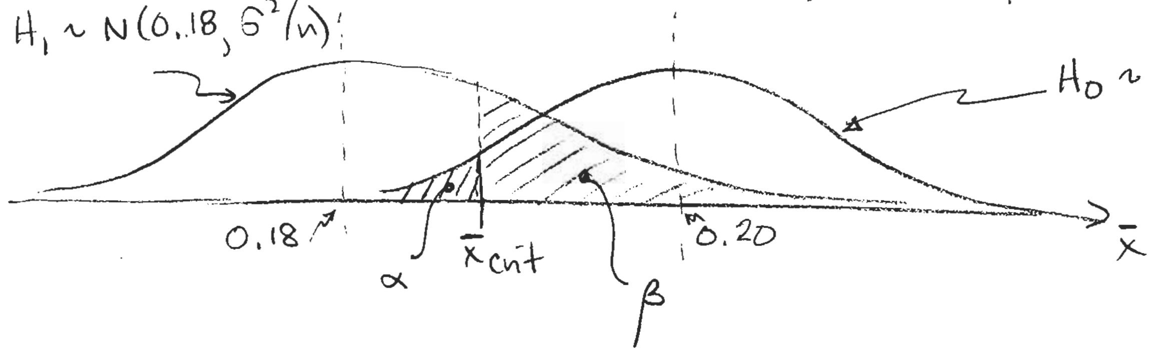
6.780 SEMICOND. MANUFACT — PROBLEM SET 4 SOLUTIONS — Spring 2003

Problem 1

As written, this problem is pretty easy (for the case where $\alpha = \beta$). In a slightly more general form so we are clear on the issues, imagine for arbitrary α and β:



NOTE: σ^2 is Total variance due to both the LW & measurement

H_0: the case where we take n measurements from the UNSHIFTED distribution. Note that we have a sampling distribution for $\bar{x} \sim N(0.20, \sigma^2/n)$ in this case.

H_1: the shifted distribution hypothesis, where $\bar{x} \sim N(0.18, \sigma^2/n)$

- We want to pick an $\bar{x}_{crit}$ where we declare H_0 if we observe $\bar{x} \geq \bar{x}_{crit}$ and H_1 if we observe $\bar{x} < \bar{x}_{crit}$. Here we are looking for a one-sided shift from 0.20 to lower values, so we put all of the α error on one side as above — that is, we have $\alpha = Pr$(falsely declare a shift when none occurred)

$$\alpha = \Phi\left(\frac{0.20 - \bar{x}_{crit}}{\sigma/\sqrt{n}}\right)$$

or

$$z_\alpha = \frac{0.20 - \bar{x}_{crit}}{\sigma/\sqrt{n}}$$

- Given the above way to chose $\bar{x}_{crit}$, we next ask what sample size n we should pick to get a β probability of missing a true shift to 0.18 μm:

$$\beta = 1 - \Phi\left(\frac{\bar{x}_{crit} - 0.18}{\sigma/\sqrt{n}}\right)$$

or

$$z_\beta = \frac{\bar{x}_{crit} - 0.18}{\sigma/\sqrt{n}}$$

- So, in general given α & β we know z_α, z_β, σ (in this problem) so we can determine both $\bar{x}_{crit}$ and n.

- In this case we are told $\alpha = \beta = 0.05$, so $z_\alpha = z_\beta$ and we find $\bar{x}_{crit}$ by symmetry is simply midway between μ_0 and μ_1 (i.e. 0.19 μm), or generally

$$\frac{0.20 - \bar{x}_{crit}}{\cancel{\sigma/\sqrt{n}}} = z_\alpha = z_\beta = \frac{\bar{x}_{crit} - 0.18}{\cancel{\sigma/\sqrt{n}}}$$

$$0.19 = \frac{0.20 + 0.18}{2} = \bar{x}_{crit}$$

- Knowing $\bar{x}_{crit}$ we can find n to get desired β error:

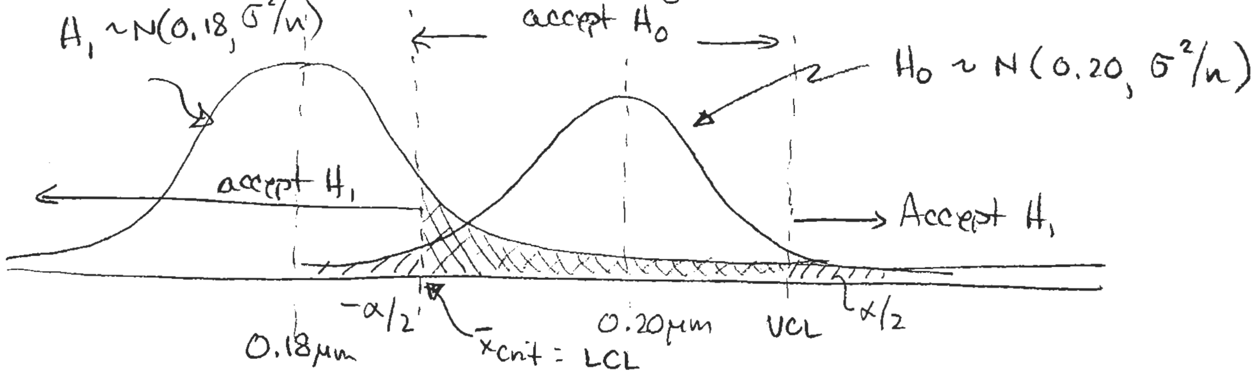
$$\underbrace{z_{0.05}}_{1.65} = \frac{0.19 - 0.18}{\sigma/\sqrt{n}}$$

