SAMPLE SIZE DETERMINATION FOR LOWER CONFIDENCE LIMITS FOR ESTIMATING MULTIVARIATE PROCESS CAPABILITY INDICES

Jeh-Nan Pan[a] and Chung-I Li[b]

[a]*Department of Statistics, National Cheng Kung University, Tainan, Taiwan, 70101, ROC*

Postal address: No.1, University Road, Tainan City 701, Taiwan (R.O.C.)

E-mail address: jnpan@mail.ncku.edu.tw ;Tel: +886 6 2757575

[b] *Center for Quantitative Sciences, Vanderbilt University, Nashville, TN* 37232, *USA*

E-mail address: chunyilee1@gmail.com

ABSTRACT

With the advent of modern technology, manufacturing processes have become very sophisticated; a single quality characteristic can no longer reflect a product's quality. In order to establish performance measures for evaluating the capability of a multivariate manufacturing process, several multivariate process capability indices, such as NMC_p and NMC_{pm} etc, have been developed over the past few years. However, the sample size determination for multivariate process capability indices has not been thoroughly considered in previous studies. Generally, larger the sample size, the more accurate an estimate will be. However, too large a sample size may result in excessive costs. Hence, the trade-off between sample size and precision in estimation is a critical issue. In this paper, the lower confidence limits of NMC_p and NMC_{pm} indices are used to determine the appropriate sample size. Moreover, a procedure for conducting the multivariate process capability study is provided. Finally, two numerical examples are given to demonstrate that the proper determination of sample size for multivariate process indices can achieve a good balance between sampling costs and estimation precision.

Keywords: lower confidence limit; multivariate capability indices; sample size determination; statistical quality control; estimation precision

1. INTRODUCTION

Good quality products are a key factor to business success. With the advent of modern technology, manufacturing processes have become very sophisticated and a single quality characteristic can no longer reflect a product's quality. Generally, the quality of an industrial product has at least two correlated quality characteristics. Recently, various multivariate statistical methods have been employed to quality characteristics that are either interdependent or correlated (Chan et al, 1991; Chen,1994; Foster et al,2005; Hubele et al,1991; Pan and Lee,2010; Taam et al,1993; Wang and Chen,1998).

Normally, quality practitioners conduct process capability studies and estimate process capability indices to determine if a process is capable. To take into account the sampling error, (Pearn et al, 2007) investigated the statistical properties of the estimated MC_p and MC_{pm} indices proposed by (Taam et al, 1993) and derived sampling distributions for those estimators by constructing lower confidence limits using a variable transformation technique. However, this technique becomes more complicated when the number of quality characteristics increase. (Pan and Lee, 2010) modified the indices proposed by (Taam et al, 1993) and proposed indices NMC_p and NMC_{pm}. They also proposed a method to derive the exact sampling distributions of the estimated NMC_p and NMC_{pm}. Those sampling distributions lead to develop the confidence intervals and lower confidence limits. In this paper, the sample size determination for a multivariate capability analysis is provided based on the lower confidence limits of the NMC_p and NMC_{pm} indices. This facilitates the trade-off analysis between sampling costs and precise estimation. Generally, the smaller the sample size, the wider the confidence interval will be; the larger the sample size, the narrower the confidence interval will be. A larger sample size may result in excessive costs. On the other hand, a smaller sample size may reduce sampling costs while resulting in less precision. Therefore, the sample size determination becomes necessary since it is directly related to the cost of data collection.

In Section 2, we briefly introduced the lower confidence limits of NMC_p and NMC_{pm} indices proposed by (Pan and Lee,2010). In Section 3, based on those lower confidence limits, two equations and their associated tables for determining the appropriate sample size are derived. In Section 4, two numerical examples are given to demonstrate how to use our proposed method in performing the trade-off assessment between sampling costs and estimation precision. Finally, conclusions and discussions are drawn in Section 5.

2. THE LOWER CONFIDENCE LIMITS OF THE NMC_p AND NMC_{pm} INDICES

Pan and Lee (2010) claimed that the multivariate process capability indices MC_p and MC_{pm} proposed by (Taam et al, 1993) may overestimate the true process performance in certain situations, when the univariate quality characteristics are not independent. They revised the modified engineering tolerance region by (Taam et al, 1993) and proposed index

$$NMC_{pm} = \frac{NMC_p}{D},$$

where $NMC_p = (|\mathbf{A}|/|\mathbf{\Sigma}|)^{1/2}$, $D = (1 + (\boldsymbol{\mu} - \mathbf{T}))^{1/2}$, $\mathbf{\Sigma}$ is the variance-covariance matrix, $\boldsymbol{\mu}$ is the mean vector and $\mathbf{T}$ is the target vector. The elements of matrix $\mathbf{A}$ is given by

$$\rho_{ij}\left(\frac{USL_i - LSL_i}{2\sqrt{\chi^2_{v,0.9973}}}\right)\left(\frac{USL_j - LSL_j}{2\sqrt{\chi^2_{v,0.9973}}}\right),\ i,j = 1,\ldots,v,$$

