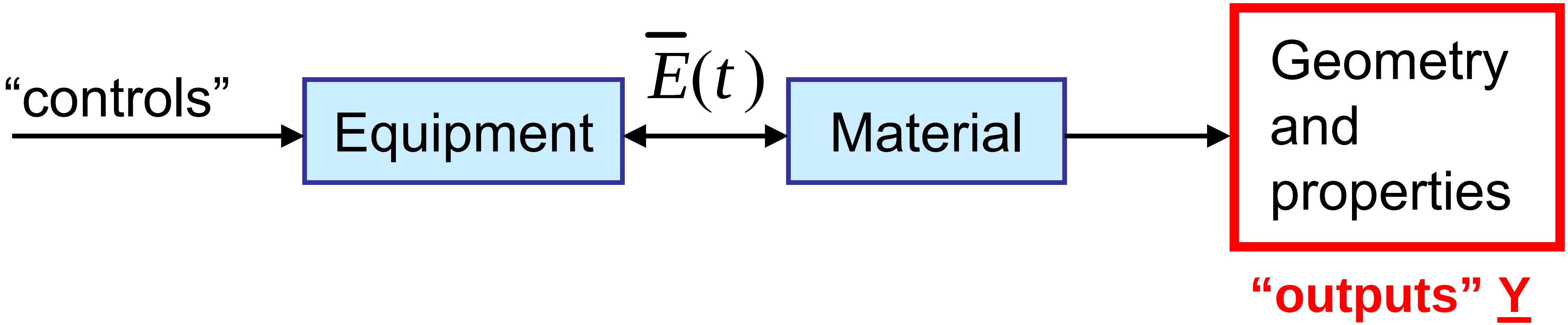
2.830J / 6.780J / ESD.63J Control of Manufacturing Processes (SMA 6303)
Spring 2008

MIT 2.830/6.780/ESD.63

Control of Manufacturing Processes

Introduction to Analysis of Variance: a tool for assessing input-output relationships

