Semiconductor Manufacturing - EE290H

Second Midterm Exam

Thursday, April 8, 1999

Costas J. Spanos and Kameshwar Poolla

University of California at Berkeley

College of Engineering

Department of Electrical Engineering and Computer Sciences

Your Name: _____*Official* _____*Solutions*_____

(last) *(first)*

Your Signature: _____

1. Print and sign your name on this page before you start.

3. Do everything on this exam, and make your methods as clear as possible.

Problem 1 _____ */25*

Problem 2 _____ */25*

Problem 3 _____ */25*

Problem 4 _____ */25*

TOTAL _____ */ 100*

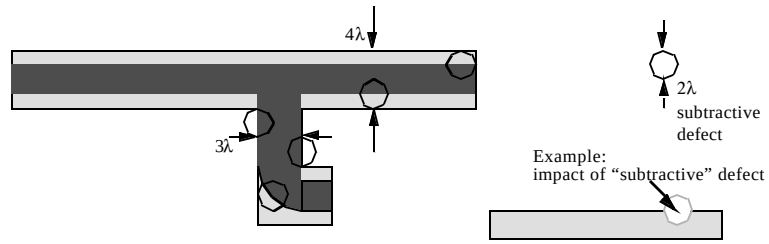
Problem 1 of 4 (Yield) (25 points)

Answer these questions as briefly as you can. Please write clearly!

What is the general relationship between yield and design rule λ? Why such a relationship holds?

Defect density has an exponential impact on yield. Defect density also increases as the square or the cube of $1/\lambda$. So, reducing the pattern size has a huge impact on yield.

Given the metal layout below, please carefully outline the critical area with respect to the center of a defect, whose presence removes material from a disk with diameter 2λ. The metal interconnection has to have at least 2λ remaining width in all places for acceptable integrity.



What is the fundamental distribution of defect counts on a wafer, and what assumption behind that distribution is most likely to fail in reality?

This is the Poisson distribution. Basic assumption is that the defects arrive on the wafer with statistical independence.

You have just created the word's best Yield simulator. It works by extracting the "critical area" given a layout, and it then proceeds to generate, randomly, simulated defects. It does this 30 times and estimates the circuit yield as the fraction of successful simulations. In one such a analysis, the yield was estimated to be 65%. Calculate approximate 95% bounds for the true yield. Explain your method.

since $0.35 \times 0.65 \times 30 = 6.825 > 6$, I can use the normal approximation to the Bernoulli distribution. The σ^2 of this distribution is approximately $0.35 \times 0.65/30 = 0.00758$, so the 95% limits are $0.65 +/- 2\sigma = [0.48, 0.82]$

Problem 2 of 4 (DOE/ANOVA) (25 points)

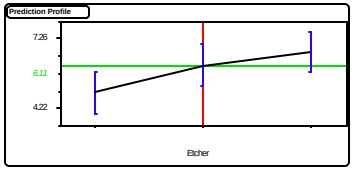
1. For a timed etch process, it is critical that etch rates be consistent among all etchers in an equipment group. An engineer responsible for a group of three of these etchers sets up an experiment in which etch rate will be measured three times in each machine. He records the following data, given in A/sec.

Etcher A	Etcher B	Etcher C
5.72	6.43	5.84
4.22	6.08	7.26
4.76	5.82	6.96

a) Compute a point estimate of the effect of each etcher.

This is simply done by taking the average for each group. (This is a good estimate assuming a normal distribution...)

mean A estimate: 4.90
mean B estimate: 6.11
mean C estimate: 6.68



b) Using ANOVA, do you find difference between the three etchers at a 0.05 level of significance?

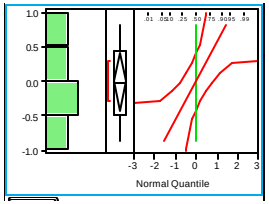
Analysis of Variance:

Source	*DF*	*Sum of Squares*	*Mean Square*	*F Ratio*
Model	*2*	*4.9888222*	*2.49441*	*6.0788*
Error	*6*	*2.4620667*	*0.41034*	*Prob>F*
C Total	*8*	*7.4508889*		*0.0361*

I.e. the difference is significant at the 0.05 level. (The significance level is about 3.6%).

c) Are the residuals normally distributed?

the residuals seem to be normal:



2. Write the generators of a 2^{5-2} fractional factorial experiment with factors A,B,C,D and E.

A=BCD

E=CD

AE = B

a) Are any of the main effects free of aliasing with one another?

