6.780 Semiconductor
Manufacturing

Spring, 2003 D. Boning

RESPONSE SURFACE MODELS

I. Factorial Design & Response Surface Models
  A. Factor effects → response functions
  B. 1st Order Models (with & without Interactions)
  C. 2nd Order Models - Center Points
  D. Fractional Factorials - Screening
  E. Process Optimization

II. Nitride Etch Case
  A. 2^4 design
  B. 2^4 + Center Points
  C. 2^{4-1} half fraction
  D. Central Composite Design

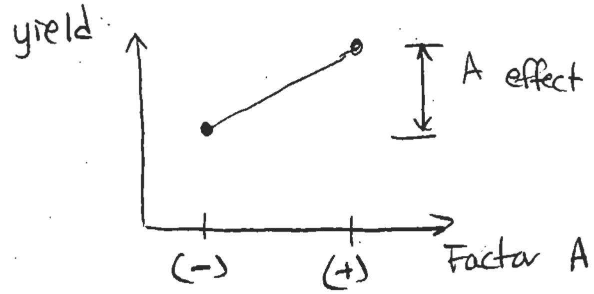
III. Regression Fundamentals
  A. 1 parameter model
  B. Polynomial model
  C. Mean-center model
  D. Multivariate model

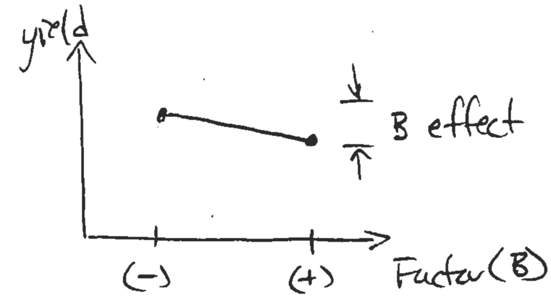
- parameter estimation
- experimental error & lack of fit
- variance in estimates, confidence intervals
- relationship to ANOVA tables

FROM FACTOR EFFECTS to RESPONSE SURFACES

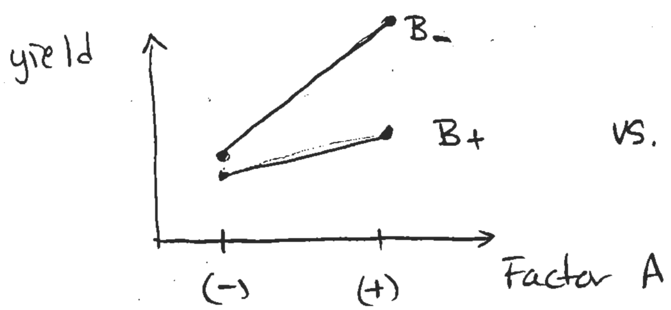
- In the basic factorial experiments discussed to this point, the factor levels could be nominal or continuous, and the effects on a response variable judged for
  - (a) significance
  - (b) relative contributions

e.g.

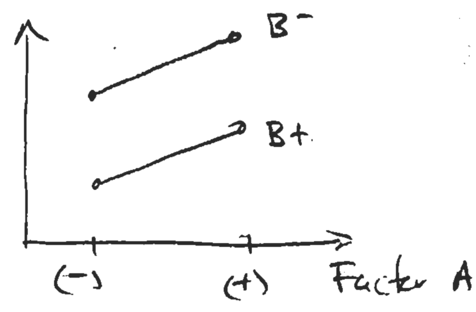




where we can also probe for interactions in the full factorial case:



vs.



SYNERGISTIC INTERACTION

NO INTERACTION (That is, effects of A and B are ADDITIVE)

MODEL:

$$\hat{y}_{ij} = \hat{\mu} + A_i + B_j + \varepsilon_{ij}$$

where we can predict $\hat{y}$ ONLY at discrete, prescribed i, j levels of factors A & B

- If factors can take on continuous values across the range of inquiry, we often prefer a RESPONSE SURFACE that captures a parametric dependency:

$$\hat{y} = \hat{\mu} + \beta_1 x_1 + \beta_2 x_2 + \varepsilon \quad , \quad x_i \in [x_{i,min} \, , \, x_{i,max}$$

APPLICATION of RSM METHODOLOGY: Process Optimization

- A typical goal of experimental design is process improvement. To accomplish this, we typically must

  - identify important factors ⇒ screening experiments
  - determine in which direction improvements lie ⇒ steepest ascent
  - move in best direction ⇒ exploration
  - improve model near optimum ⇒ confirming experiment

- A helpful observation is that
  - Far from optimum, response is often roughly linear
  - Near optimum, response is typically quadratic

APPROACH.

