國 立 交 通 大 學

統計學研究所

碩 士 論 文

容忍區間

Tolerance Intervals

研 究 生 ： 顏雅芬

指導教授 ： 陳鄰安　教授

中 華 民 國 九 十 五 年 六 月

容忍區間
Tolerance Intervals

研 究 生 ： 顏雅芬　　　Student：Ya-Fen Yen
指導教授 ： 陳鄰安　　　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

Nation Chiao-Tung University

in partial Fulfillment of the Requirements

for the degree of Master

in

Statistics
June 2006
Hsinchu, Taiwan

中 華 民 國 九 十 五 年 六 月

容忍區間

研究生： 顏雅芬　　　　　　　指導教授： 陳鄰安 教授

國立交通大學統計學研究所

摘 要

針對生產者的需求，我們定義了可容許性容忍區間的概念。由此一觀念，一般藉由 Wilks (1941) 所定義的容忍區間可能是不具可容許性的容忍區間。因此，我們證明出最常用於常態分配的 Eisenhart et al. (1947) 容忍區間是不具可容許性的。我們證明出一個隨機區間是具有可容許性的性質，若且為若它是一個由覆蓋區間所建立的信賴區間。我們更進一步地評估一些已存在的容忍區間它們的可容許性程度。最後，我們推導出某些分配的最短可容許性容忍區間。

關鍵字：信賴區間；覆蓋區間；容忍區間

Tolerance Intervals

Student : Ya-Fen Yen Advisor : Dr. Lin-An Chen

Institute of Statistics

National Chiao Tung University

Abstract

For tolerance interval, we define a concept of admissibility that is desired for manufacturer. This leads to a problem that the general concept of tolerance intervals defined by Wilks (1941) may provide in-admissible tolerance intervals. For this, we show that the most popular normal tolerance interval of Eisenhart et al. (1947) is not admissible. A theory showing that a random interval is an admissible tolerance interval if and only if a confidence interval of a coverage interval is established. We further evaluate some existed tolerance intervals for their admissibility and also derive the shortest admissible tolerance intervals for some distributions.

Key words: Confidence interval; coverage interval; tolerance interval.

誌謝

在交大統研所兩年的研究生生活轉眼間就要結束。雖然短暫，卻很充實，且獲益良多。感謝所上老師辛苦教導，特別是教授陳鄰安老師不辭辛勞指導，即使例假日亦不例外，得使學生順利完成本論文。並感謝博士班陳弘家學長協助指導使用軟體，讓我能很快上手應用；同時也感謝黃榮臣老師、許文郁老師、黃冠華老師在論文口試時提供寶貴意見與建議，使論文內容更臻完善。

很慶幸能有 93 級這群一起走過的夥伴們，無論是課業上、生活上，我們一起為課業努力，一起找尋美食，一起遊玩，兩年來點點滴滴將是腦海裡永恆的畫面，謝謝你們。另外，也很感謝從高中就一直相伴的好朋友，在我沮喪時給予我關心與鼓勵。

最後，感謝我最親愛的家人們，永遠在背後支持我，給我加油打氣以及信心。謹將此論文獻給我的家人、老師、同學、朋友，謝謝你們！

顏雅芬　　謹誌于
國立交通大學統計學研究所
中華民國九十五年六月

Contents

Tables of contents

Tolerance Intervals

Abstract

For tolerance interval, we define a concept of admissibility that is desired for manufacturer. This leads to a problem that the general concept of tolerance interval defined by Wilks (1941) may provide in-admissible tolerance intervals. For this, we show that the most popular normal tolerance interval of Eisenhart et al. (1947) is not admissible. A theory showing that a random interval is an admissible tolerance interval if and only if a confidence interval of a coverage interval is established. We further evaluate some existed tolerance intervals for their admissibilities and also derive the shortest admissible tolerance intervals for some distributions.

Key words: Confidence interval; coverage interval; tolerance interval.

1. Introduction and Motivation

Statistical theory of interval estimation mostly deals with the confidence interval to contain a parameter θ. In many applications, we require an interval to contain the future r.v. which is a prediction problem. Among the alternatives, intervals in the form of tolerance intervals are widely used in quality control and related prediction problems to monitor manufacturing processes, detect changes in such processes, ensure product compliance with specifications, etc.

In manufacturing industry, specification limits for one charateristic of an item, saying L_1 and L_2, define the boundaries of acceptable quality for an manufacturing item (component). For a manufacturer of a mass-production item, the tolerance interval is designed for a quality assurance problem. The manufacturer is interesting in an interval that contains a specified (usually large) percentage of the product and he knows that unless 90% of his production is acceptable in the sense that the item's characteristic falls in the limits, he will loss money in this production. With this interest, a hypothesis testing problem is formulated as follows:

$$H_0: \textbf{There is an interval that includes at least a certain proportion of acceptable measurements with a stated confidence.} \tag{1.1}$$

To conquer this hypothesis testing problem, it is done in literature through two steps. The first step is to consider an interval estimation problem:

$$\textbf{An interval that includes at least a certain proportion of measurements with a stated confidence,} \tag{1.2}$$

called the tolerance interval. Suppose that we have a a random sample $X = (X_1, ..., X_n)'$ from a distribution with pdf $f(x, \theta)$, what have been done in producing a tolerance interval? this problem was treated in a pioneer article by Wilks (1941) where a γ-content tolerance interval with confidence $1 - \alpha$ is an inetrval $(T_1, T_2) = (t_1(X), t_2(X))$ that satisfies

$$P_\theta\{P_\theta(X_0 \in (T_1, T_2)|X) \geq \gamma\} \geq 1 - \alpha \text{ for } \theta \in \Theta \tag{1.3}$$

where Θ is the parameter space and X_0 represents the future observation with the same distribution. The second step for solving problem in (1.1) is testing hypothesis of (1.1) based on the tolerance inteval in (1.3) with the following rule (see Bowker and Goode (1952)) as

$$\begin{array}{c}\textbf{We accept the lot of product if } t_1 \geq L_1 \text{ and } t_2 \leq L_2, \\ \textbf{i.e., we reject the lot of product if either } t_1 < L_1 \textbf{or } t_2 > L_2 \\ \textbf{or both } t_1 < L_1 \textbf{and } t_2 > L_2.\end{array} \tag{1.4}$$

In this two steps in solving the hypothesis problem of (1.1), its power of achievement completely relying on how good a tolerance interval devloped from (1.3) for need in (1.2). That is, inappropriateness of selecting a tolerance interval may provide inappropriate decision for problem (1.1).

A vast literature on tolerance intervals of (1.3) has been developed (see for example Wilks (1941), Wald (1943), Paulson (1943), Guttman (1970) and, for a recent review, Patel (1986)). There are deficiencies for the classical approaches in developing tolerance intervals. As noted by Bucchianico, Einmahl and Mushkudiani (2001), both the mathematically and the engineering oriented statistics textbooks hardly deal with this topic explicitly, and, if they do, the treatment is often confined to tolerance intervals for the normal distribution. This is partly because tolerance intervals can be difficult to construct for particular distributions (although nonparametric tolerance intervals based on order statistics can be obtained for particular values of the content) and, perhaps, partly because as Carroll and Ruppert (1991) suggest, the idea of conditional coverage probability is considered to be too difficult for beginning students. Besides the above deficincies, we consider one question regarding with a fundamental concept of interval selection.

Considering the whole class of tolerance intervals in (1.3) as a base, criterions of goodness, mainly modified from goodness of confidence intervals, have been introduced. Marshall (1949) and Wallis (1951) pointed out that a tolerance interval of (1.3) can be thought as that it provides an acceptance region for a test of the hypothesis that a new observation is drawn from the same distribution as that of the original sample. Goodman and Madansky (1962) has a similar argument. Comparing tolerance intervals based on criterion of expected length is the most popularly used selection technique. For normal tolerance interval, Eisenhart etc. (1947) constructed one, approximately the shortest. With the appealing property of shortest length, it is now popularly implemented in manufacturing industry and introduced in engineering texts. This criterion has also been a guide line for developing regression tolerance interval (see Goodman and Madansky (1962), Liman and Thomas (1988) and Mee et. al. (1991)). The goal that the manufacturer wants to know if there is an interval that includes at least a proprtion of acceptable measurements with a specifies confidence is clear. Our concern is that if the the testing rule of (1.4) applying on the shortest tolerance interval can achieves the goal of the manufacturer?

