Game theoretic computing of producer's and consumer's risks, α & β, for acceptance sampling using cost and utility

Game theoretic computing of producer's & consumer's risks

Mehmet Sahinoglu

Professor Emeritus, Auburn University and Adjunct Faculty, Troy University

mesa@aum.edu, mesa@troy.edu

KEYWORDS

mathematical statistics, non-linear smooth optimization, quality control inspection, type-I error probability, type-II error probability.

Game theoretic computing of producer's and consumer's risks, α & β, for acceptance sampling using cost and utility

ABSTRACT

When establishing a hypothesis testing procedure to ensure its credibility, the most significant of all is unquestionably to select and/or compute the optimal Type-I and the Type-II error probabilities, namely Producer's and Consumer's Risks, α & β respectively. This article as opposed to the conventionally and judgmentally picking at best a subjective Type-I error probability, outlines a Game theoretic approach, i.e. that of von Neumann, to this historically unresolved paradigm to justify optimal choices for Type-I error probability (α) and Type-II error probability (β) when cost, utility and other market-centric factors are incorporated as input data. A game theory-based algorithmic methodology and several numerical examples of practical nature with specific emphasis to company-specific Acceptance Sampling plans for Quality Assurance are illustrated. A side benefit of this method in addition to improving the Acceptance Sampling plans is to transform the traditional Hypothesis Testing process in making sound engineering decisions from a "subjective" to an "objective" stance, provided that the monetary cost and utility consequences of committing error and non-error combinations are available*.

KEYWORDS

non-linear smooth optimization, quality control inspection, type-I error probability, type-II error probability

- The author has no potential conflict of interest relevant to this article.

1. INTRODUCTION AND MOTIVATION

The objective of this article lies in optimizing a procedure to statistically test the credibility of a stated hypothesis in today's quality control-conscious and business-savvy world. Aside from the usual rule-of-thumb or best-guess or judgment-call-based choices of such as 1-out-of-20 or 1-out-of-50 etc., there have been alternative attempts to compute α (Type-I error probability) by deriving the first and second derivatives of the standard normal distribution curve. This is performed by determining the second derivative to reach maximum at $z=\pm1.732$ which corresponds to a p-value of 0.083. An alternative approach has been to find a point where the concavity in the normal distribution curve is maximal to the first derivative. That is, the maximal curvature $k(z)$ occurs when $z = \pm1.749$ corresponding to a p-value of 0.08. The p-value is used to reject H_0 for a given α. The calculus-based algebraic approaches have been recently studied by Kelley[1] (2013) and Grant[2] (2014). As Kelley[1] (2013) quoted, "No one therefore has come up with an objective statistically based reasoning behind choosing the now ubiquitous 5% level, although there are objective reasons for levels above and below it. And no one is forcing us to choose 5% either." The issue with these approaches is that they are detached from the market realities such as cost (loss) or utility (revenue) associated with varying error values (α and β), or non-error values (1-α, and 1-β) and their cross products, namely [$\alpha * \beta$], [$\alpha *$ (1-β)], [(1- α) $* \beta$] and [(1- α) $*$ (1- β)] that manifest themselves in the form of producer's or consumer's risks, or both or none. Not only are the usual judgment-call based selections subjective, those are also not attached to any joint treatment of producer's and/or consumer's risks that may occur. Simply because, millions of products are subject to producer's risk (underappreciated or declared bad by the consumers while in truth 'H_0: Good Product' costing the producer a financial loss) or consumer's risk (over appreciated or declared good by the disadvantaged consumer while in truth 'H_1: Bad

Product' but costing the company due to the ripple effects of the bad publicity circulating until the reality is discovered) or both errors or partial, occurring through partial risks where a marginal type of error is involved. Also, none may have incurred with no financial loss for the producers and consumers with a complete market satisfaction due to the cross-product of 'Power: (1-β)' and 'Confidence: (1-α)' of these hypothesis tests.

Game Theory is a branch of Mathematical Sciences devoted to the logic of decision-making in social or managerial interactions, and concerns the behaviour of decision-makers whose decisions affect each other (Sahinoglu et al.[3], 2012; Blackwell and Girschik[4], 1954). Each decision maker has only partial control. Game Theory is a generalization of *Decision Theory* where two or more decision makers compete by selecting each of the several optimal strategies, whereas *Decision Theory* is essentially a one-person game theory. A common practice is to select Type-I error probability (*Alpha*) by an existing best-judgement call, and then, given an alternative H_1 values, to compute the related Type-II error probabilities (*Beta*). The plan is to build Power Curve ("1.0-Beta" vs "Population Mean: Mu") or Operating Characteristics Curve ("Beta" vs "Population Mean: Mu") for the critical Acceptance Sampling procedure. This paper therefore outlines a Game theoretic application to a long-sought, unresolved statistical procedure toward the optimal estimation of *Alpha* (Type-I error probability: α) and *Beta* (Type-II error probability: β) while maximizing the overall sales revenue. This proposed method becomes feasible to implement when market-based cost and utility input data for composite errors (or lack-of) are deliverables. The optimal estimates, given the input cost (and utility) parameters, can be implemented to business enterprises. The objective of this research paper is to develop computationally intensive math-statistical algorithms such as a game theoretic approach to estimate Type-I and Type-II error probabilities, i.e. Producer's and Consumer's Risks, respectively, for the statistical hypothesis testing of a *wellness* of a given product such as that of a piece of a cyberware (or chip) vs otherwise. The end goal is to objectively compute the Operating Characteristic (OC) curves or Power Plots for the process of Acceptance Sampling. For sentencing incoming batches, or lots (batches) of items without doing 100% inspection, Acceptance Sampling is a well-known statistical sampling procedure to determine the quality level of an incoming shipment in order to judge whether quality level is within those limits predetermined to be desirable. However, the Acceptance Sampling gives one no idea about the process that is producing those items. Therefore, Acceptance Sampling is a form of sampling inspection applied to lots or batches of items before or after a process to judge conformance with predetermined standards. Sampling Plans are those plans that specify lot or batch size (large N), sample size (small n) and acceptance or rejection criteria. This article will deal with Single-sampling plans where Double- and Multiple- or Sequential-sampling plans are excluded because out of scope. Within the *Single-sampling* plans of the *incoming (vs outgoing) reception* and *non-rectifying (vs rectifying)* classifications, this current article will examine scenarios or examples to study sampling by *variables, usually gauge-measured for mean and standard deviation* (with continuous Normal pdf). However, batch sampling for *attributes counting the number of defectives* will also be adapted for large sample size $n \geq 100$ with Normal to approximate Binomial pdf. Acceptance Sampling has two critical levels: 1) *Acceptance Quality Level or Limit* (*AQL*) that denotes the probability or percentage of defects at which consumers are willing to accept lots or batches as "good". 2) *Rejectable Quality Level or Limit* (*RQL*) that denotes the upper limit on the percentage or probability of defects that a consumer is willing to accept lots or batches as "good". Therefore, RQL or LTPD (lot tolerance percent defective) is the poorest quality level that the consumer is willing to accept in an individually inspected lot. This brings us to the two essential concepts: A) *Producer's Risk*: The probability that a lot containing defectives, either through gauge-measurement or head-counting, between the allowed limits of AQL and RQL will wrongly be rejected. B) *Consumer's Risk*: The probability that a lot containing defectives, either through gauge-measurement or head-counting, outside the limits of *AQL* and *RQL* will be wrongly accepted. Briefly Type-I error of the Producer's Risk (5% is common) is the probability of rejecting a good lot (or batch), whereas Type-II error of the Consumer's Risk (10% is a typical value) is the probability of not-rejecting a bad lot by Baghci[5] (2012). Lastly, the 5% ($=\alpha$) or 10% ($=\beta$) cited are subjective takes but not scientific. MIL-STD-105E's (among others) standard assumption on α and β are what this article questions like Kelley[1] (2013) and Grant[2] (2015) did.

2. OBJECTIVES, AND METHODS: DECISION TABLES, RISKS, AND ERRORS

The concept of Game Theory has been brought to the attention of Hypothesis Testing in the past but at strictly theoretical (albeit not pragmatic) level by Schlag[6] (2008) involving the establishment of finite sample bounds on the general theme of statistical inference where Schlag[6] (2008) expressed," Let the data speak!" and compared tests based on sequential sampling. Schlag[6] used Game Theory to establish the minimal *Type-II* error (β: *Beta)* whereby the associated randomized test was characterized as part of Nash[7] (1950) equilibrium, also as in Osborne and Rubinstein[8] (1994). However, these attempts did not lead to an algorithmically simple, ubiquitous and practically formulated usability by the everyday practicing statistician routinely dealing with hypothesis testing at an elementary level. Sahinoglu et al.[9,10,11] (2015, 2016, 2017) instead followed up with a pragmatic approach, respectively, with a published ASA'15 proceedings paper and a Wiley Inc. textbook, and an ISI'17 proceedings paper. As pointed out by Savage[12] (1954), Game Theory can also be used to solve problems in statistics. The underlying idea is to solve worst-case problems by invoking the minimax theorem for zero-sum games developed by von Neumann[13] (1928) before the WWII, and further improved with the contributions of Neumann and Morgenstern[14] (1944) at Princeton. However, Game theoretical methods have not yet been used in hypothesis testing curriculum in layman's terms to teach the fundamental concepts at an elementary statistics level. Mainly because the applications to everyday routine hypothesis tests with pertinent costs associated to *Type-I* (α) and *Type-II* errors (β) and their cross products, and additionally, utility or revenue with respect to non-errors (Confidence = 1-α, and Power = 1-β) were not properly formulated. However in this applied research paper, the author deals with von Neumann's Game theoretic equilibrium approach. In a hypothesis testing scenario, one associates a variety of costs (money lost due to the decision errors) or a utility (revenue for the non-error) and observe what the optimal α and β will turn out to be by employing the principles of Game Theory. This is rather than sticking to a rule of thumb, such as $\alpha = 0.05$ or $\alpha \approx 0.08$ etc. by calculus algebra as pointed out and outlined by Kelley[1] (2013) and Grant[2] (2015). The new approach vs the previous one based on a subjective rule-of-thumb with no econometric parameters and devoid of cost factors, is market-friendlier. To determine whether to reject a null hypothesis based on a sample data, there is statistical hypothesis testing with four steps outlined in the statistical literature (Ostle and Mensing[15], 1975; Hogg and Ledolter[16], 1992). The two types of errors can result from testing a statistical hypothesis *Ho* as follow:

Type-I error: A *Type-I* error occurs when the analyst rejects a null hypothesis when it is true. The probability of committing a *Type-I* error is called the significance level. This probability is conventionally denoted by α. This is also known in industrial quality control or assurance sciences as the *producer's risk,* where *Ho: Good Product* vs *H_1: Bad*

Product. The probability of not committing a *Type-I* error is called the Confidence of the test (1- α). Note, if "|" denotes "given that", then the *producer's risk* is given by:

$$\boldsymbol{\alpha = P\ \{Type\text{-}I\ \textbf{error}\} = P\ \{\textbf{reject}\ H_0 \mid H_0\ \textbf{is true}\}} \qquad (1)$$

Type-II error: A *Type-II* error occurs when the analyst fails to reject a null hypothesis that is false. This is a grave error in dealing with medical tests when a false H_o*: Well-patient* vs a true H_1*: Sick patient* is not rejected. The probability of committing a *Type-II* error is β. This is known in industrial quality control or assurance sciences as the *consumer's risk*:

$$\boldsymbol{\beta = P\ \{Type\text{-}II\ \textbf{error}\} = P\ \{\textbf{fail to reject}\ H_0 \mid H_0\ \textbf{is false}\}} \qquad (2)$$

The probability of not committing a *Type-II* error is called the Power of the test (1- β):

$$\boldsymbol{(1\text{-}\beta) = P\ \{\textbf{reject}\ H_0 \mid H_0\ \textbf{is false}\}} \qquad (3)$$

Note, the Power Function is represented as [1- $\beta(\theta)$], where θ denotes the true parameter value or population mean: μ. The $\beta(\theta)$, the complement of Power Function, is known as the Operating Characteristic (*OC*) function, popularly used in quality control. Observe Table 1 for the types of errors and their cross-products accompanied by associated costs.

TABLE 1 Costs (C_{11}, C_{12}, C_{21}) and Utility (C_{22}) for the cross-products of types of errors

	β↓	(1-β)↓
α →	*C11*	*C12*
(1-α) →	*C21*	*C22*

Cost (opposite of utility) Matrix is a function of α, β and C_{ij} related to the cross product of *Type-I* and *Type-II* errors. If *cost* bears a negative value, then cost denotes *utility*. Also:

$$\boldsymbol{\alpha * \beta + \alpha * (1\text{-}\beta) + (1\text{-}\alpha) * \beta + (1\text{-}\alpha) * (1\text{-}\beta) = 1.0;\ 0< \alpha, \beta <1} \tag{4}$$

$$\boldsymbol{\Pi(\alpha, \beta, C_{ij}) = \alpha*\beta*C_{11} + \alpha*(1\text{-}\beta)*C_{12} + (1\text{-}\alpha)*\beta*C_{21} + (1\text{-}\alpha)*(1\text{-}\beta)*C_{22};\ 0< \alpha, \beta <1} \tag{5}$$

where Π (α, β, C_{ij}) is the Expected Total Cost. Let $P_{11} = \alpha*\beta$, $P_{12} = \alpha*(1\text{-}\beta)$, $P_{21} = (1\text{-}\alpha)*\beta$, $P_{22} = (1\text{-}\alpha)*(1\text{-}\beta)$ where C_{11}, C_{12}, and C_{21} are assigned total costs respectively due to products of errors, or non-errors in the case of C_{22} in Table 1, from which one derives:

$$\boldsymbol{\alpha = P_{11} + P_{12}} \tag{6}$$

$$\boldsymbol{\beta = P_{11} + P_{21}} \tag{7}$$

3. COMPOSITE-, PARTIAL- AND NON-RISKINESS, AND AN EXAMPLE

Example 1: Given the following input data e.g. about the diameter of a circular integrated circuit (IC) board of a computer chip critical to automobile manufacturing, test as follow:

One-sided Null hypothesis: H_0: Cyberware is functional (good, operating),

i.e. H_0: μ_0: *Population mean* of an integrated circuit (IC) board's diameter = 5 cm.

One-sided Alternative hypothesis: H_1: Cyberware is dysfunctional (bad, not-operating),

i.e. H_1: μ_1: *Population mean* of an integrated circuit (IC) board's diameter) $\geq$ 5 cm.

Given the input sample costs**;** C_{11} = +\$800K (cost lost), C_{12} = +\$70K (cost lost), C_{21} = +\$200K (cost lost), and C_{22} = -\$400K (utility) are the cost coefficients in the order, respectively with the following expressions, where \$K=\$1,000:

Composite Riskiness (*CR*) = $\boldsymbol{P_{11} = \alpha * \beta}$ (8.A)

Partial Riskiness *(PR₁)* due to *Type-I error probability* only = $\boldsymbol{P_{12} = \alpha * (1\text{-}\beta)}$ (8.B)

Partial Riskiness *(PR₂)* due to *Type-II error probability only* = $\boldsymbol{P_{21} = (1\text{-}\alpha) * \beta}$ (8.C)

Non-Riskiness due to *Power and Confidence* = $\boldsymbol{P_{22} = (1\text{-}\alpha) * (1\text{-}\beta)}$ (8.D)

It remains to optimize Type-I (α) and Type-II (β) error probabilities, *Producer's* and *Consumer's Risks* respectively with Game theoretic mixed-strategy solution by Neumann et al.[13,14] (1928, 1944) and Sahinoglu et al.[9,10,11] (2015; 2016, pp. 6-13; 2017). Note:

$$\boldsymbol{CR}\textbf{(Composite Riskiness) + }\boldsymbol{PR}\textbf{(Partial Riskinesses) + }\boldsymbol{NR}\textbf{(Non-Riskiness)=1} \qquad (9)$$

It is timely to formulate Neumann's Game theoretic risk computing with details including the objective function: Min *LOSS* (which is the defensive gamer's objective function transforming to Max *GAIN* when eyed through the rivalling offensive gamer's perspective) by Sahinoglu et al.[3] (2012) subject to constraints from Equations 10 to 23:

$$P_{11} * C_{11} - LOSS < 0 \qquad (10)$$

$$P_{12} * C_{12} - LOSS < 0 \qquad (11)$$

$$P_{21} * C_{21} - LOSS < 0 \qquad (12)$$

$$P_{22} * C_{22} - LOSS < 0 \qquad (13)$$

$$P_{22} \geq P_{11} \qquad (14)$$

$$P_{22} \geq P_{12} \qquad (15)$$

$$P_{22} \geq P_{21} \qquad (16)$$

$$P_{11} < 1 \qquad (17)$$

$$P_{12} < 1 \qquad (18)$$

$$P_{21} < 1 \qquad (19)$$

$$P_{22} < 1 \qquad (20)$$

$$LOSS > LOSS_{\mathrm{min}} \qquad (21)$$

$$P_{11} + P_{12} + P_{21} + P_{22} = 1 \qquad (22)$$

$$\Pi\ (\alpha, \beta, C_{ij}) = P_{11} * C_{11} + P_{21} * C_{21} + P_{12} * C_{12} + P_{22} * C_{22} < 0 \qquad (23)$$

Observe the General Representation Theorem (GRT) behind linear programming for an outline of the forward and backwards proofs, given by Lewis[17] (2008, pp. 17-22). The following spreadsheets show the data entry and outputs with NLP: Non-Linear Programming algorithm, whereas equation (23) denoting total cost ($) units accrued, for instance, shows a positive utility or overall profit. Note that (14) to (16) are optional.

If the minimum or at-least utility assumed is *–LOSS* ≤-$5 or *LOSS*≥$5 (Equations 10 to 13 and 21) per cells in Table 2 as follows, one sets up the NLP (Non-Linear Programming) problem given the game-theoretic constraints of Equations 10 to 23, for Min LOSS. The following spreadsheets show the data entry and outputs with the Game-theoretic NLP (Non-Linear Programming) for Example 1. Nonlinear means not necessarily linear.

TABLE 2 Game theoretic input spreadsheet for Example 1

Enter/Edit data: Objective function coefficients. For each constraint, enter constraint coefficients, constraint relationship (<, =, >), and constraint right-hand-side value. Do not enter nonnegativity constraints.

