

Linear Regression Error (Sec. 3.7)

- Consider the predictive model:

$$\hat{y} = c_1 f_1(x) + c_2 f_2(x) + \cdots + c_n f_n(x),$$

- Hat symbol ^ indicates that y values are generated using model parameters derived from a data set.
- $f_i(x)$ can depend nonlinearly on x

Linear Regression Error (Sec. 3.7)

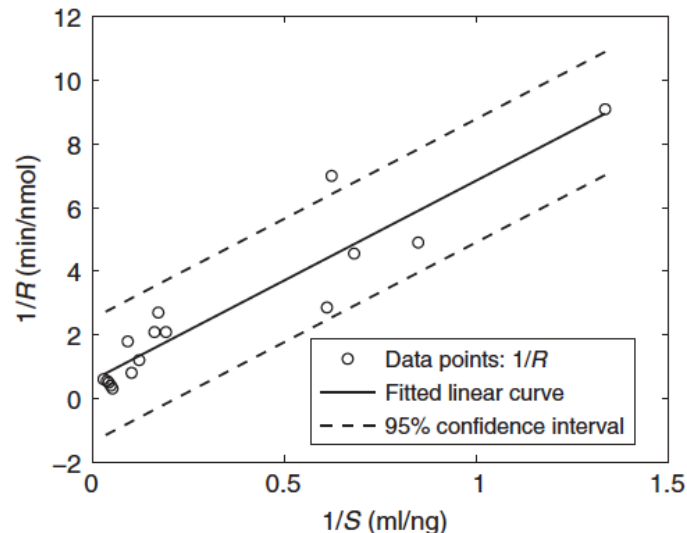
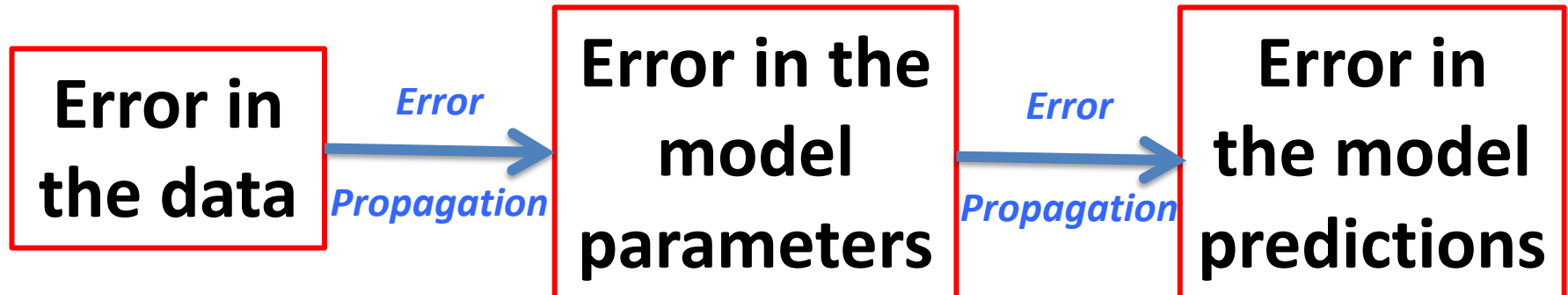
- Basic assumptions:

- (1) The measured values of the independent variable x are known with perfect precision and do not contain any error.
- (2) Every y_i value is normally distributed about its mean μ_{y_i} with an unknown variance σ^2 . The variance of y is independent of x , and is thus the same for all y_i .
- (3) The means of the dependent variable y_i , for the range of x_i values of interest, obey the proposed model $\mu_{y_i} = \beta_1 f_1(x_i) + \beta_2 f_2(x_i) + \cdots + \beta_n f_n(x_i)$.
- (4) A y value at any observed x is randomly chosen from the population distribution of y values for that value of x , so that the random sampling model holds true.
- (5) The y measurements observed at different x values are independent of each other so that their covariance $\sigma_{y_i y_j}^2 = E((y_i - \mu_{y_i})(y_j - \mu_{y_j})) = 0$.

residual, $r_i = (y_i - \hat{y}_i)$.

Linear Regression Error (Sec. 3.7)

- Road map for today, and for any linear regression error calculation:



Linear Regression Error (Sec. 3.7)

- “error in the data”

The **residual variance** measures the variability of the data points about the regression curve and is defined as follows:

$$s_y^2 = \frac{1}{m-n} \sum_{i=1}^m (y_i - \hat{y}_i)^2 = \frac{\|r\|_2^2}{m-n}, \quad (3.48)$$

where m is the number of data pairs and n is the number of undetermined model parameters. Note that $\|r\|_2^2$ is the sum of the squared residuals (SSR) (the subscript indicates $p = 2$ norm).

- next: how does this propagate into error in the (best-fit) model parameters?

Linear Regression Error

The equation that is fitted to the data,

$$y = \beta_1 f_1(x) + \beta_2 f_2(x) + \cdots + \beta_n f_n(x),$$

can be rewritten in compact matrix form $\mathbf{A}\mathbf{c} = \mathbf{y}$, for m data points, where

$$\mathbf{A} = \begin{bmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_n(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_n(x_2) \\ & & \ddots & \\ f_1(x_m) & f_2(x_m) & \cdots & f_n(x_m) \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

Linear Regression Error

- Recall the solution to the Normal Equations:

$$\mathbf{c} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{y}$$

- \mathbf{A} : $m \times n$ matrix
- \mathbf{A}^T : $n \times m$ matrix
- $\mathbf{A}^T \mathbf{A}$, $(\mathbf{A}^T \mathbf{A})^{-1}$: $n \times n$ matrix
- $(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$: $n \times m$ matrix
- \mathbf{y} : $m \times 1$ vector
- Let:

$$\mathbf{K} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T.$$

Linear Regression Error

- Now that we have defined $\mathbf{K} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$,

$$\mathbf{c} = \mathbf{K}\mathbf{y} \text{ or } c_i = \sum_{j=1}^m K_{ij} y_j, \text{ for } 1 \leq i \leq n.$$

- Note that c_i is a linear combination of m normally distributed random variables!
- Thus,

$$s_{c_i}^2 = \sum_{j=1}^m K_{ij}^2 s_{y_j}^2 = \sum_{j=1}^m K_{ij}^2 s_y^2.$$

**“Error in the
model
parameters”**


Linear Regression Error

- Now for the error in the model predictions:

Suppose the linear regression model is $\hat{y} = c_1 f_1(x) + c_2 f_2(x)$.

- Based on our “addition” error propagation formula...

$$s_{\hat{y}_0}^2 = (f_1(x_0))^2 s_{c_1}^2 + (f_2(x_0))^2 s_{c_2}^2 + 2f_1(x_0)f_2(x_0)s_{c_1 c_2}^2.$$


$$s_{c_1 c_2}^2 \neq 0.$$

Linear Regression Error

$$\sigma_{c_1 c_2}^2 = E((c_1 - \beta_1)(c_2 - \beta_2)).$$

Since $c_i = \sum_{j=1}^m K_{ij} y_j$ and $\beta_i = \sum_{j=1}^m K_{ij} \mu_{y_j}$, where $\mathbf{K} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$ and $\mu_{y_i} = E(y_i)$,

$$\begin{aligned}\sigma_{c_1 c_2}^2 &= E\left(\left(\sum_{j=1}^m K_{1j}(y_j - \mu_{y_j})\right)\left(\sum_{k=1}^m K_{2k}(y_k - \mu_{y_k})\right)\right) \\ &= E\left(\sum_{j=1}^m K_{1j} K_{2j} (y_j - \mu_{y_j})^2 + \sum_{j=1}^m \sum_{k=1, k \neq j}^m K_{1j} K_{2k} (y_j - \mu_{y_j})(y_k - \mu_{y_k})\right).\end{aligned}$$