(1) Initial design (e.g. 2^k, 2^k + center points) for screening and constructing 1st order models

(2) Take steps in best direction

(3) Map response near the optimum in more detail, e.g. via 2nd order models (factorial + star)

Case Study: Plasma Etch Experiment (SST, May '87)

- Study nitride etch on single-wafer plasma etcher, using C_2F_6 as reactant. Consider first the nitride etch rate as response of interest.

- Approach: Initial screening & exploration will be done using a factorial 2^4 experiment.

  Replications? Assume 3 & 4 factor interactions are small, so use
  → single replication
  → combine 3 & 4 factor interactions to estimate error.

- Selection of Levels:

Design Factor				
Level	Gap *A* (cm)	Pressure *B* (m Torr)	C_2F_6 Flow *C* (SCCM)	Power *D* (W)
Low (−)	0.80	450	125	275
High (+)	1.20	550	200	325

- DESIGN and RESPONSES

Run	*A* (Gap)	*B* (Pressure)	*C* (C_2F_6 Flow)	*D* (Power)	Etch Rate (Å/min)
1	−1	−1	−1	−1	550
2	1	−1	−1	−1	669
3	−1	1	−1	−1	604
4	1	1	−1	−1	650
5	−1	−1	1	−1	633
6	1	−1	1	−1	642
7	−1	1	1	−1	601
8	1	1	1	−1	635
9	−1	−1	−1	1	1037
10	1	−1	−1	1	749
11	−1	1	−1	1	1052
12	1	1	−1	1	868
13	−1	−1	1	1	1075
14	1	−1	1	1	860
15	−1	1	1	1	1063
16	1	1	1	1	729

- Qualitative Examination:
  (1) Data Values in Exp. Space (2) Interaction Plot

- Check Residuals!

  - No evidence of substantial deviation
  - Also check
    * residual vs. predicted
    * residual vs. each factor

- Quantitative Model: 1^{st} order / 1^{st} order w/. interaction MODEL

  - when factorial (or other orthogonal) experiments are used, the effect estimates provide the regression model coefficients!

$$\hat{y} = 776.0625 - \left(\frac{101.625}{2}\right) x_1 + \left(\frac{306.125}{2}\right) x_4 - \left(\frac{153.625}{2}\right) x_1 x_4$$

where x_1, x_4 are "CODED" or normalized to range $[-1, 1]$

$$x_1 = \frac{gap - 1.0}{0.2} = \frac{gap - (gap_{max} + gap_{min})/2}{(gap_{max} - gap_{min})/2}$$

$$x_4 = \frac{power - 300}{25}$$

- Table of Contrasts:

Run		A	B	AB	C	AC	BC	ABC	D	AD	BD	ABD	CD	ACD	BCD	ABCD
1	(1)	−	−	+	−	+	+	−	−	+	+	−	+	−	−	+
2	a	+	−	−	−	−	+	+	−	−	+	+	+	+	−	−
3	b	−	+	−	−	+	−	+	−	+	−	+	+	−	+	−
4	ab	+	+	+	−	−	−	−	−	−	−	−	+	+	+	+
5	c	−	−	+	+	−	−	+	−	+	+	−	−	+	+	−
6	ac	+	−	−	+	+	−	−	−	−	+	+	−	−	+	+
7	bc	−	+	−	+	−	+	−	−	+	−	+	−	+	−	+
8	abc	+	+	+	+	+	+	+	−	−	−	−	−	−	−	−
9	d	−	−	+	−	+	+	−	+	−	−	+	−	+	+	−
10	ad	+	−	−	−	−	+	+	+	+	−	−	−	−	+	+
11	bd	−	+	−	−	+	−	+	+	−	+	−	−	+	−	+
12	abd	+	+	+	−	−	−	−	+	+	+	+	−	−	−	−
13	cd	−	−	+	+	−	−	+	+	−	−	+	+	−	−	+
14	acd	+	−	−	+	+	−	−	+	+	−	−	+	+	−	−
15	bcd	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−
16	abcd	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

