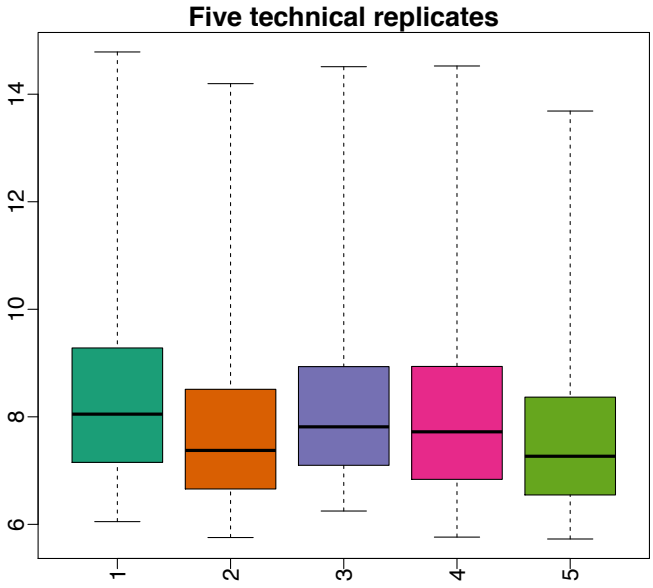
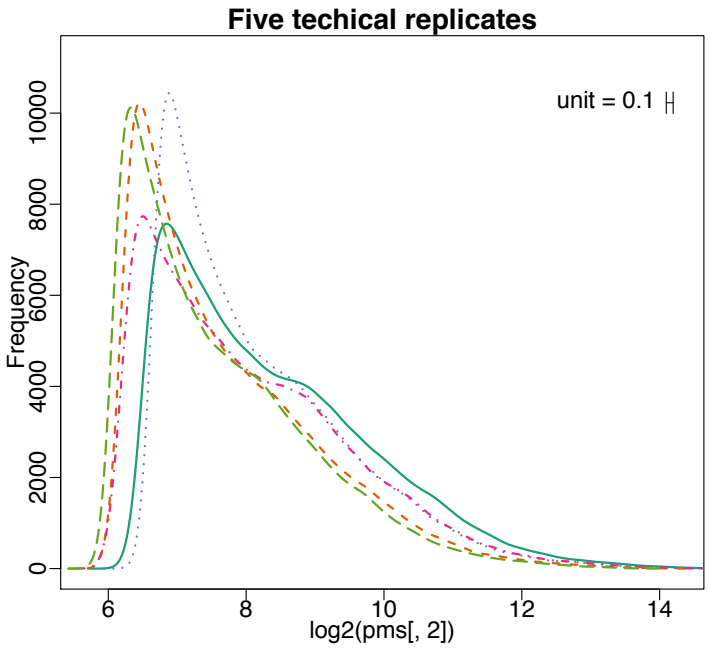
Statistical Modeling 3

