Stats Boot Camp
Nath Group Meeting, August 18 2015

1. Error propagation

2. Standard errors

3. Outliers

Error Propagation (a.k.a. Propagation of Uncertainty)

$$Z = A \cdot B$$

$$Z \pm \Delta Z = \left(A \pm \Delta A\right)\left(B \pm \Delta B\right)$$

error in derived quantity?

$$Z = f\left(A\right) \Rightarrow \Delta Z = \left(\tfrac{dZ}{dA}\right)\Delta A$$

$$Z = f\left(A,B\right) \Rightarrow \left(\Delta Z\right)^2 = \left(\left(\tfrac{\partial Z}{\partial A}\right)\Delta A\right)^2 + \left(\left(\tfrac{\partial Z}{\partial B}\right)\Delta B\right)^2 + 2\,\tfrac{\partial Z}{\partial A}\tfrac{\partial Z}{\partial B}\Delta A \Delta B$$

covariance term –
assume 0 if A, B independent

A Cheat Sheet

Table 4.1. *Combination of errors*

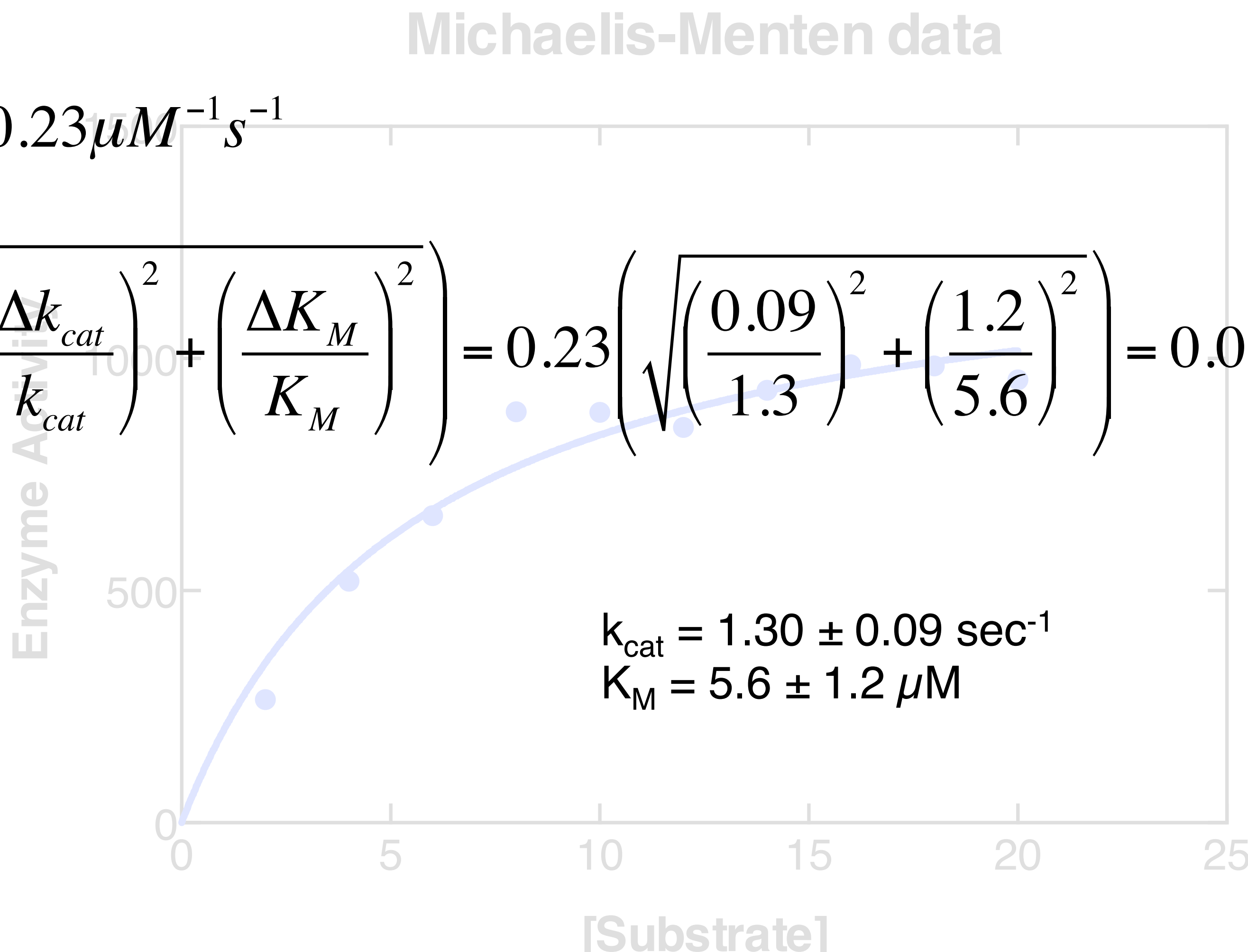
Relation between Z and A, B	Relation between standard errors	
$Z = A + B$ $Z = A - B$	$(\Delta Z)^2 = (\Delta A)^2 + (\Delta B)^2$	(i)
$Z = AB$ $Z = A/B$	$\left(\frac{\Delta Z}{Z}\right)^2 = \left(\frac{\Delta A}{A}\right)^2 + \left(\frac{\Delta B}{B}\right)^2$	(ii)
$Z = A^n$	$\frac{\Delta Z}{Z} = n \frac{\Delta A}{A}$	(iii)
$Z = \ln A$	$\Delta Z = \frac{\Delta A}{A}$	(iv)
$Z = \exp A$	$\frac{\Delta Z}{Z} = \Delta A$	(v)

