Performance Evaluation:

Simple Linear Regression Models

Hongwei Zhang
http://www.cs.wayne.edu/~hzhang

Statistics is the art of lying by means of figures.
--- Dr. Wilhelm Stekhel

Acknowledgement: this lecture is partially based on the slides of Dr. Raj Jain.
Simple linear regression models

- **Response Variable**: Estimated variable

- **Predictor Variables**: Variables used to predict the response
  - Also called predictors or factors

- **Regression Model**: Predict a response for a given set of predictor variables

- **Linear Regression Models**: Response is a linear function of predictors

- **Simple Linear Regression Models**: Only one predictor
Outline

- Definition of a Good Model
- Estimation of Model parameters
- Allocation of Variation
- Standard deviation of Errors
- Confidence Intervals for Regression Parameters
- Confidence Intervals for Predictions
- Visual Tests for verifying Regression Assumption
Outline

- Definition of a Good Model
- Estimation of Model parameters
- Allocation of Variation
- Standard deviation of Errors
- Confidence Intervals for Regression Parameters
- Confidence Intervals for Predictions
- Visual Tests for verifying Regression Assumption
Definition of a good model?
Good models (contd.)

- Regression models attempt to minimize the distance measured vertically between the observation point and the model line (or curve)
  - The length of the line segment is called residual, modeling error, or simply error
- The negative and positive errors should cancel out => Zero overall error
  - Many lines will satisfy this criterion
- Choose the line that minimizes the sum of squares of the errors
Formally,

\[ \hat{y} = b_0 + b_1 x \]

where, \( \hat{y} \) is the predicted response when the predictor variable is \( x \). The parameter \( b_0 \) and \( b_1 \) are fixed regression parameters to be determined from the data.

Given \( n \) observation pairs \( \{(x_1, y_1), \ldots, (x_n, y_n)\} \), the estimated response for the \( i \)-th observation is:

\[ \hat{y}_i = b_0 + b_1 x_i \]

The error is:

\[ e_i = y_i - \hat{y}_i \]
The best linear model minimizes the sum of squared errors (SSE):
\[ \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i)^2 \]
subject to the constraint that the overall mean error is zero:
\[ \sum_{i=1}^{n} e_i = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i) = 0 \]

This is equivalent to the unconstrained minimization of the variance of errors (Exercise 14.1)
Outline

- Definition of a Good Model
- Estimation of Model parameters
- Allocation of Variation
- Standard deviation of Errors
- Confidence Intervals for Regression Parameters
- Confidence Intervals for Predictions
- Visual Tests for verifying Regression Assumption
Estimation of model parameters

- Regression parameters that give minimum error variance are:

$$b_1 = \frac{\sum xy - n\bar{x}\bar{y}}{\sum x^2 - n\bar{x}^2}$$

and

$$b_0 = \bar{y} - b_1 \bar{x}$$

where,

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$$

$$\sum xy = \sum_{i=1}^{n} x_i y_i$$

$$\sum x^2 = \sum_{i=1}^{n} x_i^2$$
Example 14.1

- The number of disk I/O's and processor times of seven programs were measured as: (14, 2), (16, 5), (27, 7), (42, 9), (39, 10), (50, 13), (83, 20)

- For this data: \( n = 7 \), \( \Sigma xy = 3375 \), \( \Sigma x = 271 \), \( \Sigma x^2 = 13,855 \), \( \Sigma y = 66 \), \( \Sigma y^2 = 828 \), \( \bar{x} = 38.71 \), \( \bar{y} = 9.43 \). Therefore,

  \[
  b_1 = \frac{\Sigma xy - n \bar{x} \bar{y}}{\Sigma x^2 - n (\bar{x})^2} = \frac{3375 - 7 \times 38.71 \times 9.43}{13,855 - 7 \times (38.71)^2} = 0.2438
  \]

  \[
  b_0 = \bar{y} - b_1 \bar{x} = 9.43 - 0.2438 \times 38.71 = -0.0083
  \]

- The desired linear model is:

  \[
  \text{CPU time} = -0.0083 + 0.2438(\text{Number of Disk I/O's})
  \]
Example (contd.)
Example (contd.)

