

# Economic Analysis of Optimal Sample Size in Legal Proceedings

Aren Megerdichian

Correspondence: Aren Megerdichian, Compass Lexecon, 55 S. Lake Avenue, Ste 650, Pasadena CA 91101, United States. E-mail: ArenMeg@CompassLexecon.com

Received: May 6, 2019 Accepted: June 4, 2019 Online Published: June 11, 2019

doi:10.5539/ijsp.v8n4p13

URL: <https://doi.org/10.5539/ijsp.v8n4p13>

## Abstract

In many legal settings, a statistical sample can be an effective surrogate for a larger population when adjudicating questions of causation, liability, or damages. The current paper works through salient aspects of statistical sampling and sample size determination in legal proceedings. An economic model is developed to provide insight into the behavioral decision-making about sample size choice by a party that intends to offer statistical evidence to the court. The optimal sample size is defined to be the sample size that maximizes the expected payoff of the legal case to the party conducting the analysis. Assuming a probability model that describes a hypothetical court's likelihood of accepting statistical evidence based on the sample size, the optimal sample size is reached at a point where the increase in the probability of the court accepting the sample from a unit increase in the chosen sample size, multiplied by the payoff from winning the case, is equal to the marginal cost of increasing the sample size.

**Keywords:** cost-benefit analysis, economic model, legal proceeding, sample size, sampling

## 1. Introduction and Background

Legal and regulatory proceedings are often complex and can be troubled with large amounts of data or a large number of elements that require study. To adequately research, investigate, and study each and every element may require prohibitively costly resources. Thus, in a variety of contexts, relying on a statistical sample is a sensible approach to effectively adjudicate questions of causation, liability, or damages.

The need to sample can arise in many situations. For example, in copyright matters, it may not be feasible to research and investigate the copyright and royalty status of tens of millions of songs downloaded from a website. In class action suits, it may not be feasible to study the potential adverse health effects on hundreds of thousands of individuals from consuming the allegedly harmful product. As a condition for obtaining regulatory approval of a merger between two Internet service providers, it may not be feasible to individually study whether millions of households suffered a meaningful degradation of services years after the merger took place. Thus, a subset of the population of elements can be sampled and studied further, and, under certain conditions, the sample will be an effective surrogate for the full population.

Sampling in legal matters has not always been accepted by the courts. Caughey (1956) notes that surveys conducted to determine the likelihood of confusion in trademark infringement cases were typically not accepted by courts prior to the 1950s. A number of rulings helped turn around the acceptability of surveys in trademark infringement cases, including *United States v. 88 Cases*;<sup>1</sup> *Sears, Roebuck & Co. v. Johnson*;<sup>2</sup> and *Household Finance Corp. v. Federal Finance Corp.*<sup>3</sup> The use of sampling in a variety of legal contexts is now routinely accepted by the courts. In *U.S. ex rel. Martin v. Life Care Centers of America*, for instance, the court denied defendant's motion for partial summary judgement as to the Government's use of statistical sampling.<sup>4</sup> The court found that statistical sampling would not violate defendant's right to due process.<sup>5</sup> In *Blue Cross and Blue Shield of New Jersey v. Philip Morris*, the court held that statistical

<sup>1</sup> 187 F.2d 967 (3d Cir. 1951).

<sup>2</sup> 219 F.2d 590 (3d Cir. 1955).

<sup>3</sup> 105 F. Supp. 164 (D. Ariz. 1952).

<sup>4</sup> Civ. No. 1:08-cv-251, 1:12-cv-64, slip op. at 26 (E.D. Tenn., 2014).

<sup>5</sup> Civ. No. 1:08-cv-251, 1:12-cv-64, at 36 (E.D. Tenn., 2014). Noting that "courts that have considered the issue of statistical extrapolation to calculate overpayment have found that it is an acceptable practice which does not violate a defendant's due process rights" and that the "Defendant will be afforded due process by having the opportunity to depose the Government's expert, challenge the qualifications of the Government's expert, retain its own expert, and to

methods do not violate defendants' Seventh Amendment rights.<sup>6</sup> Gastwirth (2003) provides a review and discussion of legal cases using samples as evidence.

Analyzing a sample can reduce the amount of evidence and contribute toward creating a reasonable trial length. This was recognized early on in *United States v. United Shoe Machinery Corp.*, in which the court explained, "[i]f antitrust trials are to be kept manageable, samples must be used, and a sample which is in general reasonable should not be rejected in the absence of the offer of a better sample."<sup>7</sup> The potential for excessive amounts of data to get in the way of efficiently adjudicating a case was a concern for the courts even in the 1950s. Courts and regulators should continue to appreciate that time and resources are finite, so it may be more effective to investigate a relatively small but reliable sample of subjects than to unreliably analyze all subjects in the population in question.<sup>8</sup>

The current paper works through salient aspects of statistical sampling and demonstrates that sample size determination is an economic question involving the tradeoff between the marginal economic benefit and the marginal cost of analyzing a larger sample size. In Section 2, the most common type of sampling—simple random sampling—is assumed, and a measure that is often of interest in empirical studies—the proportion—is examined for the purpose of describing sample size calculations employed in Section 3. In Section 3, an economic model is developed that posits a court's preference for larger samples, which provides insight into a party's behavioral decision-making when determining the sample size to rely on in a legal proceeding. The concept of optimality here encompasses maximizing expected payoff to the party submitting statistical evidence based on a sample. The analysis reveals that the optimal sample size chosen by the party presenting statistical evidence is reached where the increase in the probability of the court accepting the sample from a unit increase in the chosen sample size, multiplied by the payoff from winning the case, is equal to the marginal cost of increasing the sample size by a unit. Section 4 concludes.

## 2. Overview of Sample Size Determination in Legal Proceedings

An important aspect of sampling design is sample size determination, which inevitably becomes a focal point when statistical evidence based on a sample is introduced in legal proceedings. A decision must ultimately be made by the party offering statistical evidence about how large a sample to draw from the population. The analysis in Section 3 shows that it is essentially an economic question about the tradeoffs between the costs and benefits of a larger sample size. This section works through some relevant aspects of sample size determination in legal proceedings, which are relied on in Section 3.

