2.830J / 6.780J / ESD.63J Control of Manufacturing Processes (SMA 6303)
Spring 2008

Name: ______SOLUTIONS____________________________

Mean: 71.6%; standard deviation 15.3%. 28 students sat the exam.

Massachusetts Institute of Technology

Department of Mechanical Engineering
Department of Electrical Engineering and Computer Science

2.830J/6.780J Control of Manufacturing Processes

Spring 2008

Quiz #2
Thursday – April 24, 2008

In all problems, please show your work and explain your reasoning. Statistical tables for the cumulative standard normal distribution, percentage points of the χ^2 distribution, percentage points of the *t* distribution, and percentage points of the *F* distribution (all from Montgomery, 5th Ed.) are provided.

Problem 1 [45%]

An experiment is designed and executed, in which the design or input variable is *x* and the output variable is *y*. The input range is normalized to [-1, +1]. Experiments are run in the order shown in Table 1 below, with the input setting and output result as given in the table.

Table 1: Full factorial DOE experiment results

Run #	*x*	*y*
1	-1	8
2	1	18
3	-1	9
4	1	19
5	-1	10
6	1	20

Part (a) [5%]

Fit a model of the form $y = \beta_0 + \beta_1 x$ to the data, and determine point estimates for β_0 and β_1.

ANSWER: Since the design is normalized to ±1 and is balanced, we can use simplified contrasts for the estimation of the offset and linear terms:

β_0 = overall average = 14

β_1 = (19-9)/2 = 5

Or $y = 14 + 5x$

Part (b) [10%]

Determine the standard error (std. err.) and 95% confidence intervals for the estimates of β_0 and β_1. Are both parameters significant to 95% confidence or better? Should you include both terms in the model?

ANSWER: First, we need an estimate of the underlying pure error. The residuals for runs 1 through 6 are -1, -1, 0, 0, +1, +1, giving a $SS_E = 4$, with degrees of freedom = 4. So the MS_E is 4/4 = 1. That is to say, our estimate for σ^2_E is 1. With this error estimate, we can now estimate the variance for both the constant and slop terms. For the constant term, $\text{Var}(\beta_0) = \frac{\sigma^2}{n} = \sqrt{\frac{1}{6}} = 0.408$ And for the slope term, $\text{Var}(\beta_1) = \frac{\sigma^2}{\sum (x_i - \bar{x})^2} = \frac{\sigma^2}{n} = \sqrt{\frac{1}{6}} = 0.408$.

Thus the standard error for each estimate is $\text{std. err.} = \sqrt{\frac{1}{6}} = 0.408$. Finally, we can formulate the 95% confidence intervals using the *t* distribution with 4 degrees of freedom (4 since we have a pooled estimate of variance using two different *x* levels): $t_{\alpha/2,\nu} = t_{0.025,4} = 2.776$, so that the confidence intervals are at $\pm 2.776(0.408) = \pm 1.13$ around the point estimates, i.e.:

$\beta_0 = 14 \pm 1.13$ or $12.87 \le \beta_0 \le 15.13$ and

$\beta_1 = 5 \pm 1.13$ or $3.87 \le \beta_0 \le 6.13$.

Part (c) [5%]

Following good practice, we next examine the residuals (differences between the model prediction values and measured values, for our data). In particular, we consider the residuals as a function of run order. What pattern in the residuals raises a concern? What modifications might you suggest to the experimental design or analysis in light of this?

ANSWER: Reviewing the residuals in time order, we see they go as -1, -1, 0, 0, +1, +1, suggesting that there may be a systematic drift in time of the experiment. The design of experiments was constructed in such a way that this possibility is "blocked" against: by alternating between the low and high input values, we do not confound a temporal drift with the main effect or dependence on the input value *x*. Thus, a systematic time drift, if present, is turned into a "noise" factor (and may, indeed, be entirely responsible for our observed residuals). However, we might want to explore this more carefully with an extended experimental design. For example, we might do multiple replicates at a center point, perhaps in sequence, to see if there is additional evidence of a temporal drift. In the analysis, if we are confident that time is indeed a significant factor, then we might include it in the model (or estimate the magnitude of the drift and subtract it from our data, before fitting to the controllable input parameter *x*).

