STAT 446 HOMEWORK 3

Leslie Gains-Germain

1. No, I do not think this would generate a SRS. Not all books are the same thickness. With this selection procedure, thicker books would be more likely to be selected than thinner books. Since not all books have equal probability of selection, this is not a simple random sample.

2. (b) The mean number of publications per faculty member is estimated to be $(0*28+1*4+2*3+3*4+4*4+5*2+6*1+7*0+8*2+9*1+10*1)/50 = 1.78$. The standard error of the estimate is calculated considering that we are drawing from a finite population. $SE(\bar{x}) = \sqrt{\frac{N-n}{N}\frac{S^2}{n}} = \sqrt{\frac{807-50}{807}\frac{2.68^2}{50}} = \sqrt{0.135} = 0.367$.

3. I will assume that s^2 is the same for each of these samples. For the first scenario, the standard error of the sample average is $\sqrt{\frac{3600}{1600000}s^2} = \sqrt{0.00225s^2} = 0.05s$. For the second scenario, the standard error of the sample average is $\sqrt{\frac{270}{9000}s^2} = \sqrt{0.03s^2} = 0.173s$. For the third scenario, the standard error is $\sqrt{\frac{299997000}{900000000000}s^2} = \sqrt{0.00033333s^2} = 0.018s$. The third scenario has the smallest standard error, and the smallest estimated variance of the sample average. The SRS of size 3000 from a population size of 300000000 will give the most precision for estimating a population mean.

4. (a) There are 100 quadrats in the figure, and 10 are selected for the sample. The probability of selecting the sample that was observed is $1/\binom{100}{10} = 5.78*10^{-14}$.

   (b) The estimate of the population total is $\hat{t} = 100\bar{y} = 100*9.2 = 920$.

   (c) The estimated variance of $\hat{t}$ is $\hat{V}(\hat{t}) = 100(100-10)15.51/10 = 13959$ and the standard error of the estimated population total is $SE(\hat{t}) = \sqrt{13959} = 118.15$.

   (d) A two sided t-based confidence interval is the estimate plus or minus the standard error times the t-multiplier. The t-multiplier is the 97.5th percentile (for a 95% confidence interval) from a t-distribution with 9 degrees of freedom, which is 2.26. The 95% confidence interval for the true population total is then $920 \pm 2.26*118.15 = (652.98, 1187.02)$.

   (e) We are 95% confident that the true total number of mines in the region is between 652.98 and 1187.02.