

Sometimes applying the optimal allocation formula in Equation (3.15) results in one or more of the “optimal” n_h ’s being larger than the population size N_h in those strata. In that case, take a sample size of N_h in those strata, and then apply (3.15) again with the remaining strata.

Neyman allocation is a special case of optimal allocation, used when the costs in the strata (but not the variances) are approximately equal. Under Neyman allocation, n_h is proportional to $N_h S_h$. If the variances S_h^2 are specified correctly, Neyman allocation will give an estimator with smaller variance than proportional allocation (see Exercise 36).

Example 3.10. The caribou survey in Example 3.3 used Neyman allocation to determine the n_h . Before taking the survey, the investigators obtained approximations of the caribou densities and distribution, and constructed strata to be relatively homogeneous in terms of population density. They set the total sample size as $n = 225$. They then used the estimated count in each stratum as a rough estimate of the standard deviation, with the result shown in Table 3.7. The investigators wanted the sampling fraction to be at least $1/3$ in smaller strata, so they used the Neyman allocations in the n_h column as a guideline for determining the final sample sizes in the last column. ■

TABLE 3.7

Quantities used for designing the caribou survey in Example 3.10.

Stratum	N_h	s_h	$N_h s_h$	$n_h = 225 \frac{N_h s_h}{\sum_l N_l s_l}$	Sample size
A	400	3,000	1,200,000	96.26	98
B	30	2,000	60,000	4.81	10
C	61	9,000	549,000	44.04	37
D	18	2,000	36,000	2.89	6
E	70	12,000	840,000	67.38	39
F	120	1,000	120,000	9.63	21
Total	699		2,805,000	225.00	211

3.4.3 Allocation for Specified Precision within Strata

Sometimes you are less interested in the precision of the estimate of the population total or mean for the whole population than in comparing means or totals among different strata. In that case, you would determine the sample size needed for the individual strata using the guidelines in Section 2.7.

Example 3.11. The U.S. Postal Service has conducted surveys asking postal customers about their perceptions of the quality of mail service. The population of residential postal service customers is stratified by geographic area, and it is desired that the precision be ± 3 percentage points, at a 95% confidence level, within each area. If there were no nonresponse, such a requirement would lead to sampling at least 1067 households in each stratum, as calculated in Example 2.12. Such an allocation is neither proportional, as the number of residential households in the population varies a great deal from stratum to stratum, nor optimal in the sense of providing the greatest efficiency for estimating percentages for the whole population. It does, however, provide the desired precision within each stratum. ■