國 立 交 通 大 學

統 計 學 研 究 所

碩 士 論 文

貨批中有γ比例良品其信心之統計推論

Statistical Inferences for Confidence of Percentage γ

Acceptable Products in Lot

研 究 生：蕭夙吟 (Su-Yin Hsiao)

指導教授：陳鄰安 博士 (Dr. Lin-An Chen)

中華民國九十七年六月

貨批中有γ比例良品其信心之統計推論

Statistical Inferences for Confidence of Percentage γ Acceptable Products in Lot

研 究 生：蕭夙吟 Student：Su-Yin Hsiao

指導教授：陳鄰安 博士 Advisor：Dr. Lin-An Chen

國 立 交 通 大 學

統計學研究所

碩 士 論 文

A Thesis
Submitted to Institute of Statistics
College of Science
National Chiao Tung University
In Partial Fulfillment of the Requirements
For the Degree of
Master
In
Statistics
June 2008

Hsinchu, Taiwan, Republic of China

中華民國九十七年六月

貨批中有γ比例良品其信心之統計推論

研 究 生：蕭夙吟　　　　　　　　　　指導教授：陳鄰安 博士

國立交通大學統計學研究所　碩士班

摘　　要

容忍區間經常被用來研究一批貨物中具有γ比例良品的信心是否達到特定水準。此篇文章說明此信心實際上是一個未知的參數，並且證明一般所使用的Eisenhart et al.(1941)所提出之最短容忍區間，其所宣稱之信心是不適切的。我們將會提出一個更適切的檢定方法來檢定實際的信心水準。最後將會根據新方法建立或決定樣本數的大小，以期當實際信心高於或是低於預期的時候，能夠平衡生產者的風險及利益。

Statistical Inferences for Confidence of Percentage γ Acceptable Products in Lot

Student:Su-Yin Hsiao   Advisor:Dr. Lin-An Chen

Institute of Statistics

National Chiao Tung University

Abstract

The tolerance interval is often used to investigate if there is γ percentage of acceptable products in a lot at some desired confidence. This paper shows that this confidence, with percentage γ fixed, is actually an unknown parameter and shows the popularly used shortest version of tolerance interval by Eisenhart et al. (1941) is not capable to serve as a test statistic for hypothesis assuming the unknown confidence to be a desired constant q_0. A new test is shown to be more capable in this purpose. The sample size determination based on this new test ensuring to protect the manufacturer's benefits and risks when the specification limits indicate true confidence well, respectively, above and below q_0 has been studied.

誌　　謝

兩年的研究生生活在老師、同學還有家人的陪伴下即將邁入尾聲了，很開心自己能夠成為統研所的一份子，在這一個大家庭裡，受到老師的指導而成長，認識了互相扶持的好朋友，一路走來充滿歡樂、關愛與祝福，讓我的生命充滿感動與幸福的樂章。

首先要謝謝我的指導教授陳鄰安博士，非常感謝老師對我的指導與包容，謝謝老師帶領我作研究，耐心的指導與鼓勵，讓我論文能夠如期完成。老師是如此的慈愛和藹，總是關心著我們的生活，對於茫茫的未來，給予我很多的方向；而與老師每次的談話，也體驗到老師傳遞給我們人生的態度與價值觀。還要感謝所辦郭姐與小陳哥給我的協助，以及口試委員許文郁教授、黃榮臣教授、彭南夫教授給我的指導與建議。還要感謝我的母親、父親以及妹妹，謝謝父母的支持與栽培，讓我能夠順利完成學業，而每當面對挫折的時候，母親與妹妹的打氣是我面對挑戰的支持動力。

最後感謝我的同學、朋友們，佩芳、姿蒨、珮琦，謝謝妳們這兩年的陪伴，一起努力討論功課，一起努力作論文，一起吃喝玩樂、一起抱怨，給我最大的支持與鼓勵。鳳如、香菱、采玲、鈺婷、郃嵐、瑜達，謝謝這些日子以來給我的加油打氣與幫忙。文廷、泰佐、翁賢、小賴學長在就算在自己也很忙碌的時候，總是願意陪我一起討論問題，給我很大的幫助。感謝昱緯、仲竹、政言、重耕、鈺玲、筱娟、怡玲、佩蓁一起奮鬥拼畢業，對我的包容與照顧。宗晟、宏瀚、勁葦、泓毅，曉怡、小方、佩玲、琳婷、大雄、博全、旻原、心瑩妹妹及心瑩一家人、瀅竹、瓊緯、育珊，謝謝你們一路上給我加油打氣。衷心感謝曾經幫助過我的人，此篇文章謹獻給我親愛的家人、師長與同學，並致上我最深的謝意。

蕭夙吟　謹誌于
國立交通大學統計研究所
中華民國九十六年六月

Contents

1. Introduction

The tolerance interval is popularly used by manufacturer and consumer for judgement of production lots. In mass-production, the manufacturer is interesting in an interval that contains a specified (usually large) percentage of the product and he knows that unless a fixed proportion (say γ) of the production is acceptable in the sense that the items' characteristics conform to specification limits LSL and USL, he will lose money in this production. On the other hand, if this claim is not true, it will entail the consumer a loss of money. With this interest, the manufacturer and consumer want to know the following:

$$\textbf{Whether there is percentage } \gamma \textbf{ of acceptable measurements in a production lot}? \tag{1.1}$$