G.L. Squires, *Practical Physics*

Example: What is the Uncertainty in Catalytic Efficiency?

$$e = \frac{k_{cat}}{K_M} = 0.23 \mu M^{-1} s^{-1}$$

$$\Delta e = e\left(\sqrt{\left(\frac{\Delta k_{cat}}{k_{cat}}\right)^2 + \left(\frac{\Delta K_M}{K_M}\right)^2}\right) = 0.23\left(\sqrt{\left(\frac{0.09}{1.3}\right)^2 + \left(\frac{1.2}{5.6}\right)^2}\right) = 0.05 \mu M^{-1} s^{-1}$$



$k_{cat} = 1.30 \pm 0.09 \text{ sec}^{-1}$

$K_M = 5.6 \pm 1.2\ \mu M$

(Sample data from Prism)

Consolidating Multiple Measurements of a Quantity

Given multiple measurements of *z*:

$$z_1 \pm \Delta z_1, z_2 \pm \Delta z_2, \ldots, z_N \pm \Delta z_N$$

The overall mean and uncertainty are given by:

$$z = \frac{\sum w_i z_i}{\sum w_i} \pm \frac{1}{\sqrt{\sum w_i}}$$

where $w_i = 1 \big/ \left(\Delta z_i\right)^2$

Stats Boot Camp
Nath Group Meeting, August 18 2015

1. Error propagation

2. Standard errors

3. Outliers

Pause to Consider Semantics

Standard deviation…

$$\sigma = \sqrt{\frac{\sum \left(x_i - \langle x \rangle \right)^2}{N}} = \sqrt{\left(\tfrac{1}{N} \sum x_i^2 \right) - \left(\tfrac{1}{N} \sum x_i \right)^2}$$

