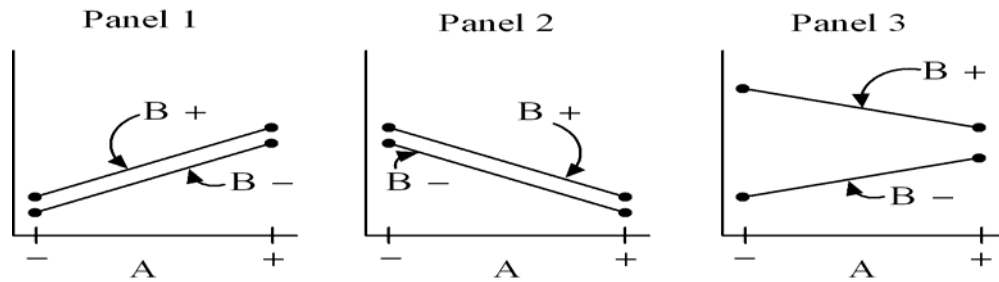
1. Replication in an experiment
 - a) requires that every set of experimental conditions be run at least twice.
 - b) provides a solid basis upon which to assess the size of experimental error or background variation.
 - c) is the basis on which one can compute a pooled standard deviation of response and make confidence intervals for factor effects.
 - d) All of a), b) and c) are true.
 - e) Exactly two of a), b) and c) are true.

2. In a two-way factorial context, if Factors A and B do not interact,
 - a) the mean response can be understood in terms of an overall mean and the separate action of A and B.
 - b) there is “parallelism” of traces for different levels of A on a plot of means against level of B.
 - c) there is “parallelism” of traces for different levels of B on a plot of means against level of A.
 - d) All of a), b) and c) are true.
 - e) None of a), b) and c) are true.

3. Suppose that in a 2×2 factorial, one is sure that factors A and B do not interact, but has data from only combinations (1), a and b.
 - a) The difference in A main effects (for levels 2 and 1 of A) can be estimated as $\bar{y}_a - \bar{y}_{(1)}$.
 - b) The difference in B main effects (for levels 2 and 1 of B) can be estimated as $\bar{y}_b - \bar{y}_{(1)}$.
 - c) The mean response for combination ab can be estimated as $\bar{y}_a + (\bar{y}_b - \bar{y}_{(1)})$.
 - d) All of a), b) and c) are true.
 - e) Exactly two of a), b) and c) are true.

4. If sample means in a 2×2 factorial are $\bar{y}_{(1)} = 13, \bar{y}_a = 7, \bar{y}_b = 7$ and $\bar{y}_{ab} = 13$
 - a) fitted A main effects are all 0.
 - b) fitted B main effects are all 0.
 - c) fitted AB two-factor interactions are all 0.
 - d) None of a), b) and c) are true.
 - e) Exactly two of a), b) and c) are true.

5. Suppose that in a 2×2 factorial $\bar{y}_{..} = 10, a_2 = 3, b_2 = -1, ab_{22} = 2$. What is $\bar{y}_{(1)}$?
 - a) 4
 - b) 10
 - c) 12
 - d) 14
 - e) None of a) through d) is correct.



The vertical axis on the plots above is “mean response” (on the same scale for all 3 plots).

6. If the panels above are interaction plots for 3 different 2×2 factorials, which ones show the *largest* A Main Effects, B Main Effects and AB interactions (in that order)?

- A Main- Panel 3, B Main- Panel 1, AB Interactions- Panel 1
- A Main- Panel 3, B Main- Panel 2, AB Interactions- Panel 2
- A Main- Panel 1, B Main- Panel 3, AB Interactions- Panel 3
- A Main- Panel 1, B Main- Panel 1, AB Interactions- Panel 3
- A Main- Panel 3, B Main- Panel 3, AB Interactions- Panel 3

7. If Panels 1 and 2 respectively show means for low and high levels of Factor C in a 2^3 factorial, are there non-zero C main effects or ABC 3-factor interactions?

- It is impossible to tell from these plots.
- (non-zero) C Main effects- YES; (non-zero) ABC 3-factor interactions- YES
- (non-zero) C Main effects- YES; (non-zero) ABC 3-factor interactions- NO
- (non-zero) C Main effects- NO; (non-zero) ABC 3-factor interactions- YES
- (non-zero) C Main effects- NO; (non-zero) ABC 3-factor interactions-NO

8. Suppose Panel 1 shows means for the low level of Factor C in a 2^3 factorial. If Panel 2 represents the high level of C are there non-zero AC 2-factor interactions? If Panel 3 represents the high level of C are there non-zero AC 2-factor interactions?

