Principios de Estadística

Leonardo Collado Torres y María Gutiérrez Arcelus

Licenciatura en Ciencias Genómicas, UNAM

www.lcg.unam.mx/~lcollado/index.php

www.lcg.unam.mx/~mgutierr/index.php

Cuernavaca, México

Febrero - Junio, 2009

1 Graphical Tools

Principios de
Estadística

Graphical
Tools

- Estas diapositivas corresponden a las 28, 29 y 30 de la presentación EDA de Jim.

boxplots

- boxplot : A plotting method for generating Tukey's boxplots.
- Excellent for comparing location shifts of k distributions of varying size.
- Assess skewness and spread of either of one or more distributions.
- Boxplots are often a much better summary of a distribution than are histograms as they do not suffer from either bandwidth choice or the need to have large data sets.

Example

What does skewness look like on a boxplot, spread? can we generate some data to exemplify these things? (hint: remember all of the random number generators which we talked about in the first lecture)

Anatomy of a boxplot

- A, B : lower/upper adjacent values:

$$
\begin{aligned}
r &\triangleq |q_{75} - q_{25}| & (1) \\
A &= \inf\{x_i : x_i > q_{25} - 1.5r\} & (2) \\
B &= \sup\{x_i : x_i < q_{75} + 1.5r\} & (3)
\end{aligned}
$$

Anatomy of a boxplot

The anatomy of a Boxplot