where v is the number of quality characteristics, ρ_{ij} represents the correlation coefficient between the ith and jth quality characteristics and $(USL_i - LSL_i)$ denotes the specification width of the ith quality characteristics. Pan and Lee [6] further derived the lower confidence limits of NMC_p and NMC_{pm}

$$N\hat{M}C_p\sqrt{w_\alpha} \quad (1)$$

and

$$N\hat{M}C_{pm}\sqrt{\frac{w^*_\alpha}{(1+\hat{\delta}^2)}}, \quad (2)$$

where $N\hat{M}C_p = (|\mathbf{A}|/|\mathbf{S}|)^{1/2}$, $N\hat{M}C_{pm} = N\hat{M}C_p\big/(1+(\bar{\mathbf{X}}-\mathbf{T}))^{1/2}$, $\bar{\mathbf{X}}$ is sample mean vector, $\mathbf{S}$ is the sample variance-covariance matrix, w_α is the α percentile of $\prod_{i=1}^{v}\chi^2_{n-i}/(n-1)^v$ distribution, χ^2_{n-i}, $i = 1,\ldots,v$, are independent chi-square distributions with $(n-i)$ degrees of freedom, $\hat{\delta} = \sqrt{(\bar{\mathbf{X}}-\mathbf{T})'\mathbf{S}^{-1}(\bar{\mathbf{X}}-\mathbf{T})}$, w^*_α is the α percentile of $\chi^2_n(\lambda)\prod_{i=1}^{v-1}\chi^2_{n-i}/(n-1)^v$ distribution, $\chi^2_n(\hat{\lambda})$ is a non-central chi-square distribution with n degrees of freedom and non-centrality parameter $\hat{\lambda} = n(\bar{\mathbf{X}}-\mathbf{T})'\mathbf{S}^{-1}(\bar{\mathbf{X}}-\mathbf{T})$. To facilitate the construction of lower confidence limits and the confidence intervals for the MC_p and the MC_{pm} indices, the values of w_α and w^*_α obtained using numerical integration techniques for various values of n, v, α and δ are provided in Table 1 and Table 2. Because the real issue is to determine how small the process index can be, we utilized the lower confidence limit to determine the sample size in next section.

Table 1. The α percentile of $\prod_{i=1}^{v} \chi^2_{n-i}/(n-1)^v$ distribution

		α							
v	n	0.01	0.025	0.05	0.1	0.99	0.975	0.95	0.9
1	10	0.23159	0.30112	0.36971	0.46420	2.41251	2.11320	1.87998	1.63065
	15	0.33292	0.40323	0.46871	0.55691	2.08475	1.86000	1.69624	1.50444
	20	0.40061	0.46692	0.53252	0.61194	1.90362	1.72633	1.58826	1.43348
	25	0.45057	0.51628	0.57720	0.65067	1.78969	1.63883	1.51233	1.38467
	30	0.49105	0.55240	0.61076	0.68302	1.70768	1.57781	1.46752	1.34815
	50	0.59127	0.64456	0.69207	0.75107	1.53053	1.43464	1.35426	1.26582
	100	0.69998	0.74131	0.77753	0.82257	1.36019	1.29663	1.24299	1.18629
2	10	0.10350	0.14632	0.19492	0.26841	3.17116	2.59150	2.13554	1.70805
	15	0.19035	0.24426	0.30022	0.38031	2.64300	2.23015	1.92500	1.61193
	20	0.25569	0.31273	0.37527	0.45514	2.36884	2.05233	1.80658	1.54004
	25	0.30917	0.36957	0.42968	0.50879	2.19604	1.91884	1.70701	1.49103
	30	0.35088	0.41082	0.47131	0.54929	2.08612	1.84051	1.64625	1.45316
	50	0.46417	0.52263	0.57762	0.64468	1.79295	1.62613	1.49573	1.35641
	100	0.59692	0.64530	0.69180	0.74705	1.53140	1.42761	1.34459	1.25465
3	10	0.04359	0.06658	0.09401	0.13838	3.30065	2.52407	1.97247	1.48066
	15	0.10554	0.14512	0.18754	0.25070	2.85477	2.32230	1.91739	1.52299
	20	0.16354	0.21167	0.26055	0.33126	2.60295	2.17407	1.83983	1.51408
	25	0.21269	0.26518	0.31701	0.38896	2.43811	2.05088	1.76918	1.48481
	30	0.25567	0.31023	0.36456	0.43649	2.29852	1.96170	1.71516	1.46163
	50	0.37130	0.43103	0.48523	0.55612	1.96227	1.74099	1.56877	1.38890
	100	0.52038	0.56962	0.62170	0.68216	1.65465	1.51728	1.40851	1.28933
4	10	0.01497	0.02492	0.03776	0.06057	2.80171	2.02128	1.49468	1.04854
	15	0.05653	0.08092	0.10917	0.15152	2.78205	2.12007	1.69730	1.27770
	20	0.10177	0.13727	0.17286	0.22902	2.61057	2.09774	1.70793	1.36522
	25	0.14327	0.18623	0.22874	0.29042	2.47479	2.01672	1.70398	1.39268
	30	0.18337	0.22964	0.27626	0.34098	2.36881	1.96983	1.67822	1.39340
	50	0.29927	0.35154	0.40540	0.47378	2.07374	1.78905	1.58455	1.37238
	100	0.45776	0.50804	0.55954	0.62241	1.73595	1.57290	1.44061	1.30004

