STAT 3611 Lab2 Procedure S2015 Maghsoodloo

1. Download data from Canvas and save as Lab2-S015 in your Lab-folder. However, presently please make the assumption that the downloaded data (Exercise 13 on p. 24 of Devore (8e)) represents the ksi during a certain period of all metallic aerospace vehicles produced by the company (i.e., there were only N = 153 vehicles produced). Thus, the 153 ksi's represent a population of size N = 153.

2. On the same Excel WS (= worksheet), column C4-D9, compute the population parameters μ, σ^2, σ, CV_x, and V($\overline{x}$) for a random sample of size n = 10, in the D4-D9 box as shown on the screen. This lab intends to illustrate the concept of random sampling variation in CIs (Confidence Intervals).

3. Then generate 10 random samples each of size 100 (in E12 through N111) from the population in column A3-A155 by first going → Data → Data Analysis → Scroll down to "Sampling" → ok. Select input range the column A3-A155, put in 100 for No. of Samples and as Output Range put in E12-E111. Repeat this 9 more times for columns F through N, obtaining a 100×10 data-matrix. You should now have 100 random samples each of size 10 in E12-N111, i.e., each row in E12-N12 represents a simple random sample of size n =10 from the population of size N = 153. The total possible random samples of size n = 10 from N = 153 is given by $_{153}C_{10}$ = 153!/(143!×10!) = Matlab's nchoosek(153,10) = $1.434461382227160\times10^{15}$. Next compute the 5 sample statistics $\mathbf{\overline{x}_i}$, $\mathbf{CF_i}$, $\mathbf{CSS_i}$, $\mathbf{S_i^2}$ and S_i (i = 1, 2, …, 100) in columns O-S.

4. In columns T through W compute the 95% confidence limits μ_L and μ_U for the case of known σ from $\mathbf{\overline{x}_i \pm z_{0.025}}\,\sigma / \sqrt{\mathbf{n}}$ and then for the case of unknown σ using $\mathbf{\overline{x}_i \pm t_{0.025,9}\, S_i / \sqrt{n}}$, bolding all 100 CIs (Confidence intervals) that do not contain μ. Note that the 2&1/2 percentage point of Gosset's t_9 in Excel is given by $\mathbf{t_{0.025,9}}$ = t.inv(0.975, 9), where CNL stands for Confidence Limit.

5. Compute the approximate E($\mathbf{\overline{X}}$), E($\mathbf{S^2}$) and E(S) in the designated box for the simulated 100 simple random samples each of size 10. Then, re-compute the approximate biases in $\mathbf{\overline{x}}$, $\mathbf{S^2}$ and S. What do you anticipate should happen to the amount of these biases? Estimate the SE($\overline{x}$) from the 100 sample mean values and also from the approximated E(S).