Statisticians seek solution to verify this problem through two steps. We start with a random sample $X = (X_1, ..., X_n)'$ from a distribution with probability density function $f_\theta(x)$ representing observations from the same production process. For the first step, the pioneer article Wilks (1941) introduced a γ-content tolerance interval with confidence $1 - \alpha$ which is defined as a random interval $(T_1, T_2) = (t_1(X), t_2(X))$ satisfying

$$P_\theta\{P_\theta[X_0 \in (T_1, T_2)|X] \geq \gamma\} \geq 1 - \alpha \text{ for } \theta \in \Theta \tag{1.2}$$

where X_0 represents the future observation from the same production process. Let (t_1, t_2) be the observed of this tolerance interval. The general rule for verifying a manufacturer's problem using the tolerance interval is as follows:

$$\begin{array}{c}\textbf{If } (t_1, t_2) \subset (LSL, USL), \textbf{the lot of product is acceptable,}\\ \textbf{because we have confidence } 1-\alpha\\ \textbf{that at least } 100\gamma\% \textbf{ of the products}\\ \textbf{conform to specification limits.}\end{array} \tag{1.3}$$

Much attention has been paid for developing tolerance intervals, for examples Wilks (1941), Wald (1943), Paulson (1943), Guttman (1970) and, for a review, Patel (1986). In general, a common effort been made in the literature is to investigate the version with minimum width, for which Eisenhart et al. (1947) constructed an approximate minimum width tolerance interval for normal random variable. This normal tolerance interval is now popularly implemented in manufacturing industries and is presented in text books of engineering statistics. The interest of this paper is to study if the tolerance interval is appropriate to deal with problem in (1.1) for the manufacturer and consumer.

To study the appropriateness of tolerance interval in this engineering problem, we need to clarify its role with classical hypothesis testing problem. The classical hypothesis testing problem set a simple hypothesis and derive a test with a specified significance level, usually a small value such as 0.05 or 0.01. However, varied philosophies set null hypothesis in different ways that may lead to completely different conclusions. The most commonly accepted rule for setting null hypothesis through the following philosophy:

$$\begin{array}{c}\text{If one wishes to prove that a hypothesis A is true,}\\ \text{one first assumes that it isn't true.}\end{array} \tag{1.4}$$

This philosophy favors right of the consumer (buyer). In a clinical trial, one wish to see if a new drug has a different effect. In general, the null hypothesis might be that the new drug is no better, on average, than the current drug. Why is this philosophy in these problems? Switching a new product (drug) or technique usually requires large initial expenditures, and a decision maker should not do so unless the new product is significantly better than the old one.

There is an approach of philosophy considers a situation employed in some engineering problems. The quality of product is distributed from background noise and abnormal noise. If the process operator adjust the process based on the tests performed periodically for hypothesis defined from rule of (1.4), it will often overacted to the background noise and to deteriorate the performance of the process. Hence, to prevent this unnecessary adjustment, a rule guiding the null hypothesis is as follows:

$$\text{Unless it is broken, we do not fix it.} \tag{1.5}$$

This rule follows based on the philosophy from the aspect of the manufacturer (producer).

In the on line quality control, without evidence of existing assignable cause, the manufacturing process is considered to be in control. For example, for detecting if there is a mean shift, one construct the $\bar{X}$ chart as

$$\begin{aligned} UCL &= \mu_0 + 3\tfrac{\sigma_0}{\sqrt{n}} \\ LCL &= \mu_0 - 3\tfrac{\sigma_0}{\sqrt{n}} \end{aligned} \tag{1.6}$$

where μ_0 and σ_0 are, respectively, the mean and standard deviation for in control process. Probability 0.9973 of acceptable region does in favor of hypothesis H_0. Why use this philosophy for process control? Suppose that the process operators adjust the manufacturing process based on tests obeying philosophy one. Only strong evidence showing in data for supporting the hypothesis of in control process will make the operators often overreacted to the random cause for unnecessary process adjustments. These unnecessary process adjustments can actually result in a deterioration of process performance. From the above discussion, the process control using the control chart does follows the philosophy of (1.5) assuming that the process is in control as the hypothesis.

Not all statistical hypotheses problems design a test with null hypothesis following a philosophy only to protect the risk of either the manufacturer or the consumer. Acceptance sampling, popularized by Dodge and Roming and was originally applied by the U.S. military, provides a rule of decision of accepting or rejecting a production lot. With a sample of size n for items selected at random from a lot, a plan as denoted as $\{n, c\}$ is conducted with lot accepted when there are defectives of number less than or equal to c. By calling the percentage defective as average quality level (AQL), this plan is generally designed to have high probability of lot acceptance when there is low AQL value and to have low probability of lot acceptance when there is large AQL value.

On the other hand, a test based on tolerance interval expects to have high probability of hypothesis acceptance when lots of production are with γ percentage of acceptable products at confidence remarkably larger than $1-\alpha$ and to have low probability of hypothesis acceptance when lots of production are with γ percentage of acceptable products at confidence remarkably lower than $1-\alpha$. Hence, this kind of test should not accompany with hypothesis following the philosophy only to protect the risk of either the manufacturer or the consumer. We will see that for lots of production, the confidence for γ percentage of acceptable products may be formulated as an unknown parameter. The test presented in (1.3) in fact deals with a hypothesis on value of this unknown confidence. Hence there are two problems we may concern that we want to deal in this paper. (a) Is the test presented in (1.3) appropriate for protecting both the risks of manufacturer and consumer? (b) Is there alternative test, also formulated based on tolerance interval, that performs better to protect the risks of manufacturer and consumer.

