United States Government Accountability Office

Human Capital Office Learning Center

Participant Manual

July 2013

Probability Sampling

GAO Accountability + Integrity + Reliability

GAO Highlights

Highlights of COAN904, an elective course for all band levels.

June 2013

THINKING CRITICALLY

Probability Sampling

Why This Course is Important

Sampling can be a powerful tool for estimating the characteristics of a population. Well designed probability samples give analysts the ability to gather information from relatively few members of a population and accurately generalize the results to the entire population.

Questions to Ask the Learner after Class

- How does probability sampling fit into GAO's work and when should we consider conducting it?
- Why might it make sense to use probability sampling? What would be the population? What attributes of a population would be selected?
- What limitations might we face if we conducted such analyses?

For more information, contact Training at training@gao.gov.

The Knowledge and Skills This Course Covers

After completing this course, you will:

- Be more familiar with sampling terminology.
- Be able to understand sources of error associated with sampling.
- Know how to specify survey objectives to select a sample design.
- Understand the use of weights.
- Be able to identify sources of error associated with surveys.
- Know how to convey the results of probability sampling.

Participants in this course learn the differences between probability and nonprobability samples and why you would choose a probability sample to collect and analyze engagement data. Discussion on the specific types of probability samples, including simple random sampling, stratified random sampling, and cluster sampling, provide the participant with a variety of examples from GAO products.

Course participants will learn how sample size, precision, and confidence interact in probability sampling. For example, they will learn the factors that affect the size of the sample and how higher levels of precision require larger sample sizes. These concepts are discussed using both performance and financial audit examples.

Participants will also learn how sample weights are used to take into account different chances of selection and to make sure that, when aggregating the sample data together, they properly reflect the population from which they were sampled. Finally, sampling and nonsampling errors are discussed.

Registrar Information

Probability Sampling (COAN904)	This course is designed to provide an overview of probability sampling, a commonly used approach that allows us to generalize from a sample to a large population. The course addresses the GAO competency of critical thinking. The course can be delivered in a team-dedicated format or open enrollment.
Who Should Enroll	Open to all who wish to learn about the principles and practices of probability sampling, and to understand when it is appropriate.
Prerequisite	None
Advanced Preparation	None
Recommended Related Courses	Non-Probability Sampling, Commonly Used Statistical Tests
Course Objectives	After completing this course, you will: • Be more familiar with sampling terminology. • Be able to discuss pros and cons of sample designs. • Know how to specify survey objectives to select a sample design. • Understand the use of weights. • Be able to identify sources of error associated with surveys. • Know how to convey the results of probability sampling.
Competencies	Thinking Critically
Instructional Method	Classroom
Length	2 hours
CPE Credits	2 (1 government-related)

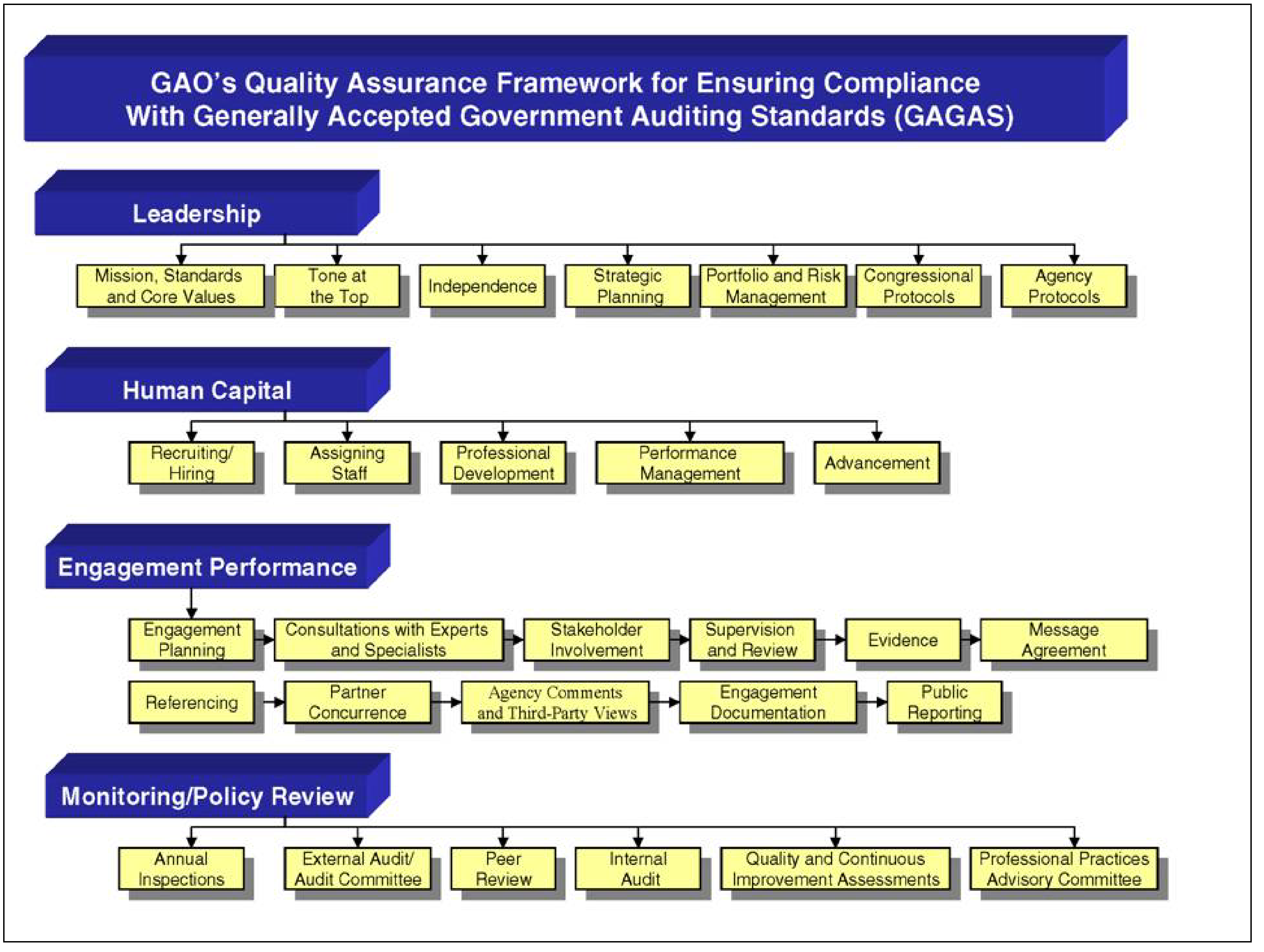
Course Manager Linda Hawkins, (202) 512-3094

Course Evaluation The first business day after the class ends, participants will receive an electronic evaluation. The first question on the evaluation asks whether or not the participant attended and completed the entire course. Marking “yes” and going on to complete the course evaluation will automatically update training records to reflect completion data and CPE credit. **Note:** Participants must attend and participate in the **entire** class to be eligible for CPE credit.

