THE UNIVERSITY OF MINNESOTA

Statistics 5401 October 5, 2005

Comparison of Bonferronized t and Ellipsoid Based Confidence Limits

This handout illustrates numerically that simultaneous confidence limits based on Bonferronized Student's *t* are shorter than simultaneous confidence limits derived from an elliptical or ellipsoidal confidence region in several important cases. Consequently Bonferronized *t* limits are almost always preferable to "ellipsoidal" limits,

Simultaneous limits for linear combinations of means

Suppose $\bar{\mathbf{x}}$ and $\mathbf{S}$ are the sample mean and unbiased sample variance matrix of a random sample of size n from $N_p(\boldsymbol{\mu},\boldsymbol{\Sigma})$. Then $\bar{\mathbf{x}}$ is $N_p(\boldsymbol{\mu},n^{-1}\boldsymbol{\Sigma})$ and $\mathbf{S}$ is distributed as multiple of a Wishart matrix, specifically $\mathbf{S} \sim f_e^{-1}W_p(f_e, \boldsymbol{\Sigma})$, $f_e = n - 1$. The estimated variance matrix of $\bar{\mathbf{x}}$ is $\hat{V}[\bar{\mathbf{x}}] = n^{-1}\mathbf{S}$.

A $1 - \alpha$ confidence region for the mean vector $\boldsymbol{\mu}$ is the ellipsoid

$$E \equiv \{\boldsymbol{\mu} \mid (\boldsymbol{\mu} - \bar{\mathbf{x}})'(\hat{V}[\bar{\mathbf{x}}])^{-1}(\boldsymbol{\mu} - \bar{\mathbf{x}}) \leq c^2\},$$

where $c^2 \equiv T^2(\alpha) = (pf_e/(f_e - p+1))F_{p,fe-p+1}(\alpha)$. E is centered at $\bar{\mathbf{x}}$ and consists of vectors $\boldsymbol{\mu}$ satisfying the specified inequality. This confidence ellipsoid can be expressed another way:

$$E = \{\boldsymbol{\mu} \mid \ |\mathbf{a}'\boldsymbol{\mu} - \mathbf{a}'\bar{\mathbf{x}}| \leq c \times (\mathbf{a}'\hat{V}[\bar{\mathbf{x}}]\mathbf{a})^{1/2} \text{ for every } \mathbf{a}\}.$$

This reflects the fact that every point $\boldsymbol{\mu}$ in the ellipsoid must lie between all pairs of parallel tangent planes (lines when $p = 2$). A pair of tangent planes orthogonal (perpendicular) to a vector $\mathbf{a}$ is defined as the set

$$\{\boldsymbol{\mu} \mid \mathbf{a}'\boldsymbol{\mu} - \mathbf{a}'\bar{\mathbf{x}} = \pm c\{\mathbf{a}'\hat{V}[\bar{\mathbf{x}}]\mathbf{a})^{1/2}\},$$

where the + sign goes with one side of the ellipsoid and the – sign with the other.

Let $\mathbf{a}$ be any p by 1 vector and let $y_{\mathbf{a}} \equiv \mathbf{a}'\mathbf{x}$ be a linear combination of the variables in $\mathbf{x}$ specified by $\mathbf{a}$. Then $\bar{y}_a = \mathbf{a}'\bar{\mathbf{x}}$ and $\hat{SE}[\bar{y}_a] = \{\mathbf{a}'\hat{V}[\bar{\mathbf{x}}]\mathbf{a})^{1/2} = \{n^{-1}\mathbf{a}'\mathbf{S}\mathbf{a}\}^{1/2} = \sqrt{\mathbf{a}'\mathbf{S}\mathbf{a}}/\sqrt{n}$. Then, because the representation of the confidence ellipsoid E in terms of tangent planes, all intervals of the form

$$\mathbf{a}'\boldsymbol{\mu} = \mathbf{a}'\bar{\mathbf{x}} \pm c\times\hat{SE}[\mathbf{a}'\bar{\mathbf{x}}], \quad c = \{(pf_e/(f_e-p+1))F_{p,fe-p+1}(\alpha)\}^{1/2}$$

simultaneously contain the true values of $\mathbf{a}'\boldsymbol{\mu}$ for *all* $\mathbf{a}$ with probability $1 - \alpha$. Thus such intervals constitute a family of *simultaneous* confidence intervals for $\mathbf{a}'\boldsymbol{\mu}$, in the sense that $\mathbf{a}'\boldsymbol{\mu}$ is contained in the interval for *every* $\mathbf{a}$ with probability $= 1 - \alpha$.