- **Error Computation**

<table>
<thead>
<tr>
<th>Disk I/O’s $x_i$</th>
<th>CPU Time $y_i$</th>
<th>Estimate $\hat{y}_i=b_0+b_1 x_i$</th>
<th>Error $e_i=y_i-\hat{y}_i$</th>
<th>Error$^2$ $e_i^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>2</td>
<td>3.4043</td>
<td>-1.4043</td>
<td>1.9721</td>
</tr>
<tr>
<td>16</td>
<td>5</td>
<td>3.8918</td>
<td>1.1082</td>
<td>1.2281</td>
</tr>
<tr>
<td>27</td>
<td>7</td>
<td>6.5731</td>
<td>0.4269</td>
<td>0.1822</td>
</tr>
<tr>
<td>42</td>
<td>9</td>
<td>10.2295</td>
<td>-1.2295</td>
<td>1.5116</td>
</tr>
<tr>
<td>39</td>
<td>10</td>
<td>9.4982</td>
<td>0.5018</td>
<td>0.2518</td>
</tr>
<tr>
<td>50</td>
<td>13</td>
<td>12.1795</td>
<td>0.8205</td>
<td>0.6732</td>
</tr>
<tr>
<td>83</td>
<td>20</td>
<td>20.2235</td>
<td>-0.2235</td>
<td>0.0500</td>
</tr>
<tr>
<td>$\Sigma$</td>
<td>271</td>
<td>66</td>
<td>66.0000</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Derivation of regression parameters?

- The error in the \( i \)th observation is:
  \[
e_i = y_i - \hat{y}_i = y_i - (b_0 + b_1 x_i)
  \]

- For a sample of \( n \) observations, the mean error is:
  \[
  \bar{e} = \frac{1}{n} \sum_i e_i = \frac{1}{n} \sum_i \{y_i - (b_0 + b_1 x_i)\}
  = \bar{y} - b_0 - b_1 \bar{x}
  \]

- Setting mean error to zero, we obtain:
  \[
  b_0 = \bar{y} - b_1 \bar{x}
  \]

- Substituting \( b_0 \) in the error expression, we get:
  \[
e_i = y_i - \bar{y} + b_1 \bar{x} - b_1 x_i = (y_i - \bar{y}) - b_1 (x_i - \bar{x})
  \]
Derivation (contd.)

The sum of squared errors SSE is:

\[ SSE = \sum_{i=1}^{n} e_i^2 \]

\[ = \sum_{i=1}^{n} \left\{ (y_i - \bar{y})^2 - 2b_1 (y_i - \bar{y}) (x_i - \bar{x}) + b_1^2 (x_i - \bar{x})^2 \right\} \]

\[ \frac{SSE}{n-1} = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2 - 2b_1 \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y}) (x_i - \bar{x}) + b_1^2 \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \]

\[ = s_y^2 - 2b_1 s_{xy}^2 + b_1^2 s_x^2 \]
Derivation (contd.)

- Differentiating this equation with respect to $b_1$ and equating the result to zero:

$$
\frac{d(\text{SSE})}{db_1} = -2s_{xy}^2 + 2b_1 s_x^2 = 0
$$

- That is,

$$
b_1 = \frac{s_{xy}^2}{s_x^2} = \frac{\sum xy - n \bar{x} \bar{y}}{\sum x^2 - n(\bar{x})^2}
$$
Least Squares Regression vs. Least Absolute Deviations Regression?

<table>
<thead>
<tr>
<th>Least Squares Regression</th>
<th>Least Absolute Deviations Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not very robust to outliers</td>
<td>Robust to outliers</td>
</tr>
<tr>
<td>Simple analytical solution</td>
<td>No analytical solving method</td>
</tr>
<tr>
<td></td>
<td>(have to use iterative computation-intensive method)</td>
</tr>
<tr>
<td>Stable solution</td>
<td>Unstable solution</td>
</tr>
<tr>
<td>Always one unique solution</td>
<td>Possibly multiple solutions</td>
</tr>
</tbody>
</table>