Part (d) [10%]

Setting aside any reservations about the existing experimental design or the model, we next consider using the model to predict some values, for further experimentation and optimization.

(i) We are interested in how well the model predicts outputs, at different values for *x*. Derive a formula for the standard error ($s_{\hat{y}}$) in the output estimate $\hat{y}_i$ as a function of the input value x_i.

ANSWER: We can use the formulas for a mean-centered model from lecture. Or re-deriving:

$$\hat{y}_i = \beta_0 + \beta_1 x$$

$$\text{Var}(\hat{y}_i) = \text{Var}(\beta_0) + (x_i - \bar{x})^2 \text{Var}(\beta_1) = \text{Var}(\beta_0) + x_i^2 \text{Var}(\beta_1) = [\text{std. err.}\, \beta_0]^2 + x_i^2 [\text{std. err.}\, \beta_0]^2$$

$$\text{Var}(\hat{y}_i) = \frac{1}{6}(1 + x_i^2)$$

$$s_{\hat{y}} = \frac{1}{6}\sqrt{1 + x_i^2} = 0.408\sqrt{1 + x_i^2}$$

(ii) Next, we consider a prediction at possible center point in the input space. Provide a 95% confidence interval prediction for the output value at the design space center point, i.e., give the point estimate and 95% confidence interval bounds for $\hat{y}(x = 0)$.

ANSWER: With the standard error in hand for *y* as a function of *x*, we combine this with the *t* distribution to generate a confidence interval with the desired level of confidence:

$$\text{conf. interval } \hat{y}_i = y(x_i) \pm t_{\alpha/2,\nu} \cdot s_{\hat{y}} = y(x_i) \pm 2.77(0.408)\sqrt{1 + x_i^2}$$

So for $x_i = 0$, we have $\hat{y}_i = 14 \pm 2.77(0.408)\sqrt{1 + 0} = 14 \pm 1.13$ consistent with our previously derived estimate for the mean.

(iii) Finally, we also consider an extrapolation of the model beyond the original range of experimentation. Provide a 95% confidence interval prediction for the output value at x equal to 3 (in normalized units), i.e., give the point estimate and 95% confidence interval bounds for $\hat{y}(x=3)$. Comment on the confidence in model outputs as a function of how far we are extrapolating from our experimental region.

ANSWER: Using the formula derived above, for $x_i = 3$, we have $\hat{y}_i = (14 + 5 \cdot x_i) \pm 2.77(0.408)\sqrt{1+3^2} = 29 \pm 1.13 \cdot \sqrt{10} = 29 \pm 3.57$. Thus, we see that the confidence interval widens considerably (from ±1.13 to ±3.57) when we extrapolate from the central region of our experiment. We should be very cautious in using extrapolated values, for example in driving toward an optimal point, outside our region of data.

Part (e) [10%]

We now perform one additional experimental run to augment Table 1. Setting the input x to 0, the output y is experimentally observed to be equal to 7. Based on this, perform and discuss a lack of fit analysis. To 95% confidence or better, does the model from part (a) show evidence of lack of fit?

ANSWER: The lack of fit analysis can be done efficiently, based on the work we have already done. From the linear model, we see that the prediction error is (14-7) = 7, for a sum of squared deviation of 49 or $MS_L = 49/1 = 49$. We compare this to the pure replicate mean square error (previously derived) of $MS_E = 1$. The F ratio is thus 49, which is much larger than $F_{1,4,0.05} = 7.71$, providing ample evidence that the linear model is a poor fit to the data.

An even faster way is to recognize that the observed value $y = 7$ is outside the 95% c.i. calculated in part d(ii), and is thus evidence of lack of fit at the confidence level.