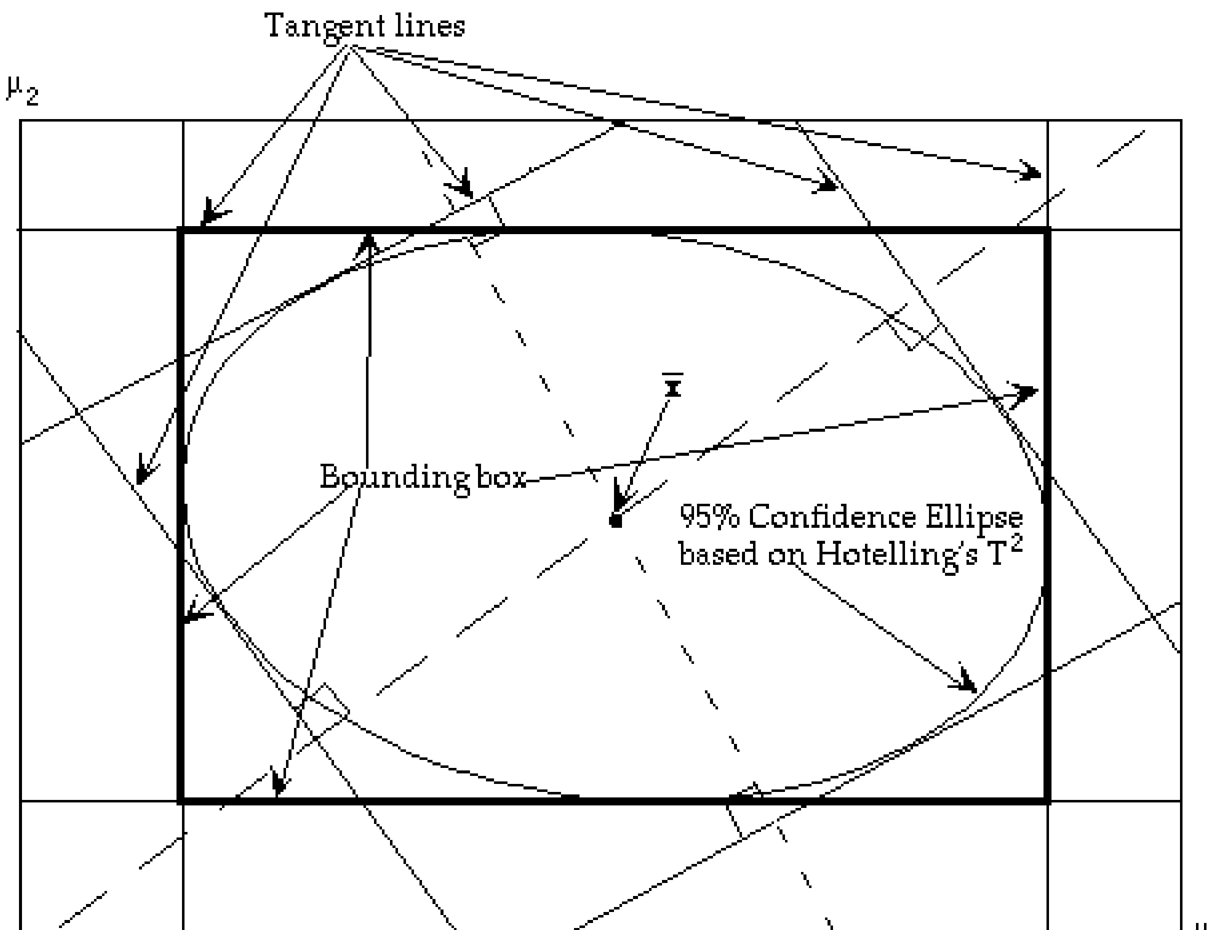
Simultaneous confidence limits for elements of $\boldsymbol{\mu}$

Now you can represent each component of $\boldsymbol{\mu} = [\mu_1, \mu_2, ..., \mu_p]'$ as $\mu_i = \mathbf{e}_i'\boldsymbol{\mu}$, where $\mathbf{e}_i = [0, . . , 0, 1, 0, . . ., 0]'$, with the 1 in row i. Also $\mathbf{e}_i'\bar{\mathbf{x}} = \bar{x}_i$, the sample mean of the i^{th} variable, and $\mathbf{e}_i'\mathbf{S}\mathbf{e}_i = s_{ii}$, the sample variance of x_i. Thus this family of confidence intervals for all linear combinations includes the following intervals for each μ_i:

$$\mu_i = \bar{x}_i \pm c\hat{\mathrm{SE}}[\bar{x}_i] = \bar{x}_i \pm c\,\{s_{ii}/n\}^{1/2},\ c = \{[pf_e/(f_e-p+1)]\mathrm{F}_{p,f_e-p+1}(\alpha)\}^{1/2}.$$

These p intervals are a set of simultaneous confidence intervals for the p components $\mu_1, \mu_2, ..., \mu_p$ of $\boldsymbol{\mu}$, in the sense that you can have confidence at least $1-\alpha$ that every interval contains its respective μ_i's.