The *unstable* property of the method of least absolute deviations means that, for any small horizontal adjustment of a data point, the regression line may jump a large amount. In contrast, the least squares solutions is *stable* in that, for any small horizontal adjustment of a data point, the regression line will always move only slightly, or continuously.
Outline

- Definition of a Good Model
- Estimation of Model parameters
- Allocation of Variation
- Standard deviation of Errors
- Confidence Intervals for Regression Parameters
- Confidence Intervals for Predictions
- Visual Tests for verifying Regression Assumption
Allocation of variation

Error variance without Regression = Variance of the response

\[
\text{Error} = \epsilon_i = \text{Observed Response} - \text{Predicted Response} = y_i - \bar{y}
\]

and

Variance of Errors without regression = \[
\frac{1}{n-1} \sum_{i=1}^{n} \epsilon_i^2
\]

\[
= \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2
\]

= Variance of y
The sum of squared errors without regression would be:

\[ \sum_{i=1}^{n} (y_i - \bar{y})^2 \]

This is called **total sum of squares** or (SST). It is a measure of \( y \)'s variability and is called **variation** of \( y \). SST can be computed as follows:

\[ \text{SST} = \sum_{i=1}^{n} (y_i - \bar{y})^2 = \left( \sum_{i=1}^{n} y_i^2 \right) - n\bar{y}^2 = SSY - SS0 \]

Where, SSY is the sum of squares of \( y \) (or \( \Sigma y^2 \)). SS0 is the sum of squares of \( \bar{y} \) and is equal to \( n\bar{y}^2 \).
 Allocation of variation (contd.)

- The difference between SST and SSE is the sum of squares explained by the regression. It is called SSR:
  \[ \text{SSR} = \text{SST} - \text{SSE} \]
  or
  \[ \text{SST} = \text{SSR} + \text{SSE} \]

- The fraction of the variation that is explained determines the goodness of the regression and is called the coefficient of determination, \( R^2 \):
  \[ R^2 = \frac{\text{SSR}}{\text{SST}} = \frac{\text{SST} - \text{SSE}}{\text{SST}} \]

Variation not explained by the regression
Allocation of variation (contd.)

- The higher the value of $R^2$, the better the regression. $R^2=1 \Rightarrow$ Perfect fit; $R^2=0 \Rightarrow$ No fit

- Shortcut formula for SSE:

$$\text{SSE} = \sum y^2 - b_0 \sum y - b_1 \sum xy$$
Example

- For the disk I/O-CPU time data of Example 14.1:

  \[
  \begin{align*}
  \text{SSE} & = \sum y^2 - b_0 \sum y - b_1 \sum xy \\
  & = 828 + 0.0083 \times 66 - 0.2438 \times 3375 = 5.87 \\
  \text{SST} & = \text{SSY} - \text{SS0} = \sum y^2 - n(\bar{y})^2 \\
  & = 828 - 7 \times (9.43)^2 = 205.71 \\
  \text{SSR} & = \text{SST} - \text{SSE} = 205.71 - 5.87 = 199.84 \\
  R^2 & = \frac{\text{SSR}}{\text{SST}} = \frac{199.84}{205.71} = 0.9715
  \end{align*}
  \]

- The regression explains 97% of CPU time's variation.
Outline

- Definition of a Good Model
- Estimation of Model parameters
- Allocation of Variation
- Standard deviation of Errors
- Confidence Intervals for Regression Parameters
- Confidence Intervals for Predictions
- Visual Tests for verifying Regression Assumption
Standard deviation of errors

- Since errors are obtained after calculating two regression parameters from the data, errors have \( n-2 \) degrees of freedom.

- \( \frac{\text{SSE}}{(n-2)} \) is called **mean squared errors** or (MSE)

\[
s^2_e = \frac{\text{SSE}}{n-2}
\]

- Standard deviation of errors = square root of MSE

**Note:**
- SSY has \( n \) degrees of freedom since it is obtained from \( n \) independent observations without estimating any parameters.
- SS0 has just one degree of freedom since it can be computed simply from \( \bar{y} \).
- SST has \( n-1 \) degrees of freedom, since one parameter \( \bar{y} \) must be calculated from the data before SST can be computed.
Standard deviation of errors (contd.)