Contents

Contents

Quality Assurance Framework

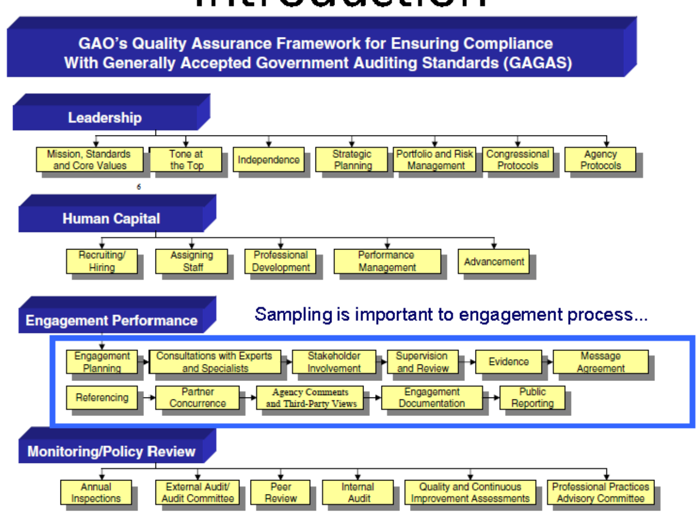


Quality Assurance Framework

Module 1 PowerPoint Slides

Probability Sampling

Introduction



Introduction (cont.)

- Sampling can be a powerful tool for estimating the characteristics of a population
- Well designed probability samples give analysts the ability to gather information from relatively few members of a population and accurately generalize the results to the entire population

Introduction (cont.)

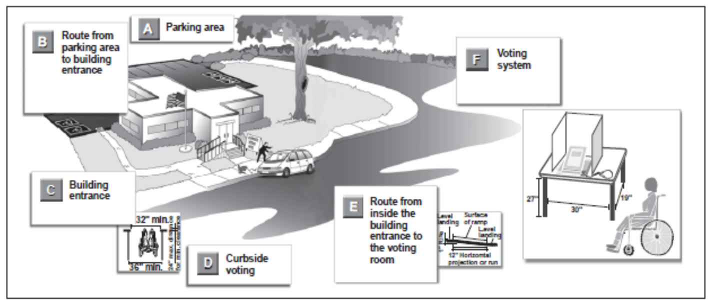
After completing this course, you will

- Be more familiar with sampling terminology.
- Be able to discuss pros and cons of sample designs.
- Know how to specify survey objectives to select a sample design.
- Understand the use of weights.
- Be able to identify sources of error associated with sampling.
- Know how to convey the results of probability sampling.

A Real Example

- Suppose the Congress wants to know the percentage of polling places which have features that impede the access to voting for people with disabilities.
  - Voting is fundamental to our democratic system, and federal law generally requires federal election polling places to be accessible to all eligible voters, including the elderly and voters with disabilities.

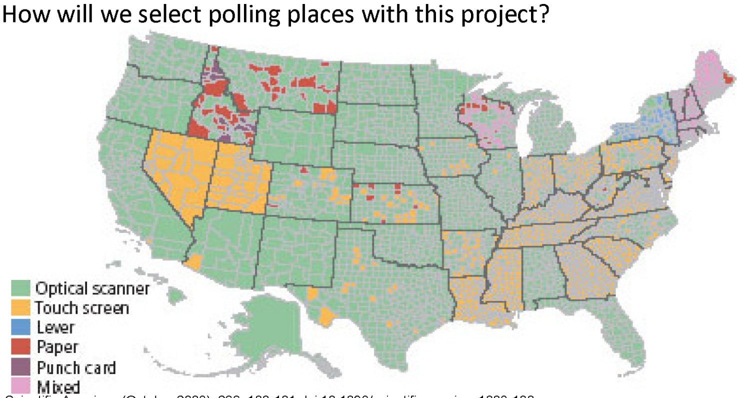
Figure 5: Key Polling Place Features That We Examined



- We have to figure out how GAO should answer this kind of question.
  - Do we visit every polling place in the United States?

A Real Example (cont.)

- How will we select polling places with this project?



Scientific American (October 2008), 299, 100-101 doi:10.1038/scientificamerican1008-100

- How might voter accessibility at polling places differ across the country? (Who oversees or administers the elections in your precinct?)

- What will be the reporting limitations, if any, if we only select polling places located in the state of Montana?

A Real Example (cont.)

Answering this audit question for the Congress will require GAO to send teams of analysts out into the field on Election Day to measure voter accessibility at polling places around the country.

Questions we might think about are

- How do we structure the design our sample so that our teams can finish the accessibility testing field work on Election Day?
- How do we create a national list of polling places that is complete, timely, and accurate?
- What are the reporting limitations if the list of polling places used for our sample is not the same as the population we want to sample from?

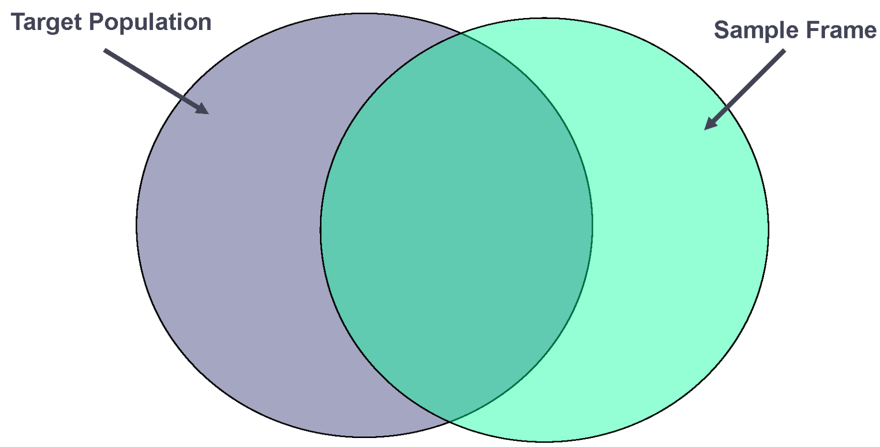
A Real Example (cont.)