- Calculation of Effects

$$
\begin{aligned}
A &= \tfrac{1}{8}[a + ab + ac + abc + ad + abd + acd + abcd - (1) - b \\
&\quad - c - d - bc - bd - cd - bcd] \\
&= \tfrac{1}{8}[669 + 650 + 642 + 635 + 749 + 868 + 860 + 729 - 550 \\
&\quad - 604 - 633 - 601 - 1037 - 1052 - 1075 - 1063] \\
&= -101.625
\end{aligned}
$$

$A =$	-101.625	$AD =$	-153.625
$B =$	-1.625	$BD =$	-0.625
$AB =$	-7.875	$ABD =$	4.125
$C =$	7.375	$CD =$	-2.125
$AC =$	-24.875	$ACD =$	5.625
$BC =$	-43.875	$BCD =$	-25.375
$ABC =$	-15.625	$ABCD =$	-40.125
$D =$	306.125		

- Which effects are significant?

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
A	41,310.563	1	41,310.563	20.28
B	10.563	1	10.563	<1
C	217.563	1	217.563	<1
D	374,850.063	1	374,850.063	183.99
AB	248.063	1	248.063	<1
AC	2,475.063	1	2,475.063	1.21
AD	94,402.563	1	99,402.563	48.79
BC	7,700.063	1	7,700.063	3.78
BD	1.563	1	1.563	<1
CD	18.063	1	18.063	<1
Error	10,186.815	5	2,037.363	
Total	531,420.938	15		

FACTORIAL + CENTER POINTS

- An important issue in the factorial design is that we only have a very limited assessment or representation for curvature in our space (i.e. via interaction terms)

W/. INTERACTION TERM

- Approach: ADD CENTER POINTS TO DESIGN
  - A relatively inexpensive addition that provides much evaluation capability
  - Run n_c replicates at center
    (1) Center runs do not impact our simple effect estimates, so analysis remains balanced

    (2) Provides an independent estimate of experimental error (i.e. in addition to that from higher order interactions)

FACTORIAL + CENTER, cont'd

- Modeling: we can now construct a SECOND ORDER MODEL e.g. for $k=2$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \underbrace{\beta_{11} x_1^2 + \beta_{22} x_2^2}_{\text{Quadratic terms}}$$

- Test: we can also explicitly check to determine if the quadratic terms are significant

$$H_0: \beta_{11} = \beta_{22} = 0$$

$$S_Q \triangleq \text{Sum of squares due to pure quadratic}$$
$$= \frac{n_F n_C (\bar{y}_F - \bar{y}_C)^2}{n_F + n_C}$$

where n_F = # factorial points
n_C = # center points
$\bar{y}_F, \bar{y}_C$ = mean for factorial, center points

$$s^2_Q = \frac{S_Q}{n_F + n_C - 2} \triangleq \text{Mean square quadratics}$$

Can now check $\dfrac{s^2_Q}{s^2_R}$ for significance

Etch Example, cont'd : Add Center Points

- Add $n_c = 4$ center points to unreplicated 2^4 design

Run	*A* (Gap)	*B* (Pressure)	*C* (C_2F_6 Flow)	*D* (Power)	Etch Rate (Å/min)
17	0	0	0	0	706
18	0	0	0	0	764
19	0	0	0	0	780
20	0	0	0	0	761

$\bar{y}_C = 752.75$

- $$SS_Q = \frac{16(4)(776.0625 - 752.75)^2}{16+4} = 1739.1 \quad \text{w/. 1 dof}$$

$$SS_E = \underset{S.S.}{\text{pure error}} = \sum_{i=17}^{20} (y_i - 752.75)^2 = 3122.7$$

So $$\frac{s^2_Q}{s^2_E} = \frac{1739.1/1}{3122.7/3} = 1.671 = F_{1,3} = t^2_{,3}$$

$$1.293 = t_{\alpha,3} \Rightarrow \alpha \approx 0.20$$

Not significant.

- An alternative often used is to utilize all residual information (SS_R), not just the "pure error" from center point, in evaluating significance.