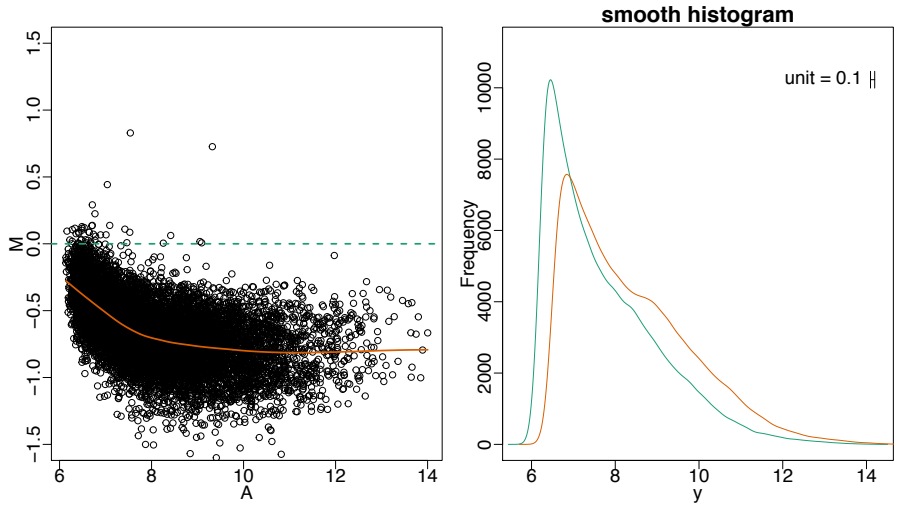
Bias correction and normalization



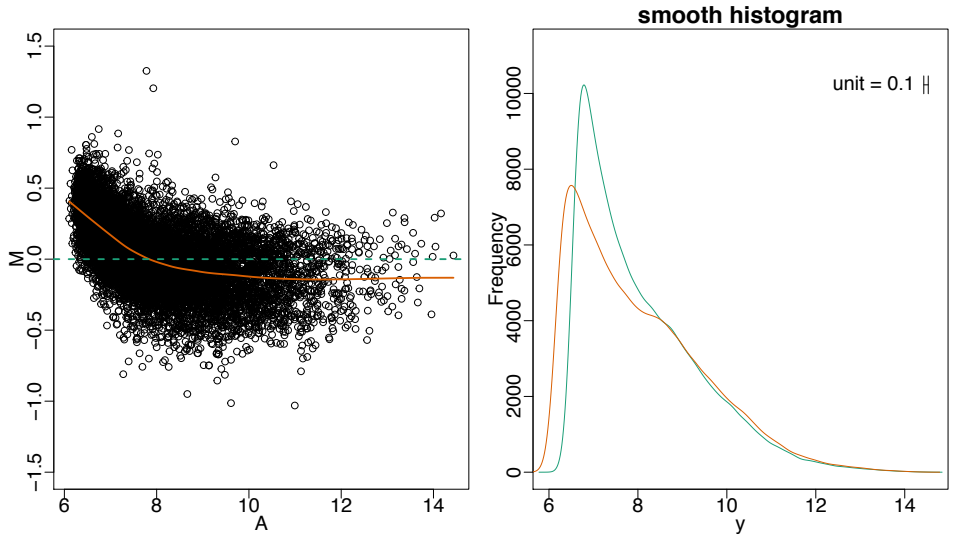
[RI]



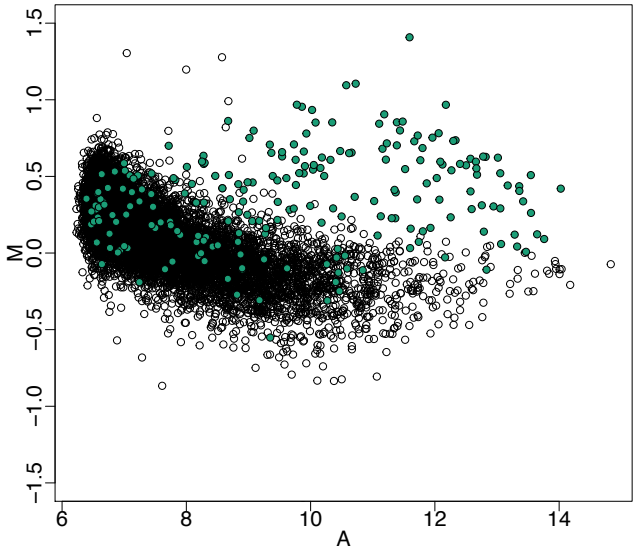
More than location and scale changes!



Median shifts do not solve the problem!



There are non-linear effects!