* “Sample deviation”, $\sigma_{sample} = \sigma \sqrt{\frac{N}{N-1}}$

S.E.M. (standard error of the mean)…

$$\sigma_m = \frac{\sigma_{sample}}{\sqrt{N}}$$

Standard error of fitting…

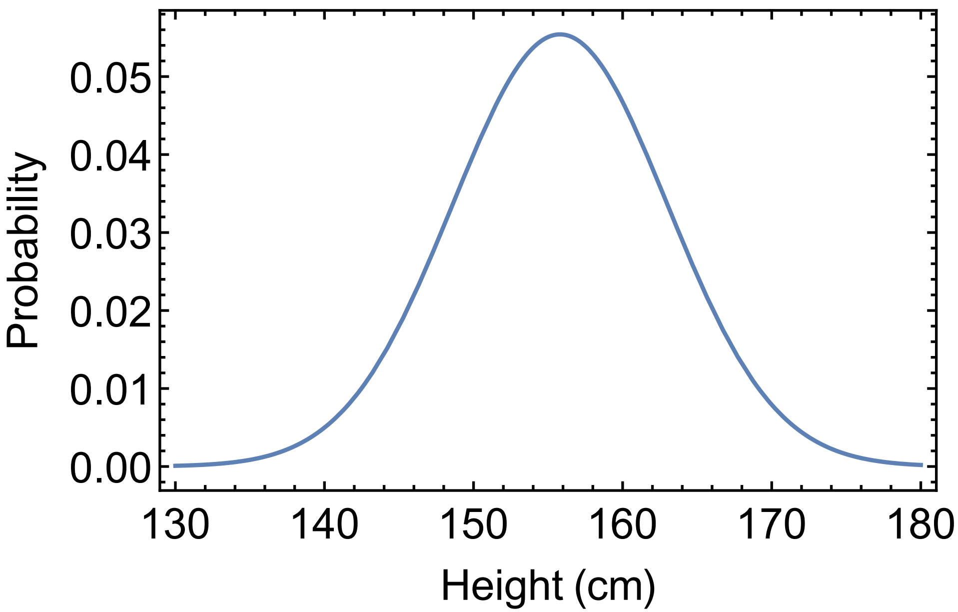
Example: SD vs. SEM

A survey of 364,538 women in 54 countries found a mean height of 155.8 cm, with a reported SD of 7.2 cm.

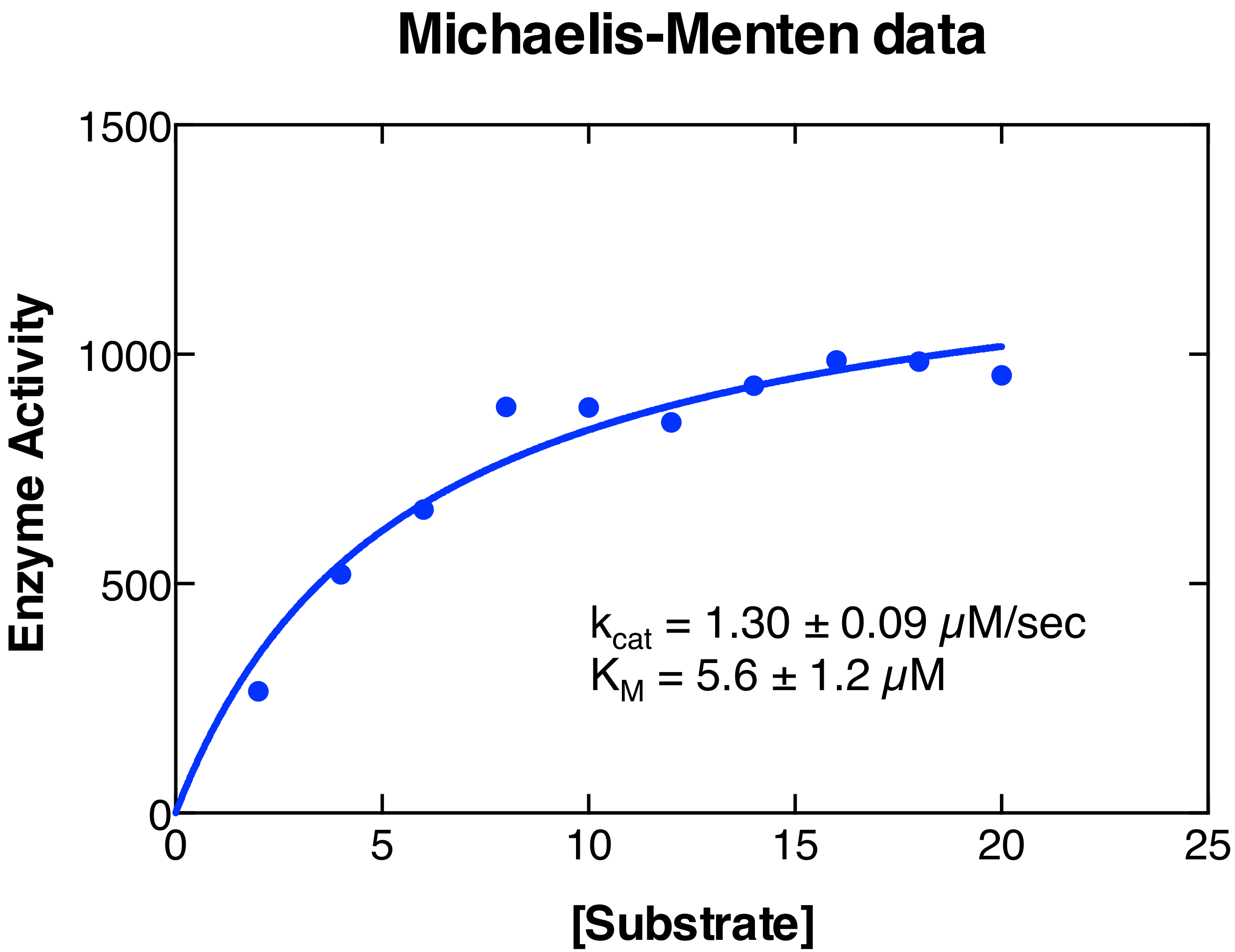
Subramaniam, Özaltin, Finlay (2011) *PLoS One* DOI: 10.1371/journal.pone.0018962