Since

$$\begin{aligned} & E\left(\sum_{j=1}^m \sum_{k=1, k \neq j}^m K_{1j} K_{2k} (y_j - \mu_{y_j}) (y_k - \mu_{y_k})\right) \\ &= \sum_{j=1}^m \sum_{k=1, k \neq j}^m K_{1j} K_{2k} E\left((y_j - \mu_{y_j}) (y_k - \mu_{y_k})\right) = 0, \end{aligned}$$

we have

$$\sigma_{c_1 c_2}^2 = E\left(\sum_{j=1}^m K_{1j} K_{2j} (y_j - \mu_{y_j})^2\right) = \sum_{j=1}^m K_{1j} K_{2j} E\left((y_j - \mu_{y_j})^2\right),$$

$$\sigma_{c_1 c_2}^2 = \sum_{j=1}^m K_{1j} K_{2j} \sigma_y^2.$$

Using Equation (3.50) to proceed with the calculation of $s_{\hat{y}_0}^2$, we obtain

$$s_{\hat{y}_0}^2 = (f_1(x_0))^2 s_{c_1}^2 + (f_2(x_0))^2 s_{c_2}^2 + 2f_1(x_0)f_2(x_0) \sum_{j=1}^m K_{1j} K_{2j} s_y^2.$$

Linear Regression Error

- But we need an error propagation approach that will work for a more general model...

$$y_i \approx \sum_{j=1}^n c_j f_j(x_i),$$

- We define a covariance matrix of the model parameters:

$$\begin{aligned}\Sigma_c^2 &\equiv E\left((\mathbf{c} - \boldsymbol{\mu}_c)(\mathbf{c} - \boldsymbol{\mu}_c)^T\right) \\ \Sigma_c^2 &= \mathbf{K}E\left(\left(\mathbf{y} - \boldsymbol{\mu}_y\right)\left(\mathbf{y} - \boldsymbol{\mu}_y\right)^T\right)\mathbf{K}^T.\end{aligned}$$

Linear Regression Error

$$\Sigma_c^2 = KE \left((y - \mu_y) (y - \mu_y)^T \right) K^T.$$

- This is just the covariance matrix of y .
- Since the indiv. observations y_i are independent of each other, the covariances are all zero.
- If all of the observations y_i have the same variance,

$$\Sigma_y^2 = I\sigma_y^2,$$

- Thus,

$$\Sigma_c^2 = \sigma_y^2 K K^T$$

$$\hat{\mathbf{y}}^{(m)} = \mathbf{A}^{(m)} \mathbf{c},$$

where $A_{ij}^{(m)} = f_j(x_i^{(m)})$ and $\mathbf{A}^{(m)}$ is a $k \times n$ matrix.
the prediction of $\hat{\mathbf{y}}^{(m)}$. We construct the matrix of

$$\Sigma_{\hat{\mathbf{y}}^{(m)}}^2 = E\left(\left(\hat{\mathbf{y}}^{(m)} - \boldsymbol{\mu}_y^{(m)}\right)\left(\hat{\mathbf{y}}^{(m)} - \boldsymbol{\mu}_y^{(m)}\right)^T\right).$$

Substituting $\hat{\mathbf{y}}^{(m)} = \mathbf{A}^{(m)} \mathbf{c}$, we obtain

$$\begin{aligned}\Sigma_{\hat{\mathbf{y}}^{(m)}}^2 &= \mathbf{A}^{(m)} E\left((\mathbf{c} - \boldsymbol{\mu}_c)(\mathbf{c} - \boldsymbol{\mu}_c)^T\right) \mathbf{A}^{(m)T} \\ &= \mathbf{A}^{(m)} \Sigma_c^2 \mathbf{A}^{(m)T}.\end{aligned}$$

Therefore

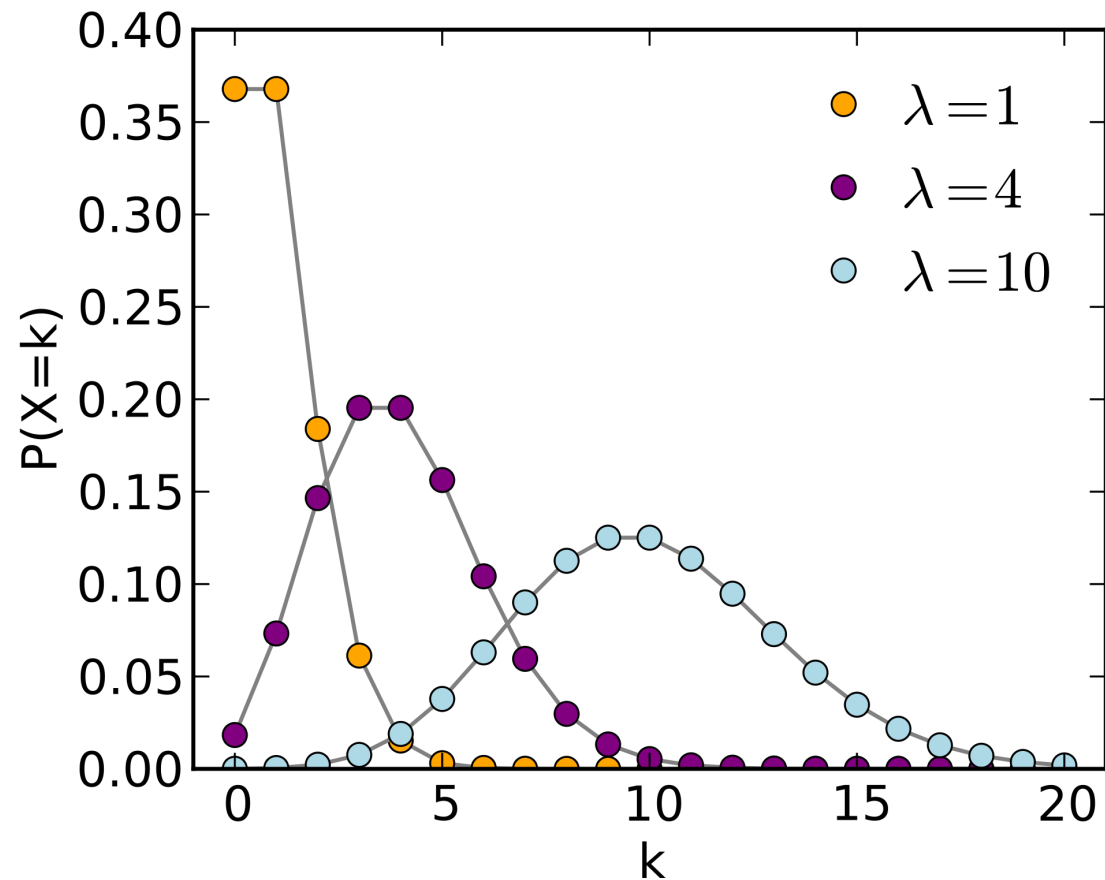
$$\Sigma_{\hat{\mathbf{y}}^{(m)}}^2 = \sigma_y^2 \mathbf{A}^{(m)} \mathbf{K} \mathbf{K}^T \mathbf{A}^{(m)T}.$$

- Only the diagonal elements are meaningful.
- See textbook **Box 3.9A,B** for a Matlab example of the whole calculation.

Q1: “I am a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time and/or space, if these events occur with a known average rate and independently of the time since the last event...

Who am I?”