The Proportion of Polling Places Without Potential Impediments Increased Since 2000

Compared to 2000, the proportion of polling places without potential impediments increased and almost all polling places had an accessible voting system. In 2008, based upon our survey of polling places, we estimate that 27.3 percent of polling places had no potential impediments in the path from the parking area to the voting area—up from 16 percent in 2000; 45.3 percent had potential impediments but offered curbside voting; and the remaining 27.4 percent had potential impediments and did not offer curbside voting. All but one polling place we visited had an accessible voting system to facilitate private and independent voting for people with disabilities. However, 46 percent of polling places had an accessible voting system that could pose a challenge to certain voters with disabilities, such as voting stations that were not arranged to accommodate voters using wheelchairs.

GAO Report : Voters with Disabilities: Additional Monitoring of Polling Places Could Further Improve Accessibility, GAO-09-941 (Washington D.C., Sept. 2009), page 12.

Some Common Terms



Some Common Terms Defined

Target Population

The population we would like to study. In the voters with disabilities example, we defined the target population by answering the question "What exactly do we mean by population of polling places?"

Sample Frame

The population we can actually select a sample from. The sample frame may or may not exactly match the target population.

Sample

A subset of cases selected from a larger population.

Some Common Terms Defined

Parameter

A parameter is a value, usually unknown, used to represent a certain characteristics of a sample frame. For example, the mean of all units in a sample frame is a parameter that is often used to indicate the average value of a quantity. Within a sample frame, a parameter is a fixed value which does not vary.

Estimate

A numerical quantity calculated from sample data and intended to provide information about an unknown parameter.

Lesson 1: Probability & Nonprobability Sampling

- Sampling in GAO Engagements
- Examples
- Probability & Nonprobability Sampling
- Sample Selection
- Summary

Sampling in GAO Engagements

Analysts conducting financial and performance audits often use sampling to meet their objectives:

- A financial audit provides evidence that a financial statement conforms with accepted accounting practices.
- A performance audit measures a program's performance and management against objective criteria. Performance audits entail various levels of analysis, research, or evaluation.

Example: Financial Audit

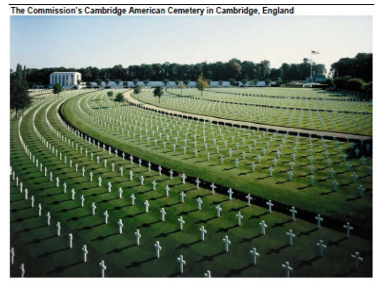
Objective:

Audits were done to determine whether

(1) the American Battle Monuments Commission's financial statements were presented fairly, and

(2) Commission management maintained effective internal control over financial reporting.

The Commission's Cambridge American Cemetery in Cambridge, England



Source: American Battle Monuments Commission.

Method:

Statistical samples of Commission payroll and nonpayroll expenditures were selected to

(1) Determine the validity of activities reported in the Commission's financial statements. The team projected errors in dollar amount to the population of transactions from which they were selected.

(2) Test effectiveness of internal controls. Attributes that indicated deficiencies were statistically projected to the appropriate populations.

Report: *Financial Audit: American Battle Monuments Commission's Financial Statements for Fiscal Years 2011 and 2010*, GAO-12-404

Example: Performance Audit

Objective:

To study the prevalence of workplace safety incentive programs as well as other workplace safety policies that may affect injury and illness reporting.

Method:

We selected a probability sample of manufacturing worksites using a systematic random sample of 1,000 manufacturers from a total of 26,552 in our sample frame of data. Our sample frame consisted of the set of manufacturers with 11 or more employees contained in a probability sample BLS establishment survey fielded in 2010

Rate-based programs
Reward workers who had few or no reported injuries or illnesses during a set time period

Behavior-based programs
Reward workers for behaviors such as reporting near-miss incidents or recommending safety improvements

Source: GAO analysis of workplace safety literature.

Report: Workplace Safety and Health: Better OSHA Guidance Needed on Safety Incentive Programs, GAO-12-329 April 2012.

Probability versus Nonprobability Sampling

- Well-designed probability samples allow analysts to make statements about an entire population, and they allow analysts to measure the accuracy of their estimates about that population.
- Nonprobability samples are more restrictive in what they allow analysts to say.

The table displays some of the differences and nonprobability samples:

Characteristic	Probability	Nonprobability
All sample frame elements have known (nonzero) probabilities of being selected	Yes	No
Estimates of sampling error can be made	Yes	No
Results can be generalized	Yes	No

Characteristics of Probability Sampling

The characteristics of a probability sample are:

- The sample is one of a number of potential samples.
- Each element in the population must belong to at least one potential sample.
- Each potential sample has a chance of being selected.
- The chosen sample is selected by a random process in accordance with the chances of selection.

Characteristics of Nonprobability Sampling

In a non-probability sample, some elements in the population have no chance, or an unknown chance, of being selected.

For example, select

- 6 employees who are "typical" Team X employees, or
- The 6 employees who have the longest lengths of service, or
- The team's 6 tallest employees.

Using Nonprobability Sampling

- Nonprobability sampling methods can be useful when:
  - ➢ Descriptive comments about the sample are needed, or
  - ➢ The analyst is trying to establish the existence of an attitude or error rather than prevalence.
- Results from a nonprobability sample can be used as a key component of a report.
- Nonprobability sampling methods may not be as useful as the sole support for findings or conclusions. This is because they give information about only the few units or cases in the sample.

Steps in Selecting Probability Samples

These steps are a reasonable approach to selecting a sample:

One →	Two →	Three →	Four →	Five →	Six
Define Target Population	Construct Sampling Frame	Determine Probability Sampling method	Determine Sample Size	Plan Sample Selection Procedure	Select Sample Units

Lesson 1: Summary

- ✓ When you plan to collect data and make estimates, your first questions are:
  - – What am I making estimates for?
    Or, what is the population?
  - – Can I get a list of the population?
    Or, what is my sampling frame?

Lesson 1: Summary (cont.)

- ✓ Your next question is about method. In GAO, to collect data for an engagement, your team may use both probability and nonprobability sampling methods.

  You have a census if you collect data on all the items you’re interested in—taxpayers, patient records, contract files.

  You’re using a sample if you collect data for only a portion of the population you're interested in.

Lesson 1: Summary (cont.)