Etch Example: w/ Center Points — ANOVA

ANOVA for Selected Model

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F VALUE	PROB > F
MODEL	521234.1	10	52123.4	31.33	0.001
CURVATURE	1739.1	1	1739.1	1.045	0.3365
RESIDUAL	13309.6	8	1663.7		
LACK OF FIT	10186.8	5	2037.4	1.957	0.3079
PURE ERROR	3122.7	3	1040.9		
COR TOTAL	536282.8	19			

ROOT MSE	40.7884	R-SQUARED	0.9751
DEP MEAN	771.4000	ADJ R-SQUARED	0.9440
C.V.	5.29%		

VARIABLE	COEFFICIENT ESTIMATE	DF	STANDARD ERROR	t FOR H0 COEFFICIENT=0	PROB > |t|
INTERCEPT	776.0625	1	10.1971		
A	-50.8125	1	10.1971	-4.983	0.0011
B	-0.8125	1	10.1971	-7.97E-02	0.9384
C	3.6875	1	10.1971	0.3616	0.7270
D	153.0625	1	10.1971	15.01	0.0001
AB	-3.9375	1	10.1971	-0.3861	0.7095
AC	-12.4375	1	10.1971	-1.220	0.2573
AD	-76.8125	1	10.1971	-7.533	0.0001
BC	-21.9375	1	10.1971	-2.151	0.0636
BD	-0.3125	1	10.1971	-3.06E-02	0.9763
CD	-1.0625	1	10.1971	-0.1042	0.9196
CENTER POINT	-23.3125	1	22.8014	-1.022	0.3365

FRACTIONAL FACTORIALS

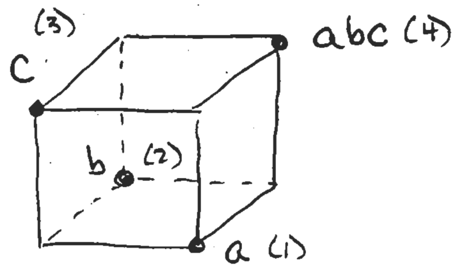
- As number of factors increases, the number of runs in a full factorial design rise dramatically, e.g. $2^5 = 32$
- In addition, we are able to sense all interactions, e.g. 2 factor, 3 factor, 4 factor, & even 5 factor!

⇒ If we only are concerned with main effects & low order interactions, we can manage with far fewer experiments

HALF FRACTION or 2^{k-1} DESIGNS

- Consider a 3 factor design (A, B, C). $2^k = 8$ runs; what if we choose our points to only make 4 runs?

i.e.



CONTRASTS

run	A	B	C
a (1)	+	−	−
b (2)	−	+	−
c (3)	−	−	+
abc (4)	+	+	+

⇒ Estimating main effects still easy, if we assume interactions don't exist.

- What if there actually are interactions? Can we estimate their effects?

	Factorial Effect							
Run	I	A	B	C	AB	AC	BC	ABC
a	+	+	−	−	−	−	+	+
b	+	−	+	−	−	+	−	+
c	+	−	−	+	+	−	−	+
abc	+	+	+	+	+	+	+	+

Confounding!

HALF FRACTION, Cont'd

- "GENERATOR"
  - Notice that the 2^{3-1} design is formed by selecting only those runs with + on ABC effect. This is the principal fraction of the design.
  - We could just as easily have selected our other half, or the alternate fraction:

	I	A	B	C	AB	AC	BC	ABC
ab	+	+	+	−	+	−	−	−
ac	+	+	−	+	−	+	−	−
bc	+	−	+	+	−	−	+	−
(1)	+	−	−	−	+	+	+	−

"I" or Identity run – positive for all four runs.

  - We can think of the fraction as "generated" by ABC, or I = ABC and I = −ABC for the two cases

- Our effect estimates thus have confounding in them:

$$\ell_A = A + BC \qquad \ell_A' = A - BC$$
$$\ell_B = B + AC \qquad \ell_B' = B - AC$$
$$\ell_C = C + AB \qquad \ell_C' = C - AB$$

NOTE: if we combine our two half fractions
(1) we have a full factorial design
(2) Main effects and interactions can be estimated separately, as expected:

$$A = \frac{\ell_A + \ell_A'}{2}, \qquad BC = \frac{\ell_A - \ell_A'}{2} \qquad \text{etc.}$$

That is, we can later run the other half fraction if we grow concerned about the interactions.

Etch Example, cont'd – HALF FRACTION DESIGN

- Suppose in our 4 factor experiment (A = gap, B = pressure, C = C_2F_6 flow, D = Power), we decided to use a 2^{4-1}?