If a sample gets to be as large as the population, then the sample contains all the information in the population and thus becomes a perfect surrogate for the population. The more-is-better approach is an intuitive way to think about sample size determination, but such an approach does not account for the cost associated with more observations. There is no one-size-fits-all rule for determining sample sizes. Courts may be reluctant to accept evidence based on samples for different reasons and in a variety of contexts. Boardman and Vining (1983) and Gastwirth and Xu (2014) provide discussions of the courts' reluctance to accept evidence based on small samples in the context of discrimination suits. Kaye and Freedman (2011) note that samples presented in legal cases have ranged from 5 to 1.7 million. Legal precedent may not be helpful in informing questions on sampling methodology and sample size determination. Instead, sampling theory must be relied on.

It may not be feasible to research and investigate a very large number of elements because there are often constraints on time, budget, or other resources that would make it practically impossible to study every element of the population. In *Blue Cross and Blue Shield of New Jersey v. Philip Morris*, the court noted that "[r]equiring individual proof as to each claim would unnecessarily intrude on the lives of hundreds of thousands of people. Examining each of the grains of

---

present all of this evidence at trial."

<sup>6</sup> 113 F. Supp. 2d 376 (E.D.N.Y.). Supporting that "[t]o accept defendants' contention that the aggregation of Empire's subrogation claims violates the Seventh Amendment would require concluding that the Amendment establishes fixed limitations on the methods of proof a *particular party* may offer in support of *its own claims*... Such a ruling would run contrary to decades of federal law in areas such as employment, copyright and patent law where the use of statistical evidence is common."

<sup>7</sup> 110 F. Supp. 295 (D. Mass. 1953), *affd*, 347 U.S. 521 (1954).

<sup>8</sup> In a related context, Cheng (2012) discusses the statistical merits for trying a sample of mass tort cases rather than trying all of the cases. Adjudicating thousands of related cases is arguably infeasible, potentially warranting the approach of trying only a sample of the cases and extrapolating the results to the population. See Lahav (2008) for a discussion of the legal arguments related to bellwether trials.

sand on the beach is too burdensome.”<sup>9</sup> Thus, in the context of litigation and regulatory proceedings, a reasonable question to first consider in determining the sample size is what resources are available for the particular analysis at hand.<sup>10</sup> Generally speaking, the more resources available to research and study a particular sample of subjects, the larger can be the sample.

Suppose it is expected to cost \$50, on average, to research and investigate a randomly drawn element from the population to determine whether it meets a specified condition of interest, and suppose there is a budget of \$100,000 available to spend on this aspect of the analysis.<sup>11</sup> This suggests that a sample size of 2,000 is going to be the largest that is economically feasible. Depending on circumstances, there may be some flexibility in the budget to increase the sample size. Obtaining additional resources would be warranted if it is thought to be the difference between a larger sample that is going to deliver results with a sufficient level of precision that is likely to be accepted by the court versus a smaller sample that would be deemed by the court to not have probative value.

Thus, there are two ways to think about sample size. First, a party may have no choice but to work with a certain sample size due to resource constraints. It is then important to understand what is the most that can be claimed about the precision with which inferences will be made from the analyses of a particular sample size. Second, if a party does not face binding budget and time constraints, one might first ask what is the least they are willing to be able to claim about the precision with which inferences will be made from a sample, and then determine the sample size that will achieve that claim.

Even when operating under constraints, it may be prudent to reallocate resources in an attempt to relax those constraints if doing so will result in a sample that will be accepted by the court. For instance, it may be that estimating a proportion with a margin of error equal to three percent is acceptable to the court, but five percent is not. Achieving the former will require a larger sample than the latter, all else equal. Obtaining a larger sample might require procuring additional budget, reallocating resources from other parts of the case to the sampling budget, reallocating personnel to research and investigate the sample, and possibly requesting that the court schedule be revised to accommodate the additional time needed to research and investigate the larger sample.

Even in matters where time and budget are not major concerns, it is often unnecessary and inefficient to draw and study an extremely large sample simply because one can afford to do so. Increasing the sample size can be thought of as purchasing additional bits of information. The larger the sample, the more information that will be available, but the incremental value of the additional information wanes as the sample gets larger. For instance, the value of an additional 100 units of information will be more valuable in going from 500 to 600, than it is in going from 1,500 to 1,600. This phenomenon, known by economists as the law of diminishing marginal returns, applies to statistical sampling. A larger sample is better than a smaller sample because there is more information that allows for more precision with which inferences can be made about the population, but each additional unit of information becomes less valuable than the preceding additional unit. This is explored in more detail in Section 3. Ultimately, relying on a sample that is too small may result in an undesirable level of precision that won't be accepted by the court; yet a sample that is too large may be an inefficient use of scarce resources.

A proportion that quantifies the share of elements in the population that meet a specified condition is relied on to explore the sampling analysis in Section 3. Assuming simple random sampling,<sup>12</sup> the sample size ( $n$ ) ensures that the

<sup>9</sup> 113 F. Supp. 2d 374 (E.D.N.Y. 2000).

<sup>10</sup> Some very large data sets with millions of observations are available and require no further research and investigation to determine if each observation meets the condition of interest, or if a continuous variable of interest is provided and no further data gathering is needed. In such cases, the relevant information for each record in the data is available, potentially kept in the normal course of business or available from third-party sources, so that analysis can commence without further research. Thus, there may be no practical need to focus on a sample in such cases, except where the data set is so voluminous that it is practically impossible to analyze given time and resource constraints. Where there will almost always be a need to sample is when some additional, costly level of research and investigation is required to determine whether each observation meets the condition of interest, or, in the context of continuous variables, to expend efforts to determine the value of the variable.

<sup>11</sup> For example, a researcher has a billing rate of \$200 per hour, and is able to research a record and determine whether it meets the condition of interest in 15 minutes. On average, it will cost \$50 to research each element in the sample.