Have focused so far on interpreting output

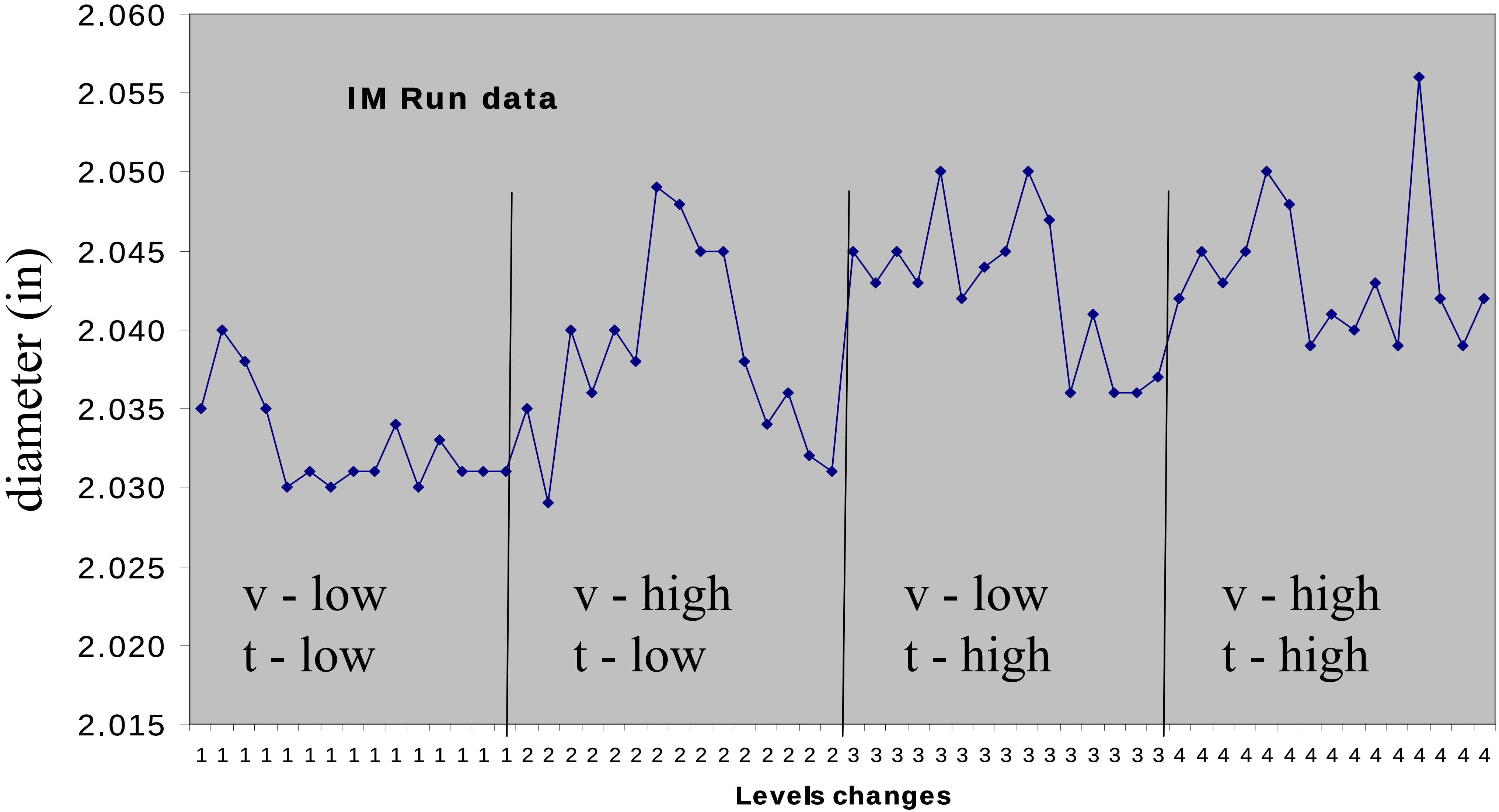


Review of tools for interpreting outputs

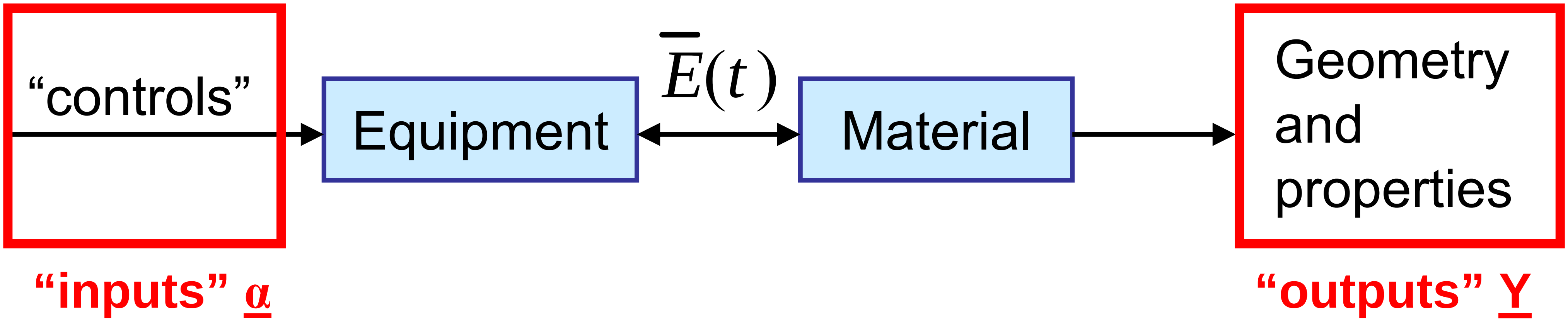
Tool
t, F tests
control charts, cusum, EWMA *etc*
χ^2, T^2 charts

Also talked about yield modeling and process capability, but need ways of *modeling* and thus *improving* processes

Injection molding data



Want to start relating input(s) to output(s)



$$\underline{Y} = \Phi(\underline{\alpha})$$

$$\underline{\alpha} \equiv \text{process } parameters$$

$$\Delta Y = \frac{\partial Y}{\partial \alpha}\Delta\alpha + \frac{\partial Y}{\partial u}\Delta u$$

What is our goal?

- **Developing a process model**
  - Relating inputs and disturbances to outputs
  - Determining significance of the input effect
    - Does it really matter?
- **Process optimization**
  - Max (Cpk) or Min (QLF)
  - Models for mean shifting
  - Models for variance reduction

Empirical Modeling

- What is the objective?
- What is the output?
- What are the input(s)?
- What do we want to vary?
- What model form should we use?
  - $\mathbf{Y}=\Phi(\boldsymbol{\alpha},\mathbf{u})$ is not specific!
- How many data can we take?

First step: determining which inputs matter

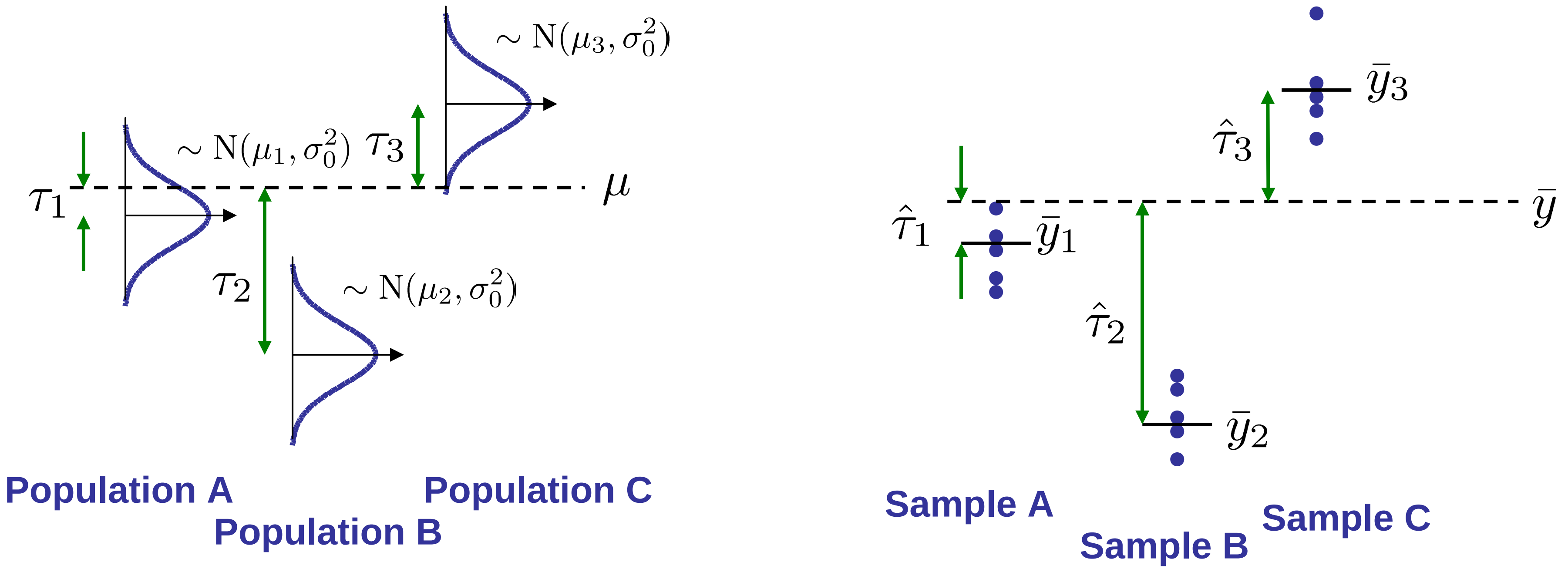
Tool	# inputs	Levels per input	# samples	# outputs
t, F tests	?	? (2?)	2	1
control charts, cusum	?	? (2?)	many	1
χ^2, T^2 charts	?	? (2?)	many	2
Analysis of variance	≥ 1	≥ 2	≥ 2	1