In statistical inferences, the general rule for determining a good technique is first setting a class of admissible, in some sense, techniques and then investigate and find the best (or an good) one with some advanced criterion from this admissible class. For examples, the admissibilities being accepted to apply in literature include unbiasedness and invariance in point estimation and restricting the type one error probability in hypothesis testing with advance criterions including variance and power, respectively. For tolerance selection, our concern come from the following reason:

Without careful determination of admissible techniques, a technique selected through any advanced criterion may be meaningless (1.5)

It is done by treating the whole class of tolerance intervals in (1.3) as an admissible class and searching the best (shortest) from it. Is this appropriate from the point of product manufacturing?

What is an appropriate sense of admissibility for tolerance interval? A $100(1-\alpha)\%$ confidence interval for a parameter θ is expected to having $100(1-\alpha)$ percentage that the sample confidence intervals in the long run will cover the unknown θ. With this, from the hypothesis testing problem in (1.1), it is not in-appropriate that a manufacturer consider the admissibility with:

A $\gamma-$ content tolerance interval with confidence $1-\alpha$ is admissible if, when H_0 is accepted through the rule in (1.4), there are at least proportion γ of acceptable measurements included with confidence $1-\alpha$. (1.6)

In this paper, we consider the following topics: (a) We introduce an explicit concept of admissibility for tolerance intervals. (b) We then show, with normal tolerance interval of Eisenhart et al. (1947) as example, that the shortest one may be in-admissible. (c) We develop a necessary and sufficient condition for a tolerance interval to be admissible, which indicates that the the confidence interval of coverage interval by Chen et al. (2005) is admissible. (d) Developing shortest admissible tolerance intervals is the final task in this paper.

2. Admissibility for Tolerance Intervals

A tolerance interval is required, for example, by a manufacturer of a mass-production item who needs to establish limits to contain at least a certain proportion of the product with high degree of confidence. Since the pioneer article by Wilks (1941), this need leads to a popular notion of tolerance interval formulated as a random interval (T_1, T_2) that satisfies

$$P_\theta\{P^\theta_{X_0}(T_1, T_2) \geq \gamma\} \geq 1 - \alpha \text{ for } \theta \in \Theta \tag{2.1}$$

which is called a γ-content tolerance interval with confidence $1-\alpha$. Our concern is, with the fact that there may have many choices, that if any proposal of interval (T_1, T_2) fulfilling (2.1) satisfies the real need of the manufacturer especially for those been popularly applied in industry? We fist consider formulating the concept of admissibility of (1.6) into an explicit form.

We say that $(a(\theta), b(\theta))$ is a γ coverage interval of variables X_0 if it satisfies

$$P_{X_0}\{(a(\theta), b(\theta))\} = \gamma \text{ for } \theta \in \Theta.$$

We introduce a concept of admissible tolerance interval.

Definition 2.1. Let (T_1, T_2) be a γ-content tolerance interval with confidence $1 - \alpha$. We call it an admissible γ-content tolerance interval with confidence $1 - \alpha$ and say that it is admissible if the following

$$P_\theta\{P^\theta_{X_0}[(T_1, T_2)] \geq \gamma, (a(\theta), b(\theta)) \subset (T_1, T_2)\} \geq 1 - \alpha \text{ for } \theta \in \Theta, \tag{2.2}$$

holds for some γ coverage interval $(a(\theta), b(\theta))$.

We have several notes in the followings:
(a) With the fact that

$$P_\theta\{P^\theta_{X_0}[(T_1, T_2)] \geq \gamma, (a(\theta), b(\theta)) \subset T_1, T_2)\} \leq$$
$$P_\theta\{P^\theta_{X_0}[(T_1, T_2)] \geq \gamma\} \text{ for } \theta \in \Theta,$$

a γ-content tolerance interval with confidence $1 - \alpha$ is not guaranteed to be an admissible γ-content tolerance interval with the same confidence. For any in-admissible tolerance interval (T_1, T_2), if its observation (t_1, t_2) is contained in specification limit interval (LSL, USL), there is no assurance with confidence that other observation intervals containing γ percentage of measurements are with acceptable measurements of γ percentage or more.

(b) If a γ-content tolerance interval with confidence $1-\alpha$ is not admissible, then it must be admissible for some other confidence smaller than $1-\alpha$.
(c) The admissibility has to be accompanied with a specified couple $\{\gamma, 1-\alpha\}$. Otherwise, every random interval is an admissible tolerance interval. When (T_1, T_2) is an admissible γ-content tolerance interval with confidence $1-\alpha$ for some γ coverage interval, there may have others (may be infinite) alternative γ coverage intervals $(a(\theta), b(\theta))$ such that (T_1, T_2) is still an admissible tolerance interval for these coverage intervals.
(d) If there is a random interval (T_1, T_2) satisfies (2.2), is it an admissible γ-content tolerance interval with confidence $1-\alpha$? The answer is yes through the the fact that

$$\{P_{X_0}[(T_1, T_2)] \geq \gamma, (a(\theta), b(\theta)) \subset (T_1, T_2)\} \subset \{P_{X_0}[(T_1, T_2)] \geq \gamma\}.$$

The general theory of developing tolerance interval in literature is fixing percentages γ and $1-\alpha$ to select a statistic $T(k)$ with factor k and search k_1^* and k_2^* such that $(T(k_1^*), T(k_2^*))$ solves the following minimization problem:

$$\text{argmin}_{0<k_1<k_2}[T(k_2) - T(k_1) : P_\theta\{P_{X_0}[(T(k_1), T(k_2)] \geq \gamma\} \geq 1-\alpha]. \quad (2.3)$$

For review of examples of choosing statistic T, see Patel (1986).

Definition 2.2. Let $(T(k_1^*), T(k_2^*))$ solves the problem of (2.3). We then call it the shortest γ-content tolerance interval with confidence $1-\alpha$.

We are arguing that a shortest tolerance interval may be meaningless for that it may be in-admissible. We will study this point with normal distribution as an example. Suppose that the normal mean μ and variance σ^2 are known. A γ-content tolerance interval with 100% confidence is

$$(\mu - z_{\frac{1+\gamma}{2}}\sigma, \mu + z_{\frac{1+\gamma}{2}}\sigma). \quad (2.4)$$

Suppose that the interval of (2.4) is contained in specification limit interval (LSL, USL). We then assure with 100% confidence that it covered at proportion γ of acceptable measurements. Actually any one with $\delta \in (0, 1-\gamma)$ in the following

$$(\mu + z_\delta\sigma, \mu + z_{\gamma+\delta}\sigma) \quad (2.5)$$

is a γ-content tolerance interval with 100% confidence. So, when interval in (2.4) is not contained in specification interval (LSL, USL), we can not be sure that there is no γ-content tolerance interval with 100% confidence since there may have other one in (2.5) contained in specification interval. Determining appropriate γ coverage is somehow more appropriate as an engineering problem.

Suppose that now we have a normal random sample $X_1, ..., X_n$ from normal distribution $N(\mu, \sigma^2)$ where mean μ and standard deviation σ are both unknown. Wald and Wolfowitz (1946) first introduced the normal tolerance interval of the form

$$(\bar{X} - kS, \bar{X} + kS) \tag{2.6}$$

where value k meets the requirement (2.1) for pre-assigned γ, $1 - \alpha$ and sample size n. The development of the shortest tolerance interval involves the distribution of $\Phi(\bar{X} + kS) - \Phi(\bar{X} - kS)$ which is extremely complicated as indicated by Guttman (1970). With this difficulty, the shortest tolerance interval $(\bar{X} - k^*S, \bar{X} + k^*S)$ has not been able to provide an explicit formulation. However, among many authors, Eisenhart et al. (1947) provides one with length approximately shortest. We will study this approximate one latter for its admissibility in Section 3.

The study of admissibility of Eisenhart et al.'s shortest tolerance interval $(\bar{X} - k^*S, \bar{X} + k^*S)$ is important since the shortest one is generally accepted the most interesting technique in literature in developing tolerance interval. The aim of the rest in this section is to show the in-admissibility of it in this normal case. To do this, we first show that among γ coverages $(\mu + z_\delta\sigma, \mu + z_{\gamma+\delta}\sigma)$ for $0 < \delta < 1 - \gamma$ the one achieves the maximum confidence when the coverage is the symmetric one with $\delta = \frac{1-\gamma}{2}$.