- ✓ You would use a probability sample to make an estimate for a population. For example,

> In 2012, from its survey, GAO estimated that 25 percent of U.S. manufacturers had safety incentive programs.

- ✓ Nonprobability samples provide results for the sample only. For example,

> Of the 30 files we reviewed, 12 contained errors

If some units in the sample frame have no chance of being selected for the sample, you have a nonprobability sample, and you will be able to discuss only the sample.

Lesson 2: Sample Size, Precision, & Confidence Levels

- Sample Size & Precision
- Confidence Intervals & Confidence Levels
- Effect of Precision & Confidence Levels
- Audit Objectives & Sample Size
- Confidence Interval Statements
- Discussion Examples
- Summary

Sample Size and Precision

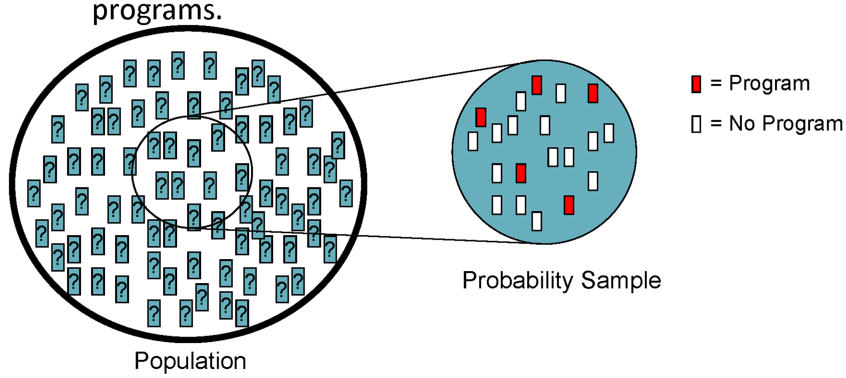
The sample size you need depends on, among other factors, the:

- Precision required,
- Confidence level required,
- Audit's objectives, and
- Resource constraints.

Precision

Precision is the extent to which estimates from different samples tend to vary.

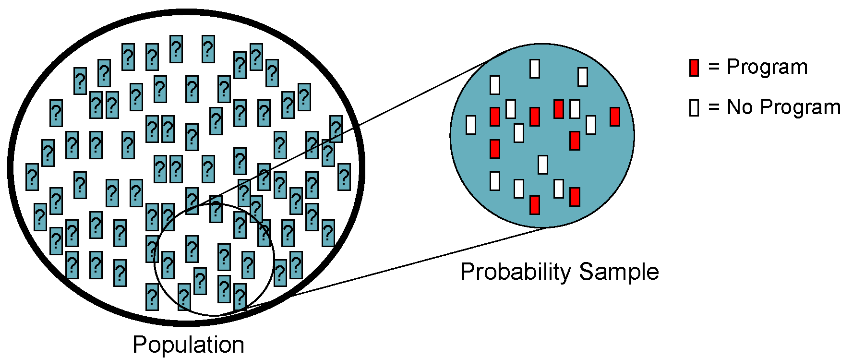
- For example, if a probability sample of 20 U.S. manufacturers found that 5 had safety incentive programs, we could estimate that 25% of *all* manufacturers had safety incentive programs.



26

Precision

– Suppose we selected a different probability sample of 20 manufacturers and found 8 that had safety incentive programs. The estimate from this sample would be 40%.



27

Precision

The variation between these 2 estimates is solely due to that fact that the estimates were based on two different samples.

Precision is the extent to which estimates from different samples tend to vary.

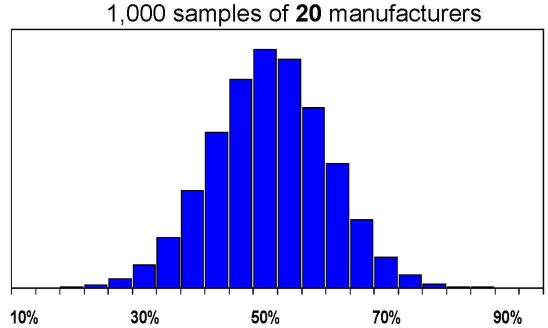
– High precision means estimates from different samples tend to be similar.
– Low precision means estimates from different samples tend not to be similar.

28

Precision

If we repeated this sample of 20 manufacturers 1,000 times and plotted the results, we would have a good idea about the extent to which estimates from different samples tend to vary (precision).

1,000 samples of **20** manufacturers

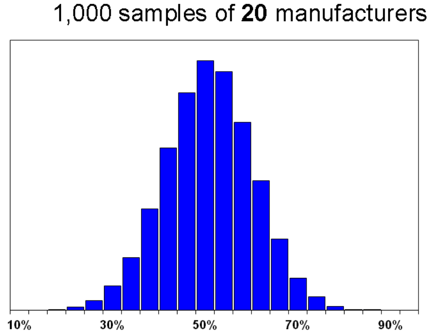


29

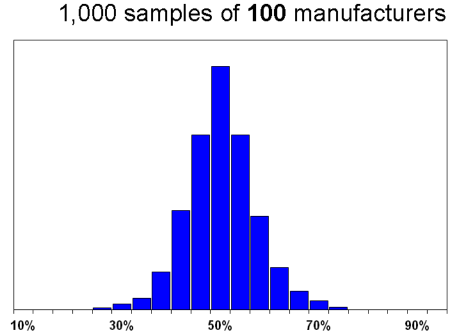
Sample Size and Precision

All else being equal, larger sample sizes results in greater levels of precision

1,000 samples of **20** manufacturers



1,000 samples of **100** manufacturers



Precision

- For probability sampling, precision increases as the sample size increases. A given estimate varies less from sample to sample as the sample size increases.
- Precision can be specified in:
  - ➢ Absolute amounts – e.g. 10 percentage points
  - ➢ Relative amounts – e.g. within 5% of the true value

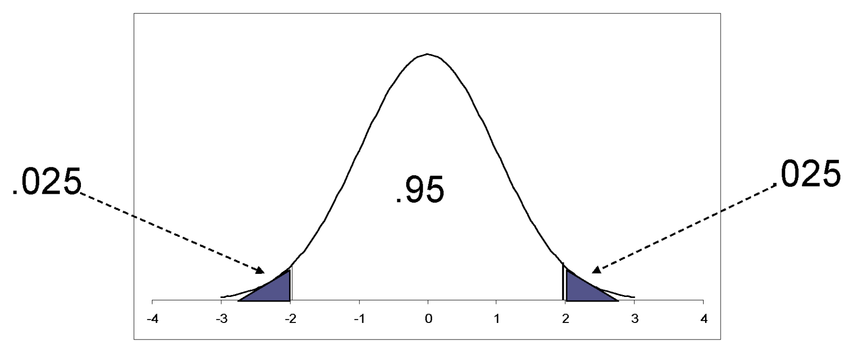
For example:

> We estimate that a total will be about 50,000.
>
> We want relative precision to be within 5% of this value, or ± 2,500.

