Journal of Statistical Planning and Inference 141 (2011) 1510–1518

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Group selection for production yield among k manufacturing lines

Chen-ju Lin [a,*], Wen Lea Pearn [b]

[a] Department of Industrial Engineering & Management, Yuan Ze University, Chungli, Taiwan
[b] Department of Industrial Engineering & Management, National Chiao Tung University, Hsinchu, Taiwan

ARTICLE INFO

Article history:
Received 14 September 2009
Received in revised form
18 September 2010
Accepted 2 November 2010
Available online 18 November 2010

Keywords:
Group selection
Production yield
Quality management
Bonferroni method

ABSTRACT

Producing qualified items or products is essential to meet the requirement preset by customers. Evaluation and selection of desired manufacturing lines become challenging tasks for decision makers. Production yield is one of the important factors in measuring production performance. The goal of this paper is to screen a group of manufacturing lines and identify the best one with the highest yield. For the production lines with extremely low fraction of defectives, the yield index, S_{pk}, is an efficient indicator for quality level. This paper considers the production selection problem by using S_{pk} to compare k $(k>2)$ manufacturing lines. A subset is constructed to contain the production lines with the highest yield. A systematic approach of test order k compares selected pairs of manufacturing lines along with the Bonferroni method is proposed to solve this problem. Each pair of production yields is compared by taking ratio. The paper provides critical values and required sample sizes of the group selection procedure. An application example on evaluating four power inductor productions is presented to illustrate the practicality of the proposed approach.

1. Introduction

Production line selection is a common problem faced by engineers or managers. Decision makers need to determine the best manufacturing lines based on preferred criteria such as quality, cost, processing time, and so on. Among these attributes, production quality is usually taken as the major consideration in decision making. The manufacturing line which produces many defectives is incapable and would incur extra material costs or even delay delivery. Evaluating multiple production lines and selecting the one that has the highest yield become critical issues in managing production systems. This paper investigates the production selection problem of comparing k $(k>2)$ production lines with a focus on production yield. The k production lines may refer to different manufacturing recipes, production procedures, or potential suppliers of the critical components.

Many industries require high quality of their products due to the features of items. For example, semiconductor, automotive, and integrated circuit assembly companies may set a minimum yield level higher than 99.9%. With the progress of manufacturing technique, the fraction of defective could be well controlled in these industries nowadays. It is clear that the conventional method of counting the frequency of nonconformities becomes inappropriate in evaluating production performance. An inspection may take thousands of samples to have one defective to be observed. Thus, a new measurement is necessary for judging the manufacturing lines with property of low defective rate.

* Corresponding author. Tel.: +886 3 4638800; fax: +886 3 4638907.
E-mail address: chenju.lin@saturn.yzu.edu.tw (C.-j. Lin).

0378-3758/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2010.11.013

When k equals two, classical hypothesis testing methods can be extended to evaluate the performance of two productions. Chou (1994) presents a likelihood ratio test to solve the selection problem. Daniels et al. (2005) uses confidence intervals to investigate process equality. Pearn et al. (2004, 2009) and Lin and Pearn (2010) apply two-phase test scheme to select the better process and examine the magnitude of the difference based on different performance measurements. When testing homogeneity of two or more production performance, Hubele et al. (2005) develops a Wald test for the products with one-sided specification. Chen and Chen (2006) examine process equality by F_{max} test followed by simultaneous confidence intervals for pairwise difference. Polansky (2006) derives a permutation method without assuming the distribution of investigated quality feature. In terms of choosing the manufacturing line with the highest capability, Tseng and Wu (1991) applies a modified likelihood ratio selection rule to select the best production line based on process variation. However, their approach is not appropriate when production mean also shifts away from target. Huang and Lee (1995) proposes subset selection approaches to identify the best process among several manufacturing lines. The authors emphasize the cost of failing customer's requirement. The square error loss is considered as the decision criteria.

The goal of this paper is to select a subset of production lines with the highest yield among k ($k>2$) manufacturing lines. The methods selecting the best subset mentioned above, however, do not directly adopt production yield as decision criterion. A yield index, S_{pk}, proposed by Boyles (1994), instead, has a one-to-one mapping to yield. Through analyzing such index, production yield can be identified. The following paper presents an efficient method to identify the best production lines among those with low fraction of defectives by using S_{pk} index. Section 2 reviews the statistical inference of S_{pk}. Section 3 proposes a systematic approach to choose a group containing the manufacturing lines associated with the highest production yield. Section 4 investigates the required sample size to eliminate the worse manufacturing lines at a pre-specified error rate and power level. Section 5 evaluates four 3D-accelerometer manufacturing lines in a factory to show practicability of the proposed method. Section 6 concludes the paper with discussions.

