

Final Quiz

Please try to answer all of the following questions.

Sampling Design

1. What is coverage? When does a sampling frame overcover or undercover a population? What implications does this have for survey design and analysis?
2. Generally speaking, if we use stratified sampling, will we need to recruit a sample that is larger or smaller than a simple random sample in order to obtain an comparable sampling variance (and thus comparable standard errors)?
3. What is a design effect?
4. What is the difference between sampling variance and item (or element) variance?

Design Weights

5. What kinds of sampling designs produce self-weighted samples?
6. What kinds of sampling designs do not produce self-weighted samples?
7. Why would we choose a sampling design that is not self-weighting? What trade-offs are involved in that decision?
8. What determines the probability that a unit is selected in *stratified* sample?
9. What determines the probability that a unit is selected in *cluster* sample?
10. Consider a survey where women are intentionally oversampled in the survey design. If we simply analyze the unweighted data, treating all respondents equally, will the results overrepresent (i.e., weight more heavily) or underrepresent (i.e., weight less heavily) responses from women, or does the design not matter for calculating sample estimates?

Nonresponse Weights

11. What is the purpose of nonresponse weights? How do they differ from design weights?
12. Do nonresponse weights solve the problem of nonresponse bias?
13. Consider a survey where women and men are sampled equally (i.e., all men and women in the sampling frame each have an equal probability of selection into the sample) and assume for purposes of this exercise, that the population as a whole is 50% women and 50% men. If our final sample after data collection includes 60% women and 40% men, will design weights correct for the overrepresentation of women in the sample?

14. To calculate nonresponse weights, we need data about respondents and nonrespondents. Why do we need information about nonrespondents? What information do we need?
15. What are some situations in which we might have the information necessary to calculate nonresponse weights?

Post-stratification

16. How do post-stratification weights differ from nonresponse weights?
17. In the table below, who is overrepresented and who is underrepresented in the sample (relative to the population)? What are the appropriate post-stratification weights for men and women?

	Men	Women
Population	50%	50%
Sample	35%	65%
Weight		

18. Imagine a survey that is designed using stratified sampling (with disproportionate allocation to strata) and assume that there are uniform response rates across strata. What kind of weights do need to calculate for this sample? Do we need to post-stratify?
19. Why should we always need to post-stratify a non-probability sample in order to estimate population quantities of interest?

Questionnaire Design

20. How do we determine the best way to measure a construct using a survey instrument?
21. What can pretesting tell us about a survey?
22. What effect does mode have on survey participation and response behavior?
23. What advantages and disadvantages do interviewers bring to a survey interview?

Budgeting

24. How do we estimate the cost of implementing a survey?
25. Compared to a simple random sample of the same population, is a stratified sample likely to cost more or less?
26. Compared to a simple random sample of the same population, is a cluster sample likely to cost more or less?

Nonresponse

27. Does a low response rate necessarily mean a survey sample is biased?
28. How do we calculate a response rate for a probability-based sample survey?
29. How do we calculate a response rate for a non-probability survey?
30. How do we calculate a response rate for a sample recruited from an online panel constructed from probability sampling?
31. What can a survey researcher (or survey interviewer) do to when a respondent refuses to participate in a survey or breaks off from participating in a survey?

Coding and Editing

32. If we have removed names, birthdates, addresses, and identification numbers from a dataset, have we guaranteed the anonymity of respondents?
33. What are some situations where it would be appropriate to edit or change the response supplied by a respondent?

Missing Data

34. What is the difference between unit nonresponse and item nonresponse?
35. Which is worse: unit nonresponse or item nonresponse?
36. What is complete case analysis?
37. What assumptions do we need to make in order to justify the imputation of missing data? How do we know if a particular missing data assumption is satisfied?
38. How does single imputation differ from multiple imputation? What are the advantages and disadvantages of each?

Codebooks

39. What information needs to be included in a codebook for each question asked on the survey?
40. What kinds of supplemental data (i.e., data not supplied by the respondent) might need to be documented in a codebook?
41. What kind of study-level metadata should be recorded?
42. Should that data be included in a codebook file, or separately?