**Working Paper Series**

**Núcleo de Normalização e Qualimetria (NNQ),**

 **Federal University of Santa Catarina, Brazil**

September 2015

WP15-002

AN ACCEPTANCE SAMPLING APPROACH TO POLLING PUBLIC OPINION

Robert Wayne Samohyl

samohyl@yahoo.com

ABSTRACT

Since the pioneering work of Dodge and Romig in the 1920’s, the risk of type I and type II error has not been used consistently for the construction of sampling plans for producers and consumers, or as proposed in this paper analogously for pollsters and candidates. We offer these new concepts as an important factor for determining sample size and risk levels. The paper evaluates and extends the area of acceptance sampling in two ways: first, by using the hypergeometric distribution to calculate the parameters of sampling plans avoiding unnecessary use of approximations such as the binomial or poisson. The conclusion is that the hypergeometric distribution should have priority in the area of acceptance sampling. Second, we elaborate acceptance sampling in terms of hypothesis testing rigorously following the original concepts of Neyman and Pearson (NP). By offering a common theoretical structure, hypothesis testing from NP can produce a better understanding of applications in areas other than industry and commerce such as public health and public opinion polling, the subject of this paper. We show that sample size can be greatly reduced and at the same time confidence increased. Compared to confidence intervals, acceptance sampling requires smaller samples and therefore a lower expenditure to produce more accurate results at less risk. The confidence interval methodology is therefore prejudicial when applied to the analysis of public opinion since it weights sampling errors equally, and forces the assumption that over and under estimation represent equal levels of risk. The construction of receiver operating characteristics (ROC) curves for this case can secure a less conflicting result for the pollster and candidate.

KEY WORDS: Acceptance sampling, hypergeometric, OCC, ROC, hypothesis test, polling

JEL: C83

[https://drive.google.](https://drive.google.com/file/d/0Bzz0QCLFzzTQZ1dHZW1qWHJlRmc/view?usp=sharing)

The views expressed in the NNQ Working Paper Series are those of the author(s) and do not necessarily reflect those of the Federal University of Santa Catarina. Working Papers have not undergone formal review and approval. Papers are included in this series to elicit feedback and to encourage debate. Copyright belongs to the author(s). **Cite Working Papers only with author approval**.

AN ACCEPTANCE SAMPLING APPROACH TO POLLING PUBLIC OPINION

Robert Wayne Samohyl[[1]](#footnote-1)

ABSTRACT

Since the pioneering work of Dodge and Romig in the 1920’s, the risk of type I and type II error has not been used consistently for the construction of sampling plans for producers and consumers, or as proposed in this paper analogously for pollsters and candidates. We offer these new concepts as an important factor for determining sample size and risk levels. The paper evaluates and extends the area of acceptance sampling in two ways: first, by using the hypergeometric distribution to calculate the parameters of sampling plans avoiding unnecessary use of approximations such as the binomial or poisson. The conclusion is that the hypergeometric distribution should have priority in the area of acceptance sampling. Second, we elaborate acceptance sampling in terms of hypothesis testing rigorously following the original concepts of Neyman and Pearson (NP). By offering a common theoretical structure, hypothesis testing from NP can produce a better understanding of applications in areas other than industry and commerce such as public health and public opinion polling, the subject of this paper. We show that sample size can be greatly reduced and at the same time confidence increased. Compared to confidence intervals, acceptance sampling requires smaller samples and therefore a lower expenditure to produce more accurate results at less risk. The confidence interval methodology is therefore prejudicial when applied to the analysis of public opinion since it weights sampling errors equally, and forces the assumption that over and under estimation represent equal levels of risk. The construction of receiver operating characteristics (ROC) curves for this case can secure a less conflicting result for the pollster and candidate.

KEYWORDS: Acceptance sampling, hypergeometric, OCC, ROC, hypothesis test, polling

JEL: C83

INTRODUCTION

Compared to confidence intervals, acceptance sampling requires smaller samples and therefore a lower expenditure to produce more accurate results at less risk. However there is a learning barrier to the approach of acceptance sampling, not as intuitively understandable as the confidence interval.

Acceptance sampling is a method for choosing sample size (n) and cutoff values for the number of bad units (Ac) in the sample that differentiate between good and bad lots from the perspective of industrial buyers/consumers and sellers/producers given the size of the lot (N). It is common to use the expression **PL(N, n, Ac) to represent the sampling plan**. The method was introduced by Dodge and Romig in the 1920’s as part of the first quality control program organized and implanted around concepts of probability and statistics.[[2]](#footnote-2) We argue that the basic concepts of **acceptance sampling can be applied to the polling of public opinion and produce better results than the traditional confidence interval procedures**.

The conceptual framework traditionally adopted for sample size determination for polling and similar kinds of market research applies the binomial distribution[[3]](#footnote-3) to estimate confidence intervals and simultaneously sample size. First, the tolerable value of the margin of error is defined, and then sample size is calculated usually through the approximation of the standard deviation to the normal distribution. The binomial assumes that population size is infinite. The use of the hypergeometric distribution is the basis of our approach here due to its more general nature which allows for the population to be finite, even though the binomial distribution in the presence of large populations (say 100,000) produces essentially the same results. In the case of political polling, for instance in primary elections in small states or for specific sectors, populations may be much smaller, even less than 5000. For small populations the hypergeometric distribution is required. We briefly outline the confidence interval approach:

To perceive the benefits of acceptance sampling for polling surveys, we need to compare it to the traditional sampling methods based on confidence intervals and the binomial distribution where $\hat{p}$$\hat{p}$ is the estimated percent of adversary votes in the sample and *p* is the true value of adversary votes in the population. The 95% confidence level based on the standard normal distribution, z value of 1.96 (2.00 in many practical applications), is common in the polling industry to calculate the margin of error.

$P\left(\hat{p}-1.96\*\sqrt{\hat{p}\*\frac{1-\hat{p}}{n}} \leq p \leq \hat{p}-1.96\*\sqrt{\hat{p}\*\frac{1-\hat{p}}{n}}\right)=95\%$ (1)

Following the precepts of random sampling and the assumptions of the normal approximation of the binomial distribution, equation 1 determines a probability of 95% that unknown population parameter *p* is contained within the confidence limits. Likewise, there is a complementary probability of 5% that parameter *p* is outside the limits. In table 3, the results based on equation 1 (right hand side of the table labelled CONFIDENCE INTERVAL) show that a margin of error of 2% requires a sample size (n) of 2400 when sample estimate $\hat{p}$ is 48%, interval limits at 46% and 50% supporting the hypothesis that the opposition candidate is losing the election campaign. Furthermore, if it were necessary to diminish the size of the sample for reasons of cost and time, the margin of error would have to be readjusted upward. This becomes clear in table 3 where a sample size of 550 produces a margin of error of 4.2% and confidence limits of 39,4% and 50.3% when estimate $\hat{p}$ is 44.9%. Even though this seemingly satisfactory result should please our candidate, the upper confidence limit being greater than 50% demonstrates that our candidate may be losing the campaign. **Small sample size and the resulting large margin of error degrade the accuracy of practical results from confidence intervals even though estimates of opposition votes seem intuitively small.** Below we will show that a small sample size may still be an effective tool for evaluating public opinion if applied in the acceptance sampling method.