$$(\sqrt{n}) = \frac{1.65}{0.01}\sigma = \frac{1.65}{0.01}(0.04) = 6.60$$

$$n = 43.56 \Rightarrow n = 44 \qquad (z_{0.025} = 1.96)$$

- NOTE: If we did a two-sided comparison (i.e. just looked for positive or negative 0.05 μm shifts) then we would need more runs (62) to avoid missing a shift, to prob. $\beta = 0.05$. Thus we see the advantage in just looking to one-side – if that's what the application is – as we can get the same β probability with a smaller sample size.

A different approach, closer to the usual two-sided hypothesis test used in control charting, gives inferior results.



- Form H_0 as the hypothesis : "distribution has not shifted positive or negative"

  This gives rise to $\alpha/2$ probably of false rejection of H_0,

Now we look for the power in detecting the one-sided shift to the low end by 0.02 μm. In this case, we also apparently would accept H_1 even if we observed $\bar{x} > \bar{x}_{UCL}$ ... which is somewhat contradictory. This IS the case if we were "control charting" H_0 and had the possibility of rejecting H_0 on the high side. But here we're only LOOKING FOR EVIDENCE OF A NEGATIVE shift in the distribution.

If one applies the above weaker scenario (as sketch above) then one needs to account for the probability of "missing" H_1 as just the region where we accept H_0, or

$$\beta = Pr\{\text{accept } H_0 \mid H_1\} = \int_{LCL}^{UCL} f_{\bar{x}|H_1}(\bar{x})\, d\bar{x}$$

$$\beta = \underbrace{\Phi\left(z_{\alpha/2} - \frac{0.02}{\sigma/\sqrt{n}}\right)}_{\int_0^{UCL} f_{\bar{x}|H_1}(\bar{x})d\bar{x}} - \underbrace{\Phi\left(-z_{\alpha/2} - \frac{0.02}{\sigma/\sqrt{n}}\right)}_{\int_0^{LCL} f_{\bar{x}|H_1}(\bar{x})d\bar{x}}$$

which has to be solved iteratively, giving $n = 52$. This is substantially MORE runs than $n = 44$ which is all that is really needed to detect the difference between H_0 and H_1.

Problem 2

Part a We can complete the ANOVA table as follows. Here we do assume a balanced design (where the same number of runs is done at each treatment level and at each block level).

source of var	sum of sq.	d.o.f.	mean square	F ratio	Pr > F
average	54,000 (d)	1 (a)	54,000 (e)	—	—
blocks	534	2	267 (e)	53.4 (f)	2.36×10^{-5} (g)
treatment	498	4 (b)	124.5 (e)	24.9 (f)	1.44×10^{-4} (g)
residual	40	8 (c)	5 (e)	—	—
total	55,072 (h)	15	—	—	—

(a) $\nu_{average} = 1$ (just one parameter -- the average itself)

(b) We're told the treatment averages and so learn there are 5 treatments, so $\nu_{treatment} = \nu_t = k-1 = 5-1 = 4$

(c) $\nu_{total} = \nu_{ave} + \nu_{blocks} + \nu_{treatment} + \nu_{residual}$

$15 = 1 + 2 + 4 + \nu_r \Rightarrow \nu_{residual} = 8$

(d) $SS_{ave} = n \cdot k \cdot \bar{y}^2$, where we know $\bar{y}$ is the average of the treatment averages, or

$\bar{y} = \frac{1}{5}(55+68+56+66+55) = 60$

Using $k = \#$ treatments
$n = \#$ blocks, we have $SS_{ave} = 3 \cdot 5 \cdot 60^2 = 54,000$

(e) The $MS = SS/\nu$ values are easily calculated from SS & ν

(f) F ratios are calculated using the $MS_{residual} = 5$

(g) $Pr > F$ using the F distribution. E.g.

$Pr(F > 53.4)$ using integrated $F_{2 \text{ or } 4, 8}$ distribution

(h) $SS_{TOTAL} = SS_{ave} + SS_{blocks} + SS_{treatment} + SS_{residual} = 55,072$

Part b Clearly, given the large values of F ratios and the resulting low probability of observing these by chance alone, we can say that both the treatments and blocks are significant.

Part c) What is the applicable statistical distribution (type and parameters) for the treatment averages?