Agenda

1. **Comparison of treatments (one variable)**
   - Fixed effects model
   - Analysis of Variance (ANOVA) technique
   - Example

2. **Multivariate analysis of variance**
   - Model forms
   - MANOVA technique

Comparison of Treatments



- Consider multiple conditions (treatments, settings for some variable)
  - There is an overall mean μ and real "effects" or deltas between conditions τ_i.
  - We observe samples at each condition of interest

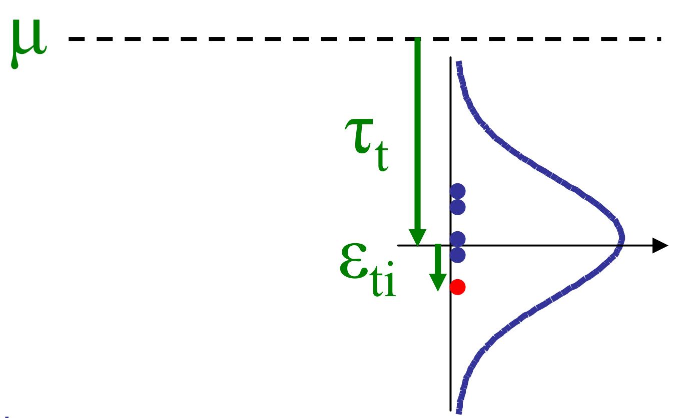
- Key question: are the ***observed*** differences in mean "significant"?
  - Typical assumption (should be checked): the underlying variances are all the same – usually an unknown value (σ_0^2)

ANOVA – Fixed effects model

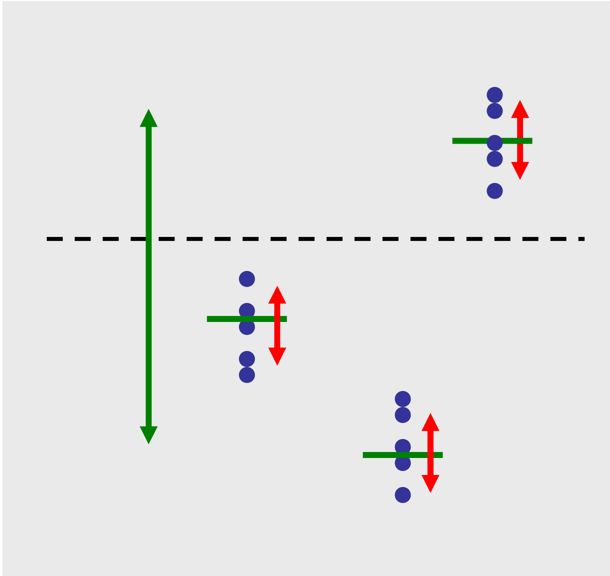
- The ANOVA approach assumes a simple mathematical model:

$$\begin{aligned} y_{ti} &= \mu + \tau_t + \epsilon_{ti} \\ &= \mu_t + \epsilon_{ti} \end{aligned}$$

- Where μ_t is the treatment mean (for treatment type t)
- And τ_t is the treatment effect
- With ε_{ti} being zero mean normal residuals $\sim N(0,\sigma_0^2)$



Steps/Issues in Analysis of Variance



1. **Within-group variation**
   – Estimate underlying population variance

2. **Between-group variation**
   – Estimate group to group variance

3. **Compare the two estimates of variance**
   – If there is a difference between the different treatments, then the between group variation estimate will be ***inflated*** compared to the within group estimate
   – We will be able to establish confidence in whether or not observed differences between treatments are significant