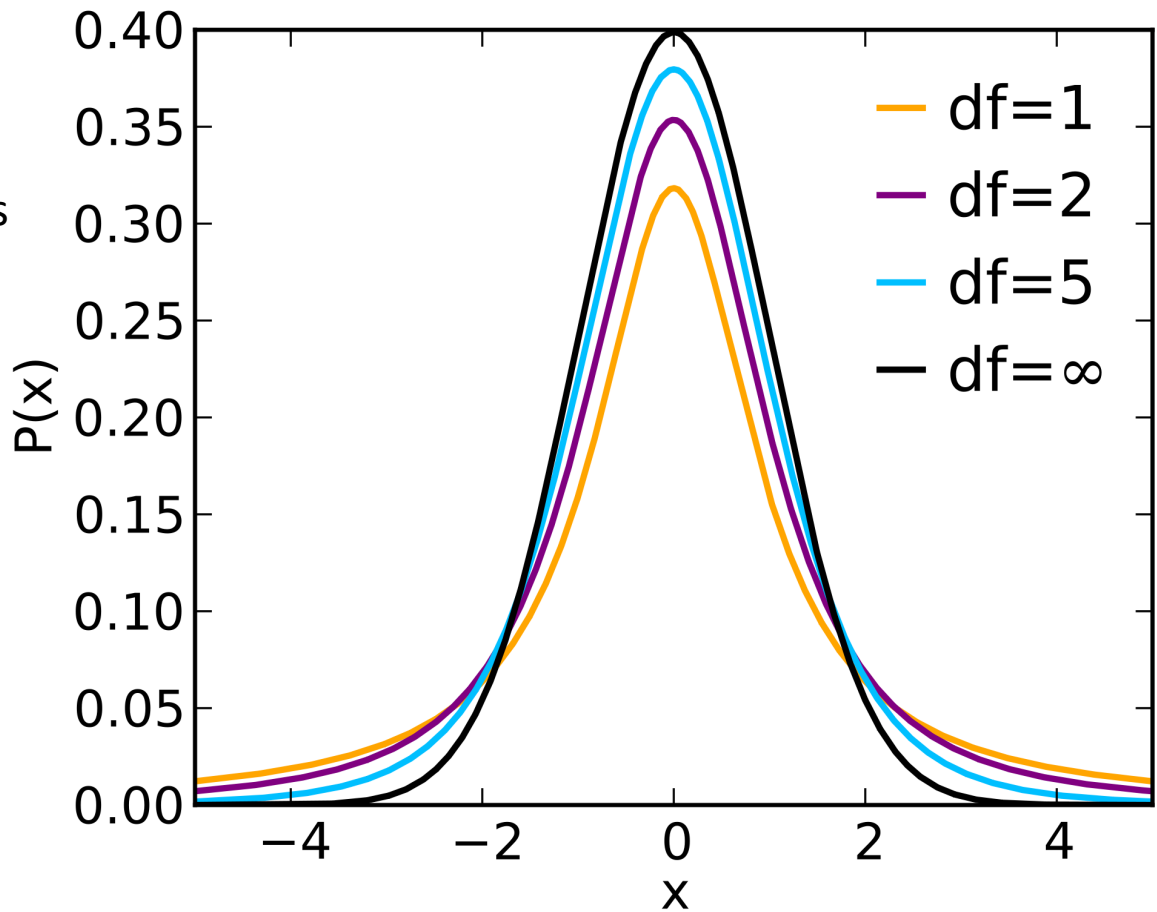
- A. Gaussian distribution**
- B. Poisson distribution**
- C. Student’s t-distribution**
- D. Chi-squared distribution**



Q2: “I am a continuous probability distribution that arises when estimating the mean of a normally distributed population in situations where the sample size is small and the population standard deviation is unknown...

Who am I?”

- A. Binomial distribution**
- B. Gaussian distribution**
- C. Student’s t-distribution**
- D. Chi-squared distribution**



Q3:

The **???** states that if a sample is drawn from a non-normally distributed population with mean μ and variance σ^2 , and the sample size n is large, the distribution of the sample means will be approximately normal with mean μ and variance σ^2/n .

A. Central limit theorem

B. Law of averages

C. Murphy's Law

D. Bayes' theorem

Enzyme Inhibition

A. Reversible

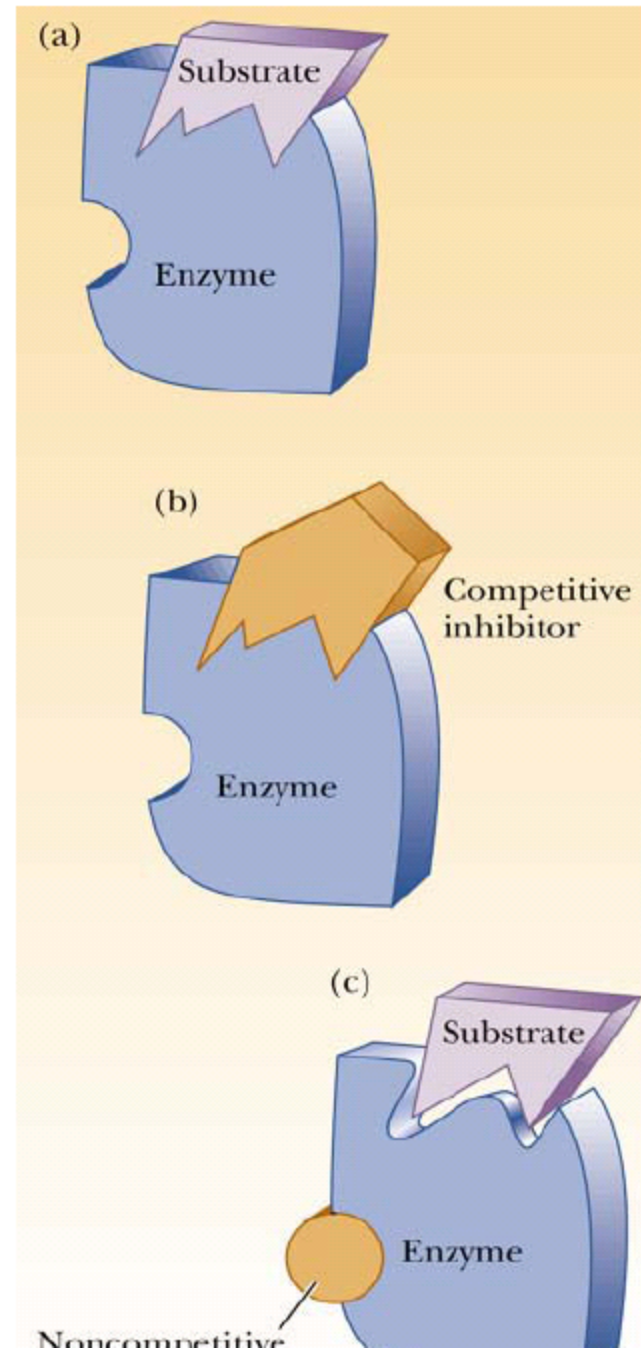
- 1) Competitive
- 2) Non-competitive
- 3) **Uncompetitive**

B. Irreversible

Enzyme inhibition

- **Reversible inhibitor:** a substance that binds to an enzyme to inhibit it, but can be released easily
 - **competitive inhibitor:** binds to the active (catalytic) site and blocks access to it by substrate
 - **noncompetitive inhibitor:** binds to a site other than the active site; inhibits the enzyme by changing its conformation

Allosteric



Irreversible inhibitor: a substance that causes inhibition that cannot be easily reversed

It is usually involves formation or breaking of covalent bonds to or on the enzyme