Theorem 2.3. Let $C_\delta(\gamma) = (\mu + z_\delta\sigma, \mu + z_{\gamma+\delta}\sigma)$. Then

$$\frac{1-\gamma}{2} = \text{argmax}_{0<\delta<1-\gamma} P_{\mu,\sigma}\{P_{X_0}[(\bar{X} - kS, \bar{X} + kS)] \geq \gamma,$$
$$C_\delta(\gamma) \subset (\bar{X} - kS, \bar{X} + kS)\}.$$

Proof. We know that

$$P_{\mu,\sigma}\{P_{X_0}[(\bar{X}-kS,\bar{X}+kS)]\geq\gamma, C_\delta(\gamma)\subset(\bar{X}-kS,\bar{X}+kS)\}$$
$$=P_{\mu,\sigma}\{\Phi(\bar{X}+kS)-\Phi(\bar{X}-kS)\geq\gamma, C_\delta(\gamma)\subset(\bar{X}-kS,\bar{X}+kS)\}$$
$$=E_{\mu,\sigma}\{P_{\mu,\sigma}[\Phi(\bar{X}+kS)-\Phi(\bar{X}-kS)\geq\gamma, C_\delta(\gamma)\subset(\bar{X}-kS,\bar{X}+kS)|S]\}.$$

We see that, for given $S=s$,

$$P_{\mu,\sigma}\{\Phi(\bar{X}+ks)-\Phi(\bar{X}-ks)\geq\gamma, C_\delta(\gamma)\subset(\bar{X}-ks,\bar{X}+ks)\} \tag{2.7}$$
$$=P_{\mu,\sigma}\{\Phi(\bar{X}+ks)-\Phi(\bar{X}-ks)\geq\gamma, \mu+z_{\gamma+\delta}\sigma-ks<\bar{X}<\mu+z_\delta\sigma+ks\}$$

where Φ is the distribution function of the standard normal distribution $N(0,1)$. Then minimizing the probability of (2.7) on $\delta\in(0,1-\gamma)$ may be done by maximizing the probability $P_{\mu,\sigma}(z_{\gamma+\delta}\sigma-ks\leq\bar{X}-\mu\leq z_\delta\sigma+ks)$ given that value k satisfies $P_{\mu,\sigma}(\Phi(\bar{X}+ks)-\Phi(\bar{X}-ks)\geq\gamma)$. This implies that $\phi_\sigma(z_\delta\sigma+ks)=\phi_\sigma(z_{\gamma+\delta}\sigma-ks)$ where ϕ_σ is the pdf of a normal distribution $N(0,\sigma^2)$ which further indicates that

$$P_{\mu,\sigma}\{\Phi(\bar{X}+ks)-\Phi(\bar{X}-ks)\geq\gamma, C_\delta(\gamma)\subset(\bar{X}-ks,\bar{X}+ks)\}\leq$$
$$P_{\mu,\sigma}\{\Phi(\bar{X}+ks)-\Phi(\bar{X}-ks)\geq\gamma, C(\gamma)\subset(\bar{X}-ks,\bar{X}+ks)\}. \tag{2.8}$$

The theorem is followed from (2.8) associated with the fact that $\bar{X}$ has a symmetric distribution. □

With this result, we evaluate the maximum confidence for that a tolerance interval of (2.5) could be requiring only to check if for γ coverage $C(\gamma)$.

Lemma 2.4. For given $k>0$,

$$P_{\mu,\sigma}\{P_{X_0}[(\bar{X}-kS,\bar{X}+kS)]\geq\gamma\}>$$
$$P_{\mu,\sigma}\{P_{X_0}[(\bar{X}-kS,\bar{X}+kS)]\geq\gamma, C(\gamma\subset(\bar{X}-kS,\bar{X}+kS)\}.$$

Proof. By letting

$$A(k)=\{\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix}:\int_{\bar{x}-ks}^{\bar{x}+ks}\phi_{\mu,\sigma}(x)dx\geq\gamma\}\text{ and}$$

$$A_{C(\gamma)}(k)=\{\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix}:\int_{\bar{x}-ks}^{\bar{x}+ks}\phi_{\mu,\sigma}(x)dx\geq\gamma, C(\gamma)\subset(\bar{x}-ks,\bar{x}+ks)\}.$$

Obviously we have $A_{C(\gamma)}(k) \subset A(k)$. For fixed $k > 0$ and (μ, σ), there exists $x_1, ..., x_n$ such that $\bar{x} - ks = \mu - z_{\frac{1+\gamma}{2}}\sigma$ and $\bar{x} + ks = \mu + z_{\frac{1+\gamma}{2}}\sigma$. Then, this vector $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ is in $A(k)$. We further let

$$A^*(k) = \{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} : \bar{x} - ks < \mu - z_{\frac{1+\gamma}{2}}\sigma \text{ and } \bar{x} + ks < \mu + z_{\frac{1+\gamma}{2}}\sigma \text{ with}$$
$$P_{X_0}[(\bar{x} - ks, \bar{x} + ks)] = \gamma\}.$$

With normal distribution, we have $P_{\mu,\sigma}(\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \in A^*(k)) = P_{X_0}[(-\infty, \mu - z_{\frac{1+\gamma}{2}}\sigma)] = \frac{1-\gamma}{2} > 0$. This implies that

$$P_{\mu,\sigma}(\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \in A(k)) > P(\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \in A_{C(\gamma)}(k)) \tag{2.9}$$

with the facts that $A^*(k) \subset A(k)$ and $A^*(k) \cap A_{C(\gamma)}(k) = \phi$. $\square$

We are now ready to state the main theorem about in-admissibility of the popularly used the shortest normal tolerance interval.

Theorem 2.5. The shortest γ-content tolerance interval of the form $(\bar{X} - k^*S, \bar{X} + k^*S)$ with confidence $1 - \alpha$ is not admissible.
Proof. From (2.9),

$$P(A_{C(\gamma)}(k)) < 1 - \alpha \text{ for } k \leq k^*.$$

This tolerance interval $(\bar{X} - k^*S, \bar{X} + k^*S)$ is not admissible in sense of (2.2). $\square$

3. Admissibility Verification of Tolerance Intervals Through Simulation

With showing that the shortest normal tolerance interval is in-admissible, it is worth in examining the popularly used normal tolerance intervals to see if they are admissible. As noted by Guttman (1970,p59), the two sided normal tolerance intervals involves the distribution of the coverage

$$\Phi(\bar{X} + kS) - \Phi(\bar{X} - kS)$$

which is exceedingly complicated so that numerical approximation or simulation are generally used to develop them. The most popularly used normal tolerance interval is the one compute by Eisenhart et al. (1947) which is considered the approximately shortest. We perform two simulations to study this normal tolerance intervals. First, we simulate its role as a tolerance interval of (2.1). Second, we simulate its admissibilty by computing its approximate confidence of a coverage interval. If the approximate confidences are smaller than $1 - \alpha$'s, they are approximately in-admissible.

We select values k from the table developed by Eisenhart et al. We perform the simulation in two steps. First, we want to evaluate the confidence that the proposed tolerance interval covers the sample space of the underlying distribution, i.e., we want to see how close that it achieves the shortest tolerance interval of (2.1). Everytime we select n random sample from normal distribution for a fixed mean μ and variance σ^2. We perform this simulation with replication $m = 100,000$. Let $\bar{x}_j$ and s_j^2 be the sample mean and sample variance for the jth sample. The simulated confidence is approximated by

$$\frac{1}{m}\sum_{j=1}^{m} I(\Phi(\bar{x}_j + ks_j) - \Phi(\bar{x}_j - ks_j) \geq \gamma). \tag{3.1}$$

where $\Phi_{\mu,\sigma}$ is the distribution function of the underlying normal distribution. For $\gamma = 0.9, 0.95, 0.99$, $1 - \alpha = 0.9, 0.95, 0.99$ and $n = 10, 30, 50$, we display the simulated results in Table 1.

Table 1. Confidence for normal tolerance intervals in covering future variable

	$1-\alpha=0.9$	$1-\alpha=0.95$	$1-\alpha=0.99$
$\gamma=0.9$			
$n=10$	0.8962	0.9491	0.9887
$n=30$	0.8976	0.9482	0.9892
$n=50$	0.8976	0.9482	0.9900
$\gamma=0.95$			
$n=10$	0.8970	0.9491	0.9899
$n=30$	0.8978	0.9432	0.9895
$n=50$	0.8985	0.9488	0.9891
$\gamma=0.99$			
$n=10$	0.9003	0.9508	0.9896
$n=30$	0.8985	0.9499	0.9897
$n=50$	0.8995	0.9489	0.9903

Table 1 indicates that there are sample intervals (t_1, t_2) with confidence close to $1-\alpha$ containing γ percentage or more of measurements, however, these specified measurements are not sured to lie in some fixed covearge interval. We further to verify this point.