- SSR, which is the difference between SST and SSE, has the remaining one degree of freedom.

- Overall,

\[
\frac{\text{SST}}{n - 1} = \frac{\text{SSY}}{n} - \frac{\text{SSO}}{1} = \frac{\text{SSR}}{1} + \frac{\text{SSE}}{n - 2}
\]

- Notice that the degrees of freedom add just the way the sums of squares do.
Example

- For the disk I/O-CPU data of Example 14.1, the degrees of freedom of the sums are:

  \[
  \begin{align*}
  SS : & \quad SST = SSY - SS_0 = SSR + SSE \\
  & \quad 205.71 = 828 - 622.29 = 199.84 + 5.87 \\
  DF : & \quad 6 = 7 - 1 = 1 + 5
  \end{align*}
  \]

- The mean squared error is:

  \[
  MSE = \frac{SSE}{DF \text{ for Errors}} = \frac{5.87}{5} = 1.17
  \]

- The standard deviation of errors is:

  \[
  s_e = \sqrt{MSE} = \sqrt{1.17} = 1.08
  \]
Outline

- Definition of a Good Model
- Estimation of Model parameters
- Allocation of Variation
- Standard deviation of Errors
- Confidence Intervals for Regression Parameters
- Confidence Intervals for Predictions
- Visual Tests for verifying Regression Assumption
Cl's for regression parameters

- Regression coefficients $b0$ and $b1$ are estimates from a single sample of size $n$. => 1) Random; 2) Using another sample, the estimates may be different.

- If $\beta_0$ and $\beta_1$ are true parameters of the population (i.e., $y = \beta_0 + \beta_1 x$), then the computed coefficients $b0$ and $b1$ are estimates of $\beta_0$ and $\beta_1$, respectively.

- Sample standard deviation of $b0$ and $b1$

$$s_{b_0} = s_e \left[ \frac{1}{n} + \frac{\bar{x}^2}{\sum x^2 - n\bar{x}^2} \right]^{1/2}$$

$$s_{b_1} = \frac{s_e}{\left[ \sum x^2 - n\bar{x}^2 \right]^{1/2}}$$
The 100(1-α)% confidence intervals for b0 and b1 can be computed using \( t[1-\alpha/2; n-2] \) --- the \( 1-\alpha/2 \) quantile of a t variate with \( n-2 \) degrees of freedom. The confidence intervals are:

\[
b_0 \pm ts_{b_0}\\
\text{And}\\
b_1 \pm ts_{b_1}
\]

If a confidence interval includes zero, then the regression parameter cannot be considered different from zero at the 100(1-α)% confidence level.
Example

- For the disk I/O and CPU data of Example 14.1, we have $n=7$, $\bar{x}=38.71$, $\sum x^2=13,855$, and $s_e=1.0834$.

- Standard deviations of $b_0$ and $b_1$ are:

$$s_{b_0} = s_e \left[ \frac{1}{n} + \frac{\bar{x}^2}{\sum x^2 - n\bar{x}^2} \right]^{1/2}$$

$$= 1.0834 \left[ \frac{1}{7} + \frac{(38.71)^2}{13,855 - 7 \times 38.71 \times 38.71} \right]^{1/2} = 0.8311$$

$$s_{b_1} = \frac{s_e}{[\sum x^2 - n\bar{x}^2]^{1/2}}$$

$$= \frac{1.0834}{[13,855 - 7 \times 38.71 \times 38.71]^{1/2}} = 0.0187$$
Example (contd.)

- The 0.95-quantile of a $t$-variate with 5 degrees of freedom is 2.015

  => 90% confidence interval for $b_0$ is:
  
  $$-0.0083 \pm (2.015)(0.8311) = -0.0083 \pm 1.6747 = (-1.6830, 1.6663)$$

  Since, the confidence interval includes zero, the hypothesis that this parameter is zero cannot be rejected at 0.10 significance level => $b_0$ is essentially zero.

