Sampling Plan and Statistical Considerations for the Long Term Field Evaluations Project

Doug Landsittel, PhD

Senior Service Fellow

Outline

- Outcome Definition
- General Approach
- Estimating the Required Sample Size
- Overall Sampling Strategy
- Additional Sampling Issues
- Secondary Analyses
- Conclusions

Outcome Definition

- **Criteria:**
  - Minimum O2 ≥ 15%
  - Maximum CO2 ≤ 4%
  - Duration ≥ rated duration
- **Formulate a pass/fail outcome**
- **Ideal Goal: 0% failures**
- **Statistical Goal: observe ≥ 1 failure if the underlying failure rate is sufficiently high**

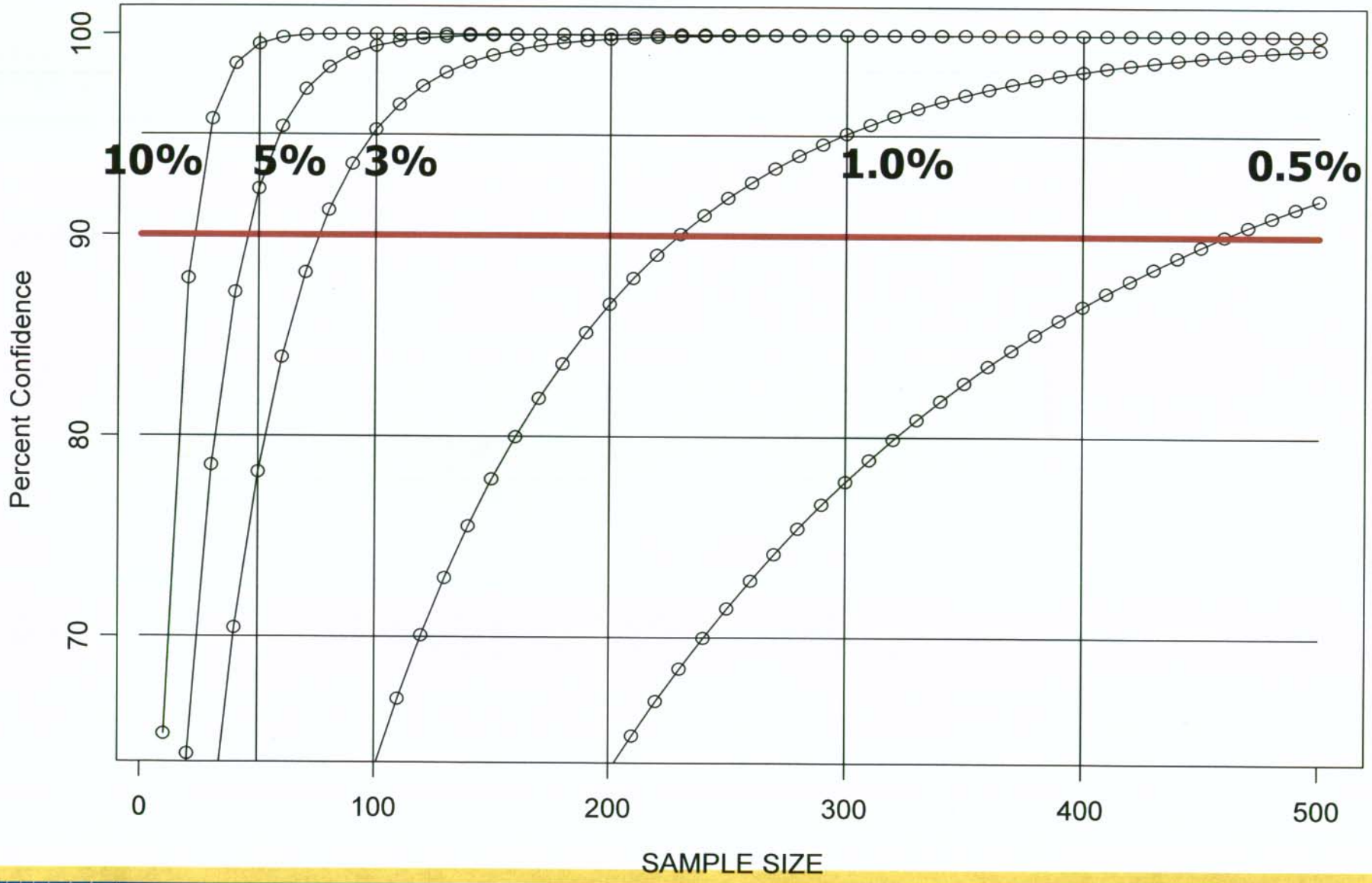
General Approach

- Assumed failure rate
- Require a certain percent confidence to observe ≥ 1 failure
- Solve for the required sample size
- Vary assumed underlying failure rate
- Alternatively, fix the sample size
- Make conclusions regarding sample size
- Other factors in the design