We now concern the question if it is appropriate treated as a γ-content coverage interval based tolerance interval with confidence $1-\alpha$? We perform the simulation in the same assumptions for results in Table 1. The approximate confidence for it playing a role of confidence interval of the coverage interval $C(\gamma) = (\mu - z_{\frac{1+\gamma}{2}}\sigma, \mu + z_{\frac{1+\gamma}{2}}\sigma)$ is defined as

$$\frac{1}{m}\sum_{j=1}^{m} I(\Phi(\bar{x}_j + ks_j) - \Phi(\bar{x}_j - ks_j) \geq \gamma, C(\gamma) \subset (\bar{x}_j - ks_j, \bar{x}_j + ks_j)).$$

Table 2. Confidence for normal tolerance intervals in covering prediction interval $(\mu - z_{\frac{1+\gamma}{2}}\sigma, \mu + z_{\frac{1+\gamma}{2}}\sigma)$

	$1-\alpha=0.9$	$1-\alpha=0.95$	$1-\alpha=0.99$
$\gamma=0.9$			
$n=10$	0.8197	0.8989	0.9765
$n=30$	0.7736	0.8648	0.9617
$n=50$	0.7595	0.8518	0.9561
$\gamma=0.95$			
$n=10$	0.8403	0.9129	0.9812
$n=30$	0.8032	0.8756	0.9700
$n=50$	0.7917	0.8773	0.9648
$\gamma=0.99$			
$n=10$	0.8627	0.9278	0.9849
$n=30$	0.8356	0.9097	0.9776
$n=50$	0.8268	0.9020	0.9762

For the simulation results in a design of given $1-\alpha$, the evaluated confidences are fluctuant in sample size n and coverage γ. The deficits could be large as the cases in $1-\alpha=0.9$ and n is large. If we expect the Eisenhart et al.'s tolerance interval to be admissible, the answer is not promissing. This verifies the result in Theorem 2.5 that the shortest tolerance interval may be in-admissible.

4. Coverage Interval Based Tolerance Intervals

With the example of normal tolerance interval of Wald and Wolfowitz (1946), it is known that not every tolerance interval of (2.1) of Wilks (1941) is admissible. Then, is there a general technique in developing tolerance interval which ensures the property of admissibility?

Chen, Huang and Welsh (2005) introduced a tolerance interval which is a $100(1-\alpha)\%$ confidence interval of a special type coverage interval (mode type interval). We extend this concept in a general setting. We say that a random interval (T_1, T_2) is a $100(1-\alpha)\%$ confidence interval of a coverage interval $(a(\theta), b(\theta))$ if it satisfies

$$1-\alpha = P_\theta\{T_1 \le a(\theta) < b(\theta) \le T_2\} \text{ for } \theta \in \Theta \tag{4.1}$$

When either $t_1(X) = a(\theta) = c$ or $t_2(X) = b(\theta) = c$ holds for some constant c including $-\infty$ or ∞ we have one sided $100(1-\alpha)\%$ C.I. of γ-content coverage interval. The interest of this confidence interval here is with γ coverage.

Definition 4.1. If a γ-content tolerance interval with confidence $1-\alpha$ is a $100(1-\alpha)\%$ confidence interval of some γ coverage interval, then we call it a γ-content coverage interval based tolerance interval with confidence $1-\alpha$.

The following theorem states a necessary and sufficient condition for that a tolerance interval is admissible which also addresses a connection of the tolerance interval with the confidence interval of a coverage interval.

Theorem 4.2. A random interval (T_1, T_2) is an admissible γ-content tolerance interval with confidence $1-\alpha$ if and only if it is a $100(1-\alpha)\%$ confidence interval of a γ coverage interval.

Proof. Let (T_1, T_2) be a $100(1-\alpha)\%$ confidence interval of γ coverage interval $(a(\theta), b(\theta))$, i.e.,

$$P_\theta\{(a(\theta), b(\theta)) \subset (t_1(X), t_2(X))\} \geq 1-\alpha. \tag{4.2}$$

For $X = x$ subjected to $(a(\theta), b(\theta)) \subset (t_1(X), t_2(X))$, we have

$$P_{X_0}\{(t_1(x), t_2(x)\} \geq P_{X_0}\{(a(\theta), b(\theta))\} \geq \gamma.$$

This leads to

$$\begin{aligned}&\{(a(\theta), b(\theta)) \subset (t_1(X), t_2(X))\}\\ &\subset \{P_{X_0}[(t_1(X), t_2(X))] \geq \gamma, (a(\theta), b(\theta)) \subset (t_1(X), t_2(X))\}.\end{aligned}$$

Henceful, from (4.2),

$$P_\theta\{P_{X_0}[(t_1(X), t_2(X))] \geq \gamma, (a(\theta), b(\theta)) \subset (t_1(X), t_2(X))\} \geq 1-\alpha$$

indicating the admissibility of (T_1, T_2).

On the other hand, let (T_1, T_2) be an admissible γ-content tolerance interval t γ coverage interval $(a(\theta), b(\theta))$ with confidence $1-\alpha$.

$$\begin{aligned}&P_\theta\{(a(\theta), b(\theta)) \subset (t_1(X), t_2(X))\} \geq\\ &P_\theta\{P_{X_0}[(t_1(X), t_2(X))] \geq \gamma, (a(\theta), b(\theta)) \subset (t_1(X), t_2(X))\} \geq 1-\alpha.\end{aligned}$$

This shows that (T_1, T_2) is a $100(1-\alpha)\%$ confidence interval of $(a(\theta), b(\theta))$ and the direction $(\Rightarrow)$ of the proof is done. $\square$

When we are interesting in admissible tolerance interval, the confidence interval of a coverage interval guranntees in achieving the goal of admissibility. With this, the confidence interval of mode coverage interval is appealing in sense of shortest length of coverage interval. If people are interesting in other type of coverage intervals such as the symmetric ones, they may be easily established from the lines of Chen, Huang and Welsh (2006).

Owen (1964) argued that most tolerance intervals developed for normal distribution are set up so that the percentage nondefective is controlled $100\gamma\%$, and hence the defectiveness could be all be in one tail. Then he consider a normal tolerance interval such that no more than the proportion $\frac{1-\gamma}{2}$ is below the lower tolerance limit and no more than the proportion $\frac{1-\gamma}{2}$ is above the upper tolerance limit. Extension from his idea, we may expect a γ-content tolerance interval (T_1, T_2) with confidence $1-\alpha$ that satisfies

$$P[P(X_0 \le T_1|X) \le \frac{1-\gamma}{2} \text{ and } P(X_0 \ge T_2|X) \le \frac{1-\gamma}{2}] \ge 1-\alpha. \quad (4.3)$$

Theorem 4.3. Let $(a(\theta), b(\theta))$ be with $P(X_0 \le a(\theta)) \le \frac{1-\gamma}{2}$ and $P(X_0 \ge b(\theta)) \le \frac{1-\gamma}{2}$. Then the γ-content tolerance interval as a $100(1-\alpha)\%$ confidence interval of $(a(\theta), b(\theta))$ satisfies the Owen's restriction (4.3).
Proof. It is induced from the followingS:

$$\begin{aligned}
&P_\theta\{P_\theta[X_0 \le T_1|X] \le \frac{1-\gamma}{2} \text{ and } P[X_0 \ge T_2|X] \le \frac{1-\gamma}{2}\} \\
&= P_\theta\{P_\theta[X_0 \le T_1|X] \le F_{X_0}(a(\theta)) \text{ and } P_\theta[X_0 \ge T_2|X] \le 1 - F_{X_0}(b(\theta))\} \\
&= P_\theta\{F_{X_0}(T_1) \le F_{X_0}(a(\theta)) \text{ and } 1 - F_{X_0}(T_2) \le 1 - F_{X_0}(b(\theta))\} \\
&= P_\theta\{F_{X_0}(T_1) \le F_{X_0}(a(\theta)) \text{ and } F_{X_0}(b(\theta)) \le F_{X_0}(T_2)\} \\
&= P_\theta\{F_{X_0}(T_1) \le F_{X_0}(a(\theta)) < F_{X_0}(b(\theta)) \le F_{X_0}(T_2)\} \\
&\ge 1-\alpha.
\end{aligned}$$

5. Rate of Confidence Accomplishment for Tolerance Intervals

For given γ and $1-\alpha$, there may have many admissible γ-content tolerance intervals all at confidence $1-\alpha$. How can we choose one from this interval class? It is now reasonable to apply the criterion of length for making decision of selecting tolerance interval. Suppose that there exists a shortest one in this class. We may call it the shortest admissible γ-content tolerance interval at confidence $1-\alpha$. At this moment, we are not going to investigate the question that if it exists or what is it? We want to introduce an index evaluating how close the admissibilities that the existed tolerance intervals are.