  => 90% Confidence Interval for $b_1$ is:
  
  $$0.2438 \pm (2.015)(0.0187) = 0.2438 \pm 0.0376 = (0.2061, 0.2814)$$

  Since the confidence interval does not include zero, the slope $b_1$ is significantly different from zero at this confidence level.
Case study 14.1: remote procedure call

<table>
<thead>
<tr>
<th></th>
<th>UNIX</th>
<th></th>
<th>ARGUS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>Time</td>
<td>Data</td>
<td>Time</td>
<td></td>
</tr>
<tr>
<td>Bytes</td>
<td></td>
<td>Bytes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>26.4</td>
<td>92</td>
<td>32.8</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>26.4</td>
<td>92</td>
<td>34.2</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>26.4</td>
<td>92</td>
<td>32.4</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>26.2</td>
<td>92</td>
<td>34.4</td>
<td></td>
</tr>
<tr>
<td>234</td>
<td>33.8</td>
<td>348</td>
<td>41.4</td>
<td></td>
</tr>
<tr>
<td>590</td>
<td>41.6</td>
<td>604</td>
<td>51.2</td>
<td></td>
</tr>
<tr>
<td>846</td>
<td>50.0</td>
<td>860</td>
<td>76.0</td>
<td></td>
</tr>
<tr>
<td>1060</td>
<td>48.4</td>
<td>1074</td>
<td>80.8</td>
<td></td>
</tr>
<tr>
<td>1082</td>
<td>49.0</td>
<td>1074</td>
<td>79.8</td>
<td></td>
</tr>
<tr>
<td>1088</td>
<td>42.0</td>
<td>1088</td>
<td>58.6</td>
<td></td>
</tr>
<tr>
<td>1088</td>
<td>41.8</td>
<td>1088</td>
<td>57.6</td>
<td></td>
</tr>
<tr>
<td>1088</td>
<td>41.8</td>
<td>1088</td>
<td>59.8</td>
<td></td>
</tr>
<tr>
<td>1088</td>
<td>42.0</td>
<td>1088</td>
<td>57.4</td>
<td></td>
</tr>
</tbody>
</table>
Case study (contd.)

UNIX:

![Graph showing the relationship between data bytes and elapsed time in milliseconds. The graph indicates a positive correlation.](image)
Case study (contd.)

- ARGUS:
Case study (contd.)

- Best linear models are:

  \[
  \text{Time on UNIX} = 0.030 \text{ (Data size in bytes)} + 24 \\
  \text{Time on ARGUS} = 0.034 \text{ (Data size in bytes)} + 30
  \]

- The regressions explain 81% and 75% of the variation, respectively.

  Does ARGUS takes larger time per byte as well as a larger set up time per call than UNIX?
Intervals for intercepts overlap while those of the slopes do not. => Set up times are not significantly different in the two systems while the per byte times (slopes) are different.
Outline

- Definition of a Good Model
- Estimation of Model parameters
- Allocation of Variation
- Standard deviation of Errors
- Confidence Intervals for Regression Parameters
- Confidence Intervals for Predictions
- Visual Tests for verifying Regression Assumption
Cl for predications

\[ \hat{y}_p = b_0 + b_1 x_p \]

- This is only the mean value of the predicted response. Standard deviation of the mean of a future sample of m observations is:

\[ s_{\hat{y}_m} = s_e \left[ \frac{1}{m} + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{\sum x^2 - n\bar{x}^2} \right]^{1/2} \]

- m = 1 ⇒ Standard deviation of a single future observation:

\[ s_{\hat{y}_1} = s_e \left[ 1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{\sum x^2 - n\bar{x}^2} \right]^{1/2} \]
CI for predications (contd.)

- \( m = \infty \Rightarrow \) Standard deviation of the mean of a large number of future observations at \( x_p \):

\[
s_{\hat{y}_p} = s_e \left[ \frac{1}{n} + \frac{(x_p - \bar{x})^2}{\Sigma x^2 - n\bar{x}^2} \right]^{1/2}
\]

- \( 100(1-\alpha)\% \) confidence interval for the mean can be constructed using a \( t \) quantile read at \( n-2 \) degrees of freedom.
CI for predications (contd.)