Some Example Results

- **Assumed failure rate of 1%**
  - Requires n = 230 to have 90% confidence that we will observe ≥1 failure
- **Assumed failure rate of 2%**
  - Requires n = 115
- **Assumed failure rate of 3%**
  - Requires n = 78
- **Assumed failure rate of 0.5%**
  - Requires n > 500

Percent Confidence by Sample Size and Failure Rate



Examples for a Sample Size of 100

- **Underlying failure rate of 1%**
  - –63% confidence to observe ≥ 1 failure
- **Underlying failure rate of 2%**
  - –87% confidence to observe ≥ 1 failure
- **Underlying failure rate of 3%**
  - –95% confidence to observe ≥ 1 failure

Overall Sampling Strategy

- Determine the total annual sample size (N)
- Use the MSHA inventory to identify all units
- Stratify by the k approved models
- Within each approved model, randomly order the list and select the 1st n units (N/k)
- Evaluate the confidence to observe ≥ 1 failure
- Consider the analysis of data over multiple years

Conclusions about Sample Size

- **Proposing 400 units tested annually**
- **Primary conclusions specific to a given approved model (currently 4)**
  - n = 100 per year per approved model
  - Assuming a sufficiently large population
  - Will modify statistical plans for smaller populations based on MSHA inventory
- **Sufficient sample to observe a failure for a given model if the true rate is 2-3%**