- If we consider the five treatment averages, we can assume they are normally distributed, but with both unknown mean and unknown variance. Since we will have to estimate both $\bar{y}$ and $\sigma^2_{\bar{y}t}$ (where $\bar{y}_t$ is the observed average for the t^{th} treatment), we must use a t-distribution:

$$\frac{\bar{y}_t - \bar{y}}{s_{\bar{y}t}} \sim t_4$$

where we have k-1 d.o.f. in our t distribution.

Estimating these parameters: $\bar{y} = \frac{1}{k}\sum_{t=1}^{k} \bar{y}_t = 60$ as in part b

$$s^2_{\bar{y}t} = \frac{1}{k-1}\sum_{t=1}^{k}(\bar{y}_t - \bar{y})^2 = 41.5$$

So we could write $s_{\bar{y}t} = \sqrt{s^2_{\bar{y}t}} = 6.44$

$$\frac{\bar{y}_t - \bar{y}}{s_{\bar{y}t}} = \frac{\bar{y}_t - 60}{6.44} \sim t_4$$

or

$$\bar{y}_t \sim 6.44 \cdot t_4 + 60$$

We can use this to calculate the probability of observing by chance alone each of the $\bar{y}_t$'s. Using $t \triangleq |\bar{y}_t - \bar{y}|/s_{\bar{y}t}$ below:

observed mean	t	$\Pr(t_4 \geq t)$	Significant w. 80% confidence
$\bar{y}_1 = 55$	0.776	0.24	NO
$\bar{y}_2 = 68$	1.242	0.14	YES
$\bar{y}_3 = 56$	0.621	0.28	NO
$\bar{y}_4 = 66$	0.932	0.20	YES (barely)
$\bar{y}_5 = 55$	0.776	0.24	NO

Notice that this lets us do more than just $F = MS_T/MS_R$, which only says if treatments as a whole are significant.

Part d How conduct the experiment to allow valid conclusions?

One needs to be careful to randomize the order of the blocks and treatments to guard against time effects. If there are any other influential parameters we would also want to block against them.

Another good thing to do is to report the actual data (not just the anova table and treatment averages), as well as the time order of the data. With actual data, one could check such details as how many runs were done at each point. More importantly, one could then check important assumptions: are there any patterns (against time or treatments/blocks) in the residuals?

If one were re-running the experiment, replications (at least at one condition if not everywhere) would help to form an estimate of "pure" or replication error. In this case, the big advantage is to let us check the assumption made and used in this experiment that interactions between the blocks and treatments are not significant.

Problem 3 Drain, pg. 222, problem 3

We have 3 measurements (3 sites) for each of 4 wafers of resistivity ρ. We will want to understand the variances wafer-to-wafer and site-to-site, including interval estimates.

Part a Estimated average resistivity?
- Nothing tricky here, just

$$\bar{\rho} = \frac{1}{WM} \sum_{i=1}^{W} \sum_{j=1}^{M} \rho_{ij} = \frac{1}{N} \sum_{k=1}^{N} \rho_k$$

where W = # wafers = 4
M = # measurements within each wafer = 3
N = total # measurements = $W \cdot M = 12$

The resulting $\bar{\rho}$ is shown on page 8 as $\bar{\rho} = 4.3796$.

Part b Compare point estimates for the wafer and site.

- First, we will build the ANOVA table for this case. Note that using JMP, if one specifies the site number (rather than just putting down replicates for each wafer), then JMP may interpret this as a "factor" and think that site #1 on wafer 1 may be related somehow to site #1 on wafer #2. One can use JMP, but it's better to omit any column numbering which "replicate". The JMP5 model fit / anova table results are shown on page 9.

  On page 8, I've also written the formulas used in respective spreadsheet columns, to arrive at the data for the ANOVA table.

- Now the calculation of variance components: One way is to use the formulas and reasoning in the discussion of Drain. Here, I'll show how I use mean square and sum of square info from the ANOVA table to arrive at the point estimates shown on page 8.

(Discussion continued at page 10).

squared deviations of wafer averages, $(\bar{w}_i - \bar{\rho})^2$

squared residuals, $(\rho_{ij} - \bar{w}_i)^2$

$W=4 \quad M=3$

$\bar{w}_i = \sum_{j=1}^{M=3} \rho_{ij}$

Drain, Problem 3

Wafer (i)	Site (j)	Rho (ρ_{ij})	W_ave	S_D	S_W	S_E
1	1	4.308	4.31200	0.00512	0.00457	0.00002
1	2	4.333		0.00217		0.00044
1	3	4.295		0.00715		0.00029
2	1	4.372	4.34700	0.00006	0.00106	0.00062
2	2	4.339		0.00165		0.00006
2	3	4.330		0.00246		0.00029
3	1	4.431	4.42967	0.00264	0.00251	0.00000
3	2	4.501		0.01474		0.00509
3	3	4.357		0.00051		0.00528
4	1	4.472	4.42967	0.00854	0.00251	0.00179
4	2	4.410		0.00093		0.00039
4	3	4.407		0.00075		0.00051
	Ave	4.3796	4.3796			
	Var	0.0042	0.0035			