- What is confounded with what?
  - Main effects are confounded with 3-way interactions:

$$I = ABCD \qquad \text{and similarly} \qquad B = ACD$$
$$A \cdot I = A \cdot ABCD \qquad C = ABD$$
$$A = BCD \qquad D = ABC$$

  - Two-way interactions are aliased with each other

$$AB \cdot I = AB \cdot ABCD \quad \text{and similarly} \quad AC = BD$$
$$AB = CD \qquad AD = BC$$

- Experimental Results:

The 2^{4-1} Design with Defining Relation I = ABCD

Run		A	B	C	$D = ABC$	Etch Rate
1	(1)	−	−	−	−	550
2	*ad*	+	−	−	+	749
3	*bd*	−	+	−	+	1052
4	*ab*	+	+	−	−	650
5	*cd*	−	−	+	+	1075
6	*ac*	+	−	+	−	642
7	*bc*	−	+	+	−	601
8	*abcd*	+	+	+	+	729

MAIN EFFECTS:

$\ell_A = -127.0 \quad \leftarrow$ GAP

$\ell_B = 4.0$

$\ell_C = 11.5$

$\ell_D = 290.5 \quad \leftarrow$ POWER

INTERACTIONS:

$\ell_{AB} = -10.0$

$\ell_{AC} = -25.5$

$\ell_{AD} = -197.5 \quad \leftarrow$ GAP & POWER (or pressure & flow, but that is unlikely)

- Comparable to full factorial, at least for screening purposes.

Etch Case, Cont'd : Process Optimization & Steepest Ascent

- The result of initial experiments is often a simple first order model $y = \beta_0 + \beta_1 x_1 + \dots \beta_k x_k + \varepsilon$

For our example, a 1st order model is

$$\hat{y} = 776.1 + 50.8\, x_1 + 153.1\, x_4 \quad \text{in Gap, Power}$$

(Note: because our design is orthogonal, we can simply drop the higher order terms without affecting our other estimates).

- GOAL: Etch rate of ~ 1100 - 1150 Å/min

$\Rightarrow$ Not achieved in our experimental space; must extrapolate and probe beyond $\Rightarrow$ steepest ascent

Etch Case, cont'd: 2^{nd} order Model

- Near optimum, surface is often curved (else it's not an optimum!) $\Rightarrow$ 2^{nd} order model typically needed

$$\hat{y} = \hat{\beta}_0 + \underbrace{\sum_{i=1}^{k} \hat{\beta}_i x_i}_{1^{st}\text{ order}} + \underbrace{\sum_{i=1}^{k} \hat{\beta}_{ii} x_i^2}_{\text{pure quadratic } 2^{nd}\text{ order}} + \underbrace{\sum_{i<j} \sum_{j=1}^{k} \hat{\beta}_{ij} x_i x_j}_{\text{interaction } 2^{nd}\text{ order terms}}$$

where $\hat{\beta}$ is a least squares estimate.

- Near optimum region: Gap = 1.2 cm, Power = 375 W

- CENTRAL COMPOSITE DESIGN
  - 2^2 factorial
  - 4 center points
  - 4 axial runs

NOTE: make distance of each point from center equal! $\rightarrow$ rotatable & the std. dev. of prediction is comparable at these points

- Experiment & Data:

Observation	Gap (cm)	Power (W)	Coded x_1	Variables x_4	Etch Rate y_1 (Å/m)	Uniformity y_2(Å)
1	1.000	350.0	− 1.000	− 1.000	1054.0	96.9
2	1.400	350.0	1.000	− 1.000	936.0	117.8
3	1.000	400.0	− 1.000	1.000	1179.0	114.4
4	1.400	400.0	1.000	1.000	1417.0	118.3
5	0.917	375.0	− 1.414	0.000	1049.0	102.6
6	1.483	375.0	1.414	0.000	1287.0	113.9
7	1.200	339.6	0.000	− 1.414	927.0	95.9
8	1.200	410.4	0.000	1.414	1345.0	125.4
9	1.200	375.0	0.000	0.000	1151.0	102.5
10	1.200	375.0	0.000	0.000	1150.0	104.5
11	1.200	375.0	0.000	0.000	1177.0	113.5
12	1.200	375.0	0.000	0.000	1196.0	108.4

Etch case, results

- RESPONSE SURFACE: Etch Rate
  - Found that 1st order + interaction terms fit the data adequately:

$$\hat{y}_{ER} = 1155.7 + 57.1\, x_1 + 149.7\, x_4 + 89\, x_1 x_4$$

- RESPONSE SURFACE: Nonuniformity
  - In addition to modeling etch rate, we're also often concerned with etch uniformity

uniformity $\triangleq$ std. dev. of remaining thickness across the wafer after the etch.