Table 2. The α percentile of $\chi_n^2(\lambda)\prod_{i=1}^{v-1}\chi_{n-i}^2/(n-1)^v$

			α							
v	n	δ	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
2	15	0.25	0.2496	0.3176	0.3885	0.4863	1.9467	2.3136	2.6758	3.1551
		0.5	0.2968	0.3779	0.4616	0.5769	2.2838	2.7096	3.1283	3.6806
		1	0.5165	0.6493	0.7854	0.9710	3.5732	4.1987	4.8085	5.6119
		1.5	0.9295	1.1519	1.3760	1.6774	5.6528	6.5693	7.4600	8.6143
		2	1.5374	1.8869	2.2351	2.6975	8.5152	9.8131	11.0628	12.6726
		2.5	2.3304	2.8457	3.3554	4.0249	12.1638	13.9377	15.6262	17.7878
		3.0	3.3073	4.0243	4.7299	5.6560	16.6152	18.9583	21.1775	24.0050
	25	0.25	0.3667	0.4366	0.5056	0.5963	1.7102	1.9577	2.1954	2.5007
		0.5	0.4353	0.5178	0.5989	0.7056	2.0059	2.2932	2.5690	2.9232
		1	0.7403	0.8733	1.0027	1.1715	3.1474	3.5715	3.9762	4.4922
		1.5	1.2951	1.5113	1.7198	1.9881	4.9960	5.6212	6.2123	6.9625
		2	2.1032	2.4336	2.7524	3.1582	7.5532	8.4372	9.2744	10.3268
		2.5	3.1543	3.6346	4.0935	4.6755	10.8151	12.0264	13.1620	14.5898
		3.0	4.4430	5.1072	5.7366	6.5356	14.7888	16.3944	17.8936	19.7556
	30	0.25	0.4078	0.4767	0.5439	0.6313	1.6445	1.8615	2.0683	2.3328
		0.5	0.4836	0.5649	0.6440	0.7465	1.9285	2.1808	2.4215	2.7263
		1	0.8183	0.9483	1.0738	1.2347	3.0287	3.4010	3.7532	4.1997
		1.5	1.4206	1.6302	1.8308	2.0855	4.8145	5.3643	5.8794	6.5311
		2	2.2924	2.6118	2.9161	3.3002	7.2827	8.0654	8.7959	9.7134
		2.5	3.4244	3.8893	4.3259	4.8746	10.4397	11.5111	12.5087	13.7470
		3.0	4.8117	5.4545	6.0554	6.8070	14.2875	15.7060	17.0248	18.6483
	50	0.25	0.5179	0.5815	0.6418	0.7179	1.4965	1.6487	1.7913	1.9695
		0.5	0.6131	0.6883	0.7588	0.8477	1.7560	1.9324	2.0982	2.3045
		1	1.0213	1.1401	1.2513	1.3900	2.7666	3.0292	3.2729	3.5764
		1.5	1.7434	1.9319	2.1063	2.3234	4.4139	4.8038	5.1631	5.6070
		2	2.7796	3.0639	3.3259	3.6493	6.6955	7.2527	7.7649	8.3934
		2.5	4.1241	4.5315	4.9055	5.3647	9.6156	10.3823	11.0844	11.9416
		3.0	5.7759	6.3327	6.8410	7.4665	13.1778	14.1952	15.1259	16.2616
	100	0.25	0.6486	0.7021	0.7511	0.8111	1.3582	1.4559	1.5456	1.6558
		0.5	0.7665	0.8293	0.8866	0.9568	1.5947	1.7084	1.8126	1.9405
		1	1.2613	1.3587	1.4476	1.5558	2.5214	2.6913	2.8461	3.0350
		1.5	2.1196	2.2711	2.4095	2.5770	4.0397	4.2918	4.5189	4.8007
		2	3.3385	3.5662	3.7711	4.0187	6.1499	6.5133	6.8427	7.2402
		2.5	4.9184	5.2395	5.5311	5.8811	8.8563	9.3579	9.8114	10.3568
		3.0	6.8544	7.2942	7.6882	8.1612	12.1559	12.8237	13.4253	14.1511

Table 2. The α percentile of $\chi_n^2(\lambda)\prod_{i=1}^{v-1}\chi_{n-i}^2/(n-1)^v$(continued)