2. Review yield index of production evaluation

Production yield is defined as the percentage of items passing inspection. For production lines with two-sided specifications, the product must have investigated characteristic falling within the manufacturing tolerance to be considered as a qualified item. Under the definition, production yield can be calculated as Yield = $F(USL)-F(LSL)$ where USL is the upper specification limit, LSL is the lower specification limit, and $F(\cdot)$ is the cumulative distribution function (cdf) of the product characteristic. When the investigated characteristic is normally distributed, the expression can be rewritten as Yield = $\Phi[(USL-\mu)/\sigma]-\Phi[(LSL-\mu)/\sigma]$ where μ is the production mean, σ is the production standard deviation, and $\Phi(\cdot)$ is the cdf of a standard normal distribution.

2.1. Yield index S_{pk}

For the manufacturing line with extremely low fraction of defectives, it is not efficient to evaluate yield by counting the number of defectives. Based on the idea of production yield mentioned above, Boyles (1994) proposed a yield index defined as follows:

$$S_{pk}=\frac{1}{3}\Phi^{-1}\left\{\frac{1}{2}\Phi\left(\frac{USL-\mu}{\sigma}\right)+\frac{1}{2}\Phi\left(\frac{\mu-LSL}{\sigma}\right)\right\} \quad (1)$$

The S_{pk} index is a convenient and effective tool for performance assurance and process improvement. The index quantifies production performance by considering production location, production variation, and manufacturing specifications all together. The S_{pk} index provides a one-to-one mapping to yield: Yield = $2\Phi(3S_{pk})-1$. For example, $S_{pk}=1$ and 1.33 imply that nonconformities in parts per million (ppm) is 2700 and 66, respectively.

2.2. Estimation of the yield index

Production parameters μ and σ are usually unknown. Lee et al. (2002) considers the general estimator of S_{pk}

$$\hat{S}_{pk}=\frac{1}{3}\Phi^{-1}\left\{\frac{1}{2}\Phi\left(\frac{USL-\bar{x}}{s}\right)+\frac{1}{2}\Phi\left(\frac{\bar{x}-LSL}{s}\right)\right\} \quad (2)$$

where $\bar{x}$ and s are the sample mean and the sample standard deviation, respectively, which can be obtained from an in-controlled manufacturing lines. The exact distribution of $\hat{S}_{pk}$ is mathematically intractable. Lee et al. (2002) applies the first order of Taylor expansion to obtain the approximate distribution of $\hat{S}_{pk}$

$$\hat{S}_{pk}\cong S_{pk}+\frac{1}{6\sqrt{n}}\frac{W}{\varphi(3S_{pk})} \quad (3)$$

where

$$W=\frac{\sqrt{n}(S^2-\sigma^2)}{\sigma^2}\left\{\frac{USL-\mu}{2\sigma}\varphi\left(\frac{USL-\mu}{\sigma}\right)+\frac{\mu-LSL}{2\sigma}\varphi\left(\frac{\mu-LSL}{\sigma}\right)\right\}-\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}\left\{\varphi\left(\frac{USL-\mu}{\sigma}\right)-\varphi\left(\frac{\mu-LSL}{\sigma}\right)\right\} \quad (4)$$

$\varphi(\cdot)$ is the probability density function (pdf) of the standard normal distribution. By Central Limit Theorem, the W statistic is asymptotically distributed as $N(0, a^2+b^2)$ where a and b can be expressed by the precision index $C_p=(USL-LSL)/6\sigma$ and the accuracy index $C_a=1-|\mu-m|/d$, $d=(USL-LSL)/2$, $m=(USL+LSL)/2$,

$$a=\frac{1}{\sqrt{2}}\{3C_p(2-C_a)\varphi(3C_p(2-C_a))+3C_pC_a\varphi(3C_pC_a)\}, \tag{5}$$

$$b=\varphi(3C_p(2-C_a))-\varphi(3C_pC_a). \tag{6}$$

Therefore, $\hat{S}_{pk}$ is asymptotically distributed as $N(\mu_s,\ \sigma_s^2)$ with

$$\mu_s=S_{pk} \text{ and } \sigma_s^2=\frac{(a^2+b^2)}{36n[\varphi(3S_{pk})]^2}.$$

3. Group selection for high production yield

Let $\Pi_1,\ \ldots,\ \Pi_k$ be mutually independent manufacturing lines. The k $(k>2)$ production lines could refer to different manufacturing recipes, production procedures, or suppliers of the investigated component. The random samples $X_{i,j}$ is the quality characteristic collected from the jth item of the ith production line, $1\le i\le k$, $1\le j\le n_i$. Assume that the collected characteristics from the same population are independent and normally distributed with mean μ_i and variance σ_i^2. Under the normality assumption, the yield index, S_{pk}, is an effective indicator of production yield. Denote S_{pki} as the yield level of production line Π_i and $\hat{S}_{pki}$ as the estimated yield of production line i derived from $X_{i,j}$. The goal of this paper is to select a nonempty group of arbitrary size containing all possibly the best manufacturing lines associated with the largest S_{pk} value.