<sup>12</sup> Whereby  $n$  sample elements are selected from the population of  $N$  elements in such a manner that all the various combinations of  $n$  elements that could be selected have the same probability of being the sample that is selected. This implies that each element of the population has the same chance ( $n/N$ ) of being selected into the sample. The analysis assumes sampling without replacement.

distance between the estimated sample proportion ( $\hat{p}$ ) and the true population proportion ( $p$ ) being less than a tolerable margin of error ( $\varepsilon$ ) is equal to a pre-specified probability ( $1 - \alpha$ ). That is, the  $n$  that satisfies  $\text{Prob}(|\hat{p} - p| < \varepsilon) = 1 - \alpha$  is the familiar sample size calculation:<sup>13</sup>

$$n = \frac{Np(1-p)}{(N-1)\left(\frac{\varepsilon}{z}\right)^2 + p(1-p)}. \quad (1)$$

The following discussion provides some useful guidance to practitioners on the inputs required in the sample size calculation in (1) and simplifying assumptions that are made in Section 3 to analyze the economics of optimal sample size.

Population parameters—for example, the proportion  $p$  here—that are necessary to determine the sample size are almost always unknown in practice.<sup>14</sup> One must either rely on past studies, expert judgement, or other guiding principles to determine an approximate value for  $p$ . Drawing a small random sample on a preliminary basis to initially research and investigate may be warranted, which will provide some guidance on the magnitude of the population parameter to use in the sample size calculation.

When no other information is available, a proportion assumed to equal 50% will result in a sample that is larger than needed for the claims about precision that are desired. Table 1 provides the necessary sample size to be 95% confident that the estimated sample proportion will be within three percentage points of the various population proportions given along the top row. A number of hypothetical population sizes are provided along the left column. As demonstrated in the table, larger samples are necessary when the assumed population proportion is 50%. For ease of exposition Section 3 assumes  $p = 0.50$ .

Table 1. Necessary sample size ( $\varepsilon=0.03$ ,  $z=1.96$ )

		p: Population Proportion				
		10%	25%	50%	75%	90%
N: Population Size	1,000	278	445	517	445	278
	5,000	357	690	880	690	357
	10,000	370	742	965	742	370
	50,000	382	788	1,045	788	382
	100,000	383	794	1,056	794	383
	1,000,000	384	800	1,066	800	384
	100,000,000	385	801	1,068	801	385

As the size of the population,  $N$ , becomes large, it becomes less important in the sample size calculation in (1). It might be counterintuitive, and even “remarkable” according to Cochran (1999), that the necessary sample size to achieve a certain margin of error is not much different whether the number of elements in the population is 50,000 or 100 million, all else equal. See Table 1. This stems from the fact that the population size appears in the sample size formula due to the finite population correction, which is applied to the variance of the sample statistic to account for the sample possibly being drawn from a finite, rather than infinite, population.  $N$  will matter for determining the sample size when it is small—suggesting that the sample size relative to the population size could be large—and so the reduction in sampling variability from having a relatively large amount of information about the population must be accounted for in the calculation. See, e.g., Lohr (2010). When the population size is infinitely large ( $N \rightarrow \infty$ ), a simpler sample size calculation obtains,  $p(1-p)/(z^{-1}\varepsilon)^2$ . For tractability, and without loss of generality, the optimal sample size example in Section 3 assumes  $N \rightarrow \infty$ .

<sup>13</sup> The finite variance of  $\hat{p}$  is given by  $\text{var}(\hat{p}) = p(1-p)(N-n)/n(N-1)$ ; dividing  $|\hat{p} - p| < \varepsilon$  by  $\sqrt{\text{var}(\hat{p})}$  and rearranging terms gives  $\text{Prob}(-z < Z < z) = 1 - \alpha$ , where  $Z = (\hat{p} - p)/\sqrt{\text{var}(\hat{p})}$ ,  $z$  is the normal  $z$ -value associated with the  $0.5\alpha$  tail of the distribution, and  $p$  is assumed to be normally distributed. Thus,  $\varepsilon/\sqrt{\text{var}(\hat{p})} = z$ ; combining with the expression for  $\text{var}(\hat{p})$  and solving for  $n$  gives the result.

<sup>14</sup> Generally speaking, it is the population variance that is needed in sample size calculations. When the measure of interest is a proportion, as in the example studied here, the variance is a function of the proportion  $p$ . For a continuous variable, the measure of interest might be the average, in which case the population variance  $p(1-p)$  in the sample size equation presented earlier is replaced with the variance of the continuous variable,  $\sigma^2$ .

The input  $z$  is the normal distribution value associated with the chosen significance level  $\alpha$ , where  $\alpha$  reflects the chance of rejecting the null hypothesis when it is true, or Type-I error.<sup>15</sup> There is no guarantee that the random sample that is drawn from the population will generate an interval that encompasses the true population parameter. Thus, the confidence level must be specified, representing the probability of not drawing such an “unlucky” sample that delivers accuracy outside the acceptable margin of error.<sup>16</sup> It is common to choose  $\alpha$  equal to 5%, corresponding to a  $z$  value of approximately 1.96.<sup>17</sup>

Table 2 provides the necessary sample size to be  $1 - \alpha$  percent confident that the estimated sample proportion will be within a given margin of error of the assumed population proportion 50%. The larger the desired confidence level, the larger the sample must be to achieve it. For simplicity, the optimal sample size example analyzed in Section 3 specifies a confidence level equal to 95%.<sup>18</sup>

Table 2. Necessary sample size ( $p=0.50$ ,  $N \rightarrow \infty$ )

		(1 - $\alpha$ ) Percent Confidence Level				
		90.0%	92.5%	95.0%	97.5%	99.0%
$\epsilon$ : Margin of Error	1%	6,764	7,926	9,604	12,560	16,588
	2%	1,691	1,982	2,401	3,140	4,147
	3%	752	881	1,068	1,396	1,844
	5%	271	318	385	503	664
	7%	139	162	196	257	339
	10%	68	80	97	126	166