This paper formulate the unknown confidence in Section 2, for a γ percentage acceptable products, in an explicit expression. We then study in Section 3 the power of the γ content tolerance interval of Eisenhart et al. (1947) with test of (1.3). We will propose in Section 4 a new test for the same purpose of studying if there is γ percentage of acceptable products at confidence $1-\alpha$. A sample size determination problem will also be studied in Section 5 that guarantees a low probability of acceptance of the hypothesis when the true specification limits are moderately shorter than the desired ones and a large probability of acceptance of the hypothesis when the true specification limits are moderately wider than the desired ones.

2. Specification Settings for Achieving Percentage γ Acceptable Products at a Fixed Confidence

Let X be a random variable having distribution function F_θ with probability density function $f_\theta(x)$ and the specification limits for product characteristic X

is $\{LSL, USL\}$. We call the probability that a product to be acceptable

$$p_{item}(\theta) = \int_{LSL}^{USL} f_\theta(x)\,dx = F_\theta(USL) - F_\theta(LSL) \tag{2.1}$$

the item reliability where F_θ is the distribution function. Suppose that the lot size is known as constant k (usually a large number). For this production lot, the number of acceptable products is with binomial distribution $b(k, p_{item}(\theta))$. Then the true confidence for having proportion γ of production lot conforming to specification limits is

$$q = \textstyle\sum_{i=[k\gamma]}^{k} \begin{pmatrix} k \\ i \end{pmatrix} p_{item}(\theta)^i (1 - p_{item}(\theta))^{k-i}. \tag{2.2}$$

This expression shows that the confidence of a γ percentage of acceptable products in lots interesting for the manufacturer actually is an unknown parameter. Hence, the interest for a manufacturer is to test the following hypothesis:

$$H^* : q \geq q_0 \tag{2.3}$$

for some specified (large) value q_0. We wouldn't call H^* a null or alternative hypothesis since it is not appropriate to consider a classical test for it.

The classical approach to test H^* is rule (1.3) through a γ-content tolerance interval (T_1, T_2) at confidence q_0 that may be re-written as

$$\text{Accept } H^* \text{ if } (t_1, t_2) \subset (LSL, USL) \tag{2.4}$$

where (t_1, t_2) is the observation of (T_1, T_2) (see Bowker and Goode (1952) and Papp (1992) for this application). With this observation that there is an unknown true confidence, it raises the question that what we may expect a test for its power function representing the probability that H^* is accepted.

Recall that the manufacturer expects to have proportion γ or more acceptable products at confidence q_0. Let's define the minimum item reliability that guarantees proportion γ acceptable products at confidence q_0 as p_{q_0} satisfying

$$\textstyle\sum_{i=[k\gamma]}^{k} \begin{pmatrix} k \\ i \end{pmatrix} (p_{q_0})^i (1 - p_{q_0})^{k-i} = q_0. \tag{2.5}$$

With product's characteristic variable having a distribution function F_θ, the manufacturer is desired to have item reliability $p_{item}(\theta)$ with

$$p_{item}(\theta) = F_\theta(USL) - F_\theta(LSL) \geq p_{q_0}. \quad (2.6)$$

From (2.5), the manufacturer will loss money most of the times when $p_{item}(\theta)$ is moderately smaller than p_{q_0}. For a given pairs (γ, q_0), we list the item reliabilities that achieve exactly proportion γ of acceptable products with confidence q_0 in the following table.

Table 1. Minimum item reliability p_{q_0}

q_0	$\gamma = 0.8$	0.85	0.9	0.95	0.99
$(k = 1,000)$					
0.8	0.8099	0.8587	0.9071	0.9549	0.9918
0.85	0.8123	0.8608	0.9089	0.9562	0.9923
0.9	0.8152	0.8635	0.9111	0.9577	0.9929
0.95	0.8196	0.8673	0.9142	0.9599	0.9938
0.99	0.8277	0.8744	0.9200	0.9638	0.9952
$k = (10,000)$					
0.8	0.8033	0.8529	0.9025	0.9518	0.9907
0.85	0.8040	0.8536	0.9030	0.9522	0.9909
0.9	0.8050	0.8545	0.9038	0.9527	0.9912
0.95	0.8064	0.8557	0.9048	0.9534	0.9915
0.99	0.8091	0.8581	0.9068	0.9549	0.9921
$k = (100,000)$					
0.8	0.8010	0.8509	0.9008	0.9505	0.9902
0.85	0.8013	0.8512	0.9010	0.9507	0.9903
0.9	0.8016	0.8514	0.9012	0.9509	0.9904
0.95	0.8021	0.8518	0.9016	0.9511	0.9905
0.99	0.8029	0.8526	0.9022	0.9516	0.9907