Quantile normalization

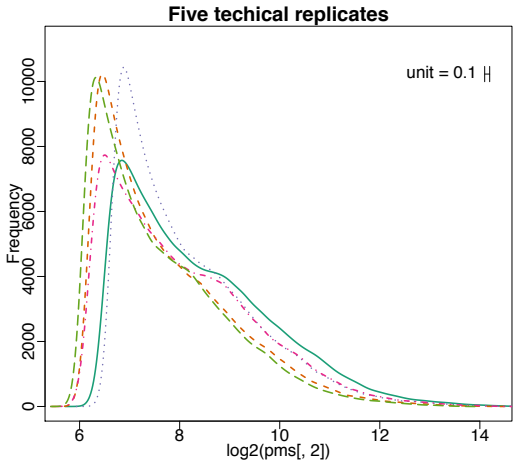
Original			
2	4	4	5
5	14	4	7
4	8	6	9
3	8	5	8
3	9	3	5

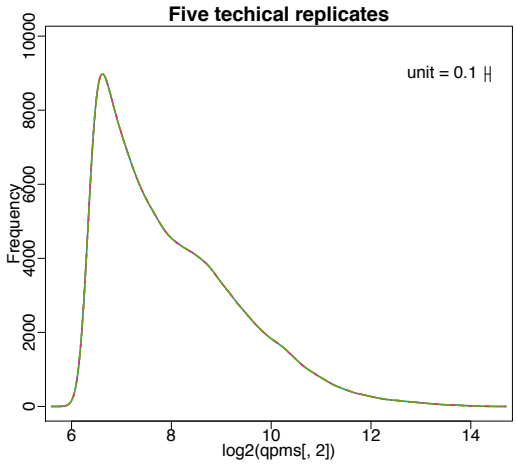
Order			
2	4	3	5
3	8	4	5
3	8	4	7
4	9	5	8
5	14	6	9

Averaged			
3.5	3.5	3.5	3.5
5.0	5.0	5.0	5.0
5.5	5.5	5.5	5.5
6.5	6.5	6.5	6.5
8.5	8.5	8.5	8.5

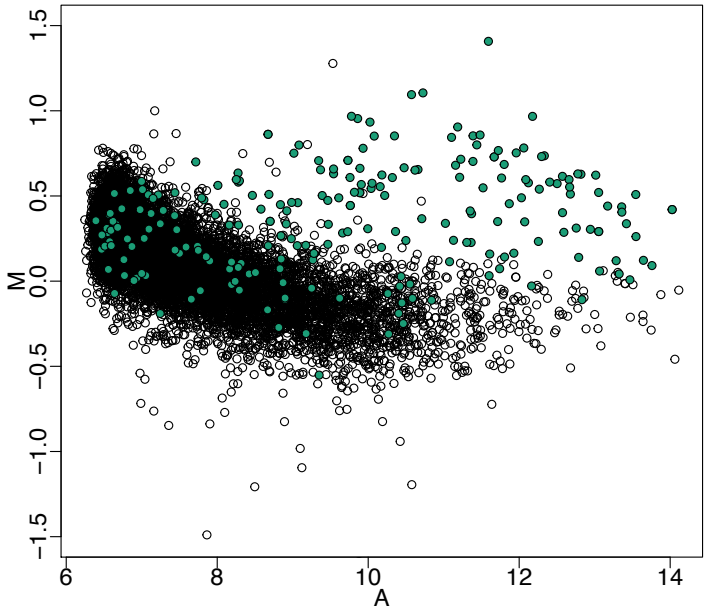
Re-order			
3.5	3.5	5.0	5.0
8.5	8.5	5.5	5.5
6.5	5.0	8.5	8.5
5.0	5.5	6.5	6.5
5.5	6.5	3.5	3.5