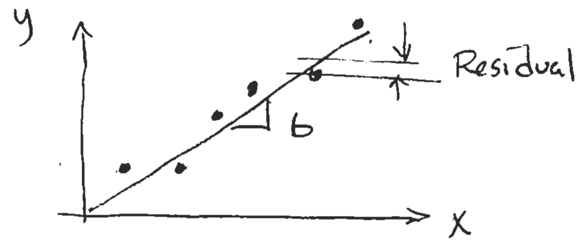
  - Required 2nd order model to fit:

$$\hat{y}_{NU} = 107.22 + 5.14\, x_1 + 7.50\, x_4 + 2.29\, x_4^2 - 4.33\, x_1 x_4$$

REGRESSION FUNDAMENTALS

- We use <u>least-squares</u> to estimate coefficients in typical response surface or regression models. A very brief overview of the ideas behind least-squares follows

① ONE-PARAMETER MODEL



$$y_i = \beta x_i + \epsilon_i, \quad i = 1, 2, \ldots n$$

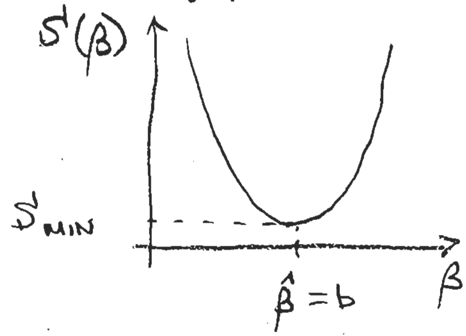
goal is to estimate β with "best" b

- How define "best"? → That b which minimizes sum of squared error between prediction and data

$$S = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \quad \text{where} \quad \hat{y}_i = \hat{\beta} x_i$$

$$S = \sum_{i=1}^{n} (y_i - \beta x_i)^2 \qquad \text{So} \quad S = f(\beta)\text{; find that } \beta \text{ which minimizes } S$$

E.g.



$$S_{MIN} = \sum_{i=1}^{n} (y_i - b x_i)^2 = S_R$$

RESIDUAL SUM OF SQUARES

- Least Squares estimation via Normal Equations
  - For linear problems, we need not calculate $S(\beta)$; rather, direct solution for b is possible
  - Vector of <u>residuals</u> with be <u>normal to</u> vector of x values at the least squares estimate:

$$\sum (y - \hat{y}) x = 0 \quad \text{or} \quad \sum (y - bx) x = 0$$

$$\sum xy = \sum b x^2$$

$$\Rightarrow \quad b = \frac{\sum xy}{\sum x^2}$$

REGRESSION, Cont'd

- Estimate of Experimental Error
  - Assuming model structure is adequate, estimate s^2 of σ^2 can be obtained:

$$s^2 = \frac{S_R}{n-1}$$

- Precision of Estimate: Variance in b

$$\hat{V}(b) = \frac{s^2}{\sum_i x_i^2} \qquad s.e.(b) = \sqrt{\hat{V}(b)}$$

Write

$$b \pm s.e.(b)$$

Why?

$$b = \left(\frac{x_1}{\Sigma x^2}\right) y_1 + \left(\frac{x_2}{\Sigma x^2}\right) y_2 + \cdots \left(\frac{x_n}{\Sigma x^2}\right) y_n$$

$$= a_1 y_1 + a_2 y_2 + \cdots a_n y_n$$

$$V(b) = (a_1^2 + a_2^2 + \cdots a_n^2)\,\sigma^2 \quad \text{assuming IID observations}$$

$$= \left[\left(\frac{x_1}{\Sigma x^2}\right)^2 + \cdots \left(\frac{x_n}{\Sigma x^2}\right)^2\right]\sigma^2$$

$$= \frac{\Sigma x^2}{(\Sigma x^2)^2}\,\sigma^2 = \frac{\sigma^2}{\Sigma x^2} \quad //$$

- Confidence Interval for β

$$t = \frac{b - \beta_*}{s.e.(b)}$$

So, to some desired $1-\alpha$ confidence limit

$$\beta = b \pm \left[t_{\alpha/2} \cdot s.e.(b)\right]$$

- Analysis of Variance: Test H_0: $\beta_* = b = 0$

$$\Sigma y_i^2 = \Sigma \hat{y}_i^2 + \Sigma (y_i - \hat{y}_i)^2$$

D.O.F. $\quad n \quad\quad p \quad\quad n-p$

p: # parameters estimated by least sq.