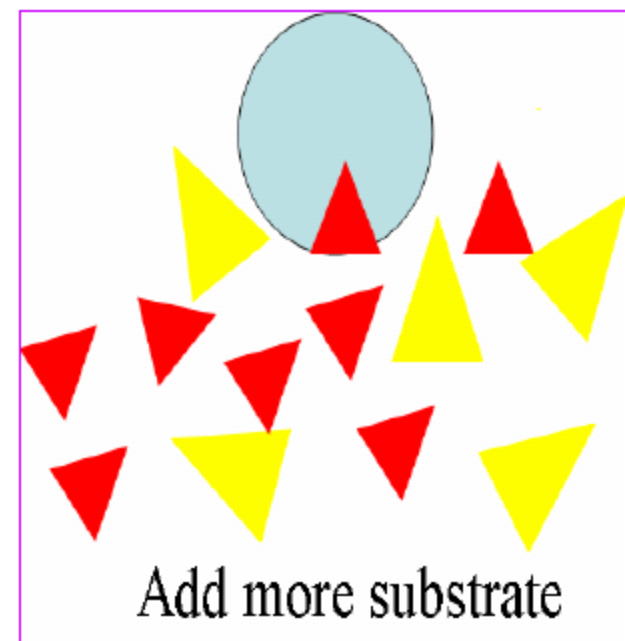
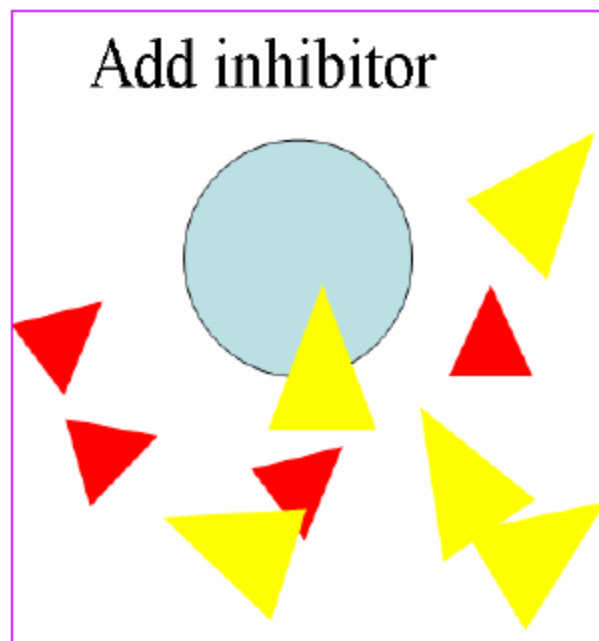
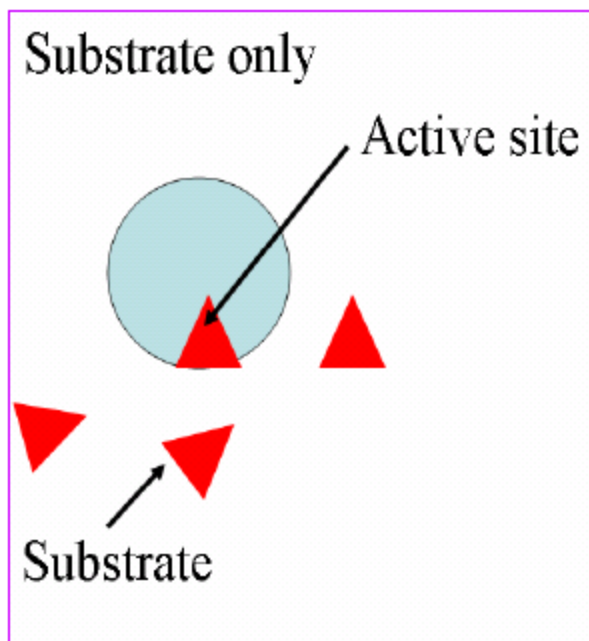
Dissociates very slowly from the enzyme as it is very tightly bound, either covalently or non covalently

Drugs: Penicillin, Aspirin

Usually modeled with first-order reaction kinetics

Competitive inhibitor

- Binds at the active site.
- Prevents binding of substrate.
- Can be reversed by adding more substrate.



Competitive inhibition

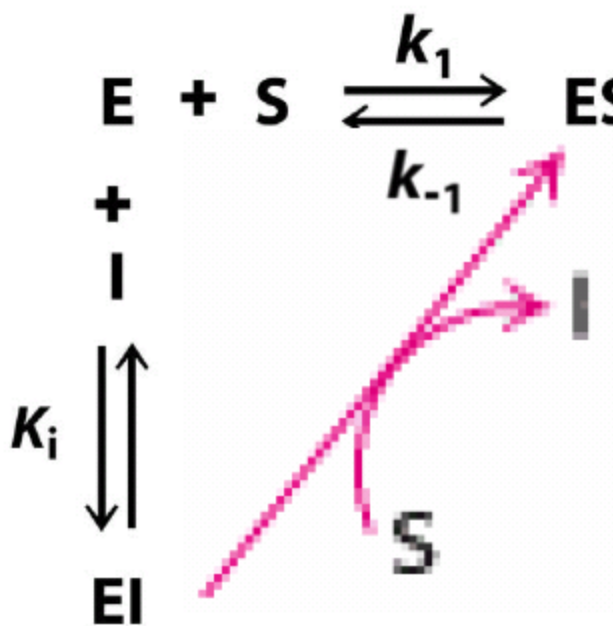
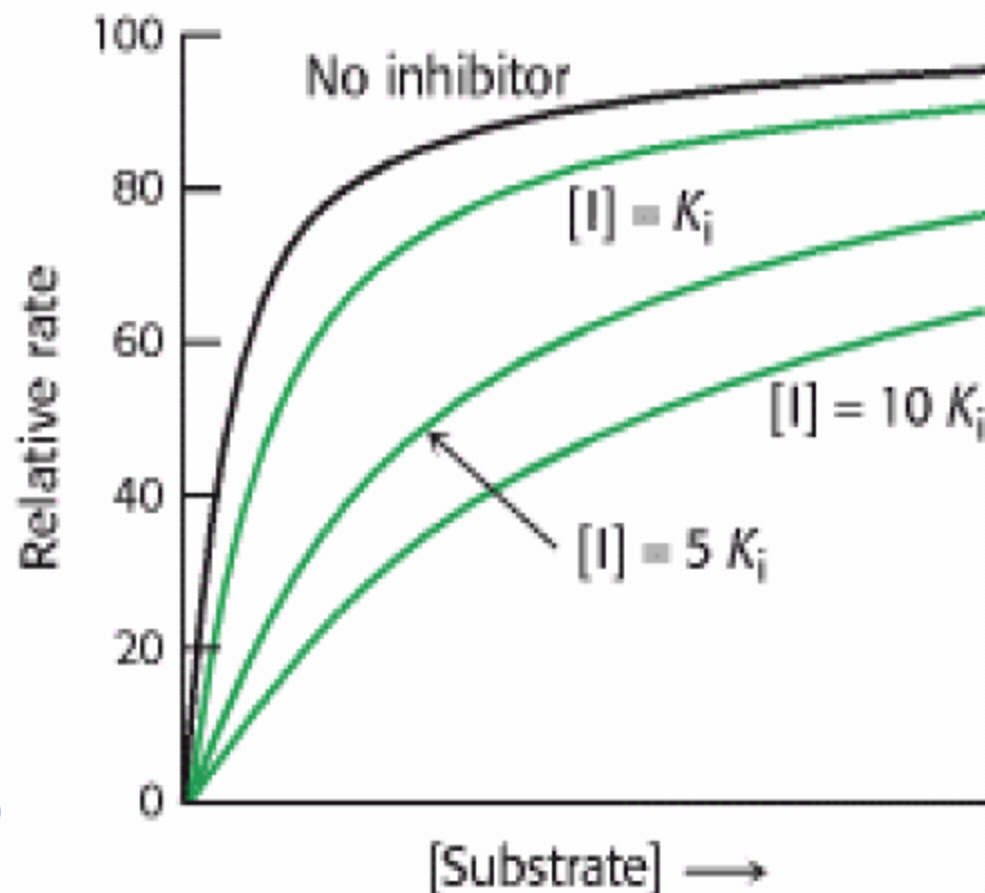


Figure 5-9a Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.