Definition 5.1. Suppose that a γ-content tolerance interval (T_1, T_2) at confidence $1-\alpha$ satisfies, for some γ-content coverage interval $(a(\theta), b(\theta))$,

$$P_\theta\{P_{X_0}(T_1 \le a(\theta) < b(\theta) \le T_2)|X) \ge \gamma\} = 1-\alpha^* \text{ for } \theta \in \Theta. \qquad (5.1)$$

Then the resulted $1-\alpha^*$ is called the retrieved confidence of this tolerance interval. We further say that a tolerance interval with retrieved confidence $1-\alpha^*$ is proper if $1-\alpha^* = 1-\alpha$, conservative if $1-\alpha^* > 1-\alpha$ and exaggerative if $1-\alpha^* < 1-\alpha$.

We consider the simple situation that probability on the left hand of (4.1) is uniformly equal to $1-\alpha^*$ for $\theta \in \Theta$. Like the confidence interval for a papramieter, this may be done if this probability is developed from two pivotal quantities one based on T_1 and $a(\theta)$ and one based on T_2 and $b(\theta)$ where the examples that we will introduce in this paper are all satisfied this restriction. Among the γ-content tolerance intervals at confidence $1-\alpha$, we classify them into three classes. This provides the manufacturer more precise information about the capability of the manufacturing process. For example of Wald and Wolfowitz's normal tolerance interval, if $(\bar{X}-kS, \bar{X}+kS)$ is a γ-content tolerance interval with retrieved confidence $1-\alpha$, then any $(\bar{X}-k^*S, \bar{X}+k^*S)$ is also a γ-content tolerance interval with retrieved confidence $> 1-\alpha$ for any $k^* > k$. We definitely do not want a tolerance interval to be too exaggerative and too conservetive. Otherwise, it losses too much information about the quality of the manufacturing process. With this concern, displaying an index measuring the degree of either conservertiveness and exaggeration is appropriate.

Definition 5.2. Suppose that we have a γ-content tolerance interval at confidence $1-\alpha$ and with retrieved confidence $1-\alpha^*$. We define the rate of confidence accomplishment for this tolerance interval as

$$RCA = \frac{1-\alpha^*}{1-\alpha}.$$

A γ-content tolerance interval at confidence $1-\alpha$ is proper if $RCA = 1$, conservative if $RCA > 1$ and exaggerative if $RCA < 1$. We are more interesting to see the size RCA for tolerance interval. We want to investigate the corresponding values of RCA for the classical and the coverage interval based tolerance intervals. From the proof of Theorem 4.2, if a γ-content tolerance interval at confidence $1-\alpha$ is a confidence interval of a γ-content coverage interval at confidence $1-\alpha$ then it is either a proper or conservative tolerance interval . Basically a conservative tolerance interval is not a serious problem if RCA is not larger than 1 too much. On the other hand, we like to investigate if a classical tolerance interval is conservative or exaggerative and how far RCA is from 1.

Suppose that we have a normal random sample from distribution $N(\mu, \sigma^2)$ parameters μ and σ unknown. Let's study RCA's for the Wald and Wolfowitz's tolerance intervals $(\bar{X}-kS, \bar{X}+kS)$. First, we consider the approximate shortest tolerance interval by Eisenhart et al. (1947) for reason of its popularity in receiving greatest attention in literature and applications. We perform the same simulation as it stated in Section 3 and the estimated RCA's are listed in Table 3.

Table 3. Retrieved confidence for Eisenhart et al.'s tolerance interval

	$1-\alpha=0.9$	$1-\alpha=0.95$	$1-\alpha=0.99$
$\gamma=0.9$			
$n=10$	0.9108	0.9462	0.9863
$n=30$	0.8595	0.9103	0.9714
$n=50$	0.8439	0.8966	0.9657
$\gamma=0.95$			
$n=10$	0.9336	0.9610	0.9911
$n=30$	0.8924	0.9217	0.9798
$n=50$	0.8797	0.9235	0.9745
$\gamma=0.99$			
$n=10$	0.9586	0.9767	0.9949
$n=30$	0.9284	0.9576	0.9875
$n=50$	0.9187	0.9495	0.9861

For this situation that parameters μ and σ are both unknown, Chen, Huang and Welsh (2005) showed that

$$(\bar{X} - t_{1-\frac{\alpha}{2}}(n-1, \sqrt{n}z_{\frac{1+\gamma}{2}})\frac{S}{\sqrt{n}}, \bar{X} - t_{1-\frac{\alpha}{2}}(n-1, -\sqrt{n}z_{\frac{1+\gamma}{2}})\frac{S}{\sqrt{n}}) \qquad (5.2)$$

is a $100(1-\alpha)\%$ C.I. for the coverage interval $(\mu - z_{\frac{1+\gamma}{2}}\sigma, \mu + z_{\frac{1+\gamma}{2}}\sigma)$ and then it is also a γ-content tolerance interval at confidence $1-\alpha$. It is then interesting to evaluate its rate of confidence accomplishmenet. We list the simulation results in the following table.

Table 4. Retrieved confidence (RCA) for coverage interval based tolerance interval when μ and σ are both uknown

	$1-\alpha=0.9$	$1-\alpha=0.95$	$1-\alpha=0.99$
$\gamma=0.9$			
$n=10$	0.9127 (1.0141)	0.9555 (1.0058)	0.9911 (1.0011)
$n=30$	0.9079 (1.0088)	0.9531 (1.0033	0.9904 (1.0004)
$n=50$	0.9076 (1.0084)	0.9523 (1.0025)	0.9905 (1.0005)
$\gamma=0.95$			
$n=10$	0.9174 (1.0194)	0.9581 (1.0085)	0.9918 (1.0019)
$n=30$	0.9099 (1.0110)	0.9550 (1.0052)	0.9908 (1.0008)
$n=50$	0.9114 (1.0126)	0.9545 (1.0047)	0.9901 (1.0001)
$\gamma=0.99$			
$n=10$	0.9213 (1.0237)	0.9609 (1.0114)	0.9924 (1.0025)
$n=30$	0.9167 (1.0185)	0.9575 (1.0079	0.9909 (1.0009)
$n=50$	0.9168 (1.0186)	0.9574 (1.0078)	0.9914 (1.0014)

Basically when a closed form of a γ-content tolerance interval at confidence $1-\alpha$ is available to derive the resulted ones are with $RCA=1$. However, when a closed form is not able to derive so that approximation or simulation is done, the resulted ones are with remarkably far from 1. That is, although there is confidence $1-\alpha$ with resulted interval (t_1, t_2) that covers X_0 with probability γ, a significant part of them do not cover a desired set of acceptable product. We alos conducted a simulation for one sided tolerance intervals where we found that the evaluated rate of accomplishments are all very close to 1. This indicates the inconsistency in developing the tolerance intervals.

Gaussian distribution with known variance

Suppose that the underlying distribution is $N(\mu, \sigma^2)$ with σ known. Owen (1964) obatined a γ-content tolerance interval at confidence $1-\alpha$ as, by letting $\hat{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$,

$$(\hat{X} - z_{\frac{1+\gamma}{2}}\sigma - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \hat{X} + z_{\frac{1+\gamma}{2}}\sigma + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}). \qquad (5.3)$$

In the following, we show that (5.3) is a $100(1-\alpha)\%$ C.I. of some γ-content coverage interval.

$$P(\hat{X} - z_{\frac{1+\gamma}{2}}\sigma - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le \mu - z_{\frac{1+\gamma}{2}}\sigma < \mu + z_{\frac{1+\gamma}{2}}\sigma \le \hat{X} + z_{\frac{1+\gamma}{2}}\sigma + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}})$$
$$= P(-z_{\frac{1+\gamma}{2}}\sigma - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le \hat{X} - \mu - z_{\frac{1+\gamma}{2}}\sigma < \hat{X} - \mu + z_{\frac{1+\gamma}{2}}\sigma \le z_{\frac{1+\gamma}{2}}\sigma + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}})$$
$$\ge P(-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le \hat{X} - \mu \le z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}})$$
$$= P(-z_{1-\frac{\alpha}{2}} \le \frac{\hat{X} - \mu}{\sigma/\sqrt{n}} \le z_{1-\frac{\alpha}{2}})$$
$$= 1 - \alpha.$$

This shows that the γ-content tolerance interval at confidence $1-\alpha$ of (5.3) is a $100(1-\alpha)\%$ C.I. of the γ-content coverage interval $(\mu - z_{\frac{1+\gamma}{2}}\sigma, \mu + z_{\frac{1+\gamma}{2}}\sigma)$. Suppose that from the specialist the interval $(\mu - z_{\frac{1+\gamma}{2}}\sigma, \mu + z_{\frac{1+\gamma}{2}}\sigma)$ contains a part of acceptable product. Then this tolerance interval of Owen (1964) is with confidence $1-\alpha$ the sample tolerance interval $(\bar{x} - z_{\frac{1+\gamma}{2}}\sigma - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \hat{x} + z_{\frac{1+\gamma}{2}}\sigma + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}})$ contains acceptable product with percentage γ. This should be satisfactory with the manufacturer.