Equation 1 can be rewritten as equation 2, affording a different perspective to the confidence interval. The conceptual difference is that equation 2 clearly shows the symmetry of the probabilities of error. The 5% probability of the parameter *p* being outside of the confidence limits is the sum of the 2.5% probability of being either above or below the interval limits, sometimes called tail probabilities.

$P\left(\hat{p}-1.96\*\sqrt{\hat{p}\*\frac{1-\hat{p}}{n}} >p\right)=2.5\% P\left(\hat{p}+1.96\*\sqrt{\hat{p}\*\frac{1-\hat{p}}{n}} <p\right)=2.5\% $(2)

Clearly, there is a strategic difference for pollsters and candidates between the upper and the lower tail probabilities. When adversary support in the voter population is actually larger than the upper confidence limit indicating strong support for the adversary, the candidate may come to the erroneous and disastrous conclusion from sampling error or simply ill applied sampling procedures that the adversary is losing the campaign ($\hat{p}<0.5 $) when in fact it is the candidate himself who is losing. This error in judgement is much worse than for the candidate who thinks he is losing when in fact he is winning. **One of the advantages of acceptance sampling is that tail probabilities may be weighted differently to reflect the differential impacts of over and under-estimation error.**

A BRIEF DESCRIPTION OF THE DODGE AND ROMIG ACCEPTANCE SAMPLING PARADIGM[[4]](#footnote-4)

Immediately after the production of a large lot of homogeneous items, but before expediting the lot to the consumer, an inspection should occur to verify its quality level (p), the percentage of nonconforming items in the lot. Usually the entire lot is not inspected due to constraints on time and resources but a much smaller subset of the lot is. The interest of the producer is to expedite lots with satisfactory quality to insure customer satisfaction. Later on we will draw an analogy between producer and pollster, and likewise between consumer and candidate. Analogously, the lot of homogeneous items will be the population (N) of voters and the quality level will be the percent (p) of opposition votes in the population. When inspection by the producer results in rejected lots, the common practice is to apply universal 100% inspection to the rejected lot still in the factory replacing all bad parts. Of course the part of the Dodge-Romig sampling procedure that considers the replacement of defective items in rejected lots does not apply to the polling of public opinion. The losing pollster may well desire the replacement of opposition voters but hopefully democratic principles will prevail and this alternative will not be available. The power of the candidate over the pollster is an important aspect of acceptance sampling that is ignored in the confidence interval method.

Once the probability distribution function is defined (in Dodge and Romig the Poisson and binomial distributions are used due to their computational ease), risk can be calculated. Here we introduce the use of the hypergeometric distribution, which differs from the binomial and the Poisson due to the inclusion of population size N. Define the number of defective parts in the sample as c and the number of defective parts in the population as *C = pN*. [[5]](#footnote-5) The probability of obtaining c defective items in a sample of size n is,

 (3)

or emphasizing the quality parameter p as

 (4)

For acceptance sampling, the probability of declaring the lot conforming should be calculated as a sum involving c = 0, 1, 2…Ac, in the framework of the cumulative hypergeometric distribution[[6]](#footnote-6). Specifically, in a sample of size n, from a population of size N, a lot is judged as acceptable even though the sample may contain up to Ac defective items. **The probability *PH(0,1,2,...Ac),* or simply *PH(Ac),* of lot acceptance** depends upon the parameters of the sampling plan PL(N, n, Ac) and the quality of the lot p. Consequently, with H representing the cumulative hypergeometric distribution:

*PH(Ac) = H{ PL(N, n, Ac), p }*  (5)

Later we will use a graphical representation of equation 5 called the operations characteristic curve (OCC) to clarify sampling risk for the candidate and for the pollster.

From the producer point of view, the major worry is the rejection of good lots. This will occur when the number of defective items (c)[[7]](#footnote-7) in the random sample is greater than a predefined cutoff level (Ac). **Producer risk** is defined as the probability of error **P(c > Ac / good lot).** We look for values for sample size (n) and the cutoff (Ac) that makes producer risk reasonably small. From the health sciences literature (RHODA, et al (2010), PEZZOLI, et al (2013), PAGANO, M. AND VALADEZ, J. J. (2010).), rejecting what is satisfactory, erroneously characterizing as a flaw or an illness something that is flawless and healthy, is called a **false positive (FP).** A medical test whose result is **positive** means that something bad has been encountered in the patient or the sample. Extending these concepts to the pollster/candidate environment, a positive result indicates that there is something wrong with the campaign; the adversary is winning. There is nothing worse and more frustrating for the producer/pollster than to see his efforts at producing high quality output (a winning campaign) being disappointed by a sampling plan that rejects his successful efforts because of sampling error, calling the producer/campaign manager to action where action is not necessary. Even though the error of erroneously accepting bad product, called a **false negative (FN)**, is less worrisome for the producer and considered a secondary risk, it will be used as one of several ingredients to calculate sampling plans in a more coherent way. Since the pioneering work of Dodge and Romig, secondary risk has not been used for the construction of sampling plans. However, we offer this new concept as an important factor for determining sample size in the examples that follow. As we will see below the producer/pollster analogy will require further elaboration.

In industry, the quality of a lot (p) is never known with certainty; all that is known with certainty is the characteristics of the sample. In a very few modern industries, items are tracked intensely along the production line, and then on to the final consumer for the entire life of the final product. In this case quality of an originating lot may be known one day in the future, but only very rarely is this true. For the pollster and the candidate however, the quality of a lot (p) is eventually known with certainty because elections are held and winners and losers defined. But even in the case of unknown lot quality in industry, this ignorance about lot quality is not a disadvantage for constructing a sampling plan. **Producers require a desired delimiting value for p that separates defective from conforming lots. Dodge and Romig called this the acceptable quality level (AQLp)[[8]](#footnote-8)**. Today´s industrial standards (INTERNATIONAL STANDARD ISO 2859-1 (1999)) usually require an AQLp of less than 1%. A pollster on the other hand might choose an AQLp of 40% or 45%, a conservative cutoff value indicating with security that the adversary´s campaign is losing. See figure 1 for an illustration of these suggested values. Sampling plans that continually reject either lots or campaigns are indicating poor quality. Producer risk can be defined quantitatively as **P(c > Ac / AQLp ≥ p),** the probability of the sample wrongly demonstrating rejection, given that the lot is conforming or the campaign successful**.** A producer or pollster will usually not tolerate risk of more than 5% or in some cases even 1% may be too much.

Finally, producer risk, based on the probability of rejection of the lot *1 - PH(Ac)* can be defined after substituting p by AQLp:

***producer risk*** *= 1 - PH(Ac) =1 - H{ PL(N, n, Ac), AQLp }*  (6)

In figure 1 AQLp is illustrated as a limiting value for lot quality defined as percent defective allowed by the producer (pollster).