			α							
v	n	δ	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
3	15	0.25	0.1554	0.2084	0.2663	0.3506	1.9877	2.4791	2.9888	3.6956
		0.5	0.1846	0.2471	0.3153	0.4148	2.3337	2.9058	3.4972	4.3174
		1	0.3161	0.4192	0.5315	0.6935	3.6816	4.5475	5.4371	6.6698
		1.5	0.5565	0.7311	0.9168	1.1835	5.8764	7.1950	8.5475	10.3886
		2	0.9053	1.1808	1.4733	1.8885	8.9238	10.8624	12.8227	15.4834
		2.5	1.3601	1.7678	2.1981	2.8033	12.8190	15.5187	18.2507	21.9327
		3.0	1.9146	2.4816	3.0778	3.9183	17.5762	21.2324	24.8988	29.8232
	25	0.25	0.2673	0.3300	0.3941	0.4817	1.7799	2.1109	2.4409	2.8799
		0.5	0.3163	0.3906	0.4661	0.5697	2.0889	2.4745	2.8586	3.3698
		1	0.5317	0.6523	0.7745	0.9400	3.2965	3.8819	4.4628	5.2305
		1.5	0.9181	1.1159	1.3150	1.5840	5.2692	6.1623	7.0358	8.1942
		2	1.4721	1.7792	2.0867	2.4995	8.0081	9.3151	10.5920	12.2714
		2.5	2.1887	2.6381	3.0859	3.6846	11.5185	13.3556	15.1438	17.4849
		3.0	3.0687	3.6912	4.3113	5.1359	15.8023	18.2749	20.6843	23.8193
	30	0.25	0.3090	0.3733	0.4379	0.5249	1.7150	2.0041	2.2883	2.6641
		0.5	0.3661	0.4418	0.5180	0.6201	2.0134	2.3505	2.6824	3.1185
		1	0.6127	0.7350	0.8571	1.0202	3.1786	3.6895	4.1891	4.8440
		1.5	1.0490	1.2497	1.4483	1.7114	5.0827	5.8650	6.6247	7.6120
		2	1.6749	1.9839	2.2906	2.6930	7.7310	8.8812	9.9893	11.4295
		2.5	2.4869	2.9369	3.3808	3.9619	11.1256	12.7332	14.2844	16.2865
		3.0	3.4812	4.1042	4.7171	5.5182	15.2646	17.4369	19.5173	22.2144
	50	0.25	0.4254	0.4901	0.5521	0.6323	1.5636	1.7651	1.9582	2.2061
		0.5	0.5034	0.5792	0.6521	0.7464	1.8359	2.0711	2.2961	2.5849
		1	0.8327	0.9535	1.0697	1.2187	2.9029	3.2604	3.6015	4.0360
		1.5	1.4072	1.6020	1.7878	2.0266	4.6554	5.2033	5.7218	6.3814
		2	2.2255	2.5245	2.8085	3.1710	7.0884	7.8938	8.6553	9.6179
		2.5	3.2837	3.7156	4.1246	4.6455	10.2127	11.3429	12.4127	13.7568
		3.0	4.5810	5.1772	5.7402	6.4539	14.0251	15.5491	16.9822	18.7904
	100	0.25	0.5709	0.6287	0.6822	0.7490	1.4119	1.5390	1.6573	1.8046
		0.5	0.6743	0.7419	0.8049	0.8832	1.6585	1.8069	1.9447	2.1167
		1	1.1037	1.2107	1.3094	1.4325	2.6296	2.8560	3.0659	3.3266
		1.5	1.8420	2.0119	2.1688	2.3637	4.2283	4.5764	4.8981	5.2962
		2	2.8875	3.1450	3.3829	3.6764	6.4549	6.9684	7.4413	8.0260
		2.5	4.2382	4.6074	4.9494	5.3691	9.3117	10.0343	10.6998	11.5192
		3.0	5.8915	6.3983	6.8662	7.4407	12.8002	13.7778	14.6758	15.7786

3. SAMPLE SIZE DETERMINATION FOR NMC_p AND NMC_{pm} INDICES

For univariate process capability indices, several researchers (Franklin, 1999 ; Pearn & Shu, 2003 ; Shu et al, 2006 ; Zimmer et al, 2001) constructed lower confidence limit that can be used to determine the sample size required for a given estimation precision. Based on a similar concept, we can use the lower confidence limits in Equations (1) and (2) to determine the sample size required for a desired estimation precision. According to the 100(1-α)% lower confidence limit in Equation (1), denoted as NMC_p^L, we have

$$NMC_p^L = N\hat{M}C_p \sqrt{\frac{u_\alpha}{(n-1)^v}} ,$$

where u_α is the α percentile of $\prod_{i=1}^{v} \chi^2_{n-i}$. Thus, the samples size needed for a given desired estimation precision ratio, $R_{NMC_p} = NMC_p / N\hat{M}C_p$, can be obtained as:

$$n = 1 + \left(\frac{u_\alpha}{R^2_{NMC_p}} \right)^v . \tag{3}$$

Moreover, Equation (3) can be rewritten as

$$\int_0^{R^2_{NMCp}} f_U(y)dy = \alpha , \tag{4}$$

where $f_U(y)$ is the probability density function of $\prod_{i=1}^{v} \chi^2_{n-i}$. Hence, given a desired estimation precision ratio R_{NMC_p} and confidence level $(1-\alpha)$, the sample size required can be obtained by solving Equation (4) via numerical integration. Based on the results of numerical integration, Table 3 displays the sample size, n, required for $R_{1-\alpha}$ (actual estimation precision ratio) > R_{NMC_p} (desired estimation precision ratio) with $v = 2, 3$, R_{NMCp} = 0.70, 0.75, 0.80, 0.85, and 0.90 and $1-\alpha$ =0.9, 0.95, 0.975 and 0.99.