Optimization Type: Minimize

Variable Names: (Change if Desired)	P11	P21	P12	P22	LOSS
Objective Function Coefficients:					1

	Coefficients				
Subject To:	P11	P21	P12	P22	LOSS
Constraint 1	1	0	0	0	0
Constraint 2	0	1	0	0	0
Constraint 3	0	0	1	0	0
Constraint 4	0	0	0	1	0
Constraint 5	1	1	1	1	0
Constraint 6	800	0	0	0	-1
Constraint 7	0	70	0	0	-1
Constraint 8	0	0	200	0	-1
Constraint 9	0	0	0	-400	-1
Constraint 10	0	0	0	0	1
Constraint 11	800	70	200	-400	0
Constraint 12	-1			1	0
Constraint 13		-1		1	0
Constraint 14			-1	1	0

TABLE 3 Feasible Vector Solution for Example 1

Optimal Solution

Objective Function Value = 5.000

Variable	Value
P11	0.006
P21	0.071
P12	0.025
P22	0.897
LOSS	5.000

Note in the EXCEL Solver, the GRG (General Regression) nonlinear engine is selected for constrained minimization problems with *differentiable* (where partial derivatives of order k are continuous), *nonlinear*, *convex* functions that are *smooth of order k* by Rapcsak[18] (1997). This includes the case where all functions are linear, i.e. the linear programming problem.

TABLE 4 Input $\{C_{ij}, i,j=1,2\}$and Output EXCEL Solution of $\{P_{ij}, i,j=1,2\}$ for Example 1

C11	800
C21	70
C12	200
C22	-400

MIN LOSS				
P11	P21	P12	P22	LOSS
0.00625	0.071428571	0.025	0.897322	5
P11	0.0062500000		<	1
P21	0.071428571		<	1
P12	0.025		<	1
P22	0.897322429		<	1
Constraint 1	-343.9289714		<	0
Constraint 2	1.00000		equal	1
Constraint 3	2.04281E-14		<	0
Constraint 4	0		<	0
Constraint 5	7.30971E-13		<	0
Constraint 6	-363.9289714		<	0
Constraint 7	5		>	5

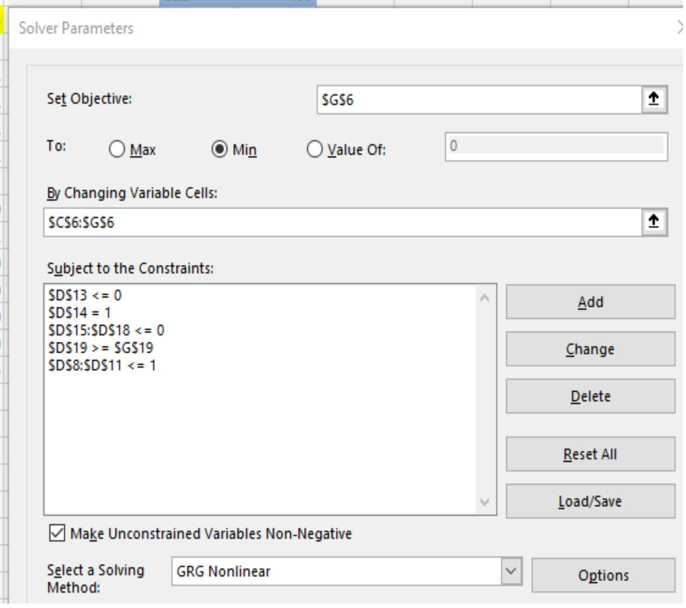
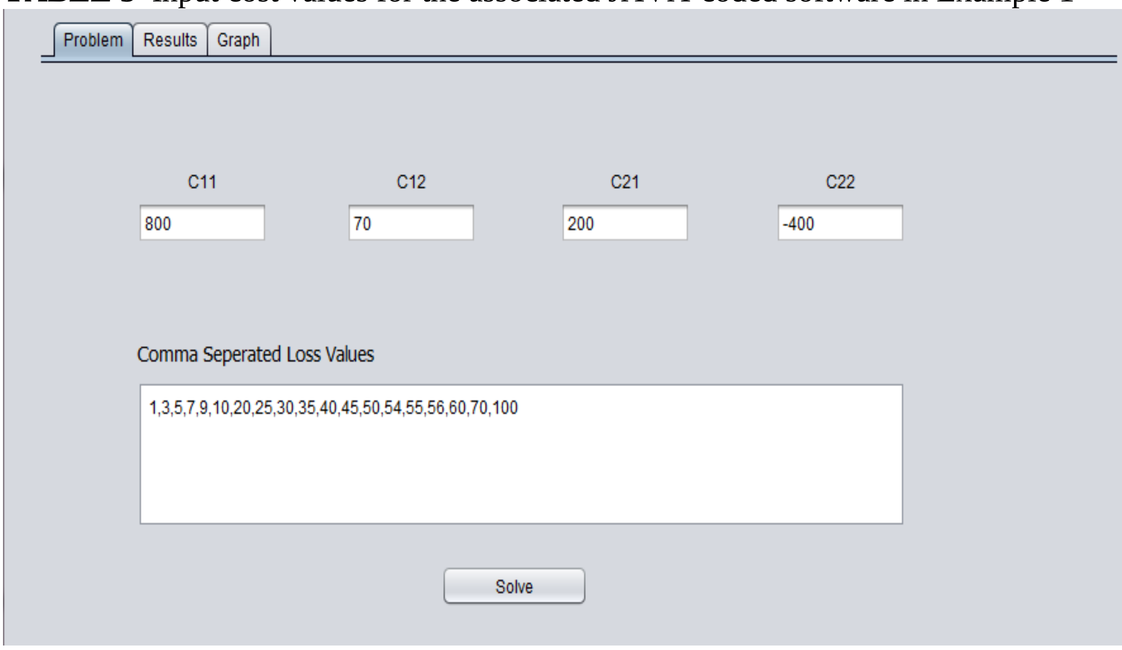


TABLE 5 Input cost values for the associated JAVA-coded software in Example 1



The results for the horizontal axis of *LOSS*: $5:

P_{11} = 0.00625, P_{12} = 0.025, P_{21} = 0.071, P_{22} = 0.8973

Alpha: 0.0313, *Beta*: 0.0777 and Expected Total Cost: -$343.92

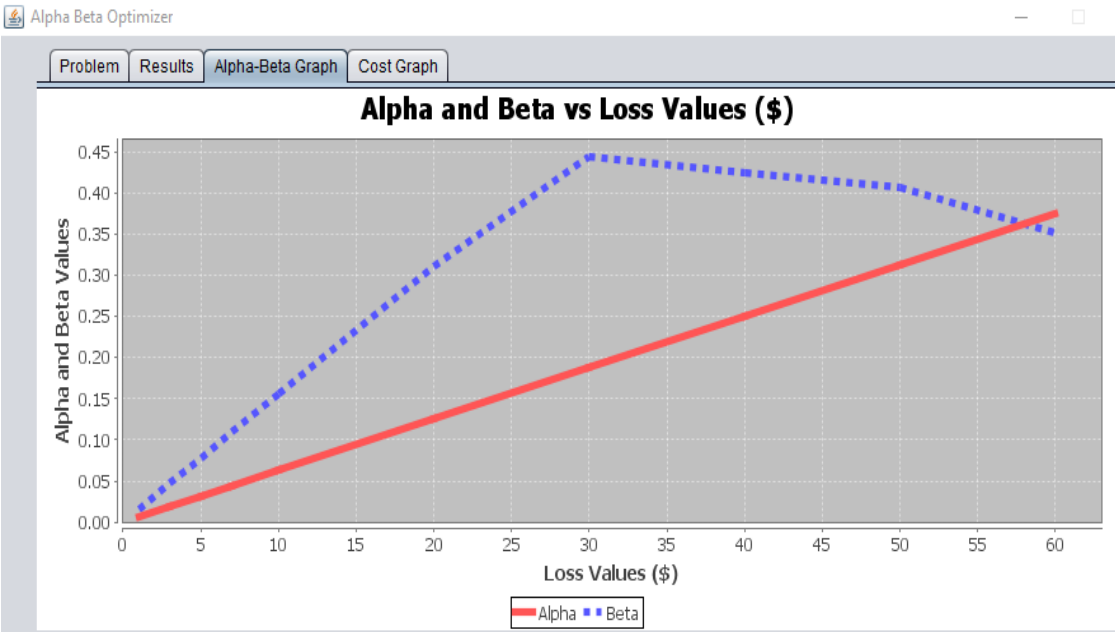


FIGURE 1 Game theoretic *Alpha* and *Beta* solutions vs *LOSS* Values in Example 1