Regression Example

- Single variable, x_i chosen at random

observation number u	age x_u	dispersion y_u	estimated dispersion $\hat{y}_u = 0.50098x_u$	residual $y_u - \hat{y}_u$
1	8	6.16	4.0079	2.1521
2	22	9.88	11.0216	−1.1416
3	35	14.35	17.5344	−3.1844
4	40	24.06	20.0393	4.0207
5	57	30.34	28.5560	1.7840
6	73	32.17	36.5718	−4.4018
7	78	42.18	39.0767	3.1033
8	87	43.23	43.5855	−0.3555
9	98	48.76	49.0963	−0.3363
		$\sum y_u^2 = 8901.31$	$\sum \hat{y}_u^2 = 8836.64$	$\sum (y_u - \hat{y})^2 = 64.669$
		S_T =	S_M +	S_R

- Regression:

$$\hat{y} = bx = 0.501x$$

- Experimental Error Estimate

$$s^2 = \frac{S_R}{n-1} = \frac{64.67}{8} = 8.0837$$

$$s = \sqrt{s^2} = 2.842$$

- Precision of Estimate – s.e.(b)

$$t_{0.05/2),8} = 2.306$$

$$\hat{V}(b) = \frac{s^2}{\sum x_i^2} = \frac{8.0837}{35208} = 0.0002296$$

$$b \pm s.E.(b) = 0.501 \pm 0.015 \quad vs. \quad 0.501 \pm 0.035$$

So 95% C.I.

- ANOVA

source	sum of squares	degrees of freedom	mean square	
model	$S_M = 8836.64$	1	8836.64	ratio = 1094
residual	$S_R = 64.67$	8	8.08	
total	$S_T = 8901.31$	9		

$$F_{1,8} = t_8^2 = \left(\frac{b}{s.e.(b)}\right)^2 = 1094$$

... VAM

REGRESSION, Cont'd

- **LACK OF FIT vs. PURE ERROR**

  - when some runs have be genuinely replicated, we have the opportunity to decompose residual error contributions:

$$S_R = S_L + S_E$$

$$\text{Residuals} = \text{Lack of Fit} + \text{Pure Error} \qquad \text{or} \qquad S_L = S_R - S_E$$

  - TEST for lack of Fit: $\dfrac{s^2_L}{s^2_E} = F_{\nu_L, \nu_E}$

② Polynomial Regression

- We may believe that a higher order model structure applies. Polynomial forms are also linear in the coefficients and can be fit with least squares.

$$\eta = \beta_0 + \beta_1 x + \beta_2 x^2 \qquad \sim \text{curvature included}$$

- Example: Growth rate data

NOTE: different numbers of replicates at different points

Growth rate data

observation number	amount of supplement (grams) x	growth rate (coded units) y
1	10	73
2	10	78
3	15	85
4	20	90
5	20	91
6	25	87
7	25	86
8	25	91
9	30	75
10	35	65

(1) First Order Model $\hat{y} = 86.44 - 0.20x$

Analysis of variance for growth rate data: straight line model

source	sum of squares	degrees of freedom	mean square
model	$S_M = 67{,}428.6$ {mean 67,404.1; extra for linear 24.5}	2 {1; 1}	67,404.1; 24.5
residual {lack of fit; pure error}	$S_R = 686.4$ {$S_L = 659.40$; $S_E = 27.0$}	8 {4; 4}	85.8 {164.85; 6.75} ratio = 24.42
total	$S_T = 68{,}115.0$	10	

- Mean significant, linear term not
- Clear evidence of LACK OF FIT

(2) Second Order Model $\hat{y} = 35.66 + 5.26x - 0.128x^2$

Analysis of variance for growth rate data: quadratic model

source	sum of squares	degrees of freedom	mean square
model	$S_M = 68{,}071.8$ {mean 67,404.1; extra for linear 24.5; extra for quadratic 643.2}	3 {1; 1; 1}	67,404.1; 24.5; 643.2
residual	$S_R = 43.2$ {$S_L = 16.2$; $S_E = 27.0$}	7 {3; 4}	{5.40; 6.75} ratio = 0.80
total	$S_T = 68{,}115.0$	10	