We will derive the specification limits that achieve the minimum item reliability for advanced study later in this paper . Suppose that the characteristic variable of interest obeys a normal distribution $N(\mu, \sigma^2)$. Then item reliability, the probability that an item conforming to specifications, is

$$p_{item}(\mu, \sigma) = \int_{LSL}^{USL} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx$$

and then the true confidence to have proportion γ acceptable products is

$$\begin{aligned} q &= \sum_{i=[k\gamma]}^{k} \binom{k}{i} p_{item}(\mu,\sigma)^i (1 - p_{item}(\mu,\sigma))^{k-i} \\ &= \sum_{i=[k\gamma]}^{k} \binom{k}{i} \left(\Phi\left(\tfrac{USL-\mu}{\sigma}\right) - \Phi\left(\tfrac{LSL-\mu}{\sigma}\right)\right)^i \quad \left(1 - \left(\Phi\left(\tfrac{USL-\mu}{\sigma}\right) - \Phi\left(\tfrac{LSL-\mu}{\sigma}\right)\right)\right)^{k-i}. \end{aligned}$$

A production lot to have proportion γ acceptable products with confidence q_0 requires that

$$p_{item}(\mu, \sigma) = \Phi\left(\frac{USL-\mu}{\sigma}\right) - \Phi\left(\frac{LSL-\mu}{\sigma}\right) \geq p_{q_0} \tag{2.7}$$

where p_{q_0} may be found in Table 1.

One interesting question is that how short the specification limits should be to achieve the minimum item reliability. Let the specification limits be $\{LSL, USL\} = \{\mu - l\sigma, \mu + l\sigma\}$ and we denote l_{q_0} as the l so that the item reliability is

$$p_{q_0} = P(\mu - l_{q_0}\sigma \leq X \leq \mu + l_{q_0}\sigma).$$

We list l_{q_0} in this design in the following table.

Table 2. Specification limits $(LSL, USL) = (\mu - l_{q_0}\sigma, \mu + l_{q_0}\sigma)$ to achieve item reliability exactly equal to p_{q_0}

q	$\gamma = 0.8$	0.85	0.9	0.95	0.99
$(k = 1,000)$					
0.8	1.3102	1.4711	1.6807	2.0043	2.6456
0.85	1.3174	1.4790	1.6898	2.0161	2.6682
0.9	1.3265	1.4890	1.7013	2.0310	2.6947
0.95	1.3397	1.5037	1.7184	2.0532	2.7372
0.99	1.3650	1.5317	1.7508	2.0955	2.8210
$(k = 10,000)$					
0.8	1.2910	1.4499	1.6569	1.9752	2.6020
0.85	1.2931	1.4525	1.6597	1.9787	2.6096
0.9	1.2961	1.4556	1.6633	1.9833	2.6187
0.95	1.3001	1.4601	1.6686	1.9901	2.6324
0.99	1.3080	1.4689	1.6786	2.0033	2.6561
$k = (100,000)$					
0.8	1.2845	1.4428	1.6488	1.9647	2.5841
0.85	1.2853	1.4436	1.6497	1.9662	2.5879
0.9	1.2861	1.4446	1.6509	1.9647	2.5906
0.95	1.2875	1.4460	1.6525	1.9698	2.5928
0.99	1.2900	1.4489	1.6554	1.9737	2.6007

This table will be used in next section to study the power of the classical tolerance interval in detection of manufacturer's confidence.

3. Power of the Classical Tolerance Interval

Suppose that we have $X_1, ..., X_n$ a random sample for the characteristic variable of interest and the specification limits for the characteristic are $\{LSL, USL\}$. Let (T_1, T_2) be a tolerance interval of Wilks (1941) constructed from the random sample. One of the popular applications of tolerance interval is to test hypothesis H^* based on rule of (2.4). With the fact that the true confidence is a parameter, it is interesting to evaluate power function of this tolerance interval, in terms of specification limits, is

$$\pi(LSL, USL) = P_\theta((T_1, T_2) \subset (LSL, USL)). \tag{3.1}$$

It provides the probability that we should conclude that there is proportion γ or more acceptable products in a lot at confidence $1-\alpha$. The optimal tolerance interval, if there is, should have power value 1 for $q \geq q_0$ and value zero for $q < q_0$. This is generally not attainable. Hence, there are two properties that we expect a tolerance interval to be satisfies:

(a) The power function is nondecreasing in terms of item reliability p_{item}.

(b) For a balance of the manufacturer's benefits and risks, the power when true confidence q is equal to q_0 is close to 0.5.

We want to simulate the powers for the Eisenhart et al.'s tolerance interval for several combinations of specification limits. Let's set replication number m and specification limits $(LSL, USL) = (-b, b)$. The simulated power of a tolerance interval (T_1, T_2) is defined as

$$\hat{\pi} = \frac{1}{m}\sum_{j=1}^{m} I((t_1^j, t_2^j) \subset (-b, b)) \tag{3.2}$$

where (t_1^j, t_2^j) is the observation of (T_1, T_2) from the jth sample. The power of (3.2) simulates the chance of (3.1) that the tolerance interval (T_1, T_2) may conclude that the production lot includes a proportion γ of acceptable products with confidence $1-\alpha$.