The preceding inputs to the sample size calculation are typically known or assumed in practice. The central question then focuses on the acceptable level of precision, or margin of error, denoted by  $\epsilon$ , between the sample estimate and the population parameter. A tolerable amount of error must be specified so that a sufficient sample size can be drawn to meet it. More precision requires larger sample size. The margin of error is often the defining input that will determine the sample size, yet there is no universally acceptable margin of error that is appropriate in all situations. Choosing a small margin of error equal to one percent compared to choosing five percent comes at the cost of requiring one to study a substantially larger sample, all else equal. See Table 2. A margin of error in the neighborhood of three to five percent, which tends to be relied on by polling organizations, is a reasonable starting point to consider.<sup>19</sup> However, a margin of error that is higher or lower may be warranted depending on the circumstances at hand. Both sides in an adversarial proceeding will present their arguments. For example, in *Re Countrywide Financial Corp. Mortgage-backed*, plaintiffs

<sup>15</sup> The distribution of the number of elements in the sample that meets the condition of interest is hyper-geometrically distributed, but it is common to instead assume that the sample proportion is normally distributed for computational simplicity. The normal approximation gets better as the sample size increases and the closer the proportion is to 50%. Mathews (2010) suggests that the normal approximation is justified if  $np > 5$  and  $np(1 - p) > 5$ .

<sup>16</sup> A common misinterpretation is that a 95% confidence level means that there is a 95% probability that the population parameter will fall within the confidence interval. Rather, the correct interpretation is that a random interval will have a 95% chance of containing the population parameter. See Larsen and Marx (2001) and Kaye and Freedman (2011). For example, a 95% confidence level means that if 10,000 random samples were drawn from the population, 9,500 of the samples are expected to generate a confidence interval,  $\hat{p} \pm \epsilon$ , that encompasses the population parameter,  $p$ .

<sup>17</sup> See Stock and Watson (2007).

<sup>18</sup> Gastwirth and Xu (2014) analyze the role of power in statistical testing in the context of legal cases. For simplicity, the analyses here and in Section 3 only consider statistical significance.

<sup>19</sup> For example, in a recent Gallup poll showing that 53% of Americans are “confident” about their financial future, Gallup explained that the random sample of 1,530 adults resulted in a margin of error equal to  $\pm 3\%$  at the 95% confidence level. See Gallup, *In U.S., slim majority confident about financial future*, June 24, 2016, available at [http://www.gallup.com/poll/193139/slim-majority-confident-financial-future.aspx?g\\_source=Politics&g\\_medium=newsfeed&g\\_campaign=tiles](http://www.gallup.com/poll/193139/slim-majority-confident-financial-future.aspx?g_source=Politics&g_medium=newsfeed&g_campaign=tiles). Another recent Gallup poll, showing that 32% of Americans have a “great deal/fair amount” of trust in the mass media, was based on a sample of 1,020 adults, resulting in a margin of error equal to  $\pm 4\%$  at the 95% confidence level. See Gallup, *Americans’ trust in mass media sinks to new low*, September 14, 2016, available at <http://www.gallup.com/poll/195542/americans-trust-mass-media-sinks-new-low.aspx>.

argued that a 10% margin of error was a reliable application of statistical methods, while defendants suggested a smaller margin of error was required.<sup>20</sup>

As shown in Table 2, samples do not have to be very large, even when the population is assumed to be infinitely large, if less precision is acceptable. Whether a particular sample size is appropriate or not cannot be determined until the precision with which one wants to make inferences about the population is established. Fixing the confidence level to some acceptable quantity, such as 95%, leaves the margin of error as the determining factor for how large the sample size will need to be. The next section presents an economic model that provides insights into how a party to litigation might think about optimal sample size, given its understanding of the court's likelihood of accepting statistical evidence based on a sample.

### 3. An Economic Model of Optimal Sample Size in a Legal Proceeding

A theoretical economic model of the sample size choice problem in the context of a hypothetical legal proceeding is presented. A logistic probability model of the court's preference for sampling precision is introduced, and, given the rewards and costs of litigation, an optimal sample size is obtainable from the resulting nonlinear expected payoff function. The analysis developed here provides guidance and insights into the thinking that is theoretically carried out when a party must decide on a sample size in a legal proceeding. As shown in this section, the optimal sample size is reached where there is equality between the marginal probability-adjusted payoff from winning the court case and the marginal cost of analyzing the sample. Thus, the question of optimal sample size in this context is fundamentally an economic question.

#### 3.1 Optimality Condition

Assume that resources to study a sample are sufficiently flexible and a sample size must be chosen by a party to conduct a study to ascertain a proportion of interest based on simple random sampling, as described in Section 2. The margin of error in (1) explicitly as a function of the sample size is

$$\varepsilon = \sqrt{\frac{p(1-p)z^2(N-n)}{n(N-1)}}. \quad (2)$$

The slope of the margin of error, provided below in (3), is negative, confirming that a larger sample size results in a smaller margin of error.

$$\frac{\partial \varepsilon}{\partial n} = \frac{-p(1-p)z^2N}{2\varepsilon n^2(N-1)} \quad (3)$$

Moreover, the curvature of the margin of error is described by

$$\frac{\partial^2 \varepsilon}{\partial n^2} = \frac{1}{2\varepsilon n^2(N-1)} - \frac{\partial \varepsilon}{\partial n} \left( \frac{2}{n} + \frac{1}{\varepsilon} \frac{\partial \varepsilon}{\partial n} \right) \quad (4)$$

The quantity in (4) is equal to  $0.75zn^{-5/2}[p(1-p)]^{1/2} > 0$  for  $N \rightarrow \infty$ , confirming that an additional sample observation results in a lower margin of error at a diminishing rate.

<sup>20</sup> 984 F.Supp.2d 1035.