How precisely do we know the mean height in these populations?

$$\mathrm{SEM} = \frac{7.2\,\mathrm{cm}}{364{,}538} = 0.012\,\mathrm{cm}$$

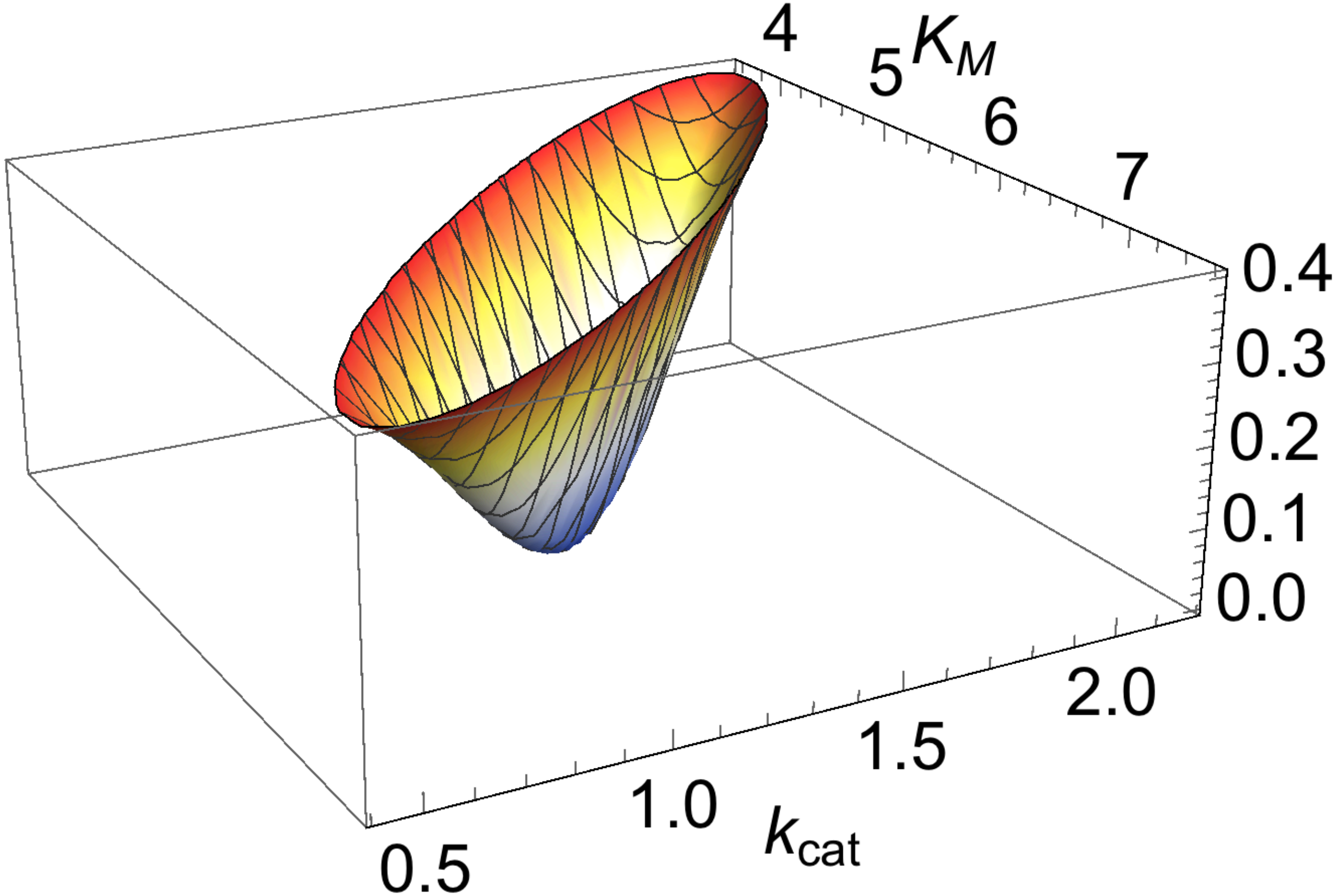


Standard Errors of Fitting

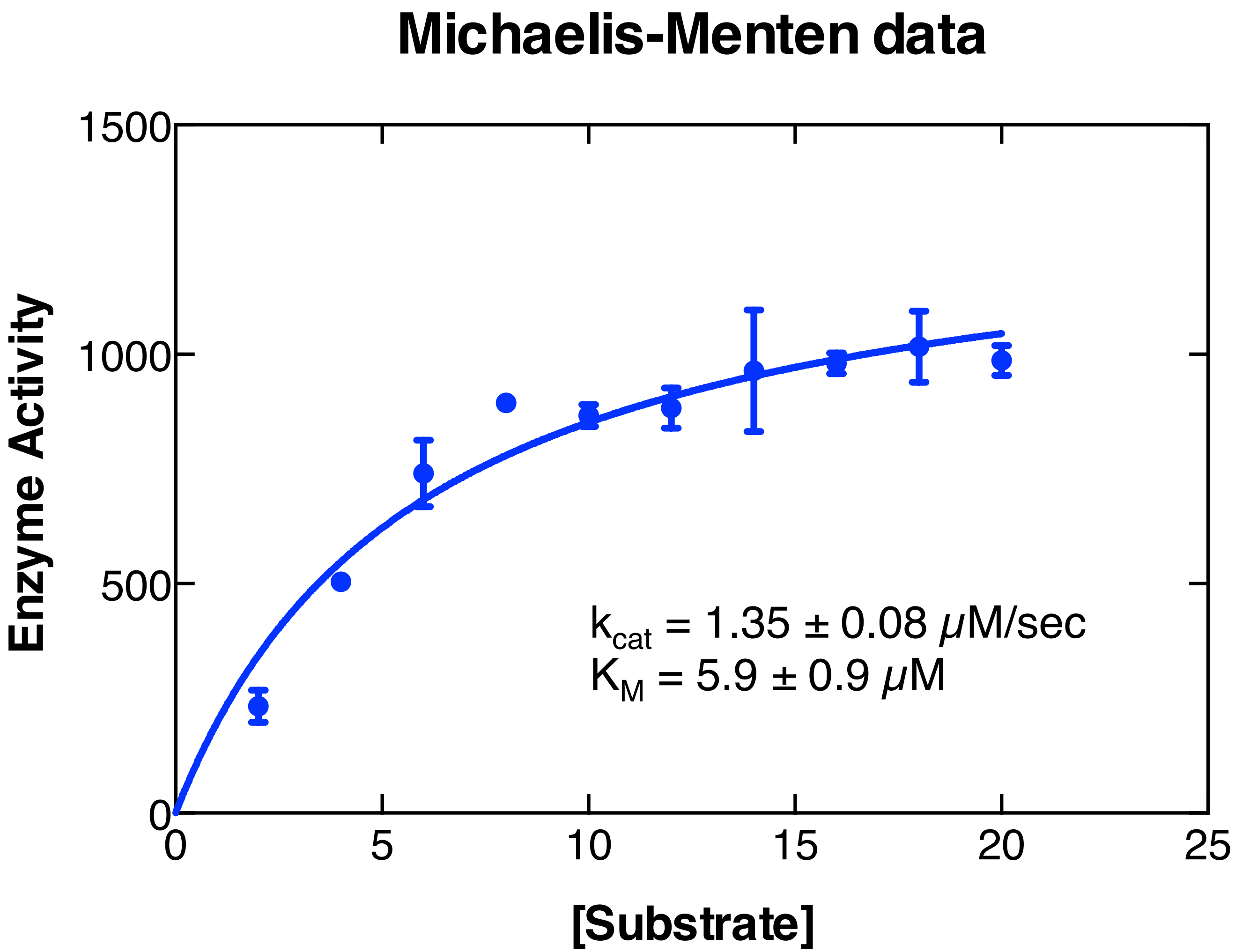


(Sample data from Prism)