Hint: we'll be using *F* tests to look at ratios of variances

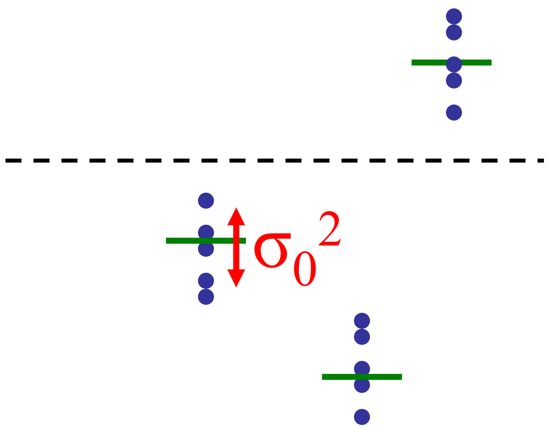
(1) Within Group Variation

- Assume that each group is normally distributed and shares a common variance σ_0^2
- SS_t = sum of square deviations within t^th group (there are k groups)

$$SS_t = \sum_{i=1}^{n_t} (y_{ti} - \bar{y}_t)^2 \text{ where } n_t \text{ is number of samples in treatment } t$$

- Estimate of within group variance in t^th group (just variance formula)

$$s_t^2 = SS_t / \nu_t = \frac{SS_t}{n_t - 1} \quad \text{where } \nu_t \text{ is d.o.f. in treatment } t$$

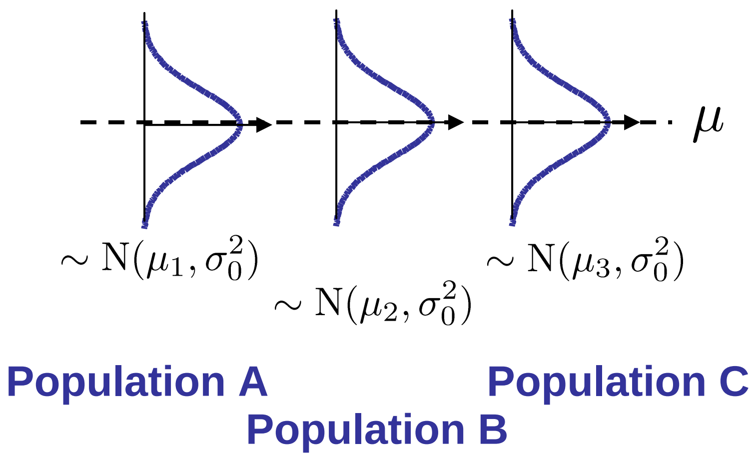


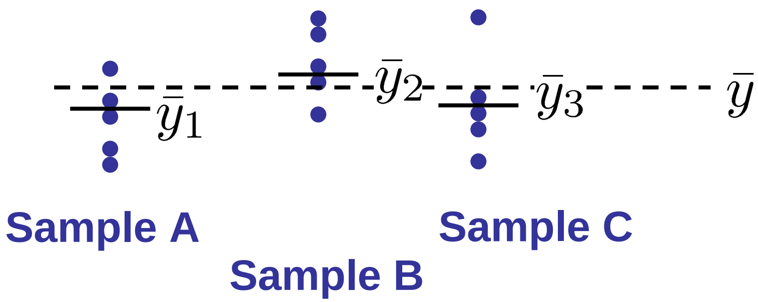
(2) Between Group Variation

- We will be testing hypothesis $\mu_1 = \mu_2 = \ldots = \mu_k$
- If all the means are in fact equal, then a 2nd estimate of σ^2 could be formed based on the observed differences between group means:

$$s_T^2 = \frac{\sum_{t=1}^{k} n_t(\bar{y}_t - \bar{y})^2}{k-1} = \frac{SS_T}{k-1}$$

where n_t is number of samples in treatment t, and k is the number of different treatments





(3) Compare Variance Estimates

- We now have two different possibilities for s_T^2, depending on whether the observed sample mean differences are "real" or are just occurring by chance (by sampling)
- Use *F* statistic to see if the ratios of these variances are likely to have occurred by chance!
- Formal test for significance:

$$\text{Reject } H_0 \; (H_0 : \text{no mean difference})$$
$$\text{if } \frac{s_T^2}{s_R^2} \text{ is significantly greater than 1.}$$

(4) Compute Significance Level

- Calculate observed *F* ratio (with appropriate degrees of freedom in numerator and denominator)
- Use *F* distribution to find how likely a ratio this large is to have occurred by chance alone
  - This is our "significance level"
  - Define observed ratio: $F_0 = s_T^2 / s_R^2$
  - If $F_0 > F_{\alpha, k-1, N-k}$
    then we say that the mean differences or treatment effects are significant to (1-α)100% confidence or better

(5) Variance Due to Treatment Effects

- We also want to estimate the sum of squared *deviations from the grand mean* among all samples:

$$SS_D = \sum_{t=1}^{k} \sum_{i=1}^{n_t} (y_{ti} - \bar{y})^2$$

$$s_D^2 = SS_D/\nu_D = \frac{SS_D}{N-1} = MS_D$$

where N is the total number of measurements

(6) Results: The ANOVA Table

source of variation	sum of squares	degrees of freedom	mean square	F_0	$\Pr(F_0)$
Between treatments	SS_T	$k-1$	$s_T^2 = \frac{SS_T}{k-1}$	$\frac{s_T^2}{s_R^2}$	table
Within treatments	SS_R	$N-k$	$s_R^2 = \frac{SS_R}{N-k}$		
Total about the grand average	SS_D	$N-1$	$s_D^2 = \frac{SS_D}{N-1}$		