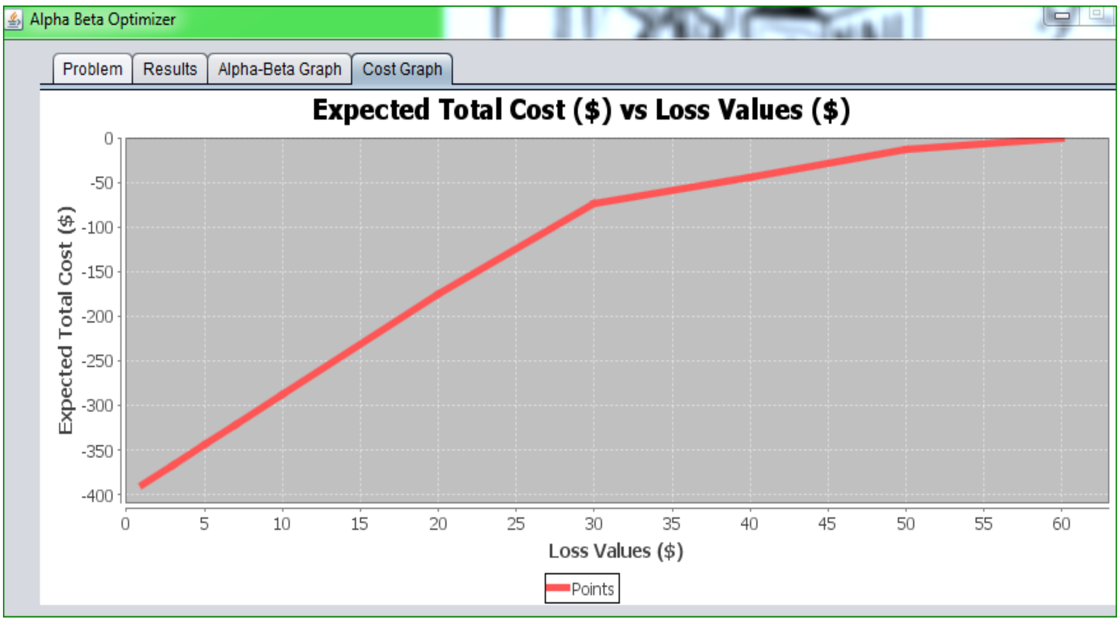


FIGURE 2 Game theoretic Expected Total Cost ≈-$344 vs *LOSS*=$5 in Example 1

Optimal Cost-Optimized Results: Utilizing Equations 6 to 9, and 10 to 23, software solutions to the unknown vector, [P_{ij}] = [P_{11}=*.0062499978*, P_{21}=*.07142863*,

P_{12}=.02499999, P_{22}=.89732146]; one derives, $\alpha = P_{11} + P_{12} = .0062499978 + .02499999 = 0.03124499$, and $\beta = P_{11} + P_{21} = .0062499978+.07142863 = 0.07767862$. For a given small n (sample size from batches) =100, σ (standard deviation) = 9; the optimized α =0.03124999 (or ≈3.13%) and β = 0.07767862 (or ≈7.77%), as computed from the input data in Table 5 demonstrates the following plan in a software-enabled Figure 3, inspired by Figures 1 and 2. They are, $0.03124999 = \alpha = P(Z \geq Z_C \mid H_0: \mu_0 = 5)$ and $0.07767862 = \beta = P(Z \leq Z_C \mid H_1: \mu = \mu_1)$ which yields: $(Z_C \mid H_0 : \mu_0 = 5) = 1.86$ and $(Z_C \mid H_1: \mu = \mu_1) = -1.42$.

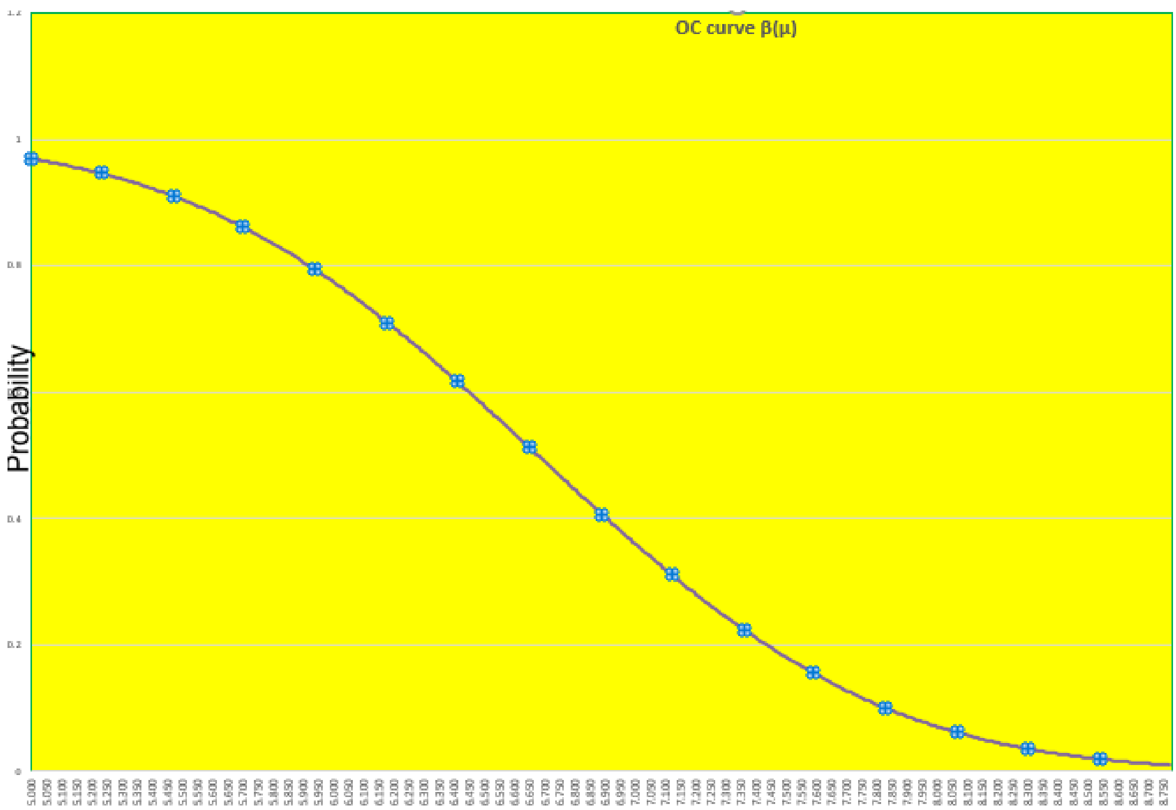
The preceding calculations will further result, assuming the standard deviation to be σ=9, in the critical value of C. That is, $C = 5+1.86*9/\sqrt{100}=6.674$ under H_0, and hence leading to $\mu_1 = 6.674+1.42* 9/\sqrt{100}=7.95$ under H_1. Note, $H_1: \mu_1 = 7.95$ is formulated by $C-Z(\beta)*\sigma/\sqrt{n} = 6.674 - (-1.42) *(9/10) = 7.95$. Hence, one is testing $H_0: \mu_0 = 5$ cm. vs $H_1: \mu_1 = 7.95$ cm. for the IC mean diameter length. Therefore, the decision plan becomes as in Figure 3 followed by its *OC* Curve in Figures 4. A. and B. Therefore, reject H_o if $\bar{x}$ (sample mean) > $C \approx 6.67$ when $H_0: \mu = 5$ to commit Type-I error, and fail to Reject H_o when $\bar{x} < C \approx 6.67$. This implies under $H_1: \mu = 7.95$ to commit Type-II error to attain optimal outcomes as dictated by the *Acceptance Sampling* Plan's OC Curve designed in Figures 3, and 4. A. and B. Therefore, executing input Table 5 subject to C_{11} = \$800K (unit cost incurred), C_{12} = \$70K (unit cost incurred), C_{21} = \$200K (unit cost incurred) and C_{22} = -\$400K (unit utility credited) under the *LOSS* constraint (21), the overall process cost without the "*K*" multiplier in Figure 2 follows as in (24):

$$\mathbf{\Sigma\{P_{ij}C_{ij}\} = .006245*800 + .07143*70 + .0245*200 + .897321*(-400) = -\$343.93} \quad (24)$$

Thus, the total negative cost (utility) that the planner is expected to profit is ≈ -\$344 given *LOSS* (max) limited to \$5. *LOSS* varies for sensitivity from \$1 to \$60 in Figures 1 and 2.

	Significance	Level	α =	0.0313				
			Z(α) =	1.86				
Defective	Proportion	p =	0.05	Std Dev(Def)	σ =	2.179449472		
Sample Size (n)		=	100					
Expected Value		(=np)	5			ATTRIBUTES→		
Standard Deviation (σ)		=	9			H1 : µ =	µ1 > µ0	9.76881825
						VARIABLES→		
	Ho :	µ =	µ0 =	5		H1 : µ =	µ1 > µ0	7.952
		C	=	6.674				
	Z(β) =	-1.42						
	β =	0.0777						

FIGURE 3 Optimal Plan with α, β, from Tables 2-5 of Example 1, n=100, σ=9, C=6.67 for both i) Variables and ii) Attributes where $np_0 \geq 5$ with Binomial to Normal approximation



µ1		OC(µ1)
AQL=5.00	Z(α) = 1.86	0.9687(=1-α=1-0.0313)
LQL =7.95	Z(β) = -1.42	0.0777(=β)

FIGURES 4 A and B Example 1's OC curve with AQL & LQL by Tables 1-5, Figures 1, 2 and 3 showing results for n (small sample size) = 100 and σ (standard deviation) = 9

Conclusive Outcomes: In Tables 1-5, Figures 1, 2, 3 and 4 A. and B., observe H_o: $\mu = 5$ (≈*AQL*) vs H_1: $\mu_0 = 7.952$ (≈*RQL*). Therefore, the *OC curve* plots H_1: $\mu_1 = 7.95$ on the x-axis with its y-axis ≈ 0.078 vs H_0: $\mu_1 = 5.0$ on the x-axis with its y-axis ≈ 1 - 0.0313 = 0.9687. Note, *RQL:* Rejectable Quality Level, whereas *AQL*: Acceptable Quality Level. Unfortunately, the ideal OC Curve can almost never be obtained in practice; whereas in theory, it could be realized by 100% inspection if the inspection were error free. This implies that the ideal OC curve can be approached by increasing the small sample size = $n \rightarrow N$ = batch (or lot) population size (Montgomery[20], 2009; pp. 631-642, pp. 670-676). For attributes, AQL=5 defectives vs RQL=9.77≈ 10 defectives from Figure 3 for large n.