$$K_i = [\text{E}][\text{I}]/[\text{EI}]$$

Smaller K_i = potent inhibitor



Competitive inhibition

$$K_i = [E][I]/[EI]$$

$$K_M^{\text{app}} = K_M(1 + [I]/K_i)$$

$$\alpha = \left(1 + \frac{[I]}{K_i}\right)$$

$$\frac{1}{v} = \frac{\alpha K_m}{V_{\max}} \frac{1}{[S]} + \frac{1}{V_{\max}}$$

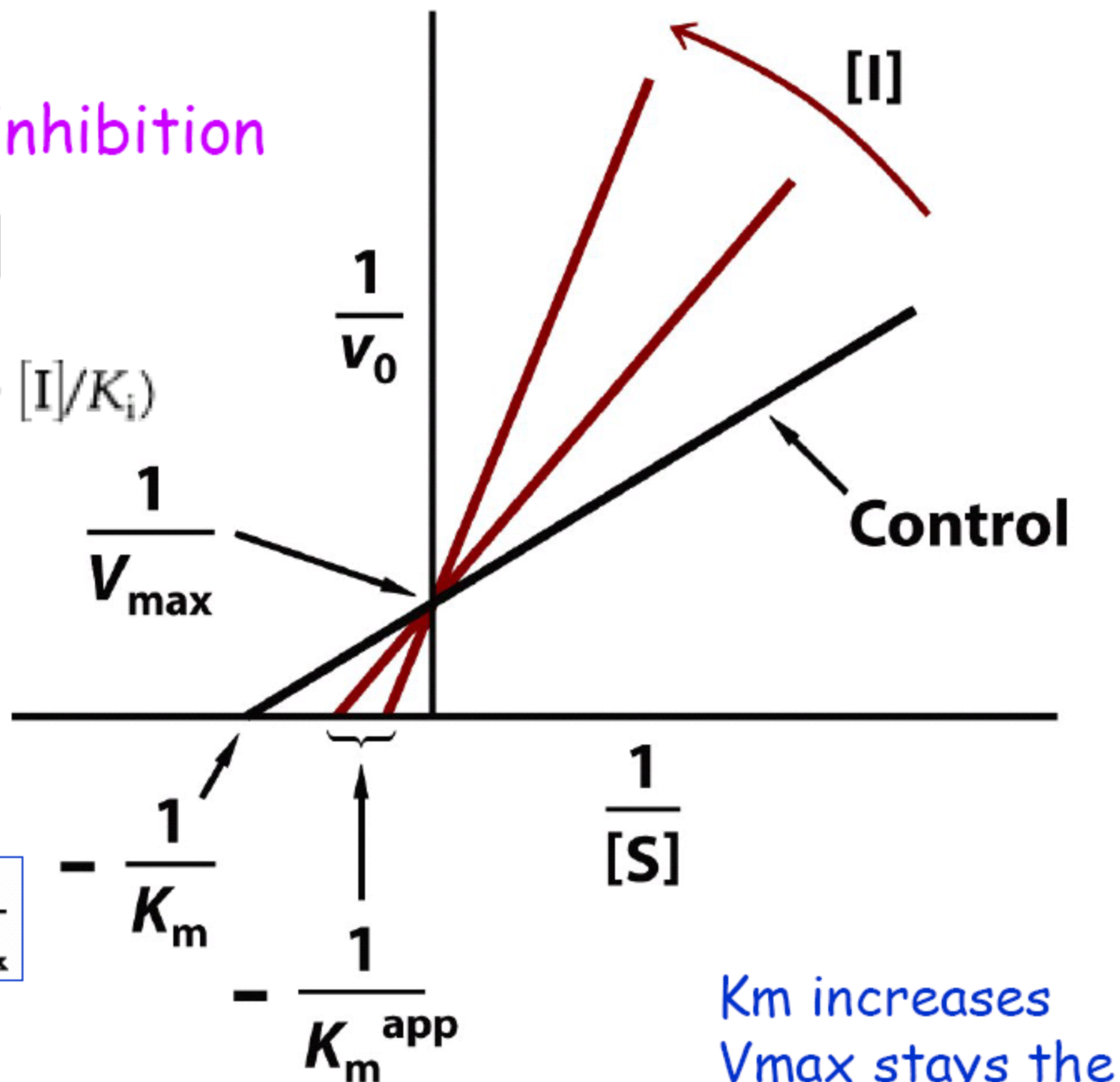
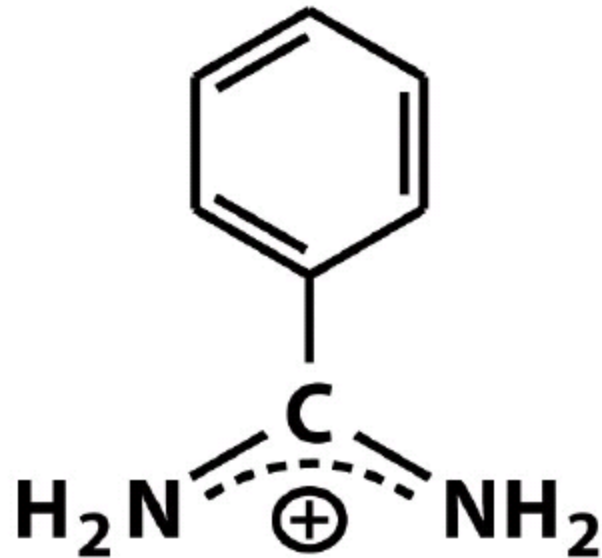
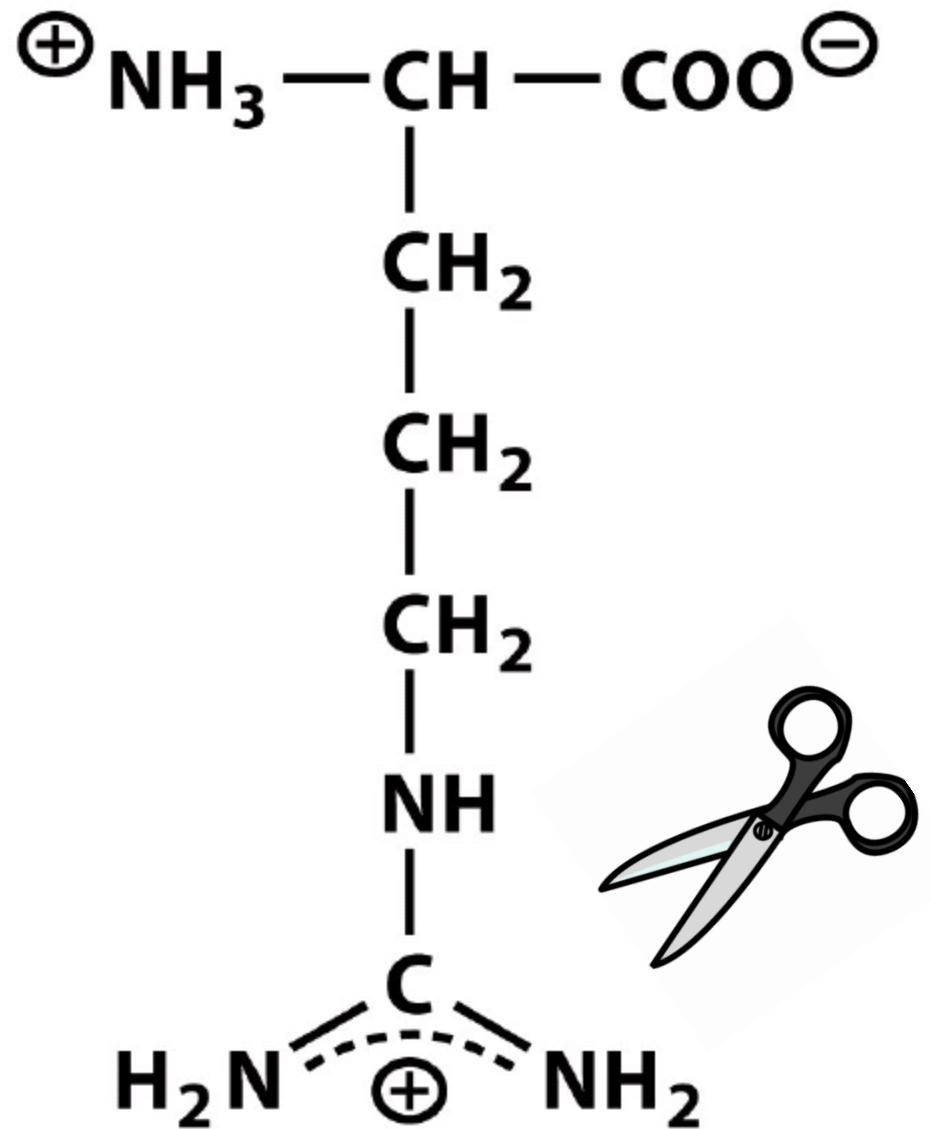


Figure 5-9b Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.