3.1. Group selection procedure for the highest yield

This paper adopts the ratio statistics to compare k $(k>2)$ yields together. Let $\hat{S}_{pk[1]}\le\hat{S}_{pk[1]}\le\ldots\hat{S}_{pk[k]}$ be the ordered estimators of S_{pk}. Define $R_i=\hat{S}_{pk[k]}/\hat{S}_{pki}$, where $\hat{S}_{pki}<\hat{S}_{pk[k]}$, as a measure of separation between the two Π's corresponding to $\hat{S}_{pk[k]}$ and $\hat{S}_{pki}$. The concept of the proposed selection procedure is to compare $(k-1)$ production lines to the line with the largest estimated yield. The proposed group selection procedure is as follows:

Step 1: Calculate sample statistics and derive estimated yield based on S_{pk}.

$$\overline{X}_i=\frac{\sum_{j=1}^{n_i}X_{i,j}}{n_i},\quad s_i^2=\frac{\sum_{j=1}^{n_i}(X_{i,j}-\overline{X}_i)^2}{n_i-1},$$

$$\hat{S}_{pki}=\frac{1}{3}\Phi^{-1}\left\{\frac{1}{2}\Phi\left(\frac{USL-\overline{X}_i}{s_i}\right)+\frac{1}{2}\Phi\left(\frac{\overline{X}_i-LSL}{s_i}\right)\right\}.$$

Step 2: Rank the estimated yields in an ascending order.

$$\hat{S}_{pk[1]}\le\hat{S}_{pk[2]}\le\ldots\le\hat{S}_{pk[k]}.$$

Step 3: Include Π_i in the selected subset if and only if the ratio statistic

$$R_i=\hat{S}_{pk[k]}/\hat{S}_{pki}<c_\alpha,\ 1\le i\le k.$$

c_α is a predetermined constant satisfying the overall error constraint and is greater than one. Since $\hat{S}_{pk[k]}/\hat{S}_{pk[k]}$ is always less than c_α, only $(k-1)$ ratio statistics need to be tested against c_α. The selected subset must contain the production line with the highest estimated S_{pk} by using the test procedure above.

Lin and Pearn (2010) investigates a simpler problem of testing H_0: $S_{pk2}/S_{pk1}\le 1$ vs. H_1: $S_{pk2}/S_{pk1}>1$. Unlike that paper which clearly specifies the test hypotheses before the experiment, here we cannot construct the hypotheses before calculating estimated yields since $\hat{S}_{pk[k]}$ is unavailable in advance. When testing more than two S_{pk}'s, two fundamental issues arise: (1) error inflation due to multiple tests and (2) unfixed population Π_i associated with $\hat{S}_{pk[k]}$. In the following sections, Section 3.2 discusses the Bonferroni method which handles the overall error rate problem. Section 3.3 calculates the critical value, c_α, under the proposed selection procedure along the Bonferroni adjustment.

3.2. The Bonferroni adjustment

The actual production yields are usually unknown in practice. Let $S_{pk[1]}\le S_{pk[2]}\le\ldots\le S_{pk[k]}$ be the ordered yield indices. Neither the values of the $S_{pk[s]}$ nor the pairing of the Π_i with the $S_{pk[s]}$ $(1\le i,s\le t)$ is known. The best production line may not be unique. A correct selection (CS) is the event that the production line associated with $S_{pk[k]}$ has been correctly chosen by the selection rule. When evaluating more than two production lines, multiple tests are required to identify the highest yield.

The Bonferroni method is an effective approach to handle multiple comparisons problem. The Bonferroni inequality states that

$$P\left(\bigcup_{i=1}^{g}E_i\right)\le\sum_{i=1}^{g}P(E_i). \tag{7}$$

Suppose that there are g tests in total. Let E_i be the event of falsely rejecting the ith test, $1 \le i \le g$. If the significance level of individual test is controlled at α/g, the probability of falsely rejecting *any* test is less than or equal to α by the Bonferroni inequality

$$P\left(\bigcup_{i=1}^{g} E_i\right) = 1 - P\left(\bigcap_{i=1}^{g} E_i^c\right) = 1 - \left(1 - \frac{\alpha}{g}\right)^g \le g \times \frac{\alpha}{g} = \alpha \tag{8}$$

The Bonferroni method controls the overall error rate by adjusting p-values and has been applied to different fields. For example, Chen and Chen (2006) examines every pair of processes to evaluate multiple manufacturing lines, Ludlow et al. (2007) applies the Bonferroni correction to medical experiments. However, the goal of this paper is to select all of the best production lines instead of ordering them, not all relationship between every two productions is of interest. Instead of making all-pairwise comparisons with $k(k-1)/2$ tests, the proposed selection rule in Section 3.1 will only carry out $(k-1)$ tests. As a result, the test order reduces from k^2 to k which is more efficient.