Is it practical to select 1,000 samples?

Precision and Confidence Intervals

For large samples, the distribution of all possible means follows the “normal” distribution.



- We construct confidence intervals based on the variation estimated from the normal distribution.
- The width of the confidence interval is a measure of precision.

32

Confidence Intervals & Confidence Levels

- A confidence interval is a range of values around an estimate. It is designed to tell us where the true, unknown population parameter lies.
- The size of a confidence interval is directly related to the assurance we need that the true population value lies within the interval.
- Confidence level is that assurance. It is the probability that an inference about an estimate says what it claims to say. Higher confidence levels give greater assurance than lower ones do that they say what they claim to say.

Effect of Precision & Confidence Levels

- GAO generally uses 95% confidence levels for performance audits. Confidence levels for financial audits vary (as detailed in GAO's Financial Audit Manual, vol. 1)[1]
- You will need a larger sample size to be 95% confident that an interval of 5 percentage points contains the true value than to be 90% confident. All else being equal, higher confidence levels require larger sample sizes, as the cells in the table show.

Confidence level \ Precision	Low	High
Low	Smallest	Higher
High	Higher	Highest

[1]See the Financial Audit Manual, Vol 1, sec. 450, GAO-08-588G

Effect of Precision & Confidence Levels (cont.)

Precision and confidence levels raise or lower sample size. In simple random sampling for an attribute that occurs about 50% of the time, the required sample size would look something like this.

Precision / Confidence level	±10 percentage points	±5 percentage points
90%	68	271
95%	97	385

Effect of Precision & Confidence Levels (cont.)

And they affect sample size for an attribute that occurs about 25% of the time, such as in the chart below.

Precision / Confidence level	±10 percentage points	±5 percentage points
90%	51	203
95%	73	289

Objectives & Sample Size

The type of objective also affects the sample size.

For example:

- How precise does the information have to be?
- Does what we're estimating exceed a threshold or a materiality?
- What is our best guess of the value of the item we're estimating?
- Do we need to compare subgroups?

Confidence Interval Statements

GAO usually provides confidence intervals with estimates in its reports, but sometimes it provides only the confidence intervals.

Some confidence interval statements are:

- We are 95% confident that the proportion of participants actually meeting the guidelines was between 13% and 31%.
- We estimate that the total revenue was $43.4 million, $6.5 million.
- We are 95% confident that at least 10% of the vouchers did not have proper supervisory approval.

Discussion Examples

Example 1

Suppose we're measuring whether voters intend to vote for Jones.

- Suppose about 50% of voters are expected to vote for Jones.
- A simple random sample of 97 will provide ± 10 percentage point precision with 95% confidence.
- 49 persons in our sample, or 51%, intend to vote for Jones.

Question:

Is this sample size and this precision sufficient to assess, with 95% confidence, whether Jones will receive more than half the votes and win the election?

Discussion Examples (cont.)

Example 2:

Suppose we're measuring control failures, where the maximum acceptable failure rate is 5%.

- Suppose we think failures occur about 50% of the time.
- A simple random sample of 97 will provide ± 10 percentage point precision with 95% confidence.
- 49 transactions in our sample of 97, or 51%, show control failures.

Question:

Is this sample size and this precision sufficient to assess, with 95% confidence, whether failure rates are acceptable?

Lesson 2: Summary

- ✓ To determine the sample size that will meet an engagement's objectives, decide on the analysis you plan to do or the estimates you will use in the report.
- ✓ Statisticians use confidence intervals and confidence levels to describe precision. Larger samples achieve better precision.
- ✓ The objective can determine the precision that is needed.
- ✓ Resource considerations can drive the precision that is needed.

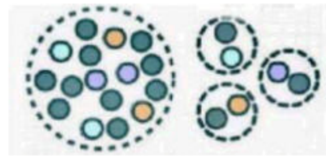
Lesson 3: Sample Design

- Simple Random Sampling
- Systematic Sampling
- Stratified Random Sampling
- Cluster Sampling
- Unequal Probability Sampling
- Summary

Simple Random Sampling

In simple random sampling, every subset of a given size has an equal chance of being selected.

Population



Random samples

Suppose that from a population consisting of Brown, Jones, and Smith, we wanted to select a simple random sample of size 2. The possibilities would be:

(Brown, Jones)
(Brown, Smith)
(Jones, Smith)

Simple Random Sampling (cont.)

- Now suppose the population is made up of 1,000 and we want a simple random sample of only 50. In the first example, we were able to list all possible samples; here, the population is too large to make a list practical.
- So how do we select a sample from a large population? We use software to assign random numbers to elements in the population and then select the sample from those elements.
- Consulting a statistician is probably best at this point.

Simple Random Sampling (cont.)

Pros

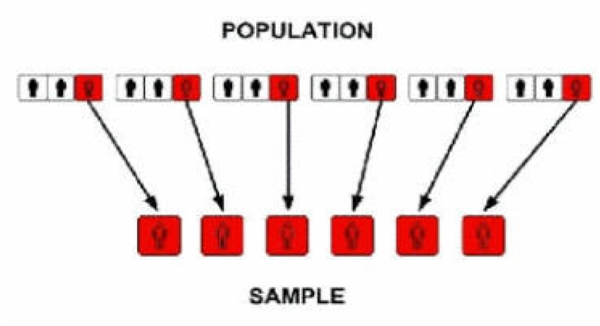
- Simple random sampling requires only a list of elements.
- It is easy to explain and implement.
- It is computationally the least demanding.

Cons

- If you have to make estimates for different subgroups, a simple random sample may not be efficient.

Systematic Sampling

Select sample elements at equal intervals throughout the population.



Systematic Sampling (cont.)

Suppose we need to select two names from the list:

Brown, Jackson, Johnson, Jones, Smith, and Williams

We randomly pick a number from 1 to 3 – for example, 2 – and select Jackson. We also choose the name that is 3 places farther down the list – Smith.

