The 'non-overlapping confidence interval' test.

I am a strong believer in including precision information in plots. If you are plotting an estimate for which you can get a confidence interval, then adding the confidence interval leads to the potential for a much better imterpretation of the plot.

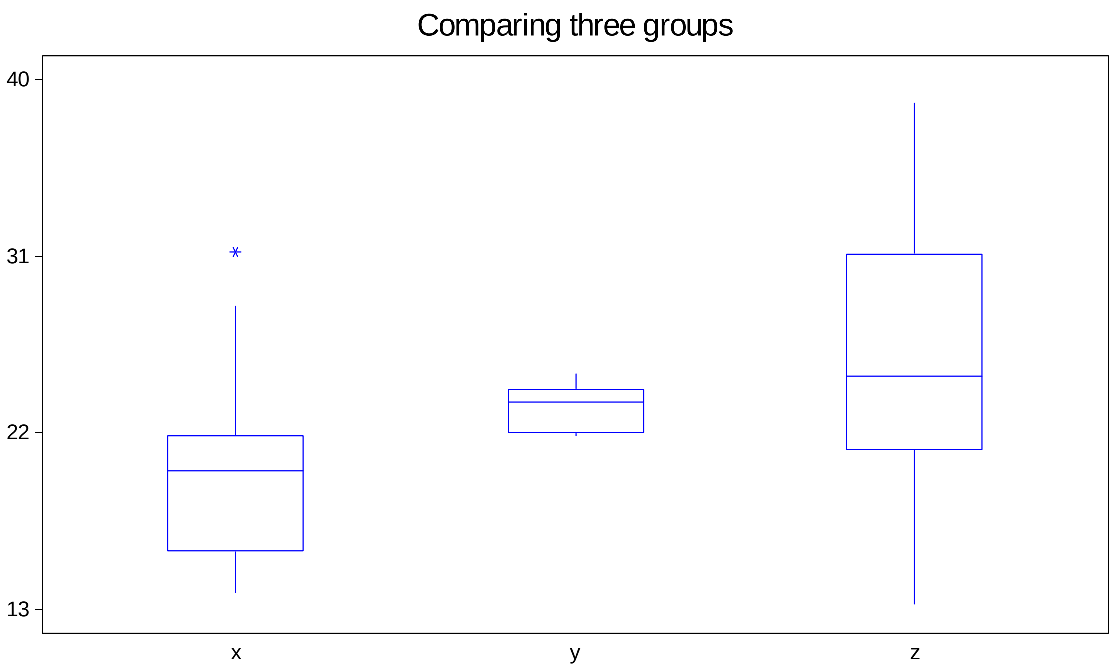
This leads to the idea.

- You have a plot showing the means of samples from two groups along with 95% confidence intervals for the means.
- If the intervals overlap, you say the two groups are not significantly different.
- If the intervals do not overlap, you declare the groups significantly different.