SS_D 0.0467

SS_W 0.0319

SS_E 0.0148

ANOVA

Source	d.o.f.	SS	MS	F	Pr > F
C TOTAL	N-1 = 11	0.0467 (SS_D)	0.0042		
WAFER	W-1 = 3	0.0319 (SS_W)	0.0106	5.7594	0.021322
ERROR	8	0.0148 (SS_E)	0.0018		

VARIANCE COMPONENTS

Variation Source		MS	# data in SS	Observed Variance	Estimated Variance	% Var
ERROR (site to site)	1	0.0018	1	0.0018	0.0018	38.66
WAFER (wafer to wafer)	2	0.0106	3	0.0035 (3)	0.0029 (4)	61.34
TOTAL		0.0042	1	0.0042	0.0048 (5)	100.00

Interval Estimates alpha: 0.05

	LOWER	POINT	UPPER
ERROR (site to site)	0.000843	0.0018	0.006784
WAFER (wafer to wafer)	0.000941	0.0029	0.040768

$\bar{\rho}$ average of wafer averages $= \frac{1}{W} \sum_{i=1}^{W} \bar{w}_i$

variance of wafer averages $\frac{1}{W-1} \sum_{i=1}^{W} (\bar{w}_i - \bar{\rho})^2$

Sum of squared deviations $\sum_{i=1}^{W} \sum_{j=1}^{M} (\rho_{ij} - \bar{\rho})^2$ (just sum of this column)

Sum of squared errors (residuals) after predicting wafer ave, or $\sum_{i=1}^{W} \sum_{j=1}^{M} (\rho_{ij} - \bar{w}_i)^2$ (so just sum of this column).

Sum of squared deviations around wafer average

$$SS\_W = \sum_{i=1}^{W} \sum_{j=1}^{M} \underbrace{(\bar{w}_i - \bar{\rho})^2}_{} = M \sum_{i=1}^{W} (\bar{w}_i - \bar{\rho})^2$$

(so just M $\times$ sum of this column).

Noting that this term is same for each set of M measurements.

Response Rho
Whole Model

Summary of Fit

RSquare	0.683524
RSquare Adj	0.564845
Root Mean Square Error	0.042993
Mean of Response	4.379583
Observations (or Sum Wgts)	12

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	3	0.03193758	0.010646	5.7594
Error	8	0.01478733	0.001848	Prob > F
C. Total	11	0.04672492		0.0213

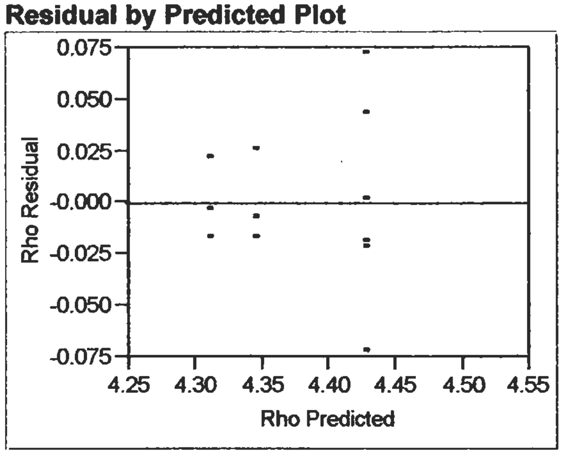
Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	4.3795833	0.012411	352.88	<.0001
Wafer[1]	-0.067583	0.021497	-3.14	0.0137
Wafer[2]	-0.032583	0.021497	-1.52	0.1681
Wafer[3]	0.0500833	0.021497	2.33	0.0482

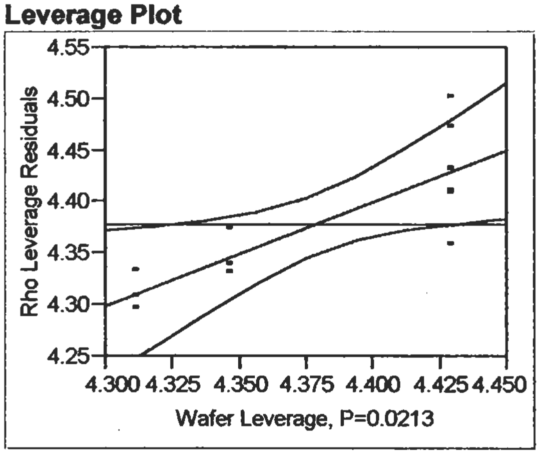
Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Wafer	3	3	0.03193758	5.7594	0.0213

Residual by Predicted Plot



Wafer
Leverage Plot



Least Squares Means Table

Level	Least Sq Mean	Std Error	Mean
1	4.3120000	0.02482215	4.31200
2	4.3470000	0.02482215	4.34700
3	4.4296667	0.02482215	4.42967
4	4.4296667	0.02482215	4.42967

Problem 3, Variance Components.