Brown, Jackson, Johnson, Jones, Smith, and Williams

Systematic Sampling (cont.)

Pros

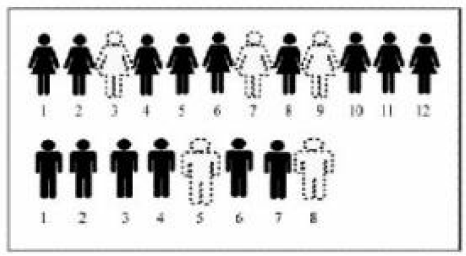
- Systematic sampling is easy to use and is easily taught.
- It guarantees a good spread throughout the list.
- It does not require an explicit sampling frame.
- Records do not have to be electronic.

Cons

- The organization of the list in systematic random sampling is critical, unlike the list in simple random sampling.

Stratified Random Sampling

Each element in the population is uniquely assigned one group. Statisticians call each group a stratum. A simple random sample is then selected from each stratum.



…Stratum 1

…Stratum 2

Estimates from a stratified random sample can be more precise than those from a simple random sample of the same size, if the characteristics being estimated are very different from one stratum to another.

Stratified Random Sampling (cont.)

Pros

- If subpopulations of interest are defined by the strata, stratified examples allow for estimates to be made for all the sub-populations.
- Stratified sampling may be convenient if the population is divided administratively into subgroups.

Cons

- Not all populations are easily classified into groups, or strata.
- Sample size may increase when using stratified sampling for sub-population estimates.

Cluster Sampling

In cluster sampling, large sample units are selected first and contain smaller sample units. The larger sampling units are the clusters. Every population unit is in one and only one cluster.

- The large sample units may be patients, grants, counties, cases and the corresponding smaller sampling units could be treatments, payments, local governments, and actions.

Suppose we need to sample temporary disability payments.

- We might use a random sample of 500 disability payments to look at under and overpayments. But, it may be difficult to sample payments individually. And, the payments for a person have a relationship over time.
- A cluster sample made up of, say, 200 people receiving disability payments might give us the same precision, while saving analyst's time. The analyst may need to review a series of payments to determine whether there is an over or underpayment.
- With cluster sampling, information is only needed for selected clusters. If the cluster is grants, lists of payments are only needed for selected grants.

Cluster Sampling (cont.)

- To decide between simple random sampling and cluster sampling, look at the availability of information and the variability within and between clusters.
- A simple random sample of payments can give us the information we need.
- But a sample of payments within cases can also give us the information we need, with savings in analysts' time.

Cluster Sampling (cont.)

Pros

- When clusters relevant to our questions are available, cluster sampling may be more practical than other designs.

Cons

- It is usually not as statistically efficient as other sample designs.

Unequal Probability Sampling

- The sampling elements have different chances of being selected.
- When the selection probabilities are related to the size of the entity we're sampling, this is called "sampling with probability proportional to size." In financial audit sampling, the same concept is called "dollar unit sampling."

For example:

- If we were sampling U.S. counties, we might give counties with larger populations a greater chance of selection, or
- If we were selecting credit card transactions for an audit sample, we might give transactions with higher dollar amounts a greater chance of selection.

Unequal Probability Sampling (cont.)

Pros

- Unequal probability sampling can be efficient if the size of the sampling element is related to what is being measured.

Cons

- It is often inefficient if the goal is to estimate the proportion of the elements that have a given characteristic.

Lesson 3: Summary

✓ Deciding on the best sample design generally depends on the available information.

For example:

- ➢ If you're selecting paper files in a storage room, you will probably use systematic sampling.
- ➢ If you're sampling from an MS Excel spreadsheet, you're likely to use a random sample or a stratified random sample.

✓ ARM's statisticians can help you determine the best way to select your sample and can select the sample for you.

Lesson 4: Weighting

- Some Basic Concepts
- Using Weights
- Discussion Example
- Summary

Some Basic Concepts

We weight sample data to ensure that estimates properly account for sampling. In weighting, it's important to remember that sample designs differ.

- In some, all elements in a population have the same chance of being selected.
- In others, different elements have different chances of being selected.

Using Weights

When we analyze sample data, we use sample weights to:

- Account for different chances of selection, and
- Ensure that the sample data properly reflects the population when they're aggregated.

Using Weights (cont.)

After we collect all the data we need, we may modify sample weights to:

- Compensate for nonresponse, or
- Ensure that certain weighted estimates from a sample (e.g. size, type, age, or location) match known population values.

ARM's statisticians can work with you to ensure that your sample data are properly weighted.

Discussion Example

Suppose first that we need to estimate the percentage of 401(K) plans where the sponsor pays for recordkeeping and admin services. We'll divide the population into:

1. **plans with >= 100 participants**
2. **plans with < 100 participants**

Next, we'll select a simple random sample of 50 units from each stratum:

Stratum	Number of 401(K) plans	Sample	Sampling Weight
1	100	50	2
2	900	50	18

Discussion Example (cont.)

The sampling weight is a factor reflecting the sample that is used to make population estimates from sample data.

Here, it is the number of 401(K) plans in the population for the stratum, divided by the number of sample units selected for the stratum.

- In stratum 1, we selected 50 of 100 401(K) plans, so every 1 sample unit represents 2 plans in that stratum.
- In stratum 2, we selected 50 of 900 401(K) plans, so every 1 sample unit represents 18 in that stratum.

Discussion Example (cont.)

- In **stratum 1**, for 40 of the 50 selected plans, plan sponsors paid for administrative services.
- In **stratum 2**, for 10 of the 50 selected plans, plan sponsors paid for administrative services.

Combining the data with the proper use of weights yields:

[(100/50) 40 + (900/50) 10] / 1,000

= [(2) 40 + (18) 10] / 1,000 = .26

Combining the data without weights yields:

(40 + 10)/100 = .5

Discussion Example (cont.)

The estimate is that in 26% of 401(K) plans plan sponsors pay for administrative services.

Why the difference between the unweighted estimate of 50% and the weighted estimate of 26%?

Discussion Example (cont.)

The sampling weight is a factor reflecting the sample design that is used to make population estimates from sample data. Here, it is the number of 401(K) plans in the population for the stratum, divided by the number of sample units selected for the stratum.

– In **stratum 1**, we selected 50 of 100 401(K) plans, so every 1 sample unit represents 2 in that stratum.

– In **stratum 2**, we selected 50 of 900 401(K) plans so every 1 sample unit represents 18 in that stratum.