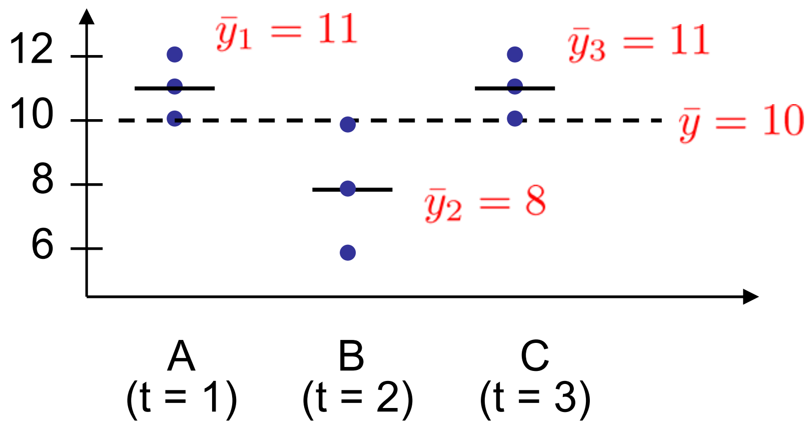
Also referred to as "residual" SS

$SS_D = SS_T + SS_R$

$\nu_D = \nu_T + \nu_R$

Example: Anova

A	B	C
11	10	12
10	8	10
12	6	11



Excel: Data Analysis, One-Variation Anova

Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
A	3	33	11	1		
B	3	24	8	4		
C	3	33	11	1		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	18	2	9	4.5	0.064	5.14
Within Groups	12	6	2			
Total	30	8				

$$F = \frac{S_T^2}{S_R^2} = \frac{9}{2} = 4.5$$

$$F_{0.05,2,6} = 5.14$$

$$F_{0.10,2,6} = 3.46$$

$$\begin{aligned} SS_1 &= (12-11)^2 + (11-11)^2 + (10-11)^2 = 2 \\ SS_2 &= 2^2 + 0^2 + 2^2 = 8 \\ SS_3 &= 1^2 + 0^2 + 1^{=}2 \end{aligned}$$

$$\begin{aligned} s_1^2 &= MS_1 = SS_1/2 = 2/2 = 1 \\ s_2^2 &= MS_2 = 8/2 = 4 \\ s_3^2 &= MS_3 = 2/2 = 1 \end{aligned}$$

$$s_R^2 = \frac{SS_1 + SS_2 + SS_3}{N - k} = \frac{12}{6} = 2$$

$$\begin{aligned} s_T^2 &= \frac{3(11-10)^2 + 3(8-10)^2 + 3(11-10)^2}{3-1} \\ &= \frac{SS_T}{\nu_T} = \frac{18}{2} = 9 \end{aligned}$$

ANOVA – residuals assumed ~N(0, σ_0^2) for every treatment

- Checks
  - Plot residuals against time order
  - Examine distribution of residuals: should be IID, Normal
  - Plot residuals vs. estimates
  - Plot residuals vs. other variables of interest

MANOVA – Two Dependencies

- Can extend to two (or more) variables of interest. MANOVA assumes a mathematical model, again simply capturing the means (or treatment offsets) for each discrete variable level:

$$\begin{array}{rcccccccc} & y_{tqi} & = & \mu & + & \tau_t & + & \beta_q & + & \epsilon_{tqi} \\ \text{^ indicates } \textbf{estimates:} & \hat{y}_{tq} & = & \hat{\mu} & + & \hat{\tau}_t & + & \hat{\beta}_q & & \end{array}$$

$$\begin{array}{rcccccc} \#\ \text{model coeffs} & = & 1 & + & k & + & n \\ & & \uparrow & & \uparrow & & \uparrow \\ \#\ \text{independent model coeffs} & = & 1 & + & (k-1) & + & (n-1) \end{array}$$

Recall that our $\hat{\tau}_t$ are *not* all independent model coefficients, because $\sum \tau_t = 0$. Thus we really only have $k-1$ independent model coeffs, or $\nu_t = k-1$.

- Assumes that the effects from the two variables are ***additive***

MANOVA – Two Factors with Interactions

- May be interaction: not simply additive – effects may depend synergistically on both factors:

IID, $\sim N(0,\sigma^2)$

$$y_{tqi} = \mu_{tq} + \epsilon_{tqi}$$

An effect that depends on both t & q factors simultaneously

t = first factor = 1,2, … k (k = # levels of first factor)
q = second factor = 1,2, … n (n = # levels of second factor)
i = replication = 1,2, … m (m = # replications at t, q^{th} combination of factor levels)

- Can split out the model more explicitly…

$$\begin{aligned} y_{tqi} &= \mu + \tau_t + \beta_q + \omega_{tq} + \epsilon_{tqi} \\ \text{Estimate by: } \hat{y}_{tq} &= \bar{y} + (\bar{y}_t - \bar{y}) + (\bar{y}_q - \bar{y}) + (\bar{y}_{tq} - \bar{y}_t - \bar{y}_q + \bar{y}) \end{aligned}$$

$$\begin{aligned} \omega_{tq} &= \text{interaction effects} = (\bar{y}_{tq} - \bar{y}_t - \bar{y}_q + \bar{y}) \\ \tau_t, \beta_q &= \text{main effects} \end{aligned}$$