No

b) Which simple interactions are aliased with one another?

All of them... Only main effects will not be confounded with other main effects.

c) What advantage does this type of experiment have over a 2^{5-1} fractional factorial experiment?

It uses fewer runs (8 vs 16).

d) What is the resolution of the 2^{5-2} design?

Three.

Problem 3 of 4 (Parameter Estimation) (25 points)

Parameter Estimation

We want to estimate the gate oxide thickness θ on chips produced at the FritoLays DRAM fab. Based on historical data, we conclude that the gate oxide thickness is normally distributed with mean $m = 100$ Angstroms and variance $\sigma^2 = 9$.

1 In the absence of measured data, the best estimate of any parameter is its mean. So we would estimate the gate oxide thickness θ to be 100 Å.

2 To revise this estimate significantly we need new information with roughly three times smaller than the *a priori* variance. Measurements from n wafers provides gate oxide thickness information with a variance of σ_ϵ^2/n. So we need

$$\frac{\sigma_\epsilon^2}{n} \approx \frac{\sigma^2}{3} \quad \text{or} \quad n \approx 5$$

3 As suggested in the hint, the joint density is

$$f(\theta) = p_\theta(\theta)p_y(y) = ce^{-\frac{(\theta-m)^2}{2\sigma^2}}e^{-\frac{(y-\theta)^2}{2\sigma_\epsilon^2}}$$

where c is some constant that is not essential for our calculations. The log-likelihood function is (within a constant)

$$J(\theta) = -\ln f(\theta) = \frac{(\theta-m)^2}{2\sigma^2} + \frac{(y-\theta)^2}{2\sigma_\epsilon^2}$$

This can be minimized by setting the derivative to zero:

$$0 = \frac{dJ}{d\theta} = \frac{\theta - m}{\sigma^2} + \frac{y-\theta}{\sigma_\epsilon^2}$$

Solving for the maximum-likelihood estimate yields

$$\theta_{\text{ML}} = \frac{\sigma^2 y + \sigma_\epsilon^2 \mu}{\sigma^2 + \sigma_\epsilon^2} = 201.95 \text{ Å}$$

Problem 4 of 4 (SPC) (25 points)

1. You are overseeing a patterning process, that is monitored by a Shewhart chart, using only the x-chart and the 3-sigma rule. Every wafer produced is measured at 20 spots, and the average across-wafer critical dimension is calculated from these 20 readings. It has been found from reliable analysis that this *average across-wafer* value is distributed according to N(180nm, $(5\text{nm})^2$). The cost of measuring one wafer is $1, and the cost of resolving a false alarm is $100. Plot the expected cost/wafer for this procedure under the following assumptions:

a. The process is in-control, the control limits are correctly calculated, and the group size is n=5 (i.e. each point on the chart is the average value of 5 wafers).

Well, if the process is in control, we have an ARL of 370. So, for every 370 good samples (at $1 each) we have a false alarm (at $100). The average per wafer cost is then

(1x5x370 + 100)/5x370 = $1.05

b. The process is in-control, but the limits have been mistakenly centered around 190nm (instead of the correct 180nm value). The group size is still 5.

This process will have an ARL of 1.075.... So for every 1.075 samples, there will be a false alarm. The average per wafer cost is then

(1x5x1.075+100)/5x1.075=$19.60

c. The process is in-control, but the limits have been mistakenly centered around 185nm (instead of the correct 180nm value). The group size is still 5.

This process will have an ARL of 4.46... So for every 4.46 samples, there will be a false alarm. The average per wafer cost is then

(1x5x4.46+100)/5x4.46=$5.48

2. (Bonus Question!)
Exponentially weighted moving average charts can be sometimes used to “filter” non-IIND data. provide the best estimate of the “forgetting” factor λ, given some historical in-control, stationary, *non-IIND* data. (Hint: a good choice of the forgetting factor is one that delivers best prediction of the next sample... Formulate the problem and its solution as a least squares estimate of the missing value...)

Given: x_t, t = 1,...,n data series, also expressed as a column vector [x] with n elements.
EWMA formulation: $z_t = \lambda x_t + (1-\lambda) z_{t-1}$

One way to do this is to choose l so as to minimize the sum of squares of the prediction error of the EWMA formulation. That is, use EWMA for a one-step ahead prediction, and then formulate the problem so that l is chosen to minimize the least-squares error of this prediction.