4. STEPS FOR INPUT DATA COLLECTION FOR COST PARAMETERS

How to obtain the input data: C_{11}, C_{12}, C_{21}, C_{22}, and *LOSS* from the corporate world's actual sales data recording with hypothetical examples keeping the input Table 5 at sight:

C_{11}: Incurred losses due to penalty arising from consumer's and producer's risk combined on the same product (a PC or an automobile or a washing machine, etc.) while the product sold was recalled or returned due to a faulty production or misconception of a defective product. Say e.g. \$800K cost of penalty due to recall/returns (unwarranted) recorded.

C_{12}: Incurred losses due to penalty arising from producer's risk while the product sold was returned due to a misconception of a hypothetical fault, knowing it truly was not a faulty product. Say e.g. \$70K cost of due to incorrect or unsubstantiated returns recorded.

C_{21}: Incurred of losses due to penalty arising from consumer's risk while the product sold was recalled or returned due to a defective production. Say e.g. \$200K cost of penalty due to recall or returns officially recorded.

C_{22}: Credited revenues due to entirety of sold products uncontested, i.e. not being returned. Say e.g. \$400K (taken -400K as utility, not cost-incurred or penalized, but a credit or revenue). These data may dictate that it is not a sound company with cash issues.

LOSS: If the minimum *LOSS* assumed is $-LOSS \leq -\$5K$ or $LOSS \geq \$5K$; then this is the tolerable, or company-paid indemnity to circumvent or intercept the damage incurred after the deductibles due to each of the three risk related constraints calculated for each of the three cubicles which in Table 5 is \$5K. Each term in the constraints of Equations 10 to 13, $P_{ij}*C_{ij}<LOSS$ for i,j=1,2 where these four constraints are bound not to exceed $\$LOSS=\$5K$, including the last constraint whose negative cost value (meaning revenue) obeys the constraint. The *\$LOSS,* is a company-paid compensation after the deductibles.

In every statistical experiment, the statement of the problem goal is the first step. The goal for input data is to estimate the C_{ij}, i,j=1,2 from the company's historical accounting data. *LOSS* parameter to be estimated (after the deductions) is a company policy, and therefore, it is a constant to be dictated by the associated company proper. Choice of response or dependent variable comes next, which in this case is the C_{ij} and *LOSS*. The goal is to first elucidate what the target parameters mean in the business (or accounting) world and then lead the way to how to collect data regarding these unknown parameters.

To determine how large a sample is needed, there are three questions to answer by Ostle and Mensing[15] (1975): i) How large a shift (i.e. range) in the parameter do you expect to detect? ii) Based on experience, how much variability do you wish to detect? iii) What sizes of risk (i.e. *Alpha and Beta*) you are willing to take. Note that these risks are what we are going to optimize, therefore we cannot flat assume them. A composite metric that considers these queries can be found in the CV: Coefficient of Variation= (Sample Standard Deviation / Sample Mean). The CV is an ideal device for comparing the variation in two "Cause & Effect" series of data that are measured in two different units, e.g. a comparison of variation in height with a variation in weight (Hicks[19], 1973). Hence for sample size (n), we can a tabulate a list of *CV*100%* vs *n* to justify where to stop. The lesser the CV, the better the outcomes stand to halt sampling.

5. COST PARAMETERS C_{11}, C_{12}, C_{21} AND C_{22} IN THE BUSINESS WORLD

Many of the larger merchandisers will break the returns down into four distinct groups by http://businessecon.org/2014/12/returns-allowances-and-discounts-in-accounting/ [12/24/2014]:

A) The first group reflects customer-based mistakes. It is solely a customer's mistake. This is therefore attributed to the producer's risk indicating the cost associated with C_{12}. As a merchandiser you would only need to monitor the growth rate for this group. If this ratio begins to increase, it might be a sign that the sales staff is forcing the wrong product onto the market and hence, the customer.

B) The other form of a return is merchandise that is broken or has a warranty issue. It is generally important to track this information as this form of return can be a clear sign of a quality issue with a particular brand or product line. That is, it's not the customer's mistake but that of the producer. This is therefore attributed to the consumer's risk indicating the cost associated with C_{21}. If the issue is brand-related, the merchant may want to consider discontinuing the brand or substituting the brand with a higher quality product. Most popular examples are large recalls from the automotive industry.

C) The two adjustments above in the business world are followed by another described as Allowances (or Incentives). These are adjustments to normal sales reflecting defective items or courtesy calls for failure in delivering the product or service in a timely fashion (merchandise fault). Then, re-education of the sales representatives is required if customers' erroneous returns increase as this relates to the wrong kind of purchase, or consumers are not educated in what they bought assuming that the product is defective while it truly is not; thus, leading to a customer fault. The combination is attributed to the intersection of the producer's and consumer's risk to define the cost associated with C_{11}.

D) *C22* is the uncontested revenue (credit or utility) not returned with a 100% customer satisfaction.

In Figure 5, Venn Diagram constituting all four sample elements (sets) of V are indicated. Note, $[(1-\alpha)*(1-\beta)]$, $\alpha*\beta + \alpha*(1-\beta) + (1-\alpha)*\beta + (1-\alpha)*(1-\beta)] = 1.0$ per constraint (22).

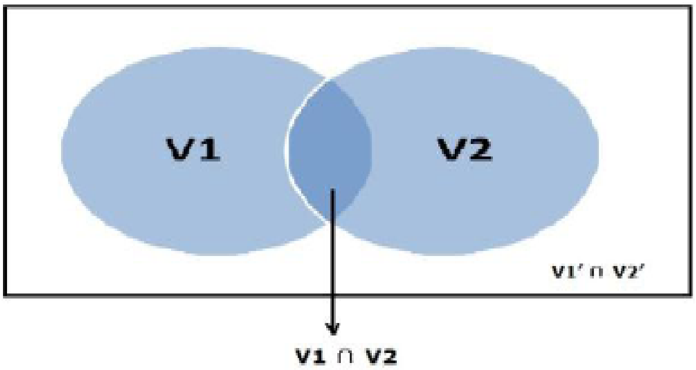


FIGURE 5 Take V_1 = Producer's Risk, V_2 = Consumer's Risk, $V_1' = V_1$ complement and $V_2' = V_2$ complement, where P $(V_1 \cap V_2')$ = Producer's Risk only by $\alpha*(1-\beta)$; P $(V_2 \cap V_1')$ = Consumer's Risk only by $(1-\alpha)*\beta$; P $(V_1 \cap V_2)$ = Intersection of both risks by $\alpha*\beta$, and lastly P $(V_1' \cap V_2')$ = No producer's and/or consumer's risks

Example 2: On how to collect the C_{ij} costs and *LOSS* limitation cost for an Acceptance Sampling plan by another company for the same hypothesis test H_o vs H_1 for Example 1:

A hypothetical *Large Automobile Production (LAP)* plant income statement for the recent year ending where company discounts excluded after 15 years of data collection:

Total Sales			$1,000,000
No-Need for Adjustments (Smooth Sale)			$800,000
Adjustments:			
Returns	$150,000		
Customer Based		$110,000	
Merchandise Based		$40,000	
Allowances	$50,000		

Based on this tabulation, one seeks what kind of an OC Curve to facilitate an Acceptance Sampling plan that the *LAP* plant is at best to undertake. The data as follow are the without *K* values for the following designed example of input Table 6, Table 7 and Table 8, Figures 6 and 7. The analyst may plan to explore e.g. 15 such large companies with their proper accounting history back to year 2000. Note these figures are updated each year. C_{ij} are: C_{22} = -\$800; C_{12}= \$110; C_{21}= \$40 and C_{11}= \$50 recorded, as company-specific input in Table 6 without the *K*. The *LAP* plant has selected *LOSS* (after deductibles) =\$5.