Morever, Jilek and Likar (1960a) established the γ-content one sided tolerance intervals at confidence $1-\alpha$ for normal distribution with known variance and unknown mean. We may also analogously proved that they are also $100(1-\alpha)\%$ C.I.'s of some γ-content coverage intervals. We combine these results and it for two sided case in the following table.

Table 5. γ-content statistical tolerance interval at confidence $1-\alpha$ as a $100(1-\alpha)\%$ C.I. of γ-content coverage interval when σ is known

γ-content coverage interval	statistical coverage interval
$(\mu - z_\gamma\sigma, \infty)$	$(\hat{X} - z_\gamma\sigma - z_{1-\alpha}\frac{\sigma}{\sqrt{n}}, \infty)$
$(-\infty, \mu + z_\gamma\sigma)$	$(-\infty, \hat{X} + z_\gamma\sigma + z_{1-\alpha}\frac{\sigma}{\sqrt{n}})$
$(\mu - z_{\frac{1+\gamma}{2}}\sigma, \mu + z_{\frac{1+\gamma}{2}}\sigma)$	$(\hat{X} - z_{\frac{1+\gamma}{2}}\sigma - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \hat{X} + z_{\frac{1+\gamma}{2}}\sigma + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}})$

Table 6. Retrieved confidence and RCA $\begin{pmatrix} 1-\alpha^* \\ RCA \end{pmatrix}$ for one sided tolerance interval when σ is known

	$1-\alpha = 0.9$	$1-\alpha = 0.95$	$1-\alpha = 0.99$
$\gamma = 0.9$ $n = 10$	0.9008 (1.0008)	0.9503 (1.0003)	0.9908 (1.0008)
$n = 30$	0.8998 (0.9997)	0.9500 (1.0000)	0.9896 (0.9996)
$n = 50$	0.8992 (0.9991)	0.9481 (0.9980)	0.9901 (1.0001)
$\gamma = 0.95$ $n = 10$	0.9014 (1.0016)	0.9488 (0.9987)	0.9896 (0.9996)
$n = 30$	0.8980 (0.9978)	0.9511 (1.0011)	0.9897 (0.9997)
$n = 50$	0.9020 (1.0023)	0.9498 (0.9997)	0.9902 (1.0002)
$\gamma = 0.99$ $n = 10$	0.8992 (0.9991)	0.9499 (0.9999)	0.9898 (0.9998)
$n = 30$	0.8997 (0.9997)	0.9514 (1.0015)	0.9902 (1.0002)
$n = 50$	0.8991 (0.9990)	0.9501 (1.0001)	0.9904 (1.0004)

Let $X_1, ..., X_n$ be a random sample from normal distribution $N(\mu, \sigma^2)$ where μ is known but σ is unknown. Denoting by $\chi^2_\alpha(n-1)$ if $P(\chi^2(n-1) \le \chi^2_\alpha(n-1)) = \alpha$ and $S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$, Jilek and Likar (1960b) and Owen

(1964) both show that

$$(\mu - z_{\frac{1+\gamma}{2}}\sqrt{\frac{n-1}{\chi^2_\alpha(n-1)}}S, \mu + z_{\frac{1+\gamma}{2}}\sqrt{\frac{n-1}{\chi^2_\alpha(n-1)}}S) \tag{5.4}$$

is a two sided γ-content tolerance interval with confidence $1-\alpha$. In the following, we show that it is a $100(1-\alpha)\%$ C.I. for some γ-content coverage interval,

$$P(\mu - z_{\frac{1+\gamma}{2}}\sqrt{\frac{n-1}{\chi^2_\alpha(n-1)}}S \le \mu - z_{\frac{1+\gamma}{2}}\sigma < \mu + z_{\frac{1+\gamma}{2}}\sigma \le \mu$$
$$+ z_{\frac{1+\gamma}{2}}\sqrt{\frac{n-1}{\chi^2_\alpha(n-1)}}S)$$
$$= P(-z_{\frac{1+\gamma}{2}}\sqrt{\frac{n-1}{\chi^2_\alpha(n-1)}}S \le -z_{\frac{1+\gamma}{2}}\sigma < z_{\frac{1+\gamma}{2}}\sigma \le z_{\frac{1+\gamma}{2}}\sqrt{\frac{n-1}{\chi^2_\alpha(n-1)}}S)$$
$$= P(-\frac{\sqrt{n-1}}{\sigma}S \le -\sqrt{\chi^2_\alpha(n-1)} < \sqrt{\chi^2_\alpha(n-1)} \le \frac{\sqrt{n-1}}{\sigma}S)$$
$$(\text{by the fact that } \frac{\sqrt{n-1}}{\sigma}S \text{ and } \sqrt{\chi^2_\alpha(n-1)}) > 0)$$
$$= P(\frac{(n-1)S^2}{\sigma^2} \ge \chi^2_\alpha(n-1))$$
$$= 1-\alpha.$$

Jilek and Likar (1960b) also considered one sided γ-content tolerance intervals with confidence $1-\alpha$ for case that μ is known and σ is unknown. We combine these results in the following table.

Table 7. Retrieved confidence for two sided tolerance interval when μ is known

	$1-\alpha=0.9$	$1-\alpha=0.95$	$1-\alpha=0.99$
$\gamma=0.9$			
$n=10$	0.9001	0.9510	0.9899
$n=30$	0.8998	0.9514	0.9896
$n=50$	0.9004	0.9515	0.9896
$\gamma=0.95$			
$n=10$	0.9002	0.9508	0.9900
$n=30$	0.9008	0.9501	0.9904
$n=50$	0.8995	0.9501	0.9895
$\gamma=0.99$			
$n=10$	0.8989	0.9500	0.9903
$n=30$	0.9007	0.9490	0.9896
$n=50$	0.9006	0.9508	0.9902

Table 8. γ-content statistical tolerance interval at confidence $1-\alpha$ as a $100(1-\alpha)\%$ C.I. of γ-content coverage interval when μ is known

γ-content coverage interval	statistical coverage interval
$(\mu - z_\gamma\sigma, \infty)$	$(\mu - z_\gamma\sqrt{\frac{n-1}{\chi^2_\alpha(n-1)}}S, \infty)$
$(-\infty, \mu + z_\gamma\sigma)$	$(-\infty, \mu + z_\gamma\sqrt{\frac{n-1}{\chi^2_\alpha(n-1)}}S)$
$(\mu - z_{\frac{1+\gamma}{2}}\sigma, \mu + z_{\frac{1+\gamma}{2}}\sigma)$	$(\mu - z_{\frac{1+\gamma}{2}}\sqrt{\frac{n-1}{\chi^2_\alpha(n-1)}}S, \mu + z_{\frac{1+\gamma}{2}}\sqrt{\frac{n-1}{\chi^2_\alpha(n-1)}}S)$

Theorem 5.3. The one sided tolerance interval of Jilek and Likar (1960b) and Odeh and Owen (1980) and the two sided tolerance interval of Jilek and Likar (1960b) are all with $RCA = 1$.

6. Shortest Admissible Tolerance Intervals

With defining the concept of admissibility of tolerance interval, it is then interesting in developing the shortest (expected shortest) one for the class of admissible tolerance intervals if it does exist. How can we accomplish this task. Let $Q(T, \theta)$ be an appropriate pivotal quantity that may be inverted for deriving the confidence interval of a coverage interval $C(\gamma) = (c_1, c_2)$

through the following

$$1-\alpha = P\{q_1 \le Q(T,\theta) \le q_2\}$$
$$= P\{Q_1(T,c_1,c_2,q_1,q_2) \le c_1 < c_2 \le Q_2(T,c_1,c_2,q_1,q_2)\}.$$

We then have choices of c_1, c_2, q_1, q_2 that minimizes the length (or expected length) $Q_2(T,c_1,c_2,q_1,q_2) - Q_1(T,c_1,c_2,q_1,q_2)$ where

$$\{(Q_1(T,c_1,c_2,q_1,q_2), Q_2(T,c_1,c_2,q_1,q_2)) : P_{X_0}(C(\gamma)) = \gamma,$$
$$1-\alpha = P\{q_1 \le Q(T,\theta) \le q_2\}\}$$

is the class of admissible $100(1-\alpha)\%$ tolerance intervals based on pivotal quantity $Q(T,\theta)$. hen, the shortest tolerance interval needs to be solved with minimizing the length (expected length) simultaneously with respect to two factors.