Table 3. Sample size (n) required for estimating NMC_p index under different combinations of the number of quality characteristic (v) and desired estimation precision ratio (R_{NMC_p})

		$1-\alpha=0.9$		$1-\alpha=0.95$		$1-\alpha=0.975$		$1-\alpha=0.99$	
v	R_{NMC_p}	n	$R_{1-\alpha}$	n	$R_{1-\alpha}$	n	$R_{1-\alpha}$	n	$R_{1-\alpha}$
2	0.70	24	0.7062	33	0.7014	43	0.7006	57	0.7006
	0.75	33	0.7536	47	0.7514	62	0.7506	83	0.7505
	0.80	49	0.8012	72	0.8005	97	0.8007	131	0.8005
	0.85	83	0.8501	126	0.8502	172	0.8504	235	0.8503
	0.90	180	0.9002	279	0.9001	386	0.9002	532	0.9001
3	0.70	38	0.7035	52	0.7020	66	0.7001	87	0.7012
	0.75	52	0.7508	73	0.7503	96	0.7520	126	0.7504
	0.80	78	0.8010	112	0.8003	149	0.8005	199	0.8005
	0.85	132	0.8502	196	0.8503	264	0.8504	356	0.8502
	0.90	282	0.9004	430	0.9002	588	0.9003	807	0.9003

Note: $R_{1-\alpha}$ is the actual estimation precision ratio.

Similarly, we can obtain the following equation based on the approximate 100(1-α)% lower confidence limit for the NMC_{pm} index

$$\int_0^{R^2_{NMCpm}(1+\hat{\delta}^2)} f_{U^*}(y)dy = \alpha\,, \tag{5}$$

where $R_{NMC_{pm}}$ is the desired estimation precision ratio for NMC_{pm} index, $\hat{\delta} = \sqrt{(\bar{\mathbf{X}}-\mathbf{T})'\mathbf{S}^{-1}(\bar{\mathbf{X}}-\mathbf{T})}$ and $f_{U^*}(y)$ is the probability density function of $\chi^2_n(\lambda)\prod_{i=1}^{v-1}\chi^2_{n-i}/(n-1)^v$. Hence, given the desired estimation precision ratio $R_{NMC_{pm}}$, $\hat{\delta}$ and confidence level (1-α), an approximate sample size required can be obtained by solving Equation (5) via numerical integration. Based on the results of numerical integration, Table 4 displays the sample size required (n) for $R_{1-\alpha} > R_{NMC_{pm}}$ with $v = 2, 3$, δ = 0.25, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, R_{NMCpm} = 0.70, 0.75, 0.80, 0.85, and 0.90 and $1-\alpha =$ 0.9, 0.95, 0.975 and 0.99. To facilitate quality practitioners in conducting the multivariate process capability study, a step-by-step procedure is listed as below.

1. Design a sampling plan based on Table 3 and Table 4.
2. Collect measurement data and perform a normality test for the collected data.
3. Construct the lower confidence limits for both NMC_p and NMC_{pm} indices using Equations (1) and (2).
4. Evaluate the performance of a multivariate manufacturing process by using the lower confidence limits obtained in Step 3.

Table 4. Sample size (n) required for estimating NMC_{pm} index under different combinations of the number of quality characteristic (v), desired estimation precision ratio ($R_{NMC_{pm}}$) and δ

			$1-\alpha=0.9$		$1-\alpha=0.95$		$1-\alpha=0.975$		$1-\alpha=0.99$	
v	δ	$R_{NMC_{pm}}$	n	$R_{1-\alpha}$	n	$R_{1-\alpha}$	n	$R_{1-\alpha}$	n	$R_{1-\alpha}$
2	0.25	0.70	18	0.7048	27	0.7009	37	0.7004	51	0.7008
		0.75	26	0.7543	40	0.7523	55	0.7518	76	0.7516
		0.80	40	0.8011	63	0.8009	87	0.8001	121	0.8001
		0.85	71	0.8503	114	0.8507	159	0.8506	222	0.8502
		0.90	161	0.9002	260	0.9001	366	0.9002	512	0.9001
	0.5	0.70	18	0.7074	27	0.7037	37	0.7034	50	0.7003
		0.75	25	0.7511	39	0.7516	54	0.7514	74	0.7504
		0.80	39	0.8010	62	0.8009	86	0.8008	119	0.8002
		0.85	70	0.8509	112	0.8506	156	0.8502	218	0.8503
		0.90	158	0.9001	255	0.9001	359	0.9001	503	0.9003
	1.0	0.70	16	0.7068	24	0.7022	33	0.7025	45	0.7006
		0.75	23	0.7555	35	0.7515	48	0.7503	67	0.7507
		0.80	35	0.8012	55	0.8007	77	0.8007	107	0.8009
		0.85	62	0.8503	99	0.8501	140	0.8505	194	0.8501
		0.90	140	0.9001	228	0.9003	321	0.9002	450	0.9003
	1.5	0.70	14	0.7084	21	0.7032	29	0.7031	40	0.7034
		0.75	20	0.7560	30	0.7501	42	0.7511	58	0.7516
		0.80	30	0.8011	48	0.8013	67	0.8012	93	0.8008
		0.85	53	0.8500	87	0.8510	121	0.8504	169	0.8503
		0.90	121	0.9004	197	0.9002	278	0.9002	391	0.9002
	2.0	0.70	12	0.7020	19	0.7048	26	0.7029	36	0.7033
		0.75	17	0.7505	27	0.7516	38	0.7523	52	0.7507
		0.80	27	0.8026	43	0.8016	60	0.8011	83	0.8004
		0.85	48	0.8518	77	0.8506	108	0.8502	151	0.8503
		0.90	108	0.9005	176	0.9003	249	0.9005	349	0.9001
	2.5	0.70	11	0.7008	18	0.7077	24	0.7017	33	0.7012
		0.75	16	0.7534	25	0.7513	35	0.7516	49	0.7523
		0.80	25	0.8028	40	0.8022	55	0.8001	77	0.8005
		0.85	44	0.8514	71	0.8505	100	0.8501	140	0.8504
		0.90	100	0.9007	162	0.9001	230	0.9003	324	0.9003
	3.0	0.70	11	0.7089	17	0.7066	23	0.7026	32	0.7037
		0.75	15	0.7522	24	0.7526	33	0.7503	46	0.7506
		0.80	23	0.8003	38	0.8024	53	0.8018	73	0.8010
		0.85	41	0.8502	67	0.8503	95	0.8505	133	0.8502
		0.90	94	0.9005	154	0.9004	218	0.9003	308	0.9004