Lesson 4: Summary

✓ The sample weights depend on the sample design.

✓ Using weights means that each sampled unit has the appropriate influence on an estimate.

Lesson 5: Sources of Error

- Sampling & Nonsampling Error
- Inaccurate Estimates
- Bias & Precision
- Summary

Sampling & Nonsampling Error

Sampling error is error associated with an estimate derived from a sample.

Nonsampling error arises from anything else.

Examples of nonsampling error are:

- Some of the target population was excluded from the sampling frame.
- Responses were not obtained from the entire sample.
- The measuring instrument was imprecise or was misunderstood or had a tendency to differ from the true value in one direction.
- Data processing was subject to errors of logic or other errors.

Inaccurate Estimates

- In the 1936 presidential election, Democratic Incumbent President Franklin Delano Roosevelt was still pushing the New Deal forward in the eighth year of the Great Depression.
- The Republicans nominated Kansas Governor Alf Landon on a platform opposing the New Deal as corrupt and wasteful.
- In its election poll before the November 3 vote, *Literary Digest* mailed 10 million questionnaires to readers and potential readers.
- By October 31, on the basis of 2,300,000 responses, *Literary Digest*–which had correctly predicted the winners of the last five elections–announced that Landon would win, with 370 electoral votes.

Inaccurate Estimates (cont.)

Roosevelt's own prediction was that he would take 270 electoral votes, Landon 178.

(An unknown at the time, George Gallup predicted, from his random sample of 50,000 people, that Roosevelt would win.)


AMERICA SPEAKS
THE NATIONAL WEEKLY POLL of PUBLIC OPINION
Today
Next Sunday
Institute Forecasts the Re-election of Franklin D. Roosevelt, Gives Him 54% of Popular Vote, Minimum of 315 Electors
Major Party Percent Is 55.7; New York in F.D.R. 'Sure' Column
Election Forecast
Election Will Test Clashing Poll Methods

Inaccurate Estimates (cont.)

In the end, FDR carried every state except Maine and Vermont, winning:

- 98.5% of the electoral vote, and
- 60.8% of the popular vote.

Question:

Why did *Literary Digest's* sample procedure get it wrong?

First, let's look at the prediction and result for the popular vote.

Candidate	Literary Digest's Oct. 31 prediction	Outcome of Nov 3 popular vote
Landon	55%	37%
Roosevelt	41%	61%

Inaccurate Estimates (cont.)

Literary Digest's sample was drawn from its subscribers and from telephone directories and car registrations.

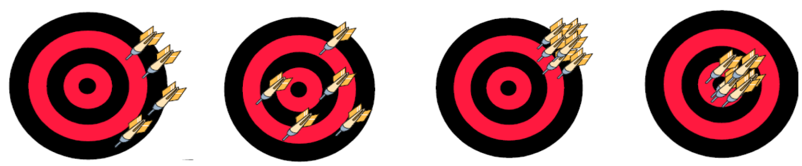
However,

- During the Depression, households that could afford subscriptions, telephones, and cars were generally more affluent than others, and
- Economic status tended to affect respondents' opinions of economic policies.

The sample frame, and therefore the sample was biased.

Bias and Precision

The contrasting skills of four different archers help illustrate bias and the precision of an estimator.



Archer 1	Archer 2	Archer 3	Archer 4
Biased, imprecise	Unbiased, imprecise	Biased, precise	Unbiased, precise

Bias and Precision (cont.)

- **Archer 1** has a tendency to miss the bull's eye to the right: he's a bad shot and biased.

- **Archer 2,** not a bad shot overall – the average position of all arrows is at the bull's eye – is unbiased.

- **Archer 3's** arrows are all close together but systematically away from the bull's eye; he's precise but not unbiased.

- **Archer 4's** arrows are all close together and near the center of the target; she's accurate.

Lesson 5: Summary

- ✓ We try to control nonsampling errors in order to mitigate sources of bias.
- ✓ In probability samples, the goal is to generate estimates that are precise enough to answer the engagement's objectives.

Lesson 6: Identifying Resources

- ARM Guidance
- Resources

ARM Guidance

ARM's web site on GAO's intranet contains several guidance papers on statistical topics.

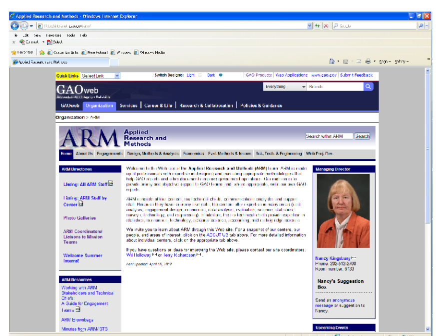
Here, among others, are four:

- Calculating and Reporting Survey Response Rates
- Documenting Sample Design and Estimates
- Reporting Results Based on Small Data Sets
- Using Probability, Nonprobability and Certainty Samples

ARM Guidance (cont.)

How to find ARM's papers?

On GAO's intranet home page, go to Organization / Applied Research and Methods (ARM). Then click on "Statistics and Sampling" under "Guidance, Tools, & Resources in the lower left corner. Scroll down to find "Statistics and Sampling."



Or type
http://intranet.gao.gov/arm/about_us/methodology_guidance_tools_and_resources/statistics_and_sampling

Types of Sampling

Primary Sources

- Cochran, William G. Sampling Techniques, 3rd ed. Hoboken, N.J.: John Wiley & Sons, 1977.
- Kish, Leslie. Survey Sampling, new ed. New York: Wiley-Interscience, 1995.
- Lohr, Sharon L. Sampling: Design and Analysis. Pacific Grove, Calif.: Duxbury Press, 1999.
- Roberts, Harry V., with W. Allan Wallis. Statistics: A New Approach. Glencoe, Ill.: Free Press, 1956.

Types of Sampling (cont.)

Secondary Sources

Simple Random Sampling

- GAO. Food Stamp Program: Farm Bill Options Ease Administrative Burden, but Opportunities Exist to Streamline Participant Reporting Rules among Programs, GAO-04-916. Washington, D.C.: Sept. 16, 2004.
- GAO. Vehicle Donations: Benefits to Charities and Donors, but Limited Program Oversight, GAO-04-73. Washington, D.C.: Nov. 13, 2003.

Types of Sampling (cont.)