To study (3.2), suppose that the random sample $X_1, ..., X_n$ is drawn from normal distribution $N(\mu, \sigma^2)$ where both μ and σ are unknown. The general form of a prediction interval for a future normal random variable is of the form

$$(\bar{X} - m^* s, \bar{X} + m^* s) \tag{3.3}$$

where the $100(1-\alpha)\%$ confidence interval (prediction interval) is the form with $m^* = t_{1-\frac{\alpha}{2}}(n-1)\sqrt{1+\frac{1}{n}}$ and where $t_{1-\frac{\alpha}{2}}(n-1)$ represents the $1-$

$\frac{\alpha}{2}$th quantile of the central t-distribution with degrees of freedom. For the Wilks' tolerance interval, Eisenhart et al. (1947) developed the shortest one which is now the most popular version of tolerance interval to deal with the manufacturer's problem when the characteristic variable does obey a normal distribution. We select values m^* corresponding with $\gamma = 0.9, 1-\alpha = 0.95$ from the table developed in Eisenhart et al. (1947).

With replication $m = 100,000$, we generate random sample of size n from distribution $N(0,1)$. Let $\bar{X}_j$ and S_j^2 be the sample mean and sample variance for jth sample. We compute this tolerance interval and study its powers of (3.2) with several sample size $n = 20, 30, 50$ and various values b where $b = 1.7184, 1.6686$ and 1.6525 corresponds, respectively, to specification limits such that their true confidences are identical to $1-\alpha = 0.95$. The simulated results are listed in Table 3,4,5.

We have several comments drawn from the these tables:

(a) As expected, the power of the tolerance interval is increasing when the specification limits are wider indicating increasing in p_{item}. For $b \geq 1.7184$ with $k = 1,000$, the corresponding confidence $q \geq q_0 = 0.95$, we see that the larger the sample size the more the chance (probability) to accept H^*.

(b) When $b = 1.7184$ for $k = 1,000$, the process does guarantee confidence 0.95 with percentage 0.9 of acceptable products. However, the simulated power values are $0.0257, 0.0289, 0.0466$, respectively, for sample sizes $n = 20, 30, 50$. These revealed little chance to observe that the lots are already $\gamma = 0.9$ percentage of acceptable products at confidence 0.95. Hence, the test of (2.3) is not satisfactory in losing benefits for the manufacturer.

4. A New Test Based on Tolerance Interval

A test for hypothesis H^* is expected to have power not too far from 0.5 when $q = q_0$ is true. The classical test based on tolerance interval does not meet this requirement. We then introduce a new test.

Suppose that we have an appropriate estimate, denoted by $\hat{\theta}$ of parameter θ and then we have estimated probability density function of characteristic variable X as $f_{\hat{\theta}}(x)$. The rule of the new test for hypothesis H^* of (2.3) is:

$$\text{Accepting } H^* \text{ if } \int_{(t_1,t_2)\cap(LSL,USL)} f_{\hat{\theta}}(x)dx \geq p_{q_0}. \tag{4.1}$$

We have two comments for setting the above rule to test hypothesis H^*:

(a) For given a γ-content tolerance interval (T_1, T_2) at confidence $1-\alpha$, if its observation (t_1, t_2) does contain percentage of products conforming with specification limits p_{q_0} or more, we conclude with confidence $1-\alpha$ that there is percentage γ acceptable products in a lot.

(b) This test sets test statistic $\int_{(T_1,T_2)\cap(LSL,USL)} f_{\hat{\theta}}(x)dx$. However, the critical point p_{q_0} is not the cut off point based on distribution of the test statistic. Hence, this test does not follows the classical hypothesis testing to ensures a specified significance level.

With this test, the power function is

$$P_\theta\{\int_{(T_1,T_2)\cap(LSL,USL)} f_{\hat{\theta}}(x)dx \geq p_{q_0}\}. \tag{4.2}$$

Consider that we have a random sample $X_1, ..., X_n$ drawn from the distribution $N(\mu, \sigma^2)$. Let $\hat{\mu} = \bar{x}$ and $\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - \hat{\mu})^2$. The rule for testing hypothesis H_0 is:

$$\text{Accepting } H_0 \text{ if } \int_{(t_1,t_2)\cap(LSL,USL)} \phi_{\hat{\mu},\hat{\sigma}}(x)dx \geq p_{q_0}, \tag{4.3}$$

and the empirical power of tolerance interval (T_1, T_2) is

$$\hat{\pi}_{Spe} = \frac{1}{m}\sum_{j=1}^m I(\int_{(t_1^j,t_2^j)\cap(LSL,USL)} \phi_{\hat{\mu},\hat{\sigma}}(x)dx \geq p_{q_0}). \tag{4.4}$$

We conduct a simulation with the same design set in Section 3 to study the power function of this new test. The simulated results for lot sizes, $k = 1,000, 10,000, 100,000$ are listed in Tables 7,8,9 (see these tables in the end of this paper). However, the tolerance intervals considered including the shortest version (STL) and the version (CITL) developed by Huang, Chen and Welsh (2007).