Densities are forced to be identical

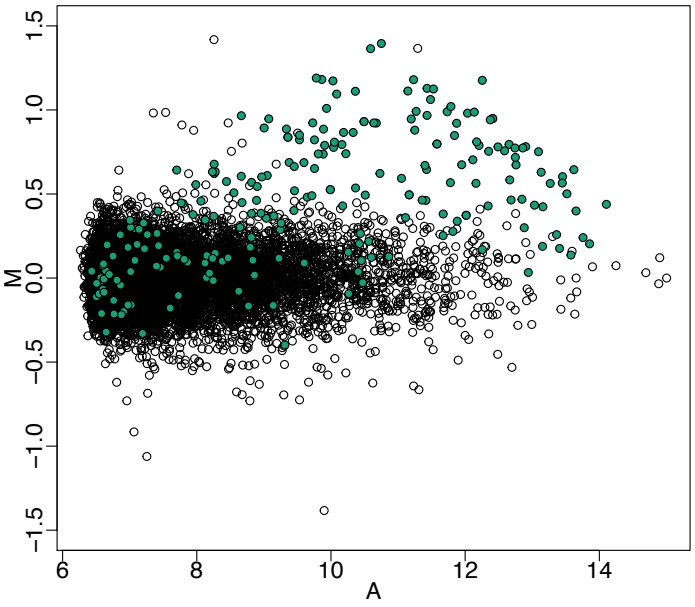


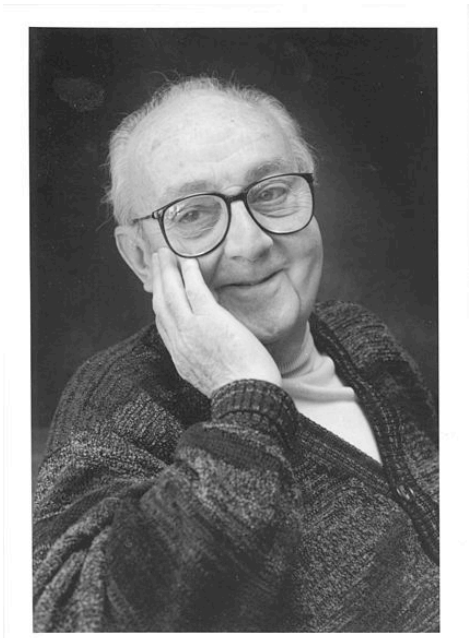


Differential expression can be preserved



Differential expression can be preserved



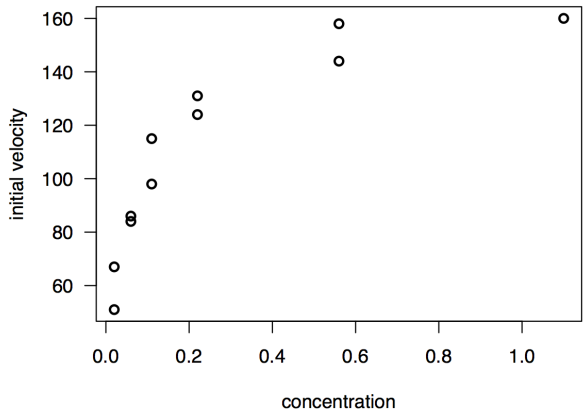


"Essentially, all models are wrong, but some are useful"

George E.P. Box

[RI]

A biochemical experiment



Michaelis-Menten equation

$$V = \frac{V_{\text{max}} \times C}{K + C}$$

V = initial velocity

C = concentration

V_{max} = maximum velocity

K = rate constant

[140.615]