(1) First, the site-to-site variance is easy to estimate. It's simply $MS_{-E} = \frac{SS_{-E}}{\nu_E} = 0.0018. = s^2_M$ (estimate for σ^2_M)

Notice what this is, and why it makes sense as site to site variance estimate. Remember that SS_{-E} was calculating the squared deviation for each measurement point <u>from its wafer average</u>. So we're just estimating the low level site to site variance within each wafer here.

(2) Now it gets trickier. We know from Drain

$$\sigma^2_W = \sigma^2_{\bar{W}} + \frac{\sigma^2_M}{M}$$

Where $\sigma^2_{\bar{W}}$ is the "observed" wafer to wafer variance. One way to calculate this is just use the variance calculated directly from the 4 wafer averages in the fourth column of our spreadsheet, or

$$s^2_{\bar{W}} = \frac{1}{W-1} \sum_{i=1}^{W} (\bar{w}_i - \bar{\rho})^2$$

In this case, that would be perfectly fine because wafers are not nested inside anything else. If they were (as in Problem 4), however, this would <u>NOT</u> work, because such a calculation would also have <u>inside of it</u> the lot to lot variation.

Instead, here's a better way. We use the MS_W value $MS\_W = \frac{SS\_W}{\nu_4} = 0.0106$. However, when SS_W was calculated, it was ALMOST BUT NOT QUITE estimating "wafer to wafer variance". Instead, it was trying to say how much of the <u>total</u> square deviations SS_{-D} could be explained by wafer to wafer differences. So it represented each $(\bar{w}_i - \bar{\rho})^2$ term <u>3</u> times in SS_W ... <u>once</u> for each measurement point (since SS_{-D} counts each measurement point). To help keep track of this, I've added a column in the VARIANCE COMPONENTS table on page 8 called "# data in SS". This tells me how many times I over-counted the wafer-to-wafer squared deviations in the SS-W sum. Thus, the "observed wafer to wafer variance" is $MS\_W/3 = 0.0029$, because

③ $$SS\_W = M \sum_{i=1}^{W} (\bar{w}_i - \bar{p})^2$$

$$MS\_W = \frac{SS\_W}{W-1} = \frac{1}{W-1} \cdot M \cdot \sum_{i=1}^{W} (\bar{w}_i - \bar{p})^2$$

but $$S_{\bar{W}}^2 = \frac{1}{W-1} \sum_{i=1}^{W} (\bar{w}_i - \bar{p})^2, \quad \text{so} \quad S_{\bar{w}}^2 = \frac{MS\_W}{M} = 0.0035$$

④ Okay, now I have S^2_M as my estimate for σ^2_M, and $S^2_{\bar{W}}$ as estimate for $\sigma^2_{\bar{w}}$. Now I can recognize that my "true" wafer to wafer variance is what I observed, less the inflation due to sampling M sites:

$$S^2_{\bar{W}} = S^2_W + \frac{S^2_M}{M}$$

$$\Rightarrow \quad S^2_W = S^2_{\bar{W}} - \frac{S^2_M}{M} = 0.0035 - \frac{0.0018}{3} = 0.0029$$

⑤ Finally, now that we have site-to-site & wafer-to-wafer variances, the total variance is just

$$S^2_T = S^2_W + S^2_M = 0.0018 + 0.0029 = 0.0048$$

Part C INTERVAL ESTIMATES on Variance Components

For estimating confidence intervals on variance components, I resort to the basic formula

$$\frac{(N-1)S^2}{\chi^2_{\alpha/2, N-1}} \le \sigma^2 \le \frac{(N-1)S^2}{\chi^2_{1-\alpha/2, N-1}}$$

where $N-1$ is the degrees of freedom in S^2 estimate.

I then simply use the d.o.f. from the MS estimate in the anova table for the respective variance. I think this probably undercounts the d.o.f. somewhat (given we're also subtracting off nested elements) and so results in somewhat broader intervals than those shown in Dram.

Site-to-site variance: $\alpha = 0.05$

$$\frac{8\,(0.0018)}{\chi^2_{0.025,8}} \le \sigma_M^2 \le \frac{8\,(0.0018)}{\chi^2_{1-0.025,8}}$$

$$0.000843 \le \sigma_M^2 \le 0.006784$$

Similarly, for

wafer-to-wafer variance:

$$0.000941 = \frac{3\,(0.0029)}{\chi^2_{0.025,3}} \le \sigma_W^2 \le \frac{3\,(0.0029)}{\chi^2_{1-0.025,3}} = 0.040768$$

Part d What would the variance $\sigma_{\bar{W}}^2$ be if one sampled 5 rather than 3 sites per wafer,

- The observed 5 wafer average would be

$$\sigma_{\bar{W}}^2 = \sigma_W^2 + \frac{\sigma_M^2}{5} = 0.0029 + \frac{0.0018}{5} = 0.00326$$

$\boxed{\text{Problem 4}}$ Problem 4 in Drain, pg. 222-223

- Now we have two measurements nested in each wafer, two wafers in each lot, and three lots

(Part a) Point estimates for variance components.