Choosing a confidence level of 95%, assuming the population proportion is 50%, and letting  $N \rightarrow \infty$ , as described in Section 2, the margin of error can be more compactly expressed as  $\varepsilon \approx 1/\sqrt{n}$ . This approximation more clearly shows that the sample size and the margin of error move in opposite directions and the slope flattens out with each additional observation, both shown in Figure 1.<sup>21</sup>

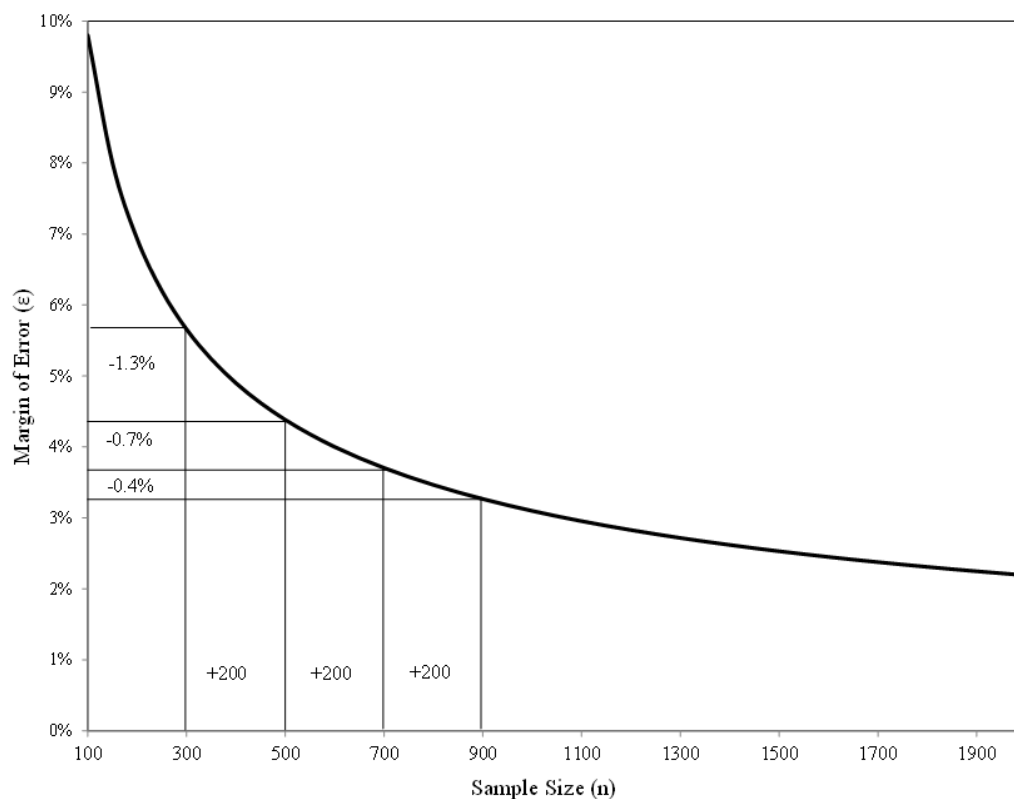


Figure 1. Sample size and margin of error ( $p=0.50$ ,  $z=1.96$ ,  $N \rightarrow \infty$ )

As the figure demonstrates, if the sample size were increased from 300 to 500, the margin of error would decline from 5.7% to 4.4% (a decline of 1.3 percentage points). If the sample size were again increased by 200 (from 500 to 700), the corresponding decline in the margin of error would be 0.7 percentage points, and if sample size were increased by another 200, the corresponding improvement in the margin of error would be 0.4 percentage points. Therefore, while increasing the sample size reduces the margin of error, the improvement diminishes with each successive increase in sample size.

To introduce a concept of optimality in the context sample size choice in a legal proceeding requires modeling the costs and benefits of a larger sample size. Let the total cost associated with conducting and analyzing a sample be comprised of a fixed cost that does not vary with the sample size and a variable cost component that does vary with the sample size. The fixed cost can be subsumed in the non-sampling costs of litigation. Assume that variable cost varies in proportion to the sample size, given by  $cn$ . This expression is without curvature and thus cannot be minimized except for the trivial solution  $n = 0$ . The model described here introduces nonlinearity to the system through the monetary reward expected by the party conducting the sampling analysis, which provides a nontrivial solution to the problem.<sup>22</sup>

Consider a monetary reward ( $R$ ) for winning the legal case and a cost ( $C$ ) of litigating the case, which does not include the sampling variable costs. Furthermore, as described above, there is an incremental cost ( $c$ ) to conduct the sampling analysis that varies directly with the sample size. Assume that the court's primary concern is whether the sample size is large enough to attain a sufficiently small margin of error to conclude that the statistical evidence is probative, and that accepting or rejecting evidence based on the sample will determine whether the court rules in favor or against the party proffering the sampling analysis.

<sup>21</sup> Numerical calculations in the figure are based on the exact relationship  $\varepsilon = 0.98/\sqrt{n}$ .

<sup>22</sup> To find the sample size that minimizes total cost in the case of taxpayer surveys that are subject to risk adjustment, Kadane (2017) introduces curvature to total cost through a risk-adjustment term equal to a function of  $1/\sqrt{n}$ .

The expected payoff of the legal case to the party conducting the sampling analysis is

$$E(\Pi) = \pi R - C - cn. \quad (5)$$

To obtain the party's expected payoff, the monetary reward from winning the case is multiplied by the probability ( $\pi$ ) that the court accepts the evidence based on the sample, delivering a favorable verdict to the party. The probability-adjusted reward,  $\pi R$ , does not come without cost, as there are non-sampling costs of litigation, including legal fees, expert fees, and other costs associated with litigating a case (including fixed sampling costs), as well as the incremental cost per study of an additional observation in the sample multiplied by the total number of observations in the sample, or  $cn$ . The expected payoff assumes that the party is already engaged in litigation, such that the non-sampling costs of litigation are not affected by the probability of the court accepting the sample (and thus a victory for the party). That is, the costs of litigation are incurred regardless of the outcome of the case; only the monetary reward is determined by the outcome of the case.