3.3. Critical values for group selection procedure

The production line with the maximal estimated S_{pk} is considered to have the highest yield among k lines. After fixing the value of $\hat{S}_{pk[k]}$, the proposed selection procedure conducts $(k-1)$ tests. For the $\hat{S}_{pki}$ slightly smaller than $\hat{S}_{pk[k]}$, the corresponding production line is also considered as one of the best lines. Suppose that $\hat{S}_{pkk}$ has $\hat{S}_{pk[k]}$, the hypotheses of comparing the yields of production lines k and i $(1 \le i < k)$ can be formed as H_0: $S_{pkk}/S_{pki} \le 1$ vs. H_1: $S_{pkk}/S_{pki} > 1$ where the test statistic $R_i = \hat{S}_{pk[k]}/\hat{S}_{pki}$ follows a distribution of $N(S_{pkk}, \sigma^2_{sk})/N(S_{pki}, \sigma^2_{si})$. By convolution, the pdf of R_i is

$$\begin{aligned} f_{R_i}(r) = {} & \frac{1}{\sqrt{2\pi}}(\sigma^2_{sk} + \sigma^2_{si}r^2)^{-3/2} \times (\sigma^2_{sk}S_{pki} + \sigma^2_{si}S_{pkk}r) \\ & \times \exp\left[\frac{1}{2}\left(\frac{S_{pki}}{\sigma^2_{si}} + \frac{S_{pkk}r}{\sigma^2_{sk}}\right) \times \left(\frac{\sigma^2_{sk}S_{pki} + \sigma^2_{si}S_{pkk}r}{\sigma^2_{sk} + \sigma^2_{si}r^2}\right) - \frac{1}{2}\left(\frac{S^2_{pki}}{\sigma^2_{si}} + \frac{S^2_{pkk}}{\sigma^2_{sk}}\right)\right] \\ & \times \left[2\Phi\left(\frac{\sigma^2_{sk}S_{pki} + \sigma^2_{si}S_{pkk}r}{\sigma_{si}\sigma_{sk}\sqrt{\sigma^2_{sk} + \sigma^2_{si}r^2}}\right) - 1\right] + \frac{\sigma_{si}\sigma_{sk}}{\pi(\sigma^2_{sk} + \sigma^2_{si}r^2)} \times \exp\left[-\frac{1}{2}\left(\frac{S^2_{pki}}{\sigma^2_{si}} + \frac{S^2_{pkk}}{\sigma^2_{sk}}\right)\right], \end{aligned} \tag{9}$$

where $-\infty < r < \infty$. Based on the sampling distribution of $R_i = \hat{S}_{pk[k]}/\hat{S}_{pki}$ and a given significance level, we reject H_0 and claim that production line i has a lower yield than line k if the test statistic $R_i \ge c_\alpha$. Otherwise, production line i may be at least as good as line k and is then selected into the subset when $R_i < c_\alpha$.

However, the test hypotheses are undetermined before data collection because the proposed selection procedure depends on the production line with $\hat{S}_{pk[k]}$, which is unknown. To control the overall error rate, the type I error rate that equals to k time the conditional type I error rate given that $\hat{S}_{pkk} = \hat{S}_{pk[k]}$ should be set to α. In another words, the probability of falsely rejecting any S_{pki}, $1 \le i < k$, given that H_0: $S_{pkk} \le S_{pki}$ and $\hat{S}_{pkk} = \hat{S}_{pk[k]}$ should be controlled at α/k. The requirement can be satisfied by adjusting the type I error of individual test to $\alpha/[k(k-1)]$

$$\Pr\{R_i \ge c_\alpha \,|\, H_0 : S_{pkk} \le S_{pki}, \hat{S}_{pkk} = \hat{S}_{pk[k]}\} \le \alpha/[k(k-1)] \tag{10}$$

Such Bonferroni adjustment then guarantees that the probability of correct selection, P(CS), is at least $1-\alpha$.

The maximal value of c_α among those investigated parameters of C_p and C_a is chosen to obtain conservative bound on the critical value for reliability purpose. Since $\hat{S}_{pk}$ has the largest variance at $C_a=1$, we calculate the critical value with the conditions of $S_{pki}=S_{pkk}=C$ and $C_{ai}=C_{ak}=1$, i.e.

$$\Pr\{R_i \ge c_\alpha \,|\, S_{pki} = S_{pkk} = C,\ C_{ai} = C_{ak} = 1, n_i, n_k\} = \alpha/[k(k+1)]. \tag{11}$$

For $C_{ai}=C_{ak}=1$ and equal sample size, the probability density function of R_i can be further simplified as follows:

$$f_{R_i}(r) = \sqrt{\frac{n}{\pi}}\frac{1+r}{(1+r^2)^{3/2}} \times \left\{2\Phi\left(\sqrt{\frac{2n}{1+r^2}} \times (1+r)\right) - 1\right\} \times \exp\left\{\frac{-n(1-r)^2}{1+r^2}\right\} + \frac{1}{\pi(1+r^2)}\exp\{-2n\} \tag{12}$$

which is a function of n only. Therefore, the value of c_0 is dependent on sample size but invariant under C.

