1.8 - Regression and Mathematical Models

A mathematical model is a function derived from a real-world situation.

How to Choose the Best Model

1. Look at the general shape of the data points by using a scatterplot.

2. Find models (using the calculator) having the general shape of the data, noting the \( R^2 \) value.

3. Consider whether the model will be used for interpolation (analyzing points within the domain of the data values) or extrapolation (analyzing points beyond the domain of the data values). If the model will be used for extrapolation, the shape of the curve beyond the data points must also be considered.

4. Keep the model as simple as possible.

Finding Mathematical Models on the Calculator

1. Standardize the data.

2. Enter the data into \( L_1 \) and \( L_2 \) by pressing [\text{STAT}] and selecting 1:Edit.

3. Graph the data points in a scatterplot.

4. Find the equation of a model you wish to consider. Press [\text{STAT}] move right and select \text{CALC} and then go down to one of the following options:
   - 4:LinReg(ax+b) for a linear model \((y = ax + b)\)
   - 5:QuadReg for a quadratic model \((y = ax^2 + bx + c)\)
   - 6:CubicReg for a cubic model \((y = ax^3 + bx^2 + cx + d)\)
   - 7:QuartReg for a quartic model \((y = ax^4 + bx^3 + cx^2 + dx + e)\)
   - 9:LnReg for a logarithmic model \((y = a + b \ln x)\)
   - 0:ExpReg for an exponential model \((y = a \cdot b^x)\)
   - A:PwrReg for a power model \((y = a \cdot x^b)\)
   - B:Logistic for a logistic model \((y = \frac{c}{1 + ae^{bx}})\)

5. Store your model in the calculator, so you can compare models to one another. To do this, after making your model selection described above, press [\text{VARS}] move right and select \text{Y-VARS}, choose option 1:Function and then choose a function name.
Ex: The average value of a single family residence (in thousands of dollars) during the 1970's is given by the following:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>House Value</td>
<td>42</td>
<td>51</td>
<td>63</td>
<td>77</td>
<td>93</td>
</tr>
</tbody>
</table>

Find the model which best fits the data and explain your selection.


<table>
<thead>
<tr>
<th>Month</th>
<th>May</th>
<th>June</th>
<th>July</th>
<th>Aug</th>
<th>Sept</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commerce</td>
<td>200</td>
<td>240</td>
<td>260</td>
<td>300</td>
<td>350</td>
<td>380</td>
<td>440</td>
<td>510</td>
<td>550</td>
<td>625</td>
<td>730</td>
<td>790</td>
<td>875</td>
</tr>
</tbody>
</table>

(a) Standardize the data and determine if a linear, quadratic, cubic, exponential, or logistic model would be the best model, if we are to extrapolate both in the past and in the future. Justify your answer.

(b) Use the model you found to determine the estimated commerce in May 2000.

(c) Use the model you found to find the AROC from May 1996 to May 1997 and compare it to the actual AROC from May 1996 to May 1997.