MANOVA – Two Factors with Interactions

$$s_T^{\ 2} = \frac{\sum_{t=1}^{k} m_t n_t \left(\overline{y}_t - \overline{y}\right)^2}{k-1}$$

$$s_B^{\ 2} = \frac{\sum_{q=1}^{n} m_q k_q \left(\overline{y}_q - \overline{y}\right)^2}{n-1}$$

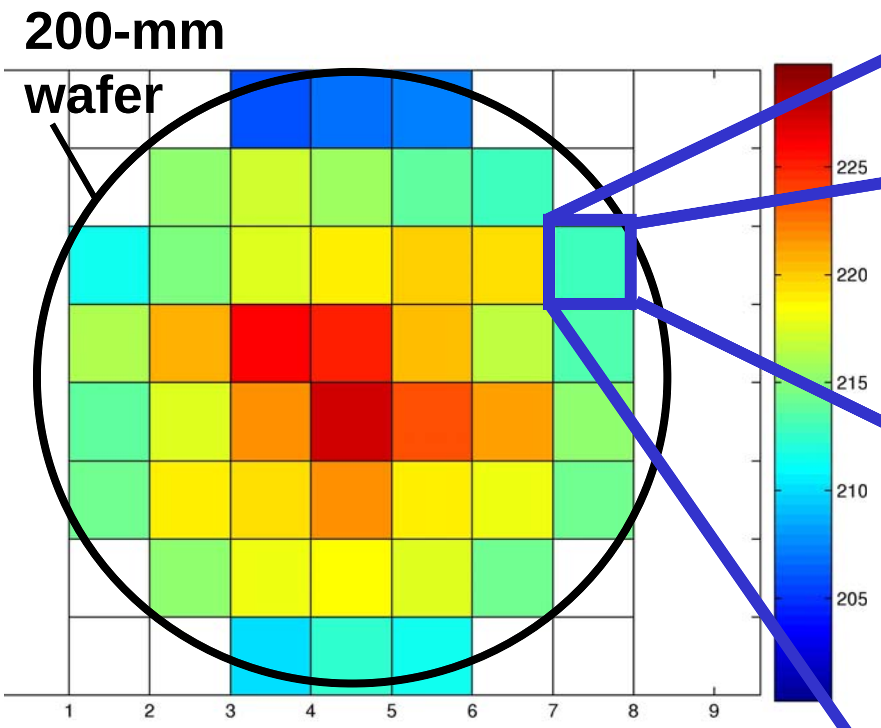
$$s_I^{\ 2} = \frac{\sum_{q=1}^{n} \sum_{t=1}^{k} m_{tq} \left(\overline{y}_{tq} - \overline{y}_t - \overline{y}_q + \overline{y}\right)^2}{(k-1)(n-1)}$$

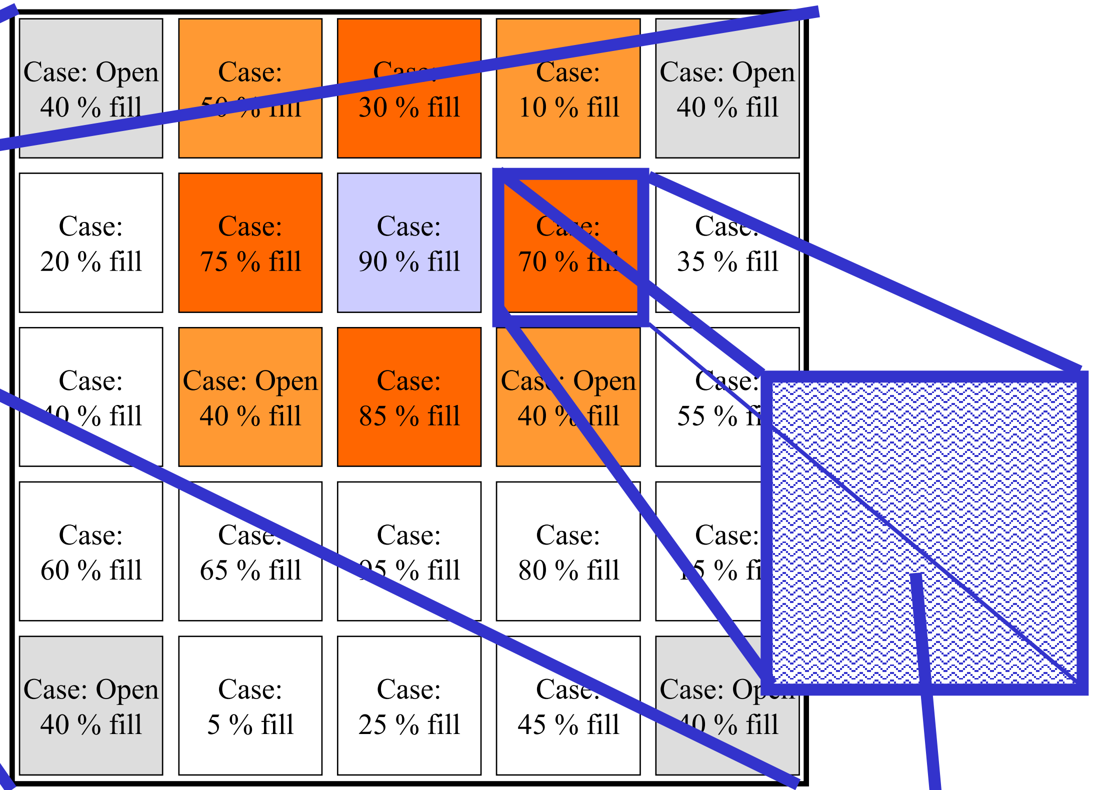
$$s_E^{\ 2} = \frac{\sum_{q=1}^{n} \sum_{i=1}^{m} \sum_{t=1}^{k} \left(y_{tqi} - \overline{y}_{tq}\right)^2}{nk(m-1)}$$