- Standard deviation of the prediction is minimal at the center of the measured range (i.e., when \( x = \bar{x} \)); Goodness of the prediction decreases as we move away from the center.
Example

- Using the disk I/O and CPU time data of Example 14.1, let us estimate the CPU time for a program with 100 disk I/O's.

  \[ \text{CPU time} = -0.0083 + 0.2438(\text{Number of disk I/O's}) \]

- For a program with 100 disk I/O's, the mean CPU time is:

  \[ \text{CPU time} = -0.0083 + 0.2438(100) = 24.3674 \]

  Standard deviation of errors \( s_e = 1.0834 \)
Example (contd.)

- The standard deviation of the predicted mean of a large number of observations is:

$$s_{\hat{y}_p} = 1.0834 \left[ \frac{1}{7} + \frac{(100 - 38.71)^2}{13,855 - 7(38.71)^2} \right]^{1/2} = 1.2159$$

- From Table A.4, the 0.95-quantile of the t-variate with 5 degrees of freedom is 2.015.

  $\Rightarrow$ 90% CI for the predicted mean

  $$= 24.3674 \pm (2.015)(1.2159)$$

  $$= (21.9174, 26.8174)$$
Example (contd.)

- CPU time of a single future program with 100 disk I/O's:

\[
s_{\hat{y}_{1p}} = 1.0834 \left[ 1 + \frac{(100 - 38.71)^2}{13,855 - 7(38.71)^2} \right]^{1/2} = 1.6286
\]

- 90% CI for a single prediction:

\[
= 24.3674 \pm (2.015)(1.6286)
= (21.0858, 27.6489)
\]
Outline

- Definition of a Good Model
- Estimation of Model parameters
- Allocation of Variation
- Standard deviation of Errors
- Confidence Intervals for Regression Parameters
- Confidence Intervals for Predictions
- Visual Tests for verifying Regression Assumption
Visual test for regress assumptions

Regression assumptions:

- The true relationship between the response variable $y$ and the predictor variable $x$ is linear.

- The predictor variable $x$ is non-stochastic and it is measured without any error.

- The model errors are statistically independent.

- The errors are normally distributed with zero mean and a constant standard deviation.
Visual test for linear relationship

- Scatter plot of $y$ versus $x \Rightarrow$ Linear or nonlinear relationship
Visual test for independent errors

- Scatter plot of $\varepsilon_i$ versus the predicted response $\hat{y}_i$

- Any trend would imply the dependence of errors on predictor variable $\Rightarrow$ curvilinear model or transformation

- In practice, dependence can be proven yet independent cannot
Visual test for independent errors (contd.)

- Plot the residuals as a function of the experiment number

- Any trend would imply that other factors (such as environmental conditions or side effects) should be considered in the modeling
Visual test for “normal distribution of errors”?

- Prepare a normal quantile-quantile plot of errors. Linear ⇒ the assumption is satisfied.
Visual test for constant standard deviation of errors

- Also known as homoscedasticity

- Trend ⇒ Try curvilinear regression or transformation
Example

For the disk I/O and CPU time data of Example 14.1

1. Relationship is linear
2. No trend in residuals ⇒ Seem independent
3. Linear normal quantile-quantile plot
Another example: RPC performance

1. Larger errors at larger responses
2. Normality of errors is questionable
Summary

- Definition of a Good Model
- Estimation of Model parameters
- Allocation of Variation
- Standard deviation of Errors
- Confidence Intervals for Regression Parameters & Predictions
- Visual Tests for verifying Regression Assumption
Exercise

The time to encrypt a $k$ byte record using an encryption technique is shown in the following table. Fit a linear regression model to this data. Use visual tests to verify the regression assumptions.

<table>
<thead>
<tr>
<th>Record Size</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>128</td>
<td>386</td>
</tr>
<tr>
<td>256</td>
<td>850</td>
</tr>
<tr>
<td>384</td>
<td>1544</td>
</tr>
<tr>
<td>512</td>
<td>3035</td>
</tr>
<tr>
<td>640</td>
<td>6650</td>
</tr>
<tr>
<td>768</td>
<td>13,887</td>
</tr>
<tr>
<td>896</td>
<td>28,059</td>
</tr>
<tr>
<td>1024</td>
<td>50,916</td>
</tr>
</tbody>
</table>