

Chapter II-13 — Graphs

tion. By default, Igor draws the whiskers to the extreme data points, but there are eight different options for whisker length.

The width of the box is not meaningful, but you can control it to make a pleasing display. The width can be expressed as a fraction, in which case it is the fraction of the width of the plot area. If the width is greater than one, it is taken to be an absolute width in points. By default the width is $1/(2^*N)$, where N is the number of datasets or box plots included in the trace.

It is common to show the actual data points when they are far from the center of the distribution. Tukey defines "outliers" and "far outliers". Igor allows you to show all data points, only outliers and far outliers, or only far outliers. You can also select different markers, sizes and colors for each category of data points.

Box Plot Fences

Some options use Tukey's "fences" to define the length of the whiskers and the boundaries defining outliers and far outliers. Tukey also uses the term "hinge" to refer to the the 25th and 75th percentiles.

Inner fences are defined as

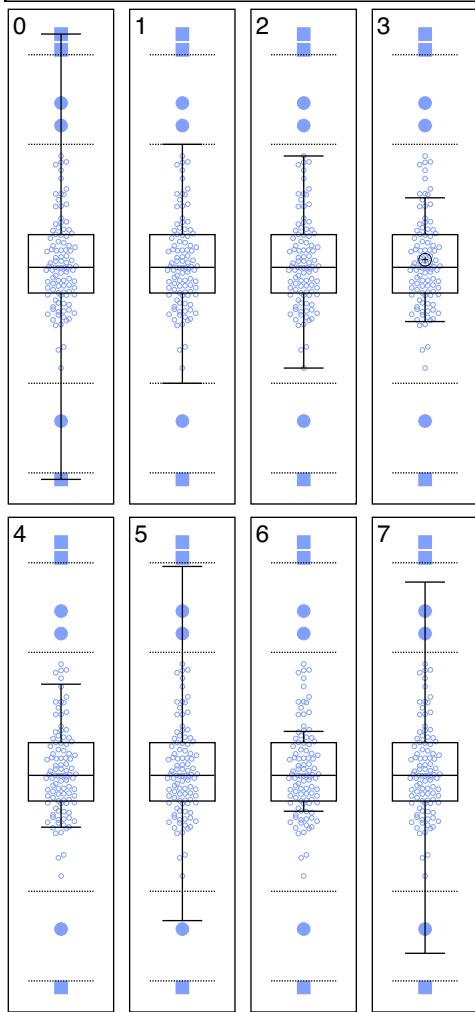
```
inner fence = upper hinge + 1.5*IQR and lower hinge - 1.5*IQR
```

Outer fences are defined as

```
outer fence = upper hinge + 3*IQR and lower hinge - 3*IQR
```

Box Plot Whisker Length

The following figure shows the options for defining the whisker length. The fences are shown for reference to Tukey's definitions, and the outlier method (see **Box Plot Outlier Methods** on page II-334) is option 0:



- Option 0: Minimum and maximum data points (the default).
- Option 1: Inner fences.
- Option 2: "Adjacent points", which Tukey defines as the most extreme data points inside the inner fences. That is, the most extreme data points that are not outliers, if you define outliers the way Tukey does.
- Option 3: One standard deviation from the mean of the data. The light circle with plus shows the mean value.
- Option 4: The 9th and 91st percentiles.
- Option 5: The 2nd and 98th percentiles.
- Option 6: Arbitrary percentiles. Here the ends are set to the 20th and 80th percentiles in order to make it look different from the other options.
- Option 7: An arbitrary factor times one standard deviation from the mean. In this case, the factor is 3.

If the data are normally distributed, the 2nd, 9th, 25th, 50th, 75th, 91st, and 98th percentiles should be equally spaced.

The examples show all the data points but it is common to show only outliers and far outliers. The data points are displayed with "jitter"—data points that would overlap are offset horizontally so that each data point can be seen. You can control the maximum offset by specifying the jitter amount in units of fractions