MANOVA Table – Two Way with Interactions

source of variation	sum of squares	degrees of freedom	mean square	F_0	$\Pr(F_0)$
Between levels of factor 1 (T)	SS_T	$k-1$	s_T^2	s_T^2/s_E^2	table
Between levels of factor 2 (B)	SS_B	$n-1$	s_B^2	s_B^2/s_E^2	table
Interaction	SS_I	$(k-1)(n-1)$	s_I^2	s_I^2/s_E^2	table
Within Groups (Error)	SS_E	$nk(m-1)$	s_E^2		
Total about the grand average	SS_D	$nkm-1$			

Example: plasma metal etch nonuniformity (lateral etch)

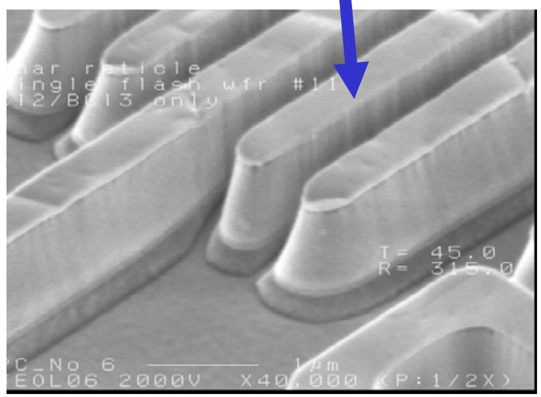




Average resistances of all structures in a 'flash' (Ohms)

Each flash has many copies of a feature set, with different 'padding' densities

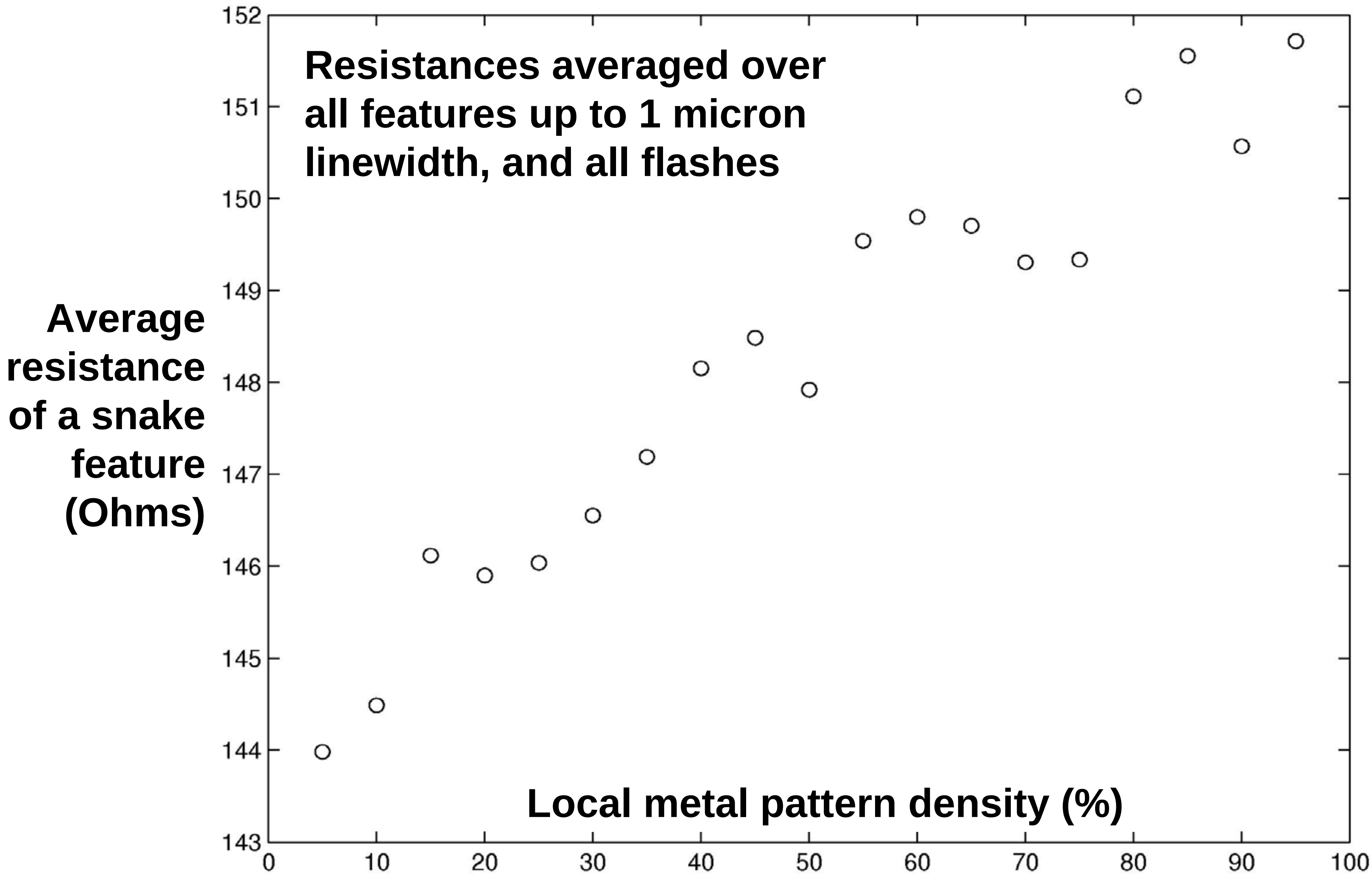
Each feature set has a range of line/space widths; can be electrically probed