The definition of optimality here is one in which a given sample size maximizes the expected payoff defined in (5). Relatively weak assumptions are required. The derivative of the expected payoff with respect to  $n$  must exist and be equal to zero at only one point on the domain of the expected payoff function. And the change in the slope of the expected payoff in (5) need be negative at the solution point. This ensures concavity, and is expected from the law of diminishing marginal returns described earlier.

The first-order condition of the expected payoff function is provided in (6). Under the preceding assumptions, the sample size that maximizes the expected payoff must satisfy the economic optimality condition that marginal benefit equal marginal cost. Here, the optimal sample size,  $n^*$ , obtains where the change in benefit from a unit change in sample size equals the change in cost from a unit change in sample size, given by

$$\frac{\partial \pi}{\partial n} R = c. \quad (6)$$

Put differently, the sample size that will result in the largest expected payoff will be at the point where the additional increase in probability of the court accepting the sample from adding one more observation to the sample size (the marginal probability), multiplied by the reward from winning, is just equal to the marginal cost of increasing the sample by one observation.

### 3.2 The Court's Probability Function

To explore a hypothetical court's probability function,  $\pi$ , assume that the court understands the negative relationship between sample size and margin of error. For tractability of the model, further assume that its only concern is that the party presenting statistical evidence based on the sample has attained a sufficiently small margin of error, and this is the determining factor whether it wins or loses the case.<sup>23</sup> The court thus has the discrete choice to either accept the evidence based on the sample, or not—that is, it must conclude whether a party is guilty or not, liable or not, or at fault or not, based on the statistical evidence presented.

Let  $U$  denote the court's utility of accepting the sample relative to not accepting it, given by  $U_i = a + b\varepsilon_i + \omega_i$ , where  $a$  and  $b$  are parameters and  $\omega$  is a continuously distributed stochastic random variable that is independent of  $\varepsilon$ .  $U$  is unobserved to the parties, but it is known that the court will accept the sample if  $U$  is greater than some threshold, such as  $U > 0$  (without loss of generality). Assuming  $\omega$  has a logistic distribution,<sup>24</sup> the likelihood of the court accepting evidence based on the sample is (omitting subscripts)

$$\pi = \text{prob}(U > 0 | \varepsilon) = \frac{\exp(a+b\varepsilon)}{1+\exp(a+b\varepsilon)}. \quad (7)$$

It is necessary to constrain  $b < 0$  to reflect the court's disutility for larger margins of error; that is, a smaller sample with a larger margin of error will result in a lower probability of the court accepting statistical evidence based on the sample.

The marginal probability from (7) is given by

<sup>23</sup> In practice, this is unlikely to be a supportable assumption, given that a number of factors would play a role in whether a party wins or loses a case. However, for the purposes of this model, one can assume that all other factors have been determined and the only factor left is the sample size.

<sup>24</sup> See Casella and Berger (2002).



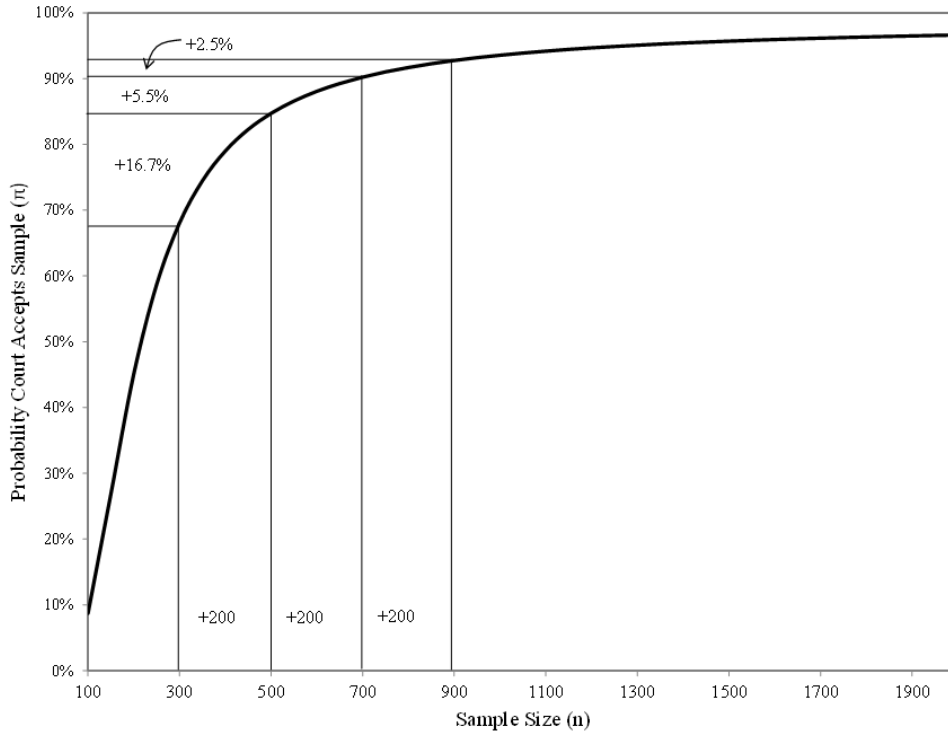


Figure 2. Probability court accepts statistical evidence ( $p=0.50$ ,  $z=1.96$ ,  $N \rightarrow \infty$ )

$$\frac{\partial \pi}{\partial n} = b\pi(1 - \pi) \frac{\partial \varepsilon}{\partial n}, \quad (8)$$

where  $\partial \varepsilon / \partial n$  is defined in (3).<sup>25</sup> Combining (2), (3), (6), (7), and (8) provides the optimal sample size that maximizes the expected payoff in (5), given values of  $a$ ,  $b$ ,  $c$ ,  $C$ ,  $R$ ,  $z$ ,  $p$ , and  $N$ .<sup>26</sup>

To simplify the analysis, let the population proportion and confidence level be 50% and 95%, respectively, and  $N \rightarrow \infty$ , as described in detail in Section 2. The probability of the court accepting evidence based on the sample simplifies to

$$\pi \approx \frac{\exp(a+b/\sqrt{n})}{1+\exp(a+b/\sqrt{n})}. \quad (9)$$

Figure 2 demonstrates the relationship is a positive one—that is, the larger the sample size, the higher the likelihood the court will accept evidenced based on the sample—for supposed hypothetical values  $a=5$  and  $b=-75$ .<sup>27</sup> For a sample size of 300, there is a 68% chance the court will accept evidence based on the 300 observations; for a sample size of 700, there is a 90% chance the court will accept evidence based on the 700 observations, and so on.<sup>28</sup> The increase in probability of the court accepting statistical evidence based on the sample from an increase in the sample size diminishes as the sample size increases.