Figure 1 AQL and LTPD in industry and polling

**Consumer risk** on the other hand is defined as the probability of accepting bad lots, **P(c ≤ Ac / bad lot).** From the health sciences literature, accepting the unacceptable is called a **false negative (FN).** There is nothing more frustrating for the consumer/candidate than to acquire inadequate product, suffering disappointment from a sampling plan that has allowed poor quality to slip into the production line or the retail shelves. Note that false positives (FP) on the other hand are less important to the consumer since rejected but conforming lots can be reinspected by the producer and any remaining bad items replaced. The appearance of false negatives however is immediately prejudicial and incorrigible in the short term. Even though false positives do represent a secondary risk for the consumer, they were not included in the original work of Dodge and Romig. The inclusion of secondary risk in sampling plans is the object of this paper. Just as with producer risk, the traditional procedure requires that consumer risk be minimized finding appropriate values for sample size (n) and the cutoff (Ac). Just as the producer defines AQLp, the consumer must define a cutoff to distinguish between good and bad lots. Lot tolerance percent defective (LTPDc)[[9]](#footnote-9) is a maximum value of p the consumer will tolerate as permissible and still define the lot as conforming. It seems reasonable to say that LTPDc from the point of view of the candidate should be 50%. Consequently, consumer risk is defined more specifically as **P(c ≤ Ac / LTPDc ≤ p),** and we can write[[10]](#footnote-10),

***consumer risk*** *= PH(Ac) =H{ PL(N, n, Ac), LTPDc }*  (7)

In traditional applications for industry, consumer risk is set at 10% and producer risk at 5% and, using the binomial distribution, n and Ac are solved for simultaneously to arrive at a sampling plan PL(n, Ac) compatible for both the producer and the consumer. **The question arises why the consumer should tolerate more risk than the producer, especially considering that even Dodge and Romer suggested that the consumer should be placed in a priority position.**

(DODGE and ROMIG, 1944) ‘**The first requirement for the method will therefore be in the form of a definite assurance against passing any unsatisfactory lot** that is submitted for inspection. [...] For the first requirement, there must be specified at the outset a value for the lot tolerance percent defective (LTPD) as well as a limit to the probability of accepting any submitted lot of unsatisfactory quality. The latter has, for convenience, been termed the Consumer’s Risk ...’ (emphasis added).

The priority given to the consumer will be an important ingredient for the discussion of hypothesis testing in the next section.

The value chosen for Ac is of pivotal importance, recognizing that the requirements of the producer (pollster) and consumer (candidate) come from different perspectives. The consumer will desire a low value for Ac, consequently committing the error of rejecting some good lots but better guaranteeing the acceptance of only good lots and likewise the rejection of bad lots. **Small Ac implies diminished consumer risk**, as illustrated in figure 2. On the other hand, the producer will want an Ac that is relatively large, admitting the possibility of accepting some bad lots so that good lots will not be rejected. **Large Ac guarantees diminished values for producer risk**. This is illustrated in figure 2 where producer risk declines as the cutoff Ac increases. Therefore, it will be difficult to find values for Ac that will serve the desires of both producer and consumer. In figure 2, we show a sampling plan that has the same sampling parameters for both producer and consumer PL(1000,10,4). The values of pollster and candidate risks are the same for this plan; however, at approximately 0.37 it is too large to be suitable for the two parties. The question is how do we find a plan that is the same for consumer and producer but contains a much smaller risk factor? And how can sampling plans take into account secondary risk? Such a plan if indeed it does exist could be agreed upon by both parties and conveniently applied only once somewhere between the factory and the store.



Figure 2 Producer/pollster risk and consumer/candidate risk for sampling plan PL(1000,10,Ac), AQLp = 0.4, LTPDc = 0.5

In the next section, we will illustrate the benefits of hypothesis testing in line with the original approach of Neyman and Pearson for developing supporting theoretical concepts in acceptance sampling, and to develop sampling plans that show explicitly the requirements and conflicts inherent in consumer and producer negotiations.

THE NEYMAN-PEARSON PARADIGM FOR HYPOTHESIS TESTING[[11]](#footnote-11)

Historically, the work of Dodge and Romig (1929) appeared in industrial applications and academic literature before the concepts of hypothesis testing (and confidence intervals) received widespread approval in practice. Their presentation depends exclusively on probability functions, and the interpretation of the concepts of producer risk and consumer risk as the probability of error, some years before Neyman and Pearson (1933) offered their seminal interpretation of type I and type II error. Our review of hypothesis testing is at most a simple sketch of this area of scientific methodology, which is better elaborated in works like Rice (chapter 9, 1995), Lehmann (1993), and the original work of Neyman and Pearson. Nevertheless, our interpretation of acceptance sampling in light of hypothesis testing is new to the literature. Here we will concentrate on the nature and definition of the null hypothesis (Ho).

Simply stated, a hypothesis is a clear statement about an objective value of a characteristic or relation among characteristics (something happens associated with something else) that may or may not be true. Truth can be supported by a test. It carries a doubt that calls for evaluation. Hypotheses are not unique but come in pairs, a dual set, of exclusive statements in the sense that if one statement is true then the other statement is false. When the decision maker judges one of the hypotheses as true, then the other hypothesis is necessarily judged as false. The lot is conforming or nonconforming. Children are vaccinated or they are not. The adversary is winning the election campaign or is losing. The accused is either innocent or guilty.

Two essential concepts in hypothesis testing are type I (erroneous rejection) and type II (erroneous acceptance) error. As is clear in the previous section, acceptance or rejection is applied to a statement that might be either true or false. That statement is called the **null hypothesis (Ho)**. There are three basic notions involved. First, a product or service in the form of a very large batch, lot or population that is to be judged relative to the null hypothesis as acceptable or not. Second, a decision-making unit supplies this item or service (factory, government health services, and election consultants/pollsters). Third, the other decision-making unit (buyer, retail outlets, a general population under government supervision, candidates) receives the item or service. The hypothesis test attempts to classify the population by accepting or rejecting the null usually by examining a small random sample drawn from the population under study.

In table 1, all of this information is organized in a contingency table, often called a confusion matrix in the area of data mining (PROVOST and FAWCETT, 2013). The decision maker indicates states of the null by examining a small sample of the population and consequently accepting or rejecting the null hypothesis. For the purpose of this article, the choice is made by analyzing the probabilities associated with the outcomes of a random sample from the relevant population, but of course, other methods might be tried like flipping a coin or throwing shells in a basket. In the population itself, the null is, in reality, either true or false. **For the majority of applications, the true state of the null will never be known. For the candidate/pollster, the truth or falsehood of the null will be known when the result of the election is revealed.** As shown in the contingency table, the result of the hypothesis test can have one of four possible results.

|  |  |
| --- | --- |
|  | Decision maker chooses between states of the null hypothesis Ho |
| Real states of the null hypothesis Ho in the population | accept Ho | reject Ho |  |
|  | true Ho | correct | type I error | sum of prob = 1 |
| false Ho | type II error | correct | sum of prob = 1 |

Table 1 Contingency table for hypothesis tests

Two quadrants are labelled as correct, and the other two are errors. In general terms, we would like to maximize the probability of falling into the correct boxes and minimize the probability of error. We can even prioritize between and within the correct boxes and the error boxes. Following Neyman and Pearson, declaring the truth of the false null is a type II error, whereas rejecting a true null is type I error. **The definition of the null hypothesis is crucial, and should be defined as the position that has the highest cost if rejected in error.** **The choice of which of the two hypotheses is to be the null depends therefore on the perspective of the decision maker, whether producer/pollster or consumer/candidate.**