A biochemical experiment

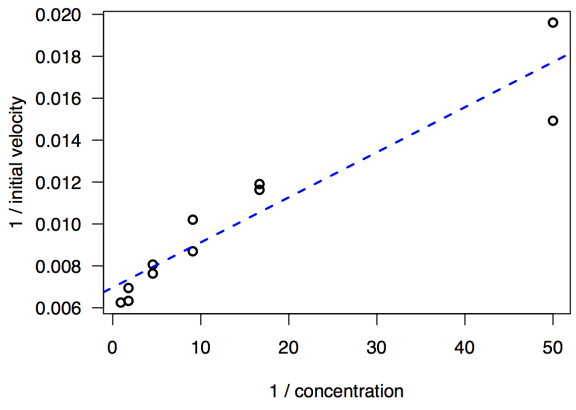
$$V = \frac{V_{\max} \times C}{K + C}$$

$$\Rightarrow \quad \frac{1}{V} = \frac{K + C}{V_{\max} \times C}$$

$$= \frac{K}{V_{\max} \times C} + \frac{1}{V_{\max}}$$

$$\Rightarrow \quad \frac{1}{V} = \left(\frac{1}{V_{\max}}\right) + \left(\frac{K}{V_{\max}}\right) \times \left(\frac{1}{C}\right)$$

A biochemical experiment



Model:

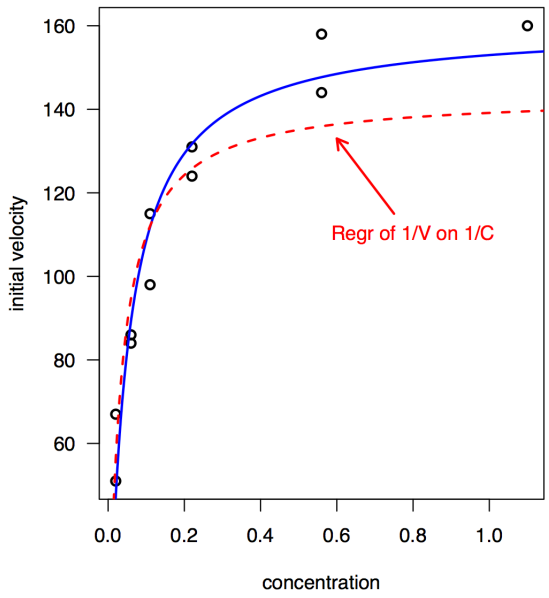
$$\frac{1}{V} = \beta_0 + \beta_1 \left(\frac{1}{C}\right) + \text{error}$$

```
Intercept    0.00697
Slope        0.00022
```

$\hat{V}_{\max}$ = 1/Intercept = 1/0.00697
= 143

$\hat{K}$ = Slope $\times$ $\hat{V}_{\max}$ = 0.031

A biochemical experiment



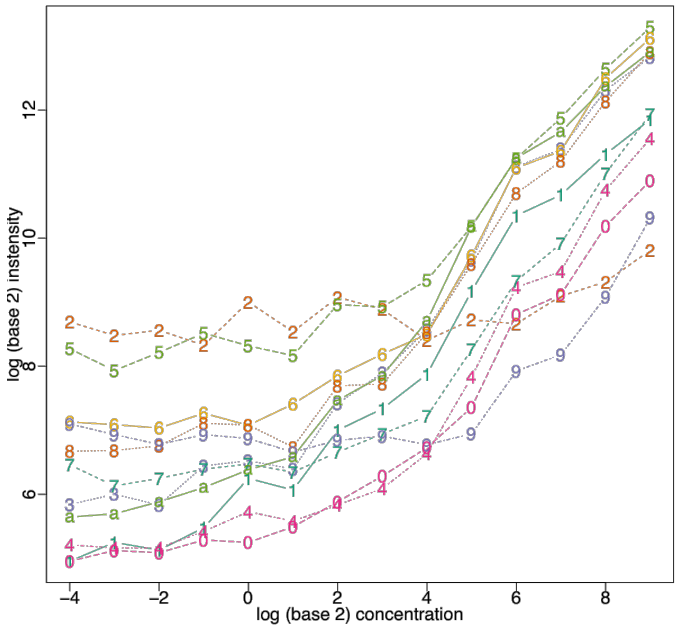
Which is more reasonable?