This problem is the most important one for understanding nested variance components. With only two levels in problem 3, it is easy to think you understand; the three levels in problem 4 really show what needs to be done.

The hardest thing (I think) is getting the degrees of freedom right in building the ANOVA, and then reconciling this with the observed and estimated variances.

Note that we should <u>NOT</u> do a "factor" based anova where we might treat the lot #, wafer #, or site # as factors. Here we are not looking for <u>fixed</u> effects (e.g. the difference between being wafer #1 vs. wafer #2). Instead, we are looking for random variation in a nested structure.

Here's our statistical model:

$$x_{ijk} = \bar{y} + N(0, \sigma_L^2) + N(0, \sigma_W^2) + N(0, \sigma_M^2)$$

where i = counter for lot, $i = 1 \ldots L$, $L = 3$
j = counter for wafer, $j = 1 \ldots W$, $W = 2$
k = counter for site, $k = 1 \ldots M$, $M = 2$

We have some formulas from Drain. Unfortunately these do not make clear the degrees of freedom to use in F statistics, or in the χ^2 estimates of variance intervals

$$s_M^2 = \frac{1}{LW} \sum_{i=1}^{L} \sum_{j=1}^{W} \sum_{k=1}^{M} \frac{(x_{ijk} - \bar{x}_{ij\cdot})^2}{M-1}$$

where $\bar{x}_{ij\cdot}$ is the average over the M sites for lot i and wafer j

$$s^2_{\bar{W}} = \frac{1}{L} \sum_{i=1}^{L} \sum_{j=1}^{W} \frac{(\bar{x}_{ij\cdot} - \bar{x}_{i\cdot\cdot})^2}{W-1}$$ where $\bar{x}_{i\cdot\cdot}$ is the average for lot i

and

$$s^2_{\bar{L}} = \sum_{i=1}^{L} \frac{(\bar{x}_{i\cdot\cdot} - \bar{x}_{\cdot\cdot\cdot})^2}{L-1}$$ where $\bar{x}_{\cdot\cdot\cdot}$ is the grand mean.

Then one can "unwrap" nested variances

$$s^2_W = s^2_{\bar{W}} - \frac{s^2_M}{M} \quad \text{and then}$$

$$s^2_L = s^2_{\bar{L}} - \frac{s_W^2}{W} - \frac{s^2_M}{MW}$$

As mentioned, these work okay but do not clarify how to build the ANOVA table, or what degrees of freedom (d.o.f.) to use in ANOVA or variance intervals. To see this, we can decompose each measurement into constituent parts as deviations from local means, and then look at sum of squared deviations (SS's):

$$x_{ijk} = \underbrace{\bar{x}}_{\text{grand mean}} + \underbrace{(\bar{x}_{i\cdot\cdot} - \bar{x})}_{\text{deviation of this lot from grand mean}} + \underbrace{(\bar{x}_{ij\cdot} - \bar{x}_{i\cdot\cdot})}_{\text{deviation of this wafer from this lot mean}} + \underbrace{(x_{ijk} - \bar{x}_{ij\cdot})}_{\text{deviation of this site from this wafer's mean}}$$

or, in "corrected" or average form:

$$x_{ijk} - \bar{x} = (\bar{x}_{i\cdot\cdot} - \bar{x}) + (\bar{x}_{ij\cdot} - \bar{x}_{i\cdot\cdot}) + (x_{ijk} - \bar{x}_{ij\cdot})$$

$$\sum_i^L \sum_j^W \sum_k^M (x_{ijk} - \bar{x})^2 = MW \sum_i^L (\bar{x}_{i\cdot\cdot} - \bar{x})^2 + M \sum_i^L \sum_j^W (\bar{x}_{ij\cdot} - \bar{x}_{i\cdot\cdot})^2 + \sum_i^L \sum_j^W \sum_k^M (x_{ijk} - \bar{x}_{ij\cdot})^2$$

$$SS'_D = SS_L + SS_W + SS_M$$

with d.o.f.:

$$LWM - 1 = L - 1 + LW - L + LWM - LW$$

$$= (L-1) + L(W-1) + LW(M-1)$$

We also need to recall the # of times each SS term is "counted", as we "undo" this multiple counting later on in variance estimation. These are clear in the above:

in SS: 1 (in SS_D) $M \cdot W$ (in SS'_L) M (in SS'_W) 1 (in SS_M)

Now we are ready to build the SS estimates, then the MS (mean squares) using the appropriate d.o.f., and finally the variance components estimates.