From the viewpoint of the decision maker, the consequences of incorrectly rejecting one of the hypotheses are more severe than those of incorrectly rejecting the other. T**he hypothesis that is wrongly judged and whose cost of error is greater should be chosen as the null hypothesis, (RICE, 1995).** As we have seen above, lots are either conforming or nonconforming, and for the consumer, incorrectly accepting the nonconforming lot assigning the false negative is a disaster. The null carries the symbol **Ho**, the alternative hypothesis **Ha**. From the consumer point of view therefore, the null hypothesis is that the lot is nonconforming. Rejecting this null when it is true has extremely high costs for the buyer. In similar fashion but from the producer point of view, the null hypothesis is that the lot is conforming, because rejecting this null has extremely high costs for the seller/pollster. In the next section, we illustrate the appropriateness of hypothesis testing from the viewpoint of the pollster and candidate.

FROM THE CANDIDATE PERSPECTIVE

This example will begin with the candidate´s perspective. We will explain below that the null hypothesis for the candidate is that he is losing the election. While it is true that in industry, supplier and buyer may share much common ground for negotiations, this is not true for the pollster and candidate. Once the pollster and the candidate witness the election results, there is no turning back, producer error and consumer error cannot be corrected by testing another sample or using 100% inspection on rejected lots. In the case of false negatives that arise from campaign surveys, the unexpected surprise of an election loss is fatal to the candidate and consequently to the pollster. There is little or nothing the pollster can do to impose his parameters for minimizing risk against the power of the candidate.

Sampling in the context of election polling usually solicits a simple response to the simple question: will you, a potential voter, vote for our candidate or not. The solicitation is made carefully to avoid response bias. Using concepts from hypothesis testing as originally developed by Neyman and Pearson (1933), in table 2 we show the results from this sampling exercise from the **point of view of the candidate**. Our candidate in this case is the consumer, client, receiver of services offered by the pollster, who is analogous to the producer in Dodge and Romig. In table 2 there are two correct results and two errors. **Define the value of Ac as a limit on opponent votes (opvotes) in the sample that indicates the cutoff between a winning and losing campaign for the adversary corresponding to the bad part in the lot or the sickness in the population.** The proportion p is the percentage of opponent votes in the population of voters. Sample size n will be calculated using the probabilities of both type I and type II errors.

**The null hypothesis from the candidate´s skeptical perspective is that his adversary is winning the election, LTPDc < p.** LTPDc defines the boundary values of p that separate winning and losing campaigns. When p is greater than LTPDc = 50% the candidate is losing the election, the null is true. The sampling procedure would verify this result if the number of opponent votes (opvotes) in the sample were greater than the predetermined cutoff value, Ac. In the health science literature, this result is a True Positive (TP). There is a chance however that the sample may produce an erroneous result with opvotes ≤ Ac, even though the opponent is in fact winning the campaign (LTPDc < p), a **False Negative (FN)**. The probability of the occurrence of the false negative is consumer/candidate risk. This is a very serious type I error for the candidate who is led to believe that the adversary is losing the campaign when in fact he is winning. The relationship between candidate risk and the cutoff Ac has been illustrated in figure 2.

|  |  |
| --- | --- |
|  | Candidate chooses between states of the null hypothesis  **Ho: opponent is winning** |
| Reality of null Ho: opponent is winning | accept Ho: opvotes > Ac | reject Ho: opvotes ≤ Ac |  |
| true Ho 50%=LTPDc<p | TP correct P(opvotes>Ac/50%=LTPDc<p) | FN type I error: **αc**=P(opvotes≤Ac/50%=LTPDc<p) | sum of prob = 1 |
| false Ho 45%=AQLc ≥p | FP type II error: **βc** = P(opvotes>Ac/45%= AQLc ≥p) | TN correct P(opvotes≤Ac/45%= AQLc ≥p) | sum of prob = 1 |

Table 2 Contingency table for candidate:

TP true positive, FN false negative, FP false positive, TN true negative

To define the alternative hypothesis, that the adversary is losing the campaign, it is necessary to decide upon a value for p that indicates our candidate is in fact not losing with a comfortable margin. The candidate should be cautiously skeptical in this regard, and therefore should approach the definition of AQLc from a prudent viewpoint. Adequate values for AQLc could be 45% as used in the contingency table 2 or even 40% representing an extremely prudent position. The secondary type II error occurs when the null in the population is false but declared true, in other words, the adversary is actually losing the campaign but the sample has produced a relatively large number of adversary votes leading to the erroneous conclusion that the adversary is winning. This type II error (the false positive) is relatively less important for the candidate.

Table 3 represents numerical values for sampling plans from the candidate viewpoint. For given values of the probability of type I (αc ) and type II (βc) error, the table presents the corresponding sampling plan PL(N, n, Ac). The choice of population size of 300,000 is arbitrary but is large enough to be considered practically infinite and results will coincide with the binomial distribution. Generally, populations in elections reach up to millions of voters. There are cases, however, where polls sometimes deal with populations that are much smaller, for example, school districts and primary elections may have only several thousand voters, in the same way that some polls are targeted for smaller segments of the population such as age groups or religious affiliation. In these situations, the use of the hypergeometric will be necessary.

In table 3, type I error is fixed at extremely low levels to prioritize the risk aversion of the candidate. **As a rule, acceptance sampling in polling offers a significant increase in accuracy and confidence when compared to the standard confidence interval method, and with much smaller samples.** As an illustration, if the candidate is only worried about avoiding his own type I error, he could set his own candidate risk at 1% and the risk of type II error at 50%, which yields sample size of only 550 and Ac = 247. In this case, if the number of opvotes in the sample is 247 or less, Ho is rejected and we conclude that the adversary is losing. Furthermore, we can compare acceptance sampling and confidence intervals by configuring several polling situations and evaluating results. The first entry in table 3 is interesting by showing a sample size that is approximately the same as the traditional sample size 2400 used by pollsters based on 95% confidence intervals, therefore with a 5% chance of error. Note that the probability of error from acceptance sampling is 0.5% and 1% for the same sample size, a substantial fall in error rates.