With the technique for developing shortest admissible tolerance interval, there is one fact interesting to investigate. Chen, Huang and Welsh (2006) introduced the confidence interval of mode interval where this coverage interval guaratees the shortest with a fixed coverage probability. This is an admissible tolerance interval. With a fixed pivotal quantity, we then may derive the shortest confidence interval for this shortest coverage interval. It is then interesting to see if this two step tolerance interval is the shortest tolerance interval. We derive the shortest tolerance intervals for several distributions and use them to investigate the desired problem.

Theorem 6.1. Let $X_1, ..., X_n$ be a random sample from normal distribution $N(\mu, \sigma^2)$ where $\sigma > 0$ is known.
(a)

$$[\bar{X} - z_{\frac{1+\gamma}{2}}\sigma - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\frac{1+\gamma}{2}}\sigma + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}] \tag{6.1}$$

is a shortest admissible γ-content tolerance interval with confidence $1-\alpha$.
(b) The shortest $100(1-\alpha)\%$ confidence interval of γ-content mode coverage interval is the shortest admissible γ-content tolerance interval with confidence $1-\alpha$.

Proof. We can modify the standard confidence interval calculation to show that

$$1-\alpha = P\{\bar{X}+c_1\sigma+q_1\frac{\sigma}{\sqrt{n}} \le \mu+c_1\sigma < \mu+c_2\sigma \le \bar{X}+c_2\sigma+q_2\frac{\sigma}{\sqrt{n}}\} \quad (6.2)$$

with constants c_i and q_i subject to $P(c_1 \le Z \le c_2) = \gamma$ and $P(q_1 \le Z \le q_2) = 1-\alpha$. This shows that the choices of admissible γ-content tolerance interval with confidence $1-\alpha$ is

$$[\bar{X}+c_1\sigma+q_1\frac{\sigma}{\sqrt{n}}, \bar{X}+c_2\sigma+q_2\frac{\sigma}{\sqrt{n}}] \quad (6.3)$$

which is a $100(1-\alpha)\%$ confidence interval of γ-content coverage interval $[\mu+c_1\sigma, \mu+c_2\sigma]$.

To obtain the shortest admissible γ-content tolerance interval with confidence $1-\alpha$, we seek to minimize

$$L = \sigma[c_2 - c_1 + (q_2 - q_1)\frac{1}{\sqrt{n}}]$$

subject to

$$\int_{c_1}^{c_2} f_Z(z)dz = \gamma \text{ and } \int_{q_1}^{q_2} f_Z(z)dz = 1-\alpha \quad (6.4)$$

where $f_Z(z)$ is the density of the standard normal distribution. Equations in (6.4) give c_2 and q_2 as functions of c_1 and q_1 respectively and parially differentiating these two equations with respect to c_1 and q_1 respectively yied

$$f_Z(c_2)\frac{\partial c_2}{\partial c_1} - f_Z(c_1) = 0 \text{ and } f_Z(q_2)\frac{\partial q_2}{\partial q_1} - f_Z(q_1) = 0.$$

To minimize L, we set $\partial L/\partial c_1 0$ and $\partial L/\partial q_1 = 0$; that is,

$$\frac{\partial L}{\partial c_1} = \sigma[\frac{\partial c_2}{\partial c_1} - 1] = 0, \text{ and } \frac{\partial L}{\partial q_1} = \frac{\sigma}{\sqrt{n}}[\frac{\partial q_2}{\partial q_1} - 1] = 0,$$

but

$$\sigma[\frac{\partial c_2}{\partial c_1} - 1] = \sigma[\frac{f_Z(c_1)}{f_Z(c_2)} - 1] = 0 \text{ and } \frac{\sigma}{\sqrt{n}}[\frac{\partial q_2}{\partial q_1} - 1] = \frac{\sigma}{\sqrt{n}}[\frac{f_Z(q_1)}{f_Z(q_2)} - 1] = 0$$

if and only if $f_Z(c_1) = f_Z(c_2)$ and $f_Z(q_1) = f_Z(q_2)$, which imply that $c_1 = -c_2$ and $q_1 = -q_2$ are the desired solutions. Morever, restrictions in (6.4) indicates that $c_2 = z_{\frac{1+\gamma}{2}}$ and $q_2 = 1 - z_{\frac{\alpha}{2}}$ which, joining with (6.3), veries the theorem.

If we plug the shortest tolerance interval of (6.1) in (6.2), we have

$$1-\alpha = P\{\bar{X}-z_{\frac{1+\gamma}{2}}\sigma-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le \mu-z_{\frac{1+\gamma}{2}}\sigma < \mu+z_{\frac{1+\gamma}{2}}\sigma \le \bar{X}+z_{\frac{1+\gamma}{2}}\sigma+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\}. \tag{6.5}$$

This formulates (6.1) as a shortest confidence interval of the shortest coverage interval $(\mu - z_{\frac{1+\gamma}{2}}\sigma, \mu + z_{\frac{1+\gamma}{2}}\sigma)$ and then, for this case, the shortest confidence interval of the shortest coverage interval is the shortest tolerance inteval. □

Theorem 6.2. Let $X_1, ..., X_n$ be a random sample drawn from the exponential distribution with density function $f(x, \lambda) = \lambda e^{-\lambda x} I(x \ge 0)$.

(a) The interval

$$(0, -\frac{2\sum_{i=1}^n X_i}{\chi^2_\alpha(2n)} \log(1-\gamma)) \tag{6.6}$$

is a shortest admissible γ-content tolerance interval at confidence coefficinet $1-\alpha$.

(b) The shortest $100(1-\alpha)\%$ confidence interval of γ-content mode interval is the shortest admissible γ-content tolerance interval at confidence coefficinet $1-\alpha$.

Proof. The exponential distribution has the quantile function $F^{-1}(u) = -\lambda^{-1}\log(1-u)$, $0 < u < 1$. By letting $0 < q_1 < q_2 < \infty$ satisfying $1-\alpha = P\{q_1 \le \chi^2(2n) \le q_2)$ and $0 < \delta < 1-\gamma$, since $2\lambda\sum_{i=1}^n X_i \sim \chi^2(2n)$, then

$$\begin{aligned}1-\alpha &= P\{q_1 \le \chi^2(2n) \le q_2\}\\ &= P\{q_1 \le 2\lambda\sum_{i=1}^{n} X_i \le q_2\}\\ &= P\{-\frac{2\lambda\sum_{i=1}^{n} X_i}{q_2}\frac{log(1-\delta)}{\lambda} \le -\frac{log(1-\delta)}{\lambda}\\ &< -\frac{log(1-(\gamma+\delta))}{\lambda} \le -\frac{2\lambda\sum_{i=1}^{n} X_i}{q_1}\frac{log(1-(\gamma+\delta))}{\lambda}\}\\ &= P\{-\frac{2\sum_{i=1}^{n} X_i}{q_2}log(1-\delta) \le -\frac{log(1-\delta)}{\lambda} < -\frac{log(1-(\gamma+\delta))}{\lambda}\\ &\le -\frac{2\sum_{i=1}^{n} X_i}{q_1}log(1-(\gamma+\delta))\}\end{aligned}$$

Since $[-\frac{log(1-\delta)}{\lambda}, -\frac{log(1-(\gamma+\delta))}{\lambda}]$ is a γ-content covergae interval, the possible choices of admissible γ-content tolerance interval with confidence $1-\alpha$ include

$$[-\frac{2\sum_{i=1}^{n} X_i}{q_2}log(1-\delta), -\frac{2\sum_{i=1}^{n} X_i}{q_1}log(1-(\gamma+\delta))] \tag{6.7}$$

in terms of $\delta, 0 < \delta < 1-\gamma$ and q_1, q_2.