Table 4. Sample size (n) required for estimating NMC_{pm} index under different combinations of the number of quality characteristic (v), desired estimation precision ratio ($R_{NMC_{pm}}$) and δ

			$1-\alpha=0.9$		$1-\alpha=0.95$		$1-\alpha=0.975$		$1-\alpha=0.99$	
v	δ	$R_{NMC_{pm}}$	n	$R_{1-\alpha}$	n	$R_{1-\alpha}$	n	$R_{1-\alpha}$	n	$R_{1-\alpha}$
3	0.25	0.70	30	0.7029	44	0.7026	58	0.7009	78	0.7007
		0.75	42	0.7503	62	0.7508	85	0.7508	115	0.7502
		0.80	65	0.8004	100	0.8014	135	0.8002	186	0.8008
		0.85	115	0.8507	177	0.8501	244	0.8503	335	0.8501
		0.90	254	0.9001	400	0.9002	560	0.9002	775	0.9001
	0.5	0.70	30	0.7044	43	0.7011	58	0.7021	77	0.7006
		0.75	42	0.7514	62	0.7502	84	0.7504	114	0.7501
		0.80	65	0.8014	98	0.8002	134	0.8006	184	0.806
		0.85	113	0.8501	175	0.8503	241	0.8503	333	0.8503
		0.90	251	0.9002	396	0.9001	552	0.9003	768	0.9003
	1.0	0.70	28	0.7034	40	0.7003	54	0.7018	72	0.7006
		0.75	39	0.7509	58	0.7504	79	0.7514	107	0.7511
		0.80	60	0.8003	92	0.8013	125	0.8008	170	0.8002
		0.85	105	0.8504	163	0.8501	224	0.8501	310	0.8503
		0.90	234	0.9002	370	0.9002	520	0.9007	713	0.9002
	1.5	0.70	26	0.7038	37	0.7002	50	0.7022	67	0.7019
		0.75	36	0.7507	54	0.7520	72	0.7501	98	0.7504
		0.80	56	0.8018	84	0.8006	115	0.8006	157	0.8003
		0.85	97	0.8503	150	0.8502	207	0.8505	284	0.8503
		0.90	215	0.9002	338	0.9003	473	0.9002	654	0.9003
	2.0	0.70	24	0.7008	35	0.7006	46	0.7024	63	0.7010
		0.75	34	0.7508	51	0.7517	68	0.7506	93	0.7514
		0.80	52	0.8004	79	0.8002	108	0.8005	147	0.8002
		0.85	91	0.8507	141	0.8504	194	0.8503	268	0.8504
		0.90	201	0.9003	318	0.9003	443	0.9003	616	0.9004
	2.5	0.70	23	0.7003	34	0.7021	45	0.7013	61	0.7020
		0.75	33	0.7524	49	0.7521	65	0.7504	89	0.7511
		0.80	50	0.8006	76	0.8002	103	0.8002	142	0.8010
		0.85	87	0.8504	135	0.8504	185	0.8502	255	0.8003
		0.90	193	0.9004	304	0.9002	422	0.9001	591	0.9007

4. NUMERICAL EXAMPLES

Example 1

Sultan (1986) discussed an example in which the Brinell hardness (*BH*) and tensile strength (*TS*) are two quality characteristics for an industrial product. The engineering tolerances for *BH* and *TS* are given by (112.7, 241.3) and (32.7, 73.3) respectively and the target vector of *BH* and *TS* is $\mathbf{T} = (177, 53)$. After collecting 25 measurements as listed in Table 5, a multivariate process capability study is conducted (assuming the process is in control).