TABLE 6 Input cost values for the associated JAVA-coded software in Example 2

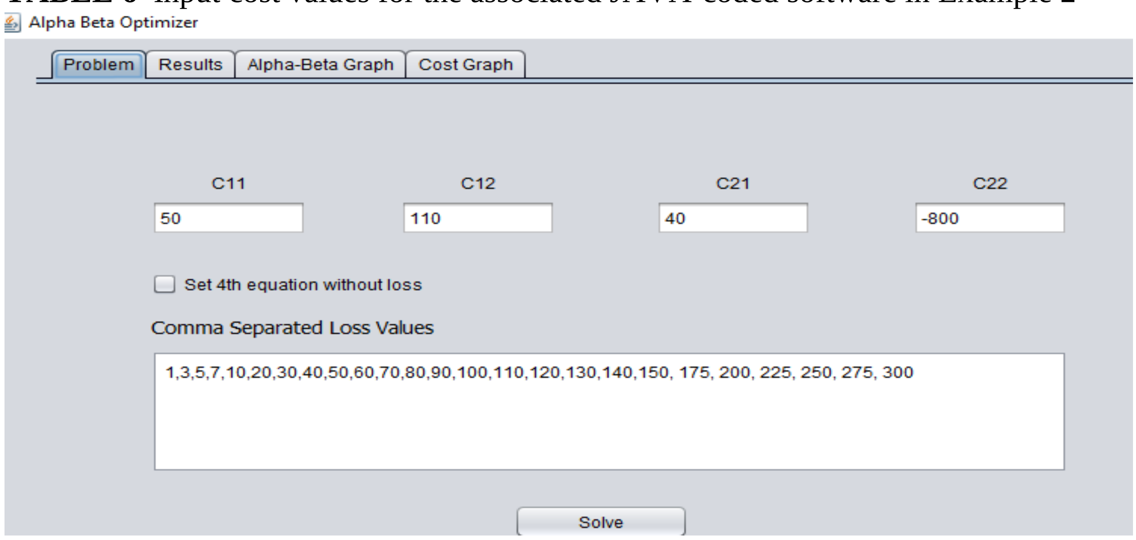


TABLE 7 Game theoretic input spreadsheet for Example 2

Enter/Edit data: Objective function coefficients. For each constraint, enter constraint coefficients, constraint relationship [<, =, >], and constraint right-hand-side value. Do not enter nonnegativity constraints.

Optimization Type: Minimize

Variable Names: (Change if Desired)	P11	P21	P12	P22	LOSS
Objective Function Coefficients:					1

	Coefficients				
Subject To:	P11	P21	P12	P22	LOSS
Constraint 1	1				
Constraint 2		1			
Constraint 3			1		
Constraint 4				1	
Constraint 5	1	1	1	1	
Constraint 6	50				-1
Constraint 7		40			-1
Constraint 8			110		-1
Constraint 9				-800	-1
Constraint 10					1
Constraint 11	50	40	110	-800	
Constraint 12	-1			1	
Constraint 13		-1		1	
Constraint 14			-1	1	

TABLE 8 Feasible vector solution for Example 2

Optimal Solution

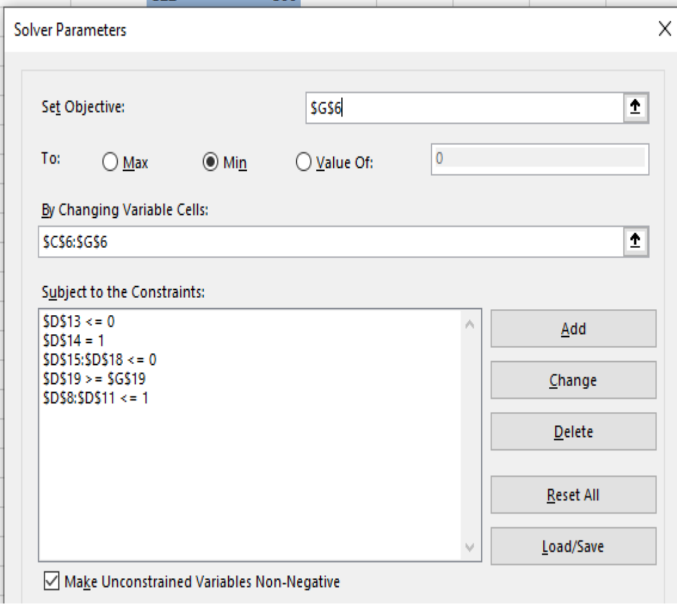
Objective Function Value = 5.000

Variable	Value
P11	0.100
P21	0.125
P12	0.045
P22	0.730
LOSS	5.000

TABLE 9 Input $\{C_{ij}, i,j=1,2\}$ and output EXCEL solution of $\{P_{ij}, i,j=1,2\}$ for Example 2

C11	50
C21	110
C12	40
C22	-800

MIN LOSS				
P11	P21	P12	P22	LOSS
0.1	0.045454545	0.125	0.729546	5
P11	0.1000000000		<	1
P21	0.045454545		<	1
P12	0.125		<	1
P22	0.729546455		<	1
Constraint 1	-568.6371636		<	0
Constraint 2	1.00000		equal	1
Constraint 3	0		<	0
Constraint 4	0		<	0
Constraint 5	0		<	0
Constraint 6	-588.6371636		<	0
Constraint 7	5		>	5



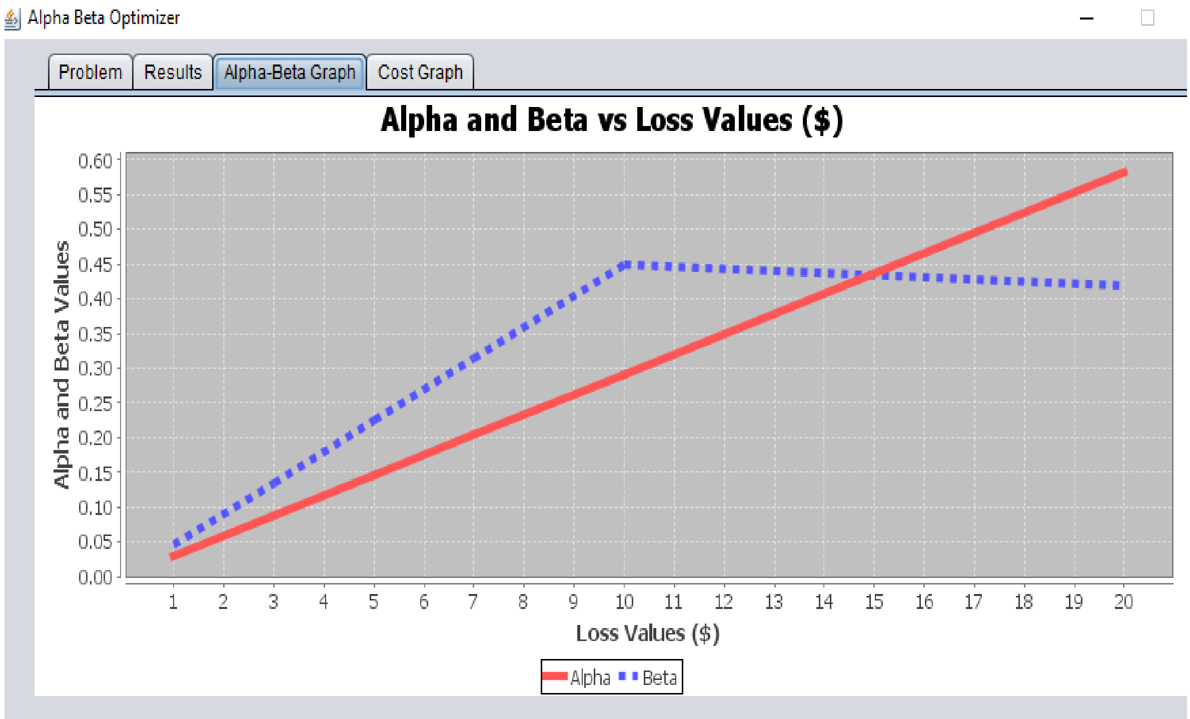


FIGURE 6 LAP Company's *Alpha*=0.145 (14.5%) and *Beta*=0.225 (22.5%) for *LOSS*=$5