Secondary Sources

Stratified Random Sampling

- GAO. 2011 Lobbying Disclosure: Observation on Lobbyists' Compliance with Disclosure Requirements, GAO-12-492. Washington, D.C.: March 30, 2012.
- GAO. Mortgage Foreclosures: Regulatory Oversight of Compliance with Servicemembers Civil Relief Act Has Been Limited, GAO-12-700. Washington, D.C.: July 17, 2012.
- GAO. Dodd-Frank Act: Hybrid Capital Instruments and Small Institution Access to Capital, GAO-12-237. Washington, D.C.: January 18, 2012.

Types of Sampling (cont.)

Secondary Sources

Cluster Sampling

- GAO. Voters With Disabilities: Additional Monitoring of Polling Places Could Further Improve Accessibility, GAO-09-941. Washington, D.C.: Sept. 30, 2009.
- GAO. Property Management: NASA's Goal of Increasing Equipment Reutilization May Fall Short without Further Efforts, GAO-09-187. Washington, D.C.: January 30, 2009.
- GAO.401(K) Plans: Increased Educational Outreach and Broader Oversight May Help Reduce Plan Fees, GAO-12-325. Washington, D.C.: April 24, 2012.

Types of Sampling (cont.)

Secondary Sources

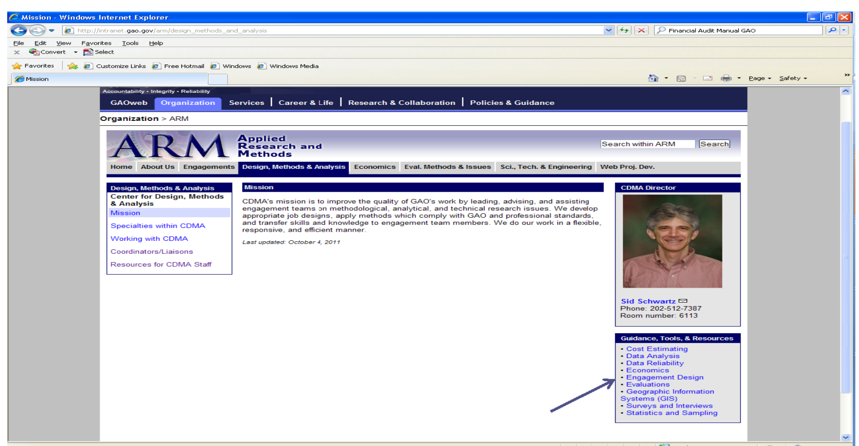
Systematic Sampling

- GAO. Private Pensions: Publicly Available Reports Provide Useful but Limited Information on Plans' Financial Condition, GAO-04-395. Washington, D.C.: Mar. 31, 2004.

Unequal Probability Sampling

- GAO. Property Management: NASA's Goal of Increasing Equipment Reutilization May Fall Short without Further Efforts, GAO-09-187. Washington, D.C.: January 30, 2009.
- GAO. Schools and Libraries Program: Application and Invoice Review Procedures Need Strengthening, GAO-01-105. Washington, D.C.: Dec. 15, 2000.
- GAO. Debt Collection Improvement Act of 1996: Department of Agriculture's Farm Service Agency Has Not Yet Fully Implemented Certain Key Provisions, GAO-02-463. Washington, D.C.: Mar. 29, 2002.

Applied Research & Methods Guidance



You can always reach the CDMA people and other ARM resources at http://intranet.gao.gov/arm/design_methods_and_analysis

Types of Sampling (cont.)

Secondary Sources

Systematic Sampling

- GAO. Private Pensions: Publicly Available Reports Provide Useful but Limited Information on Plans' Financial Condition, GAO-04-395. Washington, D.C.: Mar. 31, 2004.

Unequal Probability Sampling

- GAO. Property Management: NASA's Goal of Increasing Equipment Reutilization May Fall Short without Further Efforts, GAO-09-187. Washington, D.C.: January 30, 2009.
- GAO. Schools and Libraries Program: Application and Invoice Review Procedures Need Strengthening, GAO-01-105. Washington, D.C.: Dec. 15, 2000.
- GAO. Debt Collection Improvement Act of 1996: Department of Agriculture's Farm Service Agency Has Not Yet Fully Implemented Certain Key Provisions, GAO-02-463. Washington, D.C.: Mar. 29, 2002.

Applied Research & Methods Guidance

You can always reach the CDMA people and other ARM resources at http://intranet.gao.gov/arm/design_methods_and_analysis

Appendix 1 Confidence Intervals Example

Confidence Interval Graph

The following graph shows 60 confidence intervals from 60 different samples of the same size from the same population.

"E" denotes the sample estimate from each sample. It varies from one sample to another.

"P" denotes the population value. It is in bold if the confidence interval does not contain the population value.

```
                 [---------E--P-------]
               [------------E-P---------]
                  [-----------PE-----------]
      [----------------E------P--------]
            [-------------E---P---------]
                      [-------P-------E--------------]
                        [-----P-------E-------------]
                            [-P-----------E-------------]
            [------------E----P-------]
                 [------------P-E-------------]
            [-----------E-----P-----]
               [------------E-P---------]
                           [--P-----------E--------------]
                    [---------P--E------------]
                       [------P-------E--------------]
              [--------------EP-------------]
                   [----------P--E------------]
[--------------E-------------]P
              [------------E--P---------]
                           [--P------E--------]
    [-------------E-----------P]
                    [---------P----E--------------]
                 [------------P-E-------------]
                [-------------P-E--------------]
                       [------P------E------------]
                  [-----------P-E-------------]
                     [--------P----E-------------]
        [-------------E-------P-----]
                           [--P-----------E-------------]
       [--------------E-------P-----]
                [-------------PE--------------]
         [----------E---------P]
           [-------------E----P-------]
                             [P-----------E-----------]
      [-----------E----------]P
                   [----------PE----------]
        [--------------E------P------]
                      [-------P-----E-------------]
                    [---------P-----E---------------]
                   [----------P--E-------------]
                          [---P----------E--------------]
          [------------E------P-----]
                              P  [--------------E--------------]
                     [--------P------E--------------]
            [----------------EP--------------]
              [-------------E-P----------]
              [------------E--P---------]
                 [-----------EP---------]
          [------------E------P-----]
        [--------------E------P-------]
    [-------------E-----------P]
               [-------------EP------------]
    [-------------E-----------P]
                        [-----P--------E--------------]
        [--------------E------P------]
                    [---------P---E------------]
           [-------------E----P--------]
               [-----------E--P--------]
                      [-------P-----E------------]
                        [-----P------E------------]
```