$$\frac{1}{V} = \beta_0 + \beta_1 \left(\frac{1}{C}\right) + \text{error}$$

$$V = \frac{V_{\max} \times C}{K + C} + \text{error}$$

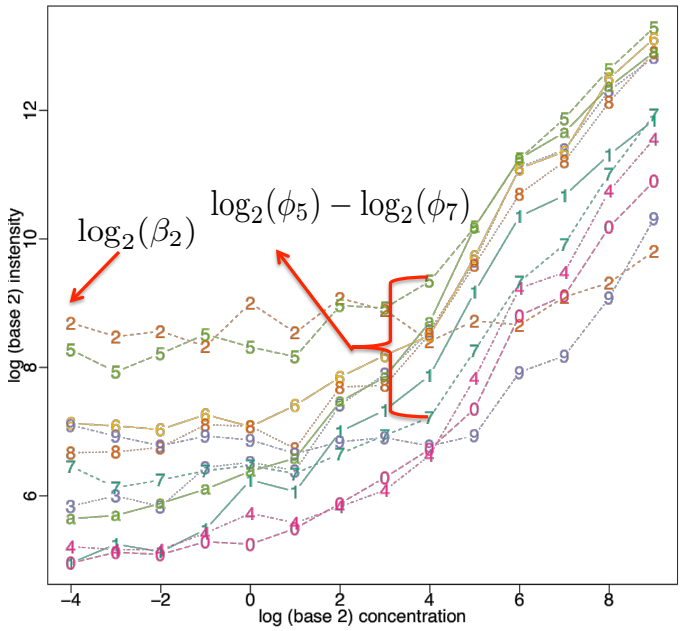
[140.615]

Eleven probes from one spiked-in gene



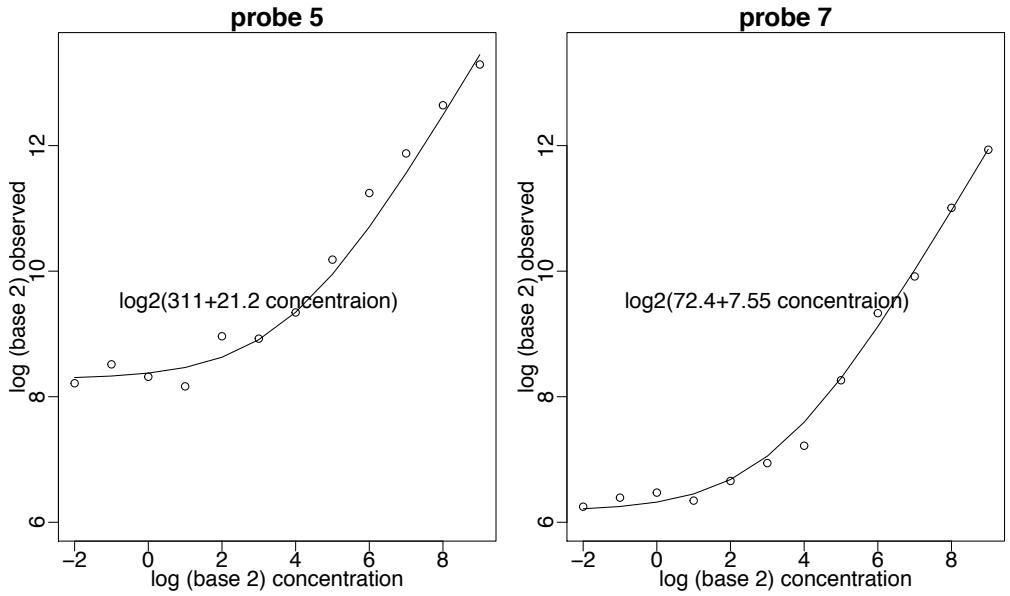
[RI]

$$Y_{ij} = \beta_j + \theta_i \phi_j + \varepsilon_{ij} \qquad \mathrm{var}(\varepsilon_{ij}) \propto \theta_i \phi_j$$



[RI]

Model fit to two probes



[RI]