The inconclusiveness of the confidence interval is apparent in table 3 since almost all of the confidence limits are represented by values above and below 0.5. The methodology does not allow the analyst to distinguish between values greater than and less than 0.5 since within the confidence interval all values are equally reliable. This means that if a given sample produces Ac + 1 opvotes, then acceptance sampling indicates that the opponent is winning but the confidence interval is almost always inconclusive in our examples.

|  |  |  |  |
| --- | --- | --- | --- |
|  | RISK |  | CONFIDENCE INTERVAL (binomial) |
|  | CANDIDATE | SECONDARY | SAMPLING PLAN (hyper) |  | LIMITS |
|  | αc = P(LTPDc) | βc= 1 - P(AQLc) | n | Ac | Ac + 1/n | LEVEL | $\hat{p}= A$$\hat{p}=$c + 1/n | LOWER | UPPER |
| 1 | 0.005 | 0.01 | 2389 | 1131 | 0.473 | 0.995 | 0.473 | 0.444 | 0.502 |
| 2 | 0.005 | 0.50 | 670 | 301 | 0.449 | 0.995 | 0.449 | 0.395 | 0.503 |
| 3 | 0.007 | 0.007 | 2393 | 1136 | 0.475 | 0.993 | 0.475 | 0.447 | 0.502 |
| 4 | 0.01 | 0.01 | 2151 | 1021 | 0.475 | 0.99 | 0.475 | 0.447 | 0.503 |
| 5 | 0.01 | 0.03 | 1765 | 833 | 0.472 | 0.99 | 0.472 | 0.441 | 0.503 |
| 6 | 0.01 | 0.05 | 1567 | 737 | 0.470 | 0.99 | 0.470 | 0.437 | 0.502 |
| 7 | 0.01 | 0.10 | 1299 | 607 | 0.467 | 0.99 | 0.467 | 0.431 | 0.503 |
| 8 | 0.01 | 0.20 | 1005 | 465 | 0.463 | 0.99 | 0.463 | 0.422 | 0.503 |
| 9 | 0.01 | 0.30 | 820 | 376 | 0.458 | 0.99 | 0.458 | 0.413 | 0.503 |
| 10 | 0.01 | 0.50 | 550 | 247 | 0.449 | 0.99 | 0.449 | 0.394 | 0.503 |
| 11 | 0.013 | 0.50 | 501 | 225 | 0.449 | 0.99 | 0.449 | 0.392 | 0.506 |
| 12 | 0.025 | 0.025 | 1534 | 728 | 0.475 | 0.95 | 0.475 | 0.450 | 0.500 |
| 13 | 0.0067 | 0.0072 | 2400 | 1139 | 0.475 | 0.95 | 0.475 | 0.455 | 0.495 |

Table 3 Candidate risk tradeoff, hypergeometric sampling plans for several values of αc, βc, LTPDc is 50%; AQLc is 45%, N=300,000, and comparable binomial confidence intervals

Calculations: R package Acceptance Sampling, Kiermeier (2008). R code in the appendix.

**The confidence interval methodology is therefore prejudicial when applied to campaign analysis since it weighs sampling errors equally.** First, there is the extra cost of larger samples. In all cases, the confidence interval method demands relatively large sample sizes relative to acceptance sampling. Secondly, the confidence of numerical results is poor. This happens because the confidence interval combines the two risks of candidate and pollster into one unique measure of error, which is the complement of the confidence level (1 – confidence level). Given that the candidate has little or no concern for pollster risk, the reliability of the confidence interval methodology is either very low or at best questionable. Candidate risk and pollster risk are distinct concepts as we have argued throughout, each one having its own determinants.

The content of table 2 and table 3 can be shown graphically in terms of the operating characteristic curve (OCC) already introduced above. We have seen in equation 5 that the probability of getting up to Ac opposition votes in a sample of size n taken from a population of N depends primarily on the percentage p of opposition votes in the population. If we use the sampling plan in table 3 with sample size 550, we can draw the OCC using these numerical values in figure 3. The two vertical lines are the AQLc and LTPDc respectively. The horizontal lines refer to candidate risk of 1% and the corresponding secondary risk of 50%. In other words following the sampling plan PL(300000, 550, 247) there is only a maximum 1% chance of committing the candidate error of not recognizing the adversary´s winning campaign.

 Figure 3 Acceptance sampling and the **CCO** PL(N,10,Ac), AQLc = 0.45, LTPDc = 0.5

R code in the appendix.

Besides working with smaller samples, another advantage of acceptance sampling is the speed and reliability of obtaining a meaningful result. For instance, as the number of opposition votes in the sample are counted sequentially, **the moment that the number of opposition votes exceeds Ac (opvotes>Ac), the sampling process indicates the acceptance of the null that our candidate is losing and counting can stop**. On the other hand, nor does the entire sample need be inspected to reject the null hypothesis that the candidate is losing the campaign.[[12]](#footnote-12) A confidence interval on the other hand could for example produce a point estimate of less than 45% support for the adversary but this result does not necessarily indicate that the adversary is losing, given the limits of the confidence interval that include the value 0.50. This is the case of sample size 550, table 3.

THE POLLSTER´S PERSPECTIVE

Even though the pollster is of minor importance in terms of bargaining power for choosing the parameters of the sampling plan, this section is dedicated to the construction of sampling plans that accommodate this point of view. Naturally a sampling plan that is agreeable for both candidate and pollster is preferable.

Given that the position of the pollster is theoretically similar to that of the producer/supplier in industry and commerce, the null from the pollster´s perspective as supplier of a service, and optimistic as to the quality of the service provided, would be that the opponent is losing AQLp ≥ p. In industry, the producer uses the null of good lots due to the high costs of incorrectly rejecting good lots including unnecessary massive inspections and even the unnecessary retooling of an entire assembly line. The result of the producer´s optimism is that he rejects relatively fewer lots and, consequently, allows some bad product to slip through inspection to the consumer. By rejecting less, the producer avoids a very costly type I error. Nevertheless, the producer cannot let an unusual amount of bad product enter the buyer´s operations and this demands that only a very small amount of bad product actually comes off the assembly line. The position of the pollster/consultant while analogous to the producer is nevertheless much more precarious in negotiations with the candidate. While it is true that nonconforming industrial lots can be identified by the buyer and rectified through 100% inspection and replacement of bad items, this alternative is not available in election campaigns. As stated earlier, once the pollster witness the election results, there is no turning back to reinspect the rejected lot (undo a lost election can only occur in computer simulations) but instead only regret. “Far easier to be forgiven if he wins then if he loses”, as a political consultant once commented in a private communication. In the end, we must conclude that the position of the pollster bears only a partial resemblance to the industrial supplier.

The question is what would be a reasonable value for AQLp? Certainly, 50% would be reasonable, but for security´s sake, the pollster could use 40%, indicating that the candidate has a substantial lead in the race. See contingency table 4. If the opponent is in fact losing the election and, in the sample, opvotes < Ac, then the sampling exercise has responded correctly, the pollster null (AQLp = 40%> p) has been verified. This is good and truthful news for our candidate (from the pollster). Furthermore, if the null is false and sampling produces opvotes > Ac, the weakness of our candidate is verified. Here the pollster brings bad news but at least it is truthful. However, when the sample exhibits opvotes>Ac, when the truth is that AQLp>p, falsely rejecting the truth that the opponent is losing, a type I error is committed by the pollster. The campaign looks weak by the sampling result when in fact it is not.

|  |  |
| --- | --- |
|  | **Pollster** chooses between states of the null hypothesis **Ho: adversary is losing** |
| Reality of null Ho: **adversary is losing** | accept Ho opvotes ≤Ac | reject Ho opvotes > Ac |  |
| true Ho 40%=AQLp≥p | **TN** correct P(opvotes≤Ac/40%=AQLp≥p) | **FP** type I error: **αp**=P(opvotes>Ac/40%=AQLp≥p)  | sum of prob = 1 |
| false Ho 50%=LTPDp<p | **FN** type II error: **βp**=P(opvotes≤Ac/50%=LTPDp<p)  | **TP** correct P(opvotes>Ac/50%=LTPDp<p) | sum of prob = 1 |