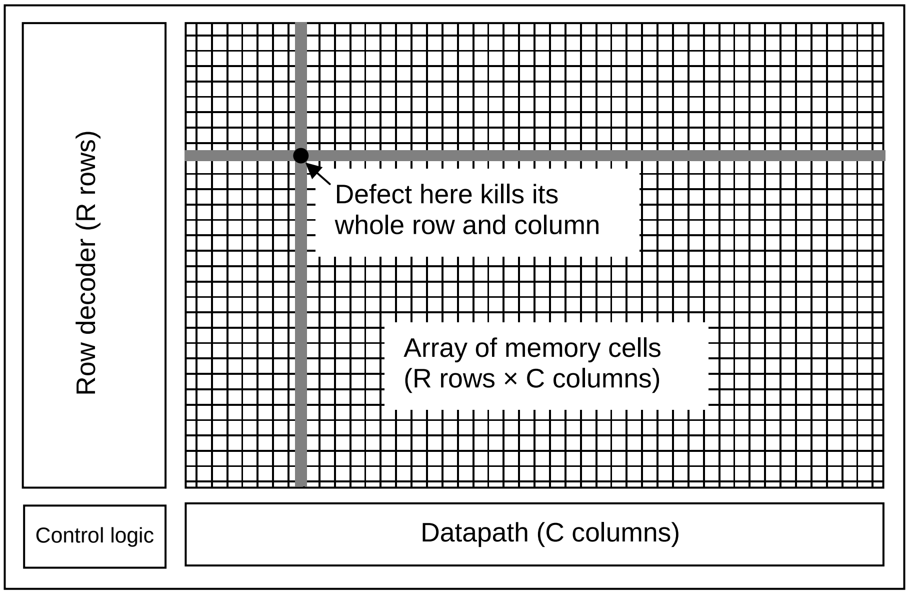
Part (f) [5%]

Fit a new model, including the center point data point (x=0, y=7). I.e., find α_0, α_1, and α_2 for a model of the form $y = \alpha_0 + \alpha_1 \cdot x + \alpha_2 \cdot x^2$. For this part, point estimates are sufficient. Note: it is possible to identify model coefficients by inspection, graphical, or other simplified means; full-scale set up and solution of regression equations is not necessary.

ANSWER: Our α_0 term is simply the observed value when $x = 0$, or $\alpha_0 = 7$. The same linear term will be estimated, since the addition of the quadratic term cannot change this value; thus $\alpha_1 = 5$. Finally, we can plug in for the observed mean at x = -1 to get $9 = 7 + 5x + \alpha_2 x^2 = 7 - 5 + \alpha_2(1)$ so $\alpha_2 = 7$.

Problem 2 [20%]

Consider the random-access memory chip shown below, having R rows and C columns:



We are interested in investigating the defect tolerance of the memory array. Small particles (which may be regarded as causing point defects) land on the memory chip during fabrication, with a per-area density of D_0. Particles landing on the datapath, row decoder or control logic do not cause faults (these circuits are very robustly designed), but when a defect lands on one of the memory cells, each cell having an area A, the entire **row** *and* the entire **column** of cells in which the afflicted cell sits are rendered inoperative.

Assume that the spatial density of defects is very tightly distributed, so that we can write the proportion of memory cells *not* hit by a particle during fabrication as $Y = \exp(-AD_0)$.

Part (a) [3%]

Write down an expression in terms of Y, R, and C for the number, F, of individual memory cells hit by particles during fabrication.

$F = (1-Y)RC$

Part (b) [10%]

Write down expressions in terms of *Y*, *R* and *C* for (i) the proportion of columns that are operative after fabrication, and (ii) the proportion of rows operative.

(i) (C–F)/C or simply (better) Y^R

(ii) (R–F)/R or simply (better) Y^C

where $F = (1-Y)RC$

The first expression in each of (i) and (ii) above assumes that two defects occupy a given row or column.

Part (c) [7%]

Write down an expression in terms of *Y*, *R* and *C* for the overall proportion, *P*, of memory cells that is available for use after fabrication.

Take the product of (i) and (ii) above:

$P = 1 - (1-Y)(R+C) + (1-Y)^2RC$: this assumes that no two defects occupy a given row or column, an assumption that appears invalid below $Y \approx 0.99$.

Or simply Y^{R+C}.

Part (d) [optional; for up to 5 bonus percentage points]

Show that the proportion of usable memory cells is maximized when the memory array is square.

RC must be constant for a memory array of a given designed capacity (assume that the design does not attempt to compensate for the anticipated yield). Y is fixed for the given memory cell design and fabrication process. The aspect ratio (R/C) is the only thing that can change. Therefore we must maximize *P* above by varying R/C.