Here is a picture displaying some of the tangent lines to a confidence ellipse and directions perpendicular to them, as well as the "bounding box" of the ellipse defined by the horizontal and vertical tangent lines.



Confidence ellipse based on Hotelling's T^2 and tangent lines

The left and right edges of the bounding box mark the ends of the "ellipsoidal" confidence interval for μ_1 and the bottom and top edges mark the ends of the "ellipsoidal" confidence interval for μ_2.

You can get a *different* set of simultaneous confidence intervals for the μ_i by *Bonferronizing* the usual univariate confidence intervals based on Student's t. These Bonferroni intervals are of the form

$$\mu_i = \bar{x}_i \pm t\times\hat{\mathrm{SE}}[\bar{x}_i] = \bar{x}_i \pm t\times\{s_{ii}/n\}^{1/2},\ t = t_{f_e}((\alpha/p)/2),\ i = 1,...,p.$$

Which of the two sets of simultaneous confidence intervals should you prefer?

Common sense says you should prefer intervals that are shorter. Here is a table of the ratio of the lengths of intervals based on Bonferronized t to the lengths of intervals based on the ellipsoid when $\alpha = .05$. Note that the Bonferroni t intervals are uniformly shorter for $p > 1$. This also holds for other α.

$$t_f(.025/p)/\{[pf/(f-p+1)]F_{p,f-p+1}(.05)\}^{1/2}$$

$p \backslash f$	20	40	60	80	100	120	140	160	∞
1	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2	0.890	0.904	0.908	0.910	0.911	0.912	0.912	0.913	0.916
3	0.805	0.833	0.841	0.845	0.847	0.849	0.850	0.851	0.856
4	0.735	0.776	0.788	0.794	0.798	0.800	0.802	0.803	0.811
5	0.674	0.729	0.745	0.753	0.757	0.760	0.762	0.764	0.774
6	0.620	0.688	0.708	0.717	0.723	0.726	0.729	0.731	0.743
7	0.570	0.652	0.676	0.687	0.693	0.697	0.700	0.702	0.717
8	0.524	0.620	0.647	0.660	0.667	0.672	0.675	0.678	0.694
9	0.480	0.591	0.621	0.635	0.643	0.649	0.653	0.655	0.674
10	0.438	0.564	0.598	0.613	0.622	0.628	0.632	0.635	0.656
11	0.397	0.539	0.576	0.593	0.603	0.609	0.614	0.617	0.640
12	0.357	0.515	0.556	0.574	0.585	0.592	0.597	0.601	0.625
13	0.318	0.493	0.537	0.557	0.569	0.576	0.581	0.585	0.611
14	0.278	0.472	0.519	0.541	0.553	0.561	0.567	0.571	0.599
15	0.237	0.452	0.503	0.526	0.539	0.548	0.554	0.558	0.587
16	0.196	0.432	0.487	0.512	0.526	0.535	0.541	0.546	0.576
17	0.152	0.414	0.472	0.498	0.513	0.522	0.529	0.534	0.566
18	0.106	0.396	0.458	0.485	0.501	0.511	0.518	0.523	0.557
19	0.056	0.378	0.444	0.473	0.489	0.500	0.507	0.513	0.548
20	0.011	0.361	0.431	0.461	0.479	0.489	0.497	0.503	0.539

When $p > 1$, ratio is always less than 1, and can be *much* less than 1. So you should never prefer simultaneous intervals for the μ_i based on T^2 to those based on Bonferronized t.

Simultaneous confidence intervals for contrasts

When the components of $\mathbf{x}$ constitute *repeated measures* on a single observational unit at different times and/or under different conditions, you are usually most interested in *comparisons* or *contrasts* among the μ_i, that is, in quantities of the form $\mathbf{a}'\boldsymbol{\mu}$, where $\sum_i a_i = \mathbf{1}'\mathbf{a} = 0$. It can be shown that the family of confidence limits of the form

$$\mathbf{a}'\boldsymbol{\mu} = \mathbf{a}'\overline{\mathbf{x}} \pm c^* \times \hat{SE}[\mathbf{a}'\overline{\mathbf{x}}],$$

where

$$c^* \equiv \{(p'f_e/(f_e-p'+1))F_{p',fe-p'+1}(\alpha)\}^{1/2},\ p' = p-1$$

constitutes a set of simultaneous $1-\alpha$ confidence intervals for all *contrasts*. These are similar to the preceding, except that $p' = p-1$ appears instead of p because there are only $p-1$ linearly independent contrasts. These limits are slightly shorter than the ellipsoid based limits that are not restricted to contrasts. It still is the case that $\hat{SE}[\mathbf{a}'\overline{\mathbf{x}}] = \{\mathbf{a}'\mathbf{S}\mathbf{a}/n\}^{1/2}$.