- No lack of fit evidence
- Quadratic term significant

REGRESSION: Mean Centered Models

- MODEL FORM $\quad \eta = \alpha + \beta(x - \bar{x})$

est. by $\quad \hat{y} = a + b(x - \bar{x}) \quad \ldots \quad y_i \sim N(\eta_i, \sigma^2)$

Minimize $\quad S_R = \Sigma (y_i - \hat{y}_i)^2$ to estimate α & β

$$a = \bar{y}$$

$$b = \frac{\Sigma (x_i - \bar{x}) y_i}{\Sigma (x_i - \bar{x})^2} = \frac{\Sigma (x_i - \bar{x})(y_i - \bar{y})}{\Sigma (x_i - \bar{x})^2}$$

$E(a) = \alpha$ & $E[b] = \beta$ ... good estimators (unbiased)

$$Var(a) = Var\left[\frac{\Sigma y_i}{K}\right] = \frac{\sigma^2}{K} \quad \& \quad Var(b) = \frac{\sigma^2}{\Sigma (x_i - \bar{x})^2}$$ (MM variance)

Confidence Limits:

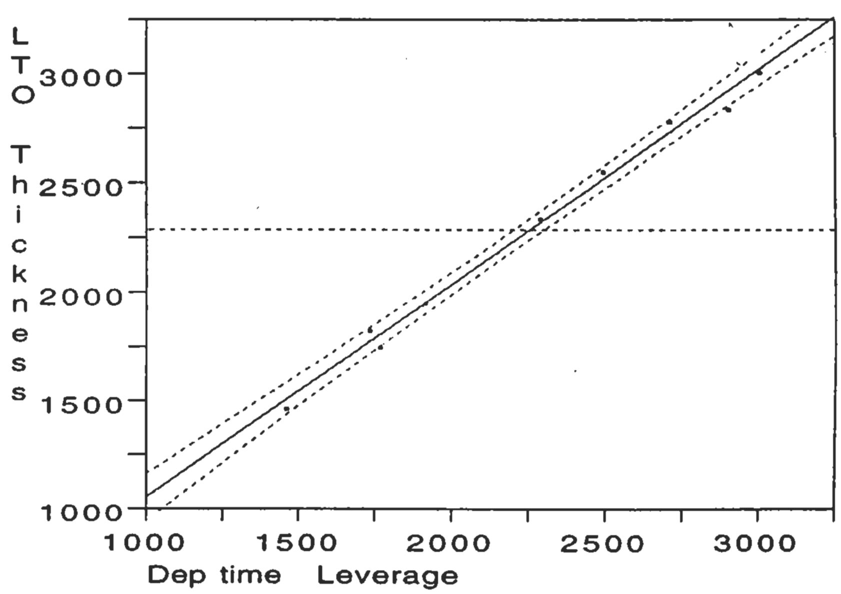
$$\hat{y}_i = \bar{y} + b(x_i - \bar{x})$$

$$Var(\hat{y}_i) = V(\bar{y}) + (x_i - \bar{x})^2 V(b)$$

$$V(\hat{y}_i) = \frac{s^2}{n} + \frac{s^2 (x_i - \bar{x})^2}{\Sigma (x - \bar{x})^2}$$ – worse as we move away from $\bar{x}$!

Confidence interval

$$\hat{y}_i \pm t_{\alpha/2} \sqrt{V(\hat{y}_i)}$$



MULTIVARIATE REGRESSION

- Response Function $\vec{\eta} = \vec{x}\vec{\beta}$

Normal equations become

$$\vec{x}^T(\vec{y} - \hat{\vec{y}}) = 0$$

$$\vec{x}^T(\vec{y} - \vec{x}\vec{b}) = 0$$

$$\vec{x}^T\vec{x}\,\vec{b} = \vec{x}^T\vec{y}$$

$$\Rightarrow \quad \vec{b} = \underbrace{[\vec{x}^T\vec{x}]^{-1}\vec{x}^T}_{\text{GENERALIZED INVERSE}}\vec{y}$$

$Var(\vec{b}) = [\vec{x}^T\vec{x}]^{-1}\,\sigma^2$ if σ^2 known

- JOINT Confidence Intervals

SUM OF SQUARES CONTOURS

- Pick S_α for some desired confidence

$$S_\alpha = S_R\left[1 + \frac{p}{n-p}F_\alpha(p, n-p\right.$$

n = # points
p = # model coeffs

- Estimates negatively correlated

individual 90% confidence intervals