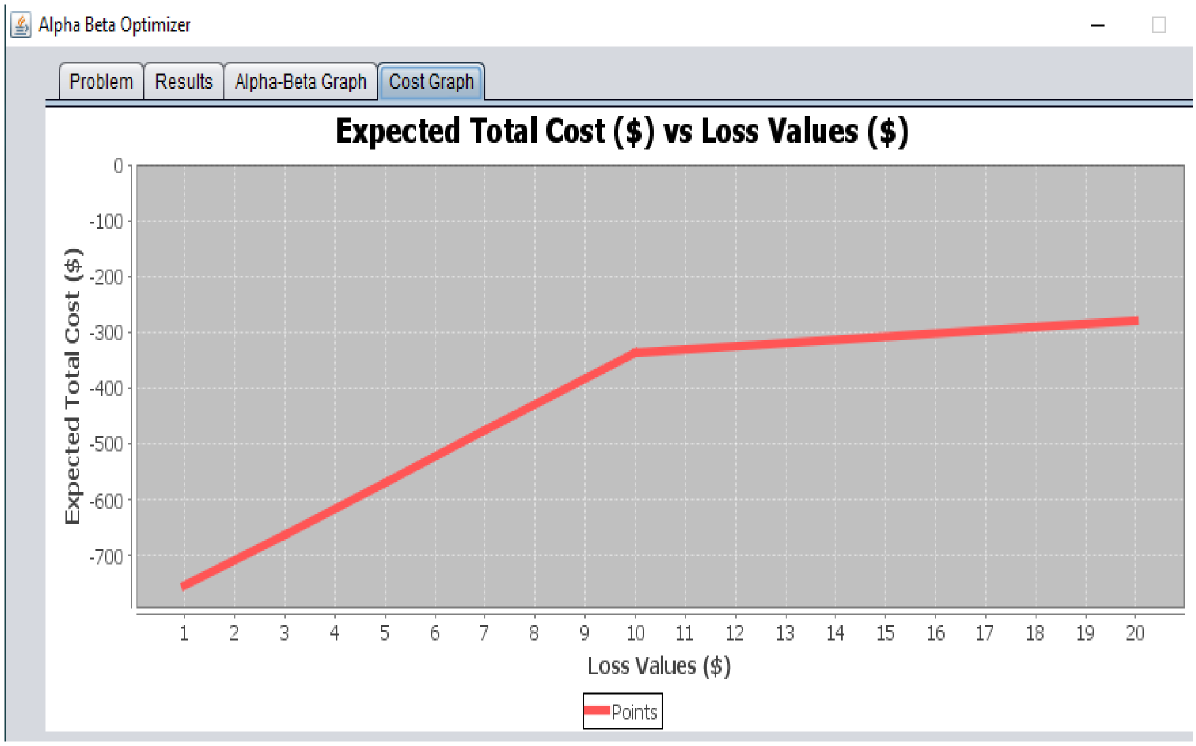
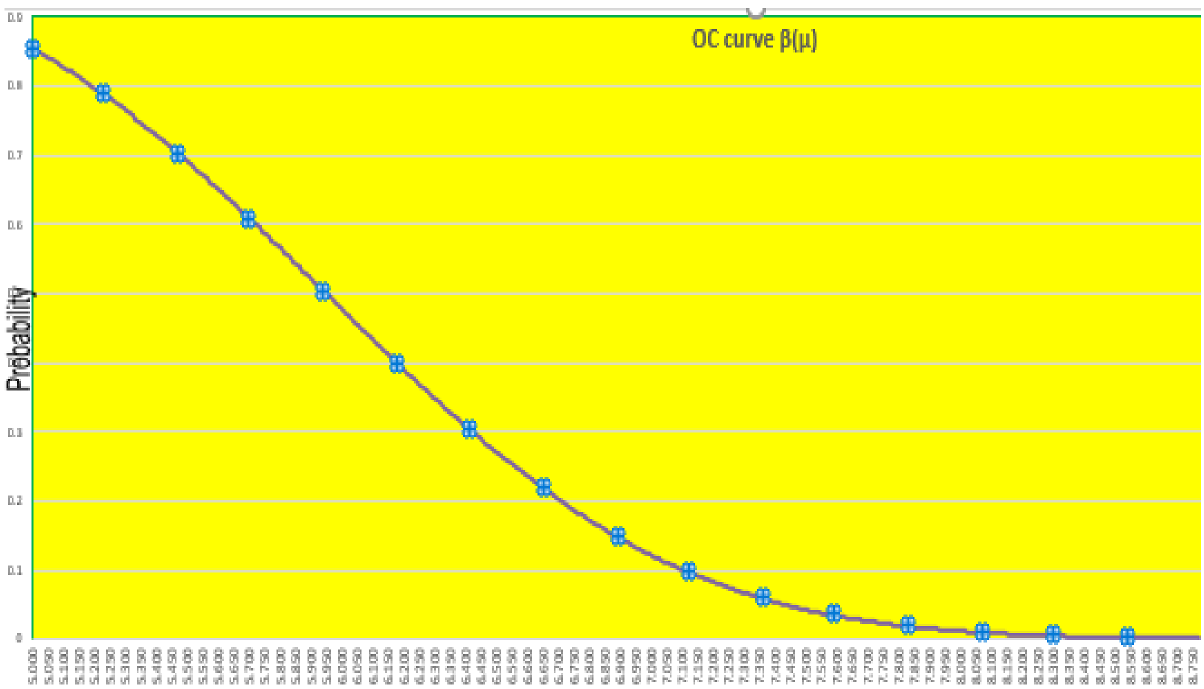


FIGURE 7 LAP Company's Expected Total Cost ≈ -$569 for *LOSS*=$5

The results for the horizontal axis of *LOSS*: $5 based on Table 6, Table 7, Table 8, and Table 9 as well as Figures 6 and 7: P_{11} = 0.1, P_{12} = 0.0454, P_{21} = 0.125, P_{22} = 0.7295 *Alpha*: 0.145, *Beta*: 0.225 and Expected Total Cost: -$568.63

	Significance	Level	α =	0.145				
			Z(α) =	1.05				
Defective	Proportion	p =	0.05	Std Dev(Def)	σ =	2.179449472		
Sample Size (n)		=	100					
Expected Value		(=np)	5			ATTRIBUTES→		
Standard Deviation (σ)		=	9			H1 : μ =	μ1 > μ0	7.579587104
						VARIABLES→		
	Ho :	μ =	μ0 =	5		H1 : μ =	μ1 > μ0	6.62
		C	=	5.945				
	Z(β) =	-0.75						
	β =	0.225						

FIGURE 8 Optimal plan with α, β from Tables 6-9 in Example 2 for n=100, σ=9, C=5.95 for both i) Variables and ii) Attributes where $np_0 \geq 5$ for Binomial to Normal approximation



μ1		OC(μ1)
AQL= 5.00	Z(α) = 1.05	0.855(=1-α=1-0.145)
RQL = 6.62	Z(β) = -0.75	0.225(=β)

FIGURES 9 A and B Example 2's OC Curve with *AQL* and *RQL* by Table 6 and Figures 6, 7 and 8 for n=100, σ=9; *RQL:* Rejectable Quality Level, *AQL*: Acceptable Quality Level

Conclusive Outcomes: In Tables 1 and 5-7, Figures 6, 7, 8 and 9 A. and B; Observe H_o: μ_0 = 5 (≈*AQL*) vs H_1: μ_1 = 6.62 (≈*RQL*). The *OC* curve plots H_1: μ_1 =6.62 on the x-axis with its y-axis = 0.225 and H_0: μ_0 =5.0 on x-axis with its y-axis=1-0.145 = 0.855. For attributes, AQL=5 defectives vs RQL=7.58≈8 defectives from Figure 8 for large n.

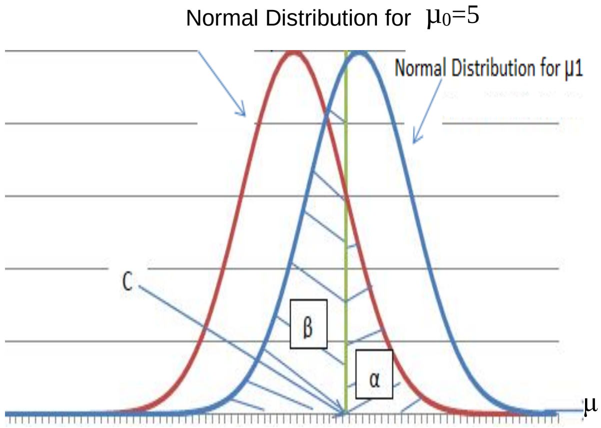


FIGURE 10 Generalized Graphical Illustration of the Acceptance and Rejection Regions, and Type-I and Type-II error probabilities, α (hatched ascending) and β (hatched descending) for the One-Sided Hypothesis Tests of Example 1 for H_o: μ_0 or np_0=5 vs H_1: μ_1=7.95 or np_1≈9.77=10 with α = .0313, β = .077, *C* = 6.67, and of Example 2 for H_o: μ_0 or np_0=5 vs H_1: μ_1=6.62 or np_1≈7.58=8 with α = .145, β = .225, *C* = 5.95.

Conclusive outcomes on Examples 1 and 2 can be illustrated respectively in Figure 10. The LAP in Example 2, if not satisfied with this plan, will for the next year try to reduce the customer-based C_{12}=\$110*K* by re-educating the customer-base and increasing the smooth-sale margin to experience less adjustments. Similar remedies for C_{21}=\$40*K* and C_{11}=\$50*K* can be undertaken to improve the cost-free C_{22}=\$800*K*. The Data Management program may have to define two sets of large company-based data contributors. The first group are the ones that have just begun to provide data (yet to meet the 15-year threshold) and the second group are those (having met the threshold) which are using the algorithms on the historic data, and they could agree to report the results on their current operations and change data. Without the incentives to participate, what would be the motivating factors for companies to join in the effort and publish information that (eventually) their competitors will see? Those companies will act with their company-specific (unique only to them proper) or identifying batch (or lot) Acceptance Sampling plans, rather than an unjustified and subjective assumption of the *Type-I* and *Type-II* error probabilities. This article suggests a Game theoretic predictive solution provided the cost parameters rather than to adopt haphazardly producer's and consumer's risks. It is critical to underscore that each enterprise will monitor to compute its own set of risks for their proper Acceptance Sampling plans to improve quality control inspections and thus raise productivity.

6. CONCLUSIONS

An innovative, Game theoretic, business-savvy and market-centric method with optimal algorithmic solution for *Type-I* and *Type-II* error probabilities (also known as *Producer's* and *Consumer's risks, or Alpha error,* and *Beta* error) is proposed in contrast to merely selecting these parameters using subjective judgment calls conventionally practiced as Kelley[1] (2013) so remarked. One labels the past practices as habitual guess-work procedures devoid of cost or utility constraints in a business-plan state-of-mind. It falls upon the author to further state that the most challenging task in this Game theoretic proposition is to generate the most-fitting rightful and authentic market-centric input data for the firmware, cyber-ware or any other commodity-based market about which the tests of hypotheses are being conducted. This will necessitate another series of "econometric" data collection study to generate the most compatible input data sets for the challenging problem proposed in this research so as to give life and meaning to the cost factors explained. The designed *OC* Curve depending upon choice, can be a business standard for a new enterprise's market-entry acceptance, such as a smart phone company in quest for an opening e,g, in USA. See Tables 1-4, Figures 1, 2, 3 and 4. A. and B., and similarly, Tables 1 and 5, Figures 6, 7, 8 and 9. A. and B. for Examples 1 and 2 respectively. The goal is to follow up with a business plan, e.g. *MIL-STD-105E* found at https://variation.com/wp-content/uploads/standards/mil-std-105e.pdf [10 May 1989] that uses popular standards such as α=0.05 and β=0.10. For more application-oriented details, see by the Sahinoglu et al.[9,10,11] (2015, 2016, 2017). In this approach to calculate the cost-optimized Type-I and II error probabilities, the author follows a Game theoretic algorithm where the probabilistic and cost-related constraints as well as the five input parameters (C_{ij} and *LOSS*) must be incorporated by the analyst to reflect the market conditions for a profitable business model. The solutions follow with Figure 10 in sight:

In *Example 1* with C_{ij} = [$800, $70, $200, -$400] from the input Tables 1-5, Figures 1, 2 and 3, and Figures 4. A. and B. and given *LOSS* = $5, the Game theoretic algorithm generates α = 0.0313 and β = 0.0777 resulting in an Expected Total Cost: -$343.92 (utility) using the company-specific input parameters. Final solutions are multiplied by K=1,000. See Figure 10 to place the constants on a plot with an illustration.