To obtain the shortest admissible γ-content tolerance interval with confidence $1-\alpha$, we seek to minimize

$$L = 2\sum_{i=1}^{n} X_i[\frac{log(1-\delta)}{q_2} - \frac{log(1-(\gamma+\delta))}{q_1}]$$

subject to

$$0 < \delta < 1-\gamma \text{ and } \int_{q_1}^{q_2} f_{\chi^2(2n)}(x)dx = 1-\alpha \tag{6.8}$$

where $f_{\chi^2(2n)}(x)$ is the density of the $chi-squre$ distribution $\chi^2(2n)$. For given q_1 and q_2 with $0 < q_1 < q_2$,

$$\frac{\partial L}{\partial \delta} = 2\sum_{i=1}^{n} X_i[-\frac{1}{q_2(1-\delta)} + \frac{1}{q_1(1-(\gamma+\delta))}] > 0$$

for $0 \le \delta < 1-\gamma$. As an increasing function of δ, L achieves minimum at $\delta = 0$ which, from (6.8), further indicates that $q_1 = \chi^2_{2n}(\alpha)$ and then (a) of the theorem is proved.

Since the γ-content mode interval is

$$(0, -\lambda^{-1}log(1-\gamma)], \tag{6.9}$$

a one sided quantile interval, which must have one sided confidence interval. On the other hand, (6.6) is the $100(1-\alpha)\%$ one sided confidence interval and then it is the shortest confidence interval of (6.9). Then, (b) of the theorem is proved. $\square$

Theorem 6.3. Let $X_1, ..., X_n$ be a random sample from Gamma distribution $Gamma(k, \beta)$ with known shape parameter $k > 1$ having density $f(x, \beta, k) = \beta^k x^{k-1} exp(-\beta x)/\Gamma(k)$.

(a) Let constants c_1, c_2, q_1 and q_2 satisfy the following conditions:

(ℓ_1) $\frac{f_{2k}(c_1)}{f_{2k}(c_2)} = \frac{q_1}{q_2}$

(ℓ_2) $\frac{f_{2nk}(q_1)}{f_{2nk}(q_2)} = \frac{c_2 q_2^2}{c_1 q_1^2}$

(ℓ_3) $\int_{c_1}^{c_2} f_{2k}(x)dx = \gamma$ and $\int_{q_1}^{q_2} f_{2nk}(x)dx = 1-\alpha$

where $f_{2k}(x)$ and $f_{2nk}(x)$, respectively, represent the densities of chi-square distributions with degrees of freedoms $2k$ and $2nk$. Then

$$[\frac{\sum_{i=1}^n X_i}{q_2} c_1, \frac{\sum_{i=1}^n X_i}{q_1} c_2], \tag{6.10}$$

stisfying conditions (ℓ_1) $-$ (ℓ_3), is a shortest admissible γ-content tolerance interval with confidence $1-\alpha$.

(b) The $100(1-\alpha)\%$ shortest confidence interval of the mode interval is the form of (6.10) that satisfies the following conditions (ℓ_2) and (ℓ_3) and

(ℓ_4) $f_{2k}(c_1) = f_{2k}(c_2)$.

Proof. Considering that c_1, c_2, q_1, q_2 satisfying condition (ℓ_3), since $2\beta\sum_{i=1}^n X_i \sim \chi^2_{2nk}$, we have

$$\begin{aligned}
1-\alpha &= P\{q_1 \le \chi^2_{2nk} \le q_2\} \\
&= P\{q_1 \le 2\beta \sum_{i=1}^n X_i \le q_2\} \\
&= \{\frac{2\beta\sum_{i=1}^n X_i}{q_2} c_1 \le c_1 < c_2 \le \frac{2\beta\sum_{i=1}^n X_i}{q_1} c_2\} \\
&= \{\frac{\sum_{i=1}^n X_i}{q_2} c_1 \le \frac{1}{2\beta} c_1 < \frac{1}{2\beta} c_2 \le \frac{\sum_{i=1}^n X_i}{q_1} c_2\} \qquad (6.11)
\end{aligned}$$

which indicates, with setting $[\frac{1}{2\beta}c_1, \frac{1}{2\beta}c_2]$ as a γ coverage interval for distribution $Gamma(k, \beta)$, that

$$[\frac{\sum_{i=1}^n X_i}{q_2}c_1, \frac{\sum_{i=1}^n X_i}{q_1}c_2] \tag{6.12}$$

is an admissible γ-content tolerance interval with confidence $1-\alpha$.

To obtain the shortest admissible tolerance interval, we seek to minimize

$$L = \sum_{i=1}^n X_i(\frac{c_2}{q_1} - \frac{c_1}{q_2}).$$

Now, partially differentiating $1-\alpha = \int_{q_1}^{q_2} f_{2nk}(x)dx$ with respect to q_1 yields

$$\frac{\partial q_2}{\partial q_1} f_{2nk}(q_2) - f_{2nk}(q_1) = 0,$$

and so

$$\frac{\partial L}{\partial q_1} = \sum_{i=1}^n X_i(-\frac{c_2}{q_1^2} + \frac{c_1}{\partial q_2^2}\frac{q_2}{\partial q_1}) = \sum_{i=1}^n X_i(-\frac{c_2}{q_1^2} + \frac{c_1}{q_2^2}\frac{f_{2nk}(q_1)}{f_{2nk}(q_2)}) = 0. \tag{6.13}$$

On the other hand, partially differentiating $\gamma = \int_{c_1}^{c_2} f_{2k}(x)dx$ with respect to c_1 yields

$$\frac{\partial c_2}{\partial c_1} f_{2k}(c_2) - f_{2k}(c_1) = 0,$$

and so

$$\frac{\partial L}{\partial c_1} = \sum_{i=1}^n X_i(\frac{\partial c_2/\partial c_1}{q_1} - \frac{1}{q_2}) = \sum_{i=1}^n X_i(\frac{f_{2k}(c_1)/f_{2k}(c_2)}{q_1} - \frac{1}{q_2}) = 0. \tag{6.14}$$

The conditions (ℓ_1) and (ℓ_2) are followed from (6.13) and (6.14) and then (a) of the theorem is proved.

For this Gamma distribution, the mode interval is an interval of the form

$$(c_1, c_2), \tag{6.15}$$

with smallest length $c_2 - c_1$ subject to $\int_{c_1}^{c_2} f_{2k}(x)dx = \gamma$. This leads that the mode interval is it of (6.15) subject to condition ℓ_4. From the derivation in (6.11), the $100(1-\alpha)\%$ shortest confidence interval of the mode interval is of the same form of (6.10) subject to conditions (ℓ_2) - (ℓ_4). $\square$

References

Bowker, A. H. and Goode, H. P. (1952). *Sampling Inspection by Variables*, McGraw-Hill Book Company, Inc.: New York.

Bucchianico, A. D., Einmahl, J. H. J. and Mushkudiani, N. A. (2001). Smallest nonparametric tolerance regions. *The Annals of Statistics*, 29, 1320-1343.

Carroll, R. J. and Ruppert, D. (1991). Prediction and tolerance intervals with transformation and/or weighting. *Technometrics*, 33, 197-210.

Chen, L.-A., Huang, J.-Y. and Welsh, A. (2005). The mode interval and its application to control charts and tolerance intervals. Submitted.

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

Goodman, L. A. and Madansky, A. (1962). Parameter-free and nonparametric tolerance limits: the exponential case. *Technometrics*. 4, 75-95.

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

Jilek, M. and Likar, O. (1960a). Tolerance limits of the normal distributions with known variance and unknown mean. *Australian Journal of Statistics*, 2, 78-83.

Jilek, M. and Likar, O. (1960b). Tolerance regions of the normal distribution with known μ, unknown σ^2. *Biom. Zeit.*, 2, 204-209.

Liman, M. M. T. and Thomas, D. R. (1988). Simultaneous tolerance intervals for the linear regression model. *Journal of the American Statistical Association*. 83, 801-804.

Marshall, A. W. (1949). A note on the power function of the Wald-Wolfowitz tolerance limits for a normal distribution. The RAND Corporation, Research Memorandum RM-271.

Mee, R. W., Eberhardt, K. R. and Reeve, C. P. (1991). Calibration and simultaneous tolerance intervals for regression. *Technometrics*. 33, 211-219.

Odeh, R. E. and Owen, D. B. (1980). *Tables for Normal Tolerance Limits, Sampling Plans, and Screening*. Marcel Dekker, Inc.: New York.

Owen, D.B. (1964). Control of percentage in both tails of the normal distrib ution. *Technometrics*, 6, 377-387. Errata (1960) 8, 570.

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

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

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

Wald, A. and Wolfowitz, J. (1946). Tolerance limits for a normal distribution. *Annals of Mathematical Statistics*. 17, 208-215.

Wallis, W. A. (1951). Tolerance intervals for linear regression. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, ed. by Neyman, University of California Press. 9, 43-52.

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