Under the proposed selection rule, we can guarantee to include all of the best production lines at a confidence level of at least $1-\alpha$. Table 1 provides the critical values for $k=3, 4, \ldots, 6$, $n_i=n=30, 40, \ldots, 200$, at $\alpha=0.05, 0.1$. The production selection procedure can be also applied to general cases with unequal sample sizes by recalculating c_α.

3.4. Correct selection under least favorable configuration

When $S_{pk1}=S_{pk2}=\ldots=S_{pkk}$, all production lines are considered as equally the best. The selection procedure will lead to a wrong decision if any of the k production lines is not selected into the group. For a fixed sample size, such setting minimizes the probability of correct selection and thus is considered as the least favorable configuration (LFC). Controlling the overall

Table 1
Critical values of the group selection procedure with. n_i=n=30, 40, …, 200, k=3, 4, 5, 6, and α=0.05, 0.1.

n	k=3	k=4	k=5	k=6
α=0.05				
30	1.577	1.660	1.722	1.771
40	1.478	1.542	1.589	1.627
50	1.415	1.469	1.508	1.539
60	1.371	1.418	1.452	1.478
70	1.338	1.380	1.410	1.434
80	1.312	1.351	1.378	1.399
90	1.292	1.327	1.352	1.371
100	1.274	1.307	1.330	1.348
110	1.260	1.290	1.312	1.329
120	1.247	1.276	1.297	1.313
130	1.236	1.264	1.283	1.298
140	1.226	1.253	1.271	1.286
150	1.218	1.243	1.261	1.275
160	1.210	1.234	1.251	1.265
170	1.203	1.226	1.243	1.256
180	1.197	1.219	1.235	1.247
190	1.191	1.213	1.228	1.240
200	1.186	1.207	1.222	1.233
α=0.1				
30	1.494	1.577	1.638	1.687
40	1.412	1.478	1.525	1.563
50	1.359	1.415	1.455	1.486
60	1.322	1.371	1.406	1.433
70	1.294	1.338	1.369	1.393
80	1.272	1.312	1.341	1.363
90	1.255	1.292	1.318	1.338
100	1.240	1.274	1.299	1.317
110	1.227	1.260	1.282	1.300
120	1.216	1.247	1.269	1.285
130	1.207	1.236	1.257	1.272
140	1.199	1.226	1.246	1.261
150	1.191	1.218	1.237	1.251
160	1.184	1.210	1.228	1.242
170	1.178	1.203	1.220	1.234
180	1.173	1.197	1.214	1.226
190	1.168	1.191	1.207	1.220
200	1.163	1.186	1.201	1.213

error rate at α=0.05, Table 2 analyzes the probability of correct selection under the LFC by using the proposed selection procedure with the critical values listed in Table 1.

It is clear that the probability requirement of $P(CS) > 1-\alpha$ is attained for all settings in Table 2. When k gets larger, the Bonferroni method gets conservative where the type I error rate is less than the preset α level and probability of correct selection is greater than 95%. This phenomenon can be further improved by using other multiple comparisons technique in the future. Note that when applying the selection procedure to compare two production lines, probability of correct selection would be exactly $1-\alpha$ since no adjustment of significant level is required for having one test.

4. Required sample size for specified power

The group selection procedure in section 3 guarantees that the probability requirement of $P(CS) > 1–\alpha$. The probability that the selected subset contains all production lines associated with $S_{pk[k]}$ should be maintained. Once the sample sizes and the α risk are specified, the ability of eliminating the inferior production line is determined. To increase the probability of excluding the line with lower yield and to maintain probability of correct selection at the same time, sample sizes need to be increased.

The proposed method separates k production lines into two parts. The lines in the selected subset are considered as having the highest yield while those that are not selected are considered as having lower yields. Suppose that the production line k has $\hat{S}_{pk[k]}$ and is selected into the subset. To exclude the line that has yield level less than S_{pkk} from the selected group, the minimal required sample size for pre-specified power q and overall error α can be obtained by recursively searching the

Table 2
The probability of correct selection under the least favorable configuration, $S_{pk1}=S_{pk2}=\ldots=S_{pkk}$, at $n_i=n=30, 40, \ldots, 200$, $k=3, 4, 5, 6$, and $\alpha=0.05$.