Benzamidine is
a competitive
inhibitor of trypsin.



Benzamidine

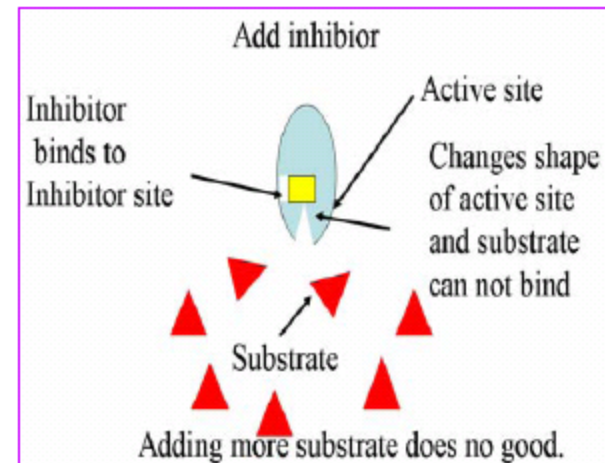
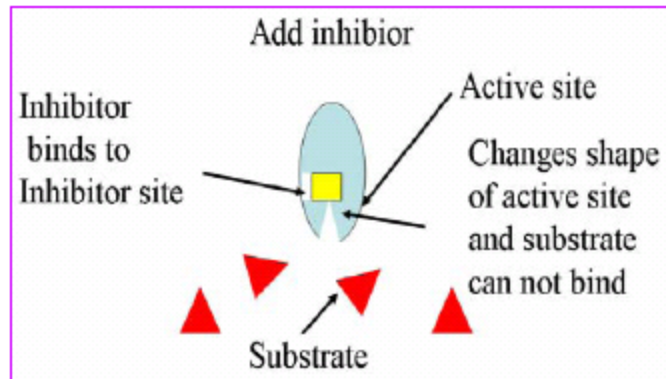
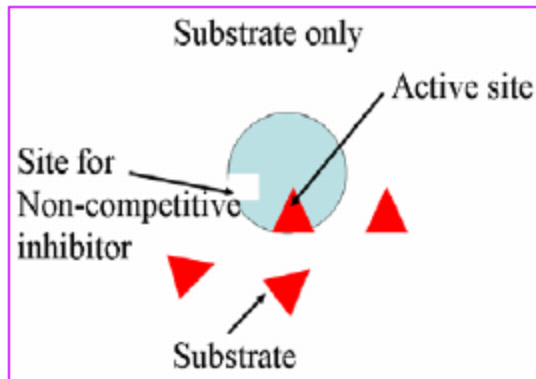


Arginine

- **Ibuprofen:** Enzymes that participate in the signaling pathways in the inflammatory response
- **Statins:** Enzyme in cholesterol biosynthesis (reduces cholesterol level)

Non Competitive inhibitor

- Binds at a different site than the active site
- Cannot be reversed by adding more substrate.



Non Competitive inhibition

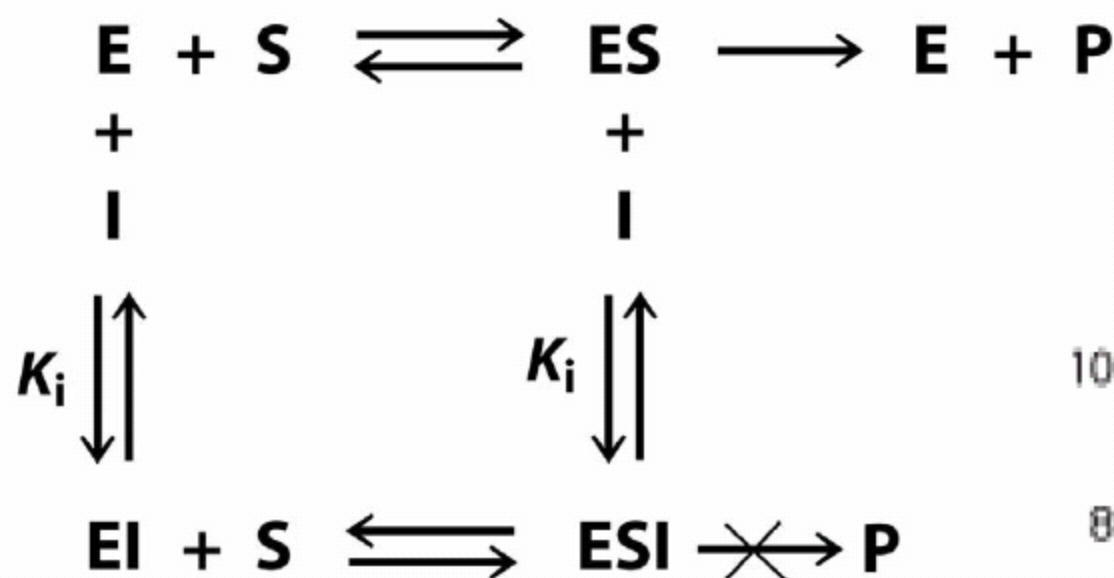
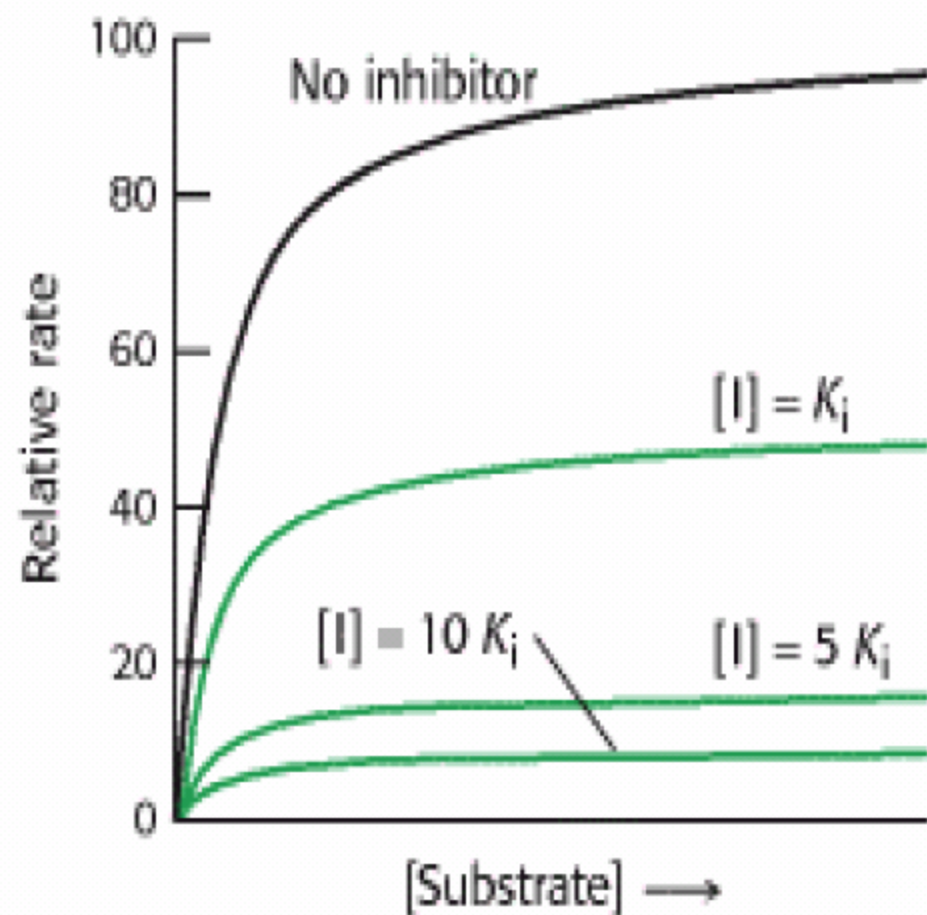


Figure 5-12a Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.

