Solutions for Assignment #2

(1) Here are the first three sources.

i)	BIPM (int'l standards group, Paris)	[6.671, 6.693]
ii)	Univ. of Colorado	[6.677, 6.696]
iii)	Univ. of Zurich	[6.673, 6.677]

This was intended to be an exercise in using our pooling procedures, generalized to work with three sources. (You can also find this interval by extending the Excel spreadsheet that pools intervals by extending it from two intervals to work with three intervals). Using the simulation method, you'll need regression. To generate the three sources, use the usual procedure, with e.g. the formula

6.682 + .011 ?normal

for the BIPM source; these intervals are formed as estimate ± one SD. The regression of

(BIPM + UC + UZ)/3 on (BIPM-UC) and (UC-UZ)

essentially reproduces the UZ interval giving 6.676 ± .0019. So, there's little to gain by adding the imprecise sources to this very precise interval. It's as though we have combined a very large survey with two much smaller surveys: adding the other two has little effect.

Root Mean Square Error	0.00192
Mean of Response	6.68097
Observations (or Sum Wgts)	250

Parameter Estimates

Term	Estimate	Std Error	t
Intercept	6.6758576	0.000197	

(2) What happens with some correlation present?

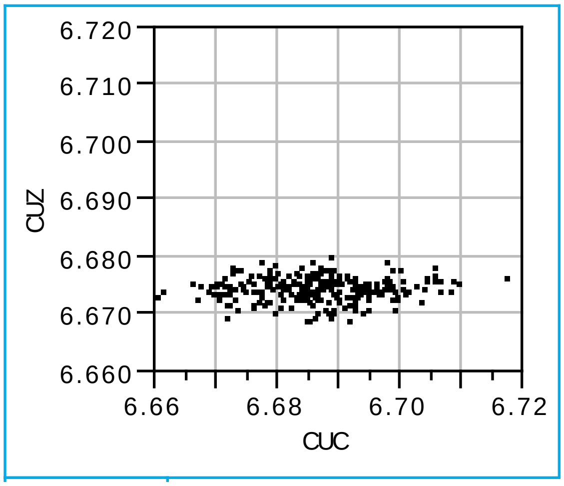
To make this answer simpler, I'll do it in just the case of two sources so that the plots will be easier to appreciate. When you do the pooling, you'll have the surprise that the intervals are wider (as expected for correlation = 0.3 than in the independent case, but narrower when the correlation is 0.8. How can this happen – correlation usually leads to less effective pooling. Well, not always. Comparing the intervals formed from Univ. of Co (CUC) and Univ. of Zurich (CUZ), I got these results

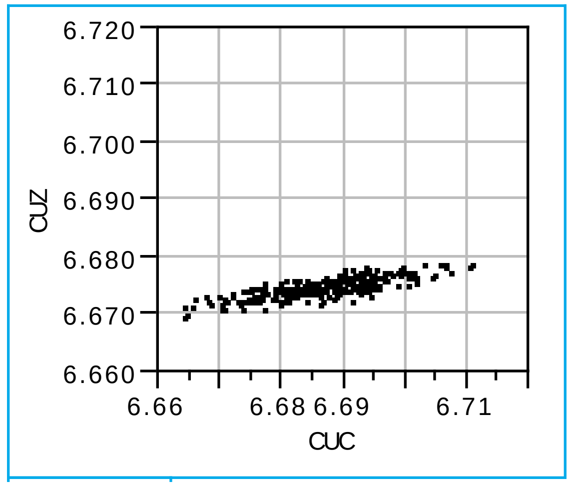
Correlation	Estimate	RMSE	Equal Scales
0	6.6756	.001966	.0066
0.1	6.6755	.0018	
0.3	6.6751	.0021	.0070
0.6	6.6746	.0018	.0080
0.9	6.673	.0013	.0092

The reason for this oddity is the large difference in scale between the two sources. If you expand the scale of CUZ to match that of CUC (make both SDs equal to 0.0095), you get the SE's in the last column shown above; these drop off with the correlation as expected. When the correlation is 0.9, the scale of the pooled information is almost back up to 0.0095, the scales of the input sources.