Table 4 Contingencies (pollster) same labels in table 2

A type II error for the pollster would occur when the opponent is winning LTPDp<p and the sample returned opvotes < Ac. The truth of the alternative hypothesis is based on the choice of a value for LTPDp. When LTPDp < p the alternative hypothesis is true (the opponent is winning) and therefore a reasonable value for LTPDp is 50%. If the pollster projects that our candidate will win (opvotes <Ac), when the candidate is actually losing, then the pollster´s alternative hypothesis is true (LTPDp<p) and in fact type II error has been committed.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | RISK |  |  |  |
|  | POLLSTER | SECONDARY | SAMPLING PLAN (hyper) |
|  | αp=1 - P(AQLp) | βp= P(LTPDp) | n | Ac | Ac/n |
| 1 | 0.005 | 0.01 | 594 | 268 | 0.451 |
| 2 | 0.01 | 0.01 | 535 | 240 | 0.448 |
| 3 | 0.012 | 0.013 | 499 | 224 | 0.449 |
| 4 | 0.013 | 0.013 | 495 | 222 | 0.448 |
| 5 | 0.01 | 0.02 | 476 | 215 | 0.452 |
| 6 | 0.02 | 0.02 | 417 | 187 | 0.448 |
| 7 | 0.025 | 0.025 | 384 | 172 | 0.448 |
| 8 | 0.05 | 0.05 | 268 | 120 | 0.448 |

Table 5 Sampling plans (pollster) for several values of

αp, βp, LTPDp is 50%; AQLp is 40%, N=300,000,

Calculations: R package Acceptance Sampling, Kiermeier (2008). R code in the appendix.

We learn from table 5 that pollster sampling plans can assure low levels of both risks and a sample size that is relatively small. Comparing table 3 sample size 550 to table 5 sample size 535 we can see similarity between

pollster sampling plan PL(300000, 535, 240)

candidate sampling plan PL(300000, 550, 247).

The ratio between Ac/n for the two plans 0.448 (= 240/535) and 0.449 (= 247/550). Both plans yield adequate protection against own errors for both sides: the probability of the pollster committing the error of declaring the adversary as winner when in fact he loses αp is a minimal 1%, and likewise, but more importantly, the minimal 1% error applies also to the candidate´s evaluation of his own risk αc. The pollster plan affords the additional satisfaction to the pollster of very small type II error βp of just 1%. The candidate´s type II error probability βc is rather large, but as argued above not as important for campaign strategy. The conclusion is that both plans yield extremely economical sampling procedures with risk levels much smaller than the traditional confidence intervals.

Table 6 presents a summary of acceptance sampling concepts and hypothesis testing for polling public opinion. We have given priority to the correct definition of the null hypothesis that depends on the point of view of the specific decision maker. This interpretation is from Neyman and Pearson (1933).

|  |  |  |
| --- | --- | --- |
| **Perspective of** | **Candidate** | **Pollster** |
| **Ho** | Adversary is winning | Adversary is losing |   |
|   | Rejection error: False negative **FN** | Rejection error: False positive **FP** |
|   | Type I error |   | Type I error |  |   |
|   | High cost |   |   | High cost |   |   |   |
| **Ha** | Adversary is losing | Adversary is winning |   |
|   | Rejection error: False positive **FP** | Rejection error: False negative **FN** |
|   | Type II error |   | Type II error |  |   |
|    | Low cost |   |   | Low cost |   |   |   |

Table 6 Summary of acceptance sampling concepts and hypothesis testing for polling public opinion

APPLYING THE RECEIVER OPERATING CHARACTERISTICS (ROC) CURVE TO POLLING

The construction of ROC curves for this case can secure a less conflicting result for the pollster and candidate[[13]](#footnote-13). The ROC curve shows the inverse relationship between the probabilities of type I and type II error as the parameter Ac is allowed to vary with a fixed sample size. In other words each point along the ROC curve represents a specific value for Ac. Recall figure 2 and related discussion. We choose n = 600 considered a small sample size and the parameter values AQLp= 0.40, AQLc= 0.43, and LTPDc= LTPDp = 0.50. Figure 4 displays the two corresponding ROC curves, one for the pollster (green) and the other for the candidate (blue)[[14]](#footnote-14). For candidate risk αc equal to about 0.8%, the candidate ROC curve in the upper right corner demonstrates secondary risk βc of about 15%. However, since candidate risk αc is much more important, the candidate would easily ignore the high secondary risk βc. The sampling plan PL(300000, 600, 270) would adequately satisfy the requirements of the candidate. Even for the pollster, whose ROC curve passes closer to the origin of the axes, the plan PL(300000, 600, 269) is based on risk factors each lower than 1%, a very advantageous situation for the pollster. In other words, this plan should please both candidate and pollster. Tables 7 and 8 present the representative sampling plans from figure 4 that could satisfy pollster and candidate. One of the important characteristics of these plans is that the risk from type I error is always less than the risk from type II error.



Figure 4 ROC curves for pollster AQLp=0.40, LTPDp=0.50 and

candidate AQLc=0.43, LTPDc =0.50

|  |  |  |  |
| --- | --- | --- | --- |
| n | Ac | αc = P(LTPDc) | βc= 1 - P(AQLc) |
| 600 | 269 | 0.006 | 0.171 |
| 600 | 270 | 0.008 | 0.151 |
| 600 | 271 | 0.010 | 0.133 |
| 600 | 272 | 0.012 | 0.116 |
| 600 | 273 | 0.015 | 0.101 |
| 600 | 274 | 0.019 | 0.087 |
| 600 | 275 | 0.023 | 0.075 |
| 600 | 276 | 0.027 | 0.064 |
| 600 | 277 | 0.033 | 0.054 |
| 600 | 278 | 0.039 | 0.046 |

Table 7 Axis values for figure 4, candidate AQLc=0.43, LTPDc =0.50

|  |  |  |  |
| --- | --- | --- | --- |
| n | Ac | αp=1 - P(AQLp) | βp= P(LTPDp) |
| 600 | 270 |  0.006  |  0.008  |
| 600 | 271 |  0.004  |  0.010  |
| 600 | 272 |  0.004  |  0.012  |
| 600 | 273 |  0.003  |  0.015  |
| 600 | 274 |  0.002  |  0.019  |
| 600 | 275 |  0.002  |  0.023  |
| 600 | 276 |  0.001  |  0.027  |
| 600 | 277 |  0.001  |  0.033  |
| 600 | 278 |  0.001  |  0.039  |
| 600 | 279 |  0.001  |  0.047  |

Table 8 Axis values for figure 4, pollster AQLp=0.40, LTPDp =0.50

PROCEDURE FOR MORE EFFICIENT SAMPLING

To illustrate an economical procedure for acceptance sampling in public opinion, we use the parameters from the preceding example using the ROC curve. The plan is PL(300000, 600, 270).

**STEP 1** Make sure that the 600 voters for the potential sample comes from a random draw of the population in question.