In *Example 2* from the input Tables 1 and 6, 7, and Figures 6, 7, and 8, and Figures 9. A. and B. with C_{ij} = [$50, $110, $40, -$800], the Game theoretic algorithm results in an Expected Total Cost: -$568.64 (utility) with the company-specific optimal α = 0.0145 and β = 0.0225 for given *LOSS*=$5. Final solutions are multiplied by K=1,000. See Figure 10.

The producer establishes a sampling plan for a continued supply of components with reference to *AQL*, which represents the acceptable level of quality for the supplier's process that the consumer would consider acceptable as a process average. The consumer may also be interested in the other end of the *OC* Curve, i.e. *RQL* or *LQL* (*Rejectable or Limiting Quality Level*), as the poorest level of quality that the consumer is willing to accept with a low probability of acceptance in an individual lot (Montgomery[20], 2009).

In summary, the publications by Sahinoglu et al.[9,10,11] (2015, 2016, 2017) significantly improved material on this topic. Therefore, it is emphasized that the business entities were traditionally not able to design their company-specific quality control goals by computing and imbedding their own Producer’s and Consumer’s Risk into their *Acceptance Sampling* plans. This was done by randomly assuming or best-guessing a Type-I error probability and continue doing a sensitivity analysis. Now they can include and benefit by quality-managing their process from the very first step without expecting to be a given any Type-I error probability. Utilizing the Game theoretic results, they can invest smarter mindfully as opposed to practicing the conventional with a subjective state of mind and therefore appreciate the meaningful quote by Kelley[1] (2013). The author believes that this technique is appropriate for pragmatic uses when batch- or lot-sampling regarding *Acceptance Sampling* Plans, because the article relates to the statistical computing and numerical optimization of hypothesis-testing parameters (α, β), including an empirical, substantively data-scientific and user-friendly albeit business-savvy application. Lastly, it would be surprising if any one theory could address such an enormous range of “games,” and in fact there is no single game theory. Several theories have been proposed, each applicable to different situations and each with its own concepts of a solution by Davis[21] (1997). The reader is recommended to refer to Figure 5 which clarifies the Venn Diagram’s sample sets for the proposed gaming optimization. The reader is strongly recommended to refer to Figure 10 that graphically illustrates and summarizes the nature of the Figures 3 and 8 for Examples 1 and 2 in a composite diagram. Last but not least, Figures 3 and 8 display tests with for i) Variables (continuous measurement), and ii) Attributes (defective counts), where $np \geq 5$ with Binomial to Normal approximation for relatively large sample size such as taken to be $n=100$ (Sahinoglu[10], 2016, pp. 20-21; Ostle and Mensing[15], 1975, pp. 84-85) in Examples 1 and 2 of Sections 3 and 5 in order.

REFERENCES

1. Kelley M. Emily Dickinson and monkeys on the stair, or: what is the significance of the 5% significance level? *ASA & RSS Significance;* **10**(5). 2013; 21-22.

2. Grant BJB. Should have been 8%, not 5%? *ASA & RSS Significance;* 2014; **11**(5); 85.

3. Sahinoglu M, Cueva-Parra L, Ang D. Game-theoretic computing in risk analysis, *WIREs Comput. Stat 2012;* **227–248**. DOI: 10.1002/wics; 1205.

4. Blackwell D, Girshick MA. *Theory of Games and Statistical Decisions.* Dover Publications, Inc. New York, 1954.

5. Bagchi TP. Acceptance Sampling, *Video #18, Indian Institute of Technology, Kharagpur, India, 2012*. https://www.youtube.com/watch?v=xJ3czkvNxpk/ [2 January 2012].

6. Schlag K. Bringing Game Theory to Hypothesis Testing; Establishing Finite Sample Bounds on Inference. *Collection of Biostatistics Research Archive,* COBRA Reprint Series, Paper 59, 2008;1-26.

7. Nash JF. Equilibrium points in N-Person Games. *Proceedings of the National Academy of Sciences of the United States of America.* **36**(1), 1950; 48-49.

8. Osborne MJ, Rubinstein A. *A Course in Game Theory.* Cambridge, MA, MIT Press, 1994.

9. Sahinoglu M, Balasuriya R, Tyson D. Game-Theoretic Decision Making for Type I and II Errors in Testing Hypotheses. *Proceedings of the JSM,* Seattle, 2015; 2976-2990.

10. Sahinoglu M. *Cyber-Risk Informatics - Engineering Evaluation with Data Science.* John Wiley and Sons, Hoboken, New Jersey, 2016.

11. Sahinoglu M, Balasuriya R, Capar S. Selecting Type-I and Type-II Error Probabilities in Hypothesis Testing with Game Theory. *61^{st} ISI World Statistics Congress, Marrakech-Morocco, Abstract Book and B08 (Methods and Theory 08)*, **18**, 2017; 68.

12. Savage LJ. *The Foundations of Statistics.* John Wiley & Sons, New York, 1954.

13. Neumann JV. Zur Theorie der Gesellschaftsspiele. *Math. Ann.* **100**, 1928; 295-320.

14. Neumann JV, Morgenstern O. T*heory of Games and Economic Behaviour.* Princeton University Press, New Jersey, 1944.

15. Ostle R, Mensing RW. *Statistics in Research.* Iowa State University Press, Ames, Iowa, 1975.

16. Hogg RV, Ledolter J. *Applied Statistics for Engineers and Physical Scientists.* Macmillan Publishing Company, New York, 1992.

17. Rapcsak T. *Smooth Nonlinear Optimization in R^n :* Springer Science, New York, 1997.

18. Lewis C. *Linear Programming: Theory and Applications.* Math Free PDF Books, 2008.

19. Hicks C. *Fundamental Concepts in the Design of Experiments.* Holt, Rinehart and Winston, Inc., New York, 1973.

20. Montgomery DC. *Introduction to Statistical Quality Control.* John Wiley & Sons, 6^{th} ed. Hoboken, New Jersey, 2009.

21. Davis M.D. *Game Theory - A Nontechnical Introduction.* Dover Publications, Inc. New York, 1997.

BIOSKETCH

Mehmet Sahinoglu, Prof. Emer., Auburn University and Adjunct, CS, Troy University

mesa@aum.edu, mesa@troy.edu

M. Sahinoglu served as the founding Director of the Informatics Institute and the founder Chairperson of the Cybersystems and Information Security Graduate Program (2008-2018) at Auburn University at Montgomery (AUM). Formerly, the Eminent Scholar and Chair-Professor at Troy University's Computer Science Department (1999-2008), he holds a B.S. from METU, Turkey (1969-1973), and M.S. from University of Manchester, UK (1974-1975), both in Electrical Engineering. He completed a joint Ph.D. in Statistics (major) and ECE (minor) from Texas A&M University, College Station, TX (1977-1981). Earlier in his career in 1992-97, he became the founder dean of the College of Arts & Sciences at Izmir's DEU (50K-strong medical university) and consequently founded the Applied Statistics Dept in 1993-1994 which celebrated its 25th anniversary in 2019. He currently conducts research in the domain of Cyber-Risk Informatics: Quantitative Risk Assesment and Management. He is the author of Trustworthy Computing (2007) and Cyber-Risk Informatics: Engineering Evaluation with Data Science (2016) both published by Wiley Inc. Dr. Sahinoglu, a Professor Emeritus from Auburn University System as of June 1, 2018 and currently an Adjunct faculty, teaching Cybersecurity-themed classes at Troy University in Troy AL. Dr. Sahinoglu is an 40-year ASA (1980-), ISI Elected (1995-), IEEE Senior Life (1978-) and SDPS: Society of Design and Process Science Fellow Member (2003-). One of the 14 Microsoft Trustworthy Computing Awardees (2006), and a silver medalist for the DAU's Hirsch Paper Competition on Sofware Assurance (2015) and Digital Forensics (2016); Mehmet is the originator of the i) S&L: Sahinoglu-Libby pdf (1981) jointly with D. Libby from the University of Iowa for repairable hardware, ii) CPSRM: Compound Poisson Software Reliability Model (1992), iii) MESAT: Stopping-Rule in Software Reliability Testing (2002), iv) SecRM: Security/ Privacy Risk Meter (2003), and v) CLOURAM: Cloud Computing Risk Assessor & Manager (2017).