

We have learned that in dealing with proportions, we must replace σ_i with $\sqrt{p_i(1-p_i)}$ in all the eqns for stratified sampling. But if we have no idea what the p_i values are, we often set all of them to $\frac{1}{2}$. Why?

First, notice that when all the σ_i 's are the same, then the eqn for the min. n is:

$$n = \frac{1}{\left(\frac{B}{1.96\sigma}\right)^2 + \frac{1}{N}} \quad \text{where } \sigma = \sqrt{p(1-p)}.$$

Now, why would we take $p = \frac{1}{2}$? Remember that the formula for n is the minimum required sample size. So, to be conservative we would want to take a sample that is as large as the eqn allows. That means, as small as the denominator becomes, i.e. as large as $\sigma (= \sqrt{p(1-p)})$ becomes. And we know from Ch 2 already that the max of $\sqrt{p(1-p)}$ occurs at $p = 1/2$.