Let RC = N

$$0 = \frac{dP}{dR} = \frac{d}{dR}\left[1-(1-Y)\left(\frac{N}{R}+R\right)+(1-Y)^2 N\right] = (1-Y)\left(\frac{N}{R^2}-1\right)$$

$$\Rightarrow R = \sqrt{N}$$

Check that P is actually maximized (not minimized) by this solution:

$$\frac{d^2P}{dR^2} = (1-Y)\left(-\frac{2N}{R^3}\right) < 0$$

So the solution does give a maximum of *P*. Therefore the optimal array is square, as required.

A similar derivation is possible if the memory cell yield is Y^{R+C}: yield is maximized for minimal R+C: implies R = C.

Problem 3 [35%]

Part (a) [20%]

Measurements are made of the threshold voltages of MOSFET devices at three randomly chosen locations on each of three wafers themselves chosen randomly from a lot. The deviations of the measured threshold voltages from their target value are shown below. Complete the nested variance analysis. Cells requiring a value to be inserted have a thick border. Some calculations have already been done for you.

Wafer #	Site #	Threshold voltage deviation (mV)	Wafer average (mV)	Squared deviations of point from grand ave. (mV^2)	Squared deviations of wafer ave. from grand ave. (mV^2)	Squared deviations of point from wafer ave. (mV^2)
1	1	1		1		1
1	2	2		0		0
1	3	3	2	1	0	1
2	1	2		0		0
2	2	2		0		0
2	3	2	2	0	0	0
3	1	4		4		4
3	2	0		4		4
3	3	2	2	0	0	0
		Grand average:	2			
			SS_D	10		
			SS_W		0	
			SS_E			10

ANOVA (in mV^2)					
Source	Degrees of Freedom	SS	MS	F_0	F_{crit} (5% level)
WAFER	W-1 = 2	0	0	0	5.14
ERROR	W(M-1) = 6	10	1.67	N/A	N/A
C TOTAL	WM-1 = 8	10	1.25	N/A	N/A

VARIANCE COMPONENTS (in mV^2)				
Variation source	MS	# data in SS	Observed variance	Estimated variance
ERROR (site-to-site)	1.67	1	1.67	1.67
WAFER (wafer-to-wafer)	0	M = 3	0	0
TOTAL	1.25	1	1.25	1.67

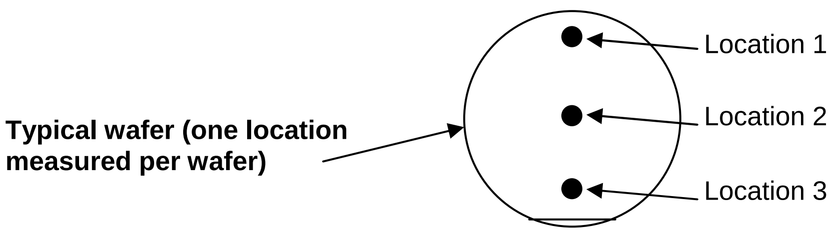
Part (b) [5%]

What do you conclude from the nested variance analysis above?

$F_0 < F_{crit,}$ so there is no evidence of wafer-to-wafer variation. All variation appears to be die-to-die, or in the measurements

Part (c) [10%]

Now you are given threshold voltage deviation data from the same process, but this time taken from nine separate wafers. From each wafer one measurement is taken at one of three specified (*not* randomly chosen) locations. The objective is to check for evidence of any systematic (fixed) relationship between threshold voltage deviation and position on the wafer. Complete the ANOVA below.



Wafer #	Site (location on wafer)	Threshold voltage deviation (mV)	Per-location average (mV)
1	1	1	7/3
2	1	2	
3	1	4	
4	2	2	4/3
5	2	2	
6	2	0	
7	3	3	7/3
8	3	2	
9	3	2	
		grand average:	2

Source of variation	Sum of squares	Degrees of freedom	Mean square	F_0	F_{crit} (5% level)
Between groups	2	2	1	0.75	5.14
Within groups	8	6	4/3	N/A	N/A
Total	10	8	N/A	N/A	N/A

Part (c), cont'd.

Is there evidence at the 95% confidence level of a significant systematic dependence of threshold voltage on site location?

$F_0 < F_{crit}$, so there is no evidence at the 95% significance level of a systematic dependence of threshold voltage on site location.