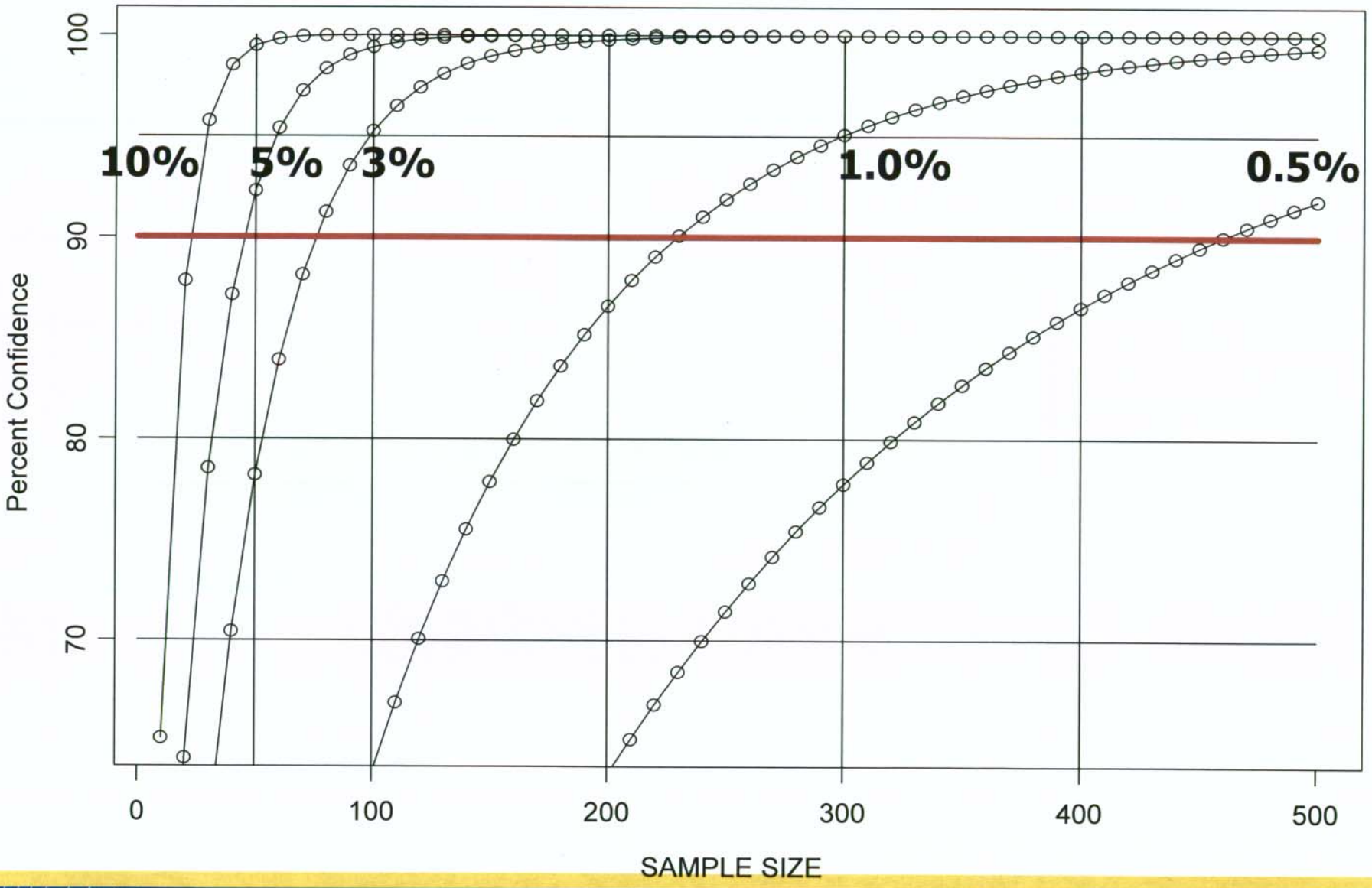
Additional Sampling Issues

- **Starting Point: k ordered lists**
  - Suspect some missing units
  - Must meet the manufacturer's inspection criteria
- **Will document all events of units not tested**
- **Continue collecting according to the randomly ordered list**
- **Sampling plans allow for programmatic flexibility**
  - EXP: may collect additional units in advance
- **Will maintain strict randomness in the units tested**

Statistical Goals for a Given Year

- **n = 100 gives high confidence to observe a failure for models with a failure rate > 2%**
  - 87% confidence given a 2% failure rate
  - 95% confidence given a 3% failure rate
- **n = 75 still gives reasonable confidence to observe a failure rate > 2%**
  - 78% confidence given a 2% failure rate
  - 90% confidence given a 3% failure rate

Percent Confidence by Sample Size and Failure Rate



Number of Units Sampled versus Percentage of the Population

- **Two main issues in specifying sample size**
  - Level of statistical confidence
  - Representativeness of the sample
- **Level of statistical confidence depends on the number of units sampled**
  - Exception: correction for smaller population
  - Will incorporate based on MSHA inventory
- **Representativeness of the sample**
  - Accomplished in this study via 1) identifying the entire population and 2) adhering to strict random sampling

Analysis for Subsequent Years

- **Accumulate data based on fixed design parameters**
- **Two main goals:**
  - Observe a failure with high confidence assuming lower failure rates
  - Observe a failure with high confidence within a subgroup of units
- **Conclusions from multiple years of data are only relevant to some parameters**
- **Employ statistical approaches to account for systematic differences across different years**

Analyzing Subgroups of Units

- **Subgroups include (but are not limited to)**
  - Deployment, age, mine height, mine size
  - Some parameters will be difficult to measure
- **Making statements about the percent confidence**
  - Depends on the proportion of units within a given subgroup
  - Depends on systematic differences across years
- **Consider as a secondary analysis**

Other Secondary Analyses

- **Collect all available information**
- **Additional outcomes**
  - Rates of units that are missing or do not meet the manufacturer's inspection criteria
  - Minimum O2, maximum CO2 and duration
- **Descriptive analysis: summaries and graphs**
- **Exploratory in nature**

Strengths and Limitations

- **Significant improvement over existing data**
- **MSHA inventory: complete enumeration of the entire population of units**
  - Random sample within each approved model
  - Calculate rates of units that are missing or do not meet the manufacturer's inspection criteria
- **Maintaining strict randomization minimizes resulting bias**

Strengths and Limitations

- **Challenges in random sampling**
  - Geographical locations or finding specific serial numbers
  - Developing strategies for programmatic flexibility while maintaining randomization
- **Future considerations might include alternative study designs**

Interpretation of Results

- **Cannot guarantee performance of every unit with 100% confidence**
- **We can say that if a given model has at least 2-3% of units that do not meet the performance criteria**
  - We will, with a high level of confidence, observe at least1 failure among sampled units

Conclusions

- **Strong experimental design**
  - MSHA inventory and randomization to minimize any systematic errors
- **Standard statistical approaches**
- **High degree of confidence to observe a failure under a range of assumptions**