We have several comments drawn from Tables 7,8 and 9:

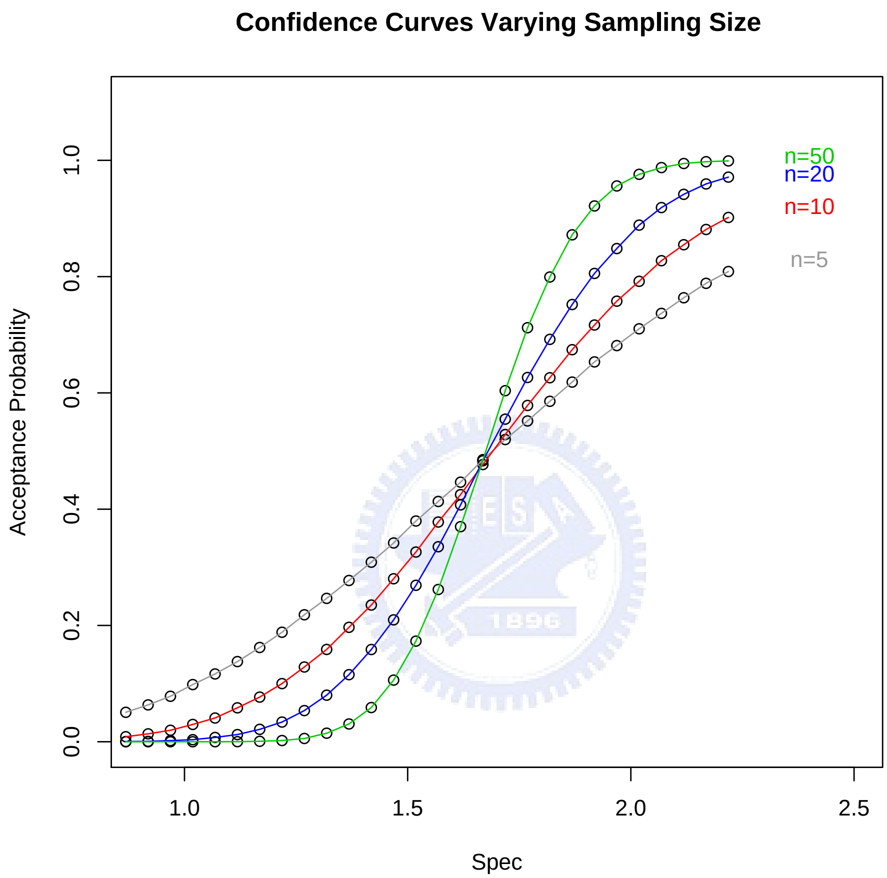
(a) Monotone power values are as our expectation. However, the power values are relatively higher than them based on test of (2.4). Hence, there are larger probabilities in all settings of specification limits and sample sizes for accepting H^*.
(b) There is no significant differences in the performance between the shortest tolerance interval and the version of Huang, Chen and Welsh.
(c) When b is the value (1.7184 for $k = 1,000$, 1.6686 for $k = 10,000$ and 1.6525 for $k = 100,000$) that the true confidence is identical to 0.95 the simulated power values are all close to 0.48. This is interesting indicating that this new test is more capable in our purpose.

5. Sample Size Determination for Tolerance Interval

The most popular technique to judge if a production lot is accepted is the acceptance sampling plan which specifies a sample size n and an acceptance number c and the lot is accepted if the number of defective items in this sample is less than or equal to c. An important question in acceptance sampling plan approach is to determine the sample size that the probabilities of acceptance when fraction of defectives are p_1 and p_2 with $p_1 < p_2$, respectively, achieve two specified values. This is to protect the manufacturer having large probability of acceptance when fraction defective is as low as p_1 and protect the consumer having small probability of acceptance when fraction defective is as large as p_2. This idea may be extended to the acceptance of lots at some confidence by tolerance intervals.

Suppose that we have a random sample $X_1, ..., X_n$ drawn from a distribution with distribution function F_θ. The probability that a product to be acceptable expressed in (2.1) and the confidence for having proportion γ of products in a lot conforming to specification limits expressed in (2.2) are both dependent on parameter θ, specification limits LSL, USL and lot size k. However, its power of (4.2) actually relies on the efficiency of the estimator of unknown parameter θ. Generally, efficiency may be improved when sample size n is increased. Hence, the sample size n should be set to satisfy the consumer with a small power of (4.2) when the interval of specification limits is shorter and to satisfy the manufacturer with a large power when the interval is wilder.

We conduct a simulation to generate the power function (4.4) for normal distribution and display the simulated power functions for several sample sizes. The resulted power functions are shown in Figure 1.



The graph that plots the confidence $q(Spec)$ versus the specification limits may be called the OC graph. It is interesting to design a tolerance interval requiring that the confidence is a desired large value when the interval of specification limits is wilder and it is a desired small value when the interval is shorter. That is, with $(LSL_1, USL_1) \subset (LSL, USL) \subset (LSL_2, USL_2)$ and

$q_1 < q_2$, find sample size n such that the followings

$$
\begin{aligned}
& P_\theta\{\int_{(T_1,T_2)\cap(LSL_1,USL_1)} f_{\hat{\theta}}(x)dx \geq p_{q_0}\} \leq q_1 \\
& P_\theta\{\int_{(T_1,T_2)\cap(LSL_2,USL_2)} f_{\hat{\theta}}(x)dx \geq p_{q_0}\} \geq q_2.
\end{aligned} \tag{5.1}
$$

We consider an empirical solution of sample size n as minimum n satisfying the following equations

$$
\begin{aligned}
& \frac{1}{m}\sum_{j=1}^{m} I(\int_{(t_1^j,t_2^j)\cap(LSL_1,USL_1)} \phi_{\hat{\mu},\hat{\sigma}}(x)dx \geq p_{q_0}) \leq q_1 \\
& \frac{1}{m}\sum_{j=1}^{m} I(\int_{(t_1^j,t_2^j)\cap(LSL_2,USL_2)} \phi_{\hat{\mu},\hat{\sigma}}(x)dx \geq p_{q_0}) \geq q_2.
\end{aligned} \tag{5.2}
$$

For the purpose of sample size determination, we further make the following assumptions:

(a) The lot size k and percentage value γ are both fixed.