Table 5. The 25 measurements of Brinell hardness (*BH*) and tensile strength (*TS*) for an industrial product

BH	*TS*	*BH*	*TS*	*BH*	*TS*
143	34.2	141	47.3	178	50.9
200	57.0	175	57.3	196	57.9
168	47.5	187	58.5	160	45.5
181	53.4	187	58.2	183	53.9
148	47.8	186	57.0	179	51.2
178	51.5	172	49.4	194	57.5
162	45.9	182	57.2	181	55.6
215	59.1	177	50.6		
161	48.4	204	55.1		

By performing a Shapiro-Wilk test, we found that the 25 collected measurements follow a multivariate normal distribution with the sample mean vector $\bar{\mathbf{X}} = (177.2, 52.316)$ and the sample covariance matrix $\mathbf{S}$ are calculated, where

$$\mathbf{S} = \begin{bmatrix} 338 & 88.8925 \\ 88.8925 & 33.6247 \end{bmatrix}.$$

Furthermore, the matrix $\mathbf{A}$ is

$$\mathbf{A} = \begin{bmatrix} 349.52131 & 92.01022 \\ 92.01022 & 34.83724 \end{bmatrix}.$$

Then, the estimator of NMC_p index can be obtained as $N\hat{M}C_p = \sqrt{|\mathbf{A}|/|\mathbf{S}|} = 1.04$. With sample size $n = 25$, the number of quality characteristics $v = 2$ and $\alpha = 0.05$, one can locate $w_{0.05} = 0.42968$ from Table 1. Thus, a 95% lower confidence limit for the NMC_p

index is

$$N\hat{M}C_p\sqrt{w_{0.05}} = 0.68.$$

and a 95% confidence interval for the NMC_p index is

$$[N\hat{M}C_p\sqrt{w_{0.025}},\quad N\hat{M}C_p\sqrt{w_{0.975}}]$$
$$= [0.63,\quad 1.44].$$

Moreover, the estimator of NMC_{pm} index can be calculated as $N\hat{M}C_{pm} = 1.01$. With $n = 25$, $v = 2$ and $\hat{\delta} = \sqrt{1.3268/25} = 0.23$, we can locate $w^*_{0.05} = 0.5056$, $w^*_{0.025} = 0.4366$ and $w^*_{0.975} = 2.1954$ from Table 2. Thus, an approximate 95% lower confidence limit and a confidence interval for the NMC_{pm} index are given by

$$N\hat{M}C_{pm}\sqrt{\frac{w^*_{0.05}}{(1+\hat{\delta}^2)}} = 0.70$$

and

$$[N\hat{M}C_{pm}\sqrt{w^*_{0.025}/(1+\hat{\delta}^2)},\quad N\hat{M}C_{pm}\sqrt{w^*_{0.975}/(1+\hat{\delta}^2)}]$$
$$= [0.65,\quad 1.46].$$

The above interval estimates for both NMC_p and NMC_{pm} indices may lack precision (i.e. the wider widths of confidence interval) because of the small sample size ($n = 25$).

Suppose that the factory production resource allows a data collection plan be implemented with a sample size $n < 100$ and the desired estimation precision ratio is more than 0.80. We can locate the appropriate sample size required $n = 72$ with $v = 2$ and $1 - \alpha = 0.95$ as well as the corresponding actual estimation precision ratio $R_{1-\alpha} = 0.8005$ from Table 3. Thus, if an estimate $N\hat{M}C_p = 1.88$ is calculated based on 72 measurements, then we can conclude that the true value of NMC_p index is no less than $1.88 \times 0.8005 = 1.50$ at a 95% confidence level. Similarly, with $v = 2$, $\delta = 0.25$, $R_{NMCpm} = 0.8$ and $1 - \alpha = 0.95$, we can locate the approximate sample size required $n = 63$ from Table 4. Thus, if an estimate $N\hat{M}C_{pm} = 1.83$ is calculated based on 63 measurements, then one may conclude that the true value of NMC_{pm} index is no less than 1.83×0.8009=1.47 at a 95% confidence level. According to the above calculated results, we can guarantee that the actual estimation precision ratio is more than 0.80 at a 95% confidence level (i.e. $R_{1-\alpha=0.95} > 0.80$) for both NMC_p and NMC_{pm} indices. However, when the index is calculated based on 25 measurements, the actual estimation precision ratio is no more than 0.70 at a 95% confidence level (i.e. $R_{1-\alpha=0.95} < 0.70$). As expected, the larger the sample size, the more accurate the estimate will be. Yet a large sample size may result in excessive costs. The results shown in Table 3 and Table 4 provide a useful reference for achieving a good balance between sampling costs and estimation precision.

Example 2

Pan and Lee (2010) used 150 measurements of experimental data provided by Pan et al.(2004) in which deposited volume, deposited area and deposited height are the three key quality characteristics of a solder paste stencil printing process. In this example, the experimental data can be served as a baseline for determining the appropriate sample size required for conducting a multivariate process capability study in the future. Based on the experimental data, an estimator of δ (i.e. $\hat{\delta} = \sqrt{(\bar{\mathbf{X}} - \mathbf{T})'\mathbf{S}^{-1}(\bar{\mathbf{X}} - \mathbf{T})}$) can be calculated as $\hat{\delta} = 0.38$. Suppose that the true value of δ is between 0.25 and 0.5. With the number of quality characteristics $v = 3$, the desired estimation precision ratio $R_{NMCp} = 0.80$ and $1 - \alpha = 0.95$, the required sample size $n = 112$ can be located from Table 3. Similarly, with $v = 3$, $\delta = 0.25$ (0.5), $R_{NMCpm} = 0.80$ and $1 - \alpha = 0.95$, the required sample size $n = 100$ (98) can be located from Table 4. Considering the maximum value of 112, 100, and 98, the appropriate sample size $n = 112$ is then determined for conducting the multivariate process capability study. Therefore, given the desired estimation precision ratio is 0.8 at a 95% confidence level, the required sample size can be reduced from 150 to 112 (i.e. approximately 25% of sampling cost savings may be expected). This numerical example further shows that our proposed method for sample size determination provides a useful guideline for quality practitioners in performing the trade-off analysis between sampling costs and estimation precision.