Given the simplifying parameter values for the confidence level and population proportion, the marginal probability in (8) becomes

$$\frac{\partial \pi}{\partial n} \approx -\frac{1}{2}b\pi(1 - \pi)n^{-3/2} \quad (10)$$

for  $N \rightarrow \infty$ , which is globally positive for  $b < 0$  such that the probability of the court accepting the evidence based on

<sup>25</sup> The second derivative is  $\frac{\partial^2 \pi}{\partial n^2} = b\pi(1 - \pi) \left[ (1 - 2\pi) \left( \frac{\partial \varepsilon}{\partial n} \right)^2 + \frac{\partial^2 \varepsilon}{\partial n^2} \right]$ .

<sup>26</sup> The functional form of the logistic probability model does not yield an analytical solution for  $n$ , which instead obtains from numerical optimization techniques given parameter values.

<sup>27</sup> These hypothetical values were obtained by calibrating the model to obtain sensible estimates for expositional purposes. As noted earlier, the key assumption is  $b < 0$ , capturing the court's disutility for larger margins of error.

<sup>28</sup> Numerical calculations are based on (7) and the exact relationship  $\varepsilon = 0.98/\sqrt{n}$ .

the sample increases as the sample size increases.<sup>29</sup> Solving (6) for  $n$  provides the optimal sample size that maximizes the expected payoff, given values for the monetary reward, legal costs, sampling costs, and parameter values of the court's logistic probability model. Figure 3 provides the relationship between the expected payoff and the sample size, supposing hypothetical values of the reward ( $R = \$5$  million), non-sampling litigation costs ( $C = \$1$  million), marginal cost of increasing the sample size ( $c = \$500$ ), and the logistic probability model described above with  $a$  equal to 5 and  $b$  equal to -75. The expected payoff in Figure 3 initially increases as the sample size increases, capturing that the court is more likely to accept the sample and rule in favor of the party conducting the sampling analysis as the sample size gets larger. A maximum expected payoff is reached at \$3.2 million where the optimal sample size is  $n^* = 869$ , and, subsequently, the expected payoff begins to decline.

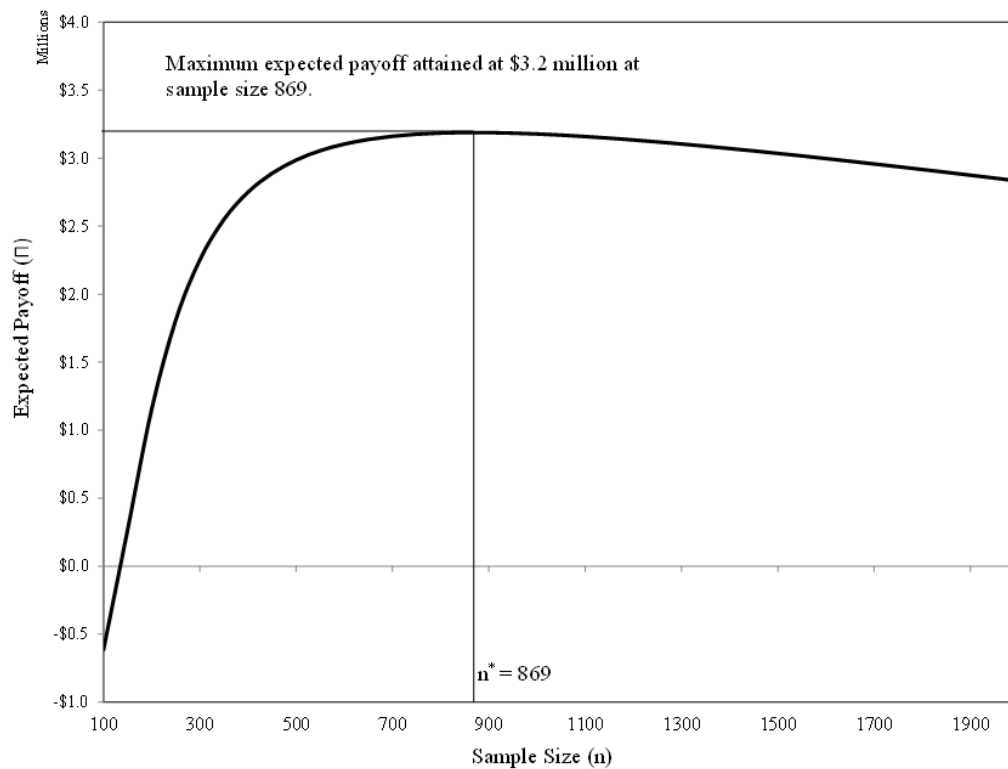


Figure 3. Expected payoff ( $p=0.50$ ,  $z=1.96$ ,  $N \rightarrow \infty$ ,  $R=\$5M$ ,  $C=\$1M$ ,  $c=\$500$ ,  $a=5$ ,  $b=-75$ )

At a sample size of 869, the marginal benefit of increasing the sample size equals the marginal cost of doing so. Thus, 869 observations yields the maximum expected payoff in this realization of the model. When the sample size is smaller than 869, the marginal benefit outweighs the marginal cost, so it pays to increase the sample size, and to keep doing so until 869 is reached. When the sample size is larger than 869, the marginal cost outweighs the marginal benefit, so it pays to reduce the sample size until the optimum is reached. The margin of error at the optimal sample size in this example is 3.3%. The probability the court accepts the sample size and concludes the statistical evidence is probative is 92.5%.