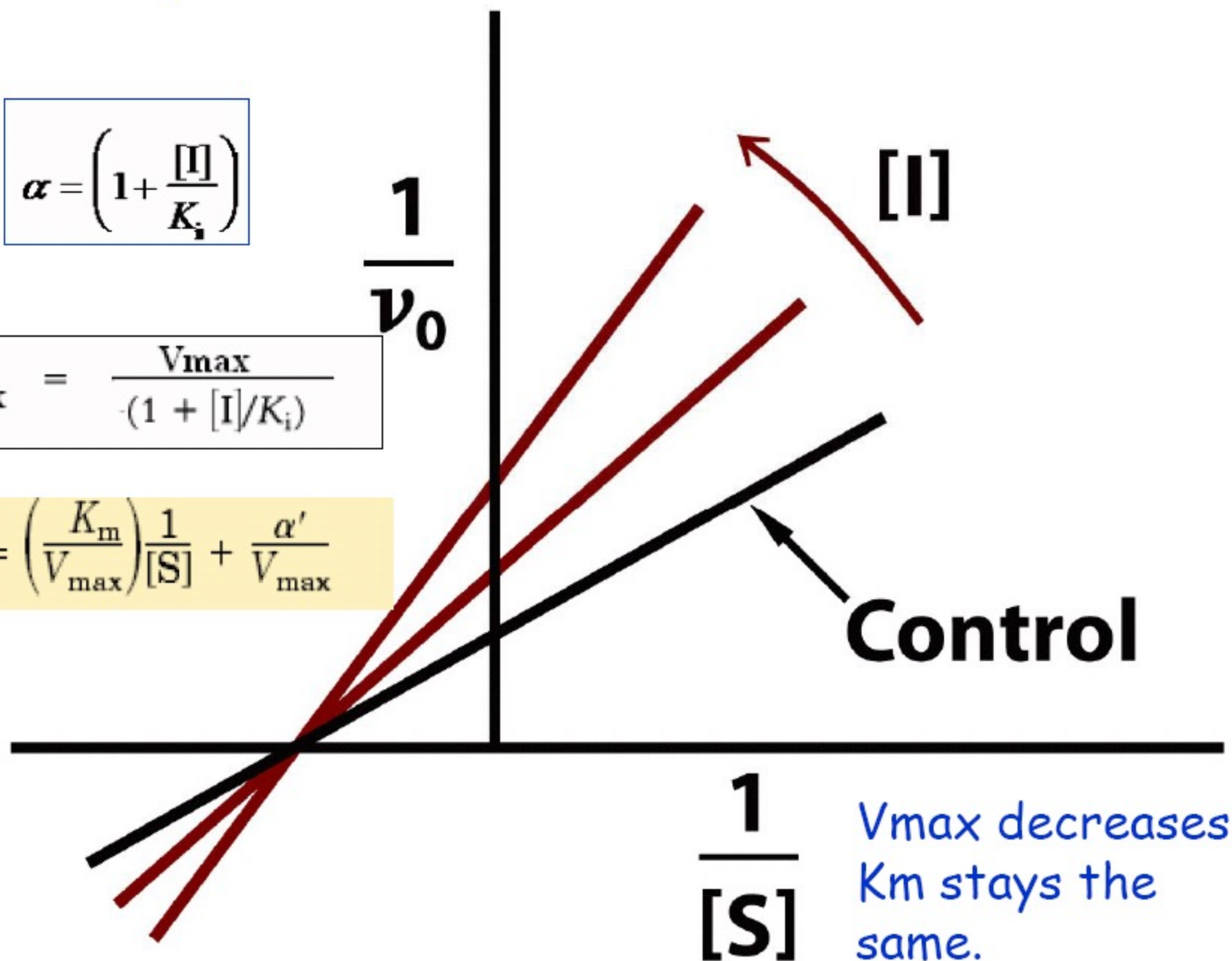


Non Competitive inhibition

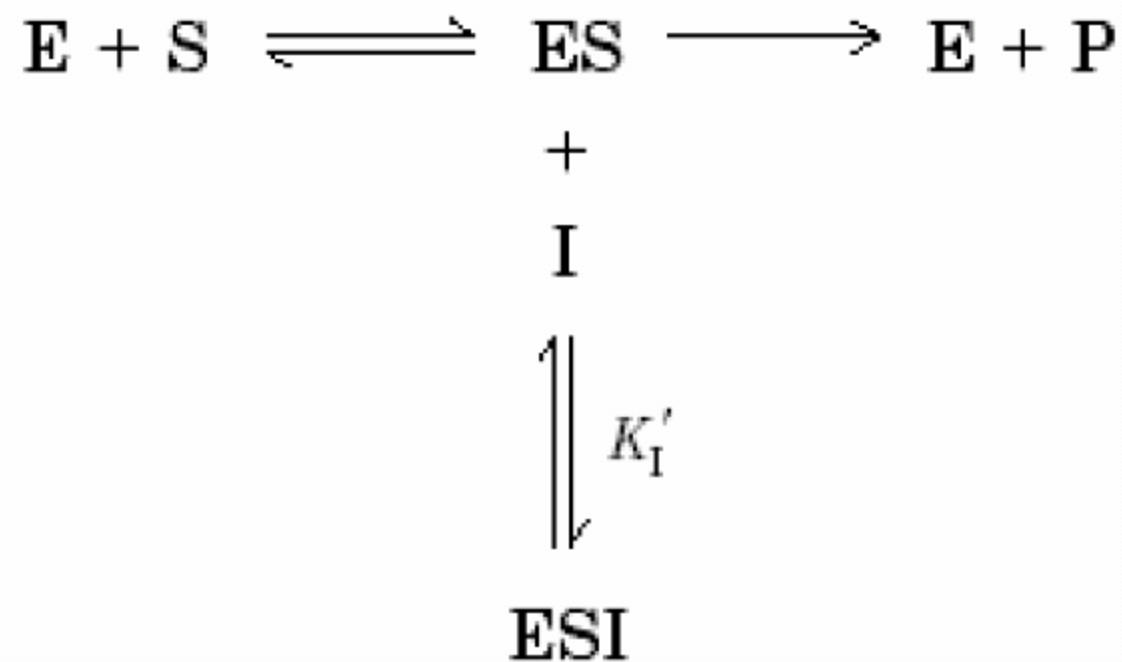
$$\alpha = \left(1 + \frac{[I]}{K_i} \right)$$