(b) The parameter(s) is assumed to be known as θ_0. For case that θ is not known, we assume that there is a training sample for us to estimate it.

Let's use the normal distribution as an example for explaining the technique of sample size determination. The sample size is determined from (5.2) with item reliability

$$p_{item}(Spec) = \Phi\left(\frac{USL - \mu_0}{\sigma_0}\right) - \Phi\left(\frac{LSL - \mu_0}{\sigma_0}\right).$$

Without lose of generality, we let $\mu_0 = 0$ and $\sigma_0 = 1$ and for our convenience, we let specification limits $\{LSL, USL\} = \{-\ell, \ell\}$. The item reliability then is

$$p_{item}(\ell) = \Phi(\ell) - \Phi(\ell).$$

We choose $k = 1,000$. From Table 2, the specification limits $(LSL, USL) = (-1.7184, 1.7184)$ guarantees to meet percentage 0.9 acceptable products at confidence 0.95. However, suppose that the manufacturer asks for probability 0.3 or lesser of accepting the lots most of the time when the specification limits are $\{LSL_1, USL_1\} = \{-1.5, 1.5\}$ and probability 0.7 or more of accepting the lots most of the time when the specification limits are $\{LSL_2, USL_2\} = \{-1.9, 1.9\}$. These conditions require sample size $n = 12$. There are other two cases of conditions that the sample sizes are also listed in Table 10.

Table 10. Desired sample sizes meet the probabilities of acceptance of lots when the interval of specification limits is wilder or shorter.

$Spec_1$	q_1	$Spec_2$	q_2	n
$(-1.5, 1.5)$	0.3	$(-1.9, 1.9)$	0.7	12
$(-1.5, 1.5)$	0.2	$(-1.9, 1.9)$	0.8	32
$(-1.5, 1.5)$	0.1	$(-1.9, 1.9)$	0.9	78

This explanation of sample size determination involves a simpler design of specification limits. For purpose of general applications, further studies are needed.

References

Eisenhart, C., Hastay, M. W. and Wallis, W. A. (1947). *Techniques of Statistical Analysis.* McGraw-Hill Book Company: New York.

Guttman, I. (1970). *Statistical tolerance regions: Classical and Baysian.* Griffin: London.

Huang, J.-Y., Chen,L.-A. and Welsh, A. (2007). The mode and its application to control charts and tolerance intervals. Submitted to *Statistics and Probability Letter* for publication (In revision).

Patel, J.K. (1986). Tolerance limits - A review. *Communications in statistics - Theory and Methods*, 15, 2719-2762.

Paulson, E. (1943). A note on tolerance limits. *Annals of Mathematical Statistics*, 14, 90-93.

Poulsen, O.M., Holst, E. & Christensen, J.M. (1997). Calculation and appli cation of coverage intervals for biological reference values (technical repo rt). *Pure and Applied Chemistry*, 69, 1601-1611.

Wald, A. (1943). An extension of Wilks' method for setting tolerance limits. *Annals of Mathematical Statistics*, 14, 45-55.

Wilks, S. S. (1941). Determination of sample sizes for setting tolerance limits. *Annals of Mathematical Statistics*, 12, 91-96.

Tables

Table 3. Powers of the minimum-width tolerance interval ($(\gamma, 1-\alpha) = (0.9, 0.95)$) $k = 1000$

Limits	$n = 20$	$n = 30$	$n = 50$
(m^*)	(2.310)	(2.140)	(1.969)
$b = 1.4$ $q = 1.371e - 08$	0.0019	0.0008	0.0004
$b = 1.5$ $q = 0.00071$	0.0043	0.0033	0.0025
$b = 1.6$ $q = 0.17899$	0.0104	0.0096	0.0111
$b = 1.645$ $q = 0.52786$	0.0157	0.0153	0.0204
$b = 1.7184$ $q = 0.949878$	0.0257	0.0289	0.0466
$b = 1.8$ $q = 0.9995$	0.0436	0.0556	0.1001
$b = 2.0$ $q = 1.0$	0.1213	0.1888	0.3777
$b = 2.2$ $q = 1.0$	0.2611	0.4224	0.7313
$b = 2.5$ $q = 1.0$	0.5528	0.7829	0.9730
$b = 3.0$ $q = 1.0$	0.9111	0.9898	1.0000
$b = 3.5$ $q = 1.0$	0.9934	0.9999	1.0000

Table 4. Powers of the minimum-width tolerance interval $((\gamma, 1-\alpha) = (0.9, 0.95))$ $k = 10000$