5. CONCLUSIONS AND DISCUSSIONS

Process capability study is a common practice for evaluating the performance of a manufacturing process in industries. The sample size determination is very important since it directly relates to the cost of data collection. In the past few years, several researchers have addressed the issue of sample size determination for univariate process capability study. However, the sample size determination for multivariate process capability study has not yet been studied. In this paper, we proposed a method for determining the required sample size in conducting the multivariate process capability study. Moreover, based on the lower confidence limits of the NMC_p and NMC_{pm} indices, two equations and their associated tables are constructed to facilitate the calculation of appropriate sample size required for estimating the new multivariate process capability indices. Finally, two numerical examples have further demonstrated that the proper sample size determination can achieve a good balance between sampling costs and estimation precision.

It is worthy to note that the two parametric estimators of NMC_p and NMC_{pm} indices shown in the numerical examples are derived based on the normal assumption of the underlying multivariate distribution. Thus, any departure from the multi-normal assumption could result in erroneous calculation of interval estimates and wrong interpretation of process capability indices. Therefore, it is suggested that the multi-normal assumption of the collected data be checked by performing a normality test, such as Shapiro-Wilk test, prior to estimate the multivariate process capability indices. Moreover, it will be questionable to evaluate the performance of a multivariate manufacturing process when the sample size is small since the two interval estimates will be wider. In this special case, it is suggested that a nonparametric approach, such as the bootstrap method proposed by (Polansky, 2001), be utilized to construct interval estimates. If the nonparametric estimators agree to the parametric estimators, then we are more confident to draw the conclusions based on the results of multivariate process capability analysis. Otherwise, the alert is posted and a more thorough analysis is required to further explore the reason behind the disagreement.

References

Chan, L. K., Cheng, S. W., and Spiring, F. A. (1991). *A multivariate measure of process capability*, Int. J. Model. Simulation, 11, 1-6.

Chen, H. (1994). *A multivariate process capability index over a rectangular solid tolerance zone*, Statistica Sin., 4, 749-758.

Foster, E. J., Barton, R. R., Gautam, N., Truss, L. T., and Tew, J. D. (2005). *The process-oriented multivariate capability index*, Int. J. Prod. Res., 43, 2135-2148.

Franklin, L. A. (1999). *Sample size determination for lower confidence limits for estimating process capability indices*, Comput. Ind. Eng., 36, 603-614.

Hubele, N. F., Shahriari, H., and Cheng, C. S. (1991). *A bivariate process of capability vector*, Statistical Process Control in Manufacturing, J.B. Keats and D.C. Montgomery, eds., Marcel Dekker, New York, 299-310.

Pan, J. N., Lee, C. Y. (2010). *New capability indices for evaluating the performance of multivariate manufacturing processes*, Qual. Reliab. Eng. Int., 26 , 3-15.

Pan, J., Tonkay, G. L., Storer, R. H., Sallade, R. M., and Leandri, D. J. (2004). *Critical variables of solder paste stencil printing for Micro-BGA and fine-pitch QFP*, IEEE T. Electron. Pa. M., 27, 125-132.

Pearn, W. L., Shu, M. H. (2003). *Lower confidence bounds with sample size information for Cpm applied to production yield assurance*, Int. J. Prod. Res., 39, 2769-2781.

Pearn, W. L., Wang, F. K., Yen, C. H. (2007). *Multivariate capability indices: distributional and inferential properties*, J. Appl. Stat., 34, 941-962.

Polansky, A. M. (2001). *A smooth nonparametric approach to multivariate process capability*, Technometrics, 43, 199-211.

Shu, M. H., Wang, C. H., and Hsu, B. M. (1986). *A review and extensions of sample size determination for estimating process precision and loss with a designated accuracy ratio*, Int. J. Adv. Manuf. Technol., 27,1038-1046.

Sultan, T. L. (1986). *An acceptance chart for raw materials of two correlated properties*, Qual. Assurance, 12, 70-72.

Taam, W., Subbaiah, P., and Liddy, J. W. (1993). *A note on multivariate capability indices*, J. Appl. Stat., 20, 339-351.

Wang, F. K., Chen, J. C. (1998). *Capability index using principal component analysis*, Quality Eng., 11, 21-27.

Zimmer, L. S., Hubele, N. F. and Zimmer, W. J. (2001). *Confidence intervals and sample size determination for Cpm*, Quality Reliab. Eng. Interna.,17, 51-68.