

**SURVMETH / SURV 625**

Applied Sampling

Winter / Spring 2025

**Homework 2**

1. A stratified random sample of graduates from the academic departments at a small university is to be selected to estimate the mean starting salary of graduates from that university. The university has five academic departments, and preliminary data are obtained from an attempted census in a recent year:

Stratum	Department	$N_h$	$\bar{Y}_h$	$S_h$
1	Humanities	20	46,800	81,200
2	Social Sciences	90	61,500	101,700
3	Natural Sciences	120	76,100	130,900
4	Engineering	140	89,500	170,100
5	Business	200	95,500	216,400

- What is the mean starting salary  $\bar{Y}$  per graduate across all departments in the population, based on the data collected from this previous attempted census?
- What is the average within-stratum element variance  $S_w^2 = \sum_{h=1}^5 W_h S_h^2$ ?
- For a sample of  $n = 100$ , what is the proportionate allocation?
- For a sample of  $n = 100$ , what is the Neyman allocation?
- Estimate the sampling variance of the mean for the proportionate c) and Neyman d) allocations.
- Estimate the total element variance  $S^2$ . (*Hint*: use the expression (2) on slide 30 of the Lecture 3 notes.)
- What are the design effects of the proportionate and Neyman allocations?
- Suppose that the cost-per-element was not the same in each stratum:  $C_1 = C_2 = C_3 = \$30$ , and  $C_4 = C_5 = \$40$ .
  - The client requesting a stratified sample design has indicated that the total available data collection budget is  $C = \$5,000$ , with the stratum-specific costs per element listed above. What allocation will minimize the sampling variance of the mean under these cost constraints?
  - Estimate the expected sampling variance and design effect of the mean starting salary under this allocation. (*Hint*: Think carefully about the definition of a design effect)