n	$k=3$	$k=4$	$k=5$	$k=6$
30	0.958	0.962	0.965	0.967
40	0.957	0.961	0.963	0.965
50	0.957	0.961	0.963	0.964
60	0.957	0.960	0.962	0.964
70	0.957	0.960	0.962	0.963
80	0.957	0.960	0.962	0.963
90	0.957	0.960	0.962	0.963
100	0.957	0.960	0.961	0.963
110	0.957	0.959	0.961	0.962
120	0.957	0.959	0.961	0.962
130	0.956	0.959	0.961	0.962
140	0.956	0.959	0.961	0.962
150	0.956	0.959	0.961	0.962
160	0.956	0.959	0.961	0.962
170	0.956	0.959	0.961	0.962
180	0.956	0.959	0.961	0.962
190	0.956	0.959	0.961	0.962
200	0.956	0.959	0.961	0.962

Table 3
The required sample sizes for detecting the yield index smaller than $1/(1+p)$ of the largest yield index at power=0.7, 0.8, 0.9, 0.95 and $\alpha=0.05$.

p	$k=3$	$k=4$	$k=5$	$k=6$	$k=3$	$k=4$	$k=5$	$k=6$
	power=0.7				power=0.8			
0.1	939	1104	1226	1322	1155	1336	1469	1574
0.15	439	515	572	617	538	623	684	733
0.2	259	304	337	364	317	367	403	432
0.25	173	203	225	244	213	246	270	289
0.3	126	148	164	178	154	179	197	211
0.35	97	115	127	136	118	137	151	162
0.4	77	91	101	109	94	109	120	129
0.45	65	76	84	90	79	90	100	107
0.5	55	63	70	76	66	76	84	90
0.55	46	55	60	66	56	66	73	77
	power=0.9				power=0.95			
0.1	1489	1694	1843	1960	1798	2023	2184	2313
0.15	694	789	859	913	838	943	1018	1077
0.2	409	465	506	538	493	554	600	634
0.25	274	311	339	360	330	371	400	424
0.3	199	227	245	261	240	269	291	307
0.35	152	173	188	200	183	207	224	236
0.4	122	139	151	161	147	165	179	188
0.45	100	115	124	132	120	136	147	155
0.5	84	96	104	111	101	115	123	131
0.55	73	83	90	96	87	98	107	113

following constraints

$$\Pr\{R_i \ge c_\alpha | \mathrm{H}_0 : S_{pkk} \le S_{pki}, \hat{S}_{pkk} = \hat{S}_{pk[k]}\} \le \alpha/[k(k-1)], \quad (13)$$

and

$$\Pr\{R_i \ge c_\alpha | \mathrm{H}_1 : S_{pkk} > S_{pki}, \hat{S}_{pkk} = \hat{S}_{pk[k]}\} \ge q \quad (14)$$

Assume that the sample sizes are equal in every production line, Table 3 shows the required sample sizes for $q=0.7, 0.8, 0.9, 0.95$ and $S_{pki}=S_{pkk}/(1+p)$ where $p=0.1, 0.15, \ldots, 0.55$ is the smallest difference to detect in the ratio of S_{pk} at $\alpha=0.05$. Taking comparisons of the four production yields as an example; to reach a 70% power of rejecting the yield index that is around 2/3 of the highest yield index ($p=0.5$), the required sample size is around 60 at a significant level of 0.05. Table 3 shows that a smaller sample size is sufficient to reach the same power level when the gap of the yields is larger. The benefit of using the ratio statistics is that the selection procedure does not require the prior information of the value of $S_{pk[k]}$. Neither the critical points nor the sample size depends on S_{pk} level.

5. An application on power inductor production

5.1. Power inductor

Power inductor is a passive electrical device composed of magnetic material surrounded by coil. The component can store energy in a magnetic field generated by the electric current passing through it. The output is measured by inductance in the units of henries. Power inductors are used in DC–DC converters that convert voltage to desired level. The DC–DC converter parts appear in many applications such as cellular phones, light-emitting diode (LED) lighting, personal digital assistants (PDAs), and digital still cameras.

Here we investigate a specific type of digital still camera whose target inductance of the power inductor is 10 micro-henries (μH) with LSL=8 μH and USL=12 μH. If the observed inductance falls outside of the tolerance (LSL, USL), the quality of pictures and lifetime of cameras will be discounted. Four power inductor production lines under different manufacturing recipes are evaluated in terms of production yield at α=0.05. Fig. 1 displays the inductance data collected from the four production lines. There are only 60 samples available from each line. Normal probability plots in Fig. 2 suggest that the data from all manufacturing lines could be normally distributed.

5.2. Select all production lines with the highest yield

For the inductance data displayed in Fig. 1, we calculate the sample means, sample standard deviations and derive estimated yield levels of all manufacturing lines: $\hat{S}_{pk1} = 1.316$, $\hat{S}_{pk2} = 1.035$, $\hat{S}_{pk3} = 1.888$, $\hat{S}_{pk4} = 1.545$. Since production line number three has the highest estimated yield, all the other lines are compared with line 3: H_0: $S_{pk3} \le S_{pki}$ versus H_1: $S_{pk3} > S_{pki}$, i=1, 2, 4. The critical value is 1.418 at α=0.05 from Table 1. By taking ratio of two estimators, only the test statistic R_4 is smaller than the critical point. It implies that the performance of line 3 and 4 cannot be separated while line 3 is superior to line 1 and 2. Therefore, both line 3 and 4 are selected into the group. We conclude that the selected group contains all production lines corresponding to the highest yield with a 95% confidence level. The calculation results and final selection decisions are listed in Table 4.