The practical relevance of this economic model is to provide a theoretical basis for how a court might potentially view statistical evidence based on a given sample size. The analysis models the behavioral decision-making a party to litigation would effectively engage in when faced with having to select a sample size, generally considering whether a larger sample size will sufficiently increase the probability of the court accepting evidence based on the sample versus the costs of studying the larger sample. Ultimately, the sample size that is selected is the solution to a problem that depends on a multitude of factors, including the court's likelihood of accepting the sample, the rewards and costs of litigation, and the costs of studying the sample.

Sample size determination is an economic question that requires an intuitive balancing act of costs and benefits. A

<sup>29</sup> The concavity of the logistic probability model depends on the sample size  $n$ . The rate of change of the slope is given by  $\frac{\partial^2 \pi}{\partial n^2} = \frac{\partial \pi}{\partial n} \left[ \frac{\partial \pi}{\partial n} \frac{1}{\pi} + \frac{b\pi}{2n^{3/2}} - \frac{3}{2n} \right]$ , which is negative for sufficiently large  $n$  and positive for sufficiently small  $n$ , given  $b < 0$ .

sample that is too small will result in too low a likelihood of the court accepting the statistical evidence as probative, but a sample that is too large is financially too costly to make economic sense given the potential payoff from winning the court case. The balance is found where the increase in the probability of the court accepting the statistical evidence from a unit increase in sample size, multiplied by the monetary reward from winning the case, is just equal to the marginal cost of increasing the sample size by a corresponding unit.

#### 4. Conclusions

Constraints on time, budget, and other resources mean that not every element in a population can be investigated to inform questions of causation, liability, and damages in legal proceedings. Thus, relying on a sample as a surrogate for the population is often an efficient means to proceed with statistical analyses. Random sampling procedures help attain samples that are representative of the population, but the key question arises of how large a sample to draw and study. Relying on a sample that is too small may result in an insufficient level of precision; yet a sample that is too large may be an inefficient use of finite resources.

A party to litigation or a regulatory matter might first determine what is desired in terms of precision and confidence, and then determine the smallest sample size that will achieve it. The tables presented in Section 2 provide some guidance. For example, if it is necessary to claim with 95% confidence that the sample estimate will be no more than three percentage points away from the true population parameter (assumed to be 50% in the example), a sample size of just over 1,000 is necessary for very large populations. In general, higher degrees of precision require larger samples.

If there is sufficient flexibility to increase resources devoted to analyzing a larger sample, the margin of error can be reduced. Thus, a decision must be made whether the additional precision afforded by a larger sample is worth the cost and resources that must be devoted to sampling and studying the additional observations. Answers to these questions will depend on the circumstances of the case at hand. As described in Section 2, there is no universal approach that will apply to all cases, nor will legal precedent provide a satisfactory answer. Sampling theory must be relied on to inform sampling procedures and sample size determination.

The economic model developed in Section 3 provides insight into the mental calculus theoretically undertaken by an economically rational party in making decisions about the sample size in a legal proceeding. The optimal sample size depends on a multitude of factors, including the court's likelihood of accepting the sample in litigation, the rewards and costs of litigation, and the cost of analyzing the sample. Relying on a logistic probability function for modeling the court's likelihood of accepting the sample confirms intuition—namely that too small a sample is unsatisfactory because it may result in a low likelihood of the court's acceptance, but too large a sample may be too costly to be economically feasible. Ultimately, the optimal sample is chosen such that these two countervailing effects balance.

#### References

- Boardman, A. E., & Vining, A. R. (1983). The role of probative statistics in employment discrimination cases. *Law and Contemporary Problems*, 46(4), 189-218. <https://doi.org/10.2307/1191599>
- Casella, G., & Berger, R. L. (2002). *Statistical Inference* (2nd ed.). USA: Duxbury.
- Caughey, R. E. (1956). The use of public polls, surveys and sampling as evidence in litigation, and particularly trademark and unfair competition cases. *California Law Review*, 44, 539-546. <https://doi.org/10.2307/3478753>
- Cheng, E. K. (2012). When 10 trials are better than 1000: an evidentiary perspective on trial sampling. *University of Pennsylvania Law Review*, 160, 955-965.
- Cochran, W. G. (1999). *Sampling Techniques* (3rd ed.). New York: John Wiley & Sons.
- Gastwirth, J. L. (2003). Issues arising in using samples as evidence in trademark cases. *Journal of Econometrics*, 113(1), 69-82. [https://doi.org/10.1016/S0304-4076\(02\)00167-7](https://doi.org/10.1016/S0304-4076(02)00167-7)
- Gastwirth, J. L., & Xu, W. (2014). Statistical tools for evaluating the adequacy of the size of a sample on which statistical evidence is based. *Law, Probability and Risk*, 13(3-4), 277-306. <https://doi.org/10.1093/lpr/mgu010>
- Kadane, J. B. (2017). Optimal sample size for risk-compensated survey assessment. *Law, Probability and Risk*, 16(4), 151-162. <https://doi.org/10.1093/lpr/mgx013>
- Kaye, D. H., & Freedman, D. A. (2011). Reference Guide on Statistics, in Federal Judicial Center, *Reference Manual on Scientific Evidence* (3rd ed.), 211-302.
- Lahav, A. (2008). Bellwether trials. *George Washington Law Review*, 76(3), 576-638.
- Larsen, R. J., & Marx, M. L. (2001). *An Introduction to Mathematical Statistics and Its Applications* (3rd ed.). New Jersey: Prentice Hall.

Lohr, S. L. (2010). *Sampling Design and Analysis* (2nd ed.). Boston: Brooks/Cole.

Mathews, P. (2010). *Sample Size Calculations: Practical Methods for Engineers and Scientists*. Ohio: Mathews Malnar and Bailey.

Stock, J. H., & Watson, M. W. (2007). *Introduction to Econometrics* (2nd ed.). Boston: Pearson.

### **Copyrights**

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).