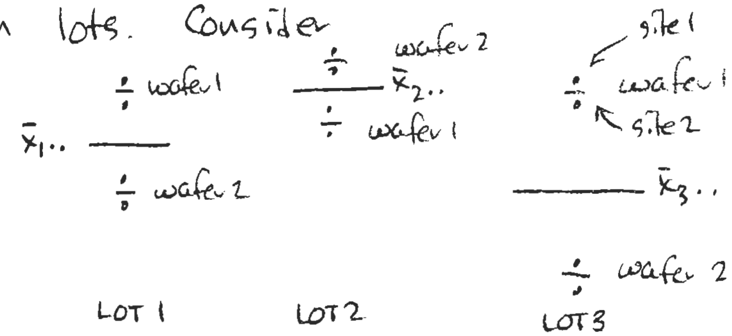
The spreadsheet on pg. 16 details these calculations. Some notes and comments (keyed to the spreadsheet):

① SITE to SITE variance. This is relatively easy... its just the

$$\frac{SS_M}{\nu_M} = \frac{0.0377}{LW(M-1)} = \frac{0.0377}{6} = \boxed{0.00628 = \hat{\sigma}^2_M} \text{ estimate}$$

so $s_M = 0.079$ which is about 1.8 % of our $\bar{x}$.

② For estimating σ^2_W we proceed as in problem 3, but we have to be careful. Now the directly calculated variance of the observed W_ave's = 0.00373 is NOT the same as the $s_{\bar{W}}^2$. The reason is that our wafers are nested within lots. Consider



So if we just take the variance across all of our wafer averages, this clearly is also getting inflated by the lot to lot differences. This is why we need to do the more careful ANOVA approach, where we look for SS deviations around local means.

However, SS_W is not quite what we want, because in it each wafer average gets counted twice in SS_W because there are two (M) points in each wafer mean (we want this when doing anova as we are trying to explain SS_D which includes all data points). So, to get an estimate for $s^2_{\bar{W}}$ we want to "undo" the extra weighting added by the "# data in SS". This gives us the right value for

Drain, Problem 4

$\bar{x}_{ij\cdot} = \frac{1}{M}\sum_{k=1}^{M} x_{ijk}$ $\bar{x}_{i\cdot\cdot} = \frac{1}{W}\sum_{j=1}^{W} \bar{x}_{ij\cdot}$

	Lot	Wafer	Rho	W_ave	Lot_ave	S_D	S_W	S_L	S_M
	1	1	4.289	4.3325	4.3183	0.0010	0.0002	0.000004	0.0019
	1	1	4.376			0.0031			0.0019
	1	2	4.320	4.3040		0.0000	0.0002		0.0003
	1	2	4.288			0.0010			0.0003
	2	1	4.149	4.2265	4.2650	0.0293	0.0015	0.00305	0.0060
	2	1	4.304			0.0003			0.0060
	2	2	4.384	4.3035		0.0041	0.0015		0.0065
	2	2	4.223			0.0095			0.0065
	3	1	4.396	4.4135	4.3775	0.0057	0.0013	0.00328	0.0003
	3	1	4.431			0.0123			0.0003
	3	2	4.404	4.3415		0.0070	0.0013		0.0039
	3	2	4.279	$\bar{x} = \bar{x}_{\cdot\cdot\cdot}$		0.0017			0.0039 $\Rightarrow (x_{ijk} - \bar{x}_{ij\cdot})^2$
						$(x_{ijk} - \bar{x})^2$	$(\bar{x}_{ij\cdot} - \bar{x}_{i\cdot\cdot})^2$	$(\bar{x}_{i\cdot\cdot} - \bar{x})^2$	
		MEAN	4.32025	4.32025	4.32025				
		VAR	0.00681	0.00373	0.00317				
		STD	0.08255	0.06104	0.05628				
					SS_D	0.0750 = Σ column			
L=		3				SS_W	0.0119 = $W \cdot \Sigma$ column		
W=		2					SS_L	0.0253 = $W \cdot M \cdot \Sigma$ column	
M=		2						S_M	0.0377 = Σ column

ANOVA

Variation	d.o.f.	d.o.f.	SS	MS	F_value	Pr > F	Error Term Used (Denominator)
LOT	L-1	2	0.0253	0.01267	3.18691432	0.18105	WAFER
WAFER	L(W-1)	3	0.0119	0.00398	0.63273157	0.62044	SITE
SITE	LW(M-1)	6	0.0377	0.00628			
C TOTAL	LWM-1	11	0.0750	0.00681			

VARIANCE COMPONENTS

Variation Source	MS	# data in SS	# data in SS	Observed Variance	Estimated Variance	% Var
SITE	0.00628	1	1	0.00628	0.00628	79.74
WAFER	0.00398	M	2	0.00199	0.00000	0.00
LOT	0.01267	MW	4	0.00317	0.00160	20.26
TOTAL	0.00681		1	0.00681	0.00788	100.00

By observed variance, we mean the SS *without* the multiplication by the number of points in the SS.