- It is impossible to tell from these plots about AC 2-factor interactions.
- Panels 1 and 2, AC interactions- YES; Panels 1 and 3, AC interactions- YES
- Panels 1 and 2, AC interactions- YES; Panels 1 and 3, AC interactions- NO
- Panels 1 and 2, AC interactions- NO; Panels 1 and 3, AC interactions- YES
- Panels 1 and 2, AC interactions- NO; Panels 1 and 3, AC interactions- NO

9. In an experimental study there are 8 experimental conditions, two sample sizes are 4 and 5 and the others are all 1. The two samples of size 4 and 5 produce corresponding sample standard deviations of 3.2 and 4.7. The pooled sample standard deviation is closest to

- 3.95
- 4.03
- 4.06
- 4.10
- 4.12

10. In the context of Question 9, 95% confidence limits for a mean response for an experimental condition with sample size 1

- can not be made because of lack of replication.
- are $y \pm 1.895s_{\text{pooled}}$
- are $y \pm 1.895s_{\text{pooled}} / \sqrt{8}$
- are $y \pm 2.365s_{\text{pooled}}$
- are $y \pm 2.365s_{\text{pooled}} / \sqrt{8}$

Use the following information in choosing your answers to Questions 11-15: Low and high levels of 3 Factors (A,B, and C) produce 8 different combinations. Samples of common size $m = 2$ then yield $s_{\text{pooled}} = 1.2$ and values from the Yates algorithm are

$$\bar{y}_{\dots} = 102, a_2 = -.2, b_2 = 5.2, ab_{22} = -3.5, c_2 = 4.1, ac_{22} = -.1, bc_{22} = .1, abc_{222} = -.1$$

11. If one temporarily ignores any factorial structure and simply determines to compare mean responses for the “all low” and “all high” combinations of A, B and C, 95% confidence limits for this difference have a “plus or minus part” closest to

- a) .3
- b) .7
- c) 1.2
- d) 2.8
- e) 3.9

12. If one uses 95% confidence limits to judge the statistical detectability of effects, the sizes of the values produced by the Yates algorithm are compared to

- a) .1
- b) .3
- c) .7
- d) 1.2
- e) 2.8

13. Suppose that by some criterion, only the B main effects, AB 2-factor interactions, and C main effects are judged to be both statistically detectable and of practical importance. Suppose further that large values of the response variable are desirable. What levels of the factors then seem “best”?

- a) A low, B low, C low
- b) A high, B high, C low
- c) A high, B low, C high
- d) A low, B high, C high
- e) A high, B high, C high

14. Suppose that, in fact, there were two more (two-level) factors beyond A, B and C studied in this 8-sample study, say factors D and E. Levels of factors D and E were chosen according to the “generators” $D \leftrightarrow AB$ and $E \leftrightarrow BC$. Levels of factors D and E corresponding respectively to “all low” and “all high” combinations of A, B and C are:

- a) impossible to determine from the given information.
- b) for “all low A,B,C” D is low; for “all high A,B,C” E is low
- c) for “all low A,B,C” D is low; for “all high A,B,C” E is high
- d) for “all low A,B,C” D is high; for “all high A,B,C” E is low
- e) for “all low A,B,C” D is high; for “all high A,B,C” E is high

15. In the context of Question 14, suppose that the values 5.2, -3.5, and 4.1 produced by the Yates algorithm are judged to represent statistically detectable sums of effects. The simplest possible interpretation of this out come is then

- a) All main effects are important (and only main effects are important).
- b) Only B,C and D main effects are important.
- c) Only B,C and E main effects are important.
- d) Only A,B and C main effects are important.
- e) Only C,D and E Main effects are important.

16. Fractional factorial studies
- provide as much information as a full factorial study, as long as proper generators are used.
 - provide a means of screening many possible factors affecting a response down to a smaller number for additional, more detailed examination.
 - allow for economical/feasible study of the effects of many factors on a response.
 - All of a), b), and c) are true.
 - Exactly two of a), b) and c) are true.
17. Normal plotting of fitted effects produced by the Yates algorithm to judge statistical detectability
- is generally preferable to making confidence intervals for that purpose.
 - relies on the principle of least squares.
 - relies on the principle of effect sparsity (that is a kind of “Pareto principle” for experimental causes).
 - All of a), b) and c) are true.
 - Exactly two of a), b) and c) are true.
18. Restricting attention to two-level factors when first studying many-way factorial experimentation and analysis
- makes sense because full factorial studies with many factors having more than two levels are typically so large as to be impractical to carry out in most engineering contexts.
 - allows the use of convenient but specialized tools like the Yates algorithm.
 - allows introduction of the basic ideas (like many-way interaction) in the simplest possible context.
 - All of a), b) and c) are true.
 - Exactly two of a), b) and c) are true.
19. A certain 2^{p-q} fractional factorial plan has defining relation
- $$I \leftrightarrow ABCDE \leftrightarrow BCDF \leftrightarrow AEF$$
- This represents a $\frac{1}{4}$ fraction of a full 2^6 factorial plan.
 - The plan represented by this defining relation confounds some main effects with some 2 factor interactions.
 - The plan represented by this defining relation has $AB \leftrightarrow CDE$ as one of its two generators.
 - All of a), b) and c) are true.
 - Exactly two of a), b) and c) are true.
20. In an experiment with $r = 5$ different sets of experimental conditions, samples of common size $m = 3$ produce $s_{\text{pooled}} = 4$ and sample means that vary from 100.1 to 101.1. Consider 95% confidence intervals for differences of mean responses under pairs of these experimental conditions.
- These are based on 10 degree of freedom t values and indicate there are statistically detectable differences in experimental conditions.
 - These are based on 10 degree of freedom t values and do not indicate the presence of statistically detectable differences in experimental conditions.
 - These are based on 4 degree of freedom t values and indicate there are statistically detectable differences in experimental conditions.
 - These are based on 4 degree of freedom t values and do not indicate the presence of statistically detectable differences in experimental conditions.
 - None of a), b), c), or d) are correct.