Suppose that the probability of detecting the smallest difference with p=0.5 is required to be around 0.7. In other words, the test procedure should be able to reject the yield index which is 2/3 of the highest yield index with a power about 0.7. Table 3 suggests that 60 samples from each production line could serve the purpose. By fixing group number, sample size,

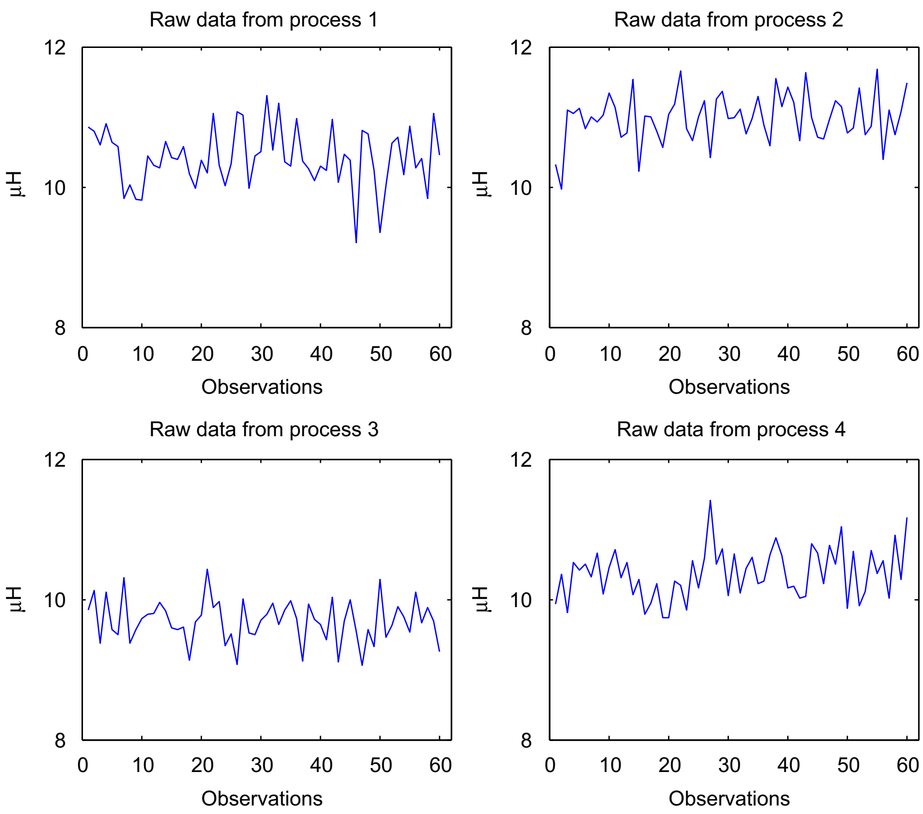

Fig. 1. The inductance data collected from the four production lines.