So, what happens when the scales differ? To see this, consider the following two plots. On the left has the two sources (same scales on the axes) with correlation 0.3, on the right with correlation 0.9. Remember, we are looking at variation in CUC or CUZ associated with the diagonal observations.

If the two sources have the same SD, then higher correlation generally implies data concentrated along the diagonal. However, with such different scales, the data start to concentrate along a line with much less slope. The correlation has led to less variability among the diagonal values, and a tighter interval. The plot also shows why the mean drops as the correlation goes up.

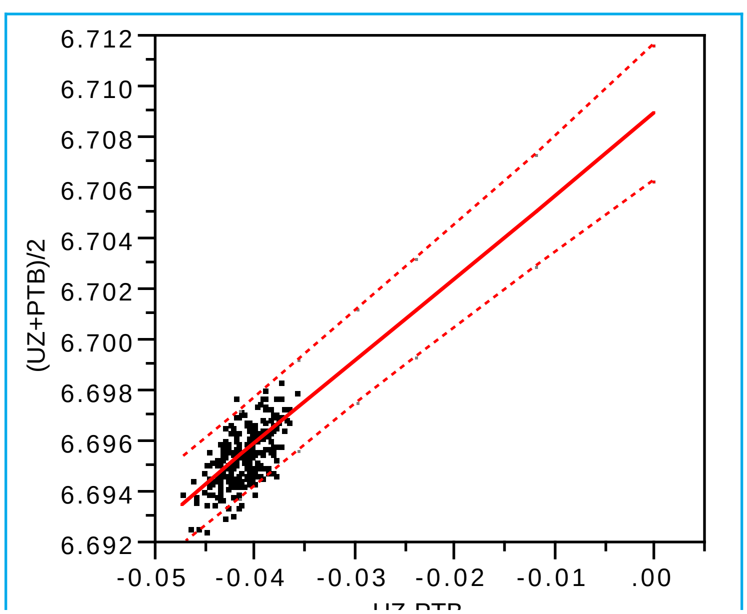




(3) Other groups have been working on the same quest. One group, located at the German standards lab PTB, obtained the interval [6.715, 6.717].

(a) What happens if you pool this interval with the first three listed in "a"?

It is best to do this part by adding the fourth source to the regression model that you used in #1. Alternatively, since that regression gives essentially the third interval, you might as well leave out the first two imprecise sources and work with just these two narrow intervals. With only two intervals, we are back in the case of a simple regression and have an easier time plotting. In the scatterplot of the average on the difference, you can see that the difference is never zero! That means we have to extrapolate very far to get an interval for this problem. To get JMP to show the extrapolation, 'subset' the JMP spreadsheet and pull out these two columns (see the Tables menu). Then add an extra row to the new spreadsheet (subsetting gets rid of the formulas so you can alter data). Add the value zero for the difference in the new row. If you use multiple regression, JMP will compute the prediction value and SE for you, namely 6.7089 ± .00136. The interval did not get shorter because of the extrapolation since the extrapolation is so far from the previous two.



(b) Should this fourth source (from PTB) be pooled with the other three?

No, the fact that the intervals do not overlap implies that the sources estimate different things. If we pool them, we get a longer interval since we have to extrapolate from either source.

(4) The recent interest in G was started by a conflict between two groups, namely PTB with the interval [6.715, 6.717] and a lab (MSL) in New Zealand that offered the interval [6.665, 6.667]. Both intervals are very short, do not overlap, and exclude the standard. Without additional calculation and ignoring the other sources listed above, how accurate would an experiment need to be in order to distinguish which of these was right? That is, if you could "build" a third source, how short an interval would this third source need to produce in order to distinguish these two?

Let's assume that we are going to have to pay someone to run an experiment to resolve the apparent difference. The cost of this experiment is higher the more precise it becomes (as in larger surveys are more accurate and more expensive). The point of the question is that we do not need to require the third experiment be so precise as either of these. Assuming that one of these is right, we roughly need the length of the 95% interval associated with our new procedure to be less than half the distance between the two. That way if our interval is centered on one or the other of these two, it will exclude the other and we'll know which has been verified. You often don't need so much data to pick among alternatives.