CONFIDENCE AND STANDARD ERRORS

The deal is ... it is NOT the standard error that determines the amount of confidence <u>that a specific CI has in including or capturing the parameter</u>. It is the number of error units you add and subtract from the statistic that determines this. To show this, consider another example similar to the book ... where the standard error is about .4. What I did below was to make a similar situation where the standard error was about half of that ... I did this by sampling from a population where the population standard deviation was 1 ... this would make the approximate standard error be: 1 / sqroot 25 = 1/5 = .2. Look at the Minitab simulation.

```
MTB > rand 200 c1-c25;
SUBC> norm 60 1.
MTB > rmean c1-c25, c26
MTB > name c26='Mean'
MTB > dotp c26


                   .        .    .
               : : :. .:..: ::   :
              .: :.:: ::::: :::..:
              :: ::::::::::::::::: :
           :  ::.::::::::::::::::: : . .
.         .:.::::::::::::::::::::::::::: .:
 ---+---------+---------+---------+---------+-Mean
59.50     59.75     60.00     60.25     60.50

MTB > desc c26

                  N     MEAN   MEDIAN   TRMEAN    STDEV   SEMEAN
Mean            200   60.014   60.010   60.013    0.186    0.013

                MIN      MAX       Q1       Q3
Mean         59.402   60.460   59.869   60.156
```

Note that the standard error of the mean in this simulation is about half the previous one ... = .186. I put that in k1.

```
MTB > let k1=.186
MTB > let c28=c26-k1
MTB > let c30=c26+k1
MTB > set c29
DATA> 200(60)
DATA> end
MTB > let c31=(c28 lt c29) and (c30 gt c29)
```

```
MTB > prin c26 c28-c31

 ROW      Mean       C28   C29       C30   C31

   1   60.2483   60.0623    60   60.4343     0
   2   60.1489   59.9629    60   60.3349     1
   3   59.9742   59.7882    60   60.1602     1
   4   59.8794   59.6934    60   60.0654     1
   5   60.2215   60.0355    60   60.4075     0
   6   60.1621   59.9761    60   60.3481     1
   7   60.0967   59.9107    60   60.2827     1
   8   60.1699   59.9839    60   60.3559     1
   9   59.8943   59.7083    60   60.0803     1
  10   59.9286   59.7426    60   60.1146     1
  11   60.1535   59.9675    60   60.3395     1
  12   59.9975   59.8115    60   60.1835     1
  13   59.8863   59.7003    60   60.0723     1
  14   59.9013   59.7153    60   60.0873     1
  15   59.8979   59.7119    60   60.0839     1
  16   59.9346   59.7486    60   60.1206     1
  17   60.1676   59.9816    60   60.3536     1
  18   59.7715   59.5855    60   59.9575     0
  19   59.8352   59.6492    60   60.0212     1

MTB > sum c31
   SUM     =      130.00
MTB > let k2=130/200
MTB > prin k2
K2       0.650000
```

In this simulation, about 65 percent of the CI's contained the parameter of 60. The fact is, it is not the standard error that drives confidence. I cut the standard error more than in half yet the confidence or percentage is still about 68.

The confusion is between width of the CI and amount of confidence. A CI in this case would look like ...

Mean +/- (# err units)(Stan err)

Clearly, the width of the CI will be affected by <u>both</u> the size of the stan err AND the number of units of err you go on either side of the mean. Bigger or smaller widths can be achieved by changing either # of err units OR size of stan err. However, what determines <u>how many intervals will tend to capture the parameter</u> is the number of err units ... NOT the stand err. While smaller stand errors will necessarily make narrower CI's

... that does not mean that narrower CI's yield more confidence!!!!!!!!!!!!!!!!! How close the statistic is to the real parameter is a function to the standard error but ... that is NOT the same question one is addressing when asking what is the liklihood that the CI itself contains the parameter.