Non-Linear Least-Squares Fitting: A Look Under the Hood

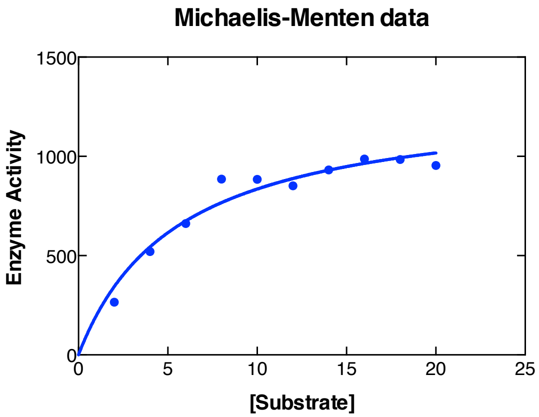


Standard Errors of Fitting

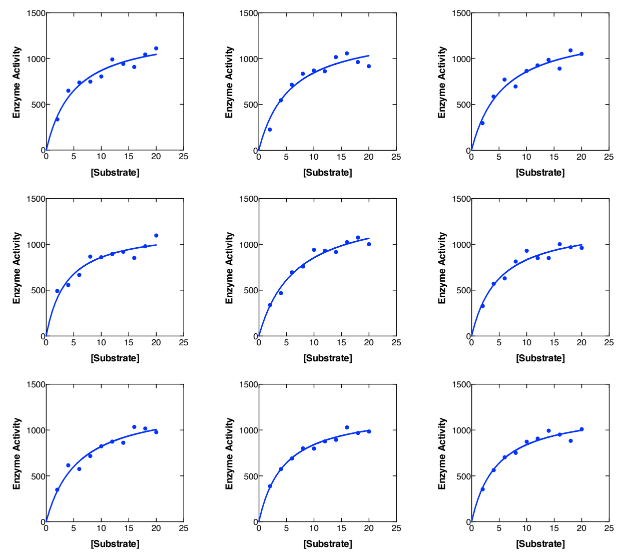


(Sample data from Prism)

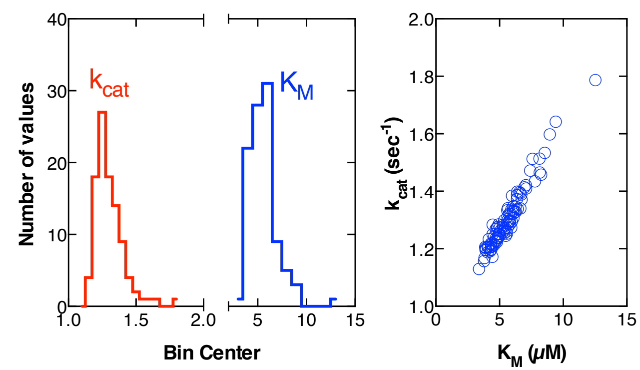
The Gold Standard of Error Propagation: Resampling (cf. Bootstrapping, Jackknifing)



Initial least-squares fit



Simulate decoy datasets



Fit each decoy

Stats Boot Camp
Nath Group Meeting, August 18 2015

1. Error propagation

2. Standard errors

3. Outliers

Detecting Outliers

Q test (a.k.a. Dixon's Q test):

$$x_1 < x_2 < \ldots < x_N$$

$$\frac{gap}{range} = \frac{x_N - x_{N-1}}{x_N - x_1} > Q_{crit}(\alpha, N)?$$

Grubbs' test:

$$\frac{\max\left|x_i - \langle x \rangle\right|}{\sigma} > G_{crit}(\alpha, N) = \frac{(N-1)}{\sqrt{N}}\sqrt{\frac{t^2_{\alpha/2N, N-2}}{N - 2 + t^2_{\alpha/2N, N-2}}}?$$

Keep in mind:

- Is *N* large enough?
- Are data normally distributed?
- No substitute for looking at the data directly

Useful Resources

Prism User Guide:
http://www.graphpad.com/guides/prism/6/curve-fitting/

G.L. Squires, *Practical Physics* – see wiki

Harvey Motulsky, *Intuitive Biostatistics* – in H-058

K. Madsen, H.B. Nielsen, O. Tingleff, *Methods for Non-Linear Least Squares Problems*,
http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/3215/pdf/imm3215.pdf