In practice, your primary interest is often in relative simple contrasts of the form $\mu_i - \mu_j$, for $i \neq j$, with little or no interest in other more general contrasts. You can express such a contrast as $\mu_i - \mu_j = \mathbf{a}_{ij}'\boldsymbol{\mu}$, where $\mathbf{a}_{ij}$ is a vector with 1 in row i, -1 in row j, and 0 in all other rows. For such an $\mathbf{a}_{ij}$, $\mathbf{a}_{ij}'\mathbf{S}\mathbf{a}_{ij} = s_{ii} - 2s_{ij} + s_{jj}$, so the estimated variance of $\overline{x}_i - \overline{x}_j$ is $\{(s_{ii} - 2s_{ij} + s_{jj,})/n\}^{1/2} = \{s_d^2/n\}^{1/2}$, where s_d^2 is the sample variance of the differences $d_{ij} = x_i - x_j$.

Therefore, the above simultaneous intervals for all contrast when used for $\mu_i - \mu_j$ become

$$\mu_i - \mu_j = \bar{x}_i - \bar{x}_j \pm c^*\{(s_{ii} - 2s_{ij} + s_{jj})/n\}^{1/2},\ i \neq j,$$

which is a set of simultaneous confidence intervals for the $p(p-1)/2$ distinct pairwise comparisons among variable means.

Each of these is essentially a confidence interval for $\mu_i - \mu_j$ based on paired t, but with confidence level $1 - \alpha^* = P(|t_{f_e}| \le c^*)$.

Alternatively, you can Bonferronize univariate confidence intervals based on a paired Student's t. Since there are $K = p(p-1)/2$ different intervals you must divide α by $p(p-1)/2$ to ensure simultaneous coverage probability of at least $1 - \alpha$. These intervals are of the form

$$\mu_i - \mu_j = \bar{x}_i - \bar{x}_j \pm t^* \times \hat{SE}[\bar{x}_i - \bar{x}_j],\ t^* = t_{f_e}(\alpha'/2),\ i \neq j,\ \alpha' = \alpha/(p(p-1)/2)$$

Which set should you prefer when you are interested *only* in pairwise contrasts of means? The following table contains ratios of the lengths of intervals based on Bonferronized Student's t to the lengths of intervals based on $F_{p,f_e-p+2}(\alpha)^{1/2}$ for α = .05.

$$t_f(.025/(p(p-1)/2))/\{[(f(p-1)/(f-p+2)]F_{p-1,f-p+2}(0.05)\}^{1/2}$$

$p \setminus f$	20	40	60	80	100	120	140	160	∞
2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
3	0.959	0.970	0.973	0.974	0.975	0.975	0.976	0.976	0.978
4	0.902	0.925	0.932	0.935	0.937	0.938	0.939	0.939	0.944
5	0.844	0.882	0.892	0.897	0.900	0.902	0.904	0.905	0.911
6	0.789	0.842	0.856	0.863	0.867	0.870	0.872	0.873	0.882
7	0.736	0.804	0.823	0.832	0.837	0.841	0.843	0.845	0.856
8	0.685	0.770	0.793	0.804	0.810	0.814	0.817	0.819	0.833
9	0.636	0.738	0.765	0.778	0.785	0.790	0.793	0.796	0.812
10	0.587	0.708	0.740	0.754	0.762	0.768	0.772	0.774	0.793
11	0.540	0.680	0.716	0.732	0.741	0.747	0.752	0.755	0.775
12	0.493	0.653	0.693	0.711	0.722	0.728	0.733	0.736	0.759
13	0.446	0.628	0.672	0.692	0.703	0.711	0.716	0.720	0.744
14	0.399	0.603	0.652	0.674	0.686	0.694	0.700	0.704	0.731
15	0.351	0.580	0.633	0.657	0.670	0.679	0.685	0.689	0.718
16	0.301	0.557	0.615	0.640	0.655	0.664	0.670	0.675	0.706
17	0.249	0.535	0.597	0.625	0.640	0.650	0.657	0.662	0.695
18	0.194	0.513	0.581	0.610	0.626	0.637	0.644	0.649	0.684
19	0.136	0.492	0.564	0.596	0.613	0.624	0.632	0.637	0.674
20	0.073	0.472	0.549	0.582	0.600	0.612	0.620	0.626	0.665

Again this shows, that simultaneous confidence intervals derived from Bonferronized t-tests are shorter than simultaneous confidence intervals derived from Hotelling's T^2. So even when testing $p(p-1)/2$ differences simultaneously, you should prefer Bonferronized Student's t to ellipsoidal confidence limits. However, if you want simultaneous limits on a considerably larger number of contrasts, at some point the ellipsoidal limits will be shorter and hence preferable.