Interval Estimates

	alpha:	0.05		
	LOWER	POINT	UPPER	d.o.f.
SITE (site to site)	0.00261	0.00628	0.03046	6
WAFER (wafer to wafer)	0.00000	0.00000	0.00000	3
LOT (lot to lot)	0.00043	0.00160	0.06306	2

$$s_{\bar{w}}^2 = \frac{\left(\frac{SS_W}{M}\right)}{d.o.f._W} = \frac{\frac{0.0119}{2}}{3} = 0.00199$$

④ Now we can "unwrap" to estimate true wafer variance:

$$s_W^2 = s_{\bar{w}}^2 - \frac{s_M^2}{M} = 0.00199 - \frac{0.00628}{2} < 0$$

This gives a value that is NEGATIVE. This is unphysical, so we set $\boxed{s_W^2 = 0}$ estimate.

Why does this happen? It is a natural outcome of sampling, particularly in cases where one of the variance components is truly small or insignificant, and when we don't have enough replicates to discern this small variance. In our case, $s_M^2 \gg s_W^2$, so it is essentially not possible to detect wafer to wafer variance in the face of the larger site to site variance.

⑤ Now we're ready for $s_{\bar{L}}^2$. Note that because lots are not nested inside anything else now our directly calculated variance in lot averages (from lot_ave column) of 0.00317 is the same as what we get by anova:

$$s_{\bar{L}}^2 = \frac{\left(\frac{SS_L}{MW}\right)}{d.o.f._L} = \frac{\frac{0.0253}{4}}{2} = 0.00317$$

So now we can unwrap to get s_L^2 without the "inflation" due to wafer and site variance:

$$s_L^2 = s_{\bar{L}}^2 - \frac{s_W^2}{W} - \frac{s_M^2}{W \cdot M}$$

$$s_L^2 = 0.00317 - 0 - \frac{0.00628}{4}$$

$$\boxed{s_L^2 = 0.0016}$$

(6) Although not asked for explicitly, it is worth looking at the ANOVA table to understand how the F ratios are formed in the nested variance case. Here again, we use the SS and appropriate d.o.f. to form mean squares MS. Now the F ratio can be done, but in each case, we form the ratio from one level of nesting to that just "inside" it.

For the wafer to wafer effect, we have

$$F_0 = \frac{\overbrace{MS_W}^{\text{w/. 3 dof}}}{\underbrace{MS_M}_{\text{w/. 6 dof}}} = 0.6327$$

Then we can look up $Pr(F > F_0)$ for $F_{3,6}$.

In this case we find $Pr = 0.62$, or 62% of the time, given the site-to-site variance we would observe our wafer to wafer variance as recorded <u>by chance alone</u>. Thus, wafer to wafer effect is NOT significant.

This is consistent with our assignment of $s^2_W = 0$.

(7) Similarly, we can now evaluate significance of lot to lot variance, in the face of wafer to wafer variance. Now

$$F_0 = \frac{\overbrace{MS_L}^{\text{w/. 2 dof}}}{\underbrace{MS_W}_{\text{w/. 3 dof}}} = 3.1869$$

In this case $Pr(F > F_0) = 0.18$, or about 81% confidence in the lot to lot variation being significant.

Actually, since we decided wafer to wafer variance was not significant, a better approach might be to recalculate by lumping all W & M values together, but we don't do this here.

Part a) To recap: point estimates

$$s^2_L = 0.0016$$
$$s^2_W = 0$$
$$s^2_M = 0.00628$$

Part b Compare estimates of variance for (observed) lot means under different sampling plans:

part i $M=6, \; W=2$

$$S_{\bar{L}}^2 = S_L^2 + \frac{S_W^2}{W} + \frac{S_M^2}{M \cdot W} = 0.0016 + \frac{0}{2} + \frac{0.00628}{6 \cdot 2} = 0.002$$

Part ii $M=2, \; W=3$

$$= 0.0016 + \frac{0}{3} + \frac{0.00628}{2 \cdot 3} = 0.0026$$

So clearly strategy (i) is more effective in this case at minimizing the observed variance in estimates of the lot averages.

Part c Costs \$5 to load a wafer
\$2 for each site measured

What is best sampling plan costing \$30 or less to estimate lot variability?

$$\text{Min } S_{\bar{L}}^2 = S_L^2 + \cancel{\frac{S_W^2}{W}}^{\to 0} + \frac{S_M^2}{MW} \quad \text{subject to } 5W + 2MW \leq 30$$

There are few enough possibilities we can enumate them:

W	M	cost
→ 1	12	29
2	5	30
3	2	27
4	1	28

In this case, we get the lowest variance when M·W is largest, or W = 1 & M = 12.

This makes sense, given that wafer to wafer variance is not significant in this problem.

Part d Find interval estimates for variance components.

We proceed as in problem 3, using the appropriate d.o.f. as went into each estimate of the corresponding S^2, e.g.

$$\frac{2(S_L^2)}{\chi^2_{\alpha/2,2}} \leq \sigma_L^2 \leq \frac{2(S_L^2)}{\chi^2_{1-\alpha/2,2}}$$

as reported on the spreadsheet of pg. 16