$$V_{\max}^{\text{app}} = \frac{V_{\max}}{(1 + [I]/K_i)}$$

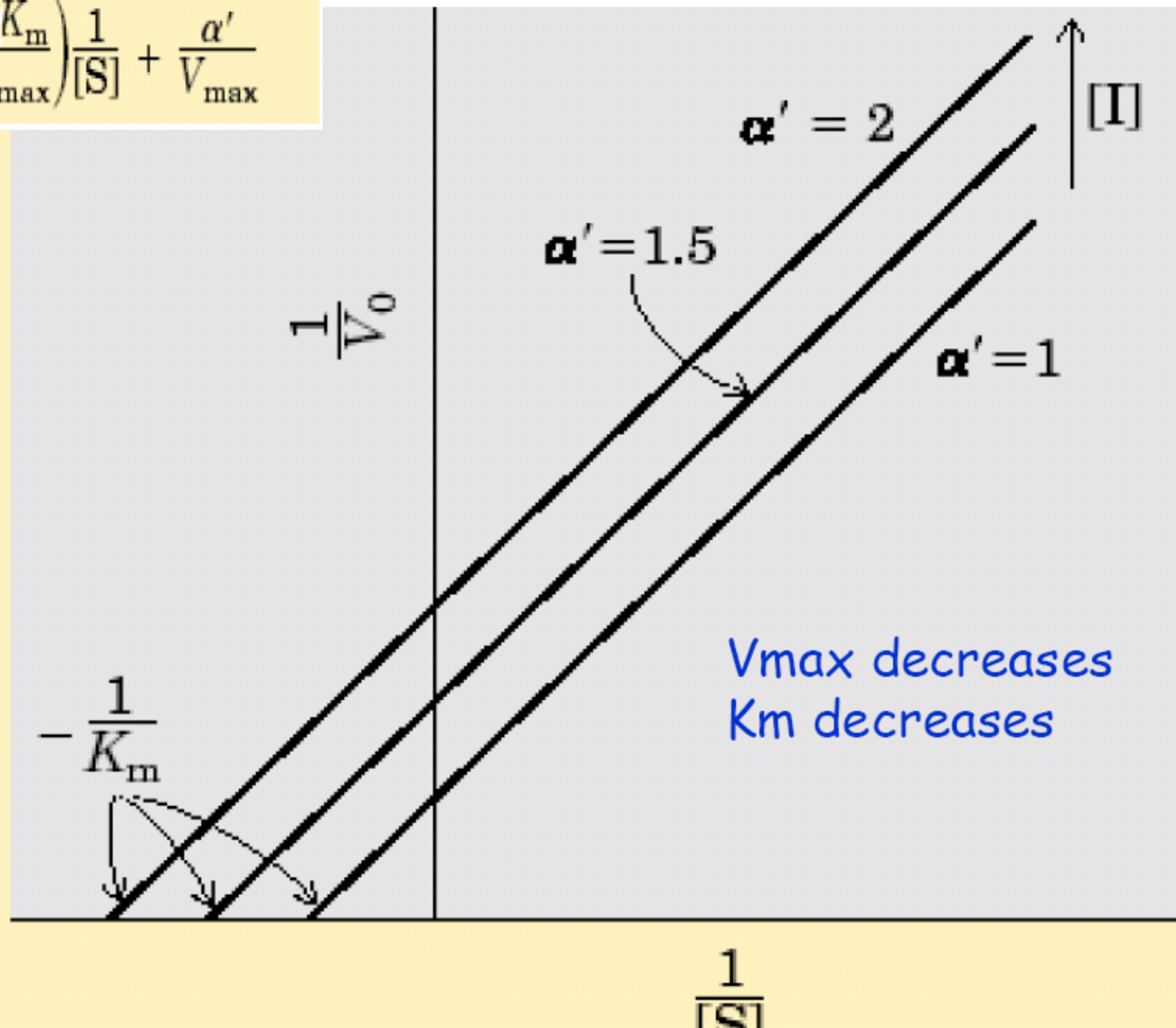
$$\frac{1}{V_0} = \left(\frac{K_m}{V_{\max}} \right) \frac{1}{[S]} + \frac{\alpha'}{V_{\max}}$$



UnCompetitive inhibition

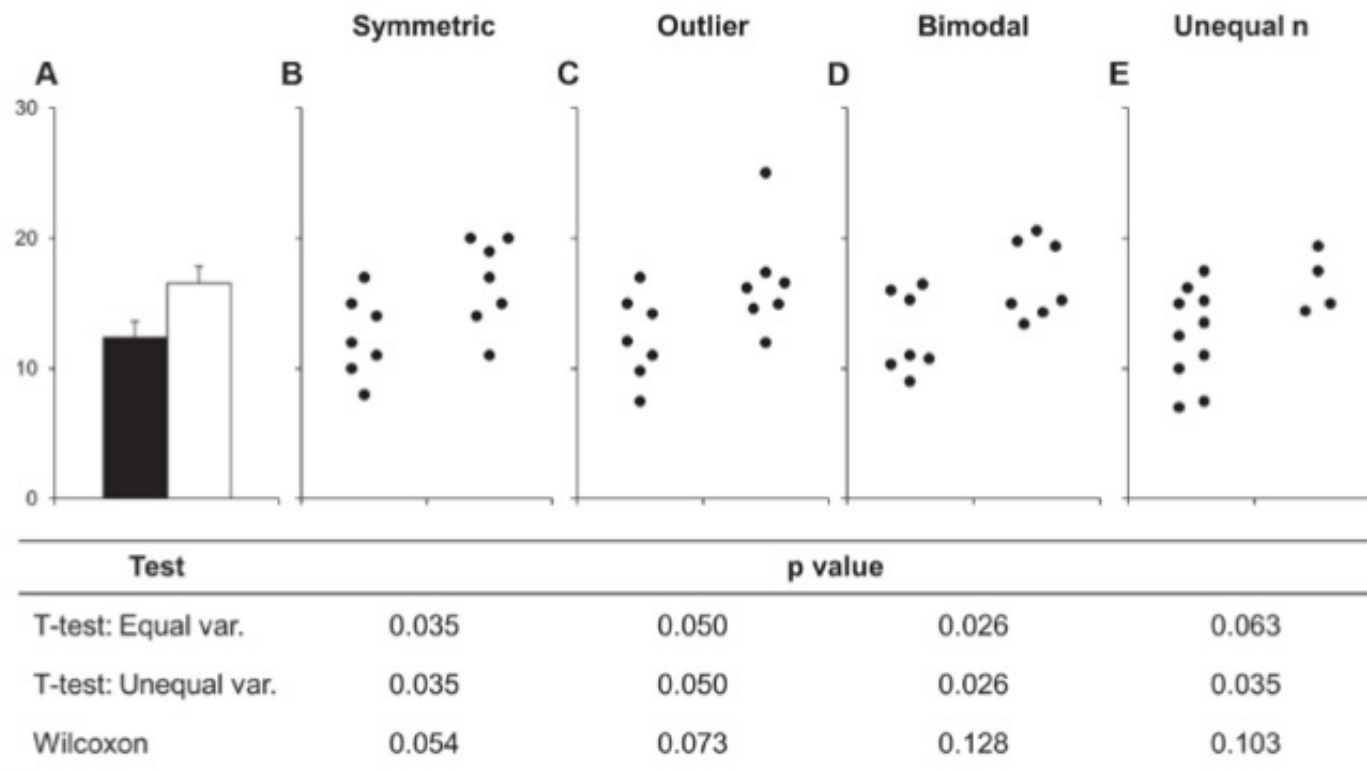


$$\frac{1}{V_0} = \left(\frac{\alpha K_m}{V_{\max}} \right) \frac{1}{[S]} + \frac{\alpha'}{V_{\max}}$$



Type of Inhibition	$K_{m_{app}}$	$V_{max_{app}}$
None	K_m	V_{max}
Competitive (Inhibitor only binds free enzyme)	Increases	V_{max}
Non-competitive (Inhibitor binds free enzyme and ES complex)	K_m	Decreases
Uncompetitive (Inhibitor binds only ES complex)	Decreases	Decreases

Many different datasets can lead to the same bar or line graph



The actual data may suggest different conclusions from the summary statistics alone