Lecture 9: Plots, Graphs, and Pictures

Statistics 10
Colin Rundel
February 8, 2012

Chapter 8
Normal Distribution

Normal probability plot

- A histogram and normal probability plot of a sample of 100 male heights.
- The points appear to jump in increments in the normal probability plot since the observations are rounded to the nearest whole inch.

Anatomy of a normal probability plot

- Empirical quantiles (based on data) are plotted on the y-axis of a normal probability plot, and theoretical quantiles (following a normal distribution) on the x-axis.
- If there is a one-to-one relationship between the empirical and the theoretical quantiles, it means that the data follow a nearly normal distribution.
- Since a one-to-one relationship would appear as a straight line on a scatter plot, the closer the points are to a perfect straight line, the more confident we can be that the data follow the normal model.
- Constructing a normal probability plot requires calculating percentiles and corresponding z-scores for each observation, which is tedious. So we generally use software to construct these plots.

Construct a normal probability plot for the data set given below and determine if the data follow an approximately normal distribution.

3.46, 4.02, 5.09, 2.33, 6.47

<table>
<thead>
<tr>
<th>Observation</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_i$</td>
<td>2.33</td>
<td>3.46</td>
<td>4.02</td>
<td>5.09</td>
<td>6.47</td>
</tr>
<tr>
<td>Percentile</td>
<td>$\frac{i}{n+1}$</td>
<td>0.17</td>
<td>0.33</td>
<td>0.50</td>
<td>0.67</td>
</tr>
<tr>
<td>Corresponding $Z_i$</td>
<td>-0.95</td>
<td>-0.44</td>
<td>0</td>
<td>0.44</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Since the points on the normal probability plot seem to follow a straight line we can say that the distribution is nearly normal.
Below is a histogram and normal probability plot for the NBA heights from the 2008-9 season. Do these data appear to follow a normal distribution?

Right Skew - If the plotted points appear to bend up and to the left of the normal line that indicates a long tail to the right.

Left Skew - If the plotted points bend down and to the right of the normal line that indicates a long tail to the left.

Short Tails - An S shaped-curve indicates shorter than normal tails, i.e. narrower than expected.

Long Tails - A curve which starts below the normal line, bends to follow it, and ends above it indicates long tails. That is, you are seeing more variance than you would expect in a normal distribution, i.e. wider than expected.

Why do we create plots?

- Tell a story
- Valid Representation of Data
- Information density
  - Remember Arbuthnot? Numbers vs Line plot
  - Boxplot vs 5 Number summary

Common Issues

- No labeling on axes
- Not starting at zero as a way to exaggerate trend
- Change(s) / Discontinuities in labeling
- Misleading units
- Using flawed data

Not in Utts:

- Data / Ink ratio
- Just because Excel lets you do it doesn’t mean you should
- Trying to convey too little or too much
Univariate Plots

- Categorical
  - “Good” - Bar plot
  - “Bad” - Pie chart, donut chart, pictogram

- Numerical
  - “Good” - boxplot (outliers), histogram (shape), violin plot (shape and outliers)
  - “Bad” - Stem and leaf plot, dot plot

Bivariate Plots

- Categorical vs Categorical
  - “Good” - Mosaic plot, stacked or grouped bar plot
  - “Bad” -

- Numerical vs Categorical
  - “Good” - Side by side boxplot or violin plot
  - “Bad” -

- Numerical vs Numerical
  - “Good” - Scatter plot, line plot
  - “Bad” -
Problem with Pie Charts

Crashes by OS Version Normalized (12/1 - 12/15)

Problem with Pie Charts

Axis Placement - What story to tell

http://junkcharts.typepad.com/junk_charts/2012/02/a-data-mess-outduels-the-pie-chart-disaster-for-our-attenting.html


Chapter 9  
Some Example from Junk Charts

Minard and Napoleon’s Russian Campaign

Florence Nightingale’s Rose Diagram

Some Example from Junk Charts

Florence Nightingale’s Rose Diagram Remixed

Hans Rosling and Gap Minder

http://www.gapminder.org/videos/