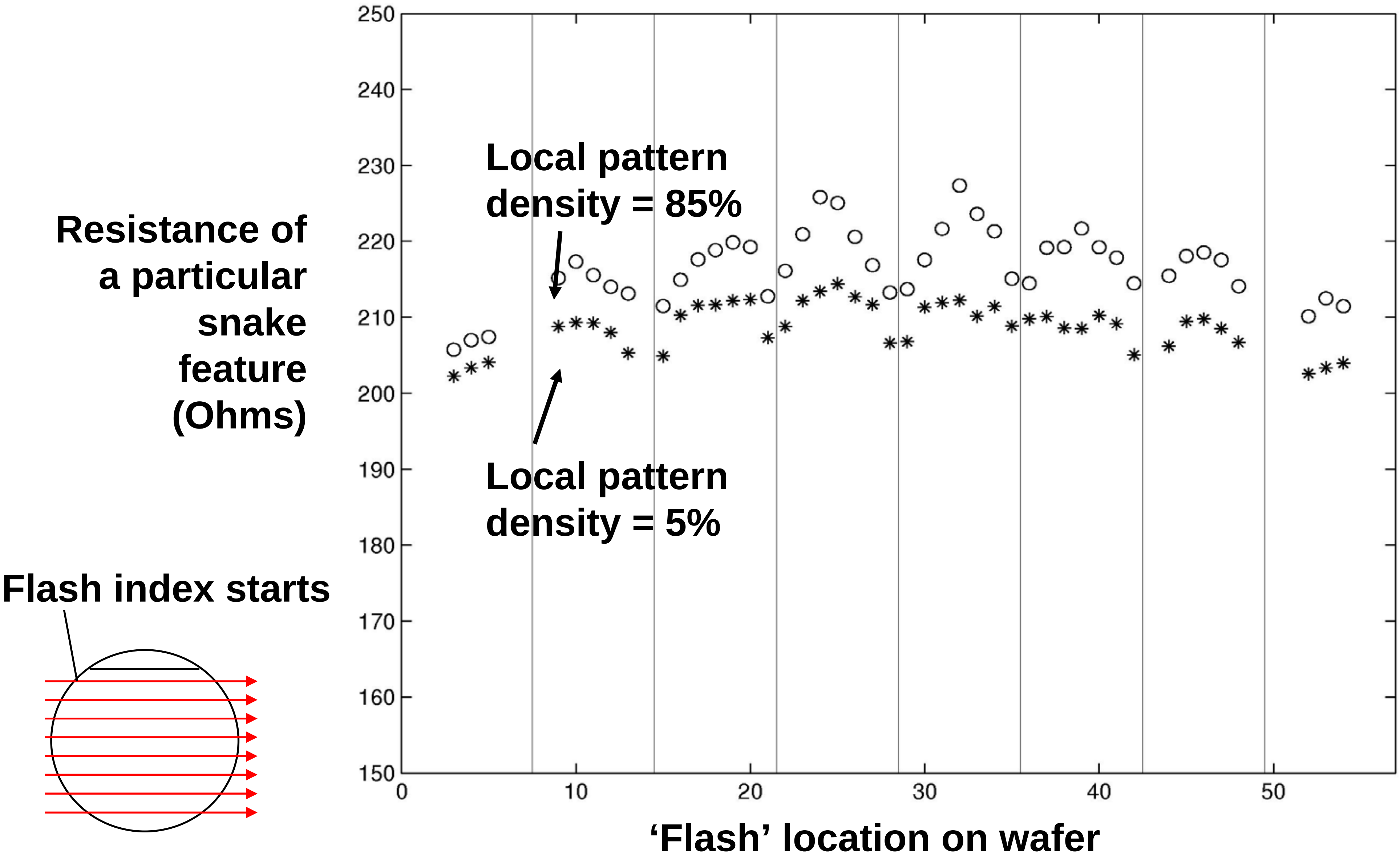
Relevant factors

- **Geometry**
  - Position on wafer
  - Locally averaged pattern density
  - Feature size and pitch
- **Physical perspective**
  - Reactant fluxes in etch chamber
  - …

Pattern density dependency



Wafer-scale nonuniformity



Next time

- Building models based on effects