**STEP 2** Start interviewing voters in the potential sample and record voter preferences. For example, after the first 10 interviews the intermediate value of opvotes might be 6. After another ten interviews, the total number of opvotes might be 13.

**STEP 3, part 1** Assume that 271 opvotes were reached with only 500 total sample voters counted, opvotes > Ac. At this point, in line with the sampling plan, the null of the candidate (the opponent is winning) is declared true (table 2) and the null of the pollster (adversary is losing) is rejected (table 4). We have sufficient evidence to conclude that our campaign is failing. Sampling can stop before reaching the total number of sample voters of 600.

**STEP 3, part 2** In another hypothetical case, this time assume that after 430 sample voters are questioned, 100 (opvotes) of these voters are favorable to the adversary. In this case, it would be impossible to surpass the cutoff Ac of 270 opvotes in the sample of size 600 since there are only 170 (= 600 – 430) voters left in the sample to be questioned. Here we already have sufficient evidence to reject the null of the candidate (Ho: adversary is winning) and accept the null of the pollster (Ho: adversary is losing).

These steps are summarized in table 9 which is a suggestion for implementing acceptance sampling for polling public opinion. As interviews are held and voter opinions tabulated in the column opvotes(i), interviews continue until either opvotes(i) > Ac concluding that the adversary is winning, or (n-i) + opvotes(i) ≤ Ac in which case the adversary is judged as losing.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| n | i interview | n-i interviews remaining | Ac | opvotes(i) | ifopvotes(i) > Ac, then | if(n-i) + opvotes(i) ≤ Acthen |
| 600 | 1 | 599 | 270 |  | .. | .. |
| 600 | 2 | 598 | 270 |  | .. | .. |
| 600 | 3 | 597 | 270 |  | .. | .. |
| 600 | 4 | 596 | 270 |  | .. | .. |
| 600 | .. | .. | 270 |  | STOP | STOP |
| 600 | .. | .. | 270 |  | adversary | adversary |
| 600 | .. | .. | 270 |  | winning | losing |

Table 9 Spreadsheet for tabulating sequential polling results. opvotes(i) the number of opvotes counted up to and including the i-th interview.

FINAL THOUGHTS

Acceptance sampling is a statistical procedure that deserves at least a tentative application in public opinion polls. Unlike confidence intervals, it takes into account both over and under estimation errors with different weights and from different viewpoints. Furthermore, it results in smaller samples and more accuracy. Theoretical merit is abundant, practical verification is underway.

**T**he paper has shown that sampling plans should be true to the data and the situation they represent. When lot sizes are finite, statistical approximations may lead to serious estimation errors. Since the process of sampling is inherently error prone, the sampling process should be applied with the greatest care in using the most accurate data available, including lot size. The priority for researchers in any area should be the utilization of the most appropriate formulations, like the hypergeometric distribution in sampling plans, so that unnecessary additions to inherent errors do not occur.

There are several directions for advancing this project. One would be to increase the dimensions of the characteristic under analysis to three or more alternatives. Polling for instance has undecided responses and the production line can offer output of medium quality.

This paper has not mentioned the extremely important area of Bayesian statistics (CHIUH-CHENG CHYU AND I-CHUNG YU, 2006),. This was not because we wanted to devalue its appropriateness for sampling plans, but rather because its contribution to the questions of this paper based entirely on traditional frequentist statistics would be doubtful. Bayesian approaches will be an important part of future research.

Finally, we would suggest that the calculations proposed here could easily be made automatically through the application of the R package Shiny (shiny.rstudio.com). Contributions to the literature have already gone in this direction (Hund, 2014).

REFERENCES

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS (1989). “Planos de amostragem e procedimentos na inspeção por atributos - ABNT NBR 5426:1985, Versão Corrigida (corrected version)”.

BIEDRON, C.; PAGANO, M.; HEDT, B. L.; KILIAN, A.; RATCLIFFE, A.; MABUNDA, S.; VALADEZ, J. J. (2010). "An assessment of Lot Quality Assurance Sampling to evaluate malaria outcome indicators: Extending malaria indicator surveys". International Journal of Epidemiology, , 39 (1): 72–79.

CHIUH-CHENG CHYU AND I-CHUNG YU, (2006), “A Bayesian Analysis of the Deming Cost Model with Normally Distributed Sampling Data.” Quality Engineering, 18:107–116. LLC ISSN: 0898-2112 print=1532-4222 online. DOI: 10.1080/08982110600567442

DEMING, W.E. (1986). Out of the Crisis. MIT Press, Cambridge, Mass.

DODGE, H.F., ROMIG, H.G. (1929). “Sampling Inspection.” The Bell System Technical Journal.

DODGE, H.F., ROMIG, H.G. (1944). Sampling Inspection Tables. New York: John Wiley & Sons. ‘

GELMAN, Andrew. 2013. “P Values and Statistical Practice.” Epidemiology (Cambridge, Mass.) 24 (1) (January): 69–72. doi:10.1097/EDE.0b013e31827886f7[. http://www.ncbi.nlm.nih.gov/pubmed/23232612.](file:///C%3A%5CUsers%5Csamohyl%5CAppData%5CRoaming%5CMicrosoft%5CWord%5CACCEPTANCE%20SAMPLING1304330541670010277%5C.%20http%3A%5Cwww.ncbi.nlm.nih.gov%5Cpubmed%5C23232612)

GONIN, H. T. (1936) "The use of factorial moments in the treatment of the hypergeometric distribution and in tests for regression," Philosophical Magazine, 7, Ser. 21, 215-226.

HEDT, BETHANY L, OLIVES, C., PAGANO, M., AND VALADEZ, J. (2008). “Large Country-Lot Quality Assurance Sampling : A New Method for Rapid Monitoring and Evaluation of…”, Health, Nutrition and Population (HNP) Discussion Paper: [www.worldbank.org/hnppublications](file:///C%3A%5CUsers%5Csamohyl%5CAppData%5CRoaming%5CMicrosoft%5CWord%5CACCEPTANCE%20SAMPLING1304330541670010277%5Cwww.worldbank.org%5Chnppublications) .

HUND, L. (2014). “New Tools for Evaluating LQAS Survey Designs.” *Emerging Themes in Epidemiology* 11: 2. doi:10.1186/242-7622-11-2. <http://www.ete-online.com/content/11/1/2>.

INTERNATIONAL STANDARD ISO 2859-1 (1999) Sampling procedures for inspection by attributes - Part 1: Sampling schemes indexed by acceptance quality limit (AQL) for lot-by-lot inspection. Second edition -11-15.

INTERNATIONAL STANDARD ISO 2859-2 (1985) Sampling procedures for inspection by attributes - Part 2: Sampling plans indexed by limiting quality (LQ) for isolated lot inspection First edition.

KIERMEIER A (2008). “Visualizing and Assessing Acceptance Sampling Plans: The R Package AcceptanceSampling.” Journal of Statistical Software, 26(6). URL http://www. jstatsoft.org/v26/i06/.

LEHMANN, D. R. (1993). “Fisher, Neyman-Pearson.”

NEYMAN, J. AND PEARSON, E. (1933). On the problem of the most efficient tests of statistical hypotheses.

