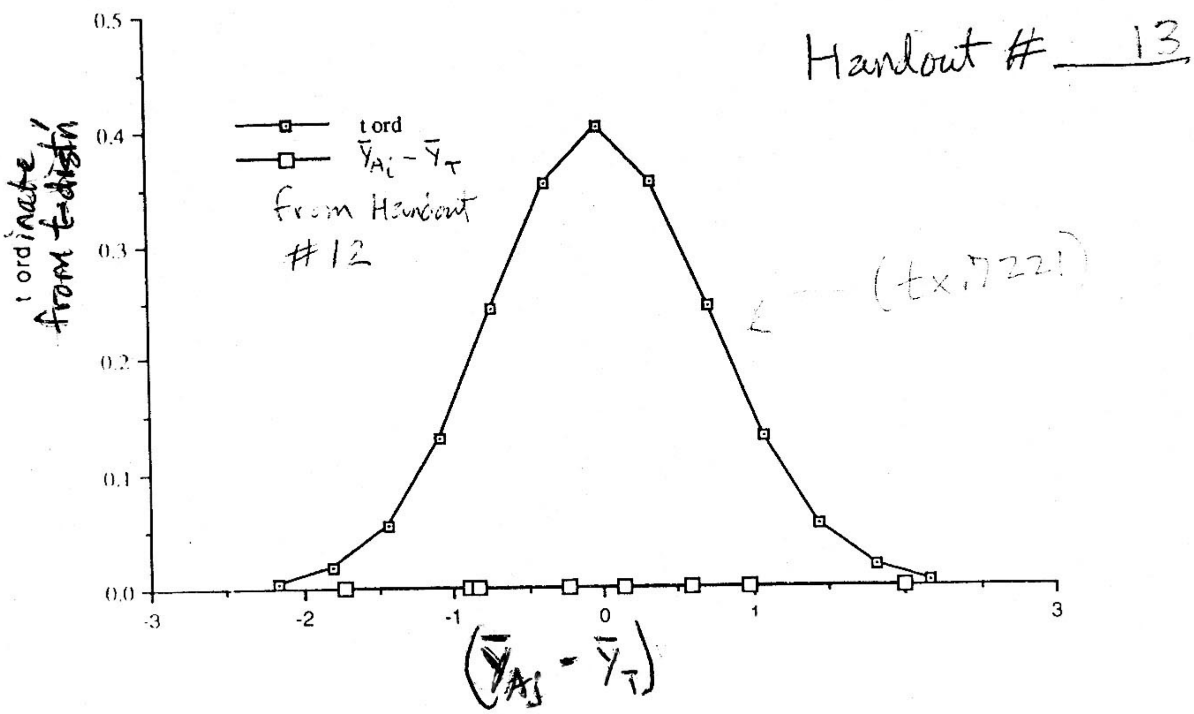
Psych 610 Handout #13
Prof. Moore

Comparison to a reference dist'n as a way of thinking about what differences between pairs of means are big enough to interpret (Based on Box et al., 1978)

The basic idea is the convert a set of means to values on a t-distribution. Then graph them across the x-axis. Then you use the estimated standard error to sketch an approximate t-distribution for your experiment. You compare the distribution of sample means with the 'reference distribution' constructed using the t-distribution and the standard error. The reference distribution is essentially a theoretical sampling distribution, given the value of your standard error. I used the data from Handout #12 to illustrate this. Here's the reference distribution based on the standard error. The boxes on the x-axis are the estimated α_j.



From this graph we see that although the two extreme means do differ significantly by a post-hoc test (as we found in Handout#12), they fall in the tails of the reference distribution, and the other 6 means fall toward the middle of the reference distribution. Hmm. Do we think the significant difference between those two extreme means is worth interpreting, or are they just part of the sampling distribution that we would expect to see if the null hypothesis is true?

Below are two other examples from Box et al.

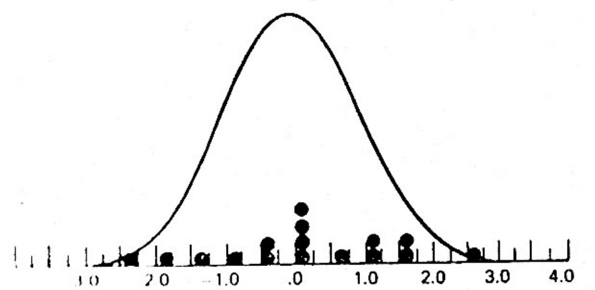
Excerpt from Box et al.: "It is easy to sketch this reference t distribution using Table B2 given at the end of this book. This table shows the ordinates of the t distribution for various values of t and for various numbers of degrees of freedom v. For the present example v = 20 and the scale factor $s/\sqrt{n}$ is 0.966. Entering Table B2 with v = 20, we obtain

value of t	0	0.5	1.0	1.5	2.0	2.5	3.0
t ordinate	0.394	0.346	0.236	0.129	0.058	0.023	0.008
t x 0.966	0	0.48	0.97	1.45	1.93	2.42	2.90

*When s is substituted for σ, the quantities $\sqrt{n}(\overline{y_1} - n_1)/s$ are not independent since they all use the same estimate s. However, this does not seriously invalidate the suggested procedure unless the number of degrees of freedom associated with s is small (say less than 10)."

Comparison to a reference distribution
From: Box, Hunter, & Hunter. *Statistics for experimenters*, pp. 190-193.

Here is one graph from an example in Box et al. that illustrates an 'iffy' significant difference. As in our class example above, two means fall in the tails of the reference distribution, and they differ significantly, but in the larger context of the rest of the means from the experiment we question whether we should interpret that significance:



[illegible] averages for which those in Figure 6.7*a* are the extremes.

FIGURE 6.7. Reference distributions and the problem of selection.

Now the next graph shows an example where a pair of means cannot be conceptualized as having come from a sampling distribution based on the estimated standard error and the assumption that the null hypothesis is true. The reference distribution is shown floating above the means. Notice that if you slide the ref dist'n back and forth on the x-axis you can't get it to capture the means of treatments A and D along with B. A and D must come from a different sampling distribution than B and C.

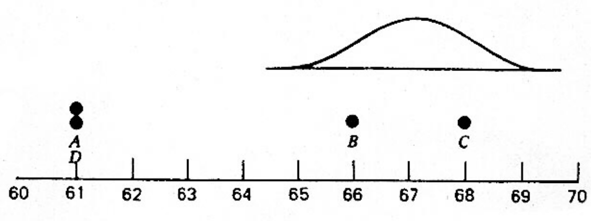


FIGURE 6.6. Sample averages in relation to a reference *t* distribution with scale factor $\sqrt{s_R^2/n} = \sqrt{5.6/6} = 0.97$, blood coagulation example.