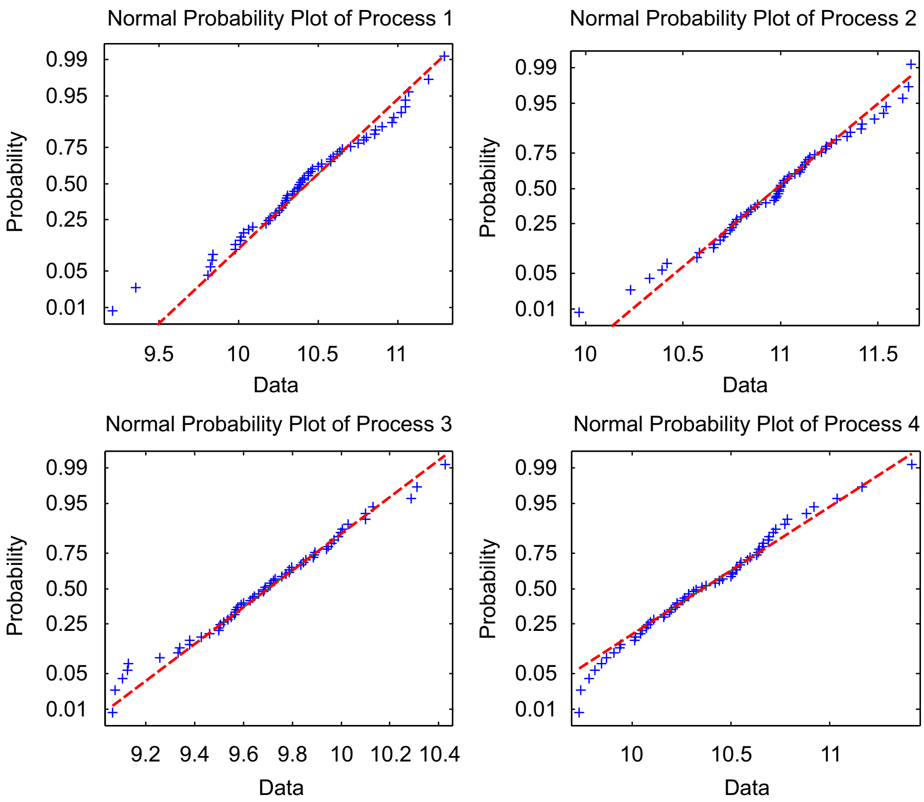

Fig. 2. The normal probability plots of the data collected from the four production lines.

Table 4
Sample statistics and selection decisions.

Production line	1	2	3	4
$\bar{x}$	10.415	10.985	9.691	10.369
s	0.419	0.351	0.305	0.363
$\hat{S}pk$	1.316	1.035	**1.888**	1.545
Estimated yield	99.992119%	99.810492%	99.999999%	99.999642%
R_i	1.435	1.823		1.222
Selection	N	N	Y	Y

Table 5
The powers of the selection procedure given p=0.5, 0.55, … , 0.75 at k=4, n=60, and c_α=1.418.

p	0.50	0.55	0.60	0.65	0.70	0.75
Power	0.67	0.75	0.82	0.88	0.92	0.95

type I error, and p, we can calculate numerical power by using eq. (9). Table 5 demonstrates the powers at given k=4, n=60, α=0.05 and p ranges from 0.5 to 0.75.

6. Conclusions

Selection of the best production lines with the highest yield is an important problem in production evaluation. The task is difficult especially when the fraction of defective is extremely low. Conventional method that examines the production lines by calculating the frequency of nonconformities is inefficient. This paper considers the production selection problem among multiple ($k>2$) two-sided manufacturing lines by adopting a yield index S_{pk}. For well-controlled normal production lines with two-sided specification limits, such index provides effective measures on production yield. Along with the Bonferroni method, a group selection approach is proposed to include all of the best production lines with the highest yield. Unlike all-pairwise comparisons with $k(k-1)/2$ tests, the selection procedure here only requires $(k-1)$ tests. The test order k, which is lower than k^2, further indicates efficiency of the proposed method. An application on selecting the best power inductor manufacturing lines among four candidates is presented to demonstrate the efficacy of the proposed approach.

Acknowledgments

The authors are grateful to the referees for their insightful comments that greatly improve the paper. This paper is supported by the National Science Council of Taiwan, Grant NSC 98-2221-E-155-008.

References

Boyles, R.A., 1994. Process capability with asymmetric tolerances. Communication in Statistics: Simulation and Computation 23 (3), 615–643.
Chen, K.S., Chen, K.L., 2006. Supplier selection by testing the process incapability index. International Journal of Production Research 44 (3), 589–600.
Chou, Y.M., 1994. Selecting a better supplier by testing process capability indices. Quality Engineering 6 (3), 427–438.
Daniels, L., Edgar, B., Burdick, R.K., Hubele, N.F., 2005. Using confidence intervals to compare process capability indices. Quality Engineering 17 (1), 23–32.
Huang, D.Y., Lee, R.F., 1995. Selecting the largest capability index from several quality control processes. Journal of Statistical Planning and Inference 46, 335–346.
Hubele, N.F., Berrado, A., Gel, E.S., 2005. A Wald test for comparing multiple capability indices. Journal of Quality Technology 37 (4), 304–307.
Lee, J.C., Hung, H.N., Pearn, W.L., Kueng, T.L., 2002. On the distribution of the estimated process yield index Spk. Quality and Reliability Engineering International 18 (2), 111–116.
Lin, C.J., Pearn, W.L., 2010. Process selection for higher production yield based on capability index S_{pk}. Quality and Reliability Engineering International 26 (3), 247–258.
Ludlow, C.L., Humbert, I., Saxon, Keith, Poletto, C., Sonies, B., Crujido, L., 2007. Effects of surface electrical stimulation both at rest and during swallowing in chronic pharyngeal dysphagia. Dysphagia 22 (1), 1–10.
Pearn, W.L., Hung, H.N., Cheng, Y.C., 2009. Supplier selection for one-sided processes with unequal sample size. European Journal of Operational Research 195 (2), 381–393.
Pearn, W.L., Wu, C.W., Lin, H.C., 2004. Procedure for supplier selection based on C_{pm} applied to super twisted nematic liquid crystal display processes. International Journal of Production Research 42 (13), 2719–2734.
Polansky, A.M., 2006. Permutation methods for comparing process capabilities. Journal of Quality Technology 38 (3), 254–266.
Tseng, S.T., Wu, T.Y., 1991. Selecting the best manufacturing process. Journal of Quality Technology 23 (1), 53–62.