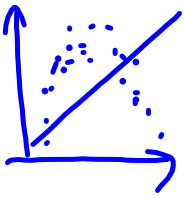
Non-Linear Regression

How would we be able to find the line of best fit for data that is supposed to be curved?



This is **non-linear** regression!

In most cases we don't know what the value of a certain point would be on a line of best fit, so we have to guess! It becomes an estimate.

Modern technology can help us out a lot by finding the info for us. It even provides us with r^2 which is how closely any type of regression curve fits the data.

r^2 is the **coefficient of Determination** defined as

$$\frac{\text{Variation in } y \text{ explained by variation in } x}{\text{Total variation in } y} = r^2$$

$$r^2 = \frac{\Sigma(y_{est} - \bar{y})^2}{\Sigma(y - \bar{y})^2}$$

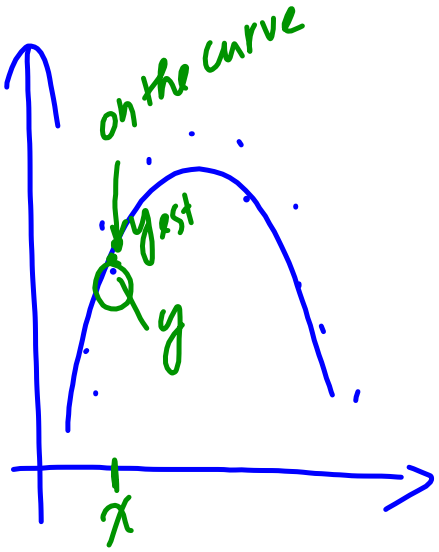
$\bar{y}$ $\bar{x}$

given by equation

$\bar{y}$ is the mean y value

y_{est} is the y value estimated by the best-fit curve for a given x value (using equation)

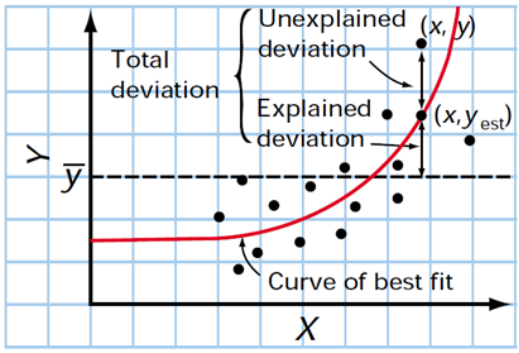
y is the actual observed value for a given x value



$\bar{y}$ → sample mean

y_{est} → point on curve of best fit given x

y_{est} → with regression → equation → substitute x to find y_{est}



The total variation is the sum of the squares of the deviations for all of the individual data points.

numerator

If the curve is a perfect fit, y_{est} and y will be identical for each value of x, and $r^2 = 1$.

If the curve is a poor fit, the total of $(y_{est} - \bar{y})^2$ will be smaller than the total of $(y - \bar{y})^2$, ← denom. therefore r^2 will be close to 0. The curve of best fit will be the one that has the highest value for r^2. r^2 can only have values from **0 to 1**.

Possible Non-Linear Regression:

Power or Polynomial Regression

•uses the line of $y = ax^b$ as the line that models the information.

$y = ax^b$ eg. $y = 2x^2$ $y = 3x^4$

Exponential Regression

•uses the line of $y = ab^x$ or $y = ae^{kx}$, where $e = 2.718\ 28\ldots$, an irrational number commonly used as the base for exponents.

$y = ab^x$ eg. $y = 2(5^x)$

eg. $y = 2e^{3x}$

There are limits to the regression curves.

It can give inaccurate results to your data. We get a different answer if we use a linear regression.

Polynomial regression could give us the wrong info if we perceive it to be exponential!

Your model should show a logical relationship between the variables with the best fit possible.

Don't just check with r^2 value !

For example:

Depreciation of the value of a car – exponential

Trajectory of a rocket – power (quadratic)