*Philosophical Transactions of the Royal Society of London, Series A, 231*, 289–337.

NIST/SEMATECH(2012). *e-Handbook of Statistical Methods*, <http://www.itl.nist.gov/div898/handbook/>.

PAGANO, M. AND VALADEZ, J. J. (2010). “Commentary: Understanding Practical Lot Quality Assurance Sampling.” International Journal of Epidemiology 39 (1) (February): 69–71. doi:10.1093/ije/dyp406. [http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=282093&tool=pmcentrez&rendertype=abstract](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2817093&tool=pmcentrez&rendertype=abstract).

PEZZOLI, LORENZO, AND SUNG HYE KIM. (2013). “Monitoring Health Interventions--Who’s Afraid of LQAS?” Global Health Action 6. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3822089&tool=pmcentrez&rendertype=abstract>.

PROVOST, F., and FAWCETT, T. (1997) “Analysis and visualization of classifier performance: comparison under imprecise class and cost distributions”. KDD 1997, American Association of Artificial Intelligence.

PROVOST, F., and FAWCETT, T. (2013). Data Science and Business. O´Reilly, Cambridge.

R CORE TEAM (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL [http://www.R-project.org/.](http://www.r-project.org/)

RICE, JOHN, (1995). Mathematical Statistics and Data Analysis, Second Edition. Duxbury Press.

RHODA, D.A|., FERNANDEZ,S.A. FITCH, D.J. AND LEMESHOW,S. (2010). “LQAS: User Beware.” *International Journal of Epidemiology* 39 (1) (February): 60–8. doi:10.1093/ije/dyn366. <http://www.ncbi.nlm.nih.gov/pubmed/20139433.>

RSTUDIO Inc (2013). *shiny: Web Application, Framework for R, R package*.

[Version 0.5.0]. http://CRAN.R-project.org/package=shiny.

SAMOHYL, R. (2013). “Audits and Logistic Regression, Deciding What Really Matters in Service Processes. A Case Study of a Government Funding Agency for Research Grants.” *Proceedings of the Fourth International Workshop on Knowledge Discovery, Knowledge Management and Decision Support*. doi:10.2991/.2013.33. <http://www.atlantis-press.com/php/pub.php?publication=eureka-13&frame=http%3A//www.atlantis-press.com/php/paper-details.php%3Fid%3D9647.>

SHMUELI, GALIT (2011). *Practical Acceptance Sampling*. Second edition.

R code

#Ph CDF function CCO

setwd("C:/Users/samohyl/Google Drive/Amostragem-Livro")

getwd()

rm(list=ls(all=TRUE))

library("AcceptanceSampling", lib.loc="~/R/win-library/3.0")

N=rep(300000,100)

n=rep(550,100)

c=rep(247,100)

p=seq(0.445,0.505,length.out=100)

Ph=phyper(c,as.integer(p\*N),as.integer((1-p)\*N),n)

Pb=pbinom(c,n,p)

data.frame(Ph,Pb,p,N,n,c,as.integer(p\*N),as.integer((1-p)\*N))

par(lwd=2)

plot(p,Ph,type="l",col="green",

ylab=

"candidate risk = 1% pollster risk = 50%",

xlab="p per cent opposition votes in population",

ylim=c(0,.52))

N=rep(1000,100)

Ph=phyper(c,as.integer(p\*N),as.integer((1-p)\*N),n)

data.frame(Ph,Pb,p,N,n,c,as.integer(p\*N),as.integer((1-p)\*N))

lines(p,Ph)

#lines(p,Pb,col="blue")

abline(v=c(.45,.50))

abline(h=c(.5,.01))

legend(.47,.3, yjust=0,

 c("hyper, N=300000, n=550, Ac=247", "hyper, N=1000, n=550, Ac=247"),

 lty=c(1,1,1,1), col=c( "green","black"))

#Polling Public Opinion Sampling Plan - iterated

setwd("C:/Users/samohyl/Google Drive/Amostragem-Livro")

getwd()

rm(list=ls(all=TRUE))

library("AcceptanceSampling", lib.loc="~/R/win-library/3.0")

AQL=0.45;LTPD=0.5;N=300000 #Pollster risk (1-P(AQL)); Candidate risk P(LTPD)

size=0;Ac=0;reject=0;PR=0;CR=0;samplepercent=0;sampleAQL=0;sampleLTPD=0

for(i in seq(1,15,1))

{

 PR[i]=i\*.01;CR[i]=.01

 plan=find.plan(PRP=c(AQL, 1-PR[i]), CRP=c(LTPD, CR[i]),N,

 type="h") #b=binomial; h=hypergeometric

 size[i]=plan$n

 Ac[i]=plan$c

 reject[i]=plan$r

 samplepercent[i]=Ac[i]/size[i]

 sampleAQL[i]=AQL\*size[i]

 sampleLTPD[i]=LTPD\*size[i]

}

result=t(as.matrix(rbind(PR,CR,size,Ac,samplepercent,sampleAQL,sampleLTPD)))

result

1. Thanks to Jan van Lohuizen, Osiris Turnes and Armin Koenig for comments on an earlier version (WP15-002). [↑](#footnote-ref-1)
2. The Hawthorne Plant/Bell laboratories are mentioned in NIST/SEMATECH(2012) and historical references given. [↑](#footnote-ref-2)
3. This paper is limited to the dual response yes and no (favorable and unfavorable) however the multi option response which would include 3 or more alternatives is the object of chapter 9. [↑](#footnote-ref-3)
4. See chapters 2 and 3 on consumer and producer risk for a more complete description of Dodge-Romig sampling plans. A popular introduction to acceptance sampling is Shmueli, (2011). [↑](#footnote-ref-4)
5. See Gonin, H. T. (1936) for the development of equation 3. [↑](#footnote-ref-5)
6. In the appendix of this book, we show that the hypergeometric distribution is at the origin of the binomial and the poisson distributions. The binomial is a specific case of the hypergeometric with infinite population size and the Poisson is a specific case of the binomial with infinite sample size. [↑](#footnote-ref-6)
7. Later on, opvotes will be used instead of c, meaning the number of opposition votes in the sample. [↑](#footnote-ref-7)
8. We have added the subscript p to represent producer/pollster. This will be an important distinction for our arguments below. [↑](#footnote-ref-8)
9. Note that we have added the subscript c to designate consumer (and candidate). Also called QL (INTERNATIONAL STANDARD ISO 2859-1, (1999)). [↑](#footnote-ref-9)
10. We have explicitly avoided using the traditional alpha (α) and beta (β) designations for producer and consumer risk, better elaborated in the next section on hypothesis testing. [↑](#footnote-ref-10)
11. Chapter 5 is devoted to more details concerning the Neyman-Pearson paradigm for hypothesis testing. [↑](#footnote-ref-11)
12. Procedures for inspection of the sample will be outlined below. [↑](#footnote-ref-12)
13. The concepts involving the ROC curve and our application in acceptance sampling is covered in more detail in chapter 5. [↑](#footnote-ref-13)
14. Complete axes for ROC curves extend from 0 to 1. Here they have been shortened to clarify the illustration. [↑](#footnote-ref-14)