Limits	$n = 20$	$n = 30$	$n = 50$
(m^*)	(2.310)	(2.140)	(1.969)
$b = 1.4$ $q = 9.71747e - 71$	0.0015	0.0009	0.0005
$b = 1.5$ $q = 5.569e - 25$	0.0045	0.0028	0.0022
$b = 1.6$ $q = 0.00099$	0.0109	0.0093	0.0108
$b = 1.645$ $q = 0.5124351$	0.0155	0.0150	0.0205
$b = 1.6686$ $q = 0.9501104$	0.0181	0.0186	0.0270
$b = 1.8$ $q = 1.0$	0.0416	0.0571	0.1002
$b = 2.0$ $q = 1.0$	0.1227	0.1890	0.3758
$b = 2.2$ $q = 1.0$	0.2622	0.4231	0.7274
$b = 2.5$ $q = 1.0$	0.5528	0.7832	0.9719
$b = 3.0$ $q = 1.0$	0.9112	0.9898	1.0000
$b = 3.5$ $q = 1.0$	0.9932	1.0000	1.0000

Table 5. Powers of the minimum-width tolerance interval ($(\gamma, 1-\alpha) = (0.9, 0.95)$) $k = 100000$

Limits	$n = 20$	$n = 30$	$n = 50$
(m^*)	(2.310)	(2.140)	(1.969)
$b = 1.4$ $q = 0$	0.0017	0.0009	0.0003
$b = 1.5$ $q = 7.83579e - 232$	0.0043	0.0031	0.0029
$b = 1.6$ $q = 3.809744e - 23$	0.0106	0.0097	0.0108
$b = 1.645$ $q = 0.5153559$	0.0154	0.0147	0.0206
$b = 1.6525$ $q = 0.9521992$	0.0160	0.0161	0.0223
$b = 1.8$ $q = 1.0$	0.0430	0.0558	0.1016
$b = 2.0$ $q = 1.0$	0.1215	0.1889	0.3762
$b = 2.2$ $q = 1.0$	0.2616	0.4222	0.7272
$b = 2.5$ $q = 1.0$	0.5512	0.7803	0.9719
$b = 3.0$ $q = 1.0$	0.9114	0.9904	1.0000
$b = 3.5$ $q = 1.0$	0.9941	1.0000	1.0000

Table 7. Powers of the minimum-width tolerance interval and coverage interval based tolerance interval ($(\gamma, 1-\alpha) = (0.9, 0.95), k = 1000$)

Limits	$n = 20$ STL	CITL	$n = 30$ STL	CITL	$n = 50$ STL	CITL
$b = 1.4$	0.1067	0.1074	0.0651	0.0659	0.0258	0.0269
$b = 1.5$	0.1954	0.1966	0.1496	0.1511	0.0892	0.0923
$b = 1.6$	0.3140	0.3154	0.2799	0.2826	0.2273	0.2339
$b = 1.645$	0.3746	0.3763	0.3529	0.3560	0.3132	0.3206
$b = 1.7184$	0.4781	0.4800	0.4813	0.4850	0.4773	0.4875
$b = 1.8$	0.5950	0.5969	0.6420	0.6273	0.6565	0.6661
$b = 2$	0.8245	0.8257	0.8781	0.8799	0.9357	0.9401
$b = 2.2$	0.9458	0.9465	0.9780	0.9786	0.9953	0.9959
$b = 2.5$	0.9956	0.9957	0.9993	0.9993	1	1
$b = 3$	1	1	1	1	1	1

Table 8. Powers of the minimum-width tolerance interval and coverage interval based tolerance interval $((\gamma, 1-\alpha) = (0.9, 0.95), k = 10000)$

Limits	$n=20$ STL	CITL	$n=30$ STL	CITL	$n=50$ STL	CITL
$b = 1.4$	0.1413	0.1419	0.0942	0.0949	0.0459	0.0468
$b = 1.5$	0.2465	0.2475	0.2022	0.2035	0.1415	0.1411
$b = 1.6$	0.3814	0.3824	0.3586	0.3604	0.3251	0.3297
$b = 1.645$	0.4434	0.4448	0.4387	0.4408	0.4258	0.4308
$b = 1.6686$	0.4781	0.4793	0.4836	0.4856	0.4788	4841
$b = 1.8$	0.6675	0.6686	0.7096	0.7116	0.7673	0.7719
$b = 2$	0.8739	0.8747	0.9232	0.9243	0.9688	0.9701
$b = 2.2$	0.9665	0.9668	0.9887	0.9889	0.9986	0.9986
$b = 2.5$	0.9978	0.9978	0.9998	0.9998	1	1
$b = 3$	1	1	1	1	1	1

Table 9. Powers of the minimum-width tolerance interval and coverage interval based tolerance interval $((\gamma, 1-\alpha) = (0.9, 0.95), k = 100000)$

Limits	$n = 20$ STL	CITL	$n = 30$ STL	CITL	$n = 50$ STL	CITL
$b = 1.4$	0.1528	0.1534	0.1067	0.1074	0.0539	0.0548
$b = 1.5$	0.2642	0.2650	0.2227	0.2242	0.1671	0.1695
$b = 1.6$	0.3995	0.4003	0.3858	0.3876	0.3603	0.3639
$b = 1.645$	0.4688	0.4699	0.4695	0.4713	0.4626	0.4668
$b = 1.6525$	0.4786	0.4797	0.4813	0.4831	4835	0.4877
$b = 1.8$	0.6898	0.6909	0.7368	0.7385	0.7965	0.8003
$b = 2$	0.8875	0.8880	0.9348	0.9357	0.9757	0.9765
$b = 2.2$	0.9724	0.9726	0.9916	0.9918	0.9991	0.9992
$b = 2.5$	0.9983	0.9983	0.9999	0.9999	1	1
$b = 3$	1	1	1	1	1	1