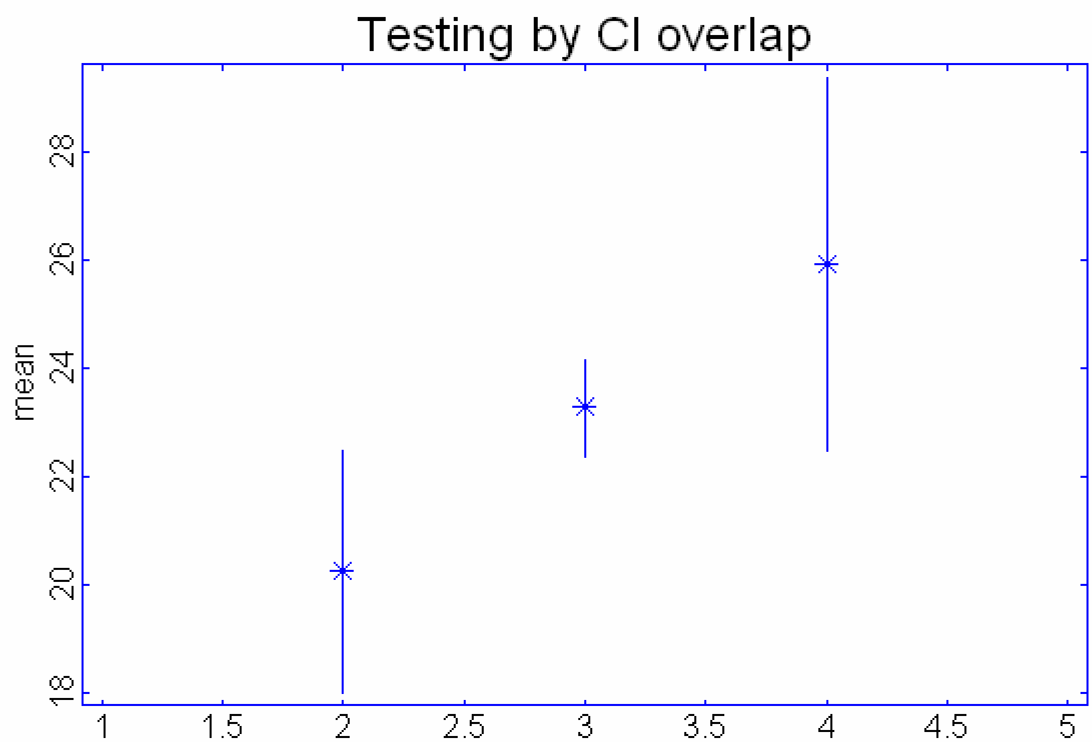
Is this method OK?

Example: three groups X, Y, Z. Summary stats

Variable	N	Mean	SD	SE Mean
x	20	20.261	4.8287	1.0797
y	9	23.281	1.1919	0.3973
z	20	25.940	7.3843	1.6512


Comparing three groups
40
31
22
13
x
y
z

20 x, 20 z, 9 y


Testing by CI overlap
mean
18
20
22
24
26
28
1
1.5
2
2.5
3
3.5
4
4.5
5

Two-Sample T Tests for y vs x

Assumption	T	DF	P
Unequal Variances	2.62	23.5	0.0150

Two-Sample T Tests for x vs z

Assumption	T	DF	P
Equal Variances	-2.88	38	0.0065
Unequal Variances	-2.88	32.7	0.0070

Two-Sample T Tests for z vs y

Assumption	T	DF	P
Equal Variances	1.06	27	0.2969
Unequal Variances	1.57	21.1	0.1323

The math. X and Y CI's overlap depending on whether $|\overline{X} - \overline{Y}| > \left(t_1 \frac{s_x}{\sqrt{n_x}} + t_2 \frac{s_y}{\sqrt{n_y}} \right)$

The two-sample *t* test accepts or rejects on the basis of whether $t_* \sqrt{\left(\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y} \right)} t_* \sqrt{\left(\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y} \right)}$

The *t* cutoffs t_1, t_2 , t_* satisfy $t_* < \min(t_1, t_2)$, though in large samples the numbers may be about equal.

Even if the t were all equal, Cauchy-Schwartz tells you that $\left(t_1\frac{s_x}{\sqrt{n_x}}+t_2\frac{s_y}{\sqrt{n_y}}\right)>t_*\sqrt{\left(\frac{s_x^2}{n_x}+\frac{s_y^2}{n_y}\right)}$. Thus the 'overlapping confidence interval test' is always conservative. How conservative? Some extremes

(a) If $s_y / \sqrt{n_y}$ is 'small' and n_x is 'large', then the overlapping confidence approach is not very conservative.

(b) If the two samples are of similar size and variability, then $\left(\frac{s_x}{\sqrt{n_x}}+\frac{s_y}{\sqrt{n_y}}\right)\approx\sqrt{2}\sqrt{\left(\frac{s_x^2}{n_x}+\frac{s_y^2}{n_y}\right)}$

If the samples are 'large' so that $t_1 \approx t_2 \approx t_*$, then the overlapping CI approach is like using a t cutoff that is too big by a factor of 1.41. So for example, 95% CIs correspond to a cutoff around 2.8, corresponding to a P value around 0.5% -- one tenth of the nominal

If the samples are 'small', then the difference can be even more extreme. For example 95% CIs when you have two groups of size 5 use $t_1 = 2.78$. The nonoverlapping CIs turn out to correspond to a P value of 0.4%.